

온라인 연속간행자료 수집 및 보존에 관한 연구*

An Investigation on the Acquisition and Preservation of Online Serials

이 치 주(Chi-Ju Lee)**

목 차

1. 서론	3.2 2차 델파이 조사 결과
2. 온라인 연속간행자료의 수집 및 보존	3.3 3차 델파이 조사 결과
2.1 디지털자료의 보존	4. 온라인 연속간행자료의 수집 및 보존을 위한 모형
2.2 OAIS 참조모형	4.1 RSS 기반 수집 모형
2.3 온라인 연속간행자료 보존 시스템	4.2 온라인 연속간행자료 장기 보존
3. 온라인 연속간행자료 수집 및 보존	5. 결론
운영 요소에 관한 델파이 분석	
3.1 1차 델파이 조사 결과	

초 록

이 연구의 목적은 온라인으로만 발행되는 연속간행자료의 수집 및 보존에 대한 운영모형을 개발하기 위한 것이다. 공공기관, 단체, 연구기관의 전문가로 패널을 구성하여 델파이 조사를 실시하여 온라인 연속간행자료 보존에 따른 운영 모형 개발에 필요한 68개 세부항목을 도출하였으며, 적합성 및 시급성을 조사·분석하여 적합성 분석에 최종 선정된 57개 세부항목을 제시하였다. 온라인 연속간행자료의 수집을 위하여 RSS 기술을 이용한 콘텐츠의 자동수집방안을 제시하고, 온라인 연속간행자료에 맞는 수집 및 보존 프로세스를 설계하였다.

ABSTRACT

This study is to develop the operation models collection and preservation of serials only distributed online. In order to identify all of the requisite items for collection and preservation of the online serials, it conducts Delphi Technique researches with panels of experts from public institutions, organizations, and research institutions. It also deduces the 68 necessary detailed items to develop the operation models for preserving the online serials, and suggests 57 detailed items selected from the compatibility analysis with the investigation and examination of compatibility and urgency. It suggests the automated collection plan of contents with the application of RSS technology of transfer protocol to provide users the updated information from the sites where the contents and data have been constantly and frequently updated. This study strives to design the collection and preservation processes suitable for the online serials.

키워드: 온라인 연속간행자료, 메타데이터, OAIS 참조모형, 디지털자료 수집, 디지털자료 보존 프로세스, 보존 기술정보, 운영체제 정보, KORMARC, RSS
Online Serials, Metadata, OAIS Reference Model, Digital Material Collection, Preservation Process, Preservation Description Information, Operation System Information, KORMARC, RSS

* 이 논문은 연세대학교 대학원 박사학위논문을 축약한 것임.

** 국립중앙도서관 자료관리부장(lcj@nl.go.kr)

논문접수일자 2007년 5월 30일

게재확정일자 2007년 6월 19일

1. 서론

1.1 연구의 필요성 및 목적

인터넷의 대중화는 정보통신 네트워크를 기반으로 한 미디어의 결합현상을 동반하여 미디어와 미디어간의 고유경계선이 무너지고 개별 매체의 독립성이 약화되는 동시에, 서로 다른 매체의 콘텐츠와 서비스가 통합되는 결합현상을 가속화 시키고 있다. 이러한 결합 현상은 자료를 발행하는데 사용되던 인쇄매체, 카세트테이프 등과 같은 전통적인 매체 영역을 인터넷을 기반으로 하는 디지털 매체 영역으로 급속히 대체시켜 세계적으로 디지털자료의 발행과 유통이 급증하게 되었다.

과거 수십 년 간 국가의 지식 정보를 대표하는 출판물을 수집하고 보존하는 역할을 담당해 왔던 영국, 호주, 프랑스, 미국, 북유럽국가 등의 국립도서관에서는 새롭게 등장하여 전 세계적으로 대중화되고 있는 인터넷상의 디지털자료에 대한 수집 및 보존에 대한 필요성을 인식하고 이에 대한 연구들을 진행하여왔다. 이들 도서관 및 관련 기관에서 수행하는 연구들에는 인터넷 자원 혹은 전자출판물의 납본에 대한 연구와 콘텐츠 표현 기술과는 독립적인 매체 보존 방안, 디지털자료의 장기보존, 자료에 대한 다양한 접근방법 제공 방안에 대한 연구들이 수행되고 있다. 특히, 디지털 자료의 장기보존 방안에 대한 연구는 많은 국가에서 수행되고 있으며 관련 연구와 실험 프로젝트를 통하여 시스템을 구축·운영하고 있다.

국내에서 발행된 모든 국가 자료의 수집과 보존 업무를 총괄하고 있는 국립중앙도서관

은 국가적인 차원에서의 온라인 디지털자원 수집 및 보존 표준모델 연구 등과 함께 온라인 디지털자원 수집·보존 사업인 OASIS(Online Archiving & Searching Internet Sources) 프로젝트를 추진하고 있다.

온라인 연속간행자료의 증가에 비하여 인터넷 자원 혹은 전자출판물을 보존하기 위한 디지털자료 보존 프로젝트에서는 인터넷 뉴스 및 웹진 등과 같이 인터넷상에서 정기적으로 발행되는 온라인 연속간행자료를 보존 대상 자료로 고려하고 있지 않고 있는 실정이다. 이는 정기적인 발행주기를 갖고 발행되는 연속간행자료의 특징을 고려하여 각 기사간의 관계를 형성하고 자료를 수집·보존하는 것에 많은 어려움이 있어 아직 온라인 연속간행자료의 수집 및 보존 시스템을 운영하지 못하고 있기 때문이다. 그러나 종이자료로 생산되지 않고 온라인 연속간행자료로 발행되는 자료가 늘어나고 있고 웹상에서 살아지는 자료의 발생으로 문제점이 도출되고 있어, 점점 온라인 연속간행자료의 보존 작업을 지속·체계적으로 수행하기 위한 온라인 연속간행자료의 수집 및 보존에 관한 연구가 필요하게 되었다.

이 연구의 목적은 국가 도서관의 역할과 기능이라는 틀 안에서 온라인 연속간행자료의 수집 및 보존을 효율적이고 체계적으로 수행하기 위해 필요한 제반요소를 도출하여 정책에 반영하고, 온라인 연속간행자료 수집 및 장기보존을 위한 모형을 제시하였다.

1.2 연구의 방법

이 연구를 위하여 온라인 연속간행자료가 가

지고 있는 고유한 특징을 분석·정리하기 위하여 관련 연구 및 문헌 고찰을 실시하고, 디지털 자료를 아카이빙하고 있는 나라의 관련 홈페이지와 자료를 조사·분석하여 국내·외에서 수행되고 있는 디지털자료 보존 프로젝트 사례들도 살펴보았다.

연속간행자료의 수집 및 보존에 필요한 제반 요소를 도출하고자 공공기관, 단체, 연구기관의 전문가를 패널로 구성하여 델파이 기법에 의한 설문 조사를 실시하였다. 온라인 연속간행자료 수집 및 장기보존을 위한 모형을 제시하였다.

설문조사는 총 3회에 걸쳐 실시하였으며, 1차는 2006년 8월 29일부터 9월 2일까지, 2차는 동년 9월 11일부터 25일까지, 3차 설문은 동년 10월 9일부터 20일까지 실시하였다. 1차 설문은 개방형 설문지를 이용하고, 2차 설문지는 개방형 설문지에서 도출된 항목에 대한 추가 및 삭제에 대한 설문내용을 조사하였다. 3차 설문지는 2차 설문지에 도출된 항목에 대한 전문 분야별 응답 빈도를 제시하고 5점 척도로 점수를 주도록 하였다. 1차 설문 조사는 전자우편을 통한 설문지와 전문가의 종합 토론을 거쳤으며, 2차와 3차 설문조사는 전자우편으로 조사하였다.

설문 대상자인 전문가 집단은 1단계에서는 디지털 아카이빙 전문가 5명을 선정하였고, 2단계와 3단계에서는 1단계 대상자를 포함하여 공공기관 10명, 단체 11명, 연구기관 10명 등 총 31명을 대상으로 진행하였다. 기관은 디지털 아카이빙을 수행하고 있거나 가능성이 있는 공공기관의 전문가로 구성하고, 단체는 디지털 아카이빙과 관련 있는 시민단체 및 업체 전문가로, 연구기관은 문헌정보학과, 컴퓨터공학과 교수들로 구성하였다.

1차 델파이에서는 총 5명의 전문가를 선정하여 100% 응답을 받았으며, 2차 델파이에서는 총 38명의 전문가 중 최종 31명에게 답변을 받았다. 3차 델파이는 2차 설문에 답변을 득한 31명 모두에게 설문을 보내어 응답을 받았다.

3차 설문에서는 2차 설문조사 결과로 조사된 전문가의 응답 빈도를 분야별로 구분하여 제시함으로써, '적합성'과 '시급성'의 5점 척도 점수로 조사하였다. 3차 조사에서는 수정 보완된 평가문항에 대하여 내용의 적합성과 시급성을 5점 척도로 구성하였다.

2. 온라인 연속간행자료의 수집 및 보존

2.1 디지털자료의 보존

각 국의 디지털로 표현된 정보의 축적을 위한 현재까지 대부분의 디지털자료의 보존 프로젝트는 OAIS 참조 모형을 기반으로 하고 있다. 이 중 디지털자료 보존에 앞장서고 있는 외국의 대표적인 프로젝트 및 한국의 보존 프로젝트는 다음과 같다.

1) 호주 PANDORA 프로젝트

호주 국립도서관(National Library of Australia: NLA)에서는 호주의 디지털정보를 수집하고 보존하여 장기적인 접근을 제공하기 위한 정책과 절차를 개발하고 전자출판물의 국가장서 구축비용과 그 성공가능성을 시험·검토하기 위한 PANDORA(Preserving and Accessing Networked Documentary Resources in Au-

stralia) 프로젝트를 1996년에 착수하였다. NLA는 디지털자료의 보존 관리 요건을 기술하는 구조적인 방법을 개발함으로써 장기적인 접근성을 보장하고자 1999년에 디지털 컬렉션에 사용할 수 있는 보존 메타데이터 규격(Preservation Metadata for Digital Collections)을 제안하였다.

PANDORA 프로젝트의 자료 선정을 위하여 NLA는 SCOAP(Selection Committee On-line Australian Publications)를 설치하고 자료 선정 지침을 개발했다. 수집된 아카이브는 2001년부터 NLA가 자체 개발한 PANDAS(PANDORA Digital Archiving System)라는 아카이브관리시스템에 의해 관리되고 있다. PANDAS는 PANDORA에 아카이브 된 타이틀에 관한 메타데이터를 관리하고, 디지털자료 수집, 품질 관리, 수집된 자원에 대한 접근관리를 수행한다. 수집된 모든 웹 자원은 MARC에 의해 편목되며, 목록은 National Bibliographic Database와 PANDORA에 참여한 각 기관 목록에 통합된다(서혜란 2003).

2) 영국의 CEDARS 프로젝트

호주 국립도서관에서 보존 메타데이터 요소 집합이 개발된 직후, 영국의 CEDARS 프로젝트에서도 보존 메타데이터의 초안이 발행되었다. CEDARS(CURL Exemplars in Digital ARchiveS)프로젝트는 JISC(the Joint Information Systems Committee of the UK education funding councils)의 지원으로 수행된 디지털 도서관 프로그램의 일부로 1998년 4월에 실시되어 2002년 3월에 완료되었다. CEDARS 프로젝트는 대학 간의 전자정보의 공유와 보존

의 필요성으로 시작된 학술 공동체를 위한 디지털 아카이브 프로젝트이다.

CURL은 디지털 보존의 전략·방법론·실제 문제를 논의하는 3 분과로 구성하고, 에플리케이션을 통한 디지털 보존 전략과 기법으로 OAIS 참조모형에 기반을 둔 분산형 아키텍처를 채택하였다(Day 1999). 이 프로젝트는 디지털자료의 보존에 관련된 전략과 실제적인 방법론 연구에 주력하고 있으며, 디지털자료의 이용을 더욱 효율적으로 제공하기 위하여 도서관에서 디지털자료의 보존을 위한 지침을 마련하였다(박은경 2002).

3) 유럽 NEDLIB 프로젝트

유럽의 여러 국가들이 주최한 NEDLIB(Networked European Deposit Library) 프로젝트는 OAIS 모형에 기반을 둔 DSEP(Deposit System for Electronic Publications) 개발의 일환으로 보존 관리를 위하여 필요한 최소한의 메타데이터를 정의하는 메타데이터 안을 2000년에 발표하였다(Lupovich, Masanes 2002). NEDLIB 프로젝트는 네덜란드 국립도서관의 주관 하에 네덜란드 국립문서협회, 핀란드, 프랑스, 이탈리아, 노르웨이, 포르투갈, 스위스 국립도서관의 협력으로 수행되었으며, 그 밖에도 Elsevier, Kluwers, Springer Verlag 등의 출판사, IT(Information Technology) 업계, 국립기록보존소 등이 참여하였다. 프로젝트는 1998년부터 2000년까지 European Commission의 지원으로 진행되었다.

NEDLIB 프로젝트에서 제안한 보존 메타데이터 또한 OAIS 모형의 일부인 정보 모형(information model)을 기초로 개발되었다. 그러

나 CEDARS, PANDORA, MINERVA의 보존 메타데이터가 디지털 정보의 장기 보존과 접근이라는 두 가지 목적을 염두에 두고 개발된 것과 달리, NEDLIB 보존 메타데이터는 엄격하게 보존에만 초점을 맞추어 개발되었다. 그런 관계로 NEDLIB이 제안한 보존 메타데이터는 변화하는 기술 환경에서 방대한 양의 디지털 정보 관리를 위한 최소한의 핵심적인 메타데이터 요소 집합이라 할 수 있다(홍재현 2004).

4) 미국 MINERVA

미국 국회도서관(Library of Congress)은 본연의 주 임무인 자료의 가용성 및 유용성을 높이고, 다음 세대를 위하여 지식자원의 보존을 위하여 MINERVA(Mapping the INternet Electronic Resources Virtual Archive) 프로젝트를 시작하였다. MINERVA 프로젝트는 인터넷 자료의 수집 및 보존을 위한 시스템 개발을 목표로 수집한 인터넷 자료를 범용적인 접근을 위하여 아카이브하고 9/11 테러 및 2000년 미국 대선에 관련된 자료를 수집하였다. MINERVA 프로젝트에서는 시범적인 프로젝트를 수행하면서 포괄적 수집 방식과 정책에 의한 선정 수집방식을 채택하였다. 먼저 도서관 직원에 의한 선정적 수집 방식은 개발된 수집 및 보존 시스템을 시험하기 위하여 자료의 수집 절차 개발을 목적으로 수행하였으며, 나중에는 포괄적 수집 방식을 통하여 실제적인 인터넷 자료를 수집하였다(Ammen 2002).

MINERVA 프로젝트에서 2002년 7월 디지털객체의 서지정보 표준 메타데이터를 위하여 개발된 MODS는 다양한 메타데이터(ONIX, IMS, TEI 등)를 분석하고 온라인 자료의 서지

정보 메타데이터 표준으로 지향하고 있는 DC의 구조적으로 단순한 문제점을 해결하고 범용서지 메타데이터를 지향하기 위한 디지털도서관의 메타데이터 표준 환경 제공을 목표로 개발되었다.

5) 한국 OASIS 프로젝트

한국에서 발행된 모든 자료를 망라적으로 수집하고 국가서지 작성을 책임지고 있는 국립중앙도서관에서는 짧은 기간 동안 공개되었다가 사라져 버리는 디지털 지적 창작물들을 영구 보존하기 위하여 2001년 온라인 디지털자원 수집·보존 사업인 OASIS(Online Archiving & searching Internet Sources) 프로젝트에 착수하였다. 이를 통하여 국립중앙도서관에서는 가치 있는 온라인 디지털자원의 체계적인 수집·보존 정책을 연구하고 실행함으로써 국가대표 도서관으로서 역할을 강화하고자 하였다.

OASIS 프로젝트는 우선 온라인 디지털자원을 정의하고 디지털자원에 해당하는 자료를 수집하고자 계획하였다. 구체적인 디지털자원의 정의는 다음과 같다(OASIS 2006).

- 인터넷을 통해 전자화된 디지털자료의 내용을 내려 받아 개인용 및 휴대용 컴퓨터 등의 정보통신 단말기에서 볼 수 있는 디지털 버전의 출판물
- 인터넷 등을 통해 내려 받아 개인용 컴퓨터, 노트북 컴퓨터, PDA 등의 정보통신 단말기에서 읽을 수 있는 전자적으로 유통되는 텍스트, 소리, 동영상정보 등을 담고 있는 컴퓨터 파일
- 웹 사이트, 웹 페이지, 문서 파일(pdf, hwp, doc, txt 등), 이미지, 동영상, 음악, 압축파일

국립중앙도서관은 OASIS 프로젝트를 통하여 가치 있는 온라인 디지털자원의 체계적 수집, 보존을 위한 중요사항의 심의 및 자문 기구인 온라인디지털자원수집위원회를 구성하였다. 온라인디지털자원수집위원회는 인문과학, 사회과학, 자연과학, 정책자료 분과 총 4 분과를 구성하여 디지털자원에 대한 심도 있는 선정 작업을 수행하고 있다. 온라인디지털자원수집위원회의 활동을 살펴보면 첫째, 총 4개 분과에 해당하는 28개의 주제 분야의 도메인 50여개에 속한 URL을 추천하였다. 둘째, 28개 주제의 전문가 84명의 정보를 데이터베이스화하였다. 셋째, 주제별로 분과위원회를 구성하고 활동 활성화 방안을 마련하였으며, 마지막으로, 유동적 디지털자료의 특성상 가치 있는 온라인 디지털자원을 상시 추천할 수 있는 제도를 제안하였다(이혜원 2004).

6) 사례 비교 분석 결과

각 국가에서 개발된 5개의 온라인 자료 보존 시스템에 대한 비교 및 분석은 해외 여러 보존 메타데이터 요소의 근간이 된 OAIS 참조모형에서 제시한 기능 모형(Functional Model)을 중심으로 실시하였다. 각 국가에서 개발된 보존 메타데이터 요소 집합은 개발 기관의 목적 및 요구 등이 상이함에 따라 도출된 요소에도 차이가 발생하였다. 이 다섯 개의 보존 프로젝트가 보존 메타데이터에 관한 사실상의 세계적

표준으로 사용되고 있으므로 우리의 전자문서 관리시스템에서도 정보공유와 시스템 호환을 위해서 이를 보존 메타데이터의 국내 도입을 위한 기초 자료로 삼고자 하였다. 특히 이 연구에서는 OAIS에서 제안한 자료 수집 및 보존 방식을 기준으로 디지털자료의 수집과 보존을 위한 각 국가별 사례들을 중심으로 비교·분석하였다.

〈표 1〉은 일반적인 디지털자료의 보존에 필요한 항목을 비교 평가한 결과를 나타내고 있다. 우선 프로젝트의 진행상황을 보면 CEDARS와 NEDLIB은 2000년에 프로젝트가 종료되었다. CEDARS는 CAMiLioN¹⁾ 프로젝트와 연계되어 후속적인 프로젝트가 진행되었다. 이외의 PANDORA, MINERVA, OASIS은 현재까지 진행 중이다. 자료의 선정 방법을 보면 PANDORA, CEDARS, MINERVA, OASIS는 각각 SCOAP(Selection Committee on Online Australian Publications), DESIRE 정책, 국회도서관, 온라인디지털자원수집보존 위원회에 의한 선정을 통하여 수집하고 있다. 그러나 NEDLIB의 경우 디지털 납본제도를 통하여 디지털자료를 수집, 보존하고 있다. 수집 방식의 경우 다섯 프로젝트 모두 웹 로봇 또는 미러링 프로그램과 같은 웹 소프트웨어를 통하여 사이트의 자료를 수집하는 방법을 사용하고 있다.

메타데이터의 경우 PANDORA, CEDARS, NEDLIB은 자체적으로 개발한 메타데이터를

1) CAMiLEON 프로젝트는 Creative Archiving at Michigan & Leeds: Emulating the Old on the New의 약자로 디지털자료의 수집보다 장기보존에 필요한 기술적 방법을 평가 및 개발에 중점을 둔 프로젝트다. CAMiLEON 프로젝트는 영국에서 주최한 CEDARS 프로젝트의 후속격에 해당하는 프로젝트로 미국의 미시건(Michigan)과 영국의 리즈(Leeds) 대학에서 각각 NSF, JISC의 지원을 받은 연합 프로젝트다. 미래의 가상시스템에서 진본 디지털자료의 재 구현 방법으로 에뮬레이션(Emulation) 기법을 제안하였다. 에뮬레이션의 장점은 콘텐츠의 정보뿐 아니라 외관과 느낌을 그대로 보존, 재 구현이 가능하다는 점이다.

MINERVA와 OASIS는 표준화된 MODS, 더블린코어(이하 DC) 메타데이터를 사용하고 있다. 표준화된 메타데이터의 경우 MARC의 단점을 보완하는 형식으로 개발되어 부분적으로 MARC와 호환할 수 있는 방법(매핑 테이블)을 제공하고 있다. 그러나 자체개발한 메타데이터의 경우 CEDARS에서 개발한 메타데이터만이 부분적으로 MARC와 상호호환 할 수 있다. 또한 디지털자료의 메타데이터 추출은 모든 프로젝트에서 전문가 또는 사서에 의하여 수행하고 있었다.

디지털 단행자료의 변경내용 또는 추가된 자료의 수집을 기존에 수집한 자료에 반영할 수 있는 자료 수집주기의 경우 CEDARS 프로젝트만이 모든 자료에 대하여 정기적으로 검사를 하고 있다. PANDORA와 NEDLIB 프로젝트는 임의적으로 변경 검사를 하였고, MINERVA의 경우는 자료의 종류에 따라 수집 주기를 다양

하게 하였다. OASIS 프로젝트의 경우는 유일하게 변화비율 검사를 통하여 수집 여부를 결정하였다. 변화비율은 한 사이트에서 기존에 수집하여 저장된 디지털자료와 비교하여 추가, 수정, 삭제된 자료의 개수를 환산하여 수치화한 값을 의미한다.

자료의 접근점 제공을 의미하는 식별기호의 경우 PANDORA 프로젝트는 PURL을 현재 사용하고 있으나 메타데이터에서는 다양한 식별기호를 포함하도록 설계되어 있다. CEDARS의 경우 또한 마찬가지로 기본적으로 CRID(Cedars Reference ID)와 같은 고유 식별기호를 제공하고 있으며 ISBN 같은 기존에 사용되고 있던 식별기호 또한 지속적으로 사용할 수 있도록 고려되어있다. 이외의 NEDLIB, MINERVA, OASIS 프로젝트 모두 메타데이터에서 모든 종류의 식별기호를 사용할 수 있도록 지원하고 있다.

<표 1> 디지털자료 보존 사례에 대한 일반적인 사항의 비교 분석

항목	PANDORA	CEDARS	NEDLIB	MINERVA	OASIS
개발 국가	호주	영국	유럽연합	미국	한국
진행상황	1996~	1998.4~2000.4	1998.1~2000.12	2000~현재	2001~현재
자료수집 방법	웹 로봇	웹 로봇	웹 로봇 (NEDLIB Harvester)	미러링 프로그램	웹 로봇
수집자료 선정	SCOAP 지침에 따른 수집	DESIRE 선정원칙에 따른 수집	납본시스템	Library of Congress 선정	온라인디지털자원 수집보존 위원회
메타데이터	자체개발	자체개발	자체개발	MODS	Dublin Core
MARC와의 호환	지원안됨	부분적	지원안됨	부분적	부분적
자료 메타데이터화	사서	사서	사서	사서	사서
자료수집 주기	부정기	정기	부정기	자료의 종류별	변화비율
식별번호	PURL	CRID	제약없음	제약없음	제약없음
장기보존 정책	Emulation	Emulation	Emulation	Migration	없음

2.2 OAIS 참조모형

OAIS 참조 모형(Reference Model for an Open Archival Information System)은 장기 간에 걸쳐 디지털자료를 보존하고 접근방법 제공에 필요한 광범위한 이해를 돕기 위하여 제안된 모형으로서 선형적으로 고안해 내었다기 보다는 각 유관 분야에서 이제까지 쌓아 온 지식과 경험을 수렴하여 각계의 일치된 의견을 하나의 완성된 틀로 정리해 낸 것으로 볼 수 있다(김연정 2004).

OAIS 참조 모형은 NASA를 비롯한 세계 각국의 우주개발기관에서 조직한 우주데이터 시스템 자문위원회(Consultative Committee for Space Data System: CCSDS)에 의하여 1999년 5월에 Red Book, Issue 1(CCSDS 1999)으로 처음 제안되었다. 이후 2001년 6월에 개정판인 Red Book, Issue 2(CCSDS 2001)가 발표되었고, 2002년 1월에 발표된 참조 모형 Blue Book, Issue 1(CCSDS 2002)의 권고가 최종적으로 ISO 14721:2002로 채택되었다.

OAIS 참조 모형은 정보 보존과 특정 사용자 집단에 정보를 제공할 수 있도록 가공하는데 책임이 있는 기관으로 구성된 아카이브 모형을 의미한다. 이 모형에서 의미하는 정보 보존은 특정한 시대에서만 의미 있는 정보를 의미하는 것이 아닌 장기 보존을 위하여 보존된 정보에 영향을 미칠 수 있는 기술, 매체, 정보 형식, 사용자 집단의 변경에 관련된 필요한 정보의 관리를 대상으로 하고 있다. OAIS 보고서는 OAIS 참조모형의 기능을 다음과 같이 설명하고 있다(CCSDS 2002).

첫째, 디지털자료의 장기 보존 및 접근에 필

요한 이해 및 인식을 위한 프레임워크를 제공한다.

둘째, 보존 과정에서 비 아카이빙 기관이 효과적으로 참여할 수 있는 개념을 제공한다.

셋째, 현존 또는 미래의 아카이빙 시스템을 설명하고 비교할 수 있는 용어, 개념을 포함하는 프레임워크를 제공한다.

넷째, 여러 장기보존 전략 및 기술을 비교 설명할 수 있는 프레임워크를 제공한다.

다섯째, 보존된 디지털자료의 데이터 모형(data model)을 비교하고, 이러한 정보모형이 시간에 따라 어떻게 변화하는지 토론할 수 있는 기반을 제공한다.

여섯째, 디지털 형식이 아닌 정보의 장기보존을 위한 기능 확장 기반을 제공한다.

이러한 기능을 위하여 OAIS 참조모형은 정보의 수집, 보존, 서비스에 요구되는 사항을 정리하고 이에 관련된 그룹을 생산자(Producer), 관리자(Manager), 소비자(Consumer)로 나누고 각 그룹이 수행해야 할 기능을 제안하였다. 우선 생산자는 보존할 정보를 제공하는 사람 또는 클라이언트 시스템을 의미한다. 둘째, 관리자는 OAIS 참조모형에서 정책의 광범위한 영역을 하나의 요소로 묶는 역할을 수행한다.

OAIS 참조 모형은 유럽, 캐나다, 브라질, 미국, 일본, 러시아와 같은 여러 국가의 기록학, 도서관, 정보통신 등의 분야별 전문가들에 의하여 서로 다른 시각에서 디지털자료의 아카이빙에 관련한 합의를 도출한 것이다. 그리고 미래에 바뀔 수 있는 표준 및 모형에 관련한 사항을 OAIS 참조 모형에 반영할 수 있도록 하고 있다(이원영, 강진영 2005). 또한 이 참조 모형을 바탕으로 보존 시스템 간 인터페이스 표준,

권고 표준, 인가 표준, 저장 매체에 대한 표준과 보존 메타데이터에 관한 후속 연구가 이루어지고 있다(김연정 2004).

디지털 아카이빙 시스템의 표준 기능과 요건은 <그림 1>과 같이 디지털문헌의 입수, 저장, 관리, 운영, 보존, 접근에 이르는 작업순서와 관련된 보존 활동에 필요한 6가지의 기능 요소(functional components)를 말한다. <그림 1>에서 설명한 정보패키지(Information Package)는 이 6가지 기능 요소를 통하여 유통되는 3가지의 패키지 유형을 말한다.

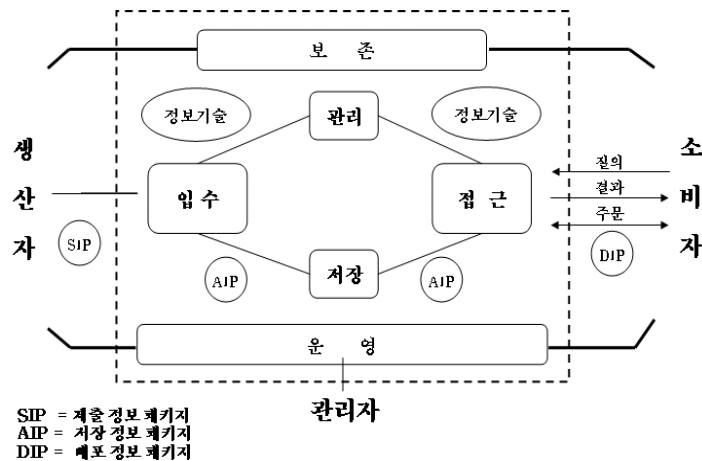
- SIP(Submission Information Package): 생산자가 정보 보존을 위해 OAIS에 제공하는 제출 정보 패키지
- AIP(Archival Information Package): OAIS에서 장기 보존을 위해 저장소에 저장되는 아카이브 정보 패키지
- DIP(Dissemination Information Package): 소비자에게 정보 서비스를 위한 배

포 정보 패키지

이러한 3가지 정보 패키지를 바탕으로 각각의 6가지 기능 요소의 구체적인 기능은 다음과 같다(CCSDS 2002).

1) 입수 단계

생산자가 제출한 자료를 수집하고 아카이빙에 포함시키기 위해 준비하는 일련의 작업 단계로서 구체적으로 생산자에 의해 OAIS로 제공된 자료를 수집한다. 즉 생산자와 운영자에게 받은 SIP 자료가 정상적인 데이터인지 검사하고 자료의 보존 시스템에서 저장 및 관리에 적합한 형태인 AIP 형식으로 변환한다. 이 AIP 형식 변환 과정 직후에 검색 및 정보 조작 도구에 필요한 자료에 대한 기술(descriptive) 메타데이터 정보를 생산 또는 추출하여 AIP와 기술 메타데이터 정보를 저장 단계로 전송하는 기능을 담당한다. 또한 기술 메타데이터 정보를 데



<그림 1> OAIS 참조 모형의 6가지 기능 요소

출처: CCSDS, 2002, Reference Model for an Open Archival Information System(OAIS).
CCSDS 650.0-B-1 BLUE BOOK: 4-1.

이더 관리 단계로 전송하는 업무를 담당한다.

2) 저장 단계

저장 단계는 AIP의 저장, 유지, 그리고 검색을 위한 서비스와 기능을 수행한다. 저장단계는 입수 단계로부터 전송받은 AIP를 장기 보존 저장소에 추가하고 운영 단계로부터 받은 저장소 관리 정책에 따라 저장 매체 관리, 에러 검사 및 로깅 업무를 수행한다. 또한 저장 매체와 백업 매체와의 연계를 통해서 장애 극복에 관여한다. 장기보존 정책의 일환인 마이그레이션을 위하여 지속적으로 AIP 가공, 유지 보수할 수 있는 공간을 마련하는 매체 교환 기능을 수행한다. 또한 AIP를 접근 단계에 전송하는 업무를 수행한다.

3) 관리 단계

관리 단계는 크게 입수 단계로부터 AIP를 전송받고, 소비자가 정보에 접근할 수 있는 제반 사항을 구축하는 업무를 담당한다. 구체적인 기능으로는 입수 단계와 운영 단계로부터 받은 자료 및 기술(descriptive) 정보에 관한 데이터베이스 업데이트 요청을 수행하며, 운영 단계가 명령한 정책(policy)에 따라 데이터베이스를 관리한다. 관리 단계는 또한 입수, 운영, 접근 단계에서 요청된 자료 목록, 사용 통계 등의 보고서 작성의 업무를 수행한다. 서비스 측면에서의 가장 중요한 기능으로는 데이터베이스를 관리하고 소비자가 요구한 질의(query)의 질의어를 처리하는 것이다.

4) 운영 단계

아카이브 시스템의 전반적인 운영을 위한 서

비스 기능을 수행한다. 운영 단계에 포함되는 가장 기초적인 업무는 생산자로부터 자료 제출 및 형식에 관한 협의를 통하여 동의를 구하고 감사(Auditing)를 통하여 자료 제출 유무를 파악한다. 또한 시스템 하드웨어와 소프트웨어 사양을 관리하는 것을 포함하여 아카이브 운영을 감독하고 개선하는 시스템 관리와 아카이브 소장내용의 보고(report), 업데이트, 배포요청 기능을 수행한다. 그러나 무엇보다도 운영 단계에서는 아카이브 시스템의 전반적인 표준과 정책을 수립하여 장기적으로 발전할 수 있는 정책을 관리한다.

5) 보존 단계

수집된 자료의 보존에 관련한 업무를 담당하는 보존 단계는 관찰(monitor)로부터 수집된 정보를 통하여 보존 계획을 수립하는 기능을 담당한다. 보존 계획 단계는 크게 4가지 기능을 수행한다. 첫째, 구성 집단 관찰기능은 조사를 통하여 생산자로부터 입수된 자료를 구현하기 위한 기술(technology) 정보와 소비자로부터 받은 서비스 요구사항을 수집하는 업무를 수행한다. 둘째, 기술 관찰기능은 디지털 기술, 정보 표준, 소프트웨어 및 하드웨어 플랫폼 등의 변화 정보를 수집하여 기술적 문제로 저장된 자료가 노후화되지 않도록 유지할 수 있는 기능을 수행한다. 셋째, 보존 정책 및 표준 개발 기능은 이전 두 기능으로부터 받은 소비자 요구 사항, 자료 생산 기술, 기술 변경 정보 등의 정보와 운영 단계에서 전송된 성능 정보, 소비자 평가 정보 등을 바탕으로 보존 정책 및 표준을 개발한다. 마지막으로 패키지 디자인 및 마이그레이션 계획 개발 기능은 운영 정책 및 지침

을 수행하기 위하여 새로운 정보 패키지 디자인, 구체적 마이그레이션 계획 업무를 수행한다. 또한 운영 단계에서 제정한 아카이브 표준, 메타데이터 표준, 문서 표준을 적용하고 AIP와 SIP 템플릿을 제공한다.

6) 접근 단계

접근 단계는 소비자가 자료를 검색하고 접근하는데 필요한 서비스 및 저장된 자료를 소비자가 이해할 수 있는 형식으로 전환하는 기능을 담당한다. 구체적으로 접근 단계는 소비자로부터 요청된 질의 및 보고서 등의 주문을 데이터 관리 단계에 전달하고 제공 받은 기술 정보와 저장 단계로부터 입수한 AIP를 소비자가 인식할 수 있는 자료로 변환하는 DIP 생성 기능을 수행한다. 또한 DIP와 과금 정보를 운영 단계에 전송하여 온/오프라인을 통해 발생하는 요금을 전송할 수 있게 한다.

2.3 온라인 연속간행자료 보존 시스템

웹을 기본으로 하는 매체들은 기존 인쇄물이 가지는 여러 지면, 시간, 독자참여 제한 등의 한계에서 벗어나 새로운 매체인 웹으로 융합되어 독창적인 새로운 분야로 발전하였다. 인쇄 매체의 단점을 극복한 웹 기반 매체들은 인터넷 이용의 폭발적인 증가와 함께 새로운 영역으로 자리 잡고 독자적인 매체로 발전하였고 극단적으로는 기존 인쇄 매체를 포기하고 웹만을 이용하는 매체의 등장으로까지 발전하였다. 즉, 일반적으로 단행자료의 성격을 갖는 온라인 디지털자료와 달리 정기 또는 비정기적으로 동일한 주제를 갖는 기사를 배포하는 온라인 연속

간행자료가 탄생하게 되었다.

Bovebschulte는 10년 내 온라인 연속간행자료가 연속간행자료 시장을 석권할 것으로 예측하였다(Lenares 1999). 그리고 2000년도 “Ulrich’s Database”에 수록된 온라인 연속간행자료는 약 20,430종에 달하며 “EBSCO Database”에 등재된 유료 잡지 5,398종 가운데 74%는 인쇄형과 온라인형의 동시 버전이며 나머지 17%는 온라인으로만 출판되었다(윤희운 2002). 이러한 온라인 연속간행자료는 양적인 팽창뿐 아니라 이용면에서도 뚜렷하게 증가하고 있는 추세이다. 또한, Lenares(1999)는 인쇄잡지와 전자잡지의 이용실태를 비교조사 한 결과, 인쇄잡지는 74%에서 65%로 감소한 반면에 전자잡지는 10%에서 15%로 증가하였다고 기술하고 있다.

국내에서는 1996년에 웹진의 창간으로 태동하여 1997~1998년에 웹진의 구체적인 활동을 보이기 시작하여 드디어 1999년에는 양적인 증가와 함께 다양화, 대중화되는 양상으로 전개되었다(윤세민 2002). 2006년 3월 현재 서비스되고 있는 인터넷 신문은 184~214 여개(184 : 네이버, 214 : 엠파스), 웹진은 1107~1027 여개(1107 : 네이버, 1027 : 엠파스) 정도로 추정되고 있으며 현재까지도 지속적인 증가와 함께 온라인 연속간행자료라는 새로운 중요 매체로 자리매김하고 있다.

디지털자료의 보존 필요성의 현실화와 함께 시작된 여러 프로젝트는 디지털자료라는 개념 하에 단행자료, 연속간행자료의 구분 없이 선정 지침에 따른 총체적 디지털자료의 수집 및 보존 업무를 수행하고 있다.

그러나 이러한 일률적인 수집, 보존 처리는 단행자료보다 복잡한 특성을 지니고 있고, 기

술적인 측면에서도 복잡한 양상을 띠고 있는 (한성국 2005) 연속간행자료를 메타데이터화 하고 보존할 수 있는 별도의 보존 프로세스가 필요하다. 또한 기존의 디지털자료의 보존 메타데이터 형식으로는 온라인연속간행자료에 대한 보존 메타데이터 형식을 명확하게 표현할 수 없으므로 온라인 연속간행자료의 보존에 관련된 메타데이터를 작성하기 위한 새로운 메타데이터 형식이 필요하다. 마지막으로 정기적인 기사(콘텐츠)의 배포와 함께 수집이 이루어져야 하는 온라인 연속간행자료의 제도적, 기술적 방안 마련이 필요하다.

3. 온라인 연속간행자료 수집 및 보존 운영 요소에 관한 델파이 분석

온라인 연속간행자료 수집 및 보존에 관한 운영요소를 도출하기 위하여 델파이 기법에 의하여 3회에 걸쳐 실시되었다. 2차까지의 설문조사로는 관련 사업들을 도출하였으며 3차 설문은 2차까지 실시한 설문조사에서 제기되었던 사항들을 종합하여 설문을 완성하고 각 사업별로 사업의 적합성과 시급성을 조사하였다. 설문조사를 시행한 이유는 공공기관 및 단체, 연구기관의 전문가 등이 연속간행자료 수집 및 보존 운영에 관련하여 어떤 인식을 하고 있는지를 조사하고 운영을 위한 제반요소를 도출하고자 하였다.

3.1 1차 델파이 조사 결과

5명의 전문가 집단을 대상으로 「온라인 연속

간행자료 보존 모형 연구」에 대하여 “거버넌스 체제 구축”과 “운영체제 구축”에 관하여 어떠한 과제들이 필요한지를 브레인스토밍의 방법으로 1차 설문조사를 하였다.

온라인 연속간행자료 보존 모형에 관하여 <표 2>와 같은 응답이 도출되었다.

3.2 2차 델파이 조사 결과

2차 델파이 조사는 1차 델파이 분석을 바탕으로 구조화된 세부 항목의 문항을 구성한 후 개방형 응답을 확보함으로써 이를 3차 델파이 조사의 구조화된 질문지를 작성하기 위하여 실시되었다. 2차 델파이 조사에서 개방형 응답으로 도출하여 문항별로 추가된 항목의 결과는 다음과 같다.

1) 온라인 연속간행자료 아카이빙의 법·제도 확립

1차 델파이 조사 설문에서 도출된 ‘온라인 연속간행물자료 아카이빙의 법·제도 확립’에 대하여 2차 델파이 조사에서 대상 전문가들로부터 다음 <표 3>과 같은 개방형 응답을 확보하였다.

2) 온라인 연속간행자료 아카이빙 장서 수집 정책 확립

1차 델파이 조사 설문에서 도출된 ‘온라인 연속간행자료 아카이빙 장서수집 정책 확립’에 대하여 2차 델파이 조사에서 대상 전문가들로부터 다음 <표 4>와 같은 개방형 응답을 확보하였다.

〈표 2〉 1차 델파이 조사 결과

구분	항목	설문 문항
제도 기반 구축	법·제도 확립	① 국가 온라인 연속간행자료 아카이빙 사업의 현황에 대하여 구체적인 조사와 파악이 이루어져야 한다. ② 온라인 연속간행자료 아카이빙 납본과 관련된 재·개정이 이루어져야 한다. ③ 온라인 연속간행자료 아카이빙 관련 저작권법 재·개정이 이루어져야 한다. ④ 온라인 연속간행자료 아카이빙 제도의 해외 모범사례에 대한 조사와 연구가 선행되어야 한다. ⑤ 온라인 연속간행자료 아카이빙에 관한 사회적 환경요인 조사·분석이 이루어져야 한다.
	장서수집 확립	① 온라인 연속간행자료 아카이빙 장서선별기준을 수립한다. ② 온라인 연속간행자료 아카이빙 분담수집정책을 수립한다. ③ 온라인 연속간행자료 수집 및 이용에 관한 표준협약서를 작성한다.
거버넌스 체제 구축	통합 기구 확립	① 온라인 연속간행자료 아카이빙과 관련하여 국가 통합기구를 설립한다. ② 지속적 온라인 연속간행자료 아카이빙을 위한 기금재단을 설립한다. ③ 온라인 연속간행자료 아카이빙을 위한 비용모형을 개발한다. ④ 지역 온라인 연속간행자료 아카이브 평가 인증제를 개발한다.
	운영 방안 구축	① 온라인 연속간행자료 아카이빙에 관한 전문 인력을 양성한다. ② 온라인 연속간행자료 보존 전문가들을 위한 정보공유체제를 구축한다. ③ 온라인 연속간행자료 아카이빙에 대한 대국민 홍보 전략을 개발한다.
	이해관계자의 협력 네트워크 구축	① 온라인 연속간행자료 이해관계자 유형을 개발하고 이해관계 내용을 조사한다. ② 온라인 연속간행자료 공공조직의 유형을 개발하고 이해관계 내용을 조사한다. ③ 온라인 연속간행자료 이해관계자 협력 네트워크 조직을 구성한다. ④ 온라인 연속간행자료 이해관계자 협력 역할분담 모형을 개발한다. ⑤ 온라인 연속간행자료 이해관계자의 체계적인 교육훈련 프로그램 개발한다.
	권역별 협력 네트워크 구축	① 온라인 연속간행자료 권역별 아카이브 역할모형을 연구 개발한다. ② 온라인 연속간행자료 권역별 아카이브 시스템의 표준모형을 개발한다. ③ 온라인 연속간행자료 권역별 아카이브 협력 네트워크를 구축한다.
전략 수립	전략 수립	① 온라인 연속간행자료 표준 보존포맷을 연구·개발한다. ② 온라인 연속간행자료 포맷 레지스트리를 연구·개발한다. ③ 온라인 연속간행자료 보존 메타데이터의 표준을 정립한다. ④ 온라인 연속간행자료 장기보존전략(마이그레이션, 에물레이션, 캡슐화 보존방식, 가상기계 프로그램)에 대한 계획을 수립한다. ⑤ 온라인 연속간행자료 아카이빙 관련 국제표준화 활동 및 국제협력 전략을 수립한다.
운영체제 구축	테스트베드 운영	① 온라인 연속간행자료 아카이빙 대상자원의 유형을 개발한다. ② 온라인 연속간행자료 아카이빙 대상자원의 패키징 모델링을 연구·개발한다. ③ 유형별 온라인 연속간행자료 아카이빙 프로세스를 정립한다. ④ 온라인 연속간행자료 아카이빙 관리 메타데이터 표준을 정립한다. ⑤ 온라인 연속간행자료 아카이빙 시스템 관리지침을 개발한다. ⑥ 온라인 연속간행자료 보존시스템 요소기술을 개발한다. ⑦ 온라인 연속간행자료 보존시스템 아키텍처를 연구한다.
	보존시스템 운영	① 온라인 연속간행자료 납본시스템을 개발한다. ② 온라인 연속간행자료 자동수집기(Harvester)를 개발한다. ③ 온라인 연속간행자료 보존자원 관리시스템을 개발한다. ④ 온라인 연속간행자료 보존자원 검색시스템을 개발한다. ⑤ 온라인 연속간행자료 보존자원 접근통제시스템을 개발한다. ⑥ 온라인 연속간행자료 보존자원 통합검색시스템을 개발한다. ⑦ 온라인 연속간행자료 보존상태 평가시스템을 개발한다. ⑧ 온라인 연속간행자료 보존시스템의 연계기능을 개발한다. ⑨ 온라인 연속간행자료 보존자원의 백업시스템을 개발한다.

〈표 3〉 '온라인 연속간행자료 아카이빙의 법·제도 확립'의 개방형 응답

- ⑥ 온라인 연속간행 아카이빙 자료의 활용방안을 조사·분석하여 법·제도에 필요한 항목을 개발한다.

〈표 4〉 '온라인 연속간행자료 아카이빙 장서 수집 정책 확립'의 개방형 응답

- ④ 온라인 연속간행자료 아카이빙 장서수집에 관한 예산 정책을 수립한다.
- ⑤ 아카이빙 대상 온라인 연속간행자료에 대한 지적 재산권을 소유하고 있는 출판사들과 공감대를 형성하여 수집기반을 마련한다.
- ⑥ 아카이빙 된 온라인 연속간행자료의 개방 및 접근 라이선스 도입방안을 추진한다.

<p>3) 온라인 연속간행자료 아카이빙 통합 기구 확립</p> <p>1차 델파이 조사 설문에서 도출된 '온라인 연속간행자료 아카이빙 통합 기구 확립'에 대하여 2차 델파이 조사에서 대상 전문가들로부터 다음 〈표 5〉와 같은 개방형 응답을 확보하였다.</p>	<p>행자료 아카이빙 운영 방안 구축'에 대하여 2차 델파이 조사에서 대상 전문가들로부터 다음 〈표 6〉과 같은 개방형 응답을 확보하였다.</p>
<p>4) 온라인 연속간행자료 아카이빙 운영 방안 구축</p> <p>1차 델파이 조사에서 도출된 '온라인 연속간</p>	<p>5) 온라인 연속간행자료 아카이빙에 이해관계자의 협력 네트워크 구축</p> <p>1차 델파이 조사 설문에서 도출된 '온라인 연속간행자료 아카이빙에 이해관계자의 협력 네트워크 구축'에 대하여 2차 델파이 조사에서 대상 전문가들로부터 추가된 개방형 응답은 없었</p>

〈표 5〉 '온라인 연속간행자료 아카이빙 통합 기구 확립'의 개방형 응답

- ⑤ 통합기구의 역할과 통합 모형을 개발한다.
- ⑥ 국가차원의 중재와 조정을 수행할 수 있는 협의기구를 설치한다.
- ⑦ 민간영역의 아카이빙에 대한 재정적 지원을 수행한다.

〈표 6〉 '온라인 연속간행자료 아카이빙 운영방안 구축'의 개방형 응답

- ④ 온라인 연속간행자료를 발행하는 기관에 대한 홍보를 수행한다.
- ⑤ 온라인 연속간행자료의 아카이빙 기술을 개발하고 보존과 활용에 대한 장기적인 전략을 수립한다.
- ⑥ Open-Archive(이용료 무료) 또는 이용료 지불에 대한 전략을 개발한다.
- ⑦ 기관 간 역할 분담 및 협력체계를 구축한다.
- ⑧ 온라인 연속간행자료 아카이빙 관련 일련의 국가표준을 수립한다.
- ⑨ 온라인 연속간행자료 아카이빙에 대한 경제적 타당성 분석을 실시한다.
- ⑩ UCI를 이용하여 온라인디지털콘텐츠에 대한 식별체계를 구축한다.
- ⑪ 문헌정보학 분야 종사자, 아카이빙 대상 저작물 관리자, 이용자들 간의 협의체 구축 등 아카이빙의 지속적인 협력 모델을 개발한다.

다. 따라서 1차 델파이 조사에서 도출된 항목을 그대로 3차 델파이 조사 항목으로 결정하였다.

6) 온라인 연속간행자료 아카이빙에 관한 권역별 협력 네트워크 구축

1차 델파이 조사 설문에서 도출된 '온라인 연속간행자료 아카이빙에 관한 권역별 협력 네트워크 구축'에 대하여 2차 델파이 조사에서 대상 전문가들로부터 다음 <표 7>과 같은 개방형 응답을 확보하였다.

7) 온라인 연속간행자료 보존 전략 수립

1차 델파이 조사 설문에서 도출된 '온라인 연속간행자료 보존 전략 수립'에 대하여 2차 델파이 조사에서 대상 전문가들로부터 다음 <표

8>과 같은 개방형 응답을 확보하였다.

8) 온라인 연속간행자료 아카이빙에 관한 테스트 베드 운영

1차 델파이 조사 설문에서 도출된 '온라인 연속간행자료 아카이빙에 관한 테스트 베드 운영'에 대하여 2차 델파이 조사에서 대상 전문가들로부터 다음 <표 9>와 같은 개방형 응답을 확보하였다.

9) 온라인 연속간행자료 보존 시스템 운영

1차 델파이 조사 설문에서 도출된 '온라인 연속간행자료 보존시스템 운영'에 대하여 2차 델파이 조사 대상 전문가들로부터 다음 <표 10>과 같은 개방형 응답을 확보하였다.

<표 7> '온라인 연속간행자료 아카이빙에 관한 권역별 협력 네트워크 구축'의 개방형 응답

- ④ 온라인 연속간행자료 권역별 아카이브의 규모·예산·인력에 대한 기준을 개발한다.
- ⑤ 온라인 연속간행자료의 유형별, 또는 주제별로 아카이빙 협력 네트워크를 구축한다.

<표 8> '온라인 연속간행자료 보존 전략 수립' 개방형 응답

- ⑥ 온라인 연속간행자료 보존 전문가의 역할을 정립한다.

<표 9> '온라인 연속간행자료 아카이빙에 관한 테스트 베드 운영' 개방형 응답

- ⑧ 온라인 연속간행자료 보존 시스템 비용 모델을 개발한다.
- ⑨ 온라인 연속간행자료에 대한 표준식별 체계 적용방안을 수립한다.

<표 10> '온라인 연속간행자료 보존 시스템 운영' 개방형 응답

- ⑩ 온라인 연속간행자료 자동 메타데이터 생성 시스템을 개발한다.
- ⑪ 온라인 연속간행자료 보안 정책시스템을 개발한다.
- ⑫ 온라인 연속간행자료 보존 자원 권한 관리 기능을 개발한다.
- ⑬ 온라인 연속간행자료에 새롭게 게재되는 논문 수집을 위하여, 업데이트가 빈번하게 발생하는 사이트의 정보 수집을 위한 RSS(Really Simple Syndication) Protocol을 개발한다.

3.3 3차 델파이 조사 결과

3차 델파이 조사에서는 각 세부항목에 대하여 적합성과 시급성을 평가하였다.

〈표 11〉에서 보듯이 3차 델파이 총 문항은 68문항이었고, 문항당 5점 척도로 평균값을 산출하였으므로 3.5이하는 적합성이 낮은 것으로 판단하고 11문항을 제외하였다. 따라서 최종

선정된 문항은 57개 문항이다.

또한 적합성 분석에서 최종 선정된 57개 항목을 기준으로 적합성 점수가 4.0점 이상으로 나온 항목을 기준으로 시급성 점수가 높은 순서를 나열하면 〈표 12〉와 같다.

델파이 분석에서 통해 도출된 온라인 연속간행자료의 수집 및 보존을 위한 운영요소는 다음과 같이 네 가지로 요약될 수 있다.

〈표 11〉 3차 델파이 조사 분석 - 문항적합성 평가결과

항 목	제고 문항	타당한 문항	최종 문항 수
법·제도 확립(6문항)	적합(3.5)이하: 없음	1,2,3,4,5,6	6
장서수집 확립(6문항)	적합(3.5)이하: 없음	1,2,3,4,5,6	6
통합기구 확립(7문항)	적합(3.5)이하: 문항2,4,5,7	1,3,6	3
운영방안 구축(11문항)	적합(3.5)이하: 문항3	1,2,4,5,6,7,8,9,10,11	10
이해관계자의 협력 네트워크 구축(5문항)	적합(3.5)이하: 문항5	1,2,3,4	4
권역별 협력 네트워크 구축(5문항)	적합(3.5)이하: 문항4	1,2,3,5	4
전략 수립(6문항)	적합(3.5)이하: 없음	1,2,3,4,5,6	6
테스트베드 운영(9문항)	적합(3.5)이하: 없음	1,2,3,4,5,6,7,8,9	9
보존시스템 운영(13문항)	적합(3.5)이하: 문항5,7,11,12	1,2,3,4,6,8,9,10,13	9
68	11	57	57

〈표 12〉 설문결과 우선순위

설문문항	적합성	시급성	순위
온라인 연속간행자료의 아카이빙 기술을 개발하고 보존과 활용에 대한 장기적인 전략을 수립한다.	4.42	4.26	1
온라인 연속간행자료 보존 메타데이터의 표준을 정립한다.	4.39	4.25	2
국가 온라인 연속간행자료 아카이빙 사업의 현황에 대하여 구체적인 조사와 파악이 이루어져야 한다.	4.33	4.23	3
온라인 연속간행자료 아카이빙 관리 메타데이터 표준을 정립한다.	4.41	4.21	4
온라인 연속간행자료 아카이빙 장서선별기준을 수립한다.	4.27	4.17	5
온라인 연속간행자료 보존 전문가들을 위한 정보공유체제를 구축한다.	4.2	4.1	6
온라인 연속간행자료 아카이빙 제도의 해외 모범사례에 대한 조사와 연구가 선행되어야 한다.	4.23	4.07	7
온라인 연속간행자료 표준 보존포맷을 연구, 개발한다.	4.29	4.07	8
유형별 온라인 연속간행자료 아카이빙 프로세스를 정립한다.	4.31	4.07	9
온라인 연속간행자료 아카이빙 관련 일련의 국가표준을 수립한다.	4.29	4.06	10
기관 간 역할 분담 및 협력체제를 구축한다.	4.1	4.03	11
온라인 연속간행자료 아카이빙 납본과 관련된 법의 제·개정이 이루어져야 한다.	4.23	4	12
온라인 연속간행자료 아카이빙에 관한 전문 인력을 양성한다.	4.1	4	13
온라인 연속간행자료 이해관계자 협력 네트워크 조직을 구성한다.	4.17	3.93	14

온라인 연속간행자료 장기보존전략(마이그레이션, 에물레이션, 캡슐화 보존방식, 가상 기계프로그렘)에 대한 계획을 수립한다.	4.24	3.93	15
온라인 연속간행자료 아카이빙 대상자원의 유형을 개발한다.	4.13	3.93	16
온라인 연속간행자료 보존자원 관리시스템을 개발한다.	4.2	3.87	17
온라인 연속간행자료 아카이빙 관련 저작권법 제·개정이 이루어져야 한다.	4.03	3.83	18
온라인 연속간행자료 납본시스템을 개발한다.	4.21	3.83	19
온라인 연속간행자료 아카이빙 분담수집정책을 수립한다.	4	3.73	20
온라인 연속간행자료 보존시스템 요소기술을 개발한다.	4.07	3.72	21
아카이빙 대상 온라인 연속간행자료에 대한 지적 재산권을 소유하고 있는 출판사들과 공감대를 형성하여 수집기반을 마련한다.	4.06	3.71	22

첫째, 제도 기반 구축은 '법·제도 확립'과 '장서수집 확립' 등의 제도 기반 구축영역으로 12문항이 조사되었으며, 문항적합성을 평가한 결과 모든 문항이 적합한 것으로 나타났다. 특히 '아카이빙 사업의 조사와 파악'은 4.33, '아카이빙 납본 관련 법 제·개정'이 4.23, '장서선별기준 수립'이 4.27 등으로 적합성이 높게 나타났다. 또한 시급성 부문도 '아카이빙 사업의 조사와 파악'은 4.23, '장서선별기준 수립'이 4.17로 높게 조사 되었다.

둘째, 거버넌스 체제 구축은 4개의 영역으로 조사되었다. 즉 '통합기구 확립', '운영방안 구축', '이해관계자의 협력 네트워크 구축', '권역별 협력 네트워크 구축' 등이 주요 영역으로 도출되었으며, 총 28문항으로 조사되었다. 문항적합성을 평가한 결과 4개 문항이 평점 3.5이하로 나타나 제외시켰으며, 최종 24개 문항을 선정하였다. 4개의 영역에서 '운영방안 구축'이 적합성과 시급성 모두 우위를 차지하였다. 특히 '아카이빙 장기전략 수립'이 적합성 4.42, 시급성 4.26으로 전체 항목 중 1위로 조사되었다.

셋째, 전략 수립은 1개의 영역으로 총 6문항으로 구성되었다. 문항적합성을 평가한 결과 6항목 모두 적합한 것으로 분석되었으며, 적합성과 시급성 평점이 높은 것으로 조사되었다. 특

히 '보존 메타데이터 표준 정립'이 적합성 4.39, 시급성 4.25로 가장 높게 나타났다.

넷째, 운영체제 구축은 2개의 영역으로 조사되었다. 즉 '테스트 베드 운영', '보존시스템 운영' 등이 도출되었으며 총 22문항이 조사되었다. 문항적합성을 평가한 결과 4개 문항이 평점 3.5이하로 나타나 제외시켰으며, 최종 18개 문항을 선정하였다. 온라인 연속간행자료의 수집 및 보존사업이 초기 단계여서 이번 조사에서는 '테스트 베드 운영'이 '보존시스템 운영'보다 적합성과 시급성 모두 우위로 나타났다. 특히 '아카이빙 관리 메타데이터 표준 정립'이 적합성 4.41, 시급성 4.21로 가장 높게 나타났으며, '아카이빙 프로세스 정립'도 적합성 4.31, 시급성 4.07로 높게 조사되었다.

4. 온라인 연속간행자료의 수집 및 보존을 위한 모형

4.1 RSS 기반 수집 모형

- 1) 디지털자료의 수집 정책의 문제점
오래전부터 각 국의 국립도서관에서는 법률

적으로 정한 납본관련 법률에 따라 그 나라에서 출판 또는 발행된 자료를 수집하고 있다. 1990년 후반부터 시작된 디지털자료의 보존에 대한 관심은 정보의 제작 및 배포의 단순성에 기인한 정보의 폭발적인 증가에 대응하여 보존 가치가 있는 디지털자료의 수집 및 보존을 연구하기 시작했다. 기존의 여러 인터넷을 통한 디지털자료 수집 프로젝트에서 채택한 방법은 크게 다음과 같이 웹 로봇을 통한 능동적 수집 정책과 제도적 납본을 통한 수동적 수집 정책으로 구분된다.

제도에 의한 강제력이 있는 납본은 웹 로봇을 이용한 수집 방법과는 달리 자국의 문화유산의 누락 위험을 배제할 수 있는 가장 좋은 정책으로 인정되고 있다. 그러나 제도적 납본을 통한 능동적 수집 정책 또한 배포기관에 따른 디지털자료 납본 여부를 감시해야하는 감시기관 설립이 필요하여 수집 시스템에 추가적인 작업단계가 필요하다는 단점이 있다. 이러한 기존 디지털자료의 수집 정책을 그대로 온라인 연속간행자료 수집에 이용할 경우 전문가에 의하여 수집하여야 하는 대상자료 인지를 선별하는 처리 시간과 추가 인력을 요구하는 문제점이 발생한다.

우선 웹 로봇을 이용한 능동적 수집 정책을 적용할 경우 또한 다음과 같은 문제점이 발생한다. 첫째, 능동적 수집정책은 자료를 개별 단위로 인정하는 단행자료에는 적합하나 자료와 자료사이의 연관 관계를 유지해야 하는 연속간행자료의 경우에는 연속간행자료가 가지고 있는 관계를 유지시키기 어렵다. 둘째, 한 웹 사이트에서 여러 종류의 연속간행자료를 배포 하는 경우 수집 후 사서 또는 전문가에 의해 연속간행자료 별로 나누어야 하는 후처리(post-processing)

과정을 거쳐야 하는 단점이 있다. 제도적 납본을 통한 수동적 수집 정책을 적용할 경우 또한 다음과 같은 문제점이 발생한다. 첫째, 온라인상의 정보는 인쇄매체와 달리 출판되는 것이 아니라 직접 자료를 공개하는 형식이 따른다. 즉, 콘텐츠의 제작 및 배포가 극도로 간편화되어 배포자의 수는 대폭 늘어나고 있는 추세이다. 이와 더불어 사서가 인터넷의 정보자원을 전문적으로 목록하거나 품질을 평가하는 것은 비효율적이다(조윤희 2003). 따라서 납본 의무자와 자료가 대폭 늘어나 막대한 시간 및 인력이 소요될 것으로 예상된다. 둘째, 신속성과 최신성이 보장되어야 하는 온라인 연속간행자료가 별도의 납본과정을 거쳐야 하므로 콘텐츠 배포와 동시에 수집이 이루어지기 힘들다. 셋째, 현재 공개되고 있는 콘텐츠의 수집이 아닌 수집기관에서 제공하는 별도의 방법을 통해 납본되므로 진본성에 대한 검사가 요구된다는 단점이 있다. 마지막으로 비전문가인 배포자가 납본을 위해 입력한 디지털자료에 대한 정보는 다시 한번 전문가에 의해 검열이 필요하다는 단점이 있다.

2) RSS 기반 온라인 연속간행자료의 수집 프로세스

RSS(Really Simple Syndication)는 뉴스나 블로그와 같이 데이터 또는 콘텐츠의 업데이트가 빈번하게 발생하는 사이트들이 업데이트된 정보를 사용자에게 제공하기 위하여 만들어진 전송규약(protocol)이다.

이 연구에서는 RSS 기술의 장점을 최대한 이용하여 온라인 연속간행자료 자동 수집 모형에 적용하고자 한다. RSS 기술을 이용한 콘텐츠 수집 프로세스는 다음과 같다.

첫째, 자료 보존 선정지침에 따라 보존 가치가 있다고 결정된 온라인 연속간행자료는 RSS 도입 정책에 따른 의무규정에 따라 배포자가 제공하는 RSS 채널을 RSS 리더 역할을 수행하는 RSS Feed Aggregator(RSS 수집기)에 등록한다. 이때 RSS 채널은 사용자가 원하는 정보만을 선택적으로 구독할 수 있도록 다양한 주제, 콘텐츠 특성에 맞게 분류하고 채널 등록 시 주제 정보를 같이 입력하여 수집기관에서 이용할 수 있게 한다.

둘째, RSS 수집기는 콘텐츠 배포자가 업데이트 시 RSS 채널을 통하여 콘텐츠 요약정보와 함께 콘텐츠 업데이트 유무를 파악한다.

셋째, RSS 수집기는 RSS 채널을 통하여 확보한 RSS 문서 파일을 해독하고 링크된 콘텐츠의 URL을 웹 로봇에 통보한다.

넷째, 웹 로봇은 RSS 수집기로 부터 얻은 URL에 해당하는 HTML 파일을 비롯하여 콘텐츠 재구현 소프트웨어, 서버 정보, IP Address 등 장기보존에 필요한 모든 정보를 확보한다.

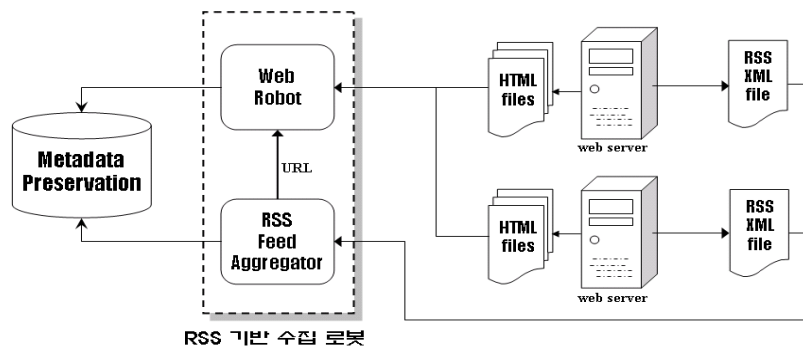
다섯째, RSS 수집기가 확보한 RSS 문서 파일과 웹 로봇이 수집한 콘텐츠를 보존 메타데이터 형식에 맞춰 메타데이터화 한다.

〈그림 2〉는 온라인 연속간행자료 콘텐츠 자동수집 모형을 개념화 한 그림이다.

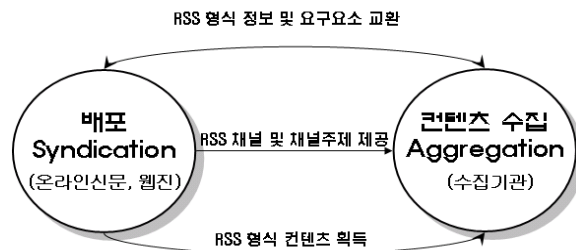
3) RSS 기반 온라인 연속간행자료 수집 모형
이 연구에서 제안하고자 하는 RSS 기반의 온라인 연속간행자료의 콘텐츠 자동 수집을 위해서 콘텐츠 배포자와 수집기관과의 기술적으로 선행되어야 할 작업은 〈그림 3〉과 같이 3단계로 구분된다.

첫 번째 단계는 RSS 형식의 합의 단계로 콘텐츠 배포자가 RSS 파일 형식에 들어가야 할 요소(Elements)를 결정하고 추가적으로 수집기관에서 요구하는 정보를 추가하여 최종적으로 RSS 메타데이터 형식을 결정하는 과정이다.

콘텐츠 수집기관에서는 RSS 제공 사이트마다 RSS 버전이 다양하고 심지어 한 사이트에서 여러 버전을 제공하는 경우가 있으므로 모든 버전에 따른 특징 정보를 알고 있어야 하며 버전에 따른 요소가 어떠한 의미인지 파악하고 있어야 한다. 이를 위하여 콘텐츠 수집 기관에서는 RSS 버전에 따른 필수 요소를 정하고 이를 배포자와 협의를 통하여 콘텐츠 자동 수집에 필요한 정보를 제공받을 수 있어야 한다.



〈그림 2〉 온라인 연속간행자료 콘텐츠 자동수집 모형



〈그림 3〉 콘텐츠 자동 수집 과정

두 번째 단계는 RSS 채널 및 채널 주제 등 채널 정보 파악 및 이용 단계이다. 일반적으로 RSS 서비스 제공 사이트에서는 RSS 채널을 다양화 하여 다양한 주제, 이슈로 분류하고 있다. 이미 위에서 예로 들었던 “오마이뉴스(<http://www.ohmynews.com>)” 사이트는 자료별, 주제별, 언어별 등 여러 분류기준으로 여러 채널을 제공하고 있다. 일반적으로 한 사이트에서 여러 온라인 연속간행자료 배포 시 이를 채널로 분류할 수 있을 뿐 아니라 콘텐츠의 주제, 자료, 장르, 언어에 따라 분류할 수 있어 자체적인 필터링 작업을 대체할 수 있다.

마지막 단계는 RSS 형식의 콘텐츠 획득 및 수집, 보존 과정이다. 최종 단계는 RSS채널을 통하여 RSS 문서파일의 콘텐츠를 분류하고 메타데이터 형식에 따라 정보를 재구성하는 일련의 작업을 수행한다. 또한 RSS 문서파일에서 제공하는 링크에 접속하여 웹사이트 콘텐츠를 수집 보존한다. 〈그림 3〉에서와 같이 온라인 연속간행자료의 배포자와 수집기관 간 협의된 RSS 채널을 RSS Feed Aggregator(RSS 수집기) 모듈에 등록한다. RSS 수집기는 일반사용자가 사용하는 RSS 리더와 동일한 기능인 RSS를 통한 콘텐츠 업데이트 유무 파악 및 채널 다양화에 따른 중복 콘텐츠의 유무를 검사하

는 기능을 한다. 또한 RSS 링크(콘텐츠 URL)를 추출하여 기존의 인터넷 자료 보존 시스템에서 사용한 동일 웹 로봇에 제공하여 해당하는 콘텐츠를 수집하는 기능을 수행한다.

4.2 온라인 연속간행자료의 장기 보존

국가의 가치 있는 지식과 정보를 다음 세대에 전승하기 위해서는 무엇보다도 체계적이고 과학적인 수집 및 보존 정책이 필요하다. 온라인 연속간행자료는 점차적으로 서지가 완성되는 특징으로 기술의 발전에 따라 서지의 새로운 규칙과 기준들이 기존의 것들에 추가됨으로써 단행자료보다 복잡한 특성을 지니고 있고, 특히 기술적인 측면에서도 복잡한 양상(한성국 2005)이 심화됨에 따라 기존의 보존 프로젝트에서 제시하는 보존 모형을 온라인 연속간행자료에 적용 시 많은 문제점을 야기 시킬 것이다.

이러한 현상은 전 세계적으로 디지털자료에 대한 의무적인 납본제도가 정착되지 못한데 기인할 뿐 아니라 온라인 연속간행자료의 정의와 이에 맞는 적합한 수집 모형의 부재에도 문제가 있다. 따라서 현재 OASIS가 담당하고 있는 디지털자료를 온라인 단행자료와 온라인 연속간행자료로 구분하고, 자료의 종류에 따라 최

적의 수집모형 개발이 필요하며 최종적으로는 국가도서관에서 국가 디지털자료의 통합적 보존 모형이 필요하다.

또한 전문가들의 델파이 조사 분석에서 나타난 것과 같이 온라인 연속간행자료 장기보존을 위하여 장서선별 기준을 마련하고, 전문 인력 양성, 보존전문가들을 위한 정보공유체제 구축, 국가 표준 보존 포맷 연구, 보존 아키텍처 연구 등이 필요한 것으로 조사되었다.

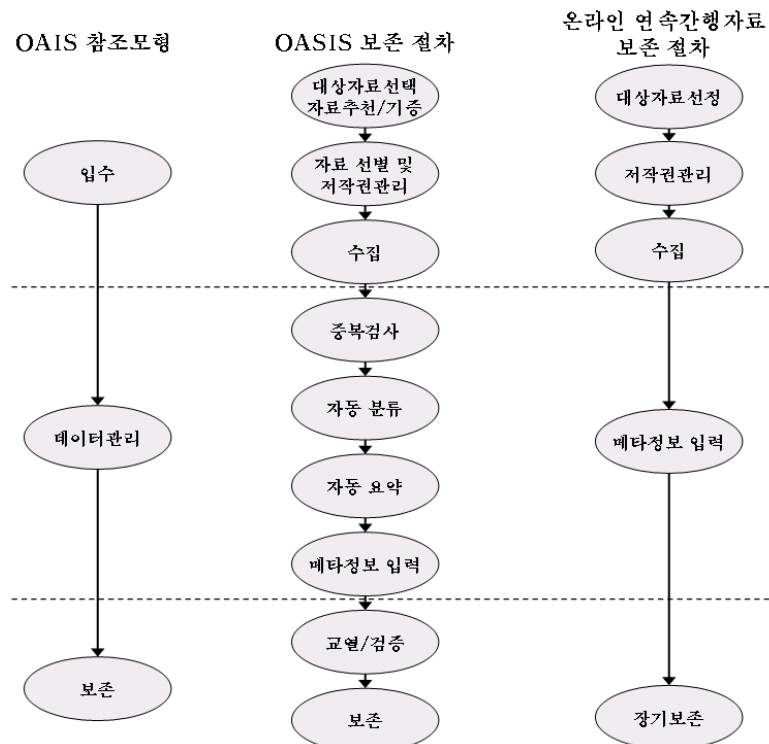
4.2.1 온라인 연속간행자료 보존 프로세스 모형

온라인 연속간행자료의 수집 및 보존은 그 특성상 웹 로봇(web robot)을 사용하여 간단하게 자료를 수집할 수 있는 자료가 아니라 주

기적으로 관리할 필요가 있다. 이런 주기적인 관리를 토대로 확보한 자료를 메타데이터의 작성을 통하여 자료와 자료 간의 연관성을 생성한다. OASIS 참조 모형의 입수-데이터관리-보존 틀 아래 OASIS 보존 절차를 분석하고 온라인 연속간행자료 보존에 적합한 절차를 설계하였다. <그림 4>는 온라인 연속간행자료의 보존 절차를 나타낸 것이다.

1) 수집대상 자료 선정

수집 대상자료 선정 단계는 온라인 연속간행자료보존 시스템이 보존을 목표로 삼고 있는 대상을 정확하게 기술(description)하는 단계이다. 즉, 온라인 연속간행자료의 의미를 정확



<그림 4> 온라인 연속간행자료의 보존 절차

하게 정의하고 이 범위에서 가치 있는 자료의 선정이 필요하다. 이를 바탕으로 이 연구에서는 보존을 위한 온라인 연속간행자료의 수집 범위를 다음과 같이 제안하고자 한다.

- 한국 또는 한국 사람에 의해 작성된 자료
- 인쇄 자료로 발행되지 않는 인터넷 자료
- 전자적 기록매체로서 동일한 제호로 계속 제작되는 정기간행자료 또는 부정기 간행 자료
- 정기간행자료(periodical) - 웹진(e-zine), 학술지, 인터넷신문, 연감류 등 출판 간격이 일정한 것
- 부정기간행자료 - 계속되는 학술회의의 회의록(proceedings), 단체나 학회의 회의록(transactions), 등 지속적으로 발행되는 자료

2) 온라인 연속간행자료 수집

선정된 온라인 연속간행자료 수집 단계에서는 대상 자료의 간행 주기에 따라 정기적으로 자료를 수집한다. 수집 기관에서는 보존 정책 및 선정 지침에 따라 생성된 수집 대상 자료의 URL 및 발행주기를 웹 로봇(Web Robot)에 등록하고 웹 로봇은 등록된 자료에 한하여 URL에 해당하는 사이트의 자료를 수집한다. 자료 수집 시, 첫째, 발행주기에 따라 신규로 배포된 자료를 우선적으로 수집하며, 둘째, 자료의 변경 이력을 위하여 기존에 수집되었던 자료와 변경 검사 후 보존 가치가 있는 변경사항의 발견 시 이력 정보로 새롭게 수집한다. 셋째, 정보의 수집과 동시에 정보 재구현에 관련된 부가적인 정보(소프트웨어, 웹서버 정보 등) 및 운영체제를 동시에 수집한다.

3) 온라인 연속간행자료 메타데이터 작성

온라인 연속간행자료 메타데이터 작성 단계는 수집한 자료에 대한 메타데이터를 추출하여 메타데이터 형식에 따라 작성하는 단계이다. 전통적인 도서관에서 서지 정보 구축을 위하여 사용되어온 MARC 형식은 1970년대의 컴퓨터 환경을 반영한 레코드구조로서, 데이터 교환 및 저장매체의 기술 변화를 반영하지 못하고 있으며(Day, Herryl 2000), 이로 인해 새로운 환경에 적용하는 데에는 유연성이 부족하다는 약점을 지니고 있다. 즉, 새로운 디지털자료에 대하여 MARC는 적절한 데이터 표현 메타데이터가 될 수 없다(김성은, 김태수 1997). 따라서 MARC의 단점을 보완하고 온라인 연속간행자료로 보존에 요구되는 정보를 기술할 수 있는 새로운 메타데이터 형식 개발이 제안되고 있다(김태수 1998a). 또한 기존의 MARC기반의 시스템에서 축적된 서지정보와 새로운 메타데이터 형식과의 통합적 운영이 요구되고 있다.

4) 장기보존 단계

전기적 신호(0과 1의 이진법으로 표현된 비트)로 표현되어 특정 매체에 저장된 디지털자료는 정보의 내용을 담고 있는 콘텐츠와 이를 저장하고 있는 매체의 이원화는 도서의 보존과 달리 비트로 표현된 콘텐츠와 이를 의미화 할 수 있는 기술(technology)까지 보존해야 한다는 특징이 있다. 그러나 이러한 콘텐츠 재구현 소프트웨어 및 저장 매체의 기술은 5년을 넘기지 못하고 퇴화된다. 보존 처리 기술은 장기 보존 단계에서 중요한 요소로서, 시스템의 보존 전략을 수행하는 기술적 접근 방식을 의미한다. 현재 가장 많이 알려진 보존전략은 크게 예를

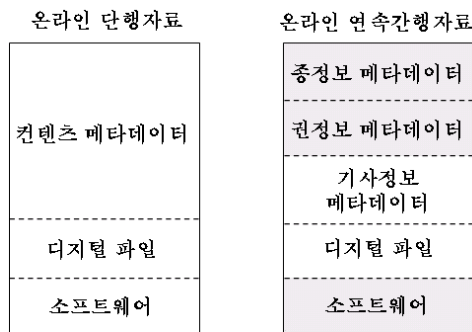
레이션(emulation), 마이그레이션(migration), 캡슐레이션(encapsulation) 방식으로 구분되며 최근 범용 가상 컴퓨터(Universal Virtual Computer: UVC)가 제안되었다.

발행기간 동안 발행되는 권 호에 따라 지속적으로 서지를 구축해야하는 연속간행자료의 경우 단순한 일회적 보존 자료를 통하여 장기 보존에 사용할 효율적인 보존 패키지를 구축할 수 없다. 보존 패키지란 <그림 5>와 같이 디지털자료, 메타데이터, 소프트웨어와 같이 실질 콘텐츠의 보존과 재구현(rendering)에 필요한 소프트웨어 환경을 통합하여 관리할 수 있는

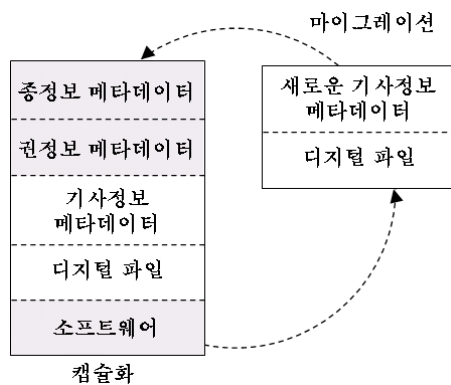
단위 요소로 정의된다.

<그림 6>에서와 같이 온라인 단행자료와 달리 온라인 연속간행자료는 같은 자료의 기사(콘텐츠)가 다른 보존 패키지로 구성될 경우 기사별로 종 정보, 권 호 정보를 비롯하여 소프트웨어 환경 등의 메타데이터를 반복하여 입력되기 때문에 중복을 피할 수 없다.

따라서 장기보존 단계에서는 온라인 연속간행자료의 장기보존을 위하여 캡슐화 및 마이그레이션 정책을 병행한 정책의 도입이 필요하다. 캡슐화에 필요한 실제 디지털자료 및 모든 환경정보에 대한 정보를 보존 패키지에 구성하여



<그림 5> 보존 패키지 구성 요소



<그림 6> 온라인 연속간행자료 장기보존 정책

장기 보존한다. 추후에 발생하는 중복적인 중 정보, 권·호 정보, 소프트웨어에 대한 메타데이터는 정기적으로 보존 패키지를 읽어 주기적으로 새롭게 배포된 기사에 관한 디지털자료를 업 데이트하여 하나의 자료에 하나의 보존 패키지가 구성될 수 있도록 마이그레이션 한다.

5. 결 론

인터넷의 대중화와 미디어의 결합현상으로 발생한 온라인상의 멀티미디어의 출현은 기존의 종이 매체나 CD-ROM 등으로 서비스되던 많은 자료들이 온라인으로 생산되고 제공되는 형태로 변경되었다. 또한 이러한 현상은 계속적으로 진행되는 연속간행자료의 경우에는 더 빠른 속도로 매체의 변경 현상이 나타나고 있다. 그러나 앞에서 본바와 같이 디지털 자료의 보존을 위하여 추진되고 있는 주요 국가의 온라인 디지털자료 아카이빙 시스템에서는 수집과 보존, 이용이 편리한 단행자료에 한하여 시범적으로 온라인 디지털자료의 아카이빙이 이루어지고 있는 실정 이었다. 단행자료에 비하여 연속간행자료는 상대적으로 생명주기가 짧을 뿐만 아니라 완벽하게 수집하여 지적문화유산으로 보존하기 위하여 많은 어려움이 있고 자료의 완결성과 활용측면을 고려하여 수집 초기부터 온라인 연속간행자료의 속성을 분석하고 수집·보존될 수 있는 시스템 구현의 연구가 필요하게 되었다.

따라서 디지털자료의 수집 및 보존에 관한 필요성을 인식하고 디지털자료의 아카이브 분야에 표준으로 인정받고 있는 OAIS 참조모형

을 활용하여 디지털자료의 상호운영이 가능한 보존 메타데이터 요소 및 전반적인 정보 수집, 보존에 대하여 연구하였다. 온라인 연속간행자료 수집 및 보존에 필요한 프로세스 모형을 개발하기 위하여 OAIS 참조모형을 분석하고 이를 기반으로 하고 있는 대표적인 보존 메타데이터 개발 프로젝트 및 국내에서 실행되고 있는 OASIS 프로젝트를 분석하였다.

온라인 연속간행자료의 수집 및 보존을 위한 정책수립 방안에 대하여 조사된 내용은 다음과 같다.

첫째, 제도 기반 구축은 '법·제도 확립'과 '장서수집 확립' 등의 제도 기반 구축영역에서는 '아카이빙 사업의 조사와 파악', '아카이빙 납본 관련 법 제·개정', '장서선별기준 수립' 등이 조사되었다.

둘째, 거버넌스 체제 구축은 4개의 영역으로 조사되었다. 즉 '통합기구 확립', '운영방안 구축', '이해관계자의 협력 네트워크 구축', '권역별 협력 네트워크 구축' 등이 주요 영역으로 도출되었으며, 총 28문항으로 조사되었다.

셋째, 전략 수립은 1개의 영역으로 '보존 메타데이터 표준 정립', 포맷 레지스트리 연구', '장기보존전략 계획 수립' 등 총 6문항으로 구성되었다.

넷째, 운영체제 구축은 2개의 영역으로 '테스트 베드 운영', '보존시스템 운영', '아카이빙 관리 메타데이터 표준 정립', '아카이빙 프로세스 정립' 등이 도출되었으며 총 22문항이 조사되었다.

또한 현재 사용자 중심의 서비스로 제공되고 있는 콘텐츠 배포기술인 RSS 기술을 이용하여 온라인 연속간행자료의 자동 수집 방법을 제안

하였다. RSS 기술의 이용은 온라인 자료의 작성 및 배포를 간편하게 수행할 수 있고, 이에 따른 잦은 콘텐츠 업데이트로 배포자에게 과중한 업무를 발생시키는 납본 제도와는 달리 배포자와 수집기관에서 동시에 이용할 수 있는 납본 기술을 제안하였다는 점에서 의미가 있다.

온라인 연속간행자료의 장기보존을 위한 기능 확장 기반을 제공하였고, 각 기능의 수행을 위한 입수 단계, 저장 단계, 관리 단계, 운영 단계, 보존 단계, 접근 단계의 자세한 프로세스를 분석하여 단계 별 수집데이터와 처리방안들을 제안하였다.

참 고 문 헌

- 구중익. 2006. 『연구장비정보의 RSS 기반 SDI 시스템 설계 및 구현』. 석사학위논문, 충남대학교 대학원, 문헌정보학과.
- 국립중앙도서관. 2006. 『한국문헌자동화목록형식』. 서울: 한국도서관협회.
- 김성은, 김태수. 1997. “KORMARC의 DTD 및 변환프로그램 개발 연구.” 『제4회 한국정보관리학회 학술대회 논문집』, 121-124.
- 김연정. 2004. 『전자문서 보존용 메타데이터 요소 개발』. 석사학위논문, 연세대학교 대학원, 문헌정보학과.
- 김준일. 2006. 『RSS 크롤러를 이용한 웹 문서의 분류』. 석사학위논문, 한양대학교 정보통신대학원, 정보통신학과.
- 김태수 1998a. “더블린코어(Dublin Core) 메타데이터.” [cited 2006. 3. 5]. <<http://dewey.yonsei.ac.kr/courses/metadata/DC.hwp>>.
- 김태수. 1998b. 메타데이터변환과 자원기술구조의 연구. 『정보관리학회지』, 15(3): 95-112.
- 김희정. 2003. 『전자저널 아카이빙을 위한 OASIS 참조모형의 적용방안에 관한 연구』. 박사학위논문, 연세대학교 대학원, 문헌정보학과.
- 박은경. 2002. 전자문서의 장기보존을 위한 표준화 연구 동향. 『한국기록관리학회지』, 2(1): 97-117.
- 서혜란. 2003. 디지털자료의 납본과 보존을 위한 각 국가의 노력. 『정보관리학회지』, 20(1): 373-394.
- 오대규. 2001. 『텔파이기법을 이용한 신중 및 재출현 전염성 질환 관리에 관한 연구』. 박사학위논문, 연세대학교 대학원, 보건학과.
- 윤세민. 2002. 웹진의 웹 미디어적 특성 연구. 『한국출판학회연구』, 44: 265-303.
- 윤희윤. 2002. 국내의 전자출판물 납본제도의 분석. 『한국문헌정보학회지』, 36(2): 185-207.
- 이경철. 2006. 『유아교육기관 평가인증제의 산, 학, 관 협력모형 개발연구』. 박사학위논문, 부산대학교 대학원, 교육학과.
- 이소연. 2002. 디지털 아카이빙의 표준화와 OASIS 참조모형. 『정보관리연구』, 33(3): 45-68.

- 이수상. 2004. 디지털 아카이빙의 현안과 과제. 『한국교육학술정보원 연구보고서』, 85-139.
- 이수상. 2004. 디지털 아카이빙의 워크플로우와 보존처리 기술에 관한 연구. 『한국도서관·정보학회지』, 35(3): 119-138.
- 이원영, 강진영. 2005. 전자문서 아카이빙 표준모델 연구. 『정보관리학회지』, 22(2): 146-164.
- 이응봉. 2000. Non-KORMARC 기반 연속간행물시스템의 문제점 및 개선방안 - 정부 기록보존소 연속간행물시스템을 대상으로-. 『사회과학연구』, 11: 199-220.
- 이중문. 2004. 디지털도서관의 정보조직화 방안 연구. 『한국도서관·정보학회지』, 35(1): 319-335.
- 이중성. 2001. 『텔파이 방법』. 서울: 교육과학사.
- 이혜원. 2004. 온라인 디지털자료 구축: 국립중앙도서관. 『디지털 유산 보존에 관한 기초연구 보고서』, 147-160.
- 전중홍. 2005. "콘텐츠 신디케이션 표준화 동향: RSS, ATOM, 블로그." *IT Standard Weekly* 2005: 252-254.
- 조운의. 2001. 『XML/KORMARC 통합시스템 구현에 관한 연구』. 박사학위논문, 중앙대학교 대학원, 문헌정보학과.
- 조윤희. 2003. 디지털콘텐츠 메타데이터 포맷의 비교 연구. 『한국문헌정보학회지』, 37(2): 135-152.
- 조현주. 1999. 『웹자원의 장기보존을 위한 국가 인터넷아카이브 모형 구축에 관한 연구』. 석사학위논문, 연세대학교 대학원, 문헌정보학과.
- 하진희. 2003. 『전자책 유통을 위한 중앙 레지스트리 모델 구축에 관한 연구』. 석사학위논문, 숙명여자대학교 대학원, 문헌정보학과.
- 한국전산원. 2003. 『기업간 워크플로우 통합 및 상호 연동성 구현』. 용인: 한국전산원.
- 한성국. 2005. 학술지 메타데이터의 표준화 체계에 관한 비교 연구. 『한국도서관·정보학회지』, 36(1): 415-440.
- 홍재현. 2004. 디지털 정보자원을 위한 보존 메타데이터 요소 설정 연구. 『한국도서관·정보학회지』, 35(3): 175-204.
- Ammen, Cassy. 2002. "MINERVA: Mapping the Internet Electronic Resources Virtual Archive." *National Diet Library Newsletter*, 125: 7-12.
- Ayers, Danny and Watt, Andrew. 2005. *Beginning RSS and Atom Programming*. Indianapolis: Wiley Publishing, Inc.
- CCSDS. 1999. "Draft Recommendation for Space Data System Standards: Reference Model for an OAIS." *CCSDS 650.0-R-1 RED BOOK*. [cited 2001. 3. 2].
 <<http://public.ccsds.org/pdf/ccsds-6500.R-1.pdf>>.
- CCSDS. 2001. "Draft Recommendation for Space Data System Standards: Reference Model for an OAIS." *CCSDS 650.0-R-2 RED BOOK2*. [cited 2002. 4. 7].
 <<http://public.ccsds.org/pdf/ccsds-6500.R-2.pdf>>.
- CCSDS. 2002. "Reference Model for an Open

- Archival Information System(OAIS).” *CCSDS 650.0-B-1 BLUE BOOK, Issue 1*. [cited 2005. 3. 11].
 <<http://public.ccsds.org/publications/Blue-Books.aspx>>.
- CEDARS. 2002. “Cedars Guide to Preservation Metadata.” [cited 2005. 4. 3].
 <<http://www.leeds.ac.uk/cedars/guideto/metadata>>.
- CEDARS. 2002. “Metadata for Digital Preservation: the Cedars Project Outline Specification.” [cited 2006. 1. 20].
 <<http://www.leeds.ac.uk/cedars/colman/metadata/metadataspec.html>>.
- CEDARS Official Website*. [cited 2005. 12. 20].
 <<http://www.leeds.ac.uk/cedars>>.
- Day, Michael. 1999. “Metadata for digital preservation: an update.” *Ariadne Issue*, 22. [cited 2006. 1. 11].
 <<http://www.ariadne.ac.uk/issue22/metadata/intro.html>>.
- Day, Michael and Heery, Rachel. 2000. *Metadata: an introduction and Practical guide*. London: Library Association.
- DCMI. 2004. “Dublin Core Metadata Element Set, Version 1.1: Reference Description.” [cited 2005. 3. 20].
 <<http://dublincore.org/documents/dces>>.
- Gartner, Richard. 2003. “MODS: Metadata Object Description Schema.” *JISC. TSW 03-06*: 1-13.
- Hammersley, Ben. 2003. *Content Syndication with RSS*. Beijing: Farnham: O’Reilly.
- Heery, R. 1996. “Review of Metadata Formats.” *Program*, 30(4): 345-378.
- Lavoie, Brian F. 2000. “Meeting the challenges of digital preservation: The OAIS reference model.” *OCLC Newsletter*, 243: 26-30.
- Lee, Kyong-ho, Slattery, Oliver, Lu, Richang Tang, Xiao and McGrary, Victor 2002. “The State of the Art and Practice in Digital preservation.” *Journal of Research of the National Institute of Standards and Technology* 107(1): 93-106.
- Lenares, Deborah. 1999. “Faculty Use of Electronic Journals at Research Institutions.” *In Racing Toward Tomorrow: Proceedings of the ACRL Ninth National Conference*-1-6.
- Lorie, Raymond. 2001. “Long-term Archiving of Digital Information.” *Proceedings of the 1st ACM/IEEE-C Joint Conference on Digital Libraries* 346-352.
- Lupovich, Chatherine and Masanes, Julien. 2002. *Metadata for long-term preservation of Electronic publications*. Amsterdam: Drukkerij Mart.spruijt.
- Meeker, Mary. 2004. “An Update from the Digital World.” *Morgan Stanley Equity Research*: 1-20.
- MINERVA. 2003. “WEB SITE CAPTURE & ARCHIVING.” [cited 2006. 4. 21.].

- 〈<http://www.loc.gov/acq/devpol/webarchive.html>〉.
- MINERVA Official Website*. [cited 2006. 4. 11]. 〈<http://www.loc.gov/minerva>〉.
- MODS. 2005. "Outline of Elements and Attributes in MODS Version 3.1" [cited 2006. 4. 21.]. 〈<http://www.loc.gov/standards/mods>〉.
- NEDLIB Official Website*. 2006. [cited 2006. 4. 21]. 〈<http://nedlib.kb.nl>〉.
- NLA. 1999. Preservation Metadata for Digital Collections. [cited 2006. 4. 21]. 〈<http://www.nla.gov.au/preserve/pmeta.html>〉.
- OASIS Official Website*. 2006. [cited 2006. 5. 7]. 〈<http://www.oasis.go.kr>〉.
- OASIS. 2006. "온라인 디지털자원 선정지침." [cited 2006. 4. 20]. 〈<http://www.oasis.go.kr>〉.
- OCLC. 2001. "Preservation Metadata for Digital Objects: A Review of the State of the Art." *OCLC/RLG Working Group on Preservation Metadata*.
- OCLC. 2002. "Preservation Metadata and the OAIS Information Model: A Metadata Framework to Support the Preservation of Digital Objects." *OCLC/RLG Working Group on Preservation Metadata*. [cited 2006. 2. 1]. 〈<http://www.oclc.org/research/pmwg>〉.
- Rainie, Lee. 2005. "The State of Blogging." *Pew Internet & American Life Project Report*. [cited 2006. 2. 1]. 〈http://www.pewinternet.org/PPF/r/113/report_display.asp〉.
- World Wide Web Consortium*. 2006. [cited 2006. 4. 1]. 〈<http://www.w3c.org>〉.
- Technology at Harvard Law. 2006. "RSS 2.0 Specification." [cited 2006. 4. 21.]. 〈<http://blogs.lav.harvard.edu/tech>〉.