

전자기록 장기보존 전략으로서의 에플레이션 사례 분석*

김 명 훈** · 오 명 진***

이 재 홍**** · 임 진 희*****

1. 머리말
2. 장기보존 전략으로서의 에플레이션 분석
 - 1) 에플레이션의 개념 및 특징
 - 2) 에플레이션의 유형
3. 에플레이션 전략 사례
 - 1) CAMILEON 프로젝트의 에플레이션 방식 분석
 - 2) KB의 에플레이션 방식 분석
 - 3) Planets 프로젝트의 에플레이션 방식 분석
 - 4) KEEP 프로젝트의 에플레이션 방식 분석
4. 시사점과 전망
 - 1) 사례 분석의 시사점
 - 2) 우리나라의 적용을 위한 전망

* 본 연구는 “안전행정부 국가기록원의 기록물 보존기술 연구 개발사업”의 일환인 “2013년도 행정기관 전자기록물 재현기술 연구 및 프로토타입 개발 연구사업”의 재정지원을 받아 수행되었음.

** 한국외국어대학교 정보·기록관리학과 겸임교수(주저자). 주요 논저 : 『전자기록 평가론』, 신진리탐구, 2009; 「전자기록 환경에서의 ‘업무친화적’ 기록관리 방향성 분석」, 『정보관리연구』 제38권 제4호, 2007.

*** 한국외국어대학교 정보·기록관리학과 강사(공동저자). 주요 논저 : 「공공영역의 2004년 대통령 탄핵사건 기록」, 『기록학연구』 제33호, 2012.

**** 명지대학교 디지털아카이빙연구소 연구원(공동저자).

***** 명지대학교 기록정보과학대학원 교수(교신저자). 주요 논저 : 「영구기록물관리를 위한 기록물 데이터베이스 스키마 개발 방향(공저)」, 『기록학연구』 제34호, 2012; 「대량기록물 처리를 위한 영구기록물관리시스템의 디지털저장소 배치형상 연구(공저)」, 『기록학연구』 제32호, 2012.

▪투고일 : 2013년 9월 23일 ▪최초심사일 : 2013년 9월 28일 ▪게재확정일 : 2013년 10월 25일.

[국문초록]

전자기록 환경을 맞이하여 필요한 기간 동안 장기적으로 전자 기록을 보존하는 것은 시대적 화두로 부상하고 있다. 특히 전자 기록이 지닌 본래의 기능적 속성 및 원래의 비트스트림을 유지하며 원래의 모습을 그대로 유지할 수 있는 방안은 전자기록의 장기보존을 위한 핵심 관건이 된다. 그동안 우리나라에서는 공공기록을 대상으로 마이그레이션 방식을 장기보존 전략으로 채택해 적용해 왔지만, 기능성 및 원래의 비트스트림을 원래 그대로 유지할 수 없다는 점에서 한계가 노정되어 왔다.

전자기록의 장기보존 전략으로서 에물레이션은 비트스트림의 변화 없이 전자기록의 원형을 그대로 재현할 수 있다는 점에서, 아울러 마이그레이션처럼 전자기록 유형별로 개별적인 전략을 적용할 필요 없이 일괄적인 보존전략을 구현할 수 있다는 점에서 상당한 강점을 지닌 보존전략이다. 특히 현재 우리나라 전자기록의 장기보존 포맷에서는 구현해 주지 못하는 전자기록의 기능적 구성요소를 재현해 줄 수 있다는 점에서, 현재 공공기관에 보유 중인 전자기록 유형 조사를 기반으로 그 실현 방안을 연구할 필요가 있다.

이에 본고에서는 전자기록의 장기보존 전략으로서 에물레이션의 적용 방안을 모색하기 위한 기초 연구로 에물레이션에 관한 서구의 최신 연구 사례를 검토하고, 에물레이션의 타당성과 대상, 방법에 대한 구체적인 사례를 고찰하려고 한다. 이를 통해 에물레이션이 지닌 장단점과 더불어 국내 적용을 위한 시사점을 모색하고자 한다. 이를 위해 본고에서는 에물레이션의 개념과 함께 장기보존 전략으로서 에물레이션이 지닌 장단점을 분석한 다음, 에물레이션에 관한 서구의 최신 선진 사례로 CAMILEON, KB, Planets, KEEP 프로젝트를 분석하였다. 이러한 분석을 기반으로

향후 우리나라 전자기록에 대한 시사점 및 적용 방향을 제시하고자 하였다.

주제어 : 에뮬레이션, 에뮬레이터, 마이그레이션, 전자기록, 장기 보존, 디지털 아카이빙, 디지털 컴포넌트

1. 머리말

전자기록은 0과 1의 비트스트림으로 이루어진 그 고유의 특성으로 인해 영구적으로 보존한다는 것은 아직 기술적으로 불가능하다. 이 때문에 전자기록의 장기보존 문제는 전 세계적으로 화두로 떠오르고 있으며, 장기보존 전략 개발을 위해 고군분투 중이다. 현재의 기술 수준에서 전자기록을 장기보존을 위한 전략은 몇 가지 활용되고 있다. 구형의 WordPerfect 파일이 있을 때 이를 두고 MS Word 파일로 변환하여 최신의 환경에서 문서를 읽을 수 있도록 하는 것이 마이그레이션이라면, 에뮬레이션은 WordPerfect가 구동되는 구형의 MS-DOS 환경을 재생산하여 원래의 환경에서 원래의 모습대로 열어 보게 하는 것이다. 또한 이를 표준화된 XML 포맷으로 변환하는 방법도 가능하다. 이것이 바로 전자기록을 장기보존하기 위한 대표적인 전략인 마이그레이션, 에뮬레이션, XML로서 이를 두고 전자기록의 3대 장기보존 전략이라고 말한다.

전자기록 영역은 상대적으로 텍스트의 비중이 높았기 때문에 지금까지 마이그레이션을 장기보존 전략으로 선택하는 사례가 일반적이었다.¹⁾ 하지만 컴퓨팅 환경은 계속적으로 진화하며 발전 속도가 매우 급

1) 네덜란드의 Digital Preservation Testbed는 기록 관리의 입장에서 보았을 때 이메일이나 텍스트와 같은 디지털 객체의 경우는 내용을 보존하는 것이 가장 중요하다 보고 마이그레이션이 가장 효과적인 전략이라고 결론짓기도 하였다. 또

속하기 때문에 그러한 모든 것을 수용하기에 마이그레이션이라는 단일 전략은 일정한 한계를 가질 수밖에 없다. 또한 장기보존 전략마다의 특성이 다르고 적용범주가 상이한 관계상 각기 처한 상황과 환경에 맞는 전략을 설정하는 것이 중요하다.²⁾

에뮬레이션(emulation)은 최근 전자기록의 잠재력 있는 장기보존 전략으로서 새롭게 주목받고 있다. 에뮬레이션은 소프트웨어의 원래 기능이 현재의 컴퓨터상에서 재현될 수 있도록 하기 위해 디지털 정보의 생산·활용된 원래의 운영 환경을 재생산하는 것을 의미한다. 이 보존 전략은 마이그레이션 및 XML 보존전략에서는 구현할 수 없는 장기보존상의 강점을 지닌다는 점에서, 양 보존전략을 대체 및 보완하는 차원에서 유용성을 지닌다. 스프레드시트와 데이터베이스와 같이 포괄적인 보존을 위한 노력이 필요한 경우에 에뮬레이션의 가능성을 확인한 사례가 이미 보고되기도 하였다.

국내의 경우 전자기록 장기보존을 위해 그동안 마이그레이션 및 XML 패키지 방법이 우선적으로 적용되어 왔다. 하지만 기술의 발달과 함께 문서 포맷 등은 지속적으로 변화하고 있으며 문서 생성 어플리케이션 중에서 더 이상 사용되지 않는 것들이 다수 존재하고 있다. 특히 디지털 객체에 포함된 기능적 요소나 원래의 외형과 느낌(look and feel)을 재현하는 데 있어서의 한계가 꾸준히 관찰되고 있다.³⁾ 아울러 국내에서 개발한 운영체제가 없는 상황에서 HW 및 SW에 비종속적으로 원문 기록물을 재현할 수 있는 에뮬레이터 개발의 필요성도 제기되고 있는 실정이다. 이제 에뮬레이션은 하나의 타당한 장기보존 전략이라는 전

한 국내의 국가기록원에서도 일찌감치 마이그레이션을 장기보존 전략으로 채택한 바 있다.

2) 오삼균 외, 「전자기록물 장기보존을 위한 에뮬레이션 전략 연구」, 국가기록원 연구결과보고서, 2009, p.3.

3) 예를 들어 보석글, 훈민정음, 하나워드부터 CAD Application, Lotus 123 등의 스프레드시트 유형 등이 이에 해당한다.

제에서 에플레이션에 대한 연구를 진행할 필요성이 시급히 제기되고 있다.

그동안 국내에서 마이그레이션에 관해서는 심도 있는 연구 성과가 제출되었지만,⁴⁾ 에플레이션에 관한 독립된 연구는 거의 진행되지 않았다. 전자기록의 장기보존 문제를 논의하는 과정에서 간략하게 소개되는 정도일 뿐,⁵⁾ 에플레이션 방식 및 사례에 대해서는 심도 있게 다루어지지 않았다. 다만 국내에서 본격적인 에플레이션에 관한 연구는 국가 차원의 장기보존 전략을 수립하기 위해 수행되었던 국가기록원의 연구 프로젝트 차원에서 이루어졌다. 먼저 오삼균 외는 국내의 공공 및 민간 기록 관리를 대상으로 장기보존 전략을 연구하였다.⁶⁾ 이 연구는 주로 마이그레이션 및 실행전략에 관한 내용이 주를 이루고 있고 에플레이션의 경우는 보존 전략의 하나로 그 개념, 특징, 적용 사례를 간략하고 소개하고 하는데 머물고 있다. 또한 오삼균 외는 장기보존의 3대 전략 중에서 에플레이션과 XML에 초점을 맞추어 연구를 수행하였다.⁷⁾ 이 연구에서는 에플레이션 전략과 관련하여 주로 외국 사례 및 관련 문헌조사를 통해 에플레이션 이론 및 접근 방법, 전략의 효과와 한계를 정리하는 한편 DPC와 Planets의 에플레이션 실험 워크플로우를 소개하고 있

-
- 4) 권도윤, 김희섭, 오삼균, 「전자기록물 장기보존을 위한 마이그레이션 전략에 관한 연구」, 『한국기록관리학회지』 제9권 제2호, 2009, 19-40쪽.
 - 5) 박은경, 「전자문서의 영구보존을 위한 표준화 연구동향」, 『한국기록관리학회지』 제2권 제1호, 2002, 97-118쪽; Cloonan, M. V., 「Shifting from Paper to Digital Records: Preservation?」, 『한국기록관리학회지』 제2권 제2호, 2002, 137-148쪽. 남성운, 윤대현, 「전자기록물의 장기보존을 위한 방안 연구: 개념을 중심으로」, 『한국기록관리학회지』, 제1권 제2호, 2001, 101-120쪽; 이원영, 강진영, 「전자문서 아카이빙 표준모델 연구」, 『정보관리학회지』 제22권 제2호, 2005, 147-164쪽; 유영수, 「전자기록관리를 위한 포맷등록시스템 개발 연구」, 『한국기록관리학회지』 제7권 제1호, 145-170쪽, 2007; 김명옥, 리상용, 「전자기록물의 장기보존을 위한 기능요소 연구」, 『한국기록관리학회지』 제10권 제2호, 2010, 101-126쪽.
 - 6) 오삼균 외, 「전자기록물 장기보존 전략 연구 및 테스트베드 구축」, 국가기록원 연구결과보고서, 2008.
 - 7) 오삼균 외, 앞의 글, 2009.

지만, 에플레이션 자체에 관한 상세한 분석에까지는 이르지 못한 한계를 지닌다.

이에 본고에서는 전자기록의 장기보존 전략으로서 에플레이션의 적용 방안을 모색하기 위한 기초 연구로 에플레이션에 관한 서구의 최신 연구 사례를 검토하고, 장기보존 전략으로서 에플레이션의 타당성과 대상, 방법에 대한 구체적인 사례 모색을 목적으로 한다. 이를 통해 에플레이션이 지닌 장단점과 더불어 국내 적용을 위한 시사점을 모색하고자 한다.

이를 위해 우선 2장에서는 에플레이션의 개념과 함께 장기보존 전략으로서 에플레이션이 지닌 장단점을 분석하였으며, 아울러 에플레이션의 기본적인 네 가지 방식에 대해 고찰하였다. 이어 3장에서는 에플레이션에 관한 서구의 선진 사례로 CAMILEON, KB, Planets, KEEP 프로젝트를 검토하였다. 이들 프로젝트는 에플레이션 전략에 관한 프로젝트 가운데 비교적 최근에 수행된 것으로 에플레이션 전략에 관심을 갖고 초기부터 다양한 연구를 수행해왔던 해외의 국립 도서관 및 아카이브 등의 문화유산기관을 비롯하여 대학 등의 연구기관, 게임회사, 소프트웨어 회사 등 유관 기관들이 참여하고 있다는 점에서 관련 동향을 고찰하는데 핵심적인 사례라고 볼 수 있다. 아울러 에플레이션에 관한 이론적 논의만이 아닌 실제 구현을 수행했다는 점에서 향후 우리나라의 적용방안 모색에 벤치마킹할 요소들을 도출할 수 있다. 그리고 4장에서는 이러한 분석을 기반으로 향후 우리나라 전자기록에 대한 시사점 및 적용 방향을 제시하고자 하였다. 다만 에플레이션의 실제 적용방안을 도출하기 위해서는 수많은 기반 연구 및 실제 사례 구현이 장기간에 걸쳐 이루어져야 한다. 이러한 점에서 본고에서는 우리나라 전자기록에 대한 구체적인 적용방안을 제시하기 보다는 향후의 방향성을 제시하는 것으로 논의의 범위를 한정하고자 하였다. 아무쪼록 향후 에플레이션 전략의 구체적 방법론 개발을 위한 토대가 되길 기대한다.

2. 장기보존 전략으로서의 에뮬레이션 분석

1) 에뮬레이션의 개념 및 특징

전자기록이 지닌 다양한 특성으로 인해 필요한 기간만큼 전자기록을 진본성 및 무결성, 신뢰성, 이용가능성을 유지하며 보존하는 것은 용이한 일이 아니다. 이는 다양한 원인에서 연유하지만, 수정·변조·복제가 용이한 전자기록의 특성과 함께 전자기록을 생성한 컴퓨터 기술의 발전 속도가 매우 급속하다는 점에서 그 이유를 찾을 수 있다. 전자기록의 장기보존을 위해서는 다양한 방법이 적용될 수 있지만, 무엇보다 전자기록이 지닌 본래의 기능성 및 원래의 비트스트림을 유지하며 원래의 모습을 그대로 유지할 수 있는 방안은 전자기록의 장기보존을 위한 핵심 관건이 된다.

에뮬레이션은 바로 이를 위한 최적의 장기보존 전략이라 할 수 있다. 에뮬레이션은 소프트웨어의 원래 기능이 현재의 컴퓨터상에서 재현될 수 있도록 하기 위해 디지털 정보가 생산·활용된 원래의 운영 환경을 재생산하는 것을 의미한다. 에뮬레이션은 현재의 컴퓨터에서 가동되면서 이를 통해 예전의 컴퓨팅 환경을 실제적으로 재현할 수 있는 프로그램인 에뮬레이터(emulator)란 소프트웨어를 필요로 하게 된다. 장기보존 전략으로서 에뮬레이션은 디지털 정보를 원본 그대로의 정확한 형태로 보존할 수 있는 가능성을 제공해줌과 아울러, 시간의 경과에 따른 시스템 내지 하드웨어·소프트웨어의 사양화에 따라 디지털 정보를 수차례 변환시킬 필요가 없는 특성을 지닌다. 이와 더불어 디지털 정보의 기타 보존 전략들에서는 디지털 정보의 포맷 및 유형, 기타 개별적 특성에 따라 분산화 된 복잡한 보존 전략들을 채택해야 함에 반해, 에뮬레이션은 일관된 보존 전략을 유지할 수 있으며 디지털 기록을 원래의 포맷

그대로 보존시킬 수 있는 잠재력을 지니고 있다.⁸⁾

현재 전자기록의 장기보존 전략으로 가장 광범위하게 사용되는 것은 마이그레이션이다. 마이그레이션은 그 방식이 상대적으로 간편하면서도 일반 상용 소프트웨어에서도 기능을 구현해주기 때문이다. 마이그레이션은 하나의 기술로부터 다른 기술로 디지털 객체를 복제 내지 변환하는 방식으로 디지털 객체에 초점을 맞춘 장기보존 전략이라면, 에뮬레이션은 하드웨어 내지 소프트웨어를 흉내 내어 원래의 디지털 객체를 재현하는 방식으로 디지털 객체의 재현 환경에 초점을 맞춘 장기보존 전략이라 할 수 있다. 일반적으로 마이그레이션 전략이 지닌 장점은 변환이 쉽고 해당 포맷의 공급자를 비롯한 다양한 관련자로부터 변환도구가 지속적으로 개발 공급되고 있기 때문에, 실행에 있어 개발 및 시행을 위한 비용이 낮다는 점을 들 수 있으며, 디지털 객체의 지적 내용을 보존하기 위한 신뢰성이 높은 방법으로서 텍스트 기반의 전자기록 장기보존에 적합한 전략이다.⁹⁾

하지만 마이그레이션은 원본의 논리적 구조 혹은 렌더링 방법을 상이하게 하여 새로 작성하는 방법으로, 빈번한 변환의 과정에서 데이터 또는 속성의 일부가 변경될 가능성을 가지고 있으며 복잡한 레이아웃의 외형이 손상될 가능성도 지니고 있다. 이는 곧 전자기록 장기보존상의 주요 이슈인 진본성과 무결성을 손상시킬 수 있다는 한계로 귀결된다. 이와 더불어 마이그레이션은 디지털 환경이 변경되는 매 시점마다 반복적으로 수행해야 할 필요가 있게 된다. 이는 장기적인 마이그레이션 전략의 수립을 어렵게 하는 한 요소로 작용하며, 마이그레이션 전략이 장기화될수록 보존 비용이 점차 증가하게 됨을 의미하게 된다.¹⁰⁾

8) Digital Preservation Testbed, *Emulation: Context and Current Status*, Digital Preservation Testbed White Paper, 2003, pp.22-25.

9) 권도윤, 김희섭, 오삼균, 「전자기록물 장기보존을 위한 마이그레이션 전략에 관한 연구」, 『한국기록관리학회지』 제9권 제2호, 2009, 24쪽.

10) 권도윤, 김희섭, 오삼균, 「전자기록물 장기보존을 위한 마이그레이션 전략에 관

에뮬레이션 전략은 마이그레이션 전략이 지닌 위와 같은 단점들을 상쇄할 수 있는 방안을 제공해준다. 에뮬레이션 전략이 지닌 장기보존 전략으로서의 장점을 여러 측면에서 도출할 수 있는데, 우선 호주 국립 기록청에서는 디지털 정보의 장기보존 전략으로 에뮬레이션이 지닌 장점을 원래 데이터 포맷의 ‘외형과 느낌(look and feel)’을 재생산할 수 있다는 점을 들고 있다.¹¹⁾ 네덜란드의 Digital Preservation Testbed에서는 에뮬레이션 전략이 지닌 강점으로 원본 형태 그대로 디지털 객체를 보존할 수 있으며, 저장되는 각 포맷 및 각각의 개별 객체에 대해서도 0에 가까운 증분비용(incremental cost)을 제공할 수 있다고 제시한다. 이와 더불어 에뮬레이션 전략은 보편적으로 적용할 수 있는 단일적이면서도 일관성 있는 보존 전략을 구현할 수 있게 해주며, 사양화된 포맷에 있는 기록물을 포기할 필요 없다는 점 역시 에뮬레이션 전략이 지닌 강점으로 꼽고 있다.¹²⁾ 한편 에뮬레이션에 관한 권위자인 Stewart Granger 역시 디지털 객체의 장기보존 전략으로 에뮬레이션이 다양한 이점을 지닌다고 보고 있는데, 우선 원본 디지털 객체를 가장 정확한 형태로 보존이 가능하다는 점과 더불어, 마이그레이션과 달리 시간의 흐름에 따라 지속적인 변환 프로세스를 거치지 않아도 된다는 점을 제시하고 있다. 그리고 마이그레이션 전략이 각 디지털 객체에 따라 서로 다른 형태를 띠게 되는 것에 비해, 에뮬레이션은 일관된 보존 전략을 수립할 수 있다는 점 역시 다양한 보존 전략들에 비해 에뮬레이션이 지닌 강점으로 파악하고 있다.¹³⁾

한 연구, 『한국기록관리학회지』 제9권 제2호, 2009, 25쪽.

11) National Archives of Australia, *Digital Recordkeeping: Guidelines for Creating, Managing and Preserving Digital Records*[Exposure Draft], 2004.

12) Digital Preservation Testbed, *Emulation: Context and Current Status*, Digital Preservation Testbed White Paper, 2003, pp.46-48.

13) Stewart Granger, “Emulation as a Digital Preservation Strategy”, *D-Lin Maganize* Vol. 6, No. 10, 2000. [cited 2013. 9. 2], (<http://www.dlib.org/dlib/october00/granger/10granger.html>).

그러나 에뮬레이션이 지닌 단점 역시 존재하는데, 호주 국립기록청(NAA)에서는 에뮬레이션의 직접적인 도구가 되는 에뮬레이터 소프트웨어의 생성이 고도의 기술력을 필요로 하고 많은 비용이 소요되며, 사유(proprietary) 소프트웨어의 에뮬레이션은 지적 자산과 저작권 문제를 발생시켜 장기보존 전략상의 효율성과 안정성을 손상시킨다는 점을 에뮬레이션 전략의 한계로 제시한다.¹⁴⁾ 또한 Stewart Granger는 에뮬레이션 전략이 고도의 기술 및 비용이 소요될 뿐만 아니라 마이그레이션에 비해 기술적으로 고려할 사항이 많기 때문에 실패할 확률이 높다고 보고 있으며, 소유권 및 지적 재산권이 존재하는 경우 에뮬레이션 자체가 지적 재산권과 관련된 문제를 야기할 수 있어 보존 전략상의 효율성과 안정성을 침해할 위험성이 있다고 파악한다. 그리고 에뮬레이션의 대상이 되는 시스템과 소프트웨어 환경 역시 다양하게 존재함으로써 지속적인 에뮬레이터 개발이 필요한 점 역시 장기보존 전략으로서 에뮬레이션이 지닌 단점 중 하나로 분석하고 있다.¹⁵⁾

장기보존 전략으로서의 에뮬레이션은 고도의 기술력 및 법적 제도적 인프라 마련이 필요한 보존전략이다. 이러한 점을 감안해 네덜란드의 Digital Preservation Testbed에서는 에뮬레이션 전략 적용 시 염두에 두어야 할 제반 문제를 다음과 같이 제시하고 있다.¹⁶⁾ 우선 에뮬레이터가 타겟 시스템의 속도, 디스플레이, 리프레쉬 율, 명암 및 색상, 사운드 등 모든 관련 기능을 완벽히 재생산할 수 있도록 하기 위해서는 현 기술력에 대한 보다 심도 있는 연구가 수행되어야 한다. 둘째는 유

14) National Archives of Australia, *Digital Recordkeeping: Guidelines for Creating, Managing and Preserving Digital Records*[Exposure Draft], 2004.

15) Stewart Granger, "Emulation as a Digital Preservation Strategy", *D-Lin Maganize*, Vol. 6, No. 10, 2000, [cited 2013. 9. 2], <<http://www.dlib.org/dlib/october00/granger/10granger.html>>.

16) 이하의 논지는 Digital Preservation Testbed, *Emulation: Context and Current Status*, Digital Preservation Testbed White Paper, 2003, pp.46-55에서 정리한 것임을 밝힌다.

효성(validation)의 확보 문제로 유효성은 현재 모든 디지털 보존전략상의 공통된 문제이지만, 에플레이션에 기반을 둔 보존전략은 유효성을 돕기 위한 자동화 방식의 형식테스트(formal test) 및 튜링테스트(turing test)에 대한 연구 수행이 필요하게 된다. 셋째는 vernacular extraction의 수행 방식 문제이다. 에플레이션이 지닌 강점 중의 하나는 원본으로부터 직접 대체 사본의 vernacular extraction이 용이하다는 점임을 감안할 때 자동적 내지 반자동적으로 수행되도록 하는 방안에 대한 연구가 필요하다. 넷째 I/O 장치 및 저장장치의 에플레이팅 문제이다. 컴퓨터 플랫폼의 에플레이팅은 단지 CPU의 에플레이팅만으로는 부족하며, 전자기록의 정확한 재현을 위해서는 I/O 장치나 저장장치와 같은 주변 장치 역시 에플레이팅을 해야 한다. 다섯째 다차원 기록의 에플레이션 문제이다. 전자기록이 멀티미디어 문서로 구성되거나 여러 컴퓨터들 간의 기술적 연계 속에 생산되는 경우에는, 모든 관련 컴퓨터 및 네트워크 하드웨어들을 에플레이션 하는 방안을 연구해야 한다. 여섯째 메타데이터의 포착 및 관리방안 문제이다. 에플레이션에서는 특정 기술 관련 메타데이터가 별도로 생산·저장될 필요가 없다고 그동안 제기되어 왔지만, 에플레이션을 통해 기록을 재생산하는데 필요한 절차를 기술하는 메타데이터들이 생산되어야 하며, 에플레이션에 기반을 둔 보존전략에서는 어떠한 종류의 메타데이터가 필요한지 향후 연구되어야 한다. 일곱째 지적 재산권 문제이다. 에플레이션의 기반이 되는 원래의 재현 소프트웨어는 대부분 사유 포맷인 관계상, 에플레이션을 기초로 한 보존전략을 구현하기 위해서는 사유 포맷과 관련된 저작권 문제를 해결하는 방안을 강구해야 한다. 마지막으로 전자기록의 장기보존은 무엇보다 경제적 요소가 성공의 기반임을 감안 할 때 에플레이션 전략상의 적절한 처리 및 비용모듈이 개발되어야 한다.

2) 에뮬레이션의 유형

(1) 누적 에뮬레이션(Stacked Emulation)

누적 에뮬레이션의 기본 모토는 에뮬레이터의 재사용으로, 최초에 개발된 에뮬레이터를 재사용하여 에뮬레이터의 중복적인 개발을 방지하는 전략이다. 즉 과거의 플랫폼을 재구성하기 위해 각기 실행되는 복합적인 에뮬레이터를 필요로 하는 플랫폼 의존적인 에뮬레이션으로, 가장 일반적인 방식의 에뮬레이션 형태라 할 수 있다.

〈그림 1〉 누적 에뮬레이션 개념도¹⁷⁾



누적 에뮬레이션의 기본 전략은 2005년도에 처음으로 개발된 에뮬레이터를 사용할 수 있는 에뮬레이터를 2010년 환경에 맞추어 개발하고,

17) Jeffrey van der Hoeven · Hilde van Wijngaarden · Remoco Verdegem · Jacqueline Slats, *Emulation: A Viable Preservation Strategy*(Version 1.4), The National Library of the Netherlands & Nationaal Archief of the Netherlands, 2005. 6. 20, p.17.

다시 2010년 에뮬레이터를 2015년도 환경에, 그리고 2015년 에뮬레이터를 2020년 환경에 맞추어 개발하는 것이다. 이는 곧 최초 에뮬레이터를 사용하기 위한 에뮬레이터를 새로운 환경에 맞추어 개발하는 전략이라 할 수 있다. 누적 에뮬레이션은 호스트 플랫폼의 한 부분에 대한 상세한 에뮬레이터에 의해 얻어지는 효율성을 지니지만, 에뮬레이터가 플랫폼 환경의 일정 부분만을 상세화하기 때문에 다른 플랫폼과의 호환성이 저하되며 영속성이 적다. 또한 각각의 에뮬레이터가 플랫폼 아래에서 매우 의존적이 되며, 에뮬레이터들 중 하나가 오류 내지 유실된다면 이 에뮬레이터에 의존하고 있는 체인은 불가피하게 손실될 것이고 어플리케이션과 문서에 접근할 수 없는 결과가 초래되게 된다.¹⁸⁾

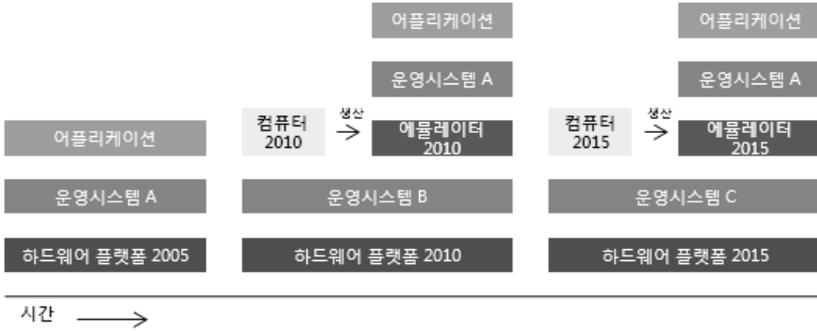
(2) 마이그레이션된 에뮬레이션(Migrated Emulation)

마이그레이션된 에뮬레이션은 에뮬레이션에 필요한 환경을 하드웨어 플랫폼이 변화함에 따라 컴파일러를 통해 에뮬레이터를 생성함으로써 누적 에뮬레이션의 단점을 보완하는 기술이다. 여기서 에뮬레이터의 논리적 구조는 변화되지 않고, 에뮬레이터와 변환된 플랫폼 사이의 매개체로써 컴파일러를 이용하게 된다. 새로운 환경으로의 에뮬레이터 변환 과정에는 X 언어로 쓰인 소스코드를 새로운 운영체제 플랫폼에서 실행될 수 있는 바이너리 코드로 전화해 줄 컴파일러가 필요하게 되며, 이를 통해 최초의 에뮬레이터는 새로운 환경에 맞추어 지속적으로 변환될 수 있다.¹⁹⁾

18) Ibid., pp.16-17.

19) Ibid., pp.17-18.

〈그림 2〉 마이그레이션된 에뮬레이션 개념도²⁰⁾



마이그레이션된 에뮬레이션은 위의 그림과 같이 도해할 수 있는데, 우선 에뮬레이터를 특정 환경에 맞추어 생성한 다음 이 환경이 변화할 경우 에뮬레이터는 새로운 환경에 맞추어 변형되게 된다. 이러한 변형은 새로운 플랫폼에서 실행될 수 있는 바이너리 형식의 언어로 소스코드를 변환하는 컴파일러를 통해 실행되며, 원본 에뮬레이터는 부분적으로 마이그레이션 되어 부가적인 환경을 생성하게 된다. 그리고 소스코드는 미래에 생성되기 때문에 환경적 문제를 야기할 수 있으며, 컴파일러의 라이브러리는 원본과 양립하지 않을 수 있기 때문에 기능 고장이 발생할 수 있다.²¹⁾

(3) 에뮬레이션 가상머신(Emulation Virtual Machine)

에뮬레이터는 시간 및 플랫폼으로부터 독립성을 확보하는 것이 가장 이상적이라 할 수 있다. 이에 에뮬레이션 전문 연구자인 Jeff Rothenberg

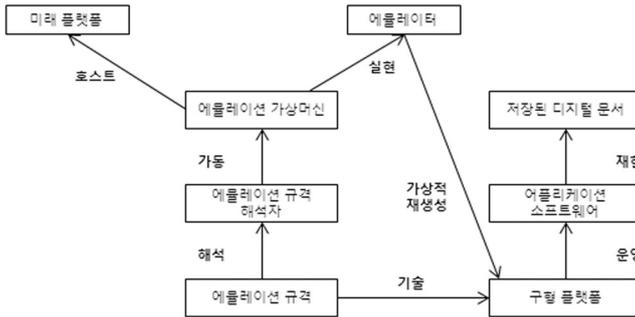
20) Jeffrey van der Hoeven · Hilde van Wijngaarden · Remoco Verdegem · Jacqueline Slats, *Emulation: A Viable Preservation Strategy*(Version 1.4), The National Library of the Netherlands & Nationaal Archief of the Netherlands, 2005. 6. 20, p.17.

21) 이소연 외, 「전자기록 관리와 보존을 위한 국제협력 아젠다 개발」, 국가기록원, 2007, pp.86-87.

는 호스트 플랫폼과 에뮬레이터 사이에 추가적인 레이어를 도입함으로써 이러한 독립성을 확보하는, 소위 에뮬레이션 가상머신(Emulation Virtual Machine: 이하 EVM으로 약칭)이라 칭하는 방식을 개발하였다.²²⁾

EVM의 개념적 도해는 아래의 그림과 같이 나타낼 수 있는데, 에뮬레이션 규격(Emulation Specification) 및 에뮬레이션 규격 해석자(Emulation Interpreter)와 함께 에뮬레이터는 EVM 상에서 가동될 수 있도록 생성된다.

〈그림 3〉 EVM 개념도²³⁾



에뮬레이션 규격은 원래의 소프트웨어가 구 컴퓨터 플랫폼에서 가동되는지를 설명해 주며, 이러한 에뮬레이션 규격은 구 플랫폼을 위한 에뮬레이터를 생성하는 에뮬레이션 해석자를 통해 해석되게 된다. 이를 통해 에뮬레이션 해석자와 생성된 에뮬레이터 양자는 EVM 상에서 가동되는데, 이는 곧 미래의 플랫폼에서 가동됨을 의미한다. EVM은 시간의 흐름과 관계없이 안정적으로 운용되며 아울러 다양한 호스트 플랫폼에

22) Jeff Rothenberg, "An Experiment in Using Emulation to Preserve Digital Publications", Koninklijke Bibliotheek: The Hague, 2000, [cited 2013. 9. 2], <<http://www.studioautomata.com/itp/indestudy/emulationpreservationreport.pdf>>.

23) Jeffrey van der Hoeven · Hilde van Wijngaarden · Remoco Verdegem · Jacqueline Slats, *Emulation: A Viable Preservation Strategy*, 2005. 6. 20, p.18.

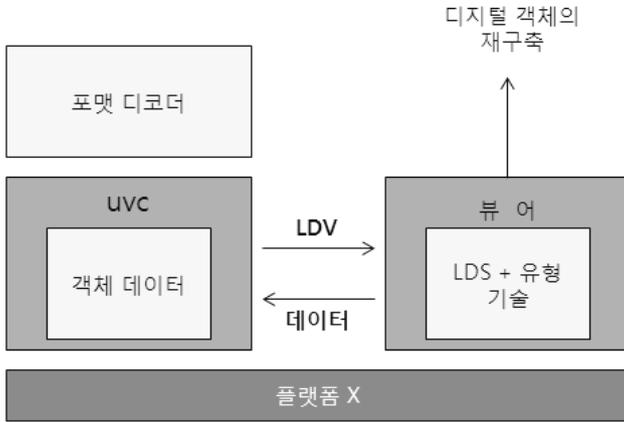
서 가동될 수 있다. 이러한 방식을 통해 애플레이터는 특정 플랫폼에 더 이상 의존적이지 않게 되며, 동일 EVM 상에서 다수의 애플레이터를 가동시킬 수 있게 된다. EVM 방식은 플랫폼 및 시간으로부터 독립성을 확보할 수 있지만, 다음과 같은 두 가지 단점을 지닌다. 우선 구 플랫폼은 단일의 규정 내에서 전체적으로 정의될 수 있는지, 게다가 복합적인 해석기는 이러한 규격을 애플레이터로 변환하는데 필요로 되는지의 문제가 아직 해결되지 못하였다. 또한 EVM은 시간의 흐름 속에서도 반드시 유지되어야 하는데 이것 역시 매우 복잡한 기술적 문제를 지닌다는 점이다.²⁴⁾

(4) UVC 기반 애플레이션

UVC는 미래의 어떠한 시점에서든 디지털 객체의 원본 재현성을 항구적으로 제공해 줄 수 있는 보존방식이다. 이 방식은 애플레이션과 마이그레이션의 결합에 기초하는데, 즉 마이그레이션은 원본 객체를 플랫폼 독립적인 포맷으로 변환할 때 수행되며 애플레이션은 플랫폼에 의존하지 않게 하는 절차에 활용된다. UVC는 일반 목적의 컴퓨터를 재생산함과 아울러 현재 및 미래의 컴퓨터 플랫폼에서든 쉽게 실행될 수 있는 소프트웨어 프로그램으로, 이러한 방식을 통해 UVC 언어로 작성된 프로그램은 플랫폼 독립적인 방식으로 UVC에서 실행되게 된다. UVC 프로그램(디코더)은 원본 객체의 포맷구조를 판독하고, 이러한 디코딩 과정은 Logical Data View(LDV)라고 불리는 XML과 유사한 구조를 제공한다. 아울러 Logical Data Schema(LDS)는 LDV에서 발생하는 요소들을 정의내리고 이 요소들이 의미하는 것에 대한 기타 정보들을 제공하며, 뷰어는 LDV 요소와 LDS 기술에 기초하여 원본 개체의 외형을 재구축한다. 이러한 절차는 아래의 <그림 4>와 같이 도식화할 수 있다.

24) Ibid., pp.18-19.

〈그림 4〉 UVC 기반 예물레이션 개념도²⁵⁾



이러한 전략을 기반으로 미래의 이용자들은 항상 원본 객체를 접근해 열람할 수 있지만, 현재 및 미래에 몇 가지 절차가 수행되어야 한다. 우선 현재 시점에서는 어떻게 UVC가 구축되어야 하는지를 정의하는 UVC 규격(UVC specification)이 보존되어야 한다. 또한 각각의 특정 파일 포맷에 대한 디코더가 각기 개발되어야 하며, LDS는 디지털 객체 각각에 필요하며 원본 객체 또한 보존되어야 한다. 검색 시점에서는 미래의 프로그래머는 UVC 규격에 기반을 둔 UVC를 재개발해야 하며, 원본 객체의 재현으로 LDV를 번역해 줄 하나 내지 그 이상의 뷰어가 생성되어야 한다.²⁶⁾

검색 및 보존 시점에서 요구되는 행위들은 이 전략의 장단점을 드러내 준다. 즉 극단적으로 광범위한 파일 포맷 및 객체 유형을 감안할 때, 만일 UVC 전략이 디지털 객체의 보존전략으로 채택된다면 수많은 포맷 디코더 및 LDS 기술이 개발될 필요가 있다는 점이다. UVC의 성패는 어

25) Ibid., pp.18-19.

26) Ibid., p19.

는 정도 범위까지 소프트웨어 및 컴퓨터 산업체들에 의해 수용되는지 여부와 함께 포맷 디코더 및 LDS 기술 활용도에 달려 있다. 하지만 일단 단일 LDS 및 디코더가 개발되면 UVC의 장점은 명백해진다. 즉 각각의 디지털 객체에 대해 마이그레이션을 수행하는 것에 비해, UVC 전략은 동일 포맷의 모든 객체들의 필요를 충족시킬 수 있다. 아울러 에물레이션 전략은 각각의 하드웨어 세대에 대해 새로운 에물레이터를 개발해야 하지만, UVC는 가능한 한 단순하게 설계됨으로 이러한 문제는 상대적으로 쉽게 해결될 수 있다.²⁷⁾

3. 에물레이션 전략 사례

1) CAMiLEON 프로젝트의 에물레이션 방식 분석

(1) CAMiLEON 프로젝트의 개요 및 목적

CAMiLEON(Creating Creative Archiving at Michigan and LEEDS: Emulating the Old on the New) 프로젝트는 미국 미시건대학교와 영국 리즈대학교와 공동으로 수행한, 에물레이션을 중심으로 한 디지털매체의 장기적 보존방법론을 개발하는 연구 프로젝트로, 디지털 자료들의 장기적 보존을 위한 기술적 전략들을 광범위하게 개발하고 평가하는 과정을 통해 상황에 따라 어떠한 장기보존 전략을 적용할 수 있을지에 대해 연구하였다. CAMiLEON이 주요 보존전략으로 채택하고 있는 에물레이션은 구형화된 시스템이 미래의 알려지지 않은 시스템 상에서도 실행될 수 있도록 하는 디지털 보존 전략으로서 제안되어 왔으며, 이러한 전략은 디지털 자

27) Ibid., p.20.

료들을 원래의 소프트웨어와 더불어 검색·디스플레이하고 이용할 수 있도록 하기 위함이다.²⁸⁾

CAMiLEON 프로젝트에서는 다음과 같은 세 가지의 핵심 목표를 제시하고 있다.²⁹⁾

- CAMiLEON은 디지털 객체의 원래의 기능성(functionality)과 ‘외형과 느낌(look and feel)’의 장기적 유지를 위한 선택사항들을 탐구
- 디지털 객체의 장기적 보존과 접근을 위한 장기적 전략으로서 기술 에뮬레이션(technology emulation)을 조사연구
- 이러한 에뮬레이션이 일련의 디지털 보존전략들과 어디서 어떻게 부합하는지(즉 어떠한 상황에서 에뮬레이션 전략이 적합한지) 여부

(2) CAMiLEON의 장기보존 전략

① 소프트웨어 수명의 장기화(software longevity)

CAMiLEON의 핵심 보존전략은 디지털 객체의 원래 바이트스트림을 해독하고 표현하는 소프트웨어의 설계와 생산을 포함한다. 소프트웨어 도구는 물론 디지털 객체로서 시간의 흐름에 따라 보존될 필요가 있으며, 이를 위해 CAMiLEON은 소프트웨어 수명의 장기화에 대한 기술을 개발하였다. 시간의 흐름에 따라 하나의 언어에서 다른 언어로의 간단한 소프트웨어 마이그레이션 통로를 생성함으로써, 소프트웨어 보존 도구는 경제적이고 정확하게 장기간동안 유지될 수 있다는 점에 서 있다.

28) <<http://www2.si.umich.edu/CAMiLEON/index.html>>.

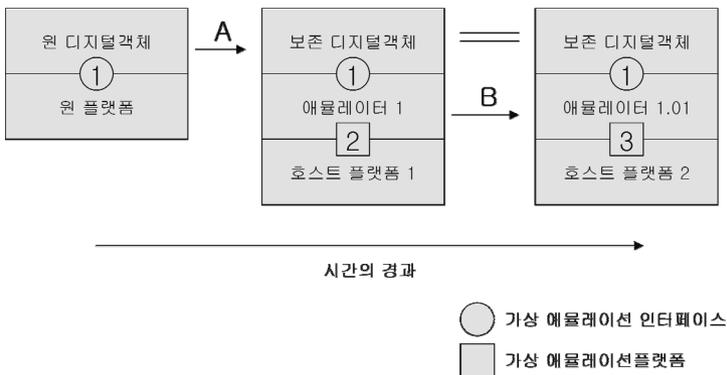
29) <<http://www2.si.umich.edu/CAMiLEON/index.html>>.

이것은 소프트웨어 도구가 하나의 언어에서 다른 언어로 경제적인 방법으로 마이그레이션 될 수 있도록 해준다.³⁰⁾

② 에뮬레이션 방식

에뮬레이션은 이전 시기에 생산된 디지털 객체를 열람하고 이용하는데 요구되는 기술적 환경을 현재의 하드웨어 상에서 재생산하는 것을 의미하는데, 그동안 에뮬레이션에 대한 실질적인 작업의 부족으로 인해 이러한 논쟁은 순전히 이론적인 수준으로만 남겨져 있는 실정이었다. 이에 CAMiLEON 프로젝트는 이러한 이론과 실제의 불균형을 시정하기 위해 에뮬레이션 전략에 관한 연구를 시도하였다. CAMiLEON 프로젝트에서 고안한 에뮬레이션 과정은 아래의 그림과 같이 도식화할 수 있다.

〈그림 5〉 에뮬레이션 개념도³¹⁾



30) [http://www2.si.umich.edu/CAMiLEON/reports/camresearch/research.html#software longevity](http://www2.si.umich.edu/CAMiLEON/reports/camresearch/research.html#software%20longevity).

31) David Holdsworth · Paul Wheatley, "Emulation, Preservation, and Abstraction". [cited 2013. 9. 2], <http://sw.ccs.bcs.org/CAMiLEON//dh/ep5.html>.

위의 다이어그램은 먼저 원본 디지털 객체를 보존된 디지털 객체로 변형하고, 원본이 지닌 중요 특성에 대한 접근이 에뮬레이션 하에서 실행될 수 있음을 의미한다. 시간의 흐름은 화살표 B에 의해 나타나는 변형을 초래하며, 그 안에서 플랫폼 1의 노후화는 다른 플랫폼 상에서 실행되는 에뮬레이터를 업데이트함으로써 조정하게 된다. 여기서 등호는 보존된 디지털 객체가 항구적으로 동일 비트스트림으로 구성되어 있음을 의미한다. 여기서 논쟁의 핵심은 두개의 화살표에 의해 표시되는 변형이 경제적으로 수행될 수 있다는 것으로 이 가운데 가장 중요한 것은 디지털 컴포넌트들 사이에 있는 접합점 상의 동심원들에 의해 표시되는 추상적인 인터페이스의 설계이다. 1번 인터페이스는 원본 디지털 객체의 성공적인 운영에 필요한 API(Application Program Interface)이고, 에뮬레이터 1은 이러한 인터페이스를 재현하며 영원히 같은 상태로 남기게 된다. 에뮬레이터 자체는 하나의 디지털 객체이며 그것의 성공적인 운영을 위해 인터페이스 2에 의존하게 된다. 핵심적인 전략적 목표는 거의 변경 없이 시간 테스트에서 생존할 가능성이 있는 유일한 특징을 포함하는 인터페이스 2를 선택함으로써 이루어지는 '미래의 증명(future proofing)'이다. 따라서 인터페이스 3은 인터페이스 2와 매우 흡사하며, 호스트 2에 대한 에뮬레이터(에뮬레이터 1.01)는 에뮬레이터 1로부터 바로 도출되게 된다.³²⁾

CAMiLEON에서는 이러한 에뮬레이션 전략을 BBC Domesday 프로젝트를 통해 실제 적용하였다. BBC Domesday 프로젝트는 11세기 영국의 인구조사 결과인 Domesday Book의 900주년을 기념하여, BBC와 Acorn Computer Ltd, Philips, Logica가 공동으로 조사한 영국 국민들의 삶에 대한 기록으로, 1984년에서 1986년 사이에 작성되어 1986년에 발표되었다. Domesday는 영국 각지의 지도, 다양한 지역사회 사람들의 일상 생활상에 관한 사진, 통계자료, 동영상 등을 수록하였는데, Domesday 프

32) Ibid.

로젝트의 저장매체는 LV-ROM(LaserVision Read Only Memory)으로 하였다. Domesday 프로젝트는 JPEG와 같은 포맷이 나오기 전에 진행되어 사양화에 대한 우려가 존재하는 대표적인 사례로 꼽힌다. 2002년 이후 Domesday 프로젝트의 보존 위기가 가시화 되었는데, LV-ROM을 사용할 수 있는 PC가 사라져 갔고 동영상의 경우 PC로는 정확한 재생이 불가능했기 때문이다. 이에 따라 CAMiLEON에서는 Domesday 프로젝트의 기록을 시간의 흐름 후에도 이용할 수 있도록 에뮬레이션 보존전략 연구를 수행하였다. CAMiLEON 연구진들은 과거의 시스템을 구현할 수 있는 에뮬레이터를 개발하여 원본 기록과 함께 에뮬레이션 소프트웨어 내에 인캡슐화하였다. 그러나 BBC의 'BeebEm' 에뮬레이터를 사용하여 Domesday의 콘텐츠를 이용 가능하게 하였지만, 원본 디스크에 대한 다양한 참여자 사이의 저작권이 복잡하게 얽혀 있어 실제 서비스는 불가능하였다.³³⁾

③ Migration on Request

시간의 흐름과 함께 하나의 포맷에서 다른 포맷으로의 마이그레이션은 보존 작업을 수행하기 위해 널리 인정된 방법이지만, 이상적인 것과는 거리가 있다. 즉 전통적인 마이그레이션은 변형하는 동안의 에러와 변형이 발생하며 고비용이 소요되는 단점을 지니고 있기 때문이다. 이에 CAMiLEON 프로젝트에서는 마이그레이션의 지속적인 수행에서 원형의 훼손을 최소화할 수 있는 마이그레이션 방법론을 개발하였고, 아울러 벡터 그래픽(vector graphics) 자원을 대상으로 테스트를 수행하였다.³⁴⁾

CAMiLEON은 이에 대해 Migration on Request라는 대체적인 접근법을 개발했는데, 이는 이용의 시점에 객체를 마이그레이션하는 도구와 함께

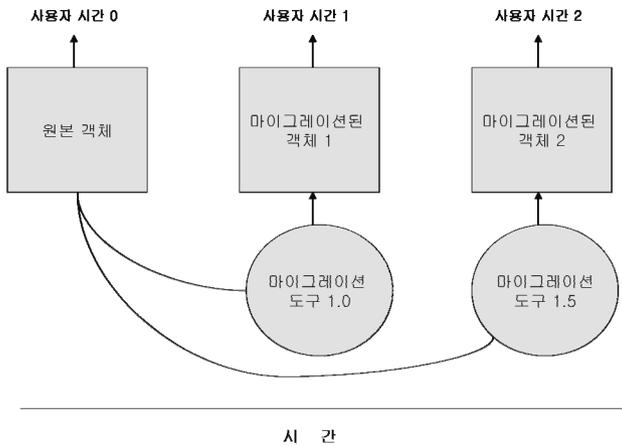
33) 국가기록원, 「전자기록물 장기보존 전략 연구 및 테스트베드 구축」, 국가기록원, 2008, p.21.

34) <<http://www2.si.umich.edu/CAMiLEON/reports/mor/index.html>>.

원래의 비트스트림을 유지함으로써 보다 정확하고 비용 효율적이 되도록 한 것이다. 전통적인 마이그레이션은 변환과정에서 파일에 원본과 차이점이 발생할 경우 향후 지속적인 변환과정에서 이 오류가 누적된다. 이것은 원본의 정확성에 큰 문제점이 발생함과 아울러 마이그레이션이 수행될 때마다 새로운 파일 포맷이 생성되게 됨을 의미한다.³⁵⁾

Migration on Request는 이러한 단점을 극복하기 위해 아래의 그림과 같이 하나의 원본파일을 보존하며, 새로운 환경이 나타날 때마다 여기에 적합한 마이그레이션 도구를 개발하는 접근방식을 취하게 된다. 따라서 마이그레이션 과정에서 발생한 오류가 누적되는 현상을 피할 수 있으며, 새로운 환경이 등장할 때마다 새로운 도구를 개발할 필요 없이 하나의 포맷을 대상으로 하는 변환도구를 개발할 수 있게 된다.

〈그림 6〉 Migration on Request 개념도³⁶⁾



Migration on Request는 소프트웨어 수명의 장기화(software longevity)

35) <<http://www2.si.umich.edu/CAMiLEON/reports/mor/conventional.html>>.

36) <<http://www2.si.umich.edu/CAMiLEON/reports/mor/onrequest.html>>.

기술을 이용하여 모든 디지털 객체들을 재현하며 오랜 시간동안 유지될 수 있다. 그리고 디지털 객체는 단순히 디지털 객체의 원래 포맷으로 보관되어 항상 원래의 디지털 포맷으로 되돌아올 수 있으므로, 변환 과정동안 초래되는 오류가능성을 크게 감소시켜 준다. 또한 원래의 포맷에서 현재의 포맷으로 변환하기 위해 항상 단 한 번의 변환과정만 거치게 되며, 어떠한 경우에도 아카이브에 보유 중인 디지털 객체가 원본이 된다.³⁷⁾ 이러한 점에서 Migration on Request은 장기보존 전략으로서 다음과 같은 장점을 지니게 된다.³⁸⁾

- 특정 파일포맷을 읽고 해석한 코드는 단지 한번만 실행됨
- 하나의 마이그레이션 단계만을 이용함으로써, 마이그레이션의 정확성 제고
- 디지털 객체가 원래 형태로 보존되기 때문에 진본성 문제들이 단순해짐
- 마이그레이션 도구가 오직 마이그레이션 요구가 있을 경우에만 실행되기 때문에, 대용량 디지털 객체가 보존될 때 대용량의 저장기능 제공

2) KB의 에플레이션 방식 분석

(1) KB 프로젝트의 개요 및 목적

네덜란드 국립도서관(Koninklike Bibliotheek, 이하 KB)과 네덜란드 국립 아카이브는 공동으로, 에플레이션 기반의 새로운 보존전략에 관한

37) 오삼균 외, 「전자기록물 장기보존을 위한 마이그레이션 전략에 관한 연구」, 『한국기록관리학회지』 제9권 제2호, 2009, p.35-36.

38) Paul Wheatley, "Migration: A CAMiLEON Discussion Paper", *Ariadne* Vol. 29, 2001. <<http://www.ariadne.ac.uk/issue29/CAMiLEON>>

연구를 2005년부터 2007년까지 수행하였다. KB는 미래의 플랫폼에서 디지털 원본을 렌더링 할 수 있는 장기보존 전략으로서의 에뮬레이션 가능성을 측정함과 아울러, 단일 에뮬레이션 접근으로 다양한 디지털 객체를 보존할 수 있는지 평가하는 것을 핵심 목표로 삼았다. 이 프로젝트 결과 누적 에뮬레이션과 마이그레이션된 에뮬레이션, 에뮬레이션 버추얼 머신, UVC 기반 에뮬레이션 방식이 제시되었고, 개념모형인 모듈화 에뮬레이션을 제안해 2007년 모듈화 에뮬레이터인 Dioscuri가 개발되어 Planets 프로젝트의 일부분으로 흡수되었다.

(2) KB의 모듈화 에뮬레이션(Modular Emulation)

현용 에뮬레이터의 보존맥락에서 전체 에뮬레이션(하드웨어 에뮬레이션)의 실현을 위해 모듈화 에뮬레이션 방식을 KB에서는 개발하였다. 모듈화 에뮬레이션은 디지털 객체를 미래 환경에서도 구현가능 할 수 있는 가능성을 고려한 것으로, 에뮬레이터와 에뮬레이션의 전 과정을 형성하는 하드웨어 아키텍처의 구성요소를 모듈로 구성하여 에뮬레이션 한 전체 에뮬레이션이다. 아울러 하드웨어 환경의 구성요소를 개별적인 에뮬레이터로 구현하고 이들의 조합으로 하드웨어 환경을 구성하여 에뮬레이션의 기반을 마련한 것이다.³⁹⁾ 이 모델은 아래의 그림에 제시된 대로 에뮬레이션 모듈식 접근 방식을 보여주는 관계상 모듈화 에뮬레이션이라 칭해지는데, 개별 에뮬레이터와 에뮬레이션 전 과정을 형성하는 하드웨어 아키텍처의 구성요소를 에뮬레이션한 전체 에뮬레이션으로 정의내릴 수 있다.⁴⁰⁾

모듈화 에뮬레이션의 구성요소는 다음과 같다.⁴¹⁾ 우선 범용 가상머

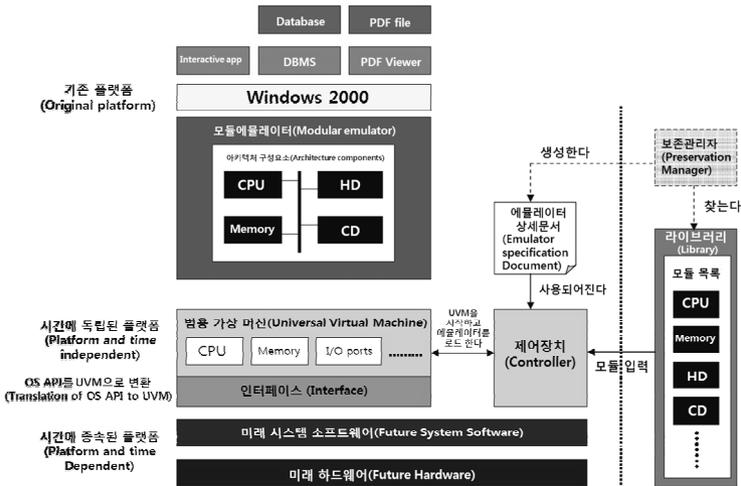
39) 이소연 외, 「전자기록의 관리와 보존을 위한 국제협력 아젠다 개발」, 국가기록원, 2007, p.88.

40) Jeffrey van der Hoeven · Hilde van Wijngaarden · Remoco Verdegem · Jacqueline Slats, *Emulation: A Viable Preservation Strategy*, 2005. 6, 20, p.21.

41) 이에 대해서는 Jeffrey van der Hoeven · Hilde van Wijngaarden · Remoco Verdegem ·

신(Universal Virtual Machine : UVM)은 모든 에뮬레이션의 기초로, UVM은 미래 호스트 플랫폼 상단의 독립적 계층으로 형성하며 UVM 프로세서와 메모리의 지원뿐만 아니라 UVC의 입력 호스트와 대상 플랫폼 사이의 주변장치와의 입출력(I/O)을 위한 추가 기능을 제공해준다. 다음은 모듈화 에뮬레이터로 UVM에 원본 소프트웨어를 실행하려면 에뮬레이터가 필요한데, 에뮬레이터는 모듈화 에뮬레이터에서 모듈식 구조로 설계되어야 한다. 원래 하드웨어 플랫폼의 구성요소는 이 구조의 모듈로 표시되며 각 모듈은 특정 하드웨어 구성 요소를 에뮬레이션 하는데, 대상 플랫폼의 재현은 다음 하드웨어 아키텍처에 따라 모든 필요한 모듈을 결합하여 실현될 수 있다.

〈그림 7〉 모듈화 에뮬레이션 개념도(42)



Jacqueline Slats, *Emulation: A Viable Preservation Strategy*, The National Library of the Netherlands & Nationaal Archief of the Netherlands, 2005. 6. 20(Version 1.4), pp.22-25의 내용을 정리.

42) Ibid., p.22.

모듈화 에뮬레이터를 구축하기 위해서는 각기 다른 모듈이 필요로 되는데, 구성요소 라이브러리(Component Library)는 모든 에뮬레이터에서 사용할 수 있는 모듈을 포함하며 가능한 모듈을 추적하고 구성요소의 유형을 조직화하게 된다. 라이브러리는 적어도 하나 이상의 대상 플랫폼을 만들 모듈의 최소 집합을 제공해야 하지만, 시간이 지남에 따라 더 많은 모듈을 확장 할 수 있다. 제어장치는 에뮬레이터 규격문서(Emulator specification document)를 기반으로 모듈을 구성하는 하드웨어 플랫폼의 규격을 해석해 사용 가능하거나 호환 가능한 모듈을 결정함과 아울러, 라이브러리에서 선택한 모듈을 검색하고 UVM에서 모듈화 에뮬레이터를 구축하기 위해 연결하는 기능을 제공함으로써 UVM에서 실제 하드웨어의 구성요소 기능을 재구성하고 원본 플랫폼을 구현한다.

모듈화 에뮬레이션은 하드웨어와 운영체제 그리고 응용 소프트웨어 등 다양한 프로세서로 구성되어 있는 컴퓨팅 환경을 모듈 구조의 에뮬레이터로 설계함으로써 특정 컴퓨팅 환경을 재현할 수 있다. 이는 컴퓨팅 환경의 기존의 하드웨어 플랫폼뿐만 아니라 시스템 소프트웨어 즉, 하드웨어를 해당 컴퓨팅 환경에서 구동할 수 있도록 재현하기 위해 하드웨어 드라이버나 추가적인 기능을 제공하는 플러그인의 에뮬레이션을 포함함으로써 하드웨어에 종속되지 않고 복잡한 하드웨어 구성이 필요한 플랫폼을 재현할 수 있으며, 플랫폼의 변형이 필요할 때는 에뮬레이션 되는 모듈의 구조를 변경하는 방식으로 하드웨어에 종속적이지 않으며 유연한 에뮬레이션의 실현이 가능해진다고 할 수 있다.

3) Planets 프로젝트의 에뮬레이션 방식 분석

(1) Planets 프로젝트의 개요 및 목적

Planets(Preservation and Long-term Access through NETworked Services)은 2006년 유럽연합의 여섯 번째 프레임워크 프로그램(FP 6-ICT)의 지원

을 받아 시작된 디지털 보존 프로젝트이다. 영국 국립도서관(The British Library)을 주관으로 16개 기관⁴³⁾이 참여했으며 1,300만 유로에 달하는 예산이 소요된 대규모 프로젝트로 2010년에 종료되었다.⁴⁴⁾

Planets 프로젝트에 참여했던 국립도서관과 아카이브들은 디지털 보존 책무를 수행함에 있어 근본적이고 장기적 전망에서 기관의 보존계획이 수립되어야 하며 이를 위한 이론적, 실제적 측면에서의 연구가 필요하다는 것에 공감하였다. 이에 Planets의 목적은 디지털 형태의 문화 및 과학 자산들에 대한 장기적인 보존과 접근을 돕기 위한 실제적인 서비스와 도구를 만드는 것에 두었고, 프로젝트를 통해 이러한 디지털 객체를 관리하기 위한 일련의 도구들을 개발하고 시험하는데 사용할 수 있는 프레임워크와 인프라를 제공하고자 하였다.

(2) Planets 프로젝트의 에몰레이션 방식

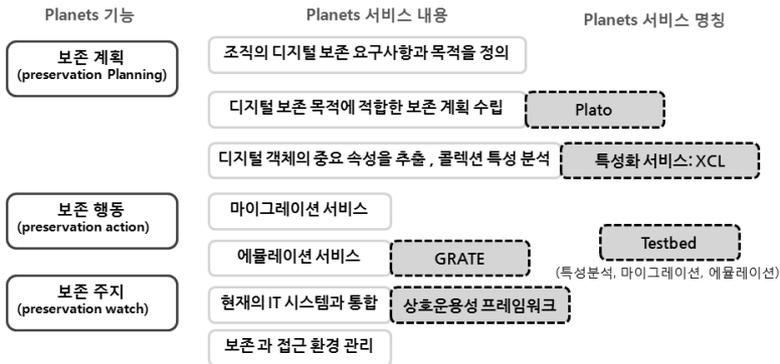
Planets은 디지털 보존을 위한 통합 서비스(integrated services)를 제공하고자 보존계획(Preservation Planning), 보존 특성화(Preservation Characterization), 보존 행동(Preservation Action), 상호운용 프레임워크(Interoperability Framework), 테스트베드(Testbed), 보급(Dissemination & Takeup)의 여섯 가지 서비스를

43) 프로젝트 기간은 2006년 6월 1일~2010년 5월 31일이며, 참여한 기관은 유럽의 국립도서관과 아카이브, 대학 등 연구기관, 주요 IT 업체로서 The British Library, The National Library of the Netherlands, The Austrian National Library, The Royal Library of Denmark, The State & University Library, Denmark, The Dutch National Archives, The UK National Archives, The Swiss Federal Archives (BAR), University of Cologne, University of Freiburg, Humanities Advanced Technology and Information Institute (HATII), Vienna University of Technology, The Austrian Institute of Technology (AIT), IBM Netherlands, Microsoft Research Limited, Tessella plc이다. <<http://www.planets-project.eu/>>.

44) Planets의 결과물들은 프로젝트 종료 이후 영국에 거점을 둔 비영리기관인 “Open Planets Foundation(OPF)”를 통해 유지, 운영되고 있다. <<http://www.openPlanetsfoundation.org>>

제공하고 있다(그림 9) 참조). 이는 OAIS 참조모델의 보존계획(Preservation Planning) 기능을 실제로 구현한 사례로 볼 수 있으며⁴⁵⁾ 조직의 디지털 보존계획을 수립하게 하고 보존 대상이 되는 디지털 객체의 특성화를 위한 서비스를 지원하며 그러한 분석결과에 근거해서 구체적인 보존행동을 계획할 수 있도록 체계화 되어 있다. 또한 이러한 서비스를 개별 사례에 맞게 시험, 평가할 수 있도록 웹 기반의 테스트 베드를 제공해줌으로써 디지털 보존을 위한 인프라를 제공하고자 하였다.

〈그림 8〉 Planets 기능과 서비스



Planets을 통합된 서비스라는 차원에서 이해하였을 때, 그 중에서 에뮬레이션 서비스는 보존 계획 및 보존 특성화 활동을 토대로 보존 대상에 적합한 구체적인 보존 행동으로 선택되는 것이다. Planets은 에뮬레이션 서비스의 방향을 결정하기 위해서 보존방식으로서 에뮬레이션과 가상화에 관한 사례연구를 진행하였다.⁴⁶⁾ 이 연구를 거치면서 Planets은 에뮬레이션과 가상화의 유의미성을 재차 확인하였으며 소프트웨어 아

45) Planets이 수행하는 기능들은 'Planets 기능모델'로서 정리되었다. PP7-D3-4. *Report on The Planets Functional Model*, 2009. 8, p.4.

46) PA/5, *Report Describing Results Case Studies*, 2008.

카이브, 원격 에뮬레이션 등의 이슈를 제기하기도 하였다. 이것은 에뮬레이션 전략을 구사하기 위해서 실제적으로 필요한 요소들을 고민하는 계기가 되었다.

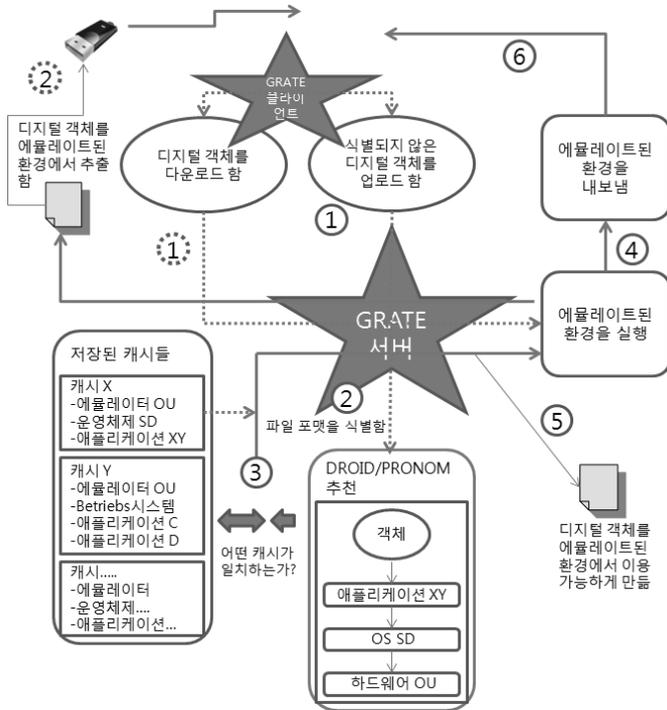
결과적으로 Planets은 GRATE(Global Remote Access to Emulation Services)라는 통합 에뮬레이션 프레임워크를 제안하였다. 이것은 웹 기반의 자동화된 에뮬레이션 서비스로서 <그림 10>의 GRATE 모델은 에뮬레이션 프레임워크가 어떻게 구성되고 있는지를 가시적으로 보여주고 있다. 이 모델의 핵심적인 요소는 에뮬레이터 및 소프트웨어 아카이브 그리고 파일 포맷 레지스트리이다.

GRATE 모델에 따르면 먼저, 이용자를 통해 구형의 디지털 객체가 GRATE 모델 안에 들어오게 되고 DROID/PRONOM⁴⁷⁾을 통해서 해당 디지털 객체의 정보가 식별된다. 다음으로 이 객체를 렌더링하기 위해서 어떤 에뮬레이터를 쓸 것인가가 결정된다. 즉 식별된 구형의 디지털 객체를 보여주는데 필요한 어플리케이션 및 그것을 작동할 수 있게 해주는 에뮬레이터가 선택된다는 것이다. 이를 통해 구형의 디지털 객체는 이용자에게 제공될 수 있다. 이를 위해서 GRATE는 다중의 에뮬레이터들을 포함하고 있는데 Dioscuri, QEMU, UVC가 그것이다. Planets에서 사용되는 에뮬레이터들은 프로젝트를 통해서 직접 개발된 것이 아니지만 그중 Dioscuri는 KB 에뮬레이션 프로젝트를 통해 개발되어 Planets 및 KEEP 프로젝트를 통해 더 다양한 환경을 포괄할 수 있도록 후속개발이 이루어졌다.⁴⁸⁾

47) PRONOM은 TNA가 디지털 자원의 장기보존을 지원하고자 파일 포맷과 해당 파일을 지원하는 소프트웨어에 대한 정보를 정의하고 저장한 파일 포맷 저장소이다. PRONOM은 DROID(Digital Record Object Identification)과 함께 사용되었다. 이소연 외, 「전자기록의 관리와 보존을 위한 국제협력 아젠다 개발」, 국가기록원, 2007. p.95.

48) KB 에뮬레이션 프로젝트에서는 모듈 방식과 내구성에 착안하여 “모듈화 에뮬레이션”이란 개념을 만들어 냈고 이를 바탕으로 만든 에뮬레이터가 바로 Dioscuri이다. Dioscuri는 2006-2008년 Tessella에 의해 개발이 주도되었고 2008년 이후로

〈그림 9〉 GRATE 모델⁴⁹⁾



이러한 Planets 활동의 핵심에는 “종합적인 보존 계획(preservation planning)”이 있다. 보존 계획이란 조직의 요구사항과 소장하고 있는 컬렉션의 특성에 맞게 적절한 보존 전략을 선별하는 과정으로 실제적으로 보존 대상의 특성이나 상황에 맞게 현존하는 보존 도구(마이그레이션 혹은 에뮬레이션 도구 등)를 포괄적으로 검토하여 필요한 보존 전략을 수립, 운영해 나가는 것이 중요하다. 하지만 포괄적인 접근을 어렵게 하는 이유들이

는 Planets, KEEP 프로젝트 내에서 네덜란드 국립도서관이 개발에 주도적인 역할을 수행하였다. 최신버전은 0.7.0.(2011. 1. 19).

49) Planets, *Frist Version of GRATE*(PA/5-D7 v.1.0), 2006, p.17.

존재하며 특히, Planets은 현존하는 보존 도구들이 독립적인 어플리케이션(stand-alone application)으로 존재함으로써 이를 어렵게 하고 있다고 보았다. 이에 Planets은 단일전략 및 단일 어플리케이션 형태의 보존 도구를 넘어서 복합적인 성격을 띤 보존 대상에 포괄적으로 대응할 수 있는 일련의 보존계획 서비스 체계가 필요하다는 판단을 한 것으로 보인다.⁵⁰⁾

결과적으로 이는 에뮬레이션 전략 자체를 대상으로 한 사례는 아니지만 보다 근본적으로 보존 계획 차원에서 에뮬레이션 전략을 구상하기 위한 의미 있는 시사점을 주고 있다. 통합된 보존계획 서비스의 시작은 우선적으로 보존 대상에 적합한 보존계획의 수립에 있으며 에뮬레이션 전략은 일련의 보존 계획 서비스 체계 속에서 일부가 될 것이다. 아울러 Planet에서 에뮬레이션 전략을 구성하기 위해 핵심적으로 필요한 에뮬레이터, 파일 포맷 레지스트리, 소프트웨어 아카이브 등의 요소들과 이들을 실제로 운영하는데 필요한 환경적 요인들에 대한 보다 심도 깊은 검토가 요구된다.

4) KEEP 프로젝트의 에뮬레이션 방식 분석

(1) KEEP 프로젝트의 개요 및 목적

KEEP(Keeping Emulation Environments Portable)은 2009년 유럽 연합의 일곱 번째 프레임워크 프로그램(FP 7-ICT)의 지원을 받아 시작된 프로젝트로, 프랑스 국립도서관(BnF)이 주관하고 에뮬레이션에 관한 다양한 유럽의 이해관계자를 포함하는 9개 기관이 컨소시엄에 참여하였으며 2012년에 종료되었다.⁵¹⁾

50) Adam Farquhar · Helen Hockx-Yu, "Planets: Integrated Services for Digital Preservation", *The International Journal of Digital Curation* Vol. 2, No. 2, 2007, p.89.

51) 이 프로젝트 기간은 2009년 2월 1일~2012년 2월 29일이며, 참여한 기관은 문화유산기관, 연구기관, 상업적 파트너와 게임회사들로서 Bibliothèque Nationale de

KEEP은 디지털 객체에 대한 보존, 접근, 재현의 과제가 존재한다고 하였을 때 에뮬레이션 대상에 대한 이러한 문제를 해결하기 위해서는 디지털 보존 도구와 환경의 관리, 보존 과정에 대한 셋업 자동화, 노후화된 저장용기로부터 데이터의 검색이라는 디지털 보존의 이슈들이 해결되어야 한다고 보았다. 그리고 다양한 선행 연구들이 존재하고 있지만 실제적인 해결책이 제한적이라고 평가하였다. 특히 웹사이트, 게임, 소프트웨어 등의 복잡한 디지털 컬렉션들의 경우는 에뮬레이션 외에 다른 대책은 현실적으로 의미가 없다고 보았다. 이러한 상황에서 KEEP은 정적(텍스트, 사운드, 이미지, 비디오 파일 등), 기능성(멀티미디어 문서, 소프트웨어, 웹사이트, 데이터베이스, 비디오게임 등) 특성을 포괄하는 디지털 객체를 정확하게 렌더링 할 수 있는 에뮬레이션 접근 플랫폼(emulation access platform)을 목표로 에뮬레이션에 기반 한 영구적인 접근 전략을 세우고 그것을 지원해줄 수 있는 도구를 개발하고자 하였다.

(2) KEEP 프로젝트의 에뮬레이션 방식

KEEP은 두 가지의 핵심 프레임워크를 제안했는데 “KEEP 전송 도구 프레임워크(KEEP Transfer Tool Framework 이하, TTF)”와 “KEEP 에뮬레이션 프레임워크(KEEP Emulation Framework 이하, EF)”로서 이를 위해 프로젝트에서는 다음의 그림과 같은 7개의 작업들(WP: Work Package)을 진행하였다.

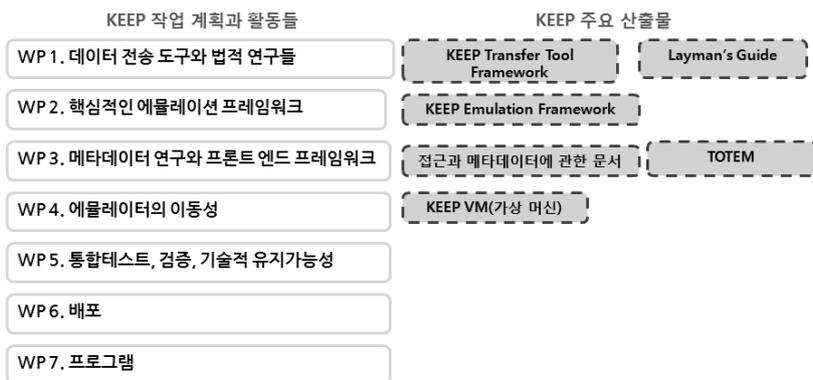
KEEP TTF는 점차 노후화되고 있는 저장매체로부터 보존 대상이 되는 디지털 객체 원래의 비트스트림을 안전하게 포착하기 위한 도구이다. 한편 KEEP EF는 에뮬레이션을 위한 자동화된 작업 체계를 제공해 주는 것이다. KEEP를 이러한 도구들을 프레임워크라고 표현하였는데 이것은

France, Joguín SAS, Koninklijke Bibliotheek, Computerspiele Museum, University of Portsmouth, Deutsche Nationalbibliothek, Cross Czech a.s, Tessella, European Games Developer Federation이다. <<http://www.keep-project.eu/>>.

해당 기능을 제공해주는 여러 개의 다중의 전송도구, 에뮬레이터를 포함하고 통합적으로 활용할 수 있게 하는 서비스로서 이해할 수 있다.

이 외에도 데이터 전송과 관련하여 발생할 수 있는 법적 측면을 분석한 보고서와 에뮬레이션에 필요한 메타데이터의 측면을 밝히기 위한 연구를 함께 진행하였으며 TOTEM(Trustworthy Online Technical Metadata Database)이라는 기술 메타데이터 데이터베이스를 시범적으로 웹상에 구축하여 운영하기도 하였다.⁵²⁾

〈그림 10〉 KEEP의 주요 활동과 산출물

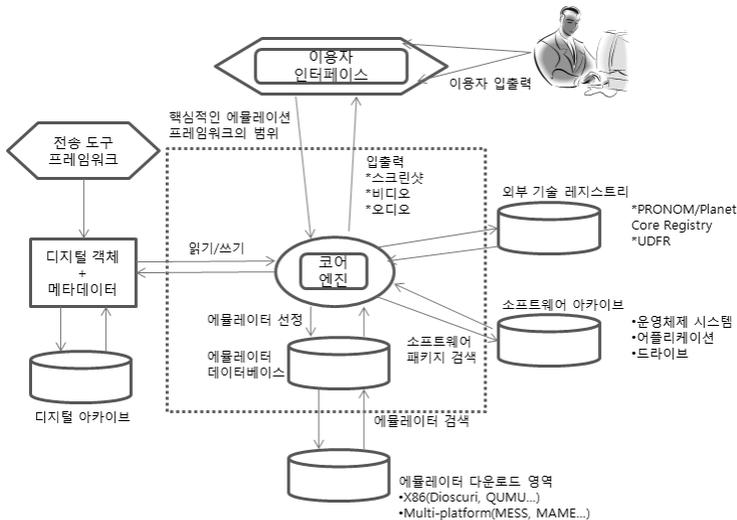


에뮬레이션 서비스의 관점에서 본다면 KEEP EF는 에뮬레이션을 위한 대상을 안전하게 캡처하기 위한 도구를 포함하면서 대상에 적합한 에뮬레이션 환경을 구성하여 서비스하는 것을 주요 기능으로 하고 있다. 다음의 그림은 KEEP EF의 구성을 보여주는데, 프레임워크의 핵심적인 요소는 에뮬레이션 코어 엔진, 소프트웨어 아카이브, 에뮬레이터 아카이브, 파일 포맷 레지스트리이다.

52) 현재는 PC x386 architecture, Commodore 64 Architecture, Cosole gaming platforms 에 관한 메타데이터를 테스트하는 것이 가능하다. <<http://keep-totem.co.uk>>를 참조.

즉 이용자의 요구로 이미 사양화된 컴퓨터 환경을 구현할 필요가 발생하게 되면 코어 엔진이 보존대상이 되는 파일 포맷에 대한 특성화 작업을 진행하는데, 이때 DROID/PRONOM을 포함하는 외부의 파일 포맷 레지스트리가 해당 디지털 객체의 정보를 식별하는데 활용된다. 이러한 결과를 토대로 에뮬레이터 아카이브 및 소프트웨어 아카이브로부터 자동적으로 해당 파일포맷에 가장 적합한 에뮬레이터를 가동 시켜 구현이 이루어지게 된다. 이를 위해서 EF는 7개 에뮬레이터와 6개 플랫폼을 포함하고 있다.⁵³⁾

〈그림 11〉 KEEP 에뮬레이션 프레임워크⁵⁴⁾



53) Qemu(x86), Dioscuri(x86), WinUAE(Amiga), Vice(c64), JavaCPC(Amstrad/Schneider), BeebEm(BBCmicro), Thomson(Thomson T07)). 또한 25개 파일들(PDF, TXT, XML, TIFF, PNG, BMP, Quark, ARJ, EXE, disk/tape images 등).

54) KEEP, *Requirements and Design Documents for Services and Architecture of Emulation Framework: Part B: Architectural Design for the Core Emulation Framework(WP2-D2.2b)*, 2009, p.43.

Planets 프로젝트가 상대적으로 보존계획이라는 근본적인 문제의식에 입각한 종합 대책의 일부 요소로서 에물레이션을 다루었다면, KEEP은 특정 유형의 디지털 객체에 있어서 에물레이션과 가상화라는 장기보존 전략의 유의미성을 인정한 가운데 본격적인 에물레이션 적용에 초점을 둔 사례이다. 특히 에물레이션을 성공적으로 진행하기 위해 적용하고 있는 도구와 환경, 자동화, 소프트웨어 아카이브, 법적 관점, 메타데이터 등의 이슈들에 대한 판단과 적용사항, 결과물들에 대해서는 보다 심도 깊은 고찰이 이어질 필요가 있다.

4. 시사점과 전망

1) 사례 분석의 시사점

3장에서 고찰한 에물레이션에 관한 연구 프로젝트들은 향후 우리나라 공공기록 및 민간기록에 대한 에물레이션 방안 도출시 벤치마킹할 수 있는 유용한 사례들을 제시해준다. 우선 CAMILEON 프로젝트는 디지털 객체의 장기적 보존을 위한 기술적 전략들을 광범위하게 개발하고 평가하는 과정을 통해 상황에 따라 어떠한 장기보존 전략을 적용할 수 있을지에 대해 연구한 프로젝트로, 디지털 객체의 원래의 기능성(functionality)과 ‘외형과 느낌(look and feel)’의 장기적 유지를 위한 각종 사항들을 모색해준다. CAMILEON 프로젝트에서는 디지털 객체의 장기보존 전략을 하나의 단일 전략만을 적용하는 것이 아닌, 디지털 기록의 유형 및 아카이브의 상황에 맞추어 적용할 수 있는 전략을 소프트웨어 수명의 장기화 전략, 에물레이션, Migration on Request란 세 가지로 구분하여 고찰하고 있는데, 특히 CAMILEON 프로젝트에서 제시

한 Migration on Request 전략은 이용의 시점에 객체를 마이그레이션하는 도구와 더불어 원래의 바이트스트림을 유지함으로써 보존전략이 보다 정확하고 비용 효율적이 되도록 하고 있다. 이와 더불어 에물레이션 전략의 경우 소프트웨어 및 하드웨어, 시스템에 대한 지적 재산권 문제의 해결이 난해한 문제임을 제시해주고 있다.

네덜란드의 KB 프로젝트는 디지털 객체의 장기보존 전략으로서 에물레이션의 다양한 가능성에 대한 연구를 통해, 누적 에물레이션과 마이그레이션된 에물레이션, 에물레이션 버추얼 머신, UVC 기반 에물레이션 등의 방법론을 제시하였고 개념모형인 모듈 에물레이션을 제안하고 있다. 누적 에물레이션의 경우 에물레이터가 플랫폼 환경의 일정 부분만을 상세화하기 때문에 다른 플랫폼과의 호환성이 저하되며 연속성이 적고, 마이그레이션된 에물레이션은 소스코드가 미래에 생성되기 때문에 환경적 문제를 야기할 수 있으며, 컴파일러의 라이브러리는 원본과 양립하지 않을 수 있기 때문에 기능 고장이 있을 수 있는 한계를 지닌다. 그리고 EVM 방식은 구 플랫폼이 단일의 규정 내에서 전체적으로 정의될 수 있는지, 게다가 복합적인 해석기는 이러한 규정을 에물레이터로 변환하는데 필요로 되는지의 문제가 아직 해결되지 못하였고, UVC 전략의 궁극적인 성공은 부분적으로는 소프트웨어 및 컴퓨터 산업 영역에 의해 채택되는 정도에 의존한다는 점을 제시해준다. KB 프로젝트에서는 에물레이션상의 이와 같은 다양한 문제점을 해결하기 위한 방안으로, 개별 에물레이터와 에물레이션 전 과정 등을 형성하는 등의 하드웨어 아키텍처의 구성요소를 에물레이트한 전체 에물레이션인 모듈화 에물레이션을 2007년 Dioscuri란 명칭으로 개발하였는데, 이 에물레이터는 Planets 프로젝트에 의해 계속 연구를 수행 중에 있다.

한편 네덜란드 국립아카이브는 Planets 프로젝트의 참여기관으로서 마이그레이션이 적절하지 않은 스프레드시트와 데이터베이스와 같은 동적인 레코드들의 보존에 있어서 에물레이션의 적정성을 살펴보았다.

그 결과 애플리케이션의 두 가지 접근법을 발견하였다. 첫 번째는 독자를 위한 애플리케이션인 렌더링이다. 이것은 데이터베이스와 같은 유형은 예외 사례 또한 존재하겠지만 애플리케이션 전략이 디지털 객체를 적절하게 구현하기 위한 최선의 방법으로 결론 내렸다. 두 번째는 마이그레이션 과정 중에 행하는 중간단계로서의 애플리케이션이다. 마이그레이션은 텍스트 등에 있어서 가장 적합한 전략이지만 마이그레이션 진행 중에 애플리케이션이 중요한 단계가 될 수 있음을 관찰하였다. 만약 디지털 객체가 그것이 생산된 이후 오랫동안 관리되지 않은 채로 넘겨져 왔을 경우, 예를 들어 개인 컬렉션과 같이 그 안에 오래된 포맷의 디지털 자료들을 포함하는 경우에 애플리케이션이 파일 컨버전 이전에 먼저 시행될 필요가 있다. 또한 애플리케이션은 디지털 저장소와 같은 시스템에 입수시키기 전에 사전적 단계로서 원래의 맥락을 재창조하기 위한 수단으로서 사용될 수 있다고 보았다.

아울러 이 사례에서는 애플리케이션이란 열린 선택을 유지하는 것을 의미한다고 결론짓고 있다. 대부분의 마이그레이션은 레코드의 특징들을 보존하며 그것은 마이그레이션 시점에 지정된 그 목적의 중요성에 따라 좌우되며 그 외의 측면들은 유실되거나 변경된다는 문제를 안고 있다. 그런데 원본을 유지할 수 있다면 미래의 보존과 접근 포인트를 극대화시킬 수 있다는 장점을 가질 수 있다는 것이다. 그리고 고고학 분야의 발전상이 발굴에서 현장 보존으로 변화하고 있다는 점에 주목하면서 애플리케이션 전략의 유의미성을 역설하고 있다. 원래의 파일을 유지하고 애플리케이션을 사용하는 것은 원래의 환경 속에 있는 객체에 대한 접근을 제공하고 그것에 대한 보존과 접근을 유지하는 열린 선택을 가능하게 하는 것이며, “디지털 고고학”에 대한 동시대적인 접근이라는 것이다.

Planets 프로젝트에서 제시한 애플리케이션에 대한 연구는 먼저 디지털 객체에 관한 장기보존 서비스, 도구를 전체적으로 제시해 줌으로써 보

존 전략 중의 하나로서 에플레이션 활용에 대한 포괄적인 시각을 제공해준다는 점에서 의미를 찾을 수 있다. 보존대상을 어떻게 정하고 대상에 적절한 보존 전략을 세우기 위한 분석을 실시하고 구체적인 보존 행위들을 진행할 수 있게 관련된 도구들로 구성된 테스트베드를 제공해주고 있다. 이와 더불어 보존 전략 중의 하나인 에플레이션이 어떻게 실제 보존 전략 체계 속에서 결합되어 구사되는 것인지를 살필 수 있게 한다는 점도 의미를 지닌다. 실제로 장기보존 전략은 단일로 구사될 수 없고 해당 환경, 대상 객체에 맞게 혼합형으로 해야 할 것이기 때문에 보존 대상이 될 디지털 객체의 특성을 토대로 다양한 보존 전략들이 결합되어야 할 것이다.

KEEP 프로젝트가 제공해주는 가장 큰 성과는 에플레이션에 기반을 둔 보존 전략을 세우고 그것을 지원해주는 실제적인 도구를 개발하였다는 점이다. 이는 Planets이 장기보존의 전체적인 계획 차원에서 진행된 프로젝트였다면, KEEP 프로젝트는 에플레이션 자체를 핵심전략으로 구사하는 것과 관련된 사례라는 점에 특징이 있다. KEEP에서는 에플레이션에 기반한 보존전략과 그것을 지원해주는 도구를 개발하였는데, 앞서 언급한 EF와 TTF가 그것이다. KEEP의 TTF는 구형의 데이터 매체로부터 디지털 콘텐츠를 안전하게 캡처할 수 있는 방법을 구현한 것이 되겠으며, Planets과 KEEP에서는 EF를 통하여 현재 사용하고 있는 디지털 아카이빙 시스템과의 통합을 지향하고 있기도 하다는 점에서 주목해볼 수 있다.

2) 우리나라의 적용을 위한 전망

전자기록의 장기보존 전략으로서 에플레이션은 비트스트림의 변화 없이 전자기록의 원형을 그대로 재현할 수 있다는 점에서, 아울러 마이그레이션처럼 전자기록 유형별로 개별적인 전략을 적용할 필요 없이

일괄적인 보존전략을 구현할 수 있다는 점에서 상당한 강점을 지닌 보존전략임은 분명하다. 특히 현재 우리나라 전자기록의 장기보존 포맷에서는 구현해 주지 못하는 전자기록의 기능적 구성요소를 재현해 줄 수 있다는 점에서, 현재 공공기관에 보유 중인 전자기록 유형 조사를 기반으로 그 실현 방안을 연구할 필요가 있다.

에물레이션 전략에 관한 이상의 선진사례의 분석결과를 기반으로 볼 때 CAMILEON 프로젝트에서 제시하는 것처럼 디지털 객체의 장기보존 전략을 하나의 단일 전략만을 적용하는 것이 아닌, 디지털 기록의 유형 및 아카이브의 상황에 맞추어 적용할 수 있는 전략을 고려할 필요가 있다. 물론 에물레이션 전략을 주축의 장기보존 전략으로 채택하더라도 전자기록의 4대 속성을 체계적으로 구현할 수 있는 보존전략을 부수적인 보존전략으로 도입할 필요가 있다. 네덜란드의 KB 프로젝트에서 수행된 텍스트, 데이터베이스, 스프레드시트의 세 유형에 대한 에물레이션, 마이그레이션, XML 보존전략에 대한 비교 분석은 다양한 유형의 전자기록 보존전략 선택 시 고려해야 할 다양한 요소들을 벤치마킹하는데 유용한 연구결과를 제공해 준다. 즉 텍스트 문서의 경우 PDF와 XML의 결합방식이 가장 효과적으로 텍스트의 장기보존을 위한 전략이며, XML은 데이터베이스의 맥락정보나 내용, 구조를 아주 잘 표현해줄 수 있으므로 현재 상태에서 데이터베이스 장기보존을 위해서는 XML을 가장 효과적인 전략으로 판단하고 있다.⁵⁵⁾

한편 Planets 프로젝트에서 제시해주는 바대로 에물레이션 전략 및 방법론 개발 시, 전체 디지털 아카이빙 아키텍처 내에서 기타 프로세스

55) 이에 대해서는 Digital Preservation Testbed, *From Digital Volatility to Digital Permanence: Preserving Databases(version 1.0)*, Dutch National Archives · The Hague, December 2003; *From Digital Volatility to Digital Permanence: Preserving spreadsheets(version 1.0)*, Dutch National Archives · The Hague, December 2003; *From Digital Volatility to Digital Permanence: Preserving Text Documents(version 1.0)*, Dutch National Archives · The Hague, December 2003을 참조.

들과 잘 융합되도록 개발해야 한다. 이와 더불어 에플레이션의 직접적인 도구가 되는 에플레이터 소프트웨어의 생산은 고도의 기술력을 필요로 하고 많은 비용이 소요되며 사유(proprietary) 소프트웨어의 에플레이션은 지적 자산과 저작권 문제를 발생시켜 장기보존 전략상의 효율성과 안정성을 손상시킨다는 점은, 추후 우리나라에서 에플레이터 개발 시 충분히 고려해야 하며 현행 우리나라 소프트웨어의 저작권법과 관련하여 그 대책 마련이 필요할 것으로 보인다. 아울러 Planets 프로젝트에서 제시한 GRATE와 같은 범용적 서비스 제공을 염두에 둔다면, 에플레이션 코어 엔진, 소프트웨어 아카이브, 에플레이터 아카이브, 파일 포맷 레지스트리 등을 개발해야 할 것으로 보인다.

현재 우리나라 공공기록을 대상으로 에플레이션 보존전략을 도입할 때 역점을 두어야 하는 분야 가운데 하나는 공공기관에서 사용 중인 디지털 컴포넌트의 분석 및 에플레이션 전략의 적용이 필요한 기능적 요소의 파악이다. 모든 시스템 및 소프트웨어를 대상으로 에플레이션을 구현하는 방안은 아직까지 기술적으로 난해할 뿐만 아니라, 막대한 예산 소요가 수반되기 때문이다. 에플레이션 전략은 고도의 기술력 및 에플레이터 개발 비용이 소요되며, 아울러 소프트웨어의 급진전 및 다양화에 따라 향후 에플레이터의 개발이 수시로 이루어져야 함을 염두에 둘 때, 현행 PDF/A를 기반으로 한 마이그레이션 전략을 기반으로 하되 PDF/A에서 구현이 어려운 기능적 요소를 파악하여 에플레이션 전략을 적용하는 것이 최선의 방안이다. 이를 감안할 때 전자기록에 포함된 디지털 컴포넌트의 포맷을 분석하여 시간이 경과한 후에도 재현이 가능하도록 PDF/A 형식의 문서보존포맷으로 변환 시 소실되는 기능적 요소를 파악해, 이들 기능적 요소를 지닌 디지털 컴포넌트에 초점을 맞추어 에플레이션 전략을 적용하는 것이 활용 가능한 인적 물적 자원으로 최대한의 장기보존 효과를 가져올 수 있는 방안이라 할 수 있다.

우리나라 공공기관에서 사용하고 있는 대부분의 응용프로그램은 마

이크로소프트사의 윈도우 환경에서 구동되고 있으나 향후에는 지금과 전혀 다른 운영체제 및 컴퓨팅 환경이 지원될 가능성도 있으므로, 기존의 응용프로그램을 더 이상 지원하지 않는 경우가 발생할 수 있다. 따라서 현재 버전의 응용프로그램을 보존하더라도 새로운 시스템 환경에서 동작하지 않을 것으로 예측되고 있다. 이를 감안할 때 현 시점에서 가장 중요한 것은 장기 보존이 필요한 전자기록에 포함된 디지털 컴포넌트의 포맷 형식을 정확하게 파악하여 레지스트리를 구축하는 것이 필요하다. PDF, JPG, MPEG 등 포맷 정보가 표준화되어 공개된 디지털 컴포넌트는 다양한 응용프로그램에서 조희가 가능하므로 변화된 컴퓨팅 환경에서 조희가 가능하지만, HWP, PPT, XLS 등 특정 사기업에 의존한 포맷을 사용하는 디지털 컴포넌트는 기업에서 지속적으로 응용프로그램을 제공해야만 조희가 가능하다.

현행 우리나라에서는 공공기록이 생성·활용되는 플랫폼이 한정적이며 특정 포맷의 디지털 컴포넌트가 대다수를 차지하는 양상을 보이고 있다. 지난 2013년 7월부터 9월동안 A공공기관을 대상으로 실시된 기록관리시스템에 등록된 디지털 컴포넌트의 유형에 대한 조사에 따르면, HWP 파일 포맷이 84.33%를 차지하여 거의 대다수 파일이 아래아한글로 작성됨을 알 수 있고, 다음으로 마이크로소프트사의 XLS 파일 포맷이 9.08%를 차지하여 전체 디지털 컴포넌트 중 2가지 포맷이 차지하는 비중(93.41%)이 거의 대부분임을 파악하였다. 그리고 PPT 파일 포맷으로 작성된 디지털 컴포넌트의 비중은 0.37%로 다른 포맷 유형보다 적은 반면 PDF 파일 포맷은 2.76%를 유지하여 비교적 많이 사용됨이 파악되었다.⁵⁶⁾ 이를 감안할 때 우리나라 공공기관에서 가장 일반적으로 사용되는 디지털 컴포넌트의 기능성을 분석한 다음, 이러한 기능성을 지닌

56) 이에 대해서는 최주호·이재영·임진희 「장기 보존 대상 디지털 컴포넌트의 포맷 유형 분석 및 에뮬레이션 사례 연구」, 한국기록학회 제94회 월례발표회 발표문, 2013, pp.5-10을 참조.

전자기록을 대상으로 에뮬레이션 전략을 적용하는 방안을 모색해볼 필요가 있다. 기능성 중 어느 하나라도 지닌 경우 일괄적으로 에뮬레이션을 적용하는 경우도 상정할 수 있고, 또한 각 디지털 컴포넌트별 기능성의 중요도(내용적, 증거적 가치 등)를 계량화하여 일정 수치를 획득한 경우 에뮬레이션 대상으로 선정하는 안도 상정해 볼 수 있다. 물론 이러한 절차 및 방식에 대해서는 별도의 방대한 연구가 필요할 것으로 사료되며, 추후의 연구를 기대한다.

ABSTRACT

An Analysis of Cases of Emulation for Long Term Electronic Records Preservation Strategy

Kim, Myung-Hun · Oh, Myung-Jin
Lee, Jae-Hong · Yim, Jin-Hee

In response to the current electronic record environment, storing electronic records for necessary long terms have been the topic of the times. Especially, the method to preserving original form such as original functional property and bit stream becomes the key to long term preservation of electronic records. Until now in Korea migration methods were chosen as long term preservation strategy for public records, but the limitations were that the functionality and the original bit stream could not be maintained.

Among the strategies for long term preservation of electronic records, emulation has significant strengths in that it can replicate the original form of electronic records without changes in the bit stream, and that unlike migration it can establish a single preservation strategy without needing to apply individual strategies according to type of electronic record. Especially because it can replicate the functional components that cannot be implemented in the current long term preservation format, there is a need to study the application method based on the studies of electronic record types currently used by public institutions.

This study, to explore the methods for applying emulation as a strategy for the long term preservation of electronic records, reviews the latest study

cases from the west about emulation as base study, and tries to analyze the specific cases about the feasibility, target, and methods of emulation. Through this the study tries to explore the implications for domestic application as well as the strengths and weaknesses of emulation. To do this, the study analyses the concept, strengths and weaknesses of emulation as a long term preservation strategy, the analyses the latest best cases of emulation in the west; CAMiLEON, KB, Planets, and KEEP project. Based on these analysis this study tries to suggest implications and application methods for electronic records in the future in Korea.

Key words : CAMiLEON, KB, Planets, KEEP, long-term preservation strategy, emulation, emulator