

연구데이터커먼즈 프레임워크 및 허브 서비스를 위한 리소스 공유 정책

이미경*, 조민희*, 송사광***, 임형준*

*한국과학기술정보연구원 연구데이터공유센터

**과학기술연합대학원대학교 응용 AI학과
{jerryis, mini, imsmallj, hjyim}@kisti.re.kr

A Study on the Resource Sharing Policy of Korea Research Data Commons Framework and Hub Service

Mikyoung Lee*, Minhee Cho*, Sa-kwang Song***, Hyung-Jun Yim*

*Research Data Sharing Center, KISTI

**Dept. of Applied AI, UST

요약

오픈사이언스와 생성형 AI의 등장으로 데이터에 대한 개방·공유의 필요성이 강조되고 있다. 글로벌 개방형 연구데이터 커먼즈를 구현하기 위해 RDA(Research Data Alliance)는 Global Open Research Commons International Model(GORC IM)을 발표하였다. GORC IM은 거버넌스 및 리더십, 참여 및 접근 규칙, 지속 가능성, 약속, 사람의 역할, ICT 인프라, 상호운용성, 표준 및 협약, 서비스 및 도구, 연구 객체, KPI 및 메트릭스의 필수 요소를 만족하는 글로벌 상호 운용 가능한 연구 커먼즈 개발을 지원하는 모델이다. 우리는 연구데이터와 관련된 리소스의 공유·활용이 가능한 KRDC 체계를 구축하고 있으며, KRDC는 허브서비스와 다수의 프레임워크로 구성된다. KRDC프레임워크는 연구기관이나 커뮤니티에서 연구데이터와 관련된 리소스의 관리·활용을 위한 것이고, 허브서비스는 프레임워크에서 관리하는 연구 리소스들을 공유·활용하는 것을 목적으로 한다. KRDC는 글로벌 개방형 연구데이터커먼즈를 실현하기 위해 GORC IM의 상호 운용성, 참여와 접근 규칙에 해당하는 정책 방안을 연구하고 있다. 본 논문에서는 KRDC 프레임워크와 허브서비스의 리소스 공유 정책을 설명하고 있으며 향후, 관련 정책을 보완한 후 글로벌 개방형 연구데이터 커먼즈로 발전시킬 예정이다.

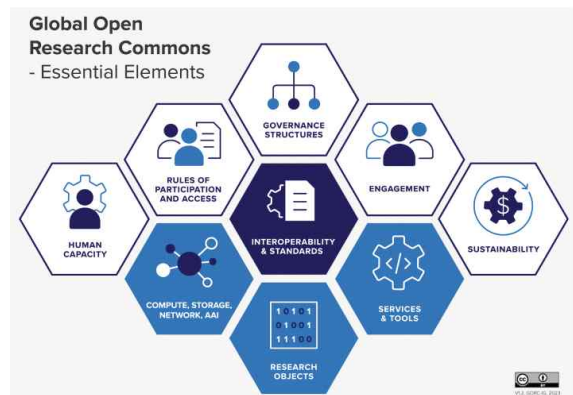
I. 서론

데이터 기반의 연구가 늘어나면서 연구 결과 및 과정을 개방·공유하는 오픈사이언스에 대한 중요성이 강조되고 있다. 유럽에서는 오픈사이언스 커먼즈(Open Science Commons)를 통해 디지털 서비스, 과학 실험, 데이터, 지식, 경험 등을 쉽게 공유하고 협업할 수 있는 연구를 수행하고 있다 [1]. 또한 ChatGPT와 같은 거대 언어모델(Large Language Model)을 이용한 생성형 AI의 등장으로 데이터에 대한 중요성이 부각되고 있다. 국내에서도 디지털플랫폼 정부를 지향하며 데이터 공유·활용을 통한 협업에 대한 필요성이 증대되고 있다. 이러한 국내의 오픈 사이언스 움직임에 맞춰 한국과학기술정보연구원에서는 연구데이터 공유·활용의 저변확대를 위한 국가연구데이터플랫폼(DataON)을 운영 중이며, 연구데이터와 관련된 리소스의 공유·활용 체계인 연구데이터커먼즈(Korea Research Data Commons, KRDC)를 구축하고 있다.

II. 관련 연구

2023년 10월 RDA에서 발표한 GORC IM은 글로벌 개방형 연구데이터 커먼즈 구현을 위해 그림 1과 같이 필수 요소들을 정의하였으며, 상호운용이 가능한 커먼즈 개발을 지원하는 것을 목표로 한다[2]. GORC IM의 구조는 거버넌스 및 리더십, 참여 및 접근 규칙, 지속 가능성, 약속, 사람의 역할, ICT 인프라, 상호운용성, 표준 및 협약, 서비스 및 도구, 연구 객체,

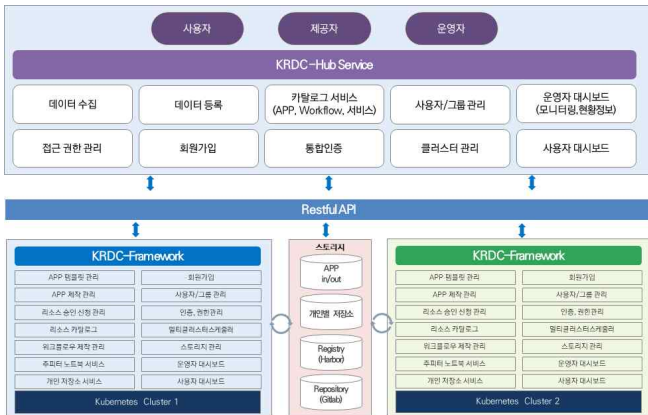
KPI 및 메트릭스로 구성된다. GORC IM의 목적은 세계적으로 상호운용 가능한 연구 서비스를 개발하려는 연구자들에게 공통의 프레임워크와 모델을 제공하는 것이다. KRDC도 글로벌 차원의 오픈사이언스를 실현하기 위해 GORC IM의 필수 요소를 충족시키는 것이 필요하다. KRDC의 상호운용성은 정보 구조화 방법에 따른 부문 상호운용성, 데이터가 동일한 방식으로 해석되는 의미 상호운용성, 기관과 협약하는 조직간 상호운용성 측면에서 GORC IM의 상호운용성에 부합한다[3].



(그림 1) RDA GORC IM

III. KRDC 프레임워크와 허브 서비스

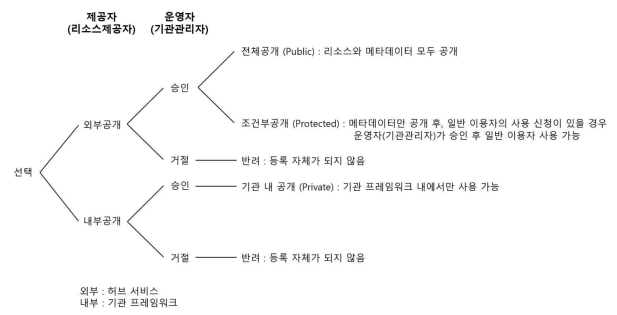
연구데이터의 분석·활용에 사용되는 리소스(모델, SW, 인프라 등)를 공유·활용하는 체계인 국가연구데이터커먼즈(Korea Research Data Commons, KRDC)는 그림 2와 같이 다수의 KRDC 프레임워크와 허브 서비스로 구성된다[4][5]. KRDC 프레임워크는 연구기관이나 커뮤니티에서 연구데이터 분석·활용을 위한 리소스를 관리·활용하는 것이다. 기관이나 커뮤니티에 소속된 연구자가 다양한 유형의 연구데이터를 분석하는데 필요한 리소스를 사용 및 재사용하기 위해 상호운용 서비스를 제공한다. 데이터 중심의 연구개발에 필요한 리소스를 자유롭게 발견, 접근, 상호운용, 재사용(FAIR)할 수 있도록 리소스를 공유·활용하는 기능을 제공하고, 연구데이터를 분석하는데 필요한 소프트웨어 및 실행 환경을 구축하는데 편리함을 제공하기 위해 연구소프트웨어 개발, 배포, 관리를 지원하는 핵심 기능을 제공한다. 프레임워크에서 제공되는 연구소프트웨어의 실행 환경을 탑재한 워크플로우 서비스를 제공하여 연구자가 연구데이터를 쉽게 분석할 수 있다. KRDC 허브 서비스는 다수의 KRDC 프레임워크를 연계하여 리소스를 수집하여 카탈로그 서비스를 통해 통합 검색 서비스를 제공한다. 프레임워크에 등록된 리소스들이 공개, 비공개 조건에 따라 검색되며, 서비스 이용자는 검색하여 필요한 리소스를 사용할 수 있다. 카탈로그에서 제공되는 리소스인 앱과 워크플로우는 단일 실행, 연계 실행이 가능하기 때문에 KRDC 허브서비스와 프레임워크 간에 리소스 공개 및 접근에 관한 정책이 필요하다.



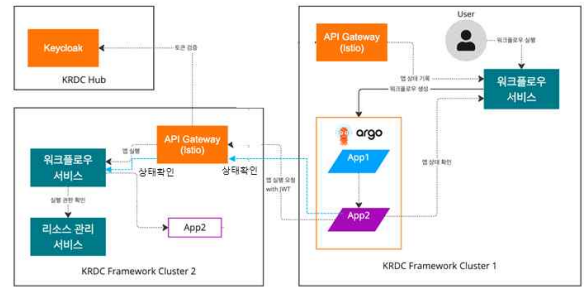
(그림 2) KRDC 프레임워크와 허브 서비스

IV. KRDC 리소스 공유 정책

이 장에서는 KRDC 프레임워크와 허브 서비스를 위해 필요한 리소스 공유 정책에 대해 설명한다. 이것은 GORC IM의 필수 요소인 참여와 접근 규칙에 해당된다. KRDC 허브서비스와 프레임워크의 구조는 그림 2와 같으며, 각 프레임워크는 멀티클러스터로 구성되어 있어서 사용자 인증/인가를 위한 통합인증 기능과 서로 간의 참여와 접근 규칙이 꼭 필요하다. KRDC 프레임워크를 운영하는 연구기관이나 커뮤니티의 경우, 허브 서비스와 연계를 통해 KRDC 체계에 참여할 수 있다. KRDC 체계는 서로 다른 기관과 커뮤니티에서 관리·운영하는 KRDC 프레임워크들 간의 리소스를 공유하고 재활용할 수 있는 것이다. KRDC 리소스 공개 및 접근 규칙은 그림 3과 같다. 리소스 제공자가 KRDC 프레임워크에 리소스를 등록할 때, 허브 서비스에 공개 또는 비공개를 선택할 수 있다. 기관의 운영자는 기관의 공개 정책에 따라 리소스 등록을 승인하게 된다. 리소스 공개 및 접근 규칙은 그림 4와 같이 허브 서비스의 통합인증을 통해 사용자의 리소스 접근 권한을 확인 후, 워크플로우 서비스를 통해 사용할 수 있다.



(그림 3) KRDC 프레임워크의 리소스 공개 및 접근 규칙



(그림 4) KRDC 프레임워크 간 통합 인증을 통한 실행

V. 결론

본 논문에서는 KRDC 프레임워크와 허브서비스를 위한 리소스 공유 정책에 대해 설명하였다. KRDC는 GORC IM의 필수 요소에 부합되는 글로벌 개방형 연구 커먼즈를 지향하고 있다. KRDC는 구문 상호운용, 의미 상호운용, 조직간 상호운용을 통해 IM 상호운용성을 만족시키고 있으며, KRDC 허브 서비스와 프레임워크의 리소스 공유 정책은 IM 참여와 접근 규칙 요소에 부합한다. 향후, KRDC는 GORC IM의 9가지의 필수 요소를 갖춘 글로벌 연구데이터 커먼즈로 발전시킬 것이다.

ACKNOWLEDGMENT

이 논문은 2024년도 한국과학기술정보연구원(KISTI)의 기본사업 “연구데이터와 인프라의 공유·활용 체제 구축”으로 수행된 연구입니다. [과제번호:(KISTI) K24L1M2C3]

참고 문헌

- [1] Andreozzi, S., Appleton, O., Coelho, S., Ferrari, T., Holsinger, S. and Legre, Y., Open Science Commons: <https://zenodo.org/records/4884748>, 2015.
- [2] The Global Open Research Commons International Model Report, Version 1: <https://zenodo.org/records/10040098>, 2023.
- [3] 송사광, 조민희, 이미경, 임형준, “글로벌 상호운용성 확보를 위한 국가 연구데이터 커먼즈 연구”, 한국통신학회 추계종합학술대회, 2023.
- [4] 임형준, 이미경, 송사광, 서동민, 조민희, “데이터 기반 연구개발 위한 국가연구데이터커먼즈 설계 및 적용 방안”, 특집호: 빅데이터 분석 및 시각화, 기법 및 응용, 제32권 5호, 한국지능시스템학회 논문지, 2022.
- [5] 이미경, 박성은, 조민희, 임형준, “국가연구데이터커먼즈 허브 서비스 설계”, 한국통신학회 추계종합학술대회, 2023.