

디지털자원의 보존 메타데이터 요소세트 개발에 관한 연구*

A Study on Preservation Metadata for Digital Resources

서 은 경(Eun-Gyoung Seo)**

초 록

디지털자원의 영구보존을 효율적으로 수행하기 위해서는 디지털자원의 고유의 속성과 변화를 설명해주는 지적 정보 즉 보존 메타데이터가 구조적으로 그리고 체계적으로 표기되어야 한다. 보존 메타데이터는 외부적/내부적 변화와 상관없이 사용자가 보다 쉽게 원하는 진본 데이터에 접근하여 검색할 수 있도록 하며 관리자는 보다 효율적으로 정보자원을 관리하고 보존할 수 있도록 한다. 따라서 2002년 OAIS 참조모형(ISO 14721)이 보존 메타데이터 표준을 권고한 이후 외국의 여러 국가 디지털도서관에서는 각 나라의 실정에 맞게 보존 메타데이터 요소세트 개발에 노력해왔다. 우리나라도 디지털도서관의 실정에 적합한 보존 메타데이터를 개발해야 할 것이다. 본 연구는 먼저 OAIS가 권고한 정보모델과 디지털자원의 장기적 보존을 위해서 외국 도서관에서 개발된 보존 메타데이터 요소세트에 대해서 비교 분석하였고, 우리나라 대학도서관에서는 어떻게 보존 메타데이터를 구축하고 활용하고 있으며, 어떠한 보존 메타데이터 요소들이 대학도서관에 적합한지를 조사하였다. 이를 기반으로 하여 OAIS 참조모델이 권고하는 기본적 구조에 맞추어 우리나라 대학의 디지털도서관에 활용될 수 있는 디지털자원의 보존 메타데이터 요소세트를 제안하였다.

ABSTRACT

The preservation metadata which is defined the property and history of changes must be describes logically and structurally for implementing long-term digital preservation. The preservation metadata makes the users to access and retrieve authentic digital resources and the managers of digital archives to manage and preserve more effectively the digital resources. his paper is review of recent developments relating to digital preservation metadata including the taxonomy of information object classes of Reference Model of OAIS and outlines some library-based projects. Next, the paper investigates how to develop and use the preservation metadata in university libraries and what is the core preservation metadata elements. Finally, the paper recommends the set of preservation metadata elements with broad applicability to digital preservation in university libraries.

키워드: 디지털자원 보존, 보존 메타데이터, 디지털자원 관리, 디지털보존,
digital preservation, preservation metadata

-
- * 이 논문은 2004년도 한성대학교 교내연구비 지원과제임
 - ** 한성대학교 지식정보학부 교수(egseo@hansung.ac.kr)
 - 논문접수일자 : 2005년 8월 20일
 - 게재확정일자 : 2005년 9월 13일

1. 서론

도서관에서의 디지털자원의 관리는 소장, 접근, 보존 면에서 종이기반의 자원관리와는 상당히 다르다. 특히 디지털자원의 활용기간을 영구적으로 연장시켜주기 위하여 이루어지는 디지털 보존 방법은 자원의 형태와 컨텐츠를 변화시키지 않고 그 매체 그대로 유지시키는 전통적인 보존 행위와 달리 상당히 난해하며 체계적 절차를 필요로 한다. 이는 상당히 빠르게 변화되고 있는 정보기술과 다양한 구축포맷에 절대적으로 의존해야 하기 때문이라 할 수 있다. 즉 소프트웨어와 하드웨어의 단명성과 다양성은 디지털자원에 대한 가독성을 저하시키거나 부분적으로 그 기능을 상실하도록 하기 때문에 디지털자원을 영구적으로 유지하고 보존하기 위해서는 주기적이며 체계적인 디지털보존 활동이 불가피하다.

OCLC/RLG는 디지털보존을 0과 1로 이루어진 bitstream을 계속적으로 유지하여 원형 그대로의 정보가 저장 매체로부터 읽혀질 수 있도록 하는 생존가능성(viability), 내용에 접근하여 이용자나 관리자가 처리 및 표기할 수 있도록 하는 실행가능성(renderability), 이용자가 디지털정보를 인지 또는 식별할 수 있게 해주는 이해가능성(understandability)의 기능 세 가지를 수행하는 활동이라고 정의하고 있다(OCLC/RLG Working Group 2002). 생존가능성과 실행가능성은 bitstream을 유지하여 내용을 보존하고 활용할 수 있도록 하는 기능을 말하고 이해가능성은 형식, 스타일, 외형, 그리고 기능 모두를 유지시키는 활동을 가리키는 것으로, 디지털보존을 디지털 환경이 변한다 할지라도 디지털자원의 진본성과 무결성을 보장하면서 원 내용을 보존하고

이와 동시에 그 디지털자원을 활용시켜주는 원문 접근기능, 검색처리기능, 원문전달기능, 원문재생기능 등과 같은 기능들이 사라지지 않게 유지시키는 작업이라고 정의할 수 있다.

디지털자원의 영구보존을 효율적으로 수행하기 위해서는 디지털보존의 핵심 기능인 디지털자원의 생존가능성, 실행가능성, 이해가능성을 유지하기 위해 필요한 지적 정보 즉 보존메타데이터가 구조적으로 그리고 체계적으로 표기되어야 한다. 메타데이터는 정보자원의 속성을 기술하는 데이터를 일컫는 용어로 “정보자원을 설명하는 요소집합” 또는 “정보자원을 보다 쉽게 검색, 활용 및 관리할 수 있도록 정보자원에 대하여 기술하고 설명하는 구조적 정보”로 정의된다(남영광 외 2005, Hodge 2001). 따라서 보존 메타데이터 항목에 따라 각각의 정보들이 표기되어질 때 외부적/내부적 변화와 상관없이 사용자는 보다 쉽게 원하는 진본 데이터에 접근하여 검색할 수 있으며 관리자는 보다 효율적으로 정보자원을 관리하고 보존할 수 있다.

일반적으로 메타데이터를 크게 설명적(descriptive) 메타데이터, 관리적(administrative) 메타데이터, 그리고 구조적(structural) 메타데이터로 세분하고 보존 메타데이터는 디지털자원의 처리 및 관리를 용이하게 해주고 지원해주는 관리적 메타데이터의 한 종류로 분류해왔으나, 현재 디지털보존의 중요성이 상당히 높게 인식됨에 따라 보존 메타데이터를 관리적 메타데이터에서 분리하여 취급하고 있다. 오히려 모든 메타데이터 요소가 보존에 관계되어있음을 강조하여 Day(1999)는 보존 메타데이터에 디지털 객체의 진본성 및 출처와 그 관계를 증명할 수 있도록 하는 것 이외에, 디지털 기술이나 자원의 발견을

도와주는 전통적인 메타데이터의 기능까지도 포함되어야 한다고 제안하였다. 일반적으로 보존 메타데이터는 다른 목적으로 사용되는 메타데이터보다 다양하고 세부적인 항목들이 많이 포함되기도 한다. 즉 정보자원의 검색을 도와주는 기술적 메타데이터, 디지털 정보자원의 관리를 지원해주는 관리적 메타데이터, 응용수준에서 특정 디지털 객체의 표기가 가능하도록 하는 구조적 메타데이터 이외에 맥락정보와 관련된 메타데이터, 디지털자원의 이력 및 출처정보를 다루는 메타데이터, 디지털자원의 진본성 인증과 관련된 메타데이터 모두를 포괄하기도 한다.

디지털자원의 장기적 보존은 정보기술적인 문제와 경제적 및 정책적 문제를 다루는 연구영역으로 다각적인 관점에서 연구가 수행되고 있으며 그 중의 하나가 바로 보존 메타데이터에 관한 연구라 할 수 있다. 보존 메타데이터에 관한 현재 연구 흐름은 보존에 관련된 메타데이터 요소를 개발하고 이에 따른 상위개념의 지침을 제안하는 정도이다. 2002년 OAIS(Open Archival Information System) 참조모형(ISO 14721)은 보존 메타데이터가 포괄해야 하는 기본적 카테고리를 제안하였고, 그 외 영국의 디지털 아카이브 프로젝트인 CEDARS, 유럽의 국가도서관 협력프로젝트인 NEDLIB, OCLC와 RLG의 PREMIS 프로젝트, 그리고 호주 국립도서관 등에서 보존 메타데이터 표준을 제안하고 사용하고 있다.

현재 우리나라 대부분의 대학도서관에서는 디지털자원을 자체적으로 구축하거나 구입하여 서비스를 제공하고 있고 이러한 디지털자원의 안정성과 기능성을 최대한 확보하고 유지하기 하기 위하여 노력하고 있는 실정이다. 이제 디지털자원을 보다 효율적이면서 장기적으로 관리하기 위

해서는 우리나라 대학도서관의 실정에 적합한 보존 메타데이터를 개발해야 할 것이다. 본 연구는 디지털자원의 장기적 보존을 위해서 개발된 현존의 보존 메타데이터에 대해서 비교 분석하고, 우리나라 대학도서관에서는 어떻게 보존 메타데이터를 구축하고 활용하고 있으며, 어떠한 보존 메타데이터 요소들이 대학도서관에 적합한지를 조사하였다. 이를 기반으로 하여 본 연구는 우리나라 대학도서관에 적합한 보존 메타데이터의 요소 집합을 권고하고자 한다.

2. 전자도서관에서의 보존 메타데이터

디지털 아카이브를 구축한 국가 대표적인 기관은 디지털자원의 영구적 보존을 지원해주는 보존 메타데이터 구축에 많은 노력을 하고 있다. 일반적으로 보존 메타데이터는 디지털도서관을 구축하여 전자정보자원을 서비스하는 전자도서관 또는 정보서비스 기관들뿐만 아니라 전자기록물 아카이빙을 담당하는 기관이나 이와 관련 프로젝트에서도 관심을 갖는 분야이어서 이와 관련된 많은 연구들이 산출되고 있다. 이 중 본 연구에서는 전자기록물이 아닌 디지털자원의 영구보존을 위하여 각국의 국립도서관이나 대표기관의 프로젝트에서 제시한 보존 메타데이터들을 중심으로 살펴보았다.

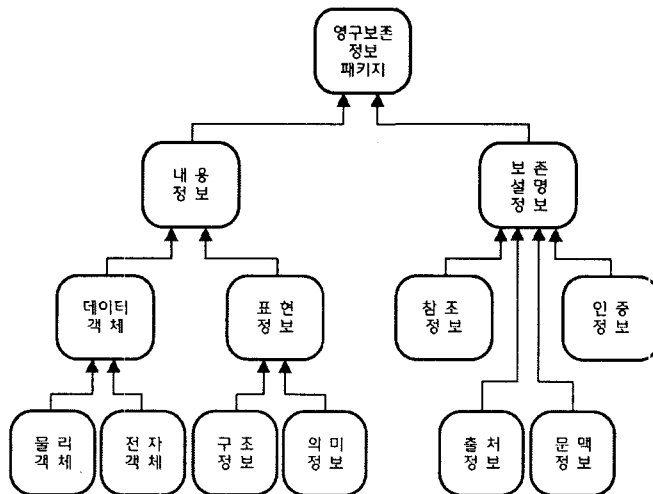
2.1 OAIS 참조모델(ISO 14721)

OAIS 참조모델은 디지털 아카이브에 대한 개념상의 프레임워크이다. 이 모델은 디지털 아카이빙에 관련된 전문용어와 개념을 정립하고, 핵심 요소

와 디지털 아카이빙 활동에서 발생하는 고유한 프로세스들을 식별하며, 디지털 객체와 관련된 다른 메타데이터들에 관한 정보모델을 제안하고 있다. 그러나 이 참조모델은 구체적 구현방식을 명시하지 않음으로 해서 디지털 객체 유형이나 기술적인 문제에 대해 중립적인 성격을 취하며 논리적 모델을 지향하고 있다(이소연 2002, CCSDS 2002).

디지털 아카이브의 핵심적인 기능의 속성을 정의하는 것에 덧붙여, OAIS 참조모형은 영구적 디지털보존을 위한 상위 개념의 메타데이터 구조와 요건을 제공하고 있다. 즉 참조모형은 디지털 환경에서 3가지 종류의 정보패키지, 즉 제공 정보패키지(Submission Information Packages), 영구보존 정보패키지(Archival Information Packages), 보급 정보패키지(Dissemination Information Packages)의 사용을 제안하였고 이 중 영구보존 정보패키지는 기본적으로 내용정보(Content Information)와 보존설명정보(Preservation Description Information)

로 구성되도록 제시하였다. 내용정보는 보존되어야 하는 정보를 기술하는 것으로 비트정보에 관한 사항을 기술하는 내용-데이터객체(Content Data Object)와 비트 열이 실제 담고 있는 정보 즉 사용자가 이해할 수 있도록 표현되는 정보를 기술하는 표현정보(Representation Information)로 나누어 기술하도록 하였다. 데이터 객체는 물리적 객체와 전자적 객체 두 가지로 나눌 수 있으며, 표현정보는 내용정보가 가지는 구조정보와 의미정보를 포함한다. 또한 내용정보의 식별자 역할을 하는 보존설명정보에는 내용정보에 부여된 식별자를 기술하고 필요한 경우 설명해주는 참조정보(Reference Information), 생성이유와 다른 정보들과의 관계를 기술해주는 맥락정보(Context Information), 그 정보의 이력사항을 기술하는 출처정보(Provenance Information), 인증 및 검증과 관련된 정보를 기술하는 인증정보(Fixity Information)를 기술하도록 하였다 (<그림 1> 참조). OAIS는 각각 메타데이터 항목



<그림 1> OAIS의 정보모델

에 기될 수는 예로서 참조정보에는 서지정보, 영구적 식별자; 맥락정보에는 이전 환경에서 생성된 문헌과의 연계정보; 출처정보에는 디지털화 과정에 대한 정보, 원 디지털 자원에 대한 정보, 버전 변환 및 변화과정 정보; 인증정보에는 디지털 사인, 인증 식별자 등을 제시하였다(CCSDS 2002).

이러한 프레임워크는 디지털 아카이브 안에서 구현되는 구조적인 데이터 모델이며 디지털자원의 유형과 보존기술, 이 두 가지 모두에 의존적이기 때문에 OAIS 참조모델은 일반적인 응용가능성이 있는 보존 메타데이터 프레임워크를 개발하는데 유용한 출발점이 될 것이다.

2. 2 CEDARS 프로젝트

CEDARS(CURL Exemplars in Digital Archives European Deposit Library) 프로젝트는 2001년에 디지털자원의 장기적 보존을 가능하게 하는 보존 메타데이터 요소의 표준 세트를 제안하였다. 대학연구도서관협의회(Consortium of University Research Libraries: CURL)에서 수행된 CEDARS 프로젝트의 목적은 디지털 보존과 관련된 전략, 실행방법, 그리고 실제적 문제를 다루는 것으로 다음과 같이 세 가지 방향으로 연구되었다: 1) 디지털보존 전략과 기술 개발, 2) 디지털 컬렉션 개발과 권한, 3) 디지털자원의 영구적 보존을 위해 필요한 메타데이터의 정의(Russell 2000).

CEDARS의 보존 메타데이터 개발연구의 목적은 CEDARS가 구축한 디지털 아카이브에서 사용될 수 있도록 하는 것이며 또한 보존 메타데이터 표준을 개발하기 위한 국제적인 노력에 동

참하는데 있다. CEDARS가 제안한 보존 메타데이터 요소는 장기적으로 보존된 디지털 객체 내용에 의미 있게 접근을 지원할 수 있도록 디지털자원에 대하여 설명적(descriptive), 관리적(administrative), 기술적(technical), 그리고 법적인(legal) 정보를 기술할 수 있도록 구성되어 있다. 또한 전자도서관에서 관리되는 디지털 자원들이 다양한 포맷을 가질 것이라는 예상 하에, 다양한 형태의 디지털자원들이 그 특성에 맞추어 기술될 수 있도록 광범위한 보존 메타데이터 요소를 고안하였다(OCLC/RLG Working Group 2001).

CEDARS 프로젝트는 OAIS 참조모델에서 권고한 최상위 분류에 따라 구조화된 보존 메타데이터 요소 세트 초안을 탄생시켰다. 다만, 출처 정보 안에 저작권관리 사항을 보다 상세하게 기술하도록 하여 디지털 콘텐츠 저작권자의 지적재산권 문제를 보다 명확하게 하여는 의도를 보여주고 있다. 그러나 일반적으로 보존 메타데이터의 기본 요소는 비교적 고수준에서 정의되었으며, 광범위한 종류의 디지털자원에 적용될 수 있도록 고안된 장점을 가진다(Day 2001). 제안된 보존 메타데이터의 요소와 구조를 5단계 안에서 살펴보면 <표1>과 같다.

2. 3 OCLC/RLG 워킹그룹

2000년에 OCLC와 RLG는 보존 메타데이터 개발을 위해 워킹그룹을 형성하여 2001년에 Preservation Metadata for Digital Object의 현황에 대한 보고서를 발간하였고(OCLC/RLG Working Group 2001), 2002년에는 실용적인 보존 메타데이터 프레임워크를 제안하였다.

〈표 1〉 CEDARS가 제시한 메타데이터 요소

1st level	2nd level	3rd level	4th level	5th level
내용정보	내용 데이터 객체 표현정보	Structure Information	Underlying abstract form description	
			Transformer objects	Platform
				Parameters
				Reeder/analyze engines
				Output format
			Input format	
		Render analysis objects	Platform	
			Parameters	
			Reeder/analyze engines	
			Output format	
		Semantic Information	Render analysis objects	Platform
				Parameters
				Reeder/analyze engines
				Output format
				Input format
보존설명 정보	참조정보	Resource Description		
		Existing Metadata	Existing records	
	맥락정보	Related Information Objects		
		출처정보	History of Origin	Reason for creation
	Custody history			
	Change history before archiving			
	Management History		Original technical environment	Prerequisites
				Procedures
				Documentation
	Right Management	Reason for preservation		
		Ingest process history		
			Administration history	Action history
			Policy history	
인증정보	Authentication Indicator	Negotiation		
		Rights information	Copy statement	
			Actors	
		Actions		

(OCLC/RLG Working Group 2002). OAIS 참조모델은 보존 메타데이터의 범위를 대분류하여 구분하였기 때문에 실제 구현시 충분한 정보의 유형을 제공하기에는 부족하다고 보고 OCLC와 RLG는 OAIS 대분류 안에서 디지털 아카이브에서의 디지털자원 보존관리를 보다 용이하도록 포괄적이며 상세한 프레임워크를 제안하였다. 특히 내용데이터 객체 항목에서 데이터간의 관계를 문서화할 수 있도록 한 점과 표현정보 항목에서 아카이브된 데이터 객체의 내용을 설명하고

렌더링될 수 있도록 한 것은 사용자들의 이해를 보장하기 위해 반드시 필요하다고 볼 수 있다. 또한 출처정보 항목에서 시간이 경과한 후에도 데이터객체의 변화를 보여줄 수 있는 사건(events)을 순서대로 기록함으로써 사용자들에게 객체의 내용이나 외관에 있어서의 변화에 대한 트랙(track) 정보를 보여줄 수 있도록 용이하게 보존관리를 할 수 있도록 지원하고 있다. 다음 <표 2>는 OCLC와 RLG가 제안한 보존 메타데이터 요소 세트를 5단계 안에서 정리한 것이다.

<표 2> OCLC/RLG가 제시한 메타데이터 요소

1st level	2nd level	3rd level	4th level	5th level
내용정보	내용 데이터 객체	Content Data Object Description	Underlying abstract form description	
			Structural type	
	Technical infrastructure of complex object			
	File description			
	Installation requirement			
	Size			
	Access inhibitors			
	Access facilitators			
	Significant properties			
	Functionality			
	Description of rendered content			
	Quirks			
	Documentation			
	표현정보	Environment Description	Software Environment	Redering programs
				Operating system
			Hardware Environment	Computation resources
				Storage
				Peripherals

1st level	2nd level	3rd level	4th level	5th level
보존설명 정보	참조정보	Archival system Identification	Value	
			Construction method	
			Responsible agency	
		Global Identification	Value	
			Construction method	
			Responsible agency	
	Resource Description	Existing metadata	Existing records	
	맥락정보	Reason for creation		
		Relationships	Manifestation	Relation type
				Identification
			Intellectual contents	Relation type
	Identification			
	출처정보	Origin	Event	Designation
				Procedure
				Date
				Responsible agency
				Outcome
				Note
				Next occurrence
		Pre-Ingest	Event	Designation
Procedure				
Date				
Responsible agency				
Outcome				
Note				
Next occurrence				
Ingest	Event	Designation		
		Procedure		
		Date		
		Responsible agency		
		Outcome		
		Note		
		Next occurrence		
Archival Retention	Event	Designation		
		Procedure		
		Date		

1st level	2nd level	3rd level	4th level	5th level
				Responsible agency
				Outcome
				Note
				Next occurrence
		Rights Management	Event	Designation
				Procedure
				Date
				Responsible agency
	인증정보	Object Authentication		Authentication type
				Authentication procedure
				Authentication date
				Authentication result

OCLC와 RLG의 보존 메타데이터 프레임워크 개발은 아카이브 시스템에서 보존 메타데이터를 실행시키기 위한 전략과 실제구현, 다양한 유형의 디지털보존 활동을 설명할 수 있는 적합한 기술(description) 수준에 대한 평가, 보존 메타데이터의 자동 생산을 위한 알고리즘 개발과 같은 후속연구의 토대가 되었다(OCLC/RLG Working Group 2002).

2. 4 NEDLIB 프로젝트

1998년부터 2000년까지 실행된 NEDLIB (Networked European Deposit Library) 프로젝트는 8개 유럽국가의 국립도서관, 출판업자, 정보기술 기관, 그리고 국가 아카이브의 컨소시엄 형태로 수행되었다. 이 프로젝트의 목적은 OAIS 모델에 기반 한 디지털 아카이브인 DSEP (Deposit System for Electronic Publications)

의 기능적 프레임워크를 개발하는 것으로 아카이브의 기본 6가지 기능(흡수, 아카이브 보존, 보존 계획, 데이터 관리, 접근, 운영)에 대해 그 구조와 정의를 제시하였고 이에 필요한 보존 메타데이터 요소를 제안하였다(Lupovici & Mazanes 2000).

NEDLIB은 변화하는 기술 환경 안에서 많은 양의 데이터의 효율적 처리와 영구적 보존을 지원해줄 수 있는 필수적인 보존 메타데이터를 제안하였다. 제안된 메타데이터의 스키마는 '핵심' 보존 메타데이터 요소를 식별하는 데 초점을 맞추었기 때문에 CEDARS 요소 세트보다 훨씬 소규모이다. 모든 요소 세트는 다양한 포맷과 타입에 적용할 수 있는 있도록 하였고 특히 기술적 쇠퇴(obsolescence)에 관한 문제를 보다 효과적으로 다룰 수 있도록 제안되었다 (OCLC/RLG Working Group 2001). 다음 <표 3>는 NEDLIB이 제안한 보존 메타데이터 요소 세트이다.

〈표 3〉 NEDLIB이 제시한 메타데이터 요소

1st level	2nd level	3rd level	4th level	5th level	
내용정보	내용 데이터 객체				
		표현정보	Specific Hardware Requirements	Specific microprocessor requirements	
	Specific multimedia requirements				
	Specific peripheral requirements				
	Operating system		Name		
			Version		
	Interpreter and compiler		Name		
			Version		
	Object format		Instruction		
			Name		
	Application		Version		
		Name			
	보존설명정보	참조정보	Creator		
			Title		
			Date of Creation		
Publisher					
Assigned Identifier			Value		
			Construction method		
			Responsible agency		
URL		Value			
		Date of Validation			
출처정보 (Change History)		Main Metadata Concerned	Date		
			Old Value		
			New Value		
			Tool	Name	
				Version	
		Reverse			
Other Metadata Concerned	Old value				
	New value				
인증정보	Checksum	Value			
		Algorithm			
	Digital signature				

2. 5 Natural Library of Australia(NLA) 표준안

디지털 보존과 관련된 문제에 대해 오랫동안 관심을 가져온 호주의 국립도서관(NLA)은 1996년에 PANDORA(Preserving and Accessing Networked Documentary Resource of Australia)라는 디지털 아카이브를 개발하였고 1999년에는 보존메타데이터의 표준을 제안하였다(Phillips et al. 1999). 이 표준안은 디지털 아카이브 운영에 필요한 보존 메타데이터 요소들의 정의뿐만 아니라 디지털자원과 메타데이터의 관계와 메타데이터 식별/검색을 명확하게 해주는 엔티티 - 관계 모델링 기반의 논리적 데이터모델도 제시하였다(Day 2001).

NLA의 보존 메타데이터는 생성 당시부터 디지털화된 자원(born digital)과 대체자원으로

나중에 생성된 디지털 자원(digital surrogate) 모두에게 적합한 메타데이터 요소를 제안하였고 보존관리만을 초점을 맞춘 결과 정보자원 검색에 필요한 메타데이터 요건을 고려하지 않았다는 점과 기술 중립적으로 모든 보존기술을 개방적으로 받아들일 수 있도록 한 점이 NLA 보존 메타데이터의 특징이라 할 수 있다. 또한 정보에 대한 설명의 상세 수준(descriptive granularity)에 따른 메타데이터 요소의 필요성 및 다양성을 강조하여 NLA 보존 메타데이터 세트는 세가지 수준의 입도(컬렉션, 객체, 하부 객체(또는 파일))에서 25개의 고수준의 요소를 정의하고 있다(OCLC/RLG Working Group 2001). 다음 <표 4>는 NLA가 제안한 보존 메타데이터 요소 세트이다.

<표 4> NLA가 제시한 메타데이터 요소

1st level	2nd level	3rd level
1. Persistent Identifier		
2. Data of Creation		
3. Structural Type		
4. Technical Infrastructure of complex object		
5. File description	Image	Format and version
		Resolution
		Dimensions
		Colour
		Tonal resolution
		Colour space
		Colour management
		Colour lookup table
		Orientation
	Compression	
	Audio	Format and version

1st level	2nd level	3rd level
		Resolution
		Duration
		Bit rate
		Compression
		Encapsulation
		Track number and type
	Video	File format and version
		Framed Dimensions
		Duration
		Frame rate
		Compression
		Video encoding Structure
	Text	Video sound
		Format and version
		Compression
		Character set
		Associated DTD
	Database	Structural Divisions
		Format and version
		Compression
Datatype		
Representation form		
Executables	Maximum size of data	
	Minimum size of data	
Code type and version		
6. Known System Requirements		
7. Installation Requirements		
8. Storage Information		
9. Access Inhibitors		
10. Finding and Searching Aids, and Access Facilitators		
11. Preservation Action Permission		
12. Validation		
13. Relationships		

1st level	2nd level	3rd level
14. Quirks		
15. Archiving Decision		
16. Decision Reason		
17. Institution Responsible for Archiving Decision		
18. Archiving Decision		
19. Decision Reason		
20. Institution Responsible for Archiving Decision		
21. Intention Type		
22. Institution with preservation responsibility		
23. Process	Description of process	
	Name of the Agency Responsible for the Process	
	Critical Hardware Used in the Process	
	Critical Software Used in the process	
	How Process was Carried Out	
	Guidelines Specified to Implement Process	
	Date and time	
	Result	
	Process Rationale	
	Changes	
Other		
24. Record Creator		
25. Other		

2. 6 National Library of New Zealand(NLNZ) 표준안

뉴질랜드 국립 도서관(NLNZ)는 디지털자원의 보존과 관리를 위해 디지털 아카이브를 설립하였고 역시 디지털자원에 대한 접근을 향상시키고자 2002년에 보존 메타데이터 구현을 위한 프레임워크를 발표하였고 2003년에 보존 메타데이

터 스키마를 제안하였다(NLNZ 2003). 특히 NLNZ는 디지털자원의 보존을 지원하는데 필요한 데이터 요소, 보존 메타데이터를 수집하고 저장을 위한 데이터베이스, 디지털자원에 대한 메타데이터뿐만 아니라 메타데이터를 관리하기 위해 필요한 여러 데이터 요소들의 정의를 상세히 설명하고 그 관계를 구조적 설명해주는 보존 메타데이터 데이터 사전 모델을 개발하였다. 또한

NLNZ의 스키마는 시스템 내에서 실행 템플릿으로 제공될 수 있도록 하였고 이와 동시에 전 세계적으로 개발되고 있는 보존 메타데이터 관련 표준들과 일치할 수 있도록 개발되었다.

NLNZ 스키마는 4개의 엔티티로 구분된다. '엔티티 1-객체'는 디지털자원을 기술하는 18개 요소로 각각의 객체를 식별하고 보존 관리에 관련된 특성을 기술한다. '엔티티 2-프로세스'는 디지털자원을 활용함으로써 일어나는 모든 과정을 기록하는 13개 요소를 포함한다. '엔티티 3-파일'

은 엔티티 1에서 식별된 논리적 객체(파일)에 대한 기술적인(technical) 정보 또는 특성을 9개의 요소에서 기술하도록 하고 있다. 특히 이미지, 오디오, 비디오, 텍스트와 관련된 기술적 특징을 상세하게 설명할 수 있도록 세부 요소를 두고 있다. '엔티티 4 -메타데이터 변화'는 디지털자원의 여러 변화 이력에 대한 정보를 기록하며, 총 5개의 요소로 구성된다(Searle & Thompson 2003). 다음 <표 5>는 NLNZ가 제안한 보존 메타데이터 요소 세트이다.

<표 5> NLNZ가 제시한 메타데이터 요소

Entity1-Object	Entity2-Process	Entity3- File	Entity4-Metadata Modification
1.1 Name of Object	2.1 Object identifier	3.1 Object identifier	4.1 Object identifier
1.2 Reference number	2.2 Process	3.2 File identifier	4.2 Metadata record modifier
1.3 Object identifier	2.3 Purpose	3.3 File path	4.3 Date and time
1.4 Group identifier	2.4 Person/agency performing process	3.4 File name and extension	4.4 Field modified
1.5 Persistent identifier	2.5 Permission	3.5 Former file name	4.5 Date modified
1.6 Preservation Master creation date	2.6 Permission date	3.6 File size	
1.7 Logical composition	2.7 Hardware used	3.7 File date and time	
1.8 Is part of Group	2.8 Software used	3.8 MIME type	
1.9 Structural type	2.9 Steps	3.9 File format	
1.10 Hardware environment	2.10 Result	3.10 File format version	
1.11 Software environment	2.11 Guidelines	3.11 Target indicator	
1.12 Installation requirements	2.12 Completion date and time	3.12 Image 3.12.1 Resolution 3.12.2 Dimensions 3.12.3 Bits per sample 3.12.4 photometric interpretation - color space 3.12.5 Photometric interpretation - ICC profile name	

Entity1-Object	Entity2-Process	Entity3- File	Entity4-Metadata Modification
		3.12.6 Colour map reference 3.12.7 Orientation 3.12.8 Compression	
1.13 Access inhibitors	2.13 Comments	3.13 Audio 3.13.1 Resolution 3.13.2 Duration 3.13.3 Bit rate 3.13.4 Compression 3.13.5 Encapsulation 3.13.6 Channels	
1.14 Access facilitators		3.14 Video 3.14.1 Frame dimensions 3.14.2 Duration 3.14.3 Number of frames 3.14.4 Frame rate 3.14.5 Codec method 3.14.6 Aspect ratio 3.14.7 Scan mode 3.14.8 Sound indicator	
1.15 Quirks		3.15 Text 3.15.1 Character set 3.15.2 Markup language	
1.16 Metadata record		3.16 Datasets	
1.17 Date of metadata record creation		3.17 System Files	
1.18 Structural composition			
1.19 Comments			

2. 7 PREMIS 워킹 그룹

PREMIS(Preservation Metadata: Implementation Strategies)는 2001년 OLCL/RLG의 보존 메타데이터 워킹 그룹이 형성되어 프로토타입 보존 메타데이터 요소를 제안한 후, 시스템 구현에 보다 적합하고 다양한 형태의 아카이브에 적합한 메타데이터 요소의 데이터 사전 모델을 개발하고자 하였다. 이 목적을 달성하고자 6개의 다른 나라(네덜란드, 독일, 미국, 영국, 캐나다, 호주)에

있는 국립도서관, 연구도서관, 박물관, 아카이브, 정부기관, 민간기업체를 대표하는 연구진은 먼저 2004년에 13개국에 퍼져있는 48기관을 대상으로 영구적 보존을 위한 정보저장소(preservation repository)에 대한 목적, 모형, 전략, 기술적 아키텍처, 메타데이터에 대하여 조사하였고 (PREMIS 2004), 조사결과물을 기초로 하여 2005년에 핵심 보존 메타데이터 요소와 데이터 사전을 개발하였다(PREMIS 2005).

OAIS 참조모델(ISO 14721)은 정보객체와

정보패키지간의 분류와 그와 관련된 메타데이터의 구조를 개념적으로 설명하고 그 개념적 구조 안에서 어떻게 보존 메타데이터를 매핑하는가에 대한 방법을 제시하였다면, PREMIS는 개념적 구조 설명보다는 실제 시스템 안에서 구현될 수 있는 핵심 의미적 요소(semantic unit) 각각을 제안하였다. 특히 PREMIS 워킹그룹은 디지털 자원의 생존가능성, 실행가능성, 이해가능성, 진본성, 식별성을 유지하는데 필요한 최상위 요소(unit)를 13개로 한정시켰으나 관리적, 구조적, 기술적(technical), 출처 및 맥락정보를 설명해주는 요소 모두가 포함되도록 하였다. 최상위 요소 13개는 객체식별자, 보존 수준, 객체범주, 객체특성, 생성/응용, 고유이름, 소장, 환경, 디지털 사인, 관계, 사건연결정보, 주제연결정보, 인가연결정보로 구성되었으며, 이중 객체특성에서는 구

성수준, 영속성, 크기, 포맷, 속성, 제한자에 대해서 그리고 환경에서는 의존정보, 소프트웨어, 하드웨어 등에 관한 상세 정보가 표기될 수 있도록 4단계의 하부 그룹을 갖도록 구조화시켰다. 각각의 의미적 요소의 구성은 <표 6>과 같다.

PREMIS가 개발한 데이터 사전에서 각각의 보존 메타데이터 요소는 10가지의 항목(즉, 요소이름, 하위 요소, 정의, 데이터 강제적 요소, 객체 종류, 응용, 예, 반복성, 필수여부, 사용법)에 따라 상세히 기술되었다. 모든 디지털 객체의 엔티티 종류를 비트(bitstream), 파일(file), 표현(representation)으로 구분하였고 각각의 엔티티는 하나 이상의 정의에 속하고 하나 이상의 사건(event)에 포함되며 하나 이상의 에이전트(agent)와 관련을 맺을 수 있는 속성을 가지도록 하였다.

<표 6> PREMIS가 제시한 메타데이터 요소

1st level	2nd level	3rd level	4th Level	
objectIdentifier	objectIdentifierType			
	objectIdentifierValue			
preservationLevel				
objectCategory				
objectCharacteristics	compositionLevel			
	fixity	messageDigestAlgorithm		
		messageDigest		
		messageDigestOriginator		
	size			
	format	formatDesignation	formatName	
			formatVersion	
		formatRegistry	formatRegistryName	
			formatRegistryKey	
	formatRegistryRole			
significantProperties				
inhibitors	inhibitorType			

1st level	2nd level	3rd level	4th Level	
		inhibitorTarget		
		inhibitorKey		
creatingApplication	creatingApplicationName			
	creatingApplicationVersion			
	dateCreatedByApplication			
originalName				
Storage	contentLocation	contentLocationType		
		contentLocationValue		
	storageMedium			
environment	environmentCharacteristic			
	environmentPurpose			
	environmentNote			
	dependency	dependencyName		
		dependencyIdentifier	dependencyIdentifierType	dependencyIdentifierValue
	software	swName		
		swVersion		
		swType		
		swOtherInformation		
		swDependency		
	hardware	hwName		
		hwType		
hwOtherInformation				
signatureInformation	signatureInformationEncoding			
	signer			
	signatureMethod			
	signatureValue			
	signatureValidationRules			
	signatureProperties			
KeyInformation	keyType			
	keyValue			
	keyVerificationInformation			
relationship	relationshipType			
	relationshipSubType			
	relatedObjectIdentification	relatedObjectIdentifierType		
		relatedObjectIdentifierValue		

1st level	2nd level	3rd level	4th Level
		relatedObjectSequence	
	relatedEventIdentification	relatedEventIdentifierType	
		relatedEventIdentifierValue	
		relatedEventSequence	
linkingEventIdentifier	linkingEventIdentifierType		
	linkingEventIdentifierValue		
linkingIntellectualEntityIdentifierType	linkingIntellectualEntityIdentifierType		
	linkingIntellectualEntityIdentifierValue		
linkingPermissionStatementIdentifier	linkingPermissionStatementIdentifierType		
	linkingpermissionStatementIdentifierValue		

2. 8 보존 메타데이터 요소들의 비교

앞에서 제안된 보존 메타데이터를 살펴보았을 때, 보존 메타데이터의 목적에 있어 공통적인 이해를 같이 하고 있음 알 수 있다. 즉 보존 관리자 들에게 의사결정을 위한 충분한 정보를 제공해줌으로써 아카이브된 디지털 객체의 관리를 지원해 주며, 계속적으로 변화하는 기술 환경에 직면하여 객체의 내용에 대한 접근을 유지시켜 주는 것을 그 목적으로 한다는 점이다.

각각의 보존 메타데이터의 구조와 요소의 특징을 살펴보면 다음과 같다. 첫째, CEDARS와 OCLC/RLG 메타데이터 접근은 OAIS 참조모델을 채택하여 기본적인 프레임워크를 제안했다는 점이다. 특히 CEDARS와 OCLC/RLG는 OAIS 참조모델에 명시되어 있는 개념과 전문용어를 그대로 사용하고 있으며 제안된 각각의 보존 메타데이터 요소는 OAIS 정보모델에서 명시되고 있는 정보 유형을 그대로 따르고 있다. 이런

의미에서, CEDARS와 OCLC/RLG의 보존 메타데이터는 고수준의 OAIS 참조모델 프레임워크의 구현 가능성을 보여주었다고 할 수 있다. 둘째, NLA와 NLNZ가 제안한 보존 메타데이터 요소 세트들이 OAIS 참조모델에 구조를 그대로 따르지 않았으나 보존메타데이터 요소를 OAIS 참조모델에서 일치하는 부분에 간단하게 대입할 수 있도록 매핑방법을 제시하고 있어 언제든지 그 구조를 변환할 수 있게 하였다. 셋째, NEDLIB 프로젝트에서 제안한 보존 메타데이터는 OAIS 참조모델의 기본 구조를 채택하였으나, OAIS 참조모델이 제시하는 모든 영역을 포함시키지는 않았다. 즉 맥락정보는 제시하지 않고 있으며, 표현정보에서 구조상의 정보와 의미상의 정보로 구분하지 않은 점이 다르다고 할 수 있다. 넷째, 가장 최근에 제안된 PREMIS 보존 메타데이터 의미적 요소는 OAIS의 개념적 구조를 따르기 보다는 시스템 구현을 보다 강조하였고 이를 위한 상세한 데이터 사전을 개발하는데 그

목적은 둔 것이 특징이라 할 수 있다.

제안된 모든 보존 메타데이터 스키마는 직접적이든 간접적이든 OAIS 참조모형을 기반으로 하여 구현되었다는 특징을 공유하고 있다는 것이다. 이러한 사실은 OAIS가 보존 메타데이터의 표준을 향한 공통의 합의를 이루어내는 시작점이 될 수 있다는 것을 시사한다. 그러나 OAIS 참조모델은 디지털자원에 대한 보존 메타데이터 개발을 하기 위한 유용한 시작점을 제공해 주지만 그 구조와 개념이 고수준(high level)이라는 한계점을 가지고 있다. 따라서 각 도서관이 구축한 디지털 아카이브 안에서 보존 메타데이터를 사용하기 위해서는 보다 그 환경에 적합하고 체계적인 요소 세트가 개발되어야 할 것이다.

3. 보존 메타데이터 요소 개발

본 연구는 외국에서 개발된 보존 메타데이터 조사와 더불어 한국의 대학도서관에서 이루어지고 있는 디지털자원 관리방식 및 보존 메타데이터 활용과 보존 메타데이터로 수용될 수 있는 요소의 선택여부를 조사하였다. 이러한 조사연구의 목적은 디지털자원 보존을 위하여 체계적으로 표기되어야 하는 보존 메타데이터 요소들을 우리나라 대학도서관에 적합하게 선택하기 위한 기초적 작업이라 할 수 있다. 이를 위하여 야후 코리아 대학도서관 카테고리에 표기된 139 대학도서관 중 체계적 표본 추출 방식을 이용하여 46개의 대학도서관을 선택한 다음, 선택된 대학도서관의 디지털자원 관리 담당자에게 온라인 질문지와 파일형태의 질문지를 전자메일로 발송하였다. 이중 37개의 도서관에서 응답을 하여 회수율 80%를

나타냈다.

3. 1 디지털자원과 메타데이터 관리 환경

본 장에서는 우리나라 대학도서관에서 이루어지고 있는 디지털자원 및 메타데이터 관리를 조사하여 디지털도서관에서 개발되어야 할 보존 메타데이터의 성격을 규명하였다. <표 7>에 나타났듯이 디지털자원의 메타데이터 관리는 전자정보실/멀티미디어실과 수서/정리를 맡고 있는 담당자가 수행하고 있었으며, 현 업무를 담당한 기간이 2년 미만인 사서가 43%를 차지하고 있으나 6년 이상 현 업무를 담당한 사서가 8명이나 있었다. 이는 응답자 중 15년 이상 한 기관에 종사한 사서가 17명이나 있음으로 해서 나타난 현상이라 볼 수 있다.

대학도서관에서 관리되고 있는 디지털자원은 구입한 것이 가장 많으며 다음, 자체적으로 구축한 자원과 기증 받은 자원순으로 구성되었고 유형별로 보면 대다수가 전자출판물과 디지털화된 자원이고 그 다음 웹자원과 오디오/비디오 자원이 많이 차지한 반면, 데이터 set은 다른 자원유형에 비교하여 매우 적은 규모로 소장되어 있음을 알 수 있다. 디지털자원을 포맷 형태로 나누어 보면, 이미지 문서가 가장 많았고 다음 데이터베이스, 비디오 파일, 텍스트 파일이 비슷한 규모로 소장되어 있고 그 다음으로 오디오 파일, 구조화 문서, 그림 파일 순으로 소장되고 있음을 알 수 있다. 또한 디지털자원의 보존을 위해 가장 중요하게 여기는 전략을 조사한 결과 '전담 보존전문가에 의한 관리'가 가장 중요하다고 답하였고(32%), 다음 '기술적 보존전략 수립'(27%), 'digital repository 설치'(22%), '보존 메타데이

터의 생성'(19%) 순으로 중요하게 생각하고 있었다. 이는 전문가가 디지털보존 업무를 부수적인 일이 아닌 주 업무로 담당해주기를 바라고 있는 것으로 나타났으며 보존 메타데이터의 생성이

가장 낮게 나타난 것은 응답도서관 중 16개의 도서관이 이미 보존 메타데이터를 별도로 생성하고 있기 때문에 나타난 결과로 보인다.

〈표 7〉 디지털자원의 보존 현황

번호	질 문	항 목	빈도수
1	부서 및 담당업무	수서 & 정리	10
		열람실 & 연속간행물	6
		전자정보	15
		멀티미디어제작 및 관리	4
		무응답	2
2	현재 업무를 담당한 기간	2년 미만	16
		2~4년 미만	8
		4~6년 미만	5
		6년 이상	8
3	귀 기간에 근무한 기간	5년 미만	11
		5~10년 미만	7
		10~15년 미만	2
		15년 이상	17
4	디지털 자원의 출처	자체적으로 구축	29
		구입	31
		연구자의 기증	9
		다른 도서관 또는 연구기관의 기증	10
		그 외	1
5	소장된 디지털 자원의 유형	전자출판물	78
		디지털화 된 자원	67
		오디오/비디오	32
		웹 자원	36
		데이터 sets	9
6	소장된 디지털자원의 포맷 형태	텍스트 문서 (.hwp, .doc...)	19
		이미지 문서 (.pdf, .tiff...)	26
		구조화 문서 (.html, .xml, .sgml...)	11
		그림 파일 (.gif, .jpg, .bmp...)	10
		오디오 파일(.avi, .mpeg, .wmv...)	12
		비디오 파일(.wav, .mps, .ac3...)	21
		Database	22
		그외	2
7	가장 중요한 디지털자원 보존 전략	전담 보존전문가에 의한 관리	12
		보존 메타데이터의 생성	7
		기술적 보존전략 수립	10
		digital repository 설치	8

〈표 8〉은 보존 메타데이터에 대한 관리자의 의견을 조사한 결과이다. 37개 대학도서관중 16 곳에서 보존 메타데이터를 별도로 생성하고 있으며 디지털자원의 보존을 위하여 표기되는 정보는 생

산자/제목/제작자/식별기호 등을 설명해주는 서지적 정보가 가장 많았고 그 다음으로는 디지털 자원의 포맷과 관련된 형태적 정보이고, 출처와 관리이력에 관한 출처정보, 권한 및 인가정보, 디

〈표 8〉 보존 메타데이터 관리 현황

번호	질문	항목	빈도수
1	디지털 자원에 대한 기술항목은?	서지적 정보	34
		형태적 정보	22
		기술적(technical) 정보	11
		관리적 정보	14
		출처정보	17
		권한 및 인가정보	16
2	디지털자원에 대한 보존 메타데이터 생성 여부	예	16
		아니오	21
3	보존 메타데이터 생성 방식은?	도서관 직원에 의하여 수동적으로 생성	10
		프로그램에 의하여 자동적으로 생성	2
		디지털자원 생산자(또는 기증자)가 생성	1
		디지털자원 생산기관에 의한 생성	4
		그 외	1
4	원하는 보존메타데이터 생성 방식은?	도서관 직원에 의하여 수동적으로 생성	4
		프로그램에 의하여 자동적으로 생성	23
		디지털자원 생산자(또는 기증자)가 생성	4
		디지털자원 생산기관에 의한 생성	2
		그 외	1
		무응답	3
5	보존메타데이터 범주의 중요순위는?	관리정보 메타데이터	62
		구조정보 메타데이터	32
		기술정보 메타데이터	31
		서지정보 메타데이터	80
		맥락정보 메타데이터	3
		출처정보 메타데이터	5
		통제정보 메타데이터	9
6	가장 적합한 보존 메타데이터 소장 및 관리방식?	관계 데이터베이스를 이용하여	16
		객체지향 데이터베이스를 이용하여	3
		XML 데이터베이스를 이용하여	15
		파일에 번들(bundle)되어	0
		파일관리를 이용하여	2
		무응답	1

디지털자산의 관리와 처리에 관한 관리적 정보, IT 및 하드웨어 환경과 관련된 기술적 정보 순인 것으로 나타났다. 이러한 양상은 보존 메타데이터 범주의 중요순위 결과에서도 나타났다. 즉 가장 중요하다고 보는 범주는 서지정보 메타데이터이고, 다음은 관리정보 메타데이터이고 구조정보 메타데이터와 기술정보 메타데이터를 세 번째로 중요하다고 보았고 영구적 보존을 위하여 다른 메타데이터에서 볼 수 없는 특수적 항목인 맥락 정보, 출처정보, 통제정보 메타데이터는 거의 중요하지 않다고 응답을 하였다. 대다수의 사서들이 여전히 어떤 형태의 메타데이터이건 간에 메타데이터의 주목적은 검색하고 관리하는 작업을 지원하는 것으로 보고 있음을 알 수 있다.

현재 보존 메타데이터를 생성하는 대다수의 대학도서관(62%)에서는 도서관 직원에 의한 수동적 방식으로 보존 메타데이터를 생성하고 있음을 알 수 있으며, 그 외 생산자나 생산기관에서 생성된 메타데이터를 그대로 사용하고 있는 것으로 나타났다. 그러나 대다수의 사서(62%)들은 프로그램에서 의하여 보존 메타데이터가 자동적으로 생성되기를 바라고 있는 반면, 생산자나 생산기관에서 알아서 생성해 주기 바라는 사서는 6명이 있었고 각각의 도서관에 적합하도록 직접 사서가 생성해야 한다고 생각하는 사서가 4명 있었다. 또한 보존 메타데이터의 소장 및 관리방식으로 가장 적합한 것은 관계 데이터베이스를 이용하거나(43%), XML 데이터베이스를 이용하는 것(41%)이라고 말하고 있다. 그 외 객체지향 데이터베이스나 파일관리를 통한 관리가 적합하다고 보는 사서는 매우 소수였다.

대체적으로 대학도서관에서 디지털자산 관리를 담당하는 사서는 보존 메타데이터의 중요성을

인지하고 있으며 다만 여전히 보존 메타데이터라 할지라도 가장 핵심적 메타데이터의 범주는 서지적 정보를 다루는 것으로 인지하고 있음을 알 수 있다. 또한 자동적으로 생성된 보존 메타데이터를 관계형 또는 XML 데이터베이스에 관리함으로써 효율적으로 그리고 영구적으로 디지털자원을 보존 할 수 있을 것이라는 사서들의 생각을 알 수 있었다.

3. 2 보존 메타데이터 요소 개발

본 연구는 우리나라 대학도서관의 실정에 적합한 보존 메타데이터 요소를 개발하기 위하여 먼저 각 대학의 디지털자원 담당자에게 연구자가 제시한 보존 메타데이터에 대한 수용여부를 조사하였다. 다음, 설문조사 결과와 외국 사례조사를 바탕으로 하여 7가지 범주 아래 17개 항목으로 구성된 전자도서관에서 활용될 수 있는 보존 메타데이터를 제안하였다.

〈표 9〉는 연구자가 제시한 7개의 메타데이터 범주(관리정보, 구조정보, 기술정보, 서지정보, 맥락정보, 출처정보, 통제정보)에 속하는 56개의 보존 메타데이터 항목에 대하여 37개 대학 소속 37명의 사서가 '필수적/선택적/불필요'라고 표시한 결과이다. 20명(54%) 이상의 사서가 '필수적'으로 표기되어야 한다고 한 항목은 생산일시정보, 갱신일시정보, 아카이브 고유번호, 보존 고유정보, 접근정보, 접근제한, 책임, 소재, 구조적 형태, 파일정보, 이미지데이터 정보, 오디오데이터 정보, 비디오데이터 정보, 텍스트 정보, 데이터베이스 정보, 크기, 하드웨어 환경, 소프트웨어 프로그램, 운영시스템, 식별자, 표제, 생산자, 생산기관, 생산일자, 내용, 주제, 언어, 형태, 인증정보, 권한

정보로 총 30개 항목이다. 관리정보 메타데이터 16개 중 8개의 항목이 과반수이상의 사서가 필수적이라 한 반면, 구조정보 메타데이터와 서지정보 메타데이터 범주인 경우에는 75%이상의 항목이 필수적이라고 답하였다. 그러나 맥락정보 메타데이터와 출처정보 메타데이터에 속하는 항목은 어느 누구도 필수적이라고 답하지 않았다. 이런 현상은 앞에서 언급한 보존 메타데이터 범주

의 중요순위에서도 나타났듯이 현 사서들은 보존 메타데이터라 할지라도 전통적으로 기술해온 서지정보와 디지털관리에 도움을 주는 관리정보와 구조정보에 관한 메타데이터 항목이 필수적으로 여기고 있는 반면, 디지털보존에는 필수적이며 특색적이라고 여겨지는 맥락정보와 출처정보는 그다지 중요하게 생각하지 않고 있음을 알 수 있다.

<표 9> 보존 메타데이터 항목에 대한 의견

메타데이터 범주	메타데이터 항목	필수적	선택적	불필요	계
1. 관리정보 메타데이터	① 생산일시정보	28	9	0	37
	② 저장일시정보	17	20	0	37
	③ 갱신일시정보	23	13	1	37
	④ 아카이브 고유번호	30	5	2	37
	⑤ 캡처 고유번호	11	24	2	37
	⑥ 보존 고유번호	25	12	0	37
	⑦ 재생산 고유번호	14	22	1	37
	⑧ 접근정보	23	14	0	37
	⑨ 접근제한	22	14	1	37
	⑩ 체크정보	14	21	2	37
	⑪ 캡처정보	9	25	3	37
	⑫ 마이그레이션 정보	16	19	2	37
	⑬ 재포맷팅 정보	10	25	2	37
	⑭ 책임	22	15	0	37
	⑮ 소재	27	10	0	37
	⑯ 예외사항	6	26	5	37
2. 구조정보 메타데이터	① 구조적 형태	26	11	0	37
	② 파일정보	32	5	0	37
	③ 이미지데이터 정보	22	15	0	37
	④ 오디오데이터 정보	21	15	1	37
	⑤ 비디오데이터 정보	21	15	1	37
	⑥ 텍스트 정보	24	13	0	37
	⑦ 데이터베이스 정보	25	12	0	37

메타데이터 범주	메타데이터 항목	필수적	선택적	불필요	계
	⑧ 크기	20	17	0	37
	⑨ 사용	13	24	0	37
	⑩ 특징	7	28	2	37
3. 기술정보 메타데이터	① 하드웨어 환경	22	14	1	37
	② 기술적 하부구조	16	21	0	37
	③ 소프트웨어 프로그램	26	11	0	37
	④ 운영시스템	22	15	0	37
	⑤ 데이터축적 환경	17	20	0	37
	⑥ 데이터변환 환경	18	19	0	37
	⑦ 설치요건	18	19	0	37
	⑧ 부수적 요건	2	35	0	37
4. 서지정보 메타데이터	① 식별자	32	5	0	37
	② 표제	33	4	0	37
	③ 생산자	29	8	0	37
	④ 생산기관	26	11	0	37
	⑤ 생산일자	24	13	0	37
	⑥ 내용	28	9	0	37
	⑦ 주제	27	8	2	37
	⑧ 요약	14	23	0	37
	⑨ 범위	9	27	1	37
	⑩ 언어	24	13	0	37
	⑪ 형태	25	12	0	37
	⑫ 영구적 식별자	16	21	0	37
5. 맥락정보 메타데이터	① 생산배경	11	26	0	37
	② 관련사항	3	33	1	37
6. 출처정보 메타데이터	① 생산유래	9	28	0	37
	② 관리내역	14	22	1	37
	③ 사용내역	8	28	1	37
	④ 보존내역	16	21	0	37
	⑤ 보존 메타데이터 내역	13	23	1	37
7. 통제정보 메타데이터	① 인증정보	28	9	0	37
	② 권한정보	26	10	1	37
	③ 인가정보	17	20	0	37
계		1081	957	34	2072

모든 응답자가 필수적이라고 또는 불필요하다고 생각하는 항목은 하나도 없었고 30명이상이 필수적이라고 답한 항목은 4개(표제(33), 파일 정보(32), 식별자(32), 아카이브 고유번호(30))이다. 또한 한 명이상이 불필요라고 표시한 항목은 20개이나 거의 한 두명 정도가 체크하였고 최대가 5명으로 그 항목은 관리정보에서 속하는 예외사항 항목이다. 다음의 <표 10>은 사서의 50% 이상이 필수적이라고 답한 항목을 정리한 것이다.

본 연구는 연구자가 제시한 모든 보존 메타데이터 중 응답자가 필수적이어야 한다는 항목을 중심으로 그러나 응답자가 간과하였으나 외국 사례에서 필수적으로 제시되는 항목을 포함하여 보존 메타데이터 요소를 개발하였다. 즉 디지털자

원의 영구적 보존을 위해서는 식별자 정보와 내용정보를 다루고 있는 서지정보 메타데이터, 고유번호 및 일시와 관련된 요소와 접근과 책임을 다루는 관리정보 메타데이터, 디지털자원의 구조, 포맷, 크기를 설명하는 구조정보 메타데이터, 디지털소장과 디지털환경을 지정해주는 기술정보 메타데이터, 이용자와 관리자의 인증 및 통제를 지정하는 통제정보 메타데이터, 디지털자원의 맥락면에서 디지털객체와 그 활동을 식별할 수 있도록 하는 맥락정보 메타데이터, 디지털정보와 메타데이터의 출처와 변화 내역을 다루는 출처정보 메타데이터가 필요하다고 보고, 이와 같은 7가지 범주아래 다음과 같이 상세 보존 메타데이터 요소를 분류하였다.

<표 10> 과반수이상의 사서가 선택한 필수적 항목

메타데이터 범주	필수적 항목	선택적 항목
관리정보 메타데이터	생산일시정보, 갱신일시정보, 아카이브 고유번호 생산일시정보, 갱신일시정보, 아카이브 고유번호, 보존 고유정보, 접근정보, 접근제한, 책임, 소재,	캡처 고유번호, 재생산 고유번호, 체크정보, 캡처정보, 마이그레이션 정보, 재포매팅 정보, 예외사항
구조정보메타데이터	구조적 형태, 파일정보, 이미지데이터 정보, 오디오 데이터 정보, 비디오데이터 정보, 텍스트 정보, 데이 터베이스 정보, 크기	사용, 특징
기술정보 메타데이터	하드웨어 환경, 소프트웨어 프로그램, 운영시스템	기술적 하부구조, 데이터축적 환경, 데이터변환 환경, 설치요건, 부수적 요건
서지정보 메타데이터	식별자, 표제, 생산자, 생산기관, 생산일자, 내용, 주제, 언어, 형태	요약, 범위, 영구적 식별자
맥락정보 메타데이터		생산배경, 관련사항
출처정보 메타데이터		생산유래, 관리내역, 사용내역, 보존내역
통제정보 메타데이터	인증정보, 권한정보	인가정보

■ 서지정보 메타데이터

- 식별자
 - 고유식별기호
 - 표제
 - 생산자
 - 생산기관
- 내용정보
 - 내용
 - 주제
 - 범주
 - 언어

■ 관리정보 메타데이터

- 고유번호
 - 아카이브 고유번호
 - 보존 고유번호
 - 영구적 식별자
- 일시정보
 - 생산일시
 - 저장일시
 - 갱신일시
- 접근정보
 - 접근
 - 제한
- 책임

■ 구조정보 메타데이터

- 구조 유형
- 포맷
 - 파일
 - 이미지
 - 오디오
 - 비디오
 - 텍스트
 - 데이터베이스

• 크기

■ 기술정보 메타데이터

- 소장
 - 소재정보
 - 매체정보
- 환경
 - 하드웨어
 - 소프트웨어

■ 통제정보 메타데이터

- 인증정보
- 권한정보

■ 맥락정보 메타데이터

- 객체관련정보
- 활동관련정보

■ 출처정보 메타데이터

- 원본내역정보
- 관리내역정보
- 보존메타데이터 내역정보

4. 결론

도서관에 디지털자원이 흡수될 때 표준화된 절차에 따라 이에 대한 정보가 상세하고 체계적으로 표기된다면 디지털자원의 관리는 보다 효율적으로 운영될 수 있고 디지털자원의 신뢰성은 더욱 높아질 것이다. 따라서 디지털관리에서 보존 메타데이터 구축은 핵심적 기능이라고 할 수 있다. 보존 메타데이터의 기능으로 : 1) 디지털자원에 대한 접근성과 이용성을 지속적으로 보장, 2) 디지털자원을 이해할 수 있도록 지원, 3) 디지털자원의 진본성, 신뢰성, 무결성을 보장할 수 있도록 지원, 4) 접근, 권한, 보안 관리, 5) 효율

적인 검색을 지원, 6) 디지털자원간의 논리적 연계를 지원하고 구조적이고 신뢰할 수 있는 방식으로 이들을 유지, 7) 디지털자원이 생산된 기술환경을 파악할 수 있게 하고 재생산할 수 있는 기술환경을 지원 등을 들 수 있다. 이러한 기능 수행함으로써 보존 메타데이터는 디지털자원 관리의 모든 영역에서 요구되는 사항을 충족시켜줄 것이다.

따라서 보존 메타데이터는 디지털자원의 흡수 시점에서 얻어지는 정보로서 디지털자원의 식별 정보와 내용정보, 디지털관리에 도움을 주는 고유번호 및 일시와 관련된 요소와 접근과 책임에 관련된 요소가 포함되어야 하며, 디지털자원이 장기 보존시에 여러 면에서 즉 구조, 포맷, 소장, 환경, 맥락 면에서 변화가 있기 마련인데 이에 대한 정보를 표기, 저장해주어야 할 것이다. 특히 현재의 상황을 설명해주는 정보도 필요하지만, 디지털자원의 구조와 기술적(technical) 속성에 있어서의 변화, 포맷과 주요 기술적 의존사항, 맥락면에서 디지털객체와 그 활동을 식별하고 찾아 낼 수 있도록 하는 정보, 변화 내역에 대한 정보 등이 포함되어야 하며, 마지막으로 디지털자원을 책임 있는 관리 환경에서 소장/보존하여 이용자에게 제공하기 위해서는 통제 및 보안 장치가 필요하므로 이에 대한 설명을 해주는 메타데이터 요소도 포함되어야 할 것이다. 본 연구는 이와 같은 내용을 담고 있는 보존 메타데이터를 제안하

였다. 제안된 보존 메타데이터는 세 단계 구조로 구조 갖도록 하였고 보존 메타데이터 요소는 36 항목으로 이루어지도록 하였다.

정보자원의 디지털화가 바로 '접근가능'을 의미하는 것이 아니다. 접근이 가능하기 위해서는 정보자원에 대하여 소재정보를 비롯하여 다양한 정보를 제공해주는 구조적 정보가 필요하다. 즉 디지털정보에 메타데이터가 표기되어야 한다. 또한 정보자원이 디지털형태로 계속해서 소장되고 있다고 해서 보존되는 것이 아니다. 장기적으로 또는 영구적으로 디지털보존을 하기 위해서는 특별한 보존관리가 필요하다. 즉 디지털자원에서의 접근 및 장기적 소장을 지원하고, 디지털보존에 대한 결정을 용이하게 하고, 디지털보존에 대한 우려를 불식시키고, 접근/재생산시 발생하는 위험을 최소화시키기 위해서는 보존 메타데이터가 구현되어야 한다. 본 연구는 OAIS 참조모델이 권고하는 기본적 구조에 맞추어 우리나라 대학의 디지털도서관에 활용될 수 있는 디지털자원의 보존 메타데이터 요소를 제안하였다. 제안된 보존 메타데이터 요소세트는 보존 메타데이터 구현의 첫걸음이라 할 수 있다. 이제 모든 디지털도서관에서 소장된 정보자원이 상호 연계되어 활용되고 또 영구적으로 보존되기 위해서는 국가적 디지털 자원의 장기적 보존계획과 정책이라 보존 메타데이터 표준 개발에 힘써야 할 것이다.

참 고 문 헌

남영광, 서태설, 황상원, 2005. ISO/IEC 11179에 따른 산업기술정보 메타데이터 표준화.

「정보관리연구」, 36(1): 57-75.
이소연 2002, 디지털 아카이빙의 표준화와 OAIS

- 참조모형. 「정보관리연구」. 33(3):45-68.
- Consultative Committee for Space Data Systems. 2002. *Reference Model for an Open Archival Information System (OAIS), Blue Book, CCSDS 650.0-B-1*. Washington, D.C.: CCSDS Secretariat, National Aeronautics and Space Administration.
- Day, Michael. 1999. Issues and Approaches to Preservation Metadata. In: *Guidelines for Digital Imaging*. London: National Preservation Office.
<http://www.rlg.org/presv/joint/day.html>.
- Day, Michael. 2001. Metadata for Digital Preservation: A review of Recent Developments. *ECDL*: 161-172.
- Hodge, G. 2001. *Metadata Made Simpler*. NISO Press.
http://www.niso.org/news/Metadata_simpler.pdf.
- Lopovici, C. and J. Masanes. 2000. *Metadata for the Long Term Preservation of Electric Publications*.
<http://www.kb.nl/coop/nedlib/results/preservationmetadata.pdf>.
- National Library of New Zealand. 2003. *Metadata Standards Framework - Metadata Implementation Schema*.
http://www.natlib.govt.nz/files/4initiatives__metaschema.pdf.
- OCLC/RLG Working Group on Preservation Metadata. 2001. *Preservation Metadata for Digital Objects: A Review of the State of the Art*.
<http://www.oclc.org/digitalpreservation/wgdeliver.htm>.
- OCLC/RLG Working Group on Preservation Metadata. 2002. *A Metadata Framework to Support the Preservation of Digital Objects*.
<http://www.oclc.org/research/pmng>.
- Phillips, M., D. Woodyard, K. Bradley, and C. Webb. 1999. *Preservation Metadata for Digital Collections*. Canberra: National Library of Australia.
- PREMIS Working Group. 2004. *Implementing Preservation Repositories for Digital Materials: Current Practice and Emerging Trends in the Cultural Heritage Community*.
<http://www.oclc.org/research/projects/pmwg/surveyreport.pdf>.
- PREMIS Working Group. 2005. *Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group*.
<http://www.oclc.org/research/projects/pmwg/premis-final.pdf>.
- Russell, K. al. 2000. *Metadata for Digital Preservation: CEDARS Project*.
<http://www.leeds.ac.uk/cedars/metadata.html>.
- Searle, Sam and David Thompson. 2003. Preservation Metadata: Pragmatic First Steps at the National Library of new Zealand. *D-Lib Magazine*, 9(4).