

디지털 자원의 웹 아카이빙을 위한 납본 프로세스 개발 및 기능 설계*

Development of Deposit Process and Function Design for Web Archiving of Digital Resources

오상훈(Sang-Hoon Oh)**

최영선(Young-Sun Choi)***

초 록

국립중앙도서관에서 인쇄 출판물을 대상으로 운영 중인 납본체계와는 달리 웹 아카이빙인 OASIS(Online Archiving & Searching Internet Sources)는 웹 사이트, 웹 자원 등의 온라인 디지털 자원을 대상으로 자원 수집, 관리 및 보존하기 위한 과정이 필요하다. 이에 본 논문에서는 웹 아카이빙을 위한 디지털 자원 납본 프로세스를 개발하기 위해 디지털 자원 납본 주체와 대상을 정의하였고, 납본 프로세스를 위한 단계별 정의와 기능을 명시하였다. 또한 디지털 납본 시스템 구성을 위한 업무 흐름도와 단위 업무에 따른 기능 정의와 정보 흐름을 위한 구조를 제시하였다.

ABSTRACT

The National Library of Korea is administering a legal deposit system for the printed-publications. Whereas, OASIS(Online Archiving & Searching Internet Sources) has to design a system to collect, manage and preserve web sites and web resources for Web Archiving. The purpose of this study is to develop a digital deposit process for digital resources. As a result, this study defines the subjects and objects for digital deposits, and describes the definitions and the functions according to digital deposit steps. Also, this study designs the data flow diagram and proposes the function definitions on unit works and the structure for the flow of information.

키워드: 디지털자원, 웹 자원, 웹 아카이빙, 디지털 납본, 디지털 납본 프로세스
digital resources, web resources, web archiving, digital deposit, digital deposit process

* 본 연구는 2007년도 국립중앙도서관 『OASIS 표준화과제 연구』의 지원으로 수행되었음.

** 한국디지털콘텐츠산업협회 사무국장(osh@dcforum.or.kr) (제1저자)

*** 한국디지털콘텐츠산업협회 연구원(youngsun@dcforum.or.kr) (공동저자)

■ 논문접수일자: 2008년 8월 19일 ■ 최초심사일자: 2008년 8월 25일 ■ 게재확정일자: 2008년 12월 1일

■ 情報管理學會誌, 25(4): 5-23, 2008. [DOI:10.3743/KOSIM.2008.25.4.005]

1. 서론

디지털 자원은 인터넷이 대중화 된 오늘날 주요한 정보자원으로 폭넓게 활용되고 있다. 온라인과 오프라인에서 망라적으로 이용되는 디지털 자원 중에서도 웹에서만 이용할 수 있는 온라인 디지털 자원은 접근의 용이성과 사용의 편리함으로 활용도가 높아졌지만, 생명주기가 짧고, 자원의 유실 또는 훼손의 가능성이 있어 안정적인 자원의 수집, 관리 및 보존을 위한 지원이 요구 되고 있다.

특히 최근에는 디지털 자원 중에서도 단순히 개개의 자산이 아닌 국가·사회의 문화 인프라로 분류될 수 있는 가치 있는 자원의 경우 국가적인 차원에서 수집·보존하여 자원의 이용을 촉진하고자 하는 움직임이 일고 있다. 이미 지난 90년대 말부터 호주, 영국 및 캐나다에서는 국립도서관 및 아카이빙 기관에서 온라인 디지털 자원을 수집해 왔으며, 법·제도적인 차원과 실질적인 자원의 수집 및 보존을 위한 기술적인 지원을 위한 연구가 지속적으로 진행되고 있다(김유승 2007). 우리나라도 예외는 아니어서 온라인자원에 대한 법·제도적인 정비와 함께 2000년 이후 국립중앙도서관을 중심으로 웹 아카이빙인 OASIS(Online Archiving & Searching Internet Sources, 이하 OASIS)를 운영하고 있다.

웹 아카이빙은 자발적 납본(voluntary deposit)을 실시하고 있는 국가에서 디지털 자원의 법정 납본을 염두에 둔 대안으로 활용될 수 있다. 실제 캐나다는 온라인 디지털 자원에 대해 자발적 납본을 실시하고 있지만 국립도서관에서 진행했던 EPPP 프로젝트를 참고하여 지속

적으로 웹 아카이빙을 운영하여 디지털 자원을 수집하고 있다(PADI 홈페이지). 웹 아카이빙을 통해 디지털 자원의 납본 과정을 실험하고 그 기반을 마련하는 것이다. 온라인 디지털 자원은 무형의 비트 형태의 파일로 존재하고 변형되기 쉬운 상황에 노출된 콘텐츠로서 전통적인 납본 방식에서 다루는 도서 출판물과는 다르다. 또한 디지털 자원을 다루는 납본체계는 납본주체가 다양할 수 있고, 자원을 수집하는 방법과 절차가 이전 방식과는 차이가 있다.

따라서 본 연구에서는 OASIS 프로젝트를 대상으로 웹 아카이빙의 디지털 자원 수집·관리·저장 및 보존하기 위한 절차와 각 절차별 요구되는 시스템의 기능을 분석하여 새로운 개념의 디지털 자원 납본 프로세스를 개발하고자 한다.

1.1 연구 목적 및 문제

본 연구는 향후 디지털 자원의 새로운 관리방법의 필요성을 인식하고 웹 아카이빙 내·외부에서 디지털 자원의 수집 주체 그리고 대상 자원의 체계적인 수집, 관리, 저장 및 서비스를 위한 납본 프로세스를 개발한 후 해당 프로세스 별 기능사항을 설계하여 제시하고자 한다. 이러한 목적을 달성하기 위해 다음의 2가지 연구 문제를 해결하고자 한다.

첫째, 국립중앙도서관에서 운영하고 있는 도서 출판물을 대상으로 하는 납본절차와는 달리 온라인 디지털 자원을 대상으로 하는 OASIS 웹 아카이빙에서 온라인 디지털 자원을 수집·보존하는 납본 체계를 위한 웹 아카이빙의 자원 수집 절차를 제시한다.

둘째, 웹 아카이빙에서 온라인 디지털 자원의 납본을 위한 범위와 대상을 새롭게 정의하고, 프로세스 단계별 기능 정의와 이를 기술하는 절차에 대한 요구사항들을 도출하여 디지털 납본 단계별 프로세스를 제시한다.

1.2 연구 수행 방법

본 연구에서는 온라인 디지털 자원을 대상으로 웹 아카이빙을 위한 절차와 방법에 대한 논의를 필요로 하였다. 따라서 기존의 납본체계와는 차별화 된 새로운 디지털 납본 프로세스 (Deposit Process)에 대한 주체, 대상, 체계 및 방법 등을 참조하기 위해 OASIS의 현황 분석 - 시스템, 정책 및 대상자원 - 그리고 웹 아카이빙과 납본에 관한 해외사례를 조사·분석하였다. 각 사례를 통해 OASIS에 적용 가능한 범위의 디지털 납본 프로세스를 개념적으로 정의하고 필요한 기능을 개발하고자 하였다. 디지털 파일의 생성에서부터 실제 보존과정에 이르기까지 디지털 파일을 다루기 위해서는 파일의 생산자, 저작권자 등 참여주체가 다양하다. 또한 권리 관계도 복잡하고 디지털 파일의 특성정보, 유통정보 등 디지털 파일을 수집하고 관리하기 위한 단계별 역할에 대한 요구사항을 정확히 정의 할 필요가 있다. 따라서 현재의 납본 체계에 대한 분석과 향후 디지털 자원을 대상으로 하는 디지털 납본 체계에 대한 프로세스 단계별 기능을 정의하고 업무 프로세스에 적용하였다.

1.3 용어 정의

일반적으로 납본이란 “자료를 발행하거나 제작한 자가 일부 부수를 법령에서 정한 기관에 의무적으로 제출하는 것”(도서관법 제2조 8항)을 말하며 자국의 출판 저작물을 완벽하게 수집·보존하고 국가서지를 작성함으로써 당대의 정보유통을 보장하고 후대에 지적·문화적 유산을 전달 계승코자 하는데 있다(서혜란 2003). 국내에서 발행 또는 제작된 자료의 경우 도서관법의해 발행 또는 제작일로 부터 30일 이내에 국립중앙도서관에 납본을 해야 하며, 현재 국립중앙도서관 소장 자료의 약 70%를 구성하고 있다.

그러나 디지털 납본에 대한 정의는 아직 명확하지 않고 다만, “디지털로만 생산된 국가 문헌이나 정부 자료 등을 종이책과 마찬가지로 국립중앙도서관에 데이터베이스화 하는 작업이 의무화 된다”(문헌정보용어사전 2004)라고 정의하고 있다. 또한 이와 관련하여 2007년 국립중앙도서관은 『온라인 디지털자료 납본 및 이용에 관한 법률안』(이하 『디지털자료 납본 법안』)을 제안하며 제1조에서 “온라인으로 출판된 디지털자료를 수집·보존하여 국가의 문화유산을 전승하고, 국민들이 도서관을 통하여 이를 효율적으로 이용하는데 필요한 사항을 규정”한다고 표현하였다. 하지만 이 법안은 현재 통과되지 못하여 우리나라는 온라인 디지털 자료에 대한 법정 납본을 실시하지 않는 여러 국가 중의 한 곳이다.

따라서 본 연구에서 사용 된 디지털 납본이라는 용어는 법적 강제는 없으나, 납본의 의미와 절차를 활용한 광의의 의미로 해석하였다.

이러한 맥락에서 기존의 납본과 디지털 납본의 대상, 주체, 방법의 차이를 기술하였다. 더불어 디지털 납본을 기존의 납본 체계와 디지털 자원의 수집 체계의 차이를 고려한 자원의 수집과 관리절차를 모두 포함하고 향후 자원의 보존까지 고려한 것으로 납본 절차와 수집 절차를 구분한다. 또한 웹 아카이빙에서 웹 크롤러에 의한 자원의 자동 수집(harvesting)과정을 납본 프로세스에서의 하나의 절차이며 직접 제출하는 것 이외시스템을 통한 자동수집에도 납본이라는 용어를 포괄적으로 사용하였다.

2. 웹 아카이빙의 개요 - OASIS를 중심으로

2.1 웹 아카이빙의 정의 및 대상자원

웹 아카이빙은 온라인 디지털 자원을 수집하여 보존하기 위한 광범위한 노력으로 주로 디지털 형태의 인터넷 사이트와 기타 문화유산

(cultural artifacts)을 수집하고 미래 세대에 전달하는 역할을 한다. 특히 온라인 디지털 자원의 양적·질적인 팽창 속도는 급격히 증가하고 있지만, 그 자원의 생명 주기가 짧아 인터넷 정보자원을 수집 및 보존하는 웹 아카이빙의 중요성을 인식하고 웹 자원을 지키기 위한 노력이 지속되고 있다.

이러한 웹 아카이빙의 수집대상자원은 웹 사이트와 그 웹사이트에 연계된 개별 자원을 포함하며, 개별 웹 아카이빙을 운영하는 기관의 자원 수집 정책에 따라 구체적인 수집 자원의 분야가 결정된다. 다만, 대부분의 웹 아카이빙에서의 수집방법은 웹 크롤러, 로봇을 이용한 자원 수집이 일반적인 경우라 웹 크롤러, 로봇의 성능에 따라 수집의 대상이 제한될 수도 있다.

다음의 <표 1>은 국제적인 사례로서 해외 아카이빙 프로젝트에서 수집하고 있는 수집대상 자원을 정리한 것이다. 현재 주요 웹 아카이빙에서는 웹 사이트의 전체 또는 일부를 스냅 샷 방식으로 저장하거나 웹 사이트에 연계된 개별 디지털 파일을 저장하고 있다.

<표 1> 웹 아카이빙 수집 자원

기관명	수집 대상 자원
European Archive	- 동영상과 음원: 파일형태의 디지털 자원 - 웹컬렉션: 유럽, 영국, 이탈리아의 웹 아카이브
UK Web Archiving Consortium	- 공공기관 웹사이트
MINERVA	- 의회 구성원(의원)의 홈페이지 - 후보, 시민, 후원단, 정보, 언론 등의 웹사이트 - 911사건을 다룬 웹 사이트
Internet Archive	- 포괄적인 자원 수집(웹, 동영상, 텍스트, 오디오, 소프트웨어 등)
WARP	- 웹사이트(국가기관, 도도부현, 행정지정도시, 지어촌합병, 법인/기구, 대학, 이벤트, 기타) - 전자저널
E-Resource	- 사회과학 중심의 웹 사이트
PANDORA	- 온라인 출판물, 웹 사이트
Canadian Information By Subject	- 인터넷 자원의 링크 정보
Intutue	- 블로그, 데이터베이스, E-books, 이미지, 법률정보, 정부 웹사이트, 통계, 소프트웨어, 동영상

2.2 국립중앙도서관 OASIS의 현황

2.2.1 개요

“OASIS는 미래 디지털 세대를 위한 현세대의 디지털 지적 문화유산의 수집·보존 프로젝트명으로 가치 있는 디지털 자원을 효과적으로 수집·보존·및 서비스하기 위한 목적으로 개발한 시스템”이다(OASIS 홈페이지). 2001년 ‘온라인저작물 수집·보존 시스템 구축 ISP 및 시범시스템 개발’을 시작으로 2005년 웹기반 OASIS 시스템 확장·개선 사업을 완료 하였으며, 2006년 OASIS 홈페이지를 통해 대국민 서비스를 실시하고 있는 우리나라를 대표하는 국립중앙도서관의 디지털 아카이빙이다.

2.2.2 수집대상자원 및 범위

OASIS는 방대한 범위의 온라인 디지털 자원들 중 가치 있는 디지털 유산의 수집을 위해 온라인 디지털 자원의 수집·선정지침을 마련하였고 그 지침에 준하여 디지털 자원을 선별적으로 선정·수집하고 있다.

OASIS의 수집대상이 되는 자원의 광의의 범위는 “한국과 관련되었거나 한국인이 생산했으며, 오리지널 디지털 형식으로 생산되어 온라인 접근이 가능한 유용성이 높은 정보자원”으로 현재 또는 미래의 이용자들의 정보요구에 부합해야 한다. 또한 오리지널 디지털 자원과 디지털화 된 자원을 구분하고 전자의 경우는 망라적 수집을 후자의 경우는 선택적 수집을 원칙으로 하며 디지털 자원의 특성을 감안하여 자원의 선별과 다양한 포맷을 관리하기 위한 기본 원칙을 명시하고 있다.

그러나 OASIS에서 수집·보존하여 후대에

물려주기 위한 자원의 범위는 실제로 비용, 인력 및 자원의 한계로 자원에 가치를 매기고 우선 수집할 자원을 선정한다. 특히 망라적인 범위를 대상으로 하는 OASIS이기에 우선 수집할 자원을 선정하는 일이 중요하다. OASIS에서 선정한 수집 우선 자료는 중앙정부와, ‘온라인 디지털자원 수집보존위원회’의 각 분야별 분과 위원회의 추천 자료, 전자저널, ~.kr의 웹사이트, 주요한 회의자료, 연구보고서 및 대학 간행물 등이 포함된다.

선정된 자원은 개별 파일이나 웹페이지 하나씩을 단위로 하는 웹문서와 사이트 전체를 미러링한 웹사이트로 나누어 웹 크롤러가 수집하고 있으며 일부 자원은 이용자 서비스를 지원하고 있다.

2.2.3 자원 수집 프로세스

OASIS는 수집자원의 유형에 따라 웹사이트와 웹자원의 특성을 고려하여 자원 수집을 위한 업무프로세스를 제공한다. 여기서 웹사이트는 특정 URL의 사이트를 그대로 미러링하여 저장하는 것이고, 웹자원은 하나 하나의 파일단위로 수집한 후 메타정보를 입력한 것을 의미하며 각각은 다음과 같은 프로세스를 구성한다.

첫째, 웹사이트 수집을 위한 업무는 관리자의 자료 선택과 기증자의 자료추천에 의해 자원 선정을 하고(1단계), 자료를 선별하여 웹사이트를 프로그램에 등록한다(2단계). 그 다음 웹 크롤러는 프로그램에 URL이 등록된 웹사이트를 1차 수집(3단계), 1판(최초의 수집 버전)을 저장하고 나면(4단계), 웹사이트의 갱신 주기를 개별적으로 확인하여 수집 주기를 선정하고(6단계), 그 수집 주기에 따라 주기적으로 웹사이트

를 수집·저장한다(7단계)(8단계). 이때 더 이상 업데이트가 없는 이벤트성 웹사이트 등은 4단계 이후 보존(9단계)이 되고, 지속적인 관리와 주기적인 업데이트가 이루어지는 웹사이트들은 7단계, 8단계의 수집과 저장이 반복이 되고 1관부터 N관까지의 웹사이트가 모두 보존 처리된다.

둘째, 웹자원의 수집을 위한 업무 프로세스는 관리자의 자료 선택과 기증자의 자료추천에 의해 선정을 하고(1단계), 자료를 선별한(2단계) 다음 해당 개별 자원을 수집한다(3단계). 수집된 자원이 기존에 수집된 자원과 중복되는지를 URL, 제목의 유사성, 파일 내용을 통해 검사하고(4단계), 자동분류(5단계)와 자동요약(6단계)을 마치면 메타정보를 입력하고(7단계), 담당 사서의 교열·검증을 마치면(8단계) 마지막으로 보존 처리된다(9단계).

그러나 OASIS에서 실제 업무가 이루어지는 과정은 정의된 프로세스에 따라 한 건, 한 건 순서대로 진행되는 것이 아니라, 선 수집 그리고 후 작업의 과정으로 이루어지기 때문에 자원을 수집한 후 제목과 URL 등 자원을 식별하기 위한 가장 기본적인 사항을 기입하고, 수집 작업이 완료된 후 전체 메타 정보를 입력하게 된다. 또한 프로세스 상으로 웹자원은 자동 분류, 자동 요약 작업이 진행되어야 하지만 현재는 생략되고 있는 등 실제 업무 프로세스와는 차이가 있다.

2.3.4 OASIS 분석의 시사점

국립중앙도서관은 구입, 구독, 기증, 제작 및 납본을 통해 자원을 수집하고 있으며 이중 OASIS는 웹크롤러를 이용한 자동 납본 방법과

저작권자의 자발적인 기증을 통해 자원을 수집하고 있다. 그러나 기증 자원의 비율이 거의 없어 대부분이 로봇에 의한 자동납본에 의한 수집이라 하겠다. 따라서 웹 자원을 대상으로 하는 디지털 납본체계에 대한 요구사항을 도출하였으며 기존 도서·출판물을 다루고 있는 납본체계와는 달리 웹 사이트와 자원을 수집 관리 보존하는 방법과 절차에 대한 업무를 부여해야 할 필요가 있다.

이에 현재 각 단계별 업무에 대한 프로세스 분석 결과 다음과 같은 한계가 나타났다.

첫째, OASIS는 웹사이트와 웹자원을 구분한 후 각각의 프로세스에 따라 자원을 수집하고 있지만 웹자원은 웹사이트의 특정 웹페이지에 첨부된 자원으로 이들을 연계해서 수집·관리해야 할 필요성이 있다. 현재 웹자원 수집 시에는 해당자원이 첨부된 웹 사이트의 정보를 수집하고 있지만 반대의 경우는 생략되는 경우가 많아 웹자원들 간의 연계 및 추가 정보에 대한 정확한 표시가 요구된다. 따라서 디지털 자원의 납본 프로세스의 자원 수집 및 목록생성 과정에서 메타데이터 항목을 생성하여 연계된 정보들이 표현되도록 해야 한다.

둘째, 웹 크롤러에 의한 자동수집에 의존하고 있어 크롤러의 성능에 의해 자원의 수집 범위가 한정될 수 있다. 물론 웹 크롤러의 성능이 향상되면 자연스레 해결될 수도 있는 문제지만 많은 비용과 시간 소요가 예상된다. 이에 본 연구에서는 수집 주체를 기존의 웹 크롤러 이외에도 자원의 생산자 또는 대행사 등으로 확대하여 한정된 자원의 수집 가능 범위를 확장시키고자 한다.

셋째, 수집한 자원을 관리·보존하여 미래세대에게까지 전달하기 위해서는 해당 자원의 객체

(object) 및 그와 관련된 모든 SW/HW 정보를 관리해야 하는데, 이러한 작업은 수집 자료의 상세한 메타데이터 작성 및 확인이 필요하며 주기적으로 시스템의 백업 및 업그레이드도 요구된다. 그러나 우선 자료의 수집에 집중하고 있는 OASIS 수집 업무에 대한 담당자들의 업무가 과중해 세부적인 디지털 납본 프로세스를 통해 각 단계별 업무 절차와 과정으로 업무를 분산하고자 한다.

3. 디지털 자원 납본 관련 사례 연구

3.1 각국의 디지털 자원 납본 현황

디지털 자원에 대해 그 가치를 인정하여 국가적인 차원에서 수집해야 할 필요를 인식하고는 있지만 법제도적 장치 및 디지털 자원 납본을 위한 기술적인 프로세스 및 시스템의 미비로 대부분의 국가에서는 저작권자들의 자발적 납본에 기대하는 수준에 있다.

일본의 납본시스템협회는 물리적인 형태가 있는 출판물과 온라인 출판물을 다르게 여겨 두 출판물을 동일한 방식으로 납본처리를 하지 않고 있다. 그렇지만 두 유형의 출판물을 모두 국가의 유산으로 인식할 수 있도록, 국가 도서관의 논리적인 확장을 요구하며, 새로운 방식의 출판물 모두를 다룰 수 있도록 적절한 메카니즘으로 납본을 고려하고 있다. 네덜란드와 스위스는 인쇄출판물의 법정 납본도 실시하지 않고 있지만, 디지털 출판물의 자발적인 납본을 위한 논쟁들이 계속되고 있다. 캐나다는 '책'의 일부

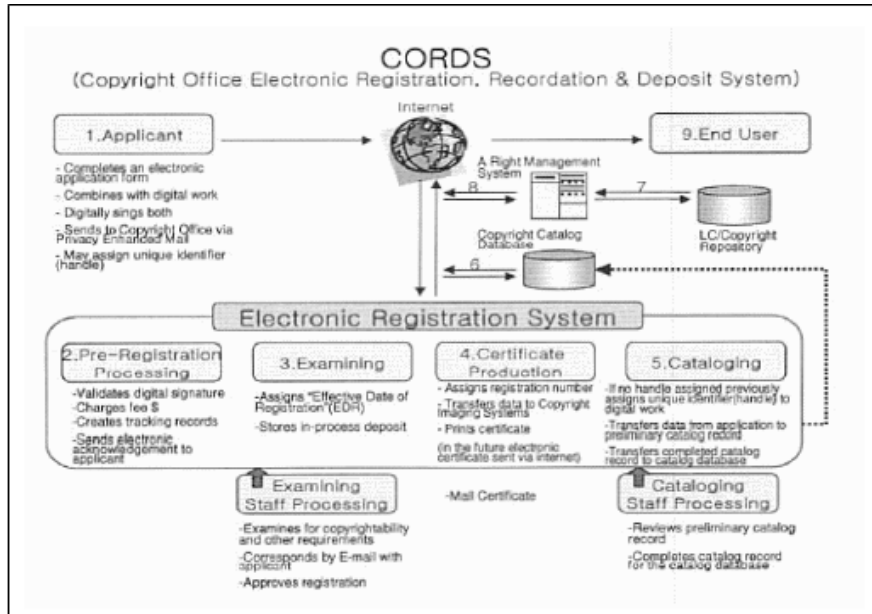
분으로 출판된 전자 자원만을 범위로 하였는데 자원에 대한 해석이 명확하지 않았다. 그러나 94년부터 95년까지 진행된 EPPP 프로젝트의 권고안을 참고로, 국가 도서관은 자발적 납본에 기반 하여 전자출판물을 수집하는 것을 계속해왔다. 특히, 다른 유형의 자원을 이용할 수 없는 출판물에 한하여 수집이 강조되었다. 오프라인 자원만을 법정납본으로 하며, 온라인 출판물을 제외한 국가는 향후 온라인 출판물을 포함하는 법률안을 제안하기 위한 과정에 있는 오스트리아와 자발적인 납본을 통해 온라인 출판물을 수집하는 중인 싱가포르가 있다. 또한 덴마크, 핀란드, 프랑스, 독일, 아이슬란드, 뉴질랜드, 노르웨이, 남아프리카 공화국, 스웨덴 등은 온라인/오프라인 디지털 출판물을 모두 포함하는 법률 제정하였거나, 법률을 제정하기 위한 과정을 진행 중에 있다(PADI 홈페이지).

3.2 디지털 자원 납본 시스템 관련 사례 및 연구

3.2.1 미국 - Copyright Office Electronic Registration, Recordation & Deposit System(CORDS)

미국 저작권사무소의 CORDS는 미국 헌법의 제1조 8항에 관한 법적 책임을 수행하기 위해 구축된 자동화 시스템으로 저작권 등록과 저작물의 저장을 가능하게 한다(오영화, 최석두 1996).

CORDS는 1993년에 인터넷을 통하여 신청과 전자등록을 위한 디지털 저작물의 납본을 자동화된 새로운 방법으로 시스템의 개발을 시작하였고, 지속적으로 CORDS에 대한 추가 연구 개발을 진행하고 있다.



〈그림 1〉 CORDS의 업무 흐름도(전자등록시스템)

또한 CORDS를 이용하여 현재 수용 가능한 파일형태(HTML, ASCII, PDF, and MP3 Audio Files)의 다양한 유형의 텍스트 작품(전자저널, 기술보고서, 전자책 등)과 음악작품에 대한 저작권 신청과 납본이 가능하며, 지속적인 CORDS 개발과 확장을 통해 다른 유형의 작품, 전자저널, 웹사이트, 멀티미디어 작품에도 적용할 예정에 있다(선명순 2003). CORDS를 통한 저작권 등록 및 납본의 처리절차는 다음과 같다.

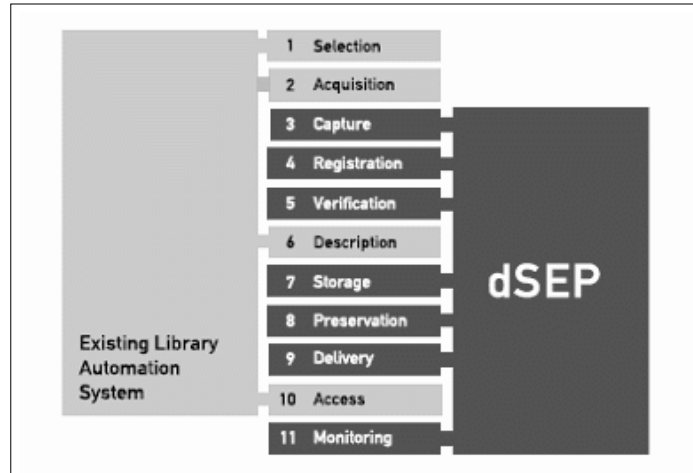
- ① 신청자는 제출서류를 접수하고 처리
- ② 전자 서명 확인 후 접수 확인서 교부
- ③ 등록일 부여 후, 납본 처리
- ④ 자원 등록 번호 부여 후 인증서를 발급
- ⑤ 목록 정보 저장
- ⑥ 외부에서 DB 검색 가능토록 준비
- ⑦ 미의회도서관으로 복제권 정보제공

⑧ 저작권 관리 시스템에 저장

3.2.4 NEDLIB Deposit System for

Elec-tronic Publications(이하 DSEP)

NEDLIB는 1999년 1월부터 2000년 12월까지 진행된 유럽의 국가 도서관들 간의 협력프로젝트로서, 네트워크화 된 유럽 보존 도서관을 실현하기 위하여 기본적인 기반을 구축하고자 하였다. 유럽 8개국 국립도서관(프랑스·노르웨이·핀란드·독일·포르투갈·스위스·이탈리아·네덜란드)이 참여하였으며, 네덜란드 왕립도서관(KB)이 프로젝트를 주도적으로 진행하였다. 주요 활동 사항으로는 전자출판물의 장기보존을 위한 보존 메타데이터 개발과 전자출판물의 납본을 위한 디지털 납본 체계 및 시스템을 개발하였다.



〈그림 2〉 DSEP 의 업무 프로세스

특히 기존의 전자도서관 시스템에 디지털 납본 체계를 결합한 형태의 DSEP은 미래 환경에서도 견디어낼 수 있는(future-proof) 시스템으로, 어떠한 디지털 형태에서도 기원이나 매체, 포맷에 상관없이 실용적 활용이 가능하다는 장점을 갖는다(Titia van der Werf 2000).

또한, DSEP의 데모 시스템을 개발하면서 ISO 14721 OAIS(Open Archival Information System) 모델의 수집(ingest), 저장(archival storage), 데이터 관리(data management), 접근(access), 운영(administration) 기능에 장기보존을 위한 보존(preservation)기능을 추가하였다. 뿐만 아니라 개별 납본도서관이 독립적인 기능을 수행할 수 있는 모델을 제공하며, 도서관 온라인 열람목록(Online Public Access Catalogue, OPAC)시스템, 도서관 수서시스템, 그리고 목록시스템과 같은 기존의 도서관시스템과 기능적 연계를 가능하게 하였다.

이러한 NEDLIB 프로젝트는 초기에 3개의 출판사 - Elsevier Science b.v., Kluwer Acade-

mic, Springer Verlag - 로부터 지원받은 전자출판물을 테스트 대상으로 진행하였다.

NEDLIB 시스템 총 11단계 중에서 DSEP 시스템 수행단계는 캡처(Capture), 등록(Registration), 검증(Verification), 저장(Storage), 보존(Preservation), 전송(Delivery), 점검(Monitoring) 단계로 총 7단계로 구성된다.

- ① 캡처: 전자출판물의 복사본이 디지털 매체 또는 네트워크를 통해 도착
- ② 등록: 수집 자원의 데이터와 시스템 환경 데이터를 보존 패키지로 시스템에 등록
- ③ 검증: 전자출판물의 무결성, 진본성을 검증
- ④ 저장: 납본시스템 저장소로 수집된 전자출판물 이전/이동
- ⑤ 보존: 적극적인 보존전략 수행 단계
- ⑥ 전송: 보존패키지의 보존과 접근을 위한 적절한 경로로 전송
- ⑦ 점검: 전 과정을 주기적으로 점검 및 통제

3.3 사례분석에 대한 시사점

몇몇 유럽의 국가들을 제외하고는 아직까지 디지털 자원의 납본과 관련된 법·제도적인 뒷받침은 부족하거나, 준비단계에 있는 경우가 대부분이다. 다만 디지털 자원의 수집 및 보존에 대한 각국의 관심이 높아 국립도서관을 중심으로 프로젝트를 운영해오고 있다.

미국은 1989년부터 물리적 형태의 전자출판물(Magnetic tape, CD-ROM, Microform 등)에 대한 납본을 받았지만 온라인 전자출판물은 법정 납본의 의무를 제외하였다. 다만 온라인 출판물의 저작권 등록을 위해서 온라인 저작권 등록 시스템인 CORDS를 통해 저작권 등록과 동시에 납본을 실시하고 있다. 온라인과 오프라인의 모든 자원에 대한 자발적 납본을 실시하고 있는 네덜란드에서는 디지털 정보의 생산자들의 자발적인 의지에 의존하는 실정이다. 각각 앞서 살펴본 CORDS와 NEDLIB의 DSEP은 온라인 디지털 자원을 수집, 관리 및 저장하고 있는 시스템을 구축한 사례로 특히 웹 아카이빙을 활용한 측면에서 NEDLIB의 DSEP에서 제시한 납본 프로세스와 각 단계별 기능은 본 연구의 선행연구로서 활용 가치가 높다.

현재까지 국내에서는 디지털 납본을 위한 법률적인 내용이나 관련 정의에 대한 근거가 부족한 상황이다. 그러나 디지털 자원을 장기적으로 보존하고 활용하고자 하는 요구사항을 반영할 수 있는 새로운 개념의 디지털 납본체계를 제시할 필요성이 있다. 따라서 본 연구에서는 디지털 자원의 지속적인 수집, 관리 및 보존을 위한 방법과 절차를 “디지털 납본체계”라는 범위로 한정하여 정의하고자 한다.

4. 디지털 납본 프로세스 및 기능 설계

4.1 개요

우리나라는 법정 납본을 통해 도서 등의 인쇄출판물 및 CD, DVD 등의 오프라인 디지털 자원을 국립중앙도서관에 제출해야 한다. 이때 납본 신청자들은 직접방문 또는 우편 접수의 방법으로 정해진 절차에 따라 양식을 작성하고 2부의 출판물과 관련 서류를 제출한 후, 관계자의 확인과 검증을 마친 다음 납본이 되는 절차를 따른다. 그러나 물리적 객체가 존재하지 않는 온라인 디지털 자원, 즉 웹 자원은 법률적인 강제가 없으며, 납본의 주체가 웹 크롤러 또는 해당 업무에 전문성이 없는 일반인이 될 수 있어 디지털 자원의 납본을 위해 체계적이고 복잡하지 않은 가이드라인이 제공되어야 한다. 또한 새롭게 제안되는 디지털 납본체계에서는 수집에서 저장까지의 단계별 구체적 업무 절차와 해당하는 세부 기능을 기술하여 담당자들의 업무 분산도 이루어야 한다. 우리보다 앞서 웹 아카이빙을 시행했던 기관들은 기존의 전자도서관 시스템을 활용하였으며(NEDLIB, MINERVA), 세계 각국의 도서관, 기록관, 대학, 연구소 등 주요 디지털 보존기관에서 디지털 아카이브를 구축하기 위한 프레임워크나 출발점으로 OASIS 참조모형을 활용하는 것이 추세이다(PREMIS survey, 2004).

특히 ISO 14721 OASIS 참조모형은 디지털 보존에 관련된 모든 활동 및 프로젝트의 개념적 기반으로 채택되어 온 디지털 보존분야의 독보적 표준이라 할 수 있으며 OASIS의 경우에도 OASIS 참조모형 준수를 표방한다.

이에 본 연구에서는 디지털 정보자원 보존시스템을 기존의 도서관 자동화시스템(Existing Library Automation System)에 디지털 납본을 위한 체제 및 시스템 구축사업으로 OASIS의 참조모형을 채택해서 기능적 모형을 플러그인 형태로 결합하여 구축한 DSEP의 프로세스를 참고하고 디지털 도서관의 업무 흐름도를 근간으로 재구성하였다. 단, 납본 이후 웹 아카이빙에서의 과정인 보존은 납본 프로세스에서 제외하였다. 그 결과 다음의 <그림 3>과 같이 6단계의 프로세스로 구성하였으며, 각 프로세스 단위별 개요, 기능 및 역할은 다음과 같다.

4.2 자원수집계획단계(1단계)

1단계는 수집 대상에 따른 납본범위와 납본 주체를 정의하는 수집 계획의 과정으로 자원을 납본하는 주체에 따라 각각의 역할이 명기되며, 이 단계에서 수립된 계획을 바탕으로 실제의 수집 작업이 이루어진다.

특히 본 연구에서는 납본에 의한 수집 프로세스가 대상과 주체에 따라 구분된다고 보고 납본 대상을 웹자원과 디지털 개별 자원으로, 납본 주체는 웹 크롤러와 납본자로 제안하였다. 웹

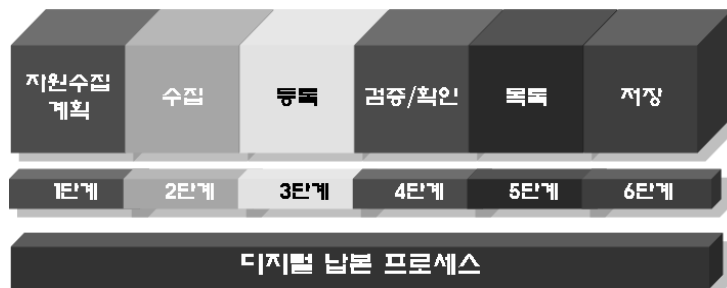
자원은 주로 웹크롤러에 의해 수집되며, 디지털 자원은 웹 크롤러와 납본자에 의해 수집된다. 웹 크롤러는 OASIS의 WebBee21과 같은 웹로봇을 일컬으며, 납본자는 출판사, 지적 재산물에 대한 권리대행사 또는 기관, 개인 및 단체를 포함한다.

가. 디지털 납본의 범위 및 주체 정의

1) 수집대상에 따른 납본범위 정의

웹 아카이빙은 디지털 자원 중에서도 웹 자원으로 그 범위가 한정된다. 웹에서 이용 가능한 자원은 크게 웹사이트, 웹문서(웹페이지), 디지털자원(개별파일)으로 구분이 가능한데 각각 서로에게 연계된 자원으로, 디지털자원은 웹문서에 첨부되어 있으며, 웹문서는 웹 사이트를 구성하는 다수의 웹 페이지를 의미한다. 다시 말해 하나의 웹 사이트는 여러 개의 웹문서로 구성되어 있으며 각각의 웹문서에는 디지털 자원이 추가될 수 있다. OASIS에서는 웹사이트와 디지털자원에 대해 다음과 같이 구분하고 있으며, 본 연구에서는 웹문서를 추가하였다

- ① 웹사이트: 독립된 URL을 갖는 하나의 사이트 단위로 OASIS에서는 해당 사이



<그림 3> 디지털 납본 시스템(Digital Deposit System) 프로세스

트의 모든 자원을 가져와 미러링 한다.

- ② 웹문서: 웹문서란 블로그, 게시판, 사이트를 포함하여, 정보를 담고 있는 웹 상의 모든 문서를 의미 한다
- ③ 디지털자원: 하나 또는 몇 개의 파일 단위로 정보의 가치를 갖게 되는 것으로, 특히 첨부된 개별 파일을 디지털자원으로 정의한다.

2) 납본주체 정의

웹 아카이빙에서 자원을 수집하는 방법은 일반적으로 2가지로 구분할 수 있다. 하나는 수집 대상에 대한 정보가 설정된 로봇이나 웹 크롤러에 의한 자동 수집이고 다른 하나는 자원의 생산자 또는 대행자 등의 신청인에 의해 직접 아카이빙에 온라인 디지털 자원을 전송하는 방법이다. 납본의 주체는 이처럼 자원을 수집하는 방법과 그것을 실행하는 주체에 따라 아래와 같이 구분하였으며 각 납본주체의 기능은 다음과 같다.

- ① 웹크롤러: 수집하고자 하는 온라인 디지털 자원의 URL을 입력하여 그 웹사이트의 내용물, 즉 웹 페이지들을 모아오는 도구이다.
- ② 납본자: 자료를 등록하는 주체이며, 납본자의 유형으로는 출판사, 대행사, 개인, 단체로 구분할 수 있다.

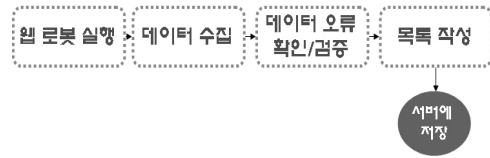
나. 디지털 납본의 납본 주체에 따른 과정

1) 웹크롤러

웹크롤러가 납본의 주체인 경우는 다음의

<그림 4>와 같은 프로세스를 갖게 된다.

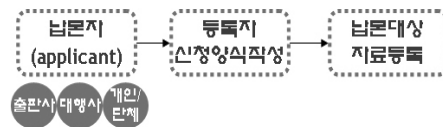
웹크롤러에 웹 아카이빙에서 수집 대상으로 선정된 자원의 웹주소(URL), 수집의 깊이(depth), 수집 파일의 형식(format) 및 수집 주기 등을 설정하면 그에 따라 해당 웹사이트에서 자원을 수집하여 가져온다. 수집된 자원은 간단한 오류 검사를 거쳐, 자동 수집된 메타데이터를 정렬하고 서버에 임시로 저장(또는 업로드)된다.



<그림 4> 웹크롤러의 납본 과정

2) 납본자

온라인 디지털 자원의 저작자(개인), 저작단체(출판사), 저작권을 위임받은 신탁단체(대행사)가 납본의 주체인 경우는 다음의 <그림 5>의 과정이 전개된다. 자원을 납본하고자 하는 납본자는 소정의 신청양식에 따라 신청서를 작성한 후, 납본 대상 자료를 시스템에 등록하면 된다. 이때 납본자의 유형은 디지털자료 생산 주체에 따라 추가 될 수도 있다.



<그림 5> 납본자의 납본 과정

4.3 수집 단계(2단계)

2단계는 웹크롤러 또는 납본자에 의해 등록 된 디지털자원이 네트워크를 통해 납본 시스템에 도착하는 과정으로 다음의 <그림 6>과 같다. 수집 단계에서는 웹크롤러가 수집해 온 파일 또는 스냅샷 한 이미지 파일 그리고 납본인이 등록한 자원이 기존의 수집 자원과 중복되지 않는지를 검사하고, 자동 수집 또는 제출된 목록의 진본성을 확인하며, 바이러스 검사를 통해 자원의 품질을 확인한다. 마지막으로 수집 자원에 관련된 시스템의 환경데이터 확인이 이루어 지는데 각각의 상세한 기능은 다음과 같다.



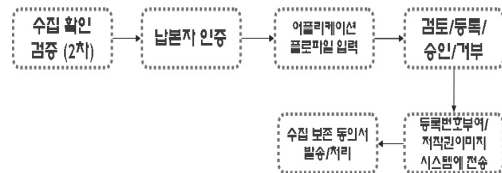
<그림 6> 자원수집 과정

- 1) 중복검사: 기존 시스템에 등록된 데이터와 중복되는 데이터가 있는지 확인한다.
- 2) 목록의 진본성 확인: 납본시스템에 등록된 납본자료와 납본자가 등록한 기본 서지정보가 일치하는지 확인한다.
- 3) 품질 확인: 바이러스 감염여부 확인(백신 프로그램으로 검증절차) 한다.
- 4) 시스템 환경 데이터 확인: 수집된 자원의 컨텐츠 데이터 및 그 데이터를 재생시킬 수 있는 시스템 환경(system configuration)과 관련된 데이터를 확인한다.

4.4 등록단계(3단계)

3단계는 수집단계와 연결되어 순차적으로 진행되는데 수집된 데이터와 시스템 환경데이터가 함께 납본 패키지로 구성되어 시스템에 등록된다.

등록에서의 과정은 <그림 7>과 같이 나타나며 상세 기능은 다음과 같다.



<그림 7> 납본자원 등록 과정

- 1) 수집 확인: 등록에 앞서 수집단계에서 수집 전에 이루어진 진본성 확인, 품질확인, 납본패키지전송, 중복검사의 결과를 다시 확인한다.
- 2) 신청인 인증: 인터넷을 통해 신청서와 예치자료를 받아, 출처를 입증하는 디지털서명을 유효화 하고, 신청인에게 신청서 도착과 전자적인 등록·기록이 완전하게 되었음을 통보한다.
- 3) 대상에 따른 입력: 지정된 어플리케이션 프로파일을 입력한다.
- 4) 시범: 등록된 저작물에 등록일을 부여하고, 처리중인 신청서와 예치자료를 저장한다. 검토 담당자는 등록된 저작물의 저작물성 여부와 다른 요구사항이 필요한지를 검토하여, 그 결과가 요건을 충족한다고 판단되면 등록일을 유효화 하여 등록을 승인하고, 요건을 충족하지 않을 경우에는

- 등록을 거부하고 거부사유를 신청인에게 전자우편으로 통지한다.
- 승인: 등록이 승인된 저작물에 등록번호를 부여하고, 이용자에게 증서를 교부하고, 데이터는 저작권 이미지 시스템에 전송한다.
 - 저작자의 웹자원 수집·보존 동의: 저작자에게 E-mail로 동의 요청 내용 발송, 온라인으로 수집 보존 동의 처리할 수 있도록 화면을 제공한다. 저작자의 동의를 얻기 위한 수집보존 허용여부 및 범위, 서비스 동의기간, 저작자 개인정보 등의 항목에 체크한다.

4.5 인증 및 진본성검증 단계(4단계)

4단계는 데이터를 검증하고 인증하는 과정으로 여기서는 납본 패키지의 유용성 확인과 논리적 무결성 확인 그리고 진본성 검증의 작업이 이루어지며, 실행 및 테스트 과정을 거친다. 각 과정의 기능은 다음의 <그림 8>과 같다.



<그림 8> 인증 및 진본성 검증의 과정

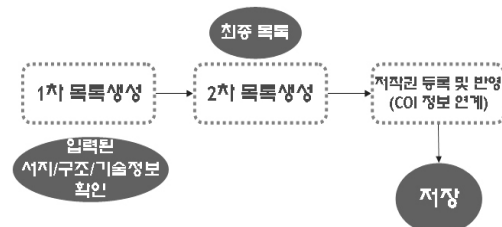
- 패키지무결성확인: 패키지의 물리적 통합성(무결성)을 확인한다. 데이터 파일과 시스템 환경 정보가 기록된 파일의 수가

일치하는지 검사한다.

- 논리적 통합 확인: 납본 패키지를 해제 후 모든 구성요소들이 존재하는지 확인한다.
- 진본부여: 납본에 대한 진본성 검증을 하며, 진본확인을 한다.
- 실행 및 테스트: 첨부된 파일을 실행해보거나 테스트를 하여 통과 시 다음단계인 저장단계로 진행되게 되며, 통과하지 못할 시에는 심의를 거쳐 자료에 대한 추가 데이터를 제출할 것을 요구하거나 등록을 무효화한다.

4.6 목록생성단계(5단계)

5단계에서는 납본된 자원의 최종 목록을 생성하며, 생성된 목록자원은 이용자들의 자원 검색과 브라우징에 활용된다. 수집단계에서 납본자에 의해 입력된 기본 서지 정보인 목록의 내용을 확인한 후, 메타정보를 추가적으로 입력하는 추가목록의 작업이 이루어진 후에, 최종 목록을 작성하게 된다. 그 다음 저작권 등록 및 반영 과정을 거쳐 저장단계로 진행된다. <그림 9>는 목록생성의 과정을 보여주며 다음과 같은 기능을 한다.



<그림 9> 목록 생성

- 1) 시스템에 메타데이터 등록: 대상 자료의 기본 서지정보, 기술정보, 관리정보 메타 데이터를 입력한다.
- 2) 등록 완료 후 식별자 부여: 자료를 등록하면 COI 식별체계 시스템과 연계하여 COI 번호를 부여 받는다
- 3) 목록레코드 작성 및 목록 데이터베이스에 전송: 기존의 자동화 시스템을 통해서 전자신청양식에서 목록레코드 작성을 위한 데이터를 전송하고, 목록레코드를 완성하여 목록데이터베이스에 전송한다. 목록 데이터베이스에 저장된 데이터는 검색과 저작물의 규명에 유용하게 사용한다.

된 웹자원을 납본 시스템의 저장소로 이동하는 과정이다. 이 단계에서는 새로운 저장매체로 정기적으로 매체 이전하거나, 안전을 위해 정기적으로 백업하는 작업이 이루어진다.

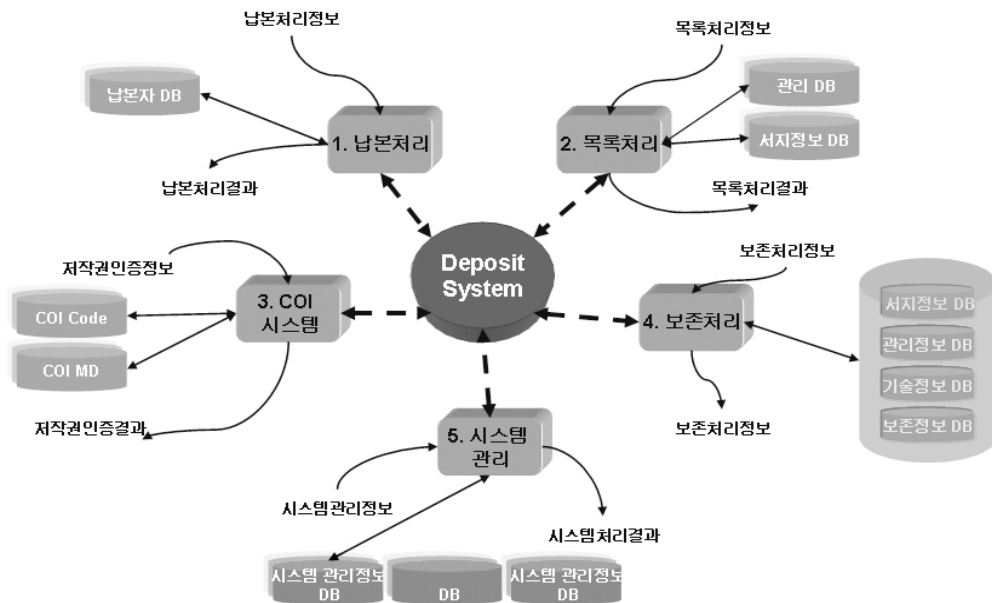
이때 납본된 자료의 훼손이나 유실에 대비하기 위한 보다 적극적인 보존전략을 수행하고자 한다면, 보존 단계로 진행하면 된다.

4.8 프로세스 관계도

다음 <그림 10>은 본 연구에서 제안한 납본 시스템 프로세스의 전체적인 관계도이다. 납본 시스템에서의 업무 처리 과정을 크게 납본처리와 목록처리, COI 시스템 연계, 보존 처리 그리고 시스템 관리로 구분하였으며, 각각의 단계별로 단위 업무를 선정하였다.

4.7 저장단계(6단계)

납본 프로세스의 마지막 단계(6단계)로 등록



<그림 10> 납본 시스템의 업무 처리 관계도

가. 납본 처리

납본 처리 과정의 순서는 다음과 같다.

- ① 납본자는 납본 등록을 하기 위한 우선 ID/PW를 신청한다.
- ② 납본자는 승인된 ID/PW로 시스템에 로그인 한다. 이때 납본자의 정보와 로그인 정보는 납본자 데이터 베이스(Database, 이하 DB)에 연계되며 이 두 정보를 제외한 납본 처리 단계에서의 모든 처리 정보는 관리 DB와 연계된다.
- ③ 로그인에 성공한 납본자는 납본자료를 등록하며 이때 등록 정보는 서지정보 DB, 구조정보 DB, 기술정보 DB 등과 연계된다.
- ④ 도서관자료제출서, 자료보상청구서, 계산서/세금계산서, 간이영수증 등의 납본을 위해 필요한 각종 관련 서류를 제출한다.
- ⑤ 저작권 동의서를 제출, 관련 서류 제출 과정과 저작권동의서 제출 과정의 관련 정보는 권리 정보 DB와 연계한다.
- ⑥ 등록된 자료에 대한 검수작업을 실시한다. 이 과정에서는 납본자가 등록해야 할 등록대상 자료(납본할 자료 및 기본 메타데이터)가 모두 등록 되었는지 확인하고, 납본자가 제출해야 할 제출대상 자료(관련 서류 및 저작권 동의서) 또한 모두 제출되었는지 확인한다. 이 두 과정을 거쳐 납본한 자료를 시스템에 등록할 것인지 등록하지 않을 것인지 결정한다.
- ⑦ 검수 과정에서 합격한 자료에 대한 등록을 인증한다.

나. 목록처리

목록 처리의 각 과정을 살펴보면 다음과 같

다. 납본 즉 등록이 결정된 자료에 대해 OASIS 장서 DB와 연계하여 중복 검사를 실시한다. 그리고, 중복 검사에서 기존 자료와 중복되지 않음을 확인 받은 자료에 한해 납본관리 DB와 연계하여 납본 처리 과정에서 납본자에 의해 입력된 기본 서지 정보의 내용이 실제 서지 정보와 일치하는지 검증한다. 그 후에, 검증에 통과된 자료에 대하여 메타정보를 추가적으로 입력하는 2차 목록 작성 과정을 거친다. 최종 목록이 작성되면 납본자료를 관리 DB와 OASIS 장서 DB에 등록 처리한다.

1) 제출 자료 검증

납본자가 입력한 목록과 실제 자료를 대조하는 목록 검증 과정을 거친 후, 바이러스의 감염 여부 등 파일의 품질을 검증한다. 이를 통해 자료의 무결성과 진본성을 확인하고 검증 결과를 처리한다. 목록 검증 과정은 구조정보DB, 기술정보DB, 서지정보DB와, 파일 검증 과정과 무결성 확인 과정은 기술정보 DB, 진본성 확인 과정은 기술정보DB와 서지정보 DB와 연계되며 마지막 검증 처리 과정은 관리DB와 연계된다.

2) 목록작성

납본 처리 시 납본자에 의해 입력된 기본 서지 정보의 내용이 실제 서지 정보와 일치하는지 확인하며 이 과정은 서지정보 DB, 기술정보 DB와 연계된다. 그 후에, 확인된 자료에 대하여 메타정보를 추가적으로 입력하여 추가목록을 작성하는데 서지정보 DB, 관리정보 DB, 보존정보 DB와 연계된다. 추가 목록이 작성되면 최종 목록이 완성된다. 이후에는 OASIS 장서 DB와 연계하여 관리한다.

다. COI 시스템 등록·연계 처리

납본 자료에 대해 연계된 COI 시스템에 등록을 신청하면 해당 자료를 검증하는 저작물 검수 과정을 거친 후에 COI 시스템에서 제공하는 어플리케이션에 따라 COI 메타데이터를 입력 및 수정한 후 확인과 검증을 거쳐 승인을 받으면 COI 식별체계가 부여된다. 이후에는 COI 식별체계 DB와 연계하여 지속적인 관리를 한다.

라. 보존처리

OASIS에 보존 중인 자료의 중복여부 및 보존처리 단계별 정보를 확인하여 보존 처리의 여부를 결정한다. 보존 처리하기로 결정된 자료에 대하여 보존 방법을 결정하고 보존 확인 및 저장 과정을 거친다. 이 과정은 관리 DB와 OASIS장서 DB와 연계되며 보존 방법 결정과 보존 확인 및 저장 과정은 보존용 MD DB(출처정보 DB, 참조정보 DB, 고정정보 DB, 문맥정보 DB)와 연계된다.

마. 시스템 관리

환경 설정 처리 과정을 통해 시스템 환경 설정을 확인 하고 파일 정보 DB에 연계되어 파일의 히스토리를 관리함과 동시에 백업 DB에서도 관리된다. 또한 이용자들의 해당 정보 검색을 위해 상세정보, 간략정보 등을 제공하는 검색 시스템의 관리도 진행된다.

5. 결론

본 연구는 디지털 자원의 웹 아카이빙을 목적으로 한 새로운 개념의 디지털 납본체계를 따르

는 프로세스를 제안 하였다. 납본제도는 “국가의 지적 문화유산의 수집 및 보존, 도서관 및 정보서비스를 위한 인프라 구축, 국가 차원의 망라적 정보 서비스를 제공, 최종 이용자를 위한 직접봉사”의 역할을 담당해왔다(윤희윤, 2003). 오늘날 국가의 지적 문화 유산은 온라인과 오프라인에 다양한 유형으로 존재하며, 특히 온라인에서만 이용 가능한 자원을 대상으로 하는 디지털 납본의 문제는 지속적인 논의와 준비를 해야 한다. 이에 이러한 디지털 납본 체계의 대상이 되는 디지털 자원이 갖는 특성과 관련 권리 관계 등 다양한 고려사항들을 충분히 검토 할 수 있는 과정이 필요하다. 또한 디지털 자원을 수집하여 등록하고 관리하는 방법과 절차가 기존의 납본체계와는 다르기에 각 단위절차에 대한 기능정의와 방법을 개발함으로써 영구적인 관리와 보존을 위한 요구사항들을 분석하여 새로운 구조를 제시한 것이다. 따라서 웹 아카이빙을 위한 디지털 자원 납본 프로세스의 결과물들은 다음과 같다.

첫째, 납본 범위, 납본대상 그리고 납본 주체를 정의하였다. 즉, 온라인 웹 아카이빙을 위한 저작물의 대상으로는 웹사이트, 웹문서, 디지털 자원으로 분류하여 정의하였고, 납본주체로는 자동수집에 의한 웹크롤러와 저작권자 또는 온라인 출판자인 신청인으로 주체를 정의하였다.

둘째, 디지털 납본 프로세스를 위한 정의와 그에 대한 기능을 명시 하였다. 세부적인 사항으로는 디지털 자원 납본을 위한 도서관 업무 프로세스 분석, 디지털 납본 프로세스 개형도 제시, 납본 시스템 프로세스 구성도와 각 단계별 디지털 납본 프로세스 정의하였다.

셋째, 디지털 납본의 국제적인 성공 프로젝트

결과물인 NEDLIB의 프로세스를 기본으로 국내 디지털 저작물에 대한 디지털 납본 프로세스의 기능을 적용하여 명시하였다.

넷째, 디지털 납본 프로세스에 대한 상관 관계도를 통해 디지털 자원 납본시스템을 위한 데이터 흐름도를 상세화 하여 단위별 업무의 기능 정의와 정보 흐름을 위한 디지털 납본 구조를 제시하였다.

본 연구를 통해 디지털 자원 납본 프로세스는

납본 업무의 범위와 납본 대상의 구체성 그리고 납본 주체의 명확한 정의를 통해 좀 더 구체화된 단계별 프로세스에 대한 정의와 기능을 향상시킬 것으로 기대한다. 디지털 자원 또한 문화유산으로서의 가치와 장기적인 보존의 대상이 될 것이다. 따라서 디지털 자원 보존을 위한 단계별 기능과 절차에 대한 연구가 지속적으로 이루어질 것을 기대한다.

참 고 문 헌

- 김유승. 2007. 웹 아카이빙의 법·제도적 문제에 대한 고찰. 『한국문헌정보학회지』, 41(3): 5-24.
- 김희정. 2003. 디지털 아카이빙 최근 연구동향 및 OASIS 참조모형에 관한 연구. 『한국기록관리학회지』, 3(1): 23-42.
- 서혜란. 2003. 디지털 자료의 납본과 보존을 위한 각 국가의 노력. 『한국정보관리학회지』, 20(1): 373-399.
- 서혜란. 2004. 웹 아카이빙의 성과와 과제. 『한국비블리아학회지』, 15(1): 5-22.
- 선명순. 2003. 한국의 온라인 전자출판물 납본제도 도입방안에 관한 연구. 석사학위논문. 연세대학교 행정대학원.
- 오영희, 최석두. 1996. 디지털도서관에서의 저작권 관리에 관한 사례연구: LC의 CORDS 프로젝트를 중심으로. 『한국정보관리학회 제3회 학술대회 논문집』 1996. 8: 111-114.
- 윤희윤. 2003. 한국의 납본제도 개선모형에 관한 연구. 『한국문헌정보학회지』, 37(4): 23-52.
- 이수상 등. 2004. 디지털 아카이빙의 현안과 과제. 『한국교육학술정보원 연구보고서』, 1-143.
- 이치주. 2007. 온라인 연속간행자료 수집 및 보존에 관한 연구. 『한국문헌정보학회지』, 41(2): 359-386.
- 최원태. 2000. 디지털 아카이빙(Digital Archiving)의 현황과 전망. 『한국문헌정보학회 학술발표 논문집』, 제9집: 59-82.
- OASIS 정보화계획 수립보고서. 2006. 국립중앙도서관.
- Andersen, Bjarne. 2006. "The DK-domain: in Words and Figures"
- Canada. Library and Archives of Canada. "Guidelines for Legal Deposit of Online Publications: Legal deposit procedures for online publications: general guidelines."
- CCSDS. 2002. Reference Model for an Open

- Archival Information System(OAIS).
CCSDS 650.0-B-1 BLUE BOOK.
E-Resources. <<http://www.nla.gov.au/apps/eresources/action/front>>.
- European Archive. <<http://www.europarchive.org/>>.
- Intute. <<http://www.intute.ac.uk/>>.
- Johan Steenbakkers. 2000. "The Deposit System for Electronic Publications: The NEDLIB Guidelines Koninklijke Bibliotheek.
- MINERVA. <<http://www.loc.gov/minerva/>>.
- National Library of Canada. 1996. Electronic Publications Pilot Project(EPPP) Final Report.
- National Library of Canada's Canadian Information By Subject service.
<<http://www.collectionscanada.ca/caninfo/eabout.htm>>.
- OASIS. <<http://www.OASIS.go.kr>>.
- PADI(Preserving Access to Digital Information). <<http://www.nla.gov.au/padi>>.
- PANDORA Archive.
<<http://pandora.nla.gov.au/>>.
- Titia van de Werf - Den Haag. 2000. "A Process Model: The Deposit System for Electronic Publications". Koninklijke Bibliotheek(NEDLIB Report series; 6). ISBN 90-62-59-1507.
- Titia van der Werf. 2000. "The Deposit System for Electronic Publications: A Process Model". Koninklijke Bibliotheek.
- UK Web Archiving Consortium.
<<http://www.webarchive.org.uk/>>.
- US Internet Archive.
<<http://www.archive.org/index.php>>.
- WARP. <<http://warp.ndl.go.jp/>>.