

연구데이터의 장기보존을 위한 기록관리 메타데이터에 관한 연구*

Research on Record Management Metadata for Long-Term Preservation of Research Data

금효진 (Hyojin Geum)**

양동민 (Dongmin Yang)***

초 록

연구개발 패러다임이 데이터 기반의 과학으로 변화하며 연구데이터의 공유 및 관리 절차의 중요성이 주목받고 있다. 특히 공적자금을 투입한 연구데이터의 경우 행정기관의 직무상 취득된 자료이기 때문에 행정정보 데이터세트로서 기록관리의 중요한 대상이다. 연구데이터는 그 특성상 기존의 기록생애주기 개념을 적용하기 어려우며, 현행 행정정보 데이터세트 기록관리방안(관리기준표)은 이러한 특성을 충분히 반영하지 못하여 연구데이터 관리와 보존에 어려움이 있다. 이에 본 연구는 연구데이터 특성을 반영하고, 향후 연구데이터의 활용까지 고려한 기록관리 메타데이터 요소를 제안하고자 하였다. 이를 위해 NAK8, NAK35, ISAD(G), PREMIS, DCAT, DataCite 등의 표준을 활용하였다. 본 연구를 통해 연구데이터 기록관리를 보다 효과적으로 운용할 수 있는 기반을 마련할 수 있을 것을 기대한다.

ABSTRACT

As the R&D paradigm shifts to data-driven science, the importance of research data sharing and management procedures is gaining attention. In particular, research data funded by public funds is an important subject of record management as an administrative information dataset because it is acquired in the course of the duties of administrative organizations. Due to the nature of research data, it is difficult to apply the existing concept of record lifecycle, and the current administrative information dataset record management plan does not fully reflect these characteristics, which makes it difficult to manage and preserve research data. Therefore, this study aims to propose record management metadata elements that reflect the characteristics of research data and consider the future utilization of research data. For this purpose, NAK8, NAK35, ISAD(G), and PREMIS data models were utilized. This study is expected to lay the foundation for more effective management of research data records.

키워드: 연구데이터, 연구성과물, 장기보존, 메타데이터, 메타데이터 설계
research data, research outcome, long-term preservation, metadata, metadata design

* 이 논문은 제31회 한국정보관리학회 하계학술대회에서 발표한 「연구데이터의 장기보존을 위한 기록관리 메타데이터 요소에 관한 연구」의 내용을 발전시킨 연구임.

이 논문은 2023년 대한민국 교육부와 한국연구재단의 인문사회분야 중견연구지원사업의 지원을 받아 수행된 연구임(NRF-2023S1A5A2A01077759).

** 전북대학교 기록관리학과 석사과정(0133sky@jbnu.ac.kr) (제1저자)

*** 전북대학교 기록관리학과 교수 / 문화융복합아카이빙연구소 공동연구원(dmyang@jbnu.ac.kr) (교신저자)

■ 논문접수일자: 2024년 11월 17일 ■ 최초심사일자: 2024년 12월 11일 ■ 게재확정일자: 2024년 12월 15일

■ 정보관리학회지, 41(4), 179-207, 2024. <http://dx.doi.org/10.3743/KOSIM.2024.41.4.179>

** Copyright © 2024 Korean Society for Information Management

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 연구의 배경과 목적

한 국가의 지식정보 능력, 즉 효율적인 연구 개발 생산성은 그 나라의 국가경쟁력에 큰 영향을 미친다. 과학 기술이 발전하며 방대한 양의 연구데이터가 축적되고 있고 이를 분석하고 활용한 연구 성과가 증가하고 있다. 이는 연구데이터 활용 역량이 연구개발의 주요 동력으로서 주목받고 있음을 의미한다. 이에 따라 연구데이터의 활용을 극대화하기 위한 체계적인 공유 및 관리 절차, 데이터 플랫폼의 중요성이 부상하고 있으며 이와 관련한 제도와 시스템에 대해서도 요구의 목소리가 높아지고 있다.

이러한 사회적 배경에 따라 미국, 영국, EU 등의 주요 해외 선진국들은 공적자금이 투입된 연구과제의 연구 과정에서 생성된 연구데이터를 다른 연구자들이 공유 및 활용할 수 있게끔 책무를 부과하고 있으며, 연구의 성과뿐만 아니라 연구의 중간 과정까지 개방하는 오픈 사이언스 정책을 적극적으로 추진하고 있다. 특히 미국의 과학기술정책실(Office of Science and Technology Policy, OSTP)은 개방적이고 공정한 연구를 위해 2023년을 ‘오픈 사이언스의 해(The Year of Open Science)’로 선포하는 등 많은 관심을 보이고 있다(The White House, 2023). 연구자는 오픈 사이언스를 통해 공유된 연구데이터를 재활용할 수 있으며 다양한 관점에서 연구 결과를 재현 및 입증하여 새로운 학술적 발견에 도움을 줄 수 있다. 또한, 연구결과의 교차 검증으로 연구의 신뢰도를 높이는 데 기여할 수 있을 뿐만 아니라 연구의 중

복 진행을 방지하여 불필요한 연구비 지출을 줄이는 데에도 유용하다.

우리나라 역시 연구데이터 관리와 관련한 노력을 지속하고 있는데, 그 일환으로 2018년에 국가연구데이터 공유·활용전략을 수립하고 2019년에 연구데이터관리계획(Data Management Plan, DMP)을 시범적으로 적용하였으며, 국가연구데이터플랫폼(DataON)을 구축하는 등 연구데이터 정책을 추진 및 시행하였다. 2022년에는 ‘국가연구데이터 관리 및 활용 촉진에 관한 법률안’이 제안되기도 하였으며 2023년과 2024년에는 각각 ‘국가R&D 연구데이터 통합 관리법’과 ‘국가연구데이터 관리 및 활용 촉진에 관한 법률안’이 발의되었다. 이를 통해 해외와 국내 모두 공적자금을 투입한 연구데이터를 국가의 주요한 자산으로서 인정하고 있다는 것을 확인할 수 있으며 현재 우리나라에서는 이와 관련하여 국가연구개발혁신법이 시행되고 있다.

전자정부법에 따르면, 행정정보란 행정기관 등이 직무상 작성하거나 취득하여 관리하고 있는 자료로서 전자적 방식으로 처리되어 부호, 문자, 음성, 음향, 영상 등으로 표현된 것을 의미한다. 우리나라의 공적자금을 투입한 연구데이터 역시 행정기관의 직무상 취득된 자료이기 때문에 행정정보로 인정된다. 또한, 공공기관은 공공기록물관리법 시행령에 따라 행정정보 데이터세트를 관리해야하므로 연구데이터는 행정정보 데이터세트로서 기록관리의 대상이 된다. 그러나 기존의 연구는 연구결과로 산출된 연구성과물과 그 효용성에 초점이 맞춰 있으며, 연구과정 중에 생성된 연구데이터 자체에 주목하는 연구는 비교적 많이 진행되어 있지 않은

상태이다. 연구데이터가 이관, 평가, 처분 등의 기본적인 기록관리 절차를 따르기 위해서는 보다 상세한 행정정보 데이터세트 기록관리방안이 마련되어야 하나, 현재로서는 연구데이터에 대한 구체적인 기록관리 방안이 부족한 실정이다. 본 연구에서는 연구데이터의 보존과 활용이 중요하게 여겨지고 있는 전세계적 추세를 반영하여, 연구데이터 보존 메타데이터에 주목하였다.

기존에도 다양한 연구데이터 메타데이터 표준이 존재하나, 이러한 표준들은 생산자, 보존일자, 처분 등의 기록관리 요소를 수용하고 있지 않으며 그 목적이 '기록관리'가 아닌 연구데이터의 보존과 재활용을 위한 검색 접근점 지원에 있다. 따라서 기존의 연구데이터 메타데이터 표준을 통해서는 제도적으로 명시된 연구데이터 기록관리 필요성을 충족시키기에 한계가 있다. 한편, 현행 기록관리 메타데이터 표준에서도 연구데이터의 다양한 특성을 반영할 수 있는 메타데이터 요소는 다소 부족한 상황이다. 마찬가지로 기존의 기록관리 메타데이터 표준을 통해서는 연구데이터 메타데이터의 본질인 연구데이터 보존과 재활용이라는 목적을 달성하기에 다소 어려움이 있다.

기록관리 메타데이터 표준인 NAK 8:2022 (v2.3)에 따르면 기록물의 특성에 따라 특별히 생산, 관리가 필요한 메타데이터는 별도로 정의하도록 하고 있다. 따라서 '기록관리 제도 준수'와 '연구데이터 보존 및 활용'이라는 두 측면을 모두 만족시킬 수 있는 합집합으로서의 새로운 메타데이터의 필요성이 요구된다고 판단할 수 있다.

이에 본 연구는 국내외의 기록관리 메타데이

터와 연구데이터 메타데이터 표준을 비교·분석하여 연구데이터 특성을 반영한 기록관리 메타데이터 요소를 도출하고자 한다. 본 연구의 연구 질문을 정리한 바는 다음과 같다.

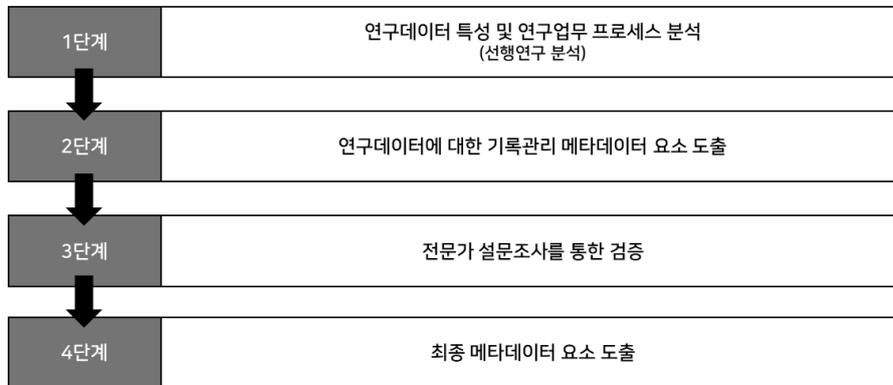
- 연구 질문 1: 연구데이터의 특성 및 연구과제 진행 프로세스 분석은 무엇인가?
- 연구 질문 2: 국내외 기록관리 메타데이터 표준 및 연구데이터 메타데이터 표준의 비교분석을 통해 도출될 수 있는 연구데이터 기록관리 메타데이터 요소는 무엇인가?
- 연구 질문 3: 전문가 설문조사를 통해 확인할 수 있는 메타데이터 요소 도출(안)의 중요도는 어느 수준인가?

1.2 연구대상 및 연구방법

본 연구는 정부 지원하에 이루어지는 공공기관 및 연구기관의 연구과정에서 생산·보존·관리되는 연구데이터를 중심으로 연구를 수행한다. 국가연구개발혁신법(법률 20057호)에 따르면 연구데이터는 연구개발과제 수행 과정에서 실시하는 각종 실험·관찰·조사 및 분석 등을 통하여 산출된 사실 자료로서 연구결과의 검증에 필수적인 데이터를 의미한다.

연구대상은 연구과제 수행 도중 획득·관리되는 원시 연구데이터 및 최종 산출된 연구데이터이다. 이는 부호, 문자, 음성, 음향, 영상, 이미지 등으로 표현된 직무상 취득 자료를 모두 포함한다.

본 연구의 목적을 달성하기 위하여 진행할 연구방법 및 절차는 다음의 <그림 1>과 같다.



〈그림 1〉 연구방법 및 절차

1단계에서는 연구방향성 및 전반적인 연구 범위를 설정하기 위해 연구데이터 기록관리를 다룬 선행연구와 메타데이터 설계 관련 문헌을 분석하였다. 이를 통해 연구데이터의 특징을 파악하고, 연구업무 진행 프로세스를 분석하여 연구데이터 기록관리 메타데이터 요소 도출 시에 고려해야 할 주의사항을 확인하였다.

2단계에서는 1단계에서 확인한 메타데이터 요소 고려사항을 기반으로 국내외의 기록관리 메타데이터 표준 및 연구데이터 메타데이터 표준을 비교·분석하고, 매핑을 통하여 연구데이터 기록관리 메타데이터 요소를 도출하였다.

본 연구의 메타데이터 요소는 연구데이터의 '기록관리'를 위함이 그 목적으로, 여러 기록관리와 관련한 메타데이터 표준을 기본 골자로 하여 설계되었으며, 설계(안)에 기존의 기록관리 메타데이터에선 수용되지 않았던 연구데이터 메타데이터 표준의 요소를 추가하는 방식으로 요소 도출이 진행되었다. 즉, 기존의 기록관리 메타데이터와 연구데이터 메타데이터 표준의 합집합으로써, 이는 연구데이터 메타데이터 표준이 생산자, 보존일자, 처분 등 기록관리 관

점의 요소를 포함하고 있지 않은 점을 고려하였기 때문이다.

활용한 기록관리 관련 메타데이터 표준은 기록관리 메타데이터 표준(NAK 8:2022(v2.3)), 행정정보 데이터세트 기록관리 기준 - 관리기준표의 작성 및 이관규격(NAK 35 2020(v1.0)), General International Standard Archival Description (ISAD(G)), PREservation Metadata: Implementation Strategies(PREMIS) 3.0 데이터모델이 있으며, 연구데이터 관련 메타데이터 표준으로는 Data Catalog Vocabulary Version 2(DCAT V2), DataCite Metadata Schema를 활용하였다. 각 메타데이터 표준을 활용하게 된 근거는 다음과 같다.

- NAK 8:2022(v2.3): 기록물의 4대 요건을 보장하기 위한 기록물의 기본적인 메타데이터 요건을 제시한다는 점에서 가장 기초적인 배경으로 활용하였다. 또한, NAK 8:2022(v2.3)에 따르면 해당 표준의 메타데이터 요소들은 결재과정에서 생산된 문서류 위주의 기록물에 적용되지만 시청각기록물, 웹기록물, 행정정

보 데이터세트 등의 기록물에도 이 표준을 적용할 수 있다고 명시되어 있다.

- NAK 35 2020(v1.0): 연구데이터가 행정정보 데이터세트에 해당하는 것을 고려하여 해당 표준을 활용하였다.
- ISAD(G): ISAD(G): 국제표준기록물기술 규칙으로서 기록물 기술의 통일성과 정확성을 보장할 수 있으며, 매체나 유형에 관계없이 모든 보존 기록에 적용되는 원칙과 요소를 제안하고 있다는 점에서 해당 표준을 활용하였다.
- PREMIS 3.0 데이터모델: 디지털 보존 메타데이터 표준으로서, 연구데이터는 대부분 디지털 객체로 생성되고 있음을 고려하였을 때 디지털 객체의 장기 보존 측면에서 해당 표준을 활용하였다.
- DCAT V2(W3C, 2019): 웹에 게시된 데이터세트 카탈로그 간의 상호운용성을 확보하기 위해 설계된 표준으로서 연구데이터 접근과 활용, 메타데이터의 통합을 지원하는 역할을 위해 해당 표준을 활용하였다.
- DataCite Metadata Schema(DataCite, 2021): 범용적 연구데이터 기술을 위한 국외 표준 메타데이터로, 인용 및 검색을 위해 연구데이터를 정확하고 일관되게 식별하기 위해 해당 표준을 활용하였다.

3단계에서는 2단계에서 도출한 메타데이터 요소의 중요도를 검증하기 위하여 기록관리 전문가, 문헌정보학 전문가, 메타데이터 전문가, 연구데이터 전문가 총 32명을 대상으로 온라인 설문조사를 실시하였다.

마지막 4단계에서는 3단계의 전문가 설문조사를 통한 요소 검증 결과를 바탕으로 2단계에

서 도출한 요소(안)에 대한 수정을 진행하고 최종 메타데이터 요소를 도출하였다.

1.3 선행연구

본 장에서는 연구데이터 기록관리 메타데이터 설계를 위해 국내외의 선행연구를 분석하여 시사점을 도출하였다. 분석한 국내외의 선행연구는 크게 연구데이터 메타데이터, 연구성과물 또는 연구데이터의 기록관리, 메타데이터 설계와 관련한 세 주제로 구분할 수 있다.

연구데이터 메타데이터와 관련한 국외 선행연구로는 Greenberg(2010), Wilson(2010), Huang, Qin(2024)의 연구가 있다. 이 논문들은 디지털 데이터의 상호운용성과 장기적 유용성을 보장하는 데 중점을 두고, 이를 위한 수단으로 메타데이터 관리를 지목하여 그 중요성을 재확인하고 있다는 점에서 공통점을 보인다.

Greenberg(2010)는 과거부터 현대까지의 과학데이터 메타데이터 관리 관행의 변화를 검토하였다. 해당 연구는 메타데이터 관리 관행이 도서관 시스템에 기반을 두고 있으며, 이는 디지털 과학데이터의 보존 및 접근 가능성을 높이는 데 이바지하므로 과학 연구의 데이터 저장, 검색, 관리에서 메타데이터의 역할이 필수적임을 강조하였다. 특히, 표준화된 메타데이터 및 데이터 수명주기 관리에 주목하여 미래의 메타데이터 개발은 저장소 기반으로 하되 데이터 상호운용성과 협력 강화를 목표로 해야 함을 제안하였다.

Wilson(2010)은 유전체 데이터의 급속한 증가와 다양한 응용 분야에서 데이터 재사용 요구가 증가함에 따라 적절한 메타데이터 스키마

개발이 주요한 과제로 떠오르고 있음을 지적하고, 유전체 데이터의 메타데이터 체계가 충족해야 할 사항을 분석하였다. 유전체 연구자들이 중요시하는 메타데이터 요구사항은 이식성, 재사용성, 상호운용성 등으로, 메타데이터 스키마가 데이터의 교환과 보존을 지원할 수 있도록 유전체 데이터의 FAIR 원칙(Findable, Accessible, Interoperable, Reusable)을 실현하는 데이터 관리 표준이 강화되어야 한다고 강조하였다.

Huang, Qin(2024)의 연구는 디지털 데이터가 증가하는 현대 환경에서 메타데이터가 데이터 보존의 핵심 요소임을 강조하며 데이터 보존의 효율성과 실질적 유용성 사이에서 균형을 잡기 위하여 어떠한 점이 고려되어야 할지, 어떤 메타데이터 요소가 필수적일지에 대한 명확한 기준이 필요함을 주장하였다. 또한, 다양한 학문 분야와 조직에서 생성되는 데이터를 통합하고 상호운용성을 촉진하기 위해 메타데이터 표준의 중요성을 강조하였으며, 기록 보존을 위한 메타데이터와 보존을 위한 메타데이터의 유사점에 주목하였다. 본 연구는 보존 우선순위 및 비용 효율성, 과학자 커뮤니티에서의 소통을 통한 지속적인 보존 전략 개선 등의 측면에서 데이터 보존 정책 수립의 방향성을 제공한다. 이 점에서 의의를 지닌다.

연구성과물 또는 연구데이터와 관련한 국내 연구로, 박동진(2006)은 연구성과물의 효율적 관리와 상호 운용성 강화를 통해 국가 과학기술 정보 인프라를 강화하고, 연구성과물에 대한 접근성을 높일 수 있다고 주장하며 과학기술 분야의 연구성과물 정보를 체계적이고 일관성 있게 관리하기 위한 표준 메타데이터를 개

발하고자 하였다. 본 연구는 연구성과물의 범위 확장 및 이에 대한 표준 메타데이터를 제시함으로써 국내 연구성과정보 관리의 기틀을 마련하였다.

장보성, 남영준(2011)은 국가 자금으로 생산된 공공기록물로서 디지털 연구성과물의 공공재적 가치를 인식하고, 체계적인 관리를 위한 장기보존 메타데이터를 설계하고자 하였다. 또한, 기존의 메타데이터는 기술적 관점에 치우쳐 있어 연구성과물의 생산 맥락, 구조, 관리 과정을 포괄하는 보존적 메타데이터 요소가 부족하기 때문에 새로운 표준 설계의 필요성을 강조하였다. 본 연구는 디지털 연구성과물의 데이터 관리 표준 정책 및 기술적 기준 설정에 대한 방향성을 제공하였다는 점에서 의미가 있다.

박동진(2006), 장보성, 남영준(2011)의 연구는 모두 국가 자산으로서의 연구성과물 및 연구기록물의 중요성을 인식하였으며 연구기록물 관리 체계의 방향성을 정립하고 개선점에 대한 의견을 개진하였다는 점에서 의의를 가진다. 그러나 연구의 결과로서 산출된 최종 결과물에 주로 주목하여 연구의 수행 과정에서 생성되는 연구데이터는 비교적 고려하지 않았다.

이미영(2014; 2015)은 연구기록의 본질적 특성과 가치를 반영하여 행정기록과는 다른 평가기준이 적용되어야 하며 평가와 연계하여 연구기록의 생산부터 보존, 활용에 이르기까지 연구자와 기록관리전문가의 협력을 통한 체계적이고 차별화된 관리 프로세스가 정립되어야 한다고 주장하였다. 본 연구에 따르면 연구기록의 평가는 유일성, 재활용 가능성, 학문적 가치, 데이터의 신뢰성과 무결성 등 다양한 요소가 포함되어야 하고, 추상적 가치판단에서 벗

어나 정보적·증거적 가치를 객관적이고 구체적으로 판단할 수 있는 기준이 요구된다. 이미영(2014; 2015)의 연구는 연구기록의 평가를 통해 단순히 보존 및 처분 여부를 결정하는 것을 넘어, 연구기록 관리 체계의 개선과 연구성과물의 가치를 극대화할 수 있는 기반을 제공하고 있다.

구찬미, 김순희(2017; 2019)는 연구기록물의 생산과 활용이 순환적으로 이루어지며, 기존 연구기록물의 재활용으로 새로운 연구의 생산성과 효율성을 증대시킬 수 있다는 점에서 연구기록물의 '활용'을 보장하는 것이 연구기록물 관리의 주요점이라고 주장하였다. 더불어 연구기록물에 대한 법적 정의와 관련 표준이 부족하여 이에 대한 체계적 관리가 어렵다는 점을 지적하고, 연구기록물과 연구개발사업 간의 결합 관계가 연구기록물의 맥락을 이해하는데 필수적인 요소임을 고려하여 DMP 및 RRMP를 통한 체계적 연구기록물 관리 방안을 구체적으로 제시하였다. 본 연구는 연구데이터, 연구물질과 같은 비정형 연구기록물을 관리대상으로 포함하여 연구기록물의 개념을 확대시켰으며, 기록관리 관점에서 고려하여야 할 연구기록물의 특성을 도출하여 그 활용성을 증대시키는 것에 기여하였다.

안채영, 김지현(2020)은 연구기록이 장기적으로 후속 연구와 정책 결정에 활용될 수 있는 중요 자산임에도 불구하고 현행 공공기록물 평가 체계를 적용하기에는 그 특수성을 반영하기 어려운 한계가 있음을 지적하고, 연구기록의 특성과 가치를 고려한 평가 기준을 제안하기 위하여 국내 정부출연연구기관에 근무하는 기록관리 전문가 11명과 인터뷰를 진행하였다. 본

연구에서 연구기록의 평가기준으로 제시된 사항은 미션과의 관련성, 연구 가치, 역사적 가치, 해당 분야의 전문가심사, 유일성, 재현불가능성, 재배포 가능성, 충분한 기술, 접근성, 이용 가능성, 경제적 상황, 비용, 대체비용 등으로 나타났다.

이미영(2014; 2015)과 구찬미, 김순희(2017; 2019), 안채영, 김지현(2020)의 연구는 앞선 연구에서 더 나아가 연구 과정 중에 생산된 원시 연구데이터도 함께 고려하였으며, 연구기록물의 '재활용성'에 주목하여 기존 관리 체계에 대한 개선 의견을 개진하였다. 특히 구찬미, 김순희(2017; 2019), 안채영, 김지현(2020)은 기존 연구기록물 관리 체계에 대한 개선 방안으로 데이터 관리 계획(DMP)을 활용하는 것을 언급하였다는 점에서 공통점을 보인다.

다음의 서지인, 노지현(2022), 이상훈 외(2022)의 연구는 각각 기록관리 메타데이터와 특정 분야 연구데이터 메타데이터 설계를 진행한 연구이다.

서지인, 노지현(2022)은 디지털화 기록관리 및 품질 유지에 필요한 메타데이터 요소안을 설계하고자 디지털화 기록의 정의 및 특성을 살펴보고, 국내 및 해외 표준을 분석하여 디지털화 기록 관리에 필요한 메타데이터 요소를 비교하였다. 이 과정에서 디지털화 기록과 원천 기록 간의 관계를 정립하고 자원 유형별 특성에 맞춘 자동 추출 메타데이터 요소를 확대하였으며 업무 과정 및 프로젝트와 관련한 메타데이터를 작성하였다.

이상훈 외(2022)는 농촌진흥청 식품성분 연구데이터 데이터베이스의 효율적인 데이터 관리 및 공유를 위해 메타데이터 요소를 도출하

고 스키마를 개발하였다. 이를 위해 식품성분 연구데이터의 특징 및 식품성분 분석 연구업무 프로세스를 살펴본 후 국제 및 국내 메타데이터 표준(DCAT, DataCite, TTAK.KO-10.0976 등)과의 비교·분석을 진행하였다. 그 결과 소재정보 메타데이터, 성분분석 메타데이터를 도출하고 필수요소, 권고요소, 선택요소를 각각 선정하였다. 해당 요소는 연구자 및 메타데이터 전문가의 검토를 통해 검증하였다.

연구성과물 또는 연구데이터 기록관리를 다룬 기존의 선행연구들은 연구성과물의 중요성을 인식하고 그 특징을 포착하여 주요 특성을 고려한 관리 체계와 방향성을 모색하였다. 그러나 연구 수행 과정 중에 생산되는 원시 연구데이터 및 최종 산출된 연구데이터에 대한 구체적인 장기보존 방안을 제시하는 데에 다소 소홀함이 있었다. 본 연구는 장기보존 방안을 고려함에 있어, 보존 메타데이터의 중요성을 강조한 여러 선행연구를 참고하여 연구데이터의 특성을 반영한 연구데이터 장기보존 메타데이터의 직접적인 설계를 진행하였다는 점에서 기존 연구와 차별성을 지닌다.

2. 연구데이터의 특성

2.1 연구데이터의 정의

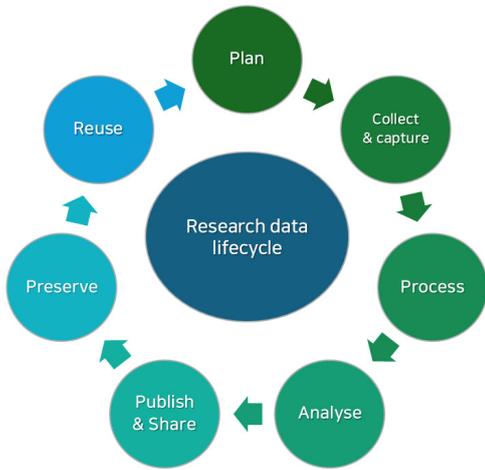
연구데이터의 정의는 국가연구개발정보처리 기준에서 제시한 내용¹⁾ 외에도 다양하게 존재한다. Longwood Medical Area Research Data

Management Working Group(LMA RDMWG, n.d.)은 연구데이터는 대학 또는 외부에서 연구비를 지원받아 연구수행과정에서 생산되거나 수집된 기록을 의미하며, 연구데이터에는 연구결과의 재현이나 평가, 연구결과의 출판에 필요한 기록들이 포함된다고 정의하였다. NC State University(n.d.)에서는 연구데이터가 광범위한 유형의 정보를 다루고 있으며, 연구 결과를 검증하는 데 필요한 경우 과학계에서 일반적으로 인정되는 기록된 사실 자료로 정의될 수 있다고 밝히고 있다.

국내 선행연구에서 연구데이터의 특성으로 언급된 바로는, 연구데이터는 과제기간 동안 생산과 활용이 연속적으로 일어나기 때문에 기존 데이터의 분석과 재활용을 보장하는 것이 중요하며, 국가연구개발사업과의 관계성이 데이터의 활용 맥락을 이해하는데 필수적이고 기존 기록생애주기 개념을 적용하기에도 어렵다(구찬미, 김순희, 2017). 또한, 안채영, 김지현(2020)은 국내 정부출연연구기관에서 근무하는 기록연구사 11명과 인터뷰를 진행한 결과, 연구기록의 주요한 평가기준으로 메타데이터의 충분한 기술, 연구기록의 재현가능성, 정부 정책이나 사회적 시대상 등의 외부적 환경, 기관의 사명, 법적 제한사항 등이 제시되었음을 밝혔다.

위의 내용을 종합한 연구데이터의 핵심 개념은 '수행 과정', '연구결과', '재활용', '검증', '재현'으로 대표될 수 있다. <그림 2>는 이러한 관리 요소를 반영한 연구데이터 생애주기(Lifecycle)를 나타낸 것이다(UK Data Service, 2019).

1) 연구과정에서 산출된 사실 자료로서 연구결과를 검증하기 위해 필수로 요구되는 데이터.



〈그림 2〉 연구데이터 라이프사이클 (Lifecycle) (UK Data Service, 2019)

즉, 이상의 내용으로 연구데이터를 정의하면 ‘연구 수행 과정 중 산출된 다양한 유형의 사실 자료로서 추후 재활용이 가능하도록 연구결과 의 재현 및 검증을 위해 요구되는 데이터’로 정리할 수 있다.

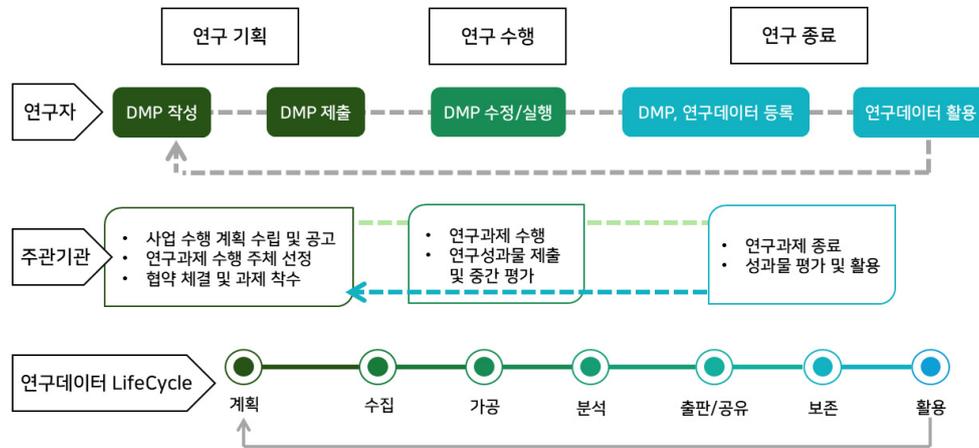
2.2 연구데이터의 보존

기존 진행된 연구들은 연구성과물에 대한 중요성을 인식하고 재활용 가능성과 같은 중요한 특성을 고려한 관리 체계를 요구하였다는 점에서 의미가 있으나, 연구데이터 보존의 주요 목적이 재활용을 위한 연구결과의 재현 및 검증 임을 명확히 하지 않아 연구 수행 과정 중에 생산되는 연구데이터에 대한 구체적인 보존 방안 논의가 다소 미비했다는 점에서 한계가 있다. 디지털 환경에서 전자기록물의 관리는 생산 시점부터 기록과 데이터를 접수하고 이에 대한 분석을 통해 이용자들이 요구하는 정보를 사전

에 확보하여 전자기록물 활용의 효과를 극대화 하는 것이 중시된다(전한역 외, 2023). 연구데이터는 디지털 환경에서 수집 및 생산되는 경우가 많고, 연구데이터를 활용하고자 하는 이용자는 보존된 메타데이터를 통해 해당 연구데이터가 어떠한 방식으로 생산되었는지에 대한 정보를 확인할 수 있어야 한다.

기록관리의 기본 원칙에 대한 표준인 ISO 15489-1:2016에 따르면, 기록은 내용과 메타데이터로 구성되며 메타데이터는 기록의 내용, 구조, 맥락과 일정기간 동안의 기록관리 이력을 기술한 데이터이다. 메타데이터는 기록의 생산 맥락 및 업무맥락을 포함하여 관련된 행위 주체에 대한 정보와 기록 자체의 내용, 외형, 구조 및 기술적 속성에 관한 정보 등을 제공할 수 있어야 한다. 특히, 디지털로 생산된 기록의 경우 기록의 맥락, 내용, 구조가 서로 분리되어 있으므로 각 요소를 동시적으로 보존하기 위해서는 메타데이터 생산과 유지가 필수적이라고 할 수 있다. 따라서 연구데이터 보존을 위한 기록관리 방안으로서 메타데이터 요소를 도출하고, 메타데이터 설계에 있어 연구데이터의 재현과 검증이라는 목적 의식을 구체화했다는 점에서 본 연구는 의의가 있다.

연구데이터의 생산 및 관리, 보존에 있어 중점적으로 확보해야 할 기록관리 메타데이터 요소를 도출하기 위하여 연구업무 프로세스를 분석하고자 했다. 〈그림 3〉은 이에 대한 도식화 이미지이며, 연구 데이터 공유, 활용 전략(안)(과학기술정보통신부, 2018), 국가연구개발정보처리기준, 한국연구재단 DMP 템플릿(한국연구재단, 2021), International Council on Archives(2010)의 연구개발 프로세스와 UK



〈그림 3〉 연구과제 프로세스

Data Service(2019)의 Data Lifecycle의 내용을 종합하였다.

연구과제 프로세스는 크게 연구 기획, 연구 수행, 연구 종료의 세 단계로 나눌 수 있으며 각 단계는 이는 연구자, 주관기관, 연구데이터 Lifecycle의 측면에서 요구되는 활동과 긴밀하게 연결될 수 있다. 주목해야 할 점은 연구과제 업무 과정에서 연구데이터의 생산은 활용과 동시에 이루어지기 때문에 전 과정이 연속적으로 순환된다는 것이다.

따라서 연구데이터의 재활용을 위해서는 연구결과 뿐만 아니라 연구과정에 대한 재현이 필수적이며, 이를 위해서는 연구에 대한 환경 정보와 해당 연구데이터가 어떠한 배경에서 생산되었는지에 대한 맥락정보가 반드시 보존되어야 한다. 미국, 영국 등의 해외 주요 선진국은 해당 사항에 대한 접근을 용이하게 하고 연구데이터 재사용을 촉진하기 위해 데이터관리계획(Data Management Plan, DMP)을 적용하여 의무화를 시행하고 있으며, 우리나라 역시 2019년부터 이를 시범적으로 적용하고 있다. 연

구데이터 관리에 DMP를 적용하면 생산 이전 단계에서부터 연구기금지원기관별이나 연구수행기관별, 연구주제 분야별로 필요에 따라 다양한 메타데이터를 확보할 수 있고 연구데이터의 관리, 보존, 공유에 필요한 메타데이터와 자산적 가치 보호에 관한 메타데이터 등도 관리대상 메타데이터로 확장하여 관리할 수 있다 (구찬미, 김순희, 2019).

기록관리 과정에서 연구데이터의 장기보존을 위해서는 위 사항들을 고려한 보존지침이 필수적이라고 할 수 있으나, 연구데이터의 관리지침으로 기능해야 할 현행 행정정보 데이터세트 관리기준표는 이를 만족하지 못하고 있어 개선이 필요하다.

2.3 행정정보 데이터세트 관리기준표

NAK 35:2020(v1.0)에 의하면 행정정보 데이터세트는 시스템을 기본단위로 하여 행정정보 데이터세트 기록관리기준표를 작성하여 관리 및 운영되어야 한다. 기록관리기준표는 기록

의 처분기준을 제시하기 위한 지침이며 기록관리 메타데이터는 기록의 내용, 구조, 맥락 및 이력을 기술한 정보로 서로 그 목적이 달라 구분되는 개념이나, 행정정보 데이터세트 기록관리 실행 매뉴얼(국가기록원, 2020)에 따르면 행정정보 데이터세트 기록관리기준표는 행정정보 데이터세트의 기록관리 요소를 정리한 표임과 동시에 메타데이터이기도 하다.

즉, 행정정보 데이터세트의 기록관리는 기존의 전자문서 기록관리체계와는 달리 목적성이 다른 두 개의 역할이 하나의 체계에 중복된 상황으로, 행정정보 데이터세트의 유연한 기록관리를 위해서는 이에 대한 분리 작업이 필요하다(김지혜, 2023).

3. 연구데이터의 기록관리를 위한 메타데이터 요소 도출

3.1 기록관리 및 연구데이터 메타데이터 표준의 요소 분석

본 장에서는 기존의 기록관리 메타데이터 표준과 연구데이터 메타데이터 표준을 조사하여 행정정보 데이터세트 및 연구데이터에 적용할 수 있는 요소를 선별하고자 하였다. 적용 대상으로 선정한 메타데이터 표준은 다음과 같은 기준에 따랐다.

- 첫째, 국제표준, 공공표준과 같이 공인성과 유효성이 검증된 메타데이터 표준
- 둘째, 기록관리 분야에서 이미 사용되고 있거나 기록관리 분야에서 사용할 수 있는

표준

- 셋째, 광범위한 유형의 데이터에 적용할 수 있는 표준

위 기준에 따라 선정된 표준은 총 6개로, 기록관리 분야 메타데이터 표준 4개(기록관리 메타데이터 표준(NAK 8:2022(v2.3)), 행정정보 데이터세트 기록관리 기준 - 관리기준표의 작성 및 이관규격(NAK 35 2020(v1.0)), General International Standard Archival Description (ISAD(G)), PREservation Metadata: Implementation Strategies(PREMIS) 3.0 데이터모델)와 연구데이터 메타데이터 표준 2개(Data Catalog Vocabulary Version 2(DCAT V2), DataCite Metadata Schema)로 이루어졌다.

NAK 8:2022(v2.3)은 가장 기초적인 기록관리 메타데이터 요소를 도출하기 위해, NAK 35 2020(v1.0)은 연구데이터가 행정정보 데이터세트로서 기록관리의 대상임을 고려하여 포함되었으며, ISAD(G)는 매체 및 유형에 관계없이 모든 보존 기록에 적용되는 원칙 및 요소를 제안하고 있기 때문에, PREMIS 3.0 데이터 모델은 디지털 객체의 장기 보존 측면에서 활용되었다. DCAT V2는 연구데이터 및 데이터 서비스의 검색 가능성을 향상시킬 수 있는 상호운용성 보장 표준으로서, DataCite Metadata Schema는 범용적 연구데이터 기술을 위한 국외 표준 메타데이터로서 요소 분석에 활용하게 되었다.

이 외에도 한국과학기술정보연구원(KISTI)에서 주관한 ‘출연(연) 연구데이터 관리·활용 방안 연구(국가과학기술연구회, 2019)’와

'CoreTrustSeal(CTS) 리포지터리 인증 요소(CoreTrustSeal, 2022)'와 같은 자료를 참고하였다.

출연(연) 연구데이터 관리·활용 방안 연구는 한국과학기술정보연구원(KISTI)에서 주관하였고 2019년 3월 국가과학기술연구회(NST)에서 발간되었다. 국내R&D 과제에서 생산된 연구데이터의 생산, 관리, 공유 현황 파악을 목적으로 하여 국가과학기술연구회소속의 출연(연) 기관 과제 책임자 301명을 대상으로 연구데이터 관리 제도, 체계, 플랫폼 현황을 조사한 연구이다.

CTS(CoreTrustSeal)는 커뮤니티 기반 비영리 조직으로, WDS-DSA 데이터 리포지터리 요구조건(DSA-WDS Core Trustworthy Data Repositories Requirements)에 기반하여 데이터 리포지터리에 핵심 수준 인증(core level certification)을 제공한다(김주섭 외, 2022). CTS 리포지터리 인증 요소는 신뢰할 수 있는 데이터 리포지터리가 갖추어야 할 특징 및 요구사항을 의미한다. 리포지터리 인증이 이루어진 체계는 정확성, 신뢰성 등을 위시한 데이터 품질을 개별적으로 확인하기에는 현실적인 한계가 존재하므로 개별 데이터의 품질 대신 데이터공유플랫폼(리포지터리)의 품질을 검증하는 방안이 대두되었기 때문이다.

3.2 메타데이터 설계 원칙 수립 및 요소 매핑

본 장에서는 연구데이터 기록관리를 위한 메타데이터 요소를 도출하기 위해 연구데이터의 특성을 고려한 설계 원칙을 수립하고, 3.1의 메타

데이터 표준과 그 요소를 매핑하여 연구데이터 기록관리 메타데이터 요소 초안을 도출하였다.

연구데이터는 주로 연구를 수행한 연구자가 직접 등록하며, 한 연구에서 생산 및 수집된 데이터들을 개별적으로 입력하는 것이 아니라 여러 데이터를 하나의 데이터셋으로 묶어서 관리하는 방식을 취하기 때문에(이미화 외, 2020) 이를 고려하여 개별 데이터가 아닌 데이터셋을 단위로 설정하였으며, 메타데이터를 상위요소-하위요소-세부요소 세 단계로 나누어 구성하였다.

또한, 연구과제정보는 연구데이터를 포함한 연구기록물 생산맥락을 보존하는데 있어 필수적인 사항이기 때문에 메타데이터 설계 시 해당 요소를 포함하였다. 작성 순서는 연구과제 진행 프로세스를 따라 연구개발사업-사업과제(성과물)-데이터(세트) 순을 따랐다. 연구과제 정보가 중요한 메타데이터 요소에 해당함을 인지하고 있으나 연구데이터 메타데이터 설계 또는 기술에 있어 해당 요소를 제외하는 경우가 존재할 수 있기 때문에(이상훈 외, 2022) 연구개발사업정보 및 연구과제정보를 기록관리 메타데이터에 필수적으로 포함시켰다.

매핑은 현행 기록관리 메타데이터 표준과의 호환성을 위하여 NAK 8을 기초로 각 기록관리 메타데이터 표준(NAK 8, NAK 35, ISAD(G), PREMIS 3.0)을 비교하며 진행되었다. NAK 8에서 한 번이라도 등장한 요소는 연구데이터 기록관리 메타데이터 요소 초안에 우선 포함하였으며, 매핑 결과 '기록물명, 생산자, 크기, 권한, 관계, 생산/관리/이용/보존이력, 무결성 체크' 요소는 기록관리 메타데이터 표준 4개에 공통적으로 등장하였기 때문에 초안에 활용할 수

있음을 확인하고 필수 요소로 지정하였다.

또한, 연구데이터의 정확성을 유지하고 재현성을 확보할 수 있도록 연구데이터 메타데이터 표준(DCAT V2, DataCite)에서 공통적으로 등장한 요소(데이터명, 수집/분석 주제, 유형, 일시, 범위, 언어)를 초안에 포함시키고 '데이터 정보'라는 상위요소로 지정하였다. 두 개의 표준 중 하나에서만 등장했거나 등장하지 않았지만 표준 외 참고사항(출연(연) 연구데이터 관리·활용 방안 연구, CoreTrustSeal(CTS) 리포지터리 인증 요소)에서 언급이 있는 경우(수집/분석 목적, 방법, 메타데이터 표준) 역시 필수 요소로 지정하여 내용을 보강하고자 하였다.

연구데이터 기록관리 메타데이터 요소 초안을 도출함에 있어 가장 중요한 원칙은 '연구데이터의 정확성', '접근성', '재현성', '재활용성의 보장'이다. 해당 원칙을 실현하기 위해 기존 메타데이터 표준 등장 요소의 매핑 이후에도 기존 표준을 보완하거나 연구데이터의 활용 및 관리 개선을 위해 신규 요소를 추가하고 불필요한 요소를 삭제하는 방식으로 조정을 진행하였고, 조정 방식을 세부적으로 정리하면 다음과 같이 나타낼 수 있다.

- 정확성 보장: 연구데이터의 신뢰성과 품질을 위해 수집 및 분석 맥락을 명확히 표현하는 요소('데이터정보' 요소의 '수집방법', '수집목적', '분석방법', '분석목적' 등)를 추가하여 데이터의 기초적인 정보와 맥락을 담아내도록 하였다.
- 생산 맥락 유지 및 통합 관리: 연구 프로젝트에서 생성된 데이터가 연구 전반의 흐름과 연계될 수 있도록 '국가연구개발사업정보', '연구개

발과제정보'를 신설하였다. 이와 더불어 '과제 진행상태'와 '산출물'을 확인할 수 있게 하여 데이터의 생성 과정을 일관성 있게 통합 관리할 수 있도록 조치하였다.

- 데이터 접근성 강화: 연구데이터의 활용성과 접근성을 높이기 위해 접근 권한을 포함한 세부 요소('URL' 등)를 추가하여, 연구자들이 데이터의 존재 여부와 위치를 손쉽게 파악할 수 있도록 하였다. 해당 요소는 국가과학기술 연구회(2019)에서 언급된 사항을 바탕으로 추가되었다.
- 데이터 해석 지원: 연구데이터 해석에 필요한 '첨부파일' 등의 부가 정보를 제공하여 데이터의 의미와 활용 가치를 높이며, 특정 데이터의 활용 목적을 명확히 할 수 있다. 첨부파일의 예시로는 통계처리를 위한 통제어휘집 또는 용어리스트 등을 들 수 있다.
- 재활용성을 위한 인용 정보 제공: 연구데이터를 재활용하려는 연구자를 위해 '라이선스' 및 인용 정보를 체계적으로 포함하여 데이터의 법적 및 윤리적 활용 범위를 명확히 하였다. 이에 대한 세부 요소는 CoreTrustSeal(2022)을 참고하여 작성하였다.
- 불필요 요소 삭제: NAK 8에서 행정정보 데이터세트에 적용하기 어렵다고 판단하거나 적용이 유의미하지 않다고 판단한 요소('기록계층', '기술', '전자기록물 여부', '유형', '매체', '분류', '보존장소', '위치 요소')는 삭제하였다.

이와 같은 상위 원칙은 기록의 일관성, 접근성, 정확성, 해석 가능성, 그리고 재활용성을 보장하여 연구데이터의 가치와 신뢰성을 유지하는 데 기여할 수 있을 것으로 기대된다.

3.3 연구데이터 기록관리 메타데이터 요소 초안

다음의 <표 1>은 연구데이터 특성에 유의하여 기록관리 메타데이터 표준인 NAK 8, NAK 35, ISAD(G), PREMIS 3.0 데이터모델, DCAT V2, DataCite의 요소를 비교·분석 후 도출된 연구데이터 메타데이터 요소(안)이다. 14개 상위요소, 57개 하위요소 그리고 61개의 세부요소로 구성되어 총 98개 요소로 이루어져 있다.

<표 1>에서 셀 배경이 음영처리 된 것은 필수 요소이며, 밑줄 처리가 된 요소는 해당 시 필수 요소이다. 기울어진 글자는 반복 가능 요소를 의미한다.

u>

‘데이터정보’ 상위요소의 ‘수집 목적’~‘분석 정보’ 하위요소의 경우, 하나의 데이터세트 내에서도 개별 데이터의 유형과 정보가 다를 수 있기 때문에 반복될 수 있는 데이터로 설정하였으며, 해당 메타데이터 요소에 대한 설명은 하나의 기술 첨부파일로 대체될 수 있다.

<표 1> 연구데이터 기록관리 메타데이터 요소 초안

요소			참조 표준의 메타데이터 요소 유무					
상위요소	하위요소	세부요소	NAK8	NAK35	ISAD(G)	PREMIS	DCAT	DataCite
기록물명(표제)	데이터세트명	대표데이터	○	○	○	○		
데이터세트 식별자	데이터식별자		○		○	○	○	○
법규정보	법령/규정			○	○			
국가연구 개발사업정보	사업번호							
	사업명							
연구개발과제정보	과제번호							
	과제명							
	과제형태	(공동/단독)						
	과제진행상태	(진행중/종료)						
	과제기간							
	과제설명							
	지원기관	기관명 지원기관 요구사항						○
	산출물	연구성과물						
데이터세트 정보	분야						○	○
	주제	주제어	○	○	○		○	○
	크기	용량	○		○	○	○	○
		단위	○					
일시	생산일시	○		○				
데이터 정보	데이터명						○	○
	데이터설명							
	데이터크기		○					
	수집 목적							
	수집 주체	책임자 기여자					○	○
	수집 유형	(실험/관측/기타)					○	○

요소			참조 표준의 메타데이터 요소 유무						
상위요소	하위요소	세부요소	NAK8	NAK35	ISAD(G)	PREMIS	DCAT	DataCite	
	수집 정보	수집 일시					○	○	
		수집 주기					○		
		수집 범위					○	○	
		수집 방법							
	분석 목적	분석 주체	책임자					○	○
			기여자						
	분석 정보		분석 일시					○	○
			분석 주기					○	
			분석 범위					○	○
			분석 방법						
	포맷		포맷명	○	○				
			포맷버전	○					
			구동SW 및 버전	○	○	○	○		
	언어						○	○	
메타데이터 표준									
암호화데이터 사용유무			○						
연계시스템 유무			○						
구성	첨부파일	첨부파일명	○						
		첨부파일 종류	○						
		첨부파일 포맷 : 구동 SW	○	○			○		
		첨부파일 저장방법	○						
생산이력	생산시스템	시스템명	○	○					
		시스템개요		○					
		구축년도		○					
		DBMS정보		○					
	생산자	기관명	○	○	○				
		기관코드	○						
		부서명	○	○	○				
		부서코드	○						
		담당자명	○	○	○				
		담당자직책	○						
생산유형		○							
이력	행위유형	(관리/이용/보존)	○		○				
		행위자명	○		○	○			
	행위자	행위자부서	○						
		행위자직책	○						
		행위일시	○		○	○			
	행위근거	○							
	행위설명	○							
	행위결과	○							
변경요소	변경요소명	○							
	변경이전값	○							

요소			참조 표준의 메타데이터 요소 유무					
상위요소	하위요소	세부요소	NAK8	NAK35	ISAD(G)	PREMIS	DCAT	DataCite
처분	보존기간		○	○				
	보존기간 사유		○	○				
	처분일시		○					
	처분방법			○				
	적용범위		○					
권한	비밀	비밀분류	○					
		비밀분류 근거	○					
		보호기간	○					
	접근	접근범위	○	○	○	○	○	○
		URL					○	
	공개	공개구분	○		○			
		비공개사유	○					
		공개제한부분	○					
		공개예정일자	○		○			
		공개관련근거	○					
		목록공개구분	○					
		목록공개제한사유	○					
		목록공개제한부분	○					
	라이선스	저자			○		○	○
		발행일			○		○	○
		타이틀			○		○	○
		버전						○
		출판사			○		○	○
		식별자					○	○
관계	관계유형		○			○	○	
	관계대상 식별자		○					
	관계설명		○					
무결성 체크	무결성검증법		○					
	무결성검증값		○					

4. 연구데이터 기록관리 메타데이터 요소에 대한 전문가 검증

4.1 전문가 설문조사를 통한 메타데이터 요소 검증

이번 장에서는 3장에서 도출한 연구데이터 기록관리 메타데이터 초안 요소에 대한 적정성

및 중요도를 검증하기 위하여 전문가 자문을 실시하였다. 이에 대한 주요 내용은 다음 <표 2>와 같다.

전문가 자문을 위해 구조화 된 조사지를 개발하여 온라인 설문조사를 실시하였다. 검증을 위한 조사지는 응답자 배경정보를 질문하는 3개의 문항과 요소와 요소별 중요도를 질문하는 14개의 문항으로 구성하였고, 각 요소별 중요도는 리커트 7점 척도를 활용하여 '전혀 중요

〈표 2〉 연구데이터 기록관리 메타데이터 요소 선정을 위한 전문가 설문조사 개요

검증 방법	주요 내용
전문가 설문조사	<ul style="list-style-type: none"> 설문 목적: 메타데이터 요소 도출(안)에 대한 적정성 검증 및 중요도 결정 설문 대상: 석사학위 이상의 기록관리, 문헌정보학, 메타데이터, 연구데이터 전문가 (기록관리 전문가(9명), 문헌정보학 전문가(8명), 메타데이터 전문가(11명), 연구데이터 전문가(4명) 총 32명) 설문 방법: KSDC DB를 활용한 온라인 설문조사 설문 기간: 2024년 11월 5일 ~ 2024년 11월 14일

하지 않음’, ‘중요하지 않음’, ‘약간 중요하지 않음’, ‘보통’, ‘약간 중요함’, ‘중요함’, ‘매우 중요함’의 보기 중에서 중요도를 선택하도록 하였다. 리커트 7점 척도 문항을 통한 중요도 검증 외의 특이사항 및 권고사항, 그 외 의견에 대해서는 마지막 자유 응답형 문항을 통해 수집하였다.

통계적으로 표본 수가 30 이상일 때 일반적으로 유의미하다고 보기 때문에, 전문가 자문은 석사학위 이상의 전문가 32명을 대상으로 진행되었다. 조사에 참여한 전문가의 구성인원은 기록관리 전문가 9명, 문헌정보학 전문가 8명, 메타데이터 전문가 11명, 연구데이터 전문가 4명이었다. 전문가 중 대학 등의 연구 경력을 포함한 전체 연구 경력이 5년 미만인 전문가는 6명(18.75%), 5년 이상 10년 미만의 연구 경력을 가진 전문가는 5명(15.63%), 10년 이상 20년 미만의 전문가는 13명(40.63%), 연구

경력 20년 이상의 전문가는 8명(25.00%)으로 나타났다. 조사는 KSDC DB를 활용하여 2024년 11월 5일부터 11월 14일까지, 총 10일간 온라인상으로 진행되었다.

다음의 〈표 3〉은 전문가 자문 결과를 ‘전혀 중요하지 않음’ 1점, ‘중요하지 않음’ 2점, ‘약간 중요하지 않음’ 3점, ‘보통’ 4점, ‘약간 중요함’ 5점, ‘중요함’ 6점, ‘매우 중요함’ 7점으로 치환하고 각 요소가 얻은 총점을 계산한 뒤 평균을 내어 이를 다시 7점 척도에 맞추어 중요도를 정량화한 것이다. 예를 들어, 세부요소 ‘대표데이터’의 경우 전문가 2명에게 5점, 8명에게 6점, 22명에게 7점을 얻어 총점 212점으로 계산되었다. 해당 요소의 평균은 총점을 32로 나눈 6.6이며, 이를 소수점 아래 첫째자리에서 반올림하여 7점으로 보았다. 7점은 보기에서 ‘매우 중요함’에 해당하므로 ‘대표데이터’ 요소의 최종 중요도는 ‘매우 중요함’으로 책정되었다.

〈표 3〉 연구데이터 기록관리 메타데이터 초안 요소 중요도 자문 결과

요소			중요도 자문 결과		
상위요소	하위요소	세부요소	총점	평균점수	최종 중요도
기록물명(표제)	데이터세트명	대표데이터	212	7	매우 중요함
데이터세트 식별자	데이터식별자		210	7	매우 중요함
법규정보	법령/규정		164	5	약간 중요함
국가연구개발사업정보	사업번호		196	6	중요함
	사업명		197	6	중요함

상위요소	요소		중요도 자문 결과			
	하위요소	세부요소	총점	평균점수	최종 중요도	
연구개발과제정보	과제번호		188	6	중요함	
	과제명		198	6	중요함	
	과제형태	(공동/단독)	156	5	약간 중요함	
	과제진행상태	(진행중/종료)	153	5	약간 중요함	
	과제기간		171	5	약간 중요함	
	과제설명		177	6	중요함	
	지원기관	기관명		180	6	중요함
		지원기관 요구사항		162	5	약간 중요함
	산출물	연구성과물	198	6	중요함	
데이터세트 정보	분야		186	6	중요함	
	주제	주제어	200	6	중요함	
	크기	용량		170	5	약간 중요함
		단위		169	5	약간 중요함
	일시	생산일시	189	6	중요함	
데이터 정보	데이터명		206	6	중요함	
	데이터설명		195	6	중요함	
	데이터크기		175	5	약간 중요함	
	수집 목적		170	5	약간 중요함	
	수집 주체	책임자		175	5	약간 중요함
		기여자		160	5	약간 중요함
	수집 유형	(실험/관측/기타)	165	5	약간 중요함	
	수집 정보	수집 일시		183	6	중요함
		수집 주기		172	5	약간 중요함
		수집 범위		176	5	약간 중요함
		수집 방법		179	6	중요함
	분석 목적		171	5	약간 중요함	
	분석 주체	책임자		163	5	약간 중요함
		기여자		151	5	약간 중요함
	분석 정보	분석 일시		173	5	약간 중요함
		분석 주기		165	5	약간 중요함
		분석 범위		168	5	약간 중요함
		분석 방법		174	5	약간 중요함
	포맷	포맷명		184	6	중요함
		포맷버전		177	6	중요함
		구동SW 및 버전		181	6	중요함
	언어		170	5	약간 중요함	
	메타데이터 표준		189	6	중요함	
	암호화데이터 사용유무		173	5	약간 중요함	
	연계시스템 유무		165	5	약간 중요함	
	구성	첨부파일	첨부파일명	179	6	중요함
첨부파일 종류			180	6	중요함	
첨부파일 포맷: 구동 SW			176	6	중요함	
첨부파일 저장방법			160	5	약간 중요함	

요소		중요도 자문 결과			
상위요소	하위요소	세부요소	총점	평균점수	최종 중요도
생산이력	생산시스템	시스템명	187	6	중요함
		시스템개요	167	5	약간 중요함
		구축년도	168	5	약간 중요함
		DBMS정보	174	5	약간 중요함
	생산자	기관명	179	6	중요함
		기관코드	173	5	약간 중요함
		부서명	153	5	약간 중요함
		부서코드	159	5	약간 중요함
		담당자명	164	5	약간 중요함
		담당자직책	132	4	보통
생산유형		161	5	약간 중요함	
이력	행위유형	(관리/이용/보존)	174	5	약간 중요함
	행위자	행위자명	164	5	약간 중요함
		행위자부서	144	5	약간 중요함
		행위자직책	133	4	보통
	행위일시		169	5	약간 중요함
	행위근거		163	5	약간 중요함
	행위설명		161	5	약간 중요함
	행위결과		167	5	약간 중요함
	변경요소	<u>변경요소명</u>	177	6	중요함
		<u>변경이전값</u>	169	5	약간 중요함
처분	보존기간		200	6	중요함
	보존기간 사유		194	6	중요함
	처분일시		193	6	중요함
	처분방법		183	6	중요함
	적용범위		176	6	중요함
권한	비밀	비밀분류	200	6	중요함
		비밀분류 근거	197	6	중요함
		보호기간	201	6	중요함
	접근	접근범위	195	6	중요함
		URL	183	6	중요함
	공개	공개구분	199	6	중요함
		비공개사유	193	6	중요함
		공개제한부분	191	6	중요함
		공개예정일자	188	6	중요함
		공개관련근거	184	6	중요함
		목록공개구분	180	6	중요함
		목록공개제한사유	182	6	중요함
		목록공개제한부분	178	6	중요함
	라이선스	저자	196	6	중요함
		발행일	187	6	중요함
		타이틀	188	6	중요함
		버전	187	6	중요함
		출판사	164	5	약간 중요함
식별자		188	6	중요함	

요소			중요도 자문 결과		
상위요소	하위요소	세부요소	총점	평균점수	최종 중요도
관계	관계유형		165	5	약간 중요함
	관계대상 식별자		175	5	약간 중요함
	관계설명		164	5	약간 중요함
무결성 체크	무결성검증법		187	6	중요함
	무결성검증값		185	6	중요함

초안 요소의 중요도를 7점 척도로 분석한 결과를 정리하면 다음 <표 4>와 같다.

본 연구에서 도출한 연구데이터 기록관리 메타데이터 초안의 98개 요소 중 '전혀 중요하지 않음', '중요하지 않음', '약간 중요하지 않음'으로 평가할 수 있는 요소는 없었으며, '보통'과 '매우 중요함'에 해당하는 요소는 각각 2개, '약간 중요함'으로 평가받은 요소는 44개, '중요함'으로 평가받은 요소는 50개로 나타났다. 98개의 요소 중 약 96%가 '약간 중요함' 또는 '중요함'에 해당하였고 '보통'으로 평가받은 요소는 선택요소로, '매우 중요함'에 해당하는 요소는 필수요소로 구분되어 있었기 때문에 본 연구에서 도출한 초안은 그 적절성을 검증 받았다고 할 수 있다.

<표 5>는 조사지의 마지막에 위치한 주관식 문항에 대한 전문가 답변으로, 연구데이터 기록

관리 메타데이터 초안에 대한 기타 의견을 내용 별로 구분하여 정리한 것이다.

전문가 기타 의견은 크게 6가지로 분류할 수 있었다. 첫 번째는 자동 입력이 가능한 항목과 수동 입력이 필요한 메타데이터 항목의 구분에 대한 내용으로, 데이터 수요자의 입장에서 필요한 항목은 무엇이 있는지에 대한 논의와 함께 진행되어야 한다는 의견이 있었다. 두 번째는 항목별 우선순위를 선별하여 이를 명시할 필요가 있음을 제시하였다. 세 번째는 계속성 연구사업에서 생성되는 연구데이터의 경우, 생산 시점 및 최종 버전에 대한 합의가 필요하다는 의견이었다. 네 번째 의견은 여러 기관이 공동으로 연구하는 과제의 연구데이터는 관리 기준에 대한 표준을 정립할 필요성이 있음을 주장하였으며 다섯 번째 의견 연구데이터의 재사용을 위해서는 파일에 대한 설명 뿐만 아니라

<표 4> 연구데이터 기록관리 메타데이터 초안 요소 중요도 자문 결과

구분	중요도	총 개수	비고
1	전혀 중요하지 않음	0	해당 요소 없음
2	중요하지 않음	0	해당 요소 없음
3	약간 중요하지 않음	0	해당 요소 없음
4	보통	2	선택요소
5	약간 중요함	44	
6	중요함	50	
7	매우 중요함	2	필수요소
총계		98	

〈표 5〉 연구데이터 기록관리 메타데이터 요소 초안에 대한 전문가 설문조사 기타 의견 정리 내용

구분	주요 내용
자동 수집/입력 가능한 메타데이터에 대한 논의	<ul style="list-style-type: none"> • 메타데이터 수집에 있어 기계 등을 통해 자동수집하지 못하면 정확성 및 메타데이터 검증 시 검수자의 전문성 여부로 인해 어려움을 겪을 때가 많으므로 데이터 수요자 입장에서 필요 데이터가 무엇인지, 어떤 항목을 메타데이터화 해야 할지에 대한 논의가 필요함 • 항목 수가 많으므로 자동입력이 가능한 항목과 직접 입력이 필요한 항목이 구분되어야 함
우선순위 선별	<ul style="list-style-type: none"> • 항목별 필수사항 선별이 필요함
계속성 연구사업에서 생성되는 연구데이터 고려	<ul style="list-style-type: none"> • 연구사업의 종료기간이 없는 경우, 연구데이터가 매시/매일 변경, 생산되는 경우가 있으므로 이 경우에는 최종 버전을 어느 시점으로 정할지, 메타데이터에는 어떻게 표기할지 등에 대한 논의가 필요함. 예를 들어 최종 버전 기준을 연도로 설정할 시, 다음 년도에는 전년도에 보존된 데이터는 제외해야 할지, 수정된 데이터는 어떻게 정리할 지 등에 대한 논의가 필요함
다수의 기관이 공동연구하는 데이터의 관리 기준 표준화	<ul style="list-style-type: none"> • 여러 기관이 공동연구를 진행하는 경우가 있으므로 연구기관들간 해당 데이터에 대한 관리 기준 표준화가 필요함
연구데이터 파일 및 파일 구성에 대한 상세 설명	<ul style="list-style-type: none"> • 연구데이터 재사용을 위해서는 파일 및 파일의 구성에 대한 상세한 설명이 필요함
특정 요소의 중요도 강조	<ul style="list-style-type: none"> • 과제 및 사업과 관련된 내용의 중요도를 높게 책정하였으며 연구데이터는 다른 기록과 달리 라이선스 부분이 중요하므로 이와 관련한 요소의 중요도를 높게 책정함
백업 항목 추가	<ul style="list-style-type: none"> • 장기보존을 위해 백업 항목이 추가될 필요가 있음

파일의 구성에 대한 상세한 설명이 모두 포함 되어야 한다고 지적하였다. 여섯 번째 의견은 연구데이터 메타데이터에 있어 과제 및 사업, 라이선스 항목이 중요함을 강조하였고 마지막으로 장기보존을 고려하여 백업 항목의 신설을 제안하는 의견이 있었다.

4.2 메타데이터 요소 수정

다음의 〈표 6〉은 전문가 설문조사 결과를 바탕으로 연구데이터 기록관리 메타데이터에 대하여 수정한 내용을 나타낸 것이다. 연구데이터 기록관리 메타데이터 초안에서 변경된 사항은 5가지로, 전문가 자문 결과 및 기타 의견의 내용을 수렴하여 수정하였다.

이상의 수정 사항을 모두 반영한 최종 연구

데이터 기록관리 메타데이터 요소는 다음 〈표 7〉과 같다.

5. 결론

21세기 지식정보사회의 연구 기술 발전으로 인해 연구데이터에 대한 중요성이 주목받고 있다. 연구데이터는 연구결과의 재현 및 검증에 요구되는 요소로서 효율적 관리를 통한 이용가치의 극대화가 필요하며, 특히 공적자금을 투입한 국가연구개발사업의 연구데이터는 행정정보 데이터셋으로서 공공기록물 관리의 대상에 해당하나 연구데이터의 특성을 반영한 기록관리 체계는 미비한 상황이다. 특히, 현행 행정정보 데이터셋 기록관리기준표는 기존의 전

〈표 6〉 연구데이터 기록관리 메타데이터 수정 내용

수정사항				근거
상위요소	하위요소	세부요소	수정 내용	
데이터정보	데이터설명		• 데이터에 대한 설명 뿐 아니라 데이터 파일의 구성 요소를 함께 기술해야 함	• 전문가 기타 의견에서 연구데이터 파일 및 파일 구성에 대한 상세 설명이 필요하다는 의견이 제시됨
데이터정보	포맷	포맷버전	• 해당 시 필수 요소로 변경	• 세부요소 '포맷버전'의 중요도: '중요함'
권한	공개	공개제한부분	• 해당 시 필수 요소로 변경	• 세부요소 '공개제한부분', '공개관련근거', '목록 공개구분', '목록공개제한사유', '목록공개제한부분'의 중요도: '중요함'
		공개관련근거		
		목록공개구분		
		목록공개제한사유		
목록공개제한사유	목록공개제한부분			
라이선스	버전		• 필수 요소로 변경	• 하위요소 '버전'의 중요도: '중요함' • 전문가 기타 의견에서 '라이선스' 요소에 대한 중요성을 강조함
데이터정보	백업	백업 유무	• 신설	• 전문가 기타 의견에서 장기보존을 위한 백업 요소 신설 의견이 제기됨
		백업 위치	• 해당 시 필수 요소	
<ul style="list-style-type: none"> 중요도가 '약간 중요함'으로 평가된 요소는 '권고' 우선순위를 부여함. 단, '데이터정보' 상위요소와 '이력' 하위요소에 한해 메타데이터 요소 수정 전 선택요소가 아니었을 경우 '필수' 또는 '해당 시 필수' 우선순위를 부여함 중요도가 '중요함'으로 평가된 요소는 '필수' 또는 '해당 시 필수' 우선순위를 부여함 				• 전문가 기타 의견에서 항목별 우선순위 선별이 필요하다는 의견이 제기됨

〈표 7〉 연구데이터 기록관리 메타데이터 요소 최종 도출(안)

요소			요소 설명			
상위요소	하위요소	세부요소	우선순위	반복가능 여부	중요도	비고
기록물명(표제)	데이터세트명	대표데이터	필수		매우 중요함	
데이터세트 식별자	데이터식별자		필수		매우 중요함	
법규정보	법령/규정		권고		약간 중요함	
국가연구개발사업정보	사업번호		필수		중요함	전문가 의견에서 중요성 강조
	사업명		필수		중요함	
연구개발과제정보	과제번호		필수		중요함	전문가 의견에서 중요성 강조
	과제명		필수		중요함	
	과제형태	(공동/단독)	권고		약간 중요함	
	과제진행상태	(진행중/종료)	권고		약간 중요함	
	과제기간		권고		약간 중요함	
	과제설명		필수		중요함	
	지원기관	기관명	필수	○	중요함	
	지원기관 요구사항	권고		약간 중요함		
데이터세트 정보	산출물	연구성과물	필수	○	중요함	
	분야		필수	○	중요함	
	주제	주제어	필수	○	중요함	
	크기	용량	권고		약간 중요함	
단위		권고		약간 중요함		

요소			요소 설명			
상위요소	하위요소	세부요소	우선순위	반복가능 여부	중요도	비고
	일시	생산일시	필수		중요함	
데이터 정보	데이터명		필수		중요함	
	데이터설명		필수		중요함	
	데이터크기		권고		약간 중요함	데이터 파일 및 파일 구성 상세 설명
	수집 목적		해당 시 필수	○	약간 중요함	
	수집 주체	책임자	해당 시 필수	○	약간 중요함	
		기여자	권고	○	약간 중요함	
	수집 유형	(실험/관측/기타)		○	약간 중요함	
	수집 정보	수집 일시	해당 시 필수	○	중요함	
		수집 주기	해당 시 필수	○	약간 중요함	
		수집 범위	해당 시 필수	○	약간 중요함	
		수집 방법	해당 시 필수	○	중요함	하나의 기술 첨부과 일로 대체될 수 있음
	분석 목적		해당 시 필수	○	약간 중요함	
	분석 주체	책임자	해당 시 필수	○	약간 중요함	
		기여자	권고	○	약간 중요함	
	분석 정보	분석 일시	해당 시 필수	○	약간 중요함	
		분석 주기	해당 시 필수	○	약간 중요함	
		분석 범위	해당 시 필수	○	약간 중요함	
		분석 방법	해당 시 필수	○	약간 중요함	
	포맷	포맷명	해당 시 필수	○	중요함	
		포맷버전	해당 시 필수	○	중요함	
구동SW 및 버전		해당 시 필수	○	중요함		
백업	백업 유무	해당 시 필수			전문가 의견에 따른 요소 신설	
	백업 위치	해당 시 필수				
언어		해당 시 필수	○	약간 중요함		
메타데이터 표준		해당 시 필수	○	중요함		
암호화데이터 사용유무		권고		약간 중요함		
연계시스템 유무		권고		약간 중요함		
구성	첨부파일	첨부파일명	해당 시 필수	○	중요함	
		첨부파일 종류	해당 시 필수	○	중요함	
		첨부파일 포맷 : 구동 SW	해당 시 필수	○	중요함	
		첨부파일 저장방법	권고	○	약간 중요함	
생산이력	생산시스템	시스템명	필수		중요함	
		시스템개요	권고		약간 중요함	
		구축년도	권고		약간 중요함	
		DBMS정보	권고		약간 중요함	
	생산자	기관명	필수		중요함	
		기관코드	권고		약간 중요함	
		부서명	권고		약간 중요함	
		부서코드	권고		약간 중요함	

요소			요소 설명			
상위요소	하위요소	세부요소	우선순위	반복가능 여부	중요도	비고
이력		담당자명	권고		약간 중요함	
		담당자직책	선택		보통	
	생산유형		권고		약간 중요함	
	행위유형	(관리/이용/보존)	필수		약간 중요함	
	행위자	행위자명	필수		약간 중요함	
		행위자부서	필수		약간 중요함	
		행위자직책	선택		보통	
	행위일시		필수		약간 중요함	
	행위근거		필수		약간 중요함	
	행위설명		권고		약간 중요함	
행위결과		권고		약간 중요함		
변경요소	변경요소명	해당 시 필수		○	중요함	
	변경이전값	해당 시 필수		○	약간 중요함	
처분	보존기간		필수		중요함	
	보존기간 사유		필수		중요함	
	처분일시		필수		중요함	
	처분방법		필수		중요함	
	적용범위		필수		중요함	
권한	비밀	비밀분류	필수		중요함	
		비밀분류 근거	필수		중요함	
		보호기간	필수		중요함	
	접근	접근범위	필수		중요함	
		URL	해당 시 필수		중요함	
	공개	공개구분	필수		중요함	
		비공개사유	필수		중요함	
		공개제한부분	해당 시 필수		중요함	
		공개예정일자	해당 시 필수		중요함	
		공개관련근거	해당 시 필수		중요함	
		목록공개구분	해당 시 필수		중요함	
		목록공개제한사유	해당 시 필수		중요함	
		목록공개제한부분	해당 시 필수		중요함	
	라이선스	저자	필수		중요함	전문가 의견에서 중요성 강조
		발행일	필수		중요함	
타이틀		필수		중요함		
버전		필수		중요함		
출판사		필수		약간 중요함		
식별자		필수		중요함		
관계	관계유형		권고		약간 중요함	
	관계대상 식별자		권고		약간 중요함	
	관계설명		권고		약간 중요함	
무결성 체크	무결성검증법		필수		중요함	
	무결성검증값		필수		중요함	

자문서 기록관리체계와는 달리 처분지침으로서의 기록관리 요소와 메타데이터의 역할을 함께 수행하고 있어 이를 분리하는 작업이 필요하다. 본 연구는 위와 같은 문제인식에 기초하여 연구데이터 기록관리 메타데이터 요소를 제안하고자 하였다.

기존 행정정보 데이터세트 기록관리기준표의 제한사항을 극복하기 위하여 문헌조사를 통해 연구데이터의 특성과 연구과정 프로세스를 분석함으로써 연구데이터 기록관리 메타데이터 요소를 도출하기 위한 원칙을 수립하였고, NAK 8, NAK 35, ISAD(G), PREMIS 데이터모델, DCAT V2, DataCite 표준의 요소를 반영하여 연구데이터 기록관리 메타데이터 요소를 제안하였다. 도출된 메타데이터 요소 초

안은 전문가 설문조사를 통해 중요도 검증을 실시하였고, 자문 결과를 바탕으로 수정을 거쳤다. 그러나 본 연구의 결과물은 과학기술 분야의 연구데이터에 더 적합하여 인문사회 분야의 연구데이터에는 적용하기 어렵다는 제한점이 있으며, 과학기술 분야 내에서도 세부분야별 특성을 고려하지 않은 일반적인 내용만을 포함하고 있다는 한계가 있다.

이런 제한사항을 극복하기 위해서는 과학기술표준분류상의 다양한 분야의 관련 연구자를 통해 분야별 연구데이터 특성을 반영할 수 있도록 후속연구가 이루어져야한다. 또한, 연구데이터의 장기보존을 위한 보존포맷에 대한 논의가 함께 진행되어야 할 것이다.

참 고 문 헌

- 과학기술정보통신부 (2018). 연구데이터 공유·활용 전략.
- 구찬미, 김순희 (2017). 기록관리 관점에서 본 연구기록물의 가치와 특성 분석. 한국기록관리학회지, 17(3), 49-70. <http://doi.org/10.14404/JKSARM.2017.17.3.049>
- 구찬미, 김순희 (2019). 과학기술분야 연구기관의 DMP를 적용한 연구기록물 관리. 한국기록관리학회지, 19(1), 1-21. <http://doi.org/10.14404/JKSARM.2019.19.1.001>
- 국가과학기술연구회 (2019). 출연(연) 연구데이터 관리·활용 방안 연구. 서울: 동연구회, 2018-09.
- 국가기록원 (2020). 행정정보 데이터세트 기록관리 실행 매뉴얼.
- 국가연구개발혁신법. 법률 제20057호.
- 기록관리 메타데이터 표준(v2.3). NAK 8:2022(v2.3)
- 김주섭, 양성준, 김선태 (2022). 데이터 리포지터리 인증 체계 분석 및 인증 전략에 관한 연구: Coretrustseal을 중심으로. 한국문헌정보학회지, 56(2), 209-229. <http://doi.org/10.4275/KSLIS.2022.56.2.209>
- 김지혜 (2023). 행정정보 데이터세트 기록관리를 위한 메타데이터 요소 연구. 석사학위논문, 전북대학

교 일반대학원 기록관리학과.

- 박동진 (2006). 과학기술분야의 연구성과물정보 표준 메타데이터 개발에 관한 연구. 한국산업경영시스템학회지, 29(4), 83-90.
- 서지인, 노지현 (2022). 디지털화 기록 관리를 위한 메타데이터 요소(안) 설계. 한국기록관리학회지, 22(4), 1-24. <http://doi.org/10.14404/JKSARM.2022.22.4.001>
- 안채영, 김지현 (2020). 정부출연연구기관의 연구기록 평가 개선방안 연구. 기록학연구, 66, 105-155. <https://doi.org/10.20923/kjas.2020.66.105>
- 이미영 (2014). 과학기술분야 연구기록의 평가에 관한 연구. 기록학연구, 41, 75-111. <https://doi.org/10.20923/kjas.2014.41.075>
- 이미영 (2015). 공적(公的) 연구기관에서의 연구기록 평가기준 연구. 기록학연구, 46, 287-323. <https://doi.org/10.20923/kjas.2015.46.287>
- 이미화, 이은주, 노지현 (2020). 연구데이터 관리를 위한 OAK 메타데이터 확장 방안 연구. 한국도서관·정보학회지, 51(3), 27-51. <http://doi.org/10.16981/kliss.51.3.202009.27>
- 이상훈, 박은지, 김주섭, 김선태 (2022). 식품성분 연구데이터의 표준화를 위한 메타데이터 설계에 관한 연구. 한국도서관·정보학회지, 53(3), 241-262. <http://doi.org/10.16981/kliss.53.3.202209.241>
- 장보성, 남영준 (2011). 디지털 연구성과물의 장기보존을 위한 메타데이터에 관한 연구. 한국도서관·정보학회지, 42(4), 281-309. <http://doi.org/10.16981/kliss.42.4.201112.281>
- 전한역, 김지혜, 김현태, 양동민 (2023). 시청각 유형 보존포맷 선정을 위한 필수보존속성 연구: 디지털 오디오를 중심으로. 디지털문화아카이브지, 6(2), 27-53. <http://doi.org/10.23089/jdca.2023.6.2.002>
- 한국연구재단 (2021. 07. 09). 연구데이터관리계획(Data Management Plan) 설명회 개최. 출처: https://www.nrf.re.kr/cms/board/general/view?nts_no=124731&menu_no=53&nts_no=&search_type=ALL&search_keyword=%EC%97%B0%EA%B5%AC%EB%8D%B0%E C%9D%B4%ED%84%B0&page=
- 행정정보 데이터세트 기록관리 기준 - 관리기준표 작성 및 이관 규격. NAK 35 2020(v1.0)
- CoreTrustSeal (2022). CoreTrustSeal Trustworthy Data Repositories Requirements. Available: <https://www.coretrustseal.org/why-certification/requirements/>
- DataCite (2021). DataCite Metadata Schema 4.4. Available: <https://schema.datacite.org/meta/kernel-4.4/>
- Greenberg, J. (2010). Metadata for scientific data: historical considerations, current practice, and prospects. Journal of library metadata, 10(2-3), 75-78. <https://doi.org/10.1080/19386389.2010.520262>
- Huang, H. & Qin, J. (2024). Metadata functional requirements for genomic data practice and

curation. *Information Research an International Electronic Journal*, 29(2), 3-29.

<https://doi.org/10.47989/ir292363>

Information and documentation - Records management. ISO 15489-1:2016.

International Council on Archives (2000). ISAD(G): General International Standard Archival Description.

Longwood Medical Area Research Data Management Working Group(LMA RDMWG) (n.d.). Data Retention: Research Records. Available:

<https://datamanagement.hms.harvard.edu/store-evaluate/data-retention>

NC State University (n.d.). Defining Research Data. Available:

<https://www.lib.ncsu.edu/do/data-management/defining-research-data>

The White House (2023). FACT SHEET: Biden-Harris Administration Announces New Actions to Advance Open and Equitable Research. Available:

<https://www.whitehouse.gov/ostp/news-updates/2023/01/11/fact-sheet-biden-harris-administration-announces-new-actions-to-advance-open-and-equitable-research/>

UK Data Service (2019). Research Data Lifecycle. Available:

<https://www.youtube.com/watch?v=-wjFMMQD3UA>

W3C (2019). Data Catalog Vocabulary (DCAT) - Version 2. Available:

<https://www.w3.org/TR/vocab-dcat-2/>

Wilson, A. (2010). How much is enough: metadata for preserving digital data. *Journal of Library Metadata*, 10(2-3), 205-217. <https://doi.org/10.1080/19386389.2010.506395>

• 국문 참고문헌에 대한 영문 표기

(English translation of references written in Korean)

Ahn, Chaeyoung & Kim, Ji Hyun (2020). A study on the improvement of research records appraisal practice of the government-funded research institutions. *The Korean Journal of Archival Studies*, 66, 105-155. <https://doi.org/10.20923/kjas.2020.66.105>

Jang, Bo Seong & Nam, Young Joon (2011). A study of metadata for long-term preservation of digital research outcome. *Journal of Korean Library and Information Science Society*, 42(4), 281-309. <http://doi.org/10.16981/kliss.42.4.201112.281>

Jeon, Hanyeok, Kim, Ji-Hye, Kim, Hyun-Tae, & Yang, Dongmin (2023). A study on significant properties for selection of audiovisual type preservation format: focused on digital audio. *Journal of D-Culture Archives*, 6(2), 27-53. <http://doi.org/10.23089/jdca.2023.6.2.002>

- Kim, Jihye (2023). A Study on Metadata Elements for Records Management of Administrative Information Dataset. Master's thesis, Jeonbuk National University.
- Kim, Ju-Seop, Yang, Seong Jun, & Kim, Sun Tae (2022). A study on data repository certification scheme analysis and certification strategy: focused on Coretrustseal. *Journal of the Korean Society for Library and Information Science*, 56(2), 209-229.
<http://doi.org/10.4275/KSLIS.2022.56.2.209>
- Koo, Chan Mi & Kim, Soonhee (2017). An analysis of the value and characteristics of research records from an archives and records management perspective. *Journal of Korean Society of Archives and Records Management*, 17(3), 49-70.
<http://doi.org/10.14404/JKSARM.2017.17.3.049>
- Koo, Chan Mi & Kim, Soonhee (2019). Research records management in scientific research institutes by applying DMP. *Journal of Korean Society of Archives and Records Management*, 19(1), 1-21. <http://doi.org/10.14404/JKSARM.2019.19.1.001>
- Korea Institute of Science and Technology Information (n.d). DMP Public Templates List. Available: <https://dataon.kisti.re.kr/dmp/open/selectDmpTemplateOpenList.do>
- Lee, Mihwa, Lee, Eun Ju, & Rho, Jee-Hyun (2020). A preliminary study on extending OAK metadata for research data. *Journal of Korean Library and Information Science Society*, 51(3), 27-51. <http://doi.org/10.16981/kliss.51.3.202009.27>
- Lee, Miyoung (2014). A study on the appraisal of research records in science and technology: focusing on foreign cases. *The Korean Journal of Archival Studies*, 41, 75-111.
<https://doi.org/10.20923/kjas.2014.41.075>
- Lee, Miyoung (2015). A study on the appraisal criteria of research records in public research institution. *The Korean Journal of Archival Studies*, 46, 287-323.
<https://doi.org/10.20923/kjas.2015.46.287>
- Lee, Sang Hoon, Park, Eunji, Kim, Ju-Seop, & Kim, Sun Tae (2022). A study on the metadata design for standardization of food composition research data. *Journal of Korean Library and Information Science Society*, 53(3), 241-262.
<http://doi.org/10.16981/kliss.53.3.202209.241>
- Metadata Standard for Records and Archives Management Version 2.3. NAK 8:2022(v2.3)
- Ministry of Science and ICT (2018). Research Data Sharing and Utilization Strategy.
- National Archives of Korea (2020). Dataset Records Management Implementation Manual.
- National Research And Development Innovation Act. Act No. 20057.
- National Research Council of Science & Technology (2019, March). Research Data Management

and Utilization for Government-Funded Research Institutes in Korea.

Park, Dong-Jin (2006). A study on the development of metadata standard for research outcomes information in science & technology. *Journal of Korean Society of Industrial and Systems Engineering*, 29(4), 83-90.

Record Keeping Criteria for Dataset - Composition of Dataset Management Reference Table & Exchange of Dataset - Version 1.0. NAK 35 2020(v1.0)

Seo, Ji In & Rho, Jee-Hyun (2022). A study on designing metadata elements for the management of digitized records. *Journal of Korean Society of Archives and Records Management*, 22(4), 1-24. <http://doi.org/10.14404/JKSARM.2022.22.4.001>

