

데이터 리포지토리의 보존 정책 프레임워크에 관한 연구*

- CoreTrustSeal 인증을 중심으로 -

A Study on the Preservation Policy Framework of Data Repository: Focusing on CoreTrustSeal Certification

이 헤 림 (Hea Lim Rhee)**

목 차

- | | |
|-------------------------------------|--------------------------|
| 1. 서 론 | 4. 디지털 보존 정책 프레임워크의 구성요소 |
| 2. 디지털 보존 정책 관련 연구 | 5. 결 론 |
| 3. CoreTrustSeal 인증기관의 디지털
보존 정책 | |

초 록

본 연구의 목적은 데이터 리포지토리가 갖추어야 할 디지털 보존 정책의 필수 구성요소를 파악하여 디지털 보존 정책의 프레임워크를 개발하는 것이다. 이를 위해 본 연구에서는 2023년도에 CoreTrustSeal에서 인증을 받은 데이터 리포지토리들의 디지털 보존 정책을 비교 분석하여 디지털 보존 정책 프레임워크에 필요한 구성요소를 추출하고, 각 구성요소에서 어떠한 내용을 포함하고 있는지를 연구 결과로 제시하였다. 본 연구에서 제시한 디지털 보존 정책 프레임워크는 CoreTrustSeal에서 인증을 받고자 하는 기관들뿐만 아니라 CoreTrustSeal 인증을 신청할 계획이 없는 국내외 기관들에서도 그들의 디지털 보존 정책을 개발할 때 기초자료로 이용할 수 있을 것으로 기대된다.

ABSTRACT

This study endeavors to delineate the essential components of digital preservation policies requisite for data repositories. Drawing from an examination of digital preservation policies from data repositories accredited by CoreTrustSeal in 2023, this study unveils the pivotal components intrinsic to a digital preservation policy framework, subsequently elucidating the content each component encompasses. The proffered framework is anticipated to serve as a foundational reference for institutions aiming to craft their digital preservation policies, transcending the confines of those solely pursuing CoreTrustSeal certification, to encompass both domestic and international institutions irrespective of their intent to seek CoreTrustSeal certification.

키워드: 데이터 리포지토리, 디지털 보존, 보존 정책, 정책 프레임워크, CoreTrustSeal
Data Repository, Digital Preservation, Preservation Policy, Policy Framework,
CoreTrustSeal

* 이 논문은 2023년도 한국과학기술정보연구원(KISTI)의 기본사업으로 수행된 연구임.
(과제번호: (KISTI)K-23-L01-C03, (NTIS)1711198423)

** 한국과학기술정보연구원 책임연구원(rhee.healim@kisti.re.kr / ISNI 0000 0004 6809 9422)
논문접수일자: 2023년 10월 25일 최초심사일자: 2023년 10월 30일 게재확정일자: 2023년 11월 8일
한국문헌정보학회지, 57(4): 119-138, 2023. <http://dx.doi.org/10.4275/KSLIS.2023.57.4.119>

* Copyright © 2023 Korean Society for Library and Information Science
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (<https://creativecommons.org/licenses/by-nc-nd/4.0/>) which permits use, distribution and reproduction in any medium, provided that the article is properly cited, the use is non-commercial and no modifications or adaptations are made.

1. 서론

1.1 연구 배경

디지털 시대에 있어 데이터는 다양한 분야에서 혁신을 촉진하고, 데이터에 입각한 의사 결정을 뒷받침하고, 정책 프레임워크를 작성하는데 도움을 주는 등 현대 학계 및 산업의 무수한 발전을 위해 구조적인 기반 역할을 하고 있다.

데이터가 학술적 연구와 과학기술의 발전에서 수행하는 역할이 커질수록, 데이터 보존의 필요성과 중요성도 함께 증가하고 있다. 디지털 보존은 보존 가치가 있는 데이터의 선정, 복원, 디지털화, 출처와 상호관련성 유지 등을 수행하는 업무이다(Brown, 2013). 이러한 디지털 보존은 필요한 기간만큼 지속적이고 안정적으로 데이터에 접근하여 이용할 수 있도록 보장하는 데 필요한 일련의 관리 작업으로 구성된다.

리포지토리는 지속적인 가치가 있는 아이템(레코드, 데이터 등)의 관리와 저장에 중점을 둔 기관을 말한다(Society of American Archivists, 2023). 즉, 리포지토리의 개념 안에는 이미 소장 아이템을 보존할 책임과 역할을 시사하고 있다. 따라서, 데이터 리포지토리는 데이터를 저장, 관리, 배포하기 위하여 중앙 집중적이고 조직적이고 접근 가능한 플랫폼 역할을 하면서 디지털 보존의 영역에서 중추적인 역할을 수행할 책임이 있다.

특히, 데이터 리포지토리는 자관의 디지털 보존 정책을 제정할 필요가 있다. 정책은 어떤 업무를 수행할 때 취해야 할 조치에 대한 로드맵을 제공하며, 어떤 프로그램을 성공적으로

실행하는 데 있어 중요한 역할을 한다. 이처럼, 디지털 보존 정책의 존재는 디지털 보존 프로그램을 수행하는 데 핵심적인 역할을 한다. 이러한 정책을 가진 리포지토리는 조직적인 방식으로 디지털 보존 활동을 수행할 수 있다(Ahmad & Rafiq, 2022). 데이터 리포지토리는 현재와 미래의 이용자들이 자관의 데이터를 신뢰하고 재이용할 수 있도록 디지털 보존의 기반과 지침이 될 수 있는 효율적이고 현실적인 디지털 보존 정책을 작성하고 그것을 실행할 필요가 있다.

미국, 영국, 호주 등 다른 선진국에 비해 상대적으로 적은 수의 데이터 리포지토리를 보유하고 있는 우리나라는 최근 몇 년 전부터 정부 출연연구원들이 중심이 되어 자관의 데이터 리포지토리를 구축하고 있다. 특히, 한국과학기술정보연구원(Korea Institute of Science and Technology Information: 이하 KISTI)에서는 정부출연연구원들을 중심으로 각 기관이 자관의 연구데이터를 직접 관리할 수 있도록 NaRDA를 통해 '기관 데이터 리포지토리'를 배포하고 있다. 그러나, 아직 국내 데이터 리포지토리들은 설립된 지 얼마 안 되어 자관의 디지털 보존 정책을 가지고 있는 경우는 사실상 거의 없다고 말할 수 있다. 실제로, 전국 4년제 대학도서관 중 디지털 콘텐츠를 구축하고 있는 102개 대학도서관을 대상으로 디지털 보존 실태를 조사한 연구 결과, 대부분의 대학도서관들이 자관의 장기적인 디지털 보존 정책을 수립하여 시행하고 있지 않은 상황을 보여주었다(서은경, 2016). 이러한 결과는 국내에서 설립된 지 얼마 안 되는 데이터 리포지토리들이 자관의 디지털 보존 정책을 가지고 있지 않을 것으로 추측하

게 한다. 더 나아가, 외국의 데이터 리포지토리와 달리, 국내 데이터 리포지토리들은 자관의 웹사이트에서 자관의 디지털 보존 정책을 제시하고 있지 않다는 점에서 그러한 추측을 뒷받침하고 있다.

국내에서는 데이터를 효율적으로 보존할 수 있는 기반이 되는 디지털 보존 정책에 관한 연구가 미비하다. 특히 '디지털 아카이빙'이나 '디지털 큐레이션'의 정책에 관한 연구는 소수 있으나, '디지털 보존'의 정책이 주제인 연구는 거의 없다. 더 나아가 데이터 리포지토리들이 자관의 데이터를 장기 보존하고 디지털 보존 활동을 향상하기 위해서는 높은 수준의 디지털 보존 정책 프레임워크가 필요하다(Beagrie, Williams, & Rettberg, 2008; Oliver & Harvey, 2016). 그런데도, 국내에서는 도서관이나 아카이브에서 전자 레코드, 디지털 문서 등 디지털 정보자원의 보존을 위한 디지털 보존에 관한 연구가 수행되었으나, 이와는 달리 데이터 리포지토리를 위한 데이터 중심의 디지털 보존 정책에 관한 연구는 현재까지도 미비한 실정이다. 더 세부적인 주제로, 디지털 보존 정책을 개발하는데 필요한 프레임워크나 그것을 구성하는 요소에 관한 연구가 해외에서는 소수 존재하나, 국내에서는 거의 이루어지지 않았다.

1.2 연구 필요성 및 목적

최근 들어 오픈 사이언스의 중요성이 부각되면서 국내 및 국제 기금 제공자들은 그들의 프로젝트에서 생산된 데이터의 장기적인 보존과 이용을 위해 데이터 관리 정책을 의무화하려는 움직임이 점차 증가하고 있다. 그러한 움직임

중의 하나는 데이터 관리 계획을 제출하고, 프로젝트 결과 생산된 데이터를 신뢰할 수 있는 리포지토리에 기탁할 것을 요구하고 있다. 더 나아가, 생산된 데이터가 재사용될 수 있도록 데이터의 공유가 중요해지고 있는 현 시점에서, 데이터 공유를 하려는 기관들은 자관의 데이터가 신뢰할 수 있는 데이터 리포지토리에서 보존되고 있음을 보여줄 필요가 있다(CoreTrustSeal, 2023a).

이러한 상황에서, KISTI의 연구데이터공유센터는 국가와 국민을 위해 운영하고 있는 국가연구데이터플랫폼인 DataON의 신뢰성을 보여주고 소장하고 있는 데이터의 우수한 품질을 보장하여 소장 데이터의 공유와 재이용을 촉진하고자 CoreTrustSeal(이하 CTS)에 주목했다. CTS는 데이터 리포지토리 서비스의 신뢰성, 지속가능성, 상호운용성과 관련된 특정 기준을 충족하는 데이터 리포지토리에 부여되는 국제적인 인증이다. Data Seal of Approval(DSA)와 World Data System(WDS)의 협력을 통해 탄생한 CTS는 데이터 기탁자와 이용자 모두가 신뢰할 수 있는 데이터 리포지토리를 쉽게 파악할 수 있도록 도와주는, 세계적으로 인정받는 표준 역할을 하고 있다. 이에 KISTI의 연구데이터공유센터는 2022년 국내 최초로 CTS에 인증 신청서를 제출하였다. 정확히 말하면, 온라인에서 제공하고 있는 CTS의 '신뢰할 수 있는 데이터 리포지토리 요구사항(Trustworthy Data Repositories Requirements)'의 각 요구사항에 대하여 신청하는 리포지토리의 현황에 대해 답변하고, 그러한 답변에 대한 근거나 답변에 대해 추가로 참고할 만한 웹페이지가 있는 경우 심사위원들이 확인할 수 있도록 링크

를 제공해야 한다. 이러한 요구사항은 본 논문의 '3. CoreTrustSeal 인증기관의 디지털 보존 정책' 부분에 상세하게 설명되어 있다.

KISTI 연구데이터공유센터는 2022년에 제출한 'CoreTrustSeal 신뢰할 수 있는 데이터 리포지토리 요구사항(CoreTrustSeal Trustworthy Data Repositories Requirements 2020-2022)'에 대한 심사결과를 2023년 2월에 받았고, 각 항목마다 2명의 심사위원들의 의견이 쓰여 있었다. 그러한 심사 결과를 반영하여 연구데이터공유센터는 새로운 버전의 'CoreTrustSeal 신뢰할 수 있는 데이터 리포지토리 요구사항(CoreTrustSeal Trustworthy Data Repositories Requirements 2023-2025)'을 사용하여 2차 신청서를 2023년 5월에 제출했고 9월에 2차 심사 결과를 받았다. 2차 심사 결과에서는 소수의 항목에 대해 보다 자세한 답변과 관련 링크를 제공해 줄 것을 요청받았다. 2차 심사 결과를 반영한 3차 신청서를 2023년 9월에 다시 보내어 현재는 심사 결과를 기다리는 중이다. 두 번의 심사 결과에서 심사위원들이 가장 많이 지적하고 해당 내용을 추가해 달라고 요구한 항목이 바로 '보존 계획' 항목이었다.

CTS는 'CoreTrustSeal 신뢰할 수 있는 데이터 리포지토리 요구사항(CoreTrustSeal Trustworthy Data Repositories Requirements 2023-2025)'의 'Introduction' 부분에서 "CoreTrustSeal 신청자는 장기 보존 서비스를 제공해야 한다"고 서술할 만큼 '보존'을 강조하고 있다. 그럼에도 불구하고, KISTI 연구데이터공유센터는 데이터의 보존을 위해 가장 중요한 기본이면서 보존 계획의 방향을 제시할 디지털 보존 정책이 없었고, 이에 디지털 보존 정책 수립의 필요성

을 인식했다. 디지털 보존 정책을 처음으로 작성할 때는 먼저 프레임워크를 만들 필요가 있으며, 이러한 프레임워크를 만들 때는 명확하고 포괄적이며 실행가능한 구성요소의 개요가 포함되어야 한다(Research Libraries Group, 2002).

이에 본 연구에서는 데이터 리포지토리가 갖추어야 할 디지털 보존 정책의 필수 구성요소를 파악하여 디지털 보존 정책의 프레임워크를 제안하고자 한다. 본 연구는 연구 결과로 생산된 프레임워크를 통해 KISTI 연구데이터공유센터뿐만 아니라 국내외 데이터 리포지토리들이 자관의 디지털 보존 정책을 수립하는 데 도움을 주고자 한다. 본 연구는 프레임워크 개발을 위하여 CTS 인증을 받은 데이터 리포지토리들의 디지털 보존 정책을 분석 비교하여 프레임워크의 필수 구성요소와 각 요소에서 서술해야 할 내용을 제시함으로써 CTS에 인증 신청을 하려는 기관들에게 도움을 줄 것이다. 또한 본 연구의 분석 비교 대상이 되는 디지털 보존 정책은 전 세계에 흩어져 있는 해외 데이터 리포지토리들의 것으로 보편성을 고려했을 때, CTS 인증과 상관없는 국내외 기관들도 그들의 디지털 보존 정책을 개발할 때 본 연구의 결과물을 참고할 수 있을 것으로 기대한다.

2. 디지털 보존 정책 관련 연구

2.1 국내 연구

디지털 보존 정책과 관련된 국내 연구는 아직 미비한 상황이다. 이러한 상황은 특히 이소연(2013)의 국내 디지털 보존 연구의 동향 분

석 연구와 그 후 약 10년 뒤에 수행된 나연희와 박희진(2022)의 국내기록보존 연구 동향 분석에서 잘 나타난다. 먼저, 이소연은 2000년부터 2013년까지 발간된 분석대상 학술지에서 디지털 아카이브와 보존 관련 키워드를 포함하는 연구논문 115편을 발견하였다. 이 중 디지털 보존에 대한 국가 차원의 정책을 다룬 논문은 4편이었고, 개별 기관의 정책을 다룬 논문은 3편이었다. 이소연의 논문에서도 밝혔듯이, 다양한 분야(기록관리학, 컴퓨터 공학, 문헌정보학 등)에서 디지털 보존 관련 연구가 수행되고 있었는데, 기본 개념이 서로 조금씩 다르므로 향후 연구에서는 보존의 정확한 용어를 학술논문에서 사용할 것을 권장하고 있다. 실제로 이 연구에서는 '디지털 아카이브와 보존'에 관련된 논문을 선별하였는데, 디지털 아카이브를 다루면서 보존 내용이 있으면, 분석 논문에 포함했기 때문에 선별된 115편 논문 중에는 '디지털 보존'보다 '디지털 아카이브'에 초점을 둔 논문들이 다수 있었으며, 이 중 디지털 보존 정책에 대해 언급하고 있으나 정책에 초점을 두고 연구한 논문은 별로 없었다.

나연희와 박희진(2022)은 2000~2021년 국내 기록관리학과 문헌정보학 분야의 학술논문 463편을 분석하여 국내 기록보존의 연구 동향을 파악하고자 하였다. 이를 위해, 학술논문의 저자가 자신의 논문에 직접 부여한 키워드를 가지고 그 키워드들의 연결 중심성 분석과 매개 중심성 분석을 수행하였다. 수집한 학술논문의 학술지 게재시기에 따라 제1구간(2000~2007년), 제2구간(2008~2014년), 제3구간(2015년~2021년)으로 나누어 분석한 결과를 발표하였다. 그 결과, 전 구간에서 '전자기록'과 '장기보

존', 제1구간에서는 'OAIS 참조 모형', 제2구간에서는 '전자기록', 제3구간에서는 '기록관리기준표'와 '장기보존'이 핵심 키워드로 나타났다. 전 구간에서 가중치 4회 이상 나타난 상위 키워드 쌍은 총 22쌍으로 '디지털 보존'과 '디지털 아카이빙'은 2위, '디지털 보존'과 '디지털 아카이브', '디지털 보존'과 '웹아카이빙', '디지털 보존'과 '디지털 큐레이션'은 모두 공동 12위를 차지하였다. 그러나 나연희와 박희진(2022)이 분석한 학술논문 중에서 '디지털 보존 정책'은 핵심 키워드가 아니었다.

앞의 두 연구에서도 추론할 수 있듯이, 우리나라 논문 중에 '디지털 보존 정책'과 관련된 논문은 소수 존재한다. 예를 들어, 디지털 보존과 관련된 정책을 주제로 하는 학술논문으로는 '디지털 콘텐츠의 무결성 유지를 위한 장기적인 보존 정책에 관한 연구(정영미, 윤화목, 김정택, 2010)'가 있다. 그 외에 '과학기술분야 디지털 콘텐츠의 아카이빙 정책 연구(곽승진 외, 2011)', '디지털 아카이빙 정책(박현영, 남태우, 2004)', '디지털 큐레이션 정책(안영희, 박옥화, 2010)'이 있다. 그러나, '디지털 아카이빙'이나 '디지털 큐레이션'은 '디지털 보존'과 관련이 있으나 연구 주제의 초점이 달라 그러한 정책이 디지털 보존 정책을 반영한다고 말하기는 어렵다.

디지털 보존 정책은 이경남과 이소연(2007)이 발표한 논문에서 우리나라의 전자기록 보존과 관리를 위한 아젠다 프레임워크에서 네 가지 영역 중 '법제 및 정책 영역'에서 그 필요성이 주장되었다. 두 저자는 이 영역에서 가장 중요한 과제는 전자기록의 장기 보존을 위한 '정책 프레임워크를 개발'하는 일이라고 주장하였다. 또한 기관 유형별 전자기록의 보존 정책 수

립을 주장하였다. 연구 영역 아젠다에서 연구 과제의 우선순위를 보면, '신뢰할 수 있는 디지털 아카이브 인증 기준'이 1위를, '전자기록 유형별 관리·보존 정책 연구'가 3위를 차지했다. 가장 시급하고 기초적인 연구 아젠다로는 '디지털 보존 현황 조사 및 보존 위험 평가 연구'라고 규정되었다.

그런데도, 디지털 보존 현황을 조사한 연구 결과는 2016년 서은경에 의해 처음 발표되었다. 서은경은 전국 4년제 대학도서관 중 디지털 콘텐츠를 구축하고 있는 102개 대학도서관을 대상으로 디지털 보존 실태를 조사했다. 그 결과, 조사 대상 사서들 대부분이 디지털 보존을 성공적으로 수행하기 위해 가장 먼저 갖추어야 할 요건으로 '디지털 보존 정책 수립'을 답했으나, 동시에 실제 디지털 보존 업무의 장애요인으로 '장기적인 정책 부재'를 두 번째로 꼽았다. 이는 많은 대학도서관이 장기적인 디지털 보존 정책을 수립하지 않고 있다는 사실을 반영하는 것이다. 조사 대상 사서들이 디지털 보존의 교육 분야 중 가장 교육받고 싶은 주제로 '보존 정책을 기획하고 개발하는 방법'이라고 답했다. 이러한 결과는 장기적인 보존 정책을 수립하여 시행하고자 하는 조사 대상 사서들의 바람을 반영하는 것이다. 서은경은 결론에서 대학도서관들이 미래지향적인 디지털 보존 정책을 수립해야 한다고 주장하였다.

위에서 언급한 연구 결과에서 알 수 있듯이, 국내에서 디지털 보존에 관한 연구는 소수 존재한다. 무엇보다도 디지털 보존을 수행하는데 기반이 되는 디지털 보존 정책에 관한 연구는 거의 없다는 것을 알 수 있다. 더 나아가 디지털 보존 정책을 개발하는 데 기반으로 이용

할 수 있는 디지털 보존 정책 프레임워크에 대한 국내 연구는 거의 알려진 바가 없다.

2.2 해외 연구

국내에서는 디지털 보존 정책에 관한 연구가 소수 존재하는 반면, 해외에서는 해당 주제에 관한 연구들이 다수 수행되었다. 디지털 보존, 더 나아가 디지털 보존 정책에 관한 연구들이 다양하고 세부적인 주제를 가지고 오래전부터 수행되었다. 먼저, 국가적 차원에서 디지털 보존 정책을 조사한 연구들이 있는데, 그러한 연구의 주제로는 디지털 보존 정책 개발을 위한 연구, 디지털 보존 정책의 수립 여부 조사, 디지털 보존 정책 계획, 디지털 보존 정책 자체에 관한 연구, 한 국가 내에서 디지털 보존 정책에 관한 문헌 조사 연구 등이 있다(예: Bodé, 2013; National Archives of Australia, 2021; Matlala, Ncube, & Parbanath, 2022; Gbaje & Mohammed, 2013).

이외에도, 대학도서관의 디지털 보존 정책 실행 현황을 조사한 연구(예: da Silva Júnior & Borges, 2017), 한 기관의 디지털 보존 정책 수립을 위한 연구(예: Archives, 2009), 디지털 보존 정책을 분석한 연구(예: Sheldon, 2013; LeFurgy, 2013), 디지털 보존 정책 개발을 위한 지침서(예: Brown, 2013; National Archives of UK, 2011) 등이 있다.

특히, 정보기관들이 자관의 디지털 보존 정책을 수립할 수 있도록 돕기 위하여 디지털 보존 정책의 프레임워크를 조사한 연구들(예: McGovern, 2004; Noonan, 2014; Xie, 2008)이 있다. 좀 더 세부적인 주제를 살펴보면, 디지

털 보존 정책 프레임워크의 개발에 관한 연구들(예: Beagrie, 1998), 대학도서관의 디지털 보존 정책 프레임워크를 조사한 연구들(예: McGovern, 2004; Noonan, 2014), 여러 기관의 디지털 보존 정책 프레임워크를 분석한 연구들(예: Dressler, 2017; Sheldon, 2013)이 있다. 또한, 한 국가의 맥락에서 디지털 보존 정책 프레임워크에 관한 연구들(예: Katre, 2012; Library and Archives Canada, 2014; Gbaje & Mohammed, 2013)이 있는데, 그러한 연구들은 대부분 프레임워크의 필요성이나 프레임워크에 관한 설명을 포함하고 있다.

위에서 언급한 연구 중 정보기관들이 자관의 디지털 보존 정책을 수립할 수 있도록 돕기 위해 수행된 연구 중 중요한 연구 몇 가지를 소개하면 다음과 같다.

첫 번째로, International Research on Permanent Authentic Records in Electronic Systems(InterPARES)는 캐나다의 School of Library, Archival and Information Studies at The University of British Columbia가 주축이 되어 여러 국가의 협력으로 수행되고 있는 국제 프로젝트이다. InterPARES의 목표는 디지털 형식으로 생성 및/또는 유지되는 진본 기록(authentic record)의 장기 보존을 위해 필수적인 지식을 개발하는 것이다. 그 중, InterPARES 2 프로젝트는 보존 정책과 직접적으로 관련이 있다. 이 프로젝트는 디지털 레코드의 장기보존을 위한 정책·전략·표준 개발을 위한 원칙 프레임워크(a framework of principles for the development of policies, strategies and standards for the long-term preservation of digital records)를 산출물로

제출했다. 이 프레임워크는 두 개의 상호보완적인 세트—하나는 레코드 생산자들, 또 다른 하나는 레코드 보존자들을 위한 원칙 세트—로 구성되어 있다. 레코드 보존자들을 위한 원칙 세트는 13개의 원칙으로 구성되어 있다(InterPARES 2 Project, 2008).

UK Joint Information Systems Committee (JISC)의 재정지원을 받은 Beagrie 등(2008)은 UK 영국의 고등교육기관들이 자관의 디지털 보존 정책을 개발할 수 있도록 실용적인 가이드를 제공하려는 목적으로 연구를 수행하였다. 이러한 목적을 위해 기관의 디지털 보존 정책을 개발하려고 할 때 고려해야 할 사항들을 제시하고, 여러 기관의 디지털 보존 정책과 관련 자료들을 분석하여 기관의 디지털 보존 정책을 위한 모델을 제시하였다. 이 모델은 두 파트로 구성되어 있다: 1) 모델 디지털 보존 정책: 정책 조항(Model digital preservation policy: Policy clauses); 2) 모델 디지털 보존 정책: 실행 조항(Model digital preservation policy: Implementation clauses).

Electronic Resource Preservation and Access Network(이하 ERPANET)(2003)는 디지털 보존 정책 도구를 제시하기 위하여 디지털 자료를 보존·유지하고 현재와 미래에 이용할 수 있게 하려고 사용 중 또는 계획 중인 정책을 조사하였다. 조사 결과, ERPANET의 도구는 디지털 보존 정책의 구성요소를 크게 혜택(benefits), 범위와 목표(scope and objectives), 필요조건(requirements), 역할과 책임(roles and responsibilities), 맥락(context), 적용 범위(areas of coverage), 비용(costs), 모니터링 및 리뷰(monitors and review), 정책 실행

(implementation of the policy)으로 나누어 각각에 대해 간략하게 서술하고 있다.

위에서 언급한 연구 결과에서 알 수 있듯이, 해외 연구 중에는 디지털 보존 정책 개발과 관련된 다수의 연구와 그 중 디지털 보존 정책 프레임워크와 관련된 소수의 연구가 있다. 그런데도, 근래 들어 디지털 자산으로 인정받는 데이터를 주로 다루는 데이터 리포지토리들이 기관의 디지털 보존 정책을 개발할 때 기초 자료로 사용할 수 있는 프레임워크에 대한 연구는 거의 없다. 이에 본 연구에서는 CTS라는 국제적인 인증을 받은 여러 나라의 다양한 데이터 리포지토리들의 보존 정책을 비교 분석함으로써 그들의 디지털 보존 정책에 기본적으로 포함되어 있는 구성요소를 파악하여, 데이터 리포지토리의 디지털 보존 정책 프레임워크에 필수적으로 포함되어야 할 구성요소를 제안하고, 더 나아가 디지털 보존 정책을 개발하는 데 도움이 되고자 한다.

3. CoreTrustSeal 인증기관의 디지털 보존 정책

디지털 보존 정책 영역의 연구들은 사례 연구 및 비교 분석에서부터 설문 조사 및 체계적 검토에 이르기까지 다양한 연구 방법론을 채택하고 있다. 조직 전반의 데이터 실무에 대한 통계 분석과 같은 정량적 방법도 점차 활용되고 있다. 이러한 방법론은 다양한 목적을 제공하는데, 예를 들어, 사례 연구는 특정 조직 과제에 대한 심층적인 통찰력을 제공하는 경우가 많고, 비교 분석은 동향 및 모범 사례에 대한 보다 광범

위한 개요를 제공한다(Patten, 2017; Newhart & Pattern, 2023). 본 연구에서는 CTS의 인증을 받은 데이터 리포지토리들의 디지털 보존 정책을 조사 분석하여 각 기관의 디지털 보존 정책 구성요소와 각 구성요소의 내용을 비교 분석하였다.

데이터 리포지토리와의 관련된 인증 체계로는 CTS, Rat für Sozialund Wirtschaftsdaten, CLARIN certificate B, DIN 31644, ISO 16363, network of expertise in long-term Storage of digital resources in Germany, Trusted Digital Repositories 등이 있다. 그 중에서도 CTS 인증은 “타 인증의 기준 및 요구사항을 충족하고 있는 핵심 인증 체계”이자, “re3data 등록 리포지토리 중에서 가장 많이 획득한 인증 체계이며 타 인증 체계와의 호환성 그리고 갱신을 통한 관리가 가능하다”(김주섭, 양성준, 김선태, 2022, 209-210).

더 나아가 CTS의 핵심 인증은 인증을 받으려는 기관들에게 디지털 보존과 관련된 혜택을 줄 수 있다. 먼저 디지털 보존의 중요성에 대한 직원들의 인식을 높일 수 있다(L'Hours, Kleemola & de Leeuw, 2019). 또한, 데이터 생산자, 자금 제공자, 이용자에게 데이터가 신뢰할 수 있는 품질로 보존되어 재이용될 수 있다는 믿음을 주고, 그들이 해당 리포지토리를 평가할 수 있는 근거를 준다(Dillo & de Leeuw, 2018). 이러한 이유로 본 연구에서는 CTS의 인증을 받은 데이터 리포지토리들의 디지털 보존 정책을 분석 대상으로 선정했다.

먼저, CTS의 인증 요구사항—특히 보존과 관련된 요구사항—을 살펴보기 위해 최신 버전인 CoreTrustSeal Trustworthy Data Repositories

Requirements 2023-2025을 검토하여 CTS 인증 요구 사항에 대한 기초 정보를 습득하였다. 이러한 요구사항은 신뢰할 수 있는 리포지토리의 특성을 반영하며, 모든 요구사항은 필수이고, 동일한 가중치를 갖는 독립형 항목으로 모두 17개이고 4개의 큰 카테고리 안에 포함되어 있다. 즉, CTS의 요구사항은 크게 배경 정보 및 맥락(R0), 조직 인프라(R01~R06), 디지털 객체 관리(R07~R13), 정보 기술과 보안(R14~R16)으로 4개의 큰 카테고리 안에 총 17개의 항목으로 구성되어 있다(〈표 1〉 참조).

KISTI가 2022년에 CTS에 처음으로 제출했던 2020-2022 버전의 요구사항과 다른 점이 있는지 파악하기 위해 CoreTrustSeal Trustworthy Data Repositories Requirements 2020-2022도 분석하였다(〈표 2〉 참조). 〈표 1〉과 〈표 2〉에서 보여주듯이, 두 버전 사이에 약간의 차이점이 있

었다.

그러나, CTS의 두 버전(2020-2022 버전과 2023-2025 버전)에서 보존 정책과 관련된 요구사항은 ‘보존 계획’ 항목으로 같았고, 요구사항의 번호(R10에서 R09)만 바뀌었을 뿐 요구사항의 항목명과 내용은 같았다. CTS 2023-2025 버전에서 ‘보존 계획(R09)’은 다음과 같이 설명하고 있다: “리포지토리는 장기 보존에 대한 책임을 지고 계획적이고 문서화된 방식으로 이 기능을 관리한다”(CoreTrustSeal, 2023b). ‘보존 계획(R09)’ 항목에서는 이러한 요구사항과 각 기관이 해당 요구사항에 대한 수준을 자가 진단하여 4개의 등급 중 하나를 선택하도록 하고, 해당 요구사항에 대하여 어떻게 답해야 하는지 설명을 제공하고 있다. CTS의 설명을 읽어보면 해당 요구사항에 대한 답변에 포함해야 할 내용 중 많은 부분이 디지털 보존 정책에 포

〈표 1〉 CoreTrustSeal 신뢰할 수 있는 데이터 리포지토리 요구사항(2023-2025) 개요

구분	항목
R0. 배경 정보 & 맥락(Background Information & Context)	
조직 인프라 (Organisational Infrastructure)	R01: 미션 & 범위(Mission & Scope)
	R02: 권리 관리(Rights Management)
	R03: 서비스의 지속성(Continuity of Service)
	R04: 법적 & 윤리적(Legal & Ethical)
	R05: 거버넌스 & 자원(Governance & Resources)
	R06: 전문성 & 지침(Expertise & Guidance)
디지털 객체 관리 (Digital Object Management)	R07: 출처 및 진본성(Provenance and Authenticity)
	R08: 기탁 & 평가(Deposit & Appraisal)
	R09: 보존 계획(Preservation Plan)
	R10: 품질 보장(Quality Assurance)
	R11: 워크플로우(Workflows)
	R12: 발견 및 식별(Discovery and Identification)
	R13: 재이용(Reuse)
정보 기술 & 보안 (Information Technology & Security)	R14: 저장 & 무결성(Storage & Integrity)
	R15: 기술 인프라(Technical Infrastructure)
	R16: 보안(Security)

〈표 2〉 CoreTrustSeal 신뢰할 수 있는 데이터 리포지토리 요구사항(2020-2022) 개요

구분	항목
R0. 맥락(Context)	
조직 인프라 (Organizational Infrastructure)	R1: 미션/범위(Mission/Scope)
	R2: 라이선스(Licenses)
	R3: 접근의 지속성(Continuity of Access)
	R4: 기밀/윤리(Confidentiality/Ethics)
	R5: 조직 인프라(Organizational Infrastructure)
	R6: 전문가 지침(Expert Guidance)
디지털 객체 관리 (Digital Object Management)	R7: 데이터의 무결성과 진본성(Data Integrity and Authenticity)
	R8: 평가(Appraisal)
	R9: 문서화된 저장 절차(Documented Storage Procedures)
	R10: 보존 계획(Preservation plan)
	R11: 데이터 품질(Data Quality)
	R12: 워크플로우(Workflows)
	R13: 데이터 발견 및 식별(Data Discovery and Identification)
	R14: 데이터 재이용(Data Reuse)
기술 (Technology)	R15: 기술 인프라(Technical Infrastructure)
	R16: 보안(Security)

함되어서 문서로 존재하며 실행되어야 한다는 것을 간접적으로 표현하고 있다.

CTS의 인증을 받은 데이터 리포지토리들의 디지털 보존 정책 구성요소를 비교 분석하기 위해 CTS의 인증을 받은 데이터 리포지토리들의 명칭과 그들의 인증 받은 신청서 PDF 파일을 CTS의 웹사이트(<https://amt.coretrustseal.org/certificates>)에서 수집하여 분석하였다.

조사 결과, 현재(2023년 9월 3일)까지 인증 받은 데이터 리포지토리는 총 91개이다. 그중에서 본 연구는 2023년에 CTS 인증을 받은 데이터 리포지토리 총 33개를 비교 분석 대상으로 선정하였다. 이러한 리포지토리들의 인증 신청서 파일을 하나씩 순차적으로 검토하여 각 리포지토리가 디지털 보존 정책을 가지고 있는지 조사하면서 그 결과를 엑셀 파일로 정리하였다. 즉, 각 리포지토리별로 디지털 보존 정책

의 유무와 그러한 정책이 있는 웹사이트 링크를 엑셀 파일로 정리하였다. 디지털 보존 정책의 유무, 정책이 있더라도 구성요소 항목별로 구분하여 정책 내용을 서술했는지, 아니면 항목별 구분 없이 정책 내용을 서술하고 있는지를 간단한 표로 정리한 것이 〈표 3〉이다. 〈표 3〉에서 디지털 보존 정책의 유무 확인이 불가능한 경우는 인증 신청서에는 정책이 있다고 대답하고 링크를 제공했으나 연결된 웹페이지에서는 에러 메시지만 떴고, 해당 리포지토리와의 상위 기관의 웹사이트까지 조사했으나 디지털 보존 정책을 찾지 못한 경우이다.

디지털 보존 정책을 가지고 있는 데이터 리포지토리의 경우에는 그 정책의 내용을 분석하고 정책의 구성요소를 추출하여 구성 요소별 명칭(예를 들어, '목적', '역할과 책임')을 엑셀 파일에서 정리하였다. 대부분의 디지털 보존

〈표 3〉 디지털 보존 정책 보유 현황

구분	리포지토리 개수	
디지털 보존 정책 있음	18	
	구성 요소별 명칭 있음	14
	구성 요소별 명칭 없음	4
디지털 보존 정책 유무 확인 불가함	1	
디지털 보존 정책 없음	14	
총합계	33	

정책은 구성요소의 명칭을 쓰고, 그 구성요소에 대한 내용을 서술하고 있었다. 이렇게 구성요소의 명칭이 있는 경우에는 각각의 명칭을 추출하고, 구성요소의 명칭과 내용이 비슷한 것들을 모아 하나의 카테고리로 만들고, 이런 방법을 반복하여 다수의 카테고리를 만들었다.

분석한 디지털 보존 정책 중에는 구성요소별로 구별하여 각각의 정책 내용을 서술하지 않고 내용만 서술한 경우도 있었다. 이러한 경우에는 그 내용을 분석하여, 미리 만들어 놓은 카테고리 중 가장 적합한 카테고리에 포함했다. 분류해 놓은 카테고리에 포함되지 않는 디지털 보존 정책 내용은 새로운 카테고리를 만들어 포함하고, 정책 내용에 따라 카테고리 명칭을 부여했다.

각 카테고리에 포함된 디지털 보존 정책의 내용은 비슷하나 구성요소의 명칭이 다른 경우에는 가장 보편적으로 쓰이는 명칭을 사용하였다. 이처럼, 각각의 카테고리에 포함되어 있는 디지털 보존 정책의 구성요소 명칭과 내용을 고려하여 각각의 카테고리에 부합하는 명칭을 부여했다.

본 연구에서는 18개 데이터 리포지토리들의 디지털 보존 정책을 분석하고, 정책 내용별로 카테고리를 만들고, 각 카테고리에 디지털 보

존 정책의 구성요소 명칭을 부여하고, 그에 해당하는 디지털 보존 정책의 내용을 가지고 있는 데이터 리포지토리들의 명칭을 병기하여 엑셀파일로 만들었다. 그러한 결과를 가지고, 본 연구에서는 18개 데이터 리포지토리들이 포함하고 있는 디지털 보존 정책의 구성요소를 파악하였다. 또한, 디지털 보존 정책 프레임워크에서 기본적으로 포함해야 하는 필수 구성요소도 파악하여, 그 결과를 '4. 디지털 보존 정책 프레임워크의 구성요소'에서 제안한다.

제안하는 구성요소들은 데이터 리포지토리의 디지털 보존 정책 프레임워크에서 기본적으로 포함해야 하는 필수 구성요소이지만, 각 데이터 리포지토리마다 상황이 다르므로 자관의 상황에 맞도록 구성요소를 변경할 필요가 있을 수 있다는 점이 본 연구의 제한점이라고 할 수 있다.

4. 디지털 보존 정책 프레임워크의 구성요소

디지털 보존 정책을 가지고 있는 CTS 인증 데이터 리포지토리들은 지리적으로 다양한 나라와 장소에 위치하고, 그들의 명칭은 '리포지

토리' 외에도 '아카이브', '도서관', '데이터 센터', '데이터 플랫폼' 등 다양하다. 예를 들어, CTS 인증을 받은 데이터 리포지토리로는 SAEON Open Data Platform, NIDDK Central Repository, Environmental Data Initiative, QSAR DataBank, Australian Ocean Data Network 등이 있다. 그리고 이러한 데이터 리포지토리를 보유하고 있는 상위 기관의 유형 또한 다양하다. 예를 들면, Cornell University, NIAID/NIH, TU Delft Library, Centerdata, Data Archiving and Networked Services가 있다. CTS 인증 데이터 리포지토리들 중에 디지털 보존 정책을 가지고 있지 않은 리포지토리들은 대부분 다른 정책이나 규정에서 디지털 보존과 관련된 내용

을 다루고 있거나 '디지털 보존 계획'을 가지고 있다. 이 중에서 소수의 리포지토리들은 자관의 디지털 보존 정책이나 계획을 작성 중이거나 작성할 예정이라고 밝히고 있다.

아래에서 제시하는 디지털 보존 정책 프레임워크를 개발할 때 포함해야 할 필수 구성요소와 그 구성요소에서 다루어야 할 내용은 2023년도에 CTS에서 인증을 받은 데이터 리포지토리들의 디지털 보존 정책을 분석 비교한 결과이다. 필수 구성요소는 분석 비교한 18개 데이터 리포지토리 중 10개 이상의 리포지토리가 공통적으로 포함하는 요소이고, 각 구성요소에 대한 설명은 분석한 디지털 보존 정책이 해당 요소에서 서술하고 있는 내용이다(〈그림 1〉 참조).

- 정책의 목적과 범위(Purpose and Scope of the Policy)
- 목표, 사명, 임무(Objective, Mission, and Mandate)
- 역할과 책임(Roles and Responsibilities)
- 작업 흐름(Workflow)
- 보존 대상의 범위(Content Scope)
- 협력(Cooperation)
- 필요조건(Requirements)
 - 법적 및 윤리적 측면(Legal and Ethical Aspects)
 - 재정적 측면(Financial Aspects)
 - 기술적 측면(Technical Aspects)
- 표준 준수(Standards Compliance)
- 절차적 적합성(Procedural Suitability)
- 평가, 선정, 획득(Appraisal, Selection, and Acquisition)
- 접근과 이용(Access and Use)
- 데이터 품질 관리(Data Quality Management)
- 저장(Storage)
- 위기관리(Risk Management)
- 보안(Security)
- 정책 실행(Policy Implementation)
- 정책 검토(Policy Review)
- 용어집(Glossary)
- 다른 관련 정책과 문서(Other Related Policies and Documents)
- 참고 자료(References)

〈그림 1〉 디지털 보존 정책 프레임워크의 필수 구성 요소

본 연구의 목표가 다른 데이터 리포지토리들이 본 연구 결과를 바탕으로 자관의 디지털 보존 정책을 작성할 수 있게 돕기 위한 것이므로, 각 구성요소에서 포함하고 있는 내용을 일반화하여 아래와 같이 간략하게 정리하였다. 이는 데이터 리포지토리들이 아래의 내용을 참고하여 그들의 현황에 맞도록 해당 구성요소에서의 내용을 자유롭게 작성하게 하기 위함이다.

- 정책의 목적과 범위(Purpose and Scope of the Policy): 해당 디지털 보존 정책의 목적이 무엇인지 간단명료하게 서술하고, 그 정책이 적용되는 데이터 포맷, 메타데이터의 유형, 그 정책이 사용될 맥락을 정의하고 있다. 예를 들면, 소장 데이터를 장기적으로 보존하고 지속적인 접근을 제공한다, 보존된 데이터의 진본성과 무결성을 보장한다, 필요에 따라 데이터의 포맷을 바꾼다 등의 내용을 포함하고 있다.

- 목표, 사명, 임무(Objective, Mission, and Mandate): 해당 리포지토리의 목표, 사명, 임무를 서술하고 있다.

- 역할과 책임(Roles and Responsibilities): 디지털 보존과 관련된 업무를 파악하여 그에 따른 역할과 책임을 정의하고 있다. 디지털 보존 업무를 수행하는 직원들의 이름은 정책에서 밝히지 않고, 대신 리포지토리의 전체적인 조직과 사업 구조를 개략적으로 설명하면서 디지털 보존에 대한 역할과 책임을 기술하고 있다.

- 작업 흐름(Workflow): 디지털 보존 업무가 처음부터 끝날 때까지 거치는 일련의 단계를 순서대로 상세하게 서술하거나 워크플로우 차트로 그려 두었다. 정책에 따라 디지털 보존을 위한 워크플로우의 복잡한 정도가 다르며,

효율적으로 디지털 보존 실무를 수행하기 위해서는 잘 정의된 일련의 단계가 정책에 포함되어야 한다는 것을 보여준다.

- 보존 대상의 범위(Content Scope): 해당 리포지토리의 장서 개발 정책과 연결하여 보존할 데이터에 대한 기준과 범위를 명확하게 설정하고 있다. 더 나아가 그러한 데이터 중 우선적으로 보존해야 하는 순위를 설정해 놓은 정책도 있다.

- 협력(Cooperation): 디지털 보존이 공동체의 공동 책임임을 인정하고, 협력 대상과 방법, 협력하기 위해 취해야 할 단계, 협력에 대한 요구사항이나 기준을 명시하고 있다. 또한 다른 리포지토리 또는 조직과의 잠재적인 협업을 고려하여 잠재적인 협력 방안을 제시하고, 협력 조건과 상호 이익을 정의하고 있다. 분석한 정책 중 소수는 해당 리포지토리들이 디지털 보존과 관련된 자관의 이해관계자(데이터 기탁자, 이용자, 다른 리포지토리 등) 리스트를 작성하고, 각 이해관계자의 역할과 책임을 규정하고 있다는 것을 간접적으로 나타내고 있다.

- 필요조건(Requirements)

- 법적 및 윤리적 측면(Legal and Ethical Aspects): 해당 리포지토리-특히 정부기관이나 공공기관의 데이터 리포지토리-의 디지털 보존 정책이 작성되는 정치적 맥락과 법적 환경에 관해 기술하고 있다. 디지털 보존과 관련된 데이터의 지적 재산권, 데이터 프라이버시와 보호, 데이터 소유권, 라이선스 등 관련 법률을 준수하고 있음을 보여준다. 더불어 민감한 데이터 또는 개인 데이터를 관리하기 위한 절차를 설명하고, 데이터를 보존하는 데 있어 윤

리적 표준을 준수하고 있음을 기술하고 있다. 몇몇 리포지토리의 정책에서는 법적 측면과 윤리적 측면을 각각의 요구사항으로 나누어 기술하고 있다.

- 재정적 측면(Financial Aspects): 기술 교육, 표준, 컨설턴트 이용, 협력 또는 공유 스토리지/접근/보존 시설 등에 들어가는 비용을 어떻게 안정적으로 유지할 것인지에 관해 서술하고, 재정 계획에 관한 내용도 포함하고 있다. 장기 자금 조달 전략과 출처를 설명하고, 자금의 지속가능성을 위해 잠재적인 파트너십이나 협력에 관해 서술하고 있다.

- 기술적 측면(Technical Aspects): 장기 데이터 보존을 위해 리포지토리가 준수하는 기술 사양 및 표준과 하드웨어, 소프트웨어, 네트워크 요구사항 등을 상세히 설명하고 있다. 기술 예측의 정의, 유지 관리 절차의 정의, 보존 전략의 정의, 데이터 재포맷, 새로 고침, 마이그레이션, 에뮬레이션 등을 위한 정확한 지침과 관련된 내용도 기술하고 있다.

- 표준 준수(Standards Compliance): 표준은 다른 조직들과의 협력(시스템 호환 등)을 촉진하고 다른 이니셔티브의 지식과 경험을 공유할 수 있으므로, 디지털 보존 정책에서 해당 리포지토리에서 사용하고 있는 표준을 나열하고 있다. 예를 들면, 보존 메타데이터를 위한 표준이나 데이터에 장기적으로 접근가능하게 만드는 표준 데이터 포맷에 관해서 서술하고 있다.

- 절차적 적합성(Procedural Suitability): 리포지토리의 보존 목적과 표준에 일치하는 보존 활동과 그러한 활동의 절차를 서술하고 있

다. 여기에서 언급되는 보존 활동으로는 데이터 수집, 입수 준비, 변환, 폐기 등 다양한 활동이 있다. 해당 리포지토리는 디지털 보존 절차에 대한 상세하고 구조화된 로드맵을 제공하여 그 리포지토리가 디지털 보존 업무를 명확하고 일관성 있게 수행하고 있음을 보여준다. 즉, 리포지토리들은 자관의 보존 활동이 적합한 절차에 따라 수행되고 있음을 알림으로써 자관의 보존 활동에 대한 투명성을 향상시켜 데이터 기탁자, 이용자, 기타 이해관계자들과의 신뢰를 구축하려는 노력을 보여주고 있다.

- 평가, 선정, 획득(Appraisal, Selection, and Acquisition): 해당 리포지토리의 장서 개발 정책과 연결하여 보존할 데이터를 획득하기 위한 데이터 평가, 선정, 획득과 관련된 기준, 방법, 과정에 관해 기술하고 있다. 또한 해당 리포지토리에 데이터를 기탁할 수 있는 사람의 자격, 기탁가능한 데이터의 종류와 선호하는 파일 포맷, 기탁가능한 데이터의 품질 수준, 기탁 절차, 기탁자가 제공해야 할 메타데이터, 데이터의 엠바고 여부, 엠바고 기간 등에 관한 내용을 포함하고 있다. 데이터의 기탁을 도와주는 해당 리포지토리의 지원 메커니즘도 설명하고 있다.

- 접근과 이용(Access and Use): 데이터의 접근 방법, 접근 제한의 수준, 접근 통제 권한 등 데이터 접근에 관한 내용을 서술하고 있다. 또한 데이터의 현재 이용뿐만 아니라 미래의 이용과 잠재적 이용자에 관해서도 예측하는 바를 서술하고 있다. 소장 데이터의 접근과 이용을 도와줄 수 있는 해당 리포지토리의 지원 메커니즘도 설명하고 있다.

- 데이터 품질 관리(Data Quality Management): 소장 데이터의 진본성, 무결성, 신뢰성 등 데이

터 품질을 보장하기 위해 사용하는 도구, 절차, 방법을 밝히고 있다.

- 저장(Storage): 리포지토리의 디지털 저장 매체를 어떤 방식과 절차를 통해 지속적으로 유지하고 보수하는 지, 그리고 백업 시스템과 방법 등에 대해서도 설명하고 있다.

- 위기관리(Risk Management): 디지털 보존과 관련한 위기(자연재해, 기술 변화 등)에 미리 대처하고, 데이터 손실이나 저장 매체의 노후화 등 잠재적인 위험을 파악하고, 파악된 위험에 대한 대처 전략 및 계획을 준비하고 있음을 보여 준다.

- 보안(Security): 권한이 없는 사람이 리포지토리의 소장 데이터에 접근, 이용, 변경, 공개, 삭제하는 것을 어떻게 방지하는지 설명하고 있다. 분석한 정책에 따라 시스템 보안, 물리적 보안, 네트워크 보안, 파일 보안 등 다루고 있는 내용이 약간씩 다르다.

- 정책 실행(Policy Implementation): 디지털 보존 정책이 존재한다고 해서 그 정책이 실행된다는 것을 의미하는 것은 아니므로, 그 정책이 실제 리포지토리에서 실행될 수 있는 방안을 명시하고 있다. 예를 들어, 정책 실행을 위해서 리포지토리의 안정적인 재정과 추가적인 활동이 필요하다는 점을 서술하고 있다.

- 정책 검토(Policy Review): 작성된 디지털 보존 정책을 리포지토리의 내외부 변화(기술 변화, 새로운 표준 등)에 반응하면서 정기적으로 검토하며, 검토 시기가 아니더라도 필요에 따라 수시로 검토하고, 또한 필요할 때마다 해당 정책을 수정하고 업데이트해야 한다고 밝히고 있다.

- 용어집(Glossary): 디지털 보존 정책에서

사용된 주요 용어에 관해 해당 리포지토리의 직원들이 일관된 개념을 가지고 일할 수 있도록 주요 용어에 대한 정의를 포함하고 있다.

- 다른 관련 정책과 문서(Other Related Policies and Documents): 해당 보존 정책 내에서 언급되거나 관련이 있는 기타 문서와 정책 등의 목록과 링크를 제시하고 있다.

- 참고 자료(References): 정책 작성 시 이용한 참고 자료 목록과 링크를 제시하고 있다.

연구 결과, CTS 인증을 받은 데이터 리포지토리들은 그들의 신청서에서 ‘보존 전략’에 대한 답변에 그들의 디지털 보존 정책과 관련된 내용을 직접 답변하고, 추가로 그들의 답변에 대한 근거로서 다양한 관련 링크를 제공하고 있다. 그러한 링크들이 보여주는 문서로는 디지털 보존 정책, 보존 메타데이터 표준, 위기관리 계획서, 선호하는 파일 포맷, 백업 절차와 방법에 관한 문서 등이 있다.

5. 결 론

생산되는 데이터양의 증가와 함께 데이터의 장기 보존을 위한 데이터 리포지토리의 수도 증가하고 있다. 그럼에도 불구하고, 많은 데이터 리포지토리들이 자관의 디지털 보존 정책을 가지고 있지 않다. 그중의 하나가 KISTI 연구 데이터공유센터의 DataON으로, CTS 인증 신청 준비를 하는 과정에서 디지털 보존 정책의 필요성을 인식하고 정책 개발을 할 계획을 세우고 있다. 이를 위한 준비 단계로서 본 연구는 디지털 보존 정책 프레임워크를 개발하려는 목

적으로 시작되었다. 왜냐하면 데이터 리포지토리를 위한 디지털 보존 정책 프레임워크를 작성하는 것은 데이터의 장기적인 지속가능성, 접근가능성, 이용가능성을 보장하는 데 필수적인 단계이기 때문이다.

본 연구는 2023년도에 CTS에서 인증을 받은 데이터 리포지토리들의 디지털 보존 정책을 비교 분석하여 디지털 보존 정책 프레임워크에 필요한 구성요소를 추출하고, 각각의 구성요소에서 어떠한 내용을 포함하고 있는지를 '4. 디지털 보존 정책 프레임워크의 구성요소'에서 연구 결과로 제시하였다. 본 연구에서 제시한 구성요소는 다음과 같다: 정책의 목적과 범위; 목표, 사명, 임무; 역할과 책임; 작업 흐름; 보존 대상의 범위; 협력; 필요조건 (법적 및 윤리적 측면, 재정적 측면, 기술적 측면); 표준 준수; 절차적 적합성; 평가, 선정, 획득; 접근과 이용; 데이터 품질 관리; 저장; 위기관리; 보안; 정책 실행; 정책 검토; 용어집; 다른 관련 정책과 문서; 참고 자료.

본 연구의 결과로 제시한 구성요소와 내용은 데이터 리포지토리들이 자관의 정책을 개발하는 데 어떤 구성요소를 포함하고 그 구성요소에 어떤 내용을 포함해야 하는지를 보여준다. 또한 이러한 연구 결과는 정책의 직접적인 내용뿐만 아니라 해당 정책의 데이터 리포지토리들이 현재 어떻게 디지털 보존 실무를 수행하고 있는지, 그러한 보존 실무를 향상하기 위해 어떤 노력을 하고 있는지를 직간접적으로 보여주고 있다.

그러므로 본 연구의 결과는 KISTI 연구데이터공유센터처럼 CTS 인증을 신청하려는 데이터 리포지토리들이 자관의 디지털 보존 정책

프레임워크를 개발하고, 그것을 바탕으로 정책을 만들고, 더 나아가 '보존 계획'을 세우는 데 도움이 될 것이다. 또한 CTS를 신청할 계획이 없는 데이터 리포지토리들도 본 연구의 결과를 이용하여 자관의 디지털 보존 정책을 세울 수 있다. 왜냐하면 본 연구의 비교 분석 대상인 데이터 리포지토리들의 지리적 위치, 유형 등이 다양하기 때문에 본 연구에서 제안하는 프레임워크는 다양한 리포지토리들의 정책을 만드는 데 보편적으로 사용될 수 있다. 그러나 동시에 각 데이터 리포지토리는 다른 상황에 있으므로, 리포지토리에 따라서는 자관의 상황에 맞도록 본 연구에서 제안하는 프레임워크의 구성요소를 가감하거나 변경할 수 있고, 각 요소에 서술하는 내용이나 내용의 상세한 정도가 다를 수 있다.

데이터 리포지토리들은 본 연구 결과에서 제시한 디지털 보존 정책 프레임워크의 구성요소를 자세하게 서술함으로써 소장하고 있는 데이터의 장기 보존, 접근가능성, 이용가능성을 보장하는 포괄적인 디지털 보존 정책을 수립할 수 있을 것이다. 그리고 그러한 디지털 보존 정책은 리포지토리의 운영을 모범 사례 및 이해관계자의 기대에 맞게 조정하는 로드맵을 제공하여 데이터 보존의 일관성, 신뢰성, 장기적인 효율성을 보장할 수 있을 것이다.

본 연구는 데이터 리포지토리를 위한 디지털 보존 정책 프레임워크의 구성요소에 대한 것으로, 향후 연구에서는 디지털 보존 정책 프레임워크 또는 더 나아가 그러한 정책 자체에 대한 보다 심층적인 연구를 수행하기를 기대한다. 예를 들어, 향후 연구에서는 본 연구에서 제시한 프레임워크를 기반으로 작성된 디지털 보존

정책을 성공적으로 실행한 사례 연구를 할 수 있을 것이다. 이러한 사례 연구는 해당 데이터 리포지토리의 디지털 보존과 관련된 도전, 전략 등에 대한 실제 지침 및 실제 교훈을 다른

기관들에게 제공할 수 있을 것이다. 또한 디지털 보존 정책의 효과와 영향력을 평가하기 위한 매트릭스를 개발하는 연구도 향후 기대되는 연구 중 하나이다.

참 고 문 헌

- 곽승진, 최재황, 배경재, 정영미 (2011). 과학기술분야 디지털 콘텐츠의 아카이빙 정책 연구. *정보관리연구*, 42(4), 115-136. <http://dx.doi.org/10.1633/JIM.2011.42.4.115>
- 김주섭, 양성준, 김선태 (2022). 데이터 리포지터리 인증 체계 분석 및 인증 전략에 관한 연구: Coretrustseal을 중심으로. *한국문헌정보학회지*, 56(2), 209-229. <https://doi.org/10.4275/KSLIS.2022.56.2.209>
- 나연희, 박희진 (2022). 국내 기록보존 연구동향 분석: 2000~2021년 학술논문을 중심으로. *한국기록관리학회지*, 22(4), 175-196. <https://doi.org/10.14404/JKSARM.2022.22.4.175>
- 박현영, 남태우 (2004). 디지털 아카이빙 정책에 관한 연구. *한국정보관리학회 학술대회 논문집*, 11, 69-76.
- 서은경 (2016). 대학도서관에서의 디지털보존 실태분석 연구. *정보관리학회지*, 33(1), 269-292. <http://dx.doi.org/10.3743/KOSIM.2016.33.1.269>
- 안영희, 박옥화 (2010). 디지털 큐레이션 정책을 위한 프레임워크 개발. *한국도서관·정보학회지*, 41(1), 167-186. <https://doi.org/10.16981/kliss.41.1.201003.167>
- 이경남, 이소연 (2008). 전자기록의 장기적 보존과 관리를 위한 아젠다 개발. *한국기록관리학회지*, 8(1), 211-234. <https://doi.org/10.14404/JKSARM.2008.8.1.211>
- 이소연 (2013). 국내 디지털 보존 연구의 동향 분석. *한국기록관리학회지*, 13(2), 247-283. <https://doi.org/10.14404/JKSARM.2013.13.2.247>
- 정영미, 윤화목, 김정택 (2010). 디지털 콘텐츠의 무결성 유지를 위한 장기적인 보존 정책에 관한 연구. *정보관리연구*, 41(4), 205-226.
- Ahmad, R. & Rafiq, M. (2022). Global perspective on digital preservation policy: a systematic review. *Journal of Librarianship and Information Science*, 0(0). <https://doi.org/10.1177/09610006221111572>
- Beagrie, N. (1998). Developing a policy framework for digital preservation. *Proceeding of the Sixth DELOS Workshop on Preservation of Digital Information*, Tomar, Portugal.

- Beagrie, N., Williams, P., & Rettberg, N. (2008). Digital reservation policy: a subject of no importance?, *iPRES*, 1-4.
- Bodé, E. C. (2013). Digital preservation policy of the chamber of deputies: methodology for its development. *Proceedings of The Memory of the World in the Digital Age: Digitization and Preservation*, 26-28.
- Brown, A. (2013). *Practical Digital Preservation: A How-to Guide for Organization of Any Size*. London: Facet Publishing.
- CoreTrustSeal (2023a). Why certification. Available: <https://www.coretrustseal.org/why-certification/>
- CoreTrustSeal (2023b). CoreTrustSeal Trustworthy Data Repositories Requirements 2023-2025. Available: <https://zenodo.org/record/7051012>
- da Silva Júnior, L. P. & Borges, M. M. (2017). Digital preservation policies of the institutional repositories at Brazilian Federal Universities. *The Electronic Library*, 35(2), 311-321. <https://doi.org/10.1108/EL-09-2015-0170>
- Dressler, V. A. (2017). The state of affairs with digital preservation at ARL member libraries: a survey and analysis of policy. *Digital Library Perspectives*, 33(2), 137-155. <https://doi.org/10.1108/DLP-08-2016-0030>
- Electronic Resource Preservation and Access Network (ERPANET) (2003). *Erpa Guidance: Digital Preservation Policy Tool*. Available: <https://www.erpanet.org/guidance/docs/ERPANETPolicyTool.pdf>
- Gbaje, S. E. & Mohammed, Z. (2013). Digital preservation policy in National Information Centres in Nigeria. *The Electronic Library*, 31(4), 483-492. <https://doi.org/10.1108/EL-01-2012-0011>
- Katre, D. (2012). Need of legislation and digital preservation policy framework in Indian context. *DESIDOC Journal of Library & Information Technology*, 32(4), 321-326. <https://doi.org/10.14429/djlit.32.4.2526>
- L'Hours, H., Kleemola, M., & de Leeuw, L. (2019). CoreTrustSeal: From academic collaboration to sustainable services. *IASSIST Quarterly*, 43(1), 1-17. <https://doi.org/10.29173/iq936>
- LeFurgy, B. (2013). Analysis of Current Digital Preservation Policies: Archives, Libraries and Museums [Blog]. Available: <https://blogs.loc.gov/thesignal/2013/08/analysis-of-current-digital-preservation-policies-archives-libraries-and-museums/>
- Library and Archives Canada (2014) *Stewardship Policy Framework*. Canada: Library and Archives

- Canada.
- Matlala, M. E., Ncube, T. R., & Parbanath, S. (2022). The state of digital records preservation in South Africa's public sector in the 21st century: a literature review. *Records Management Journal*, 32(2), 198-212. <https://doi.org/10.1108/RMJ-02-2021-0004>
- McGovern, N. Y. (2004). *Cornell University Library Digital Preservation Policy Framework*. NY: Cornell University Library. Available: <https://hdl.handle.net/1813/11230>
- National Archives of Australia (2021). *Preservation Policy 2021-2025*. Available: <https://www.naa.gov.au/about-us/our-organisation/accountability-and-reporting/archival-policy-and-planning/preservation-policy-2021-2025>
- Newhart, M. & Patten, M. L. (2023). *Understanding Research Methods* (11th ed.). UK: Routledge.
- Noonan, D. (2014). *Digital preservation policy framework: a case study*. *EDUCAUSE Review*. Available: <https://er.educause.edu/articles/2014/7/digital-preservation-policy-framework-a-case-study>
- Oliver, G. & Harvey, R. (2016). *Digital Curation* (2nd ed.). Chicago: Neal-Schuman.
- Patten, M. L. (2017). *Understanding Research Methods: An Overview of the Essentials* (10th ed.). UK: Routledge.
- Research Libraries Group (2002). *Trusted Digital Repositories: Attributes and Responsibilities*. CA: RLG.
- Sheldon, M. (2013). *Analysis of Current Digital Preservation Policies: Archives, Libraries, and Museums*. Available: <https://blogs.loc.gov/thesignal/2013/08/analysis-of-current-digital-preservation-policies-archives-libraries-and-museums/>
- Xie, S. L. (2008). A foundation for developing digital preservation policy: the InterPARES policy framework. *New Technology of Library and Information Service*, 24(1), 1-12.

• 국문 참고자료의 영어 표기

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

- Ahn, Young-Hee & Park, Ok-Wha (2010). Development of a framework for digital curation policy. *Journal of Korean Library and Information Science Society*, 41(1), 167-186. <https://doi.org/10.16981/kliss.41.1.201003.167>
- Jung, Young-Mi, Yoon, Hwa-Mook, & Kim, Jeong-Taek (2010). A study on the preservation policy for maintaining the integrity of digital contents. *Journal of Information Management*,

41(4), 205-226.

- Kim, Juseop, Yang, Seongjun, & Kim, Suntae (2022). A study on data repository certificate scheme analysis and certification strategy: focused on CoreTrustSeal. *Journal of the Korean Society for Library and Information Science*, 56(2), 209-229.
<http://dx.doi.org/10.4275/KSLIS.2022.56.2.209>
- Kwak, Seung-Jin, Choi, Jae-Hwang, Bae, Kyung-Jae, & Jung, Young-Mi (2011). A study on the policy on digital contents archiving in the field of science and technology. *Journal of Korean Information Management*, 42(4), 115-136.
<http://dx.doi.org/10.1633/JIM.2011.42.4.115>
- Lee, Kyung-Nam & Lee, So-Yeon (2008). Developing agenda for electronic records long-term preservation and management. *Journal of Korean Society of Archives and Records Management*, 8(1), 211-234. <https://doi.org/10.14404/JKSARM.2008.8.1.211>
- Lee, So-Yeon (2013). Trends analysis of digital preservation research in Korea. *Journal of Korean Society of Archives and Records Management*, 13(2), 247-283.
<https://doi.org/10.14404/JKSARM.2013.13.2.247>
- Na, Yonwhee & Park, Heejin (2022). A study on the research trends of archival preservation papers in Korea from 2000 to 2021. *Journal of Korean Society of Archives and Records Management*, 22(4), 175-196. <https://doi.org/10.14404/JKSARM.2022.22.4.175>
- Park, Hyun-Young & Nam, Tae-Woo (2004). A study on digital archiving policy. *Conference of the Korean Society for Information Management*, 11, 69-76.
- Seo, Eun-Gyoung (2016). An analysis on digital preservation practices in academic libraries. *Journal of Information Management*, 33(1), 269-292. <http://dx.doi.org/10.3743/KOSIM.2016.33.1.269>