

디지털 정보자원을 위한 보존 메타데이터 요소 설정 연구

A Study on Preservation Metadata Elements for Digital Information Resources

홍재현(Jae-Hyun Hong)*

< 목 차 >

- | | |
|----------------------------|---|
| 1. 서론 | 3.2 OAIS 참조 모형 |
| 2. 디지털 정보자원의 특성과 보존의 개념 | 3.3 디지털 정보자원의 보존 메타데이터종류 |
| 2.1 디지털 정보자원의 특성 | 4. 디지털 정보자원의 보존 메타데이터 요소 비교 분석 |
| 2.2 디지털 보존의 개념 | 5. 우리나라 디지털 정보자원의 보존 메타데이터 기본 요소 세트 설정 방안 |
| 3. 디지털 정보자원의 보존 메타데이터 | 6. 결론 및 제언 |
| 3.1 디지털 정보자원의 보존 메타데이터 필요성 | |

초 록

디지털 정보자원은 장기적 보존 및 다음 세대로의 접근을 보장하기 위한 상세한 메타데이터를 필요로 한다. 본 연구는 먼저 디지털 정보자원의 특성과 디지털 보존의 개념, 디지털 정보자원의 장기 보존과 접근을 위한 메타데이터의 필요성을 조명하였다. 디지털 정보자원을 위한 상호운용적인 보존 메타데이터 개발에 근거가 되고 있는 OAIS 참조 모형을 살펴보았다. 이어 실제 OAIS 참조 모형을 적용한 CEDARS, NLA, NEDLIB, OCLC/RLG 디지털 정보자원의 보존 메타데이터를 대상으로 하여 그 개발 과정, 목적 및 메타데이터 요소 세트를 각각 구체적으로 비교 분석하였다. 본 분석 결과를 기초로 하여, 향후 우리나라의 디지털 정보자원의 장기 보존과 접근을 위한 상호운용성을 갖춘 보존 메타데이터의 기본적인 요소 세트 설정 방안을 제시하였다.

주제어 : 디지털 정보자원, 보존 메타데이터, OAIS 참조 모형, CEDARS, NLA, NEDLIB, OCLC/RLG

Abstract

Digital information resource requires detailed metadata to ensure its long-term preservation and accessibility for future generations. This paper illuminated the characteristics of digital information resources, the concept of digital preservation, and the necessity of metadata for long-term preservation of digital information resources. And this paper examined the 16 elements of RLG metadata for preserving digital information and the Reference Model for an Open Archival Information System(OAIS) to support the interoperability of digital archiving. And then this study investigated concretely element sets of the preservation metadata for digital information resources in the Cedars and OCLC/RLG embodied within the OAIS framework. As result of the study, the paper suggested basic plan required to provide preservation metadata elements to preserve digital information resources in Korea.

Key Words : digital information resource, preservation metadata, OAIS reference model, CEDARS, NLA, NEDLIB, OCLC/RLG

* 중부대학교 문헌정보학 전공 부교수(jhhong@joongbu.ac.kr)

· 접수일 : 2004. 8. 26 · 최초심사일 : 2004. 9. 1 · 최종심사일 : 2004. 9. 8

1. 서 론

카드 목록의 시대에서 기계 가독의 시대를 거쳐 인터넷 시대가 되자 이전에 인쇄 형태로 생산되었던 막대한 규모의 자료들이 디지털로 변환되고 있다. 뿐만 아니라 아예 처음부터 디지털 형태로 생산되고 있는 디지털 출판물이 증가하고 있다. 또한 웹을 통하여 유통되고 접근할 수 있는 전자저널, 색인·초록데이터베이스, 웹 페이지 등의 웹 정보자원과 인터넷 링크자료 등의 양이 급증하고 있다.

이러한 변화에 대처하고자 도서관계에서는 디지털 정보의 수집·조직화·제공 등을 위한 다양한 논의와 연구를 거듭해 왔다. 그렇지만 대부분 디지털 정보를 구축하여 어떻게 이용에 제공할 것인가가 목전의 과제로 우선함으로써, 디지털 정보자원의 보존과 관련된 문제는 도서관으로부터 그다지 주목을 받지 못했다.

그러다가 최근 디지털 정보자원의 장기 보존과 접근에 관련한 문제점들이 부각되고 이에 이목이 집중되기 시작했다. 그 가장 큰 이유는 디지털 정보자원 자체가 가변성과 단명성 등의 자생적인 취약성을 가지고 있기 때문이다. 그로 인하여 디지털 정보자원은 보존에 따른 포맷이나 매체의 변환, 매체 재생, 마이그레이션(migration), 에뮬레이션(emulation) 등에서 비롯되는 여러 가지 불안정성을 노출시키고 있다. 또한 디지털 정보는 인쇄 자료에 비해 불법적인 복제와 변조가 매우 용이하다. 아울러 유달리 급속하게 진전하는 테크놀로지나 정보 기술에 의존하는 관계에서 생기는 매체나 기술의 노화 속도는 종이나 마이크로형태 자료보다 훨씬 빠르다. 뿐만 아니라 디지털 정보는 CD-ROM 등의 소위 패키지계 디지털 정보도 있지만, 대부분이 파도와 같이 나타났다가 사라져 가는 네트워크계 디지털 정보라는 사실에서 그 보존에 우려가 증폭되고 있다.

모든 디지털 정보자원이 국가 정보 자원으로 보존할 가치가 있는 것은 아니지만, 국가의 문화유산과 지식 유산이 보존되지 못한 채 이미 사라졌거나 사라지고 있다면 그 문제의 심각성은 매우 크다. 디지털 환경에서도 우리 후손들이 인류의 지식 문화의 소산인 디지털 자료들을 장기적으로 접근하여 이용할 수 있도록 하는 사회적 보존 계승 장치로서의 도서관의 역할은 아날로그 환경에서의 역할과 결코 다를 바 없다.

현재 우리 도서관들이 당면한 디지털 정보자원의 장기 보존과 접근에 관한 문제는 인쇄 자료와 다른 특성에서 기인되는 새로운 요소 및 항목이 여러 가지에 걸친 복잡한 새로운 과제이다. 도서관들은 디지털 정보자원의 장기적 보존과 접근에 필요한 관련 정보를 완벽하게 기술하고 효율적으로 계속 관리할 수 있도록 하기 위한 새로운 도구의 개발이 필요하다. 나아가 상호운용성(interoperability)을 갖춘 디지털 보존 메타데이터의 중요성이 점점 증대되고 있다.

미국과 유럽에서는 1990년대 중반 이후 디지털 정보자원의 장기 보존과 접근을 위한 메타데이터에 관한 기초 연구가 시작되었고, 여러 연구 프로젝트/그룹에 의해 그 개발이 활발하게 수행되

어 왔다. 그 주요 연구 프로젝트/그룹으로는 PANDORA (Preserving and Accessing Networked Documentary Resources of Australia), CEDARS 프로젝트, OCLC/RLG 보존 메타데이터에 관한 워킹 그룹, 유럽 보존 도서관 네트워크(Networked European Deposit Library, 이하 NEDLIB라 한다), OCLC 디지털 아카이브 서비스(OCLC Digital Archive Service), 오스트레일리아 국립 도서관(National Library of Australia, 이하 NLA라 한다), EAD(Encoded Archival Description) 등을 들 수 있다. 그러나 디지털 정보자원의 장기 보존과 접근을 위한 메타데이터에 대한 국제적인 표준은 아직 확립되어 있지 않다.

국내의 경우 디지털 정보자원의 접근과 검색, 아카이빙 정책, 기록관리 메타데이터 및 기술요소 등에 관한 연구는 비교적 활발한 편이다. 반면에 디지털 정보자원을 대상으로 하여 그 장기 보존과 접근을 위해 필요한 메타데이터의 개발에 초점을 맞춘 연구는 아직 미비한 것이 사실이다.¹⁾ 현재 디지털 정보자원의 장기 보존과 접근을 위한 메타데이터와 관련해서는 Dublin Core (이하 DC라 한다)와 MARC을 사용하고 있거나 자체 포맷을 가지고 있으면서 서비스를 제공받는 기관에서 요청할 경우 MARC로 제작하여 제공²⁾하고 있는 수준에 머물러 있다.

본 연구는 이점에 착안하여 우리나라 디지털 정보자원의 장기 보존과 접근을 위한 상호운용성을 갖춘 보존 메타데이터의 요소 설정을 위한 기초 작업으로서, 먼저 디지털 정보자원의 특성과 디지털 보존의 개념, 디지털 정보자원의 보존 메타데이터의 필요성을 구체적으로 조명하고자 한다. 또한 본 연구는 디지털 정보자원의 상호운용적인 보존 메타데이터 개발 시 정보 패키지 구조의 모델이 되고 있는 OAIS 참조 모형(Reference Model for an Open Archival Information System)을 살펴보고자 한다. 이어서 OAIS 참조 모형을 기반으로 하여 개발된 CEDARS, NLA, NEDLIB, OCLC/RLG 메타데이터를 중심으로 하여 그 개발 과정과 목적, 구조를 검토한다. 또한 OAIS 참

1) 본 연구와 관련해서 국내에서 이루어진 연구는 다음과 같다 :

① 최원태는 기술적, 경제적, 국가적인 관점에서 디지털 아카이브의 현황과 디지털 정보의 장기 보존과 관련된 배경과 이슈, 디지털 아카이브의 구성 요소로서 OAIS 참조 모형과 전망에 대해 고찰하였다.

최원태, “디지털 아카이브의 현황 및 구성 요소에 관한 연구,” 한국문헌정보학회지, 제35권, 제2호(2001), pp.23-40.

② 김희정은 디지털 아카이빙의 최근 연구 프로젝트(InterPares, NEDLIB, CEDARS, CAMLEON, PRIMSM)를 중심으로 연구 동향과 OAIS 참조 모형의 주요 기능에 관 해 다루었지만, 디지털 정보자원의 장기 보존과 접근을 위한 메타데이터 요소를 연구 범위에 포함시키지 않았다.

김희정, “디지털 아카이빙 최근 연구 동향 및 OAIS 참조 모형에 관한 연구,” 한국기록관리학회지, 제3권, 제1호(2003, 3), pp.23-41.

③ 또한 김희정은 전자저널 아카이빙을 위한 OAIS 참조모형의 적용방안의 일부로서 OAIS 참조모형의 정보 패키지의 구성요소의 하나인 보존기술정보만을 중심으로 국내 디지털 정보자원의 장기 보존을 위한 메타데이터 기술요소를 간단하게 제시하였다. 그렇지만 본 연구는 OAIS 참조 모형을 적용하여 개발된 디지털 정보자원의 보존 메타데이터 요소를 대상으로 한 분석은 시도하지 않았다.

김희정, “전자저널 아카이빙을 위한 OAIS 참조모형의 적용방안에 관한 연구,” 한국 기록관리학회지, 제3권, 제2호(2003, 9), pp.124-125.

2) 김희정(2003, 9), *ibid*, p.123.

조 모형의 정보 패키지를 구성하는 콘텐츠 정보와 정보 기술 정보를 기준으로 하여 각 디지털 정보자원의 보존 메타데이터 요소 세트를 비교 분석하고자 한다.³⁾ 그 분석 결과를 기초로 하여 본 연구는 우리나라의 디지털 정보자원의 장기 보존과 접근을 위한 상호운용성을 갖춘 메타데이터의 구조와 그 기본 요소 세트 설정 방안을 제시하고자 한다.

본 연구의 결과는 우리나라 디지털 정보자원의 보존 메타데이터 요소 설정에 필요한 지식 강화에 도움이 될 것이다. 아울러 향후 우리나라 디지털 정보자원의 상호운용성을 갖춘 표준적인 보존 메타데이터 요소 설정을 위한 기초 자료로 활용될 수 있을 것이다.

2. 디지털 정보자원의 특성과 보존의 개념

2.1 디지털 정보자원의 특성

현재 정보자료로서의 디지털 정보자원은 그 양적 증가와 함께 다양화가 진전됨에 따라 도서관 자료의 중요한 구성부분이 되어 가고 있다. 그러나 동시에 디지털 정보자원의 장기적 접근과 보존에 관한 문제는 종래의 인쇄 자료에서는 경험하지 못한 새로운 도전이 되고 있다. 이에 디지털 정보자원의 장기 보존이라는 새로운 문제를 제기시키고 있는 디지털 정보자원의 특성을 정리하면 다음과 같다.

- 인쇄 자료는 종이 등에 기록되어 있고 매체에서 직접 읽을 수 있는 정보 자료인 반면, 디지털 정보자원은 그 외형(매체)에서 직접 내용을 읽을 수 없다.
- 디지털 정보자원은 디지털 형태로 저장되고 접근되며, 동일 자원이 다양한 형식으로 표현될 수 있다.
- 디지털 정보자원은 급속하게 진전하는 테크놀로지나 재생기에 의존한다. 그러한 관계로 보존대상은 매체자체 외에 재생장치, 서버시스템, 데이터베이스, OS, 소프트웨어 등의 동작 환경 및 이용시스템, LAN, 통신회선 등의 인프라, 이들의 기술과 유지·관리를 위한 시스템 전체를 모두 포함해야 한다.
- 디지털 정보자원의 경우 개개의 매체나 기술의 노화 속도가 종이나 마이크로 형태 자료보다 훨씬 빠르다. 현재 1960년대, 1970년대의 중요한 디지털 정보가 재생기술이 사라졌기

3) PANDORA는 수집된 웹 자원을 MARC에 의해 기술하고 있고, EAD는 아카이브 기술을 위한 표준이자 디지털 정보자원의 장기 보존과 접근을 위한 상호운용적인 디지털 보존 메타데이터로 개발된 것이 아니다. 그래서 본 연구는 이들을 분석 대상으로 삼지 않았다.

때문에 이용할 수 없게 되었다고 하는 사례가 각국에서 보고되고 있다.⁴⁾ 따라서 디지털 정보자원은 재생을 위한 기술과 장치가 동시에 보존되어 있지 아니할 경우 정보의 접근이 끊겨져 버릴 위험성이 매우 크다.

- 디지털 정보자원은 재생기기 및 네트워크 고장, 조작 실수 등의 사고에 의해서 기록의 소실이 발생할 수 있다.
- 디지털 정보자원의 물리적 수명은 아직 검증되어 있지 않아 불확실하지만, 종이 자료는 그 평균 수명이 약 100년 단위인데 비하여, 패키지계 디지털 정보자원은 대략 10년에서 30년 정도로 수명이 짧다. 웹 정보자원은 그 내용에 따라 차이가 있을 수 있으나 반감기가 약 2년 정도로서 그 수명이 훨씬 더 짧다.
- 특히 네트워크계 디지털 정보자원의 개개 정보자원은 인터넷상의 웹사이트에 분산되어 존재하고 있고, 대부분 파도와 같이 나타났다가 사라져 가는 단명성을 나타내는 등 그 보존에 있어서 매우 불안정한 상태에 놓여 있다.
- 디지털 정보자원은 장기보존을 위한 매체와 포맷의 변환 또는 조작이 쉽다.
- 디지털 정보자원은 매체의 복제와 내용의 변경이 쉽다. 때문에 진본성(authenticity)과 저작권 침해의 문제를 일으킬 소지가 다분하다.

2.2 디지털 보존의 개념

디지털 보존이라는 말은 “원래 디지털 형태로 생산(born digital)되어 인쇄 형태나 아날로그 형태로 존재하지 않은 전자 레코드 자료의 보존과, 또 접근 및 보존을 위하여 아날로그 자료를 디지털 대응물(digital surrogate)로 만들어진 자료의 보존을 모두 지칭한다. 아예 처음부터 디지털 형태로 만들어진 것이든 또는 디지털 형태로 변환된 것이든 간에 관계없이 모든 디지털 자료들은 기술의 노화와 물리적인 열화를 겪고 있다.⁵⁾ 또한 디지털 보존은 디지털 객체의 모양과 느낌(look and feel) 물론 콘텐츠를 보존하는 것으로 아카이브 관리의 한 부분이다.⁶⁾ 뿐만 아니라 디지털 보존이란 디지털 형태로 수록된 데이터의 유용성, 내구성, 및 지적 내용의 무결성을 보장할 수 있도록 보존하는 과정⁷⁾을 의미한다.

4) 大島 薫, “電子出版物の保存,” 情報の科學と技術, Vol.50, No.7(2000), p.383.

5) Margaret Hedstrom, *Digital Preservation : A Time Bomb for Digital Libraries*, <<http://www.uky.edu/~kiernan/DL/hedstrom.html>> [cited 2004. 6. 25].

6) Gail M. Hodge, *Best Practices for Digital Archiving*(January 2000), <<http://www.press.umich.edu/jep/05-04/hodge.html>> [cited 2004. 6. 25].

7) Peter S. Graham, *Intellectual Preservation : Electronic Preservation of the Third Kind*(March 1994), <<http://www.clir.org/pubs/reports/graham/intpres.html>> [cited 2004. 6. 25].

이를 좀더 구체적으로 말하면 디지털 보존은 디지털 객체/자료를 장기간에 걸쳐 저장하고, 관리하고 접근하는 것이다. 여기에는 기술 보존, 에뮬레이션(emulation) 기술 또는 디지털 정보 이송(migration)을 포함한 하나 이상의 디지털 보존 전략을 수반할 것이다.⁸⁾ 이처럼 현재 디지털 정보의 보존에 대해서는 다양하게 정의가 내려지고 있다. 본 연구에서는 디지털 보존이란 첫째, 바이트 스트림을 장기적으로 유지하며 둘째, 기술이 변화하고 시간이 경과하더라도 지속적으로 콘텐츠의 유용성을 보장하는데 필요한 제반 관리 활동이라는 의미로 사용하기로 한다.

디지털 보존의 주요 방법으로는 매체 재생, 매체 변환, 포맷 변환, 디지털 정보 이송, 에뮬레이션, 인캡슐레이션(encapsulation) 등이 채용되고 있다. 그러나 현재 디지털 보존 방법에 대한 확실한 해결책은 마련되어 있지 않다

문제는 불행하게도 방대한 양의 디지털 정보가 디지털 보존에 주의를 기울이지 못하고 있는 사이에 이미 영원히 사라졌다는 점이다. 관리되지 않은 디지털 역사는 개인이나 기관에 의해 재창조되기 쉽다. 전자 시대로 이전하고 있는 관문에서 새로운 미개인이 있다는 것과 현재 디지털 형태로 작성되고 코드화 된 것 중의 많은 것을 영원히 잃어버리게 될 것이라는 것⁹⁾을 아는 것은 중요하다. 따라서 중요한 디지털 정보의 보존은 지금 보장되어야 한다.

도서관들은 디지털 정보의 손실을 줄이고 장기간의 보존을 위해 서지 기술과 매체의 특징 및 시스템 운영상의 기술, 보존을 위한 작업, 변경이력 등의 상세한 기록을 지속적으로 유지 관리해야 할 중요한 새로운 책무를 떠맡고 있음을 상기해야 한다. 과거 중세 때 수사가 책을 보존하고 배포하는데 중요한 역할을 한 것과 마찬가지로 우리 시대의 역사와 발행된 지식 유산을 보존하고 관리하는 것은 사서들과 아키비스트들의 몫이다.

일반적으로 인쇄 자료의 보존에서는 매체와 기술 및 내용의 보존은 통상 일체 된 것으로서 생각되어 왔다. 그러나 디지털 정보자원은 그 자체의 종류 및 형태가 다양할 뿐만 아니라 인쇄 자료와 근본적으로 다른 특성으로 인하여 보존 전략 수립 시 매체의 보존, 기술의 보존, 지적 내용의 보존이라는 3가지 수준을 모두 고려할 필요가 있다.¹⁰⁾ 뿐만 아니라 이 3가지 수준은 디지털 정보자원의 보존 메타데이터 개발 시에도 기본적으로 반영되어야 한다. 이에 그 내용을 정리하면 다음과 같다.

① 매체 보존

매체 보존은 적정 환경에서 디지털 정보자원을 수록한 매체를 보존하는 것을 의미한다. 매체

8) Kelly Russell, *CEDARS : Long-term Access and Usability of Digital Resources*(last updated December 1998), <<http://www.ariadne.ac.uk/issue18/cedars>> [cited 2004. 6. 25].

9) Terry Kuny, *The Digital Dark Ages? Challenges in the Preservation of Electronic Information* (May 1998), <<http://www.ifla.org/VI/4/news/17-98.htm#2>> [cited 2004. 6. 25].

10) Peter S. Graham, *op. cit.*

보존의 목적은 매체가 계속해서 이용될 경우 매체 노화로 인하여 발생하는 데이터 손실을 방지하기 위함이다.

② 기술 보존

디지털 정보자원은 기록매체, 포맷, 소프트웨어, 하드웨어, 통신 네트워크 등의 기술에 의존해서 존재한다. 멀티미디어와 같은 유형의 정보는 소프트웨어 및 하드웨어 기술에 훨씬 더 밀접하게 연관되어 있어 이러한 기술을 소유하지 않은 환경에서는 이용할 수 없다. 정보기술 변화의 역동성은 장기간에 걸친 하드웨어와 소프트웨어의 지속성을 예측 불가능하게 하여 물리적 매체의 노화보다 더 큰 도전이 되고 있다. 그러므로 빠르게 세대 교체되거나 노후화 되는 이들 기술을 보존함으로써 내용의 접근이 보장되도록 해야 한다.

③ 지적 내용 보존

지적 내용 보존은 최초로 기록된 디지털 정보자원의 무결성(integrity)과 진본성을 유지한 상태로 보존하는 것을 의미한다. 이것은 오리지널 디지털 정보의 내용 자체를 그대로 보존하는 것은 물론 디지털 정보자원의 ‘모양과 느낌’(look and feel)의 영역까지를 포함한 보존을 말한다.

3. 디지털 정보자원의 보존 메타데이터

3.1 디지털 정보자원의 보존 메타데이터 필요성

도서관에서는 일반적으로 디지털 정보자원을 기술하기 위한 메타데이터로 크게 전통적인 MARC와 최근 메타데이터의 표준으로 인지되고 있는 DC 메타데이터를 적용해 왔다. 그렇다면 검색용 기술도구로 사용되고 있는 MARC와 DC 메타데이터가 디지털 정보자원의 보존용 메타데이터로 사용하기에 적절한지를 먼저 살펴보고자 한다.

MARC는 인쇄 기반 환경에서 전 세계적으로 널리 사용되고 있는 메타데이터이다. 그러나 1990년대 이후로 디지털 정보자원이 증가하고 인터넷이 확산됨에 따라 기존의 MARC의 기술요소 만으로는 한계가 있어 그 성능을 향상시키려는 노력이 시도되었다. 최근에 발표된 MARC 21은 856(electronic location and access) 태그를 설정하여 네트워크 자원의 기술을 위한 부분을 많이 수용하고 있다. 그럼에도 불구하고 MARC 21은 주로 소재위치와 접근에 관한 정보제공을 목적으로 네트워크 자원을 기술하고 있을 수준에 머물러 있을 뿐, 디지털 정보자원의 장기 보존에 필요한 제반 기술 부분을 충분히 수용하고 있지 못하다.

DC 메타데이터(15개의 요소 세트로 구성)는 웹 정보자원의 관리에 널리 적용되고 있다. 현재 DC가 디지털 정보자원의 단순한 기술과 상이한 메타데이터 간의 최소 호환성을 확보하는 차원에서 적절한 도구로 간주되고 있음은 사실이다. 그러나 DC 메타데이터도 디지털 정보자원의 검색에 유용한 수단으로서 개발된 것이지, 보존이라는 관점에서 개발된 메타데이터는 아니다. 따라서 MARC과 DC 메타데이터가 디지털 정보자원의 장기 보존에 필요한 요구사항을 충실히 수용한 메타데이터로 사용하기에는 적절하다고 보기 어렵다.

이에 반하여 디지털 정보자원의 보존 메타데이터는 디지털 보존과 관련된 전 과정을 지원하는 정보 인프라를 제공한다. 이를 좀더 구체적으로 말하자면 디지털 정보자원의 보존 메타데이터는 장기간에 걸쳐 디지털 정보자원을 이용할 수 있도록 하는 능력을 제공한다. 또한 디지털 정보자원을 기술하고 식별할 수 있도록 하며, 내용을 이해할 수 있도록 풍부한 모든 정보를 가장 효율적으로 담아내는 도구로서의 기능을 한다. 뿐만 아니라 디지털 보존 과정에서의 입력물로서의 역할과 아울러 그 보존 과정에서의 출력물로서의 역할도 충실히 수행한다.¹¹⁾

이에 디지털 정보자원의 장기 보존과 접근을 위한 메타데이터 개발의 필요성을 정리하면 다음과 같다.

① 디지털 정보자원의 보존 메타데이터는 디지털 정보자원의 특성과 관련하여 인쇄 정보와 크게 다른 디지털 정보자원의 장기 보존과 접근을 보장할 수 있도록 관련된 모든 정보를 상세하고 완벽하게 기술·식별할 수 있게 한다.

② 디지털 정보자원의 보존 메타데이터는 변화하는 환경에서도 디지털 정보를 계속 이용할 수 있게끔 그 소재에 지속적으로 접근할 수 있도록 다른 객체와의 관계나 디지털 객체의 이력(history) 등에 관한 정보를 기록하고 관리할 수 있도록 한다.

③ 디지털 정보자원의 보존 메타데이터는 매체나 기술의 노화가 발생하더라도, 이용자가 그 내용을 해석하고 이해할 수 있도록 데이터 객체 자체와 이를 표현하는 하드웨어/소프트웨어 환경 등에 대한 상세한 정보를 기술할 수 있도록 한다.

④ 디지털 정보자원의 보존 메타데이터는 디지털 정보의 진본성과 무결성을 유지할 수 있도록 디지털 정보자원의 발생, 관리 상황, 객체 인증 등을 기록하고 관리할 수 있도록 한다.

⑤ 디지털 정보자원의 보존 메타데이터는 디지털 정보자원과 관련된 저작권에 관한 정보를 상세하게 기술하여 저작권 처리 관계를 효율적으로 관리할 수 있도록 한다.

11) OCLC/RLG Working Group on Preservation Metadata, *Preservation Metadata and the OAIS Information Model : A Metadata Framework to Support the Preservation of Digital Objects*(June 2002), <http://www.oclc.org/research/projects/pmwg/pm_framework.pdf> p.1. [cited 2004. 6. 25].

3.2 OAIS 참조 모형

OAIS 참조 모형은 디지털 정보자원을 위한 상호운용적인 보존 메타데이터 정보 구조의 모델이 되고 있다.

OAIS 참조 모형은 『오픈 아카이브 정보시스템을 위한 참조 모형』(Reference Model for an Open Archival Information System)을 줄여 표현한 것이다. OAIS 참조 모형은 디지털 정보의 장기 보존을 위한 상호운용적인 아카이브 시스템의 구축을 위한 국제적인 개념적 틀(framework)을 의미한다.

OAIS 참조 모형은 NASA를 비롯한 세계 각국의 우주개발기관에서 조직한 우주데이터 시스템 자문위원회(Consultative Committee for Space Data System)에 의해 개발되었다. 이 참조 모형은 1999년 5월에 Red Book, Issue 1¹²⁾이라는 형태로 권고 초안이 발표되었다. 그 후 2001년 6월에 개정판인 Red Book, Issue 2¹³⁾가 발표되었고, 2002년 1월에 발표된 참조 모형 Blue Book, Issue 1¹⁴⁾의 권고(Recommendation)가 드디어 국제 표준 ISO 14721:2002로 채택되었다.

OAIS 환경에서 정보는 하나의 정보 패키지(Information Package)의 형태로 존재한다. OAIS 참조 모형에서 유통되는 하나의 아카이브 정보 패키지(AIP: Archive Information Package)는 4가지 유형의 정보 객체의 집합체로 구성되어 있다. 첫째, 보존의 원 대상이 되는 정보의 집합체인 ‘내용 정보’(Content Information)와 둘째, 내용 정보에 대한 일종의 메타데이터인 ‘보존 기술 정보’(Preservation Description Information, 이하 PDI라 한다) 셋째, 내용 정보와 보존 기술 정보를 묶거나 식별하기 위해 부여된 패키지화 정보(Packaging Information)와, 넷째 그 패키지를 검색할 수 있도록 패키지를 기술한 목록 정보인 기술 정보(Descriptive Information)로 이루어져 있다. OAIS 참조 모형에서의 아카이브 정보 패키지를 구성하는 4가지 유형의 정보 객체를 도식화하면 <그림 1>과 같다.

OAIS 참조 모형에서 디지털 정보자원의 보존 메타데이터와 직접적으로 관련된 중요한 두 가지 요소는 ‘Content Information’과 ‘PDI’이다. 여기에서 Packaging Information과 Descriptive Information은 그 대상에서 제외되었다. 그 이유는 Packaging Information이 데이터 객체의 보존과 직접적으로 연관되어 있지 않고 단지 보존 대상 데이터와 메타데이터를 묶는 것에 불과하다고 보았기 때문이다. 또한 Descriptive Information도 자원 검색을 위한 메타데이터로 보존용 메

12) CCSDS, "Reference Model for an Open Archival Information System(OAIS)," CCSDS 650.0-R-1. Red Book, Issue 1(May 1999),

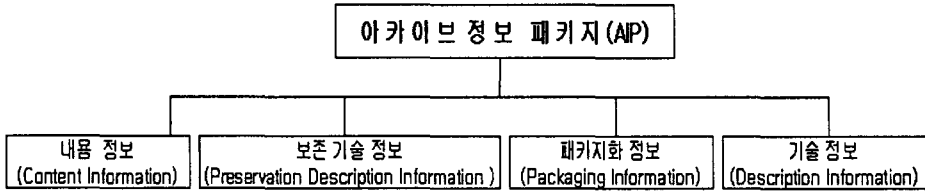
<<http://ssdoo.gsfc.nasa.gov/nost/wwwclassic/documents/p2/CCSDS-650.0-R-1.pdf>> [cited 2004. 6. 27].

13) CCSDS, "Reference Model for an Open Archival Information System(OAIS)," CCSDS 650.0-R-2. Red Book, Issue 2(June 2001),

<<http://ssdoo.gsfc.nasa.gov/nost/wwwclassic/documents/p2/CCSDS-650.0-R-2.pdf>> [cited 2004. 6. 27].

14) CCSDS(Jan. 2002), *op. cit.*, [cited 2004. 6. 27].

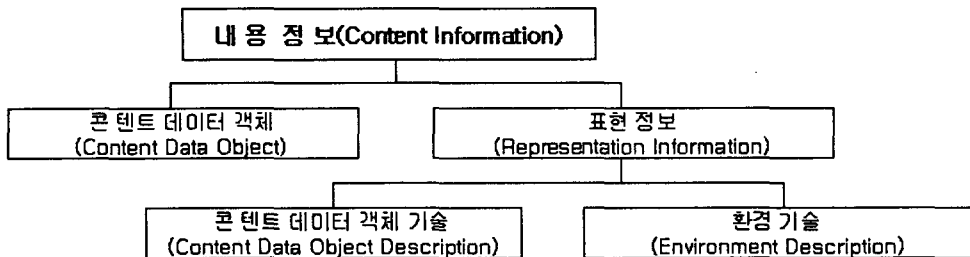
타데이터의 영역에 해당되지 않는다고 보았기 때문이다.



<그림 1> OAIS 참조 모형의 '아카이브 정보 패키지' 구조

디지털 정보자원의 보존 메타데이터와 관련된 두 가지 핵심 요소는 Content Information과 PDI이다. 이 요소들은 장기 보존과 접근이라는 두 가지의 중요한 기능을 각각 지원한다.

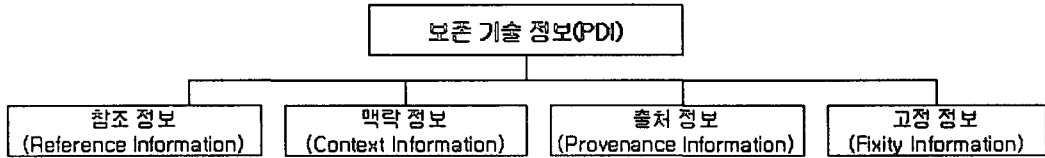
그 첫 번째 요소인 Content Information은 '콘텐츠 데이터 객체'(Content Data Object)와 이와 관련된 '표현 정보'(Representation Information) 요소로 구분된다. Content Data Object는 아카이브화 되는 원 데이터인 비트 스트림(bit stream)이나 비트 스트림 세트에 관한 사항을 기술한다. Representation Information은 Content Data Object의 비트 스트림을 접근 가능한 의미 있는 지식으로 만드는데 필요한 메타데이터로서, '콘텐츠 데이터 객체 기술'(Content Data Object Description)과 '환경 기술'(Environment Description) 요소로 이루어져 있다. Content Data Object Description 요소는 콘텐츠 데이터 객체의 내용을 제공하거나 이해하는데 필요한 콘텐츠 데이터 객체 자체의 특징에 관한 정보를 상세하게 표시하기 위한 부분이다. 이와 함께 Representation Information을 구성하는 또 하나의 요소인 Environment Description 요소는 현재 아카이브에 존재하는 형태로 Content Data Object를 디스플레이 하거나 접근하는데 필요한 하드웨어/소프트웨어 환경을 기술한다. <그림 2>는 이러한 Content Information의 구조를 나타낸 것이다.



<그림 2> OAIS 참조 모형의 'Content Information' 구조

또한 보존 메타데이터를 구성하는 두 번째 중요한 구성 요소는 'PDI'이다. PDI는 특수한 내용 정보를 적절하게 보존하는데 필요한 정보를 의미한다. 특히 PDI는 내용 정보를 고유하게 식별할

수 있게 하고 그 정보를 은밀하게 변경할 수 없도록 내용 정보의 과거 상태와 현재 상태를 상세하게 기술하는데 초점을 두고 있다. 이러한 PDI는 네 가지 유형의 카테고리로 대분된다. <그림 3>은 PDI의 구조를 나타낸 것이고, 다음의 <표 1>은 이 네 가지 유형의 내용을 정리한 것이다.



<그림 3> OAIS 참조 모형의 PDI 구조

<표 1> OAIS 참조 모형의 PDI를 구성하는 4가지 유형

보존 기술 정보의 유형	내 용
참조 정보(Reference Information)	내용 정보에 식별자(identifier)를 지정하기 위해 사용된 하나 이상의 메커니즘을 기술한다(예, ISBN, URN).
맥락 정보(Context Information)	내용 정보의 작성 이유 및 다른 내용 정보와의 관계를 기술한다.
출처 정보(Provenance Information)	내용 정보의 기원, 변화 내력, 보존 내력 등을 기술한다.
고정 정보(Fixity Information)	특수한 내용 정보가 입증되지 않은 방법으로 변경되지 않았음을 보장하기 위해 사용된 인증 메커니즘과 인증 키를 기술한다(예, 체크섬, 디지털 서명).

그러나 실제로 Content Information과 PDI 간의 차이는 개념적인 정의만큼 뚜렷하지 않다. 왜냐하면 PDI에 부여된 메타데이터는 디지털 객체의 내용을 나타내고 이해하는데 사용될 수 있다. 아울러 내용 정보에 부여된 메타데이터는 입력물로 사용될 수도 있고 아카이브의 보존 처리에 의해 출력물이 될 수도 있기 때문이다.

현재 OAIS 참조 모형은 디지털 아카이브 분야에서 표준으로 인정되고 있다. 디지털 정보의 보존 프로젝트들은 디지털 정보자원의 상호운용적인 보존 메타데이터의 요소를 개발하는데 실제로 OAIS 참조 모형의 구조를 활용하고 있다. OAIS 참조 모형의 정보 구조에 입각해서 디지털 정보자원의 보존 메타데이터를 개발할 경우, 도서관들은 보존 메타데이터의 상호운용성을 쉽게 확보할 수 있을 것이다. 아울러 이러한 노력은 디지털 정보자원의 보존 메타데이터의 표준에 관한 합의를 이끌어 내려는 기반 조성을 촉진시킬 것이다.

3.3 디지털 정보자원의 보존 메타데이터 종류

3.3.1 CEDARS 보존 메타데이터

CEDARS 프로젝트는 디지털 정보자원의 보존을 지원하기 위한 메타데이터 방식 초안¹⁵⁾을 2000년 4월에 발표하였다. CEDARS(CURL Exemplars in Digital Archives) 프로젝트는 디지털 정보자원을 위한 보존 메타데이터를 구체적으로 규정한 대표적인 선도 프로젝트의 한 예이다. 이 프로젝트는 영국의 CURL (Consortium of University Research Libraries)¹⁶⁾에 의해 추진되었다.

CEDARS 프로젝트의 보존 메타데이터 초안의 작성 목적은 실험적인 디지털 아카이브 구축을 지향하기 위해 보존 메타데이터의 기본 요소 세트를 제안하고, 국제적인 협력에 의해 디지털 정보자원을 위한 보존 메타데이터의 표준 개발을 도모하자는 것이었다. 그리하여 CEDARS가 제안한 디지털 정보자원의 보존을 위한 메타데이터는 보존을 포함한 완벽한 아카이브 기능을 제공하기 위한 관리적 정보, 기술적 정보 및 법적 정보를 비교적 충실하게 수록하고 있다. CEDARS 보존 메타데이터 방식은 OAIS의 참조 모형의 정보 구조를 따르고 있다.

3.3.2 NLA 보존 메타데이터

NLA의 디지털 보존 메타데이터는 호주의 PANDORA 프로젝트를 거쳐 자체적으로 개발되었다. PANDORA 프로젝트가 호주의 온라인 출판물의 장기 보존과 활용을 목적으로 NLA의 주도하에 1996년부터 수행되었지만, PANDORA 아카이브 관리 시스템(PAMS)은 다양한 디지털 정보자원의 보존을 지원하는 메타데이터를 담아내지 못하였다.¹⁷⁾ 그러자 NLA는 디지털 정보자원의 장기 보존을 위하여 별도의 디지털 보존 메타데이터의 개발에 관심을 갖게 되었다. NLA는 디지털 정보자원의 보존을 위한 메타데이터 개발에 독자적으로 착수하게 되었다. 그 결과로서 1999년 10월 15일에 디지털 정보자원을 위한 보존 메타데이터 초안을 발표하였다.¹⁸⁾

NLA의 디지털 보존 메타데이터는 디지털 형태로 생산된 객체와 디지털 대용물로 만들어진 디지털 객체의 장기 보존을 모두 지원하는 메타데이터 개발에 목적을 두었다. NLA의 디지털 보존 메타데이터의 구조는 CEDARS 보존 메타데이터와 마찬가지로 OAIS 참조 모형을 따르고 있다.

15) The Cedars Project Team and UKOLN, *Metadata for Digital Preservation : The CEDARS Project Outline Specification* (Draft for Public Consultation)(March 2000), <<http://www.leeds.ac.uk/cedars/MD-STR5.pdf>> [cited 2004. 7. 1.].

16) CEDARS 프로젝트를 추진한 CURL은 영국의 Leeds, Cambridge, Oxford 대학교로 구성되었다.

17) National Library of Australia, *Preservation Metadata for Digital Collections*(2000), <<http://www.nla.gov.au/nla/staffpaper/2000/webb1.html>> [cited 2004. 7. 2.].

18) OCLC/RLG Working Group on Preservation Metadata, "Preservation Metadata for Digital Objects : A Review of the State of the Art," *White Paper* (January 31, 2001), <http://www.oclc.org/research/pmwg/presmeta_wp.pdf>, [cited 2004. 7. 2.].

3.3.3 NEDLIB 보존 메타데이터

NEDLIB 보존 메타데이터는 유럽의 NEDLIB 프로젝트에 의해 개발된 메타데이터이다. 그 초안은 2000년 7월 21일에 발표되었다.¹⁹⁾

NEDLIB 프로젝트는 전자출판물의 장기 보존을 위한 프로젝트로서 1998년 1월 1일에 시작되어 2001년 1월 31일에 종료되었다. 이 프로젝트는 네덜란드 왕립도서관의 주관 하에 네덜란드 국립아카이브와 핀란드 국립도서관, 프랑스 국립도서관, 이탈리아 국립도서관, 노르웨이 국립도서관, 포르투갈 국립도서관, 스위스 국립도서관의 협력 하에 수행되었다. 그 외에 주요 출판사로서는 Elsevier, Kluwers, Springer Verlag가 참가하였다.²⁰⁾

전술한 바와 같이, NEDLIB 보존 메타데이터의 구조도 OAIS 참조 모형에 기초하여 개발되었다. 그러나 CEDARS, NLA, OCLC/RLG의 보존 메타데이터가 디지털 정보의 장기 보존과 접근이라는 두 가지 목적을 염두에 두고 개발된 것과 달리, NEDLIB 디지털 정보자원의 보존 메타데이터는 엄격하게 보존에만 초점을 맞추어 개발되었다. 그런 관계로 NEDLIB가 제안한 보존 메타데이터는 변화하는 기술 환경에서 방대한 양의 디지털 정보 관리를 위한 최소한의 핵심적인 메타데이터 요소 세트라 할 수 있다.

3.3.4 OCLC/RLG 보존 메타데이터

OCLC와 RLG는 2000년 3월 상호 협력하여 디지털 객체의 장기적 보존의 업무를 규명하고 지원하기 위하여 「OCLC/RLG 보존 메타데이터에 관한 워킹 그룹 (OCLC/RLG Working Group on Preservation Metadata)」을 발족시켰다. 이 워킹 그룹은 그 다음 해인 2001년에 디지털 객체의 보존 메타데이터의 최근 개발 동향을 리뷰한 백서²¹⁾를 발표하였다. 또한 이 워킹 그룹은 2002년 6월 드디어 OAIS 참조 모형에 근거한 디지털 객체의 보존을 지원하는 메타데이터 제안 보고서²²⁾를 발표하였다.

이 보고서가 제안한 보존 메타데이터 요소는 CEDARS, NLA, NEDLIB, OCLC에 의해 개발된 4개의 디지털 정보를 위한 보존 메타데이터를 종합적으로 검토한 것과, OCLC/RLG의 워킹 그룹(이하 WG라 한다)이 권고한 독자적인 요소가 추가되어 완성된 것이다. 현재 OCLC/RLG의 보존 메타데이터는 테스트를 통해 향후 개선·보완될 점들도 있지만, OAIS 참조 모형을 따른 디지털

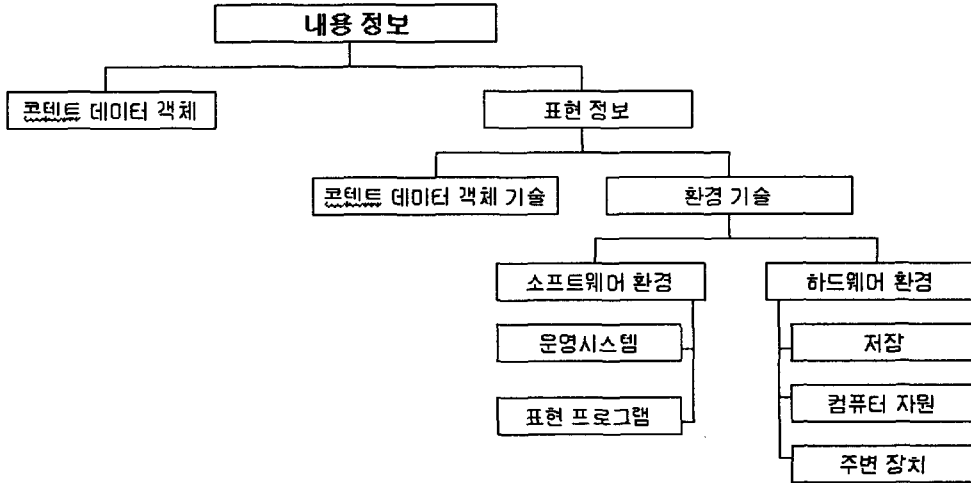
19) *Ibid.*, [cited 2004. 7. 2.].

20) Catherine Lupovici, *Technical Data and Preservation Needs*, 67th IFLA Council and General Conference August 16-25, 2001, <<http://www.ifla.org/IV/ifla67/papers/163-168e.pdf>> [cited 2004. 7. 2.].

21) OCLC/RLG Working Group on Preservation Metadata, *op. cit.*, [cited 2004. 7. 3.].

22) OCLC/RLG Working Group on Preservation Metadata, *Preservation Metadata and the OAIS Information Model : A Metadata Framework to Support the Preservation of Digital Objects*(June 2002), <http://www.oclc.org/research/projects/pmwg/prm_framework.pdf> [cited 2004. 7. 2.].

정보자원의 보존 메타데이터로 개발된 것 중에서 가장 종합적이고 진보적인 메타데이터라 할 수 있다. 또한 OCLC/RLG의 보존 메타데이터는 도서관계의 요구 사항에 부응하는 디지털 정보의 보존 메타데이터 표준에 관한 합의를 이루어내기 위한 미래의 기반을 마련했다는 점에서도 그 의의가 크다.

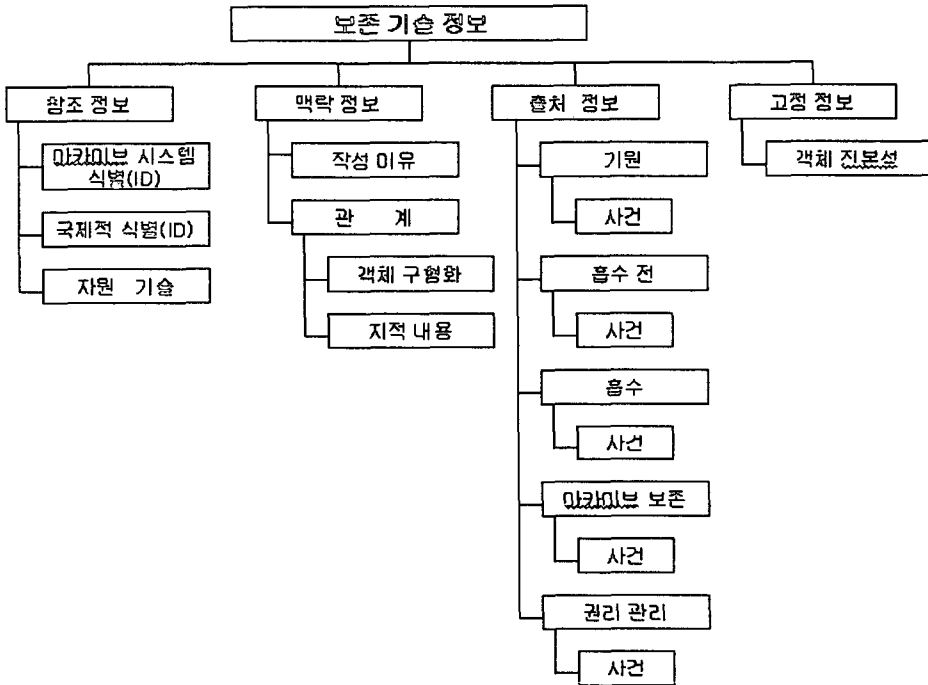


<그림 4> OCLC/RLG 메타데이터에서의 'Content Information' 구조

OCLC/RLG의 디지털 정보자원을 위한 보존 메타데이터의 구조는 OAIS 참조 모형을 따라 Content Information과 PDI로 이루어져 있다. OCLC/RLG 워킹 그룹이 OAIS 모형에 근거하여 규정한 디지털 객체의 보존을 지원하는 'Content Information' 요소의 전체 구조²³⁾는 <그림 4>와 같다. 또한 <그림 5>는 OAIS 모형에 근거하여 규정된 OCLC/RLG 디지털 보존 메타데이터 중 PDI의 전체 구조²⁴⁾이다.

23) *Ibid*, Figure 6

24) OCLC/ RLG Working Group on Preservation Metadata. Figure 12



<그림 5> OCLC/PLG 메타데이터에서의 'PDI' 구조

4. 디지털 정보자원의 보존 메타데이터 요소 비교 분석

본 장에서는 CEDARS, NLA, NEDLIB, OCLC/RLG의 디지털 정보자원의 보존 메타데이터의 요소를 비교 분석하고자 한다. 디지털 정보자원의 각 보존 메타데이터 요소는 상호운용의 틀을 제공하는 OAIS 참조 모형의 구조를 기준으로 하여 아카이브 정보 패키지를 구성하는 2가지 핵심 요소인 Content Information과 Preservation Description Information을 대상으로 분석하였다.

4.1 내용 정보(Content Information)

디지털 정보자원의 경우 현재의 접근성을 보장함은 물론 장기적인 이용이나 보존을 보장하기 위해서는 보존용의 주 대상의 상세한 기술이 필요하다.

4.1.1 콘텐츠 데이터 객체(Content Data Object)

콘텐츠 데이터 객체를 기술하기 위한 요소로 CEDARS 메타데이터는 Data Object 요소를, OCLC/RLG 메타데이터는 Content Data Object를 제안하고 있다.

4.1.2 표현 정보(Representation Information)

1) 콘텐츠 데이터 객체 기술(Content Data Object Description)

- 기본적인 초록형 기술. 이 요소는 Content Data Object에 대해 기본적인 정보를 우리가 읽을 수 있는 초록형식으로 기술하기 위해 필요하다. 이를 기술하기 위해 CEDARS와 OCLC/RLG 메타데이터는 Underlying abstract form description 요소를 두고 있다.

- 구조 유형. Content Data Object로 표현된 디지털 객체의 종류 즉 스틸 이미지, 사운드, 텍스트, 데이터베이스, 웹 문서 등을 기술하기 위한 테크니컬 메타데이터 요소로서, NLA와 OCLC/RLG 메타데이터에 설정된 Structural Type 요소가 이에 해당된다.

- 복합 매체의 테크니컬 하부 구조. 이 요소는 복합 디지털 객체들의 내부 구조를 기술 즉, 복합 객체를 구성하고 있는 객체들을 열거하고 그 관계를 기술(예, 웹 페이지는 1개의 HTML 파일과 함께 내장된 3개의 GIF 파일 및 1개의 WAV 파일로 구성됨)하기 위해 필요하다. NLA와 OCLC/RLG 메타데이터가 제안한 Technical Infrastructure of Complex Object 요소가 이에 해당된다.

- 파일 기술. Content Data Object를 구성하는 파일에 대한 테크니컬 명세서를 기술하기 위한 요소이다. ZIP 또는 TAR 파일과 같은 축적의 편리함을 위해 사용된 파일 포맷을 기술하기 위함보다는 콘텐츠를 직접 접근하고 표현하는데 사용된 파일 포맷(예, GIF 이미지 파일의 경우 픽셀 크기, 해상도, 컬러 팔레트, 압축 알고리즘 등)을 기술하기 위해 설정된다. NLA와 OCLC/RLG 메타데이터는 이를 기술하기 위해 File Description 요소를 두고 있다. 더욱이 NLA 메타데이터는 파일 기술 요소를 Image, Audio, Video, Text, Database, Executables라는 하위요소로 구분한 다음 각각 또 하위 요소들로 세분하고 있다. 이는 다양한 유형의 디지털 정보자원의 특성을 반영하여 기술하기 위함이라고 판단된다. NEDLIB 메타데이터는 객체 포맷을 기술하기 위해 Object Format(Name과 Version이라는 하위 요소로 구성)요소를 설정하고 있다.

- 설치 조건. 이 요소는 객체를 설치하는데 필요한 특수한 절차를 기술하기 위해 필요하다. Installation Requirements 요소명으로 NLA와 OCLC/RLG 메타데이터에 설정되어 있다.

- 크기. 이 요소는 객체의 크기(예, 1.3 MB)를 기술하기 위한 요소로서, 아카이브 시스템 내에서 객체를 관리하는데 필요한 테크니컬 메타데이터이다. Size 요소는 WG의 제안을 받아들여 OCLC/RLG 메타데이터가 새로이 설정한 요소이다.

- 접근 통제 수단. 이 요소는 Content Data Object로의 접근이나 복제, 마이그레이션을 통제

하기 위해 사용된 수단 즉 암호화, 디지털 서명, 비밀번호 보호 등을 기술하기 위해 필요하다. NLA와 OCLC/RLG 메타데이터는 이를 기술하기 위해 Access Inhibitors 요소를 설정하고 있다.

- 접근 촉진 수단. 이는 다음 세대로의 보존을 보장하기 위해 Content Data Object에 수록된 정보에 대한 접근을 향상시키는데 사용된 시스템이나 수단(예, 오디오 또는 비디오 파일의 타임 마커(time markers), 하이퍼텍스트 문서의 링크)을 기술하는 요소이다. NLA 메타데이터는 Finding and Searching Aids, and Access Facilitators를, OCLC/RLG 메타데이터는 Access Facilitator 요소를 설정하여 이를 기술하고 있다.

- 중요 속성. 지속적인 보존 과정 동안에 보존되어야 할 Content Data Object의 속성(예, PDF 문서의 경우 색상표, 내장된 이미지, 페이지 레이아웃, 내부 하이퍼링크 등)을 기술하기 위한 요소이다. OCLC/RLG 메타데이터는 WG의 제안을 받아들여 새로이 설정한 Significant Properties 가 이에 해당된다.

- 기능 속성. 보존 과정의 현 단계와 관련하여 Content Data Object의 기능적 속성 또는 “모양과 느낌” 속성을 기술하기 위해 필요하다. OCLC/RLG 메타데이터가 WG의 제안을 받아들여 새로이 설정한 요소인 Functionality 요소가 이에 해당된다.

- 기능 변화. 아카이브가 실시한 보존 과정과 절차에 의해 발생된 Content Data Object의 기능 속성의 상실 또는 모양과 느낌의 변화를 기술하기 위해 필요한 요소이다. 이를 나타내기 위해 NLA와 OCLC/RLG 메타데이터는 Quirks 요소를 설정하고 있다.

- 제공된 콘텐츠 기술. 이 요소는 Content Data Object의 내용에 대해 기술하는 요소로서, 아카이브 이용자에게 객체의 내용에 대한 적절한 이해와 해석을 보장한다. OCLC/RLG 메타데이터에 새로이 설정된 Description of rendered content 요소가 이에 해당된다.

- 기타. OCLC/RLG 메타데이터는 WG의 제안을 받아들여 Documentation이라는 요소를 추가 설정하여 Content Data Object를 디스플레이 하거나 설명하는데 필요한 도큐멘테이션을 기술할 수 있도록 한다.

한편 NEDLIB 메타데이터는 Content Data Object에 대한 기술을 가능하게 하는 어떤 요소도 제안하고 있지 않다.

2) 환경 기술(Environment Description)

(1) 소프트웨어 환경(Software Environment)

① Rendering Program

이 요소는 디지털 객체 내용을 나타내거나, 디스플레이 하거나, 접근할 수 있도록 디지털 객체 상에서 직접 작동되는 프로그램을 기술하기 위해 필요하다.

- 변환 과정 프로그램. 이 요소는 Content Data Object의 비트 스트림을 기본적인 초록형 기술로 자동 변환시키기 위한 소프트웨어 메커니즘을 기술하기 위한 요소이다. 이를 위해 CEDARS 메타데이터는 Transformer Object를, OCLC/RLG 메타데이터는 Transformation process를 설정하고 있다. OCLC/RLG 메타데이터는 이를 더욱 상세하게 기술할 수 있도록 하기 위해 Transformer engine(CEDARS의 경우 Render/Analyze Engines)하위 요소를 두고 있고, 이 하위 요소 또한 Parameters, Input format, Output format, Location, Documentation으로 세분하고 있다.

- 디스플레이/접근 응용프로그램. 이 요소는 Content Data Object를 디스플레이 하거나 그 내용에 접근하는데 필요한 프로그램(예, Internet Explorer 6.0, Adobe Acrobat Reader 4.0)을 기술하기 위한 요소이다. OCLC/RLG 메타데이터는 WG의 제안을 받아 들여 Display/Access Application 요소를 두고 있으며, 이를 다시 세분하여 기술할 수 있도록 Input format, Output format, Location, Documentation이라는 하위 요소를 두고 있다. CEDARS 메타데이터는 디지털 객체의 지적 내용에 접근하는데 필요한 소프트웨어 메커니즘을 기술하기 위해 Render/Analyze/Convert Objects 요소를, 또한 특수한 디지털 객체의 내용에 접근하는데 필요한 플랫폼과 독특한 소프트웨어 메커니즘을 기술할 수 있도록 하기 위해 Render/Analyze Object를 설정하고 있다. Render/Analyze/Convert Objects와 Render/Analyze Object 요소들은 각기 하위 요소를 두어 상세한 기술을 가능하도록 도모하고 있다.

한편 NEDLIB 메타데이터는 인터프리터와 컴파일러를 기술하기 위한 Interpreter and compiler(name, version, instruction으로 세분)요소와 응용프로그램을 기술하기 위한 Application(name, version으로 세분) 요소로 구분하고 있다. 그러나 이들과 달리 NLA 메타데이터는 Rendering Program의 기술이 가능하도록 고려하고 있지 않다.

② Operating System

Rendering Program을 작동시키는데 필요한 소프트웨어를 기술하기 위한 필요한 요소이다. 이를 위해 NEDLIB 메타데이터는 Operating System 요소(OS name, OS version 하위 요소 포함)를 설정하고 있다. 이를 채택한 OCLC/RLG 메타데이터도 Operating System 요소(OS name, OS version, Location, Documentation 하위 요소 포함)를 설정하고 있다.

그러나 CEDARS 메타데이터는 Operating System에 대해 기술할 수 있는 요소를

전혀 제안하지 않았고, NLA 메타데이터는 그나마 Known System Requirements에서 부분적으로 이를 기술할 수 있도록 하고 있다.

(2) 하드웨어 환경

객체의 내용을 나타내거나 접근하는데 필요한 물리적인 장치를 기술하기 위해 필요한 요소이다.

NEDLIB 메타데이터는 비표준적인 플랫폼 배치(configuration)나 하드웨어 장치를 기술하기 위해 3개의 하위 요소로 구성된 Specific Hardware requirements를 제안하고 있다. 그 3개의 하위 요소는 특수한 마이크로프로세서를 기술하는 Specific micro-processor requirements와, 비표준적인 멀티미디어 하드웨어 장치를 기술하는 Specific multimedia requirements, 비표준적인 주변장치를 기술하는 Specific peripheral requirements로 이루어져 있다.

OCLC/RLG 메타데이터는 현재와 다음 세대에서도 내용 정보의 이용을 보장하기 위해 8개의 요소를 설정하여 하드웨어 장치를 더욱 세부적으로 기술하도록 하고 있다. 즉, 계산 장치를 기술하기 위한 요소인 Microprocessor requirements, Memory requirements, Documentation과, 저장 장치를 기술하기 위한 요소인 Storage information, Documentation, 주변 장치를 기술하기 위한 요소인 Peripheral requirements, Documentation이 그에 해당된다.

그러나 CEDARS와 NLA 메타데이터는 모두 하드웨어 환경을 기술할 수 있는 요소를 전혀 제안하지 않았다. <표 2>는 보존 메타데이터별로 Content Information에 해당되는 요소와 그 하위 요소를 나타낸 것이다.

4.2 보존 기술 정보(PDI)

PDI는 보존 메타데이터를 구성하는 두 번째 핵심 요소로서, 각 메타데이터는 다음과 같은 네 개의 하위 요소로 대분하여 조직하고 있다. 그 내용을 각 메타데이터별로 비교하여 살펴보면 다음과 같다.

4.2.1 참조 정보(Reference Information)

Reference Information은 자원의 식별자를 기술하고 부가적으로 자원을 기술한 정보를 식별할 수 있도록 하는 2개의 카테고리로 조직된 요소이다.

CEDARS 메타데이터는 객체의 고유한 식별자를 기술하는 Resource Description과, 그 객체를 작성한 저자와 버전 등에 대한 기존의 메타데이터 기록을 나타내는 Existing Metadata로 이루어져 있다. NLA 메타데이터에서 이와 대응되는 요소로는 객체를 고유하게 정의하는 Persistent Identifier-type and identifier와, 기존의 자원 기술을 위한 요소인 Date of Creation을 포함하고 있다.

<표 2> 디지털 정보자원의 보존 메타데이터별 'Content Information' 요소

OAIS	CEDARS	NLA	NEDLIB	OCLC/RLG
Representation Information	Structure Information	Structural Type	Specific Hardware requirements	Content Data Object Description
	- Underlying Abstracts Form Description	Technical Infrastructure of Complex Object	- Specific microprocessor requirements	Underlying abstract form description
	- Transformer Objects	File Description	- Specific multimedia requirements	Structural type
	- Platform	- Image	- Specific peripheral requirements	Technical infrastructure of complex object
	- Parameters	- Image Format and Version	Operating System	File description
	- Render/Analyze Engines	- Image Resolution	- Name	Installation requirements
	- Output Format	- Image Dimensions	- Version	Size
	- Input format	- Image Color	Interpreter and compiler	Access inhibitors
	- Render/Analyze/Convert Objects	- Image Tonal Resolution	- Name	Access facilitators
	- Platform	- Image Color Space	- Version	Significant properties
	- Parameters	- Image Color Lookup Table	- Instruction	Functionality
	- Render/Analyze Engines	- Image Orientation	Object Format	Description of rendered content
	- Output Format	- Compression	- Name	Quirks
	- Input Format	- Audio	- Version	Documentation
	Semantic Information	- Audio Format and Version	Application	Environment Description
	- Render/Analyze Objects	- Audio Resolution	- Name	Software Environment
	- Platform	- Duration	- Version	Rendering Programs
	- Parameters	- Audio Bit Rate		Transformation process
	- Render/Analyze Engines	- Compression		- Transformer engine
	- Output Format	- Encapsulation		- Parameters
	- Input Format	- Track Number and Type		- Input format
		- Video		- Output format
		- Video File Format and Version		- Location
		- Frame Dimensions		- Documentation
		- Duration		Display/Access Application
		- Frame Rate		- Input format
		- Compression		- Output format
		- Video Encoding Structure		- Location
		- Video Sound		- Documentation
		- Text		Operating System
		- Text Format and Version		OS name
		- Compression		OS version
		- Text Character Set		Location
		- Text Associated DTD		Documentation

<표 2> 디지털 정보자원의 보존 메타데이터별 콘텐츠 정보 요소

OAIS	CEDARS	NLA	NEDLIB	OCLC/RLG
Representation Information		- Text Structural Division		Hardware Environment
		- Database		Location
		- Database Format and Version		Computational Resources
		- Compression		Microprocessor requirements
		- Datatype and Representation Category		Memory requirements
		- Representation Form and Layout		Documentation
		- Maximum Size of Data Element Values		Storage
		- Executables		Storage information
		- Code Type and Version		Documentation
		Known System Requirements		Peripherals
		Installation Requirements		Peripheral requirements
		Storage Information		Documentation
	Content Data Object		Access Inhibitors	
		Finding/Searching Aids and Access Facilitators		
		Quirks		
Data Objects				Content Data Object

NEDLIB 메타데이터에서는 Assigned Identifier(Value, Construction method, Responsible agency 하위 요소로 세분)와 URL(Value, Date of validation 하위 요소로 세분)이 그 식별자에 해당되며, Creator, Title, Date of Creation, Publisher는 객체를 검색하는데 사용된 기존의 기술적 메타데이터를 나타내기 위한 요소로 설정되어 있다. OCLC/RLG 메타데이터는 해당 아카이브 시스템 내에서의 식별을 위한 요소인 Archival system identification과, 아카이브 외부 시스템에서의 식별을 위한 요소인 Global identification(예: ISBN, persistent URL), 자원 검색이나 검색을 지원하기 위해 아카이브가 생성한 정보를 기술하는 요소인 Resource description으로 구성되어 있다. Archival system identification과 Global identification은 모두 Value, Construction method, Responsible agency 하위 요소로 세분하여 그 내용을 더욱 상세하게 기술할 수 있도록 하고 있다. <표 3>은 보존 메타데이터별로 Reference Information에 해당되는 요소와 그 하위 요소를 나타낸 것이다.

<표 3> 디지털 정보자원의 메타데이터별 Reference Information 요소

OAIS	CEDARS	NLA	NEDLIB	OCLC/RLG
Reference Information	Existing Metadata	Persistent identifier	Assigned Identifier	Archival system identifier
	- existing records	Date of Creation	- Value	- Value
	Resource Description		- Construction method	- Construction method
			- Responsible agency	- Responsible agency
			URL	Global identifier
			- Value	- Value
			- Date of validation	- Construction method
			Creator	- Responsible agency
			Title	Resource description
			Date of Creation	- Existing metadata
		Publisher	- Existing records	

4.2.2 맥락 정보(Context Information)

Context Information은 Content Data Object가 왜 만들어졌으며 그것이 다른 형식으로 존재하는 동일 객체와 어떠한 관계를 가지고 있는지를 제시하기 위해 필요하다. Context Information은 ‘작성 이유’와 ‘관계’ 라는 2개의 카테고리로 조직하는 것이 바람직하다.

CEDARS 메타데이터는 작성 이유를 Context Information 카테고리가 아닌 출처 정보(Provenance Information) 내에서 이를 기술하도록 하고 있지만, OCLC/RLG 메타데이터는 Context Information 내에 Reason for creation 요소를 포함시키고 있다. 반면에 NLA와 NEDLIB 메타데이터는 모두 이를 기술할 수 있는 요소를 배제시키고 있는 것으로 나타났다.

특히 디지털 정보자원은 동일 자원이 다양한 형식으로 표현될 수 있으므로, 보존되는 내용 정보가 기존의 다른 형식으로 존재하는 Content Data Object와 어떠한 관련을 가지고 있는가를 기술하는 것은 매우 중요하다고 본다. 이를 위해 CEDARS 메타데이터는 Related Information Objects를, NLA 메타데이터는 Relationships를, OCLC/RLG 메타데이터는 Relationships를 이를 기술하기 위한 요소로 제안하고 있다. 더욱이 OCLC/RLG 메타데이터는 Relationships 내에 하위 요소를 두어 그 관계를 상세히 기술하도록 하고 있다. 즉 아카이브 객체와 다른 형식의 동일 관련 객체와의 관계를 기술하는 Manifestation 요소(Relationship type, Identification 하위 요소로 세분)를 통해 객체의 변화 내력을 유지하거나 현 객체의 다른 버전과의 관계에 대해 상세하게 기술할 수 있도록 하고 있다. 또 아카이브 객체와 형식이 다른 관련 객체간의 지적 내용의 관계를 기술할 수 있도록 Intellectual content(Relationship, Identification 하위 요소로 세분)를 설정하고 있다. 그러나 NEDLIB 메타데이터는 너무 간략하다 보니 일체 이러한 관계를 기술할 수 있는 요소를 고려하고 있지 않다. 따라서 실제로 디지털 환경에서 이를 구현되도록 하기 위해서는 이러한 미흡한 사항은 추후 보완될 필요가 있다. <표 4>는 보존 메타데이터별로 Context

Information에 해당되는 요소와 그 하위 요소를 나타낸 것이다.

<표 4> 디지털 정보자원의 메타데이터별 'Context Information' 요소

OAIS	CEDARS	NLA	NEDLIB	OCLC/RLG
Context Information	Related Information Objects	Relationships		Reason for creation
				Relationships
				- Manifestation
				- Relationship type
				- Identification
				- Intellectual content
			- Relationship type	
			- Identification	

4.2.3 출처 정보(Provenance Information)

Provenance Information은 내용 정보의 내력(history)을 제공한다. 즉 내용 정보의 기원, 그것이 보존용으로 작성된 때부터 내내 발생한 변화들의 내력, 그 보존에 책임을 지고 있는 개인이나 기관에 대한 정보 등을 제공한다.

CEDARS 메타데이터는 Provenance Information을 3개의 카테고리 즉, 흡수(ingest) 이전의 오리지널 객체 정보의 내력을 기술하기 위한 History of Origin 요소, 디지털 아카이브가 보존을 위해 수집한 디지털 객체에서 일어난 모든 변화의 내력을 기술하기 위한 Management History 요소, 디지털 객체와 관련된 지적재산권에 관한 정보를 제공하는 Rights Management 요소로 구분하고 있다. 그 중 첫 번째 카테고리에 해당되는 History of Origin 요소는 Reason for Creation, Custody History(보존에 책임을 지고 있는 개인이나 조직을 기술하기 위해 설정), Change History Before Archiving, Original Technical Environment, Reason for Preservation 이라는 하위 요소를 두고 있다. 그러나 Reason for Creation은 Provenance Information에서 기술하기보다는 OCLC/RLG 메타데이터처럼 Context Information 내에서 기술하여 그 작성 이유와 다른 형식으로 존재하는 동일 객체와의 관계에 대한 정보를 함께 제공할 수 있도록 하는 것이 더 타당하다고 생각된다. 그 두 번째 카테고리에 해당되는 Management History는 디지털 아카이브 내에서 디지털 객체의 저장을 준비하는 과정에서 발생한 변화를 기록하기 위한 Ingest Process History 하위 요소와, 흡수가 완료된 후 디지털 객체에서 발생한 변화를 기술하기 위한 Administration History 하위 요소로 구분하여 상세히 기술하도록 한 점은 바람직하다고 생각된다. 끝으로 Rights Management는 보존용의 디지털 객체를 제출하도록 이끈 권리 협상 정보를 기술하는 Negotiation History와, 디지털 객체와 관련된 지적재산권에 관한 정보를 기술하는 Rights Information라는 하위 요소로 구분하여 조직하고 있다. Right Information 하위 요소는 ① Copyright Statement(Name of Publication, Date of Publication, Place of Publication,

Rights Warning, Contacts or Rights Holder라는 하위 요소로 또 세분됨), ② Actor(허락 받은 이용자를 기술), ③ Actions(허락 받은 행위를 기술하기 위한 하위 요소로서, 법령에 의한 허락을 기술하기 위한 Permitted by Statute와 라이선스에 의한 허락을 기술하기 위한 Permitted by License라는 하위 요소로 또 세분됨)라는 세 개의 하위 요소로 또 구성되어 있다. 이처럼 CEDARS 메타데이터는 다른 메타데이터에 비해 디지털 객체의 권리 처리 관계를 구체적으로 명시할 수 있도록 지적재산권 관계 요소를 매우 충실하게 세분하여 조직하고 있음을 나타냈다.

NLA 메타데이터는 Provenance Information을 12개의 요소로 조직하고 있다. 저작물 자체의 보존 내력과 그 보존용의 구현물에 대한 내력 정보를 구분하여 제공할 수 있도록 이를 2개의 카테고리 나누어 조직하고 있다. 첫 번째 카테고리는 저작물 자체의 보존 내력을 기술하는 요소로서는, ① 저작물의 보존 결정 여부와 그 결정일(보유기간 또는 검토일도 포함)을 기술하는 Archiving Reason(work), ② 저작물을 보존하기로 결정한 이유를 기술하는 Decision Reason(work), ③ 저작물의 보존 결정에 책임을 지고 있는 기관명을 기술하는 Institution Responsible for Archiving Decision(work) 요소가 이에 해당된다. 두 번째 카테고리는 보존용 구현물의 내력 정보도 상세하게 기술하기 위한 요소로서는, ① 보존용의 특수한 구현형을 보존/보유하기로 한 결정과 그 결정일을 기술하는 Archiving Decision (manifestation), ② 그 구현형을 보존/보유하기로 결정한 이유를 기술하는 Decision Reason (manifestation), ③ 그 구현형을 보존/보유하기로 결정을 내리는데 책임을 지고 있는 기관명을 기술하는 Institution Responsible for Archiving Decision (manifestation) 요소가 그것이다. 또한 NLA 메타데이터는 보존용으로 만들어진 특수한 구현형의 사용에 대해 기술할 수 있는 Intention Type 요소와 보존에 책임을 지고 있는 기관에 관한 정보를 제공하는 Institution with preservation responsibility를 설정하고 있다. 특히 Process 요소는 11개의 하위 요소(Description of Process, Name of Agency Responsible for Process, Critical Hardware Used in the Process, Critical Software Used in the Process, Date and Time 등)를 두어 디지털 객체에 적용된 보존 과정을 상세하게 기술하도록 고려하고 있는 것으로 보인다. 한편 권리 관리와 관련해서는 보존용 복제물의 합법성에 관한 정보를 먼저 제공하기 위해 그 복제물을 만드는데 허락을 받았는지의 여부를 기술하는 Preservation Action Permission 요소를 설정하고 있다. 때문에 NLA 메타데이터는 CEDARS 메타데이터에 비해 지적재산권의 관리에 필요한 정보를 구체적으로 명시할 수 없다는 문제점을 노출시키고 있다.

NEDLIB 메타데이터는 Provenance Information도 가장 간략하게 요소를 설정하고 있음을 나타냈다. 디지털 객체에서 발생된 모든 변화에 대한 정보를 기술하기 위해 Change History라는 하나의 요소만을 설정하고 있다. 그리고 Change History는 디지털 객체에서 발생된 변화를 대한 정보를 기술하는 Main metadata concerned와, 디지털 객체의 다른 측면에서 발생된 변화에 대한 정보(이러테면 마이그레이션에 대한 체크섬 값의 변화)를 기술하는 Other metadata concerned라는 하위 요소로만 단순하게 조직되어 있는 것으로 나타났다.

OCLC/RLG 메타데이터는 Provenance Information을 Content Data Object의 생명 주기의 단계별로 구분하여 변화 이력을 기술할 수 있도록 즉 기원(Origin), 흡수 전(Pre-ingest), 흡수(Ingest), 아카이브 보존(Archival Retention)이라는 4개의 요소와, 객체의 합법적인 사용을 명시하기 위한 Rights management라는 요소로 구분하고 있다. 이들 요소들은 모두 Content Data Object에 영향을 미친 변경 이력을 기술하기 위해 사건(Event)이라는 하위 요소를 두고 있다. 또 Event는 객체에 어떤 변화가 언제 일어났으며 그에 대한 책임을 누가 지고 있는 지 등에 관한 정보를 제공하기 위해 Designation, Procedure, Date, Responsible agency, Outcome, Note라는 하위 요소를 또 두고 있다. 이와 같이 디지털 객체에 영향을 미친 보존 과정을 세분하여 기술하도록 하는 점은 NLA의 Process 요소와 유사하다. OCLC/RLG 메타데이터는 보존 단계에서의 독립된 기술에 그치지 않고 Content Data Object의 생명 주기의 과정을 고려하여 전 과정을 포괄하여 변화의 이력을 명시할 수 있도록 한 점은 통합 관리를 강조하는 것이라고 평가된다.

4.2.4 고정 정보(Fixity Information)

고정 정보는 특수한 콘텐츠 정보 객체가 입증되지 않은 방식으로 변경되지 않았음을 보증하는 메커니즘을 기술하기 위해 설정된다. 따라서 이 카테고리에 기술된 정보는 디지털 객체의 진본성을 입증하는데 유용하게 사용될 수 있다.

CEDARS 메타데이터는 Fixity Information 카테고리 내에 Authentication Indicator 요소를 설정하여 디지털 객체의 진본성을 보증하기 위해 사용한 체크섬 또는 디지털 서명과 같은 메커니즘을 기술하도록 하고 있다. NLA 메타데이터는 이에 대응하는 요소로 Validation 요소를, NEDLIB 메타데이터는 Checksum 요소(Value, Algorithm의 하위 요소로 세분)와 Digital Signature 요소를 설정하고 있다. 그리고

<표 5> 디지털 정보자원의 메타데이터별 'Provenance Information' 요소

OAIS	CEDARS	NLA	NEDLIB	OCLC/RLG
Provenance Information	History of Origin	Archiving Decision(work)	Change History	Origin
	- Reason for Creation	Decision Reason(work)	- Main metadata concerned	- Event
	- Custody History	Institution Responsible for Archiving Description(work)	- Date	- Designation
	- Change History Before Archiving	Archiving Decision (manifestation)	- Old Value	- Procedure
	- Original Technical Environment	Decision Reason (manifestation)	- New Value	- Date
	- Prerequisites	Institution Responsible for Archiving Decision (manifestation)	- Tool	- Responsible agency
	- Procedures	Intention Type	- Name	- Outcome
	- Documentation	Institution with preservation responsibility	- Version	- Note
	- Reason for Preservation	Process	- Reverse	- Next occurrence
	Management History	- Description of Process	- Other metadata concerned	Pre-ingest
	- Ingest Process History	- Name of Agency Responsible for Process	- Old Value	- Event
	- Administration History	- Critical Hardware Used in the Process	- New Value	- Designation
	- Action History	- Critical Software Used in the Process		- Procedure
	- Policy History	- How Process was Carried Out		- Date
	Rights Management	- Guidelines Specified to Implement Process		- Responsible agency
	- Negotiation History	- Date and time		- Outcome
	- Rights Information	- Result		- Note
	- Copyright Statement	- Process rationale		- Next occurrence
	- Name of Publication	- Changes		Ingest
	- Date of Publication	- Other		- Event
	- Place of Publication	Preservation Action Permission		- Designation
	- Rights Warning	Record Creator		- Procedure
	- Contacts or Rights Holders	Other		- Date
	- Actors			- Responsible agency
	- Actions			- Outcome
	- Permitted by Statute			- Note
	- Legislation Pointers			- Next occurrence
	- Permitted by License			Archival retention
	- License Text Pointer			- Event

<표 5> 디지털 정보자원의 메타데이터별 'Provenance Information' 요소

OAIS	CEDARS	NLA	NEDLIB	OCLC/RLG
Provenance Information				- Designation
				- Procedure
				- Date
				- Responsible agency
				- Outcome
				- Note
				- Next occurrence
				Rights management
				- Event
				- Designation
				- Procedure
				- Date
				- Responsible agency
			- Outcome	
			- Note	
			- Next occurrence	

OCLC/RLG 메타데이터는 Object Authentication 요소를 설정하고 그 안에서 Authentication type, Authentication procedure, Authentication date, Authentication result라는 하위 요소를 둬으로써 콘텐츠의 진본성 체크 또는 인증키를 제공하는데 필요한 최소한의 조건에 관한 정보를 다른 메타데이터보다 구체적으로 기술할 수 있도록 하고 있다. <표 6>은 보존 메타데이터별로 고정 정보에 해당되는 요소와 그 하위 요소를 나타낸 것이다.

<표 6> 디지털 정보자원의 메타데이터별 'Fixity Information' 요소

OAIS	CEDARS	NLA	NEDLIB	OCLC/RLG
Fixity Information	Authentication Indicator	Validation	Checksum	Object Authentication
			- Value	- Authentication type
			- Algorithm	- Authentication procedure
			Digital Signature	- Authentication date
				- Authentication result

끝으로 CEDARS, NLA, NEDLIB, OCLC/RLG 보존 메타데이터 요소의 반복 사용 여부와 최소한의 공통 필수 요소를 제시하고 있는 지를 정리하면 다음과 같다.

CEDARS 보존 메타데이터는 데이터 요소의 반복 사용 여부를 규정하고 있지만, 데이터 요소가 필수 또는 선택인지의 여부를 규정하고 있지 않다. 대신에 각 데이터 요소의 중요도를 매우 중요, 중요, 그다지 중요하지 않음이라는 3단계로 구분하고 있는 것으로 나타났다. NLA 보존 메타데이터는 각 데이터 요소의 반복 여부를 명시하고 있다. 데이터 요소의 사용 의무는 3단계로 구분하되 매우 중요, 중요, 재량이라는 용어를 사용하고 있음을 나타냈다. NEDLIB 보존 메타데

이터는 각 데이터 요소의 반복 여부를 제시하고 있지만, 최소한의 핵심적인 요소로만 구성되어 있음을 나타냈다. OCLC/RLG가 제안한 디지털 정보자원을 위한 보존 메타데이터는 상이한 보존 메타데이터 요소의 상호운용성을 위해 필요한 최소한의 공통된 필수/선택 데이터 요소를 제시하고 있지 않으며 데이터 요소의 중요도도 제시하고 있지 않다. 뿐만 아니라 메타데이터 요소의 반복 사용이 가능한지의 여부도 제시하고 있지 않다. 따라서 OCLC/RLG가 제안한 보존 메타데이터가 가장 진보적이지만 결정판은 아니므로, 이러한 점들은 후속 연구를 통해 개선·보완되어야 할 과제를 지적한다.

5. 우리나라 디지털 정보자원의 보존 메타데이터 기본 요소 세트 설정 방안

본 연구는 이상에서의 분석 결과를 기초로 하여, 향후 우리나라 디지털 정보자원의 장기 보존과 접근을 위한 메타데이터의 기본적인 요소 세트 설정 방안을 다음과 같이 제시한다.

1) 디지털 정보자원의 장기 보존과 접근을 위한 메타데이터에 대한 국제적인 표준은 아직 확립되어 있지 않다. 따라서 우리는 국제적인 표준 제정의 움직임을 예의 주시하면서 우리의 디지털 정보자원을 위한 보존 메타데이터의 개발을 추진하도록 한다.

2) 향후 개발될 우리나라의 디지털 정보자원을 위한 표준적인 보존 메타데이터는 포괄성이 있고, 고도로 구조화되어야 하며, 광범위한 적용 가능성이 있어야 한다. 아울러 디지털 정보자원을 위해 외국에서 개발된 보존 메타데이터와 상호운용성을 갖추어야 한다.

3) 우리나라의 디지털 정보자원을 위한 보존 메타데이터는 OAIS 참조 모형의 개념 틀에 입각하여 데이터 요소를 구조화함으로써 상호운용성을 도모한다. 디지털 정보자원의 장기적 보존과 접근 기능을 지원할 수 있도록 Content Information과 Preservation Description Information라는 핵심 요소로 대분한 뒤 필요한 요소와 하위 요소로 구성한다.

4) 우리나라의 디지털 정보자원을 위한 보존 메타데이터 요소는 OCLC/RLG가 제안한 디지털 보존 메타데이터의 요소들을 중점적으로 고려한다. 동시에 이상에서 분석한 각 메타데이터들의 장점을 취하고, 우리의 실정에 부합하면서 디지털 정보자원의 장기 보존과 접근이라는 특수한 목적에 적합한 요소 세트로 설정한다.

5) Content Information 요소는 Content Data Object와 Representation Information으로 카테고리화를 조직한다. Representation Information 요소는 Content Data Object Description 요소와 Environment Description으로 구분하고 그 안에서 요소들을 다음과 같이 세분한다.

(1) Content Data Object Description 요소는 ‘기본적인 초록형 기술’, ‘구조 유형’, ‘복합 매

체의 테크니컬 하부 구조, '파일 기술(텍스트·이미지·오디오·비디오 등의 정보 유형별로 세분)', '설치 조건', '크기', '접근 통제 수단', '접근 촉진 수단', '중요 속성', '기능 속성', '기능 변화', '제공된 콘텐츠 기술' 등으로 구성한다.

(2) Environment Description 요소는 소프트웨어 환경과 하드웨어 환경으로 대분한다. 소프트웨어 환경 요소로는 '변환 과정 프로그램', '디스플레이/접근 응용 프로그램', '운영 체제' 요소로 구성된 뒤 각각 구체적으로 기술할 수 있도록 세분한다. 하드웨어 환경 요소로는 '특수한 마이크로프로세서 조건', '특수한 멀티미디어 조건', '특수한 주변장치 조건' 요소들로 구성한다.

6) Preservation Description Information 요소는 Reference Information, Context Information, Provenance Information, Fixity Information라는 4가지의 핵심적인 카테고리로 조직한다. 이를 위해 다음과 같은 요소들을 설정한다.

(1) Reference Information 요소는 먼저 '아카이브 시스템 식별', '국제적인 식별' 요소와 함께 디지털 정보자원의 검색을 지원할 수 있도록 기존의 메타데이터를 기술하는 '자원 기술' 요소로 구성한다.

(2) Context Information 요소는 아카이브화 되는 Content Data Object의 '작성 이유'와 '관계' 요소로 구성한다. 더욱이 관계 요소는 아카이브 객체와 다른 형식의 동일 객체와의 관계 및 아카이브 객체와 형식이 다른 관련 객체와의 관계도 기술할 수 있도록 세분한다.

(3) Provenance Information 요소는 생명 주기의 단계별로 구분하여 변화 이력을 기술할 수 있도록 '기원', '흡수 전', '흡수', '아카이브 보존'이라는 요소로 대분 한다. 그리고 그 안에서 각각 객체에서 일어난 변화의 내용과 시기, 책임기관 등을 상술할 수 있도록 세분한다. 또한 기원, 흡수 전, 흡수, 아카이브 보존 요소와 함께 디지털 정보가 불법적으로 복제 또는 변조되는 것을 방지하고 저작권의 합법적인 권리 처리 관계와 권리 관리를 명확히 기술할 수 있도록 '저작권' 요소를 설정한다. 저작권 요소는 '권리 관리', '협상 내력', '저작권 표시', '허락 받은 이용자', '법령에 의한 허락', '라이선스 계약에 의한 허락' 요소 등으로 구성하며, 이러한 각 요소들은 그 안에서 구체적인 저작권 기술이 가능하도록 세분한다.

(4) Fixity Information 요소로는 '객체 인증' 요소를 설정한다. 객체 인증 요소는 인증 유형, 인증 절차, 인증 날짜, 인증 결과를 기술할 수 있도록 세분한다.

6. 결론 및 제언

디지털 정보자원은 그 자체의 종류 및 형태가 다양할 뿐만 아니라 인쇄 자료와 근본적으로 다른 특성으로 인하여 장기적 보존 및 다음 세대로의 접근을 보장하기 위한 상세한 메타데이터를

필요로 한다. 그러나 현재 우리나라에서는 디지털 정보자원의 장기 보존과 접근을 위한 메타데이터는 DC나 MARC을 사용하는 수준에 머물러 있다. 본 연구는 향후 우리나라 디지털 정보자원의 장기 보존과 접근을 위하여 상호운용성을 갖춘 표준적인 보존 메타데이터의 요소를 설정하기 위한 기초 연구로서 수행되었다.

이를 위하여 본 연구는 새로운 과제로 부각되고 있는 디지털 정보자원의 장기 보존과 접근을 위하여 디지털 정보자원의 특성과 디지털 보존의 개념을 검토하였다. 이와 함께 디지털 정보자원의 장기보존과 접근을 위한 보존 메타데이터의 필요성을 구체적으로 규명하였다. 또한 OAIS 참조 모형의 정보 구조를 검토함으로써 이모형이 디지털 정보자원을 위한 상호운용적인 보존 메타데이터 구조의 근거가 되고 있음을 확인하였다. 이어서 실제로 OAIS 참조 모형을 적용하여 개발된 CEDARS, NLA, NEDLIB, OCLC/RLG의 디지털 정보자원을 위한 보존 메타데이터를 대상으로 그 개발 과정과 목적을 살펴보았다. 그리고 나서 OAIS 참조 모형의 정보 구조를 기준으로 하여 아카이브 정보 패키지를 구성하는 2가지 핵심 요소인 Content Information과 Preservation Description Information을 대상으로 CEDARS, NLA, NEDLIB, OCLC/RLG의 디지털 정보자원을 위한 보존 메타데이터 요소를 구체적으로 비교 분석하였다. 끝으로 본 분석 결과를 기초로 하여, 향후 우리나라 디지털 정보자원의 장기 보존과 접근을 위한 메타데이터의 기본적인 요소 세트 설정 방안을 5장에서 제시하였다.

아울러 우리의 디지털 정보자원을 위한 상호운용성을 갖춘 보존 메타데이터 개발을 위해서는 우선적으로 도서관계에서의 이에 대한 필요성 및 공감대 확대가 필요하다고 본다. 그 개발은 학계와 도서관 전문가, 기록보존 전문가 및 시스템 전문가 등의 공동의 참여 하에 컨소시엄 형태로 수행되어야 할 것이다. 또한 국립중앙도서관이 주도적인 역할을 수행하면서 민간부문과 협력하여 활동을 전개해 나가야 할 것이다.

한편 본 연구의 제한점이라 할 수 있는 우리나라 디지털 정보자원의 장기 보존과 접근을 위한 메타데이터요소의 필수/선택 구분과 그 반복 사용 여부는 후속 연구를 통해 제시되기를 기대한다.

<참고문헌은 각주로 대신함>