

# DAF(Data Asset Framework)를 활용한 임산공학 분야 연구자들의 연구데이터 관리 개선 방안\*

- 국립산림과학원을 중심으로 -

## The Study on Research Data Management of Researchers in the Field of Forestry Engineering using DAF(Data Asset Framework) - Focused on National Institute of Forest Science -

김주섭(Juseop Kim)\*\* · 한연중(Yeonjung Han)\*\*\* · 유원재(Won-Jae You)\*\*\*\*  
전예린(Yerin Jeon)\*\*\*\*\* · 김선태(Suntae Kim)\*\*\*\*\*

### 〈목 차〉

I. 서론	IV. 임산공학 분야 연구자들의 연구 데이터 관리 현황
II. 이론적 배경	V. 결론
III. 임산공학 분야 연구데이터 특성	

**요 약:** 본 연구는 임산공학 연구자들의 연구데이터 관리 현황 파악을 목적으로 하였다. 연구 목적 달성을 위해 DAF(Data Asset Framework)라는 도구를 이용하여 설문조사를 수행하였다. DAF는 기관이 연구데이터를 어떻게 관리하고 있는지 식별, 위치, 설명 및 평가할 수 있는 수단을 제공하는 조사 도구이다. 이러한 DAF를 활용하여 국립산림과학원 내 임산공학 분야 연구자들을 대상으로 연구데이터 관리 현황을 분석한 결과, 연구데이터 생성 방법 및 유형, 공유, 저장, 보존 그리고 재사용 등과 같은 5가지의 카테고리의 현황 및 문제점이 도출되었으며, 또한 문제점 관련하여 해결책을 제시하였다. 이번 연구는 DAF와 같은 체계적인 도구를 활용한 기초 조사로서 특정 분야의 RDM 시스템 설계 시 연구데이터의 현황 및 문제점을 분석하는데 참고자료가 될 수 있을 것이다.

**주제어:** 연구데이터, 연구데이터 관리, DAF, 임산공학, 국립산림과학원

**ABSTRACT:** This study was started with the aim of grasping the current status of research data management of forestry engineering researchers. In order to achieve the research purpose, the survey was conducted using a tool called DAF (Data Asset Framework). DAF is an investigative tool that provides a means to identify, position, describe and evaluate how the agency manages research data. Using this DAF, the research data management status was analyzed for researchers in the field of forestry engineering at the National Institute of Forest Science. As a result of analysis, the current status and problems of the five categories such as the method and type of research data creation, sharing, storage, preservation, and reuse were identified, and solutions were presented in relation to the problems. This study is a basic investigation using a systematic tool such as DAF, and can be used as a reference for analyzing the current status and problems of research data when designing RDM system in a specific field.

**KEYWORDS:** Research Data, Research Data Management(RDM), Data Asset Framework(DAF), Forestry Engineering, National Institute of Forest Science

\* 본 연구는 2019년도 국립산림과학원의 연구지원에 의해 이루어진 것임

\*\* 전북대학교 문헌정보학과 강사 (jusup.kim@gmail.com / ISNI 0000 0004 7492 1806) (제1저자)

\*\*\* 국립산림과학원 목재이용연구부 임업연구사 (yeonjung@korea.kr / ISNI 0000 0004 5947 1958) (공동저자)

\*\*\*\* 국립산림과학원 목재이용연구부 임업연구사 (sngkgk@korea.kr / ISNI 0000 0004 8032 4581) (공동저자)

\*\*\*\*\* 전북대학교 문헌정보학과 석사과정 (nirye96@gmail.com / ISNI 0000 0004 7936 5183) (공동저자)

\*\*\*\*\* 전북대학교 문헌정보학과 조교수 (kim,suntae@jbnu.ac.kr / ISNI 0000 0004 6492 6355) (교신저자)

• 논문접수: 2020년 5월 21일 • 최초심사: 2020년 6월 03일 • 게재확정: 2020년 6월 24일

• 한국도서관정보학회지, 51(2), 103-131, 2020. [http://dx.doi.org/10.16981/kliss.51.2.202006.103]

## I. 서론

2017년 10월 정부는 「4차산업혁명위원회의 설치 및 운영에 관한 규정」에 의하여 ‘4차 산업혁명 위원회’를 설립하였다. 해당 위원회의 역할은 4차 산업혁명 관련 정책들을 심의·조정하는 것으로 그 중심에는 인공지능과 빅데이터가 핵심 추진 동력으로 손꼽히고 있다(대통령직속 4차산업혁명위원회 2020). 여기에 맞추어 2018년 과학기술정보통신부는 국가 R&D 추진 과정에서 축적되는 연구데이터를 관리 및 공유하고자 「연구데이터 공유·활용 전략」을 수립하였다. 연구데이터는 연구 과정에서 생성되는 모든 데이터를 말하며 연구 패러다임이 데이터-기반 연구로 변화함에 따라 더욱 중요하게 인식하고 있다. 여기에서 말하는 연구데이터에는 일반적으로 관측 데이터, 실험 데이터, 시뮬레이션 데이터, 파생되거나 컴파일된 데이터 그리고 참조 또는 표준 데이터를 포함한다(University of Leeds n.d.).

이러한 데이터-기반 연구의 배경으로 언급되고 있는 오픈 사이언스(Open Science)는 연구를 가속화하기 위한 수단으로서 공개적으로 자금을 지원하는 연구 결과의 주요 산출물을 디지털 형식으로 접근 가능하도록 만드는 연구자, 정부, 연구비 지원 기관 또는 과학 커뮤니티 자체의 노력을 의미한다(OECD 2015, 7). 이러한 오픈 사이언스의 개념이 전 세계적으로 공유되면서 각 국가별로 오픈 사이언스 정책의 범위도 넓어지고 있다. 즉, 2000년 전후로 오픈 사이언스 정책은 오픈 액세스를 가리키고 그 이후로는 공공 영역을 중심으로 공적 자금으로 수행된 연구 논문에 대한 공개토록 하는 정책으로 확대되었다(신은정 외 2017).

하드웨어의 발전, 고성능 네트워크, 그리고 센서 장비의 발달은 많은 학문 분야에서 다양한 연구데이터를 생산하고 있다(김선태, 이태영, 김용 2012). 분야별로 작은 센서부터 거대한 장비까지 실시간으로 연구데이터를 생산하고 있는 지금 연구데이터 관리(Research Data Management, 이하 RDM)는 효과적인 연구를 위한 필수적인 요소가 되었다. 또한, 4차 산업의 주요한 변화 동인인 데이터 분석 및 활용 증대, 연구 환경 패러다임의 변화 그리고 오픈 사이언스로 주목되는 연구결과물에 대한 자유로운 접근은 국내에서도 연구데이터 관리를 태동시킬 수 있는 배경이 되었다. 최근에는 연구데이터 정의와 데이터 관리 계획 도입 및 DMP 정보 등록 근거가 법령에 신설되었다(국가연구개발사업의 관리 등에 관한 규정 2019). 해외의 경우, 2013년을 시작으로 미국, 영국, EU 그리고 일본 등은 공공 연구를 통해 산출된 연구데이터의 선제 관리 및 접근 확대 지침을 제정한 반면에 우리의 현실은 상당히 미흡한 실정이다(신은정 2017). 특히, 해외의 경우, 대학 도서관이나 국가의 연구기관을 통해 다분야 연구데이터를 관리하거나 또는 특정 학문 분야의 연구데이터를 관리하기 위한 시스템을 마련하여 서비스하고 있다. 이외는 다르게, 우리는 연구데이터가 얼마나 생산되고 있으며 제대로 관리되고 있는지 최근에 연구가 진행되었을 뿐이다.

덧붙여 연구데이터의 관리가 미흡함에 따라 많은 시간과 노력 그리고 비용을 투자하여 생산한 데이터가 사장되고 있는 현실이다.

이에 따라 본 연구에서는 국내에서 생산되는 연구데이터의 체계적인 관리를 위해 실제 연구데이터가 어떻게 관리되고 있는지를 살펴보고자 한다. 모든 학문 분야를 대상으로 연구데이터 관리를 조사하는 건 현실적으로 불가능하므로 본 연구에서는 임산공학 분야를 선택하였다. 또한 해당 연구가 임산공학 분야를 대상으로 하였기 때문에 국립산림과학원의 연구데이터 관리 정책 등에 대한 논의를 포함하였다. 본 연구를 수행하기 위하여 국립산림과학원 내 임산공학 연구자들을 대상으로 설문조사를 실시하였다. 설문지는 ‘DAF(Data Asset Framework, 이하 DAF)<sup>1)</sup>’ 설문지를 국내 상황에 맞게 수정하여 온라인으로 연구자들에게 배포하였다.

## II. 이론적 배경

### 1. 연구데이터

이번 절에서는 연구데이터의 의미와 범위에 대해서 살펴보고자 한다. 다음의 <표 1>은 해외 대학도서관 및 연구기관 등에서 확인한 연구데이터에 대한 정의를 나타낸 것이다.

<표 1> 연구데이터의 정의

출처	연구데이터의 정의
University of Leeds	연구데이터는 원래 연구 결과를 검증하기 위해 수집, 관찰, 발생 또는 생성된 정보를 말함
University of Sheffield	연구데이터는 여러 가지 형태를 취할 수 있지만 본질적으로 연구 결론을 알리거나 뒷받침하는 데 사용되는 증거를 가리킴
Yale University Library	연구데이터는 독창적인 연구를 위해 분석 목적으로 수집, 관찰 또는 생성된 정보를 말함
University of Turku	연구데이터는 연구 과정에서 생성된 데이터를 말하며 연구 결과를 정당화하거나 재생산하는 데 사용될 수 있음
CGIAR	연구데이터는 과학 연구의 주요 출처로 사용되는 실제 레코드를 말함
OECD	과학 연구의 주요 출처 및 연구 결과를 검증하는데 사용되는 사실 기록을 말함

영국의 리즈대학, 쉐필드 대학, 미국의 예일 대학, 핀란드의 투르크대학, 세계 최대의 글로벌 농업 혁신 네트워크인 CGIAR 마지막으로 OECD 등 해외 대학 및 연구 기관 등에서 정의한 연구

1) DAF는 조직이 자신의 연구데이터를 어떻게 관리하고 있는지 식별, 위치, 설명 및 평가할 수 있는 수단을 제공하며, 영국의 고등 교육 기관에서 생산된 연구데이터가 보존되고 장기적으로 접근 가능한 상태를 유지하도록 돕는다.

데이터를 살펴본 결과 연구데이터는 연구 중 발생하는 모든 데이터를 말하며 여기에는 디지털 또는 비-디지털의 모든 형태를 포함한다고 말할 수 있다.

## 2. 선행 연구

이번 절에서는 연구데이터 관리와 관련한 국내외 논문 11편에 대하여 살펴보고자 한다.

〈표 2〉 선행 연구 목록

구분	연도	제목	비고
해외	2019	이라크 대학에서 연구데이터 관리의 도전과 사례	모든 학문 분야
해외	2019	생명과학 및 사회과학 연구 학자 간의 연구데이터 관리 및 데이터 공유	생명과학 및 사회과학
해외	2019	TU 델프트의 연구데이터 관리 및 공유에 문화적 장애	모든 학문 분야
해외	2019	슬로베니아 과학의 연구데이터 관리 및 연구데이터 활용 능력	슬로베니아 과학
국내	2018	과학기술분야 출연연구기관 연구데이터 관리 및 공유 사례 분석 연구	과학기술 분야
해외	2017	대학도서관의 연구데이터 관리 개발: 연구데이터 서비스 성숙도에 대한 이해	대학도서관
해외	2017	화학 연구데이터 관리 및 공유의 필요성에 대한 조사	화학 분야
국내	2017	국내 과학기술분야 연구기관의 과학데이터 관리 현황	과학기술 분야
해외	2017	대기 분야 과학자 및 공학 교수의 데이터 관리 사례 및 전망	대기 및 공학 분야
국내	2012	대학 내 연구자들의 연구데이터 관리에 관한 연구	모든 학문 분야
해외	2011	과학자들의 데이터 공유: 사례와 인식	과학 분야

Mahadi와 Rafea(2019)는 연구데이터 관리의 문제점을 확인하기 위하여 이라크의 5개 대학 155명의 대학원생과 20명의 교수진을 대상으로 설문을 실시하였다. 연구 결과 전반적으로 연구데이터 관리가 부족한 것으로 나타났다. 특히, 저자들은 연구데이터 관리의 문제점으로 연구데이터 관리 가이드라인 부족, 적절한 인적 자원 부족, 기술 노후화, 비효율적인 인프라, 재정 부족, 연구데이터 관리 정책 부재 그리고 해당 기관과 연구자의 협조 부족 등을 파악하였다. 이에 대한 해결책으로 대학원생과 연구자들은 연구데이터 리포지토리를 구축하고 타 대학 및 연구 기관과의 협력을 권장하였다.

Naushad와 Sidra(2019)는 연구데이터 관리 및 공유에 대한 연구자의 인식을 조사하기 위하여 Aligarh Muslim 대학 352명의 생명과학 및 사회과학 연구자를 대상으로 설문을 진행하였다. 조사 결과 사회과학 연구자들이 생명과학 연구자들보다 연구데이터를 더 기꺼이 공유하고 있음을 파악하였다. 또한 과학적 발전에 기여하고 연구 인용 및 가시성을 높이는 것이 연구자들이 데이터를 공유하도록 동기를 부여한 주요 요인으로 확인하였다.

Plomp 등(2019)은 델프트 공과 대학(TU 델프트) 내 많은 연구자들의 연구데이터 관리 인식 제고와 실무 개선을 위하여 모든 교수진을 대상으로 데이터 관리자를 임명함으로써 문화적 장벽을 해결하였다. 이러한 Data Stewardship 프로젝트는 전문가의 조언과 인식 제고를 통해 현재의 데이터 및 소프트웨어 관리 및 공유의 점진적 개선에 중점을 둔다. 여기에 저자들은 학문별 커뮤니티 그리고 문화의 변화를 위한 적절한 보상과 인센티브의 필요, 나아가 이해관계자의 공동 노력 등이 필요하다고 지적하였다. 특히, 커뮤니티 구축이 TU 델프트 대학 내에서 규정 및 데이터 관리 문화를 변화시키는데 필수적임을 기술하였다.

Vilar와 Zabukovec(2019)는 슬로베니아의 과학 분야에 대한 연구데이터 리터러시(RDL)와 연구데이터 관리(RDM)에 대하여 조사하였다. 조사 결과, 사회과학과 자연과학 그룹은 현재의 연구데이터 관리가 명확하지 않고, 모든 대학이 데이터 관리 계획이 필요치 않다고 판단하였다. 인문학, 공학, 의학 및 보건 과학은 연구데이터 관리에 만족 중이며 각 대학에 관한 제도적 데이터 관리 계획 수립에 관하여 긍정적인 반응을 보였다. 또한, 의학, 자연과학 그리고 인문학의 경우, 메타데이터에 태그를 달지 않는 경우가 많았고, 공학 및 사회과학 등은 공공 도메인 데이터를 사용하기도 하였다.

박미영, 안인자, 남승주(2018)는 과학기술분야의 연구데이터 관리 및 공유 사례를 분석하기 위하여 27개의 정부출연기관을 대상으로 인터뷰를 진행하였고 그 중에서 연구데이터 관리가 우수한 9개 기관의 사례를 구체적으로 기술하였다. 제시된 우수 사례는 연구데이터 수집 단계, 수집·관리 고도화 단계 그리고 공유·활용 시작 단계 등 3단계로 나누었다. 연구결과, 저자들은 다음과 같이 5가지의 시사점을 제시하였다.

- 연구데이터 수집·관리를 위한 데이터 관리 규정 및 DMP 마련 등
- 연구데이터 관리를 위한 인력 양성
- 연구자의 적극적인 참여를 위한 보상체계 마련
- 연구자 인식 제고
- 국제표준에 부합하는 데이터 품질관리 및 표준화

Cox 등(2017)은 고등 교육 도서관의 연구데이터 관리 활동, 서비스 및 기능을 조사하기 위하여 호주, 캐나다, 독일, 아일랜드, 네덜란드, 뉴질랜드 및 영국의 고등 교육 도서관에 대한 설문 조사를 실시하였다. 연구 결과, 도서관은 연구데이터 관리의 지지 및 정책 개발에서 리더십을 발휘하였다. 구체적으로 서비스 개발의 경우 제한적이지만 기술 서비스(데이터 카탈로그 및 현용 데이터 큐레이션)보다는 자문 및 컨설팅 서비스(DMP 지원 및 데이터 관련 교육)에 중점을 두었다고 기술하였다.

Xiujuan과 Ming(2017)은 119개 항목으로 구성된 설문지를 바탕으로 중국과학원의 화학과 연

구원과 대학생들에게 화학의 데이터 관리 지원 필요성 및 실태를 밝혔다. 데이터 생성 및 수집, 데이터 기록 및 처리, 데이터 보존과 백업, 데이터 공개 및 공유, 데이터 관리 및 공유 서비스의 필요성 등의 5가지 관점에서 수행되었으며, 이를 통하여 화학 분야의 연구데이터 관리 및 공유 촉진, 연구데이터 서비스 설계 등에 활용할 수 있도록 하였다.

최명석, 이승복, 이상환(2017)은 과학기술 분야 연구기관의 과학데이터 관리 현황을 조사하기 위하여 22개의 정부출연연구기관과 20개 대학의 연구자를 대상으로 조사를 진행하였다. 이를 통해 다음과 같은 과학데이터 관리·활용 체계 개선 방향을 제시하였다.

- 데이터 관리 계획 의무화 정책 추진
- 과학데이터 성과 인정 및 인센티브 방안 마련
- 개인정보보호 등 데이터 활용 규제 완화 방안 마련
- 과학데이터 품질 관리 체계 수립
- 국가 차원의 과학데이터 관리·공유 인프라 구축
- 과학데이터 공유·활용 인식 제고
- 역할과 책임: 다양한 이해관계자간의 협력 등

Wiley와 Mischo(2016)는 연구 활동과정에서 데이터 관리 사례 및 요구를 파악하기 위해 일리노이 대학교 어바나-삼페인 캠퍼스의 공학 및 대기과학 분야 교직원 21명을 대상으로 심층 인터뷰를 진행하였다. 인터뷰는 최근 연구 프로젝트, 편딩 소스, 데이터 타입 및 포맷, 기관 및 분야 리포지토리의 이용, 데이터 공유, 대학 도서관 데이터 관리 및 보존 서비스에 대한 인식, 편딩 기관 검토 패널 경험, 연구데이터 관리 관련 애로사항 또는 문제점 등을 주로 다루었다. 인터뷰 결과, 연구자들은 연방 기관 및 저널 출판사의 데이터 공유 지시에 대해 인식하고 있으며 이를 수행하기 위한 매커니즘 또는 스킬이 부족하다는 반응, 데이터 공유로 인한 위험을 우려하고 있었다.

김지현(2012)은 대학 내 연구데이터 관리 현황을 조사하기 위하여 한국연구재단의 연구과제 책임자를 대상으로 설문조사를 실시하였다. 연구 결과, 연구데이터 보관은 개인 PC와 이동식 매체를 많이 선호하였으며 연구데이터 유지 연한은 10년 미만으로 나타났다. 데이터 공유에 있어서는 외부에서 요청하는 경우를 제외하고 소속 연구팀 내에서 이루어지며 데이터 수집은 대부분이 출판된 논문에서 추출하거나 해당 연구자와 개인적으로 연락을 취한다고 하였다. 이것을 통해 저자는 대학 내 연구데이터의 체계적인 관리가 필요하다고 언급하였다.

Tenopir 등(2011)은 데이터 공유의 장애물에 대한 인식과 최근 데이터 공유 사례를 연구하기 위해 1329명의 과학자를 대상으로 설문조사를 진행하였다. 조사결과, 과학자들은 시간 및 자금의 부족 등 다양한 이유로 자신의 데이터를 타인이 전자적으로 이용할 수 없게 하였다. 대부분의

응답자는 데이터 또는 연구 라이프사이클의 초기 및 단기적 부분(연구데이터 수집, 탐색, 기술 또는 카탈로그, 분석, 데이터의 단기적 저장)에 대한 현재 프로세스에 만족하고 있으나 데이터 장기 보존 프로세스에 불만족하고 있다고 답변하였다. 많은 기관들이 연구원들에게 단기 및 장기 데이터 관리 관련 지원을 제공하지 않고 있으며 응답자들은 특정 조건(인용 등)을 만족할 경우 데이터를 공유할 용의가 있다고 응답하였다. 또한 주요 펀딩 기관, 주제 분야, 연령, 업무 중점, 지역에 따라 데이터 관리 사례에 있어 상당한 차이와 접근방식이 존재하였다.

이상 11편의 국내외 선행연구를 분석한 결과 연구자들은 연구데이터 관리의 문제점에 대하여 데이터 관리 규정(매커니즘, 보존 프로세스 등) 및 가이드라인, 인적 자원 및 인프라, 재정, 공유로 인한 위험 그리고 연구자의 인식 등의 부족을 확인하였다. 이에 따른 해결책으로 연구 인용 및 가시성을 높이는 동기 부여, 리포지토리 구축 및 관련 기관의 협력, 학문별 커뮤니티 구축, 변화를 위한 적절한 보상과 인센티브의 필요, 관리 인력 양성, 나아가 이해관계자의 공동 노력 등을 제시하였다. 이와 같이 국내외 선행연구는 대학, 도서관 그리고 학문분야를 중심의 연구자를 대상으로 연구데이터 관리 현황을 분석하였다. 본 연구에서는 이제까지 선행연구에서 다루지 않은 임산공학 분야의 연구자들을 대상으로 연구데이터 관리 현황을 파악하여 문제점과 해결방안을 제시하고자 한다.

### 3. DAF(Data Asset Framework)

DAF는 이용자 피드백을 고려하여 'Data Audit Framework'에서 변경된 명칭으로 조직이 연구데이터 자산을 어떻게 관리하고 있는지 식별, 위치, 설명 및 평가할 수 있는 수단을 제공하며, 이를 위해 연구데이터 지원 서비스의 범위를 넓히는 데 필요한 정보, 견해 및 경험을 수집하는 일련의 방법을 수집한다(DCC n.d.).

다음은 DAF의 역할을 정리한 것이다(JISC, University of Glasgow and DCC 2009).

- 데이터의 생성 및 보유 현황 파악
- 데이터의 저장, 관리, 공유 및 재사용 방법 탐색
- 위험 식별(예: 데이터 오용, 손실 또는 검색 불가능)
- 데이터 생성 및 공유에 대한 연구자의 태도를 알 수 있음
- 지속적인 데이터 관리를 개선할 수 있는 방법 제안

DAF 모델은 데이터 조사를 수행하기 위해 4단계 접근 방식을 제안한다. 다음은 각 단계별 주요 내용을 정리한 것이다.

- 1단계 - 조사의 목적과 범위를 계획하고 정의하며 예비조사를 실시

- 2단계 - 어떤 데이터 자산이 존재하는지 파악하고 이를 분류하여 보다 심층적인 분석을 위한 노력을 어디에 집중해야 할지를 결정
- 3단계 - 연구자의 워크플로우를 이해하고 데이터 생성 및 큐레이션 사례의 취약점을 파악하기 위한 정보 라이프 사이클을 고려
- 4단계 - 수집된 정보를 종합하여 데이터 관리 개선을 위한 권장사항을 제공

첫 번째 단계는 평가를 계획하는 단계로서 설문조사의 목적과 범위를 기획하기 위한 예비조사를 실시하는 단계이다. 두 번째는 기관 내부의 데이터를 확인하고 분류하는 단계로서 데이터의 현황 파악과 심층 분석을 위한 선택을 집중하는 단계이다. 세 번째 단계는 데이터 자산을 평가 관리하는 단계로서 연구자의 작업흐름과 큐레이션 사례를 통한 정보 라이프 사이클을 고려하는 단계이다. 마지막 단계는 설문조사를 통해 수집된 정보를 종합하여 데이터 관리 현황 개선을 위한 권장사항을 제시하는 단계이다.

이러한 DAF의 설문 프로세스에는 정보전문가, 연구자 그리고 관리자와 같은 다양한 이해 관계자가 참여해야 한다.

첫 번째, 설문조사를 수행하는 정보전문가는 인터뷰를 통해 연구자들이 필요한 지원을 알 수 있으며 리포지토리와 아카이브의 콘텐츠 식별 그리고 기관 내에서 데이터 관리의 모범 사례에 대한 인식을 제고하는 이점을 가진다.

다음으로 설문조사에 참여하는 연구자는 기존 데이터 지원 확인, 효과적인 데이터 큐레이션이 가능한 서비스 구체화, 데이터 관리 모범 사례에 대한 인식 제고 그리고 장기적인 데이터 선택 및 관리 방법에 대하여 배울 수 있는 이점이 있다.

마지막으로 데이터 큐레이션 개선을 위한 지원을 담당하는 고위 관리자 및 전략가는 중복 노력 식별 및 기존 자원의 최적 사용, 데이터 손실 등의 위험 인식, 정보 전략 개발 지원, 접근 및 재사용을 위한 데이터 자산 식별 마지막으로 협업 및 학제 간 연구를 촉진할 수 있는 이점을 가진다.

즉, DAF는 정보전문가, 연구자, 관리자 측면에서 얻을 수 있는 이점은 다양하며 해당 프레임워크를 통하여 궁극적으로 기관의 연구데이터의 관리 현황 파악 및 개선방안을 제시할 수 있도록 도와준다. 본 연구에서는 DAF에서 제시하는 설문지를 토대로 국립산림과학원의 연구자를 대상으로 설문을 실시하였으며 설문 내용에서 연구자의 요구사항 및 문제점을 도출하였다.

### Ⅲ. 임산공학 분야 연구데이터 특성

#### 1. 국립산림과학원의 연구데이터 관리 정책



임산공학 분야의 연구데이터 특성을 살펴보기 전에 연구 대상 기관인 국립산림과학원에 대한 소개와 해당 기관의 연구데이터 관리와 관련한 정책 및 인프라에 대하여 확인해보고자 한다.

국립산림과학원은 산림과학기술의 개발과 보급이라는 고유 미션을 갖고, 산림자원의 보전 및 육성, 임산업 활성화, 산림재해 예방 및 관리 등 산림분야 R&D를 수행하는 국가연구기관이다. 그 중 목재이용연구부는 「최적 가공기술을 활용한 목재 이용 증진」을 목표로 “첨단 가공기술을 활용한 목재의 부가가치 증진연구”를 수행하는 목재공학연구과, “도시목조화를 통한 친환경 목재 수요 확대 연구”를 수행하는 목재이용연구과, “목질계 바이오연료, 펄프·제재 및 신소재화 기술 개발”을 수행하는 목재화학연구과로 구분된다. 다음의 <그림 1>은 국립산림과학원의 조직도 일부를 나타낸 것이다.



<그림 1> 국립산림과학원의 조직도(국립산림과학원 2020)



<그림 2> 목재이용연구부의 연구추진 체계도(국립산림과학원 2018)

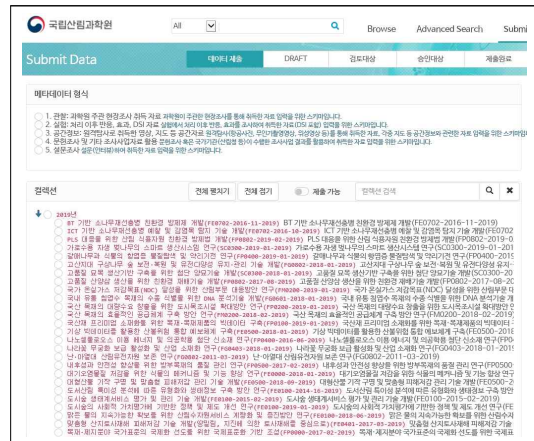
최근에는 기관 전략과제의 하나로서 목재 생산 및 이용 원천기술을 활용한 산업 활성화를 목표로 친환경 목재 가공 및 건축재 이용기술을 개발하고 있다. 즉, 목재 재질정보 빅데이터 구축과 맞춤형 가공기술고도화를 통하여 목재자원의 활용도를 증진시키고 목재의 건축부재 이용기술 개발 및 목조건축산업 활성화를 위한 신수요를 창출하며, 목질재료의 친환경 고부가가치 소재화 기술을 개발하고 있다.

국립산림과학원은 “연구데이터 관리 기본계획”의 수립을 위하여 2018년 3월부터 TFT를 구성하여 「국립산림과학원 연구사업 관리 규정(예규 제307호)」의 일부 개정(2019.2.11.)을 통하여 연구데이터 관리 의무화 조항을 신설하였다. 또한 연구데이터 아카이빙 플랫폼을 도입하여 2019년 4월에 내부망(인트라넷)에 연구데이터 리포지토리를 설치하였다. 2019년 7월부터 국립산림과학원에서 수행되는 일반연구과제별 연구데이터관리계획(Data Management Plan, 이하 DMP)를 수립하여 제출하도록 하였으며, 이러한 상황은 2020년에도 지속되고 있다. 즉, 국립산림과학원은 연구데이터 관리를 위한 기술적(리포지토리) 및 정책적(연구데이터 관리 기본계획) 인프라가 마련되어 있음을 확인할 수 있다. DMP 제출 등 연구데이터 관리에 관한 구체적 시행이 최근 시작되어 해당 효과 및 연구자들의 만족도를 파악할 수는 없지만 연구데이터 관리에 관한 기관의 의지를 확인할 수 있다는 점에서 그 의미를 파악할 수 있다.

다음의 <그림 3>과 <그림 4>는 국립산림과학원의 리포지토리의 메인 화면과 해당 리포지토리의 컬렉션 구성 화면을 나타낸 것이다.



<그림 3> 국립산림과학원의 리포지토리



<그림 4> 리포지토리의 컬렉션 구성 화면

## 2. 임산공학의 정의 및 분류

이번 연구에서는 임산공학 분야의 연구자를 대상으로 연구데이터 관리 현황을 살펴보는 것이

주된 목표이므로 임산공학 분야의 정의 등을 확인하기 위하여 국가과학기술표준분류체계(과학기술정보통신부, 한국과학기술기획평가원 2018)를 참고하였다. 다음은 해당 분류체계에 기술되어 있는 임산공학(Forest Products)에 대한 정의와 세부 범위를 나타낸 것이다.

- 임산공학: 환경 친화적인 Biomass자원인 목재를 중심으로 삼림에서 생산되는 각종 임산자원의 합리적·효율적 이용과 용도개발, 재료 및 제품개발과 그 연구발전에 요구되는 기초적인 기술
- 목재 조직/분류: 목재의 이화학적 성질을 고려하여 목재를 조직 또는 분류하는 기술
- 목재 물리/목구조/목재역학: 목재의 응력과 변형률, 탄성, 점탄성, 강도, 경도 및 기타 역학적 성질, 목재의 생장 응력과 기계적 성질의 경년 변화 등을 파악하는 기술
- 목재 절삭/목공: 목재 절삭 기구와 구성 요소 및 성질과 방법에 대하여 파악하는 기술
- 목재 건조/보존: 목재 건조 및 보존 시 목재 내 응력 변화 조절 및 건조 환경 조절 원리 등을 파악하는 기술
- 목재 화학: 임산물의 화학적 성질과 펄프, 제지 관련 기술
- 목질 복합재료: 목재의 재료인 집성재, 합판, 파티클보트, MDF, 섬유판, LVL 등에 대한 제조 이론 및 실제 각각의 물리·기계적 특성을 검토하여 2차 가공 및 개질 능력 함양 등을 이루는 기술
- 펄프/종이: 목재로부터 얻어지는 각종 펄프 및 종이의 제조 원리와 제반 펄프 제조 기구의 구조와 기능을 다루고 펄프의 물리, 화학적 특성 등에 대한 기술
- 임산 미생물/버섯: 산림 생태계에서의 물질의 순환, 수목의 병 발생, 목재의 부후 및 변색, 제지 산업 등에 이용되는 미생물 기술
- 임산 에너지: 바이오매스(biomass)를 연료로 하여 얻어지는 에너지를 이용하는 기술
- 달리 분류되지 않는 임산공학

해당 분류체계에서 나타난 임산공학은 바이오매스인 목재의 이용, 개발 및 연구에 요구되는 기초 기술을 말하는 것으로 세부 범위에는 목재 조직/분류, 목재 물리/목구조/목재역학, 목재 절삭/목공, 목재 건조/보존, 목재 화학, 목질 복합재료, 펄프/종이, 임산 미생물/버섯 그리고 임산 에너지 등 9개의 분류로 전개되어 있다.

### 3. 임산공학 연구데이터의 특성

본 절에서는 임산공학 분야의 연구데이터 특성을 이해하기 위하여 해당 분야의 학문적인 특성과 연구 대상기관인 국립산림과학원의 데이터 유형 분류표를 통해 해당 분야의 연구데이터 특성을

을 파악하고자 한다. 다음은 산림과학 분야의 학문적 특성을 정리한 것이다(이경준 외 2014).

- 장기성: 산림과학 연구는 산림, 즉 나무의 수명부터 5~100년 이상을 넘나들기 때문에, 필수적으로 장기적인 연구 기간이 요구됨. 이러한 특성으로 국가가 연구를 주도하고 있음
- 다양성: 산림과학은 자연과학, 사회과학, 인문학, 공학 등 넓은 분야를 망라한 학문으로 산림생물 분야(자연과학), 산림 경영 분야(사회과학), 산림공학 분야 그리고 임산가공 및 이용 분야로 구성되어 있음
- 경험성: 수목학의 경우, 분류학적 형질에 대하여 1년 중 관찰할 수 있는 기간이 매우 짧기 때문에 꽃이나 열매가 없는 시기에 식별해야 하는 임학 및 산림과학에서는 실용성이 떨어짐. 따라서 이와 같은 분류학적 형질보다 경험에 의한 자신만이 알 수 있는 변이에 대한 선택, 즉 경험성도 산림과학에서는 중요하게 여겨짐

즉, 산림과학은 장기성, 종합성 및 경험성이라는 특성에 따라 장기간에 걸쳐 인문학, 사회과학, 자연과학 그리고 공학 등 다양한 전공 분야를 망라한 연구가 이루어짐을 알 수 있다.

다음의 <표 3>은 국립산림과학원에서 생산되는 데이터 유형을 정리한 것이다.

<표 3> 국립산림과학원의 데이터 유형 분류표(국립산림과학원 2019)

데이터 유형	공유, 재활용, 보존에 적합한 파일 포맷	호환 가능 포맷
<b>광범위한 메타데이터를 지닌 정략적 테이블 데이터</b> - 다양한 유형의 데이터 레이블, 코드 레이블, 결측치 정의 등이 포함된 데이터 매트릭스	• 정형화된 텍스트 데이터 • SAS, SPSS 등 통계패키지 입력용 데이터 • 메타데이터 정보를 포함하는 구조화된 텍스트 또는 마크업 파일 (e.g. XML: Extensible Markup Language)	• 마이크로소프트 액세스 (e.g. mdb, accdb) 등
<b>최소한의 메타데이터를 지닌 정략적 테이블 데이터</b> - 다양한 유형의 데이터 레이블, 코드	• 콤마, 탭, 공백 등으로 구분된 텍스트 데이터 (CSV, PRN, txt 등) • SQL (Simple Query Language) 데이터 정의 구문이 포함되고 특정 문자 세트로 구분된 텍스트	• 데이터에 사용되지 않은 문자로 구분된 텍스트 데이터 • 범용 포맷 (e.g. 마이크로소프트 엑셀(xls, xlsx), 액세스 (mdb, accdb), dBase(dbf), 오픈 다크먼트 스프레드시트 (ods) 등)
<b>공간지리정보</b> - 벡터 및 래스터 데이터	• ESRI 셰이프 파일 (필수: shp, shx, dbf 옵션: prj, sbx, sbn) • 좌표계에 투영된 TIFF (tif, tfw) 파일 • CAD 데이터 (dwg, dxf) • 테이블 형태의 GIS 속성 (attribute) 데이터	• ESRI Arcinfo 커버리지 파일 (폴더 형태) • ESRI Arcinfo 그리드 파일 (폴더 형태) • Imagine 이미지 파일 (img)

데이터 유형	공유, 재사용, 보존에 적합한 파일 포맷	호환 가능 포맷
정성적 데이터 - 원문 데이터	<ul style="list-style-type: none"> <li>문서의 유형이 적절히 정의된 XML</li> <li>서식있는 텍스트 포맷 (rtf: Rich Text Format)</li> <li>일반적 텍스트 데이터 (UTF-8, EUC-KR 등)</li> </ul>	<ul style="list-style-type: none"> <li>아스키 (ASCII) txt</li> <li>HTML (Hyper Text Mark-up Language)</li> <li>범용 문서 포맷 (hwp, doc, docx 등)</li> </ul>
디지털 이미지 데이터	<ul style="list-style-type: none"> <li>비 압축 TIFF (tif)</li> </ul>	<ul style="list-style-type: none"> <li>JPEG (jpeg, jpg), TIFF (tiff 등)</li> <li>JPEG2000 (jp2)</li> <li>Adobe pdf</li> </ul>
디지털 오디오 데이터	<ul style="list-style-type: none"> <li>무손실 코덱 (flac)</li> <li>웨이브 오디오 (wav), MPEG-1 오디오 레이어 (mp3)</li> </ul>	<ul style="list-style-type: none"> <li>오디오 교환 파일 포맷 (aif)</li> </ul>
디지털 비디오 데이터	<ul style="list-style-type: none"> <li>MPEG-4 mp4</li> <li>motion JPEG2000 (jp2)</li> </ul>	<ul style="list-style-type: none"> <li>JPEG2000 (jp2), flv, mkv, mov, wmv</li> </ul>
화학 데이터 - 피크 포지션 및 강도 (intensity), 등치선 표현을 필요로 하는 분광 데이터 및 유사 데이터	<ul style="list-style-type: none"> <li>JCAMP 포맷 파일 (뷰어: JspecView, Chemdoodle)</li> </ul>	<ul style="list-style-type: none"> <li>NMR, IR, Raman, UV, Mass Spectrometry 파일 등을 공유가 쉬운 JCAMP 포맷으로 변환</li> </ul>

데이터 유형은 크게 8가지 유형으로 메타데이터를 포함한 테이블 데이터, 공간지리정보, 정성적 데이터, 디지털 이미지 데이터, 디지털 오디오 데이터, 디지털 비디오 데이터 그리고 화학 데이터 등으로 구분된다. 또한 각 데이터 유형에 따라 공유, 재사용 및 보존에 적합한 파일 포맷과 호환 가능 포맷이 제시되었다.

위에서 제시된 국립산림과학원에서 생산되는 데이터 유형에 따라서 임산공학 분야의 연구데이터 특성을 정리하면 다음과 같다.

- 전문장비를 이용하여 측정 및 실험이 수행되므로 장비 의존적이라고 할 수 있음
- 이미지, 오디오, 비디오 등 비정형 데이터가 많음
- 실험 등을 통해 측정값이 수치로 도출되는 수치 데이터가 많음
- 수치 데이터를 모아 그래프 형태로 가공한 데이터가 많음

즉, 임산공학 분야의 연구데이터는 장비 의존적이며 이미지, 수치 데이터 등 정형 및 비정형 데이터가 많음을 알 수 있다. 최근에는 AI에게 현미경 사진을 교육시키는 딥러닝 기술을 이용한 수종식별 방법에 대한 연구가 진행되고 있어 이미지 데이터에 대한 저장과 재사용이 매우 중요한 것으로 판단하고 있다.

## IV. 임산공학 분야 연구자들의 연구데이터 관리 현황

본 연구의 목적은 임산공학 분야 연구자들의 연구데이터 관리 현황을 파악하기 위함이다. 따라서, 해당 분야의 연구자들을 대상으로 2장에서 기술한 DAF를 이용하여 설문조사를 실시하였다. 다음은 설문조사의 개요를 나타낸 것이다.

- 조사 목적: 임산공학 분야 연구자들의 연구데이터 관리 현황 조사
- 조사 대상: 국립산림과학원 목재이용연구부 내 모든 연구자(32명)
- 조사 방법: 웹 설문조사
- 조사 기간: 2019년 9월 19일 ~ 2019년 10월 07일
- 응답률: 100%
- 설문지 응답 URL: <http://ksurv.kr/?a=26420>

설문지는 2009년 10월에 작성된 실행 가이드의 내용을 최대한 준용하였으며 국내 실정과 맞지 않는 부분은 수정하여 배포하였다. 응답자는 총 32명이며, 국립산림과학원 목재이용연구부 내 모든 연구자를 대상으로 하였다. 조사 대상은 실제 연구데이터를 생산하는 연구자를 대상으로 하였으며 설문지 발송 전 해당 부서 연구자들에게 설문조사의 의의를 전달하여 모든 대상자가 해당 조사에 참여할 수 있도록 하였다. 설문지의 모든 문항은 항목별로 단/복수 응답이 가능하도록 설계하였다. 설문조사 문항은 크게 연구데이터의 유형 및 생산방법, 공유, 저장, 보존 및 재사용에 관련한 항목으로 구성하였다.

### 1. 연구데이터 생성 방법 및 유형

임산공학 분야 연구자들의 연구데이터 생성 방법 및 유형 등에 대해 질문한 결과 다음의 <표 4>와 같은 결과를 얻을 수 있었다.

먼저, 연구데이터 생성 방법 중 가장 많이 차지한 것은 실험으로서 52%로 절반 이상이 실험을 통해 데이터를 얻는 것으로 나타났다. 다음으로는 관찰, 시뮬레이션, 파생 그리고 참조 순으로 확인되었다. 다음으로 생성되는 연구데이터 유형을 질문한 결과, 엑셀 시트가 17%로 가장 많은 유형을 차지하였고 다음으로 사진 모음 그리고 실험실 노트 그 다음으로 이미지, 스캔 또는 X선 순서로 나타났다. 마지막으로 데이터 수집 및 생산에 연구비를 지원한 기관이 어디인지를 파악한 결과, 소속기관이 가장 많고 그 다음으로 소속기관 이외의 정부기관 그 다음 민간기업 순으로 나타났다. 기타 의견으로 웹사이트를 통해 데이터를 수집한 경우도 확인되었다.

〈표 4〉 연구데이터 생성 방법 및 유형

연구데이터 생성 방법(복수)			연구데이터 유형(복수)			데이터 수집 및 생산 지원 기관(복수)		
	빈도	%		빈도	%		빈도	%
관찰	14	24%	컴퓨터 프로그램에서 자동으로 생성되는 데이터	13	8%	정부 (소속기관 외)	15	36%
실험	30	52%	센서 또는 계측기에서 수집된 데이터(질문표 포함)	16	9%	소속기관	22	52%
참조	1	2%	이미지, 스캔 또는 X선	18	11%	민간기업	3	7%
파생	4	7%	웹사이트	2	1%	기타	2	5%
시뮬레이션	9	16%	MS Word 파일	13	8%			
기타	0	0%	엑셀 시트	29	17%			
			SPSS 파일	4	2%			
			디지털 오디오 파일	3	2%			
			디지털 비디오 파일	3	2%			
			현장 작업 데이터	1	1%			
			실험실 노트	19	11%			
			사진 모음	22	13%			
			비디오 테이프	1	1%			
			오디오 테이프	0	0%			
			슬라이드	6	4%			
			텍스트 말뭉치	1	1%			
			문서 또는 보고서	18	10%			

임산공학 분야 연구자들의 연구데이터는 대다수가 관찰과 실험을 통해 생성되며, 생성되는 데이터 유형은 엑셀 시트와 이미지 및 실험실 노트 등으로 파악됨을 알 수 있다. 특히, 해당 분야의 특성상 정형 및 비정형 데이터가 많은 관계로 해당 연구데이터를 효율적으로 관리할 수 있는 RDM 시스템이 권고될 수 있다.

## 2. 연구데이터 공유

임산공학 분야 연구자들의 공동 작업 시 연구데이터의 공유 방법 및 발생하는 문제에 대하여 질문한 결과 다음의 〈표 5〉와 같이 나타났다.

임산공학 분야 연구자들의 공동 작업 현황을 파악한 결과 78%가 동료 연구자들과의 협업이 이루어진다고 응답하였다. 또한, 데이터 공유 방법에서는 43%가 이메일, 24%는 소속기관 파일 스토리지 서비스, 22%는 CD, DVD, USB와 같은 저장장치를 이용하는 것으로 나타났다. 기타 응답으로는 인쇄 형태로 작성된 데이터를 공유하거나 SNS를 통해 파일을 공유한다는 의견이 있었다. 마지막으로 데이터 공유 시 문제점을 질문한 결과, 없다는 의견이 26%로 가장 많았지만 파일 이름 생성 규칙 부재가 17%, 소속기관의 파일 스토리지 서비스 기능 한계가 15%, 파일 버전 관리의 어려움이 13% 등 스토리

〈표 5〉 연구데이터 공유

공동 작업 여부			데이터 공유 방법(복수)			데이터 공유 시 문제점(복수)		
	빈도	%		빈도	%		빈도	%
네	25	78%	이메일	23	43%	없음	12	26%
아니오	7	22%	외부 파일 스토리지 서비스	3	6%	소속기관의 파일 스토리지 서비스 기능 한계	7	15%
			소속기관 파일 스토리지 서비스	13	24%	외부 파일 스토리지 서비스 기능 한계	4	9%
			CD, DVD, USB 등	12	22%	파일 이름 생성 규칙 부재	8	17%
			기타	3	6%	파일 버전 관리의 어려움	6	13%
						데이터 전송 관련 법적 문제	4	9%
						데이터 소유권 문제	4	9%
						기타	2	4%

지 서비스의 기능, 파일명 네이밍 그리고 버전 관리에 어려움을 겪는 것으로 나타났다. 기타 의견으로 다음이나 네이버와 같은 상용 이메일 사용이 불가한 점을 문제점으로 제시한 의견도 있었다.

위의 내용을 살펴본 결과, 임산공학 연구자들은 동료와의 협업이 진행되고 있으며 데이터 공유는 이메일이나 파일 스토리지 서비스 그리고 이동식 저장장치를 많이 사용함을 알 수 있었다. 또한 데이터 공유 시 문제점으로 파일명 네이밍, 데이터 버전관리 그리고 스토리지 서비스 기능 한계 등과 같은 기술적인 문제를 제시한 점으로 보아 이 부분은 향후 연구자를 대상으로 한 RDM 교육 및 가이드라인 제시 등으로 해결책을 제시할 수 있다.

### 3. 연구데이터 저장

다음의 〈표 6〉은 임산공학 분야 연구자들의 연구데이터 저장과 관련된 설문 결과를 정리한 것이다.

먼저, 임산공학 연구자들의 연구데이터 저장 위치를 확인한 결과, 50%가 사무실의 개인 PC에 저장한다고 응답하였고, 23%는 USB와 같은 이동식 저장장치, 12%는 출력물 형태로 그리고 10%는 소속기관 파일 스토리지 서비스에 저장한다고 답하였다. 보유하고 있는 연구데이터의 크기에 대하여 질의하였을 때 가장 많은 비율을 가진 크기는 1-50GB이며 다음으로 100-500GB 그리고 연구자 본인이 가지고 있는 연구데이터 크기를 모르는 경우도 16%로 나타났다. 위의 표에서는 표현하지 않았지만 파일 크기로 인해 연구데이터 저장에 문제가 생긴 적이 있느냐는 질문에는 ‘아니오’가 63%로 ‘네’라는 응답한 연구자가 38%로 나타났다. 이러한 저장 공간 문제의 해결방법에 대해서는 75%의 대부분의 연구자가 추가 저장 장치를 구입하는 것으로 나타났다.

위의 연구데이터 저장에 대한 설문결과를 분석한 결과 대부분의 연구자가 사무실의 개인 PC에 연구데이터를 저장하고 보유하고 있는 연구데이터 크기는 1-50GB가 가장 많은 것으로 나타났다.



〈표 6〉 연구데이터 저장

연구데이터 저장 위치(복수)			연구데이터의 크기(복수)			저장 공간 문제 해결방법(복수)		
	빈도	%		빈도	%		빈도	%
출력물	7	12%	1 GB 미만	3	9%	외부 파일 스토리지 서비스	1	6%
개인 PC(집)	1	2%	1 - 50GB	11	34%	소속기관 파일 스토리지 서비스	2	13%
외부 파일 스토리지 서비스	1	2%	50 - 100GB	3	9%	추가 저장 장치 구입	12	75%
소속기관 파일 스토리지 서비스	6	10%	100-500GB	6	19%	소속기관에게 저장 공간 요청	1	6%
개인 PC(사무실)	30	50%	500GB - 1TB	4	13%	기타	0	0%
이동식 저장장치	14	23%	1 - 50TB	0	0%			
CD/DVD/외장 하드 드라이브	1	2%	50 - 100TB	0	0%			
기타	0	0%	100TB 초과	0	0%			
			알 수 없음	5	16%			

하지만, 본인이 보유하고 있는 연구데이터 크기를 알 수 없다는 의견도 16%나 되어 기관차원에서 연구자가 보유하고 있는 연구데이터 현황 파악이 시급함을 알 수 있다. 또한, 연구데이터 저장 문제 해결방법으로 추가 저장 장치를 구입한다는 의견이 대다수 인 것으로 보아 기관 차원의 RDM 서비스 운용 시 개인별 스토리지 확대를 해결방안으로 제시할 수 있다.

#### 4. 연구데이터 보존

다음의 〈표 7〉은 임산공학 분야 연구자들의 연구데이터 보존과 관련된 설문 결과를 정리한 것이다.

〈표 7〉 연구데이터 보존 - 1

연구데이터 보존 기간			연구데이터 사본 백업 위치(복수)			리포지토리 제출		
	빈도	%		빈도	%		빈도	%
1년 이하	0	0%	출력물	6	11%	전부 제출	0	0%
1 - 2년	1	3%	개인 PC(집)	1	2%	일부 제출	7	23%
2 - 5년	5	17%	외부 파일 스토리지 서비스	1	2%	제출하지 않음	24	77%
5 - 10년	8	28%	소속기관 파일 스토리지 서비스	6	11%			
10 - 20년	4	14%	개인 PC(사무실)	14	25%			
20 - 100년	4	14%	이동식 저장장치	19	35%			
100년 이상	0	0%	CD/DVD/외장 하드 드라이브	8	15%			
연구 보관	4	14%	기타	0	0%			
모르겠음	3	10%						

연구자들의 연구데이터의 적절한 보존 기간에 대하여 질문한 결과, '5-10년 동안'이라고 응답한 비율이 28%로 가장 많았고, 다음으로 '2-5년'이 17%, '10-20년'와 '20-100년' 그리고 '영구 보관'이 동일하게 14%로 나타났다. 다음으로 연구데이터의 사본 백업 위치에 대해서는 전체의 60%가 이동식 저장장치와 개인 PC(사무실)에 백업하는 것으로 나타났다. 마지막으로 기관 및 외부 리포지토리에 데이터를 제출하느냐는 질문에 대해서는 전체의 77%가 제출하지 않는다는 의견을 보였으며 23%만이 전부가 아닌 일부를 제출한다고 응답하였다.

〈표 8〉 연구데이터 보존 - 2

기관 보존 의무			데이터 관리 담당 직원(복수)			데이터 업데이트 주기		
	빈도	%		빈도	%		빈도	%
네	26	81%	본인	25	50%	전혀 안함	2	7%
아니오	6	19%	연구 보조원	6	12%	매일	5	17%
			연구 프로젝트 관리자	9	18%	주간	5	17%
			프로젝트에서 담당자로 지정된 사람	6	12%	월별	8	27%
			외부 프로젝트 파트너	0	0%	분기별	3	10%
			소속기관 IT 직원	3	6%	매년	1	3%
			외부 리포지토리	0	0%	모르겠음	6	20%
			관리 하지 않음	0	0%			
			모르겠음	1	2%			

위의 〈표 8〉은 기관에서 생산된 데이터를 해당 기관이 보존해야 하는지 누가 데이터를 관리하는지 마지막으로 데이터를 얼마나 자주 업데이트 하는지에 대한 설문결과를 정리한 것이다. 우선, 연구자 본인이 생산한 데이터에 대하여 기관이 보존해야 하는지에 대해서는 81%가 '그렇다'고 응답하였다. 다음으로 누가 연구데이터 관리를 담당하고 있는지에 대해서 본인이 담당한다는 비율이 50%로 가장 많았고 다음으로 연구 프로젝트 관리자 18%, 연구 보조원과 프로젝트 내에서 담당자로 지정한 직원은 12%로 동일하게 응답하였다. 마지막으로 프로젝트 기간 동안 연구데이터를 얼마나 자주 업데이트하는지에 대한 질문에 대해서는 월별이라고 답한 의견이 27%로 가장 많았고, 다음으로 '모르겠음'이라는 응답이 20%, 다음으로 매일과 주간이 17%로 동일한 비율로 나타났다.

위의 〈표 7〉과 〈표 8〉을 분석한 결과, 임산공학 분야 연구자들의 연구데이터 보존과 관련하여 먼저 보존 기간은 5-10년을 선호하는 것으로 나타났고 연구데이터 사본은 이동식 저장장치와 개인 PC(사무실)에 백업하는 것으로 확인되었다. 또한, 기관 내/외부 리포지토리에 데이터를 제출하는지에 대한 답변으로 대부분이 제출하지 않은 것으로 나타났다. 연구데이터의 기관 보존 의무에서는 81%가 기관이 보존하는 거에 대해 찬성하는 것으로 나타났으며 데이터 관리 담당은 본인

이 직접 한다는 의견이 50%로 확인되었다. 마지막으로 연구데이터 업데이트는 월별로 한다는 의견이 27%로 가장 많았지만 모르겠다는 의견이 20%로 본인의 연구데이터 현황 파악이 제대로 되어 있지 않음을 시사하였다. 즉, 연구데이터 보존과 관련하여 보존 기간에 대해서는 기관 내 가이드라인 및 데이터 정책을 수립하는 것이 해결방안으로 제시될 수 있으며, 데이터 사본 백업은 RDM 시스템으로 리포지토리에 데이터 제출은 RDM 홍보와 교육으로 문제점을 해결할 수 있다. 마찬가지로 연구데이터 관리와 업데이트도 RDM 시스템을 통해 수행될 수 있음을 연구자들에게 주지할 수 있을 것이다.

### 5. 연구데이터 재사용

다음의 <표 9>는 임산공학 분야 연구자들의 연구데이터 재사용 등과 관련된 설문 결과를 정리한 것이다.

<표 9> 연구데이터 재사용

연구데이터 재사용 여부		재사용할 수 없는 이유			연구데이터 관리에 필요한 지원(복수)			
	빈도	%		빈도	%		빈도	%
모두 재사용 가능	11	34%	기밀성 또는 데이터 보호 문제	10	36%	교육	10	19%
일부만 재사용 가능	20	63%	라이선스 계약으로 공유 금지	0	0%	가이드라인	9	17%
재사용 불가능	1	3%	타 연구자가 이해할 수 없는 데이터임	15	54%	DMP 컨설팅	17	33%
			데이터가 더 이상 읽을 수 없는 형식으로 되어 있음	1	4%	스토리지 추가	9	17%
			기타	2	7%	필요 없음	7	13%

연구자들의 연구데이터 재사용 여부를 확인한 결과, 일부만 재사용 가능하다는 의견이 63%로 가장 많았고, 다음으로 모두 재사용 가능한 응답비율도 34%를 차지하였다. 다음으로 연구데이터를 재사용할 수 없는 이유에 대해서는 '타 연구자가 이해할 수 없는 데이터임'이라는 의견이 54%로 가장 많았고, 다음으로 '기밀성 또는 데이터 보호 문제'가 36%를 차지하였다. 기타로는 엑셀 데이터를 재사용하는 경우 실험방법에 따라 기술되어지는 것이 상이하게 때문에 해당 데이터를 설명하는 자료가 필요하다는 의견도 있었다. 마지막으로 연구데이터를 관리하기 위해 필요한 것이 무엇인지에 대한 질문한 설문결과를 분석한 결과, DMP 컨설팅이 33%로 가장 많았고, 다음으로 교육이 19%로 나타났고 가이드라인과 스토리지 추가가 17%로 동일하였다.

연구데이터 재사용과 관련한 응답을 분석한 결과, 대부분이 연구데이터의 일부 또는 전부를

재사용할 수 있다고 응답한 것으로 보아 수치 및 이미지 등 가독성이 뛰어난 연구데이터가 생산됨을 알 수 있다. 하지만, 재사용할 수 없는 이유에 대해서는 연구자가 이해할 수 없는 데이터이거나 기밀성 문제로 인한 부분을 지적하였는데 이 부분은 해당 데이터셋을 업로드 시 추가 기술(description) 자료를 첨부하도록 하면 해결될 수 있을 것으로 파악된다. 다음으로 연구데이터 관리에 필요한 것으로 DMP 컨설팅과 교육, 가이드라인 및 스토리지 추가와 같은 기본적인 RDM 시스템이 필요한 것으로 나타났다.

또한, 추가적으로 의견을 제시해달라는 질문에 대하여 연구데이터의 효율적인 관리 방안에 대한 지원과 최종 이용자가 해당 연구데이터를 활용 시 데이터를 이해하고 신뢰할 수 있도록 연구 윤리와 교육 그리고 가이드라인이 필요하다는 의견을 제시하기도 하였다.

## V. 결론

최근 데이터를 기반으로 하는 연구가 새롭게 출현함에 따라 천문, 항공, 우주, 유전자 등 첨단 분야에서 센서와 같은 실험 장비의 발전으로 데이터 급격히 증대되고 있다(한국과학기술정보연구원 2019). 또한, 해외 대학 및 연구기관에서는 연구기금으로 생산된 연구결과물이 국가 및 기관 자산으로 재사용될 수 있도록 체계적인 관리 및 보존되어야 한다는 환경이 만들어졌다. 국내에서도 늦었지만 법령에 연구데이터의 정의와 DMP 적용 대상에 대한 근거를 마련하였다(국가연구개발사업의 관리 등에 관한 규정 2020.3.17. 개정).

국립산림과학원에서는 2019년 10월에 「국립산림과학원 연구데이터 관리 기본계획」이 수립되었다. 해당 계획이 수립된 배경은 규정이나 지침과 같은 연구데이터 관리정책과 연구데이터 관리 프로세스의 부재 그리고 연구데이터에 대한 개인관리로 공유 및 재활용이 어렵다는 점에 있었다. 따라서 국립산림과학원은 효율적인 데이터 관리와 공유를 위한 DMP와 관리 시스템의 도입을 검토하였고 연구데이터 관리 실행내용을 정책에 명시하도록 하였다. 이러한 기관의 노력으로 현재 국립산림과학원 연구데이터 관리 기본계획이 수립되었고 연구데이터 아카이빙 플랫폼(DataNest)<sup>2)</sup>을 도입하여 운영 중에 있다. 실제로 연초에 연구과제에 대한 DMP 제출이 의무화되었고 연말에 DMP에 따라 연구데이터를 플랫폼에 업로드하고 있다.

본 연구의 목적은 임산공학 연구자들의 연구데이터 관리 현황을 파악하기 위함이다. 이러한 목적을 수행하기 위하여 먼저, 임산공학 분야의 학문 분류와 해당 분야의 실제 실험 내용을 파악하여 해당 분야의 데이터 특성에 대하여 확인하였다. 또한 설문조사를 실시하였으며 설문조사에 사용된

2) 현재는 DataFos로 명칭이 바뀌었음

질문지는 DAF의 설문문항을 최대한 준용하되 국내 현황과 맞지 않은 문항은 수정하여 연구자들에게 온라인으로 배포하였다. 설문 조사 결과로 확인된 연구데이터 관리 현황은 다음과 같이 연구데이터 생성 방법 및 유형, 공유, 저장, 보존 그리고 재사용 등 5개의 카테고리로 분석하였다.

- 연구데이터 생성 방법 및 유형: 연구데이터는 대부분 관찰과 실험을 통해 생성되며 데이터 유형은 엑셀 시트와 이미지 및 실험실 노트가 가장 많음. 특히 정형 및 비정형데이터를 관리할 수 있는 RDM 시스템 도입이 필요함
- 연구데이터 공유: 연구데이터 공유 수단은 이메일, 파일 스토리지 서비스 그리고 이동식 저장장치를 가장 많이 사용하며 데이터 공유 시 문제점으로 파일 네이밍 그리고 데이터 버전관리 등과 같은 기술적인 문제를 제시하고 있음. 이러한 부분은 RDM 교육 및 연구데이터 가이드라인(규정)의 수립을 통해 해결될 수 있음<sup>3)</sup>
- 연구데이터 저장: 연구데이터는 대부분의 연구자들이 사무실의 개인 PC에 저장하고 있고 보유하고 있는 연구데이터 크기는 1-50GB가 가장 많음. 특히 개인 연구자들은 스토리지 문제에 대하여 개인별 스토리지를 따로 구입한다고 응답한 바 RDM 시스템 운용 시 개인별 스토리지 확대가 필요함
- 연구데이터 보존: 연구데이터 보존 연한은 5-10년을 가장 많이 선호하고 연구데이터 사본은 이동식 저장장치에 백업하고 있음. 또한 데이터 보존을 위해 리포지토리는 거의 사용하고 있지 않으며 데이터 관리는 연구자 본인이 하는 것으로 파악되었음. 이러한 부분은 기관의 연구데이터 관리 규정 내에 세부적인 지침 사항을 마련하여 연구자들의 RDM에 대한 이해를 확보해야 함
- 연구데이터 재사용: 연구자 대부분 본인의 연구데이터의 재사용에 대하여 일부 또는 전체가 가능하다고 하였음. 또한, 재사용할 수 없는 이유에 대하여 연구자가 이해할 수 없는 데이터이거나 기밀성 또는 데이터 보호 문제가 가장 많았음. 연구지원 부서 내에 RDM에 대한 교육 그리고 홍보를 전담할 직원을 두어 해당 연구자들에 대한 온라인 및 오프라인 교육이 강화되어야 함

위에서 분석한 설문결과 내용 중 문제점으로 제시된 부분에 대해서는 연구데이터 저장, 공유, 보존 및 재사용을 위한 RDM 시스템의 도입을 통해 해결할 수 있을 것으로 판단된다. 특히 RDM 시스템에는 연구자를 대상으로 한 RDM 교육, 기관의 데이터 정책 및 가이드라인 제시, 개인별 스토리지 확대, 리포지토리 안내 그리고 연구 윤리 교육 등이 포함된 구체적인 실행 방안이 포함되어야 할 것이다. 여기에는 구체적이고 실행가능한 단계별 로드맵을 구성하여 연구자들이 RDM

3) 2019년 10월에 국립산림과학원 연구데이터 관리 규정이 마련되어 있지만 세부 지침(파일 네이밍 등)이 추가적으로 필요함

의 이해부터 수행까지를 포괄할 수 있는 연구데이터 관리 정책이 추가되어야 할 것이다.

4차 산업혁명 등 데이터가 중요시되고 있는 사회에 있어 이러한 연구데이터에 대한 체계적인 관리 및 보존에 대한 연구가 많이 수행되어야 한다. 여기에는 학문 분야별 특징이 다르므로 각 전문 영역에서 다루어지는 연구데이터의 특징과 소속 연구자의 연구데이터 관리 현황에 대한 파악이 이루어져야 함은 마땅할 것이다. 본 연구가 국립산림과학원이라는 특정 연구 기관을 대상으로 연구데이터 관리 현황을 파악한 것은 연구의 제한점으로 제시될 수 있지만 타 기관에서도 DAF와 같은 공식적이고 체계적인 도구를 이용하여 해당 분야의 연구데이터 관리 현황을 파악할 수 있을 것으로 기대한다. 또한, 기관에서는 DAF와 같은 도구를 이용하여 연구데이터 관리를 위한 구체적인 정책이나 서비스를 제시하기 위한 기초 조사로서 연구데이터 관리에 대한 요구사항과 문제점을 파악할 수 있기를 희망해본다. 마지막으로 연구데이터 관리에 대한 요구사항과 문제점을 해결하기 위한 RDM 기능 및 구성요소에 대한 연구가 진행되어 실질적인 연구데이터 관리 프로세스를 구현하여야 할 것이다.

## 참고문헌

- 국립산림과학원. 2019. 『국립산림과학원 연구데이터 관리 기본계획』. 서울: 국립산림과학원.
- 국립산림과학원. 2018. 『제2차 중장기 기술개발계획』. 서울: 국립산림과학원.
- 국립산림과학원. 2020. 조직·직원. <[https://nifos.forest.go.kr/kfsweb/kfi/kfs/cms/cmsView.do?cmsId=FC\\_001661&mn=UKFR\\_04\\_02](https://nifos.forest.go.kr/kfsweb/kfi/kfs/cms/cmsView.do?cmsId=FC_001661&mn=UKFR_04_02)> [cited 2020. 6. 14].
- 과학기술정보통신부, 한국과학기술기획평가원. 2018. 『2018년 국가과학기술표준분류체계 연구분야 해설서』. 한국과학기술기획평가원(KISTEP): 충북 음성.
- 국가연구개발사업의 관리 등에 관한 규정. 대통령령 제29625호 2020.
- 김선태, 이태영, 김용. 2012. 해양관측데이터 관리를 위한 결정요소 도출. 『정보관리연구』, 43(3): 97-115.
- 김지현. 2012. 대학 내 연구자들의 연구데이터 관리에 관한 연구. 『한국도서관·정보학회지』, 43(3): 433-455.
- 대통령직속 4차산업혁명위원회. <<https://www.4th-ir.go.kr/>> [cited 2020. 2. 5].
- 박미영, 안인자, 남승주. 2018. 과학기술분야 출연연구기관 연구데이터 관리 및 공유 사례 분석 연구. 『한국비블리아학회지』, 29(4): 319-344.
- 신은정 외. 2017. 『오픈 사이언스 정책의 도입 및 추진 방안』. 세종: 과학기술정책연구원, 정책연구 2017-08.
- 신은정. 2018. 『오픈사이언스시대 연구데이터의 공유·활용 동향과 시사점』. 과학기술정책연구원, 2018년

- 6월 25일. <[http://www.krnet.or.kr/board/data/dprogram/2295/C3-1\\_%BD%C5%C0%B A%C1%A4.pdf](http://www.krnet.or.kr/board/data/dprogram/2295/C3-1_%BD%C5%C0%B A%C1%A4.pdf)> [cited 2020. 4. 2].
- 최명석, 이승복, 이상환. 2017. 국내 과학기술분야 연구기관의 과학데이터 관리 현황. 『한국콘텐츠학회논문지』, 17(12): 117-126.
- 한국과학기술정보연구원. 2019. 『출연 연구데이터 관리·활용 방안 연구』. 대전: 한국과학기술정보연구원.
- CGIAR. 2017. *RESEARCH DATA MANAGEMENT Good Practice Note SMB-ARC 13/9/17*. <<https://storage.googleapis.com/cgiarorg/2018/01/GPN-Research-Data-Management-Sept-2017.pdf>> [cited 2020. 3. 12].
- Cox, Andrew et al. 2017. "Developments in Research Data Management in Academic Libraries: Towards an Understanding of Research Data Service Maturity". *Journal of the Association for Information Science and Technology*, 68(9): 2182-2200.
- DCC. n.d. *Data Asset Framework*. <<http://www.dcc.ac.uk/resources/tools/data-asset-framework>> [cited 2020. 2. 21].
- JISC, University of Glasgow and DCC. 2009. Data Asset Framework Implementation Guide. <[https://www.data-audit.eu/docs/DAF\\_Implementation\\_Guide.pdf](https://www.data-audit.eu/docs/DAF_Implementation_Guide.pdf)> [cited 2020. 2. 21].
- Mohammed, Mahdi & Ibrahim, Rafea. 2019. "Challenges and Practices of Research Data Management in Selected Iraq Universities". *DESIDOC Journal of Library & Information Technology*, 39: 308-314.
- OECD. 2015. "Making Open Science a Reality." *OECD Science, Technology and Industry Policy Papers*. No. 25. <<https://doi.org/10.1787/5jrs2f963zs1-en>> [cited 2019. 3. 8].
- Pilat, D. and Y. Fukasaku. 2007. "OECD Principles and Guidelines for Access to Research Data from Public Funding". *Data Science Journal*, 6: 4-11.
- Plomp, Esther et al. 2019. "Cultural obstacles to research data management and sharing at TU Delft". *Insights the UKSG journal*, 32.
- PM, Naushad and Sidra Saeed. 2019. "Research Data Management and Data Sharing among Research Scholars of Life Sciences and Social Sciences". *DESIDOC Journal of Library & Information Technology*, 39: 290-299.
- Tenopir C, et al. 2011. "Data Sharing by Scientists: Practices and Perceptions". *PLoS ONE*. 6(6): e21101.
- The Pennsylvania State University. 2020. *Forest Science Research*. <<https://ecosystems.psu.edu/research/areas/forest-science>> [cited 2020. 5. 14].
- The University of Sheffield. 2020. What is research data management?. <<https://www.sheffield.ac.uk/library/rdm/whatisrdm>> [cited 2020. 3. 14].
- University of Leeds. 2020. Research data management explained. <[https://library.leeds.ac.uk/info/14062/research\\_data\\_management/61/research\\_data\\_management\\_explained](https://library.leeds.ac.uk/info/14062/research_data_management/61/research_data_management_explained)> [cited

2020. 3. 12].

University of Pittsburgh. 2020. Research Data Management @ Pitt. <<https://pitt.libguides.com/managedata>> [cited 2020. 3. 19].

University of Turku. 2020. What is research data?. <<https://utuguides.fi/friendly.php?s=researchdata>> [cited 2020. 3. 15].

Vilar, P. and V. Zabukovec. 2019. "Research data management and research data literacy in Slovenian science." *Journal of Documentation*, 75(1): 24-43.

Wiley, C., & W. H. Mischo, 2016. "Data management practices and perspectives of atmospheric scientists and engineering faculty". *Issues in Science and Technology Librarianship*, 85.

Xiujuan Chen and Wu Ming. 2017. Survey on the Needs for Chemistry Research Data Management and Sharing. *The Journal of Academic Librarianship*, 43(4): 346-353.

Yale University Library. 2020. What is research data?. <<https://guides.library.yale.edu/datamanagement>> [cited 2020. 3. 14].

#### 국한문 참고문헌의 영문 표기

(English translation / Romanization of reference originally written in Korean)

Kim, Sun-Tae, Lee, Tae-Young, & Kim, Yong. 2012. Deriving the Determining Factor for the Management of Oceanographic Data. *Journal of Information Management*, 43(3): 97-115.

Kim, Jihyun. 2012. A Study on University Researchers' Data Management Practices. *Journal of Korean Library and Information Science Society*, 43(3): 433-455.

Park, M., Ahn, I., & Nam, S. 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.

Choi, M.-S., S.-B. Lee, & S. Lee. (2017). Research Data Management of Science and Technology Research Institutes in Korea. *The Journal of the Korea Contents Association*, 17(12): 117-126.



## 설문지

1. 현재 연구데이터를 보유하고 있습니까?

- 네
- 아니요 [GO TO END]

2. 보유하고 계신 1차 데이터는 어떤 종류의 데이터입니까?

[복수 선택 가능]

※ 1차 데이터는 연구자가 직접 수집 및 생산한 데이터입니다.

- 1차 데이터를 보유하지 않음 [GO TO QUESTION 4]
- 컴퓨터 프로그램에서 자동으로 생성되는 데이터
- 센서 또는 계측기에서 수집된 데이터(질문표 포함)
- 이미지, 스캔 또는 X선
- 웹사이트
- MS Word 파일
- 엑셀 시트
- SPSS 파일
- 디지털 오디오 파일
- 디지털 비디오 파일
- 현장 작업 데이터
- 실험실 노트
- 사진 모음
- 비디오 테이프
- 오디오 테이프
- 슬라이드
- 텍스트 말뭉치
- 문서 또는 보고서
- 기타

기타 경우, 내용을 자유롭게 기술하십시오.

2-1 데이터의 생산 방법을 선택 하십시오. [복수 선택 가능]

- 관찰
- 실험
- 참조
- 파생
- 시뮬레이션됨
- 기타

기타 경우, 내용을 자유롭게 기술하십시오.

3. 보유하는 1차 데이터 수집 및 생산에 누가 연구비를 지원했습니까? [복수 선택 가능]

- 정부(소속기관 외)
- 소속기관
- 민간기업

기타

기타 경우, 내용을 자유롭게 기술하십시오.

4. 연구에 활용하는 2차 데이터는 누가 수집했습니까? [복수 선택 가능]

※ 2차 데이터는 제 3자에 의해 제공되는 연구데이터입니다.

2차 데이터를 활용하지 않음 [GO TO QUESTION 6]

소속기관 외 연구자

소속기관

민간기업

기타

기타 경우, 내용을 자유롭게 기술하십시오.

5. 2차 데이터의 유형은 무엇입니까? [복수 선택 가능]

※ 2차 데이터는 제 3자에 의해 제공되는 연구데이터입니다.

컴퓨터 프로그램에서 자동으로 생성되는 데이터

센서 또는 계측기에서 수집된 데이터(질문표 포함)

이미지, 스캔 또는 X선

웹사이트

MS Word 파일

엑셀 시트

SPSS 파일

디지털 오디오 파일

디지털 비디오 파일

현장 작업 데이터

실험실 노트

사진 모음

비디오 테이프

오디오 테이프

슬라이드

텍스트 말뭉치

문서 또는 보고서

기타

기타 경우, 내용을 자유롭게 기술하십시오.

6. 나머지 질문은 현재 보유하고 있는 모든 연구데이터와 관련이 있습니다. 데이터를 사용하거나 생성할 때 다른 사람과 공동 작업을 했습니까?

네.

아니요. [GO TO QUESTION 9]

7. 공동 작업을 할 때 어떻게 데이터를 공유합니까? [복수 선택 가능]

동료에게 파일을 이메일로 전송

외부 파일 스토리지 서비스(Dropbox, Google Drive, NDrive 등)

소속기관 파일 스토리지 서비스

CD, DVD, USB 등 사용

DAF(Data Asset Framework)를 활용한 임산공학 분야 연구자들의 연구데이터 관리 개선 방안

기타

기타 경우, 내용을 자유롭게 기술하십시오.

8. 공동 작업할 때 발생하는 문제는 무엇이었습니까? [복수 선택 가능]

없음

소속기관 파일 스토리지 서비스 기능 한계

외부 파일 스토리지 서비스(Dropbox, Google Drive, NDrive 등) 기능 한계

파일 이름 생성 규칙 부재

파일 버전 관리의 어려움

데이터 전송 관련 법적 문제

데이터 소유권 문제

기타

기타 경우, 내용을 자유롭게 기술하십시오.

9. 데이터를 어디에 저장하십니까? (백업 복사본 제외) [복수 선택 가능]

출력물(On Paper)

개인 PC(집)

외부 파일 스토리지 서비스(Dropbox, Google Drive, NDrive 등)

소속기관 파일 스토리지 서비스

개인 PC(사무실/연구실)

이동식 저장장치(USB/SD 카드 등)

CD/DVD/외장 하드 드라이브

기타

기타 경우, 내용을 자유롭게 기술하십시오.

9-1 현재 보유하고 있는 연구데이터 크기를 선택하십시오.

1 GB 미만

1 - 50GB

50 - 100GB

100-500GB

500GB - 1TB

1 - 50TB

50 - 100TB

100TBs 초과

알 수 없음

10. 파일 크기로 인해 연구데이터를 저장하는 데 문제가 생긴 적이 있습니까?

네

아니오 [GO TO QUESTION 12]

11. 이러한 저장 공간 문제를 어떻게 극복하셨습니까? [복수 선택 가능]

외부 파일 스토리지 서비스(Dropbox, Google Drive, NDrive 등)

소속기관 파일 스토리지 서비스

추가 저장 장치 구입(외장 하드디스크, USB 등)

한국도서관·정보학회지(제51권 제2호)

- 소속기관에게 저장 공간 요청
  - 기타
- 기타 경우, 내용을 자유롭게 기술하십시오.

12. 현재 보유하고 있는 연구데이터가 백업되어 있습니까?

- 전부 백업
- 일부 백업
- 백업하지 않음 [GO TO QUESTION 14]

12-1 데이터 보유기간을 선택하세요.

- 1년 이하
- 1 - 2년
- 2 - 5년
- 5 - 10년
- 10 - 20년
- 20 - 100년
- 100년 이상
- 영구 보관
- 모르겠음
- 기타

기타 경우, 내용을 자유롭게 기술하십시오.

13. 어디에 데이터 사본을 백업하십니까? [복수 선택 가능]

- 출력물(On Paper)
- 개인 PC(집)
- 외부 파일 스토리지 서비스(Dropbox, Google Drive, NDrive 등)
- 소속기관 파일 스토리지 서비스
- 개인 PC(사무실/연구실)
- 이동식 저장장치(USB/SD 카드 등)
- CD/DVD/외장 하드 드라이브
- 기타

기타 경우, 내용을 자유롭게 기술하십시오.

14. 기관 리포지토리 혹은 외부 리포지토리에 데이터를 제출하십니까?

- 전부 제출 [GO TO QUESTION 14-1]
- 일부 제출 [GO TO QUESTION 14-1]
- 제출하지 않음

14-1 이용하고 있는 리포지토리의 이름은 무엇입니까?

15. 귀하의 데이터는 국립산림과학원이 보존해야 한다고 생각하십니까?

- 예
- 아니오. [GO TO QUESTION 16]

DAF(Data Asset Framework)를 활용한 임산공학 분야 연구자들의 연구데이터 관리 개선 방안

15-1 누가 데이터 관리를 담당하고 있습니까? [복수 선택 가능]

- 본인
- 연구 보조원
- 연구 프로젝트 관리자
- 프로젝트에서 담당자로 지정된 사람
- 외부 프로젝트 파트너
- 국립산림과학원 IT 직원
- 외부 리포지토리
- 관리하지 않음
- 모르겠음
- 기타

15-2 프로젝트(과제) 기간 동안 데이터를 얼마나 자주 업데이트하십니까?

- 전혀 안함
- 매일
- 주간
- 월별
- 분기별
- 매년
- 모르겠음

16. 귀하가 보유하고 있는 데이터를 다른 사람이 재사용할 수 있습니까?

- 모두 재사용 가능 [GO TO QUESTION 18]
- 일부만 재사용 가능
- 재사용 불가능

17. 귀하의 데이터를 재사용하거나 공유할 수 없는 이유는 무엇입니까? [[복수 선택 가능]

- 기밀성 또는 데이터 보호 문제
- 라이선스 계약으로 공유 금지
- 타 연구자가 이해할 수 없는 데이터임
- 데이터는 더 이상 읽을 수 없는 형식으로 되어 있음
- 기타

기타 경우, 내용을 자유롭게 기술하십시오.

18. 귀하의 연구데이터를 관리하기 위해 필요한 것은 무엇입니까? [복수 선택 가능]

- 교육(Training)
- 지침서(Guideline)
- 데이터 관리 계획(DMP) 작성 지원 컨설팅
- 스토리지 추가
- 필요 없음
- 기타

기타 경우, 내용을 자유롭게 기술하십시오.

19. 위의 답변 중 하나를 더 자세히 설명하거나 더 자세히 언급하고 싶은 경우, 여기에서 자유롭게 기술하십시오.