

# 해외 전자저널의 디지털 아카이브 구축 전략에 관한 연구

## A Study on the Strategies for Building a Digital Archive of Electronic Journals

최 호 남(Ho-Nam Choi)\*

이 응 봉(Eung-Bong Lee)\*\*

### 목 차

- |                   |                          |
|-------------------|--------------------------|
| 1. 서론             | 3. 1 디지털 아카이빙의 목적과 필요성   |
| 1. 1 연구의 목적 및 필요성 | 3. 2 전자저널 도입과 디지털 아카이빙   |
| 1. 2 연구의 범위 및 방법  | 3. 3 전자저널 아카이빙을 위한 선결 과제 |
| 2. 선행연구           | 4. 아카이브 구축을 위한 전략 및 제안   |
| 2. 1 국내           | 4. 1 구축 모델링과 콘텐츠 전략      |
| 2. 2 국외           | 4. 2 기술전략                |
| 3. 디지털 아카이빙과 선결과제 | 5. 결 론                   |

### 초 록

본 연구에서는 현재의 라이선스 기반의 전자저널 구독모델 하에서 서비스 중단 등 미래의 예기치 못한 불확실성을 해소하고 영속적인 원문 액세스의 보장을 확보하기 위해 반드시 필요한 해외 전자저널 원문자료에 대한 디지털 아카이브의 구축과 보존을 수행하는 데 예상되는 제반 위험요소와 핵심 성공요인을 분석함으로써 우리나라 도서관 현실에 맞는 디지털 아카이브 구축 모델과 전략을 제시하고자 한다.

### ABSTRACT

Current electronic journals subscription model based on license purchasing faces a dilemma of uncertainty in securing the perpetual access to the e-journal content that was once licensed and paid. In order to avoid future risks of suspended services this study addresses a digital archiving model and recommends a practical strategies for building a digital archive suitable for Korean library society by analyzing all of both expected problematic issues and success factors in performing the project.

키워드: 전자저널, 아카이빙, 아카이브, 디지털 아카이빙, 디지털 보존, 아카이빙 전략, 디지털 아카이브 구축  
Electronic Journal, Archiving, Preserving, Preservation, Digital Archiving, Digital Preservation, Archiving Strategy, Digital Archive

\* 한국과학기술원 과학기술전자도서관 책임연구원(hnchoi@kaist.ac.kr)

\*\* 충남대학교 사회과학대학 문헌정보학과 부교수(eblee@cnu.ac.kr)

논문접수일자 2005년 5월 15일

게재확정일자 2005년 6월 10일

## 1. 서론

### 1.1 연구의 목적 및 필요성

우리나라의 지식 축적량은 OECD 국가 중 최하위 그룹에 속하며 해외 학술정보에 대한 의존도가 매우 높다. 특히, 과학기술분야 연구자들의 경우 해외 학술저널에 의존하는 비율이 매우 높기 때문에 해외 학술자료의 원활한 유통 및 적시적 공급이 연구생산성 제고 및 나아가서 국가 경쟁력 강화에 많은 영향을 미치고 있다.

전통적으로 도서관이 인쇄형태의 도서나 저널을 구매함으로써 이들 정보자료를 축적하고 조직하여 필요로 하는 이용자에게 정보서비스를 제공하고, 이들 자료를 영원히 소유하여 이용자들에게 일정한 기준에 따라 자료를 영속적으로 이용하게 하던 것과 달리, 전자저널의 경우에는 특정 이용 조건하에서 특정 기간 동안만 이들 자료에 접근할 수 있는 권리가 허락되어진다. 또한 허락된 이용권의 내용 혹은 조건에 따라 이용자들의 이용방식이 달라지게 되며, 이는 이용자들이 미래에 동일한 정보자원을 계속 접근 활용할 수 있게 하는 자료의 영속적 접근(perpetual access) 기능에 제약을 초래하고 있음을 뜻한다.

현재, 라이선스 기반의 전자저널 구독 모델 하에서는 정보자원을 각 기관이 도입하는 것이 아니고 단지 계약 기간 동안 이용만을 할 수 있기 때문에 이를 안정적으로 관리 할 수 없다. 따라서 출판사의 도산, 예기치 못한 서비스의 중단 등 미래에 있을지 모르는 어떠한 위협에 항상 노출되어 있는 상태이다.

따라서, 이를 해결하기 위해서는 최소한의

안전장치로서 원문자료의 확보 및 이를 영구적으로 보존할 수 있는 제도 및 장치 마련이 절실히 요구된다. 그러나, 해외에서 생산되는 디지털 원문자료의 아카이브 구축을 위해서는 제도적 측면, 법적 측면, 그리고 기술적 측면에서 해결해야 할 과제들이 산적해 있다.

본 연구에서는 국내 각급 기관에서 필요로 하는 해외 전자저널 원문자료에 대한 디지털 아카이브 구축 및 운영에 따르는 제반 위험요소와 핵심 성공요인을 분석함으로써 현실적으로 가장 바람직한 디지털 아카이브 구축 모델을 제시하고자 한다.

### 1.2 연구의 범위 및 방법

본 연구에서는 현재 카이스트가 운영하고 있는 해외 전자저널 공동구매 콘소시엄인 KESLI (Korea Electronic Site License Initiative)에서 유료로 라이선스하고 있는 해외 전자저널을 아카이브 대상으로 고찰하였으며, 전자책이나 기타 유형의 전자정보는 연구대상에서 제외하였다. 따라서 2004년 12월말 현재 299개로 구성된 KESLI 참여도서관이 44개 해외 공급사로부터 구독하는 총 8,863종의 전자저널이 아카이브 대상이다. 국내에서는 아직까지 디지털 아카이브 구축 사례가 없기 때문에 주로 해외의 디지털 아카이브에 대한 조사와 연구에 중점을 두었고, 연구방법으로서 외국의 디지털 아카이브 구축 기관에 대한 현지 출장 조사(National Library of Netherlands: 2003.10, LOCKSS, Stanford: 2003.10, National Library of Australia: 2004.6, Ithaka: 2004.6, OCLC: 2004.6)와 해당 전문가와의 세미나(Archive

and Preservation Issues and Approaches for Digital Libraries, 2004.5, Archiving and Preserving of e-Journals: 2004.8, EJOS 시연을 위한 원격세미나: 2004.9)를 통해 얻은 지식과 기타 관련 문헌 조사를 통해서 연구를 수행하였다.

## 2. 선행연구

### 2.1 국내

국내의 경우, 디지털 콘텐츠에 대한 아카이빙 움직임은 시작단계라 할 수 있으며, 현재 국회도서관, KERIS, KISTI 등의 국가기관과 북토피아, DBPIA 등의 개인기업체에서 디지털 콘텐츠를 저장 관리하고 있다. 그리고 국가기록원은 기록물관리법 시행령 제30조에 의거 자료관시스템(정보공개청구의 접수 등 기록물의 관리가 전자적으로 수행)과 전문관리기관시스템(중앙기록물관리기관, 특수기록물관리기관 등에서의 기록물의 수집, 평가, 재분류, 보존)의 구축을 진행 중이다. 하지만, 이들 디지털 아카이브는 장기 보존 및 액세스에 대한 체계적 접근이 없는 상태로서 진정한 의미의 아카이브로 보기는 어려운 실정이다.

### 2.2 국외

해외에서는 디지털 아카이빙에 관한 연구가

활발하게 진행되고 있으며, 현재까지 진행된 주요 연구 사례를 살펴 보면 다음과 같다.

#### 2. 2. 1 Yale 대학교의 디지털 아카이빙<sup>1)</sup>

2001년도에 Mellon 재단의 지원을 받아 Yale 대학 도서관과 Elsevier사의 공동연구로 수행된 이 연구에서는 첫째, 주 정보자원의 장기적 손상, 권리보유자 또는 승계자의 책임포기 등으로 인한 이용성의 손상, 합의된 기간의 경과 등 디지털 아카이빙을 촉발할 수 있는 사안들에 대한 연구, 둘째, 아카이브 구축에 비용이 소요되는 주된 요인들을 밝혀내는 비용요인에 관한 연구, 셋째, 이러한 비용조달 모델 및 라이선스 모델에 관한 연구와 기타 아카이브 대상자료 선정과 메타데이터에 관한 연구를 수행하였다.

#### 2. 2. 2 Harvard 대학의 디지털 아카이빙<sup>2)</sup>

2001년 Harvard 대학은 Blackwell, Wiley, University of Chicago Press의 3개 출판사와의 협력을 바탕으로 추진한 디지털 아카이빙 연구를 통해 저널의 구성요소별 아카이브 대상요소의 선정기준을 만들었고, 비즈니스 모델 및 아키텍처 설계에 관한 연구를 마쳤으며, 아카이브 추진 주체들별 역할과 임무를 정의하였다.

#### 2. 2. 3 CEDARS(Curl Exemplars in Digital ARchiveS) project<sup>3)</sup>

영국의 Leeds, Oxford, Cambridge 등 3개 대학이 공동으로 수행한 e-Library 일환인 CEDARS 프로젝트에서는 디지털 정보자원 보존정책의

1) <http://www.library.yale.edu/~okerson/yea/frontmatter.pdf>

2) <http://xml.coverpages.org/harvardEJournalArchive.html>

3) <http://www.leeds.ac.uk/cedars/OurPublications/CedarsProjectReportToMar01.pdf>

수립과 보존 메카니즘의 개발을 목표로, OAIS (Open Archival Information System) 참조 모형을 토대로 한 디지털 정보자원의 보존에 대한 기술적, 조직적 전략 연구, 콘텐츠 관리정책에 대한 가이드라인 제시, 장기보존에 필요한 메타데이터 연구를 수행한 바 있다.

#### 2. 2. 4 네덜란드 국립도서관(Koninklijke Bibliotheek: KB)의 디지털 아카이브4)

1996년 Elsevier사의 전자저널 270종을 KB에 보존하기로 아카이브 계약을 맺고 1998년 IBM과 공동으로 OAIS 참조모형을 응용한 DIAS (Digital Information Archiving system) 개발을 시작하여 2000년말에 완성, 운영중이다. 현재에는 KB가 Kluwer, Springer, Blackwell, Wiley 등의 전자저널 원문자료를 아카이빙하기로 계약을 맺은 상태이다. 아울러, 디지털 자료의 장기적 보존문제를 해결하기 위한 솔루션으로 UVC(Universal Virtual Computer) 개념을 도입하였다.

#### 2. 2. 5 PADORA 프로젝트5)

호주 국립도서관이 추진하는 디지털 아카이브 구축 과제로서 호주 정부간행물 및 관련 웹 정보를 수집하여 보존하는 PANDORA Digital Archive System(PANDAS)을 개발 운영하고 있다.

#### 2. 2. 6 OCLC의 ECO(Electronic Collection Online)6)

ECO는 출판사가 비용부담으로 어려워하는 아카이빙의 임무를 OCLC가 대신 수행하기 위한 모델로서, 전 세계 70여 출판사의 약 5,000여종의 전자저널을 웹으로 서비스 하고 있다. OAIS 모형을 기본 아키텍처 프레임워크로 구축되었으며, 아카이빙 구축을 위해 콘텐츠 수집 및 메타데이터 입력용 도구인 'Harvest', 데이터 무결성 체크 결과보고서 자동 생성 도구인 'Harvest Report', 콘텐츠 권리, 서비스 레벨, 콘텐츠 그룹 선택 등의 정보를 아카이브 단계로 보내 주는 'Ingest Request' 의 3개 도구를 자체 개발하여 사용하고 있다. 특히할만한 점으로는 아카이브 자료의 주기적인 무결성 테스트를 통해 현재의 매체를 갱신해야 하는 데이터를 자동점검하는 장치가 들어있다.

#### 2. 2. 7 미국의 Ithaka에서 추진중인 e-Archive 프로젝트7)

참여출판사 및 도서관에 제공할 아카이빙 서비스에 대한 정의, 아카이브를 지탱할 수 있는 경제적인 모델 개발, 해당 출판사와의 유대관계 개발, 디지털 아카이브가 출판사 및 도서관에 미치는 재정적 효과에 관한 연구를 수행하고 있으며, 부속기관으로 JSTOR와 ARTstor가 있다.

해외 디지털 아카이브 유형은 <표 1>에서와 같이 추진주체별, 목적별, 대상 콘텐츠별, 사용자

4) Johan F. Steenbakkers, *Archiving and preserving of e-journals: The 5th KESLI Annual Meeting Kyeongju, Korea, September 2004.*

5) <http://pandora.nla.gov.au/about.html>

6) <http://www.oclc.org/electroniccollections/default.htm>

7) <http://www.ithaka.org/e-archive/>

〈표 1〉 해외 디지털 아카이브 유형

추진주체별	국가주도	개별기관	출판사	Trusted 3rd Party	
	<ul style="list-style-type: none"> <li>▼KB(네덜란드)</li> <li>▼NLA(호주)</li> <li>▼BL(영국)</li> <li>▼NDIIPP(미국)</li> </ul>	<ul style="list-style-type: none"> <li>▼Harvard</li> <li>▼Yale</li> <li>▼Cornell</li> <li>▼CEDARS(영국)</li> </ul>	<ul style="list-style-type: none"> <li>▼Johns Hopkins</li> <li>▼Blackwell</li> <li>▼Highwire</li> <li>▼Elsevier</li> </ul>	<ul style="list-style-type: none"> <li>▼OCLC</li> <li>▼JSTOR</li> <li>▼ITHAKA</li> </ul>	
목적별	<ul style="list-style-type: none"> <li>▼KB(네덜란드)</li> <li>▼NLA(호주)</li> <li>▼BL(영국)</li> </ul>	<ul style="list-style-type: none"> <li>▼Harvard</li> <li>▼Yale</li> <li>▼Cornell</li> </ul>	<ul style="list-style-type: none"> <li>▼CEDARS</li> </ul>	<ul style="list-style-type: none"> <li>▼Johns Hopkins</li> <li>▼Blackwell</li> <li>▼Highwire</li> <li>▼Elsevier</li> </ul>	<ul style="list-style-type: none"> <li>▼OCLC</li> <li>▼JSTOR</li> </ul>
대상 Contents별	<ul style="list-style-type: none"> <li>▼KB(네덜란드)</li> <li>▼NLA(호주)</li> <li>▼BL(영국)</li> <li>▼NDIIPP(미국)</li> </ul>	<ul style="list-style-type: none"> <li>정부기록물</li> </ul>	<ul style="list-style-type: none"> <li>전자출판물</li> </ul>	<ul style="list-style-type: none"> <li>문화유산</li> </ul>	
사용자 그룹별	<ul style="list-style-type: none"> <li>▼Harvard</li> <li>▼Yale</li> <li>▼Cornell</li> <li>▼CEDARS</li> </ul>	<ul style="list-style-type: none"> <li>학술연구자</li> </ul>	<ul style="list-style-type: none"> <li>▼Harvard</li> <li>▼Yale</li> <li>▼Cornell</li> <li>▼CEDARS</li> <li>▼Blackwell</li> <li>▼JSTOR</li> <li>▼ITHAKA</li> <li>▼Blackwell</li> </ul>	<ul style="list-style-type: none"> <li>일반인</li> </ul>	

추진 주체별로 볼 때, 국가 주도 및 Trusted 3rd Party 그리고 출판사별 Archiving을 상형하였고, 개별기관의 경우 Archiving 연구만을 함.

이항 목적별로 볼 때, 공공 목적을 위한 프로젝트에서 상형한 경우가 많았으며, 기관의 목적을 위해서는 연구프로젝트가 주를 이룸.

대상 Contents별로 볼 때, Trusted 3rd Party 는 주로 학술전자 출판물, 국가 주도형은 정부기록물 및 문화유산물 Archiving함.

사용자 그룹으로 비교해 볼 때, 개별기관, 출판사, Trusted 3rd Party는 학술 연구자를, 국가 주도형은 일반인을 주로 대상으로 함.

그룹별로 구분될 수 있으며 각각의 사례는 아래와 같다. 추진주체별로 살펴 보면, 국가주도형 아카이브로서 KB, NLA(호주국립도서관)가 대표적 예이고, 개별기관이 추진하는 예로서, 예일이나 하버드대학교를 들 수 있다. 그 외에 Elsevier나 Blackwell 등과 같이 출판사 자체에서 추진하는 경우와 제 3자를 통한 아카이브로서 OCLC나 JSTOR 등이 있다. 이들 아카이브는 구축 목적별로 크게 공익적 목적과 상업적 목적으로 구분할 수 있다. KB, NLA, CEDARS가 공익적 목적으로 구축된 데 비해, OCLC, JSTOR는 상업적 목적으로 구축되었다. 또한 KB나 NLA가 정부기록물과 문화유산을 아카이브 주 대상자료로 삼고 있는데 비해 OCLC, JSTOR, Harvard, Yale 등은 전자출판물을 대상으로 하고 있다.

### 3. 디지털 아카이빙과 선결과제

#### 3.1 디지털 아카이빙의 목적과 필요성

디지털 아카이빙이란 한마디로 가치있는 디지털 객체를 장기적으로 이용할 수 있도록 보존, 관리하는 제반 활동을 의미한다. 즉, 선별된 디지털 정보자원과 이에 대한 메타데이터를 보존 관리하여 미래 이용자들이 이 디지털 정보를 현재와 같이 활용할 수 있도록 수행하는 모든 서비스 활동이다.

전자출판물의 등장 이후, 전자출판물의 이용도는 지속적으로 확대되어 가고 있으며, 그에 따른 전자출판물에 대한 라이선스 계약의 한계성 극복을 위한 디지털 원문의 영구적 액세스 확보 및 보존의 필요성에 대한 인식이 증가하고 있다. 1990년대 중반부터 해외 선진 각국에서는 정부 차원에서 디지털 자료의 영구 보존을

위한 프로젝트를 시작하였으며 가시적인 성과를 이루었고 1999년에는 디지털 아카이브간의 상호운용성과 자원의 공동활용을 제고시키기 위해 표준적인 아카이브 모형으로서 OAIS 참조모형이 제안되기도 했다.

특정 기술에 의존해야 콘텐츠를 이용할 수 있는 특징을 지닌 전자저널의 경우에는 분야별로 상이한 목적과 표준을 추구하고 있어, 각 전자저널을 생성, 저장, 배포 및 이용하는 정보기술 환경이 상이하다. 전자저널을 비롯한 전자 학술정보의 지속적 이용을 위해서는 정보기술 환경과 변화에 무관하게 관리되어야 한다.

전자저널 이용자는 전자정보의 이용권을 구매하게 되므로 원문을 보유하지 못하여 라이선스 계약 종료 후의 이용 가능성이 매우 불투명하다. 또한, 전자저널의 원문 파일을 보유하고 있는 출판사의 경영 상태에 따라 전자저널의 이용이 예기치 못한 상황에서 중단될 수 있으며, 출판사의 상업적 가치 판단에 따라 일부 전자저널의 보존이 소홀히 다루어 질 수 있다.

나아가서, 과학 기술 분야의 연구자들이 이용하는 학술정보의 대다수가 해외에서 생산된 정보이며, 특히 학술저널에 많이 의존하고 있는 상황에서, 전자저널의 지속적 이용권 확보는 지식정보시대에 국가적 차원의 주요 지식정보자원 관리라는 관점에서 필수불가결한 일이다.

### 3. 2 전자저널 도입과 디지털 아카이빙

학술저널이란 “번호 혹은 연호 표시를 가지고 무한정, 연속적인 분책으로 발행되고 독창적인 학문에 기반한 상당 부분의 논문들을 포함하여 일반적으로 자료의 입력이 사전에 결정되지

않은 출판물”로 정의하고 있다. 이러한 학술저널을 전자적인 방법으로 이용자들에게 전달하는 수단이 전자저널로서 아래와 같은 특징을 가진다:

- 대학, 학회 등에서 생산한 정보 중에서 학문과 과학의 발전에 사용되는 정보
- 일반 정보보다 전문화되어 있으며, 한정된 이용자가 이용함.
- 정보활동의 주체인 대학, 연구소, 기업 등에게 정보활동을 진작시키는 역할
- 국가의 공공재로서, 국가경쟁력과 직결됨.
- 연구개발 및 교육활동을 지원하는 정보자원의 원천
- 교육, 연구, 학습 등의 목적으로 사용되는 정보를 담고 있는 매체로서, 내재적 가치, 이용가치, 보존적 가치를 지니고 있음.
- 커뮤니케이션이 용이하며 연구자와 독자간의 상호반응이 즉각적임
- 하이퍼링크, 멀티미디어 등의 최신 정보통신기술을 활용하여 다양한 검색 지원 및 이용의 편리성을 제공
- 이용자의 요구에 따른 맞춤서비스를 제공할 수 있으며, 다양한 형태의 상품개발과 요금체계 구성이 가능함.

우리나라가 전자저널을 본격적으로 이용하기 시작한 것은 2000년도부터이다. KAIST가 정보통신부 지원을 받아 추진해 온 NDSL 프로젝트의 일환으로 KESLI라는 대규모 국가적 컨소시엄을 구성하여 전국적으로 전자저널 이용을 확산함으로써 매년 전자저널의 구독량과 투입비용이 늘어나고 있는 추세이다. <표 2>는 2000년부터 2005년까지 국내에서 도입한 연도

별 전자저널 구독량을 나타낸 것이다.

또한, 이러한 해외 전자정보에 대한 의존도가 매년 높아지고 있음은 다음의 통계에서 알 수 있다(표 3 참조). 국내 대학도서관의 연간 평균 자료구입비는 증가하지 않으면서 자료 구입비 중

해외 전자정보 구독료의 비율은 매년 증가하고 있다(그림 1 참조). 하지만 국내 50% 이상의 대학도서관은 5종 이하의 해외전자 정보만을 구독하고 있어, 많은 대학도서관이 예산문제로 자료구입에 어려움을 겪고 있는 것으로 나타났다.

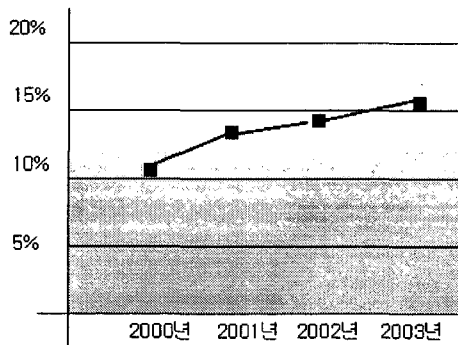
〈표 2〉 연도별 해외 전자저널 라이선스 현황

2005.4.30 현재

구분	2000	2001	2002	2003	2004	2005
컨소시엄 수	6	12	27	31	44	57
총 전자저널 수	2,102종	3,585종	4,813종	5,877종	8,610종	7,766종
참가기관 수	160	223	259	276	299	319
총 전자저널 비용	-	-	\$10,021,642	\$12,433,834	-	-
총 계약 수	383	903	1,431	1,515	1,750	1,976
평균 계약 수	2.39	4.04	5.52	5.48	5.85	6.19

〈표 3〉 대학별 해외전자정보 구독비 및 구독종수 현황

구분	2000	2001	2002	2003
평균 자료구입비 (억원)	8.5	9.5	8.2	8.25
평균 해외전자정보 구입비 (억원)	0.94	1.16	1.2	1.28
자료구입비 중 전자정보 구독비의 평균비율 (%)	10.96	12.20	14.60	15.50
평균 전자정보 구독종수 (종)	8.04	9.37	11.27	15.5



〈그림 1〉 자료비 대비 전자정보 구독비율

전자저널의 아카이빙을 위해서는 먼저 전자저널의 구성요소를 정의하고 이 요소들 중에서 아카이빙 대상이 되는 콘텐츠 요소를 정의하는 것이 필요하다. 하버드대 Mellon 프로젝트팀은 과학자들을 대상으로 인터뷰를 실시하였으며, 이 인터뷰 결과를 바탕으로 아래 <표 4>와 같이 보존 대상이 되는 콘텐츠 요소<sup>8)</sup>를 정의하였다.

가장 어려운 부분은 연구 결과를 보여주는 데 필수적인 타 자료와의 연결을 보존하는 문제인데, 이 연결이 동적인 특성을 가진 경우에는 이를 유지하기가 매우 어려우며, 이 작업을 출판사가 할 것인지 아니면 아카이빙 기관이 할 것인지도 명확히 결정되고 이에 따른 역할과 책임이 기술되어야 한다.

<표 4> 전자저널의 보존대상 요소

구분	콘텐츠 요소	비고
보존 대상	<ul style="list-style-type: none"> <li>• 논문</li> <li>• 부록 / 보강 내용</li> <li>• 저자가 제공한 참조</li> <li>• 외부 자원과의 연결</li> <li>• 요약(초록)</li> <li>• Table of Contents</li> <li>• 보관되지 않은 객체에 대한 위치를 기록한 파일</li> <li>• 기타 편집 내용</li> </ul> <ul style="list-style-type: none"> <li>• 서지 정보</li> <li>• 편집위원회</li> <li>• 편집자</li> <li>• 토론 논지</li> <li>• 저작권 기술 및 정보</li> <li>• 편집 정책</li> <li>• 평론가 목록</li> <li>• 잡지 설명</li> <li>• 인쇄판의 표지 이미지</li> </ul>	<p>텍스트와 논문을 구성하는 그래프, 도형, 표, 사진 같은 부속파일</p> <p>데이터, 음향 및 영상 파일, 컴퓨터 프로그램 같은 논문에 관련된 디지털 객체를 저자가 출판사에 보낸 경우, 이 객체들은 대상에 포함</p> <p>연구, 보고서, 칼럼, 사설, 편지, 독자투고, 논평, 편집자에의 편지, 주해, 등</p> <p>논문과 기타 편집 내용을 기술한 메타데이터</p>
제외 대상	<ul style="list-style-type: none"> <li>• 저자에 대한 정보</li> <li>• 예약 구독 정보</li> <li>• 광고</li> <li>• 기타 비즈니스 정보</li> <li>• 추가 정보</li> </ul>	<p>저작권 양도 계약, 필사본의 준비 및 제출에 관한 지침 등</p> <p>재판번호 부여정보, 광고계제 정보, 연락처, 고객 서비스 정보 등</p> <p>경력 / 직업 정보 등</p>

8) <http://xml.coverpages.org/harvardEJournalArchive.html>



### 3. 3 전자저널 아카이빙을 위한 선결 과제

#### 3. 3. 1 기술적 측면

디지털 아카이빙은 정보기술과 긴밀히 연관되므로 기술의 변화에 민감하다. 전자저널은 정보기술을 사용하여 생성, 저장, 보관, 이용되므로 정보기술 환경이 제공되지 않으면 유통될 수 없다. 정보기술은 시간에 따라 변화, 발전하므로 장기적 보존을 위해서는 기술 노후화에 의한 이용의 위험이 발생하기 전에 이에 대한 대안 수립과 필요한 조치가 이루어져야 한다.

아카이빙에 필요한 여러 기능을 완전하고 지속적으로 수행하기 위해서는 보안, 인증 뿐 아니라 기술표준과 각 기능이 수행되는 기술환경에 대한 이해를 통해 기술적 위험을 사전에 파악할 필요가 있다.

출판업계에서는 전자저널에 대한 표준 형식이 정착되지 않아 다양한 파일 형식이 존재한다. 대부분의 전자저널은 PDF와 HTML 형식으로 동시에 제공되나, Postscript, TeX, DVI 등과 같은 형식으로 제공되는 경우도 있다. 이미지, 그림, 수식, 기호 등은 GIF, 또는 JPEG 파일로 제공된다. 일반적으로 부속 파일의 기술적 형식에 대해 통제가 잘 이루어지지 않으며, 저자가 어떤 파일 형식을 따르는 것이 좋은지에 대한 지침을 제공하고 있지 않다.

전자저널 아카이빙 관련 표준으로 EAD(Encoded Archival Description), OAI-PMH(Open Archives Initiative - Protocol for Metadata Harvesting), METS(Metadata Encoding and Transmission Standards) 등이 있다. EAD는

1993년 California, Berkeley 대학도서관 프로젝트를 통해 개발되었고 SGML 형식의 암호화 저장검색을 위한 표준이다. 참고로 이 표준은 미국의회도서관의 'Network Development and MARC Standards Office'에 의해서 관리되어지고 있다<sup>9)</sup>.

OAI-PMH는 메타데이터 수집용 프로토콜 표준으로서 웹 기반 학술지의 관리 및 운영의 e-Print 커뮤니티에서 성장하였고, 이들의 목적은 아카데미 커뮤니티에 빠르고, 자유롭게 접근하는 것을 만드는 것인데, 학술지의 출판과 참조를 위한 시간을 최소화하는 데 있다. OAI의 초기 목적은 e-Print 아카이브에 접근하기 위해서 참가하거나 연합할 수 있도록 상호운영 프레임워크를 개발하는 것이었고, OAI 구조는 리포지토리 안의 XML 형식의 메타데이터를 추출 또는 수집하기 위한 메커니즘을 정의하고 있다.

METS 스키마는 WWW 컨소시움의 XML 스키마 언어의 사용 표현 및 전자도서관내의 오브젝트에 관하여 구조적 메타데이터와 관리 그리고 표현의 암호화를 위한 표준이다. 미국의회도서관의 'Network Development and MARC Standards Office'에 의해서 관리되어지며, DLF(Digital Library Federation)에 의해서 개발되었다. 기타 전자저널 아카이빙에 필요한 표준으로 MARC, Dublin Core, Open URL 등에 대한 고려도 필요하다.

아카이브는 안전성이 보장되어야 한다. 저장, 보관된 정보가 안전하게 유지, 이용될 수 있도록 외부로부터의 침입에 대한 방지 및 유사시에 대비한 원문의 안정성이 확보되어야 한다. 또한 이

9) <http://www.lcweb.loc.gov/ead/>

용에 있어서도 권한의 인증 절차를 거치도록 하여 인가되지 않은 사용으로부터 보호하여야 한다.

인터넷 정보 통신에 대한 보안 문제와 저작권 문제가 중요한 이슈로 대두됨에 따라 디지털 콘텐츠의 보호 및 불법 방지를 위한 기술 개발이 점차 중요시되고 있다. 워터마킹은 콘텐츠를 보호하기 위해 특별한 형태의 워터마크를 감추고 추출하는 모든 기술적 방법으로 초기에는 저작물 자체에 은닉시키는 방법을 연구했지만, 현재는 마크의 인지, 강인성 제고, 삽입 및 검출 방식, 마크의 삽입 영역으로 분류하여 다양한 기술적 변환 방법을 이용한 강력한 워터마킹 기술이 개발되고 있다. 현재의 기술로는 두 가지 유형이 있다. 하나는 시스템 기반형으로서 디지털 콘텐츠 전체 유통 프로세스의 암호화, 네트워크, 정보관리 등 핵심 정보기술을 결합한 기술 유형이고, 다른 하나는 요소기술 기반형으로서 암호화, 워터마킹 등 요소기술을 활용한 기술 유형이다.

보호기술에 필요한 요소로는 다음의 여섯 가지를 들 수 있다. 1) content encryption : 암호화 알고리즘을 이용하여 콘텐츠를 암호화하여 사용자에게 안전하게 전달하고 재생시 복호되어 제공되며, 복호키는 인증된 사용자에게만 제공한다. 2) usage rule : 사용자의 요구 및 권한에 의한 콘텐츠 사용 제한, 콘텐츠 제공자의 다양한 마케팅 전략에 의한 다양한 콘텐츠 규칙을 적용한다. 3) persistent protection : 콘텐츠 이용은 전 유통과정을 통해 business rule의 조건에 따라 지속적으로 제어되어지고 사용권한으로 자체를 변경 및 복제, 그리고 사용권한 제어정보 변경으로부터 보호한다. 4) trusted environment : 저작권을 피해가기 위

해 DRM(Digital Rights Management)의 모듈 일부를 변경 및 대체로부터 보호하고 정상적인 제품임을 보증하는 인증 절차와 모듈이 손상되거나 대체되지 않았음을 증명하는 기술이다.

5) super distribution : 사용자가 자신이 구입한 콘텐츠를 이메일, CD-ROM, 디스켓 등을 통해 다른 사람에게 반복적으로 배포할 수 있게 하여 콘텐츠의 급속한 확산을 가능하게 하는 유통기술이다. 6) value-chain support : 생산자, 소유자, 배급자, 판매자가 콘텐츠의 전자상거래를 위해 복잡한 유통구조를 시스템화하기 위한 value chain 방식이 적용되는 기술 요소이다.

최근 디지털 콘텐츠의 중요성이 커지면서 DRM에 대한 관심이 더욱 커지고 있다. 현재 콘텐츠 DRM 시장은 콘텐츠 유료화 추세에 따른 기대 효과로 인해 과열되어 있는 양상을 보인다. 과거에는 콘텐츠 저작권 보호, 유통 솔루션 관련 업체들이 워터마킹 솔루션 업체와 DRM 관련 업체가 구별되었으나 워터마킹 업체들이 점차 DRM 분야로 사업 영역을 확장하고 있는 추세이다. 또한, 통합 솔루션 수요가 크게 증가함에 따라 업체들 역시 어느 한 제품이나 솔루션에 국한되는 것이 아니라 다양한 요소기술을 통합해 콘텐츠 보호 및 유통 등 전체 과정을 총괄 지원하는 형태를 지향하고 있다. 이 기술은 장차 전자저널 아카이빙 서비스를 제공하는 경우 제공 전자저널 콘텐츠가 적법한 사용범위를 벗어나지 않게 보호해 주는 장치로 반드시 필요한 기술이다.

전자저널 아카이빙에 있어서 가장 중요한 요소기술 중의 하나가 저장장치에 대한 기술이다. 대용량의 저장장치에 대한 저장 아키텍처에서 요구되는 기술 요소는 다음의 <표 5>와 같다.<sup>10)</sup>

〈표 5〉 저장 아키텍처에서 요구되는 기술 요소

기술 요소	내 용
High Availability	Storage 내부의 Power, Disk, 전원, Adaptor 등이 모두 이중화 또는 N+1 형태로 이중화되어 있으며, 무 중단 교체가 가능함
High Performance	Internal Storage 대비 Read와 Write시 성능을 최대한 발휘할 수 있도록 설계됨
High Scalability	Storage의 최대 확장성이 수십 tera byte 이상 장착 가능해야 하며, 최대 용량을 장착한 이후에도 성능에 지장이 없어야 함
Redundant Connections	Server와 Storage간의 연결이 N+1 경로로 이중화되어 단위 연결요소의 장애시에도 영향이 최소화 될 수 있어야 함
Scalable Performance	Peak Time시의 I/O 부하를 수용할 수 있는 충분한 I/O Path와 확장성을 제공해야 함
Non-disruptive backups	원본 데이터의 복제본을 이용하여 tape 백업을 수행할 수 있는 기능을 제공함으로써, 백업으로 인한 production 시스템의 성능을 최소화하고 down time을 “zero”화 할 수 있음
Rapid Restore	원본 데이터의 복제본을 이용하여 데이터의 손실이나 훼손시 수 시간이 소요되는 전통적인 tape 기반의 복구가 아닌 “복제본”을 이용하여 분 단위 복구를 수행할 수 있는 기능의 제공
Instant Test Environment	원본 데이터의 복사본을 이용하여 Application의 테스트, Stress 테스트, Upgrade 테스트 등을 수행할 수 있는 기능을 제공하며, 이것이 Production 시스템에 영향이 없어야 함
Business Continuance	Production 시스템에 최소한의 영향을 주면서 원격지에 데이터 복사본을 생성할 수 있도록 하며, 이 복사본 데이터를 이용하여 백업, 재난 복구, 테스트 용도로 사용할 수 있는 기능
Advanced Recovery	주 전산센터 장애시 원격지의 데이터를 이용하여 Business를 재개할 수 있는 기능

### 3. 3. 2 법적/제도적 측면

일반적으로 현재의 지적재산권 및 저작권법은 인쇄저작물이나 음반, 영상 등 과거의 기록 매체를 대상으로 입안된 것으로 디지털 환경의 저작물, 특히 전자저널의 생성에서 소멸까지 관련된 권리와 의무를 규정하기에는 적합치 않다.

전자저널은 저자와의 협약에 의해 출판사가 이를 제작하고 그 이용권(license)을 기관이나 개인에게 직접, 또는 대행사를 통해 제공함으로써 유통된다. 이러한 사용자 계약은 오래 전부터 시장에서 일반화되어 있지만, 디지털 아카이빙은 새로이 등장한 사업모델로서 전자저널의 장기적 보존을 위한 복제, 이전 활동과 저장된 전자저널의 용도를 명확히 규정하기에는 현행

지적재산권법과 저작권법은 적합하지 않다. 따라서 아카이브와 관련한 계약은 당해 계약서에서 아카이브된 전자저널의 용도와 계약 당사자간의 권리, 의무가 명확히 기술될 필요가 있다.

예를 들면, 아카이브 기관이 아카이빙을 하는데 있어 출판사는 어떤 의무를 가지는지, 아카이브 기관은 콘텐츠의 복제를 어떤 용도로 어떤 환경에서 몇 번이나 할 수 있으며, 새로운 기술 환경으로 콘텐츠를 이전했을 경우, 이전된 콘텐츠의 소유권과 이용권은 어떻게 바뀌는지, 콘텐츠의 이용에 따르는 의무는 무엇이며, 이 의무가 지켜지지 않았을 때, 이에 대한 책임은 누가 어떻게 지는지, 등 가능한 여러 경우에 대해 명확한 언급이 있어야 한다. 아카이브 기관

10) 한국과학기술원. 2004. NDAC 정보전략계획 최종보고서

의 활동의 합법성이 계약에 의해 성립된다는 점을 감안하면, 계약서 상에 쌍방간의 권리와 의무가 상세히 기술되어 모호한 점이 없도록 해야 할 것이다.

영국처럼 법률적으로 전자저널의 납본을 강제하는 것은 아카이브에 소극적인 출판사에게 최소한의 협력을 하게 하는 효과가 있다. 그러나 자국의 출판사가 아닌 해외 출판사의 전자저널을 강제적 납본(legal deposit)의 범주에 두기에는 어려움이 많다.

전자자료의 액세스 통제와 저작권 침해 우려 문제를 해결하기 위해 개발되는 기술로 인해 전자자료의 보존활동이 방해받을 수 있다. 따라서 저작권 당사자와의 협상과 보존문제가 선결되어야 한다.

아울러, 전자저널 구독자는 출판사가 해당 전자저널에 대한 아카이빙과 보존을 고려하고 있는지 걱정이 된다. 왜냐하면 이미 라이선스를 샀던 저널에 대한 영속적 액세스가 불가능하게 되는 경우의 문제에 대한 제도적 장치가 아직까지 마련되지 않았기 때문이다.

마지막으로 프라이버시와 비밀보장 문제를 들 수 있다. 디지털 아카이브내에서 관리되는 정보는 정보보호법이나 이와 유사한 프라이버시 관련 법에 의해 규제될 수 있다. 또한 경우에 따라 비밀보장에 관한 협약에 따라 아카이빙되고 있는 정보가 규제받을 수 있다. 이러한 프라이버시 및 비밀보장 관련사항들이 아카이브내 정보관리 방안과 외부 액세스 제공 요령에 미치는 영향을 고려해야 한다.

### 3. 3. 3 운영 측면

전자저널 아카이브 구축을 위해 기술적, 법적 문제 이외에 운영상의 문제도 극복해야 할 과제중의 하나이다.

현실적으로 가장 중요한 문제중의 하나가 비용문제이다. 기본적으로 비용은 정보의 양과 액세스 수준에 따라 좌우되는 것이 원칙이다. 현재로서는 전자저널 아카이빙을 위한 정확한 비용 산출을 할 수 있는 표준 비용모델이 없다.

그러나 인쇄저널보다 전자저널의 아카이빙 비용이 훨씬 더 든다는 가정은 다음과 같은 네 가지 이유에서 생각할 수 있다<sup>11)</sup>. 첫째, 일정한 주기로 끊임없이 기술변화를 능동적으로 관리해야 할 필요성, 둘째, 정보자원 자체와 출판사와의 라이선스 계약 모두가 표준화되지 않은 관계로 규모의 경제를 실현하기 어렵게 만든다는 점, 셋째, 기술변화 이후에도 정보 본질의 손실이 없도록 전자자료를 안정적이고 정밀하게 랜더링해주는 방법이 개발되지 않은 점, 넷째, 특별한 비용절감 방안이 개발되지 않는다면 디지털 보존비용이 전통적 콜렉션 비용보다 더 많은 비용을 요구할 것이라는 점 때문에 그러한 가정이 성립될 수 있다.

전통적 자료와 디지털 자료 모두를 책임지고 있는 납본도서관 같은 경우에는 두 개의 상반된 요구사항 사이에서 더욱 난감한 처지에 빠질 수 있다. 이러한 기관들은 단기 액세스 지원 보다도 장기 보존에 더 중점을 두게 될 것이다. 따라서 장기적으로는 비용 공동 분담 모델에 대한 범위가 설정되고 이러한 모델이 비용 대비 효과 면에서 가장 바람직한 형태로 판명될 것이다.

11) <http://www.dpconline.org/graphics/index.html>

디지털 아카이빙과 보존에 대한 연구는 1990년대 중반부터 시작되었기 때문에 아직까지 이 분야 직무에 대한 전문성이 확립되지 않았다. 게다가 끊임없이 변화하는 기술에 부응하여 직무를 수행할 수 있는 전문 직원을 육성하는 일 자체가 어려운 일일 뿐만 아니라 이 일이 중요한 일임에도 불구하고 이에 대한 교육 프로그램조차 없다는 점도 간과할 수 없는 문제이다.

특정 주제분야나 특정 포맷별로 아카이빙 책임을 떠 맡아 추진하는 사례를 찾아 볼 수 있다<sup>2)</sup>. 영국의 경우 'ADHS(Arts and Humanities Data Service)'가 사회과학과 인문학 연구 정보 아카이빙에 대한 책임을 맡고 있는 반면에 'National Sound Archive'는 'sound recordings'에 대한 아카이브 책임을 맡고 있다. 게다가 디지털 보존 사업에 대한 기관간 협력모델을 수립하여 진행하고 있는 나라들도 있다.

기관내에서 또는 기관간의 역할 분담도 변할 수 있기 때문에 디지털 자료의 보존에 대한 책임을 부여하는 일은 조직내 여러 다른 부서들의 참여가 필연적으로 요구된다.

## 4. 전자저널 아카이브 구축을 위한 전략 및 제안

### 4.1 구축 모델링과 콘텐츠 전략

현재 운영되고 있는 해외 전자저널에 대한 아카이브 구축모형으로는 크게 'Safe Place' 모형과 'LOCKSS' 모형을 들 수 있다(Drimmelen,

2004). 'Safe Place' 모형은 집중형 모형으로서 국립도서관이나 국가정보센터가 구축 주체가 되는 모형이다. 따라서 공공이익을 추구하기 위해 적합한 모형이며 장기 보존을 본래의 목적으로 하고 있다. 보안이나 정치적인 이유로 단일 아카이브 보다는 복수개의 아카이브 구축을 권고하고 있는 이 모형은 네덜란드 국립도서관이 대표적인 예이다. 반면에 'LOCKSS' 모형은 분산형 모형으로서 개별기관이 구축 주체가 되는 모형으로서 미래의 액세스 문제해결에 주안점을 두고 있다. 'Safe Place' 모형보다 초기비용 면에서 강점이 있으나 장기 보존 문제 해결을 위한 고려가 미흡하다는 단점이 있다.

우리나라의 경우, 개별기관은 해외 전자저널을 거의 대부분 KESLI 콘소시엄을 통해 구입하고 있다. 따라서 경제적, 기술적, 사업적 측면에서 볼 때, 대부분의 전자저널 구입을 국가 콘소시엄에 의존하고 있는 우리나라와 같은 환경 하에서는 'Safe Place' 모형이 더 효과적이라고 판단된다.

과거에 인쇄저널 구독의 경우에 개별 기관별로 자체 아카이브를 가지는 것이 자연스러운 현상이었으나, 전자저널의 경우에는 저널 구독료 이외에 디지털 아카이브 구축과 운영을 위한 추가적인 경제적 부담과 극복해야 할 기술적, 제도적 과제가 단일 기관으로서는 감당하기 벅찬 수준이 될 것이기 때문에 국가적 차원이나 공공성을 목적으로 하는 기관이 이를 담당하여 수행할 필요가 있다. 다시 말하면, 전자저널의 아카이빙은 학술연구의 중요한 원천이기 때문에 국가 전체를 대표하여 학술 연구를 뒷받침할 수

12) <http://www.dpconline.org/graphics/digpres/presissues.html>

있는 기관이 아카이빙 주체가 되어야 한다.

또한, 아카이빙 주체는 전자저널의 원문을 소유하고 있는 해외 출판사들과의 원만한 관계를 유지하고, 아카이빙 자료를 제공받기 위한 협상을 원활하게 수행할 수 있는 협상능력을 가지고 있어야 하며, 해외 출판사들로부터 받은 전자저널 원문 데이터를 손상 없이 영구히 보존하기 위한 시스템을 구축하고 이를 효율적으로 운영할 수 있어야 한다. 이를 위해 정보기술의 생명 주기가 짧으므로 이를 지속적으로 조망할 수 있는 기술조망 프로세스를 수립해야 하며, 이를 통해 필요한 경우 데이터를 마이그레이션하고 그 주기를 면밀히 관찰하여야 한다.

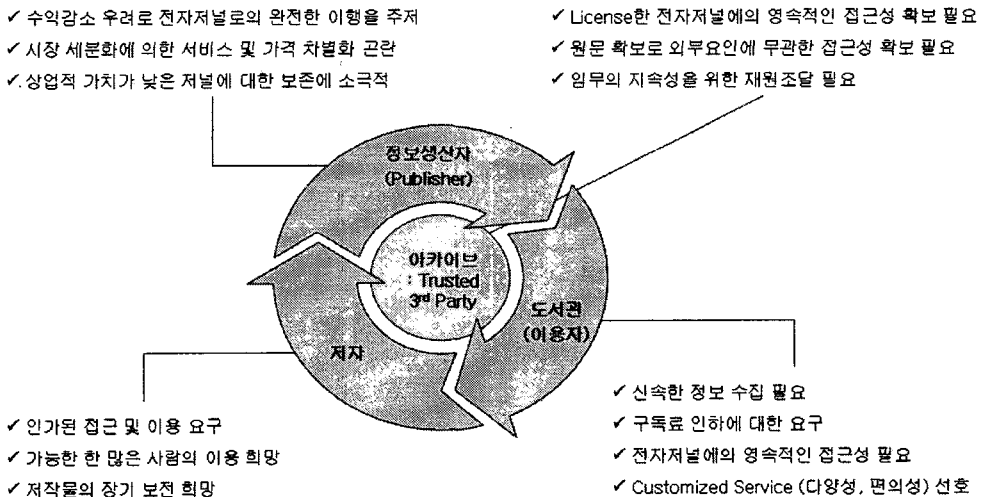
전자저널의 아카이빙을 위해서는 먼저 학술정보의 생성에서 소멸까지 관련된 주요 이해 당사자인 저자, 출판사, 아카이빙 주체, 도서관(이용자)의 이해관계를 고려하여야 한다.

출판사는 상업적 가치를 우선으로 하기 때문

에 그 가치가 낮다고 판단되는 전자저널의 아카이빙을 소홀히 할 수 있는데 반하여 아카이빙 주체는 외부 요인에 무관하게 전자저널의 영속적 액세스를 확보하는데 더 큰 관심을 가진다. 전자저널 구매자인 도서관은 과거 인쇄저널 구매시와 같이 한 번 구매한 전자저널에 대해서는 영구적 액세스를 보장받기를 원한다(그림 2 참조).

전자저널 제공자에 따라 아카이빙에 대한 정책과 조건이 상이하므로 전자저널 제공자와의 계약내용에 따라 달라질 수 있지만, 어떤 전자저널을 우선적으로 아카이빙해야 할 것인가에 대한 결정은 아카이빙 활동의 시발점이 된다.

대상 저널별로 아카이빙 우선순위를 도출하기 위해 아래 <표 6>와 같이 네 가지 평가기준과 각각에 대한 측정수단을 제시하고자 한다. '활용의 영속성', '콘텐츠 중요도', '이용량 규모', '투자비용'을 우선순위 평가기준으로 정하고, '저널 반감기', 'SCI Impact Factor', '이용 빈



<그림 2> 전자저널 아카이빙 이해관계 분석표

〈표 6〉 아카이빙 우선순위 도출을 위한 저널별 평가기준

평가 기준	측정 수단	비 고
활용의 영속성	반감기	반감기가 긴 저널은 Aging 대비 수요 탄력성이 낮으므로 장기 보존이 필요함.
콘텐츠 중요도	SCI Impact Factor	SCI Impact Factor가 큰 저널은 연구 활동에의 기여도가 높다고 추론함.
이용량 규모	이용 빈도수	이용 빈도수가 많은 저널은 해당 이용자가 많은 것이므로 수혜 범위가 큼.
기 투자 비용	저널 당 라이선스 비용	KESLI 해당 컨소시엄의 라이선스 비용을 저널수로 나눈 금액

도수', '저널 당 라이선스 비용'을 각각의 측정 수단으로 정하여 전자저널 아카이빙 우선순위를 산정하였다.

전자저널의 최근호와 과월호에 대한 아카이빙을 모두 하도록 계약이 된 경우, 무엇을 먼저 하는 것이 바람직한지 고려할 필요가 있다. 최근호와 과월호 중 어느 것을 먼저 아카이빙할 것인가는 중요한 논점은 아니지만, 제한된 시간 내에서 어느 작업을 먼저 할 것인가는 아카이브 용도에 따라 결정되어야 할 것이다. 일반적으로 고려하여야 할 요소는 보존의 시급성과 활용의 중요성이다. 아카이브 용도가 Dark Archive<sup>13)</sup>만 허용된다면 오래된 전자저널부터 보존할 필요가 있는데, 이는 정보생산자가 상업적 가치가 낮은 자료는 아카이빙을 소홀히 할 수 있기 때문이다. 반대로 Light Archive<sup>14)</sup>가 허용되어 국내 이용자를 대상으로 직접 전자원문 서비스를 할 수 있는 경우는 활용도가 높은 최근의 전자저널부터 아카이빙할 필요가 있다. 후

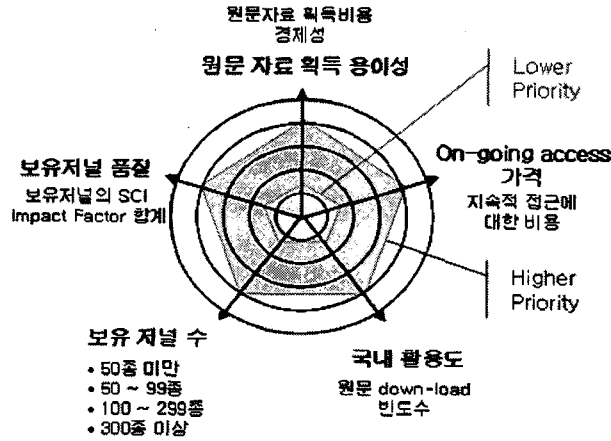
자의 경우라도 보존에 대한 시급성이 요구되는 전자저널이 있다면 그것부터 가장 먼저 수집, 저장, 보존해야 하는 것은 당연하다.

전자저널 제공자(출판사, 학협회, 서비스 제공자 등)는 전 세계적으로 많을 뿐만 아니라 지속적으로 그 수가 증가하고 있다. 따라서 전자저널 아카이빙에 투여되는 자원의 한계와 제공자 측의 정책에 따라 아카이빙 협상대상 출판사의 우선순위를 결정하는 일이 중요하다. 우선협상 대상 출판사의 판단 기준으로는 1) 원문 획득 용이성, 2) 지속적 접근에 대한 추가비용 여부, 3) 저널 활용도, 4) 보유 저널 수, 5) 보유 저널 품질 등 〈그림 3〉의 다섯 가지 판단기준을 가지고 협상대상 출판사의 우선순위를 정해야 한다.

협상전략으로는 첫째, 출판사에게 줄 수 있는 혜택과 아카이빙 주체가 얻을 수 있는 이익을 준비한다. 둘째, 출판사에게 신뢰를 줄 수 있어야 한다. 필요한 경우 출판사의 지적 재산권

13) 사용자의 액세스를 제한하는 아카이브 형태로서 관리자나 허가된 자 이외는 액세스 할 수 없다. dark archive의 목적은 재난이나 불가피한 서비스 중단에 대비해 만들어진 정보 저장소임

14) 사용자의 액세스를 제한하지 않는 아카이브 형태로서 정보제공자와의 협약에 따라 아카이브된 데이터를 가지고 부가적인 서비스도 가능함



〈그림 3〉 아카이빙 협상대상 출판사의 우선순위 판단기준

보호를 위해 필요한 DRM과 같은 기술 투자를 명시해야 한다. 셋째, Dark Archive 계약시에는 한정된 용도의 이용을 통해 보관된 콘텐츠의 무결성과 이용성을 점검할 수 있도록 해야 한다. 넷째, 아카이빙 데이터에 대한 대가를 지불해야 하는 경우에는 이의 활용도와 비례하도록 비용산출 방식을 고안해야 한다.

전자저널 콘텐츠를 선별적으로 아카이빙 할 수 밖에 없는 이유는 첫째, 전자저널 콘텐츠가 대량의 정보이고, 동시에 지속적으로 양적 팽창이 이루어지고 있다는 점, 둘째, 전자저널의 품질과 이용 빈도가 아주 다양하다는 점, 셋째, 아카이빙을 위해 투입되는 자원의 한계성 때문에 선별적으로 아카이빙을 추진하는 것이 필연이다.

아카이브 대상 콘텐츠의 선정은 콘텐츠의 장기보존 및 이용 가치에 따라 결정되어야 하며, 수집된 콘텐츠는 무결성 검증 과정을 거쳐 저장되어야 한다. 콘텐츠 선정 기준은 아래와 같다:

- 콘텐츠 선정시 이용의 시급성 보다 장기적 보존 및 이용성에 우선순위를 둔다.

- 보존 여부에 대한 결정은 수집시 이루어져야 한다.
- 출판사로부터 수집하는 콘텐츠의 저장 형식은 수집 시점에 필요한 보존 문제에 한정하여 결정한다.
- 아카이빙 주체별로 누가, 어떤 콘텐츠를, 언제까지 아카이빙 할 것인지를 먼저 결정하고 그에 따른 콘텐츠 선정을 해야 한다.
- 콘텐츠 수집시 저작권과 지적소유권에 대한 내용을 검토하여 권리와 의무를 명확히 한다.
- 모든 선정 프로세스에 대한 다큐멘테이션을 유지해야 한다.

참고로 수서 단계에서 적용되는 콘텐츠 선정 정책의 예로 Berkeley 대학 Digital Library에서는 Archived; Served; Mirrored; Linked의 네 가지 단계<sup>15)</sup>를 채택하고 있고, AHDS<sup>16)</sup>는 Archived; Served; Brokered; Linked; Finding Aids의 다섯 단계를 채택하고 있으며



캐나다 국립도서관<sup>17)</sup> Archived: Served: Linked의 세 단계 정책을 유지하고 있다.

#### 4. 2 기술전략

전자저널 아카이빙은 아카이브 구축 업무와 디지털 보존 업무로 분리하여 접근할 필요가 있다. 먼저 아카이브 구축을 위한 기술적 전략에 대해 살펴보기로 한다.

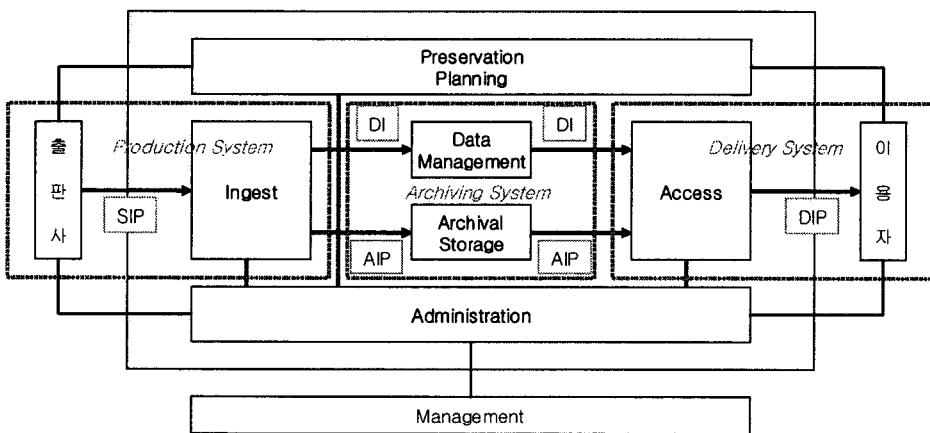
##### 4. 2. 1 아카이빙 시스템 모델

〈그림 4〉의 아카이빙 시스템 구축 모형은 ISO14721:2002 표준인 OAIS 모델을 근간으로 하여 각 기능에 대한 정의, 프로세스간의 연관관계 그리고 입력, 프로세스, 출력에 대한 관

리와 도큐멘테이션에 대해 기술하고 있다.

현재 운영되고 있는 네덜란드 국립도서관의 DIAS나 영국의 CEDARS, 미국 OCLC ECO의 아카이빙 시스템은 모두 OAIS를 기반으로 하는 시스템 아키텍처를 갖추고 있다. 따라서, 전자저널 아카이빙 시스템 구축시 필요한 기능 모델은 반드시 OAIS 표준을 채택하는 것이 바람직하다.

전자저널 아카이빙을 위한 시스템 구성은 〈그림 5〉과 같이 OAIS의 여섯 개 영역으로 구분하여 구성하는 것이 효과적이고, 그 기능별 특성을 표현하기 위해서 각각의 시스템을 표시하는 것이 좋다. 이 시스템들은 어플리케이션 특성 및 부하에 따라 하나의 시스템에서 통합 운영될 수 있도록 설계하는 것이 바람직하다.



DI : Descriptive Information  
 SIP : Submission Information Package  
 AIP : Archival Information Package  
 DIP : Dissemination Information Package

〈그림 4〉 OAIS 참조모형

15) <http://europa.eu.int/ISPO/dlm/documents/gdlines.pdf>

16) <http://heds.herts.ac.uk/resources/papers/drh99.pdf>

17) <http://www.nationalarchives.gov.uk/recordsmanagement/selection/appraisal.htm>

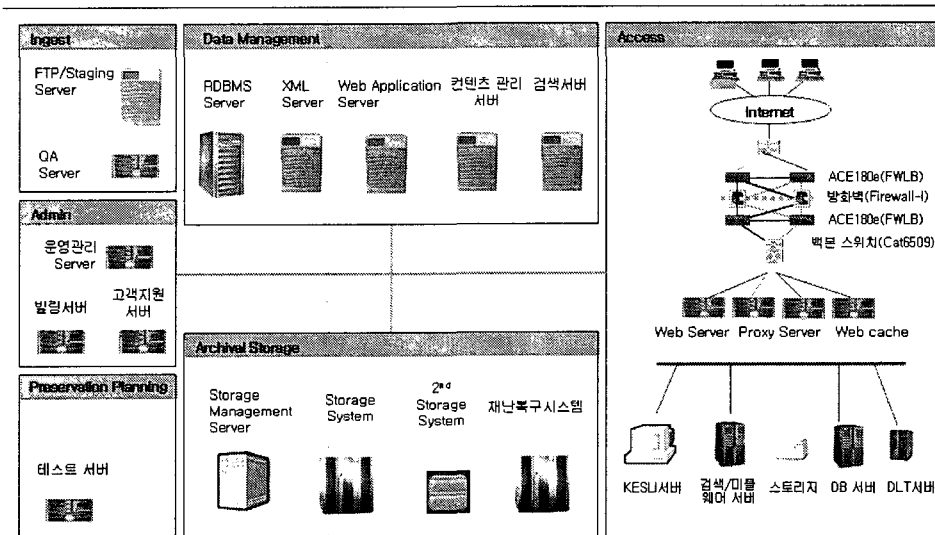
아카이빙 시스템에서 어플리케이션 개발은 주로 출판사로부터 데이터를 받는 부분과 사용자에게 서비스의 제공과 요청을 담당하는 부분이며, 전자저널 데이터에 대한 생성, 입수에서부터 내부의 표준에 입각한 저장과 활용 그리고 최종 서비스에 대한 모든 부분을 콘텐츠 관리

소프트웨어가 담당하고 있다(그림 6 참조).

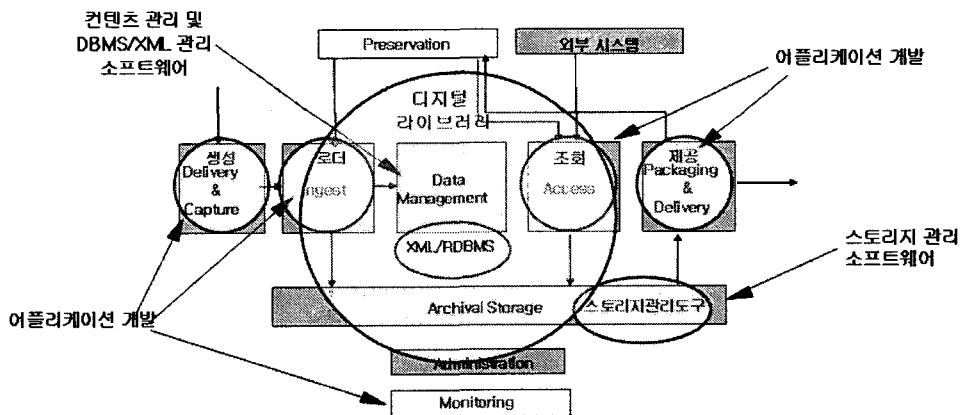
4. 2. 2 디지털 보존 전략

전자저널의 디지털 보존을 위한 기술전략으로는 아래와 같이 크게 세 가지 전략이 있다<sup>18)</sup>.

- 기술보존(technology preservation) 전략:



<그림 5> 아카이빙 시스템 구성(안)



<그림 6> 소프트웨어 구성(안)

18) <http://www.ukoln.ac.uk/interop-focus/gpg/Preservation/>

정보를 담고 있던 소프트웨어 및 하드웨어를 그대로 보존하는 방법

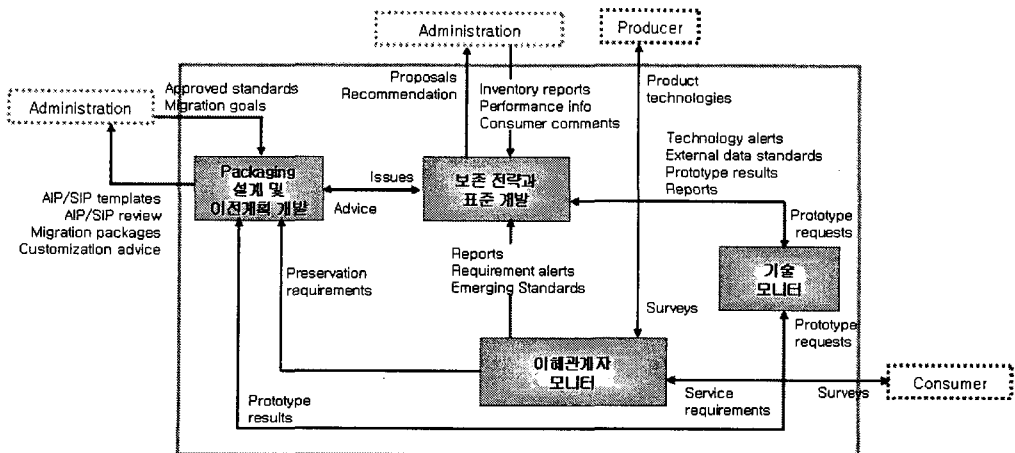
- 기술 에뮬레이션(technology emulation) 전략: 기존의 프로그램을 대신할 강력한 새 프로그램을 개발하여 디지털 정보를 보존하는 방법
- 디지털 정보 마이그레이션(digital information migration) 전략: 기존의 정보를 새로운 포맷으로 다시 인코딩 하는 방법

이러한 디지털 보존을 위한 기술전략이 정상적으로 작동되기 위해서는 정보시스템 전략으로서 표준화와 기술변화에 대한 기술조망 프로세스를 가지고 있어야 한다. 아울러, 기술변화와 H/W, S/W의 대체 및 유지를 위한 체크리스트와 함께 사라진 포맷으로 된 데이터 액세스를 찾아내기 위한 '디지털 고고학(digital archaeology)'도 고려해야 한다.

보존 계획(Preservation Planning) 프로세스는 OAIS 환경을 감독하고 이용자 커뮤니티

의 장기적 접근이 가능하도록 하는 권고 안 또는 계획을 제안하는 기능과 최근의 기술변화를 점검하고 새롭게 출현하는 디지털 기술 정보표준, 컴퓨팅 플랫폼등을 점검하는 기능을 담당한다. 또 정보 생산자, 이용자 커뮤니티, 정보기술에 대한 최근 경향과 변화를 조사 파악하여 현 보존시스템을 수정하고 업그레이드한다. 또, 향후 중장기적으로 적절한 보존전략과 표준을 제안하고 새로운 정보패키지를 디자인하며, 상세한 이전계획과 프로토타입(prototype) 개발을 수행한다(그림 7 참조).

과거 인쇄저널 구독은 개별기관별로 이루어졌고, 콘텐츠가 물리적으로 각 도서관에 수집되어 서가에 배치됨으로써 도서관은 컬렉션의 소유자로서 자연스럽게 저널의 관리와 보존을 수행해 왔다. 그러나, 전자저널이 등장하면서 구매 패턴도 컨소시엄 기반의 라이선스 공동구매 형태로 바뀌었고 '정보의 소유자'에서 '정보의 액세스 지원자'로 바뀌면서 라이선스한 전자저널의 영속적 액세스에 대한 보장요구가 커지게



〈그림 7〉 보존계획 프로세스(안)

되었으며, 자연스럽게 디지털 아카이빙 구축의 필요성이 크게 대두되었다.

그러나, 전자저널 아카이빙은 인쇄저널과 달리 개별 도서관에서 추진할 수 있는 일이 아니다. 따라서 국립도서관이나 컨소시엄 센터에서의 전자저널 아카이빙을 추진해야 한다.

전자저널의 아카이빙 구축과 디지털 보존을 위해서는 관련 표준의 연구와 적용, 콘텐츠 보호기술, 콘텐츠 관리 및 보존시스템 구축 등 기술적인 문제와 지적재산권 및 저작권 문제, 그리고 비용조달과 전문가 육성, 아카이빙 역할분담 문제 등 제도 및 운영차원의 과제들을 해결해야 한다.

이러한 문제점들을 직시하고 전자저널의 아카이빙과 장기적 보존을 위해 우리 상황에 가장 적합한 전략과 제안을 종합하면 다음과 같다.

- 해외 전자저널의 디지털 아카이빙 및 보존 문제는 개별기관의 문제가 아니라 국가적 과제이므로 국가적 차원에서 추진되어야 한다.
- 전자저널 공동구매 컨소시엄 센터가 아카이빙의 추진 주체가 되어야 한다.
- 아카이빙 구축 모형은 KB 추진 모형과 같은 'Safe Place' 모형으로 구축한다.
- 디지털 아카이브 가치에 대한 평가체계가 반드시 필요하다.
- 전자저널의 디지털 보존문제는 포맷이나 기술적인 문제보다도 비즈니스적 관련성이 더 많으므로 이에 대한 고려가 선행되어야 한다.
- 해외 전자정보의 안정적이고 지속적인 이용을 위해서 디지털 아카이빙 및 보존에 대한 표준이나 국가적 전략 수립이 시급히

시행되어야 한다.

- 동일한 콘텐츠의 중복 보존작업을 피하고 디지털 보존의 효율성을 제고하기 위해 보존 주체별 분담과 협력을 가능케 하는 국가적 보존 프로그램이 필요하다.
- 디지털 보존 프로세스는 능동적이지 않으면 성공할 수 없기 때문에 기술이 변화함에 따라 기술추이를 모니터링하여 주기적으로 디지털 아카이빙 기술환경의 변경 필요성이 있는지 검토할 수 있는 기술조망 프로세스를 가져야 한다.
- 디지털 보존을 위해서 기술보존 전략, 기술 에플리케이션 전략, 디지털 정보 마이그레이션 전략의 세 가지를 적절히 사용한다.
- 시간을 초월한 보편적 액세스(universal access)가 가능하려면 디지털 보존에 대한 영구적인 연구개발 노력이 필요하다.
- 아카이빙 시스템 아키텍처는 ISO 표준인 OAIS 참조모형을 따른다.
- 시스템 환경에 대한 형상관리를 하여야 한다.
- 저장 및 보존을 위한 전자저널의 파일 형식에 대한 지침을 출판사에 제공하여 파일 형식을 통제할 수 있도록 한다.
  - Text Document : PDF, HTML
  - Metadata : XML
  - Color Images : GIF / JPEG 등
- 보존 메타데이터(preservation metadata)와 원문정보(content information)는 하나의 파일에 저장하지 않고 분리하여 관리한다.
- 논문의 외부 링크 연계는 자체적으로 구현하지 않고 출판사에서 주기적으로 보내주

는 링크 정보에 의해 구현한다.

- 업무 처리용 서버는 분산형으로, 데이터베이스 서버는 집중형으로 구성하며, 저장 장치는 메모리 디스크, 디스크, 테잎의 세 단계로 분산하여 이용 빈도에 따라 데이터를 저장한다.
- 아카이브 용도의 원문 데이터는 무결성이 검증된 수집 본, 보존 본, 이용 본의 세 가지로 보관하되 보존 본과 이용 본은 물리적으로 분리하여 보존 본은 원격지에 저장 관리한다.
- 기술인력 조달 및 육성 방안의 하나로 해외의 아카이빙 기관과 제휴관계를 맺어 아카이빙 관련 기술을 흡수한다.
- 주기적으로 아카이빙 서비스를 일정 테스트 도메인에 제공하고 이에 대한 평가와 피드백을 실행한다.
- 아카이빙 우선순위 도출을 위한 저널별 평가기준과 우선협상 대상 출판사 선정을 위한 판단기준을 수립해야 한다.
- 전자저널 아카이빙의 관건은 대상자료의 수집에 있다. 자료의 소유자인 출판사와의 협상이 가장 중요하기 때문에 출판사별로 가장 합리적인 협상전략을 개발하여야 한다.

## 5. 결 론

우리나라의 경우 전자저널에 투입되는 비용이 연간 100억원을 훨씬 상회하고 그 증가율이

매년 약 20% 정도로 투자비용 비중이 높아가고 있다. 그러나 일부 도서관에서 인쇄저널의 중복구독을 통해 전자저널 아카이빙의 방편으로 삼는 경우가 있지만 전자저널 아카이빙에 대한 근본적인 대책이 전혀 마련되어 있지 않은 실정이다. 또 전자출판물은 인쇄출판물과 달리 단순히 텍스트와 그래픽으로만 출판되는 것이 아니라 여러 다른 매체를 결합시키는 복잡한 데이터 세트로 출판되기도 하며, 다양한 링크를 통해 외부 소스와 연계되는 형태로 출판되기도 한다. 나아가서 전자출판물은 상호반응하는 오디오 및 비디오 컴포넌트, 운영용 스프레드시트, 데이터베이스 프로그램, 그리고 특정 소프트웨어 컴파일용 프로그램 등을 수록하기도 한다. 장래에는 이러한 전자출판물에 냄새나 촉각에 의해 반응하는 에뮬레이터와 같은 첨단 장치 등도 포함될 것이다. 따라서 전자출판물에 대한 기술이 발전하면 할수록 디지털 아카이빙 및 보존활동이 그만큼 어려워질 것이다. 앞서 기술한 바와 같이 전자저널의 아카이빙은 인쇄저널과 달리 개별기관이 독자적으로 수행할 수 없을 뿐만 아니라 그렇게 하는 것은 국가 차원에서 비효율적이고, 자원낭비이다.

따라서 외국의 경우처럼 전자저널의 디지털 아카이브 구축은 정부의 정책적 지원하에 국가적 차원에서 수행해야 하며, 이를 위해서 필요한 법적 제도적 장치를 마련해야 하고, 아카이빙 시스템 구축과 대상 콘텐츠의 장기 보존을 위한 주도면밀한 기술전략 수립과 심층연구가 필요하다.

## 참 고 문 헌

- 김희정. 2003. 전자저널 아카이빙을 위한 OAIS 참조모형의 적용방안에 관한 연구, 연세대 대학원 석사학위논문.
- 서은경. 2003. 디지털 아카이브의 영구적 보존을 위한 개념적 모형 설계에 관한 연구. 『한국문헌정보학회지』, 38(1): 13-34.
- 신은자. 2001. 전자저널의 아카이빙에 관한 연구, 『정보관리학회지』, 18(3): 139-158.
- 윤중현. 2002. 정보자원 활용도 영향요인들에 관한 연구. 『한국정보관리학회지』, 19(2): 31-48.
- 이란주, 황신혜. 2003. 전자저널의 관리 및 서비스에 관한 연구, 『도서관』, 58(2): 3-40.
- 이소연. 2002. 디지털 아카이빙의 표준화와 OAIS 참조모형, 『정보관리연구』, 33(3): 45-68.
- 황옥경. 2004. 전자저널 라이선스 계약모델에 관한 연구, 『한국문헌정보학회지』, 38(1): 199-228.
- 한국과학기술원. 2004. NDAC 정보전략계획 최종보고서.
- 한혜영. 2004. 해외학술 구독현황 분석 및 주요 시사점. 『국회도서관보』, 2004년 4월호: 40-49.  
 <<http://ssdoo.gsfc.nasa.gov/nost/wwwclassic/documents/pdf/CCSDS-650.0-B-1.pdf>>
- CEDARS Project Team. 2001. "The CEDARS Project Report: April 1998-March 2001".
- CEDARS. 2001 "CEDARS Guide to Preservation Metadata, CEDARS Project".
- Commission on Preservation and Access & The Research Libraries Group. 1996. "Preserving Digital Information: Report of the Task Force on Archiving of Digital Information."
- Consultative Committee for Space Data System(CCSDS) 2002. "Reference Model for an Open Archival Information System."
- DLM Forum. 1997. Guidelines on the Best Practices for Using Electronic Information.  
 <<http://europa.eu.int/ISPO/dlm/documents/gdlines.pdf>>
- E-Journal Archive DTD Feasibility Study, IneraTM Incorporated for Harvard Univ. Library, 2001.
- Good Practice Guide for Developers of Cultural Heritage Web Services.  
 (<<http://www.ukoln.ac.uk/interop-focus/gpg/Preservation/>>)
- Harvard University Mellon Project Team, 2002. "Report on the Planning Year Grant For the Design of an E-journal Archive".  
 <<http://xml.coverpages.org/harvardEJournalArchive.html>>
- Lupovici, C., and J.Masanés. 2000. "Metadata for the Long Term Preservation of Electronic Publications: NEDLIB report".
- Lyon, Liz. 2003. "eBank UK: Building the

- links between research data, scholarly communication and learning”, *Ariadne*, 36.  
 <<http://www.ariadne.ac.uk/issue36/lyon/>>
- Mark Bide & Associates. 2000. “Standards for Electronic Publishing: an Overview”, NEDLIB.
- NLA, 1999. “Logical Data Model”.  
 <<http://www.nla.gov.au/dsp/rft/model.doc>>
- NLA, 2002. “Electronic Information Resource Strategies & Action Plan 2002-2003.”  
 <<http://www.nla.gov.au/policy/electronic/resourceplan2002.html>>
- OCLC. 2003 “OCLC Environmental Scan: Pattern Recognition.”  
 <<http://www.oclc.org/membership/escan/default.htm>>
- OCLC/RLG Working Group on Preservation Metadata. 2002. “Preservation Metadata and the OAIS Information Model: A Metadata Framework to Support the Preservation of Digital Objects.”  
 <<http://www.oclc.org/research/pmwg/>>
- Owen, J.M., M.Coutts, and S.Veling. 2001. CEDARS Summative Evaluation: Final Report.
- Pandora Business Process Model,  
 <<http://pandora.nla.gov.au/#obje>>
- Proposal for a Study of Electronic Journal Archiving for Andrew W. Mellon Foundation, 2000.
- RLG, 2002. “Trusted Digital Repositories: Attributes and Responsibility”.
- Steenbakkers, Johan F. 2004. Archiving and preserving of e-journals: The 5th KESLI Annual Meeting, Kyeongju, Korea, September.
- van Drimmelen, Wim. 2004. “Universal Access Through Time: Archiving Strategies for Digital Publications”. *Libri*, 54(2): 98-103.
- YEA: The Yale Electronic Archive, One Year of Progress Report on the Digital Preservation Planning Project, 2002.  
 <<http://www.library.yale.edu/~oke/rson/yea/frontmatter.pdf>>  
 <<http://www.dpconline.org/graphics/index.html>>  
 <<http://www.dpconline.org/graphics/digpres/presissues.html>>  
 <<http://heds.herts.ac.uk/resources/papers/drh99.pdf>>  
 <<http://www.loc.gov/standards/mets/>>  
 <<http://www.openarchives.org>>  
 <<http://europa.eu.int/ISPO/dlm/documents/gdlines.pdf>>  
 <<http://www.curl.ac.uk/projects/elib.pdf>>  
 <<http://www.lcweb.loc.gov/ead/>>  
 <<http://www.nationalarchives.gov.uk/recordsmanagement/selection/appraisal.htm>>  
 <<http://www.oclc.org/electroniccollections/default.htm>>  
 <<http://www.ithaka.org/e-archive/>>