

기계공학분야 연구자들의 연구데이터 생산과 관리에 관한 연구

A Study on Research Data Creation and Management by Researchers in Mechanical Engineering

박윤미(Yunmi Park)¹, 김지현(Jihyun Kim)²

E-mail: pympym1203@gmail.com, kim.jh@ewha.ac.kr



¹ 제 1저자 한국기계연구원, 이화여자대학교 일반대학원 문헌정보학과 기록관리학전공 박사과정
² 교신저자 이화여자대학교 사회과학대학 문헌정보학과 부교수

논문접수 2021-10-19
최초심사 2021-10-20
게재확정 2021-11-15

ORCID

Yunmi Park
https://orcid.org/0000-0003-4714-5728
Jihyun Kim
https://orcid.org/0000-0002-3087-3553

© 한국기록관리학회

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.

초 록

본 연구는 기계공학 분야 연구자들의 연구데이터 생산과 관리에 대한 인식과 경험을 조사하고 해당 분야의 연구데이터 관리와 서비스를 위한 시사점을 제안하는 것을 목적으로 한다. 국내외 연구기관의 연구데이터 관리 및 서비스에 대해 알아보고, 국내 기계공학 분야 연구기관의 소속 연구자들을 대상으로 심층 면담을 실시하여 ‘연구데이터, 책임있는 연구수행과 연구윤리 준수, 연구데이터 관리의 효율성 및 효과성, 연구데이터 공유의 가치’ 등 4개의 주요범주로 연구데이터 생산과 관리에 대한 인식과 행태를 분석하였다. 기계공학분야 연구데이터 관리와 서비스를 위해서는, 생산과정과 유형, 형태에 대한 데이터 조사를 실시하여 명시적 메타데이터와 암시적 맥락정보의 수집이 필요하며 데이터학술지에 데이터 논문을 출판함으로써 연구실적으로 인정하는 방안을 제안하고 안전한 데이터 관리와 연구자들간의 소통을 지원하는 클라우드 기반 시스템 등의 인프라 마련이 필요하다. 또한 연구 현장의 다양한 관계자들이 조직적 차원의 연구데이터 관리와 서비스에 대한 역할과 책임을 배분하는 것이 중요함을 제안하였다.

ABSTRACT

This study aimed to examine the perception and experience of researchers in the field of mechanical engineering on research data creation and management, and suggest implications for research data management and services in the field. Research data management and services of domestic and foreign research institutes were investigated, and in-depth interviews were conducted with researchers belonging to domestic mechanical engineering research institutes to analyze the perception and conduction of research data creation and management according to four major categories: “research data, accountable conducting of research and compliance with research ethics, utility and effectiveness of research data management, and the value of sharing research data.” To ensure effective research data management and services in mechanical engineering, it is necessary to conduct a data investigation on the process, type, and form of production to collect explicit metadata and implicit contextual information. It is also necessary to propose a plan to recognize research results using the publication of data journals and to prepare infrastructure such as a cloud-based system that supports safe data management and communication between researchers. In addition, it suggests that it is important for various officials in the research field to allocate roles and responsibilities for research data management and services at the organizational level.

Keywords: 연구데이터, 기계공학, 연구데이터 관리, 데이터 생산, 데이터 공유
research data, mechanical engineering, research data management,
data creation, data sharing

<https://jksarm.koar.kr>

1. 서론

1.1 연구의 목적

법적·제도적·기술적 장벽 없이 모든 이용자에게 연구 수행과정에서의 다양한 형태의 결과물들에 대한 접근이 가능하도록 제공하는 오픈 사이언스(Open Science) 운동이 전 세계적으로 전개되고 있다. 오픈 사이언스는 연구자, 정부기관, 연구비지원기관 혹은 과학 커뮤니티를 중심으로 공적 자금이 투입된 연구의 주요 결과물인 논문 등의 출판물과 연구데이터를 디지털 형태로 누구나 이용 가능하도록 공개하는 활동을 의미한다. 이는 연구 활성화를 목적으로 이루어지는 활동이며 연구의 투명성과 협력을 강화하고 혁신을 촉진하는데 주안점을 두고 있다(OECD, 2015, 7). 이러한 관점에서 오픈 사이언스는 연구출판물의 접근성을 높이는 오픈 액세스, 연구데이터 접근과 이용을 지원하는 오픈 데이터 뿐만 아니라 개방형 연구협력과 소통을 강조하는 오픈 콜라보레이션(open collaboration)까지 포괄하는 활동으로 정의된다. 국내의 오픈 사이언스 환경은 오픈 액세스 논문이 점차 증가하는 추세에 있다는 점에서 긍정적이나, 오픈 데이터를 위한 연구데이터 관리와 공유의 기반이 취약하고 온라인 플랫폼을 통한 연구협력에 대한 논의 역시 부족한 실정이다(신은정 외, 2017).

특히 오픈 데이터는 연구재현성 확보와 학문 발전의 필수 조건이면서 이를 통해 인용 등 연구 영향력이 향상될 수 있다는 측면에서 가장 중요한 오픈 사이언스 영역 중 하나로 간주되고 있다(Burgelman et al., 2019). 해외에서는 오픈 데이터의 실행을 의미하는 연구데이터 공유, 그리고 이를 지원하는 연구데이터 관리에 대해 연구자들의 인식과 관행을 조사한 연구가 다수 수행되어 왔다(정지윤, 이강현, 강혜원, 2020). 또한 Fecher, Friesike, Hebing (2015)가 수행한 연구데이터 공유에 대한 체계적 문헌고찰에 따르면 조사대상 문헌의 약 60%에서 과학·기술·공학·수학(STEM) 분야의 데이터 공유를 논의하고 있는 것으로 나타났다. 따라서 STEM 분야를 중심으로 데이터 관리와 공유에 대한 연구가 주로 이루어져 왔음을 알 수 있다.

그러나 STEM 분야 중에서도 공학 분야 연구자들은 전통적으로 산업계와 긴밀히 협력하면서 상업적인 이익을 추구하는 연구를 수행하는 경우가 많아 데이터 공유를 고려하지 않는 경향이 높다(Research Data Alliance, 2019). 이와 같이 다소 폐쇄적인 연구 문화를 가지는 공학 분야에서도 체계적인 연구데이터 관리는 반드시 필요하다는 인식이 높아지고 있으며 국제적인 연구데이터 연맹인 Research Data Alliance(RDA)에서는 2019년 공학 분야의 연구데이터 관리를 위한 관심 그룹(interest group)을 결성하여 이를 지원하고 있다. 또한 인공지능과 데이터 기반으로 운영되는 제조 기술에서의 혁신을 의미하는 Industry 4.0을 중심으로 스마트 공장 등 지능형 설비와 제조 과정에 대한 다양한 논의가 전개되면서 산업계에서의 데이터 공유 필요성이 강조되고 있다(Müller, Veile, & Voigt, 2020). Industry 4.0과 밀접히 연관된 기계공학 분야에서도 데이터 관리 및 교환과 공유를 촉진하는 플랫폼 개발과 빅데이터 관리 및 분석에 대한 교육이 필요하다는 주장이 제기되고 있다(Fernández-Miranda et al., 2017). 생명과학이나 사회과학 분야에 비해 공학 분야에서의 연구데이터 공유에 대한 인식을 조사한 연구는 드물게 이루어져 왔다(Joo & Kim, 2017). 해외에서는 재료공학이나 지진공학 등 일부 분야의 연구자 인식을 조사한 연구(Suhr, Dungal, & Stocker, 2020; Wu & Worrall, 2019)와 영국과 노르웨이의 공학자들을 대상으로 한 질적 연구(Mallasvik & Martins, 2020)가 수행되었으며 국내에서 공학 분야에 초점을 맞춘 데이터 관련 연구는 부재한 상황이다.

연구수행의 환경과 연구자들의 복잡한 상황을 알아보고 그 맥락을 탐구하는 것은 연구 활동에서 생산되는 데이터 및 데이터 공유 행위를 이해하는데 있어 필수적이다(Kurata, Matsubayashi, & Mine, 2017). 해외에서는 대학을 중심으로 연구데이터 관리서비스의 요구사항을 파악하기 위한 목적으로도 데이터 공유와 관리에 대한 연구자 인식을 조사한 연구들이 많이 수행되고 있다. 국내의 경우 대학에서의 연구데이터 관리서비스 사례는 찾아보기 어렵지만 정부출연연구기관을 중심으로 연구데이터 관리를 위한 플랫폼을 구축한 사례들이 존재한다.

더불어 정부출연연구기관에서 연구데이터, 연구관리기록과 성과물을 포함하는 연구기록을 기록관리 영역에서 관리할 필요성에 대해서도 논의되어 왔으며 이는 연구기록과 연구과정에 대한 깊이 있는 이해와 기관 소속 연구자들과의 협력에 기반을 두어야 함을 강조하고 있다(이미영, 2017; 안채영, 김지현, 2020).

연구자들의 관점에서 데이터 관리와 공유의 맥락을 이해하기 위한 연구는 그들의 요구에 맞는 서비스를 제공하기 위한 실무적인 차원에서도 필요하며, 특히 그간 관련 연구가 미진했던 공학 분야 연구자들을 대상으로 데이터 관리와 공유에 대한 인식을 조사하는 연구가 필요함을 알 수 있다. 그러므로 본 연구는 정부출연연구기관의 기계공학 분야 연구자들을 대상으로 연구 활동에서 이루어지는 연구데이터 생산과 관리의 전체적인 과정과 행태 및 인식을 조사하고 이를 토대로 기계공학 분야의 연구데이터 관리와 서비스를 위한 시사점을 제시하는 것을 목적으로 한다. 본 연구에서 살펴본 연구 문제는 다음과 같다.

- 1) 연구데이터 생산과 관리에 대한 연구자들의 인식과 행태는 어떠한가?
- 2) 기계공학분야 연구데이터는 어떤 특징을 갖는가?
- 3) 연구데이터 관리에 대한 방해요인과 동기요인은 어떤 것이 있는가?
- 4) 연구데이터 관리를 위해 고려해야 할 사항과 실행 가능한 방안은 무엇인가?

1.2 연구 방법

본 연구에서는 연구자들의 연구데이터 생산과 관리의 과정, 구체적인 행태와 인식을 면담을 통해 살펴보는 질적 연구를 진행하였다. 연구 문제에 대한 국내 선행연구가 많지 않기 때문에 탐험적 연구가 필요하며 질적 연구를 통해 현상에 대한 상세한 점검을 하고 이를 토대로 시사점을 도출하는 방법이 타당하다고 판단하였다(임진희, 2010). 따라서 질적연구 자료 생성을 위해서 선행연구 분석 및 해외사례 조사를 통한 문헌조사를 실시하고 면담을 수행하였으며, 녹음자료를 바탕으로 녹취록을 생성하고 해당 내용을 엑셀로 코딩 분석하여 추적과정을 거쳤다. 이렇게 다중적 방법을 사용하는 것을 흔히 삼각기법(triangulation)이라고 하는데(Glesne, 2017) 이를 통해 연구자들의 연구데이터 생산과 관리 행태에 대한 현상을 객관적이고 입체적으로 관찰하고 다면적으로 분석하고자 하였다. 문헌 조사 및 사례 분석 내용과 면담 분석 내용 중 연결고리가 있거나 공통적으로 제시된 부분을 도출하여 기계공학 분야의 특성에 따른 연구데이터 관리와 서비스를 위한 시사점을 제시하였다.

먼저 문헌연구를 통해서 연구데이터 관리 현황을 논의한 연구와 과학 및 공학 분야 연구자 인식을 조사한 연구들을 살펴보았다. 또한 국내 연구데이터 관리 및 서비스의 법적·정책적 기반과 국내 정부출연연구기관에서 수행한 연구데이터 관리 플랫폼 구축 사례들을 문헌에 제시된 내용을 중심으로 제시하였다.

이와 함께 해외 연구기관의 분석 사례로 호주의 연방과학산업연구회(Commonwealth Scientific and Industrial Research Organisation, 이하 CSIRO)와 독일의 프라운호퍼연구소를 조사하였다. 이들 기관은 과학기술분야의 연구활동과 그 연구를 토대로 한 기술사업화가 활발히 이루어지는 연구기관이면서 본 연구의 면담 참여자들이 속한 기계공학분야 연구기관의 성격과 매우 유사하므로 사례조사 대상기관으로 선정하였다. 해외 사례조사 기관들의 웹사이트를 중심으로 해당 국가의 관련 제도, 관리되는 연구데이터 대상, 관리 수행 주체 등을 살펴보았다.

다음으로 연구자들을 대상으로 심층 면담을 수행하였다. 문헌조사 및 선행연구분석, 해외사례 조사를 바탕으로 질문지를 작성하고 사전에 질문지에 대한 답변을 토대로 반구조화된 면담을 실시하였다. 기계공학 분야 연구기관에서 연구를 수행 중인 공학 분야 연구자들로 면담 대상자를 선정하였고, 면담내용을 전사하여 녹취록을 작성하고 코딩 체계 구축을 통해 내용을 범주화하여 분석하였다.

1.3 선행연구

본 연구에서는 국내에서 연구데이터 관리 현황과 시사점을 논의한 연구들과 과학 및 공학 분야 연구자 인식을 분석한 국내외 연구들을 중점적으로 살펴보았다. 먼저 국내에서 연구데이터 관리 현황을 분석하여 개선방안을 제시한 연구들 중 대학의 연구자들을 대상으로 한 연구로 강주연(2017)은 생명공학분야 연구데이터의 특징, 관리 및 보존, 공유와 접근에 대한 특성을 파악하기 위해서 대학 실험실 소속 연구자를 대상으로 설문조사를 수행하였고 그 결과를 바탕으로 연구데이터 리포지터리 설계 시 필요한 요구사항을 분석하여 도출하였다. 김지현(2012)은 대학 소속 연구자 대상 설문조사를 통해 인문사회과학 분야와 자연과학 의학 및 공학분야 연구자들간에 생산, 수집되는 연구데이터의 유형에 차이가 있음을 제시하였고 장기적인 데이터 보존 및 메타데이터 작성에 대한 지원이 필요함을 제안하였다.

정부출연연구기관을 중심으로 한 연구로 박미영, 안인자, 남승주(2018)는 정부출연연구기관 27개 데이터 업무 담당자를 대상으로 과학기술분야 학문분야별 연구데이터 공유와 활용에 관하여 반구조화된 인터뷰를 실시하고 그 중 9개 우수기관을 선정하여 연구데이터 수집 및 관리현황 사례를 분석하였다. 또한 박미영, 안인자, 김준모(2018)는 미국 NISO와 영국 UK 아카이브의 연구데이터 관리지침을 조사하여 연구데이터 관리계획의 주요 단계를 도출하고 데이터 활용도가 가장 높은 생명공학분야 국내외 7개 기관의 연구데이터 공유 활용 사례를 비교분석하여 국내 생명공학분야 연구데이터 관리 방안을 제시하였다. 이 외에도 최명석, 이승복, 이상환(2017)은 정부출연연구기관과 국내 대학 소속 연구자들을 대상으로 설문조사 및 인터뷰를 실시하고 국내 과학기술분야 연구기관의 과학데이터 생산, 관리, 활용 현황 조사를 위해 과학데이터 관리·활용체계 개선방향을 제시하였다.

이와 같이 국내에서는 정부출연연구기관 및 대학을 주요 연구대상으로 하여 데이터의 생산에서 보존과 재이용에 대한 전반적 측면에서 나타난 연구데이터 관리상의 문제를 살펴보고 이에 대한 시사점을 도출하는 연구가 주로 이루어졌다. 특히 최근에 세부분야별 연구데이터 관리 현황을 분석하려는 연구들(강주연, 2017; 박미영, 안인자, 남승주, 2018)이 점진적으로 수행되고 있음을 알 수 있었다.

연구데이터 관리와 공유에 대한 연구자들의 인식과 관련하여 해외에서는 다수의 관련 연구들이 축적된 관계로 이를 토대로 하는 체계적 문헌고찰(systematic review) 연구가 이루어지고 있었다. Zuiderwijk, Shinde, & Jeng (2020)은 2004년에서 2019년 사이에 출판된 32편의 오픈 데이터 관련 연구를 분석하여 데이터 공유를 촉진 혹은 방해하는 요인을 연구자 배경, 공유 의무, 개인적 동기, 공유 촉진 조건, 신뢰, 기대 효과, 사회적 영향, 노력, 연구자의 경험과 능력, 법제도, 데이터의 특성 등 11개 범주로 제시하였다. Perrier, Blondal, & MacDonald(2020)의 연구에서는 2003년부터 2018년 사이 출판된 45편의 연구와 3편의 보고서를 중심으로 데이터 공유 및 재이용에 대한 연구자들의 견해와 경험을 분석하였다. 데이터 및 데이터에 대한 도큐멘테이션의 품질을 포함하는 데이터 무결성, 책임 있는 연구 수행, 데이터 공유에 필요한 인프라, 시간 및 노력과 데이터 공유에 필요한 기술적 지식, 데이터 공유의 가치를 언급하였다.

이와 더불어 van Panhuis 외(2014)는 보건학 분야의 데이터 공유 연구들을 중심으로 체계적 문헌고찰을 통해 데이터 공유의 방해요인을 동기부여, 경제적, 정치적, 법적, 윤리적 측면의 다섯 가지 범주에서 제시하였다. 동기부여 측면에서는 인센티브의 부재, 데이터에 대한 비판, 기회 비용, 데이터가 재이용되는 것이 적절하지 않다는 연구자들의 인식이 포함되었다. 경제적 측면의 경우 경제적 손실 가능성과 데이터 공유를 위한 인적·기술적 자원 부족을 언급하였다. 정치적 측면에는 신뢰 부족, 데이터 공유를 제한하는 정책 및 데이터 공유를 지원하는 가이드라인의 부족이 제시되었으며 법적 측면에서는 소유권과 저작권, 개인정보 보호 문제가 논의되었다. 윤리적 측면과 관련하여 공유하는 데이터의 양과 유형에 따른 위험과 혜택이 데이터 재이용을 통한 잠재적인 연구 영향력에 비례하는지의 문제와 데이터 공유자가 재이용자에 비해 손해를 보는 것이라는 인식이 존재함을 설명하였다.

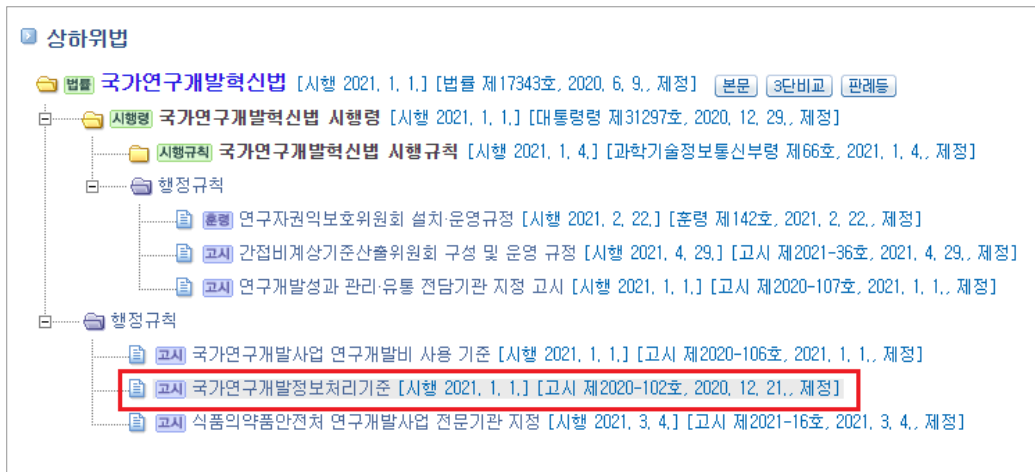
국내 연구로 권나현, 이정연, 정은경(2012)은 생명 및 나노과학기술자들을 대상으로 심층 면담을 통해 아이디어

생성 및 개발, 연구지원비 확보, 실험 및 분석, 성과 창출, 평가로 구성되는 과학기술 R&D 전 주기 과정을 제안하였다. 면담 참여자들은 실험 과정에서 생산되는 연구노트를 개인의 기억보조장치, 연구실의 기억, 정보공유의 도구, 실험실 작업의 연속성을 돕는 측면에서 그 가치를 높게 평가하고 있었다. 그러나 연구노트를 실험실 간에 공유하는 측면에 있어서 보안의 문제와 다른 실험실의 연구자들이 알아보기 어려울 수 있다는 이유로 꺼리는 경향이 많았으며 실험실 내에서의 공유가 유용하다는 입장을 보였다. 김은정, 남태우(2012)는 2010년 SCI 논문을 출판한 국내 정부출연연구기관 소속 연구자들을 대상으로 연구데이터의 제출과 공유에 영향을 미치는 요인을 정보인프라, 동기부여, 장애요인의 세 가지 측면에서 분석하였다.

본 연구와 가장 연관된 공학 분야의 연구자들을 대상으로 한 연구는 소수 이루어졌으며 모두 반구조화된 면담을 기반으로 하는 질적 연구로 진행되었다. Wu & Worrall(2019)은 지진공학 연구자들을 대상으로 데이터 공유의 영향요인을 조사하였으며 연구비지원기관의 데이터 공유 요구와 연구분야 내에서의 사회적 친분이나 같은 연구팀 혹은 소속기관 등의 관계가 데이터 공유에 영향을 미치는 것으로 나타났다. Suhr, Dungal, & Stocker(2020)는 재료공학 연구자들의 데이터 공유를 조사하였는데 마찬가지로 연구비지원기관의 요구가 데이터 공유의 가장 큰 영향요인이었으나, 부분적으로 산업계의 연구지원을 받는 경우 법적 문제가 발생할 수 있다는 우려가 존재하였다. 이들은 데이터와 데이터 생산에 대한 정확하게 상세한 기술(description)을 제시하는데 많은 시간이 소요된다는 점을 데이터 공유의 어려움으로 언급하였다. Mallasvik & Martins(2020)는 영국의 셰필드 대학 소속의 연구자 3명과 노르웨이 과학기술대학교 소속 연구자 4명을 대상으로 그들이 인식하는 데이터 공유의 가치와 방해요인을 조사하였다. 연구자들은 데이터의 생산맥락과 연구의 배경에 대한 상세한 메타데이터를 제공하는 것이 데이터 공유에 있어 필수적임을 인식하고 있었다. 그러나 산업계의 지원을 받는 연구를 많이 수행하는 공학 분야 연구의 특성상 비밀유지와 상업적 이익의 보호를 위해 데이터 공유에 많은 제약이 존재하는 것으로 나타났다.

2. 국내 연구데이터 관리 법체계 및 플랫폼 사례

인공지능, 클라우드 기술 등 디지털 기술이 발전하면서 연구데이터의 양과 종류가 기하급수적으로 축적되고 연구데이터 공유 및 활용에 대한 수요가 증가하면서 연구데이터의 가치와 중요도는 점점 높아지고 있다. 국내에서도 이러한 흐름에 따라 연구 분야별 연구데이터 구축사업을 진행한 바 있으며, 2017년 하반기부터 과학기술정보통신부 주관으로 본격적으로 연구데이터에 대한 종합 대책을 마련한 바 있다(신은정, 2018). 2018년 국가과학기술심의회 운영위원회 안전(2018. 1. 19.)으로 ‘연구데이터 공유·활용 전략(안)’이 상정·의결되었고 세부 추진 과제로 (1) 연구데이터 관리체계 구축 및 공유커뮤니티 형성 촉진, (2) 국가연구데이터플랫폼 구축 및 서비스 제공, (3) 데이터 및 컴퓨팅 활용 R&D 인재 성장 지원, (4) 연구데이터 관리·공유·활용에 관한 법·제도 마련, (5) 연구데이터의 산업적 활용 촉진 및 일자리 창출이 제안되었다(국가과학기술심의회 운영위원회, 2018). 당시의 정책적 추진과제를 통해서 국내 유관 법령 중 「국가연구개발사업의 관리 등에 관한 규정(이하, 공동관리규정)」에 연구데이터의 정의 및 데이터관리계획(Data Management Plan, 이하 DMP)의 도입을 포함하여 개정된 바 있다. 이는 2019년 9월부터 시행되었으며 당시 유관 기관들은 과학기술정보포럼 등을 개최하여 연구데이터 관리 등에 관한 대응방안을 논의하기도 하였다(과학기술정보포럼, 2019). 그러나 중앙행정기관별로 다르게 운용되고 있는 국가연구개발사업이 통합적·체계적으로 운영될 수 있게 하고, 국가연구개발사업을 추진하는 과정 전반의 비효율과 불필요한 부담을 제거함과 동시에, 자율적이고 책임 있는 연구개발 환경을 조성하기 위하여 국가연구개발사업의 추진에 대한 범부처 공통규범으로 「국가연구개발혁신법(이하, 혁신법)」이 2020년 제정·시행되었다. 그 과정에서 공동관리규정은 혁신법에 귀속되어 폐지되었고, ‘연구데이터’는 혁신법 제3조(연구개발성과)에 포함되지 않으며 관련 내용들도 일제히 삭제되었다. 대신하여, 과학기술정보통신부 고시 「국가연구개발정보처리기준」에 연구데이터에 대한 상기의 내용 등을 포함하고 있다(<그림 1> 참조).



〈그림 1〉 국가연구개발혁신법 법령체계도

혁신법의 제정으로 제도의 권위와 포괄성이 강화되었다고 볼 수 있지만 한편으로는 여전히 각 부처별 법령들로 인해 국가연구개발사업의 발주 부처에 따라 제도의 일관성을 기대하기 어려울 수 있다. 대부분의 연구들은 과기부, 산업부, 환경부 등 여러 부처에서 공모하여 실제 연구기관(국공립 연구소, 출연연, 법인연구기관 등)에 의해 수행된다. 연구기관은 발주 부처의 소관 법령 및 관련 규정을 준수해야 한다. 혁신법 제4조(다른 법률과의 관계)에 따르면, 해당 법률이 다른 법률에 우선하여 적용한다고 되어 있어 이는 연구기관에서의 혼란을 야기할 여지가 있다. 연구데이터 관리에 대한 여러 가지 정책적 과제들이 추진되었고, 현재도 추진 중에 있지만 구체적인 연구데이터 정의와 범위, 관리 및 서비스 방안, DMP 도입 및 이행에 대한 권위 있고 실효성 있는 법률적 근거는 여전히 미흡함을 알 수 있다.

법제도적으로는 개선의 여지가 많으나 정부출연연구기관을 중심으로 연구데이터 관리와 공유를 위한 플랫폼을 구축하는 노력은 지속적으로 이루어지고 있다. 대표적인 사례로 한국생명공학연구원에서는 생명자원분야에 대한 연구성과 관리 및 유통 전담기관으로 지정되어 국가생명연구자원정보센터(KOBIC)의 임무를 수행하고 있으며 ‘국가생명연구자원 통합정보시스템(KOBIS)’을 운영 중이다. 과학기술정보통신부 생명연구자원 포털(ARIS), 농림축산식품부 생명자원서비스(BRIS), 환경부 국가생물 다양성정보공유체계(CBD-CHM KOREA), 해양수산부 해양생명자원통합정보시스템(MBRIS), 보건복지부 질병관리본부 국가병원체자원은행(NCCP)의 정보를 연계하여 범부처 통합서비스를 제공한다.

또한 한국화학연구원에서는 2000년 한국화학물은행을 설립하고 2008년 화학물 관리 및 유통 전담기관으로 선정되어 운영 중이다. 국내에서 합성되는 화학물(유기화학물 및 천연물) 및 관련 정보를 범국가적 차원에서 통합·관리하고 공동으로 약효시험에 활용함으로써 국내 신약개발 및 바이오 연구를 지원하고 있으며 67만종 이상(2020년 기준)의 화학물을 수집 및 관리하고 있다. 한국화학물은행은 기탁기관에서 기탁한 화학물을 체계적으로 관리하여 약효시험기관에 제공하는 기탁기관과 약효시험기관을 연결해주는 역할을 하고 있다.

한국지질과학연구원에서는 지질자원데이터센터를 통해 지질학 분야의 연구데이터 관리 및 서비스를 ‘지오빅데이터 오픈플랫폼’을 통해 제공하고 있다. 2019년 연구데이터 관리규정을 제정하여 연구활동을 통해 획득·생산한 연구데이터의 수집, 관리, 보존, 개방 및 활용에 관한 사항을 규정하고 있다(과학기술분야 정부출연연구기관 기록관리협의회, 2021).

이 외에도 한국과학기술연구원에서는 데이터 집약적 연구를 위한 웹기반 연구인프라(KiRI)를 구축하여 온라인 연구노트 작성과 연구데이터와 분석 틀을 기관 내부 구성원간 공유할 수 있는 플랫폼을 지원한다. 기관 내 촉매제

연구자에 한정하여 데이터 공유 플랫폼을 운영하고 있지만 데이터 공개 및 방식에 대한 판단은 연구책임자가 내릴 수 있도록 위임하고 있다. 연구자들은 데이터를 통해 목적인 바를 우선 달성할 때까지 우선적·배타적으로 데이터를 활용할 수 있다. 이와 같이 연구분야나 기관 특성에 따라 연구데이터의 범주와 속성이 다르게 나타남을 알 수 있다(신은정 외, 2019).

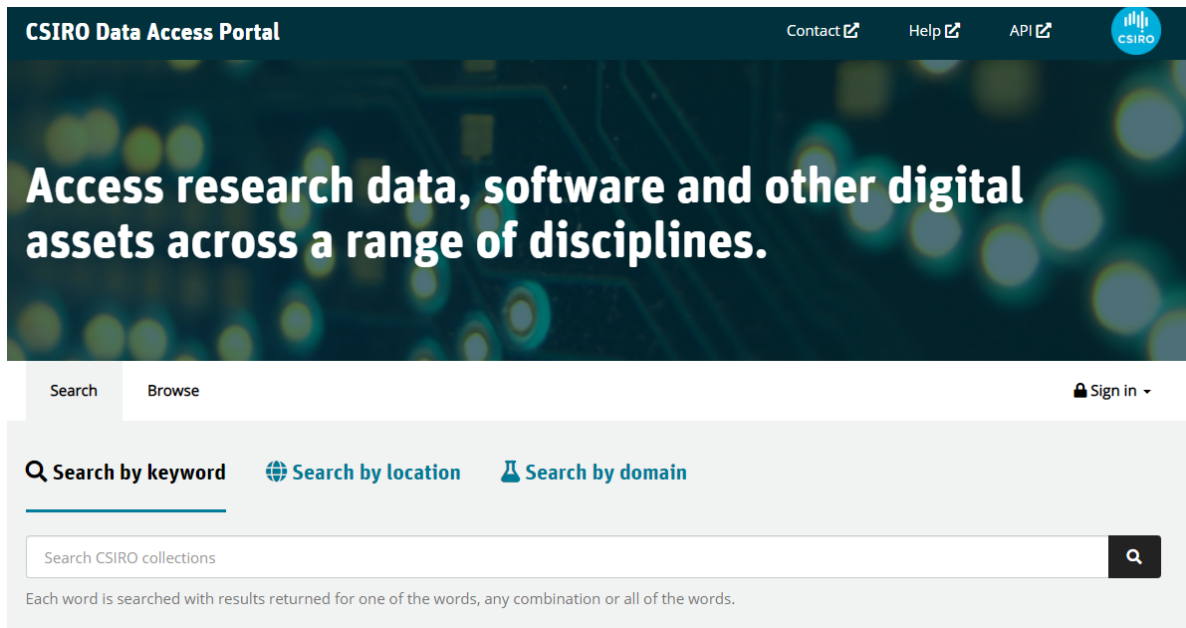
3. 해외 기계공학 분야 연구기관 사례 분석

해외에서는 2000년대 초반부터 공공연구에서 생산된 연구데이터 접근성 제고를 위한 논의를 시작하여 OECD에서 2006년 「공공기금으로 수행된 연구의 연구데이터에 대한 접근을 포함하는 위원회의 권고(Recommendation of the council concerning access to research data from public funding)」와 「공공기금으로 수행된 연구의 연구데이터에 접근하기 위한 원칙과 지침(Principles and guidelines for access to research data from public funding)」을 발표하면서 현재까지 국제기구 및 세계 주요국에서 관련 규정 및 지침을 구체화하고 있다. 최근 들어 연구출판물과 연구데이터의 연계, 연구데이터의 공유 및 활용, 보다 개방적인 연구협력 및 소통을 강조하는 광의의 오픈 사이언스 정책을 본격화하고 있다(신은정, 2018). 본 연구에서는 여러 선진사례 중에서 본 연구와 관련있는 기계공학분야를 포함하는 과학기술 연구기관인 호주의 ‘CSIRO’와 독일의 ‘프라운호퍼연구소’를 선정하여 살펴보았다.

3.1 호주: CSIRO

호주 연방과학산업연구회(Commonwealth Scientific and Industrial Research Organisation; 이하 CSIRO)는 1949년 「Science and Industry Research Act(이하, 과학 및 산업 연구법)」에 따라 설립되었고 이에 따라 운영된다. 산업과학자원부(DIST) 산하에 있는 호주 최대의 과학기술분야 공공연구기관으로 환경, 기후기술, 정보기술, (재생)에너지, 광공업, 자원재활용 등을 연구한다. 5,000명 이상의 인력과 전 세계 56개의 지소(Sites)를 운영하며, 연간 20억 달러 이상의 예산으로 50여 개국의 3,600여개 이상의 산업분야 파트너와 연구하고 있다(Australia’s National Science Agency, 2020). 호주의 연구수행을 뒷받침하는 광범위한 원칙과 책임에 대한 사항은 ‘Australian Code for the Responsible Conduct of Research’를 따르는데 이 원칙은 국가 보건 및 의료 연구위원회(National Health and Medical Research Council; 이하 NHMRC), 호주 연구위원회(Australian Research Council) 및 Universities Australia와 공동 개발한 것으로, 모든 연구 분야에 적용된다. 그 중 ‘연구데이터 및 정보관리(Management of Data and Information in Research)(NHMRC, 2019)’ 부분에 대한 가이드는 책임 있는 연구수행과 데이터 및 정보의 적절한 생성, 수집, 접근, 사용, 분석, 공개, 저장, 보존, 폐기, 공유, 재사용을 포함하며 강령의 목적을 달성할 수 있도록 호주 정부의 후원하에 연구를 수행하는 모든 기관에 적용하도록 권고하고 있다. 또한 이를 달성하기 위해서 기관과 연구자의 책임 사항을 명시하고 있다.

CSIRO의 연구데이터는 연방 정부차원에서 호주 국가 기록으로 포함되고 ‘과학적 연구’ 기능에 따른 기록의 유형으로 관리되며, 처분에 대한 사항은 호주 국립기록원(National Archives of Australia; NAA)에서 정하고 있다. 국립기록원의 CSIRO 기록처분지침 ‘Scientific Research’ 기능에서 ‘발표, 실험 및 관찰, 절차, 보고, 연구분석, 연구설계, 연구협력, 표준화’ 등으로 구분하여 기술(설명) 및 보존기한을 명시하고 있다(NAA, 2020). CSIRO는 연구데이터, 소프트웨어 및 기타 디지털 자산에 대한 접근을 제공하기 위해서 ‘CSIRO Data Access Portal’을 운영 중이다(<그림 2> 참조).



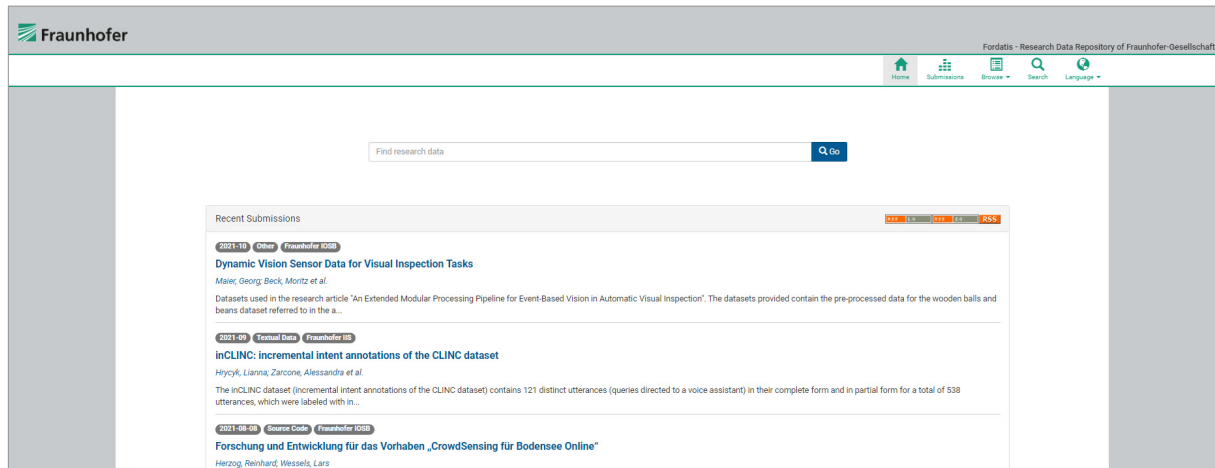
〈그림 2〉 CSIRO Data Access Portal

CSIRO는 연구데이터 서비스를 통해 혁신, 협업 및 과학적 무결성을 지원하는 데이터를 수집, 관리 및 공개하며 연구데이터의 안전한 저장과 편리한 검색, 적절한 재사용을 지원하는 것이 목적이다. CSIRO Data Access Portal에 제공되는 연구데이터들은 CCSDS 650.0-M-2 ‘REFERENCE MODEL FOR AN OPEN ARCHIVAL INFORMATION SYSTEM(OAIS) Magenta book’을 기반으로 수집, 저장, 데이터 및 메타데이터 관리, 보존 계획, 접근을 위한 기록 보유 일정을 포함하고 있다(CSIRO, 2018). 연구데이터의 상세정보에는 ‘제목, 컬렉션 서술, 저자, ORCID ID, 생성 시작일과 종료일, 접근 권한, Creative Commons License, 관련 연구프로젝트 설명, DOI’ 등을 제공하며 이용자들은 검색 기능을 통해서 접근할 수 있고 다운로드를 통해 재이용할 수 있다.

3.2 독일: 프라운호퍼연구소

독일은 유럽연합의 중심 회원국으로 R&D 지출 비중이 국내총생산(GDP) 대비 3% 초반대를 기록한다. 유럽연합의 평균이 2.2%인 것을 감안하면 독일은 이를 웃도는 높은 수준의 R&D 투자를 유지하는 과학기술 강대국으로 꼽힌다. 프라운호퍼연구소는 1949년 뮌헨에 설립되었고 독일 전역에 75개의 연구소를 두고 있으며 29,000여명의 직원이 연간 28억 유로(2020년 기준, 한화 약 3조8천억)의 예산으로 연구 활동을 수행하고 있는 유럽 최대의 응용과학연구소이다. 연구소는 산업계에서 활용될 수 있는 응용기술에 대한 연구를 수행하여 국가 산업 발전에 이바지 하는데 그 사명을 두고 있다. 총 8개의 프라운호퍼 그룹으로 이루어져 있으며, 75개의 연구소들은 1개의 그룹에 속하여 운영되고 있다. 연구분야는 바이오경제, 디지털 헬스케어, 인공지능, 차세대 컴퓨팅, 양자 기술, 자원 및 기후기술, 수소기술 등이다. 위탁연구 수행, 기술 지원, 검사 및 인증 등의 연구활동을 명시하고 있어 우리나라 과학기술분야 출연연구기관들과 비슷한 임무를 수행하고 있음을 알 수 있다. 2003년 과학 지식의 개방적 접근에 대한 ‘베를린 선언’에 최초로 서명했으며 2008년부터 개방적 접근 정책을 시행했다. 이러한 정책을 기반으로 2015년 10월 ‘Fraunhofer Open Access Strategy 2020’을 발표하였고 연구결과와 관련된 연구데이터도 자유롭게 접근할 수 있도록 하였다. 주요 내용은 연구결과에 대해 가능한 직접적이고 즉각적인 오픈엑세스를 추구하며 연구자의 과학적·법적 이익을 존중하고 개인정보 보호와 고객 및 협력파트너와 같은 제 3자에 대한

의무를 준수하는 것을 포함한다. 이와 같은 오픈액세스 정책을 바탕으로 연구과정에서 생산된 모든 정보자료들에 대해 웹을 통해 서비스하고 있다. ‘Fraunhofer-Publica 데이터베이스’는 정보자료 서비스를 위한 하나의 플랫폼으로 볼 수 있는데 여기에는 ‘모노그래프, 저널 기사, 도서, 소송 절차, 학위논문, 연구보고서, 특허, 회색문헌’ 등이 포함된다. 연구데이터 검증 및 오픈액세스를 위해서 2019년 9월부터 연구데이터 리포지터리와 서비스 플랫폼의 역할을 하는 ‘Fordatis’를 운영 중이다(<그림 3> 참조).



<그림 3> Fordatis 메인화면

Fordatis에서 제공하는 연구데이터에는 프라운호퍼연구소의 연구 프로젝트에서 생성된 모든 데이터를 포함하며, 표준화된 메타데이터를 사용하고, OAI-PMH(Open Archives Initiative-Protocol for Metadata Harvesting) 인터페이스를 통해 Google에서 색인을 생성하기 때문에 모두에게 공개되어 있다. 공개된 연구데이터는 검색 기능을 통해 자유롭게 접근할 수 있고 다운로드하여 재이용이 가능하다.

연구데이터의 상세정보에는 ‘제목, 저자, 초록, 분류, 키워드, 관련 자료, 용량, DOI, 기술적 정보(Technical Information)’ 등으로 이루어지며 전체 메타데이터 레코드 항목은 Dublin Core 필드에 따라 제공한다. 또한 연구데이터의 라이선스는 Creative Commons License에 따라 미리 선택된 다양한 라이선스 옵션을 사용하며 게시자는 리포지터리에 업로드함으로써 발행된 라이선스에 따라 연구데이터를 제 3자가 사용할 수 있게 하는데 동의한다.

독일 프라운호퍼연구소는 오픈액세스 정책에 따라 활용을 위한 저작권을 명시하여 게시된 데이터에 대해 최소 10년 동안 저장 및 제공한다. ‘Research Services & Open Science Competence Center’를 통해 데이터 품질관리, 데이터 검토를 통한 플랫폼에서의 삭제 권한 보유, 데이터 저장, 백업, 보관 및 장기보존 등을 수행하고 있다.

4. 기계공학분야 연구데이터 생산과 관리

4.1 자료 수집

기계공학분야 연구자들의 연구수행과정에 따른 연구데이터 생산과 관리에 대한 행태와 인식을 자세히 알아보고자 기계공학 분야 연구기관의 연구자 8명을 대상으로 반구조화된 면담을 수행하였다. 면담 질문지 작성에 있어 연구자 인식에 대한 부분은 김지현(2012)의 연구를, 연구데이터 생애주기와 관련된 부분은 Corti 외(2020)의 가이

드를 참고하였다. 질문지 검토를 위하여 사전 인터뷰를 1회 실시하였고, 이를 바탕으로 질문지를 수정, 보완하였다. 질문지는 크게 5개 부분으로 1) 면담자 정보; 2) 연구데이터 정의와 범위; 3) 연구수행과정과 연구데이터 생애 주기; 4) 연구데이터 관리와 서비스에 대한 연구자의 인식; 5) 연구데이터 공유와 공개로 구분하여 설계하였다. 질문지는 면담 대상자에게 면담 수행 전 메일로 전달하여 1차 회신을 받은 뒤 그 내용을 바탕으로 개별적으로 직접 면담을 수행하였다(<표 1> 참조).

<표 1> 면담 수행 요약

면담자 번호	1	2	3	4	5	6	7	8
면담 수행 일자	2021.8.11	2021.8.12	2021.8.13	2021.8.26	2021.8.27	2021.9.2	2021.9.8	2021.9.10
면담시간(분)	120	40	100	70	70	70	70	80
연령대	30	30	40	40	50	50	50	40
성별	남	여	남	여	남	남	남	여
직급	선임	선임	선임	책임	책임	책임	책임	책임
연구분야	에너지 기계	첨단기계	에너지 기계	나노응용 기계	기계시스템	나노응용 기계	첨단기계	환경기계
전공	기계공학	기계공학	기계공학	기계공학	기계공학	기계공학	기계공학	화학공학
세부전공	열유체, 극저온 공학	반도체 전자 패키징	열전달, 전산설계	나노기술, 에너지 생산	동역학, 신호처리	고체역학, 파괴역학	광학레이저, 측정가공	유동화학, 폐기물연소
연구수행경력(년)	5	3	11	16	26	18	17	17
수행연구(개)	7	4	6	5	22	5	5	5

인터뷰 및 면담은 2021.8.11. ~ 2021.9.10. 약 한 달 동안 총 8회, 620분으로 약 10시간 20분 동안 수행되었다. 면담 수행 시 녹음동의서를 모두 수취하여 녹음 파일을 바탕으로 녹취록을 작성하였고 엑셀을 통해 코딩 체계를 구축하고 면담내용을 구성 및 범주화하였다. 또한 면담 종료 후에도 면담 내용과 관련된 참고자료가 필요한 경우 메일을 통해서 확보하는 등 최대한 면담 참여자의 목소리를 구체적이고 자세하게 반영하여 분석하고자 하였다. 면담자 선정 방식은 연구자가 주관적으로 표본을 선정하는 비확률적 표본추출 방법 중 판단표본추출 방법을 사용하여 연구대상이 되는 현장이나 대상에 대해 상세한 정보를 수집하여 특정하고 구체적인 사실을 명료하게 드러내 고자 하였다. 이는 연구자가 모집단에 대한 사전지식을 많이 가지고 있어 연구의 목적에 부합할 수 있는 표본을 추출할 수 있을 경우에 적합한 방법이다(한국행정학회, 2001). 면담 참여자들은 모두 공학자이며 8번 대상자 1명 을 제외하고는 기계공학 전공자들이고, 기계공학 관련 분야 연구기관에서 3년에서 26년의 경력을 가진 연구자들 로 이루어졌다. 표본의 모집단 대표성을 보완하기 위하여 연령, 성비, 직급을 최대한 골고루 포함하고자 노력하였 다. 기계공학 전공자 특성상 남성비율이 높은 것을 감안하여 5명의 남자와 3명의 여자로 구성하였고 연령대는 30~50대, 직급은 선임~책임으로 골고루 포함하였다. 2번 대상자는 기업체에서의 2년간의 연구 경력을 포함하면 면담자 모두 5년 이상의 연구수행 경력을 가지고 있기 때문에 본 연구의 참여자들로 적절하다고 보았다. 또한 면담자 모두 다수의 연구과제를 수행하고 있기 때문에 연구데이터 생산과 관리에 대한 구체적인 이야기를 들을 수 있었다.

4.2 분석 결과

면담을 통해 A4 72장의 녹취록을 작성하였고 엑셀을 통해 녹취록의 면담내용을 구성 및 범주화한 결과 주제별 로 4개의 주요범주와 11개의 하위범주로 구분할 수 있었다. 주요범주는 ‘연구데이터, 책임있는 연구수행과 연구윤

리 준수, 연구데이터 관리의 효용성 및 효과성, 연구데이터 공유의 가치'이다. 범주에 따라 연구자들의 관련 인식과 행태 측면에서 분석한 내용은 <표 2>로 요약하였다. 표에 제시된 주요범주와 각 범주별 하위범주에 따라 면담 내용 분석 결과를 제시하였다.

<표 2> 면담 내용의 주요범주 및 하위범주

범 주	하위범주	분석측면	내용
연구데이터	정의와 범위	인지	관련 법률내용에 대한 인지 부족 연구수행 전반적인 과정에서 도출되는 모든 형태의 데이터로 포괄적 인지
		행태	'생산 및 획득-저장-수정-재생산'의 과정 계속적이고 반복적 과정
	품질 및 신뢰성 확보	인지	데이터 생산과정에 대한 신뢰도 확보 필요
		행태	백라정보 획득 및 관리 노력
	가치 / 이용가능성	인지	현용 및 미래가치와 잠정적 가치
		행태	생산자(연구자 자기자신)와 연구팀이 추후에도 해석가능한 수준으로 데이터 정리 및 관리
특성	현황	실험, 분석 및 해석, 관찰 데이터가 대부분	
	행태	동시다발적, 대량 생산으로 관리의 필요성과 어려움이 동시에 존재	
책임있는 연구수행과 연구윤리 준수	데이터 재이용	인지	이용가능성에 대한 두려움 해석가능성에 대한 의구심
		행태	출판된 논문에 포함된 연구데이터만 인용 배타적이고 보수적인 공개와 공유
	데이터 소유권	인지	나의 것, 나의 연구 자산, 우리 연구팀 것
		행태	개인적 관리 또는 팀단위 관리 및 공유
	수탁기관과의 협약 사항	인지	연구수행 시 준수해야하는 가장 중요한 기준
행태		협약서에 명시된 연구성과물만을 제출 연구과정과 결과 공개시 협약처와 사전 협의	
연구데이터 관리의 효용성 및 효과성	데이터 관리 수행	인지	데이터 관리는 필수적이고 중요함 기계가독형으로 잘 정리된 데이터의 재이용이 용이
		행태	인프라, 인력, 시간, 연구환경의 한계가 존재 거의 전적으로 연구자 개인의 역량에 의존함
	보관방식	인지	인프라 부족 심각 소속기관의 보안규정
		행태	개인적 또는 팀단위로 보관 구식으로 관리
	보존기간	인지	보존기간에 대한 기준이 모호 5년 ~ 영구보존으로 인식
		행태	획득한 데이터는 무기한 보관
연구데이터 공유의 가치	공개와 공유의 가치	인지	보안성, 기밀성, 유출의 위험도가 공개와 공유의 이점보다 크게 인지
		행태	공개가능한 인원에게, 필요한만큼만 공개 및 공유

4.2.1 연구데이터

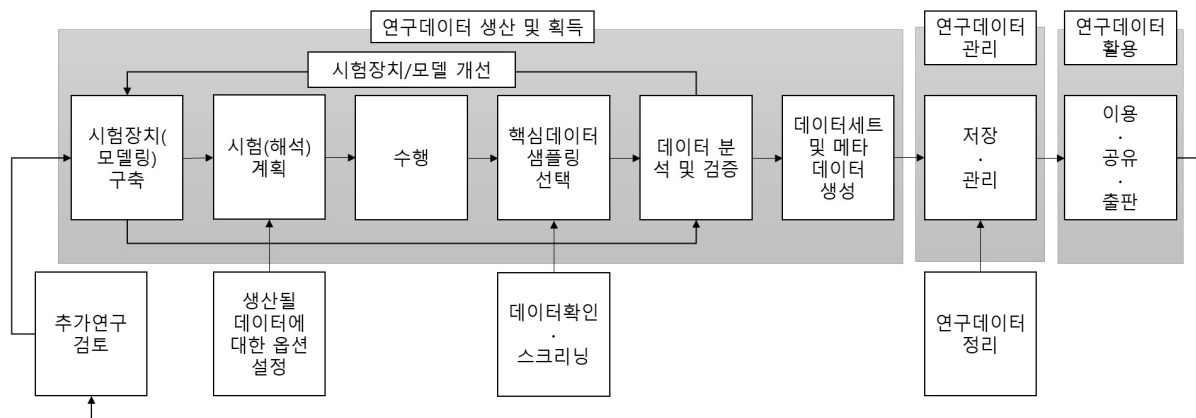
<표 2>에 제시된 첫 번째 주요범주인 '연구데이터'의 하위범주인 '정의와 범위'와 관련하여 면담 참여자 대부분이 연구데이터 및 DMP와 관련된 법률 내용을 알지 못하고 있었다. 최근 연구에서 대학도서관의 연구지원서비스를 담당하면서 연구데이터 관련 동향을 알고 있는 사서들도 해당 법률의 개정 내용을 인지하는 경우가 드문 것으로 나타났다(박형정, 김지현, 2021). 따라서 연구자들이 이를 인식하는 것은 더욱 요원한 일임을 알 수 있다.

“사실 저는 잘 모르겠어요. 굳이 연구데이터와 관련된 법이라고 하면 연구노트? 아마 이런 것들은 들어봤는데... 연구노트 안에 기재되는 것들이 다 데이터라고 생각하면 그런 거 정도만 들어봤지...” (면담자 5)

또한, 과기부 고시에 명시된 연구데이터와 DMP에 대한 정의를 알려준 뒤 의견을 물었을 때에는 해당 내용에 대부분 동의하지만 실제로는 상식적으로 훨씬 포괄적인 범위로 인지하고 관리하고 있다고 응답하였으며 정확한 의미에 대해서는 잘 모르겠다는 의견이 많았다.

“법률상 정의에 전반적으로 동의하지만 검증과정 뿐 아니라 연구 수행 일련의 과정에서 실시하는 실험, 관찰, 조사 및 분석 등을 통해 산출된 사실 자료를 포함하는 모든 데이터로 좀 더 포괄적으로 볼 필요가 있다고 봅니다.” (면담자 3)

면담 참여자들은 나름의 연구 관행을 기반으로 연구데이터의 생산·관리·보존을 수행하고 있었으며 연구 목표를 달성하기 위해서 대부분 실험, 해석, 분석 과정을 일반적으로 수행하고 있었다. 면담을 통해 기계공학분야 연구자들의 연구활동에서 도출된 연구데이터의 생애주기는 다음과 같다(<그림 4> 참조).



<그림 4> 실험(해석)과정의 연구수행에 따른 연구데이터 생애주기

<그림 4>에서 첫 번째 단계인 ‘시험장치(모델링)를 구축하는 단계’와 두 번째 ‘시험(해석) 계획 단계’는 사전 준비 단계로 물리적인 실험 장치를 준비하거나 해석 모델링을 설계·구축하여 실제 연구수행을 위한 준비를 하는 단계이다. 어떤 환경에서 어떤 과정으로 어떤 데이터를 얼마만큼 생산하고 획득할 것인지에 대한 결정이 이루어진다. 다음으로 계획한 대로 실제 수행이 이루어지며 그 과정에서 생산·획득한 데이터를 1차적으로 확인하고 스크리닝하여 핵심 데이터 샘플링을 선택한다. 이렇게 획득된 데이터를 분석 및 검증하여 해당 시험장치나 모델에 대한 개선이 이루어지고 이는 몇 번의 사이클을 거치면서 계속해서 업데이트 될 수 있다. 최종적으로 연구자가 확인하고자 하는 목표치를 얻게 되면 데이터세트 및 메타데이터를 생성하여 기록을 남기고, 이를 저장하고 관리하게 되며 이 과정에서 연구데이터에 대한 2차적인 정리가 이루어진다. 정리된 데이터를 기반으로 공유 및 출판, 공개의 단계로 나아가며 이를 활용한 최종 연구산출물에 대한 검토를 통해서 후속 연구나 새로운 연구에 착수하게 된다.

두 번째 하위범주 ‘품질 및 신뢰성 확보’와 관련하여 연구자들은 연구데이터가 도출된 과정에 대한 신뢰도 확보가 매우 중요하다고 인식하고 있었고 그렇기 때문에 맥락 및 배경정보를 획득하기 위해 노력하고 있음을 알 수 있었다. 맥락 및 배경정보는 두 가지 유형으로 요약될 수 있는데, 하나는 메타데이터와 같은 ‘명시적 맥락정

보'이고 다른 하나는 암묵적 지식을 포함하는 '암시적 맥락정보'이다(Kowalczyk & Shankar, 2011).

첫째, '명시적 맥락정보'는 연구데이터에 대한 메타데이터로 실험상황을 예로 들면, 실험 주제, 실험 장비, 수행자, 작업 흐름 및 데이터 수집 대상 등 연구결과 도출을 위해 연구자와 연구팀원들 간에 공유되는 정보이고 이는 연구과정의 신뢰도나 재현의 정확도를 향상시키기 위해 관리되는 정보이다. 면담자 6은 “온도에 따른 물성변화, 입력에 따른 물성변화 이런 것도 다 데이터고, 그래서 그 물성을 어떤 환경에서 어떤 방법으로 어떻게 측정했는지”가 모두 연구의 신뢰성과 연결되어 있기 때문에 실험 “과정과 결과값”을 “검토”하고 “정리”하는 작업이 필요하다고 언급하였다.

둘째, '암시적 맥락정보'는 '노하우'적 성격을 가진 것으로 이에 대한 중요성은 익히 알려져 있지만 암묵지를 표현하는 구체적인 방법에 대한 부분은 여전히 불분명하다. 연구활동은 비공식적 지식에 의존하는 경우도 많은데, 이는 연구 활동이 다양한 수준의 공동체 사이에서 암묵적 지식으로 공유되는 경우가 많기 때문이다(Kowalczyk & Shankar, 2011). 면담자 5는 데이터에 대한 기본적인 정리 뿐만 아니라 팀원들과 공유하기 위해 본인이 “이해한 부분, 해석한 부분 등에 대한 설명까지 적는 것”이 필요하다고 설명하였다.

“특허를 내거나, 논문을 내거나 하는 것 말고 노하우 같은 거 있잖아요 노하우는 머릿속에 있는 건데 그런 걸 정리하는 거죠 표현해서 정리해서 자료를 만들어 놓는건데, 제가 연구하고 있는 것들이 우리 기관의 자산이라고 생각해요...(중략) 이런 것들은 논문에서는 알아낼 수 없는 우리 연구팀의 노하우이고, 실험할 때 실수를 줄이고 효율을 높이는 중요한 정보가 될 수 있어요.” (면담자 5)

세 번째 하위범주로 연구데이터의 '가치와 이용가능성'에 대해 연구자들은 연구데이터의 가치가 무궁무진하다고 보았으며 현재 가치와 잠정적인 미래 가치에 대해 대부분 포용적이고 긍정적인 반응을 보였다. 이들은 미래의 어느 시점에서든 해석 가능한 수준으로 연구데이터를 정리하고 관리하려고 노력하고 있음을 알 수 있었는데, 면담자 4와 면담자 7의 경우 PPT를 활용하여 연구데이터가 생산된 실험조건을 '목적, 재료, 방법' 등으로 정리·작성하여 관리하고 팀내 공유에 활용한다고 하였다(<그림 5> 참조).

실험목적
구형, 신형, 3층 coater를 이용한 후 Imprint 레진 두께와 Etch 상태 확인 하기

실험 재료
glass/FTO, Imprint Mold(200nm, 400nm, depth 250nm), LOR 2A, LV600

실험 방법: Imprint

1. Glass/FTO 기판에 taping
2. Plasma 표면처리(30s)
3. LOR 2A Spin coating → 3000rpm, 60s
4. Baking(Hot plate) → 200도, 30min
5. Mold(200nm, 400nm, depth 250nm, hexa pattern)를 뜬다 → uv경화 90s, 2번 → 진공uv 경화 3min 1번
6. LV600 coating → 2000rpm, 40s
7. Baking(Hot plate) → 120도, 45s
8. 가압 UV경화 → 3bar가압, 5min대기 후 95s uv 진행

** LOR 2A, LV600을 coating 시, 구형/신형/3층 사용

실험 방법: Etching

1. He valve open
2. 전원 On(spare 제한된 레버 모두 on한 후, 컴퓨터를 켜다) → sorona 시스템 작동
3. Program: 박사님 → LV400(40s, 50W) → O2(80s, 100W)

<그림 5> 실험메모

면담자 8의 경우에도 실험과정이 매우 어렵고 변수가 많기 때문에 실험조건에 대해서 최대한 많은 정보를 메모하려고 한다고 설명하였다. <그림 6>에서와 같이 '실험시트'라는 것을 엑셀로 작성하여 날짜, 시간, 재료,

방법, 비교 등으로 정리·작성하여 실험과정과 조건 등의 맥락 정보를 기술하고 있었다.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
		NO	SO2	O2	CO2	CO	CH4	H2	투입	출탕	송풍	ASR	코크스	레벨	비고		
2		코크스 투입량:300kg.(레벨 4.5)															
3	10:20																
4	11:22										175, 33				착화		
5	11:45													4.7	ADF 가동, 출탕구 폐쇄		
6	11:50										160,33				송풍량 160m3/hr 변경		
7	12:44									o					출탕구 개폐, 세정		
8	13:04														세정작업 완료, 폐쇄		
9	13:05													4.9			
10	13:18								1			52	6	4.4			
11	13:25	237	229	0	7.94	48.25	0.78	2.86									
12	13:40								2				7	4.6			
13	13:48								3			52		4.3			
14	13:56													4.4			
15	14:00									o							
16	14:05	316	357	0	6.11	47.02	1.45	6.12						4.5	출탕시작, 세정		
17	14:16								4			52	7	4.3	출탕구 폐쇄		
18	14:25	440	556	0	6.99	44.92	1.94	8.6									
19	14:35	432	767	0	7.49	44.25	2.24	8.46	5			52	5	4.4			
20	15:00														열교환기 를 0으로 setting		
21	15:15													4.6			
22	15:35									o					출탕시작, 세정		
23	15:41													5	출탕구 폐쇄		
24	15:47								6		143,26	52	10	4.6	ASR 52kg, 코크스 10kg 투입		
25	16:07								7			52	5	4.5			
26	16:29								8			52	7	4.5			
27	16:51								9			52	4				
28	17:00									o					출탕시작		

<그림 6> 실험시트

그러나 면담 참여자들은 로우데이터(raw data)의 가치와 이용가능성은 거의 없다고 보았으며 그렇기 때문에 데이터의 가치와 이용가능성을 높이기 위해서는 충분한 맥락정보에 대해 파악하고 검토하는 것이 꼭 필요하다고 강조하였다. 면담자 7은 같은 연구를 수행하는 팀원들도 데이터 자체만을 보고 이해하거나 해석하는 것은 불가능하며 로우데이터만 제공하는 것은 “날생선을 주고 먹으라는 것과 똑같다”고 비유하기도 하였다.

네 번째 하위범주로 기계공학 분야 연구데이터의 ‘특성’에 대해 살펴보면, 기본적으로 숫자의 집합인 수치데이터를 기반으로 하여 그것들을 해석한 데이터, 모델링 데이터 등의 형태를 가지며 엑셀, 캐드(dwg, gds), SEM(Scanning Electron Microscope) 또는 TEM(Transmission Electron Microscope) 분석을 통한 이미지 파일, 그래프, ASCII포맷의 텍스트, 초고속 카메라 촬영영상 등에 따라 다양한 파일 확장자를 가지고 있는 것으로 나타났다. 대량의 데이터가 동시다발적으로 생산되는 경우가 많기 때문에 관리의 어려움이 존재하는 것을 알 수 있었다. 예를 들어 면담자 5의 경우 주파수를 데이터로 다루는데 이는 “1초에 2만 번이 움직이는 데이터”이므로 짧은 시간에 “엄청나게 많은 용량의 데이터들이 들어오는” 상황이라고 설명하였다. 그렇기 때문에 데이터에 대한 상세한 기술 등 데이터 관리가 더욱 중요해질 수밖에 없으며 이를 자동화하는 방법에 대해서 고민하고 있다는 의견도 있었다.

“저희 분야는 장비를 개발하는 연구가 많은데 그럼 특히 새롭고 다양한 것들을 많이 해요. 그래서 일일이 세세하게 적어놓지 않으면 나중에 내가 봐도 몰라요. 그 결과값들에 대해 알아보기도 힘들고 나중에 재현하기도 힘들고.. (중략) 데이터 관리는 힘들고 어려운 일이지만 해야만 하는 거예요. 그래서 최대한 자동화시키는 방법을 고민은 하는데 이것 저것 하다보면 우선순위에 밀리게 돼요.” (면담자 7)

4.2.2 책임있는 연구수행과 연구윤리 준수

<표 2>의 두 번째 주요범주인 ‘책임있는 연구수행과 연구윤리 준수’ 범주 아래 ‘데이터 재이용’과 관련해서

면담 참여자들은 타인의 데이터 오용가능성에 대한 우려를 가지고 있었으며 잘못된 해석 가능성에 대한 의구심도 나타내었다. 이들은 누가 자신의 데이터를 어떤 목적으로 어떻게 사용하는지 알고 싶어하였으며 데이터를 재이용하고자 하는 사람은 사전에 생산자의 승인을 받거나 공적 또는 사적인 네트워크 내에 포함된 사람으로 제한하려는 경향을 보였다. 면담자 4는 “논문에 들어간 그림”을 일종의 데이터로 인식하고 있었고 이를 재이용할 때 본인의 동의를 받을 수 있게 하고 있다고 설명하였다.

반면에, 논문 출판 통과와 논문 인용도를 높이기 위해서 데이터 재이용을 긍정적으로 인식하는 면담 참여자도 있었다. 논문 투고 시 보충 자료를 제출할 때 논문 내에 삽입된 그림의 기반이 되는 수치데이터를 함께 제출하여 자유롭게 활용하도록 한다고 하였고 논문에 공개되는 내용의 토대가 되는 데이터는 공개 가능하다고 설명하였다.

“최근에 논문 투고시 데이터를 함께 제출하기를 원하는 저널들이 있는데 논문의 accept와 신뢰도를 높이기 위해서 모두 출판사에 제출하였습니다. 논문에는 정리되고 정제된 내용으로 작성하고, Supporting materials를 제출할 때는 좀 더 자유롭게 관련 자료를 링크시키고 논문에 사용된 그래프 이미지에 대한 수치데이터 세트들을 첨부하는 등 유연하게 작성이 가능합니다. 논문에 공개되는 내용들이라면 해당되는 데이터들도 공개가 가능하다고 생각합니다.” (면담자 8)

두 번째 하위범주로 ‘데이터 소유권’에 대해서는 면담 참여자 대부분이 ‘나의 연구 자산, 나의 가장 소중한 자산, 우리 연구팀 것’ 등으로 설명하였으며 이를 통해 개인 또는 팀 단위로 소유권을 인식하고 있음을 알 수 있었다. 연구팀 협업을 통해서 생성된 데이터의 소유권이 개인에게 있다고 명확히 정하기는 힘들지만 보편적으로 1차 생산자가 소유자인 것으로 인식하고 있었다.

‘데이터 소유권’과 논의가 이어질 수 있는 부분은 세 번째 하위범주 ‘수탁기관과의 협약 사항’이다. 연구비를 지원하는 수탁기관과의 협약서에 명시된 계약 사항은 연구자 대부분이 가장 중요하게 인지하고 준수하는 기준이다. 협약서에 명시된 연구성과물은 필수적으로 제출하는데 대표적 형태가 보고서이며, 장비, 논문, 특허 등 지식재산권 등이 있다. 면담자 5는 “데이터가 언제, 어떻게 기술이전이 될지 모르기 때문에 보안을 잘 유지해서 기술화한 다음에 공개나 공유를 생각”한다고 응답하였다. “연구팀에서 결과를 잘 내야 후속 연구도 지속적으로 이루어질 수 있고 계속해서 과제를 수주”할 수 있기 때문에 “공개보다는 관리량 보안에 더 치우치게” 된다고 설명하였다. 또한 연구데이터는 연구성과물에 포함되지 않기 때문에 일반적으로 협약서에 명시된 결과물만 제출할 뿐 원천기술을 담고 있는 핵심데이터나 정보들은 연구팀 내 노하우로 남겨두는 경향을 보였다.

“연구데이터 관련한 수행사항을 수탁기관에서 전달받아본 적이 없습니다. 대부분 정보 유출이나 보안등에 기준을 엄격히 하는 편이기 때문에 공개나 공유에 대한 부분은 특별히 없었던 것 같습니다.” (면담자 1)

면담자 7은 수탁기관에서 DMP 작성을 지시받은 적이 있었으며 면담 참여자 중 유일하게 DMP 작성 경험을 가지고 있었다. 그러나 DMP 작성을 수탁기관의 협약 사항으로 포함하여 제도적으로 강제하는 측면에 대해서는 부정적인 인식을 갖고 있음을 알 수 있었다.

“최근 과기부 수행과제 중에 데이터관리계획을 제출하라는 요구가 있어서 작성하여 제출한 경험이 있습니다. 연구과제 종료 후 제출하도록 되어 있으므로 아직 연구데이터를 실제로 제출하지는 않았습니다. 어떤 데이터를 몇 개 생성하여 제출할 것인지에 대한 내용을 기재하긴 했지만 내라고 했기 때문에 형식적으로 적었지 실효성이나 가치가 없다고 생각하는 편입니다. 그리고 언제 어떤 방식으로 제출해야 하는지에 대한 구체적인 안내는 없었어요. 이런 부분을 규정으로 해결하려고 하면 역으로 연구를 저해하는 요소가 될 수도 있고, 형식적인 과정이 될 수도 있다고 생각합니다.” (면담자 7)

실제로 면담자 7이 작성하여 제출한 DMP를 살펴보면 <표 3>과 같다. 최명석과 이상환(2020)이 제안한 DMP 구성요소는 1) 연구과제 개요(연구과제명, 수행기간, 주관기관, 책임자, 연구개요); 2) 연구데이터 개요(데이터명, 유형, 포맷, 용량, 데이터 관련 표준, 메타데이터 관련 표준, 연구데이터 생산 및 수집 방법, 사용된 소프트웨어 및 장비); 3) 연구데이터의 저장 및 보존계획(연구수행 중 데이터 저장 및 관리 방안, 연구종료 후 장기적인 데이터 보존계획, 연구데이터 보존 필요기간); 4) 연구데이터 공유계획(데이터 공개시기, 공개방법, 데이터 식별자 정보, 접근 및 이용 조건, 연구데이터 공개 및 공유 제한 사항); 5) 연구데이터 관리책임자 이다.

<표 3> DMP 작성 사례

연구데이터 관리계획(DMP: Data Management Plan)															
사업 명	OO 및 OO 기술개발사업														
과제 명	OOOO 상용화 기술 및 OOOO 계획 수립														
주관연구기관명	OOOO	주관연구책임자 성명	김OO												
총 연구기간	2021. 04. 00 ~ 2023. 12. 00														
주관연구책임자 연구자등록번호	000000	과제번호													
I. 연구데이터 생산	※ 실험데이터와 메타데이터를 하나로 묶어서 하나의 데이터로 표시 - DMP작성 가이드라인 참고 1. 응용분야(복수 선택 가능) <input type="checkbox"/> 에너지·환경 <input checked="" type="checkbox"/> 스마트·IT(반도체, 전기전자 등) <input type="checkbox"/> 안전·구조 <input checked="" type="checkbox"/> 자동차·항공·우주 <input type="checkbox"/> 바이오헬스 <input type="checkbox"/> 기타() 2. 데이터 분류(복수 선택 가능) <input checked="" type="checkbox"/> 실험데이터 <input type="checkbox"/> 계산데이터 3. 데이터 형태 및 건수 <table border="1" style="width: 100%; margin-top: 10px;"> <thead> <tr> <th>형태</th> <th>건수</th> <th>비고 * 메타데이터가 없는 경우만 사유 작성</th> </tr> </thead> <tbody> <tr> <td>결합 이미지</td> <td>10</td> <td></td> </tr> <tr> <td>적층 이미지</td> <td>5</td> <td></td> </tr> <tr> <td>결합 측정 데이터</td> <td>5</td> <td></td> </tr> </tbody> </table>			형태	건수	비고 * 메타데이터가 없는 경우만 사유 작성	결합 이미지	10		적층 이미지	5		결합 측정 데이터	5	
형태	건수	비고 * 메타데이터가 없는 경우만 사유 작성													
결합 이미지	10														
적층 이미지	5														
결합 측정 데이터	5														
II. 연구데이터 제공	② 연구데이터 관리 및 제공 계획	▪ 관리 계획 ※ 데이터 제공 전/후의 데이터 관리 계획 선택(복수선택 가능) <input checked="" type="checkbox"/> 자체 저장공간에 저장 <input type="checkbox"/> 지정기관의 연구자 저장공간에 저장 <input type="checkbox"/> 기타() ▪ 제공 계획 (변경 불가) 연차실적계획서 및 최종보고서 제출 기간에 제공													
III. 연구데이터 공유 및 제한	③ 연구데이터 공유 및 제한 계획	※ 제공 후 공중에 공개까지 필요한 embargo 기간 선택 <input type="checkbox"/> 6개월 <input type="checkbox"/> 9개월 <input checked="" type="checkbox"/> 12개월 <input type="checkbox"/> 기타(embargo 기간 및 사유 기재)													
IV. 연구데이터 등록	④ 연구데이터 등록	▪ 등록 지정기관(000,000,000,000 또는 000)에 접속 후 업로드(변경 불가) ※ 메뉴얼은 등록사이트에서 다운로드													
V. 연구데이터 관리책임자		▪ 데이터 실무 책임자 정보(소속/이름/연락처(email, 전화번호))													

이를 기준으로 <표 3>의 실제 작성된 DMP의 구성요소를 살펴보면 연구데이터 관리와 공유에 있어 중요한 요소들이 상당수 누락되어 있음을 알 수 있다. 누락된 요소에는 연구개요 등 연구 맥락에 대한 정보, 데이터명, 포맷, 용량, 데이터 및 메타데이터 관련 표준, 사용된 소프트웨어 및 장비 등 데이터 생산 관련 정보, 연구종료 후 장기적인 데이터 보존계획과 보존 필요기간, 데이터 공유와 관련된 데이터 식별자 정보와 접근 및 이용조건이 포함된다. 면담자 7이 언급하였듯이 DMP 작성이 연구자들에게 “언제 어떤 방식으로 제출하는지에 대한 구체적인

안내” 없이 “형식적인 과정”으로 인식될 가능성이 높으므로 이에 대한 개선이 필요함을 알 수 있다.

보안과제를 수행하는 경우에는 연구데이터의 소유권이나 관리 주체, 관리 가능성에 대한 좀 더 세밀한 주의를 기울여야 한다. 군 관련 비밀과제를 수행해본 경험이 있는 면담자 5에 따르면 “보안등급에 해당되는 데이터를 생산”하는 경우 해당 데이터를 등록해서 생산하고 생산된 보안데이터 목록을 작성하여 수탁기관에 제출하며 보안데이터의 보관기간이 지나면 폐기한 후 보안데이터 목록에 서명을 하고 폐기한 사실을 수탁기관에 알리는 절차를 진행한다고 하였다. 이를 위해 보안데이터는 “비밀을 취급하는 방에 특정 PC에서만 작업”을 하다가 과제가 종료되고 보관기간이 지나면 PC에서 삭제하는 등 별도로 관리한다고 하였다. 데이터가 “공유되고 이어지면서 보관”되는 것이 바람직하지만 보안이 필요한 데이터는 수탁기관에서 요구하는 대로 관리할 수밖에 없다고 설명하였다.

4.2.3 연구데이터 관리의 효용성 및 효과성

<표 2>의 세 번째 주요범주인 ‘연구데이터 관리의 효용성 및 효과성’의 하위범주인 ‘데이터 관리 수행’에 대해서는 면담 참여자들 모두 필수적이고 중요하다고 인식하고 있었으며, 연구팀 내에서의 데이터 공유 및 전달을 위해서는 디지털 형태로 잘 정리된 데이터를 선호함을 알 수 있었다. 면담자 2는 “자동 또는 수동으로” 데이터 관리를 수행하는데 “수기로 연구노트 등에 필기를 한 다음 엑셀에 데이터를 하나씩 기록”하는 방식이라고 설명하였다. 개별 연구자의 관련 소프트웨어 활용역량에 따라 데이터 조직 및 정리에 대한 편차가 존재한다는 의견도 있었다.

“데이터의 획득과 생산은 DAQ(Data Acquisition System) 장치를 통해 방대한 양의 데이터를 샘플링하여 획득할 수 있어 요. 그런데 획득된 데이터에 대한 조직화나 후처리 작업들은 연구자 개인이 수동 또는 자동화 tool로 하기 때문에 개인에 따라 편차가 있을 수밖에 없습니다. 재이용의 관점에서 보면 자동화된 기계가독형 데이터가 용이하고 이렇게 정리되어 보존된 데이터들을 통해서 이전에 연구를 어떻게 했는지 그런 것들을 파악하고 판독할 수 있습니다.” (면담자 1)

또한 관행적으로 필요에 의해서, 필요한 만큼 데이터가 관리되고 있기 때문에 개인 또는 연구팀별 편차도 큰 것으로 나타났다. 면담 참여자들 중 3년 이상 경력의 연구자의 경우 최소 5개에서 많게는 22개의 연구과제를 동시에 수행하고 있었다. 이들은 연구과정에서 많은 규제와 기준을 준수하면서 최선의 연구결과를 도출하고자 노력하고 있지만 인프라, 인력, 시간, 연구환경의 한계 등으로 체계적으로 연구데이터 관리를 수행하기 어렵다고 응답하였다.

인프라는 두 번째 하위범주인 ‘보관방식’과도 연계되는데 데이터 보관방식에 대해서 대다수 면담 참여자들이 여러 가지 한계가 많다고 답변하였다. 가장 많이 언급된 부분은 보안규정과 충돌과 물리적 인프라 마련에 대한 지원이다. 일반적으로 업무용 PC나 백업 용도의 저장장치인 하드디스크에 연구데이터를 개별적으로 보관하고 있었으며 소속기관의 보안규정에 따라 보안매체만을 사용하는 경우도 있었는데 불편함이 많다는 의견이었다.

“기관의 보안규정에 따라서 개인이 그냥 보안 USB, 보안 외장하드 같은 것만 사용합니다. 엄청 불편하고 구식의 방식이라고 생각하는데 팀원한테 전달할 때도 거의 메일로 하고, 용량이 크면 보안매체를 사용하고는 합니다.” (면담자 2)

따라서 면담 참여자들은 기관에서 연구데이터 관리를 위한 시스템을 구축하는 것이 필요하다고 제안하였다. 면담자 6은 시스템 구축을 통해 “기관의 자산이 쌓여가는 거고 어떤 연구가 어떻게 되었는지 히스토리도 다 알 수가 있고 추적”이 가능하다는 것을 장점으로 꼽았다. 면담자 4의 경우 연구자들이 연구비로 개별적인 연구데이터 관리시스템을 구축하는 것은 한계가 있으므로 소속기관에서 이를 지원해주는 것이 필요함을 언급하였다.

“수탁기관에서 지원되는 연구비를 통해서 연구데이터 관리에 필요한 물리적 인프라를 마련하는 것은 직접적으로 연구과제에 필요한 것인지 증빙해야 하므로 연구비 처리가 매우 까다롭습니다. 연구팀이 수행하는 과제가 다수인 경우 그것들을 전부 분리해서 보관하는 게 현실적으로 한계가 있어요. 이런 것들은 기관차원에서 관리와 저장을 위한 인프라 지원이 필요하다고 봅니다.” (면담자 4)

이러한 한계에도 불구하고 연구팀 내에서 자체 DB를 만들어서 관리하는 사례도 확인할 수 있었다. 22개 과제를 진행하는 면담자 5에 따르면 자체 DB는 “PC 한 대에 지정이 돼있고 암호가 있어서” 연구팀 소속연구자만 “폐쇄적으로 접근할 수 있고 폐쇄적으로 운영”하는 방식으로 관리된다고 하였다. 연구데이터는 연구팀의 “연구자산”이며 데이터를 “계측한 결과를 DB”로 구축하여 “나중에 다른 과제할 때 DB를 찾아서 데이터를 확인”하는 등의 용도로 사용한다고 하였다. 앞서 면담자 4가 언급한 대로 이러한 DB 구축은 국가연구개발사업으로는 어려우나 면담자 5는 민간 수탁기관에 해당 DB를 만들 수 있는 소프트웨어를 납품한 후 이를 활용하는 방식을 취했기 때문에 가능하였다고 설명하였다.

연구데이터 관리와 저장에 대한 인프라 지원은 보안상 문제가 없는 클라우드 방식을 선호한다는 의견이 있었다. 면담자 6에 따르면 대규모 연구 수행에 대한 요구가 높기 때문에 다수의 참여연구원들과 함께 연구를 진행해야 하는데 각자 수행하는 연구 내용과 데이터를 효율적으로 공유하기 위해서는 보안이 해결된 클라우드 시스템 등의 인프라가 바람직하다는 의견이었다.

“제가 하는 연구에 참여연구원이 30명 가까이 되요. 각자 뭘 하는지 알 수가 없는데 이런 것들을 알려면 클라우드나 어떤 공유되는 플랫폼이 있어야 되거든요. 프로젝트라게 a-b-c-d 연결해서 시스템을 만들어서 연계가 되어야 되는데 이럴 때 공유는 되게 필요하죠. 뭔가 협업을 하려면 우리가 갖고있는 정보를 서로 내놓아서 공유가 되어야 하는데 그걸 어떻게 체계화된 방법으로 효율적으로 할 거냐가 문제인거죠. 대단위 규모의 연구를 하라고 기관에서 원해요. 우리가 컴퓨터를 활용하는 것처럼 우리 조직도 제2의 뇌가 필요해요. 내가 하는 업무뿐 아니라 다른 사람이 나랑 어떻게 연결되는지 이해해야 하는데 이런 정보들이 공통적으로 활용할 수 있는, 보안이 해결된 공유할 수 있는 어떤 시스템이 필요하죠. 인프라가 없는거예요.” (면담자 6)

세 번째 하위범주인 ‘보존기간’에 대해서는 면담자들의 분야마다 조금씩 다르게 인지하고 있지만 영구보존에 대한 의견이 다수인 것으로 나타났다. 또한 데이터의 유용성과 재이용의 가치가 유지되는 기간과 실제 데이터 폐기를 수행할 수 있는 기간에 대해서는 다르게 인지하고 있음을 알 수 있었다. 일반적으로 데이터의 유용성이 유지되는 기간은 10년에서 20년 정도로 인지하고 있었으며 폐기 기한은 거의 고려하지 않는 것으로 나타났다.

“연구분야나 연구데이터에 따라 다르겠지만 보통 3년 단위로 연구과제들이 수행되니까 초기 연구과제 이후 2번 정도의 후속연구를 진행하면 연구기간이 약 9년 정도 됩니다. 그래서 데이터의 유용성은 10년 정도라고 보는데, 그렇다고 그 이후에 바로 폐기하지는 않으니까 20년 정도 되지 않을까 생각합니다.” (면담자 2)

연구데이터 폐기를 고려하지 않는 면담 참여자들이 대다수였으므로 획득·저장된 연구데이터는 무기한 보존한다는 답변이 많았다. 면담자 5는 “영구보존이 원칙”이며 “100년 전 데이터”의 활용성에 의문을 제기할 수도 있겠지만 “현재 데이터와 비교해서” 변화 추이를 분석하려면 100년 전 데이터도 의미가 있을 것으로 보았다. 면담자 8은 연구데이터를 “하나 얻기 힘들고 비싸고 귀한 몸”이며 “진리가 숨겨져 있기 때문에 두고두고 잘 살펴봐야” 하는 대상이라고 언급하였다. 면담자 8은 특히 논문에 출판된 연구데이터의 경우 “영구보존이 필수적”이라고

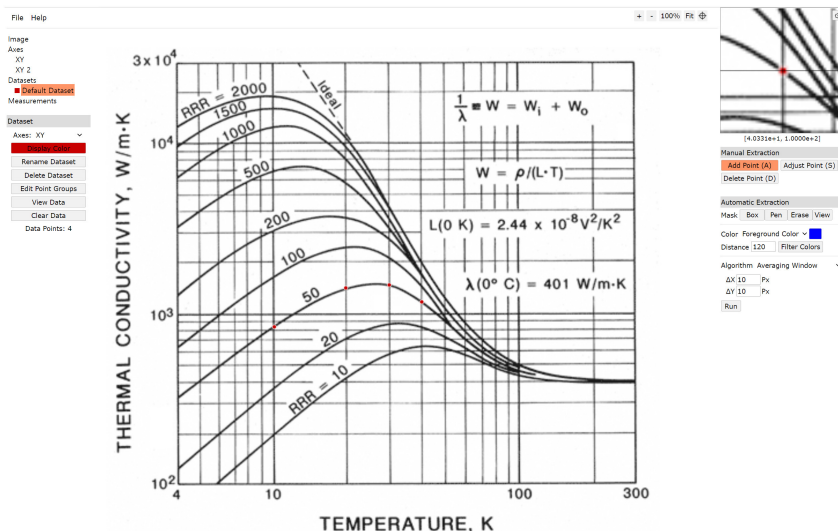
하였는데 “계속 인용되고 이용될 가치”가 있기 때문이라고 하였으며 “관리된 연구데이터는 자산이고 요즘은 저장 매체의 성능도 좋아졌기 때문에 보존기간은 무한하다”고 설명하였다.

기록관리 업무에서는 생산 시 업무에 대한 중요도에 따라 기록의 보존기간을 부여하고 역사적, 심미적 가치가 있는 소수의 기록을 제외하면 대부분의 기록이 업무가치에 따라 평가되고 업무가치가 종료하면 폐기된다. 그러나 면담 참여자들은 연구데이터의 현용 가치뿐 아니라 미래의 이용가능성과 그에 따른 잠재적 가치를 높게 평가하고 있었고 이에 따라 보존기간 역시 10년 이상 혹은 영구보존으로 제시하고 있었다. 이에 따라 면담 참여자 중 획득·저장한 데이터에 대해 폐기를 수행한 경우는 없는 것으로 나타났다.

4.2.4 연구데이터 공유의 가치

<표 2>의 네 번째 주요범주인 ‘연구데이터 공유의 가치’에 대해서는 데이터의 유출 또는 오남용으로 인해 발생할 문제들에 대한 우려가 공개와 공유의 가치와 이점에 대한 인지보다 높은 것으로 파악되었다. 또한 데이터를 사용해서 올바른 해석과 재이용이 가능한 자격이 있는 전문가로 식별되는 경우에 한해서만 데이터가 공유되어야 한다는 답변이 다수를 차지하며 데이터 공개와 공유에 대한 면담 참여자들의 인식은 매우 보수적인 것을 알 수 있었다. 면담자 5에 따르면 최근 연구데이터의 중요성이 높아지면서 연구팀마다 로우데이터는 공개하지 않고 공개하더라도 정제된 데이터만을 공개하는 경우가 많다고 하였으며, 특히 “데이터를 다 활용하지 못한 상황에서 오픈”하지는 않는다고 하였다. 심지어 공개된 “보고서 내에서도 필요한 부분만” 제시하고 논문 출판이나 기술 이전 등의 목적으로 활용될 수 있는 데이터는 감추는 경향이 있다고 설명하였다. 이와 같이 누구나 볼 수 있게 데이터를 공개하는 것에 대한 거부감은 높았지만 신뢰할 수 있는 연구자들에게는 데이터를 “100% 다 공개하는 게 업무의 효율을 높이는 방법”(면담자 6)이라는 의견도 있었다.

면담 참여자들은 타인의 데이터를 재이용하는 경우 출판된 논문이나 보고서를 참고한다는 답변이 대부분이었고, 데이터 공개 및 공유 역시 논문과 보고서를 통해 하고 있다고 응답하였다. 논문에서 공개된 그래프나 이미지에서 참고할만한 수치를 확인하기 위해서 연구자들은 ‘Digitizer Tool’이라는 소프트웨어를 사용해서 수치데이터를 추출하여 참고하고 있었다. 예를 들어 <그림 7>은 면담자 1이 구리 재료의 극저온 물성치를 획득하기 위해 논문의 그래프를 활용하여 데이터를 추출하는 과정을 보여준다. 이와 같이 해당 물성치를 이용하여 알고자 하는 현상의 수치해석에 활용하거나 실험결과 분석 등에 활용한다고 하였다.



<그림 7> Plot Digitizer를 활용한 그래프 수치 추출

그렇지만 일부 면담 참여자들의 경우 사적인 네트워크를 활용하여 데이터를 전달받은 경험이 있었으며 저널에 논문 투고 시 연구데이터 제출을 요구받거나 제출한 연구자들도 있었다. 또한 데이터의 분석 결과가 아닌 데이터의 생산 및 획득 방법과 데이터 특성 등 데이터에 대한 사실만을 기술하는 데이터 논문(data papers)을 출판하는 데이터 학술지(data journal)에 대한 질문에는 면담 참여자 중 1명만이 이를 인지하고 있었다. 데이터 학술지는 데이터에 대한 설명만을 논문의 형식으로 출판하고 데이터 자체는 데이터 리포지터리에 기탁하여 공개하는 방식으로 운영된다. 이를 통해 연구자들이 데이터 공유를 연구성과로 인정받을 수 있다는 점에서 데이터 학술지는 데이터 공유를 촉진할 수 있는 출판 모형으로 주목받고 있다(Kim, 2020). 면담 중에 데이터 학술지를 알게 된 참여자 중에는 데이터 학술지 관련 정보를 전달해달라고 요청하는 경우도 있었다. 연구성과에 대한 압박이 높은 국내 연구환경에서 데이터 학술지는 데이터 공유 방식 중 하나로 수용될 수 있으며 실제로 한국지질자원연구원에서는 Geo Data라는 데이터 학술지를 창간하여 운영 중이다(Geo Data Council, 2019).

연구데이터 공유에 대한 세계적 흐름에 따라 기계공학 분야의 해외 저널들에 논문 투고 시 관련 연구데이터 제출에 대한 요구도 늘고 있으며 연구윤리와 연구 신뢰도를 강화하는 측면에서도 연구데이터 공유가 더욱 강조되고 있는 실정이다. 그러나 일부 면담 참여자들은 논문을 통해 어떤 설계도를 개략적으로 공개하고 그에 대한 기능설명, 명칭 제시 정도만 설명하였는데도 이러한 개략도만을 바탕으로 무차별하게 아이디어를 도용하고 제품을 만드는 일들이 발생하고 있다고 하였다. 연구데이터의 공개와 공유는 기술 유출과 아이디어 도용을 야기할 수 있고 연구자들의 땀과 노력을 한순간에 물거품으로 만들어 버릴 수 있는 여지도 분명히 존재한다. 이들은 연구데이터를 전부 공개하고 내어놓으라는 것은 도둑질과 같은 행위라고 강력하게 성토했기도 하였다.

이상과 같이 기계공학 분야에서 생산되는 연구데이터는 단시간에 대량으로 생산되는 수치데이터가 대부분이며 이를 바탕으로 하는 해석 데이터와 모델링 데이터가 주로 생산됨을 알 수 있었다. 면담 참여자들은 현재와 이후의 활용을 위해 데이터를 이해하고 해석하는데 필요한 맥락 정보를 기술하는 것이 데이터의 신뢰도와 품질을 결정하는 핵심 요소라고 인식하고 있었으며 이를 기술하고 정리하는데 많은 노력을 기울이고 있었다. 면담 참여자들은 수탁기관의 요구사항을 철저히 준수하고 있었으며 보안을 유지해야 하는 연구과제가 많은 기계공학 분야의 특성상 데이터의 공개나 공유는 연구팀 내의 공유 등으로 제한적인 경우가 대부분이었다. 이러한 결과들은 상세한 메타데이터 기술을 강조하고 상업적 이익 보호를 위해 보안을 유지하는 등 데이터 공개와 공유의 제약조건이 많다는 공학 분야의 선행연구 결과와 맥을 같이 한다(Mallasvik & Martins, 2020; Suhr, Dunl, & Stocker, 2020). 또한 면담 참여자들은 소속기관에서 연구데이터 관리를 지원하는 시스템을 구축하는 등 기술적인 인프라에 대한 필요성을 강조하였으며 연구데이터는 10년 이상 혹은 영구적으로 보존해야 하는 기관의 가치 있는 자산으로 인식하고 있었다. 면담 참여자들은 연구데이터의 오용 및 기술 유출에 대해 우려를 표시하였으며 공유된 데이터를 적절하게 활용할 수 있는 자격을 갖춘 믿을 수 있는 연구자들에게만 데이터를 공유하겠다는 인식이 강한 것으로 나타났다. 이는 면담 참여자들이 Zuiderwilk, Shinde, & Jeng(2020)과 Van Panhius et al.(2014)가 공통적으로 제시한 데이터 공유의 영향 요인 중 신뢰와 관련된 사항을 중시하고 있음을 보여준다.

5. 결론 및 제언

본 연구에서 조사한 해외사례와 면담내용을 바탕으로 서론에서 언급한 연구문제 1, 2, 3에 대한 답을 정리하고, 연구문제 4에서 제기한 기계공학 분야 연구데이터 관리 및 서비스를 위한 고려사항을 다음과 같이 제시한다. 첫째, 기록관리를 위해 사전에 기록조사를 실시하듯이, 연구데이터 관리를 위해서는 생산되는 데이터들의 유형과 형태에 대한 조사를 실시하여 이에 대한 공통적인 메타데이터를 구성하는 ‘표준 메타데이터 레지스트리’ 등을 제시하는 방안을 고려할 수 있다. 하지만 연구문제 2에서 다룬 기계공학분야 연구데이터 특징을 살펴보면 실험,

분석 및 해석, 관찰을 통해 동시다발적으로 대량 생산된 수치데이터의 형태를 가지며 다양한 파일확장자를 가지는 특성이 있기 때문에 공통적인 메타데이터를 일괄적으로 확정하는 것은 매우 어려운 일이며 불가능할 수 있다. 따라서 연구팀 단위 정도의 미시적 수준으로 연구 수행과정에서 관행적으로 연구데이터에 대한 설명을 제시하는 기본적인 메타데이터 항목을 파악할 수 있을 것이다. 연구데이터를 저장하고 관리하는 DB에서는 최대한 유연하게 이를 수용할 수 있어야 한다. 명시적 메타데이터 항목들에 대해서는 호주 CSIRO의 Data Access Portal에서 컬렉션 데이터에 대한 기술항목으로 제시된 ‘설명, 생성기간(날짜), 접근권한, 관련 자료, 연구과제정보(연구명, 연구책임자, 수행기관, 설명), 이용된 소프트웨어 정보’ 등으로 이루어져 있음을 알 수 있었고 <그림 5>, <그림 6>과 같은 실험메모, 실험시트, 기술노트 등을 참고할 수 있다. 또한 암시적 맥락정보는 자동화된 툴을 사용하여 확인 및 도출할 수 없는 정보이므로 기관의 기록관리 차원에서 고정력 연구자들이나 퇴직을 앞둔 연구자들을 대상으로 연구데이터 관리를 위한 기록화 사업 등을 추진하여 암묵지 공유를 지원하는 것도 하나의 방안이 될 수 있다. 이런 작업들을 통해서 더욱 풍부한 연구데이터 관리 및 서비스 DB를 설계할 수 있다.

연구문제 3에서 다룬 연구데이터 관리에 대한 방해요인으로는 관리방법의 어려움, 연구자 개인의 전적인 책임, 물리적 및 제도적 인프라의 한계 등이 있고, 동기요인으로는 연구데이터의 가치, 관리된 연구데이터의 올바른 해석과 재이용의 가능성, 데이터 학술지 출판을 통한 성과 인정 및 확산 등으로 파악되었다. 동기요인을 강화하기 위해서는 연구데이터 공개와 공유를 위한 연구자 인센티브로서 데이터 학술지에 데이터 논문으로 단독 출판하는 경우 이를 논문의 한 형태로 보고 연구실적으로 인정하는 방안을 고려해볼 수 있다. 출판 시에 생산자 및 기여자에 대한 저자 정보를 포함하고 내외부적인 연구실적으로 인정하는 방향이 된다면 연구데이터 공개와 공유에 대한 인식과 참여도를 높일 수 있을 것이며 연구데이터 관리와 재이용의 편리성 또한 개선할 수 있을 것이다. 현재 기계공학 분야를 포함하는 데이터 출판이 가능한 저널에는 Elsevier 사의 ‘Data in Brief’와 SpringerNature 사의 ‘Scientific Data’가 있다. ‘Data in Brief’는 2014년에 창간된 학제적 저널로 계간 발행되고 있으며 오픈액세스 저널 형태로 발간된다. 2020년까지 약 6,200여건의 데이터 기사가 출판되었으며 공학분야 약 2,100여건 중 기계공학 분야는 약 500여건 정도이다. 데이터 기사의 구성항목은 ‘초록, 키워드, introduction, Material and methods, Results, Discussion and conclusions, 부록(보충자료)’로 이루어진다. 기사에 제출된 고해상도 이미지, 수치값으로 이루어진 표나 그래프 등에 대한 다운로드를 제공한다(Elsevier, 2021). ‘Scientific Data’는 2014년 창간된 오픈액세스 저널로 주요 출판 기사유형은 ‘Data Descriptor’ 형태로 데이터에 대한 선별되고 구조화된 메타데이터와 전통적 서술 기사를 결합한 형태이다. 구성항목은 ‘초록, 배경 및 요약, 방법, 데이터레코드, 기술적 유효성, 재이용 참고사항, 코드 가용성(접근제한, 데이터 도출에 사용된 SW버전, 데이터세트를 생성 또는 처리하는데 사용된 특정 변수 또는 매개변수등에 대한 정보), 참고문헌, 관련된 데이터 세트에 대한 설명’ 등으로 이루어지며 이들은 기계가독형 데이터로 제공된다. 2021년 5월부터 공학분야를 포함하도록 분야를 확장시켰으며 2021년 8월까지 약 7,100여건의 공학분야 기사가 발표되었으며 그 중 기계공학 분야는 약 1,500여건 정도이다(Springer Nature, 2021).

또한 연구데이터에 대해 ‘모두에게 모든 것을 공유한다’는 목표보다는 작은 단위에서부터, 예를 들어 연구팀 단위에서 시작하여 연구기관 단위, 연구협의체 단위, 연구분야 단위와 같이 점점 범위를 넓혀나갈 수 있도록 하는 것이 효과적일 것이다. 연구데이터를 이해하고 접근할 수 있는 사람들부터 단계별로 시작하는 것이 연구자들의 연구데이터 공개와 공유에 대한 인식을 높이고 거부감을 낮출 수 있을 것이다. 작은 단위에서부터의 데이터 공유를 시작하는 것이 첫 단계이며 이를 위한 기관의 인프라 지원이 요구된다. 이러한 인프라 지원은 방해요인으로 파악된 물리적 한계를 극복하기 위한 방안으로도 제시될 수 있는데 기관에서는 보안에 문제가 없는 클라우드 등을 활용하여 연구데이터 관리와 서비스를 지원할 수 있는 인프라 마련을 고려해야 한다. 여기에는 연구데이터 업로드, 관리, 저장을 연구자가 개별적이며 자율적으로 할 수 있어야 하며 연구팀 내 공유를 지원하는 기능이 함께 제공되어야 한다. 전담 및 관리 조직에서는 리포지터리에서 공개 및 공유가능한 부분에 대해 현재 존재하는 정보서비스 창구(플랫폼 형태 또는 홈페이지 형태)에 연계하여 서비스를 제공할 수 있는 기준을 마련할 필요가 있고, 이에 따른 DB 백업 및 장기보존, 대외 서비스를 지원해야 한다. 독일 프라운호퍼연구소의 ‘Fraunhofer-Publica 데이터베이스’

와 'Fordatis'를 참고하여 안전하게 연구자들의 연구데이터 관리를 수행할 수 있고, 이를 통해 협업을 지원하는 편리하고 원활한 시스템이 제공된다면 연구자와 기관, 나아가 과학기술계의 지식정보 전수를 위한 초석을 마련할 수 있을 것이다.

제도적으로 강제하는 연구데이터 의무 제출과 공개·공유는 당연히 연구자들의 반발을 일으킬 수 있으며 현 상황에서 그것을 강제할 수는 없는 실정이다. DMP와 관련해서도 연구데이터 관리에 있어 중요한 정보인 데이터 생산, 보존 및 공유계획에 대한 내용 중 일부만이 연구자들에게 요청되고 있었으며 이에 대한 안내나 교육이 부족한 상황에서 의무적인 DMP 제출 요청은 형식적인 과정으로 인식될 가능성이 높다. 연구문제 1을 통해 살펴본 기계공학분야 연구자들의 연구데이터 생산과 관리에 대한 인식과 행태를 보면 연구데이터를 소중한 자산으로 인식하고 있으며 가능한 한 효율적이고 효과적으로 생산과 관리 전 과정을 수행하고자 하는 의지가 있다. 그러나 연구데이터 관리는 연구자만의 책임이 아니고 연구를 설계하고 감독하는 연구책임자, 연구를 수행하고 데이터를 수집, 분석하는 참여연구자, 연구 자금 관리를 지원하는 직원, 데이터 저장 및 보안 백업을 제공하는 IT 담당자, 데이터 관리와 공유를 위한 데이터센터 및 기관 리포지터리 운영 조직 등 다양한 사람들이 연구 과정에 참여하고 연구데이터를 보호하고 공유를 촉진하는 역할을 할 수 있다. 따라서 연구 현장에서 다양한 분야의 전문가들과 다각도로 검토하며 조직적 차원에서 연구데이터 관리와 서비스에 대한 역할과 책임을 배분하는 것이 중요하다.

연구데이터의 공개와 공유, 개방의 움직임은 오픈 사이언스 운동의 확산과 함께 앞으로 계속될 것으로 예상된다. 그러나 오픈 데이터를 촉진하는 현재의 움직임은 산업계와 연계되는 기계공학 분야의 연구 특성상 데이터 공유와 공개에 있어 자유롭지 못한 연구자들을 불편하게 하는 측면도 있다. 과학기술계는 패러다임 전환을 맞이했지만 아직 준비되지 않았고 변화된 환경에 충분히 적응하지 못하고 있다. 이러한 상황에서 연구데이터를 어떻게 규정하고 생성된 정보를 어떻게 수집하여 처리할 것인지를 세밀하게 고려해야 하며 내실 있는 연구데이터의 관리와 서비스를 위해서는 관련 이해관계자들의 참여와 합의에 기반을 두어 제도-인프라-합의의 삼박자를 갖추는 것이 필요하다. 정부출연연구기관의 핵심 정보자산인 연구데이터 관리와 서비스가 활성화 될 수 있도록 다양한 학문분야에 대한 지속적인 후속연구와 실무적인 노력이 필요할 것이다.

참고문헌

- 강주연 (2017). 생명공학분야 연구데이터 관리 방안 연구. 석사학위논문, 전북대학교 일반대학원 기록관리학과.
- 과학기술분야 정부출연연구기관 기록관리협의회 (2021). 과학기술분야 연구기관 연구기록관리의 현재와 과제. 2021 KIRD LearningLab Program. 국가과학기술인력개발원.
- 과학기술정보통신부 (2020). 국가연구개발정보처리기준. 과학기술정보통신부고시 제2020-102호.
- 과학기술정보통신부 (2020). 국가연구개발혁신법. 법률 제 17343호.
- 국가과학기술심의회 운영위원회 (2018). 연구데이터 공유 활용체계 구축전략. 출처: https://www.pacst.go.kr/jsp/m/council/councilPostView.jsp?post_id=1096&etc_cd1=COUN01&cpage=6&board_id=11#this
- 권나현, 이정연, 정은경 (2012). 과학기술분야 R&D 전주기 연구. 한국문헌정보학회지, 46(3), 103-131.
<http://dx.doi.org/10.4275/KSLIS.2012.46.3.103>
- 김은정, 남태우 (2012). 연구데이터 수집에 영향을 미치는 요인 분석. 정보관리학회지, 29(2), 27-44.
<https://doi.org/10.3743/KOSIM.2012.29.2.027>
- 김지현 (2012). 대학 내 연구자들의 연구데이터 관리에 관한 연구. 한국도서관·정보학회지, 43(3), 433-455.
<http://dx.doi.org/10.16981/kliss.43.3.201209.433>
- 박미영, 안인자, 김준모 (2018). 생명공학분야의 연구데이터 공유사례에 관한 연구. 한국비블리아학회지, 29(1), 393-416.
<http://dx.doi.org/10.14699/kbiblia.2018.29.1.393>

- 박미영, 안인자, 남승주 (2018). 과학기술분야 출연연구기관 연구데이터 관리 및 공유 사례 분석 연구. 한국비블리아학회지, 29(4), 319-344. <http://dx.doi.org/10.14699/kbiblia.2018.29.4.319>
- 박형정, 김지현 (2021). 대학도서관의 데이터관리계획 서비스 고려사항에 관한 연구. 한국비블리아학회지, 32(3), 187-215. <https://doi.org/10.14699/kbiblia.2021.32.3.187>
- 신은정 (2018). 과학기술 연구데이터 공동 활용 기반 마련을 위한 연구. 과학기술정보통신부 과제 보고서. 과학기술정책연구원.
- 신은정, 손수정, 서지영, 김영린 (2019). 공공부문 연구데이터의 소유, 활용제도 개선방향, 정책연구, 2019(16), 1-216. 과학기술정책연구원.
- 신은정, 안형준, 양현채, 최병삼, 양승우, 정원교, 김수연 (2017). 오픈사이언스정책의 도입 및 추진 방안. 정책연구, 1-194.
- 안채영, 김지현 (2020). 정부출연연구기관의 연구기록 평가 개선방안 연구. 기록학연구, 66, 105-155. <https://doi.org/10.20923/kjas.2020.66.105>
- 이미영 (2017). 정부출연연구기관의 기록분류 현황과 개선과제. 기록학연구, 53, 229-259. <https://doi.org/10.20923/kjas.2017.53.229>
- 임진희 (2010). 설명책임 메커니즘을 통한 공공기관 업무정보의 추구 및 제공 과정에 관한 연구. 박사학위논문, 연세대학교 정보대학원 정보시스템 통합·관리 전공.
- 정지훈, 이강현, 강혜원 (2020). 국내·외 연구데이터 분야 연구 동향에 관한 연구. 제27회 한국정보관리학회 학술대회 논문집, 27-32.
- 최명석, 이상환 (2020). 데이터 관리 계획의 국내 현황 및 과제. 한국콘텐츠학회논문지, 20(6), 220-229. <https://doi.org/10.5392/JKCA.2020.20.06.220>
- 최명석, 이승복, 이상환 (2017). 국내 과학기술분야 연구기관의 과학데이터 관리 현황. 한국콘텐츠학회논문지, 17(12), 117-126. <https://doi.org/10.5392/JKCA.2017.17.12.117>
- 최준혁 (2014). 혼합적 연구방법을 활용한 상황분석의 일례. 홍보학연구, 18(4), 169-186. <http://dx.doi.org/10.15814/jpr.2014.18.4.169>
- 한국행정학회 (2001). 행정학전자사전. 출처: <https://kapa21.or.kr/bbs/dictionary/6626>
- Australia's National Science Agency (2020). CSIRO 2020 Year in Review. Available: https://www.csiro.au/-/media/About/Files/YearInReview2020_WEB.pdf
- Borgman, C. L. (2015). Big data, little data, no data: Scholarship in the networked world (1st ed.). Cambridge, MA: MIT Press.
- Burgelman, J. C., Pascu, C., Szkuta, K., Von Schomberg, R., Karalopoulos, A., Repanas, K., & Schouppe, M. (2019). Open science, open data, and open scholarship: European policies to make science fit for the twenty-first century. *Frontiers in Big Data*, 2(43). <https://doi.org/10.3389/fdata.2019.00043>
- Corti, L., Van den Eynden, V., Bishop, L., & Woollard, M. (2020). *Managing and Sharing Research Data: A Guide to good Practice* (2nd edition). California: SAGE Publications Ltd.
- CSIRO (2018). CSIRO Data Access Portal Notes Before Completing the Application. Available: <https://www.coretrustseal.org/wp-content/uploads/2018/10/CSIRO-Data-Access-Portal.pdf>
- CSIRO [n.d.]. CSIRO Data Access Portal. Available: <https://data.csiro.au/collections/>
- Elsevier (2021). Data in Brief. Available: <https://www.journals.elsevier.com/data-in-brief>
- Fecher, B., Friesike, S., & Hebing, M. (2015). What drives academic data sharing? *PLoS ONE* 10(2): e0118053. <https://doi.org/10.1371/journal.pone.0118053>
- Fernández-Miranda, S. S., Marcos, M., Peralta, M. E., & Aguayo, F. (2017). The challenge of integrating Industry 4.0 in the degree of mechanical engineering. *Procedia Manufacturing*, 13, 1229-1236. <https://doi.org/10.1016/j.promfg.2017.09.039>
- Fraunhofer [n.d.]. Fraunhofer-Fordatis. Available: <https://fordatis.fraunhofer.de>
- Fraunhofer [n.d.]. Fraunhofer-Publica. Available: <http://publica.fraunhofer.de/starweb/pub09/en/index.htm>

- Fraunhofer [n.d.]. Über Fraunhofer. Available: <https://www.fraunhofer.de>
- Geo Data Council (2019). 저널 소개. 출처: <https://geodata.kr/about/aimScope>
- Glesne, C. (2017). 질적 연구자 되기 5판. 안혜준 옮김. 서울: 아카데미프레스.
- Joo, Y. K. & Kim, Y. (2017). Engineering researchers' data reuse behaviours: a structural equation modelling approach. The Electronic Library, 35(6), 1141-1161. <https://doi.org/10.1108/EL-08-2016-0163>
- Kim, J. (2020). An analysis of data paper templates and guidelines: types of contextual information described by data journals. Science Editing, 7(1), 16-23. <https://doi.org/10.6087/kcse.185>
- KISTI, '제1회 과학기술정보포럼' 개최 (2019. 5. 22.). 대전일보, 출처: http://www.daejonilbo.com/news/newsitem.asp?pk_no=1370586
- Kowalczyk, S. & Shankar, K. (2011). Data sharing in the sciences. Annual Review of Information Science and Technology, 45(1), 247-294. <https://doi.org/10.1002/aris.2011.1440450113>
- Kurata, K., Matsubayashi, M., & Mine, S. (2017). Identifying the complex position of research data and data sharing among researchers in natural science. Sage Open, 7(3), <https://journals.sagepub.com/doi/pdf/10.1177/2158244017717301>
- Mallasvik, M. L. & Martins, J. T. (2020). Research data sharing behaviour of engineering researchers in Norway and the UK: uncovering the double face of Janus. Journal of Documentation, 77(2), 576-593. <https://doi.org/10.1108/JD-08-2020-0135>
- Müller, J. M., Veile, J. W., & Voigt, K.-I. (2020). Prerequisites and incentives for digital information sharing in Industry 4.0 - an international comparison across data types. Computers & Industrial Engineering, 148, 106733. <https://doi.org/10.1016/j.cie.2020.106733>
- NAA (2020). Records Disposal Authority CSIRO. 73-80. Available: <https://www.naa.gov.au/sites/default/files/2019-12/agency-ra-2002-04926193.pdf>
- NHMRC (2019). Management of Data and Information in Research: A guide supporting the Australian Code for the Responsible Conduct of Research. Available: <https://www.nhmrc.gov.au/sites/default/files/documents/attachments/Management-of-Data-and-Information-in-Research.pdf>
- OECD (2015). Making Open Science a Reality. OECD Science, Technology and Industry Policy Papers, 25, Paris: OECD Publishing. <http://dx.doi.org/10.1787/5jrs2f963zs1-en>
- Perrier, L., Blondal, E., & MacDonald, H. (2020). The views, perspectives, and experiences of academic researchers with data sharing and reuse: A meta-synthesis. PLoS ONE, 15(2), e0229182, <https://doi.org/10.1371/journal.pone.0229182>
- Research Data Alliance (2019). Research Data Management in Engineering IG Charter Statement. Available: <https://rd-alliance.org/group/research-data-management-engineering-ig/case-statement/research-data-management-engineering-ig>
- Springer Nature (2021). Scientific Data. Available: <https://www.springernature.com/gp/authors/research-data>
- Suhr, B., Dungal, J., & Stocker, A. (2020). Search, reuse and sharing of research data in materials science and engineering—A qualitative interview study. PLoS ONE, 15(9), e0239216. <https://doi.org/10.1371/journal.pone.0239216>
- Van Panhuis, W. G., Paul, P., Emerson, C., Grefenstette, J., Wilder, R., Herbst, A. J., Heymann, D., & Burke, D. S. (2014). A systematic review of barriers to data sharing in public health. BMC Public Health, 14(1), 1144. <https://doi.org/10.1186/1471-2458-14-1144>
- Wu, S., & Worrall, A. (2019). Supporting successful data sharing practices in earthquake engineering. Library Hi Tech, 37(4), 764-780. <https://doi.org/10.1108/LHT-03-2019-0058>
- Zuiderwijk, A., Shinde, R., & Jeng, W. (2020). What drives and inhibits researchers to share and use open research data? A systematic literature review to analyze factors influencing open research data adoption. PLoS ONE, 15(9), e0239283, <https://doi.org/10.1371/journal.pone.0239283>

• 국문 참고자료의 영어 표기

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

- Ahn, Chae-Young & Kim, Jihyun (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>
- Choi, June-Hyock (2014). How to Conduct Mixed Methods in Situation Analysis. *Journal of Public Relations*, 18(4), 169-186.
<http://dx.doi.org/10.15814/jpr.2014.18.4.169>
- Choi, Myung-Seok & Lee, Sang-Hwan (2020). Current Status and Issues of Data Management Plan in Korea. *The Journal of the Korea Contents Association*, 26(6), 220-229. <https://doi.org/10.5392/JKCA.2020.20.06.220>
- Choi, Myung-Seok, Lee, Seung-Bock, & Lee, Sang-Hwan (2017). Research Data Management of Science and Technology Research Institutes in Korea. *The Journal of the Korea Contents Association*, 17(12), 117-126.
<https://doi.org/10.5392/JKCA.2017.17.12.117>
- Gang, Ju-Yeon (2017). A Study on the Methods for Biotechnology Research Data Management. Master's thesis, Chonbuk National University.
- Glesne, Corrine (2017). *Becoming Qualitative Researchers: An Introduction* (5th edition). New York: Pearson College Div.
- Jeong, Ji-Yoon, Lee, Kang-Hun, & Kang, Hye-Won (2020). A Study on Research Data trends through Keyword Network Analysis. 27th Proceedings of the Korean Society for Information Management Conference, 27-32.
- Kim, Eun-Jeong & Nam, Tae-Woo (2012). Factor Analysis of Effects on Research Data Collection. *Journal of the Korean Society for Information Management*, 29(2), 27-44. <https://doi.org/10.3743/KOSIM.2012.29.2.027>
- Kim, Jihyun (2012). A Study on University Researchers' Data Management Practices. *Journal of Korean Library and Information Science Society*, 43(3), 433-455. <http://dx.doi.org/10.16981/kliss.43.3.201209.433>
- Kwon, Na-Hyun, Lee, Jung-Yeon, & Chung, Eun-Kyung (2012). Understanding Scientific Research Lifecycle: based on Bio- and Nano- Scientists' Research Activities. *Journal of the Korean Library and Information Science*, 46(3), 103-131.
<http://dx.doi.org/10.4275/KSLIS.2012.46.3.103>
- Lee, Mi-Young (2017). A Study on Current Status and Improvement Tasks of Records Classification in Government-funded Research Institutes. *The Korean Journal of Archival Studies*. 53. 229-259. <https://doi.org/10.20923/kjas.2017.53.229>
- Park, Hyeong-Jeong & Kim, Jihyun (2021). A Study on the Considerations for Data Management Plan Services in University Libraries. *Journal of the Korean Biblia Society for Library and Information Science*, 32(3), 187-215.
<https://doi.org/10.14699/kbiblia.2021.32.3.187>
- Park, Mi-Young, Ahn, In-Ja, & Kim, Jun-Mo (2018). A Study on Use Case of Research Data Sharing in Biotechnology. *Journal of the Korean Biblia Society for Library and Information Science*, 29(1), 393-416.
<http://dx.doi.org/10.14699/kbiblia.2018.29.1.393>
- Park, Mi-Young, Ahn, In-Ja, & Nam, Seung-Joo (2018). A Study on the analysis of Research Data Management and Sharing of Science & Technology Government-funded Research Institutes. *Journal of the Korean Biblia Society for Library and Information Science*, 29(4), 319-344. <http://dx.doi.org/10.14699/kbiblia.2018.29.4.319>
- Science and ICT (2020). National Research and Development information processing standards. Standards No. 2020-102
- Science and ICT (2020). National Research and Development Innovation ACT. Act No. 17343
- Shin, Eun-Jung (2018). A Study on Fundamentals of Research Data Co-utilization for Science and Technology. MSIT Research Paper. STEPI.
- Shin, Eun-Jung, An, Hyung-Jun, Yang, Hyeon-Chae, Choi, Byong-Sam, Yang, Seung-Woo, Jung, Won-Kyo, & Kim, Soo-Youn (2017). Policy Measures to Promote Open Science in South Korea. *Policy Studies*, 1-194.
- Shin, Eun-Jung, Sohn, Soo-jung, Suh, Ji-Young, & Kim, Yeong-Lin (2019). Considerations for Data Rights in Publicly Funded

Research, Policy Studies, 2019(16), 1-216. Science and Technology Policy Institute.

The Korean Association for Public Administration (2001). Public Administration Electronic Dictionary. Available:
<https://kapa21.or.kr/bbs/dictionary/6626>

Yim, Jin-Hee (2010). A Study on the Process of Public Information Seeking and Providing through Accountability Mechanism in Korea. Doctoral dissertation, Yonsei University.