

E-ARK 기반 디지털 아카이빙 모델에 관한 연구*

A Study of a Digital Archiving Model based on E-ARK

박희진 (Heejin Park)**

한상은 (Sangeun Han)***

오삼균 (Sam-Gyun Oh)****

초 록

본 연구는 E-ARK(European Archival Records and Knowledge Preservation)를 기반으로 한 디지털 아카이빙 모델을 제안하는 것을 목적으로 한다. 이를 위해 디지털 아카이빙 관련 국제 표준과 기술 사항을 분석하여 디지털 아카이브의 핵심 기능 영역을 도출하고 기능 모델 구축에 필요한 기술, 정보패키지와 메타데이터 구현을 위한 요구사항을 파악하였다. 디지털 아카이브의 전반적인 프로세스를 고려하여 정보패키지 구성에 있어서 상호운용성을 확보할 수 있는 디지털 아카이빙 모델을 제안하였다.

ABSTRACT

This study aims to present a model of digital archiving based on E-ARK. It analyzed the international standards and technological specifications designed for digital archiving. The analysis employed in study explored the common specifications including core processes, information packages, and metadata structure needed for digital archiving. Based on the analysis and reviews, this study developed a model for digital archiving, in order to achieve interoperability of information packages throughout the process.

키워드: 디지털 아카이브, OAIS, E-ARK, 정보패키지, 기능모델
digital archives, OAIS, E-ARK, information package, functional model

* 본 연구는 2017년 한국교육학술정보원의 지원을 받아 수행된 '교육자원 디지털 아카이빙체제 구축방안 개발' 과제 영역의 일부 내용을 수정·보완한 것임.

** 한성대학교 크리에이티브인문학부 조교수(papermod@hansung.ac.kr) (제1저자)

*** 성균관대학교 문헌정보학과 박사과정(silver86eun@gmail.com) (공동저자)

**** 성균관대학교 문헌정보학과 교수(samoh21@gmail.com) (교신저자)

■ 논문접수일자: 2018년 2월 18일 ■ 최초심사일자: 2018년 3월 9일 ■ 게재확정일자: 2018년 3월 20일

■ 정보관리학회지, 35(1), 83-101, 2018. [http://dx.doi.org/10.3743/KOSIM.2018.35.1.083]

1. 서론

1.1 연구목적

디지털 아카이빙이란 “지속적인 가치를 가졌다고 판단할 수 있는 디지털 객체를 장기간 관리하는 활동(이소연, 2002, p. 46)”을 말한다. 최근 국내에서는 이러한 디지털 아카이빙의 필요성을 인식하고 많은 기관이 아카이브 시스템 개발에 예산을 투자하고 있다(안대진, 임진희, 2016). 기관마다 한정된 예산에 이용자가 원하는 모든 정보를 보존하는 것은 불가능 한 일이다. 따라서 아카이브 시스템 개발과 운영에 소요되는 비용을 절감하고 자원을 효율적으로 보존하고 공유하기 위해서는 아카이브 기관간의 상호협력이 절실하다. 하지만 디지털 아카이빙의 상호운용성에 대한 국제적 수준의 실무 지침이 아직 마련되어 있지 않아 협력이 이루어지기 어렵다.

디지털 아카이빙은 또한 정보자원의 생산부터 수집, 기술, 저장, 보존 및 접근까지 다양한 활동이 포괄적으로 고려되어야 하는 종합적인 프로세스이다. 디지털 아카이빙이 효율적으로 수행되려면 전체적인 프로세스에서 이루어지는 보존과 관련된 일련의 활동이 상호 연결되어 유기적으로 이루어져야 한다. 이를 위해서는 각 기능에서 수행되는 정보의 관리와 기술에 있어서 상호운용성 확보가 매우 중요하다. 그러나 아카이빙에서 널리 사용되고 있는 METS 및 PREMIS 등의 메타데이터 표준들도 실질적으로 상호운용성을 확보하는데 있어서 요구되는 구현 수준의 구체적인 지침이나 세부 사항이 부족한 실정이다(E-ARK, 2014).

이에 본 연구는 디지털 아카이빙의 프로세스에 있어서 자원의 장기적 보존을 위해 유기적 협력과 운영을 지원할 수 있는 아카이빙 모델을 제안하는 것을 목적으로 한다. 이를 위해 유럽 아카이브 프로젝트 E-ARK(European Archival Records and Knowledge Preservation)을 기반으로 디지털 아카이브의 기능 모델 구축에 필요한 기술과 정보패키지의 요구사항을 정의하고자 한다. OAIS 참조모형에서 제시한 디지털 아카이브의 핵심 기능 영역을 중심으로, 각 영역에서 상호 유기적으로 적용 되는 업무 프로세스와 단계에 적용되는 기술사항을 제안하고자 한다. 제안한 디지털 아카이브 모델은 디지털 정보자원의 생산단계에서부터 보존, 활용에 있어서 이루어지는 정보의 교환, 공유, 접근 등 다양한 아카이빙의 측면에서 기관끼리의 협력을 지원할 수 있는 지침으로 활용될 수 있을 것이다.

1.2 선행 연구

본 연구와 관련된 디지털 아카이빙 분야의 선행연구는 크게 디지털 아카이빙 기능모델에 관한 연구와 장기보존전략 및 아카이빙 자원의 기술사항에 관한 연구로 구분할 수 있다. 디지털 아카이빙 기능모델에 관한 연구로는 주로 OAIS 참조모형을 기반으로 하여 적용할 수 있는 장기보존전략을 검토하는 연구들이 다수 수행되어 왔다. 이소연(2009)은 OAIS 참조모형의 디지털 아카이브 기능모델과 PLANETS 프로젝트를 중심으로 디지털 아카이브의 장기 보존 기능을 종합하여 제시하고 국내 디지털 아카이브의 향후 과제를 제안하였다. 신재민과

곽승진(2013)은 체계적인 디지털 아카이빙 전략을 수립하기 위해 국내외 디지털 아카이빙의 선진 사례를 조사하고 국내 디지털 아카이빙 시스템에 적용할 수 있는 전략 방안을 제안하였다. 안대진과 임진희(2016)는 OAIS 참조모형을 중심으로 하여 디지털 아카이브 시스템 구축에 필요한 기능요건을 도출하고 공개 소프트웨어를 활용하는 방안을 제안하였다.

디지털 아카이빙에 적용할 수 있는 메타데이터 요소를 구축하기 위한 연구 또한 다수 수행되고 있으며 대표적인 연구를 살펴보면 다음과 같다. 박옥남(2012)은 PREMIS를 기반으로 하여 보존 메타데이터 스키마 개발 및 기존의 메타데이터에 보존 메타데이터를 재사용할 수 있는 가이드라인을 제안하였다. 최근 이승민(2015)은 디지털 객체의 보존을 위해서 디지털 아카이빙의 핵심적인 단계에 따라 최적화된 방식으로 제공할 수 있는 메타데이터 패키지를 구축하였다. 이외에도 많은 디지털 아카이빙의 전략과 방안 마련을 위한 연구가 진행되고 있는데, 이들 대부분은 디지털 아카이빙 프로세스 전반에 걸쳐서 상호운용성을 확보할 수 있는 방안을 제시하고 있지 못하는 한계가 있다.

2. 디지털 아카이브 관련 표준과 기술

2.1 ISO 14721 OAIS 표준

OAIS(Open Archival Information System) 참조모형은 아카이브가 디지털 정보를 장기간 보존하는 다양한 기관들의 의사소통과 협력을

활성화하기 위해 개발된 기술적 권고안으로 2002년 ISO 14721로 공포되었다. OAIS 참조모형은 디지털 정보를 장기 보존하기 위한 사항을 기술하고 장기 보존에 필요한 접근 및 관리에 관한 개념을 이해시키는 기본적인 틀로 널리 적용되고 있다. OAIS는 장기적인 정보 보존과 접근 기능을 수행하는 시스템이나 기관으로, 생산자(Producer), 관리(Management), 사용자(Consumer)로 구성된 외부환경과의 상호작용을 통해 주요 기능을 수행한다. OAIS 환경에서 정보는 보존의 목적이 되는 내용정보(Content Information)와 이의 보존을 돕는데 필요한 보존기술정보(Preservation Description Information)를 담은 정보패키지(Information Package, 이하 IP) 형태로 존재한다. IP는 OAIS 환경에서 각 기능의 단계에 따라, 제출 정보패키지(Submission Information Package: SIP), 보존정보패키지(Archival Information Package: AIP), 배부정보패키지(Dissemination Information Package: DIP) 형태를 취한다. 정보생산자가 SIP 형식을 사용하여 OAIS 시스템으로 입수되면, OAIS 환경 내에서 보존, 관리를 위한 AIP 형태로 보존한다. 이용자의 요구가 있을 경우, 시스템 내의 존재 여부나 접근 권한을 확인하여 DIP 형태로 제공된다.

OAIS는 디지털 아카이브의 기능을 입수, 아카이브 저장, 데이터 관리, 운영, 보존 계획, 접근의 여섯 가지로 정의하고 있다. 입수는 정보생산자의 SIP를 입수하여 품질을 관리하고 AIP를 생성하는 기능을 수행한다. 아카이브는 AIP의 저장 및 유지관리, 검색을 위한 기능과 서비스를 제공한다. 데이터 관리는 아카이브를 관리하는데 사용되는 데이터베이스와 관련된

여 데이터 입력, 접근, 유지 등과 같은 기능을 수행한다. 행정(운영)은 아카이브 시스템의 전반적인 운영을 관리한다. 보존계획은 요구사항 검토, 보존전략과 표준을 개발하는 기능 등을 수행하며, 운영을 관리한다. 접근 및 이용은 OAIS에 저장된 정보의 존재, 기술, 저장장소, 이용 가능성 등을 결정하고 이를 지원하고 사용자들이 정보생산물을 요청하고 입수할 수 있도록 DIP를 제공한다.

OAIS는 디지털 아카이빙을 수행하고 있는 다양한 기관의 의사소통과 협력을 위해 마련된 국제적 표준으로 이를 준수하는 것은 신뢰할 수 있는 디지털 아카이브 구축의 기본 조건이라 할 수 있다. 그러나 OAIS 참조모형은 최상위차원에서 디지털 아카이빙을 정의한 것으로, 실제 디지털 아카이브 구축과 운영에 적용할 수 있는 구체적인 방안을 제시하고 있지 못하다. OAIS에서 제시한 아카이브 기능모델과 정보모델을 구현하기 위한 구체적인 지침과 기술 사항 마련이 요구되고 있다.

2.2 보존전략과 메타데이터 표준

OAIS 참조모형은 디지털 자원의 보존전략으로 마이그레이션과 에물레이션 기법을 제안하고 있다. 마이그레이션은 보존하고자 하는 정보를 OAIS 내에서 이전하는 것을 말하며, 에물레이션은 원래의 하드웨어, 매체, 운영체제의 운용을 그대로 재현하는 프로그램을 마련하는 전략을 말한다. 이러한 보존전략과 함께 디지털 자원의 장기보존을 위해서는 포맷 정보를 체계적으로 저장, 관리할 수 있는 포맷 레지스트리가 필요하다. 포맷 레지스트리는 기술보존

의 직접적인 방법이자 기술정보요소를 지속적으로 수집 및 관리하여 마이그레이션, 에물레이션과 같은 보존 전략을 효율적으로 수행할 수 있도록 지원한다(박지화, 2016). 디지털자원의 장기보존과 장기접근성 유지에 필요한 기술(technical) 정보, 즉 포맷정보와 관련된 하드웨어/소프트웨어 정보, 매체정보를 소장한 포맷 레지스트리들이 활용되고 있다(유영수, 2006; 유영수, 2007; 박지화, 정영미, 2015).

또한 디지털 아카이빙에서는 디지털 객체가 수록하고 있는 내용정보에 대한 세부적인 사항을 포괄적으로 제공하며 자원의 무결성과 신뢰성을 유지할 수 있는 보존 메타데이터의 구축이 필요하다. 아카이빙 자원을 효과적으로 기술하고 관리, 검색할 수 있도록 수록하고 있는 디지털 객체에 대한 세부적인 사항을 포괄하는 메타데이터 구조들이 개발되고 있다(이승민, 2015). PREMIS(Preservation Metadata: Implementation Strategies)는 디지털 아카이빙 시스템에서 보존 메타데이터를 구현하기 위한 실제적인 도구로 보존 메타데이터의 관리 및 활용을 위한 지침과 함께 핵심적인 보존 메타데이터 요소들을 정의한 표준이다. PREMIS는 OAIS 참조모형의 IP 개념에 기반하여 보존 메타데이터의 관리 및 활용에 필요한 핵심 요소들을 상위요소로 정의하고 데이터 모델과 데이터 사전으로 구성되어 있다. 데이터 모델은 데이터 사전 내에 정의되어 있는 요소들을 논리적으로 조직하기 위한 것으로 2017년 최신 버전에는 객체(object), 이벤트(event), 행위주체, (agents), 권한(rights)으로 구성되어 있다. 데이터 사전은 총 13개의 최상위 요소와 보존 대상이 되는 디지털 객체에 대한 의미적 단위를

기술하고 있다.

METS(Metadata Encoding and Transmission Standard)는 디지털 저작물에 대한 구조적, 기술적(descriptive), 관리적 메타데이터를 입력하기 위한 XML 언어 스키마로 디지털 정보자원의 교환을 표준화하기 위해 개발되었다. 이 표준은 디지털 객체의 내용을 구성하는 파일들을 구별하고, 그 내용의 구조를 표현하며 관련된 메타데이터를 연결하여 데이터의 원활한 전송, 교환, 보존을 지원한다. METS는 오디오, 비디오 자원을 포함하여 다양한 디지털 정보자원의 구조에 대한 기술을 제공하고 외부 자원에 대한 링크 기능 기술 및 관리 메타데이터에 더 많은 확장성을 제공하도록 하고 있다. 따라서 METS를 사용하여 분산된 환경에서 생성된 다양한 형식의 메타데이터를 변환하지 않고 사용할 수 있다. 이러한 장점으로 METS는 OAIS 참조모형을 지원하여 SIP, DIP를 인코딩하고 전달할 수 있는 AIP를 인코딩할 수 있는 프레임워크로 널리 활용되고 있다.

3. E-ARK 프로젝트

3.1 개발 배경과 목적

E-ARK는 디지털 자원의 장기보존에 있어서 국제적 상호운용성 확보를 위해 유럽연합에서 출자하여 2014년부터 2017년까지 3년간 진행된 프로젝트이다. 이 프로젝트의 참여기관은 유럽의 국가 기록 보존 핵심그룹 5개 기관, 5개의 선도적인 연구기관, 아카이빙 소프트웨어 솔루션 및 서비스를 제공하는 3개의 업체, 2개

의 정부기관 및 11개의 유럽 국가회원으로 구성되었다. E-ARK는 범 유럽 수준에서 공통으로 적용할 수 있는 프로세스 모델을 모색하고 다양한 디지털 정보 환경에서 생산된 데이터를 장기적으로 보존할 수 있는 지침 및 기술사항을 마련하고 있다(Paulic, 2015; E-ARK, 2014; Wilson et al., 2016).

2014년 E-ARK 프로젝트 초기단계에서 파악한 아카이브 시스템들의 현황은 다음과 같다. 첫째, 현재 디지털 아카이브들은 개별적인 필요에 의해 만들어져 운용되고 있으며, 아카이브 기관마다 규제력과 보존 실무 경험에서 큰 차이를 보인다. 둘째, 데이터에 대한 아카이브가 필요하나 현재 찾아보기 어렵다. 디지털 리포지터리는 다수의 아카이브 기관과 협력하고 싶어 하나 디지털 자원의 신뢰성 확보와 이용에 대한 체계적 접근이 미흡하다. 셋째, 현재 보존 도구와 시스템은 단기적 사용과 장기적 관리를 위한 시스템으로 구분되어 있다. 데이터 소유자 및 시스템 관리자들은 장기간 보관해야 하는 정보를 아카이브 시스템을 비롯한 여러 시스템 간에 교환하되, 가능한 한 효과적으로, 정보의 신뢰성과 무결성을 해치지 않고, 정보를 발견하고 재사용 할 수 있는 가능성을 제한하지 않는 아카이브를 요구하고 있다(DLM Archival Standards Board, 2016).

이에, E-ARK는 모든 디지털 보존 시스템이 정보의 수집, 추적 및 저장, 접근을 제공하되 정보의 크기, 유형, 포맷에 상관없이 합의된 원칙에 따라 정보를 식별, 확인, 검증 시 단일화된 절차를 따르도록 하는 것을 목표로 하고 있다. 나아가 E-ARK는 아카이브 기관마다 다른 경험에 따라 생길 수 있는 아카이빙 관련 기술의

격차를 줄이고 아카이브 간 협력을 지원하며 상호운용성을 확보할 수 있도록 지원하고자 한다. 디지털 기록물의 상호운용성 확보를 위한 표준화 및 도구 개발과 디지털 정보에 대한 전송, 보존 및 접근에 소요되는 비용 절감이 필요하다. 이를 위해, E-ARK는 기관이 디지털 아카이브에 정보를 내보내고 가져오는 방법을 표준화하고, 다양한 콘텐츠의 장기 보존을 위한 방법과 아카이브에 접근하는 이들을 위한 인터페이스와 서비스 구축에 주력하고 있다(E-ARK, 2014).

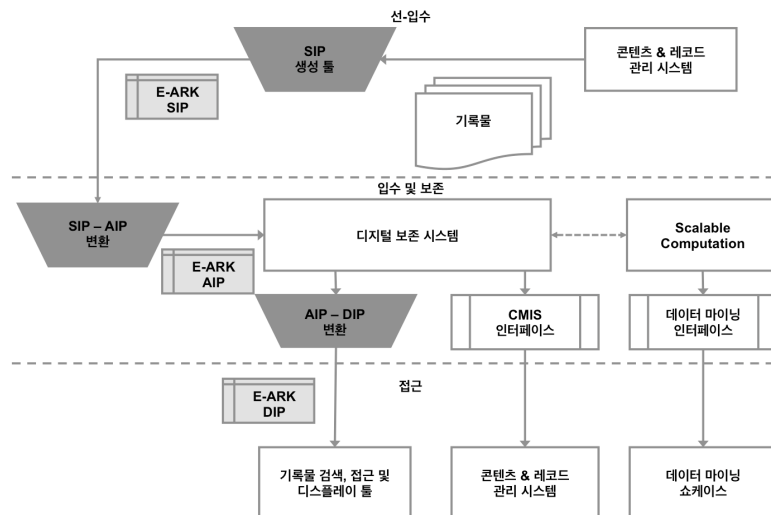
E-ARK는 궁극적으로 디지털 아카이빙에 있어서 의미적, 기술적으로 상호운용성을 확보하기 위해 OAIS 참조모형에 근거한 공통명세서(Common Specification)와 디지털 자원의 내용정보에 대한 세부적인 사항을 포괄할 수 있는 기술명세서를 개발하였다. 현재 E-ARK에서 개발한 보존 도구와 지침에 대하여 다양한 측면에서 아카이빙 자원의 진본성과 신뢰성을

시험하는 파일럿 테스트를 진행 중에 있다. 이러한 활동을 통해 디지털 아카이빙의 프로세스(입수, 저장 및 보존, 접근)에 있어서 자원의 크기, 유형, 포맷과 상관없이 합의된 원칙에 따라 정보를 식별, 확인, 검증에 필요한 표준 체계를 마련하고자 한다. 이러한 E-ARK 프로젝트의 성과는 개별 아카이브 간의 보존 도구의 경계를 없애 아카이빙 협력에 도움이 되리라 기대한다.

3.2 주요 개념과 내용

3.2.1 E-ARK 구조

E-ARK는 디지털 객체의 보존을 위해 다음의 <그림 1>과 같은 정보구조를 제안하고 있다. E-ARK의 정보구조는 디지털 정보자원의 생애주기에 기반 하여 아카이빙의 생산부터 보존, 활용까지의 프로세스에서 정보자원의 신뢰성과 무결성을 담보하기 위한 것이다. OAIS 참조모형의 디지털 자원의 아카이브로의 입수로부터



<그림 1> E-ARK 정보구조

이루어지는 프로세스와 비교하면 E-ARK 정보 구조는 선입수(Pre-Ingest) 단계가 추가된 것이 특징이다.

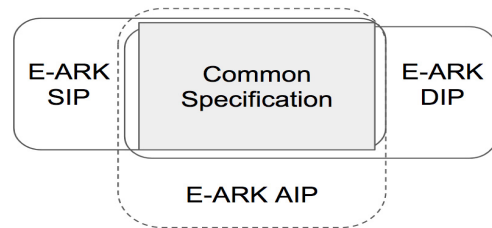
E-ARK의 선입수 단계는 디지털 아카이빙에 있어서 정보생산자와의 원활한 정보교환, 입수를 지원하기 위한 핵심적인 부분을 차지하고 있다. 선입수 단계에서는 E-ARK SIP가 생성된다. E-ARK SIP는 생산자들이 제공하는 모든 유형의 포맷 및 메타데이터를 수용할 수 있도록 확장이 가능해야 한다. 또한 디지털 아카이빙의 전 단계에 걸쳐 필요한 IP로 변환이 용이하도록 대용량의 내용정보는 분할 처리하도록 하고 있다. E-ARK SIP는 구조적 메타데이터 METS를 기반으로 하여 콘텐츠 파일의 메타데이터와 SIP 자체에 관한 메타데이터를 구분하도록 하고 있다(DLM Archival Standards Board, 2017).

E-ARK 프로젝트는 오픈소스 도구인 E-ARK Web(<http://github.com/eark-project/earkweb>)을 개발하여 IP의 생성과 변환을 지원하고 있다. E-ARK Web은 SIP 생성 틀을 활용하여 아카이빙 자원의 입수 과정에서 자원의 무결성을 검증하도록 하고 있다. E-ARK Web은 SIP 추출 뿐 아니라, 각 단계에서 AIP 저장, AIP에서 DIP로 변환 등을 지원한다. E-ARK의 IP의 요구사항에 따라 필요한 정보를 추출하고, 유효성을 검증하는 등의 일련의 작업을 수행한다.

3.2.2 IP 공통명세서(Common Specification)

E-ARK는 OAIS 참조모형에 기반 하여 모든 IP에 공통으로 적용할 수 있는 공통명세서와 아카이빙 단계별 확장, 적용할 수 있는 구체

화된 IP 기술명세서를 제안하였다. 공통명세서는 아카이빙의 단계별 생성되는 E-ARK SIP, E-ARK AIP, E-ARK DIP에 공통으로 적용되는 기술사항을 말한다. <그림 2>와 같이 공통명세서는 모든 IP에서 요구되는 핵심적인 내용들로, 디지털 아카이빙의 전체적인 프로세스에서 상호운용성을 확보하여 어떠한 기능 영역에서도 다루어지고 구현될 수 있도록 한다. 그리고 세 가지 유형의 IP는 각 기능의 요구사항에 따라 공통명세서를 바탕으로 확장이 가능하다.



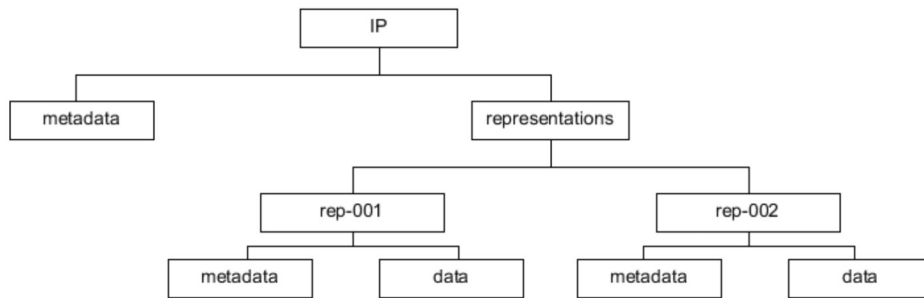
<그림 2> OAIS 참조모형과 E-ARK 모델의 범위

E-ARK의 공통명세서는 정보의 교환, 식별, 확인, 이용, 재이용을 위해 필요한 IP를 구현하기 위한 지침이다. 공통명세서는 1)일반적 요구사항, 2)정보 패키지의 식별을 위한 요구사항, 3)정보 패키지의 구조를 위한 요구사항, 4)정보 패키지의 메타데이터를 위한 요구사항으로 구성되며 내용은 <표 1>과 같다.

IP 공통명세서에 따른 E-ARK의 IP 모델은 <그림 3>과 같이 폴더 구조로 사람과 기계 모두가 이해할 수 있게 논리적, 물리적 모델로 설계되었다. IP는 그 자체로 최상위 폴더가 되고, 그 하위에 metadata와 representations폴더를 필수로 하는 구조를 갖는다. 그리고 동등한 위치에 METS파일을 가지고 있으면서 IP 모델의

〈표 1〉 E-ARK IP 공통명세서

구분	내용
일반적 요구사항	<ul style="list-style-type: none"> - 유형이나 형식에 관계없이 모든 데이터 또는 메타데이터를 포함할 수 있어야 함 - 정보를 교환하기 위한 수단, 방법 또는 도구를 제한해서는 안 됨 - 정보 패키지를 구성하는 데이터 및 메타데이터의 범위를 제한해서는 안 됨 - 확장성(Scalable)이 높아야 함 - 정보 패키지는 기계가독형(machine-readable)으로 표현되어야 함 - 정보 패키지는 인간가독형(human-readable)으로 표현되어야 함 - 데이터에 가장 적합한 방법으로 제시해야 함
정보 패키지의 식별을 위한 요구사항	<ul style="list-style-type: none"> - 정보 패키지의 유형(SIP, AIP, DIP)은 명확히 식별되어야 함 - 정보 패키지에 포함된 데이터와 메타데이터의 콘텐츠 정보 유형을 제시해야 함 - 정보 패키지는 저장소 내에서 고유하고 영구적인 식별자를 가져야 함 - 정보 패키지는 국제적으로 통용되는 고유하고 영구적인 식별자를 가지는 것을 권고
정보 패키지 구조를 위한 요구사항	<ul style="list-style-type: none"> - 정보 패키지는 데이터와 메타데이터가 논리적으로나 물리적으로 서로 분리되어야 함 - 정보 패키지의 구조는 다른 유형의 메타데이터로 구분 - 정보 패키지의 구조는 다른 유형의 메타데이터를 허용해야 함 - 정보 패키지의 구조는 반드시 데이터와 메타데이터 representations의 분리를 허용해야 함 - 정보 패키지의 구조는 추가 논리 구성 요소를 추가 할 수 있는 가능성을 제공해야 함 - 정보 패키지의 구조는 기술적 구현에 구애받지 말고 제시되는 구조를 따라야 함 - 공통 사양 정보 패키지는 어느 시점에서든 단 하나의 구현으로 구현되어야 함
정보 패키지의 메타데이터를 위한 요구사항	<ul style="list-style-type: none"> - 정보 패키지의 메타데이터는 반드시 표준을 기반으로 함 - 공통 사양 정보 패키지의 메타데이터는 명확하게 사용되어야 함 - 공통 사양 정보 패키지는 메타데이터의 추가를 제한하는 것을 지양



〈그림 3〉 E-ARK 일반적 IP 모델

구조를 설명한다. representations 폴더는 하위에 객체와 그 객체에 대한 메타데이터를 가질 수 있도록 하는 구조를 갖는다.

IP 구현을 위한 상세한 요구사항은 〈표 2〉와 같이 정리할 수 있다.

E-ARK는 IP 공통명세서를 통해 아카이빙

의 단계에 따라 갖추어야 할 세분화된 요구사항과 기술내용을 정의하고 있다. 공통명세서는 선입수 단계에서 E-ARK SIP를 생성하고, 입수와 보존 단계에서 입수한 SIP를 AIP로, AIP를 DIP로 변환하기 위한 사전점검사항, 상세한 메타데이터 구조 및 요소 등 IP 구축에 적용할

〈표 2〉 E-ARK IP 구현 요구사항

구분	요구사항
IP 폴더	- 전체 구조를 설명하고 식별할 수 있게 하는 METS.xml 파일을 포함함 - 반드시 metadata와 representations 두 개의 하위 폴더를 가짐
metadata 폴더	- 보존 메타데이터가 있을 경우, 하위에 “preservation” 폴더를 생성하여 해당 메타데이터 파일을 preservation 폴더에 저장함 - 서술 메타데이터가 있을 경우, 하위에 “descriptive” 폴더를 생성하여 해당 메타데이터 파일을 descriptive 폴더에 저장함 - 그 외의 메타데이터가 있을 경우 “other”과 같은 폴더를 만들어서 해당 파일을 저장함
representations 폴더	- 반드시 하위에 “representation” 폴더를 가져야 함 - representation 폴더는 여러 개 생성될 수 있으므로 식별 할 수 있도록 폴더명을 지정(예: rep-001, rep-002)
	representation - 반드시 하위에 data 폴더를 포함 - representation의 구조를 설명하기 위한 METS.xml파일을 가질 수 있음 - 하위에 metadata 폴더를 가질 수 있음

수 있는 방안을 설명하고 있다. 공통명세서에서 E-ARK AIP를 아카이빙에서 있어서 자원의 생애주기 동안 수행되는 보존 조치에 따라 지속적으로 업데이트 되는 ‘진화하는 객체’로 강조하고 있다. E-ARK SIP, DIP, AIP에 대한 세부 요구사항은 다음과 같다.

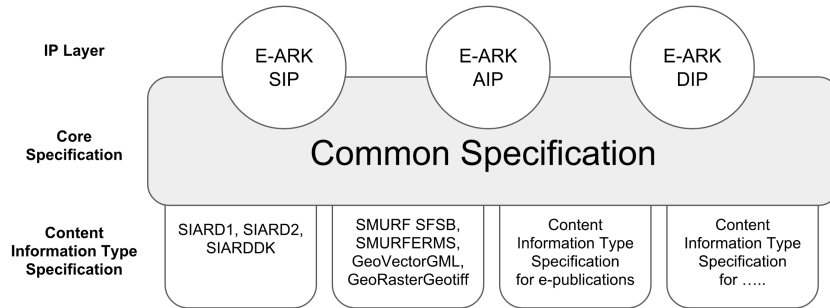
- E-ARK SIP: 공통명세서를 기반으로 하되 정보생산자와 제출 협의서와 제출과정 행위자에 관한 정보 기록에 대한 요구사항 추가
- E-ARK DIP: 다양한 이용자 접근환경을 고려하여 DIP의 정보를 재활용할 수 있도록 필요한 정보 제공
- E-ARK AIP: 공통명세서의 속성을 모두 상속하고 특정 AIP 요구사항에 따라 확장할 수 있는 포맷 명세서 제공

이처럼 E-ARK의 공통명세서는 디지털 보존의 생애주기 전반에 걸쳐 보존과 관리, 검색을 지원할 수 있는 기반으로서 IP간의 상호운

용성을 보장하기 위한 지침이다. E-ARK는 3년간의 프로젝트 진행을 통해 디지털 아카이빙 실무에 적용할 수 있는 구체적인 요구사항과 구현원칙을 개발한 것이 가장 큰 성과라고 할 수 있다.

3.2.3 내용정보 유형 명세서(Content Information Type Specification)

E-ARK SIP, AIP, DIP는 공통명세서를 기반으로 하고, 공통명세서 내에서 내용정보 유형 명세서(Content Information Type Specification)를 제공한다. 내용정보 유형 명세서는 OAIS의 IP 개념에서 보존의 목적이 되는 일련의 정보로, 디지털 정보환경에서 보존해야하는 다양한 내용정보 유형을 이해하기 위해서 필요한 요구사항을 정리한 것이다. 2017년 E-ARK는 2월 디지털 아카이빙에서 입수, 보존, 활용되는 모든 유형의 내용정보를 포괄할 수 있는 정보 유형 명세서를 제안하였으며, 그 세부 내용은 〈그림 4〉와 같다.



〈그림 4〉 공통명세서와 내용정보 유형 명세서와의 관계

내용정보 유형 명세서에는 관계형 데이터베이스 또는 지리정보 데이터와 같은 특정 콘텐츠 유형에 따라 메타데이터, 문서 및 콘텐츠를 어떻게 다루어야 하는지에 대한 세부 요구사항을 포함하고 있다. 현재 E-ARK 프로젝트에서는 공통명세서 내에서 적용 가능한 총 7개 유형의 내용정보에 대한 명세서를 개발하고 검증 단계를 거쳤다.

SIADR1, SIARD2, SIRDDK는 SIARD (Software Independent Archiving or Relational Data Base) 계열의 내용정보에 관한 것으로, 관계형 데이터베이스의 아카이브, 보존, 재사용을 다루고 있다. SMURF SFSB(Semantically Marked-Up Record Format for Simple-File Based records Systems)는 폴더 구조로 구성된 간단한 컴퓨터 파일과 EAD 인코딩 표준 사용법에 대해 설명한다. SMUF EMS(Semantically Marked-Up Records Format for Electronic Records Management Systems)는 기록시스템에서 반출된 레코드의 아카이브에 대해 다룬다. GeoVectorGML과 GeoRaterGeoTIFF는 SMURF SFSB를 기반으로 GML4, GeoTIFF5 형식의 지리 공간 정보를 저장하기 위한 구조 및 메타데이터 요구사항을 다룬다.

3.3 E-ARK 적용 사례

주로 유럽 국가인 덴마크, 노르웨이, 에스토니아, 슬로베니아, 헝가리를 중심으로 다양한 디지털 아카이브 환경에서 E-ARK가 개발한 도구와 요구사항의 적용가능성을 진단하기 위한 사례연구들이 진행되고 있다. 공공과 민간의 아카이빙 영역에서 테스트베드활동을 통해 E-ARK 도구와 지침들에 대하여 실무 관점에서 필요한 보존처리 과정과 기술사항, IP 구현과 관련한 광범위한 경험과 의견을 공유하고 있다.

덴마크 국가기록원(Dannish National Archives)은 데이터 및 정보를 내보내는 형식에 관한 가이드만 제공하고 있으며, 모든 기술 표준이 준수 되어 들어온 데이터와 정보만 입수할 수 있다. 이에 덴마크 국가기록원은 공공기관에서 소유하고 있는 데이터 및 정보들을 기관 시스템과 연계하여 표준을 지켜 쉽게 SIP를 생성 할 수 있도록 할 필요성을 느끼고 E-ARK 프로젝트를 통해 개발된 E-ARK Tool을 테스트 하였다. E-ARK Tool을 사용하는 경우 즉시, 그리고 보존할만한 가치가 있는 모든 데이터를 저비용으로 쉽고 효율적으로 국가기록원으로 전송할 수 있다는 것을 확인하였다.

슬로베니아 국가기록원(National Archives of Slovenia)은 덴마크 국가기록원의 도움을 받아 디지털 지형학적 지도(geographical maps)를 보관하기 위한 지침과 도구를 개발하고 테스트 하였다. 지침과 도구를 개발하는 과정에서 기존 오픈 소스 툴과 E-ARK Tool을 통합하여 아카이빙 프로세스를 지원하고 데이터에 접근할 수 있도록 하였다. 헝가리 국가기록원(National Archives of Hungary) 또한 덴마크 국가기록원의 지원을 받아 데이터를 장기보존하고 접근할 수 있도록, 데이터가 이동하는 과정에 손실을 최소화하고 일관성을 유지하는 테스트를 수행하였다. 이 테스트를 위해 관계형 데이터베이스에서 데이터웨어하우스로 전환되는 과정에서 기능 손실이 없게 설계된 E-ARK ESSArch Tool을 사용하였다.

노르웨이 국가기록원(National Archives of Norway)은 E-ARK 프로젝트 파트너사들 중 하나인 ES Solutions AB사가 개발한 소프트웨어를 사용하여 SIP를 생성하고 레코드를 반입하는 테스트를 진행하였다. 이 SIP는 국제 메타데이터 표준인 OAIS, METS, PREMIS를 따르도록 하였다. 에스토니아 국가기록원(National Archives of Estonia)은 정부 기관의 기록관리 시스템에서 정보를 이관 받고 있다. 에스토니아 국가기록원은 기관이 콘텐츠와 관련 메타데이터를 표준에 맞추어 국가기록원에 제공할 수 있도록 자체 프로세스를 개발하도록 설득하는 것이 어렵다고 판단하고 E-ARK 정보패키지를 구현하는 E-ARK Tool을 테스트 하였다.

민간영역에서도 E-ARK의 지침과 도구를 실제 디지털 아카이브에 적용해보고, 그 가능성에 대한 피드백을 제공하고 있다. 에스토니아

국가기록원은 민간 부분 서비스를 제공하고 요구사항에 대한 풍부한 경험을 가지고 있는 에스토니아의 비즈니스 기록원(Business Archives National Archives of Estonia)과 협업을 통해 사장될 수 있는 상용 또는 개인용 데이터를 보관할 수 있도록 하는 테스트를 진행하였다. 포르투갈의 민간 기업인 Keep Solutions은 영구보존에서 레코드의 생애주기 동안 데이터를 어떻게 패키징해야 하는지에 대해 정의하는 E-ARK IP 명세서를 평가하였다.

이처럼 많은 기관과 국가에서 E-ARK를 적용하여 디지털 아카이브를 개발, 운영하는 사례가 늘어남에 따라, 장기적으로 디지털 아카이빙의 상호운용성을 보장할 수 있는 국제 표준을 개발하는 기초로 채택할 것이라고 기대하고 있다.

4. E-ARK 기반 디지털 아카이브 기능 모델

디지털 아카이빙을 위해서는 수집, 관리, 보존, 이용 등 단계에서 정보자원의 무결성을 확보하고 지속적인 접근을 위해서는 각 단계의 작업이 유기적으로 이루어질 수 있는 체계적인 디지털 아카이빙 모델이 필요하다. 본 연구는 실제 디지털 아카이브 구축 시, 효율적으로 활용할 수 있는 기능 모델 K-ARK를 제안하고자 한다. K-ARK 디지털 아카이빙 모델은 디지털 자원의 생애주기에 기반하여 생산, 수집, 기술, 저장, 보존, 접근 등을 포괄적으로 고려하는 종합적인 프로세스를 제시하고 있다. 각 프로세스에서는 국제 표준 OAIS 참조모형과 E-ARK 프로젝트의 경험에서 도출된 사안들을 종합하

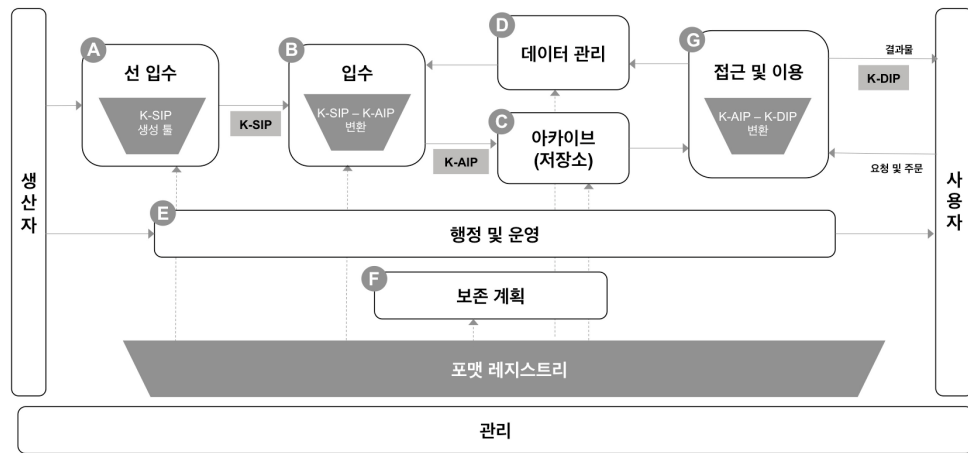
여 실질적으로 디지털 아카이브 구축 시 적용할 수 있는 구체적인 지침을 제공하고자 한다.

4.1 기능 모델과 프로세스

K-ARK에서 제시하는 디지털 아카이빙의 프로세스는 <그림 5>와 같이 일곱 단계로 선입

수, 입수, 보존(아카이브), 데이터 관리, 행정 및 운영, 보존 계획 수립, 접근 및 이용으로 이루어진다.

K-ARK는 OAIS 참조모형과 E-ARK를 기반으로 하여 디지털 아카이브에 공통적으로 적용할 수 있는 모델로, 각 기능 엔티티의 역할은 <표 3>과 같이 정리할 수 있다.



<그림 5> K-ARK 모델 - 개요

<표 3> K-ARK의 기능 엔티티의 역할

엔티티	역 할
A. 선입수	<ul style="list-style-type: none"> - 아카이빙 대상 자원을 식별 - 아카이빙 정책 확인 - 생산자의 제출정보패키지(SIP) 작성이 용이한 정보와 도구(K-SIP 생성물)를 제공 - K-SIP 생성물은 선입수된 SIP의 품질 검수 후 K-SIP 생성 - K-SIP는 아카이빙 전 단계에 걸쳐 데이터공유를 지원할 수 있는 E-ARK의 표준 프레임 워크를 따름
B. 입수	<ul style="list-style-type: none"> - 생성된 K-SIP를 인수하여 데이터 포맷 표준 및 문서화 표준을 준수하는 K-AIP로 변환 - K-AIP는 METS 기반의 메타데이터 프레임워크를 포함하여, 시스템 간의 다양한 유형의 디지털 객체의 전송 및 아카이빙을 지원
C. 보존(아카이브)	<ul style="list-style-type: none"> - K-AIP의 저장 및 유지관리, 검색을 위한 기능과 서비스 제공
D. 데이터 관리	<ul style="list-style-type: none"> - 아카이브 자원을 확인하고 기록하여 문서화하는 기술적인 정보와 아카이브를 관리하는데 사용되는 업무 관리 데이터를 입력, 유지관리, 접근하는 기능 및 서비스 제공
E. 행정 및 운영	<ul style="list-style-type: none"> - K-ARK 시스템의 전반적인 운영 관리
F. 보존 계획 수립	<ul style="list-style-type: none"> - K-ARK 환경을 모니터링하고 저장된 정보에 접근할 수 있도록 권고하며 보존 계획에 대한 서비스와 기능을 제공 - 아카이브 내용을 평가하고 주기적인 업데이트, 마이그레이션, 표준 및 정책, 권장사항 개발, 위험 분석 보고서, 기술 환경 및 요구사항에 대해 다룸
G. 접근 및 이용	<ul style="list-style-type: none"> - K-ARK에 저장된 정보의 존재, 기술, 저장장소, 이용가능성 등을 결정할 뿐만 아니라 이를 지원하고 이용자들이 아카이빙 자원 정보를 요청하고 받을 수 있도록 서비스 함

선입수는 정보 생산자에 관한 정보를 식별하고 확인하는 기능으로, 아카이빙 대상 자원의 콘텐츠 유형과 생산주기, 저작권 협의사항, 콘텐츠 관리 시스템의 특징, 콘텐츠 이름 및 입수 여부 등을 관리한다. 이때 K-SIP 생성 툴을 활용하여 생산자가 콘텐츠 제공을 최대한 쉽고 간결하게 할 수 있도록 도우며 관리자는 아카이빙 자원의 중복과 누락, 저작권 등 추후 발생할 수 있는 문제를 최소화하도록 한다. K-SIP는 E-ARK의 표준에 따라 생성하여 아카이빙 전 단계에 걸쳐 IP간의 상호운용성을 보장하도록 한다. 입수 기능은 K-SIP를 인수하여 그 품질을 확인하고 아카이브 내부에서 저장과 관리를 위하여 K-AIP로 변환한다. K-AIP는 METS 구조메타데이터를 기반으로 하여 아카이브 내에서의 디지털 객체의 전송 및 아카이빙을 지원한다. 보존(아카이브)은 K-AIP의 저장 및 유지 관리, 검색을 위한 기능과 서비스 기능을 수행한다. 데이터 관리는 아카이브 자원을 확인하고 문서화하는 기술정보(Descriptive information)를 관리하고 접근할 수 있게 하는 서비스와 기능을 수행한다. 데이터 관리 기능은 아카이브 데이터베이스 기능을 관리하고, 데이터베이스를 갱신하고 데이터 관리 데이터에 질의를 수행한다. 행정 및 운영은 아카이브 시스템의 전반적인 운영을 위한 서비스와 기능을 수행한다. 보존계획은 환경을 모니터링 하고 아카이빙 자원의 지속적인 접근을 지원할 수 있도록 보존 계획을 수립하는 기능을 수행한다. 접근 및 이용은 K-ARK에 저장된 자원의 존재, 기술, 소재, 입수 가능성을 확인하고 이용자의 요청에 따라 지원하는 서비스와 기능을 수행한다. K-AIP를 이용자의 요청에 따라 필요한 기술정보를 추가

한 K-DIP로 변환한다. 콘텐츠에 최적화된 서비스 플랫폼을 통한 검색 및 이용이 가능하도록 지원하며, 이용자들이 아카이빙 자원을 활용하는데 문제가 없도록 콘텐츠의 저작권 정보를 제시하고 자원의 무결성을 보장하기 위해 이용자의 접근권한과 이력을 관리한다.

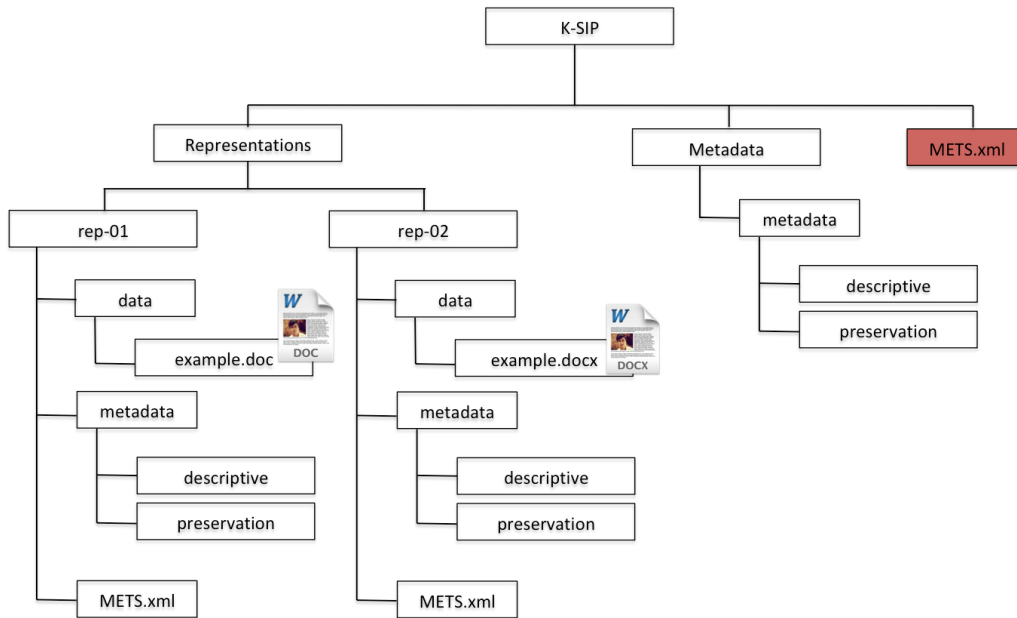
4.2 정보패키지 구성

K-SIP 데이터 모델과 K-AIP 데이터 모델은 3장에서 언급한 E-ARK IP 공통명세서를 기반으로 작성되었다. IP 공통명세서에 따라 필수 요구사항인 representations, metadata 폴더를 가지고 IP의 전체적 구조를 설명하는 METS.xml 포함하는 것을 기본으로 하였다.

4.2.1 K-SIP 데이터 모델

본 연구에서 예시로 작성한 데이터 모델은 Example 문서가 두 가지 형식으로 입수되는 경우에 대해 다루었으며 K-SIP 데이터 모델 구조 예는 <그림 6>과 같다.

Example 문서가 doc와 docx의 두 가지 형식으로 입수되었기 때문에 두 가지 형식에 대한 각각의 Representation이 생성된다. rep-01 폴더는 example.doc 파일에 대한 정보를, rep-02 폴더는 example.docx 파일에 대한 정보를 가진다. 각 Representation 폴더(rep-01, rep-02)에는 data, metadata 폴더를 생성하고 Representation 구조를 설명하는 METS.xml 파일을 생성하였다. 실제 입수된 파일은 data 폴더에 저장되며, 입수된 파일에 대한 보존, 기술 메타데이터는 metadata 폴더에 저장될 수 있도록 하였다.



〈그림 6〉 K-ARK K-SIP 데이터 모델 구조 예

K-SIP 모델의 최상위 폴더에 있는 METS 파일은 아래의 그림과 같이 구성될 수 있다. METS의 ①<dmdSec>에는 K-SIP에 대한 서술 메타데이터 파일 위치 정보, ②<fileSec>에는 각각의 representation을 설명하는 METS파일 즉,

rep-01, rep-02 폴더 내 최상위 METS 파일에 대한 위치 정보를 작성한다. ③<structureMap>에는 K-SIP 데이터 모델의 위계 정보를 작성한다(〈그림 7〉 참조).

```

<mets:dmdSec ID="K-SIP_1_dmd1">
  <mets:mdRef CHECKSUM="a7a50ddad4c741034cf72f676266504c3be14076cd75446352392d535087e91e" CHECKSUMTYPE="SHA-256"
    CREATED="2012-04-26T12:00:00+01:00" LOCTYPE="URL" MDTYPE="EAD" MIMETYPE="text/xml" SIZE="1109" xlink:href="metadata/descriptive/ead.xml"
    xlink:type="simple"/>
</mets:dmdSec>
<mets:fileSec>
  <mets:fileGrp>
    <mets:file CHECKSUM="TODO" CHECKSUMTYPE="SHA-256" CREATED="2015-05-01T01:00:00+01:00" ID="rep_01" MIMETYPE="application/xml" SIZE="0" USE="representation">
      <mets:Flocat LOCTYPE="URL" xlink:href="representations/rep_01/METS.xml" xlink:type="simple"/>
    </mets:file>
    <mets:file CHECKSUM="TODO" CHECKSUMTYPE="SHA-256" CREATED="2015-05-01T01:00:00+01:00" ID="rep_02" MIMETYPE="application/xml" SIZE="0" USE="representation">
      <mets:Flocat LOCTYPE="URL" xlink:href="representations/rep_02/METS.xml" xlink:type="simple"/>
    </mets:file>
  </mets:fileGrp>
</mets:fileSec>
<mets:structMap ID="sm1" LABEL="Profilestructmap" TYPE="physical">
  <mets:div DMDID="K-SIP_1_dmd1">
    <mets:div LABEL="representations">
      <mets:fptr FILEID="rep_01"/>
      <mets:fptr FILEID="rep_02"/>
    </mets:div>
  </mets:div>
</mets:structMap>
  
```

〈그림 7〉 E-ARK K-SIP 최상위 METS 파일 예시

4.2.2 K-AIP 데이터 모델

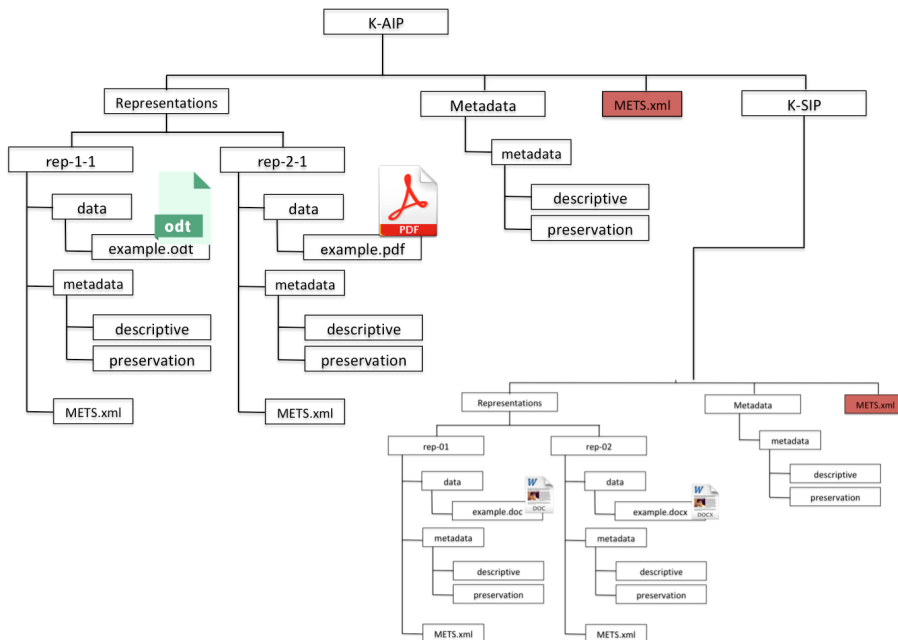
K-AIP는 K-SIP에서 입수된 example.doc, example.docx 파일에 대한 장기보존을 위해 생성되며, K-SIP 데이터 모델에 포함된 메타데이터 구조 정보 등을 제공해 주어야 한다. 따라서 K-AIP 데이터 모델은 Representation, Metadata 폴더, METX.xml을 필수적으로 가지고 K-SIP 폴더를 그대로 상속한다. K-AIP 데이터 모델 구조는 <그림 8>과 같다.

K-SIP에서 입수된 두 가지 형식의 문서가 각각 odt와 pdf 형식으로 변환되었을 때, 이 두 가지 형식에 대한 각각의 Representation이 생성된다. rep-1-1은 example.doc 파일에서 example.odt로 변환된 정보를 rep-2-1은 example.docx 파일에서 example.pdf 파일로 변환된 정보를 가진다. K-SIP와 동일하게 각각의 Representation

폴더 하위의 data 폴더에는 변환된 odt, pdf 형식의 파일이 저장되고, metadata 폴더에는 변환된 파일에 대한 서술, 보존 메타데이터가 저장된다.

K-SIP데이터 모델을 바탕으로 작성된 K-AIP 최상위 METS 파일은 <그림 9>와 같다.

K-AIP 최상위 METS 파일의 ①<dmdSec>에는 K-AIP에 대한 서술 메타데이터 위치 정보를 ②<amdSec>에는 K-AIP에 대한 보존 메타데이터 파일 위치 정보를 작성한다. ③<fileSec>에는 K-SIP정보를 가지고 있는 K-SIP 최상위 METS 파일, 보존을 위해 변환한 자원의 정보를 가지고 있는 rep-1-1, rep-2-1 폴더 내 최상위 METS 파일에 대한 위치 정보를 작성하고, ④<structureMap>에는 K-AIP의 데이터 모델 위계 정보를 작성한다.



<그림 8> K-ARK K-AIP 데이터 모델 구조 예시

```

< mets:dmdSec ID="K-AIP_dmd1">
  < mets:mdRef CHECKSUM="a7a50ddad4c741034cf72f676266504c3be14076cd75446352392d535087e91c" CHECKSUMTYPE="SHA-256" CREATED="2012-04-26T12:00:00+01:00" LOCTYPE="URL" MDTYPE="EAD" MIMETYPE="text/xml" SIZE="1109" xlink:href="metadata/descriptive/ead.xml"
    xlink:type="simple"/>
</ mets:dmdSec>
1

< mets:amdSec ID="K-AIP_ami1">
  < mets:digiprovMD ID="ID4566af74-0f7b-11e5-a6c0-1697f925ec7c">
    < mets:mdRef CHECKSUM="TOD0" CHECKSUMTYPE="SHA-256" CREATED="2012-04-26T12:00:00+01:00" LOCTYPE="URL" MDTYPE="PREMIS" MIMETYPE="text/xml" SIZE="1109"
      xlink:href="metadata/preservation/premis.xml" xlink:type="simple"/>
  </ mets:digiprovMD>
</ mets:amdSec>
2

< mets:fileSec>
  < mets:fileGrp>
    < mets:file CHECKSUM="TOD0" CHECKSUMTYPE="SHA-256" CREATED="2015-11-30T01:00:00+01:00" ID="K-AIP_K-SIP" MIMETYPE="text/xml" SIZE="2432" USE="submission">
      < mets:Flocat LOCTYPE="URL" xlink:href="K-SIP/METS.xml" xlink:type="simple"/>
    </ mets:file>
  </ mets:fileGrp>
  < mets:fileGrp>
    < mets:file CHECKSUM="TOD0" CHECKSUMTYPE="SHA-256" CREATED="2015-05-01T01:00:00+01:00" ID="rep_1-1" MIMETYPE="application/xml" SIZE="0" USE="representation">
      < mets:Flocat LOCTYPE="URL" xlink:href="representations/rep_1-1/METS.xml" xlink:type="simple"/>
    </ mets:file>
  </ mets:fileGrp>
  < mets:fileGrp>
    < mets:file CHECKSUM="TOD0" CHECKSUMTYPE="SHA-256" CREATED="2015-05-01T01:00:00+01:00" ID="rep_2-1" MIMETYPE="application/xml" SIZE="0" USE="representation">
      < mets:Flocat LOCTYPE="URL" xlink:href="representations/rep_2-1/METS.xml" xlink:type="simple"/>
    </ mets:file>
  </ mets:fileGrp>
</ mets:fileSec>
3

< mets:structMap ID="sm_2" LABEL="Profilestructmap" TYPE="physical">
  < mets:div ADMID="K-AIP_ami1" DMDID="K-AIP_dmd1" />
  < mets:div LABEL="submission">
    < mets:fptr FILEID="K-AIP_K-SIP" />
  </ mets:div>
  < mets:div LABEL="representations">
    < mets:fptr FILEID="rep_1-1" />
    < mets:fptr FILEID="rep_2-1" />
  </ mets:div>
</ mets:structMap>
4

```

<그림 9> E-ARK K-AIP 최상위 METS 파일 예시

4.3 포맷레지스트리 활용방안

포맷레지스트리는 K-ARK 모델에서 디지털 아카이빙의 선입수, 보존 및 접근 등 전반적인 단계에 걸쳐 안전하고 신뢰할 수 있는 보존 전략으로 활용될 수 있다. 선입수 단계에서는 하드웨어, 소프트웨어 요소들이 결정되는 중요한 시점으로 생산자가 데이터의 생산에 앞서 기록의 장기보존 지속가능성을 고려하여 파일 포맷을 선택하도록 할 수 있다. 보존 및 접근 단계에서는 입수-저장-접근에 이르는 기본적인 프로세스와 파일 포맷레지스트리를 연계하여 SIP, AIP 정보패키징 변환을 지원하고 관련 기술메타데이터를 추출하여 레지스트리에 등록 관리하도록 하여 활용할 수 있다. 보존계획 기능 수행 시에는 포맷 레지스트리를 통해 자원의 기술변화와 노후화를 모니터링하고 아

카이빙 자원에 미칠 영향력을 평가한다. 이런 과정을 기반으로 마이그레이션, 에뮬레이션 등 그에 맞는 적절한 보존전략을 수립하여 수행함과 동시에 새로 생성되는 기술메타데이터를 레지스트리에서 등록, 관리함으로써 효율적인 전략으로 활용될 수 있을 것이다.

5. 결론 및 제언

OAIS 참조모형은 디지털 정보를 장기적으로 보존하는데 필요한 개념적 프레임워크로 OAIS를 준수하는 것은 신뢰할 수 있는 디지털 아카이브 구축의 기본조건이다. OAIS의 기능 모델과 정보모델은 디지털 정보를 보존하는데 기반이 되는 표준적인 정보 모델링으로 널리 사용되고 있다. 그러나 OAIS 참조모형은 개념적

프레임워크이기 때문에 디지털 아카이브의 실질적인 보존 실무분야에서 적용하기 위해서는 구체적인 지표와 기준이 요구된다. 이에 본 연구는 E-ARK을 기반으로 실제 디지털 아카이브 구축 시, 효율적으로 활용할 수 있는 기능 모델을 제안하고자 하였다.

본 연구에서 제안한 디지털 아카이빙 모델은 디지털 자원의 보존과 관련된 여러 단계가 상호 연결되어서 종합적으로 이루어지는 프로세스를 제시하며, 실제 디지털 아카이브 구축 시 각 프로세스에서 적용할 수 있는 기술사항을 제안하고자 하였다. 디지털 아카이브의 핵심 기능 영역을 중심으로, 각 영역에서 디지털 아카이브의 세부적인 프로세스와 단계에 적용되는 기술사항(메타데이터패키지)을 구체화 시키는 데 초점을 두고 있다.

제안한 K-ARK 아카이빙 모델은 E-ARK에서 제시한 선입수(Pre-Ingest) 단계를 추가 반영함으로써 생산자와 아카이브 관리자 간 데이터와 메타데이터의 전달을 용이하게 하고, 디지털 아카이빙의 전 단계에서 생성되는 IP간의 상호운용성을 확보하고자 하였다. 메타데이터와

파일을 패키징하는데 최적화된 표준 METS 컨테이너를 활용함으로써 구체적인 보존패키지 생성 방안을 제시하여 아카이브에서 빈번하게 발생하는 파일의 변경 및 이력변화를 효과적으로 관리할 수 있게 하였다. 또한 K-ARK 모델 내의 각 기능적 엔티티에서 보존 전략으로써 선입수단계로부터 보존 및 접근까지 전체 프로세스에서 포맷 레지스트리의 활용방안을 제안하였다.

본 연구에서 제안하고 있는 K-ARK 모델은 디지털 아카이빙의 전체적인 프로세스가 원활하게 수행될 수 있도록, 각 단계에서 실제로 IP를 구축하면서 고려해야할 구체적인 사항들을 종합한 포괄적 모델이라는 의의가 있다. 추후 실질적인 디지털 아카이브 구축을 통해서 자원의 생산, 관리, 활용, 보존 등 다양한 아카이빙의 측면에서 모델에 대한 평가가 수행되어야 할 것이다. 제안한 디지털 아카이브 모델은 기관간의 정보자원의 교환, 공유, 접근 등 아카이빙 프로세스에 있어서 유기적인 협력을 도모할 수 있는 지침으로 활용될 수 있을 것이다.

참 고 문 헌

- 국가기록원 (2008). 디지털 포맷 및 애플리케이션 기술정보은행(DFR) 프로토타입 개발. 대전: 국가기록원.
- 국가기록원 (2013). 행정기관 전자기록물 재현기술 연구 및 프로토타입 개발. 대전: 국가기록원.
- 박옥남 (2012). PRMIS 기반 보존 메타데이터 요소 개발에 관한 연구: 국립중앙도서관 디지털 자료물 중심으로. 한국문헌정보학회지, 46(2), 83-113. <http://dx.doi.org/10.4275/KSLIS.2012.46.2.083>
- 박지화 (2016). 국내 포맷 레지스트리 운영 현황 및 개선에 관한 연구. 석사학위논문, 동의대학교 대학

원, 문헌정보·사회과.

- 박지화, 정영미 (2015). 포맷 레지스트리 운영 사례 분석. 한국도서관·정보학회 하계 학술발표대회, 135-148.
- 신재민, 광승진 (2013). 디지털 콘텐츠 아카이빙 정책수립을 위한 문헌 및 사례 고찰. 사회과학연구, 24(1), 305-330. <http://dx.doi.org/10.16881/jss.2013.01.24.1.305>
- 안대진, 임진희 (2016). 디지털 아카이브 시스템 구축을 위한 공개 소프트웨어 활용방안 연구. 정보관리학회지, 33(3), 345-370. <http://dx.doi.org/10.3743/KOSIM.2016.33.3.345>
- 유영수 (2006). 전자기록의 장기보존을 위한 포맷등록관리 시스템 개발 방안 연구: 영국 PRONOM Service를 중심으로. 석사학위논문, 명지대학교 대학원, 기록관리학과.
- 유영수 (2007). 전자기록관리를 위한 포맷등록시스템 개발 연구. 한국기록관리학회지, 7(1), 145-170.
- 이소연 (2002). 디지털 아카이빙의 표준화와 OAIIS 참조모형. 정보관리연구, 33(3), 45-68.
- 이소연 (2009). 디지털 아카이브의 장기보존 기능에 대한 연구. 한국정보관리학회 학술대회 논문집, 16, 73-78.
- 이승민 (2015). 디지털 아카이빙을 위한 보존 메타데이터 패키지 구축. 정보관리학회지, 32(3), 21-47. <http://dx.doi.org/10.3743/KOSIM.2015.32.3.021>
- DLM Archival Standards Board (2016, December 21). Common Specification For Information Packages Ver 0.17. Retrieved from <http://www.eark-project.com/resources/specificationdocs/67-e-ark-draft-common-specification-ver-017/file>
- DLM Archival Standards Board (2017, February 1). E-ARK DIP Specification version:1.0. Retrieved from http://www.dasboard.eu/images/Specifications/DIP/DIP_10_v2.pdf
- E-ARK (2014). Retrieved from <http://www.eark-project.com/>
- E-ARK General Model v2.3 (n.d.). DLM forum foundation. Retrieved from <http://kc.dlmforum.eu/gm>
- Paulic, A. (2015). European Archival Records and Knowledge Preservation (E-ARK) Project. Goals and Achievements: an Overview. Retrieved from <http://www.ias-trieste-maribor.eu/fileadmin/atti/2015/Paulic.pdf>
- Wilson, A., Ferreira, J. M., Kuldar, A., Nielsen, A. B., & Tommerholt, P. M. (2016). E-ARK: Harmonising pan-European archival processes to ensure continuous access to e-government records and information. Retrieved from <https://repositorium.sdum.uminho.pt/bitstream/1822/42478/1/E-ARK%20eGov%20Conference%20-%202011.pdf>

• 국문 참고문헌에 대한 영문 표기
(English translation of references written in Korean)

- An, Dae-Jin, & Yim, Jin-hee (2016). A study on using open software for building a digital archive system. *Journal of the Korean Society for Information Management*, 33(3), 345-370. <http://dx.doi.org/10.3743/KOSIM.2016.33.3.345>
- Lee, Seungmin (2015). Construction of preservation metadata package for digital archiving. *Journal of the Korean Society for Information Management*, 32(3), 21-47. <http://dx.doi.org/10.3743/KOSIM.2015.32.3.021>
- Lee, So-Yeon (2002). Standardization of digital archiving and OAIS reference model. *Journal of Information Management*, 33(3), 45-68.
- Lee, So-Yeon (2009). Long-term preservation functionality in digital archives. *Proceedings of the Korean Society for Information Management*, 16, 73-78.
- National Archives of Korea (2008). DFR prototype development for digital formats and applications. Daejeon: National Archives of Korea.
- National Archives of Korea (2013). A study on the reproduction technology and the prototype for the electronic records of administrative agency. Daejeon: National Archives of Korea.
- Park, Ji Hwa (2016). A study on the operating status and improvement of format registry in Korea. M. A. thesis, Dong Eui University. Busan, Korea.
- Park, Ji Hwa, & Jung, Young Mi (2015). A study on the format registry operation practice. *Proceedings of Korean Library and Information Science Society*, 2015(5), 135-148.
- Park, Ok Nam (2012). A study on developing preservation metadata based on PREMIS focusing on digital data in national library of Korea. *Journal of the Korean Library and Information Science Society*, 46(2), 83-113. <http://dx.doi.org/10.4275/KSLIS.2012.46.2.083>
- Shin, Jaemin, & Kwak, Seungjin (2013). A review of literature and cases for developing digital content archives. *Journal of Institute for Social Sciences*, 24(1), 305-330. <http://dx.doi.org/10.16881/jss.2013.01.24.1.305>
- Yu, Young-Soo (2006). Study on the development of format registry system for long-term preservation of electronic records-focusing on PRONOM service by the UK national archives. M. A. thesis, Myungji University. Seoul, Korea.
- Yu, Young-Soo (2007). A study on the development of format registry for ERM. *Journal of Records Management & Archives Society of Korea*, 7(1), 145-170.

