

연구 분야별 특성에 기반한 연구데이터 기록관리 메타데이터의 모듈형 설계 방안 연구*

Modular Metadata Design for Research Data Record Management: A Discipline-Specific Approach

금 효 진 (Hyojin Geum)**

양 동 민 (Dongmin Yang)***

초 록

본 연구는 연구데이터의 장기보존 및 재 활용을 위한 기록관리 메타데이터 설계의 방향성을 제시하고자, 연구 분야별 특성을 반영한 모듈형 설계 방안을 제안한다. 기존 연구데이터 기록관리 메타데이터는 과학기술 분야를 중심으로 설계되어, 인문사회 분야의 연구 특성을 충분히 반영하지 못하는 한계가 있었다. 이에 과학기술 및 인문사회 분야 연구기관 4곳의 연구보고서와 데이터 관리지침을 분석하고, 공통 요소와 분야별 특수 요소를 비교함으로써 연구 재현성 확보 방식에서의 차별된 특성을 도출하였다. 분석 기준은 기존 메타데이터 표준 및 선행연구를 바탕으로 도출한 6가지 항목(데이터 유형, 연구 조건 기록, 도구 및 소프트웨어 정보, 연구 절차, 지침상 요구 사항, 보존 및 접근성 조치)을 적용하였으며, 이를 통해 모듈형 기록관리 메타데이터 요소 체계를 설계하였다. 제안된 설계 방안은 분야 간 공통 요소를 기반으로 하되, 분야별 특수성을 반영한 요소들을 선택적으로 확장할 수 있는 구조로 구성되어 있어, 연구데이터의 다양성과 재현성 확보 요구에 유연하게 대응할 수 있다. 본 연구는 시범적 사례 분석을 통해 분석 틀의 범용성과 구조적 확장 가능성을 제시함으로써, 향후 다양한 연구 분야와 데이터 유형에 대한 후속 연구의 기반을 제공한다는 점에서 의의를 지닌다.

ABSTRACT

This study proposes a modular design framework for records management metadata aimed at ensuring the long-term preservation and reuse of research data, with particular attention to the distinct characteristics of different research domains. Existing metadata frameworks have primarily been developed with a focus on the science and technology fields, often failing to adequately reflect the unique requirements of research in the humanities and social sciences. To address this limitation, this study analyzed research reports and data management guidelines from four representative institutions across these domains. By comparing common elements with field-specific components, the study identified key differences in approaches to ensuring research reproducibility. The analysis employed six criteria—data type, documentation of research conditions, tools and software, research procedures, mandated requirements, and preservation and accessibility measures—drawn from existing metadata standards and prior studies. Based on this, a modular metadata element system was developed, allowing for selective expansion of field-specific elements while maintaining a common core structure. This flexible design accommodates the diversity of research data and varying reproducibility requirements. Through a pilot case analysis, the study demonstrates the versatility and scalability of the proposed framework, offering a foundation for future research across diverse disciplines and data types.

키워드: 연구데이터, 기록관리 메타데이터, 장기보존, 메타데이터 설계, 재현성

Research Data, Record Management Metadata, Long-term Preservation, Metadata Design, Reproducibility

* 이 논문은 2024년도 전북대학교 연구기반 조성비 지원에 의하여 연구되었음.

이 논문은 2023년 대한민국 교육부와 한국연구재단의 인문사회분야 중견연구자지원사업의 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 : 2025년 5월 24일 논문심사일자 : 2025년 5월 24일 게재확정일자 : 2025년 6월 10일
한국비블리아학회지, 36(2): 235-270, 2025. <http://dx.doi.org/10.14699/kbiblia.2025.36.2.235>

© Copyright © 2025 Korean Biblia 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 및 목적

공적자금을 투입한 국가연구개발사업의 결과물은 전자정부법에 의거하여 행정기관의 직무상 취득된 자료로서 행정정보로 인정된다. 따라서 국가연구개발사업을 수행하며 생성된 연구데이터 역시 국가의 주요한 공공자산으로, 책임 있는 보존과 활용 체계를 필요로 한다. 우리나라는 「국가연구개발혁신법」(이하 ‘혁신법’)을 통해 국가연구개발사업을 관리하고 있으며, 이 과정에서 생성된 연구데이터는 「전자정부법」과 「공공기록물 관리에 관한 법률」에 따라 행정정보 데이터셋으로 분류되어, 국가와 수행기관은 이를 기록관리 해야 할 법적·정책적 의무를 지닌다.

그러나 기존에 존재하는 행정정보 데이터세트 기록관리기준표는 연구데이터의 장기 보존에 있어 가장 중요하게 다루어져야 할 연구 결과 검증 가능성과 연구의 재현성 확보를 고려하지 못하고 있고, 처분지침으로서의 기록관리 요소와 메타데이터의 역할을 함께 겸하고 있어 이를 분리할 필요성이 있다.

금효진과 양동민(2024)의 선행연구(이하 ‘기존 연구’)는 이러한 제도적, 실무적 타당성에 따라 연구데이터의 장기보존을 위한 기록관리 메타데이터를 설계하였다. 연구 목적을 달성하기 위해 문헌 조사를 통해 연구데이터의 특성 및 연구 과정 프로세스를 확인하고, 연구데이터 기록관리 메타데이터 요소 도출 원칙을 수립하였다. 이후 4개의 기록관리 관련 표준(NAK 8, NAK 35, ISAD(G), PREMIS 데이터모델)과

2개의 연구데이터 표준(DCAT V2, DataCite)을 비교·통합하여 기록관리 관점에서의 연구데이터 보존 메타데이터를 직접 설계하고, 설계된 초안은 전문가 설문조사를 통한 중요도 검증을 통해 요소 수정을 거쳐 유효성을 확보하였다.

그러나 이렇게 설계된 연구데이터 기록관리 메타데이터는 단일형 구조로, 과학기술 분야의 일반적 요소를 중심으로 설계되었기 때문에 인문사회 분야의 연구데이터 특수성을 반영하기에는 한계가 있었으며 실사례 적용 및 확장성에도 제한점을 보였다. 이는 인문사회 분야 또한 연구데이터의 생성, 활용, 재현 방식에서 고유의 특성과 독립성을 지니고 있음에도, 이에 대한 체계적인 기록관리 논의가 상대적으로 미흡하였음을 시사한다. 지금까지 연구데이터의 관리 및 기록관리 논의는 과학기술 분야에 집중됐으며, 그 결과 관련 법제도 역시 과학기술 분야를 중심으로 설계되어 있었다.

본 연구는 이러한 제도적·실천적 불균형을 문제로 인식하고, 인문사회 분야를 포함한 전 분야로의 연구데이터 기록관리 체계 확장이 필요함을 제기한다. 최근 혁신법에서 국가연구개발사업의 개념이 과학기술 중심에서 전 분야로 확장된 것은 이러한 문제의식을 제도적으로 뒷받침해주는 변화라 할 수 있다. 과학기술정보통신부(2025)에 따르면 국가연구개발사업은 중앙행정기관이 법령에 근거하여 연구개발을 위해 예산 또는 기금으로 지원하는 사업을 의미한다. 이는 2025년 이전까지 동일한 규정에서 ‘과학기술 분야로 한정하였던 것과 달리, 해당 용어의 삭제를 통해 인문사회 분야의 연구 역시 국가연구개발사업의 범주에 포함되었음을 시사한다.

이상의 배경을 바탕으로 과학기술 분야뿐만 아니라 인문사회 분야의 연구데이터에도 적합한 기록관리 메타데이터를 설계하기 위한 기반을 마련하는 것이 본 연구의 목적이다. 이를 위해선 기존 연구의 기록관리 메타데이터 요소를 비판적으로 검토하고, 연구데이터의 학제 간 차이를 반영한 기록관리적 접근이 필요하다. 분야별 특성을 반영한 기록관리 메타데이터 설계 방법을 제안하기 위하여 연구 질문을 정리한 바는 다음과 같다.

- 연구 질문 1: 과학기술 분야와 인문사회 분야 연구데이터의 특성은 기록관리 메타데이터 설계에 어떤 영향을 미치는가?
- 연구 질문 2: 분석대상 기관의 연구성과물과 내부 관리규정은 분야별 연구데이터 특성을 어떻게 반영하고 있는가?
- 연구 질문 3: 분야별 연구 재현성과 검증 가능성을 확보하기 위한 연구데이터 기록관리 메타데이터 요구 사항은 무엇인가?

1.2 연구의 대상 및 범위

기존 연구에서 연구데이터 기록관리 메타데이터를 설계한 이유는 공적자금이 투입된 연구개발과제에서 생성된 연구데이터의 장기보존을 위함이었다. 본 연구는 기존 연구의 학문적 논의점을 계승하여 공적 자금을 투입한 국가연구개발사업의 진행 과정 중 생산·보존·관리되는 연구데이터를 기록관리 대상으로 설정하였다. 이는 해당 데이터가 행정기관의 직무상 취득된 정보로서 행정정보 데이터세트에 해당하며, 공공기록물로서 장기보존의 필요성이 인

정되기 때문이다.

국가연구개발사업의 추진 체제는 기본적으로 혁신법을 따르며, 이에 따라 대학, 정부출연 연구기관, 지방자치단체출연 연구원 등에서 국가연구개발사업 추진을 위해 연구개발과제를 수행할 수 있다. 또한, 연구데이터는 연구개발과제 수행 과정에서 실시하는 각종 실험·관찰·조사 및 분석 등을 통하여 산출된 사실 자료로, 연구 결과의 검증에 필수적인 데이터를 의미한다(국가연구개발정보처리기준). 따라서 최종 산출된 연구데이터 뿐만 아니라 연구 초기에 획득한 원시 연구데이터 역시 기록관리의 대상이 될 수 있으며, 전자정부법에 따라 직무상 취득하여 부모, 문자, 음성, 음향, 영상, 이미지 등으로 표현된 자료를 모두 포함한다.

1.3 연구 방법 및 절차

연구 목적을 달성하기 위하여 설계한 연구 방법과 절차는 다음의 <그림 1>과 같다.

먼저, 문헌 조사를 통해 분야별 연구데이터 메타데이터의 설계 및 연구데이터를 포함한 연구기록물의 기록관리에 관한 선행연구를 검토한다. 이를 통해 기존 연구에서 제시된 분야별 메타데이터의 설계 과정과 적용 사례를 확인하고, 연구데이터 기록관리에 대한 학문적 논의의 범위와 한계를 고찰한다. 다음으로, 국내 중앙행정기관과 공공연구기관 중에서 과학기술 분야와 인문사회 분야를 대표할 수 있는 기관을 각 두 곳씩 선정하고, 기관별 사례를 분석한다. 중앙행정기관은 정부조직법에 따라 설치된 부·처·청을 말하며, 국가 행정사무를 담당하기 위하여 설치된 행정기관으로서 그 관할권의

수행방법	연구 내용
문헌조사	1. 선행연구 검토: 분야별 메타데이터 설계 관련 선행연구 검토 연구성과물 기록관리 관련 선행연구 검토
분야별 대표 기관의 분석을 통한 실증적 확인	2. 각 분야를 대표하는 분석대상 기관 선정: 과학기술(2), 인문사회(2) 3. 연구데이터 관련 자료 분석을 통한 재현성 확보 방식 차이 파악
기록관리 메타데이터 요구사항 도출	4. 자료 분석 시사점을 통해 연구데이터 기록관리 메타데이터 요구사항 도출
연구데이터 기록관리 메타데이터 설계 제언	5. 분야별 요구사항과 기존 연구의 연구데이터 기록관리 메타데이터 설계 원칙을 고려한 분야별 연구데이터 기록관리 메타데이터 설계안 제언

〈그림 1〉 연구 방법 및 절차

범위가 전국에 미치는 기관이다.

선정된 기관을 대상으로, 이들이 수행한 국가개발연구과제의 최종보고서, 내부 연구데이터 관리지침 및 매뉴얼을 검토하여 각 분야에서 연구데이터의 재현성을 확보하는 방식을 비교한다. 이후 각 기관의 사례 분석을 통해 확보한 시사점을 바탕으로, 연구데이터의 장기보존과 재현성을 보장하기 위한 기록관리 메타데이터의 필수 요소와 요구 사항을 정리한다. 마지막으로, 도출된 요구 사항을 바탕으로 실무 적용 가능성과 유연성을 갖춘 분야별 연구데이터 기록관리 메타데이터 설계 방안을 논의한다. 금효진과 양동민(2024)의 선행연구에서 제시된 연구데이터 기록관리 메타데이터 설계 원칙을 기반으로 하여 기존 연구 결과와의 일관성을 유지함과 동시에, 기관별 사례 분석을 통해 도출한 분야별 연구데이터 관리 요구 사항을 반영할 수 있도록 한다.

2. 이론적 배경

2.1 연구데이터의 개념과 기록관리 메타데이터의 역할

연구데이터는 국가연구개발과제의 수행 중 산출된 다양한 유형의 사실 자료로, 추후 재활용이 가능하도록 연구 결과의 재현 및 검증을 위해 요구되는 데이터이다(금효진, 양동민, 2024). 따라서 국가연구개발사업과의 관계성이 연구데이터의 활용 맥락을 이해하는 데 필수적이지만, 과제 기간 생산 및 활용이 연속적으로 일어나 기존의 기록생애주기 개념을 적용하기에 어려움이 있다(구찬미, 김순희, 2017).

한편, 기록관리 메타데이터는 기록의 내용과 구조, 맥락, 그리고 일정 기간 내의 기록관리 이력을 기술한 데이터이다. 기록관리 메타데이터는 기록의 생산맥락과 업무맥락을 포함하여 관련 행

위 주체를 나타내는 정보와 기록 자체의 내용, 외형, 구조 및 기술적 속성에 관한 정보 등을 제공할 수 있어야 한다(International Organization for Standardization, 2016, ISO 15489-1:2016). 연구데이터는 특히나 디지털 방식으로 수집 및 생산되는 경우가 많아 쉽게 내용, 구조, 맥락이 분리될 수 있으므로 기록관리 메타데이터를 통한 관리가 필수적이다. 연구데이터의 향후 활용을 위해서는 기록 생성 주체, 생성 일시, 버전, 분석 도구 및 환경, 접근 권한, 보존 기간 등의 메타데이터 요소를 통해 연구데이터 생산 정보를 확인할 수 있도록 해야 한다.

DCAT와 같은 연구데이터 메타데이터 표준이 존재하나, 이를 기록관리에 그대로 적용하기에는 다소 어려움이 있다. 이러한 표준들은 데이터 자체 정보를 담을 수는 있으나, 기록관리가 필요로 하는 요구 사항을 충족하지 못하기 때문이다. 따라서 연구데이터의 기록관리를 위해서는 맞춤형 메타데이터를 구축하여 연구데이터의 내용, 구조, 관련 국가연구개발사업 맥락, 연구 조건, 생산자, 생산시스템 정보 등을 포함함으로써 연구 재현율과 재이용 가능성을 높여야 한다.

2.2 연구데이터의 재현성 변동에 따른 기존 설계안의 한계

기존 연구에서 제안한 연구데이터 기록관리 메타데이터 요소는 연구의 재현성 확보에 중점을 두어 기록관리 표준과 연구데이터 표준의 비교·분석을 통해 단일형 표준으로 설계되었다. 이렇게 설계된 메타데이터 요소는 분야별 차이를 반영하는데 다소 미흡하여, 보편적이고

일반적인 요소를 포함하는 것에 그쳤다는 제한점이 있다. 이것이 문제가 되는 이유는, 분야별 특성에 따라 연구의 재현성 개념이 다르기 때문이다.

분야별 특성에 기반하여 연구 재현성 개념이 변동될 수 있다는 것은 국내 및 해외 연구에서 그 근거를 찾을 수 있다. 최형욱과 정은경(2017)은 사회학 분야 연구를 논문 기반 연구와 데이터 기반 연구로 구분하였는데, 데이터 기반 연구는 논문 기반 연구보다 계량적인 연구를 진행하지만 동일한 실험을 반복하여 동일한 결과를 얻는 과학기술 분야에서의 재현성과는 다른 개념을 갖고 있다고 설명하였다. 이는 사회적 맥락, 시기적 특성, 데이터 수집 과정이 모두 연구 결과에 영향을 미치기 때문이다. National Academies of Sciences, Engineering, and Medicine(2019)은 사회과학 분야의 연구는 자연과학 연구보다 실험 조건의 제어가 어렵고, 윤리적 또는 법적 이유로 데이터 공유가 제한될 수 있어 연구의 재현이 더 어려워질 수 있다고 논하였다. Joyeux-Prunel(2024)는 디지털 인문학과 계산 과학(Computational Sciences) 간의 재현성 차이와 그에게 영향을 주는 분야별 요인들을 정리하였다. 디지털 인문학은 정성적 데이터를 포함할 수 있으며, 역사 또는 예술과 관련될 경우 저작권과 희귀성으로 인해 디지털화 과정에서 주관적 해석이 개입될 수 있다. 이에 따라 복수의 해석 가능성이 있어, 완벽한 재현보다는 해석의 과정을 추적할 수 있도록 함으로써 '정당한 해석 가능성'을 확보하는 것이 중요하다. 이에 반해, 계산 과학은 대부분 정량적 데이터를 활용하기 때문에 동일한 입력과 알고리즘으로 일관적인 연구 결과를 도출하

는 것에 중점을 둔다.

이러한 선행연구들은 인문사회 분야가 정성적·비정형 데이터를 중심으로 연구되며, 실험 조건의 반복이나 통제가 어렵고 해석의 다양성이 존재해 연구 결과의 일관적 재현에 한계가 있음을 공통적으로 지적하고 있다. 그러나 다양한 해석의 과정에서 새로운 데이터 또는 맥락이 추가되면 이전보다 더 정교한 해석을 도출할 수 있으며, 이것이 오히려 해당 분야의 학문적 신뢰성과 공신력을 높이는 데 기여할 수 있다(Peels, 2019). 따라서 인문사회 분야에서의 재현성은 동일한 데이터를 통한 결과의 일관성 보다 데이터의 맥락적 해석과 해석 타당성에 초점을 맞추는 것이 적절하다. 반면, 과학기술 분야는 대개 구조화된 정량적 데이터를 활용한다. 또한, 인문사회 분야보다 실험 조건의 제어가 원활하기 때문에 실험의 반복이 가능하고, 주관성보다는 객관성이 두드러진다. 이 때문에 연구 조건과 절차에 대한 정확하고 자세한 정보가 주어진다면 높은 재현성을 확보할 수 있다. 특히, 분석 소프트웨어 버전, 실험의 매개변수, 데이터 출처와 같은 연구 조건을 상

세히 기록한 메타데이터는 자동화 도구와의 연계를 통해 시간이 지나도 타인이 데이터를 이해하고 재사용하는 데 유용하게 작용하므로, 과학기술 분야에서 데이터 공유를 촉진하는 중요 요소이다(Leipzig et al., 2021). 다음의 <표 1>은 과학기술 분야와 인문사회 분야의 연구 재현성 비교를 위하여 앞서 언급한 선행연구의 내용을 기반으로 분야별 연구 특징을 정리한 것이다.

지금까지 논의한 사항을 종합하면, 과학기술 분야와 인문사회 분야는 데이터 생성 방식과 주요 유형, 연구 방법론 등에서 본질적으로 차이를 보이기 때문에 분야에 따라 재현성 개념이 상이하고, 재현성의 확보방법도 달라진다. 이는 분야별로 연구데이터 관리 요구 사항과 연구 재현을 위한 필요 메타데이터 구성에도 변동이 있을 수 있다는 의미이다. 즉, 연구데이터의 분야별 특성은 재현성 확보방법에 영향을 끼치고, 기록요소에 실질적인 차이를 불러올 수 있다. 따라서 연구데이터 기록관리 메타데이터의 설계는 분야별 특수성을 반영한 차별화가 필요하다.

<표 1> 과학기술, 인문사회 분야별 연구 특징 비교

구분	과학기술 분야	인문사회 분야
연구 목적	일반적 법질 발견 및 가설 검증	특정 개념, 사회 현상에 대한 해석적 이해
연구 방법	모델/통계기반 분석이 일반적, 수치적 정확성과 객관성 요구	실험보다는 탐색적 접근이 일반적 상대적으로 맥락 중시, 주관성 높게 개입
데이터 유형	정량적, 구조화된 데이터	정성적, 비정형 데이터
실험 조건 제어	비교적 쉬움	비교적 어려움
연구 반복 가능성	높음	낮음
재현성 개념	동일한 연구 조건의 반복을 통한 일관적 연구 결과 도출	타당한 해석에 대한 설명 가능성, 타당성 해석 과정의 추적 가능성
재현성 확보방법	정량적, 계산적 접근법 실험 프로토콜과 분석 절차의 명확한 기록	정성적, 맥락 위주 접근법 연구 맥락과 해석 과정의 상세한 설명

본 연구는 금효진과 양동민(2024)의 선행연구에서 설계된 연구데이터 기록관리 메타데이터 요소 체계를 기반으로 하되, 분야별 특성에 따라 연구 재현성의 개념이 변화하는 것을 고려하여 분야별 연구데이터의 맞춤형 기록관리 메타데이터 설계를 논의하고자 한다.

2.3 선행연구

본 장에서는 각 분야의 특성에 기반한 연구데이터 기록관리 메타데이터 설계를 위해 관련 선행연구를 분석하여 시사점을 도출하고자 한다. 본 연구의 주제인 '기록관리 관점의 연구데이터 메타데이터 설계'와 관련된 기존 연구는 다음의 두 축으로 나누어볼 수 있다. 한 그룹은 특정 분야의 연구데이터에 특화된 메타데이터 요소를 설계한 연구들이며, 다른 연구는 연구기록물 전반에 대한 관리와 관련된 기록관리 연구이다.

다음의 세 논문은 특정 분야에 특화된 연구데이터 메타데이터 요소를 직접 설계한 선행연구이다.

예상준 외(2019)는 한의학 연구데이터의 관리 및 공유를 위한 메타데이터 요소 설계를 목적으로 연구를 수행하였다. 한의학 연구데이터가 고문헌 분석, 천연물 연구, 임상 연구 등 다양한 형태를 포함한다는 특징에 유의하며, 기존의 메타데이터 표준(TTAK.K0-10.0976, DataCite) 및 국가 연구 데이터 플랫폼 스키마를 분석하여 적합한 메타데이터 요소를 도출하였다.

김주섭 외(2020)는 임산공학 분야에서 생산되는 연구데이터를 관리하기 위한 메타데이터 항목을 도출하는 것을 목표로 기존의 다양한

메타데이터 표준(FGDC-CSDGM, ISO 19115, NEFIS, INSPIRE, ANZLIC, DataCite 4.3, TTA KO-10.0976)을 분석하여 임산공학 분야에 적합한 메타데이터 요소를 선정하기 위한 매핑 작업을 수행했다. 매핑 작업 결과, 필수 6개, 권고 13개, 선택 요소 9개의 메타데이터 요소가 도출하였다.

백재은(2022)은 산업유산자원의 효율적인 이용, 관리, 보존을 위한 메타데이터 요소 설계를 위해 기존의 산업유산 관련 메타데이터를 분석하고, 산업유산자원의 특성을 반영한 메타데이터 요소를 설계하였다. 최종적으로 관리, 기술, 보존, 정보기술, 이용의 5가지 유형으로 분류된 25개의 상위 요소와 86개의 하위 요소가 도출되었다.

이들의 연구는 기존 메타데이터 표준의 비교·분석을 통해 각 세부분야의 연구데이터 특성을 고려한 메타데이터 요소를 설계했다는 점에서 공통점을 가진다. 특정 분야 맞춤형 연구데이터 메타데이터 관리를 위한 구체적 수단을 제시하였다는 점에서 의의가 있으나 기록관리 관점을 반영하지 않았기 때문에 해당 연구에서 제시한 메타데이터는 연구데이터의 기록관리 메타데이터로써 적용하기에 다소 제한점이 있다.

이후로 제시하는 선행연구는 연구데이터를 포함한 연구기록물의 관리를 다룬다.

김상준(2008)은 연구기록물이 연구성과의 공개 및 활용, 연구 결과의 재현성 확보, 연구 윤리 준수 측면에서 중요하다고 강조하였다. 그 중에서도 연구노트가 연구 결과의 공식적·비공식적 검증과정에서 독창성과 객관성을 인정받을 수 있는 주요 수단임에 주목하여, 연구노트의 활성화를 위해 디지털화 추진, 생산·수

집·정리·분류·평가·보존·이용 등에 대한 구체적 규정 마련 등의 방법을 주장하였다.

김수진과 정은경(2012)은 정부출연연구기관의 연구기록물 관리 개선을 위해 영국 정보시스템 합동위원회(JISC), 멜버른 대학교 연구기록물 관리지침, 캐나다 연구도서관협회(CARL) 보고서를 분석하고, 총 49개 정부출연연구기관 사서를 대상으로 설문조사를 실시하였다. 그 결과 연구기록물 제출 의무화, 연구기록물 관리규정 정비, 연구기록물 관리시스템 개발 등의 개선 방안을 도출했다. 특히, 시스템은 국가 기록원 기록관리시스템 혹은 전자문서 시스템을 그대로 적용하는 것이 아닌, 연구기록물에 대한 수집, 분류, 보존, 활용이 가능하도록 보장되어야 함을 강조하였다.

김로사와 장우권(2016)은 과학기술 분야 정부출연연구기관에서 생산하는 연구관리기록물의 관리 필요성을 제기하고, 기록물 관련 규정을 수집하였다. 또한, 연구관리 담당자를 대상으로 설문조사와 인터뷰를 시행하여 연구관리 및 기록물 관리 실태를 분석하였다. 연구 결과, 기록물관리 전문요원 및 기록관리부서 배치율 저조, 기록물관리규정 미비, 연구관리기록물 관리의 미흡함 등의 문제점이 도출되었으며, 이를 대응하는 방안으로 공공기록물을 체계적으로 보존·관리할 수 있는 관리 매뉴얼(안)을 제시하였다.

김혜미와 이영학(2019)은 연구관리 전문기관에서 수집하는 핵심 연구기록물의 현황을 파악하고 문제점을 분석하여 효율적인 관리 개선 방안을 제시하고자 했다. 이를 위해 19개 연구관리전문기관의 일반기록물 관리 현황 및 연구기록물 수집 현황을 분석하였다. 그 결과, 기관 내 연구기록물 관리규정 및 프로세스의 부재,

기관마다 핵심 연구기록물 유형이 상이함, 통일되지 않은 연구기록물 표준 서식 등이 문제점으로 제기되었다. 해당 연구는 문제점 해결을 위해서 법률의 별지 서식 추가, 표준화된 R&D 관리규정 제정 등을 통해 연구기록물의 관리, 활용, 보존에 있어서 체계적인 프로세스를 모색해야 함을 역설하였다.

이상의 네 가지 선행연구는 모두 연구기록물에 대한 중요성을 인식하고, 연구기록물의 관리를 위해 현황을 분석하고 개선 방안을 도출했다는 점에서 의미가 있다. 그러나 제안한 개선 방안이 정책적 시사점에 그쳐 연구기록물 관리 활성화를 위한 구체적인 수단을 실제로 도출해내지 못하였고, 기록관리 분야 관점에서 연구기록물을 논하면서도 그 일부인 연구데이터에 관한 연구는 다소 미흡하였다. 즉, 선행연구들은 대체로 연구성과물의 기록관리 원칙과 유의점에 중점을 두어, 연구데이터의 장기보존을 위해 기록관리 관점을 반영한 메타데이터를 직접 설계하는 것은 현재 학계에서 거의 다루어지지 않았음을 확인할 수 있었다.

기존에 진행한 금효진과 양동민(2024)의 연구는 연구데이터의 장기보존을 위해 기록관리 관점을 도입하고, 실제 메타데이터 요소를 직접 도출함으로써 선행연구에서 다소 간과된 실무적·구조적 접근을 시도하였다. 본 연구는 이와 같은 학문적 시사점을 계승하되 분야별로 상이한 연구 재현성 개념을 연구데이터 기록관리 메타데이터 설계에 반영함으로써 기존 연구의 한계를 보완하고자 하였다는 점에서 의의가 있으며, 분야 간 차이를 구조화하여 메타데이터 설계의 개선방안을 제시한다는 점에서 기존 연구와 차별화된다.

3. 분야별 연구기관 사례 분석

3.1 분석대상 기관 선정 기준 및 분석 방법

본 장에서는 기관별 사례 분석을 통해 국가연구개발사업 연구데이터의 분야별 연구데이터 특성 및 재현성 확보 요구 사항을 확인하고 그 차이를 비교한다. 분야별 연구데이터 특성을 비교하기 위한 기관별 사례 분석에 앞서, 다음 기준에 따라 과학기술 분야와 인문사회 분야를 대표할 수 있는 중앙행정기관 또는 공공연구기관을 선정하였다.

첫째, 각 기관은 해당 분야의 대표성을 가지며, 주요 연구과제가 분야 고유의 데이터 특성을 반영하고 있는 기관으로 한정하였다. 각 기관의 웹사이트를 통해 기관의 설립 목적, 최근 수행한 연구과제의 주제, 소속 부처, 연구보고서 유형 등을 종합적으로 검토하여 각 기관이 과학기술 분야 또는 인문사회 분야에 대표성을 가진다고 판단하였다.

둘째, 연구성과물(보고서, 데이터 세트 등)에 대한 접근이 가능하며, 실질적 분석이 가능한 기관을 우선하였다. 연구성과물은 연구자가 실제 어떤 데이터를 생성했는지, 어떤 방법으로 연구를 진행하였는지를 가장 직접적으로 확인할 수 있는 주요 현장 결과물이다. 따라서 연구성과물에 대한 직접 확인이 가능한 기관을 선정하였다.

셋째, 연구과제 관리규정 또는 연구데이터 관리지침 등을 통해 데이터 관리의 제도적 기반을 갖춘 기관을 포함하였다. 본 장에서 분석하고자 하는 대상 기관은 국가연구개발사업을 진행하여 혁신법의 영향을 받지만, 혁신법의 존재에도

불구하고 자체적으로 연구과제 관리규정을 두는 기관이 존재한다. 혁신법은 국가연구개발사업의 운영 및 관리를 위한 통합적이고 포괄적인 프레임워크를 제시하는 상위법으로서 작용하지만, 기관 고유의 업무 및 역할까지는 고려하지 못하는 한계를 지니고 있기 때문이다. 이에 따라 각 기관은 자체 규정을 통해 조직에 적합한 관리 방식을 구체화하고 있다고 판단할 수 있으며, 이는 기관 업무 및 분야별 특성에 따라 요구되는 연구데이터 관리사항이 상이함을 뒷받침한다.

이상의 세 가지 기준에 따라 선정한 분야별 분석대상 기관은 다음의 <표 2>와 같다.

분야별로 2개 기관을 선정하였으며, 과학기술 분야를 대표하는 분석대상 기관은 질병관리청과 한국지질자원연구원, 인문사회 분야를 대표하는 분석대상 기관은 한국교육개발원과 한국보건사회연구원이다.

질병관리청은 정부조직법 제39조에 근거한 보건복지부 산하 기관으로, 국가 감염병 연구 및 관리와 생명과학연구를 수행하는 중추 기관이다. 과학적 근거에 기반하여 공중 보건 정책을 수립하고, 감염병·만성질환·건강위해 요인에 관한 연구 및 예방관리 정책을 수행한다.

한국지질자원연구원은 과학기술분야 정부출연연구기관 등의 설립·운영 및 육성에 관한 법률 제8조에 근거하여 설립된 정부출연연구기관이다. 지질과학 연구를 통해 지질자원 기반 정보를 구축 및 제공하고, 지반·지하 공간의 효율적 이용, 지하 에너지자원확보, 지진·지질재해 및 지구환경변화 대응을 위한 연구개발 사업을 진행하고 있다.

〈표 2〉 분야별 분석대상 기관 리스트

분야	기관명	분야 대표성	연구성과물 접근성	데이터 관리 규정
과학기술	질병관리청	<ul style="list-style-type: none"> 정부조직법 제39조(보건복지부) 근거 질병관리청과 그 소속기관 직제 보건복지부 산하의 기관 	<ul style="list-style-type: none"> 웹페이지를 통해 정책 연구 결과 보고서와 연구성과(국립의과학지식센터 연계) 공개 	<ul style="list-style-type: none"> 질병관리청 연구노트 관리 지침(행정규칙) 시행 중
	한국지질자원연구원	<ul style="list-style-type: none"> 과학기술분야 정부출연연구기관 등의 설립·운영 및 육성에 관한 법률 근거 과학기술정보통신부 산하의 기관 	<ul style="list-style-type: none"> 지오빅데이터 오픈플랫폼(기관 리포지터리)을 통해 연구데이터 제공 	<ul style="list-style-type: none"> 연구데이터의 수집 지침, 관리지침, 보존 지침, 윤리·저작권·라이선스 지침을 각각 명시
인문사회	한국교육개발원	<ul style="list-style-type: none"> 정부출연연구기관등의 설립·운영 및 육성에 관한 법률 근거 국무조정실 산하의 기관 	<ul style="list-style-type: none"> 웹페이지를 통해 연구보고서 및 통계조사자료 공개 	<ul style="list-style-type: none"> 한국교육개발원 조사자료 관리 및 공개 지침
	한국보건사회연구원	<ul style="list-style-type: none"> 정부출연연구기관등의 설립·운영 및 육성에 관한 법률 근거 국무조정실 산하의 기관 	<ul style="list-style-type: none"> 웹페이지를 통해 각종 보고서 및 정기간행물 공개 	<ul style="list-style-type: none"> 한국보건사회연구원 조사자료 관리지침

한국교육개발원은 주무 기관으로 국무조정실을 두고 있으며, 정부출연연구기관 등의 설립·운영 및 육성에 관한 법률에 따라 출범하였다. 한국교육의 현안에 관한 종합적·과학적인 연구 수행으로 당면과제를 합리적으로 해결할 수 있는 혁신적인 미래 교육체제를 개발하는 것이 설립 목적이다. 유·초·중등교육, 고등교육, 평생교육 정책 전반에 관련된 연구·개발 및 사무가 기관의 주요 기능이다.

한국보건사회연구원은 한국교육개발원과 마찬가지로 국무조정실 산하의 기관이며, 근거 법률도 같다. 대한민국 국민의 인구·사회·경제 상황을 조사·분석하고, 사회정책 및 사회보장 제도를 수립하고 지원하기 위해 설립되었다. 주로 보건 의료·사회복지 분야에 대한 제도 평가 및 정책개발, 교육 및 홍보를 수행한다.

이상의 선정된 네 개의 기관을 대상으로, 기

관별 연구데이터 관련 자료를 살펴보고자 한다. 분석할 자료는 총 두 종류로 분석대상 기관이 수행한 연구과제의 최종보고서, 기관 내부의 연구데이터 관리지침 및 매뉴얼이다.

연구과제 최종보고서는 연구데이터가 실제로 생산되고 정리된 최종 결과물을 보여주는 문서로, 연구자가 어떠한 데이터를 생산했는지, 어떻게 기술했는지 확인할 수 있다. 이를 통해 해당 기관의 연구는 어떤 메타데이터 요소를 필요로 하는지, 실무적 맥락 내에서 도출해낼 수 있으므로 기관별 연구데이터 특성 파악에 있어 유효하다고 판단된다.

기관의 연구데이터 관리지침 또는 매뉴얼은 실질적으로 기관 내부에서 연구데이터를 어떻게 관리하는지에 대한 세부 가이드라인을 제공하는 문서이다. 실제 현장에서 구체적인 프로세스로서 연구자의 실무 수행을 직접 규율하기

때문에, 기관이 주요하게 다루는 연구데이터 정보가 무엇인지 확인할 수 있다.

한편, 연구데이터 관리지침과 함께 분석대상 기관 선정 기준 중 하나였던 연구과제 관리규정은 분석 자료에서 제외하였다. 연구데이터 관리에 대한 요구 사항을 직접적으로 확인할 수 있는 관리지침과는 달리, 연구과제 관리규정의 중심 내용은 연구 개발비 집행, 과제 평가 등 주로 행정적 절차를 다루고 있어 연구데이터 기록관리 메타데이터 설계와 직접적인 관련성이 낮다고 판단했기 때문이다.

분석 자료는 주로 웹 검색을 통해 수집하였다. 연구보고서는 각 기관의 공식 웹사이트 또는 사이언스온(ScienceOn)에 게시된 자료를 활용하였으며, 연구데이터 관리에 관한 내부 지침 및 메뉴얼은 공공기관 경영정보 공개시스템인 알리오(ALIO)를 통해 확보하였다. 단, 질병관리청은 중앙행정기관으로서 예외적으로 국가법령정보센터를 통해 관련 현행규칙을 확인하였다.

기관별 연구데이터 관련 자료의 분석은 '데이터 유형', '연구 조건 기록', '도구 및 소프트웨어 정보', '연구 절차 설명', '지침상의 요구 항목', '보존 및 접근성 조치'의 여섯 가지 항목을

기준으로 진행한다. 해당 항목들은 연구데이터의 재현성 확보와 기록관리 측면에서의 장기보존 및 활용이라는 두 가지 목적에 기반하여 도출하였다.

연구 재현성은 연구데이터의 검증 및 재활용과 밀접하게 연관된 개념으로, 이를 확보하기 위해서는 연구에 사용된 데이터의 유형이 명확히 구분되어야 한다. 연구데이터 관리 및 공유를 위한 국내 메타데이터 표준인 TTA.KO-10.0976(이하 'TTA.KO')는 연구데이터 재사용을 지원하기 위해 dataType, dataFormat 요소를 포함하고 있으며, 국제 표준인 DataCite Metadata Schema(이하 'DataCite') 역시 resourceType, formats 요소를 통해 이를 강조한다. 한편, Leipzig et al.(2021)는 과학기술 분야 데이터 재사용을 위해 분석에 사용된 소프트웨어 버전, 실험 변수, 데이터 출처 등의 연구 조건을 포함한 메타데이터의 필요성을 제시하였고, 최형욱과 정은경(2017)은 사회학 분야 연구 결과가 사회적 시대상 등의 외부 환경, 데이터 수집 과정의 영향을 받는다고 언급하였다. 이는 단순한 데이터 자체뿐 아니라, 데이터 생

〈표 3〉 연구데이터 자료 분석 항목 기준

분석 항목	연구보고서
데이터 유형	<ul style="list-style-type: none"> • 데이터 유형에 따라 보존 전략이 상이함 • 연구데이터의 재현성 확보와 기록관리 측면에서의 장기보존 및 활용이라는 두 가지 목적에 기반하여 도출
연구 조건 기록	<ul style="list-style-type: none"> • 연구 재현성 및 연구데이터 정확성 확보를 위한
도구 및 소프트웨어 정보	<ul style="list-style-type: none"> • 기존 표준 TTA.KO-10.0976, DataCite Metadata Schema 요소 참고
연구 절차 설명	<ul style="list-style-type: none"> • Leipzig(2021), 최형욱과 정은경(2017)은 연구 조건 및 데이터 생산맥락 정보가 필요함을 언급함
지침상의 요구 항목	<ul style="list-style-type: none"> • 기록정보의 신뢰성 확보를 위해 각 기관이 자체적으로 설정한 지침과 규정에 따라 기록을 수집·보존·기술하여 기록 생성 맥락을 명확히 해야 할 필요성이 있음(ISO 15489)
보존 및 접근성 조치	<ul style="list-style-type: none"> • 보존 방식, 접근권, 활용제한 등을 통해 연구 재사용 측면에서 주요 역할 수행

성과 분석과 관련한 전 과정에 걸친 맥락 정보가 함께 제공되어야 연구 재현성이 확보될 수 있음을 시사한다. 이에 따라 데이터 유형, 연구 조건 기록, 도구 및 소프트웨어 정보, 분석 절차 설명의 항목을 연구 재현성 확보를 위한 분석 기준으로 설정하였다.

기록관리 관점의 핵심 요소는 데이터 유형, 지침상의 요구 항목, 보존 및 접근성 조치 항목이다. 기록관리는 단지 기록의 보존뿐만 아니라, 해당 기록이 추후 이해되고 재사용될 수 있도록 맥락을 보존하는 데 목적이 있다. 이를 위해서는 데이터의 본질적 속성은 물론 보존 기간, 접근 권한과 같은 요소들이 함께 고려되어야 한다. 현용 및 준현용 기록의 기록관리 국제 표준인 ISO 15489는 기록정보의 신뢰성 확보를 위해 각 기관이 자체적으로 설정한 지침과 규정에 따라 기록을 수집·보존·기술할 것을 강조하였으며, 이는 기록관리 메타데이터가 기관 지침의 요구 사항을 충실히 반영하여 기록 생성 맥락을 명확히 해야 함을 의미한다.

한편, 보존 및 접근성 조치 항목은 기록관리 관점에서뿐만 아니라 연구 재현성 측면에서도 중요한 역할을 수행한다. DataCite는 rights, publisher, identifier 요소를, TTA.KO는 accessRights, useLimitation, preservationStatement 요소를 명시함으로써 보존 방식, 접근권, 활용제한 등의 중요성을 제시하고 있다.

본 연구에서 설정한 6가지 분석 항목은 기존의 메타데이터 표준 및 선행연구를 기반으로 도출된 것으로, 특정 데이터 형태나 연구 분야에 종속되지 않는다. 이 분석 기준은 연구데이터 메타데이터 설계의 핵심 목적이라 할 수 있는 재현성 확보와 기록의 장기보존, 활용 가능

성 보장을 중심으로 구성되었으며, 다양한 분야와 데이터 유형에 적용 가능한 범용적인 분석 틀로 기능할 수 있다.

3.2 과학기술 분야 기관의 사례

3.2.1 질병관리청

질병관리청은 감염병 예방 및 관리, 그리고 보건의료·역학 조사를 기반으로 정책 수립을 위한 근거 자료를 생산하는데 그 목적을 두고 있는 중앙행정기관으로, 주로 감염병 백신·치료제 핵심 기술 확보와 질병의 대비·대응 역량 강화, 감염병 위험분석·평가, 예측모델 개발 및 위험요인에 관한 연구를 다룬다. 연구사업유형은 총 4가지로, 진행 연구량이 많은 유형 차례대로 연구개발, 정책연구, 일반과제, 민간경상보조사사업이다. 연구보고서의 주제분류는 기초의학, 임상의학, 치의학, 간호학, 예방의학/보건학, 약학, 한의학, 수의학, 기타의학으로 구분된다.

질병관리청은 과학기술 분야를 다루는 기관이지만 보건의료를 주제로 한 다양한 연구를 수행하며, 설문조사나 심층 면접 등을 활용한 지역 단위 건강조사, 정책 전략 수립과 같은 인문사회적 성격의 연구도 함께 진행하고 있다. ‘결핵으로 인한 재난적 비용 추계 관련 연구’, ‘마이코플라스마 페렴균 감염증 진료지침 마련’, ‘역학조사관 교육 표준교재 개발’ 등의 연구를 그 예시로 들 수 있다. 그러나 본 연구는 분야별 연구데이터의 특성을 비교하고 그에 따른 시사점을 도출하는 것을 목적이기 때문에 분석 자료 연구보고서 선정 시 인문사회적 성격의 연구는 제외하고 과학기술 분야의 특성이 뚜렷한 연구만을 대상으로 하였다.

분석 자료로 선정된 연구개발 사업유형의 연구보고서 3편은 'miR-204 조절을 통한 알츠하이머 치료제 후보물질 개발', '전분화능줄기세포 유래 분화세포의 유전체 안정성 및 종양원성 평가', '전장 유전자 정보 기반 SARS-CoV-2 항원성 변이 예측 시스템 개발 및 HTS 기반 에피토프 특이적 항바이러스 후보물질 발굴'이다. 또 다른 분석 자료인 연구데이터 관리지침의 경우, 국가법령정보센터에서 현행 행정규칙인 '질병관리청 연구노트 관리지침'을 활용하였다.

다음의 <표 4>는 4개의 질병관리청 자료를 6개 분석 항목 기준으로 분석한 결과를 정리하여 나타낸 것이다.

질병관리청의 연구보고서는 질병 및 치료제 개발을 위한 정량적 실험 데이터 기반 연구로

서, 실험 환경, 장비, 분석 방법, 반복 조건 등이 체계적으로 기술되어 있다. 정교한 조건 기록과 실험 절차 관리가 매우 강조되고 있음을 알 수 있으며, 연구노트 관리지침 역시 실험 장비 및 도구 기록, 실험 실패 및 예외 사항 포함, 접근 제한 및 보존 기간 등 실무적·법적 요건을 갖춘 기록으로써 기준을 명확히 제시하고 있다. 또한, 질병관리청은 의생명과학 분야의 연구를 다루는 기관으로서 IRB 또는 유사한 윤리 심의 절차가 필수적으로 요구된다. 이로 인해 메타데이터 설계 시 생명윤리, 조사 기반 분석 등이 병행되어야 함을 시사한다.

3.2.2 한국지질자원연구원(KIGAM)

한국지질자원연구원은 국내 유일의 지질자원

<표 4> 질병관리청 자료 분석 결과

분석 항목	연구보고서	연구노트 관리지침
데이터 유형	<ul style="list-style-type: none"> 유전체 데이터, 전자체 데이터, 실험 결과 데이터, 세포·동물 모델 기반 분석 자료, 고속 스크리닝(HTS) 결과 등 정량적 생물 의학 실험데이터 활용 	<ul style="list-style-type: none"> 실험, 조사, 임상시험 등에서 생성되는 연구노트 및 연구기록물을 포함 전자 및 서면 형태 모두 허용
연구 조건 기록	<ul style="list-style-type: none"> 실험 대상(세포주, 생쥐 모델), 실험 환경(온도, 처리 조건, 유전자 편집 조건 등), 분석 조건(Library 구성, 실험 반복 횟수 등)이 상세히 기록됨 	<ul style="list-style-type: none"> 연구명, 실험 목적, 배경, 조건, 진행 과정, 실험 실패, 변경 사항 포함 시료 상태, 장비 설정, 환경 조건 등
도구 및 소프트웨어 정보	<ul style="list-style-type: none"> NGS 분석 소프트웨어, 유전자 발현 분석 도구(예: RNA-Seq), HTS 분석 시스템, ImageJ, GraphPad 등 실험 및 시각화 도구 사용 명시 	<ul style="list-style-type: none"> 실험 사용 장비, 소프트웨어의 버전 및 설정값 기록 권장
연구 절차 설명	<ul style="list-style-type: none"> 실험 디자인 → 시료 준비 → 유전자 발현/변이 분석 → 약물 처리 및 반응 관찰 → 데이터 분석 → 후보물질 제안 등의 단계로 절차 기술 	<ul style="list-style-type: none"> 실험 절차와 분석 단계, 측정 방법 등을 단계별로 구체적 기술해야 함 연구노트는 단순 기록이 아닌 재현 가능한 실험 흐름의 기록 수단으로 간주
지침상의 요구 항목	<ul style="list-style-type: none"> '연구노트', 'IRB', '생명윤리 심의', '표준작업지침(SOP)' 등 내부 규정에 따른 절차 언급 	<ul style="list-style-type: none"> 행정적 요구 사항: 연구노트 작성 의무자, 확인자 서명, 날짜 표기, 연구책임자의 정기 점검 전자노트: 위·변조 방지 기능과 접근 로그 기록을 의무적으로 포함
보존 및 접근성 조치	<ul style="list-style-type: none"> 일부 실험 데이터는 '추후 논문 게재 후 공개 예정' 등으로 기재 별도 DOI, 공개 시스템 연계는 명시되어 있지 않음. 대부분 제한적 내부 활용 중심 	<ul style="list-style-type: none"> 연구노트의 보관: 최소 5년 이상 보존 연구 종료 후 기관이 지정한 방식으로 이관 및 보관 열람 제한, 접근 등급 설정 가능 필수 전자노트: 백업과 무결성 관리 요구됨.

분야 정부출연기관으로서 광물자원의 탐사·개발·활용과 관련된 연구개발을 수행한다. 국내외 육상·해저 지질조사, 지하자원의 탐사·개발·활용, 지질재해 및 지구환경변화 대응 등 지질 조사 및 부존자원 관련 연구가 주요 목적이다. 또한, 국가 차원의 지질자원 연구데이터 관리, 보존체계 구축을 통한 연구데이터 자산화 및 빅데이터화를 위하여 기관 리포지터리인 지오빅데이터 오픈플랫폼(Geo Big Data Open Platform)에 지질자원 연구데이터를 무료로 공개하고 있다.

분석 자료로 사용한 연구보고서 3편은 '3차원 다공성 나노시트 촉매 제조 및 이의 바이오

플라스틱 분해 반응 연구', '온실가스 방출원으로서의 한반도 대륙붕 해저메탄 자연분출 규명 및 정량화 연구', '한국 해양분지 내 능동적 수소생산 가능성에 대한 예비 연구'로, 사이언스온(ScienceOn)에서 수집하였다. 또 다른 분석 자료인 연구데이터 관리지침은 지오빅데이터 오픈플랫폼에 게시된 자료를 활용하였고, '연구데이터 수집 지침', '연구데이터 관리지침', '연구데이터 보존지침', '윤리·저작권·라이선스 지침'의 4개 지침을 대상으로 하였다.

다음의 <표 5>는 한국지질자원연구원 연구보고서 3편과 관리지침 4개를 분석한 결과를 분석 항목별로 정리한 것이다.

<표 5> 한국지질자원연구원 자료 분석 결과

분석 항목	연구보고서	연구노트 관리지침
데이터 유형	<ul style="list-style-type: none"> • 실험 장소(위도/경도), 수집, 시기, 기온, 음향 반사 신호, 지질 분석, 가스 농도, 나노소재 특성값 등 고정밀 실험데이터 • 실험 장비에서 수집된 수치 기반 정량 데이터 중심 	<ul style="list-style-type: none"> • AGS, LAS, SGY, XTF 등 분야별 파일 형식 지정 • 구조화된 계층 데이터 위주로 데이터 유형 분류 및 메타정보 연계 권고
연구 조건 기록	<ul style="list-style-type: none"> • 실험 지점, 수집, 온도, 시간, 회전 속도, 촉매 조건 등 상세 변수 명시 • 실험 반복 조건 및 환경 제어 요소 기술 	<ul style="list-style-type: none"> • 측정 환경(위치, 장비, 시간 등)을 필수 입력 항목으로 설정 • 연구계획서 및 DMP에서 실험 조건의 사전 정의의 요구
도구 및 소프트웨어 정보	<ul style="list-style-type: none"> • EK60 음향기, 회전형 유동반응기, TEM, AFM, XRD, HPLC 등 실험 장비와 분석 도구 명시 • FlareHunter, ESP3 등 분석 소프트웨어 	<ul style="list-style-type: none"> • 도구명, 버전, 제조사, 활용 목적 등 메타데이터에 포함 권고 • 알고리즘 및 자동 분석 툴 사용 시 프로세스 명시 요건 포함
연구 절차 설명	<ul style="list-style-type: none"> • 수집 → 신호 처리 → 분석 → 정량화 → 시각화 등 분석 흐름 단계별 기술 • 특정 기준값과 판단 조건 명시로 반복 실험 가능성 확보 • 나노촉매 시트: 기계적 공정 설계와 반응 조건 최적화 흐름 구체적으로 제시 	<ul style="list-style-type: none"> • 데이터 수집 → 처리 → 보존 단계별 흐름도 • 각 단계별 책임과 점검 절차 명시
지침상의 요구 항목	<ul style="list-style-type: none"> • SOP 및 계획서 기반 실험 설계 간접적으로 언급됨 	<ul style="list-style-type: none"> • 생산자, 위치 좌표, 생성일, 품질관리 등 메타데이터 필수 항목 지정 • 연구데이터 품질관리를 위한 FAIR 원칙 기반 체크리스트 제공 • DMP 제출·검토·보완 절차 포함
보존 및 접근성 조치	<ul style="list-style-type: none"> • 원자료 보존 위치·형식은 보고서 내 명시되지 않음 • 일부 보고서는 데이터 공개 여부 불명확 	<ul style="list-style-type: none"> • 장기보존을 위한 리포지터리 권고 • DOI, IGSN 발급 권장 • 접근 등급 관리, 메타데이터 일괄 백업, 기계 판독 포맷 사용 지침 포함

한국지질자원연구원은 정량적 실험 조건의 명시와 계측 장비 기반의 실험 데이터 기록을 중점적으로 수행하고 있다. 특히 연구보고서에 서는 수심, 온도, 반응 시간, 촉매 조건 등 세부 실험 변수를 반복 가능하게 기록하고 있으며, 분석에 사용된 도구와 소프트웨어, 처리 알고리즘도 명확히 기술되어 있었다. 또한 연구데이터 관리지침을 통해, 실험데이터의 형식, 메타데이터 입력 항목, FAIR 원칙 기반 점검 항목 등을 제도적으로 규정하고 있으며, 장기보존을 위한 DOI·IGSN 발급과 기계판독 가능한 포맷 사용을 권장하고 있다.

3.3 인문사회 분야 기관의 사례

3.3.1 한국교육개발원

한국교육개발원은 한국교육이 당면한 현장 문제를 파악하고, 데이터를 바탕으로 교육 관련 과학적 분석과 연구를 통해 제반 문제에 대한 합리적인 해결방안을 탐색하고자 하는 교육 정책 전문연구기관이다. 주요 업무는 교육체제 및 교육 정보 공시에 관한 연구개발, 학교에 관한 종합적 평가 지원, 교육 통계조사이다. 연구보고서의 분류 체계는 기본연구보고, 현안보고, 수탁보고, 이슈페이퍼, 기술보고, 교수자료, 연구자료, 통계자료 등 다양하다. 본 연구는 연구보고서의 대표성을 확보하기 위해, 한국교육개발원이 발간한 총 6,736편의 보고서 중 '기본연구보고'로 분류된 1,497편을 분석대상으로 선정하였다.

분석 자료로 사용한 연구보고서 3편은 '디지털 전환 시대 교육격차 변화 양상 연구', '중등 글로벌 인문교양교육 교육지표 개발 연구', '지

역-대학 연계 강화를 위한 데이터 실태 분석 및 활용 방안 연구'로, 한국교육개발원의 공식 웹페이지에서 수집하였다. 사실상 기관의 연구데이터 관리지침으로 활용되는 '조사자료 관리 및 공개 지침'은 알리오에서 제공하는 내부 규정 게시판에서 확인하였다.

〈표 6〉은 한국교육개발원의 연구보고서와 조사자료 관리 및 공개 지침을 분석하여 그 특징을 분석 항목별로 분류하여 정리한 것이다.

한국교육개발원은 교육 통계, 정책 평가, 지표 개발 등 인문사회 분야 특유의 맥락 중심적 연구를 수행하는 기관으로, 연구보고서들은 조사 설계, 대상 설정, 지표 정의 과정에 대한 상세한 설명을 제공하고 있으며, 앞서 분석한 과학기술 분야 기관의 사례와는 달리 데이터 처리 도구와 실험 조건은 제한적으로 제시되고 있다. 이를 통해 한국교육개발원은 인문사회 분야 기관으로서 조사 기반의 반복 가능성을 위해 분석 기준의 일관적 유지와 정책 적용 가능성을 중심으로 재현성을 인식하고 있음을 확인할 수 있다.

한편, 조사자료 관리 및 공개 지침에서는 조사자료의 법적·윤리적 요건을 명확히 제시하고 있으며, IRB 승인, 설계서 제출, 출처 명시, 비식별화 요구, 외부 이용 신청 절차 등은 인문사회 분야 특유의 윤리성과 맥락성을 반영한 재현성 확보 노력이라고 판단된다.

3.3.2 한국보건사회연구원

한국보건사회연구원은 복합적인 사회 변화에 능동적으로 대응하는 데 필요한 정책 대안을 제시함으로써 국민의 삶의 질 향상에 이바지하는 것을 목적으로 한다. 이를 위해 크게 인

〈표 6〉 한국교육개발원 자료 분석 결과

분석 항목	연구보고서	연구노트 관리지침
데이터 유형	<ul style="list-style-type: none"> 정량적: 국가 통계(KESIS, PISA, 서울서베이 등) 활용 정성적: FGI, 델파이, 전문가 자문, 지표 개발 등 혼합자료 기반 연구가 일반적 	<ul style="list-style-type: none"> 설문, 인터뷰, 코딩자료 등 조사자료 전반을 포괄하며, 정량/정성 자료 모두 포함 메타데이터와 원자료 구분 저장 요구
연구 조건 기록	<ul style="list-style-type: none"> 교육대상, 학교급, 지역, 정책 시기 등 연구 맥락 설명 충실 조건 변수에 대한 명확한 정의 존재(예: 디지털 인프라 수준, 교육기회 불균형 등) 	<ul style="list-style-type: none"> 조사 목적, 대상, 기간, 항목 수 등 기술 IRB 심의 및 조사승인 절차 포함 조사설계서 필수 제출 및 관리 의무화
도구 및 소프트웨어 정보	<ul style="list-style-type: none"> 구체적 도구 언급은 제한적 조사 도구로는 구조화된 설문지와 전문가 검토 기반 지표구성 방식 포함 	<ul style="list-style-type: none"> 자료는 엑셀, SPSS 등 표준 포맷으로 제출 권장 분석에 사용된 도구/시스템 명시는 제출양식에 포함되어야 함
연구 절차 설명	<ul style="list-style-type: none"> 문헌 조사 → 지표 초안 구성 → 전문가 검토 및 정제 → 지표 확정 등 분석 절차 비교적 명확히 기술됨 데이터 수집 → 정제 → 분석 → 정책 제언의 흐름 유지 	<ul style="list-style-type: none"> 조사자료 등록 → 기초정보 기입 → 심의 → 공개 승인 단계로 관리됨 등록일, 검수일 등 연구 절차별 메타데이터 요구
지침상의 요구 항목	<ul style="list-style-type: none"> 조사 설계의 정합성, IRB 승인 여부, 교육정책 연계 필요성 등 간접적으로 드러남 조사자료에 대한 제출 의무는 보고서 내 직접 언급되지 않음 	<ul style="list-style-type: none"> 조사자료 등록 시 조사명, 목적, 설계서, 조사자 정보 필수 조사자료 관리카드와 메타데이터 작성 서식 양식 제공 비식별화 요구 포함
보존 및 접근성 조치	<ul style="list-style-type: none"> 자료 공개 여부나 보존 정책은 명확히 언급되지 않음 대부분 결과 보고서 수준에서 공유 제한적이며, 데이터 자체는 내부 보관 전제로 작성된 경우 많음 	<ul style="list-style-type: none"> 최소 5년 이상 보존 의무 메타데이터와 원자료 분리 보관 자료 이용 신청서 및 서약서 요구 외부 공개 시 출처 명시 의무화 및 제한 규정 포함

구와가족, 사회서비스, 소득보장, 보건의료, 사회보장일반의 5가지 사회정책 분야를 대상으로 정량적 조사와 정책 분석을 수행하고 있다.

분석 자료로 사용한 연구보고서는 3편으로, '일본의 인구감소 대응 정책 사례 연구', '한부모 가족의 시간 및 경제적 자원과 정책', '행복 및 삶의 질 측정과 정책 활용성 제고 방안 연구'이다. 본 기관의 연구데이터 관리지침의 역할을 하는 '조사자료 관리지침'은 알리오에서 수집하였다.

〈표 7〉은 한국보건사회연구원의 연구보고서 3편과 조사자료 관리지침을 분석하여 그 결과를 분석 항목별로 구분하여 작성한 것이다.

해당 기관의 연구보고서에서는 통계청 마이크로데이터, 사회통합실태조사, 자체 조사 등을 활용한 구조화된 정량 데이터가 주로 산출되며, 변수 정의서(코드북), 정책 지표, 문헌 기반의 정성적 분석 자료도 병행하여 생성된다. 이를 통해 한국보건사회연구원의 연구데이터는 정량-정성 혼합형임을 확인할 수 있다. 또한, 연구데이터 관리지침의 역할을 하는 조사자료 관리지침은 한국교육개발원의 사례와 유사하게 조사자료의 등록, 메타데이터 작성, 비식별화, 보존 및 공개 절차 등을 체계화하고 있어, 연구데이터의 정책적 활용성과 공공 접근성 확보를 주요 관리목표로 삼고 있음을 확인할 수 있다.

〈표 7〉 한국교육개발원 자료 분석 결과

분석 항목	연구보고서	연구노트 관리지침
데이터 유형	<ul style="list-style-type: none"> • 조사실험, 통계분석 중심 연구로 인해 정량적 조사 자료와 이를 기반으로 도출된 정리된 통계지표 및 정책 분석 결과, 보조적인 정성적 자료로 구성됨 	<ul style="list-style-type: none"> • 설문, 면접, 코드북, 조사표 등 조사 기반 데이터 전반을 포함 • 정량·정성자료 모두 관리 대상으로 명시
연구 조건 기록	<ul style="list-style-type: none"> • 응답자 특성, 기구 유형, 시간대, 정책 구간 등 분석 조건이 구체적으로 명시됨 • 특히 삶의 질 연구에서는 조사 요일, 조사자 선정 방식까지 기술됨 	<ul style="list-style-type: none"> • 조사 개요서, 조사 설계 문서 제출 필수 • IRB 심의, 조사자 정보, 대상자 선정 방식 등을 메타데이터로 기록해야 함
도구 및 소프트웨어 정보	<ul style="list-style-type: none"> • 분석 도구의 구체적 명시 드뭄 • 조사 설계와 측정 도구(질문지, 지표 등)에 대한 설명은 비교적 상세 	<ul style="list-style-type: none"> • 데이터 포맷은 엑셀, SPSS 등 표준형식을 권장 • 분석 도구 및 시스템 정보는 조사자료 등록 시에 서식에 포함됨
연구 절차 설명	<ul style="list-style-type: none"> • 문헌 조사 → 설문 설계 → 자료 수집 → 통계분석 → 정책 제언으로 이어지는 일관된 흐름 유지 • 설문과 분석 절차 간 논리적 연결성 중시 	<ul style="list-style-type: none"> • 조사자료 등록 절차를 단계별로 구분(등록 → 심의 → 공개 승인) • 등록일, 접수일, 메타정보 작성 일자 등 절차 기반 메타데이터를 요구
지침상의 요구 항목	<ul style="list-style-type: none"> • 조사 설계와 윤리적 책임(사전 동의, 비식별화, 통계법 준수 등) 언급 	<ul style="list-style-type: none"> • 조사명, 목적, 조사자, 설계서, 비식별화 조치 등 필수 기재 • 조사자료 관리카드 및 메타데이터 작성 양식 제공
보존 및 접근성 조치	<ul style="list-style-type: none"> • 보고서 자체는 공개되나, 조사자료의 보존 위치나 접근 조건은 명시되지 않음 	<ul style="list-style-type: none"> • 책임연구자가 최소 5년 이상 보존 • 원작성 비식별화 상태로 관리 • 외부 공개 시 서약서, 출처 명시 등 이용 절차 명시 • 메타데이터와 원자료 분리 보관 • 폐기 기준 포함

3.4 분야별 연구 재현성 확보 방안 비교를 통한 메타데이터 요구 사항 파악

본 장에서는 3.3장에서 확인한 분야별 대표 기관의 실제 자료를 토대로 과학기술 분야와 인문사회 분야의 연구 재현성 확보 방안을 비교하고, 기록관리 메타데이터 설계 요구 사항을 파악하고자 한다. 이를 위해 각 기관의 자료 분석 결과를 분야별로 구분하여 공통점을 추출하고, 해당 공통점을 바탕으로 분야별 연구 재현성의 개념과 그 확보 방안을 고찰하였다.

다음의 〈표 8〉은 과학기술 분야 기반의 연구 기관인 질병관리청과 한국지질자원연구원의 자료 분석 결과에서 분석 항목에 따라 공통점을

추출하여 정리한 것이다.

질병관리청과 한국지질자원연구원의 연구데이터는 주로 실험 및 관측을 통해 수집된 정량적 데이터와 이를 분석해 도출한 2차 데이터로 구성되어 있으며, 표준 포맷에 따라 구조화되어 체계적으로 관리·저장되고 있다. 두 기관의 연구보고서는 정량 변수의 상세한 기록과 장비 및 도구에 관한 정보를 명시함으로써 동일한 실험 환경의 재현이 가능하도록 하고 있다. 또한, 분석 알고리즘과 프로세스를 함께 기록하여 계산적 반복 가능성도 확보하였다. 연구 절차는 단계별 흐름에 따라 구조화되어 있어, 제삼자도 동일한 경로로 실험을 반복할 수 있도록 설계되어 있다. 아울러, 연구노트와 관리지침을 통해

〈표 8〉 과학기술 분야 대표 기관(질병관리청, 한국지질자원연구원)의 분석 항목별 공통점

분석 항목	공통적으로 나타난 분석 결과
데이터 유형	<ul style="list-style-type: none"> • 실험·관측 기반의 정량적 구조화 데이터 • 측정값 등 상세 수치 기반 자료 생성 • 표준 포맷 기반의 데이터 수집·관리 추구
연구 조건 기록	<ul style="list-style-type: none"> • 실험 조건을 정량적으로 상세하게 기록 • 실험의 반복 가능성과 조건 통제를 위한 정보 제공 필수
도구 및 소프트웨어 정보	<ul style="list-style-type: none"> • 실험 장비명, 버전, 설정값, 분석 소프트웨어 정보 등 기술적 메타데이터 기입 강조 • 동일 실험 반복 가능성 확보를 위한 알고리즘, 장비 세팅 정보 포함
연구 절차 설명	<ul style="list-style-type: none"> • 수집 → 처리 → 분석 → 결과 도출의 일관된 절차 흐름이 보고서에 명확히 기술 • 각 단계별 수행 내용과 조건 구조화
지침상의 요구 항목	<ul style="list-style-type: none"> • 지침을 통해 작성자, 날짜, 실험 목적, 서명, 변경 이력 등의 기록 요건 명시 • 장비 및 도구 정보 기재, 정기 점검 등 관리규정 포함
보존 및 접근성 조치	<ul style="list-style-type: none"> • 접근 권한 설정, 백업 및 무결성 점검 등 기술적 보존 조치 강조 • DOI와 같은 자료 식별자 발급 권장을 통해 외부 공유를 고려한 항목 존재

작성자, 작성 시점, 수정 이력 등 행위 기반의 메타데이터를 기록함으로써 연구의 검증 가능성을 높이고 있으며, DOI와 같은 식별자 발급 및 접근 권한 관리를 통해 데이터의 추적성과 활용 가능성 역시 확보하고 있다. 공통적으로 실험·관측·분석의 과정과 정밀한 통제를 위

한 연구 조건, 자료 흐름 등을 포함한 전체 과학적 과정이 상세하게 드러날 수 있는 요소를 중시함을 확인할 수 있다.

한편, 〈표 9〉는 인문사회 분야를 대표하는 기관으로서 한국교육개발원, 한국보건사회연구원을 분석한 결과 나타난 공통점을 보여준다.

〈표 9〉 인문사회 분야 대표 기관(한국교육개발원, 한국보건사회연구원)의 분석 항목별 공통점

분석 항목	공통적으로 나타난 분석 결과
데이터 유형	<ul style="list-style-type: none"> • 정량적 설문조사 데이터(서베이)와 정성적 자료(FGI, 인터뷰 등)를 혼합하여 활용 • 사회통계, 정책자료, 2차자료 기반 분석이 일반적 • 산출물은 SPSS, Excel 등 구조화된 데이터와 지표, 코드북 등 문서 기반 결과물이 병행됨
연구 조건 기록	<ul style="list-style-type: none"> • 응답자 특성, 조사 시점, 지역, 조사 범위 등 조사 설계에 기반한 변수 정의 기술 • 연구 맥락, 조사 목적, 분석 단위 등 정성적 맥락 정보 기록 강조
도구 및 소프트웨어 정보	<ul style="list-style-type: none"> • 분석 도구를 직접적으로 명시하지 않는 경우도 있음 • 조사 설문지, 지표 개발 문서, 코드북 등 분석 도구의 기반 자료는 비교적 풍부
연구 절차 설명	<ul style="list-style-type: none"> • 문헌 조사 → 조사 설계 → 자료 수집 → 분석 → 정책 제언의 전형적인 사회조사 흐름이 유지됨 • 정량/정성 데이터를 함께 설명하고, 설계-분석 간 해석적 일관성 유지 노력 존재
지침상의 요구 항목	<ul style="list-style-type: none"> • IRB 심의, 조사 설계서, 조사자 정보, 비식별 조치 등 법적·윤리적 요구 사항을 지침에 명시 • 메타데이터 작성 서식, 조사자료 등록카드 등 제도화된 관리 양식 사용
보존 및 접근성 조치	<ul style="list-style-type: none"> • 최소 5년 이상 보존 원칙 • 원자료와 메타데이터를 분리 보관 • 외부 공개 시 출처 명시, 이용 신청, 서약서 제출 등 절차적 통제 필요 • 비식별화 조치, 열람 등급 설정 등 개인 정보 보호 조치 포함

한국교육개발원과 한국보건사회연구원은 모두 사회조사 기반의 정량·정성 혼합자료를 다룬다. 그러나 과학기술 분야와는 달리 정량화된 수치가 아닌 변수 정의로서 연구 조건을 기술하는 경우가 많았으며 조사 설계를 위한 정성적 맥락 정보를 강조하는 경향이 뚜렷했다. 또한, 분석 도구 자체에 대한 정보와 분석 도구를 통한 분석 절차를 명시하지 않는 경우가 존재하였다. 대신 분석 도구의 기반이 되는 코드북, 조사 설문지 등의 자료는 상대적으로 풍부하게 제시하고 있었다. 또한, 과학기술 분야 기관의 사례보다 상대적으로 법적·윤리적 사항에 대한 요구 사항이 명확하였으며 접근 조건 제한에서도 보다 엄격한 절차가 요구된다. 이를 통해 인문사회 분야 연구에서는 정성·윤리 기반의 요소가 연구 검증에 있어 핵심적으로 작용함을 확인할 수 있다.

재현성이라는 개념은 모든 학문 분야에서 연구의 신뢰성과 타당성을 확보하기 위한 핵심적 기준이지만, 앞서 살펴본 바와 같이, 과학기술 분야와 인문사회 분야에서 연구의 검증과 재현성이 분야에 따라 다르게 적용될 수 있다. 분야별 기초에 놓인 학문적 성격과 연구 방법론의 차이로 인해 연구의 신뢰성을 보장하는 방법이 상이하기 때문이다.

과학기술 분야에서는 정량화된 수치가 실험 결과이자 그 자체로 과학적 성과물로 인정받을 수 있다. 따라서 동일한 정량적 수치의 반복 산출 가능성이 곧 해당 연구의 신뢰도로 직결되고, 재현성의 핵심 기준이 된다. 이에 따라 실험 과정에서 수치에 영향을 줄 수 있는 모든 변수를 정확하게 기록하는 것이 재현성 확보를 위한 핵심 과제로 작용한다. 질병관리청과 한국

지질자원연구원의 사례에서도 확인되듯, 해당 기관들은 실험 대상, 반복 조건, 장비 정보, 소프트웨어 설정값, 알고리즘 절차 등을 상세히 기록하고 이를 메타데이터 요소로 체계화함으로써 정량적 재현성을 뒷받침하고 있다.

반면, 인문사회 분야에서도 설문조사 결과, 통계치와 같은 정량화된 수치가 연구데이터로 산출될 수는 있지만, 수치 자체만으로는 연구의 목적을 충족하거나 해석 가능한 연구성으로 기능하지 않는다. 오히려 그 수치가 어떤 맥락에서 도출되었는지, 어떤 개념적 기준과 분석틀에 따라 해석되어야 하는지가 연구의 핵심으로 작용한다. 따라서 동일한 조사를 반복한다고 하더라도, 조사 설계의 논리 구조와 변수 정의, 지표구성 방식 등 해석의 일관성과 맥락 유지 여부가 재현성의 주요 기준으로 작용한다. 한국교육개발원과 한국보건사회연구원은 분석 도구나 소프트웨어의 설정값보다는 설문 문항, 코드북, 조사 설계서 등 연구 맥락과 설계 과정을 충실히 기록하고 있으며, 이는 인문사회 분야에서의 재현성이 해석 가능성을 중심으로 정의된다는 점을 실증적으로 보여준다.

이와 같은 과학기술-인문사회 분야 간의 차이점은 각 분야의 재현성 개념에 상응하는 기록관리 메타데이터 요구 사항의 차이를 발생시킨다. 과학기술 분야에서는 실험 조건, 장비 정보, 분석 알고리즘 등 정량적 실험 정보를 중심으로 한 메타데이터가 필수적으로 요구되며, 인문사회 분야에서는 조사 설계 문서, 변수 정의, 조사 목적 등 절차적·해석적 맥락 정보를 중심으로 한 메타데이터가 핵심 요소로 설정될 필요가 있다. 이는 각 분야의 학문적 재현성 개념이 기록관리 요구 사항에도 직접적인 영향을

미친다는 점에서, 분야별 맞춤형 메타데이터 설계의 필요성을 정당화한다.

〈표 10〉은 과학기술 분야 및 인문사회 분야의 재현성과 그 확보 방안을 고찰한 현재까지의 내용을 정리하여, 분야별로 필요로 하는 메타데이터 요구 사항을 나타낸 것이다.

과학기술 분야의 연구 재현성은 단순히 동일 조건의 연구 재현이 아닌 '실험 조건의 통제와 과정을 조건을 정량화하여 반복 실현 가능성을 보장하는 기술적 구조의 구현'으로 판단할 수 있다. 재현성은 실천적이고 기술적인 개념으로, 이를 확보하기 위해선 정량적 수치에 직접 영향을 미칠 수 있는 변수에 대한 상세한 기록이 요구된다.

인문사회 분야의 연구 재현성은 동일한 개념, 사회 현상 또는 문제에 대하여 동일한 연구 설계와 해석 기준을 적용하였을 때 유사한 분석 결과와 정책적 해석을 도출할 수 있는 해석적 일관성과 타당성의 확보를 의미한다. 과학기술 분야와는 달리 결과 자체보다는 연구가 수행되는 조건과 해석 구조의 맥락이 연구의 핵심으로 인식되기 때문이다. 따라서 조사 설계의 정합성을 입증할 수 있는 맥락 정보를 일

관적으로 보존하는 것이 중요하다. 이에 따라 절차적 정보와 해석적 기준을 기록하는 것이 필수적이다.

4. 연구데이터 기록관리 메타데이터의 분야별 설계 방안

4.1 연구데이터 기록관리 메타데이터 설계 기본 원칙

본 장에서는 연구데이터 기록관리를 위한 분야별 메타데이터 요소를 도출하기에 앞서, 기존 연구에서 수립한 기록관리 메타데이터 설계 기본 원칙을 명확히 함으로써 이후 분야별 설계에 있어서도 일관성을 유지하고자 한다.

기존 연구에서 설계한 요소안의 메타데이터는 데이터세트를 기본 단위로 하였으며, 이는 국가연구개발사업 도중 생성된 데이터 하나하나가 시스템에 개별적으로 입력되는 것이 아니라 여러 데이터가 하나의 데이터세트로 묶어서 관리되기 때문이다. 메타데이터의 구성은 상위 요소-하위요소-세부요소의 세 단계로 나뉘어

〈표 10〉 분야별 재현성 개념에 따른 메타데이터 요구 사항 비교

구분	과학기술 분야	인문사회 분야
연구성과물 도출 의의	정량적 수치를 통한 일반 법칙과 가설 검증	맥락의 논리적 연결을 통한 특정 개념, 현상 해석
재현성 개념	실험, 분석, 관측 조건의 통제를 통한 정량적 수치의 반복 산출 가능성	해석 가능한 설계 맥락의 반복 가능성, 타당성
재현성 확보 수단	자동화된 텍스트, 코드 공유, 환경 고정 측정 기기, 방법, 실험 조건	설문지, 지표, 코드북, 틀, 설계서, 데이터 세트
메타데이터 요구 사항	실험 조건, 분석 코드, 장비 정보 등 동일 환경의 통제를 위한 변수 정보	참여자 정보, 연구 배경, 분석 틀, 설계 문서, 윤리적 고려사항 등 절차 정보/해석의 기반 자료
분야별 중점 사항	통제 가능성	설명 가능성
공통 사항	정확성, 반복 가능성, 접근성, 재사용 가능성	

있으며, 연구데이터를 포함한 연구기록물 생산 맥락을 보존하기 위하여 국가연구개발사업 정보와 과제정보를 포함하였다. 기술의 순서는 연구과제 진행 프로세스를 따라 국가연구개발사업-연구개발과제(성과물)-데이터(세트)순으로 설정하였다. 이러한 맥락 정보는 기록의 주요 요건으로서, 과학기술 분야와 인문사회 분야에 구분 없이 연구데이터 기록관리에서 공통적으로 포함되어야 할 사항에 해당한다.

연구데이터의 재현성과 정확성을 확보하는데 있어 핵심적인 역할을 하는 상위요소는 '데이터 정보'에 해당하며, 각 연구데이터의 주요 정보를 포함할 수 있어야 한다. 이를 위해 연구데이터 메타데이터 표준인 DCAT V2, DataCite의 비교·분석을 통해 두 표준에서 모두 등장한 요소를 데이터 정보의 하위요소-세부요소로 지정하였다.

또한, 기록관리 메타데이터로서 현행 표준과의 호환성을 위하여 기록관리 메타데이터의 4가지 표준(NAK 8, NAK 35, ISAD(G), PREMIS 3.0)을 비교·분석하여 공통적으로 활용할 수 있는 필수 기록관리 요소를 지정하였다.

연구데이터의 보존가치와 신뢰성을 유지하는 데 가장 중요하다고 판단한 사항은 연구데이터의 정확성과 접근성, 연구 결과의 재현성, 그리고 재활용성의 보장이었다. 메타데이터 설계안에 이를 실현하기 위해, 연구데이터의 기초적인 정보와 맥락을 담아내도록 하여 정확성을 보장하였고 접근 권한을 포함하여 접근성을 강화하였다. 추가로, 첨부파일 등의 부가 정보를 제공하여 연구데이터의 올바른 해석을 지원하고 재활용을 위한 라이선스 및 인용 정보를 체계적으로 포함하도록 하였다.

이상의 기록관리 메타데이터 설계 기본 원칙을 기반으로, 3장에서 확인한 분야별 재현성 확보 및 연구데이터 관리 요구 사항을 반영하여 각 분야에서 필요한 연구데이터 기록관리 메타데이터 요소를 도출하고자 한다.

4.2 연구데이터 기록관리 메타데이터의 분야별 핵심 요소 도출

본 장에서는 기존 연구의 메타데이터 설계 원칙과 더불어 <표 10>을 반영하여 연구데이터 기록관리 메타데이터에 필요한 핵심 요소를 분야별로 분리하여 제시하고, 본 연구에서 분석하였던 분야별 대표 기관 사례에 근거하여 예시 메타데이터 요소를 도출하고자 한다.

질병관리청과 한국지질자원연구원은 과학기술 분야의 대표 기관으로서 재현성을 단순한 결과 재현이 아닌 실험 조건의 통제와 과정의 반복 가능성을 전제로 한 기술적 실현 가능성의 문제로 접근하고 있다. 이는 연구 환경의 동기화와 연구 절차 구조화를 통해 연구 전반의 과학적 과정을 상세하게 확인할 수 있게끔 기록관리 메타데이터가 지원해야 함을 시사한다. 따라서 해당 분야의 기록관리 메타데이터 설계에서는 정량적 조건 정보, 실험 환경, 실험 흐름 정보, 기술적 식별자 등이 핵심 요소로 포함되어야 한다. <표 11>은 이러한 시사점을 바탕으로 과학기술 분야의 연구데이터 기록관리 메타데이터 예시 핵심 요소를 도출한 것이다.

질병관리청과 한국지질자원연구원이 실제로 운영하는 정량 실험 기반 재현성 확보 방식을 반영하여 기록관리 메타데이터 요소로서 다음과 같이 구조화하였다. 정량 수치 기반 데이터

〈표 11〉 과학기술 분야 연구데이터 기록관리 메타데이터 핵심 요소 예시

분석 항목	확인된 정보	메타데이터 요소	기록관리 목적
데이터 유형	유전체 데이터, 임상시험 데이터, 실험 결과 수치, 센서 기반 수치, 지질·자원 분석값, 고속시퀀싱, 이미지 기반 측정자료 등	데이터 유형, 포맷, 생성 방식, 단위, 샘플 유형, 데이터 구조	데이터의 정체성과 형식 명확화 장기보존 및 호환성 확보
연구 조건 기록	샘플 채취 위치(위도·경도), 수심, 채취 시점, 실험 반복 조건, 환경 설정값	실험 조건(온도, 압력, 농도 등), 위치정보, 실험 반복 횟수, 환경 변수, 샘플 특성 정보, 조건 제어 변수	수치 생성 맥락의 명확화 실험 반복성 보장 조건 복원 기반 마련
도구 및 소프트웨어	실험 장비 명칭(예: NGS 장비), 버전, 소프트웨어 분석 도구, 설정값 및 알고리즘	장비 정보, 버전, 분석 도구, 설정값, 알고리즘	분석 환경의 기술적 재현 실험/분석 장비 등 시스템 기반 검증 및 이식 가능성 확보
연구 절차 설명	자료 수집 → 가공 → 분석 → 결과 도출의 흐름 기술, 각 단계별 수행 방법 기술	실험 절차 단계, 소요 시간, 입력/출력 관계, 수행 주체, 실험 버전, 시험 계획서 정보	실험 흐름의 추적 가능성 확보 정량적 과정의 투명성 확보 과학적 검증 가능성 보장
지침상의 요구 항목	연구노트 기입 의무, 실험자 및 책임연구원 서명, 작성일자, SOP 준수 기록, 메타데이터 등록 등	작성자, 작성 일시, 책임연구자, 프로젝트명, SOP 번호, 실험 목적, 실험노트 번호	책임성과 검증 가능성 확보 법적·제도적 요건 충족 근거 확보
보존 및 접근성 조치	저장 포맷 지정, 실험노트·연구데이터 이중 보관, DOI 발급을 통한 외부 공개 고려, 접근 등급 설정, 백업 시스템 운영	보존 기간, 접근 권한, 백업 주기, 공개 여부, DOI 등 자료 식별자, 메타데이터 위치, 폐기 조건	장기보존 보장 접근성 및 무결성 확보 공공 정보 접근성 균형 유지

를 기록하기 위한 기술적 메타데이터의 상세화가 핵심으로, 이에 따라 ‘실험 조건’, ‘환경 변수’, ‘장비 정보’, ‘분석 도구’ 등 수치에 영향을 줄 수 있는 모든 요소를 명확히 기술하도록 설계되어야 한다. 구체적으로, 데이터 유형 및 형식, 실험 조건의 변수, 사용된 장비 및 소프트웨어의 버전과 설정값, 실험 절차의 단계별 흐름과 수행 주체 등은 모두 기술적 재현성을 확보하기 위한 핵심 정보로 간주된다. 또한, 실험 수행자의 정보, 작성 시점, SOP 번호와 같은 책임성 인증 관련 요소 역시 메타데이터 항목에 포함되어야 한다. 한편, 질병관리청의 사례에서 확인할 수 있듯 의학, 보건학과 같은 과학기술

분야는 생명윤리 심의 등 윤리적 맥락을 포함하는 경우가 존재한다. 이러한 특성을 가진 연구 분야 및 연구기관의 경우, 메타데이터 설계 시 일부 인문사회적 속성이 공존할 수 있음을 고려해야 한다.

인문사회 분야 기관의 예시로 선정되었던 한국교육개발원과 한국보건사회연구원의 분석 결과는 인문사회 분야의 재현성이 과학기술 분야처럼 기술적 반복 가능성에 초점을 맞추기보다는, 사회 현상 또는 개념을 해석하고 설명할 수 있는 맥락의 보존, 연구 설계의 정합성 및 일관성 유지, 데이터 해석의 투명성 확보를 중심으로 설정된다는 점을 보여준다. 따라서 기

록관리 메타데이터 설계 시에도 자료 수집 배경 등 맥락적 요소, 변수 해석 기준, 비식별 처리 절차 등 절차 중심의 요소가 중시되어야 한다. 다음 <표 12>는 이를 반영하여 인문사회 분야의 연구데이터 기록관리 메타데이터 핵심 요소를 도출한 예시이다.

인문사회 분야에서는 주로 설문조사, 인터뷰, 통계자료 등 다양한 정량·정성 데이터를 혼합하여 활용하며, 해당 데이터의 해석 가능성을 담보하기 위해 조사 설계의 정합성과 맥락 정보가 특히 중요하게 다루어진다. 이러한 특징은 메타데이터 설계에도 반영되며, 실험 장비나 분석 소프트웨어 자체에 대한 정보보다는 '설문 문항 구성', '샘플링 방식', '조사 일정표'와 '지표구성 항목' 등 설계기반 정보의 상세 기

록이 핵심 요소로 나타난다. 실제로 분석된 기관 사례에서는 분석 도구의 기반이 되는 문서들이 체계적으로 제공되었으며, 이를 통해 수치의 해석 가능성과 일관된 분석 기준의 유지가 가능하게 된다. 또한, 인문사회 분야는 응답자 정보 보호와 연구 윤리 확보의 필요성이 매우 높으므로, '비식별화 여부', '이용 제한 조건' 등의 법적·윤리적 요구 사항이 메타데이터 작성 과정에서 함께 포함되어야 한다.

4.3 분야별 연구데이터 기록관리 메타데이터 설계 방안 제시

과학기술 분야와 인문사회 분야의 대표 기관 사례를 비교한 결과, 각 분야는 데이터의 성격,

<표 12> 인문사회 분야 연구데이터 기록관리 메타데이터 핵심 요소 예시

분석 항목	확인된 정보	메타데이터 요소	기록관리 목적
데이터 유형	설문조사 원자료(정량), 인터뷰·FGI 전사자료(정성), 정부통계, 정책 지표 데이터 등	데이터 유형(정량/정성 구분), 수집 방식, 포맷 등	접근 방식 구분 및 정리 기준 설정 활용 목적별 분류 보존 전략 마련
연구 조건 기록	조사 시점, 응답자 속성, 조사 방식, 설계 기준, 샘플링 등	조사 시기, 대상자 속성 정보, 조사 방식, 변수 정의, 샘플링 방식	조사 설계 정합성 보존 데이터 해석 기준 및 맥락 제공
도구 및 소프트웨어	대개 분석 도구를 명시하지 않음, 문항지, 코드북, 지표 설명서 등 분석의 기반 자료를 제공	설문 문항 구성, 코드, 지표구성 항목, 분석 도구(선택)	분석 기반 문서 확보 수치 해석 가능성 보장
연구 절차 설명	문헌 조사→설계→조사→분석→해석이라는 구조적 흐름 유지, 설계와 분석의 연계성 강조	연구 절차, 설계목적, 지표 개발 단계, 조사 일정표 등	설계기반 재현성 확보 절차적 일관성 확보
지침상의 요구 항목	조사자료 등록 시 설계서, 조사표, 메타데이터 서식 작성 필수, 비식별화 조치, IRB 심의, 이용자 서약 등 법적 절차 강조	조사 목적, 비식별화 여부, 책임자, IRB 심의, 자료등록서식, 이용 제한 조건 등	민감 정보 보호 → 윤리성 확보 법적 책임 대응 공공 공개 기준 충족
보존 및 접근성 조치	조사자료 5년 이상 보존, 외부 공개 시 라이선스 명시, 접근등급 구분, 메타데이터와 원자료의 분리 보관, 외부 요청 시 서약서 요구 등	보존 기간 접근 권한, 외부 이용 조건 공개 여부, 폐기 기준, 메타데이터 연계 정보	장기보존 필요성 및 활용성 확보 접근 통제 및 투명성 확보

재현성 개념, 관리 방식에 있어 상이한 요구 사항을 지니고 있음을 확인할 수 있었다. 본 장에서는 이러한 분야별 차이점을 바탕으로 실무에 적용 가능한 연구데이터 기록관리 메타데이터 설계 방법을 논의하고자 한다. 연구데이터 기록관리 메타데이터 설계를 위한 방안으로는 크게 두 가지 접근 방법이 고려될 수 있다.

첫 번째는 분야별 단일형 설계 방식이다. 이 방식은 과학기술 분야와 인문사회 분야의 메타데이터를 각각 별도로 설계함으로써, 각 분야의 특수성을 보다 정밀하게 반영할 수 있다는 장점이 있다. 그러나 세부분야별로 메타데이터 요구 사항이 다양하게 존재할 경우, 모든 세부 분야에 맞춘 별도의 메타데이터를 일일이 설계해야 하므로 설계 및 운영 측면에서 확장성과 효율성이 저하될 수 있다. 특히, 메타데이터 구조의 중복 설계와 관리 자원의 낭비가 발생할 우려가 있어, 실제 적용 및 유지보수 단계에서 비효율적일 수 있다.

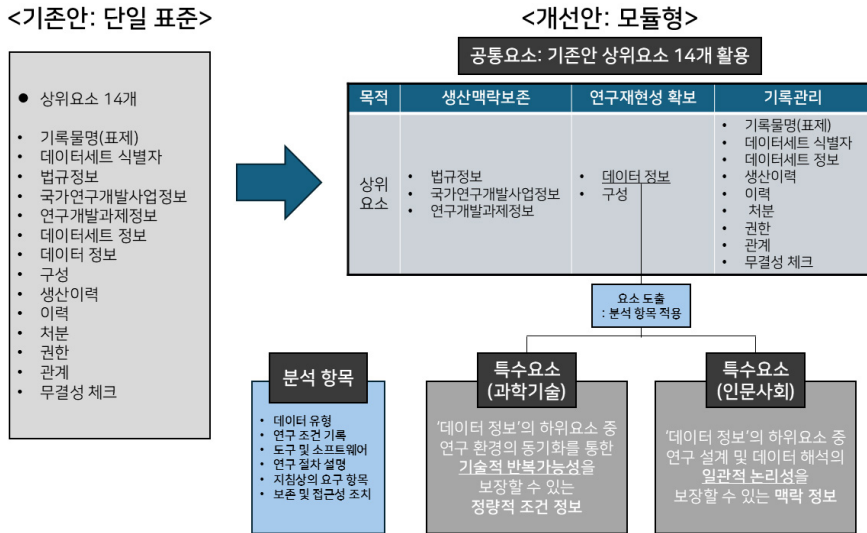
두 번째는 모듈형 설계 방식이다. 이 방식은 과학기술 분야와 인문사회 분야에서 공통적으로 요구되는 메타데이터 요소는 하나의 통합된 기본 구조로 유지하되, 분야별 특성을 반영하는 특수 요소는 모듈(Module)로 분리하여 설계하는 것이다. 모듈형 방식은 공통 요소를 재사용함으로써 설계 및 관리의 일관성을 확보하고, 중복 작업을 최소화할 수 있다는 점에서 효율성이 높다. 또한, 기관별 업무 특수성으로 인해 세부분야별로 새로운 요구 사항이 발생하더라도 모듈 단위로 메타데이터 요소의 수정 및 추가가 가능하므로 유연성과 확장성 측면에서 상대적으로 우수하다. 시스템 전체에 영향을 주지 않고 개별 요구를 반영할 수 있어, 실무에

유연하게 적용할 수 있다.

각 기관은 소속 분야의 특성과 연구 목적, 시스템 환경, 정책적 요구 등을 고려하여 메타데이터 설계 방식을 유연하게 선택할 필요가 있다. 그러나 분야별로 요구되는 기록관리 메타데이터 요소가 실질적으로 상이하다는 점을 고려할 때, 특정 분야에 특화된 단일형 메타데이터를 별도로 설계하기보다는 공통 요소와 특수 요소를 분리한 모듈형 설계 방식이 보다 합리적이라 판단된다.

모듈형 설계 방식이 가지는 구체적인 장점은 다음과 같다. 첫째, 개발 효율성 측면에서 여러 팀이 독립적인 모듈을 병렬적으로 설계하고 동시에 개발할 수 있어 시간과 자원을 절약할 수 있다. 둘째, 유지보수의 용이성이 높아 각 모듈의 오류 수정이나 업데이트 시 전체 시스템에 미치는 영향을 최소화할 수 있다. 단일형 설계 방식을 사용할 시, 메타데이터 오류 수정 및 업데이트가 시스템 전체의 중단을 초래할 위험성이 상대적으로 높다. 셋째, 공통 요소를 통한 표준화 설계로, 서로 다른 시스템 간 메타데이터의 호환성을 높일 수 있다. 기관 간 데이터 공유 및 통합이 용이하기 때문에 상호운용성 측면에서 강점을 가진다. 이와 같은 이유로, 본 연구는 연구데이터 기록관리 메타데이터의 분야별 설계를 위한 가장 적합한 방안으로 모듈형 설계 방식을 제안한다. <그림 2>는 기존의 연구데이터 기록관리 메타데이터 설계안과 새롭게 제안하는 개선안을 비교하여 나타낸 이미지이다.

모듈형 방식을 통해 실제로 연구데이터 기록관리 메타데이터를 설계할 시, 구체적인 요소명은 금효진과 양동민(2024)의 기존 연구에서도 출현한 '연구데이터 기록관리 메타데이터 요소



〈그림 2〉 연구데이터 기록관리 메타데이터 설계 개선안

최종 도출(안)의 요소와 〈표 11〉, 〈표 12〉에서의 예시를 활용할 수 있다. 기존 연구의 연구데이터 기록관리 메타데이터의 상위요소는 총 14개로 '기록물명(표제)', '데이터세트 식별자', '법규정보', '국가연구개발사업정보', '연구개발과제정보', '데이터세트 정보', '데이터 정보', '구성', '생산이력', '이력', '처분', '권한', '관계', '무결성 체크'가 있으며, 모두 개선안에서 공통 요소로 활용되었다.

'법규정보', '국가연구개발사업정보', '연구개발과제정보'는 연구기록물로서 생산맥락을 보존하기 위함이며, '기록물명(표제)', '데이터세트 식별자', '데이터세트 정보', '구성', '생산이력', '이력', '처분', '권한', '관계', '무결성 체크'는 기록관리 관점을 반영하기 위하여 공통 요소로 지정하였다.

'데이터 정보'와 '구성'은 모두 연구의 재현성 확보를 위한 공통 요소로, 이 중에서도 특히 연구데이터의 분야별 특성이 가장 명확하게 드러

나는 상위요소는 '데이터 정보'이다. '수집 정보', '분석 정보', '포맷', '메타데이터 표준'과 같은 하위요소를 포함하고 있어 해당 연구데이터의 직접적인 정보를 가장 많이 담고 있기 때문이다. 분야별 연구데이터의 재현성은 연구 수행 과정 내에서의 수집, 실험, 조사, 분석의 기계적 반복 가능성 또는 해석 맥락의 지원을 통해 확보되기 때문에, 재현 가능성의 보장을 위해서는 이러한 하위요소의 구분이 필수적이다. 따라서 모듈형 설계 시 분야별 특수 요소로 지정되는 것은 '데이터 정보'의 하위요소 및 세부 요소가 적합하다고 판단되며, 특히 기관별 사례 분석 항목 기준 중 '연구 조건 기록', '도구 및 소프트웨어 정보', '연구 절차 설명'에 해당하는 하위요소에 해당한다.

본 연구에서는 단순 개별 사례 조사에 그치지 않고, 선행연구 및 기존 메타데이터 표준을 토대로 도출한 6가지 분석 항목 기준을 활용하여 '데이터 정보'의 하위요소를 분류하고 각 기

준을 충족하는 새로운 하위요소 및 세부요소를 추가함으로써 기존안보다 연구 재현성 확보를 강화하고자 하였다. 이 분석 기준은 특정 분야나 데이터 형태에 종속되지 않고, 연구데이터의 재현성, 장기보존, 활용이라는 기록관리 목적을 중심으로 설정되었으므로 정성적 또는 혼합형 데이터를 포함하는 인문학, 예술, 교육학 등 인문사회 분야 내의 다양한 세부 영역에도 적용할 수 있다. 세부분야에 따라 도출되는 특수 요소는 달라질 수 있으나, 분석 항목 자체는 공통적으로 기능할 수 있으므로 범용적인 분석 틀로서 차후 활용성이 있다.

예를 들어, 시각자료 기반 연구에서는 ‘도구 및 소프트웨어 정보’ 항목에서 시각화 환경과 파일 포맷이 중요 요소로 작용하여 특정 요소가 도출될 수 있으며, 질적 인터뷰 데이터의 경

우 ‘연구 조건’이나 ‘연구 절차’ 항목에서 인터뷰 구조 및 상황 맥락의 기술과 관련된 요소가 핵심 요소로서 다뤄질 수 있다. 즉, 특정 세부분야 또는 데이터 유형 분석을 통해 새로운 요소가 추가되더라도 6가지 분석 항목 기준 내에 포함된다.

따라서 본 장에서 제시하는 메타데이터 설계 방안은 본 연구에서 분석한 4개 기관 사례에만 국한된 고정형 모델이 아니라, 동일한 분석 틀을 적용함으로써 다양한 연구 분야의 특수성을 반영한 모듈형 설계를 도출할 수 있는 구조적 유연성을 갖추고 있다. 다음의 <표 13>은 연구데이터 기록관리 메타데이터의 ‘데이터 정보’ 상위요소를 분석 항목 기준에 따라 분류하고, 공통 요소에 해당하는 하위요소 및 세부요소를 나타낸 것이다.

<표 13> 상위요소 ‘데이터 정보’ 중 공통 요소에 해당하는 하위요소 및 세부요소

분석 항목 기준	공통 요소		분석 항목 기준	공통 요소		
	하위요소	세부요소		하위요소	세부요소	
(기본 정보)	데이터명		도구 및 소프트웨어	(분야별 특수 요소)		
	데이터설명					
	데이터크기					
데이터 유형	자료유형	(정량/정성구분)	연구 절차 설명	연구절차		
	포맷	포맷명	지침상의 요구 항목	IRB심의	심의 여부 승인번호 심의기관명	
		포맷버전				
		구동SW및버전				
(분야별 특수 요소)						
연구 조건 기록	생산목적		보존 및 접근성 조치	백업	백업유무	
	생산주체	책임자 기여자		언어	백업위치	
	생산범주	(실험/조사/분석/수집/관측 등)		메타데이터표준		
	생산정보	생산일시			암호화데이터 사용유무	
		생산주기			연계시스템 유무	
		생산범위				
		생산방법				

기존안의 '데이터 정보' 하위요소를 기관별 사례 분석 항목에 따라 구분한 결과, 기존안의 하위요소를 그대로 적용하여도 기준을 충족할 수 있는 경우는 기존안의 요소만을 활용하였으며, 기존안만으로는 기준을 충족하는 하위요소가 부족하여 연구 재현성의 확보가 부족하다고 생각될 경우 신규 요소를 추가하는 방식으로 공통 요소를 조정하였다. 조정 결과를 세부적으로 정리한 결과는 다음과 같다.

- '데이터명', '데이터설명', '데이터크기': 분석 항목 기준으로 구분되지는 않지만, 데이터의 기본적인 정보를 담고 있어 공통 요소에 포함함
- 데이터유형: 기존 안의 '포맷' 하위요소를 활용함. '자료형식' 하위요소를 새롭게 추가하여 정량, 정성, 혼합 연구데이터의 형식을 명확히 하고 장기보존 전략을 마련하기 위한 정보를 제공하고자 함
- 연구 조건 기록: 기존안에서는 연구 조건 및 과정 정보를 확보하기 위하여 하위요소로 '수집 목적', '수집 주체', '수집 유형', '수집 정보', '분석 목적', '분석 주체', '분석 정보'를 포함하고 있었으나 이것은 과학기술 분야에 다소 편중된 경향이 있어, 인문사회 분야의 연구 조건 기록에는 적합하지 않았음. 또한, 데이터 생산 방법에 따라 목적, 주체 등을 모두 따로 기록해야 하는 불편함이 있었음. 이에 기존안의 하위요소를 삭제하고, 신규 요소인 '생산목적', '생산주체', '생산범주', '생산정보'를 추가함. '생산정보'의 경우 '생산일시', '생산주기', '생산범위', '생산방법' 세부요소를 추가하여

분야와 관계없이 기본적인 연구데이터 정보를 기술할 수 있도록 조치함. 이 외의 세부요소는 분야간 재현성 확보 방식에 따라 요구되는 요소가 달라, 해당 기준을 만족하는 요소는 특수 요소로 별도 구분함

- 도구 및 소프트웨어: 과학기술 분야와 인문사회 분야 간 재현성 확보 방식에 차이가 있어, 해당 기준을 만족하는 하위요소는 특수 요소로 별도 구분함
- 연구 절차 설명: '연구절차', '연구일정' 하위요소를 추가함
- 지침상의 요구 항목: 과학기술 분야와 인문사회 분야간 재현성 확보 방식에 차이가 있어, 해당 기준을 만족하는 하위요소는 특수 요소로 별도 구분함. 단, IRB 심의의 경우 과학기술 분야에서 실험, 임상, 유전체 연구 등을 실시하는 의생명과학 분야, 설문조사, 인터뷰 등을 시행하는 사회과학 분야에서 필수적으로 요구됨. 따라서 개인 식별 정보 및 민감 정보를 다루는 모든 연구는 분야를 불문하고 IRB 심의 절차를 요구받으므로 공통 메타데이터 요소로 추가함
- 보존 및 접근성 조치: 기존안의 하위요소인 '백업', '언어', '메타데이터표준', '암호화데이터사용유무', '연계시스템유무'를 활용함

본 연구에서는 기존안의 요소를 활용하되 조정 방안을 통해 재현성 확보 방안을 반영하지 못하는 요소를 공통 요소에서 제외하고, 분야별 특성에 부합하는 하위 요소는 별도로 구분하고자 하였다. 다음의 <표 14>는 이상의 논의를 토대로 연구데이터 기록관리 메타데이터의 공통 요소에 포함되는 상위요소, 하위요소, 세

〈표 14〉 연구데이터 기록관리 메타데이터의 공통 요소

요소			요소				
상위요소	하위요소	세부요소	상위요소	하위요소	세부요소		
기록물명(표제)	데이터세트명	대표데이터	생산이력	생산시스템	시스템명		
데이터세트 식별자	데이터식별자				시스템개요		
법규정보	법령/규정				구축년도		
국가연구개발사업정보	사업번호				DBMS정보		
	사업명				기관명		
연구개발 과제정보	과제번호			기관코드			
	과제명	(공동/단독)		부서명			
	과제형태	(진행중/종료)		부서코드			
	과제진행상태			담당자명			
	과제기간			담당자직책			
	과제설명		생산유형	(관리/이용/보존)			
	지원기관	기관명 지원기관 요구사항	행위유형				
	산출물	연구성과물	행위자	행위자명 행위자부서 행위자직책			
데이터세트 정보	분야	주제어	이력	행위일시	행위일시		
	주제	용량 단위			행위근거		
	크기	단위			행위설명		
일시	생산일시	행위결과					
	데이터명			변경요소	변경요소명 변경이전값		
	데이터설명			보존기간			
	데이터크기			보존기간 사유			
	자료유형	(정량/정성/혼합) 포맷명	처분일시				
	포맷	포맷버전 구동SW및버전	처분방법				
	생산목적		적용범위				
데이터 정보	생산주체	책임자 기여자	처분	비밀	비밀분류		
	생산범주	(실험/조사/분석 /수집/관측/설문 등)			비밀분류 근거		
	생산정보	생산일시				보호기간	
		생산주기				접근	접근범위 URL
		생산범위				공개구분	공개구분 비공개사유
	연구절차		공개	공개	공개제한부분		
	연구일정				공개제한일자		
	IRB 심의	심의 여부				공개관련근거	
		승인번호				목록공개구분	
	백업	백업유무 백업위치				목록공개제한사유	목록공개제한부분
언어		관련	라이선스	저자			
메타데이터표준				발행일			
암호화데이터				타이틀			
사용유무				버전			
연계시스템유무				출판사			
구성	첨부파일	첨부파일명	관계	관계유형	식별자		
		첨부파일종류			관계대상 식별자		
		첨부파일포맷 : 구동SW			관계설명		
	첨부파일저장방법		무결성체크	무결성검증법 무결성검증값			

부요소를 정리하여 나타낸 것이다. 셀 배경이 짙게 칠해진 요소는 기존에는 존재하지 않았으나, 본 연구에서 새롭게 추가된 요소를 의미한다. 본 표에서는 공통 요소 도출에 중점을 두어, 요소의 우선순위 표기는 제외하였다.

연구데이터의 분야와는 관계없이 기술되어야 하는 공통 요소와는 달리, 특수 요소의 경우 과학기술 분야와 인문사회 분야 각각에서 생성하는 데이터의 성격과 재현성 확보 방식이 다르기 때문에 구분되어야 하는 요소이다. 특히 연구 조건 기록과 도구 및 소프트웨어 정보에서 구체적인 메타데이터 요소가 상이하다. IRB 심의 요소의 경우, 연구데이터에 포함된 정보의 민감성, 개인정보 여부, 대상자의 권리 보호와 관련된 사항으로, 연구 분야와 무관하게 공통적으로 요구되는 윤리적 절차이다. 다만 인문사회 분야에서는 이러한 윤리 절차가 연구 설계 전반 및 데이터 수집 과정에 긴밀하게 연결되는 경향이 있으며, 이로 인해 메타데이터 항목 중 윤리적 승인 및 관련 정보의 상세 기술이 특히 강조되는 특성이 있다. 즉, 공통 요소이지만 적용 강도 및 기술 방식에서 분야별 특성이 나타날 수 있음을 고려해야 한다.

다음의 <표 15>는 앞서 논의한 모듈형 설계 방식의 논리를 바탕으로, <표 11>, <표 12>의 과학기술·인문사회 분야 메타데이터 요소 예시와 <표 13>의 '데이터 정보' 공통 요소를 활용하여 이를 실제 분야별 연구데이터 기록관리 메타데이터 설계에 적용했을 경우의 구성 예시를 나타낸 것이다. 배경에 음영이 있는 셀은 해당 요소가 특수 요소임을 의미한다. 해당 표는 '데이터 정보' 상위요소에 해당하는 기록관리 메타데이터 작성 예시를 공통 요소와 분야별

특수 요소로 구분함으로써 모듈형 설계의 구조와 개념을 보다 구체적으로 이해할 수 있도록 하였으며, 본 표에서는 설계 구성의 이해에 중점을 두고자 요소의 우선순위 표기는 제외하고, 구체적인 예시를 추가하였다.

연구 분야에 관계없이 기술되어야 하는 공통 요소와는 달리, 특수 요소는 각 분야에서 생성하는 데이터의 성격과 재현성 확보 방식에 따라 그 값이 상이하다. 특히 세부분야 및 연구 조건 기록, 도구 및 소프트웨어 정보를 기준으로 구체적인 메타데이터 요소에 차이가 생길 수 있으며, 위 표는 과학기술 분야에서 수집을 통해 생성된 연구데이터와 인문사회 분야에서 설문을 통해 생성된 연구데이터를 비교함으로써 그 차이를 가시적으로 나타내하고자 하였다. 이러한 설계 방안은 공통 요소와 함께 분야별 특성에 부합하는 특수 요소를 모듈형으로 구성함으로써, 기록관리를 위한 필수 속성을 확보하는 동시에 연구 재현성 확보에 필요한 핵심 정보를 각 분야의 특성과 데이터 유형에 따라 유연하게 적용하고 보존할 수 있도록 지원한다. 다만, 공통 요소와 기관의 시스템 환경, 관리 인력의 수준, 연구 세부분야, 데이터 유형에 따라 설계 적용 방식은 조정이 필요할 수 있으며, 향후에는 이를 고려한 단계별 설계 가이드라인이 추가로 마련되어야 할 것이다.

5. 결론

공적자금이 투입된 국가연구개발과제의 최종 산출물뿐만 아니라 연구 수행 과정 중에 생성된 연구데이터는 주요 공공자산이며, 법률에 따라

〈표 15〉 분석 대상 기관 연구데이터 적용 예시: 공통 요소 및 분야별 특수 요소

과학기술 분야			인문사회 분야		
하위요소	세부요소	요소 설명 또는 예시	하위요소	세부요소	요소 설명 또는 예시
데이터명		화석(필석류) F-175	데이터명		청소년 인식조사 원자료
데이터설명		국토지질 시료	데이터설명		중학생 대상 글로벌 시민 의식 조사
데이터크기		1.00 mb	데이터크기		△△ mb
자료유형		정량	자료유형		혼합
포맷	포맷명	JPG	포맷	포맷명	SPSS(.sav)
생산목적		지질자원주제도 작성을 위한 시료 수집	생산목적		글로벌 시민의식 정책 기반 자료 작성 목적
생산주체	책임자	홍○○	생산주체	책임자	홍○○
	기여자	박○○, 최○○		기여자	김○○
생산범주		수집	생산범주		설문
생산정보	생산일시	2019.12.31	생산정보	생산일시	2022.05.01
	생산주기	단발성(1회)		생산주기	매 3년 주기 반복조사
	생산범위	-		생산범위	-
	생산방법	Manual(수작업)		생산방법	온라인 설문을 통한 조사
	수집조건	하부 오르도비스 산출		조사설계	증화확률추출 방식
	샘플크기	7.3		지표구성	○○ 지표는 5개 문항으로 구성하여 7척도 사용 ...
	샘플크기단위	cm		변수정의	'가구소득', '학업스트레스', ...
	샘플유형	화석 표본		대상자속성	응답자의 연령, 성별, 거주지, 교육수준 등 인구통계학적 정보
샘플특징	그물 모양 또는 방사성이거나 선형으로 균락을 이룸	샘플링방식	증화추출, 표본오차 기재, ...		
장비	장비정보	-	설문조사	설계목적	교육격차 실태 파악 및 정책 개선 기초 자료
	장비설정값	-		문항구성	인적사항, ...
소프트웨어	소프트웨어정보	○○	분석도구	도구정보	SPSS
	버전	1.0		도구설정값	-
	사용목적	이미지화		사용목적	통계처리
연구절차		첨부파일 참조	연구절차		선행연구 조사, 지표 초안 구성, 전문가 검토 및 정제, 지표 확정, ...
연구일정		첨부파일 참조	연구일정		2022.05.01. 1차 설문, 2022.05.10. 2차 설문, ...
IRB 심의	심의 여부	-	IRB 심의	심의 여부	승인
	승인번호	-		승인번호	제2022-IRB-000호
	심의기관명	-		심의기관명	-
SOP		SOP번호			
			비식별화여부		비식별화진행
			데이터공개조건		외부 신청 시 제공

행정정보 데이터세트의 일부로서 관리되어야 한다. 연구데이터가 기록으로서 연구적, 교육적, 재정적, 증거적 가치를 보존하기 위해서는 연구의 재현성과 검증 가능성의 확보가 중요하며, 이에 따라 연구데이터 기록관리 메타데이터의 설계는 각 연구의 재현성을 극대화할 수 있는 방향으로 이루어져야 한다.

과학기술 분야와 인문사회 분야의 연구데이터는 분야별 특성에 따라 재현성 개념이 다르며, 이는 필요 메타데이터 요소와 연구데이터 관리 요구 사항의 변동까지 불러올 수 있는 중대한 사안이다. 그러나 기존 연구에서 설계한 연구데이터 기록관리 메타데이터는 과학기술 분야의 재현성 개념에 상대적으로 더 집중하여, 인문사회 분야의 연구데이터에 적용하기 어려움이 있다는 한계가 존재하였다. 기록관리 메타데이터는 연구데이터 기록관리에 있어서 연구데이터의 재사용성, 보존성, 접근성에 큰 영향을 미치기 때문에 기존 설계안의 한계점을 극복할 필요가 있다.

따라서 본 연구는 연구데이터 기록관리 메타데이터의 분야별 설계를 위한 기초가 될 수 있는 기준과 요소들을 도출하여 향후 연구데이터 관리 및 정책에 기여하고자 했다. 이를 위해 각 분야별 대표 기관을 선정하고 대표 기관의 최종 보고서를 비롯한 연구성과물과 기관 내부 연구데이터 관리지침 및 메뉴얼을 분석하였다. 두 가지 자료 유형은 각각 연구데이터의 생산 실태, 제도적 관리 기준, 그리고 실무적 적용 방식을 보여주는 핵심 자료이므로, 분석대상 연구기관이 어떤 방식으로 연구데이터를 생성·기술·보존하고 있는지를 체계적으로 드러낸다. 특히 지침 등의 규정은 데이터 형성 및 구조에 직간

접적인 영향을 끼칠 수 있기 때문에, 각 자료가 담고 있는 정보를 분석하여 기관별 연구데이터 특성을 파악할 수 있었다. 두 가지 자료 유형의 분석을 통해, 국내 기관이 분야별 연구 재현성을 확보하기 위해 무엇을 필요로 하는지와 연구데이터 관리 요구 사항을 확인하였다.

최종적으로는 기관 분석 내용과 기존 연구의 연구데이터 기록관리 메타데이터 설계 원칙을 모두 반영하여 과학기술 분야와 인문사회 분야 각각에 적합한 연구데이터 기록관리 메타데이터 요소를 파악하였다. 과학기술 분야는 실험 환경의 완전한 복원을 통한 정량적 수치를 통제 및 반복 가능성을 재현성으로 간주한다. 이에 따라 과학적 과정 자체의 동기화를 전제하기 위하여 '환경 변수', '실험 조건', '장비 정보'와 같은 메타데이터 요소를 필요로 하였다. 인문사회 분야의 재현성 개념은 해석 가능한 설계 맥락의 반복 가능성으로 정의할 수 있다. 수집한 데이터 자체만으로는 그 연구적 의미가 희미하므로, 어떠한 맥락과 논리 구조에서 해당 데이터가 구성되었는지를 나타낼 수 있는 절차적·맥락적 정보가 필요하다. 이에 따라 '변수정의', '설계목적'과 같은 메타데이터 요소가 필수적이다.

이는 기존 연구의 한계를 보완하여 연구데이터의 장기보존 및 책임 있는 관리 체계를 강화하는데 기여할 수 있으며, 더 나아가 모듈형 설계 방법론을 제안하고 기관 사례에 기반한 요소 설계 예시를 보임으로써 현실적 적용 가능성과 선택의 유연성을 확보하였다는 점에서 연구 의의가 있다.

본 연구에서 제시한 '분석 항목' 기준은 소수 기관을 대상으로 도출된 결과이지만, 특정 기

관에 국한되지 않고 다양한 분야와 데이터 유형에 범용적으로 적용 가능한 분석 체계로 설계되었다. 다만, 본 연구는 몇몇 대표 기관의 사례 분석에 제한되었기 때문에, 향후 연구에서는 분석대상 기관의 범위를 확대하고, 다양한 연구 분야 및 비정형 데이터 유형에 본 분석 틀

이 어떻게 적용될 수 있을지에 대한 논의가 추가로 이루어져야 한다. 아울러, 메타데이터 설계의 실질적 유효성을 확보하기 위해 전문가 평가, 실제 연구데이터 적용, 사용자 기반 후속 연구 등이 병행되어야 할 것이다.

참 고 문 헌

과학기술정보통신부 (2025). 2025년도 국가연구개발혁신법 매뉴얼.

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>

국가기술표준원 (2020). TTAK.KO-10,0976: 연구데이터 메타데이터 요소. 한국정보통신기술협회 (TTA).

국가연구개발정보처리기준. 과학기술정보통신부고시 제2020-102호.

국가연구개발혁신법. 법률 제20057호.

금효진, 양동민 (2024). 연구데이터의 장기보존을 위한 기록관리 메타데이터에 관한 연구. 정보관리학회지, 41(4), 179-207. <http://doi.org/10.3743/KOSIM.2024.41.4.179>

김로사, 장우권 (2016). 연구관리기록물 관리 매뉴얼 연구: 과학기술분야 정부출연연구기관을 중심으로. 한국기록관리학회지, 16(3), 179-207. <http://doi.org/10.14404/JKSARM.2016.16.3.179>

김상준 (2008). 연구기록물 중 연구노트의 중요성과 관리 제도화. Journal of Information Science Theory and Practice, 39(2), 45-74. <http://doi.org/10.1633/JIM.2008.39.2.045>

김수진, 정은경 (2012). 정부출연연구기관의 연구기록물 관리를 위한 수집 개선방안에 관한 연구. 한국비블리아학회지, 23(4), 109-124. <http://dx.doi.org/10.14699/kbiblia.2012.23.4.109>

김주섭, 한연중, 유원재, 김선태 (2020). 임산공학 분야 연구데이터 관리를 위한 메타데이터 설계에 관한 연구. 한국문헌정보학회지, 54(4), 169-194. <http://doi.org/10.4275/KSLIS.2020.54.4.169>

김혜미, 이영학 (2019). 연구관리 전문기관 기록물관리 현황 및 개선 방안: 핵심 연구기록물을 중심으로. 한국기록관리학회지, 19(2), 173-191. <http://doi.org/10.14404/JKSARM.2019.19.2.173>

백재은 (2022). 산업유산자원의 이용·관리·보존을 위한 메타데이터 요소 설계에 관한 연구. 한국문헌정보학회지, 56(2), 231-254. <http://doi.org/10.4275/KSLIS.2022.56.2.231>

예상준, 장호, 김선태 (2019). 한의학 연구 데이터 관리 및 공유를 위한 메타데이터 요소 설계. 한국문

- 헌정보학회지, 53(2), 223-246. <http://doi.org/10.4275/KSLIS.2019.53.2.223>
- 질병관리청 (2023). miR-204 조절을 통한 알츠하이머 치료제 후보물질 개발.
- 질병관리청 (2024). 전분화능줄기세포 유래 분화세포의 유전체 안정성 및 증양원성 평가.
- 질병관리청 (2024). 전장 유전자 정보 기반 항원성 SARS-CoV-2 변이 예측 시스템 개발 및 기반 에피토프 특이적 항바이러스 HTS 후보물질 발굴.
- 질병관리청 연구노트 관리지침. 질병관리청예규 제157호.
- 최형욱, 정은경 (2017). 사회학 분야의 연구데이터 특성과 지적구조 규명에 관한 연구. 정보관리학회지, 34(3), 109-124. <https://doi.org/10.3743/KOSIM.2017.34.3.109>
- 한국교육개발원 (2024). 디지털 전환 시대 교육격차 변화 양상 연구 (2024-13).
- 한국교육개발원 (2024). 조사자료 관리 및 공개 지침.
- 한국교육개발원 (2024). 중등 글로벌 인문교양교육 교육지표 개발 연구 (2024-18).
- 한국교육개발원 (2024). 지역-대학 연계 강화를 위한 데이터 실태 분석 및 활용 방안 연구 (2024-23).
- 한국보건사회연구원 (2024). 일본의 인구감소 대응 정책 사례연구 (2024-04).
- 한국보건사회연구원 (2024). 조사자료 관리지침.
- 한국보건사회연구원 (2024). 한부모가족의 시간 및 경제적 자원과 정책 (2024-35).
- 한국보건사회연구원 (2024). 행복 및 삶의 질 측정과 정책 활용성 제고 방안 연구 (2024-36).
- 한국지질자원연구원 (2019). 화석(필석류(Graptolite, Dictyonema flabelliforme)_F-175). 지오빅 데이터 오픈플랫폼. 출처: <http://hdl.handle.net/10273/KGM2021000102>
- 한국지질자원연구원 (2023). 3차원 다공성 나노시트 촉매 제조 및 이의 바이오 플라스틱 분해 반응 연구.
- 한국지질자원연구원 (2023). 한국 해양분지 내 능동적 수소생산 가능성에 대한 예비연구.
- 한국지질자원연구원 (2024). 온실가스 방출원으로서의 한반도 대륙붕 해저메탄 자연분출 규명 및 정량화 연구.
- 한국지질자원연구원. (발행년불명) 연구데이터관리지침.
- 한국지질자원연구원. (발행년불명) 연구데이터보존지침.
- 한국지질자원연구원. (발행년불명) 연구데이터수집지침.
- 한국지질자원연구원. (발행년불명) 윤리 · 저작권 · 라이선스지침.
- DataCite (2021). DataCite Metadata Schema 4.4. Available: <https://schema.datacite.org/meta/kernel-4.4/>
- International Organization for Standardization (2016). ISO 15489-1:2016 Information and documentation - Records management - Part 1: Concepts and principles.
- Joyeux-Prunel, B. (2024). Digital humanities in the era of digital reproducibility: towards a fairest and post-computational framework. *Int J Digit Humanities* 6, 23-43.

<https://doi.org/10.1007/s42803-023-00079-6>

Leipzig, J., Nüst, D., Hoyt, C. T., Ram, K., & Greenberg, J. (2021). The role of metadata in reproducible computational research. *Patterns* (New York, N.Y.), 2(9), 100322.

<https://doi.org/10.1016/j.patter.2021.100322>

National Academies of Sciences, Engineering, and Medicine (2019). *Reproducibility and Replicability in Science*. Washington, DC: The National Academies Press.

<https://doi.org/10.17226/25303>.

Peels, R. (2019). Replicability and replication in the humanities. *Res Integr Peer Rev* 4, 2.

<https://doi.org/10.1186/s41073-018-0060-4>

• 국문 참고자료의 영어 표기

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

Baek, Jae-Eun (2022). A study on metadata schema development for the use, management and preservation of industrial heritage resources. *Journal of the Korean Society for Library and Information Science*, 56(2), 231-254. <http://doi.org/10.4275/KSLIS.2022.56.2.231>

Choi, Hyong-Wook & Chung, Eunkyung (2017). An investigation on characteristics and intellectual structure of sociology by analyzing cited data. *Journal of the Korean Society for Information Management*, 34(3), 109-124. <https://doi.org/10.3743/KOSIM.2017.34.3.109>

Geum, HyoJin & Yang, Dongmin (2024). Research on record management metadata for long-term preservation of research data. *Journal of the Korean Society for Information Management*, 41(4), 179-207. <http://doi.org/10.3743/KOSIM.2024.41.4.179>

KIGAM (2019). Graptolite, *Dictyonema flabelliforme*_F-175. Geo Big Data Open Platform. Available: <http://hdl.handle.net/10273/KGM2021000102>

Kim, Hye-mi & Lee, Young-Hak (2019). Current status and improvement plan for records management at research and management institutions: focused on core research records. *Journal of Korean Society of Archives and Records Management*, 19(2), 173-191. <http://doi.org/10.14404/JKSARM.2019.19.2.173>

Kim, Ju-Seop, Han, Yeonjung, Youe, Wonjae, & Kim, Sun-Tae (2020). A study on the design of metadata for research data management in forestry engineering. *Journal of the Korean Society for Library and Information Science*, 54(4), 169-194. <http://doi.org/10.4275/KSLIS.2020.54.4.169>

Kim, Rosa & Chang, Woo-kwon (2016). A study on the record management manual for R

- & D management: a focus on government: supported science research institutes. *Journal of Korean Society of Archives and Records Management*, 16(3), 179-207.
10.14404/JKSARM.2016.16.3.179
- Kim, Sang-Jun (2008). Importance and management of the laboratory notebooks in the research record. *Journal of Information Science Theory and Practice*, 39(2), 45-74.
<http://doi.org/10.1633/JIM.2008.39.2.045>
- Kim, SuJin & Chung, Eunkyung (2012). Toward research records management in government-funded research institutes. *Journal of the Korean Biblia Society for Library and Information Science*, 23(4), 109-124. <http://dx.doi.org/10.14699/kbiblia.2012.23.4.109>
- 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>
- Korea Disease Control and Prevention Agency (2023). Development of Candidate Substances for Alzheimer's Disease Treatment by Regulation of miR-204.
- Korea Disease Control and Prevention Agency (2024). Development of The Genome-Based Serological System for in Advance Tracing of SARS-CoV-2 Antigenic Evolution and Epitope-Specific Antiviral Candidates Using High Content Analysis Method.
- Korea Disease Control and Prevention Agency (2024). Genetic Abnormality and Occurrence of Abnormal Tissue after Transplant of Pluripotent Stem Cell Derivatives.
- Korea Disease Control and Prevention Agency Research Note Management Guidelines. Korea Disease Control and Prevention Agency Regulations No. 157.
- Korea Institute For Health And Social Affairs (2024). A Case Study on Japan's Policy Responses to Population Decline (2024-04).
- Korea Institute For Health And Social Affairs (2024). Guidelines for Managing Research Data.
- Korea Institute For Health And Social Affairs (2024). Measuring Happiness and Quality of Life for Social Policy-Making (2024-36).
- Korea Institute For Health And Social Affairs (2024). Time and Economic Resources of Single-Parent Families and Policies (2024-35).
- Korea Institute of Geoscience and Mineral Resources (2023). A Preliminary Study of Active Hydrogen Production Potential from Korean Marine Basins.
- Korea Institute of Geoscience and Mineral Resources (2023). Preparation of 3D Porous Nanosheet Catalyst and Its Bioplastic Degradation.

- Korea Institute of Geoscience and Mineral Resources (2024). A Study on The Identification and Quantification of Methane as A Source of Greenhouse Gas Venting from The Seafloor of The Continental Shelf of The Korean Peninsula.
- Korea Institute of Geoscience and Mineral Resources. (n.d.) Ethics · Copyright · Licensing Guidelines.
- Korea Institute of Geoscience and Mineral Resources. (n.d.) Research Data Collection Guidelines.
- Korea Institute of Geoscience and Mineral Resources. (n.d.) Research Data Management Guidelines.
- Korea Institute of Geoscience and Mineral Resources. (n.d.) Research Data Preservation Guidelines.
- Korean Agency for Technology and Standards (2020). TTAK.KO-10,0976: The Integrated Metadata for The Scientific Data. TTA.
- Korean Educational Development Institute (2024). A Study on the Data Status and Utilization Strategies to Strengthen Regional-University Collaboration (2024-23).
- Korean Educational Development Institute (2024). A Study on the Development of Educational Indicators for Secondary Liberal Arts and Sciences Education at the Global Level (2024-18).
- Korean Educational Development Institute (2024). Changing Dynamics of Educational Inequality in the Digital Transformation Era: An Analytical Study (2024-13).
- Korean Educational Development Institute (2024). Guidelines for Managing and Disclosing Research Data.
- Ministry of Science and ICT (2025). National Research and Development Innovation Act Manual. National Research and Development Information Processing Standards. Ministry of Science and ICT Public notice No. 2020-102.
- National Research and Development Innovation Act. Act No. 20057.
- Yea, Sang-Jun, Jang, Ho, & Kim, Sun-Tae (2019). Metadata element design for korean medicine research data management and re-use. Journal of the Korean Society for Library and Information Science, 53(2), 223-246. <http://doi.org/10.4275/KSLIS.2019.53.2.223>