

디지털 아카이빙 최근 연구동향 및 OAIS 참조모형에 관한 연구

Research Trends in Digital Archiving and the OAIS Reference Model

김 희정(Hee-Jung Kim)*

◁ 목 차 ▷

- | | |
|---------------------|------------------|
| 1 서론 | 4 OAIS 참조모형 및 적용 |
| 2 전자기록물과 장기보존정책 | 5 향후방향과 제언 |
| 3 디지털아카이빙 최근 연구프로젝트 | <참고문헌> |

<국문초록>

디지털 아카이빙과 관련하여, 특히 전자기록물의 영구보존을 중심으로 국내외에서 완료되었거나 현재 수행 중인 프로젝트를 중심으로 연구동향을 정리한다. 특히, 2002년 1월에 국제표준으로 확정된 OAIS 참조모형을 중심으로 모형 내 주요기능에 대하여 알아본다. OAIS 참조모형이 전자기록물 영구보존에 어떠한 역할을 하는지를 살펴보고, 국내의 향후 연구방향에 대하여서도 살펴본다.

<ABSTRACT>

This paper reviewed research trends in Digital Archiving and the OAIS Reference Model. Well renowned International Projects are also reviewed. InterPares, NEDLIB, CEDARS, CAMILEON, PRISM are included. Several key functions of the OAIS Reference Model are presented and further research direction of digital preservation in Korea is also suggested.

1 서론

디지털정보가 중심이 되는 정보사회에서는 전자기록물에 대한 비중과 의존도가 현격히 높아지는 만큼, 전자기록물에 대한 구체적인 관리방안의 필요성도 더욱 증가하고 있다. 전자기록물의 관리방안은 생산·정리·유통·활용·보존 및 재현(재생산)에 이르는 기록물의 전 생애주기와 관련하여 각 단계별로 이루어지게 되는데, 특히 최근

* 연세대학교 문헌정보학과 박사과정 / heejung@lis.yonsei.ac.kr

주요한 쟁점으로 대두되는 사안 중의 하나가 적절한 보존정책의 수립이다. 적절한 보존정책이 수립되었을 때에, 세대를 거쳐서 기록물의 지속적이고도 유용한 활용이 가능해지며, 이러한 활용을 통하여서 고부가가치의 정보와 지식이 재창출되는 지식의 순환이 이어질 수 있기 때문이다.

특히 전자기록물의 경우, 종이기록물과 달리 변형되거나 훼손되기 쉽고, 유실되기 쉬운 휘발적 성격을 갖고 있으므로, 적절한 보존정책의 수립이 더욱 더 필수적으로 요청된다. 이러한 요청에 따라, 미국과 유럽에서는 전자기록물의 보존과 관련된 연구가 1990년대 중반 이후 활발히 수행되어 오고 있다.

그런데, 이러한 전자기록물의 장기보존과 관련된 연구들을 살펴보면, 대부분 특정한 한 기관이나 분야에서 수행하기보다는, 정부·도서관·기록관·컴퓨터공학분야·문헌정보학분야 등을 위시한 공동연구체제가 필연적으로 구성되어 있음을 확인할 수 있다. 또한, 전자기록물의 장기보존이라는 공통적인 주제를 놓고도 다각적인 관점에서 접근하여 다양한 연구가 수행되고 있으며, 상당수의 프로젝트는 아직 뚜렷한 해답을 제시하지 않고, 지금까지도 계속과제로 수행되어 오고 있음을 확인할 수 있다.

이러한 양상은, 그만큼 전자기록물의 장기보존이라는 주제가 여러 기관과 여러 분야에 공통적으로 연구과제가 될 수 있는 주제임과 동시에, 단기간에 해답을 제시하기에는 매우 어려운 영역이라는 점, 그럼에도 불구하고, 현 시점에서 관련 기초연구가 수행되어야 하며, 미래에 더욱 활발한 연구가 수행될 것으로 예측되는 분야라는 점 등의 다양한 해석을 가능하게 한다.

이에 본 고에서는 전자기록물의 장기보존과 관련하여 국내외에서 현재 수행중인 연구들을 전자기록물의 장기보존과 관련된 내용을 중심으로 상세히 정리해 보고, 2002년 1월에 ISO 표준 초안으로 인증 발표된 OAIS 참조모형의 내용과 함께 주요 프로젝트에서의 적용 내용에 대하여 정리해보고자 한다. 또한, 국내 전자기록물의 장기보존과 관련해서는 어떠한 방향으로 연구가 수행되어야 할 것인지에 대하여서도 정리해보고자 한다.

2 전자기록물과 장기보존정책

2.1 종이기록물과 전자기록물

전자기록물은 종이기록물과는 많은 부분에 있어서 다른 특성을 갖는다. 우선, 전자기록물은 종이기록물에 비하여 매체에 대한 의존도가 매우 크다. 그런데, 전자기록물의 내용을 담고 있는 디지털매체는 IT기술의 급속한 발전으로 말미암아 대부분 물리적 단명성을 갖고 있으며, 이에 따라 전자기록물의 경우 종이기록물보다 장기보존에 있어서 상대적으로 매우 불안정하고, 매체의 특성상 손실의 위험도도 크다.

또한, 전자기록물의 경우 내용과 매체가 구분되어 있다는 점이다. 종이기록물의 경우 내용과 매체가 하나의 물리적 객체를 이루고 있다. 내용적 구조와 형식이 종이에 의해 물리적으로 고정되어 있어서 이를 분리하여 기술되어서는 안된다. 즉, 매체의 영구보존은 곧 내용의 영구보존으로 이어지고, 내용의 영구보존은 곧 접근성의 보존으로 이어짐으로써, 가독성(readability) 및 해석성(interpretability)을 유지해준다. 또한, 종이매체의 특성에 적합한 환경을 중심으로 한 이상적인 저장장소를 구축한다면, 단 한번의 일회적인 환경 구축으로 반영구적인 장기보존이 가능해진다.

반면, 전자기록물의 경우에는 종이기록물과는 아주 다르다. 매체의 영구보존과 내용의 영구보존, 접근성 유지가 모두 개별적으로 진행이 되어야 한다. 아무리 매체가 잘 보존되고, 내용, 즉 매체 내에 저장된 콘텐츠가 영구·보존된다고 할지라도, 그 전자기록물을 둘러싼 하드웨어적인 환경이 달라진다면, 그 변화에 따라서 접근성이 불가능해짐으로써, 실질적인 보존이 이루어지지 않는다. 따라서, 전자기록물의 경우의 보존이라는 의미는, 항상 현재 시점에서의 접근이 가능함으로써 열람 및 해석, 이해가 가능해야 하는 것이다. 즉, 전자기록물의 보존은 단회적인 환경구축으로 이루어질 수 있는 성격이 아니며, 계속적으로 항상 이루어져야만 하는 것이다.

2.2 전자기록물과 장기보존정책

앞서 종이기록물과 전자기록물의 특성 및 각 특성에 따른 보존의 의미를 살펴보았다. 전자기록물의 경우, 장기보존과 관련된 연구는 여러 관점에서 이루어져왔다.

전자기록물을 저장하는 매체 자체의 물리적 내구성에 대한 연구도 진행되어 왔으며, 내용의 통합성(integrity)과 진본성(authenticity)에 관한 연구도 이루어져왔다. 전자기록물의 장기보존과 관련하여 이전(migration)이나 에뮬레이션(emulation) 등과 같은 기술적인 전략에 대한 연구도 진행이 되어왔으며, 전자기록물의 해석과 관련하여 기원(provenance)이라든지 맥락(context)과 관련된 메타데이터에 관한 연구도 보존과 관련하여 중요한 축을 자리잡고 있다. 이 외에도 법·제도적인 차원에서 해결해나가야 할 문제도 산적하여 있는 것이 사실이다.

또한, 앞서 언급하였듯이, 전자기록물의 보존환경이 일회적으로 구축하여 완료할 수 없으므로, 종이기록물과 대비하여 소요되는 경비 및 전문인력의 규모가 막대하다. 이에 따라 전자기록물의 장기보존 관련 전문가들이 최근에는 전자기록물의 장기보존과 관련된 비용이익 분석모델에 관한 연구에도 많은 관심을 나타내는 듯 하다. 이러한 연구에서는 비용절감과 경제성 분석을 고려하여, 실제로 많은 전자기록물의 보존작업이 전자기록물이 생성되는 시점부터 이미 함께 이루어져야 함을 강조하고 있다.

이와 같이 전자기록물의 장기보존정책은, 다양한 관점에서의 접근과 함께, 특정 주제별로 관련 전문가집단 및 기관을 중심으로 한 협동연구의 필요성이 특히 부각된다.

3 전자기록물 장기보존 관련 선행 연구 프로젝트

미국과 유럽에서는 이미 1990년대 중반부터 전자기록물 장기보존과 관련된 기초연구가 활발히 수행되어 왔다. 이에 비하여 국내에서는, 전자기록물이 아닌 종이기록물이나 희귀본의 보존과 관련된 연구는 체계적으로 연구가 수행되어 오고 있으나, 전자기록물, 특히 전자기록물의 장기보존과 관련된 연구는 아직까지는 매우 미미한 것이 사실이다. 이에 본 단원에서는 외국의 대표적인 연구 프로젝트들을 중심으로 전자기록물의 장기보존 연구의 현황을 살펴보고자 한다.

3.1 InterPares

InterPares는 International Research on Permanent Authentic Records in Electronic Systems의 약어로서, 전자기록물의 영구보존을 다루고 있으며, 특히 진본성(authenticity)에 그 초점을 맞추고 있다. 1994년부터 1997년까지 전자문서의 진본성에 초점을 맞추어 수행된 캐나다의 UBC(the University of British Columbia) 프로젝트가 그 모체였으며, UBC 프로젝트에서는 active 전자레코드의 진본성 보존(preserving authentic active electronic records)을 연구 대상으로 했다면, InterPares 프로젝트는 inactive 전자레코드의 진본성 보존(long-term preservation of authentic inactive electronic records)이라는 데에 그 차이점이 있다. 프로젝트 내용에 의하면, 여기에서 active 전자레코드란, 업무수행상 정기적으로 매일매일 필요한 전자레코드를 의미하고, inactive 전자레코드란 매일 매일의 업무수행상 필요한 레코드는 아니더라도 법적으로 또는 역사적 이유로 보존되어야만 하는 레코드를 의미한다. 즉, 보다 장기적 보존이 필요한 전자기록물을 inactive electronic records로 표현하고 있다.

InterPares 프로젝트는 미국, 캐나다, 이탈리아를 중심으로, 주요 기록관리학과 관련 교수와 연구원·컴퓨터공학자·시스템디자이너·국립기록보존소 연구원 및 정책결정자 등 관련분야의 전문가들이 대거 참여하는 세계적인 규모의 정부와 산학 공동연구프로젝트로서, 1999년부터 2002년까지 이미 1단계 프로젝트가 완료되었고, 2002년부터 현재 계속 2단계 프로젝트가 진행중이다. InterPares 프로젝트에서는, 전자문서의 영구보존을 위한 정책 개발 및 모형구축을 위하여, 전자문서 보존과 관련된 대표적인 세계 13개 기관을 대상으로 심층적인 인터뷰와 설문조사를 통하여 보편적인 전자문서 보존현황을 조사한 후 방향을 제시하였다. 특히, 이 프로젝트의 특징은, 이러한 실증적이고도 귀납적인 연구방법과 함께, 고문서학에 근간을 둔 이론적이고 연역적인 방법론도 함께 병행했다는 데에 있다. 즉, 전통적인 기록보존과학으로부터 도출된 방법론과 현재의 사례연구를 함께 고려하여 연구하였다.

3.2 NEDLIB

NEDLIB(Networked European Deposit Library)은 유럽의 국가 도서관들 간의 협력프

로젝트로서 수행되었는데, 네트워크화된 유럽 보존 도서관을 실현하기 위하여 기본적인 기반구조를 구축하는 데에 그 목적이 있다. 특히, 이 프로젝트의 목적은 도서관에서 활용될 디지털 정보의 장기적인 보존에 있는데, EC의 Telematics Application Programme의 재정적 지원 하에서 프랑스·노르웨이·핀란드·독일·포르투갈·스위스·이탈리아·네덜란드를 중심으로 수행되었으며, 1998년에 시작하여 2000년도에 종료되었다.

NEDLIB 프로젝트의 가장 큰 특징은, 디지털 정보자원 보존시스템을 스탠드 얼론 형식의 단일시스템(monolithic system) 형태로 구축한 것이 아니라, 기존에 개발된 디지털 도서관 기반구조(Infrastructure)에 끼워 넣어질 수 있는(embedded) “플러그 인(plugin)” 모델로 구축되었다는 데에 있다.

이러한 구상은, 미래 환경에서도 견디어낼 수 있는(future-proof) 시스템으로 구축함으로써, 어떠한 디지털 형태에서도, 기원이나 매체, 포맷에 상관없이 실용적으로 활용할 수 있는 시스템을 만들고자 한 노력에서 기인한 것이다. 즉, NEDLIB 프로젝트에서는 기존의 도서관 자동화시스템(Existing Library Automation System)에 DSEP(Deposit Systems for Electronic Publications) 기능적 모형을 플러그 인 형태로 결합하여 구축하였다. 이에 따라, 총 11단계로 이루어지는 과정 중, 기존의 도서관 자동화시스템에서 전자출판물을 대상으로 ① 선정(Selection), ② 수서(Acquisition) 단계를 수행하며, 이어서, DSEP 시스템을 중심으로 ③ 캡처(Capture), ④ 등록(Registration), ⑤ 검증(Verification) 단계가 수행된다. 다시 도서관자동화시스템에서, ⑥ 기술(Description) 단계를 수행하며, 이어서 DSEP 시스템에서 ⑦ 저장(Storage), ⑧ 보존(Preservation), ⑨ 전송(Delivery) 단계를 수행하고, 도서관자동화시스템에서 ⑩ 접근(Access), DSEP 시스템에서 ⑪ 점검(Monitoring)을 수행함으로써, 전 단계 중, 특히 디지털 출판물의 장기보존과 관련된 단계에 한하여서 DSEP 시스템이 플러그 인 형태로 결합되어 있음을 확인할 수 있다.

이 중, 특히 디지털 출판물의 장기보존과 관련된 DSEP 시스템 수행단계를 기능을 중심으로 간략히 살펴보면 다음과 같다.

■ 캡처(Capture) - 전자출판물의 복사본이 출판사 시스템으로부터 도서관 deposit system으로 고정된 매체(CD Rom, 플로피디스크), 또는 네트워크를 통해 도착하는 과정인데, 이 과정에는 세가지 주요 작업이 수행되게 된다.

- ① 진본성 확인(Authentication) : 도서관에 등록된 출판사항과 대조하여 진본성을 확인함.
- ② 품질 확인(Quality Scan) : 전자출판물의 기본적인 바이러스 감염여부 등 확인
- ③ 보존(납본)을 위해 전송(Transfer to Deposit) : 새로 입수된 전자출판물과 그 전자출판물을 재생시켜 해독하는 데에 필요한 시스템 환경(system configuration)이 함께 deposit package로서 보존됨. deposit package는 미래에 사용할 수 있도록 콘텐츠 데이터 및 시스템 환경과 관련된 데이터가 함께 저장소(container)로 전송이 됨.

■ 등록(Registration) - 기존 도서관자동화 시스템 2단계에서 수행된 수서(Acquisition) 작업과 맞물려 진행이 되며, 수집된 데이터와 시스템 환경데이터가 함께 deposit package로서 시스템에 등록이 된다.

■ 검증(Verification) - 새롭게 입수된 전자출판물의 물리적·논리적 통합성(integrity, 무결성)을 검증하기 위한 통제과정(control routine)이 수행된다. 세가지 작업이 수행되며, 내용은 다음과 같다.

- ① 패키지 유용성 확인(validate package) : 표준을 비롯한 패키지의 물리적 통합성(무결성,integrity)을 확인한다.
- ② 논리적 무결성(logical integrity) 확인 : deposit package를 unpack하여, 모든 필요한 구성요소들이 존재하는지를 확인한다.
- ③ 진본성 검증(establish authenticity) : deposit copy가 보존을 위하여 deposit system으로 이전(migrate)되었다는 것을 입증하기 위한 납본에 대한 진본 확인 도장(deposit authenticity stamp)을 찍는다.
- ④ 실행과 테스트(installation and testing) : 적절하게 작동(proper operation)하는지를 확인

■ 저장(Storage) - 새로운 전자출판물이 deposit system 의 저장소로 이동하며, 두가지 작업이 수행된다.

- ① 정기적 매체이전(periodic refreshment) : 새로운 저장매체로 정기적으로 이전시킴.
- ② 지속적인 백업(continuous backup) : 안전을 위해 정기적으로 백업화한다.

■ 보존(Preservation) - 전자출판물의 훼손이나 유실에 대비하기 위한 보다 적극적인 보존전략을 수행함. 다음 세가지로 요약됨.

- ① 포맷 특정화(format specification) : 미래에 포맷을 디코딩할때를 대비하여 함께 효율적인 작업이 가능하도록, HTML 또는 XML 포맷으로 통일하여 특정화함.
- ② 포맷 이전(format migration) : 전자출판물을 한 포맷에서 다른 포맷으로 이전함. 특히, DSEP는 정보손실에 대비하여, 대규모의 이전절차와 자동화된 품질통제 절차를 지원하고 있음.
- ③ 하드웨어 에뮬레이션(hardware emulation) : 쇠퇴된 컴퓨터 환경을 재생시키는 것을 포함함.

■ 전송(Delivery) - deposit package를 보존과 접근을 위해 각각 적절한 경로로 전송함.

■ 점검(Monitoring) - 전 과정을 주기적으로 점검하고 통제함.

3.3 CEDARS

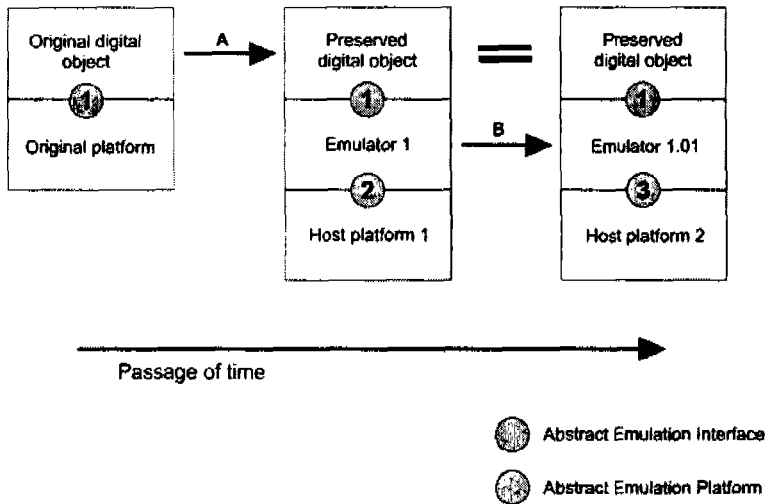
CEDARS(CURL Exemplars in Digital Archives) 프로젝트는 1998년 4월에 시작하여 2002년 3월에 종료된 사업으로서, 디지털 보존과 관련된 주제들을 발굴하고 디지털 아카이브즈 관련 향후 방향을 제시하고자 하는 데에 그 목적이 있다. 영국의 Leeds, Oxford, Cambridge 세 대학의 협력체인 CURL(Consortium of University Research Libraries)를 중심으로 진행되었으며, JISC(Joint Information Systems Committee)의 재정적 지원을 통하여 수행되었다.

CEDARS 프로젝트는 다섯가지 주요 영역을 중심으로 연구가 수행되었는데, 메타데이터 보존·지적재산권·장서관리·기술적 전략·디지털 아카이빙 프로토타입이 주요 주제이다. 이 중 디지털 아카이빙 프로토타입은 분산 디지털 아카이빙 프로토타입(distributed digital archiving prototype)이며, OAIS 모델을 기초로 하여 구축하였다. CEDARS에서는 대표적인 기술적 보존전략으로 일컬어지는, 기술보존, 에뮬레이션, 이전 중에서 이러한 세 전략의 혼합(arnalgam)적인 전략을, 특히 에뮬레이션을 중심으로 제안하고 있다. 이에 대한 내용을 정리하여 발전시킨 연구내용은 협력 수행하였던 CAMiLEON 프로젝트에서 잘 나타내고 있다.

3.4 CAMiLEON

CAMiLEON(Creating Creative Archiving at Michigan and Leeds: Emulating the Old on the New)은 에뮬레이션 기법을 중심으로 디지털 보존 방법론을 연구한 프로젝트이다. 이 프로젝트는 미국 미시간대학의 문헌정보학과와 영국의 Leeds 대학의 협동연구로 수행되었으며, 미국의 국립과학재단(National Science Foundation)과 영국의 JISC(Joint Information Systems Committee)의 재정적 지원으로 이루어졌으며, 1999년에 시작하여 2002년 말까지 완료예정으로서 수행 중이다. 에뮬레이션은 쇠퇴된 소프트웨어를 미래의 알지못하는 새 시스템에 구동시킴으로써 디지털 문서의 검색, 제시, 재사용을 가능하게끔 하는 보존전략으로 제안되어 왔다.

<그림 1> 시간경과에 따른 접근성 보존방법



CAMiLEON 프로젝트 팀에서는 전자기록물의 지적내용, 구조, 그리고 서로 다른 형태의 디지털 문서들의 “look-and-feel”을 보존하기 위하여 에뮬레이션이 얼마나 효과가 있고, 또 구현가능한지를 검증하였다. 그동안은 이러한 에뮬레이션 기법에 대하여, 너무 이론적이고 가설·중심적인 기법이라는 논쟁이 있어 왔다. CAMiLEON 프로젝트에서는 이러한

상황에서 균형을 잡기 위한 연구를 수행하였으며, 에뮬레이션을 실질적인 보존전략으로 활용하기 위한 실험을 지속적으로 수행하고 있으며, 프로젝트 남은 기간까지도 지속할 예정이라고 한다. 구체적으로는, 1970년대와 1980년대에 생산된 자료들을 중심으로 어떻게 급속히 변하는 아키텍처로 매핑할 수 있는지를 실험하고 있다.

<그림 1>은 디지털 원본을 디지털 객체로 보존하고, 에뮬레이션을 통하여 지속적으로 원본에의 접근성을 보장하는 과정을 나타낸 그림이다. 처음에는 화살표 A로 1차적 보존이 이어지지만, 실제로 시간이 흐름에 따라 플랫폼 1은 쇠퇴해가고 화살표 B를 통한 새로운 플랫폼으로 교체되면서 접근이 이어져가게 된다.

3.5 PRISM

PRISM 프로젝트는 미국 코넬대학에서 1999년에 시작하였는데, 지금까지도 계속 과제로 수행중인 연구 프로젝트이다. PRISM은 Preservation, Reliability, Interoperability, Security and Metadata을 의미하는데, 보존(Preservation)은 디지털형태에 담겨진 정보의 영구보존(long term survivability)을 의미하며, 신뢰성(Reliability)은 정보자원과 서비스의 유용성에 초점을 맞춘 개념이다. 상호운용성(Interoperability)은 정보제공자와 이용자간의 보다 광범위한 정보공유를 위한 개방형 표준을 의미하며, 보안(Security)은 정보이용자의 프라이버시와 관련된 권리와 콘텐츠 작성자의 지적재산권 모두를 보호하기 위한 것이다.

메타데이터(Metadata)는 디지털도서관에서의 정보 통합성을 가능하게 해주는 구조화된 정보를 위한 것이다. PRISM 프로젝트는 이와 같이 디지털도서관에서의 정보통합성(information integrity)에 필요한 정책이나 메커니즘을 개발하는 데에 초점을 맞추고 있다. 사서, 컴퓨터과학자, 시스템실무자들이 협력하여 수행 중인 프로젝트이며, 특히, 디지털 보존과 관련해서는 기술적(technical) 전략과 적절한 정책(policy) 수립을 강조하고 있다. 디지털 보존 분야에는 사서·아키비스트·정보전문가 등이 함께 연구를 수행하고 있으며, PRISM 프로젝트에서 개발한 CRADL(the Cornell Reference Architecture for Distributed Digital Libraries) 아키텍처 내에서 보존서비스를 어떻게 기술적으로 구현할 것인가를 연구 중에 있다. 최근에는, PRISM 프로젝트를 수행해 나가면서 파생된 관련 주제들을 중심

으로 새로운 하위 프로젝트들이 수행 중에 있다.

코넬대학도서관(CUL, Cornell University Library)을 중심으로 한, CUL PRISM 연구 프로젝트로서, 주제는 웹 자원의 리스크관리와 보존, 디지털 정보의 영구보존에 관한 연구, 웹 자원 리스크관리 실험 등의 주제를 중심으로 하며, 연구주체는 코넬대학 도서관과 컴퓨터과학과를 중심으로, 2003년까지 수행예정으로 있다.

4 OAI 참조모형 및 적용

4.1 OAI 참조모형

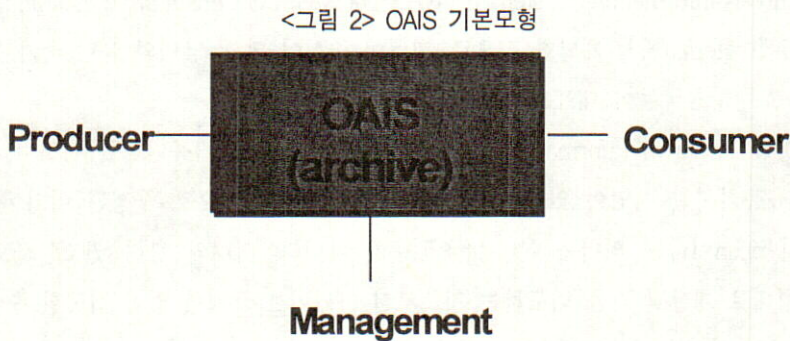
2002년 1월 ISO 표준(14721:2002)으로 확정된 OAI 참조모형(Reference Model for an Open Archival Information System)은 NASA의 CCSDS(Consultative Committee for Space Data Systems)에서 개발한 것으로, 미국이 중심이 되고, 남미와 유럽 국가 등 10여 개 국가의 우주항공국들이 대표기관으로 참여하고 있다.

OAI(Open Archival Information System)는, 정의된 특정 커뮤니티를 위하여, 정보를 보존하고 유용화시킬 책임이 있는 사람들과 시스템들의 조직으로 구성된 아카이브이며, 여기에서의 open이라는 의미는 본 기본안 뿐만 아니라, 미래의 관련 권고안들도 open forum의 형태로 개발될 것을 내포하는 것으로서 아카이브에 제한 없이 접근할 수 있음을 의미하는 것이 아님을 확인할 필요가 있다.

OAI 참조모형은 디지털 정보의 영속적인 장기보존을 위하여 요구되는 사항들에 대한 광범위한 합의(consensus)를 중심으로 개발된 기술적인 권고사항이다. 또한, OAI 참조모형의 확정은 첫째, 전자기록물의 영구 보존과 관련하여 각계 각 분야의 전문가들의 서로 다른 시각 속에서 합의점을 도출해 냈다는 데에 큰 의미가 있으며 둘째, 향후 표준과 보존을 위한 기초모형을 마련했다는 점에서 그 의의를 찾을 수 있을 것이다. 보고서에서는 OAI 참조모형의 유용성을 여러 가지로 정리하고 있는데, 그 내용은 다음과 같다.

- ① 디지털 정보의 장기보존과 접근에 필요한 archival 개념의 이해와 인식에 필요한 개념들 제공
- ② 보존과정에 non-archival 조직들도 효율적으로 참여할 수 있도록 기초개념 제공
- ③ 현존 아카이브즈와 미래 아카이브즈의 구조와 기능을 기술하고 비교하기 위한 용어와 개념을 포함한 기본 틀 제공
- ④ 서로 다른 장기보존 전략과 기법을 기술하고 비교하기 위한 개념들 제공
- ⑤ 아카이브즈에 보존된 디지털 정보의 데이터 모델들을 비교할 수 있는 기초 개념들 제공
- ⑥ 장기간 보존에 필요한 기본요소에 대한 파악 가능

전자기록물의 증가와 디지털 형태의 보편화는, 이전에는 아카이빙 기능을 수행해야 한다는 생각을 전혀 하지 못했을 많은 기관들조차 영구보존과 아카이빙에 대한 장기적인 계획의 필요성을 상기시켜 주었다. OAIS를 둘러싼 가장 기본적인 환경을 그림으로 나타내면 다음의 <그림 2>와 같다.



위의 그림에서 생산자(Producer)를 통하여 장기보존 되어야 할 정보가 제공되면, OAIS를 통하여 장기보존에 필요한 절차와 요소들이 체계화된다. 그리고, 이후 미래의 이용자(Consumer)에게 장기보존된 정보를 제공하게 된다.

이 때에 소통하게 되는 정보는 정보패키지 형태로 이동하게 되는데, OAIS에서 유통되는 정보패키지의 구성을 그림으로 나타내면 다음의 <그림 3>과 같다.

<그림 3> OAI 내에서 유통되는 정보패키지(Information Package)



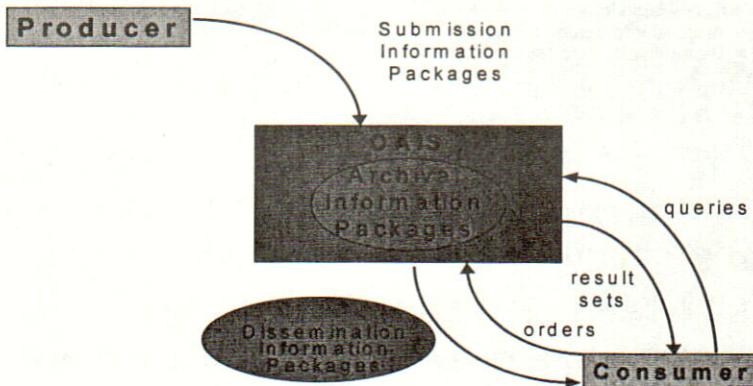
- An Information Package is a conceptual container holding two types of information
 - Content Information
 - Preservation Description Information (PDI)

OAI에서 유통되는 정보패키지는 정보컨텐츠(CI)와 보존기술정보(PDI)로 구성되는데, 보존 기술정보(PDI)는 정보컨텐츠(CI)에 대한 일종의 메타데이터로서 참조정보(Reference)·출처정보(Provenance)·문맥정보(Context)·인증정보(fixity) 등으로 구성된다.

OAI 모형에서는 이러한 메타데이터의 대상이 되는 정보를 세가지 유형으로 구분해서, 각각 유형에 따른 참조정보·출처정보·문맥정보·인증정보 등의 보존기술정보(PDI)에 대해서 기술하고 있는데, 구분한 세가지 유형은, 과학데이터(Space Science Data), 디지털 도서관컬렉션(Digital Library Collection), 소프트웨어 패키지이다.

OAI에서의 정보패키지의 흐름을 한단계 더 자세히 나타내면 다음의 <그림 4>와 같다.

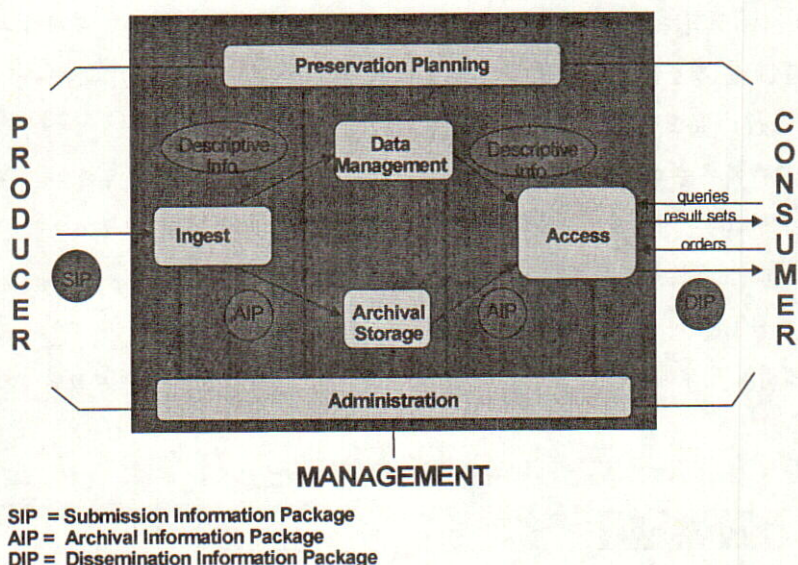
<그림 4> OAI 내에서의 정보흐름도



그림을 통하여, 앞에서 설명한 정보패키지(Information Package)는 크게 세가지로 구분되어서 OAIS 모형 내에서 유통되고 있음을 확인할 수 있다. 먼저, 생산자에서 OAIS로 유통되는 정보패키지는 SIP(Submission Information Package)이며, OAIS 내에서 영구보존되는 정보패키지는 AIP(Archival Information Package)이다. 또한, OAIS 내에서 보존된 정보가 이용자에게 제공될 때에는 DIP(Dissemination Information Package)형태로 제공되고 있음을 확인할 수 있다.

OAIS 내에서 수행되는 기능을 좀 더 구체적인 기능으로 표현할 수 있는데, 이러한 구체적인 여섯 가지 세부 기능을 중심으로 한 OAIS의 모형은 다음의 <그림 5>와 같다.

<그림 5> OAIS 모형 내의 주요 여섯가지 기능



OAIS에서 수행되는 여섯가지 세부 기능에는 흡수(Ingest)·보존(Archival Storage)·데이터관리(Data Management)·운영(Administration)·보존계획(Preservation Planning)·접근(Access) 등이 포함된다. 그리고, 이러한 여섯 가지 기능들을 중심으로 앞에서 설명한 세가지 형태의 정보패키지가 유통된다.

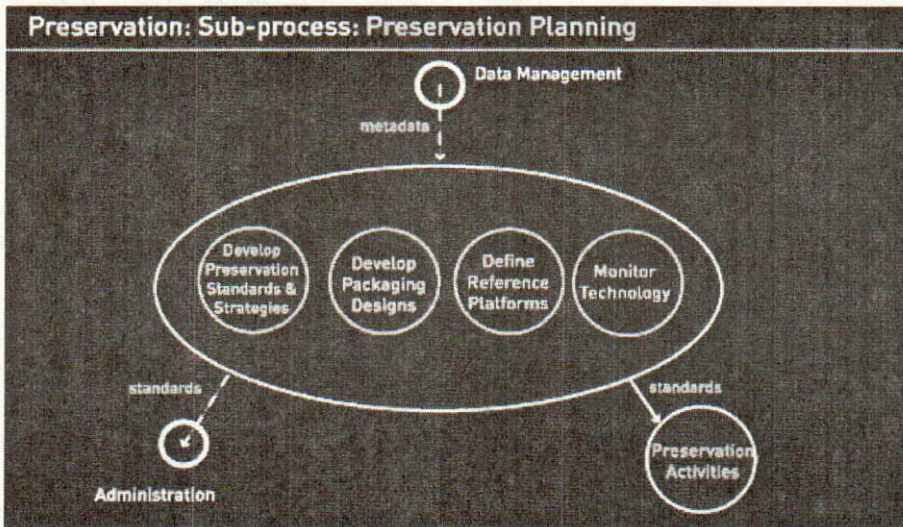
먼저, 생산자로부터 OAIS에는 SIP 정보패키지가 전달되며, OAIS 내에서는, 흡수(Ingest)에서 보존(Archival Storage)으로, 그리고 보존(Archival Storage)에서 접근(Access) 방향으로

4.2 OAIS 참조모형의 적용

이상 기능을 중심으로 간략히 설명한 OAIS 참조 모형은, 전자기록물의 장기보존을 중심으로 한 상당수의 프로젝트에서 각기 성격에 따라 적용하고 있는 것을 확인할 수 있다. 특히, CEDARS 프로젝트와 NEDLIB 프로젝트에서는 기본적인 OAIS 모형을 상당히 많이 활용하고 있으며, PANDORA, InterPares 프로젝트에서도 기본적인 개념을 채택하여 활용하고 있음을 확인할 수 있다. OAIS 보고서에서도, 실제로 전자기록물의 장기보존과 관련된 다양한 기관에서, 다양한 목적에 따라 폭넓게 적용할 수 있도록 참조모형을 개발했음을 밝히고 있다.

다음의 <그림 7>은 NEDLIB 프로젝트의 DSEP 시스템에서의 보존과정의 일부분을 표현한 것이다.

<그림 7> NEDLIB의 DSEP시스템에서의 보존계획(Preservation Planning)



<그림 7>을 통하여, NEDLIB의 DSEP 시스템에서도 <그림 6>에서 살펴본 OAIS의 주요 보존계획(Preservation Planning)의 핵심기능을 그대로 채택하고 있음을 확인할 수 있다. 즉, 보존 표준과 전략을 개발하고, 패키지 디자인을 개발한다. 또한, 참조 플랫폼을 정

의하고 있으며, 다른 기능들과의 상호교류를 통하여 정기적으로 기술을 모니터링함으로써 신기술이 출현하였을 때 반영하고 있음을 확인할 수 있다.

5 향후 방향과 제언

지금까지 간략하나마, 디지털 아카이빙과 관련하여 현재 수행 중이거나 최근에 완료된 주요 프로젝트들을 중심으로 살펴보았다. 인상적인 것은, 상당히 다양한 배경의 전문가들이 적절한 역할분담 속에서 디지털 아카이빙이라는 공통주제에 대한 연구를 협력하여 수행해 나가고 있다는 점이다. 뿐만 아니라, 국가간의 영역을 뛰어넘어 전세계적인 규모로 활발히 연구가 진행되고 있음도 확인할 수 있었다.

특히, 전자기록물의 장기(또는 영구)보존이라는 주제는, 수년전부터 그 중요성을 인식해 왔음에도 불구하고, 단기간에 검증이 되기 어렵고, 한 세대가 지나가서야 비로소 어느 정도 확실히 확인할 수 있는 성격이기에, 이론적이고 이상적인 방향제안에 그치는 것이 아닐까 하는 우려 속에서 상당수의 연구가 진행되었던 것도 사실이다. 그러나, 이제는 한 걸음씩 구체화되고, 또 검증될 수 있는 영역으로 개척되어 나가고 있는 듯하다.

우리 나라에서는, 실제로 기록물에 대한 법적 마련도 2000년에 들어와서야 비로소 마련되기 시작한 만큼, 아직은 전자기록물의 장기보존이라든지, 전자기록물의 보존을 포괄하여 전반적인 시스템까지도 포함한 주제인 디지털 아카이빙에 대한 연구를 언급하기에는 여러 가지 상황과 여건이 미흡한 것은 사실이다.

그러나, 선행연구들을 통하여 확인하였듯이, 효율적인 전자기록물의 적절한 보존정책 수립을 위해서는, 전자기록물의 창출 시점인 첫 단계에서부터 함께 고려되어야 할 것으로 믿는다. 본격적인 전자정부시대의 개막과 더불어, 첨단 IT 기반시설 및 활용도에 있어서는 우리나라가 조금도 뒤떨어짐이 없기에 더욱 더 그러하다. 우리나라에서도 이제는 국제표준으로 자리잡아가는 OAIS에 대한 검토와 더불어, 관련 학계와 연구소를 중심으로 한 전문가들의 협력 속에서, 전자기록물의 보존과 관련된 기본적인 체계를 틀을 시급히 마련함으로써 디지털 아카이빙 관련 연구가 더욱 더 활발히 수행될 것을 기대해 본다.

<참고문헌>

- 박은경. 2002. "전자문서의 영구보존을 위한 표준화 연구 동향". 한국기록관리학회지 2(1): 97-118.
- 신은자. 2001. "전자저널의 아카이빙에 관한 연구". 정보관리학회지 18(3): 139-158.
- 이소연. 2002. "디지털 아카이빙의 표준화와 OAIS 참조모형". 정보관리연구 33(3):45-68.
- 최원태. 2001. "디지털 아카이브의 현황 및 구성 요소에 관한 연구". 한국문헌정보학회지 35(2): 23-40.
- Cloonan, V. Michele. 2002. "Shifting from paper to Digital Records - Preservation?" 한국기록관리학회지 2(2): 137-148.
- Beagrie, Neil. and Greenstein, Daniel. 1998. A Strategic Policy Framework for Creating and Preserving Digital Collections. [cited 2002. 10. 7].
<<http://www.ukoln.ac.uk/services/papers/bl/framework/framework.html>>.
- CCSDS. 2002. Reference Model for an Open Archival Information System(OAIS). CCSDS 650.0-B-1 Blue Book. [cited 2002.10.24].
<http://www.ccsds.org/blue_books.html>.
- Granger, Stewart. 2001. Digital Preservation & the CAMILEON PROJECT. [cited 2002.10.24].
<<http://dSPACE.dial.pipex.com/stewartg/cam-london.htm>>.
- Hofman, Hans. 2002. ERPANET:Principles of Digital Preservation. [cited 2002.10.24].
<http://www.erpanet.org/www/content/documents/Digitalpreservationcharterv4_1.pdf>.
- Holdsworth, D. 2001. Architecture of CEDARS Demonstrator. [cited 2002.10.24].
<<http://www.personal.leeds.ac.uk/~eclhdh/cedars/architecture.html>>.
- InterPares Project Team. 2001. The Long-term Preservation of Authentic Electronic Records :Findings of the InterPARES Project. [cited 2002.10.7].
<<http://www.interpares.org/book/index.cfm>>.
- NEDLIB Homepage. [cited 2002.10.7]. <<http://www.kb.nl/coop/nedlib/index.html>>.
- Open Archives Initiatives Homepage. [cited 2002.10.7].

- <<http://www.openarchives.org>>.
- Rieger, O. 2000. Project Prism : Preservation Metadata Research. Information Infrastructures for Digital Preservation Conference, [cited 2002. 10.7].
<<http://www.rlg.org/events/pres-2000/rieger/index.htm>>.
- RLG. 2002. Open Archival Information System (OAI) Resources. [cited 2002. 10. 24].
<<http://www.rlg.org/longterm/oais.html>>.
- Russell, Kelly., and Weinberger, Ellis. 2000. Cost Elements of Digital Preservation. [cited 2002.11.10].
<<http://http://www.leeds.ac.uk/cedars/documents/CIW01r.html>>.
- Russell, Kelly. 2000. Digital Preservation and the Cedars Project Experience. An International Conference on the Preservation and Long Term Accessibility of Digital Materials, [cited 2002.10.24].
<<http://www.rlg.org/events/pres-2000/russell.html>>.
- Wheatley, Paul. 2001. Migration - a CAMiLEON discussion paper. [cited 2002.10.7].
<<http://www.ariadne.ac.uk/issue29/camileon/>>.

* 논문접수일 : 2003년 1월 10일

* 논문심사일 : 2003년 1월 15일