

디지털 보존의 비용요소에 관한 연구

A Study on Costs of Digital Preservation

정혜경(Hye-Kyung Chung)*

초록

디지털 자료에의 장기적인 접근을 보장하기 위해서는 체계적인 디지털 보존이 이루어져야 하며, 디지털 보존에 대한 지속적인 예산 지원에는 비용요소에 대한 구체적인 파악이 선행되어야 한다. 본 연구는 기존 관련연구를 토대로 디지털 보존의 필수 비용요소들을 추출하였으며, 이를 보존과의 관련성 정도에 따라 직접보존비용과 간접보존비용으로 범주화하였다. 디지털화 사업이 대규모로 진행되고 있는 국내 도서관을 사례분석 대상으로 선정하여 디지털 보존 현황을 분석하고 보존비용을 추정하였다. 사례분석 결과, 기본적인 디지털 보존은 이루어지고 있으나 디지털 자료에의 장기적인 접근을 보장하기 위한 보다 포괄적이고 체계적인 보존기능이 미흡한 것으로 나타났다. 현재 사례분석 대상 도서관에서 수행되고 있는 디지털 보존을 토대로 분석했을 때, 도서관 A와 B에서는 각각 원문DB 구축비용의 11.8%와 8.6%에 달하는 디지털 보존비용이 매년 투입되어야 하는 것으로 추정되었다. 그러나 본 연구의 결과는 매우 보수적인 추정치라고 할 수 있다. 이는 대상 도서관의 경우 디지털 서고의 구축과 메타데이터 생성 등의 보존기능이 수행되지 않고 있어 이에 소요되는 비용이 추정에서 배제되었기 때문이다. 본 연구는 디지털 보존이 디지털 자료의 생산 단계부터 접근까지의 모든 활동에 직간접으로 관련되어 있는 종합적인 활동이며, 디지털화 사업을 착수하는 단계에서부터 반드시 고려되어야 하는 사업임을 입증해 주고 있다.

ABSTRACT

To guarantee the long-term access to digital material, digital preservation needs to be systemized, and detailed investigation on cost elements of digital preservation should be done for the continued support of budget. To meet the needs in this area, this paper categorized the digital preservation cost into direct and indirect cost through deriving common elements used in prior research on this issue. For case analysis, two institutions, representing domestic University Library and National Library of Korea under large-scale digitization currently, are selected to analyze the current status of digital preservation and estimate the preservation cost. The case analysis shows the systematic preservation function should be performed to guarantee the long-term access digital material, even though a basic digital preservation is currently conducted. It was projected that the digital preservation cost for the two libraries, accounting for 11.8% and 8.6% of digitization cost, respectively, should be injected every year. However, the estimated figures are very conservative, because the cost for estimating the preservation function, such as installing digital repository and producing meta data, was excluded in the estimation. This proves that digital preservation is a synthetic activity linked directly and indirectly to various activities from production to access of digital object and an essential costs that should be considered from the beginning stage of digitization project.

키워드: 디지털 보존, 비용분석, 비용요소, 직접보존비용, 간접보존비용, 보존비용
digital preservation, cost analysis, cost elements, direct preservation cost, indirect preservation cost, preservation cost

-
- * KDI 국제정책대학원 정보자료실장(hkc@kdischool.ac.kr)
 - 논문접수일자 : 2005년 2월 2일
 - 게재확정일자 : 2005년 3월 5일

1. 서론

1. 1 연구의 필요성 및 목적

인터넷이 현대인의 생활에서 차지하는 비중이 커짐에 따라 디지털 자료의 중요성이 부각되고 있으며, 이의 생산량은 큰 폭으로 증가되고 있다. 인쇄 자료를 디지털로 매체 변환하여 이용자의 접근을 용이하게 하며, 원본을 이용으로부터 격리 보존하는 사례도 늘어가고 있다. 이는 원본인 인쇄 자료의 이용을 제한함으로써 언제든 무결성을 지닌 진본에 접근 가능하게 하기 위한 것이다.

디지털 자료는 인쇄 자료에 비해 장기적 접근이 어렵다. 디지털 자료는 특정한 소프트웨어나 하드웨어로 생산, 관리, 저장됨으로써 컴퓨터 환경의 발전과 더불어 기존 시스템 및 보존 매체의 노후화로 인하여 순식간에 망실되어 질 수 있다. 따라서 디지털 자료에의 장기적 접근을 위해서는 기술 환경의 빠른 변화, 매체의 비영구성, 환경변화로 인한 여러 가지 문제에 적절히 대응할 수 있도록 지속적이고 체계적인 디지털 보존 전략을 수립해야 한다.

디지털 보존은 개별적이고 독립적인 기능이 아니라 여러 기능들과 밀접하게 연관되어 있는 종합적인 기능이다. 디지털 보존은 생산 당시의 포맷과 파일 등의 기술적 환경에 의해 영향을 받기 때문에 생산 단계에서부터 고려되어야 하며 접근과도 불가분의 관계에 있다. 보존 없이 검색하는 것은 불가능하며, 미래의 접근을 보장하지 못하는 보존 또한 의미가 없기 때문이다.

따라서 디지털 보존비용은 관련된 비용의 총합으로 나타나야 한다. 디지털 보존비용 연구는

크게 개별 보존비용에 대한 연구와 이를 종합적으로 파악하려는 연구로 나누어진다. 전자는 Rothenberg(1999)의 보존기술에 대한 연구로부터 시작되었다. Rothenberg는 에뮬레이션이 원본의 모양과 느낌을 보존하는 데 가장 강력한 기법이나 위험발생률과 비용을 고려한다면 한계가 있는 기법이라고 주장하였다. 그리고 CAMiLEON(Creative Archiving at Michigan and Leeds: Emulating the Old on the New) 프로젝트에서 Granger(2000)는 에뮬레이터 개발관련 조사 및 사례연구 등을 실시하여 에뮬레이션의 구현가능성과 비용효과성을 검증하면서 보존기술이 디지털 보존의 핵심적인 비용요소임을 입증하였다.

한편 Lawrence 등(2000)은 디지털 보존기술의 위험관리에 대한 연구에서 대표적인 보존기술인 마이그레이션 처리시 발생하게 될 위험을 양적으로 평가하였다. 마이그레이션의 위험 평가에는 소프트웨어 전환, 다른 포맷으로의 전환, 파일전환 및 메타데이터 작성과 관련된 위험이 포함되어 있다. 또한 Lotus 1-2-3와 wk1 포맷을 위한 사례연구가 마이그레이션의 위험관리를 분석하기 위하여 수행되었다. 이와 유사한 연구로 서은경(2004)은 대학도서관을 대상으로 위험관리 분석을 수행하였다. 마이그레이션과 에뮬레이션 기술의 사용정도 및 이들에 대하여 디지털 정보관리 담당자들이 생각하는 위험가능성과 위험영향력을 조사하였다. 그 결과 대학도서관에서는 포맷변환이 가장 빈번히 사용되었으며 에뮬레이션은 전혀 사용되지 않는 기법으로 나타났다. 위험이 발생할 확률과 데이터 손실 영향력이 가장 낮은 방법은 매체재생인 것으로 판명되었고, 가장 높은 것은 에뮬레이션

기법과 정보전환인 것으로 나타났다.

디지털 보존의 비용요소 중 디지털 서고(repository)의 구축비용에 대한 연구는 Crespo and Garcia-Molina(2001)에서 볼 수 있다. 이 연구에 의하면, 서비스의 질과 종류에 따라서 크게 달라지는 디지털 서고의 구축비용에 대한 정확한 예측은 보존비용에 대한 합리적이고 경제적인 결정을 하는 데 매우 중요한 요소라고 주장하고 있다.

디지털 보존비용의 산출모형이 Hendley(1998)와 Sanett(2002)에 의하여 제시되었다. Hendley는 보존비용요소를 일회적 비용과 지속적 비용으로 범주화하였다. 전자에는 수서, 목록 등 단기적 투입에 필요한 기능이, 그리고 후자에는 디지털 객체의 수명주기 동안 장기적으로 투입되어야 하는 기능이 포함되어 있다. 한편 Sanett은 디지털 보존의 비용요소를 자본비용, 직접운영비용과 간접운영비용으로 범주화하였다. 자본비용에는 디지털 보존의 인프라를 구축하는 데 소요되는 비용이 포함되어 있다. 직접운영비용에는 자료선별, 평가/시험(지적재산권 협의), 입수, 인벤토리, 보존준비(진본성, 고결성 확인), 메타데이터 생성, 보존, 저장, 관리(보존기술처리), 모니터 등에 소요되는 비용이 포함되며, 간접운영비용에는 간접인건비, 시설, 관리비용 등이 포함되었다.

디지털 보존비용의 산출을 위한 종합적인 비용요소의 파악은 Russell과 Weinberger(2000), Chen(2001)에서 찾아볼 수 있다. 그러나 보존비용요소는 보존환경 -도서관의 목적과 사명, 비전, 이용자들의 요구, 보존대상-에 따라 매우 다양하므로 한 연구에서의 결과를 일반화하는 데 어려움이 있다. 따라서 어떠한 환경에도 적용될 수 있는

객관화된 비용요소를 파악해야 할 필요성이 대두되었다. 이러한 필요성에 따라 본 연구가 수행하고자 하는 연구의 목적은 다음과 같다.

첫째, 선행연구를 토대로 디지털 보존의 공통 비용요소를 파악하며, 이를 보존과의 관련성 정도에 따라 직접보존비용과 간접보존비용으로 범주화한다.

둘째, 추출된 디지털 보존의 비용요소와 관련된 기능이 사례분석 대상에서 수행되고 있는 지 여부를 파악한다.

셋째, 디지털 보존비용을 산출하기 위한 산출식을 제시한다.

넷째, 사례분석 대상들의 보존비용을 산출함으로써 디지털 자료의 지속적인 접근을 보장하기 위하여 고려해야 하는 장기적인 예산 규모를 추정한다.

1. 2 연구의 방법

1) 연구대상

두개의 도서관(국립대학도서관, 공공도서관)을 연구대상으로 선정하였다. 이들은 자관의 설립목적에 적합한 다양한 학술정보의 콘텐츠를 디지털화하고 디지털 콘텐츠 시스템을 개발하여 학술정보 서비스를 실현하고 있는 대표적인 도서관이다.

연구대상 자료는 영구보존대상이 되는 원문 DB자료로서 저작권이 도서관에 속해 있는 원문DB만을 포함하였다. 초기에 인쇄매체로 생산되었으나 이후에 디지털로 매체 변환된 형태뿐만 아니라 원래 디지털 형태로 생산된 전자저널 및 웹DB들을 소정의 구독료를 지불하여 출판사로부터 구독하고 있었다. 그러나 이들은 구독

기간 동안 사용할 수 있을 뿐 영구적으로 보유할 수 있는 권한은 없으므로 본 연구의 대상 자료에서 배제하였다.

적용되는 전산 및 시설장비의 유지보수비는 도서관에 따라 차이가 있을 수 있다. 유지보수비의 추적이 가능한 도서관의 경우는 실제 사례를 적용하였으나, 그렇지 않은 경우에는 일반적 기준인 8%를 적용하였다. 디지털화를 수행하는데 필요한 하드웨어 및 소프트웨어 등의 시스템은 매 5년마다 교체하며 초기와 동일한 비용으로 재투자비가 발생하는 것으로 가정하였다. 이는 시간이 지나면서 장비의 질이 향상되어 동일한 비용으로 향상된 품질의 장비구입이 가능해지기 때문이다.

연구기간은 10년으로 설정하였다. 이는 구축한 원문DB의 보존에 소요되는 비용을 한 순간이 아닌 축적되고 종합적인 차원에서 분석하기 위한 것이다. 그 이상으로 장기화하지 않은 것은 국내에 아직 디지털 보존에 대한 개념이 확립되어 있지 않으므로 예기치 않은 변수가 존재할 수 있기 때문이다.

2) 디지털 보존비용의 산출식

Hendley(1998)는 <표 1>과 같이 일회적 비용의 절반과 지속적 비용의 전체를 포함한 보존비용의 산출식을 제시하였다. 한편 Sanett(2002)은 디지털 보존비용을 자본비용, 직접운영비용, 간접운영비용으로 범주화하고 <표 2>와 같이 이들의 합을 보존대상 자료의 건수로 나누어 보

존비용의 단가를 산출하는 비용모형을 제시하였다. Sanett이 제시한 산출식은 보존비용을 지나치게 상세하게 나열하고 있어서 특별히 모형 개발에 관여되어 있는 사람만이 이해할 수 있으며, 보존활동들이 서로 겹쳐 있어서 범주화하는 것이 쉽지 않은 단점이 있다. 그러나 직접비용과 간접비용으로 범주화하는 Sanett의 방식은 디지털 아카이빙의 모든 기능들과 직간접으로 연관되어 있는 보존비용을 구분하기에 적합한 것으로 분석되어진다.

본 연구에서 제시한 산출식은 Sanett의 범주화 방식을 토대로 디지털 보존비용을 직접보존비용과 간접보존비용으로 범주화하였다. 전자에는 디지털 객체를 기술적 퇴화의 위협으로부터 보존하는 데 직접적인 연관이 있는 보존 계획, 보존기술처리, 품질관리, 위기관리에 투입되는 비용을 포함하였으며, 그 외의 디지털 아카이빙 기능을 위한 비용들은 후자에 포함하였다. 디지털 보존의 비용 요소는 관련연구 3개 이상에서 공통으로 제시된 경우에 한하여 추출되었다.

또한 직접보존비용과 간접보존비용 각각의 비율은 영국국립도서관 프로젝트에서 그 타당성이 입증된 바 있는 Hendley의 이론을 도입하였다. 따라서 본 연구에서 제시한 디지털 보존비용의 산출식은 영구적으로 투입되어야 하는 직접보존비용의 전체와 필요시 일시적으로 적용되는 간접보존비용의 절반(50%)을 포함하고 있다(표 3 참조).

<표 1> Hendley의 디지털 보존비용 산출식

$\text{디지털 보존비용} = \text{일회성 비용} / 2 + \text{지속적 보존비용}$

〈표 2〉 Sanett의 디지털 보존비용 단가 산출식

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

〈표 3〉 디지털 보존비용 산출식

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

2. 디지털 보존과 비용요소

2.1 디지털 보존의 개념 및 현황

한국국가기록연구원에 의하면, 디지털 보존은 “디지털 콘텐츠의 무결성(integrity)을 유지하면서 디지털 매체의 손상 및 하드웨어, 소프트웨어의 문제로 인한 기술적 퇴화의 위험으로부터 디지털 객체를 안전하게 보호하는 일련의 관리활동”으로 정의되고 있다. 인쇄 자료의 보존이란 보존대상의 형태와 콘텐츠를 변화시키지 않고 있는 그대로 계속 유지하는 것을 말하나, 디지털 보존은 데이터의 포맷과 매체의 변환을 통하여 콘텐츠를 장기적으로 유지시키는 활동을 의미한다. 디지털 환경에서는 형태적 변화 없이 원래의 모습으로 소장하여 보존하는 것이 불가능하기 때문이다.

장기적 보존 가치가 있는 디지털 콘텐츠의 특성을 규명하는 연구가 잇따라 나왔다. Duranti(1999)와 NARA(2000)는 디지털 콘텐츠의 신뢰성, 진본성, 무결성 및 활용성이 보존되어야 한다고 주장하고 있으며, Michele(2000)은 기록보존학적 관점에서의 보존은 전자기록물의 진본성을 장기간에 걸쳐 보장하는 것이라고 주장한다. 이를 토대로, 본 연구는 디지털 보존을 “지속적인 가치를 가졌다고 판단되는 디지털 자

료의 장기적 접근을 위하여 기술적 퇴화의 위험으로부터 디지털 콘텐츠의 특성을 장기간 안전하게 보호하는 제반활동”으로 정의하였다. 장기간은 가능한 긴 기간, 무한의 미래로까지 의미 확장이 가능한 시간을 의미한다.

보존 기술적 처리에 초점을 두는 디지털 보존은 디지털 아카이빙의 하위개념이라고 볼 수 있다. 디지털 아카이빙은 “지속적인 가치를 가졌다고 판단되는 디지털 자료를 장기간 관리하는 활동” 또는 “입수부터 배포까지의 수명주기 동안에 필요한 모든 활동”으로 정의된다.

디지털 아카이빙에 대한 관심이 고조되어 있는 선진국에서는 국가적, 국제적 차원에서 디지털 아카이빙이 수행되고 있다. 영국의 DPC(Digital Preservation Coalition, 2001년 시작)는 관련 기관들이 연합체를 결성하여 디지털 아카이빙을 추진하는 대표적인 사례로서 독일의 Nestor 프로젝트(2003년 시작)의 본보기가 되었다. 한편 2001년에 시작된 스위스의 e-Helvetica 프로젝트는 국가도서관이 주관하여 관련 기관들과 파트너십을 유지하는 형태로 진행되고 있다. 그리고 미국의 LC-NDIPP는 의회도서관이 정책주체가 되는 중앙집중식 정책 추진방식의 대표적인 사례이다.

국내에서는 KESLI가 전자저널의 아카이빙을 위한 NDAC(National Digital Archiving

Center)을 국가적 차원에서 추진하고 있다. 이 프로젝트를 위해서 KESLI는 컨소시엄으로 구성된 44개의 출판사들과 협의해 그들이 가지고 있는 전자저널의 메타데이터를 국내 서버에 저장하여 이용자들이 장기적으로 이용할 수 있도록 협의 중이다.

한편 기관 중심의 디지털 아카이빙 사업은 디지털 서고를 중심으로 추진되고 있다. 현재 국내·외 여러 가지 유형의 디지털 서고가 개발되어 운영되고 있는데, 이 중 유일하게 보존기능을 갖추고 있는 대표적인 디지털 서고 시스템은 DSpace이다. 그 외 영국에서 개발한 ePrints나 국내의 유일한 사례인 KERIS의 dCollection은 아직 보존과 관련된 기능을 제공하지 못하고 있다.

2.2 디지털 보존의 비용요소

영국국립도서관에서 수행된 디지털 보존 프로젝트에서 Hendley(1998)는 디지털 자료의 수명주기를 바탕으로 디지털 보존의 비용요소를 7개 항목으로 구분하였다. 항목에는 데이터 선정 및 평가, 데이터 생산, 데이터 관리, 메타데이터 생성, 데이터 사용, 데이터 보존, 그리고 저작권 관리가 포함되어 있다.

Russell과 Weinberger(2000)는 업무흐름을 토대로 디지털 보존의 비용요소를 추출하였다. 비용요소는 보존대상 선정, 저작권 관리, 접근점 제공에 대한 권리, 보존전략, 디지털 객체의 완전성 여부 점검, 메타데이터 생성, 파일저장, 아카이브 관리 등의 8개 항목으로 구분되었다. 이들의 연구결과는 다음과 같이 요약할 수 있다. 첫째, 디지털 자료는 급변하는 기술로 인하여 지속적인 보존처리가 요구되기 때문에 장

기적인 시간 프레임이 요구된다. 둘째, 디지털 자료 구축 당시의 포맷과 파일 등의 기술 환경이 디지털 보존정책에 영향을 미치게 되므로 보존정책은 디지털 자료의 생산단계부터 고려되어야 한다. 그러나 최근의 비용모형이 이러한 디지털 아카이빙의 복잡한 현실을 반영하지 못하고 있다. 셋째, 투입하고자 하는 보존비용이 그 자료의 가치에 적합한 수준인가에 대한 철저한 분석이 필요하다.

Chen(2001)은 디지털 보존의 비용요소를 디지털 자료의 포맷, 스타일, 내용, 저장매체, 시스템 기술, 업무흐름 절차, 메타데이터 정책으로 구분하였으며, 이들의 복잡한 연관관계를 세부적으로 분석해 볼 필요가 있다고 주장하였다. 그는 또한 디지털 내용을 그대로 유지하면서 환경적인 변화(매체의 단명성, 시스템의 노후화 등)로 인한 매체의 지속적인 변화를 허용해야 하는 점이 디지털 보존의 양면성이라고 지적하였다.

한편 Sanett(2002)은 InterPares 작업팀에 의하여 개발된 IDEF(Integration DEfinition for Function Modeling) 프로젝트에서 디지털 보존의 비용요소를 크게 세 개의 범주 - 보존비용, 기관 사용비용, 이용자 정보 제공에 소요되는 비용-로 구분하였다. 각각의 범주는 크게 자본비용, 직접운영비용, 간접운영비용으로 나누어진다. 자본비용은 인프라에 구축에 필요한 비용을 말한다. 직접운영비용에는 자료선별, 평가/시험(지적재산권 협의), 입수, 인벤토리, 보존준비(진본성, 고결성 확인), 메타데이터 생성, 보존, 저장, 관리(보존기술처리), 모니터 등에 소요되는 비용이 포함된다. 한편 간접운영비용에는 간접인건비, 시설, 관리비용 등이 포함된다. Sanett은 기관의 목적과 사명, 비전, 이용자

요구, 보존대상 등에 따라 매우 다양하나, 여기에는 공통 비용요소가 존재한다고 주장하였다.

OAIS(Open Archival Information System) 참조모형은 2002년에 ISO 표준으로 인증되어 디지털 정보를 장기적으로 보존하는데 기본 지침이 되고 있다. OAIS 참조모형은 NASA와 CCSDS에 의하여 제안되었으며 디지털 아카이빙을 장기적으로 유지 및 관리하기 위하여 흡수, 보존, 데이터 관리, 운영, 보존계획, 접근과 같은 6가지 기본적 기능이 유기적으로 이루어져야 한다고 제시하였다(CCSDS 2002).

디지털 보존 비용분석 관련연구에서 나타난 디지털 보존의 비용요소는 <표 4>에 정리되어 있으며 관련연구에는 동일한 의미의 비용요소들이 상이한 용어로 규정되어 있다. 본 연구는 관련연구 3개 이상에 공통으로 제시된 요소를 디지털 보존의 비용요소로 추출하였다. 또한 이들에 대한 의미의 통일을 시도하여 디지털 보존의 공통 비용요소로 보존, 관리, 대상선정, 메타데이터 생성, 접근기능을 추출하였다.

3. 디지털 보존 비용분석 연구척도

관련연구에서 추출된 디지털 보존의 비용요

소에 보존계획, 인프라 구축, 원문DB구축을 추가하고, 관리를 품질관리와 위기관리로 분리하여 총 9가지 비용요소로 구분하였다. 이렇게 구분한 이유는 다음과 같다.

첫째, 관련연구로부터 추출된 필수 비용요소를 OAIS 참조모형이 제시하는 디지털 아카이빙 업무흐름의 관점에서 재구성하였다. 둘째, 원문DB 구축은 주로 외주로 이루어지는 과정이나, 여기에 사용된 기술적 환경이 향후 보존활동에 영향을 미치므로 원문DB구축을 추가하였다. 셋째, 인프라 구축은 디지털 아카이빙의 가장 기본적인 비용요소이므로 추가하였다. 넷째, 관리는 디지털 콘텐츠의 품질관리와 매체를 보호하기 위한 위기관리로 구분하였다. 다섯째, 디지털 보존은 궁극적으로 접근과 밀접하게 연관되므로 접근기능을 추가하였다. 여섯째, 디지털 보존을 체계적으로 수행하기 위해서는 보존계획이 반드시 선행되어야 하므로, 이를 보존요소에 추가하였다.

직접보존 비용요소는 OAIS 참조모형(보존 모듈)을 토대로 보존기술처리, 품질관리와 위기관리를 포함하였다. 또한 보존계획은 직접보존의 비용요소와 직접적으로 관련이 있기 때문에 직접보존 비용요소에 포함하였다. 간접보존 비용요소에는 직접보존 비용요소 외에 디지털 보존과 관련되어 있는 원문DB대상선정, 인프라

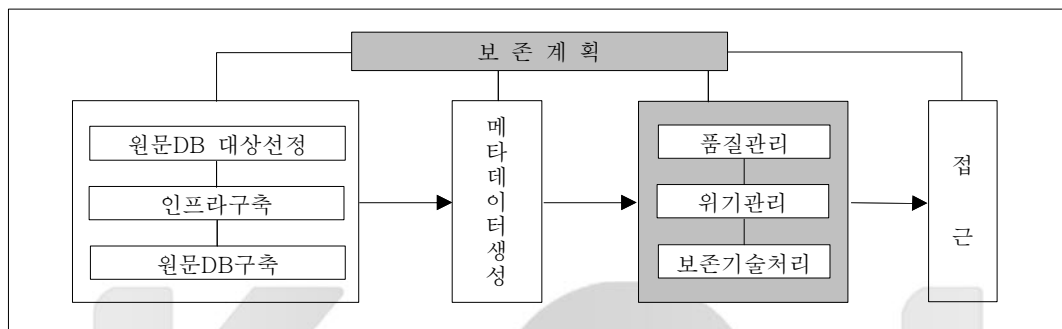
<표 4> 디지털 보존 비용분석 관련연구

Hendley(1998)	데이터 선정 및 평가, 데이터 관리, 저작권 관리, 메타 데이터 생성, 데이터 사용, 데이터 보존,
Russell and Weinberger(2000)	보존대상 선정, 저작권 관리, 접근점 제공에 대한 권리, 보존전략, 디지털 객체의 완전성 여부 점검, 메타데이터 생성, 파일저장, 아카이브관리
Chen(2001)	아카이브 가치 및 행정적 가치 평가, 검증, 기록관리시스템 관리, 메타데이터 생성, 접근, 폐기
Sanett(2002)	자본비용: 인프라 구축(시스템 구입비용) 직접운영비용: 자료선별, 평가/시험, 입수, 인벤토리(진본성/고결성 확인), 메타데이터 생성, 보존, 저장, 모니터 간접운영비용: 간접 인건비, 시설, 관리비용 등
OAIS참조모형	흡수, 보존, 데이터 관리, 운영, 보존계획, 접근

구축, 원문DB구축, 메타데이터 생성, 접근을 포함하였다(그림 1 참조).

직접보존비용은 디지털 보존에 직접적으로 연관이 되는 기능을 하는 데 소요되는 비용이다. 직접보존 비용요소에 관련이 있는 기능으로는 보존계획, 품질관리 및 위기관리가 포함된다(표 5 참조).

3. 1 직접보존 비용요소



<그림 1> 디지털 보존의 비용요소

<표 5> 직접보존 비용요소

구분	비용요소	세부비용요소	
보존계획	관심공통체의 요구사항 점검	인건비, 국내의 보존관련기술 모니터링 비용, 표준개발비용 등	
	국내의 보존관련기술 검토		
	보존전략 및 표준개발		
	포맷전환 계획 발표		
보존기술처리	보존기술처리(마이그레이션, 에뮬레이션, 기술보존) 중 대상 자료에 적합한 기법에 대한 협의	인건비, 전산장비의 유지관리비, 에뮬레이터 개발비, 시스템 보존비, 변환처리비용	
	보존기술처리 수행		
관리	품질관리	컨텐츠 품질 관리	인건비, 소프트웨어 개발비
		링크체크작업	
		소프트웨어 개발	
		서비스의 질(화질, 속도, 접속시간 등) 체크	
		이용자 그룹별 접근허용과 정도에 대한 규정책정	
		모니터링 리포트	
		반품 및 폐기에 필요한 서류작성 및 협의	
		보존 CD-ROM의 관리	
위기관리	백업전용서버 구입	인건비, 인프라 구축비, 전산장비의 유지관리비	
	백업전용 프로그램 라이선스 구입비		
	전산장비(서버, 하드웨어, 소프트웨어) 유지보수		
	전산장비(서버, 하드웨어, 소프트웨어) 교체		
	위험분석		
	대체방안 수립 및 위기관리 프로그램 개발		

1) 보존계획

디지털 자료를 장기적으로 보존하기 위해서는 국내·외 관련기술과 관심 공동체의 요구사항과 각종 기술변화의 모니터링을 토대로 한 보존계획이 수립되어야 한다. 보존계획에는 보존전략 수립, 표준개발, 패키징 설계, 포맷전환 계획 개발 및 보존기술처리의 기준 설계를 위한 인건비, 국내외 관련기술 검토비용, 표준개발비용 등이 포함된다.

2) 보존기술처리

기술적 퇴화에 따른 데이터 손실을 예방하기 위한 보존기술처리는 보존활동에서 가장 핵심 요소이다. 보존기술로는 기술보존, 에뮬레이션, 마이그레이션의 3가지 방식이 알려져 있다. 이중 가장 많이 사용되는 방식인 마이그레이션 기법에는 매체재생, 매체변환, 포맷변환, 버전업그레이드 등이 있다. 보존기술처리를 위해서는 인건비, 유지관리비, 에뮬레이터 개발비, 시스템 보존비 및 변환처리비용 등이 소요된다.

3) 품질관리

장기적 보존가치가 있는 자료에 대해서는 원문DB 콘텐츠의 품질상태를 정기적으로 파악하고 콘텐츠에 링크하는 데 걸리는 시간과 이상 여부, 콘텐츠가 제대로 업데이트 되고 있는 지 등에 대한 품질관리가 요구된다. 품질체크과정에서 이상이 발견된 원문DB를 대상으로 보존기술처리 여부를 결정한다. 또한 원문DB의 복사본(CD)에 대한 품질관리에서는 반품 및 폐기를 위한 서류작성 및 협의, 그리고 보존 CD-ROM의 품질을 관리하는 데 인건비가 소요된다.

4) 위기관리

위기관리는 예기치 못한 돌발사태로 인해 데이터 및 매체에 발생하게 될 위기를 방지하기 위한 것이며, 다음의 작업이 포함된다. 첫째, 백업 전용 서버와 이를 위한 프로그램 라이선스 및 바이러스 점검 소프트웨어를 구입하고, 이를 정기적이고 지속적으로 점검한다. 둘째, 갑작스런 재난에 대비하기 위해 원본을 동일한 매체 또는 보다 장기적 접근이 가능한 다른 매체로 변환 한다. 셋째, 하드웨어, 소프트웨어 및 시설 장비의 유지보수 계약을 체결한다. 여기에는 백업전용 서버 구입비, 프로그램 라이선스 구입비, 바이러스 체크업 소프트웨어, 그리고 전산장비의 유지보수비, 교체비, 위기관리를 위한 위험분석과 위기관리 프로그램 개발 등의 연구 개발비 및 인건비 등이 소요된다.

3. 2 간접보존 비용요소

간접보존비용은 디지털 보존에 간접적으로 관여된 기능을 하는 데 소요되는 비용으로 원문DB 대상선별, 인프라 구축, 원문DB구축, 메타데이터 생성, 그리고 접근비용이 포함된다 (표 6 참조).

1) 원문DB 대상선정

원문DB 대상선정을 위한 준비작업에는 저작권 협의, 자료 수집을 위한 공문발송, 구축방안 검토, 구축업체 선정, 선정위원회 위원 위촉, 구축 요구서 작성 그리고 계약의뢰 등이 포함된다. 선정위원회를 구성하여 원문DB 구축의 대상 자료를 선정하며, 선정기준은 기관의 유형과 보존정책에 따라 달라질 수 있다. 원문DB

〈표 6〉 간접보존비용

가능	세부기능	비용요소
원문DB 대상선정	저작권 협의	인건비, 기타
	학내 전 기관 자료수집 공문발송	
	수집자료 현황 및 구축방안 검토	
	국내 구축업체 설명회 개최	
	대상자료 선정위원회 위촉을 위한 검토	
	선정위원회 위원 위촉	
	구축 요구서 작성 및 계약의뢰	
	기준설정	
	대상평가	
	대상선정	
인프라 구축	저작권법 검토	하드웨어 구입비, 소프트웨어 구입비, 시설 구입비
	저작권자와의 협의	
	서버와 디스크 구입(확장성 고려)	
	컨텐츠관리시스템, 변환 프로그램, 보안 프로그램, 백업 프로그램, 바이러스 체크업 프로그램 등 소프트웨어 구입	
원문DB 구축	시설구입	원문구축비
	목차생성	
	스캔(원문스캔, 검수)	
	백업용 CD 제작	
	이미지 보정(양면, 단면) 및 검수	
	목차입력 및 교정	
	검수	
업로드		
메타데이터 생성	연동점검 및 납품	인건비
	서지 및 보존 메타데이터 생성 및 관리	
접근	메타데이터 수정 또는 삭제의 갱신	인건비, 시스템도입비
	아이템의 위치식별, 검색요청 처리, 아이템의 제공 등에 대한 서비스관리	

대상선정에는 기준설정, 대상평가, 저작권법의 검토 및 저작권자와의 협의 등의 작업을 위한 인건비 및 기타 비용이 소요된다.

것이다. 디지털 보존에는 하드웨어, 소프트웨어 및 시설장비 등의 인프라 구축을 위한 비용이 투입되어야 한다.

2) 인프라 구축

디지털 보존에는 반드시 하드웨어, 소프트웨어, 시설 등의 인프라 구축이 선행되어야 한다. 하드웨어에는 서버와 디스크, 확장성을 고려한 확장형 서버가 포함된다. 또한 보안을 위한 보안벽, 백업전용 소프트웨어, 메타데이터 등 여러 가지 용도의 소프트웨어가 필요하다. 보안벽은 디지털 자료의 진본성을 보존하기 위한

3) 원문DB구축

디지털 자료의 생산시점에서 어떠한 기술과 소프트웨어를 이용하는가는 향후 보존활동에 커다란 영향을 주므로 원문DB의 보존은 이를 구축하는 단계부터 고려되어야 한다. 일반적으로 원문DB의 구축은 주로 외부 용역으로 진행되고 있다. 단, 원문DB 구축비용과 디지털 보존비용을 비교하는 것이 이 연구의 목적이므로

원문DB 구축비용은 비용 산출에서 제외된다.

4) 메타데이터 생성

메타데이터는 디지털 자원의 특성을 기술하는 구조화된 데이터를 말한다. 메타데이터는 가장 기본적인 용도인 자원의 관리 및 탐색을 포함하여 안전, 내용등급, 저작권관리, 보존 등의 다양한 용도에 사용될 정보들로 구성되어 있다. 또한 보존에 관련된 각종 정보를 기술하는 보존 메타데이터는 보존과 직접적인 관련이 있으며 기술적인 데이터, 저작권 관련 데이터, 디지털 문헌의 진본을 유지하는 데 필요한 데이터 등이 기술되어 있다. 보존 메타데이터에 필요한 요소를 추출하기 위한 연구와 관리, 그리고 생성하는 데 인건비와 소프트웨어 도입에 비용이 투입된다.

5) 접근

저장되어 있거나 보존 작업에 의해 유지되고 있는 디지털 자료는 이용자의 요청이 있을 때 접근을 제공해야 한다. 이용자는 다양한 검색 방식(일반검색, 고급검색, 브라우징, 통합검색 등)을 이용하여 메타데이터를 검색하여 자료에 접근하게 된다. 검색된 디지털 자료에 접근하기 위해서는 콘텐츠의 식별자에 대한 해석 작업과 다운로드 작업, 필요시 반출작업 등이 필요하다. 또한 이용자에게 디지털 자료의 지적 내용에 접근하도록 하기 위하여 저장/보존된 디지털 파일을 그대로 제공하거나 포맷 전환하여 제공, 또는 공통의 뷰어를 통한 제공 등의 방식을 선택해야 하며, 이러한 작업이 보존 작업과 연관되어 있다. 그 외 저작권 제어방식, 표준화된 유통방식(OAIS기반의 메타데이터

및 원문 유통), 일괄반출 기능 등에 대한 고려도 필요하다. 이용자에게 장기적으로 접근을 제공하기 위해 필요한 모든 기능을 수행하는데 필요한 비용이 여기에 포함된다.

4. 디지털 보존 비용분석 사례연구

4. 1 도서관 A

4. 1. 1 원문DB 보존현황

도서관 A는 국립대학도서관으로서 「BK21」 정보화 사업의 일환으로 원문DB 구축사업을 착수하였으며 one-stop 서비스의 실현을 위하여 전자도서관 시스템을 구현하고 있다. 본 연구의 대상인 원문DB구축자료에는 학내간행물, 행정업무 관련자료, 대학의 기록과 관련된 사진 및 그림자료 등과 저작권이 해결된 고문헌 및 웹자원 등이 포함되어 있다.

원문DB로 구축된 자료는 총 8,326,947면이며, 이를 보존하기 위해서 집행된 예산은 총 5,483,965천원으로 집계되었다. 원문DB구축을 위한 예산으로는 2002년과 2003년에 각각 23억, 22억9천여만원이 집행되었으며, 2005년에 890,766천원의 예산이 책정되어 있다. 사업의 특성 때문에 이 비용에는 향후 몇 년간의 시스템 유지관리비가 포함되어 있다. 그러나 이를 원문DB의 구축비용에서 분리해 내는 것이 쉽지 않았다(표 7 참조).

원본은 Tiff와 JPG 포맷으로, 서비스용은 PDF와 DjVu 포맷으로 구축하였다. 원문DB 구축을 위한 준비작업으로 구축대상 파악, 콘텐츠 유형 파악, 표준 형식과 메타데이터 표준

〈표 7〉 도서관 A: 원문DB 구축현황

1 차			2 차			3 차		
2002년			2003년			2005년		
면수	건수	비용(천원)	면수	건수	비용(천원)	면수	건수	비용(천원)
6,944,833	214,982	2,300,000	891,980	323,672	2,293,199	512,846	57,259	890,766
면수: 8,326,947			건수: 595,913			비용: 5,483,965(천원)		

안 등의 표준화 연구와 저작물에 대한 권리 및 저작물 보호에 대한 연구가 선행되고 있다.

이 도서관에는 디지털 보존을 담당할 별개의 센터나 부서가 설립되어 있지 않았으며 별도의 예산이 책정되어 있지 않았다. 또한 디지털 보존업무가 종합적으로 수행되고 있지 않아 디지털 보존의 세부요소에 대한 집행내역을 추적하는 것이 불가능하였다.

전산실의 시스템 담당자는 데이터 백업, 바이러스 체크업 등의 위기관리와 업로드(연동)에 이상 여부, 그리고 보다 용이한 이용자 접근 방법 검토 등 지속적인 데이터 품질관리를 수행하고 있었다.

시스템의 노후화나 매체의 손상 등으로 인해 디지털 자료에의 이용자 접근에 문제가 발생

했을 때 주로 사용하는 보존처리기술은 마이그레이션 기법이며, 대표적인 예로는 고문헌 콘텐츠 구축사업과 학위논문을 들 수 있다.

고문헌 콘텐츠 구축사업은 고문헌을 디지털화 하여 동시대 연구자들 뿐만 아니라 일반인들에게 시간적, 공간적 제약 없이 손쉽게 이용할 수 있게 하고자 실시한 사업이다. 초기에 고문헌 콘텐츠를 위해서 구축된 PDF파일은 고문헌의 용량을 감당하기에 역부족이어서 지속적으로 이용자의 접근에 불편을 초래하였다. 따라서 이미 PDF로 구축된 소수의 콘텐츠를 포함하여 고문헌 모두를 DjVu파일로 전환하게 되었다.¹⁾ 고문헌 콘텐츠 구축사업은 60일 동안 추진되었으며 구체적인 구축대상 및 구축내역은 〈표 8〉과 같다. 학위논문은 1999년부

〈표 8〉 고문헌 콘텐츠 구축사업

구 분	내 역
구축대상	* 일제강점기자료(단행본)
자료형태	* 인쇄자료
입력형태	* 목차 입력 후 TOC 생성 * 권 단위 입력
이미지 구축형태	* JPEG(보관용 양면), * JPEG(보관용 단면) * DJVU(서비스용)
기타파일	* 파일관리정보, 저작권관리정보, 순서목차, 해제
총합	* 22,712면(76책 x 300면), A4 칼라 오버헤드 스캐닝
집행예산	* 약 43,199천원(단가 1902원)

1) DjVu는 1996년 AT&T사에서 개발한 이미지 압축기술이다. PDF 파일에서 불가능하였던 칼러스캐닝을 가능하게 하였고 JPEG, GIF, TIFF와 같은 이미지 파일 보다 3-10배 이상의 높은 압축률을 보여 높은 해상도의 이미지를 파일 크기에 부담 없이 웹상에서 서비스 할 수 있는 것이 특징이다.

터 2002년까지 주로 Tiff로 서비스 되어 왔는데 2002년에 KERIS의 지원으로 45,079건의 학위논문을 PDF 파일로 변환하는 작업을 수행하였다. 총 대상건수(49,746건)에서 완료되지 못한 건들은 파일 누락 및 원문과 불일치, 목차파일의 누락 등의 문제에 의한 것이다.

4. 1. 2 원문DB 보존비용

구축된 원문DB(8,349,659면)에 소요된 비용은 총 5,483,965천원으로, 면당 단가는 약 657원으로 집계되었다.²⁾ 직접보존비용에는 시스템의 유지관리비, 매체변환비용 및 인건비가 포함되어 있다. 하드웨어와 소프트웨어의 유지관리비로 각각 2,893,862천원, 821,350천원이 투입되었고, 이용자 접근의 용이함을 위하여 파일형식을 Tiff에서 PDF로 변환한 학위논문 처리비용(85,500천원)은 별개의 매체변환비용에 포함하였다. 학위논문처리비용은 별도의 예산으로 집행되었기 때문에 이를 매체변환비용으로 구분한 것이다. 그 외의 매체변환 등은 유

지관리비에 포함되어 있다.

한편 고문헌 콘텐츠 구축사업에 들어간 비용은 원문DB 구축비용에 포함하였다. 이는 PDF에서 DjVu로 파일 변환된 고문헌은 매우 적은 비중을 차지하고 대부분은 인쇄 고문헌에서 DjVu로 원문DB화되었기 때문이다. 그리고 보존에 직접적으로 관여하고 있는 정규직원 1명의 급여를 포함하여 직접보존비용은 4,166,412천원으로 집계되었다.³⁾

한편 간접보존비용에는 인프라 구축비와 인건비가 포함되어 있다. 하드웨어에 2,378,400천원, 소프트웨어에 635,000천원, 그리고 UPS와 항온항습기의 시설구입에 87,600천원이 투입되었다. 인건비 항목에는 원문DB의 보존활동에 간접적으로 관여한 4명의 정규직원의 급여가 포함되어 있다. 간접보존비용은 총 4,611,780으로 집계되었다(표 9 참조).

디지털 보존에 집행된 예산이 다른 기능과 혼합되어 있는 경우에는 담당자의 제언을 토대로 원문DB의 보존에 관계되는 비용만을 분리해 내었다.

〈표 9〉 도서관 A: 직접보존비용과 간접보존비용

(단위: 천원)

		구분	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	합계	
직접 비용	보존 기술 처리	유지 관리비	H/W				84177	136204	1208100	1245000		84177	136204	2893862
			S/W				36025	38100	38100	635000		36025	38100	821350
		매체변환비용		85500										85500
	인건비	1명	36570	36570	36570	36570	36570	36570	36570	36570	36570	36570	365700	
	합계		122070	36570	36570	156772	210874	1282770	1916570	36570	156772	210874	4166412	
간접 비용	인프라 구축비	H/W	2378400											
		S/W	635000											
		시설	87600		8000	8000	8000	87600		8000	8000	8000	135600	
	인건비	4명	146278	146278	146278	146278	146278	146278	146278	146278	146278	146278	1462780	
	합계		3247278	146278	154278	154278	154278	233878	146278	154278	154278	154278	4611780	

2) 원문DB 면당단가 = 5,483,965천원 / 8,349,659면 = 약 657원/1면
 3) 도서관 A의 정규직원의 1인당 평균급여 36,569천원을 기준으로 책정되었다.

본 연구가 제시한 보존비용 산출식에 의한 10년간의 보존비용은 총 6,472,302천원으로 추정되어진다. 따라서 매년 647,230천원의 보존비용이 원문DB의 보존을 위해서 투입될 것으로 추정되며, 이는 구축비용의 11.8%에 해당된다⁴⁾(표 10 참조).

4. 2 도서관 B

4. 2. 1 원문DB 보존현황

도서관 B는 공공도서관으로서 원문DB 구축 사업을 통하여 정보를 신속하게 제공하고 지역의 정보 빈부 격차를 해소하며 정보 활용도를 향상시키는 것을 목적으로 하고 있다. 원문으로 구축된 자료는 총 67,553,190면이며, 이에 집행된 예산은 11,973,120천원으로 집계되었다(표 11 참조).

발행된 지 5년이 경과된 소장자료 중 학술적

정보적 가치가 있는 자료만을 선별하여 원문DB를 구축하였다. 이는 학문별 전문가 10명으로 구성된 원문DB 구축위원회 및 관내 자문위원에 의하여 선정되었다. 보존 대상 자료는 고서, 1997년 이전 단행자료, 문화관광부 발간자료원, 한글편 고전소설, e-book 등 총 14개의 범주로 구분되어 있다.

디지털 자료의 보존을 위하여 별도로 구분된 예산항목은 없다. 원문DB에 이용자들이 지속적으로 접근할 수 있도록 하기 위해 매년 시스템 유지보수계약을 체결하여 하드웨어, 소프트웨어, 통신기술, 소장매체, 보존기술 등을 체계적이고 지속적으로 관리하고 있다. 또한 원문정보의 링크서비스를 자동적으로 체크할 수 있는 소프트웨어를 개발하여 이용자가 발견한 원문정보의 오류여부가 관리자에게 실시간 전달되도록 하였다. 초기에는 원문정보를 저작권 및 데이터 보안을 위하여 single Tiff로 구축하

〈표 10〉 도서관 A: 보존비용

$$\begin{aligned} \text{보존비용} &= \{ \text{직접보존비용} + (\text{간접보존비용} \div 2) \} \\ &= \{ 4,166,412\text{천원} + (4,617,780\text{천원} \div 2) \} \div 10\text{년} = 647,230\text{천원(년)} \end{aligned}$$

〈표 11〉 도서관 B: 주요자료의 원문DB 구축현황

사업연도	대상자료	면수	건수	비용(천원)	면당 단가(원)
2000년	주요자료단행본(흑백)	34,000,000	170,000	4,418,000	177
2001년	주요자료단행본(칼라), 한국관련 자료, 고대소설, 한국고서 등	4,750,440	16,980	2,210,000	
2002년	한국관련자료, 연속간행물, 한국관련외국자료 등	17,035,750	59,282	2,122,120	
2003년	1945년 이전 발간자료	7,935,000	23,000	1,338,000	
2004년	귀중본, 한국고서, 고문서, 대학기본학습서 등	3,832,000		1,885,000	
	소 계	67,553,190		11,973,120	

4) 647,230천원(보존비용) ÷ 5,483,965천원(원문DB 구축비용) × 100% = 11.8%

고 전용 뷰어로 국가전자도서관 데이터베이스를 이용하였으나, 최근에는 현재의 추세인 아크로벳 리더기로 읽을 수 있는 PDF 파일로 전환하고 있다.

4. 2. 2 원문DB 보존비용

직접보존비용에는 시스템 유지관리비와 데이터 백업, 품질관리 및 시스템관리를 담당하는 1명의 정규직원(전산직 6급 1명)에 대한 인건비가 포함되어 있다.⁵⁾ 시스템 유지관리비에는 매체변환, 버전 업그레이드, 기술 환경 전환 등의 보존기술처리에 집행된 비용이 포함되어 있다. 시스템 유지보수비율은 구체적으로 명시되어 있지 않으므로 초기 구입비용의 8%를 구입 1년 후부터 일괄적으로 적용하였고 매 5년마다 시스템 교체를 위해서 초기와 동일한 구입비용이 소요되는 것으로 가정하였다. 직접 보존비용은 6,727,180천원으로 추정되었다.

간접보존비용에는 인프라 구축비용, 소프트웨어 개발비용 및 인건비가 포함되었다. 간접 보존비용에 포함되어 있는 인건비는 사서직 6급 1명과 전산직 7급 1명의 평균 급여가 포함되어 있다. 소프트웨어 개발비용에는 원문DB 구축과 가장 밀접한 관련이 있는 국가 전자도서관 경량화 사업에 6억3천만원, 유니코드체계 구축 정보화전략계획 수립 및 신규 개발에 7억 8천만원과 이에 대한 유지보수비가 포함되어 있다. 도서관 B의 간접보존비용은 7,761,473천원인 것으로 추정되었다(표 12 참조).

보존비용 단가 산출식에 의한 도서관 B의 보존비용은 10,607,917천원으로 매년 1,060,792천원에 해당하는 비용이 원문DB의 보존을 위하여 소요될 것으로 추정되었다(표 13 참조). 이는 원문DB구축에 소요된 총 예산(11,973,120,000원)의 8.6%에 해당하는 비용이다.⁶⁾ 원문DB 구축비용은 보존비용의 산출에서 배제된다. 이는

〈표 12〉 도서관 B: 직접보존비용과 간접보존비용

(단위: 천원)

구 분			2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	합계	
직접 비용	보존기술처리	유지관리비	H/W	68343	121812	126646	953528	1703274	146497	121812	126646	953528	1703274	6025360
			S/W	1994	10536	11507	11507	171859	21348	10536	11507	11507	171859	434160
		인건비	1명	26766	26766	26766	26766	26766	26766	26766	26766	26766	26766	267660
		합계		97103	159114	164919	991801	1901899	194611	159114	164919	991801	1901899	6727180
구 분			구입년도	2003	2004	2005	2006	2007	2008	2009	2010	2011	합계	
간접 비용	인프라구축비	H/W	2717549										2717549	
		S/W	190048										190048	
		시설	85900	4100	6272	6272	42796	49376	4100	6272	6272	42796	254156	
	인건비	2명	53532	53532	53532	53532	53532	53532	53532	53532	53532	53532	535320	
	소프트웨어 개발비		630000	50400	50400	50400	630000		50400	50400	50400	630000	2192400	
				780000		62400	62400	62400	780000		62400	62400	1872000	
	합계	3677029	888032	110204	172604	788728	165308	888032	110204	172604	788728	7761473		

5) 공무원 정규직원의 평균직급을 7급10호봉으로 설정하여, 이 직급의 연평균 급여 26,766천원을 적용하였다. 이는 기본급 11,424천원(기본급)에 각종 수당 및 보조비(정근수당, 가계지원비, 명절휴가비, 급식비, 직급보조비, 교통보조비 등)를 포함하여 산출한 것이다.
 6) 1,060,802천원(보존비용) ÷ 11,973,120천원(원문DB 구축비용) x 100% = 8.6%

〈표 13〉 도서관 B: 보존비용

$\begin{aligned} \text{보존비용} &= \{ \text{직접보존비용} + (\text{간접보존비용} \div 2) \} \\ &= \{ 6,727,180\text{천원} + (7,761,473\text{천원} \div 2) \} \div 10\text{년} = 1,060,802\text{천원(년)} \end{aligned}$
--

본 연구가 원문DB 구축비용에 대한 보존비용의 비율을 파악하는 데 목적을 두고 있기 때문이다.

4. 3 사례분석 결과

사례분석 결과에 의하면, 대상 도서관들이 실시하고 있는 디지털 보존은 문서화된 보존계획을 토대로 총체적으로 이루어지는 것이 아니라 파일을 읽거나 접근하기 어려운 상황이 발생했을 때 ad hoc 방식으로 이루어지고 있었다. 또한 이들 도서관에서는 디지털 보존의 공통 비용요소 중 몇 가지 요소만이 수행되고 있었고, 집행된 예산 중에는 개별항목으로 분리되어 있지 않은 부분이 있어 추적이 쉽지 않았다.

도서관 A와 B에서는 각각 원문DB 구축비용의 11.8%와 8.6%에 해당하는 보존비용이 매년 소요되고 있는 것으로 추정되었다. 그러나 도서관 A의 경우에는 원문DB구축비용에 추후 몇 년간의 시스템 유지보수비가 포함되어 있어, 이 결과를 일반화하는 데 무리가 있는 것으로 분석되어진다.

한편 사례분석 결과에 따른 디지털 보존비용의 추정은 다소 과소평가 되었을 가능성이 크다. 디지털 보존에서 가장 큰 비중을 차지하는 것으로 나타난 디지털 서고비용과 보존 메타데이터 관련 비용 등이 배제되어 있기 때문이다. 이는 대상 도서관에서의 디지털 보존이 아직 부분적으로 이루어지고 있음을 입증하는 좋은 예이다.

5. 결 론

디지털 보존은 인쇄 보존과 달리 자료가 훼손 되었을 당시에 이루어지는 일시적인 활동이 아닌 보존전략에 따라 지속적으로 이루어져야 하는 고가(高價)의 활동이다. 디지털 자료는 시스템의 기술적인 노후화와 매체의 손상으로 인하여 장기적인 접근을 보장하기 어렵기 때문에 지속적인 가치를 가졌다고 판단되는 디지털 자료의 장기적인 접근을 위해서는 체계적이고 지속적인 보존활동이 수반되어야 한다. 적절한 보존활동이 수반되지 않는 디지털 자료는 컴퓨터 환경의 비약적인 발전과 더불어 많은 문제에 노출되어 순식간에 망실되어질 수 있다.

디지털 보존은 독립적인 기능이 아니라 생산부터 접근까지의 여러 기능들과 관련이 있다. 선행연구를 토대로 디지털 보존의 공통 비용요소를 추출하여, 이를 관련성 여부에 따라 직접 보존비용과 간접보존비용으로 범주화하였다. 직접보존 비용요소로 보존계획, 보존기술처리, 품질관리, 위기관리, 그리고 간접보존 비용요소로는 원문DB 대상선정, 인프라 구축, 원문DB구축, 메타데이터 생성, 접근이 추출되었다.

국내 디지털 보존 현황을 분석하고 디지털 보존비용을 추정하기 위하여 디지털화 사업이 활발하게 추진되고 있는 대표적 도서관을 대상으로 사례분석을 수행하였다. 또한 관련연구를 토대로 디지털 보존비용의 산출식을 제시하였다. 디지털 보존에 직접적으로 관여된 비용전체와

간접보존비용의 50%를 합산하도록 되어 있는 산출식을 적용한 결과, 두 도서관의 연간 보존 비용은 각각 원문DB 구축비용의 11.8%, 8.6%에 이르는 것으로 나타났다.

한편 본 연구에서 추정된 보존비용은 과소평가 되었을 가능성이 높다. 이는 선행연구에서 추출된 디지털 보존 비용요소의 일부분만이 연구대상 도서관에서 수행되고 있었기 때문이다.

즉 대상 도서관에서 수행되지 않는 디지털 서고 및 보존 메타데이터 등의 비용이 추가된다면 지금의 추정치보다 훨씬 높게 나타날 것이다.

한편 국내에 실제로 디지털 보존을 수행하고 있는 도서관에 의존하여 실증분석을 수행하기에는 절대적인 표본 수가 부족하여 여기에서 나타난 결과를 단순 비교하거나 그대로 적용하는데는 한계가 있음을 밝힌다.

참 고 문 헌

- 이란주, 황신혜. 2003. 전자저널의 관리 및 서비스에 관한 연구. 『도서관』, 58(2): 3-39.
- 이란주, 황신혜. 2003. 전자저널의 효과적인 관리 및 이용자 서비스에 관한 연구: 대학도서관을 중심으로. 『정보관리학회지』, 20(2): 135-156.
- 이소연. 2002. 디지털 아카이빙의 표준화와 OAIIS 참조모형. 『정보관리연구』, 33(3): 45-68.
- 서은경. 2004. 디지털 아카이브의 영구적 보존을 위한 개념적 모형 설계에 관한 연구. 『한국문헌정보학회지』, 38(1): 13-34.
- 신은자. 2001. 전자저널의 아카이빙에 관한 연구. 『정보관리연구』, 18(3): 139-157.
- 한국국가기록연구원, 편. 2003. 레코드관리를 위한 ISO 표준 해설.
- CCSDS. 2002. *Reference Model for an Open Archival Information System (OAIS)*. [cited 2004.12.30]. <<http://www.ccsds.org/document/pdf/ccsds-650.0-B1.pdf>>.
- Chen, Su-Shing. 2001. "The Paradox of Digital Preservation." In *Computer Online* 34(3), no.3. [cited 2004.12.30]. <<http://www.gseis.ucla.edu/us-int/erpare/pdf/ParadoxOfDigitalPreservation.pdf>>.
- Cloonan, Michele Valerie and Shelby Sannett. 2000. "Comparing Preservation Strategies and Practices for Electronic records." *New Review of Academic Librarianship*, 6: 205-16.
- Crespo, Arturo and Hector Garcia-Molina. 2001. "Routing Indices for Peer-to-Peer Systems." [cited 2004.12.30]. <<http://newdbpubs.stanford.edu/pub/2001-48>>.
- Crow, R. 2000 "The Case for Institutional Repositories: A SPARC Position Paper". p.27. <<http://www.arl.org/sparc/IR/ir.html>>.
- "dCollection". [online]. [cited 2004.12.30].

- <<http://www.dcollection.net/project/index.htm>>.
- “DSpace” [online]. [cited 2004.12.30].
<<http://www.dspace.org>>.
- Duranti, L. 1999. The Long term preservation of the uthenticity of Electronic records, 4th General Conference of EASTICA, Hong Kong, 8-12 November.
- “ePrints” [online]. [cited 2004.12.30].
<<http://www.eprints.org>>.
- Granger, S. 2000 “Emulation as a Digital Preservation Strategy.” *D-Lib Magazine*, 6(10). [cited 2004.12.30].
<<http://www.dlib.org/dlib/october00/granger/10granger.html>>
- Hendley, Tony. 1998. “Comparison of Methods & Costs of Digital Preservation.” [cited 2004.12.30].
<<http://www.ukoln.ac.uk/services/elib/papers/tavistock/hendley/hendley.htm>>.
- Hodge, Gale H. 2000. “Best Practices for Digital Archiving: An Information Life Cycle Approach.” *D-Lib Magazine*, 6(1). [cited 2004.12.30].
<<http://www.dlib.org/dlib/january00/01hodge.html>>.
- Lawrence, G.W. et al. 2000. “Risk Management of Digital Information: a File Format Investigation.” Council on Library and Information Resources.
<<http://www.diglib.org/preserve.htm>>.
- Michele, V. Cloonan and Shelby Sanett. 2000. Comparing Preservation Strategies and Practices for Electronic records, University of California. [cited 2004.12.30].
<<http://www.rlg.org/eventspres-2000/cloonan.html>>.
- NARA, 2000. Records Management guidance for Agencies Implimenting Electronic Signature technologies. [cited 2001.7.9].
<<http://www.nara.org/records/policy/gpea.html>>.
- NEDLIB. “NEDLIB Glossary” 2000.[cited 2004.12.10].
<<http://www.kb.nl/coop/nedlib/glossary.pdf>>.
- Rothenberg, J. 1999. “Ensuring the Longevity of Digital Documents.” *Scientific American*, 272(1): 46-47.
- Russell, Kelly and Ellis Weinberger. 2000. “Cost Elements of Digital Preservation.” [cited 2004.12.30].
<<http://www.leeds.ac.uk/cedars/documents/CIW01r.html>>.
- Sanett, Shelby. 2002. “Toward Developing a Framework of Cost Elements for Preserving Authentic Electronic Records into Perpetuity.” *College & Research Libraries*, September: 388-404.
- Waters, Donald J. 1998. “Transforming Libraries Through Digital Preservation.” *Collection Management*, 22(3): 104.