

다크 아카이브 운영 효율화를 위한 정보패키지 구축

Construction of Information Packages for the Operational Efficiency of Dark Archives

박 효 은 (Hyo Eun Park)*

이 승 민 (Seungmin Lee)**

목 차

- | | |
|---------------|-----------------------|
| 1. 서론 | 4. 다크 아카이빙 프로세스 구조 구축 |
| 2. 이론적 배경 | 5. 결론 |
| 3. 다크 아카이브 사례 | |

초 록

다크 아카이브를 통해 다양한 유형의 전자기록물을 장기적으로 보존하는 활동은 그 중요성이 점차 증대하고 있으나, 현재 다크 아카이브에서는 전자기록물 장기보존을 위한 최적화된 정보패키지 구조가 마련되어 있지 않다. 이에 본 연구에서는 다크 아카이브의 핵심적인 프로세스를 중심으로 OAIS 참조모델 정보패키지를 재구성하여 4개의 카테고리를 제안하였다. 각 카테고리의 세부 기술항목은 OAIS 참조모델, ISO 23081, 기록관리 메타데이터 표준, ISAD(G), ISAAR(CPF), ISDF, ISDIAH를 기반으로 총 4개의 상위요소와 27개의 하위요소로 구성하였다. 이는 다크 아카이빙에 최적화된 정보패키지 구성의 기반으로 활용될 수 있으며, 전자기록물의 장기 보존을 보다 효율적으로 지원할 수 있을 것으로 기대된다.

ABSTRACT

The importance of long-term preservation of various types of electronic records through dark archives is gradually increasing. However, the current dark archive does not have an optimized information package structure for long-term preservation of electronic records. In order to address these problems, this research proposed four element categories by re-organizing the OAIS reference model information package based on the core process of the dark archiving. The detailed descriptive items of each category consist of a total of 4 upper-level elements and 27 sub elements based on the OAIS reference model, ISO 23081, Records Management Metadata Standard, ISAD(G), ISAAR(CPF), ISDF, and ISDIAH. This structure can be used as a basis for constructing an information package optimized for dark archiving, and is expected to support the long-term preservation of electronic records more efficiently.

키워드: 다크 아카이브, 다크 아카이빙, 정보패키지, OAIS 참조모델, 보존 메타데이터
Dark Archives, Dark Archiving, Information Package, OAIS Reference Model,
Preservation Metadata

* 중앙대학교 일반대학원 문헌정보학과 기록관리학전공 석사과정(graphe314@gmail.com) (제1저자)
** 중앙대학교 사회과학대학 문헌정보학과 부교수(ableman@cau.ac.kr / ISNI 0000 0004 6418 7521) (교신저자)
논문접수일자: 2020년 10월 26일 최초심사일자: 2020년 11월 6일 게재확정일자: 2020년 11월 18일
한국문헌정보학회지, 54(4): 261-281, 2020. <http://dx.doi.org/10.4275/KSLIS.2020.54.4.261>

1. 서론

진화하는 정보기술에 기반한 정보환경의 정착과 함께, 현재 인쇄 기반 기록물 환경과는 비교할 수 없을 정도로 수많은 전자기록물이 생산되고 있다. 이들 전자기록물은 모두 고유한 가치를 지니고 있으며, 후대에 전승, 보존되어야 하는 정보적, 사회적, 문화적 유산이라고 볼 수 있다. 따라서 디지털 형태 기록물의 장기적인 보존은 향후의 지식정보사회에서 매우 중요한 자원으로 활용될 것이며, 이에 최근 여러 분야에서는 디지털 아카이브를 구축하여 생산되는 전자기록물, 디지털 기록물의 장기 보존에 노력을 기울이고 있다.

디지털 아카이브에서는 다양한 기록물의 기술을 위해 서지적 도구들을 적용하고 있으며, 국내외 기술 표준을 기반으로 기록물에 대한 메타데이터를 구축하고 있다. 현재 많은 디지털 아카이브에서는 메타데이터의 구축을 위해 OAIS 참조모델(OAIS Reference Model)을 적용하고 있으며, 이는 정보패키지(information package)를 기반으로 운영되는 시스템으로 설명할 수 있다. OAIS 참조모델의 정보패키지는 장기간에 걸쳐 기록물을 보존하기 위해 입수, 보존, 접근 및 활용 등의 프로세스에 기반을 두고 구축되고 있으며, 기록물의 보존을 위한 구조와 기능을 제시하고 있다. 이는 기록물 보존을 위한 시스템이나 서비스를 구축하기 위한 기반을 제공해 주는 정보 모델링(information modeling)으로서의 의미를 지니고 있으며, 전세계적으로 널리 적용되고 있다(이승민 2015).

현재 디지털 아카이브는 기록물의 보존과 활용이라는 두 가지 측면을 모두 반영하고 있으

나, 디지털 형태의 기록물이라는 특성으로 인해 디지털 아카이브는 다양한 상황에서 발생하는 재난에 매우 취약하다는 문제를 지니고 있다. 이러한 문제에 대처하기 위해 디지털 아카이브와는 다른 차원에서의 아카이빙이 이루어지고 있으며, 이는 다크 아카이브의 형태로 구축이 되고 있다. 다크 아카이브(dark archive)는 기록물의 원 제공기관에서 기록물 서비스를 지속할 수 없는 상황이 발생하는 것에 대비하여 기록물을 비공개로 보존하는 비공개 아카이브를 의미한다. 즉, 다크 아카이브는 중요하고 방대한 양의 전자기록물을 보호하고 보존하는데 집중하기 위한 목적으로 구축되는 아카이브로 설명할 수 있다.

방대한 양의 전자기록물이 저장, 보존되는 다크 아카이브는 기존의 아카이브와 달리 기록물의 활용에 중점을 두지는 않고 있다. 장기보존 및 보호에 집중한다는 특성을 지니고 있다. 다크 아카이브는 기록물의 장기보존 및 보호에 집중하고 있기 때문에 기록물이 군 단위로 저장되는 방식을 취하고 있다. 다만 저장되는 기록물 군의 보존을 위한 기술 방식에서는 기록물 건 또는 철 단위로 기술이 이루어지는 경향을 보이고 있다. 다크 아카이브가 지닌 이러한 특성으로 인해, 다크 아카이브 프로세스에 기존 디지털 아카이브의 모델이나 구조를 적용하는데 있어서는 많은 한계가 발생하고 있다. 하지만 디지털 아카이브와 다크 아카이브는 활용과 보존 가운데 어느 측면에 중점을 두는 지에서만 차이를 보이고 있을 뿐, 모두 기록물을 대상으로 하는 공통점을 지니고 있다. 따라서 다크 아카이브의 효율적인 운용을 위해서는 기존의 디지털 아카이빙의 기반 위에서 다크 아카

이빙에 최적화된 아카이빙 프로세스를 마련할 필요가 있다.

이에 본 연구에서는 디지털 아카이빙의 정보 모델링을 중심으로 다크 아카이빙에 적용할 수 있는 정보패키지 구조를 제안하고자 한다. 이는 기존 OAIS 참조모델을 기반으로 하고 있으며, 다크 아카이빙에 적용할 수 있는 구조 및 기술항목을 재구성하고 다크 아카이브의 특성에 맞는 기술항목을 반영한 정보패키지의 구조를 제안하고자 한다.

2. 이론적 배경

2.1 다크 아카이브의 정의와 특성

SAA 용어사전(Dictionary of Archives Terminology)에서는 다크 아카이브(dark archives)를 “향후 사용을 위해 기록정보자원을 저장하지만 관리자만 접근할 수 있는 저장소 혹은 향후 사용을 위해 보존하지만 현재 접근할 수 없는 기록물 컬렉션”으로 정의하고 있다(Society of American Archivists 2020). 국내에서는 국립중앙도서관에서 SAA 용어사전의 다크 아카이브 정의를 번역한 것 이외에 국가기록원이나 법률 및 표준에서 공식적으로 다크 아카이브를 정의하지는 않고 있다. 하지만 일반적으로 다크 아카이브는 기록물의 원 제공기관에서 서비스를 더 이상 제공할 수 없는 상황이 발생할 것에 대비하여 대량의 기록물을 비공개로 보존하는 것을 의미한다. 기록물을 비공개로 보존한다는 측면에서 ‘비공개 아카이브’, ‘보존용 아카이브’, 혹은 ‘어두운 아카이브’라는 용어를 사용

하기도 한다. 이에 본 연구에서는 SAA 용어사전에서의 정의를 기반으로, ‘향후의 활용을 위해 보존되었지만, 현재 일반 대중이 접근할 수 없는 비공개 아카이브’로 다크 아카이브를 정의하고자 하며, 다크 아카이브와 관련한 여러 가지 명칭들을 ‘다크 아카이브’로 통일하여 사용하고자 한다.

다크 아카이브는 일반적인 디지털 아카이브와는 다른 의미를 지니고 있으며, 그 수행하는 역할에 있어서도 디지털 아카이브와는 차이를 보이고 있다. 이에 다크 아카이브가 지닌 고유한 특징을 종합하여 정리하면 다음과 같다.

첫째, 다크 아카이브는 기록물 군을 단위로 기록물을 보존한다. 일반적인 아카이브가 기록물 철 또는 건 단위로 기록물을 관리하는 것과 달리, 기록물 군을 기본 단위로 하고 있기 때문에 장기간의 보존에도 기록물에 대한 기능 및 맥락 파악이 용이하다는 특성을 지니고 있다.

둘째, 다크 아카이브는 복본을 생산하여 보존한다. 다크 아카이브는 수많은 복본을 생산하여 각각의 아카이브에 저장해 기록물을 보관한다. 이는 ‘복사본이 많을수록 안전하다’는 Lots of Copies Keep Stuff Safe(LOCKSS)의 전략에 따른 것이다. 복본을 다크 아카이브에 분산 저장하기 때문에 자연 재해, 기술적 문제 등으로 인해 기록이 손실되더라도 나머지 다크 아카이브들의 기록을 비교 분석하여 문제가 발생한 아카이브를 백업하고 복구하는 역할을 수행해 기록물에 대한 안전한 보존이 가능하다.

셋째, 다크 아카이브는 권한을 가진 관리자만 접근이 가능하다. 일반 대중에게 공개되어 수집 및 활용이 이루어지는 기존의 아카이브와 달리, 다크 아카이브는 관리 권한이 승인된 관

리지만 접근이 가능하다. 이러한 접근의 통제 는 기록물의 원활한 보존을 효율적으로 지원할 수 있게 된다.

넷째, 다크 아카이브는 공공의 이익을 보장 한다. 현재는 접근 권한이 있는 관리자만 접근 할 수 있는 아카이브이지만, 기록물이 손실되 었을 때 이용자의 접근을 허용해 해당 기록물 의 이용가능성을 보장함으로써 이용자의 권리 보호의 역할을 수행할 수 있다. 또한, 기록물의 이용가능성을 지속적으로 보장함으로써 정부 기관이나 학계에서 참고 혹은 연구 목적으로 이용될 수 있다.

이와 같이, 일반적인 아카이브 및 디지털 아 카이브가 기록물의 관리와 활용에 중점을 두는 것과는 다르게, 다크 아카이브는 기록물의 장 기적인 접근 가능성 및 상호운용성을 보장하고 지속적인 이용가능성 보장을 하는 아카이브이 며, 영구적인 보존을 위한 기능이 강조되는 아 카이브이다(Caplan 2005). 따라서 기록물의 장 기보존을 통한 기록물의 정보적, 사회적, 문화 적 가치 전승을 위해 다크 아카이브 구축 및 운 영의 중요성에 대한 인식이 점차 높아지고 있 으며, 다크 아카이브 자체 또한 디지털 아카이 브와는 다른 고유한 가치를 지니고 있다.

이를 종합해 보면, 다크 아카이브는 무수한 복본 생산을 통한 분산 저장과 기록물 군 단위 로 보존하는 특징을 지니며, 중요하고 방대한 양의 전자기록물을 보호하고 보존하는 데 집중 하기 위한 목적으로 구축된다. 하지만 다크 아 카이브에 보존되는 전자기록물의 진본성 등의 품질을 보장하고 기록의 증거적 가치를 효과적 으로 보존하는 프로세스를 마련하기 위해서는 다크 아카이브의 특성을 반영한 정보 모델링의

구축이 필요하다는 요구가 대두되고 있다.

2.2 기록물 기술표준의 개념

지난 20년간 International Council on Archives (ICA)는 다중개체 구조의 기록물 기술을 위해 서 4개의 기술표준인 ISAD(G), ISAAR(CPF), ISDF, ISDIAH를 개발, 발표하였다. 이 표준들 은 국제적·국가적으로 일관성 있는 표준 형식 에 따라 기록물을 기술하기 위해 마련되었으며, 여러 국가들이 이들 표준을 각국의 국가 기술 표준으로 채택함에 따라 기록관리 발전에 크게 기여하였다는 의의를 지니고 있다.

국제 기록물 기술규칙인 International Stand-ard Archival Description (General) (ISAD(G)) 은 일관성 있고 체계적인 기술 목록의 생산과 통합을 원활히 수행하기 위하여 ICA가 제정한 표준이다. ISAD(G) 제2판에서는 7개 영역과 26개의 기록물 기술요소를 제시하고 있으며, 필 수 요소로는 참조코드, 제목, 일자, 기술계층, 생 산자명, 범위와 내용 등 6가지가 지정되어 있다 (한국기록학회 2008).

국제 전거레코드 기술표준인 International Standard Archival Authority Record for Corporate Bodies, Persons and Families (ISSAR(CPF))는 전거레코드의 구성 및 요 소 기술을 위한 국제 표준이다. ISSAR(CPF) 제2판에서는 식별영역, 기술영역, 관계영역, 통 제영역 등의 4개의 영역과 27개의 세부요소를 제시하고 있다(김성희 2005). 필수 요소는 개 체 유형, 전거명, 존속 기간, 전거레코드 식별자 등 4가지 요소가 지정되었다.

국제 기능 기술표준인 International Standard

for Describing Functions(ISDF)는 기록물 생산 단계에서의 기능에 중점을 둔 표준으로, 식별영역, 맥락영역, 관계영역, 제어영역 등의 4개 영역과 22개의 요소를 제시하고 있다. 필수 요소로는 유형, 인가 형식의 명칭, 기능 기술 식별기호 등 3가지 요소가 지정되어 있다.

국제 기록소장기관 기술표준인 International Standard for Describing Institutions with Archival Holdings(ISDIAH)는 기록물을 관리하고 일반 대중에게 제공하는 기록물 소장기관에 대한 기술사항을 일관성 있게 생성하기 위해 마련된 별도의 표준이다. ISDIAH에서는 식별영역, 연락영역, 기술영역, 이용영역, 서비스영역, 제어영역 등의 6개 영역과 31개의 요소를 제시하고 있으며, 필수 요소로는 식별기호, 전거명, 위치 및 주소 등의 3가지 요소를 제시하고 있다(이혜영 2020).

ISAD(G), ISAAR(CPF), ISDF, ISDIAH는 기록물을 표준적인 형식으로 기술하기 위해 국제적으로 정한 표준으로, 전 세계 대부분의 국가들은 이 표준들을 기반으로 하여 기록물을 기술하고 있다. 이들 표준은 다양한 유형의 기록물과 관련된 사항들을 일관성 있는 방식으로 기술할 수 있는 환경을 제공해 주고 있으며, 기록물을 관리, 보존, 활용하는 아카이브에서도 이들 표준을 기반으로 하여 기록물 기술사항을 생성하고 있다(Perrin, Winkler and Yang 2015). 따라서 디지털 아카이빙뿐만 아니라 다크 아카이빙에서도 이들 기술 표준을 준수하게 되면 거시적인 차원에서의 기록물 기술사항의 일관성을 확보할 수 있는 환경을 조성할 수 있다. 이에 본 연구에서는 다크 아카이브 프로세스에서 기록물 기술을 위한 정보패키지의 구조 구

축에 이들 표준을 적용하고자 하며, 이를 통해 다크 아카이브의 상호운용성의 수준을 높이고자 한다.

2.3 기록관리 메타데이터의 개념

현재와 같이 기록의 내용, 맥락, 구조가 각각 분리되어 있는 디지털 환경에서 기록의 4대 속성을 보장하고 기록을 체계적으로 관리하고 보존하기 위한 방법으로 메타데이터의 중요성이 지속적으로 대두되고 있다. 이에 기록관리 메타데이터에 대한 논의가 지속적으로 이루어지고 있는데, 기록관리 메타데이터는 기록의 증거적 가치를 지원하고, 기록물의 4대 속성인 진본성, 무결성, 신뢰성 및 이용가능성을 보장하며, 기록에 대한 이해를 촉진하는 등 기록관리의 전반적인 과정을 지원하는 역할을 한다(임진희 2013).

기록관리 메타데이터의 표준화를 위해 ISO 15489가 발표되었으며, 현재 ISO 15489-1: 2016이 적용되고 있다. 하지만 이는 기록관리 메타데이터에 대한 지침을 직접적으로 제시하지 못하고 있다는 한계를 보이고 있으며, 이에 ISO 15489를 기반으로 메타데이터를 실행하고 사용하기 위한 지침으로 ISO 23081이 제정되었다(Caplan 2007). 기록관리 메타데이터 표준인 ISO 23081은 현재 기록관리 메타데이터의 근간이 되는 국제표준으로 인식되고 있다. 이 가운데 ISO 23081-1은 ISO 15489를 기반으로 메타데이터를 실행하고 사용하기 위한 지침으로 제정되었으며, ISO 23081-2는 ISO 23081-1에서 제시한 원칙과 실행의 고려사항을 준수하기 위해 관련된 메타데이터 요소를 정의하고 있다

(Formenton et al. 2017)

국내에서는 ISO 23081를 기반으로 KS X ISO 23081-1(2007)과 KS X ISO/TS 23081-2(2008/2010)을 국가표준으로 제정하였다. KS X ISO 23081의 개념 모델에는 기록 자체에 대한 업무규정이나 정책과 범규에 관한 메타데이터, 행위주체 관한 메타데이터, 업무활동이나 과정에 관한 메타데이터, 기록관리 과정에 관한 메타데이터 개체 등이 주요 개체 유형으로 제시되어 있다.

이와 함께, 2007년에 국가기록원은 국내 기록관리 메타데이터 공공 표준인 NAK/S 8:2016(v2.1)을 제정하였다. 이 표준은 기록물의 4대 속성을 보장하기 위해 공공기관이 생산 혹은 접수하는 기록물에 대한 기술요소 및 기록관리 생애주기 전반에서 기록물 관리기관이 관리해야 할 사항을 일관성 있는 방식으로 기술하기 위해 제정되었다. 이를 통해 공공기관의 기록관리 메타데이터 요소를 표준화하고 메타데이터의 이관, 메타데이터 통합 및 연계 등을 지원하고 있으며, 기록물 중심의 단일 개체 모형을 바탕으로 기록물 기술요소를 제시하고 있다. 이 표준에 수록된 기록관리 메타데이터는 23개 상위요소, 66개 하위요소, 58개 세부요소의 3계층으로 구성되어 있으며, 하위요소 중 세트로 작성되어야 할 요소들은 컨테이너로 묶어 하위요소-세부요소로 계층화되어 있다. 현재 우리나라 공공기관의 현용 및 준현용 기록물은 국가기록원이

제정한 기록관리 메타데이터 표준인 NAK/S 8:2016(v2.1)에 의거해 관리되고 있다.

2.4 OAIS 참조모델의 개념과 구조

OAIS 참조모델(Reference Model for an Open Archival Information System)은 디지털 정보를 영구적으로 보존하는 디지털 아카이브를 위한 국제표준이다. 1999년 초안이 발표된 이후, 2002년에는 국제표준인 ISO 14721로 확정되었다. 2012년에 제2판이 발행되었으며, 2019년에 제3판 초안이 발표되었다. OAIS 제3판 초안에서는 디지털 아카이브 내의 모든 정보 객체가 내용정보(Content Information: CI) 역할을 할 수 있다는 점을 강조하면서, 내용정보와 보존기술정보(Preservation Description Information: PDI) 사이의 관계를 명확히 정립하였다. 또한 데이터 객체에서 정보객체로 해석이 되는 과정에서의 일관성을 보장하기 위해 정보패키지(information package)의 정의를 보완하였다(〈그림 1〉 참조).

OAIS 참조모델에서의 정보는 〈그림 1〉에서 도식한 과정을 거쳐 해석되는 정보 객체(Information Object)를 기반으로 하는 정보패키지 형태로 존재한다(임진희 2013). 데이터 객체는 표현 정보(Representation Information)를 통해 해석되며, 표현 정보는 정보 객체를 산출한다. 정보 객체를 보존하기 위해서는 데이터 객



〈그림 1〉 데이터객체에서 정보 객체로 해석되는 과정(CCSDS 2019)

체와 관련 표현 정보를 명확하게 식별하고 이해하는 것이 매우 중요한 의미를 지니게 된다 (Beagrie 2004).

이러한 내부적인 프로세스를 기반으로 하여, OAIS 참조모델에서는 아카이브의 기능을 정보의 입수, 아카이브 저장, 데이터 관리, 운영, 보존계획, 접근 등의 6가지 영역으로 정의하고 있다. 이들 각 기능영역들 사이의 상호관계를 도식하면 다음 <그림 2>와 같다.

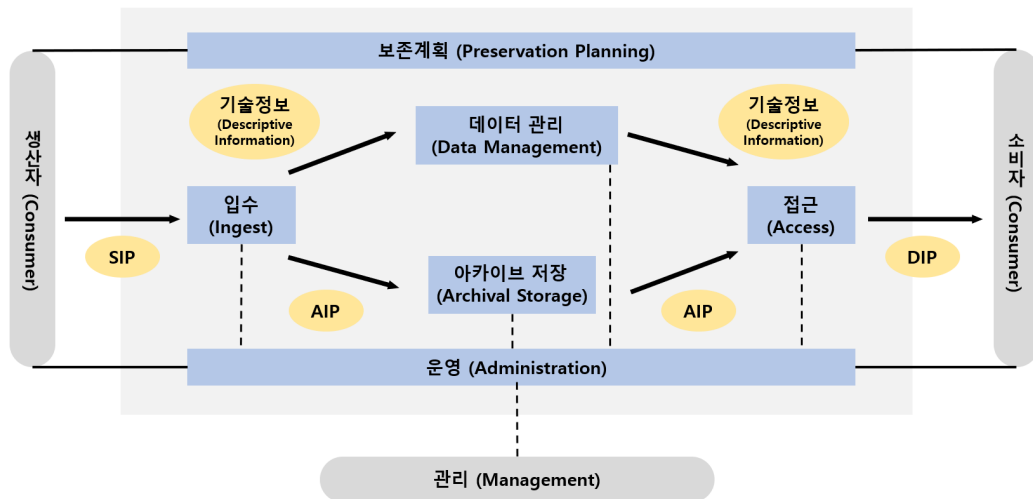
<그림 2>에서 제시한 각각의 영역에서는 보존의 대상이 되는 기록물을 각 보존 단계의 목적에 맞게 기술함으로써 정보자원의 조직 및 관리, 운영, 보존 활동을 지원하게 된다. 여기에는 디지털 아카이브의 구축 및 운영을 위한 필수적인 과정들이 모두 포함되며, Submission Information Package(SIP), Archival Information Package(AIP), Dissemination Information Package(DIP) 등의 정보패키지를 이용하여 디지털 아카이빙의 각 단계에 맞는 기술사항을 생

성하고 있다(이승민 2017, 133-134).

2.5 정보패키지의 구성과 유형

정보패키지는 기록물의 장기보존을 지원하는 개념적 구조이며, 정보 객체(Information Object), 패키징 정보(Packaging Information), 보존기술정보(Preservation Description Information: PDI)로 구성되어 있다. 이외에도 정보패키지는 정보 객체를 위한 컨테이너로 패키징 정보 및 패키지 기술과 연결이 되어 있다.

이러한 구조를 통해서, 정보패키지는 생산자로부터 OAIS 참조모델로 필요한 정보를 전송하거나 사용자가 요청한 정보를 전송할 때 사용될 수 있으며, OAIS 참조모델 내에서의 구조 및 저장에도 사용될 수 있다. 각각의 정보패키지에는 각기 다른 기능이 존재하며, 정보패키지 별로 필수적으로 포함되는 정보와 클래스 간의 연결이 다양하기 때문에 해당 정보패키지들을



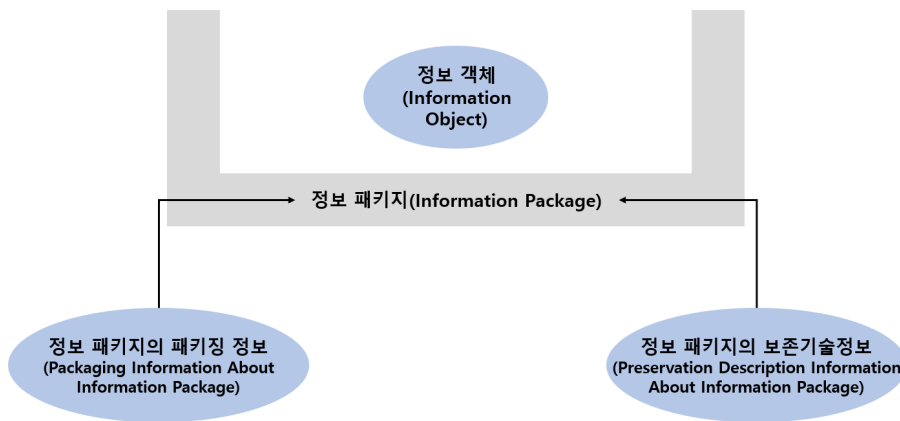
<그림 2> OAIS 참조모델 기능 엔티티(CCSDS 2019)

구별하는 것이 중요하다. 이와 같은 정보패키지의 구성을 살펴보면 다음 <그림 3>과 같다.

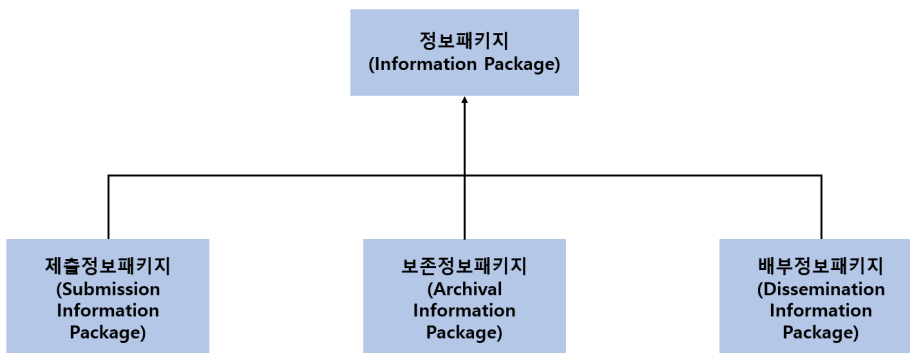
전통적인 아카이브에서 내용정보(CI)를 이해하기 위한 기술규칙이 마련되었던 것과 마찬가지로, 전자기록물의 장기보존을 위해서는 보존기술정보(PDI)가 반드시 필요하다(CCSDS 2002). 보존기술정보는 내용정보에 대한 보존 메타데이터의 역할을 수행하며, 출처정보, 맥락정보, 인증정보, 참조정보, 접근권한정보 등의 세부요소로 구성되어 있다. 이외에도 보존기술

정보는 내용정보가 임의로 변형되지 않도록 보호하고 지원하는 역할 또한 수행한다. 특히 전자기록물의 경우 매체와 환경에 대한 의존도가 높기 때문에 보존기술정보의 구체적인 요소들을 통하여 맥락정보를 파악할 수 있다.

OAIS 참조모델에서는 정보패키지를 Submission Information Package(SIP), Archival Information Package(AIP), Dissemination Information Package(DIP)의 3가지 유형으로 구분하고 있다(<그림 4> 참조).



<그림 3> 정보패키지의 개념과 관계(CCSDS 2019)



<그림 4> 정보패키지의 유형(CCSDS 2019)

OAIS 참조모델에서의 정보패키지 중 보존 정보패키지(AIP)는 내용정보와 보존기술정보로 구성된 정보패키지로서 OAIS 참조모델 내에 보존되어 있다. 하지만 현재의 OAIS 참조모델 기반의 정보 모델링 방식에서는 개념적인 측면으로 정보패키지의 구조가 정의되어 있으며, 다양한 유형의 기록물 보존, 관리, 활용을 지원하기 위한 일련의 프로세스로서의 구조적인 환경만을 제공해 주고 있다. 따라서 현재의 OAIS 참조모델 기반의 프로세스는 다크 아카이빙을 위한 프로세스로 적용되는데 있어서는 많은 한계를 보이고 있으며, 다크 아카이빙의 기능과 역할을 명확하게 반영하기 어려운 구조라고 볼 수 있다.

다크 아카이브는 기록물의 영구적인 보존이 가능한 디지털 아카이브로 설명할 수 있다. 따라서 다크 아카이빙을 위한 프로세스는 현행 OAIS 참조모델의 개념적 구조에 기반을 두고 운영되는 것이 바람직하며, 다크 아카이빙의 고유한 특성과 기능을 충분히 지원하기 위해서는 OAIS 참조모델의 개념적 구조와 기록물의 장기보존에 필요한 모든 정보를 포함하는 AIP를 기반으로 하여 기록물의 보존에 필요한 정보패키지의 구조를 구축하는 것이 하나의 효율적인 방안이 될 수 있다.

이에 본 연구에서는 국제적으로 사용되는 기록물 기술표준, 즉 ISAD(G), ISAAR(CPF), ISDF, ISDIAH를 통해 기존 아카이브의 기술요소들을 분석하고, 이를 OAIS 참조모델에 적용하여 다크 아카이브를 위한 기술항목을 도출하고자 한다. 이를 기반으로 다크 아카이빙에 최적화된 정보패키지 구조를 제안하고자 한다.

3. 다크 아카이브 사례

전자기록물의 가치에 대한 인식이 높아지면서 다크 아카이브의 구축 및 운영에 대한 중요성 또한 점차 증가하고 있다. 하지만 다크 아카이브를 실제로 구축하고 운영하고 있는 사례에 대한 연구 혹은 다크 아카이브에 대한 논의는 충분하게 이루어지지 않고 있는 실정이다. 이 가운데 국내외의 대표적인 다크 아카이브 사례를 살펴보면 다음과 같다.

국내의 사례 가운데 하나로, 국립중앙도서관은 국내 학술지 전자저널의 장기보존을 위해서 다크 아카이브를 구축하였다. 다크 아카이브 구축을 위해 누리미디어, 한국학술정보, 학술교육원, 교보문고, 학지사, 코리아스칼라 등의 전자저널 유통사와 전자저널 보존을 위한 업무협약을 체결하였으며, 총 400만여 건에 달하는 학술 논문을 아카이빙하였다.

다크 아카이브는 다양한 유형의 전자기록물을 최적화된 메타데이터를 이용해 영구히 보존·관리해야 한다. 하지만 현재 국립중앙도서관의 다크 아카이브는 전자저널 유통사와의 이해관계로 인해 외부와 물리적으로 차단된 별도의 보존 장치에서 암호화된 원본 파일을 저장하는 데 그치고 있으며, 추가적인 메타데이터를 통한 관리가 이루어지지 않는 상황이다. 다크 아카이브가 메타데이터 없이 보존만 된다면 영구적인 보존을 위한 관리가 이루어지기 어려우며, 다양한 환경에서 생산된 기록물에 대한 상호운용성을 보장하기 어렵다는 문제를 초래할 수 있다.

국내에서는 다크 아카이브가 구축, 운영되는 사례를 확인하기 어려운 실정이지만, 해외에서는 다크 아카이브와 관련한 일부 사례를 확인할

수 있다. 이 가운데 다크 아카이브의 대표적인 사례로 Controlled Lots of Copies Keep Stuff Safe(CLOCKSS)를 들 수 있다. CLOCKSS는 글로벌 연구 커뮤니티를 위해 지리적으로 분산된 지속가능한 다크 아카이브를 구축하는 것을 목적으로 학술출판사와 연구도서관이 공동으로 설립한 비영리단체이다. 분산된 인프라를 통해 관리되는 다크 아카이브는 기술적, 경제적, 환경적, 정치적 원인 등과 같은 다양한 위험으로부터 안전하게 디지털 기록을 보존할 수 있다. CLOCKSS는 웹 기반 학술 콘텐츠의 장기적인 보존을 위한 다크 아카이브를 운영하고 있으며, 현재 200개 이상의 출판사, 700관 이상의 도서관이 CLOCKSS의 학술정보자원 협동 보존 프로젝트에 참여하고 있다(Rosenthal 2014).

CLOCKSS 이외에도 일본에서 구축·운영하고 있는 PORTICO의 사례를 들 수 있다. 2018년 일본 과학기술진흥기구(Japan Science and Technology: JST)는 2,105종의 저널을 대상으로 J-STAGE 내에 다크 아카이브 PORTICO를 구축했다. J-STAGE는 일본 내 학회 및 협회 등이 발행하는 약 2,700종의 저널, 약 470만 편의 논문(2018년 기준)이 등재되어 있는 전자저널 출판·제공 웹사이트이다. PORTICO은 학술논문의 장기 보존과 안정적인 제공을 위해 구축되었으며, J-STAGE가 자연재해 등의 이유로 학술논문을 제공할 수 없게 될 경우를 대비하여 학술논문의 데이터를 수집부터 복구까지 본래의 사이트를 대신하여 제공해 주게 된다. 현재 PORTICO는 일본 학회 및 협회 등에서 발행하는 전자저널뿐만 아니라 전자책까지 보존하고 있다.

PORTICO의 내용 모델과 프로세스는 OAIS와

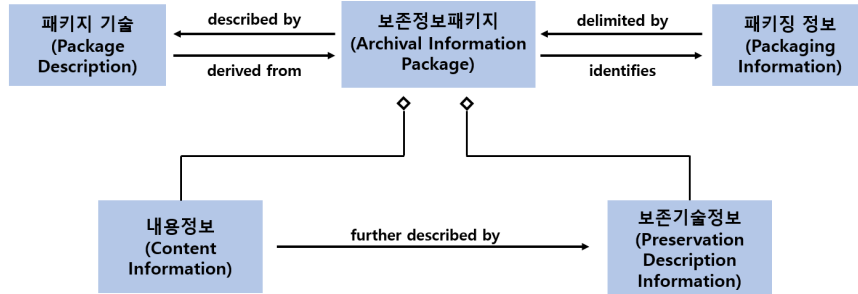
Metadata Encoding & Transmission Standard (METS) 표준을 기반으로 구축되었으며, 전자기록물의 이용가능성, 진본성, 발견가능성, 접근가능성을 보장한다. 또한 PORTICO에 보존되는 모든 기록물에 서지 메타데이터를 부여하여 언제든지 기록을 찾을 수 있는 환경을 구축하고 있으며, 관리자에게 보존된 기록과 그 상태를 검토할 수 있는 권한을 부여한다. 그리고 이용자들이 원본 기록을 이용할 수 없는 상황이 발생할 때, 학생, 직원 및 교수진을 대상으로 연구 및 교육 목적의 향후 접근 가능성을 보장한다.

4. 다크 아카이빙 프로세스 구조 구축

4.1 OAIS 참조모델 기반 보존정보패키지

OAIS 참조모델의 정보패키지는 정보 객체와 정보패키지에 대한 패키징 정보, 정보패키지의 보존과 설명에 필요한 메타데이터인 보존기술정보로 구성되어 있다. 정보 객체는 기록물의 원형인 내용정보와 보존기술정보로 구성되며, 보존기술정보는 다시 5가지 세부요소들로 이루어진다.

OAIS 참조모델의 보존정보패키지는 정보 객체의 영구 혹은 장기 보존에 필요한 정보를 기술한 패키지이다. 보존정보패키지에는 지정된 정보 객체가 존재하며, 이를 내용정보로 보기도 한다. <그림 5>에서 도식한 바와 같이, 보존정보패키지는 보존기술정보에서 추가적으로 설명하는 내용정보와 보존기술정보로 이루어지며, 패키지 기술이 보존정보패키지를 설명하



〈그림 5〉 보존정보패키지와 관련 패키지 기술 및 패키징 정보(CCSDS 2019)

고 패키징 정보에 의해 보존정보패키지가 설명된다. 보존정보패키지의 보존기술정보는 반드시 존재해야 한다는 점에서 일반 정보패키지와 차별화된다.

다크 아카이브는 OAIS 참조모델을 기반으로 정보패키지의 형태로 기록물을 보존·관리하는 것이 바람직하다. 하지만 OAIS 정보패키지는 일반 아카이브를 바탕으로 제안된 기술적 권고안이기 때문에 표준화된 구조가 미비한 다크 아카이브에 그대로 적용하는 데는 한계가 존재한다.

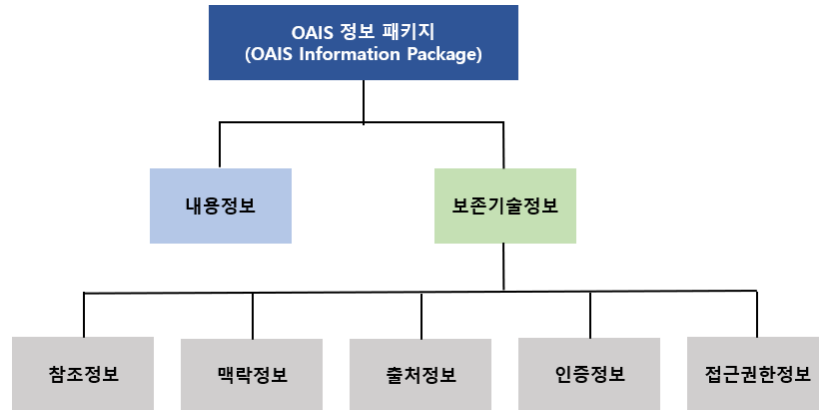
이에 본 연구에서는 여러 유형의 정보패키지 중 기록물의 영구적인 보존을 위한 정보패키지인 보존정보패키지를 기존의 구조에서 보완 및 확장하여 다크 아카이브에 특화된 정보패키지인 다크 아카이브 정보패키지(Dark Archive Information Package: DAIP) 구조 구축을 제안하고자 한다. 제안한 다크 아카이브 정보패키지를 통해 보존 환경 변화에 관계없이 전자 기록물을 장기간 보존할 수 있는 지속가능성을 보장할 수 있으며, 지속적인 관리를 바탕으로 전자기록물의 진본성 및 무결성을 유지함과 동시에, 마이그레이션, 에물레이션 등을 통해 새로운 기술변화에 쉽게 대응할 수 있는 유연성을 보장하고자 한다.

4.2 다크 아카이브를 위한 정보패키지 구조 구축

다크 아카이브는 기관 및 부서 단위로 기록물이 이관되어 저장되기 때문에 기록물군을 단위로 기록이 보존되며, 수많은 복본을 생산하여 분산 저장한다. 그리고 접근 권한에 따른 철저한 통제를 통해 기록물의 장기적인 보존을 지원한다. 이러한 특성을 지닌 다크 아카이브에 기존의 정보패키지 구조를 그대로 적용하여 활용하기에는 한계가 있으며, OAIS 참조모델에서도 정보패키지의 특정한 형식이나 구성 방법에 대해서는 제시하고 있지 않은 상황이다. 이에 다크 아카이브 정보패키지에서는 기존의 정보패키지 구조 중 AIP를 바탕으로 다크 아카이브 특성에 최적화된 정보패키지의 세부 요소들을 추가하고자 한다.

현재 OAIS 참조모델은 내용정보와 보존기술정보로 정보패키지가 구성되어 있으며, 보존기술정보는 참조정보, 맥락정보, 출처정보, 인증정보, 접근권한정보의 하위요소로 구성되어 있다(〈그림 6〉 참조).

〈그림 6〉에서 도식한 구조는 OAIS 참조모델의 기본적인 정보패키지 구조이며, 본 연구에서는 OAIS 참조모델의 정보패키지 구조에



〈그림 6〉 OAIS 참조모델의 정보패키지의 구조(CCSDS 2019)

다크 아카이브에 최적화된 기술요소를 보완하여 구축하고자 한다.

다크 아카이브는 이미 한 번 정리되어 보존하고 있는 기록물에 대한 보존을 진행하기 때문에 국제기록물 기술 표준 및 기록관리 데이터 표준에 따라 관리되어야 기존에 생산되어 보존되고 있는 기록물과의 상호운용성을 보존할 수 있다. 따라서 다크 아카이브 정보패키지도 입수되는 기록물의 속성 정보를 포괄할 수 있도록 설계되어야 한다.

이를 위해 OAIS의 정보패키지 구조에 기록

물 기술 표준 및 기록관리 메타데이터 표준에 정의된 요소들을 반영하였으며, 다크 아카이브 정보패키지 구성에 필요하다고 판단되는 요소를 추가하여 다크 아카이브에 보존되는 기록물을 기술할 수 있는 기술항목을 도출하였다. 기존 현용·준현용 기록에 부여된 메타데이터 요소는 국내 기록관리 메타데이터 표준(NAK 8:2016(v2.1))을 참고하였으며, 다크 아카이빙에 필요한 메타데이터 요소는 ISAD(G), ISAAR(CPF), ISDF, ISDIAH와 다크 아카이브 사례로부터 추출하였다(〈표 1〉 참조).

〈표 1〉 다크 아카이브 정보패키지의 세부 요소(안)

기록관리 메타데이터 표준(NAK 8:2016(v2.1)) 메타데이터 요소		다크 아카이브 정보패키지 메타데이터 요소			
상위요소	하위요소	상위요소	하위요소	추출한 요소의 출처	
1 생산자	1.1 생산자 유형	내용정보	참조코드	ISAD(G)	
	1.2 기관명			제목	ISAD(G)
	1.3 기관코드				등록 일자
	1.4 부서명		유형	ISAAR(CPF), ISDF	
	1.5 부서코드			포맷	
	1.6 개인명				
	1.7 개인코드				
	1.8 직위(직급)명				
	1.9 직위(직급)코드				

기록관리 메타데이터 표준(NAK 8:2016(v2.1)) 메타데이터 요소		다크 아카이브 정보패키지 메타데이터 요소		
상위요소	하위요소	상위요소	하위요소	추출한 요소의 출처
2 기록계층	-			
3 기록식별자	3.1 기본식별자	매체와 크기	ISAD(G), 기록관리 메타데이터 표준	
	3.2 시스템식별자			
	3.3 보조식별자			
4 기록물명	4.1 제목	기록식별자	OAIS(참조정보), 기록관리 메타데이터 표준	
	4.2 기타제목			
5 기술	5.1 기술 유형	기록계층	ISAD(G), 기록관리 메타데이터 표준	
	5.2 기술 내용			
6 주제	6.1 주제 유형	전거명	ISAAR(CPF), ISDIAH	
	6.2 주제명			
7 전자기록물 여부	-	기관 고유식별자	ISAAR(CPF), ISDIAH	
8 유형	8.1 기록 유형			
	8.2 사본 유형			
9 매체	-	존속 기간	ISAAR(CPF)	
10 크기	10.1 용량			
	10.2 단위			
11 분류	11.1 분류체계 유형	수집/이관의 직접적 출처	ISAD(G)	
	11.2 분류값			
12 일시	12.1 생산일시	기술	ISDF, 기록관리 메타데이터 표준	
	12.2 종료일시			
13 생산이력	13.1 생산시스템	관계	OAIS(맥락정보), ISAAR(CPF), ISDF	
	13.2 생산부서			
	13.3 비치기록물			
14 보존기간	14.1 보존기간	상호연관성	OAIS	
	14.2 보존기간책임사유			
15 보존장소	-	보존 위치	다크 아카이브 특성, ISDIAH	
16 권한	-			
17 위치	17.1 소장처 필수	보존 이력	다크 아카이브 특성, ISDIAH, OAIS(출처정보)	
	17.2 소장위치			
18 관리이력	18.1 관리 유형	보존 전략	다크 아카이브 특성	
	18.2 관리설명			
	18.3 관리일시	원본	ISAD(G), 다크 아카이브 특성	
	18.4 관리행위자			
	18.5 변경요소			
19 이용이력	19.1 이용 유형	보존 정보	다크 아카이브 특성	
	19.2 이용내용설명			
	19.3 이용일시	관리자	OAIS	
	19.4 이용자			
20 보존이력	20.1 보존처리 유형	접근 권한	OAIS(접근권한정보)	
	20.2 보존처리설명			
	20.3 보존처리일시	관리 이력	ISDF, 기록관리 메타데이터 표준	
	20.4 보존행위자			
21 관계	21.1 관계 유형	이용 이력	기록관리 메타데이터 표준	
	21.2 관계대상식별자			
	21.3 관계설명			
22 무결성체크	22.1 무결성체크법	변환	다크 아카이브 특성	
	22.2 무결성체크값			
		복구	다크 아카이브 특성	
		무결성 점검	OAIS(인증정보)	

구축한 다크 아카이브 정보패키지 구조의 내용정보, 맥락정보, 보존정보, 관리정보를 메타데이터 요소의 상위요소로 보고, 다음의 <표 1>에서 도출한 기록물의 메타데이터 요소와 다크 아카이빙에 필요한 요소를 결합하여 하위요소로 재정리하였다. 이를 구체적으로 살펴보면 <표 2>와 같다.

<표 2> 다크 아카이브의 내용정보(안)

상위요소	하위요소
내용정보	참조코드(Reference code(s))
	제목(Title)
	등록 일자(Registration Dates)
	유형(Type)
	포맷(Format)
	매체와 크기(Medium and Extent)
	기록식별자(Record Identifier)
기록계층(Aggregation Level)	

<표 2>에서 제시한 내용정보는 기존의 OAIS 참조모델 구조를 기반으로 하여 생성된 구조이다. 기존 OAIS 참조모델과 달리 기록물의 원본과 첨부파일 대신 기록물군의 내용과 관련한 하위 요소들이 작성된다는 점에서 차이점을 가진다. 또한 이관된 기록물의 메타데이터 요소 중 포맷, 매체와 크기, 기록식별자, 기록계층의 하위요소들을 기록관리 메타데이터 표준에서 추출하였으며, 철, 건 단위로 관리되는 기록관리 메타데이터 표준과 달리 군 단위로 관리된다는 점에서 기존 요소들과 차이가 존재한다. 이와 함께, ISAD(G)를 기반으로 하여 참조코드, 제목, 등록 일자 등의 하위요소를 추출하였으며 ISAAR(CPF)와 ISDF를 기반으로 유형을 추출하였다. 그리고 내용정보의 하위요소들 중 기록식별자는 OAIS 참조모델의 참조정보

를 반영한 요소이다.

맥락정보는 내용정보의 출처와 생산 배경과의 관계를 기술하여 전자기록물의 진본성 확인을 지원해 준다. ISAAR(CPF)와 ISDIAH를 기반으로 전거명, 기관 고유식별자 등의 요소를 추출해 작성기관이 채택한 규약이나 규칙에 따라 표준화된 이름 형식을 기입하고 기관을 식별하기 위한 숫자나 숫자와 문자가 혼합된 식별자를 기입한다. ISAAR(CPF)를 바탕으로 기관의 설립 및 해체일을 기입하는 존속 기간 요소를 추출하였으며, ISAD(G)를 기반으로 수집·이관의 직접적 출처 요소를 제시하였다(<표 3> 참조).

<표 3> 다크 아카이브의 맥락정보(안)

상위요소	하위요소
맥락정보	전거명(Authorized form of name)
	기관 고유식별자(Identifiers for corporate bodies)
	존속 기간(Dates of existence)
	수집/이관의 직접적 출처(Immediate source of acquisition)
	기술(Description)
	관계(Relation)
	상호연관성(Interconnected Relationship)

이관된 기록물의 메타데이터 기술요소를 ISDF와 기록관리 메타데이터 정보를 바탕으로 추출해 기록물의 내용에 대한 정보를 제공함으로써 제목으로 표현하지 못하는 기록물 내용에 대한 검색을 지원하고자 한다. 그리고 ISAAR(CPF)와 ISDF, OAIS 참조모델의 보존기술정보 중 맥락정보를 바탕으로 다크 아카이브에 저장된 기록물과 원자료와의 관계에 대한 요소를 추가

하였다. 하위요소들 중 상호연관성은 OAIIS 참조모델을 기반으로 한 요소로 내용정보가 생산된 배경, 해당 내용정보가 다른 내용정보와의 상호연관성 등에 대한 설명들이 포함된다.

보존정보는 ISAD(G)를 기반으로 다크 아카이브의 특성을 반영하여 원본과 복본의 존재와 저장 정보에 대해 기술한다. 웹 기반 학술정보 다크 아카이브인 CLOCKSS의 프로세스를 바탕으로 다크 아카이브에 복본을 생산하여 분산 저장하고 ISDIAH를 참고하여 이에 대한 보존 이력과 보존 위치를 기술한다. 또한 전자기록물 기술 동향을 주기적으로 모니터링하고 동향에 맞춘 보존 전략 요소를 기술하고 전자기록물 보존에 있어 진본성, 이용가능성을 보장하고 데이터베이스의 무결성을 유지하기 위한 요소 등을 제안하였다(〈표 4〉 참조).

〈표 4〉 다크 아카이브의 보존정보(안)

상위요소	하위요소
보존정보	보존 위치(Preservation Location)
	보존 이력(Preservation History)
	보존 전략(Preservation Strategy)
	원본(Original)
	복본(Duplicate Copy)

관리정보는 아카이브에 오류가 발생하였는지 주기적으로 점검하여 기록물의 품질을 유지한다. 다크 아카이브의 관리자는 보존된 기록과 그 상태를 검토할 수 있는 권한이 부여된다. 따라서 OAIIS 참조모델을 바탕으로 관리자 요소를 추가하여 기록물의 보존 과정에 관여한 주체를 식별하였다. 이와 함께, 주기적인 검사를 통해 기록물을 지속적으로 관리해야 하는 다크 아카이브의 특성을 반영해 변환 요소를 추가하여

마이그레이션과 에뮬레이션 등의 사항을 기술할 수 있도록 하였다(〈표 5〉 참조).

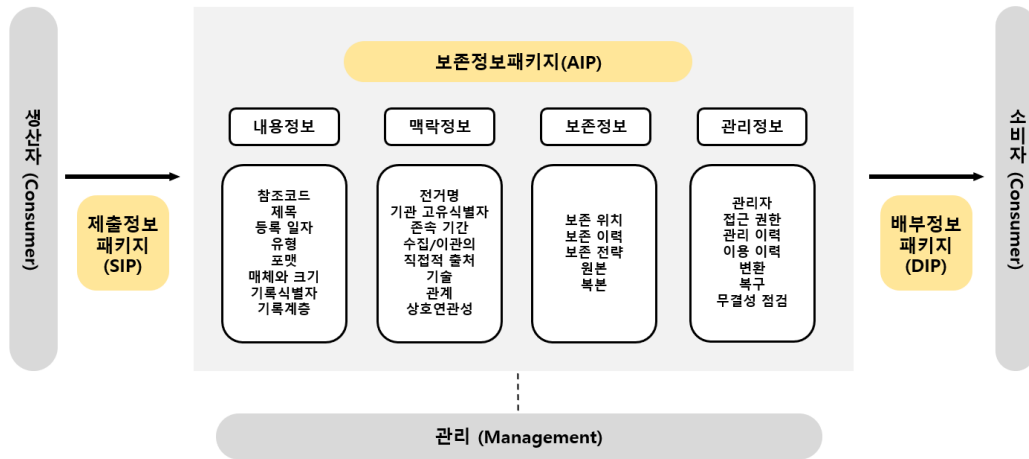
〈표 5〉 다크 아카이브의 관리정보(안)

상위요소	하위요소
관리정보	관리자(Manager)
	접근 권한(Access Rights)
	관리 이력(Management History)
	이용 이력(Use History)
	변환(Conversion)
	복구(Restoration)
무결성 점검(Integrity Check)	

이와 같이, 본 연구에서는 다크 아카이브의 구축과 운영에 있어 기존의 정보패키지에 비해 효율적으로 적용될 수 있는 정보패키지의 구성을 제안하였다. 상기와 같이 구축된 다크 아카이브 정보패키지 구조를 도식하면 다음 〈그림 7〉과 같다.

〈그림 7〉에 나타난 바와 같이, 다크 아카이브 정보패키지는 보존의 대상이 되는 정보 객체와 관련된 내용정보와 해당 정보 객체의 장기보존과 관련된 기술사항인 맥락정보, 보존정보, 관리정보로 구성된다. 각각의 구조는 요소에 따라서 개별적으로 구성되지만, 전체적인 다크 아카이브 프로세스 안에서 연결되어 기록물의 장기보존을 지원한다.

본 연구에서 제안한 다크 아카이브 정보패키지에서는 기존의 OAIIS 참조모델, 기록관리 메타데이터 표준, 그리고 국제 기록물 기술 표준을 바탕으로 하여 기존의 메타데이터 레코드와의 상호운용성을 확보하였으며, 동일한 내용에 대한 중복 기술을 방지하여 업무 효율성을 높이고자 하였다. 하지만, 이 구조는 정보 객체를



〈그림 7〉 다크 아카이브 정보패키지 구조

장기적으로 보존하는 데 필요한 최소한의 상위 요소들만을 제시해 주고 있으며, 정보 객체를 실제적으로 기술할 수 있는 세부적인 구조를 제공해 주지 못하고 있다. 이에 본 연구에서는 제안한 정보패키지 구조를 다크 아카이빙을 위한 핵심 보존 프로세스로 보고, 각 프로세스에서 필요한 세부적인 메타데이터 요소를 제안하고자 한다.

4.3 정보패키지를 위한 메타데이터 요소 제안

디지털 아카이빙에서 보존 메타데이터는 디지털 객체 자체에 대한 기술사항 및 해당 객체의 보존과 관련된 기술사항이 결합된 정보패키지의 형태로 존재한다(이승민 2015). 본 연구에서는 디지털 객체 자체에 대한 기술정보인 내용정보와 해당 객체의 보존과 관련된 기술사항인 맥락정보, 보존정보, 관리정보를 제안하였으며, 이를 기반으로 다크 아카이빙에 최적화된 다크 아카이브 정보패키지(DAIP) 구조를 제안하였다. 하지만 제안한 다크 아카이브 정

보패키지는 개념적인 구조이며, 다크 아카이빙에 필요한 요소들을 정보패키지로 구성하여 실질적으로 보존 활동에 적용될 수 있는 구체적인 기술항목을 제시해 주지 못했다는 한계를 보이고 있다. 이에 다크 아카이빙에 적용할 수 있는 세부적인 메타데이터 요소를 도출하고 이를 기반으로 보존의 대상이 되는 정보 객체를 실제적으로 기술할 수 있는 다크 아카이빙 메타데이터 요소를 제안하고자 한다.

메타데이터 요소는 기록물을 기술 및 보존하는 데 필요한 기본 단위로서, 기록물이 지닌 다양한 측면들은 각각의 메타데이터 요소를 통해서 기술된다. 따라서 메타데이터 요소는 기록물을 기술하는 데 있어 필수불가결한 요소이며, 기록물이 지닌 다양한 측면들은 모두 메타데이터 요소를 중심으로 하여 기술되어야 한다. 메타데이터가 지닌 이러한 기능을 반영함과 동시에, 앞서 제안한 정보패키지의 구조와 상위요소들을 바탕으로 다크 아카이브의 운영 효율화를 위한 세부적인 메타데이터의 요소를 제안하였다(〈표 6〉 참조).

〈표 6〉 다크 아카이브 정보패키지의 메타데이터 요소(안)

상위요소			하위요소		
요소명	필수여부	반복여부	요소명	필수여부	반복여부
내용정보 (Content Information)	필수	-	참조코드(Reference code(s))	필수	-
			제목(Title)	필수	-
			등록 일자(Dates)	필수	-
			유형(Type)	필수	-
			포맷(Format)	해당	반복
			매체와 크기(Medium and Extent)	해당	반복
			기록식별자(Record Identifier)	필수	-
			기록계층(Aggregation Level)	필수	-
맥락정보 (Context Information)	필수	-	전거명(Authorized form of name)	필수	-
			기관 고유식별자(Identifiers for corporate bodies)	필수	-
			존속 기간(Dates of existence)	해당	-
			수집·이관의 직접적 출처(Immediate source of acquisition)	해당	-
			기술(Description)	선택	반복
			관계(Relation)	필수	-
			상호연관성(Interconnected Relationship)	해당	-
보존정보 (Preservation Information)	필수	반복	보존 위치(Preservation Location)	필수	반복
			보존 이력(Preservation History)	필수	반복
			보존 전략(Preservation Strategy)	해당	반복
			원본(Original)	필수	-
			복본(Duplicate Copy)	필수	반복
관리정보 (Management Information)	필수	반복	관리자(Manager)	필수	반복
			접근 권한(Access Rights)	필수	-
			관리 이력(Management History)	필수	반복
			이용 이력(Use History)	해당	반복
			변환(Conversion)	해당	반복
			복구(Restoration)	해당	반복
			무결성 점검(Integrity Check)	해당	반복

〈표 6〉에서는 다크 아카이브에서의 기록물 장기보존에 최적화된 메타데이터 요소를 제안하고 있다. 이는 일반적인 디지털 아카이브와는 다른 기록물의 장기보존 측면에 중점을 두고 있으며, 기록물의 활용에 대한 부분이 아닌 기록물 보존 관리에 특화된 메타데이터 요소 세트가 구

성되어 있다. 제안한 메타데이터는 총 4개의 상위요소와 27개의 하위요소로 구성되어 있으며, 하위요소들은 13개의 '필수요소'와 14개의 '해당시 필수요소' 혹은 '선택요소'로 구성하였다.

필수요소는 다크 아카이브의 보존 대상 기록물의 기술에 반드시 포함되어야 하는 요소로,

다크 아카이빙 프로세스상에서 꼭 필요한 요소이다. 해당 시 필수요소 혹은 선택요소는 기술할 수 있는 경우에 선택적으로 기술하는 요소로, 다크 아카이브의 기록을 관리하기 위해 필요하다라고 판단되는 요소들을 추가하여 이들 요소를 도출하였다.

내용정보의 참조코드, 제목, 등록일자, 유형, 기록식별자, 기록계층의 경우 기록물군에 대한 구체적인 설명을 작성하는 항목들이기 때문에 필수요소로 설정하였으며, 이들 요소 가운데 참조코드, 제목, 등록일자, 기록식별자는 수정이 불가능한 요소이다. 이외의 포맷과 매체와 크기 요소의 경우 마이그레이션, 에뮬레이션 등의 변환 프로세스를 통해 변동될 수 있기 때문에 해당 시 필수요소로 설정하였으며 반복이 가능하다.

맥락정보에서는 전거명, 기관 고유식별자, 관계 요소가 필수요소로 설정되었다. 기술의 대상이 되는 기록물에 따라 존속기관, 수집·이관의 직접적 출처, 상호연관성은 해당 시 필수요소로 설정하였으며, 기술은 선택 요소로 설정하였다.

다크 아카이브 정보패키지(DAIP)에서 가장 중요한 보존정보의 경우, 보존 위치, 보존 이력, 원본, 복본 등 대부분의 요소를 필수로 설정하여 보존과 관련된 사항을 상세하게 기술할 수 있도록 하였으며 반복을 통해 보존 프로세스를 확인할 수 있도록 하였다. 보존 전략의 경우, 기록물 대상에 따라 해당되는 보존 전략이 있을 경우에만 기술하며 반복할 수 있다.

관리정보의 경우, 무결한 보존 환경을 유지하기 위해 관리자, 접근 권한, 관리 이력을 필수요소로 설정하였다. 또한 이용이력 요소를 해당 시 필수 및 반복 가능 요소로 설정하였으며, 이

를 통해 다크 아카이브 정보패키지 구조 내에서 기록물을 이관, 입수하거나 이용자의 요청에 따라 제공할 경우를 충분히 기술할 수 있도록 하였다. 이외의 변환, 복구, 무결성 점검의 요소를 통해 다크 아카이브 보존정보패키지의 기록물의 원활한 장기 보존이 이루어질 수 있도록 하였다.

이와 같이 4개의 상위요소 및 27개의 하위요소로 구성된 메타데이터 요소 세트는 다크 아카이브의 고유한 목적을 충족시키는데 최적화된 방식으로 구성된 것이다. 일반 대중의 접근이나 활용이 아닌 기록물의 장기 보존 및 관리를 위한 프로세스 환경을 지원하고 기록물의 보존적인 측면을 기술하는 데 중점을 둔 메타데이터 요소 세트의 구축을 통해, 다크 아카이빙을 지원하기 위한 정보패키지를 보다 효율적으로 구축할 수 있는 환경을 마련할 수 있다. 특히 OAIS 참조모델의 내용정보와 보존기술 정보를 구성하는 항목들을 다크 아카이브 정보패키지의 생성에 맞도록 4개의 상위요소로 재구성함으로써 다크 아카이빙 전반에 걸친 운영 효율화를 지원할 수 있는 환경을 마련할 수 있을 것으로 보인다. 이와 함께, 디지털 아카이빙에 일반적으로 적용되는 OAIS 참조모델의 정보패키지에 기반을 두고 있으므로, 향후 디지털 아카이브와 다크 아카이브의 상호연계를 위한 환경을 구축하는 데 있어서의 효율성도 확보할 수 있다는 장점을 지니고 있다. 이는 다크 아카이브의 정보패키지 구축을 위한 메타데이터 요소 세트 마련이라는 측면 이외에도, 아카이빙의 생태계를 보다 확장성 있게 구현할 수 있다는 점에서 기록물 관리에서의 핵심적인 기반으로 활용될 수 있을 것으로 기대된다.

5. 결 론

방대한 양의 전자기록물이 저장, 보존되는 다크 아카이브는 기존의 아카이브와는 다른 특성과 목적을 지니고 있으며, 기록물의 저장과 활용보다는 장기보존 및 보호 관리에 중점을 지니게 된다. 따라서 다크 아카이브에서는 기록물을 군 단위로 저장 관리하는 방식을 적용하고 있으며, 이에 따라 다크 아카이브에 소장되는 기록물을 기술하고 이를 위한 정보패키지를 생성하는 데 있어서도 기존의 아카이빙과는 다른 방식이 적용되어야 한다. 하지만 디지털 아카이브와 다크 아카이브는 동일한 기록물을 대상으로 한다는 공통점을 지니고 있기 때문에, 기존의 디지털 아카이빙의 기반 위에서 다크 아카이빙에 최적화된 아카이빙 프로세스를 마련하는 것이 보다 적절한 방식이 될 수 있다.

이에 본 연구에서는 다크 아카이빙에서 생성되는 정보패키지의 효율성을 확보하는 데 최적화된 메타데이터 요소들을 제안하였으며, 이는 기록물의 장기 보존 및 군 단위의 기록물 기술에 적합하도록 재구성 및 새로운 요소들을 추가한 요소 세트이다. 이 요소 세트는 기록물 기술에 적용되는 국제 표준인 ISAD(G), ISAAR(CPF), ISDF, ISDIAH로부터 핵심적인 기술 항목들을 추출하여 구성하였으며, OAIS 참조모델의 구조를 기반으로 하여 다크 아카이브의 기술 목적을 충족시킬 수 있는 다크 아카이브 정보패키지(DAIP)의 구조를 구축하였다.

제안한 다크 아카이브 정보패키지 구조에서는 기록물 기술 표준 및 기록관리 메타데이터 표준에 정의된 요소들이 다크 아카이브의 구조에 하위요소로 포함되어 있다. 즉, 내용정보, 맥

락정보, 보존정보, 관리정보가 메타데이터 구조상에서 상위요소를 구성하고 있으며, 기존 기록물 메타데이터 요소 및 다크 아카이빙에 필요한 요소들을 재구성하여 상위요소의 하위에 배치되는 하위요소로 재정리하였다.

이렇게 구성된 메타데이터 요소 세트는 다크 아카이브의 운영 효율을 증가시키는데 주된 목적을 두고 구축한 것이며, 총 4개의 상위요소와 27개의 하위요소로 구성되어 있다. 하위요소들은 13개의 '필수요소'와 14개의 '해당 시 필수요소' 혹은 '선택요소'로 구성하였다.

내용정보의 참조코드, 제목, 등록일자, 유형, 기록식별자, 기록계층의 경우 기록물군에 대한 구체적인 설명을 작성하기 때문에 필수요소로 설정하였으며, 이 요소들 중 참조코드, 제목, 등록일자, 기록식별자는 수정이 불가능한 요소이다. 이외의 포맷과 매체와 크기 요소의 경우 마이그레이션, 에뮬레이션 등의 변환 프로세스를 통해 변동될 수 있기 때문에 해당 시 필수요소로 설정하였으며 반복이 가능하다.

맥락정보 가운데 전거명, 기관 고유식별자, 관계 요소는 필수요소로 설정하였으며, 기술의 대상이 되는 기록물에 따라 존속기관, 수집·이관의 직접적 출처, 상호연관성 등의 요소는 해당 시 필수 및 반복 가능으로 설정하였다.

보존정보의 경우, 보존 위치, 보존 이력, 원본, 복본 등 대부분의 요소가 필수로 설정되어 있으며, 반복 사용을 통해 보존 프로세스를 확인할 수 있도록 하였다. 보존 전략의 경우, 기록물 대상에 따라 해당되는 보존 전략이 있을 경우에만 기술하며 반복할 수 있다.

관리정보의 경우, 무결한 보존 환경을 유지하기 위해 관리자, 접근 권한, 관리 이력을 필수

요소로 선정하였다. 이용 이력 요소를 통해 반복 가능한 해당 시 필수 요소로 선정하여 다크 아카이브 정보패키지 구조 내에서 기록물을 이관, 입수하거나 이용자의 요청에 의해 배부될 경우를 기술할 수 있게 하였다.

본 연구에서 제안한 다크 아카이브 정보패키지에서는 기존의 OAIS 참조모델, 기록관리 메타데이터 표준, 그리고 국제 기록물 기술 표준을 바탕으로 하여 기존의 메타데이터 레코드와의 상호운용성을 확보하였으며, 동일한 내용에 대한 중복 기술을 방지하여 업무 효율성을 높이고자 하였다. 하지만 이 구조는 정보 객체를 장기적으로 보존하는 데 필요한 최소한의 상위

요소들만을 제시해 주고 있으며, 정보 객체를 실제로 기술할 수 있는 세부적인 구조를 제공해 주지는 못하고 있다. 이에 본 연구에서는 제안한 정보패키지 구조를 다크 아카이빙을 위한 핵심 보존 프로세스로 보고, 각 프로세스에서 필요한 세부적인 메타데이터 요소를 제안하는 데 중점을 두고 있다. 제안된 메타데이터 요소는 다크 아카이브 정보패키지를 운영하는 데 있어서의 효율성을 높여줄 수 있는 방안 가운데 하나가 될 수 있을 것이며, 향후 각 기관에서 전자기록물의 장기적 보존을 위해 다크 아카이브를 구축하고자 할 때에 기초자료로서 활용될 수 있을 것으로 기대된다.

참 고 문 헌

- [1] 남태우, 이승민. 2018. 『정보자원의 기술과 메타데이터』. 개정판. 서울: 한국도서관협회.
- [2] 이승민. 2015. 디지털 아카이브를 위한 보존 메타데이터 패키지 구축. 『정보관리학회지』, 32(3): 21-47.
- [3] 이승민. 2017. 디지털 아카이빙을 위한 보존 기술항목 프레임워크 구축. 『한국도서관·정보학회지』, 48(4): 129-151.
- [4] 이해영. 2020. 『기록의 분류·기술과 검색도구』. 서울: 선인.
- [5] 임진희. 2013. 『전자기록관리론』. 서울: 선인.
- [6] 한국기록관리학회. 2018. 『기록관리의 이론과 실제』. 서울: 조은글터.
- [7] Beagrie, N. 2004. "The Continuing Access and Digital Preservation Strategy for the UK Joint Information Systems Committee (JISC)." *D-Lib Magazine*, 10(7/8). [online] [cited 2020, 10, 20.] <<http://www.dlib.org/dlib/july04/beagrie/07beagrie.html>>
- [8] Caplan, Priscilla. 2005. "Building a Dark Archive in the Sunshine State: A Case Study." In *Proceedings of Archiving 2005: Final Program and Proceedings*, 9-13.
- [9] Caplan, Priscilla. 2007. "The Florida Digital Archive and DAITSS: A Working Preservation "Repository Based on Format Migration."" *International Journal on Digital Libraries*, 6(4):

305-311.

- [10] Formenton, D. et al. 2017. "Metadata Standards as a Technological Resource to Guarantee Digital Preservation." *Biblios*, 68: 82-95.
- [11] Perrin, J. M., Winkler, H. M. and Yang, L. 2015. "Digital Preservation Challenges with an ETD Collection: A Case Study at Texas Tech University." *Journal of Academic Librarianship*, 41(1): 98-104.
- [12] Rosenthal, D.S.H. 2014. "Architectural Choices in LOCKSS Networks." *Library Hi Tech*, 32(1): 2-10.
- [13] Society of American Archivists, 2020. *Dictionary of Archives Terminology*. s.v. "dark archives." [online] [cited 2020. 10. 21.] <<https://dictionary.archivists.org/entry/dark-archives.html>>
- [14] CCSDS, 2019. Reference Model for an Open Archival Information System (OAIS): Preservation Metadata and the OAIS Information Model: Draft Recommended Practice, Issue 3. [online] [cited 2020. 10. 19.] <https://cwe.ccsds.org/moims/_layouts/15/WopiFrame.aspx?sourcedoc=/moims/docs/MOIMS-DAI/Draft%20Documents/OAIS%20v3/OAIS%20final%20v3%20draft%20with%20changes%20wrt%20OAISv2%2020190924-rl.docx&action=default>

• 국문 참고자료의 영어 표기

(English translation / romanization of references originally written in Korean)

- [1] Nam, Taewoo and Lee, Seungmin, 2018. *Resource Description and Metadata*. Revised Edition. Seoul: Korea Library Association.
- [2] Lee, Seungmin, 2015. "Construction of Preservation Metadata Package for Digital Archiving." *Journal of the Korean Society for Information Management*, 32(3): 21-47.
- [3] Lee, Seungmin, 2017. "Construction of Preservation Description Framework for Digital Archiving." *Journal of Korean Library and Information Science Society*, 48(4): 129-151.
- [4] Rieh, Hae-Young, 2020. *Records Classification, Description and Tools for Retrieval*. Seoul: Seonin Pub.
- [5] Yim, Jin-Hee, 2013. *Theory of Electronic Record Management*. Seoul: Senin Pub.
- [6] Korean Society of Archives and Records Management, 2018. *Theory and Practices of Records Management*. Seoul: JoEunGeulTeo.

