정보관리연구, vol.41, no.4 2010, pp.205-226

# 디지털 콘텐츠의 무결성 유지를 위한 장기적인 보존 정책에 관한 연구*

## A Study on the Preservation Policy for Maintaining the Integrity of Digital Contents

정 영 미** · 윤 화 묵*** · 김 정 택****
Young-Mi Jung · Hwa-Mook Yoon · Jeong-Taek Kim

차 례

1. 서론
2. 이론적 배경
3. 장기적인 보존 정책과 전략
4. 장기적인 보존 전략의 세부 항목과 요건
5. 결론 및 제언
· 참고문헌

초 록

디지털 콘텐츠의 장기적인 보존 정책에서 디지털 콘텐츠의 무결성 유지를 보장하는 것은 가장 중요한 부분 중 하나이다. 그래서 본 연구는 장기적인 관점에서 디지털 콘텐츠의 무결성 유지를 보장할 수 있는 명문화된 보존 정책의 기틀을 마련하기 위해, 실행 수준의 장기적인 보존 정책에 포함되어야 하는 항목들과 세부적인 요건들을 도출해보았다. 특히 해외의 디지털 아카이빙 사례들 중에서 무결성 유지에 초점을 맞춘 장기적인 보존 정책들을 조사하였고 저장 매체와 파일 포맷, 매체와 포맷 변환, 재난 방지와 복구 정책, 관리와 취급, 검사와 모니터링, 그리고 보안의 관점에서 분석하였다.

키 워 드

무결성, 디지털 아카이빙, 보존 정책, 디지털 콘텐츠, 보존 전략

* 본 논문은 KISTI의 정책과제인 '수명주기 기반 디지털 콘텐츠 아카이빙 정책 연구'의 일부 내용임.
** 동의대학교 문헌정보학과 조교수
(Assistant Professor, Dept. of Library and Information Science, Dong-Eui University, yomjung@deu.ac.kr)
*** 한국과학기술정보연구원 정보기술연구실 책임연구원
(Principal Researcher, Dept. of Information Technology Research, KISTI, hmyoon@kisti.re.kr)
**** 배재대학교 도서관 사서(교신저자)
(Corresponding Author, Librarian, Central Library, Paichai University, kjt@pcu.ac.kr)
· 논문접수일자: 2010년 8월 25일
· 최종심사일자: 2010년 9월 30일
· 게재확정일자: 2010년 10월 11일

**ABSTRACT**

Assuring that the integrity of digital contents is reliably maintained is an important component of long-term digital preservation. The purpose of this study was to examine and analysis the best practices of digital preservation policy and provide the basis theory for establishing a preservation policy for maintaining the integrity of digital contents. In order to achieve the purpose of the study, we examined the digital archiving practices for maintaining the integrity of digital contents. Especially, we focused on the storage media and file formats, the media and format conversion, disaster prevention and recovery, management and handling, inspection and monitoring, and security policies.

**KEYWORDS**

Integrity, Digital Archiving, Preservation Policy, Digital Contents, Preservation Strategies

## 1. 서 론

디지털 콘텐츠는 공간을 초월한 접근성, 배포의 용이성, 그리고 즉시성 등의 강력한 장점을 지닌다. 하지만 이것을 정보자원의 한 부분으로 유지·관리하기 위해서는 휘발성과 기술의존성과 같은 심각한 약점을 지니고 있어 많은 문제점들이 유발된다. 그래서 장기적인 관점에서 정보자원을 유지·관리하여 서비스하고 후대에까지 보존해서 전달해야 하는 도서관과 정보서비스 기관의 측면에서 디지털 콘텐츠의 수집과 보존은 매우 큰 도전 과제이다.

세계의 많은 나라들이 국가적 또는 조직적 차원에서 디지털 콘텐츠의 장기적인 보존을 위해 지난 십여 년 간 노력해왔다. 미국의 NDIIPP(National Digital Information Infrastructure and Preservation Program), 영국의 DCC(Digital Curation Centre), JISC에서 지원한 CEDARS(CURL Exemplars in Digital Archives)와 DPC(Digital Preservation Coalition) 등 디지털 아카이빙과 관련한 많은 대규모의 프로젝트들이 진행되어 왔다.

우리나라의 경우는 국가기록원이나 국립중앙도서관 등을 중심으로 디지털 콘텐츠의 장기 보존을 위한 정책이나 전략 수립에 관한 몇 차례의 연구(이소연 2004; 최원태 2005; 한국국가기록연구원 2005; 최원태 2006 등)들을 수행하였다. 그리고 초기단계이긴 하지만 현재 국립중앙도서관은 디지털 자원을 효과적으로 수집/보존/서비스하기 위해 개발한 OASIS(Online Archiving & Searching Internet Sources) 시스템을 구축·운영하고 있다. 그러나 이들 기관을 제외한 디지털 자원들을 대량으로 수집·관리하고 있는 많은 기관들은

명문화된 장기적인 보존 정책을 가지고 있지 않고 현실적인 정책 수립에 유용할 만한 토대 연구들은 여전히 부족한 실정이다. 더군다나 정책의 보다 실행적인 수준에서 장기적인 보존 방법 및 전략에 관한 연구는 더욱 미비하다.

따라서 본 연구에서는 디지털 콘텐츠의 장기적인 보존 정책 수립을 위한 기초적인 토대로서 디지털 콘텐츠의 무결성을 유지할 수 있는 실행적인 수준의 장기적인 보존 전략에 포함되어야 하는 항목들을 도출하고자 하였다. 이를 위해 본 연구에서는 디지털 콘텐츠의 장기적인 보존 정책과 전략적인 관점에서 무결성 개념과 이론들을 살펴보고 이를 위한 선행연구들을 살펴보았다. 그리고 세계의 기관들에서 제공하는 디지털 콘텐츠의 무결성 유지를 위한 명문화된 장기적인 보존 정책 사례를 조사·분석하였다. 특히 정책의 실행적 수준에서의 디지털 콘텐츠의 축적 및 저장 방법과 디지털 콘텐츠의 장기적인 유지·관리와 운영에 관한 요소들을 조사·분석하였다. 그 결과 무결성 유지를 보장하는 디지털 콘텐츠의 장기적인 보존 정책을 위해 실행적 수준에서 포함되어야 하는 세부 항목들과 요건들을 도출하였다.

## 2. 이론적 배경

### 2.1 디지털 콘텐츠와 무결성 개념

디지털 콘텐츠는 부호, 문자, 음성, 음향, 영상 정보 등을 디지털 포맷으로 가공·처리한 정보 자원으로 일반적으로 디지털 저장매체나 온라인상의 정보통신망을 통해 전달된다. 디지털 콘텐츠는 크게 아날로그로 생산된 이력없이 처음부터 디지털로 생산된 것(born digital)과 아날로그 형태의 자원이 디지털 포맷으로 변환된 것(digitization)을 모두 포함한다(DPC 2008, 24).

디지털 콘텐츠의 범주에는 논문, 보고서, 회의록, 도서 등의 기존의 일반적인 학술 정보원의 모든 유형뿐만 아니라 각종 멀티미디어 자원들과 더불어 웹 페이지들도 포함된다. 단, 디지털 콘텐츠의 장기 보존하고자 하는 해당 기관의 정의에 따라 구체적인 디지털 콘텐츠의 범위 및 대상이 제한될 수 있을 것이다.

Higgins(2009, 2)는 디지털 콘텐츠를 장기 보존하는 궁극적인 목적을 다음의 세 가지라고 언급하였다. 첫째는 현재와 미래의 이용자를 위해 일정 기간 동안 이용 가능하도록 디지털 콘텐츠의 지속성을 제공하는 것, 둘째는 데이터의 진본성이 보장되고 신뢰할 수 있도록 무결성을 보장하는 것, 그리고 마지막으로는 미래의 잠재적 이용자에게 적합한 방식으로 데이터를 배치하고 이용할 수 있도록 접근성을 유지하는 것이라고 하였다. 디지털 콘텐츠의 지속성과 접근성도 장기적인 보존 전략에서 중요하지만 이를 통해 실제로 접근하게 되는 디지털 콘텐츠 자체의 무결성이 보장되지 않으면 나머지 요소들도 의미 없는 것이 될 수 있다. 그래서 디지털 콘텐츠의 장기적인 보존 전략에

서 디지털 콘텐츠의 무결성 유지를 보장하는 것은 가장 중요한 부분이라 할 수 있다. 만일의 경우에 대비한 디지털 아카이빙 시스템에 요구되는 위험관리, 철저한 복구와 회복 기능도 결국은 디지털 콘텐츠의 무결성 유지를 보장하기 위한 하나의 방법이다. 즉 디지털 콘텐츠의 무결성을 보장한 장기적인 보존 전략의 수립은 신뢰성 있는 디지털 아카이빙의 필수적인 부분이다.

디지털 아카이빙에서 중요한 개념인 디지털 콘텐츠의 무결성은 다양하게 해석될 수 있다. 완전한 전자기록의 장기보존을 위한 국제적인 연구 프로젝트인 InterPARES (International Research on Permanent Authentic Records in Electronic System) 에서는 무결성을 모든 본질적인 측면에서 완전하고 변경되지 않은 상태(InterPARES 2 Project 2008, 120)라고 정의하고 있다.

Factor 등(2009, 4-5)은 디지털 콘텐츠의 무결성을 장기적인 보존의 근본적인 개념 중 하나로 자원의 모든 측면에서 완전하고 깨끗한 것을 의미한다고 언급하였다. 또한 데이터의 물리적인 무결성, 즉 예를 들면, 원래의 비트 스트림(bit stream)은 일부 변경될 수도 있지만 내용의 구조와 필수적인 구성요소는 완전하게 보존되어야 한다고 제안하고 있다.

미네소타대학(University of Minnesota)의 디지털 보존 정책에서는 지적 콘텐츠로의 접근은 디지털 객체의 무결성과 유용성 유지에 의존하며 그래서 디지털 콘텐츠의 무결성은 결국 모든 디지털 객체의 영속성(fixity)유지라고 명시하고 있다. 이 대학에서는 디지털 객체의 영속성 유지를 위해 체크섬 알고리즘을 이용하여 최초의 저장 당시의 디지털 객체의 비트 스트림과 현재의 비트 스트림을 검사하고 있다.

이와 같이 디지털 콘텐츠의 무결성 유지는 디지털 콘텐츠의 저장방법과 그것의 유지에 관한 실질적이고 기술적인 전략들과 관련된다. 특히 디지털 콘텐츠의 저장 방법은 접근과 보존 둘 다에 영향을 미칠 수 있는 중요한 부분이다. 바꾸어 말하면 디지털 보존은 완전성(authenticity), 무결성, 그리고 기능성을 유지하고 있는 디지털 콘텐츠에 계속적인 접근을 보장하는 모든 활동들을 포함한다(DPC 2008, 103). 다음은 장기적인 보존 정책의 관점에서 디지털 콘텐츠의 무결성 유지에 관한 선행 연구들을 살펴보았다.

### 2.2 장기적인 보존 정책과 무결성 유지

디지털 콘텐츠의 무결성 유지를 위한 장기적인 보존 정책에서 실행적인 수준의 보존 방법 및 전략 수립에는 어떤 세부적인 활동들이 포함되는지, 어떤 결정 요소들이 필요한지를 살펴보기 위해 관련 선행연구들을 살펴보았다.

장기적인 보존과 관련된 구체적이고 세부적인 전략 및 활동 제시에 앞서, 디지털 콘텐츠를 보존하고자 하는 기관은 우선적으로 디지털 콘텐츠의 보존 주체가 어디인지, 어떤 자원들이 보존 될 것인지, 그리고 얼마동안 보존할 것인지와 같은 질문에 대한 명확한 결정을

내린 뒤에 구체적인 보존 전략과 활동들을 수립할 수 있을 것이다(OCLC 2006, 3-4). 이와 관련하여 디지털 콘텐츠의 장기 보존 정책에 관한 다수의 선행 연구들이 거시적인 관점에서 정책의 프레임워크나 포함되어야 할 요소들을 제시하고 있다(Beagrie 2008; NLA 2008; OCLC 2006; 안영희, 박옥화 2010 등).

거시적 관점의 디지털 아카이빙 정책은 보존의 목적 및 목표, 범위, 협력기관의 역할, 보존 전략, 품질관리, 용어 정의나 정책의 버전 정보와 같은 기타 정보 등으로 구성될 것이다. 본 연구는 이 중에서도 실행적인 수준에서의 정책에 접근하기 위해 무결성 유지를 위한 보존 전략에 초점을 맞춘 것이다.

일단 거시적인 관점에서 디지털 아카이빙의 대상과 범위가 결정되면 장기적으로 무결성과 접근성을 보장하기 위한 구체적인 보존 전략이 수립될 수 있는데, 이를 위해 OCLC(2006, 5)는 다음과 같은 전략들을 포함하고 있다.

첫째, 일반적으로 사용되는 파일 포맷과 소프트웨어와 같은 기술적인 변수들에 의해 콘텐츠 손실의 위험을 측정하는 것이다.

둘째, 포맷 변환이나 다른 보존 활동의 유형이나 수준을 결정하기 위해 디지털 콘텐츠 객체를 평가하는 것이다.

셋째, 각 객체 유형을 위해 필요한 적절한 메타데이터와 그것을 어떻게 객체와 연계할 것인가를 결정하는 것이다.

그리고 마지막으로 콘텐츠로의 접근을 제공하는 것이다.

Beagrie(2008)은 대규모의 디지털 보존 정책들을 조사·분석하여 디지털 콘텐츠의 장기적인 보존 전략 수립을 위한 정책의 기술적인 이행과 전략 결정 과정에서 보존 전략에 필요한 세부적인 활동들을 제시하였다.

이를 살펴보면 〈표 1〉과 같이 장기적인 보

〈표 1〉 장기적인 보존 전략과 세부 활동

| 보존 전략 | 세부 활동 |
|---|---|
| 보존 방법 | • 비트 스트림 보존 방법 • 오픈 포맷으로 변환, 번역, 에뮬레이션 방법<br>• 원본 유지 방법 등을 결정 |
| 정보 수집 | • 원본 버전과 새로운 버전을 함께 수집해야하는지 결정<br>• 포맷 변환으로 인한 새로운 객체에 대한 저작권의 정당성<br>• 고유한 ID/명명의 부여 방법 • 리포지터리에 객체 수집 방법 결정 |
| 축적 | • 리포지터리에 축적 또는 미러 사이트에 축적할 것인지 결정<br>• 갱신은 얼마나 정기적으로 수행할 것인지 결정<br>• 백업은 언제 수행할 것인지 결정<br>• 다크 아카이브 또는 정규적인 접근 아카이브를 구축할 것인지 결정 |
| 데이터 관리 | • 어떤 메타데이터 표준과 버전을 사용할 것인지 결정<br>• OAIS 권고안을 따를 것인지 여부 |
| 행정 | • 데이터 선택과 폐기에 대한 기준 결정 • 질적 통제를 위한 점검 횟수와 시기 결정 |
| 접근 | • 접근 권한 결정 • 기타 다른 언어 제공 여부 |

[자료: Beagrie 2008, 31-34의 내용을 재구성함]

존을 위한 전략과 세부 활동에는 보존 방법, 데이터 관리, 행정, 접근 등의 수집된 데이터의 유지·관리뿐만 아니라 디지털 콘텐츠의 수집 단계에서부터 축적까지 디지털 콘텐츠의 수명주기에 따른 모든 단계가 포함되어 있다.

이윤주와 이소연(2008)은 InterPARES 성과물에 기초한 진본 전자기록의 장기보존을 위한 정책 프레임워크를 제안하였고 이것은 전자기록의 장기보존을 위한 원칙, 정책수립, 진본성, 그리고 기록 관리 과정의 네 가지 범주로 구성되었다. 특히 세 번째 범주의 진본성에 무결성 유지를 위한 원칙과 지침이 포함되어 있는데 그것은 ① 접근 권한 설정 및 실행, ② 기록의 손실과 손상에 대한 예방조치와 문제 발견 조치 및 복구절차 확립 이행, ③ 허가받지 않은 행위로부터 전자기록을 보호, ④ 예기치 않은 손실 및 수정이 발생하지 않도록 전자기록을 보호, ⑤ 매체 노화에 대한 대비와 절차를 수립 및 이해, 그리고 마지막으로 ⑥ 하드웨어 및 소프트웨어의 노화에 대비한 조치를 취하라는 것으로 구성되어 있다(이윤주, 이소연 2008, 221).

앞에서 살펴본 선행 연구들에 의하면 무결성 유지를 위해 디지털 콘텐츠의 실행적 수준의 보존 전략에 포함해야 하는 세부적인 요소에는 저장 매체와 파일 포맷, 매체와 포맷의 변환, 재난 방지와 복구 정책, 관리와 취급, 검사와 모니터링, 그리고 보안 등으로 요약할 수 있다. 이와 같은 선행 연구의 검토 결과는 다음에서 조사한 디지털 아카이빙 정책의 실례를 통해 보다 현실화되고 구체화될 것이다.

## 3. 장기적인 보존 정책과 전략

세계 각국은 국가적 또는 조직적 차원에서 디지털 콘텐츠의 장기적인 보존을 위한 다양한 프로젝트들을 수행해왔고 그 결과 디지털 콘텐츠의 장기적인 보존을 위한 명확한 정책들이 수립되어 왔다. 그 중에서 다수의 장기적인 보존 정책들이 거시적인 관점에서 보존 전략 및 방법을 논의하였고 보다 미시적 관점에서 무결성 유지를 위한 실행적인 수준의 보존 전략들을 포함하고 있는 명문화된 정책들은 많지 않다.

디지털 콘텐츠 아카이빙 사례는 컨소시움형, 기관주도형, 단일 국가주도형 등 매우 다양하고 많지만 본 논문에는 포함하지 않았다. 자세한 내용은 본 논문이 기반을 두고 있는 정책과제 보고서인『수명주기 기반 디지털 콘텐츠 아카이빙 정책 연구』의 제2장을 참조하면 된다. 그래서 본 논문에는 이들 중에서 명문화된 정책을 표방하거나 실운영하고 있는 사례를 살펴보고자 한다. 특히 우선적으로 거시적인 관점의 장기적인 보존 정책 중 보존 주체별 대표적인 사례를 간단하게 살펴보면 다음과 같다.

CEDARS 프로젝트는 영국의 디지털 보존 관련 연구를 선도한 대규모 프로젝트로 JISC의 지원과 대학도서관 컨소시엄인 CURL(Con-

sortium of University Research Libraries)의 관리로 영국의 주요 대학을 중심으로 대학들 간의 디지털자료 공유와 보존의 필요성으로 1998년부터 시작되었다. 이후 CEDARS는 디지털 자료의 장기보존에 관련된 전략과 실제적인 방법론 개발에 초점을 맞추고 도서관에서의 디지털자료 장기보존의 지침을 마련하였다. 이들의 지침은 크게 메타데이터 보존, 지적 재산권, 장서 관리, 기술적 전략, 디지털 아카이빙 프로토타입의 주요영역으로 구성되어 있다. 특히 본 연구에서 다루고자 하는 디지털자료의 장기적인 보존 전략과 관련된 4번째 지침(Cedars Project 2002, 2-14)은 정책의 필요성, 디지털자료의 수집에 관한 첫 번째 단계, 수집과 조직, 표준 포맷과 포맷 마이그레이션, 매체의 수명, 수집 매체와 포맷, 중요성 평가, 비용, 보존 메타데이터, 저장, 제적에 관한 전반적이고 포괄적인 항목들을 포함하고 있다. 특히 이것은 책임 소재의 명확한 분담에 대해 중요하게 다루고 있다.

코넬대학도서관(Cornell University Library 2004)의 디지털 보존 정책 프레임워크는 디지털 보존에서 중요한 모든 단계를 체계적으로 포함하고 있다. 하지만 이 정책은 기술적인 전략과 같은 세부적인 내용은 많이 포함하고 있지 않아 상세하진 않지만 전반적인 많은 정보로 이루어져 있다. 여기에는 정책의 목표와 목적, 관계자들의 책임과 역할에 대지만 령, 범위, 운영 원칙, 자료의 선택과 수집, 보존활동, 접근/이용, 협력에 관한 전반적인 정보를 제공한다. 특히 이것은 관계자들의 역할과 책임에 관한 섹션과 법적 가이드라인 등이 강화되어 있다.

영국왕립도서관(British Library 2006)은 대규모의 매우 다양한 자원들에 대한 보다 포괄적이고 광범위한 디지털 보존 정책을 제시하고 있다. 그래서 대규모의 이질적인 자원을 포함하고자 하는 디지털 아카이브 기관들에게 좋은 사례가 될 수 있다. 여기에서 제공하는 디지털 보존 전략에는 사명, 목적, 원칙들, 디지털 자료를 수집하고 유지하는 지침, 디지털 보존팀의 역할, 디지털 보존을 이행하기 위한 DOM 프로그램에 관한 설명, 위험관리에 대한 내용을 포함하고 있다.

이 외에도 디지털 콘텐츠의 장기적인 보존 정책 수립을 위한 거시적 관점을 제공하는 다수의 정책들이 존재한다. 하지만 본 연구의 목적에 초점을 맞추기 위해 실행적 수준에서 무결성 유지를 위한 장기적인 보존 전략을 제공하는 정책을 중심으로 살펴보았다.

호주국립도서관(National Library of Australia 2008)은 디지털 보존 정책을 통해 보존을 위한 핵심 항목뿐만 아니라 상세한 실행 단계를 포함하고 있다. 특히 비트 스트림의 무결성 강화를 위한 전략으로 ① 비트 스트림에 관한 복사본 유지, ② 데이터 보안과 백업 관리, ③ 디지털 자원의 질적 관리(감사)와 문서화, ④ 복본을 소유할 수 있는 권리 및 보존을 가능하게 하는 권리 확보를 제시하고 있다.

OCLC(2006)는 그들의 디지털 아카이브 보

존 정책에서 보다 기술적이고 세부적인 정책을 제공한다. 여기에서는 디지털 콘텐츠 접근 환경을 정의하고 기술 변화를 포함하는 보존 계획을 권고하면서 각 파일 포맷은 그것의 특징에 따른 각각의 보존 계획이 필요하고 위험성 평가가 이루어져야 한다고 명시하고 있다. 또한 이러한 작업은 연속적이어야 한다고 제시하고 있다. 그리고 OCLC는 구체적인 보존 전략과 더불어 수집된 콘텐츠의 데이터 무결성과 접근의 지속성을 보장하기 위한 보다 실행적인 정책을 제시하고 있다. 그것은 데이터 관리, 백업 정책, 저장 시설과 보안, 재난방지와 복구, 정책 업데이트에 관한 정보를 포함하고 있다.

예일대학(Yale University 2005)의 디지털 보존 정책은 전반적인 정책과 더불어 세부적이고 명확한 정책을 포함하고 있다. 여기에는 보존되는 디지털 콘텐츠의 식별자, 파일 포맷, 보존될 디지털 콘텐츠의 소스와 수집 수준, 예산, 접근, 그리고 인증 등에 관한 정책을 제공한다.

그리고 미네소타대학(University of Minnesota 2007)의 디지털 보존 정책은 조직적인 차원에서의 정책보다는 기술적인 전략을 강조한 정책의 대표적인 사례이다. 이 정책에서는 디지털 콘텐츠의 장기적인 안정성을 위해 체크섬 알고리즘을 사용하여 디지털 객체의 비트스트림을 검사하도록 하고, 접근성을 위해 안정된 메타데이터와 보안과 백업 서비스를 제공하도록 하고 장기적인 유용성을 위해 마이그레이션이나 에뮬레이션 전략을 통해 기술적인 변환을 시행하도록 명시하고 있다. 또한 개별적인 파일 포맷을 위한 세부적인 보존 전략을 제공하고 있다.

Higgins(2009)는 DCC의 큐레이션 수명주기 모델에서 디지털 콘텐츠의 무결성 보장과 접근성을 유지하기 위한 정책으로 ① 충분한 표현정보의 기록, ② 지적 재산권과 기타 권리 관리, ③ 디지털 자원을 신뢰성 있게 유지할 수 있는 능력, ④ 접근성에 영향을 미치는 기술변화에 대해 모니터링해야 한다는 것을 명시하고 있다.

국내의 디지털 아카이빙 기관 중 국가기록원은 그들의 홈페이지 상에서 장기보존을 위한 전략으로 장기보존 포맷, 문서보존 포맷, 매체변환, 보안, 인증기술에 대한 다소 간단한 항목들을 명문화해놓고 있다. 장기보존 포맷은 기록물과 관련된 모든 설명정보를 함께 포함하고 검색이 용이하도록 전체는 XML로 캡슐화하고, 문서보존 포맷으로는 PDF/A를 선택하고 있다. 보존 매체는 두 가지로 구분하여 관리하고 전자서명을 이용하여 보안 및 인증을 제공하도록 하고 있다.

앞에서 살펴본 바와 같이 디지털 콘텐츠의 장기적인 보존 정책들이 모두 같은 수준에서의 정책을 제공하지는 않는다. 이들 정책들은 필요에 따라 포함하고 있는 내용 구성도 상이할 뿐만 아니라 기술하고 있는 내용의 수준도 때로는 거시적이고 때로는 미시적이다. 거시적인 관점의 정책을 제외하고 보다 세부적인

접근점을 제공하는 정책들을 대상으로 무결성 유지를 위한 실행적 수준의 장기적인 보존 전략의 항목들을 내용에 의해 대구분해서 정리해보면 다음의 〈표 2〉와 같다. 정책에서 접근성 유지를 위한 메타데이터 관리와 디지털 객체의 저작권 확보에 관한 내용들은 표에 포함하지 않았다.

다음 장에서는 선행연구와 정책 사례를 통해 제시된 실행 수준의 장기적인 보존 전략에 포함되어야 하는 대항목의 세부적인 요소들과 요건을 도출하기 위해 더 많은 관련 정책들과 가이드라인, 그리고 디지털 아카이빙 관련 보고서들을 조사·분석하였다.

## 4. 장기적인 보존 전략의 세부 항목과 요건

### 4.1 저장 매체와 파일 포맷

디지털 콘텐츠의 저장 매체와 파일 포맷은 장기적인 보존을 위한 근본적인 조건이며 보다 효율적인 보존 전략을 위해서 자원의 수집 단계에서부터 결정되어야 한다.

#### 4.1.1 저장 매체

다양한 저장 매체들에 대한 이해는 각 매체들이 서로 다른 저장 조건과 보존 요구사항들을 지니고 있고 또한 접근을 위해 서로 다른 소프트웨어와 하드웨어 장비를 요구하기 때문에 중요하다. 저장 매체들은 제각기 저장 능력과 장단점이 다르며 보존과 접근에 있어 다양한 수준의 안전성을 지닌다.

물리적인 저장 매체에 대한 변환 작업은 매우 단순하고 쉽지만 대부분의 전자적인 매체들이 접근을 제공하는 하드웨어와 소프트웨어의 빠른 기술적 퇴화에 따라 매우 빨리 사용불가능하게 된다. 그래서 저장 매체에 대한 적절한 선택은 모든 보존 전략을 위한 아주 기초적이면서도 여전히 필수적인 요소이다.

현재 가장 많이 사용되고 있는 저장 매체는 마그네틱 매체와 광학 매체이다. 마그네틱 매

〈표 2〉 무결성 유지를 위한 장기적인 보존 전략 내용

| 정책 사례 / 보존 전략 내용 | NLA | OCLC | DCC | Yale Univ. | Minnesota Univ. | 국가기록원 |
|---|---|---|---|---|---|---|
| 저장매체와 파일 포맷 | ○ | ○ | ○ | ○ | ○ | ○ |
| 매체와 포맷 변환 | ○ | ○ | ○ | | ○ | ○ |
| 재난방지와 복구정책 | ○ | ○ | ○ | | ○ | |
| 관리와 취급 | | ○ | | | | |
| 검사와 모니터링 | ○ | ○ | ○ | | ○ | |
| 접근과 보안 | ○ | ○ | ○ | ○ | | ○ |

체는 마그네틱 테이프 릴, 카트리지, 카세트와 디스크 종류인 하드 디스크, 플로피 디스크가 포함된다. 이것은 다방면에 사용될 수 있고 경제적이며 저장 능력과 유지 능력이 뛰어나다. 반면 광학 매체는 마그네틱 매체에 비해 대용량의 정보를 저장할 수 있고 반영구적이고 복제비용이 저렴하며 분배가 용이한 특징을 지닌다. 그래서 그림, 사운드, 동영상 등의 대용량 데이터를 취급하는 멀티미디어 자원을 저장하기 편리하다.

보다 효율적인 장기 보존을 위해 저장 매체의 수명에 대한 이해는 매우 중요하며 인쇄 매체인 종이보다 그 수명이 훨씬 짧고 환경조건에 민감하기 때문에 정기적으로 재 가공될 필요가 있다. 저장 매체는 다양한 온도와 습도, 사용 빈도에 따라 일정한 수명을 지니는데 DPC(2008, 154)에서 밝힌 저장 매체의 환경적 조건에 따른 수명주기는 일반적으로 〈표 3〉과 같다. 매체의 보관 환경이 상대습도 25와 10℃인 경우 마그네틱 테이프 카트리지와 CD/DVD는 75년이고 CD-ROM은 최장 30년에 불가하다. 저장 매체에 대한 수명은 장기 보존을 위한 위험 관리를 위해 중요한 기준점이 될 수 있다.

#### 4.1.2 파일 포맷과 표준

파일 포맷의 선택 범위는 매우 다양하며 그 기준 또한 다양할 수 있다. 파일 포맷 선택은 수집 단계에서부터 결정되며, 수집한 자료가 포맷이 맞지 않는 경우에는 정책에서 선택된 파일 포맷으로 자료들을 변환시켜 저장할 필요가 있다.

디지털 콘텐츠의 휘발성이나 기술 의존성 등과 같은 장기적인 보존 관점에서의 주요 문제점들이 파일 포맷과 관련성이 있다. 그래서 보다 안정적이고 효율적인 보존을 위해서는 파일 포맷의 선택이 매우 중요하다. 많은 학자들에 의해 장기적인 보존을 위한 파일 포맷의 선택 기준들이 제시되었다.

Todd(2009, 13-14)는 장기적인 보존을 위한 디지털 콘텐츠의 파일 포맷 선택 기준으로 포맷의 사용 범위에 해당하는 채택성, 기술적인 의존 여부, 공공성, 투명성, 메타데이터의 지원 여부를 제시하였다.

〈표 3〉 저장 매체의 환경적 조건에 따른 수명

| 조건 / 매체의 종류 | 25RH 10℃ | 30RH 15℃ | 40RH 20℃ | 50RH 25℃ | 50RH 28℃ |
|---|---|---|---|---|---|
| D3 마그네틱 테이프 | 50년 | 25년 | 15년 | 3년 | 1년 |
| DLT 마그네틱 테이프 카트리지 | 75년 | 40년 | 15년 | 3년 | 1년 |
| CD/DVD | 75년 | 40년 | 20년 | 10년 | 2년 |
| CD-ROM | 30년 | 15년 | 3년 | 9개월 | 3개월 |

National Archives의 Brown(2003, 5)은 개방적인 표준, 보편성, 안정성, 메타데이터 지원, 특징, 상호운용성, 그리고 에러감지 기능을 파일 포맷의 선택 기준으로 제안하였다. 이와 별도로 마이그레이션을 위한 파일 포맷의 선택에 있어서 고려해야 할 기준으로 확실성, 처리능력, 표현력의 세 가지 항목을 추가로 제시하고 있다.

앞에서 살펴본 바와 같이 파일 포맷의 선택과 관련된 많은 고려사항들이 존재한다. 하지만 모든 경우에 모든 기준에 부합하는 파일 포맷을 선택하기란 쉽지 않다. 또한 분명한 것은 파일 포맷은 계속 발전하여 새로운 버전들이 나타나고 이전의 버전들은 기술적으로 퇴화될 것이다. 그렇지만 보다 최장의 기간 동안 지속적이고 안정적으로 사용 가능한 파일 포맷의 선택이 보존 담당자나 기관에게는 효율적인 일이 될 것이다.

이런 이유로 보존의 대상이 되는 디지털 객체나 주관 기관에 따라 파일 포맷의 선택 기준은 다소 상이할 수 있지만, 디지털 콘텐츠의 장기적인 보존을 위한 파일 포맷은 적절한 표준 포맷을 따름으로 어느 정도 도달할 수 있다. 표준 포맷을 채택함으로 마이그레이션에서 발생하는 문제점들을 줄일 수 있고 관련 기관과 각 부문들 간의 디지털 아카이빙에 대한 협력을 단순하고 쉽게 해 줄 수 있다. 그리고 비교적 적은 수의 표준 포맷들을 채택하는 것이 단기적 그리고 장기적인 관리에 더 쉬울 것이다.

DPC(2008, 103)는 표준 포맷을 선택함과 동시에 디지털 콘텐츠 제작이나 수집 시 장기적인 관점에서 파일 포맷을 선택할 때 좋은 실제는 ① 공개되고 소유권이 없는 파일 포맷을 사용, ② 자원에 대한 문서화와 표준과 일치하는 메타데이터 제공, ③ 온라인 디지털 자원에 대한 영구적인 식별자 할당과 같은 요소들을 포함해야 한다고 제시하였다.

또한 ICPSR(2010)의 디지털 보존에 대한 정책은 장기적인 보존을 위한 파일 포맷의 선택 시 고려해야 할 사항으로 파일 포맷의 광범위한 채택, 이전 버전과의 호환성, 좋은 메타데이터 지원, 복잡하지 않으면서 많은 기능을 지원, 내부에 에러 체킹 기능, 그리고 합리적인 업그레이드 주기를 제시하고 있다.

앞에서 언급한 장기적인 보존을 위한 파일 포맷의 선택 기준을 종합해보면 다음과 같은 결론을 얻을 수 있다. 즉, 장기적인 보존을 위한 파일 포맷의 선택 기준으로는,

첫째, 기본적으로 표준 포맷
둘째, 무료로 사용가능한 공공성
셋째, 광범위한 보편성
넷째, 메타데이터 지원
다섯째, 호환성
여섯째, 에러 감지 기능
일곱째, 기타의 추가 기능이 포함될 수 있다.

미네소타 대학의 디지털 보존 관리 위원회(UDC: University of Minnesota Digital Conservancy)는 디지털 콘텐츠를 위해 사용가능한 표준적인 파일 포맷들과 이들 포맷들

의 장기 보존을 위한 포맷으로의 지원기술 정도에 따라 〈표 4〉와 같이 세 가지의 수준으로 분류해 놓았다. 장기적인 보존에 대한 기술 지원이 비교적 우수한 파일 포맷으로는 텍스트를 위해 PDF, Plain Text, SGML, 그리고 XML가 제안되었고, 이미지 자원을 위해서는 JPEG, TIFF 파일 포맷, 오디오를 위해서는 Wave 파일 포맷이 제안되었다.

현 시점에서는 제시된 포맷이 장기적인 보존 관점에서 다른 파일 포맷과 비교하여 우수하지만 어떤 포맷이라도 장기적인 보존 관점에서 완전하게 안정적으로 무결성 유지를 보장할 수 있는 단일의 파일 포맷은 존재하지 않는다. 즉, 이것은 저장매체와 같은 이유로 복수의 파일 포맷에 디지털 콘텐츠가 유지될 필요가 있음을 뜻한다. 그래서 다수의 디지털 아카이빙에서 디지털 콘텐츠에 대해 복수의 저장 매체나 다양한 포맷의 복사본을 유지 관리하고 있다. 이것은 매우 소모적인 일이긴 하지만 내·외부의 위험으로부터 디지털 콘텐츠의 무결성을 유지할 수 있는 가장 안전하고 바람직한 방법이다. 또한 이들 포맷도 계속해서 기술적인 진화와 동시에 퇴화가 발생하기 때문에 표준 포맷들을 채택하더라도 장기적인 무결성 유지와 위험 관리를 위해서는 기술적인 추이변화를 계속적으로 주시하여 적절한 시점에 적절한 매체와 포맷으로 변환해야 한다.

앞의 내용을 종합하여 저장 매체와 포맷에 대한 실행 수준의 정책에 포함되어야 하는 장기적인 보존 전략 항목과 요건은 다음의 〈표 5〉와 같다.

〈표 4〉 파일 포맷의 장기적인 보존 지원 수준

| 구 분 | 지원 기술 |
|---|---|
| Level 1(완전한 지원) | • 마이그레이션, 에뮬레이션, 또는 규격화를 포함한 장기적인 파일의 유용성 유지를 위한 알맞은 활동들을 지원함.<br>• 접근과 데이터 영속성을 보장함. |
| Level 2(제한된 지원) | • 퇴화된 포맷에서 다른 포맷으로 파일을 변환시키는 활동과 같이 파일의 유용성 유지를 위한 제한된 활동을 지원함.<br>• 접근과 데이터 영속성을 보장함. |
| Level 3(최소 지원) | • 오직 해당 파일 포맷의 아이템에 접근을 제공함.<br>• 데이터 영속성을 보장함(단, 파일의 물리적인 비트스트림이 변화지 않는 경우 보장됨). |

〈표 5〉 장기적인 보존 전략 항목과 요건: 저장 매체와 파일 포맷

| | | |
|---|---|---|
| **저장 매체와 파일 포맷** | • 효율적이고 안정적인 저장 매체와 파일 포맷의 선택<br>• 복수의 저장 매체와 파일 포맷 유지 | |
| | 저장 매체 | • 저장 매체의 환경적 조건에 따른 수명을 주시 |
| | 파일 포맷 | • 선택기준 결정: 표준포맷, 공공성, 보편성, 메타데이터 지원, 호환성, 에러 감지 기능, 기타의 추가적인 기능<br>• 파일 포맷의 기술적 진화와 퇴화를 주시 |

### 4.2 매체와 포맷 변환

현재 이용하고 있는 그리고 미래에 이용 요구가 있을 자원들은 원본을 유지하면서 장기적인 접근이 가능하도록 지속적인 관리가 필요하다. 이를 위한 저장 매체의 재가공과 재포맷팅 활동은 장기적인 보존 전략을 편리하게 이행할 수 있고 매체의 퇴화를 막기 위한 필수적인 관리 요소이다.

DPC(2008, 104-105)는 매체와 포맷 변환에 대한 세부적인 보존 전략 가이드라인을 제공하는데, 이것은 매우 상세하고 실질적이다. 매체와 포맷 변환이 필요한 시점으로 바이러스나 버그에 걸린 소프트웨어가 발생했을 때, 환경조건에서 공급자가 정의한 매체의 최소 수명이 끝나갈 때, 새로운 저장 장치가 설치될 때, 데이터 자원에 중요한 일시적인 에러가 발견되었을 때 등이 제시되고 있다. 또한 추가적인 이유로 매체의 결점을 보완하기 위해 이에 상응할만한 또 다른 매체에 복사본을 저장하라고 권고하고 있다.

매체와 포맷 변환을 통해 복수의 저장 매체와 파일을 통해 안정성을 유지하는 것과 더불어 매우 중요한 것 중의 하나는 변환을 통해 위협받을 수 있는 디지털 객체 원본의 무결성을 유지하는 것이다. 이를 위해 매체나 포맷 재가공후에는 디지털 콘텐츠의 완전성과 무결성을 보증하기 위해 원 데이터와 비교하는 체크섬 알고리즘을 통해 자동적인 질적 통제가 반드시 필요하다. 또한 혹시 모를 데이터 손실을 위해 원래의 포맷으로 원본의 비트 스트림을 유지하는 것이 바람직하다. 변환과 관련된 이러한 모든 과정은 문서화하거나 보존 메타데이터의 한 부분으로 포함하는 것이 필요할 것이다.

기술적인 변화에 따라 모든 매체나 파일 포맷의 변환은 주기적인 요구사항이고 변환할 매체나 파일 포맷의 선택은 4.1에서 언급한 선택 기준들이 모두 포함될 수 있다. 특히 그 중에서도 매체나 파일 포맷의 장기간 사용은 재포맷팅의 간기를 줄일 수 있기 때문에 무엇보다 중요하다.

매체와 포맷 변환과 관련하여 장기적인 보존 전략에 포함되어야 하는 세부 항목과 요건은 다음의 〈표 6〉과 같다.

### 4.3 재난 방지와 복구 정책

재해로부터의 복구 계획은 디지털 콘텐츠의 수집단계에서부터 포함되어야 하고 우선적인

〈표 6〉 장기적인 보존 전략 항목과 요건: 매체와 포맷 변환

| | |
|---|---|
| 매체와 포맷 변환 | • 기술적인 변화에 따라 매체나 파일 포맷의 변환은 주기적으로 시행<br>• 매체와 파일 포맷의 변환 시기 결정<br>• 매체와 포맷 재가공후에는 원본과의 무결성 검사 방법<br>• 원 데이터의 원래 포맷으로 비트스트림 유지 |

관리 방침에 명시되어야 하는 중요 항목이다. 또한 이 활동은 훈련된 직원에 의해 이루어져야 한다. 이러한 명확하고 준비된 계획만이 재해와 사고로부터의 위험을 최소화할 수 있을 것이다. 재해 방지의 목표는 디지털 아카이브 내의 콘텐츠와 메타데이터를 보호하고 디지털 아카이브의 소프트웨어와 시스템을 보호하는 것이다. 재해 방지와 복구에서 모든 데이터는 동일한 가치를 지닌다.

British Library(2006, 2-3)는 디지털 컬렉션에 대한 백업 파일에 대한 관리 지침으로 "지리적으로 떨어진 장소에(적어도 3곳) 디지털 객체의 복수의 복사본을 저장하여 백업을 쉽게 하고 무결성을 유지하라"고 명시하고 있다. 또한 디스크의 퇴화, 에러나 재난 등으로 데이터를 잃을 수 있는 위험을 위험관리의 한 요소로 정의하고 이를 예방하기 위해 "각 디지털 객체의 복사본은 다수의 사이트에 중복하여 저장하고, 백업 시키고, 무결성을 조사하라"고 비트스트림의 보존을 보장하기 위해 세부적인 재난으로부터의 복구 계획을 수립하고 있다.

Parliamentary Archives(2009)의 디지털 저장 관리 정책 또한 모든 디지털 자원들에 대해 다수의 중복적인 복사본을 유지관리하고 적어도 원본이 보존된 건물 밖의 100km 이상 떨어진 안전한 시설에 하나 이상의 완전한 백업 파일을 유지 관리하라고 명시하고 있다. 백업 파일은 지속적으로 업데이트되고 잘 유지되고 있는지 규칙적인 검사를 수행하고 있다.

그리고 Columbia University Libraries (2006)의 디지털 자원 보존에 대한 정책에서는 디지털 자원의 저장 버전은 언제라도 사용 가능한 공개된 파일 포맷으로 손실 없이 보존할 계획으로 주요한 자료를 위한 상세하고 구체적인 저장 계획을 밝히고 있다. 최근 사용중인 자원들은 규칙적인 백업, 재포맷, 그리고 마이그레이션으로 온라인 접근을 유지할 계획이고 온라인 또는 오프라인 접근 자원의 모든 저장 버전(높은 해상도, 풍부한 캡쳐, 무손실 압축)은 공인된 저장 매체에 축적되고 도서관 디지털 프로그램 부서내의 오프라인에 저장할 계획이다. 이들은 규칙적인 재포맷과 마이그레이션 주기를 지닌다. 최근 온라인이 아닌 저장 버전 자원은 중복적인 오프라인 복사본을 다른 사이트에 저장하고 이러한 모든 온라인과 오프라인 버전들은 콜롬비아 대학 도서관의 로컬 정보제공 관리 시스템을 통해 탐지될 수 있도록 하고 있다.

Economic and Social Research Council (ESRC)과 JISC의 지원에 의해 운영되는 UK Data Archive(2010, 9. 11)는 디지털 콘텐츠에 대한 다양한 복사본을 제작함으로써 재해 복구가 가능하도록 계획하고 있다. 모든 데이터에 대한 각 파일은 원본에 해당하는 Main copy, 독립적인 테이프에 복사해서 저장하는 Shadow copy, 데이터의 세트 당 하나씩 만들어지는 CD-ROM copy, 오프라인 상에서 운영되는 Off-line copy의 적어도 네 개 이상의 복사본을 유지하도록 명시하고 있다. 보다

포괄적으로 비트 스트림을 보존하기 위한 호주국립도서관의 가이드라인 역시 복사본을 통한 재해 방지와 복구 계획을 포함하고 있다. 구체적으로 이것은 복사본 허락과 보존을 이해시키기 위한 저작권자와의 협상, 물리적인 매체의 안정적인 환경조건에 저장, 마그네틱 매체로부터 가능한 한 빨리 데이터를 백업, 정기적인 백업과 복구 방법을 통해 디지털 저장 시스템으로 자료를 복사, 디지털 객체에 영구적인 식별자 할당, 적절한 데이터 보안과 백업, 에러 체크와 정기적인 새로운 미디어로 복사를 포함하여 고품질의 IT 실제에 맞도록 지속적으로 데이터 보호를 관리하라는 항목을 포함하고 있다.

백업의 적절한 시기나 간격에 대해서는 OCLC (2006, 9-10)의 디지털 아카이브 보존 정책에서 언급하고 있는데 저장소 내의 새로운 그리고 변화된 콘텐츠와 메타데이터는 즉시 그리고 중복적으로 디스크에 기록하고 디스크 매체가 아닌 반복되는 백업은 매일 수행하고 로컬에 축적하도록 되어 있다. 그리고 매일 시행되는 백업 작업은 장기간에 걸쳐 유지되고 로컬에 저장되고 이러한 백업 데이터는 다시 시스템 소프트웨어와 오프라인 사이트 시설에 매주 백업되어 저장된다.

앞에서 살펴본 바와 같이 재난으로부터 디지털 콘텐츠의 무결성을 유지하고 재난 발생 시 원 데이터로의 회복을 보장하기 위해 다수의 장기적인 보존 정책에서는 재난 방지 및 복구 정책에 관한 항목을 중요하게 다루고 있음을 볼 수 있었다. 특히 세부적인 전략으로 물리적으로 떨어진 장소에 복수의 복사본을 유지·관리하고 정기적이고 주기적인 백업 파일을 유지할 것을 권고하고 있다. 앞에서 살펴본 재난 방지와 복구 정책에 포함되어야 하는 장기적인 보존 전략의 세부 항목과 요건은 다음의 〈표 7〉과 같다.

### 4.4 관리와 취급

디지털 콘텐츠는 물리적으로 디지털 저장 매체에 저장되어 있다. 앞에서 밝힌 바와 같이 디지털 저장 매체는 전통적인 종이 매체보다 훨씬 더 취급하기 어렵고 다양한 재해로부터 손실을 쉽게 입을 수 있다. 앞에서는 다양한 재해로부터의 방지와 데이터 복구 계획에 대

〈표 7〉 장기적인 보존 전략 항목과 요건: 재난 방지와 복구 정책

| | |
|---|---|
| **재난 방지와 복구 정책** | • 재해로부터의 복구 방안을 계획하고 개발<br>• 복구를 위한 복사본의 개수와 저장 매체의 결정<br>• 복사본의 물리적으로 독립적인 저장 장소, 온라인과 오프라인 사이트 운영<br>• 백업의 시기와 간격 결정<br>• 복구를 위한 전문 인력의 양성 |

해 언급하였고 여기에서는 물리적인 저장매체에 발생할 수 있는 일반적인 관리와 취급을 다루었다. 일반적인 관리상의 소홀에 의한 손실을 막고 보다 지속성을 증가시키기 위해 저장매체는 최적의 환경적 조건을 유지하는 것이 중요하다.

일반적으로 저장매체의 수명은 온도와 상대습도(RH: Relative Humidity)에 크게 영향을 받는다. 다음의 〈표 8〉은 British Standard 4783의 저장 매체에 대한 적절한 환경적 조건을 보여준다(DPC 2008, 108). 각각의 저장 매체에 따라 장기적인 저장이 가능한 온도는 다소 상이하긴 하지만 주로 18~22℃이고 상대습도는 35~45%임을 알 수 있다. 매체에 대한 최적의 환경적 조건을 유지함으로써 저장 매체와 더불어 디지털 콘텐츠가 확보할 수 있는 최장의 수명을 보장받을 수 있다.

일반적으로 유지·관리해야 하는 저장 매체 보관의 온도와 습도 외에도 최적의 환경적 조건을 만족시키기 위한 몇 가지 가이드라인이 DPC(2008, 108)에 의해 제시되었다. 그것을 살펴보면 ① 온도와 같은 중요한 변수가 변하는 이동이 있다면 마그네틱 테이프를 적응시키기 위한 절차와 가이드라인을 세워라, ② 환경적 조건을 관리·감독하는 절차를 수립하라, ③ 먼지와 다른 공기 오염으로부터의 손실 위험을 최소화하라, ④ 저장 구역에서 흡연과 음식 섭취를 금하라, ⑤ 직접적인 햇빛으로부터 멀리 저장하라, ⑥ 매체를 위해 방어벽과 같은 추가적인 보호막을 제공하라, ⑦ 화재나 홍수와 같은 자연적인 재해로부터 위협을 최소화하고 자기장으로부터 저장 매체를 보호하는 저장 시설물을 제공하라, ⑧비디지털 수행 자료들(예를 들면, 코드북) 또한 적절한 환경적 조건에 저장하라는 항목을 포함한다. 이와 같은 저장 매체에 대한 환경적 조건을 꾸준히 유지하기 위해서는 관리 책임자와 역할이 정책에 포함되어야 할 것이다.

〈표 8〉 저장 매체에 대한 적절한 환경적 조건

| 저장 매체 | 운영 | 운영되지 않음 | 장기적인 저장 |
|---|---|---|---|
| 마그네틱 테이프<br>카세트 12.7mm | 18 to 24℃<br>45 to 55% RH | 5 to 32℃<br>5 to 80% RH | 18 to 22℃<br>35 to 45% RH |
| 마그네틱 테이프<br>카트리지 | 10 to 45℃<br>20 to 80% RH | 5 to 45℃<br>20 to 80% RH | 18 to 22℃<br>35 to 45% RH |
| 마그네틱 테이프 4 &<br>8mm helical scan | 5 to 45℃<br>20 to 80% RH | 5 to 45℃<br>20 to 80% RH | 5 to 32℃<br>20 to 60% RH |
| CD-ROM | 10 to 50℃<br>10 to 80% RH | -10 to 50℃<br>5 to 90% RH | 18 to 22℃<br>35 to 45% RH |

디지털 저장 매체의 저장 시설에 대한 관리와 취급 항목과 요건들을 정리해보면 다음의 〈표 9〉와 같다.

### 4.5 검사와 모니터링

디지털 콘텐츠에 대한 검사와 모니터링은 일반적으로 마이그레이션이나 재포맷에 따라 자원이 비의도적 또는 의도적으로 변하지 않았다는 것을 보증하거나 시간의 흐름에 따른 데이터의 무결성과 가독성을 검사하기 위해 필요하다.

이 때 정기적으로 데이터의 가독성을 조사하는 것이 좋으며 조사를 원활하게 수행하기 위해서는 시스템 운영자에 의해 매개변수 세트에 따라 대량의 저장 시스템에서 자동적으로 수행하는 것이 좋다. 체크섬 절차를 이용하여 정기적으로 데이터 파일의 무결성을 조사하는 것도 필요하다. 특히 마이그레이션이나 재포맷과 같은 장기적인 보존 처리과정을 거쳤다면 작업 후 이런 과정은 필수적이다. 이런 무결성 조사 역시 자동적으로 시스템에서 수행될 필요가 있다.

OCLC(2006, 9)의 디지털 아카이브 보존 정책은 데이터의 무결성과 지속적인 접근성을 제공하기 위한 데이터 관리 정책으로 검사와 모니터링의 구체적인 시점을 명시하고 있다. 모든 콘텐츠는 처음에 저장소에 수집될 때 불변성과 바이러스 조사를 받는다. 두 번째로 객체 확인 절차에 의해 수집된 후 며칠 뒤 불변성과 바이러스 검사를 받는다. 두 번째 검사는 객체와 완전한 메타데이터 유지를 위해 필요한 최적의 방법으로 주어진 객체의 모든 구성요소들이 저장되었는지와 모든 객체의 구성요소를 위한 적절한 메타데이터가 존재하는지를 확인하는 것이다. 또한 OCLC는 지속적인 불변성 조사, 바이러스 검사, 그리고 객체 확인을 위해 처음 객체가 수집된 일을 기준으로 연4회 실시하고 있다.

즉, 디지털 콘텐츠에 대해 정기적이고 규칙적인 검사와 모니터링은 무결성 유지를 위해 중요하고 이 뿐만 아니라 파일 포맷이나 저장 매체와 같은 디지털 콘텐츠의 상태에 변화가 있을 때는 상시적인 검사와 모니터링이 필요하다. 이에 관련된 장기적인 보존 전략의 세부 항목과 요건은 〈표 10〉과 같다.

〈표 9〉 장기적인 보존 전략 항목과 요건: 관리와 취급

| | |
|---|---|
| **관리와 취급** | • 저장 매체의 시설을 위한 최적의 환경적 조건에 대한 체크리스트를 개발<br>• 온도, 습도유지, 청결, 햇빛 및 자기장 차단<br>• 저장 시설에 대한 관리 책임자와 역할 명시 |

〈표 10〉 장기적인 보존 전략 항목과 요건: 검사와 모니터링

| | |
|---|---|
| **검사와 모니터링** | • 무결성 검사와 모니터링 방법<br>• 검사 항목 및 내역 결정<br>• 검사와 모니터링 시점와 횟수 |

### 4.6 접근과 보안

모든 디지털 자원이 같은 수준의 보안을 필요로 하는 것은 아니다. 물론 엄격한 보안 절차는 디지털 콘텐츠의 무결성 유지를 위한 기본적인 요소이다. OCLC(2006, 10)는 네트워크와 시스템 보안만 담당하는 정보시스템 보안 전문가인 직원을 따로 두고 있다. 여기에서는 보안을 위해 백업 테이프를 직원만 들어갈 수 있는 통제 구역 내에 보존하고 있다. 그리고 백업 테이프가 보존되어 있는 컴퓨터실에 접근하는 모든 사람과 시간에 대해 기록하고 있다. 또한 화재로 인한 재해를 방지하기 위해 헬론 가스 제압 시스템을 갖추고 있고 기온 조절 시스템도 갖추고 있다. 이러한 물리적인 공간에 대한 보안 외에도 다양한 측면에서의 보안 활동들이 있는데 우연한 또는 의도적인 변화로부터 디지털 자원들을 보호, 디지털 자원의 완전성을 보호, 잠재적이고 내부적인 보안 파괴를 제지, 도난이나 손실로부터 보호, 어떤 법적 그리고 규제 조건들의 준수를 보장, 의무 조건을 만족시키기 위한 검사 흔적을 제공하는 것 등이 포함될 수 있다.

디지털 콘텐츠의 장기적인 보존 정책의 사례는 아니지만 접근과 보안에 관한 장기적인 보존 전략의 세부 항목 및 요건을 도출하기 위해 BS 7799 정보 보안 관리 표준(정성헌 2005)을 참조할 수 있는데 이는 보안에 대한 다음과 같은 상세한 가이드라인을 제공하고 있다. 여기에는 ① 재해 복구 계획을 수립하라, ② 저장 시설물과 처리 영역으로의 접근을 통제하라. 가능한 지역적으로 분산하여 저장하라, ③ 승인되지 않은 접근으로부터 보호하라, ④ 대규모의 저장 시스템들은 검사기능을 설계하고 물리적인 접근 통제를 자동화하라. 자동적인 검사가 수행되지 않는다면 정규적인 랜덤 체크를 수행하라, ⑤ 우연히든 의도적인 변화이든 발생하지 않도록 보장하는 절차들을 수립하라, ⑥ 모든 법적 요구사항을 적합하도록 하라, ⑦ 완전성을 보장하는 절차를 수립하라, ⑧ 이용자 아이디와 패스워드를 사용하고 다른 네트워크 보안 절차를 이용하라, ⑨ 직원의 시스템과 영역 접근에 대한 권한을 정의하라, ⑩ 데이터 보안과 저장 시설물에 대한 특정한 직원의 책임 소재를 할당하라와 같은 항목이 포함된다.

이와 같이 장기적인 보존 정책 내에는 디지털 콘텐츠의 무결성 유지를 위해 저장 매체의 보관 시설에 대한 보안 항목뿐만 아니라 이용자의 접근 권한 제한, 인증과 같은 철저한 보안 계획이 수립되어 있어야 한다. 또한 보안이 무너져서 발생하게 될 데이터 손실에 대비한

〈표 11〉 장기적인 보존 전략 항목과 요건: 접근과 보안

| | | |
|---|---|---|
| **접근과 보안** | • 저장 매체의 보관 시설에 대한 보안<br>• 디지털 콘텐츠 사용의 인증 절차<br>• 재해 복구 방법 및 절차 | • 이용자의 접근 권한 제한<br>• 관리 담당자와 역할 명시 |

복구 정책 또한 명확하게 명시되어 있어야 한다. 접근과 보안에 관한 실행적인 수준의 장기적인 보존 전략에 포함될 세부 항목과 요건은 다음의 〈표 11〉과 같다.

## 5. 결론 및 제언

본 연구에서는 디지털 콘텐츠의 장기적인 보존 정책 수립의 한 부분으로서 무결성 유지 방안을 위한 실행적인 수준에서 장기적인 보존 전략에 포함되어야할 세부 항목과 요건을 도출하였다. 이를 위해 무결성 유지를 중점으로 다루고 있는 다수의 장기 보존 정책을 조사하였고 이를 위한 세부적인 전략 및 활동으로 저장 매체와 파일 포맷, 매체와 포맷 변환, 재난 방지와 복구 정책, 관리와 취급 정책, 검사와 모니터링 정책, 그리고 접근과 보안 정책과 관련된 내용들을 검토해 보았다. 디지털 콘텐츠의 비트 스트림과 메타데이터와 같은 관련 정보를 완전하게 보존하는 것이 무결성 유지를 위한 기본 토대이며 무결성 유지는 곧 디지털 아카이빙의 핵심 요소이기도 하다.

연구의 결과를 정리해보면, 무결성 유지를 위한 디지털 콘텐츠의 장기적인 보존 전략에 포함해야하는 대항목은 6개로, 이것은 다시 저장 매체와 파일 포맷에 5개, 매체와 포맷 변환 정책에 4개, 재난 방지와 복구 정책에 5개, 관리와 취급 정책에 3개, 검사와 모니터링 정책에 3개, 그리고 접근과 보안에 5개의 세부 항목과 요건을 포함한다. 물론 이들 항목들은 적용되는 기관의 특징이나 디지털 콘텐츠의 종류에 따라 다수의 항목으로 구체화되고 명시될 수 있을 것이다. 본문의 내용에는 세부 항목들을 위해 참조 가능한 실질적이고 세부적인 지침과 권고 사항들의 사례들이 포함되어 있다.

본 연구에서는 디지털 콘텐츠의 장기적인 보존 정책을 거시적인 관점에서 분석하지는 않았다. 만약 디지털 콘텐츠를 장기적으로 보존하고 있거나 계획하고 있는 기관에서 보다 명확한 장기적인 보존 정책을 수립하고자 한다면, 우선적으로 디지털 콘텐츠의 보존의 목적과 목표, 보존 주체, 보존 될 자원의 특징과 범위, 요구되는 보존 연한, 요구되는 무결성 유지 수준 등과 같은 거시적 관점에서의 보존의 정의와 특성을 명확하게 정의해야 할 것이다. 그 다음 단계로 실행적인 수준에서 정책의 세부적인 전략과 방법을 수립할 수 있을 것이다. 이 단계에서 본 연구의 결과는 무결성 유지에 대한 기관의 장기보존 전략의 세부항목 및 요건 결정시 유용한 기초 자료가 될 것이다.

## 참고문헌

설문원. 2005.『국가 디지털 아카이빙 체제 구축에 관한 연구』. 대전 : 한국과학기술정보연구원.

안영희, 박옥화. 2010. 디지털 큐레이션 정책을 위한 프레임워크 개발.『한국도서관·정

보학회지』, 41(1): 167-186.

이소연. 2004.『디지털유산 보존을 위한 국가정책수립방안』. 서울 : 국립중앙도서관.

이윤주, 이소연. 2009. 진본 전자기록의 장기보존을 위한 정책프레임워크.『기록학연구』, 19: 193-249.

이재윤, 최원태, 이수상. 2004.『디지털 아카이빙의 현안과 과제』, 서울 : 한국교육학술정보원. RM 2004-14.

정성헌. 2005. "BS 7799(Information Security Management System)". [인용 2010.07.12].
〈www.kab.or.kr/Upload/CWorld_Data/18/01.pdf〉.

최원태. 2005.『디지털 유산 보존과 국가대표도서관의 정책방안』. 서울 : 국립중앙도서관.

최원태. 2006.『OASIS 중장기 정보전략계획 세부과제 도출』. 서울: 국립중앙도서관.

국가기록원 홈페이지.
〈http://www.archives.go.kr〉.

국립중앙도서관 OASIS 홈페이지.
〈http://www.oasis.go.kr〉.

Beagrie, Neil et al. 2008. "Digital Preservation Policies Study". [cited 2010.05.24].
〈http://www.jisc.ac.uk/.../preservation/jiscpolicy_plfinalreport.pdf〉.

British Library. "Digital Preservation Strategy". [cited 2010.05.21].
〈http://www.bl.uk/aboutus/stratplprog/ccare/introduction/digital/digpresstrat.pdf〉.

Brown, Adrian. 2003. "Selecting File Formats for Long-Term Preservations". [cited 2010.06.04].
〈http://www.nationalarchives.gov.uk/documents/selecting_file_formats.pdf〉.

Cedars Project. 2002. "Cedars Guide to Digi- tal Preservation Strategies". [cited 2010.08.20].
〈http://www.imaginar.org/dppd/DPPD/146%20pp%20Digital%20Preservation%20Strategies.pdf〉.

Columbia University Libraries. 2006. "Policy for Preservation of Digital Resources". [cited 2010.05.24].
〈http://www.columbia.edu/cu/lweb/services/preservation/dlpolicy.html〉.

Cornell University Library. 2004. "Cornell University Library Digital Preservation Policy Framework". [cited 2010.06.03].
〈http://commondepository.library.cornell.edu/cul-dp-framework.pdf〉.

Cornell University. 2007. "Digital Preservation Management: Implementing Short-term Strategies for Long-term Problems". [cited 2010. 06.03].
〈http://www.icpsr.umich.edu/dpm/〉.

DPC. 2008. "Preservation Management of Digital Materials: The Handbook". [cited 2010.05.22]. 〈http://dpconline.org/advice/organisational-activities-storage-and-preservation.html〉.

Factor, Michael et al. 2009. "Authenticity and Provenance in Long Term Digital Preservation: Modeling and Implementation in Preservation Aware Storage". [cited 2010.05.14]. 〈http://www.research.ibm.com/haifa/projects/storage/datasrores/papers/Auth_Prov_CamReady_sent.pdf〉.

Higgins, Sarah. 2009. "Applying the DCC Curation Lifecycle Model". [cited 2010.05.21]. 〈http://www.ieee-tcdl.org/Bulletin/v5n1/Higgins/higgins.htm〉.

ICPSR. 2010. "Digital Preservation Management". [cited 2010.07.01]. 〈http://www.icpsr.umich.edu/dpm/dpm-eng/oldmedia/obsolescence1.html〉.

InterPARES 2 Project. 2008. "InterPARES 2: Experiential, Interactive and Dynamic Records". [cited 2010.09.28]. 〈http://www.interpares.org/ip2/book.cfm〉.

James, Hamish. 2003. "Feasibility and Requirements Study on Preservation of E-Prints". [cited 2010.05.21]. 〈http://www.sherpa.ac.uk/documents/feasibility_eprint_preservation.pdf〉.

National Library of Australia. 2008. "Recommended Practice for Digital Preservation". [cited 2010.06.12]. 〈http://www.nal.gov.au/preserve/digipres/digipraspractices.html〉.

OCLC. 2006. "OCLC Digital Archive Preservation Policy and Supporting Documentation, Last Revised: 8 August 2006". [cited 2010.06.01]. 〈http://www.oclc.org/asiapacific/zhtw/support/documentation/digitalarchive/preservationpolicy.pdf〉.

Parliament Archives. 2009. "A Digital Preservation Policy for Parliament". [cited 2010.06.03]. 〈http://www.parliament.uk/documents/u.uoad/digitalpreservationpolicy1.0.pdf〉.

Todd, Malcolm. 2009. "File Formats for Preservation". [cited 2010.05.26]. 〈http://www.dpconline.org/newsroom/file-formats-for-preservationtechnology-watch-report.html〉.

University of Minnesota. 2007. "University Digital Conservancy Preservation

Policy". [cited 2010.05.24]. 〈http://conservancy.umn.edu/pol-preservation.jsp〉.

UK Data Archive. 2010. "UK Data Archive Preservation Policy". [cited 2010.08.11]. 〈http://www.data-archive.ac.uk/media/54776/ukda062-dps_preservationpolicy.pdf〉.

Yale University. 2005. "Yale University Library Digital Preservation Policy". [cited 2010.05.28]. 〈http://www.library.yale.edu/iac/DPC/final1.html〉.