

전주기 연구활동 지원을 위한 연구데이터 리포지터리 개발

- 전주기 연구활동 지원을 위한 연구데이터 리포지터리는 연구활동 주기별로 연구데이터 관리를 위한 기능을 제공한다. 연구 활동 주기는 크게 보면, 연구데이터관리계획 수립, 연구수행을 통한 연구데이터 생성, 연구수행완료 후, 생성된 연구데이터의 등록 및 공유로 나뉜다.
- 따라서, 각 연구데이터를 다루는 연구활동에 필요한 기능을 연구데이터 리포지터리에서 지원한다. 첫째, 연구데이터 관리계획수립 부분에서는 연구데이터관리계획양식을 작성하고, 양식에 따라 연구데이터 관리계획서를 제출할 수 있는 기능을 지원한다. 둘째, 연구수행 중 생산되는 연구데이터에 대해서 자유롭게 기술할 수 있도록, Key/Value 형태의 메타데이터를 연구자가 작성할 수 있는 기능과, 연구자의 계정에 개인 저장소를 할당하고 연구데이터를 업로드/다운로드, 그리고 분석환경과 연계할 수 있는 기능을 제공한다. 셋째, 연구수행 완료 후, 연구데이터 등록을 위해, 연구데이터등록양식을 작성하는 기능과 양식에 따라 연구데이터를 등록 및 부서장 또는 승인/검토 연구자에 의해 검토/승인할 수 있는 기능을 제공한다. 또한 연구데이터관리계획관리 기능을 통해 작성된 연구데이터관리계획서를 메타데이터의 항목으로 연결할 수 있는 기능을 제공함으로써, 어떠한 연구데이터관리계획서로부터 산출된 연구데이터인지를 확인할 수 있도록 한다. 연구데이터관리계획서에는 관련된 연구과제 또한 기술할 수 있기 때문에 연결된 연구데이터관리계획서를 통해서 연구데이터와 관련된 연구과제 또한 파악할 수 있다.

□ 현재, 당해연도에 개발한 전주기 연구 활동 지원을 위한 연구데이터 리포지터리의 전체적인 시스템 개념도는 다음과 같다.

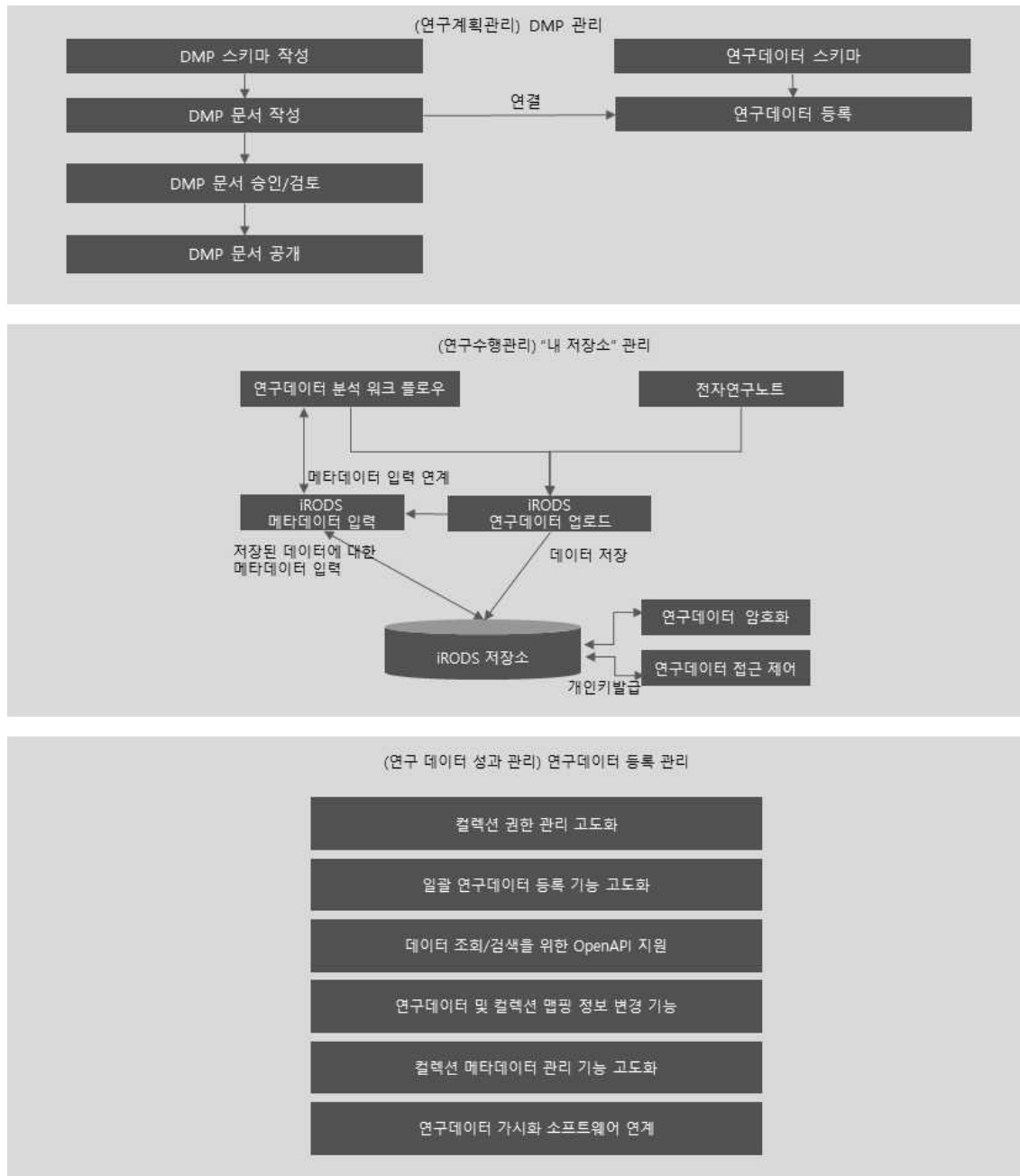


그림 1. Overall System Architecture of Research Data Repository for supporting research data life-cycle

□ 연구데이터리포지터리는 2018년도부터 개발을 시작하였으며, 당해년도에는 전주기 연구 활동 지원을 위해 연구계획관리(DMP:Data Management Plan) 기능과 “내 저장소” 관리 기능을 개발하고, 연구데이터 등록 및 검색 공유 기능을 고도화하였다.

○ 연구계획관리 기능에서는 연구데이터관리계획 양식을 작성하는 기능을 개발하였다. 연구데이터관리계획은 기관마다 양식이 다를 수 있기 때문에, 기존 연구데이터의 메타데이터관리양식을 편집하는 기능을 활용하여 연구데이터관리계획양식을 편집할 수 있도록 지원한다. 연구데이터관리계획양식이 작성되면, 양식에 따라 연구데이터관리계획을 작성할 수 있다. 연구데이터관리계획서는 원내 또는 대외 공개를 할 수 있으며, 연구데이터 등록 시 메타데이터의 항목으로써, 작성된 연구데이터관리계획서를 추가 및 연결할 수 있는 기능을 제공한다.

○ 연구 수행 중의 데이터 관리를 위한 “내 저장소” 관리 기능에서는 Python, R 과 같은 프로그래밍 환경에서 작동하는 분석 워크 플로우에서 커넥터를 통해 연구데이터를 업로드/다운로드/편집할 수 있는 기능을 제공한다. 또한 전자연구노트와 연계를 통해 전자연구노트에 등록된 데이터를 “내 저장소”로 반입할 수 있다. 연구 수행 중에는 연구데이터에 대하여 실험 방법, 실험 단계, 기타 필요한 정보를 기술해야 될 필요가 있을 경우에는 메타데이터를 key/value 형태로 작성할 수 있다. 연구데이터 파일은 명칭을 통해 검색 가능하다. 또한 연구 수행 중의 연구데이터 보호를 위해 개인 암호화/복호화 기능을 제공한다.

○ 연구데이터 등록관리 기능에서는 기존 컬렉션 권한 관리를 세분화하여 관리자 계정뿐만 아니라, 컬렉션 관리자로 지정된 사용자 계정은 컬렉션을 생성 및 관리할 수 있다. 일괄 연구데이터 등록 기능을 고도화하여 기존 EXCEL 형식의 메타데이터 양식만 일괄 반입 할 수 있었던 부분에서, XML 형식의 계층화된 메타데이터 양식 또한 일괄 반입 할 수 있다. 또한, 파일의 명칭을 작성하면, 해당 파일 명칭을 내 저장소에서 탐색 후 매칭을 한다. 만약 매칭이 안 되어 있는 파일 명의 경우에는 웹 화면을 통해 업로드하면 메타데이터의 파일명과 매칭을 할 수 있다. 데이터 조회/검색을 다양한 다른 프로그램 및 서비스에서 연동할 수 있도록 OpenAPI를 제공한다. 권한 설정이 되어 있다면, OpenAPI를 통해 메타데이터뿐만 아니라 파일 데이터 또한 다른 서비스에서 가져갈 수 있다. 연구데이터 제출 시에는 해당하는 컬렉션을 지정할 수 있다. 하지만 컬렉션을 변경하고 싶은 경우, 이를 변경할 수 있도록 기능을 제공한다. 컬렉션 메타데이터를 연구자가 key/value 형태로 추가할 수 있다. 연구데이터 등록 과정에서 검토를 위해 가시화가 필요한 경우가 있다. 기존 썸네일 등을 제공하는 형식 외에 연구 데이터 또는 과학 데이터 형식에 대하여 데이터를 가시화할 수 있는 기능을 개발하였으며, 다양한 형식의 연구 데이터 형식에 대해서도 가시화 툴을 연계할 수 있도록 모듈화 하였다.

□ 연구 계획 관리는 연구데이터관리계획양식 생성/수정/삭제, 연구데이터관리계획서 제출/검토/승인, 연구데이터관리계획서 검색의 기능으로 구성된다.

○ 연구데이터관리계획양식은 아래의 그림에서 “DMP 관리” => “DMP 템플릿 관리”를 통해 확인할 수 있으며, “DMP 템플릿 생성” 메뉴를 클릭하면 DMP 템플릿을 기존에 있는 DMP 양식을 복사하여 생성할 수 있다. “DMP 템플릿 표시 순서” 메뉴에서는 DMP 양식이 많아질 경우에 순서를 정해서 사용자가 많이 활용하는 양식을 상단에 배치할 수 있다.

#	DMP 템플릿명	설명	발표여부	최종수정일	작업자
1	DMP_SAMPLE_TEMPLATE		완료	2021-01-12	
2	KISTI_DMP_TEMPLATE	KISTI에서 활용하는 DMP 양식을 바탕으로 데모용으로 작성된 DMP 양식	완료	2021-01-13	
3	삭제		미완료	2020-11-27	

그림 2. user interface of DMP schema management

DMP 템플릿 표시 순서
제출 화면에 보여지는 DMP 템플릿의 표시 순서를 변경합니다

DMP_SAMPLE_TEMPLATE	↑ 위로 ↓ 아래로	사용안함
KISTI_DMP_TEMPLATE	↑ 위로 ↓ 아래로	사용안함

저장 닫기

그림 3. Display order for DMP template/Schema



그림 4. Generating new DMP scheme/template

- 연구데이터관리계획양식의 편집은 연구데이터 양식의 편집화면과 동일하며, DMP 템플릿 형식을 출력하는 화면에서 좌측의 “수정” 버튼을 클릭하여 편집가능하다. “반출” 메뉴는 작성된 DMP 템플릿 양식을 EXCEL 형식으로 반출하는 기능을 제공한다.

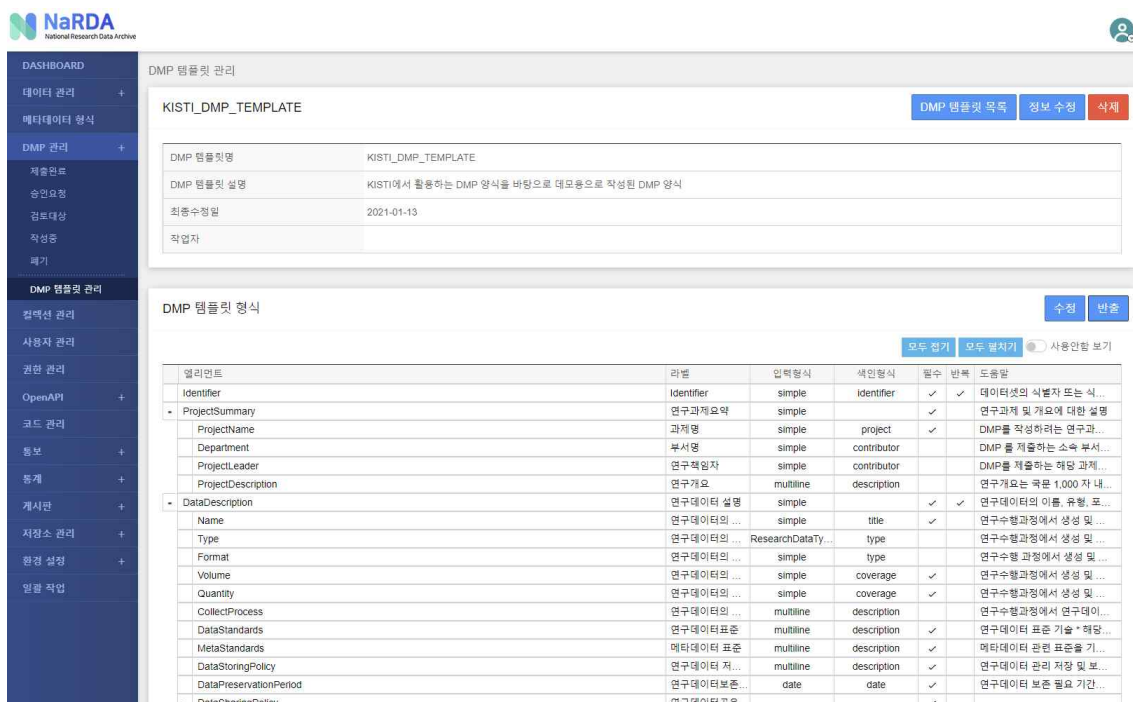


그림 5. display of elements description for the selected DMP template/scheme

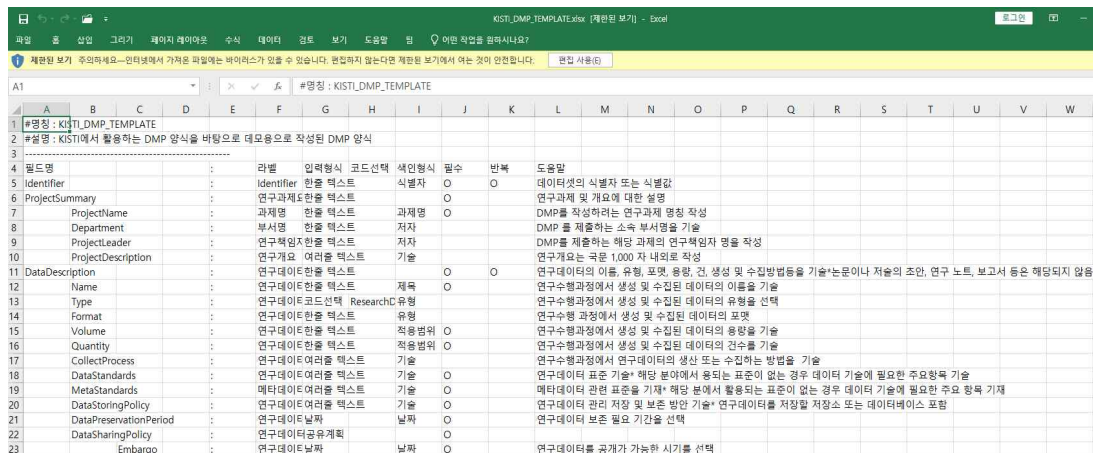


그림 6. Example of exporting a DMP scheme to an EXCEL file

- 연구데이터관리계획양식 편집화면에서는 “입력항목추가” 버튼을 통해 입력 양식의 항목(element)을 추가할 수 있으며, 항목에 마우스 오버를 하면, 해당항목의 “하위항목을 추가”, “위”, “아래”의 순서 변경, “사용안함”, “삭제”의 항목 삭제 기능을 활용할 수 있다. “모두접기”와 “모두 펼치기” 버튼은 하위항목이 존재하는 항목들에 대해서 하위항목의 표시할지 여부를 결정한다. “입력항목추가” 버튼 또는 특정 항목에서 마우스를 클릭하면, 입력항목 편집 화면이 팝업으로 표시된다.



그림 7. editing interface of DMP template/scheme

- 연구데이터관리계획양식 편집화면에서 “입력항목추가” 버튼을 누르면, 아래와 같은 화면이 표시된다. “필드명”은 시스템 내부적으로 활용하는 명칭이며, “라벨”은 연구데이터관리계획을 등록하는 연구자에게 표시되는 명칭이다. “입력 형식”은 유형에 따라 “한줄 텍스트”, “여러줄 텍스트”, “날짜”, “코드선택”으로 선택할 수 있다. “코드선택”을 선택한 경우에는 관리자 메뉴의 “코드선택” 메뉴에서 정의된 코드들 중에서 그 유형을 선택한다. “색인유형”의 경우에는 DC(Dublin Core)에 대응되는 표준 메타데이터 요소와 매칭하기 위한 요소이다. 이는 향후에 OpenAPI 또는 OMI-PMH 등의 표준 연계 프로토콜을 통해 수집 시스템에 데이터를 제공 시 필요한 부분이다. “필수/반복” 부분은 해당 항목(요소)을 필수요소로 선택할지와 그리고 반복 사용은 해당 요소를 반복해서 입력할지에 대해 선택하는 항목이다. “도움말”은 연구데이터관리계획을 등록하는 연구자에게 해당 항목을 어떻게 작성하는지 가이드 또는 지침을 기술하는 부분이다.

필드 정보 수정

필드명

Identifier

라벨

Identifier

▶ 사용자 입력 항목으로 표시되는 라벨입니다. 입력하지 않을 경우 엘리먼트명이 사용됩니다.

입력형식

☒ 한줄 텍스트
☐ 여러줄 텍스트
☐ 날짜
☐ 코드선택

코드선택

색인형식

☐ 없음
☐ 제목
☐ 부제목
☐ 과제명

☐ 기술
☐ 언어
☐ 저자
☐ 날짜

☒ 식별자
☐ 형식
☐ 출판사
☐ 관계

☐ 저작권
☐ 출처
☐ 주제
☐ 유형

☐ 적용범위
☐ 좌표(WKT)

필수/반복

☒ 필수 입력
☐ 반복 사용

도움말

데이터셋의 식별자 또는 식별값

저장

닫기

그림 8. Editing interface of an element

필드 정보 수정

필드명

Element

라벨

▶ 사용자 입력 항목으로 표시되는 라벨입니다. 입력하지 않을 경우 엘리먼트명이 사용됩니다.

입력형식

☐ 한줄 텍스트
☐ 여러줄 텍스트
☐ 날짜
☒ 코드선택

코드선택

자료유형

국가코드

IdentifierType

PersonnelRoleEnum

BooleanTypeEnum

country-3

OrganizationTypeEnum

RepositoryType

CC 라이선스

GranuleSpatialRepresentationEnum

lang

ResponsibilityType

ContributorType

CoordinateSystemEnum

CollectionAccessType

OrganizationPersonnelRoleEnum

ISO TopicCategoryEnum

SpatialCoverageTypeEnum

DurationUnitEnum

PersistentIdentifierType

PhoneNumberType

색인형식

필수/반복

도움말

그림 9. Example of Code(Controlled Vocabulary) selection

- 연구데이터관리계획양식이 작성되면, 연구자는 로그인하여 리포지터리의 “DMP” => “작성” 메뉴를 통해 연구데이터관리계획(DMP)를 제출할 수 있다. 또는 첫 화면 상단에 보이는 “DMP 제출”을 통해서도 가능하다.

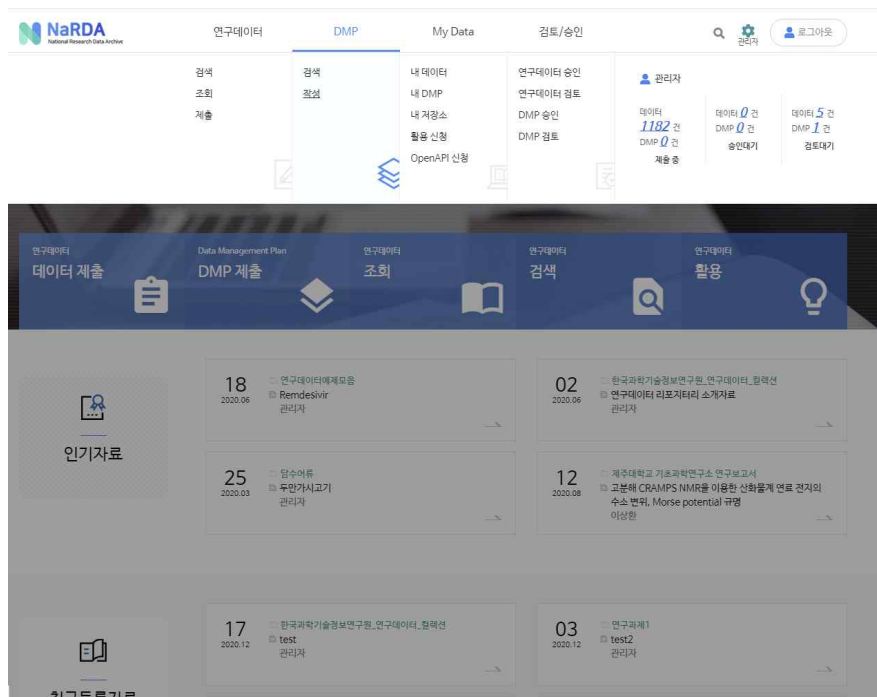


그림 10. DMP submission menu

- “DMP 제출” 또는 “작성” 메뉴를 통해 연구데이터관리계획서 작성 기능을 사용하면, 기존 제출된 DMP 계획서 목록과 “DMP 신규” 라는 메뉴가 있다. 해당 메뉴를 클릭하여 DMP 계획서를 작성할 수 있다.



그림 11. View of DMP application list and new submission of DMP application

- “DMP 신규” 메뉴를 선택하면, DMP 템플릿(양식)을 선택하는 화면으로 전환되고, “미리보기” 버튼을 통해, 양식 제출에 어떠한 항목이 있는지를 조회할 수 있다. 해당 양식을 확인한 후, “DMP 제출 시작” 버튼을 통해 연구데이터관리계획서 작성을 진행할 수 있다.

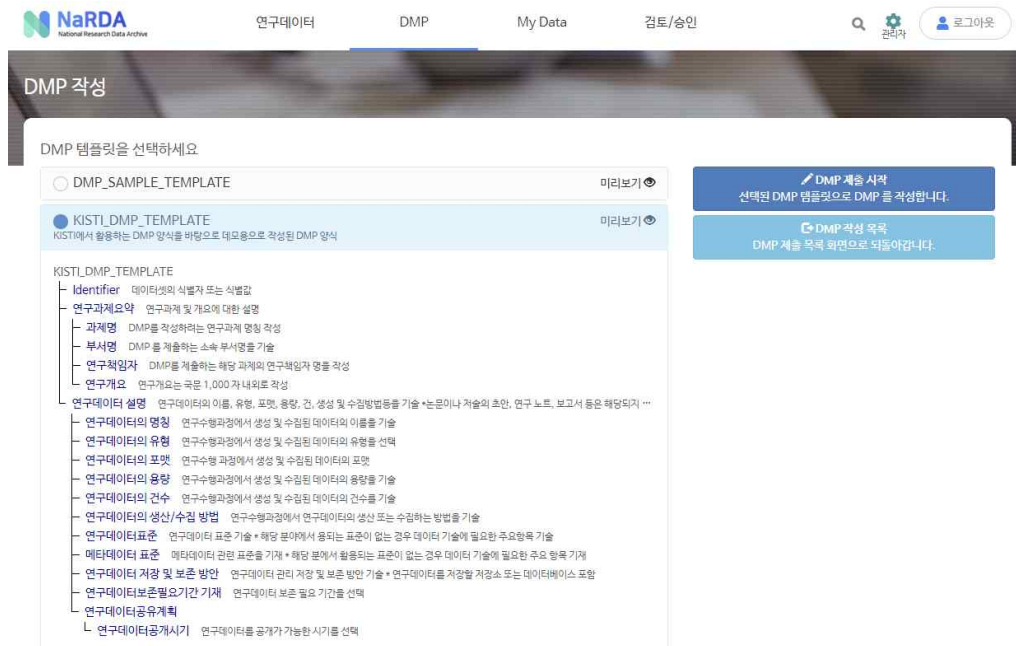


그림 12. Display of DMP template and DMP submission menu

- DMP 제출 양식에 따라 DMP 입력 화면이 표시되고, 해당 입력 화면에는 입력 형태에 맞게 DMP 자료를 입력한다. 또한 해당 DMP의 공개여부는 “원내공개” 또는 “원외공개”를 선택할 수 있다. “원내공개” 선택 시에는 사이트에 로그인하는 사용자만 확인할 수 있으며,

“원외공개” 선택 시에는 모든 이용자들이 작성된 DMP를 조회할 수 있다. “임시저장”은 작성된 자료를 저장하며, “다음단계”에서는 작성된 DMP를 제출하는 연구자가 확인 및 검토할 수 있다.

DMP 작성

DMP 템플릿: KISTI_DMP_TEMPLATE KISTI에서 활용하는 DMP 양식일 바깥으로 배포용으로 작성된 DMP 양식

메타데이터

Identifier	Identifier	+ -
연구과제요약	연구과제요약	
과제명	과제명	
부서명	부서명	
연구책임자	연구책임자	
연구개요	연구개요	
연구데이터 설명	연구데이터 설명	+ -
연구데이터의 명칭	연구데이터의 명칭	
연구데이터의 유형	연구데이터의 유형	
연구데이터의 포맷	연구데이터의 포맷	
연구데이터의 용량	연구데이터의 용량	
연구데이터의 건수	연구데이터의 건수	
연구데이터의 생산/수집 방법	연구데이터의 생산/수집 방법	
연구데이터 표준	연구데이터 표준	
메타데이터 표준	메타데이터 표준	
연구데이터 저장 및 보존 방안	연구데이터 저장 및 보존 방안	
연구데이터 보존필요기간 기재	연구데이터 보존필요기간 기재	
연구데이터 공유계획	연구데이터 공유계획	
연구데이터 공개시기	연구데이터 공개시기	

설명 및 공개 구분

설명:

공개 구분: ☐ 원외공개 ☐ 원내공개

임시 저장 다음 단계

그림 13. The view of DMP submission form

- 작성된 DMP는 검토 사용자에게 제출되며, 검토사용자는 연구자가 제출한 DMP를 검토하여 검토 승인/반려를 선택할 수 있다. 승인 절차 또한 연구데이터 검토 절차와 동일하게 진행되며, 승인은 관리자 설정에 따라 최대 3차 승인까지 지정할 수 있다.

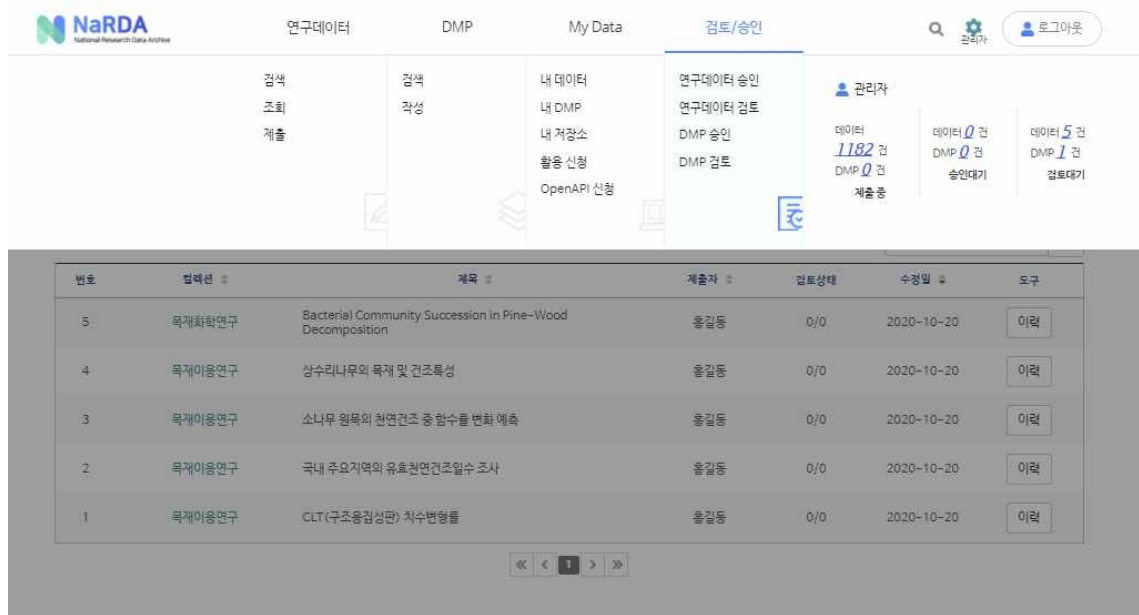


그림 14. Review/Approval menu for DMP application



그림 15. Review list of DMP application

NaRDA
National Research Data Archive
연구데이터
DMP
My Data
검토/승인
로그아웃

DMP 검토

DMP 기본정보

DMP 템플릿

DMP_SAMPLE_TEMPLATE

제목

테스트 명칭

과제

메타데이터

Identifier

33333

연구과제요약

과제명

테스트 과제

부서명

한국과학기술원

연구책임자

송용건

연구개요

개요 테스트

연구데이터 설명

연구데이터의 명칭

테스트 명칭

연구데이터의 포맷

테스트 포맷

연구데이터의 용량

100MB

연구데이터의 건수

30000

연구데이터의 생산/수집 방법

센서를 통한 데이터 수집

연구데이터 표준

데이터 표준 없음

메타데이터 표준

메타 표준 없음

연구데이터 저장 및 보존 방안

미보존

연구데이터보존필요기간 기제

2030-03-01

연구데이터공유계획

연구데이터공개시기

2025-05-01

DMP 검토

테스트 명칭

검토 결과

승인

반려

닫기

설명 및 공개 구분

설명

비공개

공개 구분

원래공개

목록

이력

검토

그림 16. Approved or Companion selection menu for the DMP application

- 작성된 연구데이터관리계획서가 최종 검토 및 승인이 완료되면 공개여부에 따라 원내 또는 원외로 공개되며, 아래 그림과 같이 게시된 내용을 확인할 수 있다. 아래 그림의 예제는 2020년도 동과제의 사업계획서 작성 시 첨부되었던 DMP 계획서를 양식에 맞게 작성한 것이다.

연구데이터 공유·확산체제 구축

메타 데이터 데이터 닫기

Identifier KISTI-DMP-01

연구과제요약

과제명 연구데이터 공유·확산체제 구축

부서명 연구데이터공유센터

연구책임자 송사광

연구개요 본 연구과제는 스마트 시티 구현을 위한 관련 연구데이터를 수집, 관리하는 시스템을 개발함. 또한, 다양한 스마트 시티 데이터를 통합 관리 및 검색시스템을 개발하며, 해당 데이터를 분석 및 모니터링하는 기술도 개발하여 스마트 시티 관련 데이터를 활용하는 생태계를 구현하는 것이 목표임.

연구데이터 설명

연구데이터의 명칭 대기질 측정 데이터

연구데이터의 유형 수치데이터

연구데이터의 포맷 CSV,JSON

연구데이터의 용량 50MB

연구데이터의 건수 100

연구데이터의 생산/수집 방법 대기질 측정 센서 장비로부터 이동 차량 2대에 탑재하여 하루 8시간 측정하여 생성하고 무선통신을 사용하여 서버에 저장

연구데이터표준 CSV,JSON

메타데이터 표준 자체표준

연구데이터 저장 및 보존 방안 - 본 연구과제를 통해 생산된 데이터는 기관 리포지터리에 등록하여 장기 보관하여 관리
- 기관 내부 시스템(IoT 리포지터리)에 연구데이터 저장소

연구데이터보존필요기간 기재 2025-01-01

연구데이터공유계획

연구데이터공개시기 2021-03-01

연구데이터 설명

연구데이터의 명칭 도로영상 데이터

연구데이터의 유형 동영상파일

연구데이터의 포맷 MPEG-4, LIDAR

연구데이터의 용량 10TB

연구데이터의 건수 1

연구데이터의 생산/수집 방법 GoPro 센서를 활용하고 이동 차량 2대에 탑재하여 하루 8시간 측정하여 생성하고 내부 저장장치에 저장

연구데이터표준 MPEG-4, LIDAR

메타데이터 표준 자체 메타데이터 사용

연구데이터 저장 및 보존 방안 - 본 연구과제를 통해 생산된 데이터는 기관 리포지터리에 등록하여 장기 보관하여 관리
- 기관 내부 시스템(IoT 리포지터리)에 연구데이터 저장소

연구데이터보존필요기간 기재 2025-01-01

연구데이터공유계획

연구데이터공개시기 2021-03-01

이전 화면

제출자 관리자

최종 수정 2021-01-12

그림 17. Example of a DMP document in the proposal of this project

- 연구데이터 제출 시, 작성된 DMP 목록을 검색하여 메타데이터 항목으로 추가할 수 있다. 해당 DMP 항목은 연구데이터 메타데이터 양식에 대한 항목(요소) 유형을 선택할 때, DMP를 선택하면 리포지터리에 제출된 DMP 항목을 메타데이터로 추가할 수 있다.

파일 데이터 ● 파일 업로드 제한

파일 데이터 #1

메타데이터 제출 연구노트 파일 추가 내지참소 파일 추가 사용자 파일 추가 +

데이터 파일

파일 설명

자료 유형

파일 공개 여부

파일 공개

파일 공개 공개 파일이 이용자에게 표시되며 다운로드 가능합니다.

파일 열 공개 파일이 이용자에게 표시되지만 다운로드 제공하지 않습니다. 파일을 요청/승인 후에 다운로드 할 수 있습니다.

열리고 이 간행됩니다.

3월 6월 1년 2년 기간설정 (~이후 공개: 최대 2년) 연도-월-일 열리고 기간 후에는 파일공개로 공개권한

파일 비공개 파일이 이용자에게 표시되지 않으며 파일을 요청/승인 할 수 없습니다.

메타데이터 필수 항목 보기 설정 강하기

제목

Identifier

생산자

출판사

출판 (예정) 년도

생성일자

개요

주제

기여자

키워드

NTIS연구과제

국가선택

관련 DMP 정보

데이터위치정보

DMP 검색

연구데이터에 검색

연구데이터 총 1건의 DMP가 있습니다.

과제 제목 제출자

연구데이터 공유-확산체계 구축 관리자

선택 완료 선택 배경 취소

공개 및 라이선스 설정

공개 구분

원외공개 원내공개 **특수공개** 열리고 기간설정 (~이후 공개: 최대 2년) 연도-월-일

DOI 출판

예 아니요

라이선스

CC BY-NC 저작자표시-비영리

저작권

이 데이터의 저작권은 (연구자 기관/그룹/사용자)에게 있습니다.

자료 사용시 주의사항

그림 18. Related DMP document selection menu when submitting research data

메타데이터 형식

메타데이터 형식을 편집 중입니다.

본 정보는 새로운 서비스로 임시 메타데이터의 형식과 저장되어 기존 메타데이터 형식을 영향을 주지 않습니다. 또한 다음을 다시 로그인하여 메타데이터 형식 수정에 참여하면 편집 완료 버튼을 클릭하여 작업을 종료하시기 바랍니다.

메타데이터 형식 편집

필드명

Title

Identifier

Creator

Publisher

Publication/year

Date

Description

Subject

Contributor

Keyword

Funder

Nation

DMP

GeoLocation

필드명

Title

Identifier

Creator

Publisher

Publication/year

Date

Description

Subject

Contributor

Keyword

Funder

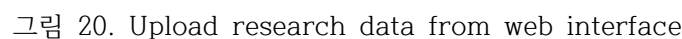
Nation

DMP

GeoLocation

그림 19. DMP metadata type selection menu in the editor of metadata scheme for research data

○ 연구데이터의 업로드/다운로드는 웹 인터페이스, WEBDAV, FTP를 활용할 수 있다. 웹을 통한 데이터의 업로드/다운로드는 웹에서 통신할 수 있는 크기의 데이터일 경우, 주로 활용될 수 있으며, WEBDAV는 http 와 동일한 방식으로 설정하여 사용할 수 있다. FTP의 경우 FTP를 활용하는 포트들이 오픈되어 있는 경우, FTP를 통한 데이터 업로드/다운로드를 지원받을 수 있다. 이처럼 다양한 프로토콜을 통해 연구자는 연구데이터를 선호하는 방식 또는 효율적인 방식으로 리포지터리의 “내 저장소”에 연구 수행 중인 데이터를 자유롭게 업로드/다운로드할 수 있다.



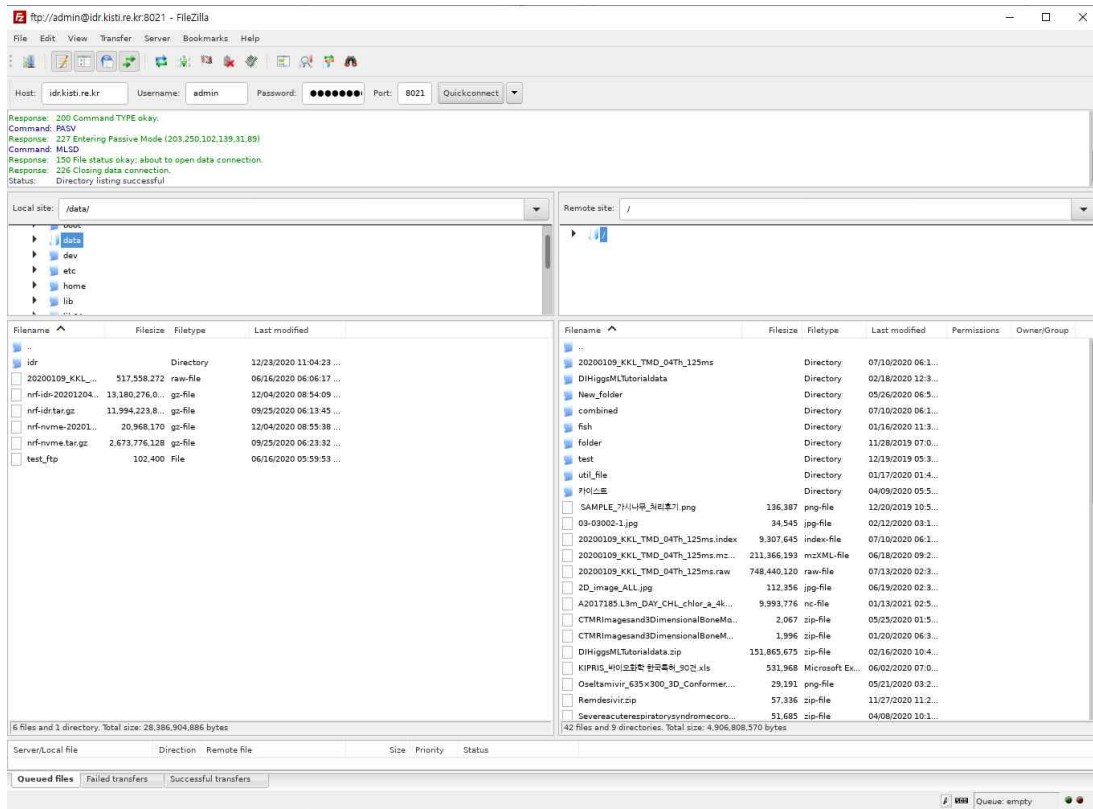


그림 23. client my storage folder connecting by FTP protocol

- 연구 수행 중에 업로드한 연구데이터에 대하여, 실험 절차, 설명 등 관련 메타데이터를 속성 명칭(key) 및 값(value) 형태로 작성할 수 있다. 아래 그림에서와 같이, “속성추가” 를 누르면 새로운 속성을 추가할 수 있고, “속성저장” 을 누르면 새로운 속성을 저장할 수 있다. 이와 같이 기술된 메타데이터는 연구데이터 등록 시 연구데이터의 특정 메타데이터와 연결할 수 있다. 해당 자료의 연결은 연구데이터를 업로드하고, “메타데이터 추출” 에서 “내 저장소 메타데이터” 를 선택하면 된다.

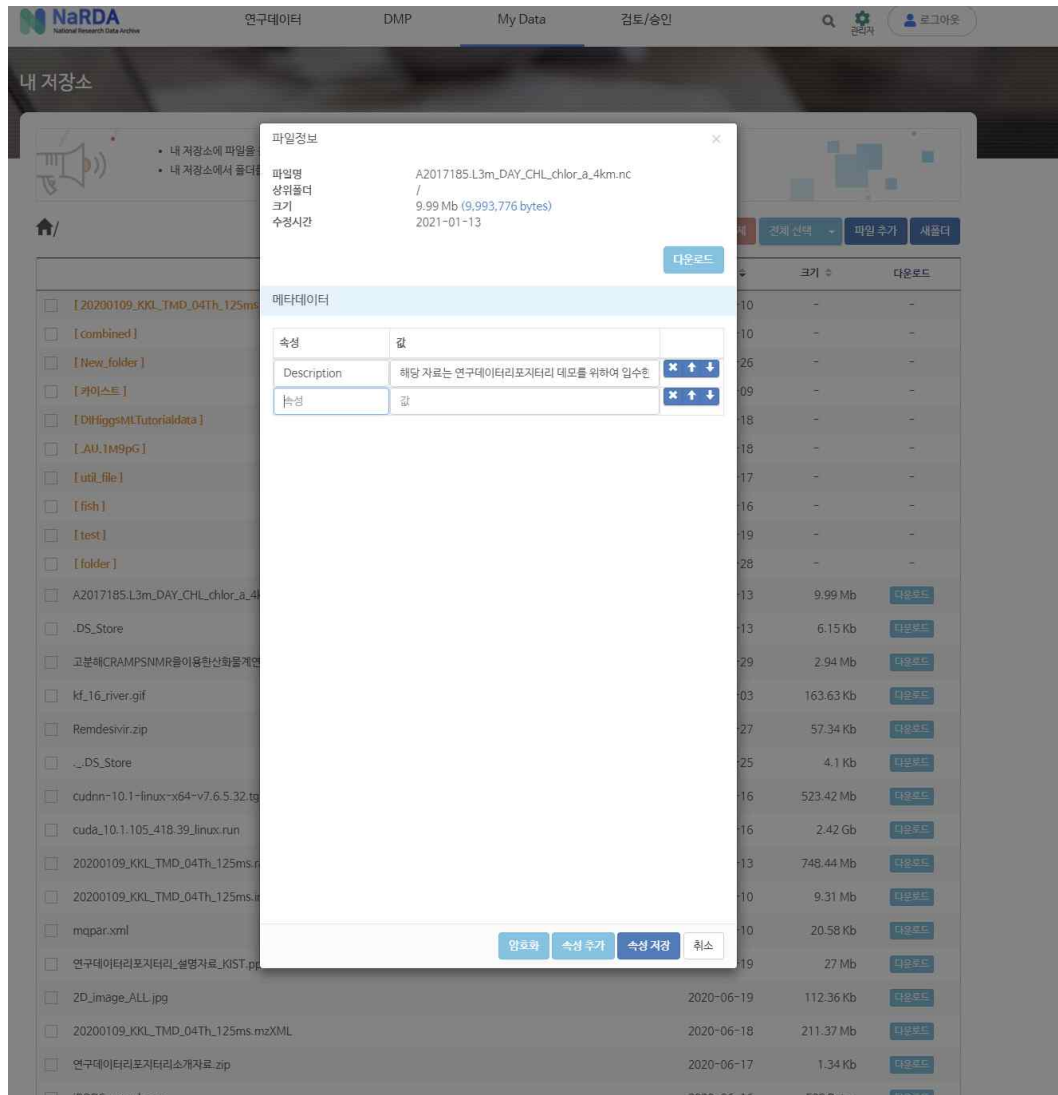


그림 24. Adding attributes and value to the files in my private storage space

- 연구기관에서 활용되고 있는 전자연구노트에 저장된 연구데이터를 연구데이터 등록 시에 업로드 할 수 있도록 시스템 간 연계를 통해 반입 기능을 개발하였다. 개발된 연구데이터 리포지터리(IDR)에서는 전자 연구 노트를 위해 본 연구원(KISTI)의 전자연구노트시스템과 연계를 수행하였다. KISTI 전자연구노트에서는 그림과 같이 REST 방식과 Java API 방식으로 데이터를 송수신하고, KISTI 내 SSO 인증을 활용하여 추가 인증 없이 IDR 과 전자연구노트 간 데이터 교환이 가능하도록 구현하였다.



그림 26. System process between research data repository and KISTI electronic laboratory notebook

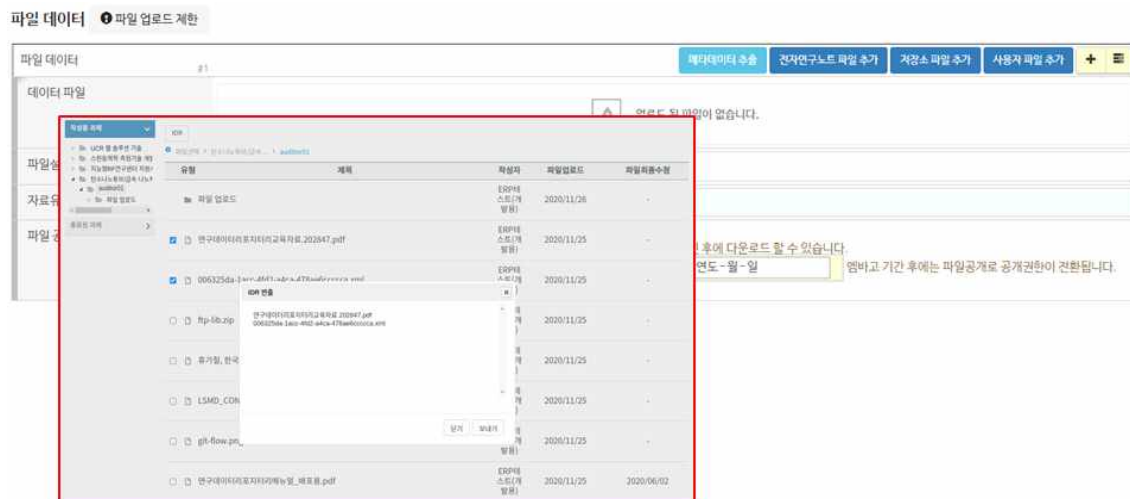


그림 27. Example of upload research data from KISTI electronic laboratory notebook

- 연구데이터 암호화는 연구데이터 등록 시까지 지적재산권 확보를 위한 연구활동(특허 또는 논문 작성 등)으로 인해 데이터를 보호해야 할 사항이 발생하였을 경우 이를 위해 지원하는 기능이다. 연구데이터 암호화는 파일 단위로 개인이 암호를 설정하여 암호화하여 보호할 수 있다. 마찬가지로 복호화 또한 암호를 입력하는 방식을 통해 복호화 할 수 있다. 여기서는 개인 암호화를 위해 비밀번호 방식을 이용하기 때문에 비밀번호를 유출 또는 분실에 주의가 필요하다.

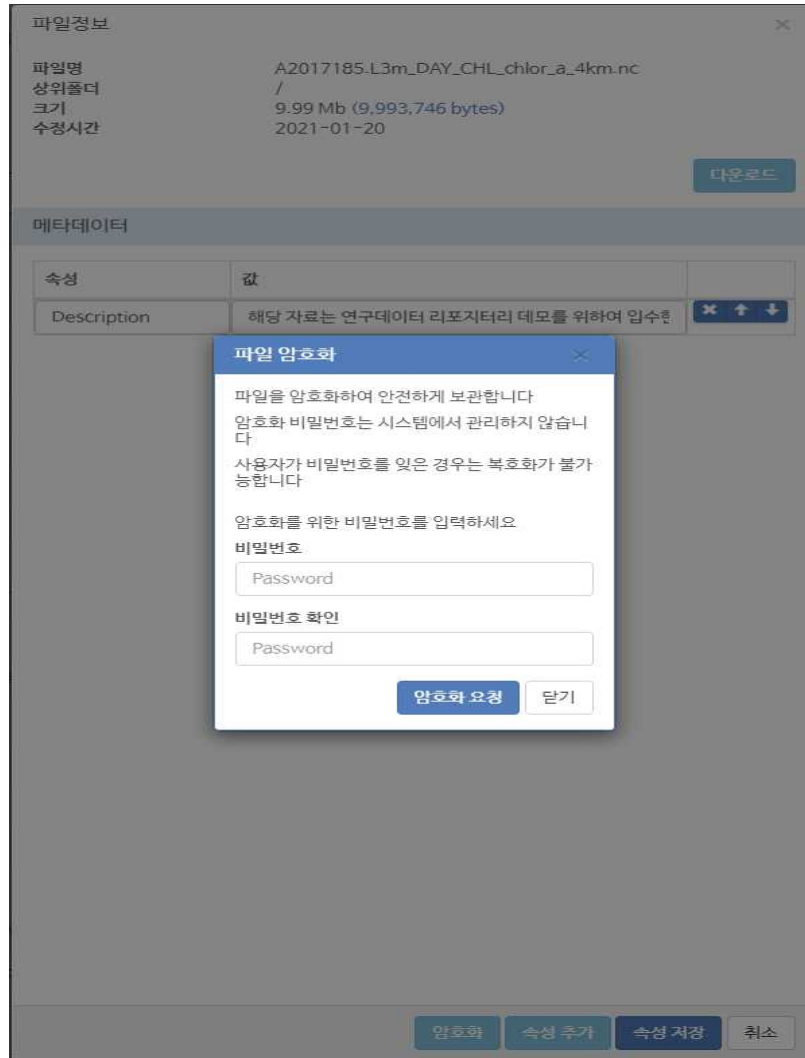


그림 28. Example of Encryption for research file data

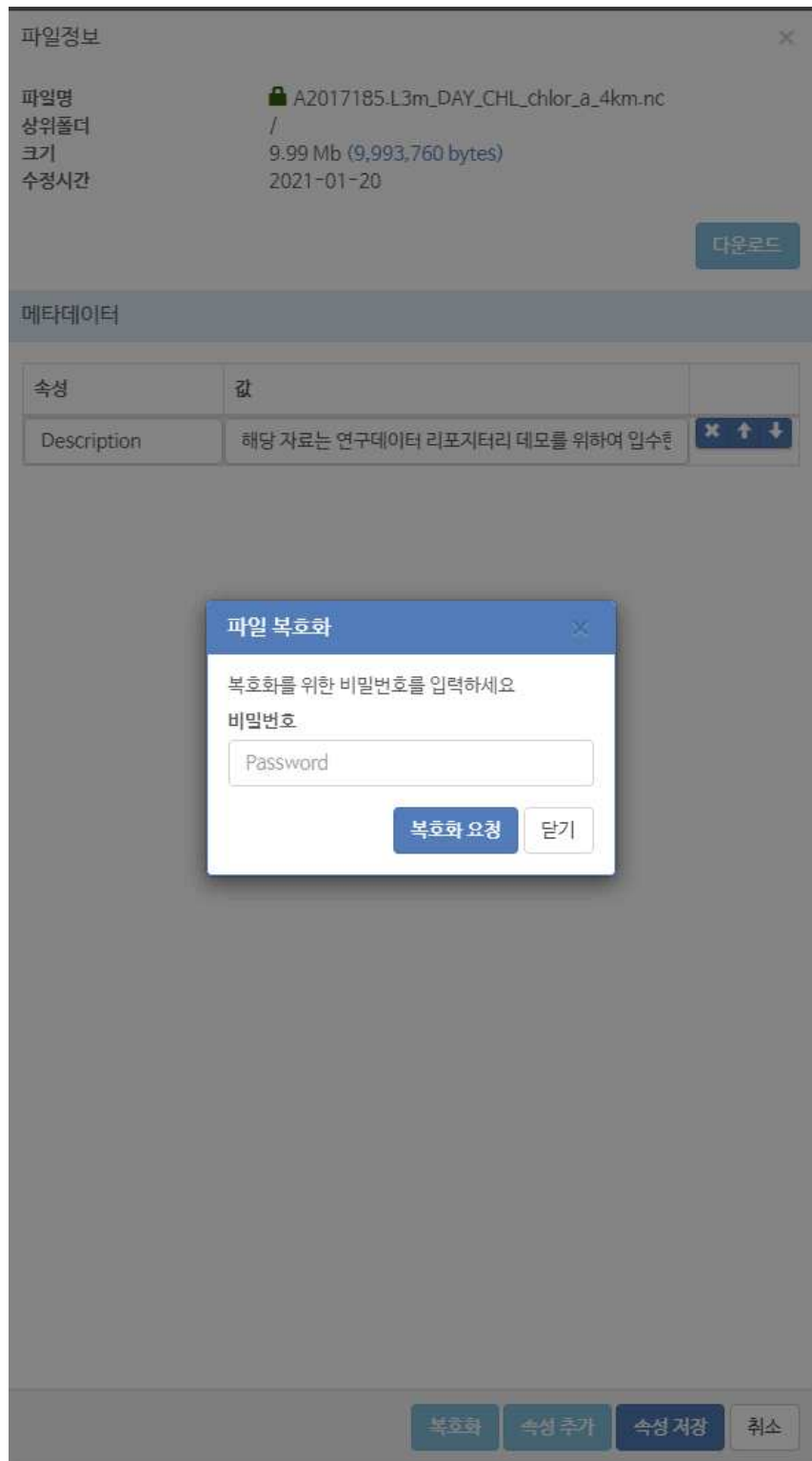


그림 29. Example of decryption for research file data

- 연구데이터리포지터리에서는 데이터 분석에 자주 활용되는 python에서 iRODS 기반 저장소를 제공하는 연구데이터리포지터리에 연결하여 데이터를 활용할 수 있도록 기능 테스트를 진행하였다. 이를 통해, 연구자들의 용이한 리포지터리의 데이터 활용을 제공할 수 있게 되었다.

- python-irodsclient: python 에서 내 저장소에 접근하기 위해서는 python-irodsclient 가 설치되어 있어야 한다.

표 1. instruction set for installing python-irodsclient

\$pip install python-irodsclient

- 테스트 데이터 예제: 내저장소에 아래의 그림과 같이 기상데이터 폴더를 생성하고 데이터 파일 20191129115500.csv 가 저장되어 있다.



그림 30. Example research data for testing of connecting data to python

- 내저장소 연결: python-irodsclient 를 사용하여 내저장소에 연결한다. 리포지터리 계정과 저장소를 생성할 때 입력한 비밀번호 또는 자동 생성된 비밀번호를 사용하여 내저장소에 연결한다. zone 은 기본 설치시에는 rdss 이다.

표 2. Example code for connecting my storage from python client

```
# iRODS 세션 생성
import os
from irods.session import iRODSSession
with iRODSSession(host='리포지터리 IP또는 호스트명',port=1247,user='bvztcfqxfpx@test.idr',password='JadQqFJh',zone='rdss') as session:
    pass
```

- 폴더 접근: 내저장소의 기본 경로는 /rdss/home/{리포지터리 계정} 이며 irods에서는 collection 으로 표현된다. 위에서 생성한 세션으로 collection 을 조회한다.

표 3. Example code for accessing my storage folder from python client

```
# 컬렉션(폴더) 접근
coll=session.collections.get("/rdss/home/bvztcfgxfepx@test.idr")
for c in coll.subcollections:
    print(c.name)
```

```
In [34]: # 컬렉션(폴더) 접근
coll = session.collections.get("/rdss/home/bvztcfgxfepx@test.idr")
for c in coll.subcollections:
    print(c.name)

테스트
기상데이터
```

그림 31. Accessing my storage folder from python client

- 데이터파일 읽기 예제: 내저장소에서 기상데이터 폴더 csv 파일 읽기 예제

표 4. Example code for lookup my storage files from python client

```
import pandas as pd # 데이터를 저장하고 처리하는 패키지
# 데이터 파일 불러오기
data_obj=session.data_objects.get("/rdss/home/bvztcfgxfepx@test.idr/기상데이터/20191129115500.csv")
f=data_obj.open("r")
df=pd.read_csv(f,encoding="euc-kr")
```

```
In [46]: import pandas as pd # 데이터를 저장하고 처리하는 패키지
# 데이터 파일 불러오기
data_obj = session.data_objects.get("/rdss/home/bvztcfgxfepx@test.idr/기상데이터/20191129115500.csv")
f = data_obj.open("r")
df = pd.read_csv(f, encoding="euc-kr")
df # 데이터 내용 표시
```

	지점	일시	기온 (°C)	기온 QC플래그	강수량 (mm)	강수량 QC플래그	풍속 (m/s)	풍속 QC플래그	풍향 (16방위)	풍향 QC플래그	습도 (%)	습도 QC플래그	중기압 (hPa)	이슬점 온도(°C)	현저기압 (hPa)	현저기압 QC플래그	해면기압 (hPa)	해면기압 QC플래그
0	133	2018-01-01 01:00	-2.5	0	NaN	NaN	1.5	0	320	0	67	0	3.4	-7.7	1018.1	0	1027.0	0
1	133	2018-01-01 02:00	-2.9	0	NaN	NaN	0.9	0	290	0	63	0	3.1	-8.9	1018.2	0	1027.1	0
2	133	2018-01-01 03:00	-3.0	0	NaN	NaN	0.7	0	270	0	64	0	3.1	-8.8	1018.6	0	1027.5	0
3	133	2018-01-01 04:00	-3.4	0	NaN	NaN	0.9	0	250	0	73	0	3.5	-7.5	1018.5	0	1027.4	0
4	133	2018-01-01 05:00	-4.0	0	NaN	NaN	1.0	0	270	0	78	0	3.6	-7.2	1018.6	0	1027.5	0
5	133	2018-01-01 06:00	-4.9	0	NaN	NaN	0.2	0	0	0	77	0	3.3	-8.3	1018.5	0	1027.5	0

그림 32. Lookup research data in my storage at the python client

- 읽은 데이터를 python 라이브러리를 사용하여 그래프를 표시

표 5. Example code for plotting research data file from python client

```
import matplotlib as mpl          # 그래프를 그리는 패키지
import matplotlib.pyplot as plt   # 그래프를 그리는 패키지

d=df[["일시","기온(°C)","습도(%)"]]# 필요한 데이터만

# Windows 한글 폰트 설정
#font_name =
mpl.font_manager.FontProperties(fname='C:/Windows/Fonts/malgun.ttf').get_name()
mpl.rc('font',family='NanumGothic')

ax=d.plot(kind='line',title='기온변화',figsize=(12,4),fontsize=12)
ax.set_xlabel('일시',fontsize=12)      # x축 정보 표시
ax.set_ylabel('기온/습도',fontsize=12)  # y축 정보 표시
ax.legend(['기온','습도'],fontsize=12)
```

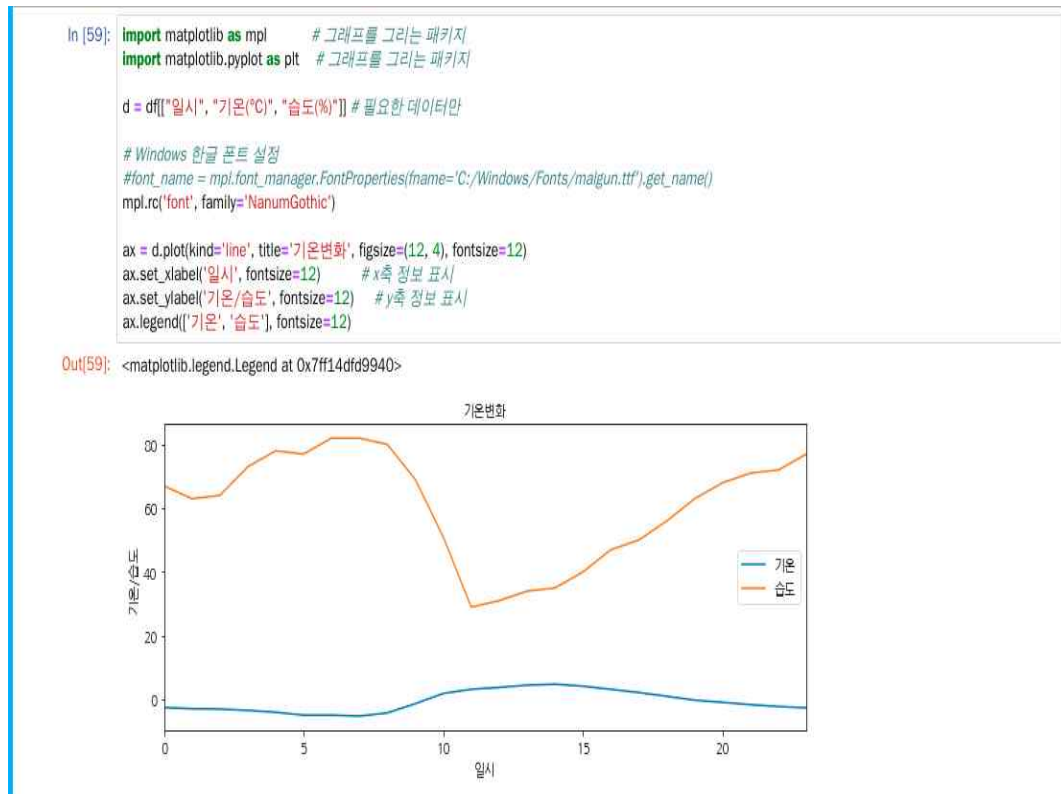
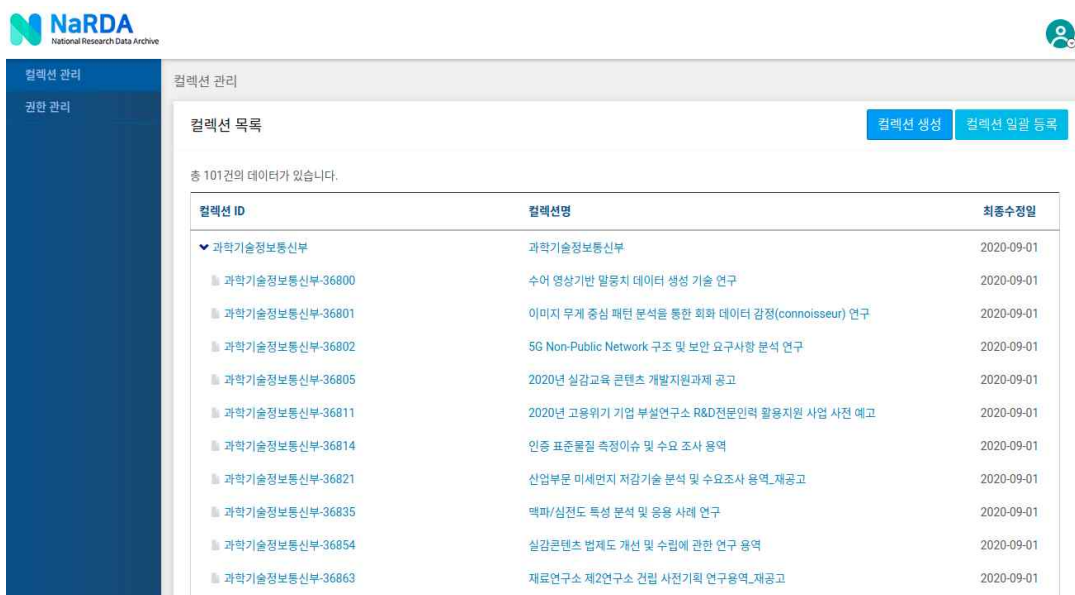


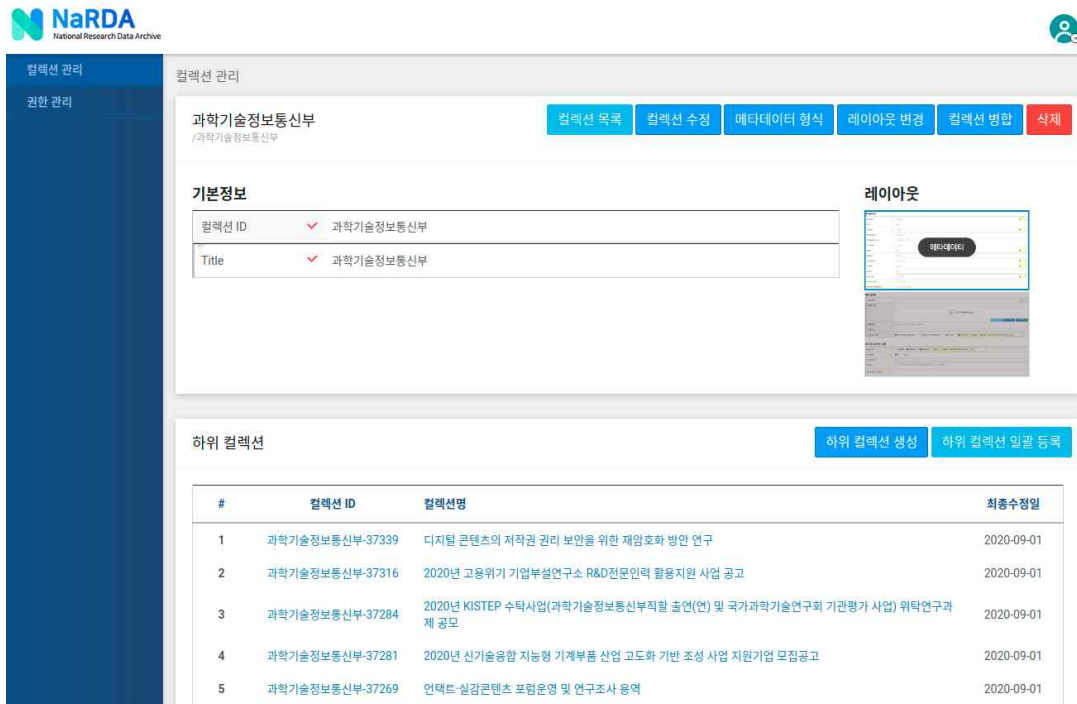
그림 33. plotting research data in my storage at the python client

- 연구를 수행한 후, 연구성과관리 및 연구데이터 공유·활용을 위해 연구데이터를 등록한다. 연구데이터 등록 관리 고도화를 위해 기존 개발한 연구데이터 리포지터리에서 다음의 사항을 개발 및 고도화 하였다. 첫째, 리포지터리 관리자 이외에도 컬렉션 관리자로 지정된 연구자 계정에서 컬렉션을 생성/편집할 수 있는 권한을 부여하였다. 둘째, 일괄 연구데이터 등록 기능을 계층적 메타데이터 구조 입력이 가능하도록 XML형태의 입력이 가능한 형태로 변경하였으며, “내 저장소”에 존재하는 파일 데이터와의 매칭을 통해 데이터 등록도 한꺼번에 이루어지도록 고도화하였다. 셋째, 메타데이터 및 연구데이터 공유를 위한 OpenAPI를 개발하였다. 넷째, 연구데이터와 컬렉션의 맵핑 정보를 변경 가능하도록 개발하였다. 다섯째, 컬렉션 메타데이터를 추가할 수 있도록 고도화하였다. 마지막으로 한국기초과학지원연구원과의 협업을 통해 mzXML 형태의 단백질 실험 데이터에 대한 가시화 기능을 개발하였다.
- 컬렉션 관리 권한 고도화를 위해, 컬렉션 관리를 하는 연구자 계정에서 컬렉션의 생성/수정/삭제 및 컬렉션을 이용하는 사용자에게 제출/검토/승인 권한을 부여할 수 있다. 또한 컬렉션 관리자는 해당 컬렉션의 하위 컬렉션까지 모두 관리 가능하다.



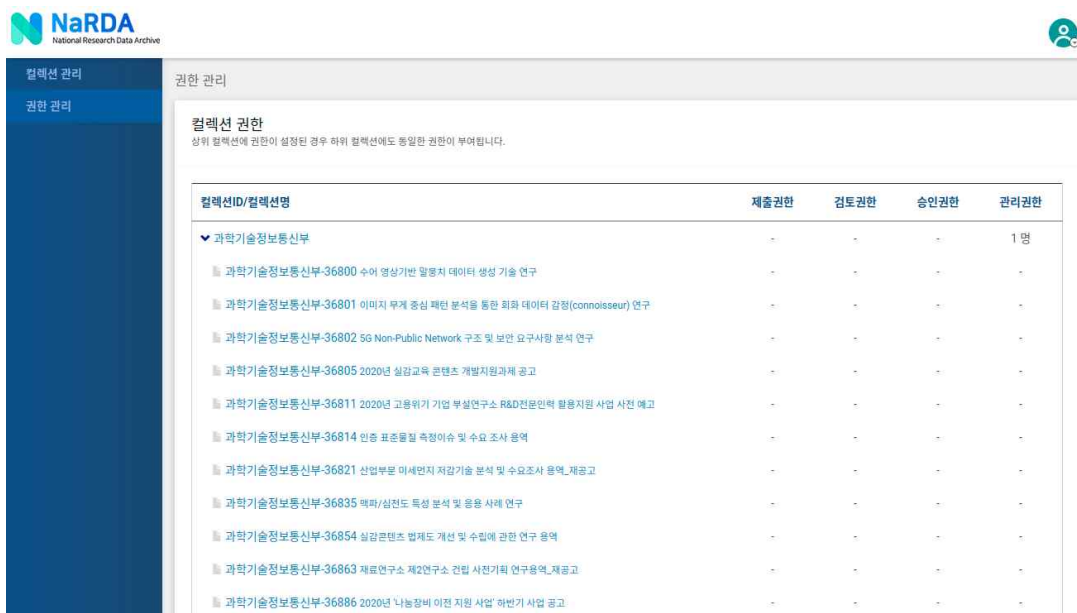
컬렉션 ID	컬렉션명	최종수정일
▼ 과학기술정보통신부	과학기술정보통신부	2020-09-01
■ 과학기술정보통신부-36800	수어 영상기반 발음지 데이터 생성 기술 연구	2020-09-01
■ 과학기술정보통신부-36801	이미지 무계 중심 패턴 분석을 통한 회화 데이터 감정(connoisseur) 연구	2020-09-01
■ 과학기술정보통신부-36802	5G Non-Public Network 구조 및 보안 요구사항 분석 연구	2020-09-01
■ 과학기술정보통신부-36805	2020년 실감교육 콘텐츠 개발지원과제 공고	2020-09-01
■ 과학기술정보통신부-36811	2020년 고용위기 기업 부설연구소 R&D전문인력 활용지원 사업 사전 예고	2020-09-01
■ 과학기술정보통신부-36814	인증 표준물질 측정이슈 및 수요 조사 용역	2020-09-01
■ 과학기술정보통신부-36821	산업부문 미세먼지 저감기술 분석 및 수요조사 용역_재공고	2020-09-01
■ 과학기술정보통신부-36835	맥파/심전도 특성 분석 및 응용 사례 연구	2020-09-01
■ 과학기술정보통신부-36854	실감콘텐츠 법제도 개선 및 수립에 관한 연구 용역	2020-09-01
■ 과학기술정보통신부-36863	재료연구소 제2연구소 건립 사전기획 연구용역_재공고	2020-09-01

그림 34. Display of Collection List logged in collection administrator



#	컬렉션 ID	컬렉션명	최종수정일
1	과학기술정보통신부-37339	디지털 콘텐츠의 저작권 권리 보안을 위한 재암호화 방안 연구	2020-09-01
2	과학기술정보통신부-37316	2020년 고용위기 기업부설연구소 R&D전문인력 활용지원 사업 공고	2020-09-01
3	과학기술정보통신부-37284	2020년 KISTEP 수탁사업(과학기술정보통신부직할 출연(연) 및 국가과학기술연구회 기관평가 사업) 위탁연구과제 공모	2020-09-01
4	과학기술정보통신부-37281	2020년 신기술융합 지능형 기계부품 산업 고도화 기반 조성 사업 지원기업 모집공고	2020-09-01
5	과학기술정보통신부-37269	언택트 실감콘텐츠 포럼운영 및 연구조사 용역	2020-09-01

그림 35. Display of Collection management menu



컬렉션ID/컬렉션명	제출권한	검토권한	승인권한	관리권한
▼ 과학기술정보통신부	-	-	-	1명
■ 과학기술정보통신부-36800 수어 영상기반 알맹치 데이터 생성 기술 연구	-	-	-	-
■ 과학기술정보통신부-36801 이미지 무게 중심 패턴 분석을 통한 회화 데이터 감청(connoisseur) 연구	-	-	-	-
■ 과학기술정보통신부-36802 5G Non-Public Network 구조 및 보안 요구사항 분석 연구	-	-	-	-
■ 과학기술정보통신부-36805 2020년 실감교육 콘텐츠 개발지원과제 공고	-	-	-	-
■ 과학기술정보통신부-36811 2020년 고용위기 기업 부설연구소 R&D전문인력 활용지원 사업 사전 예고	-	-	-	-
■ 과학기술정보통신부-36814 인종 표본질 측정이슈 및 수요 조사 용역	-	-	-	-
■ 과학기술정보통신부-36821 산업부분 미세연지 저감기술 분석 및 수요조사 용역_재공고	-	-	-	-
■ 과학기술정보통신부-36835 맥파/삼전도 특성 분석 및 응용 사례 연구	-	-	-	-
■ 과학기술정보통신부-36854 실감콘텐츠 법제도 개선 및 수립에 관한 연구 용역	-	-	-	-
■ 과학기술정보통신부-36863 재료연구소 제2연구소 건립 사전기획 연구용역_재공고	-	-	-	-
■ 과학기술정보통신부-36886 2020년 '나눔장비 이전 지원 사업' 하반기 사업 공고	-	-	-	-

그림 36. Granting privileges to collection users

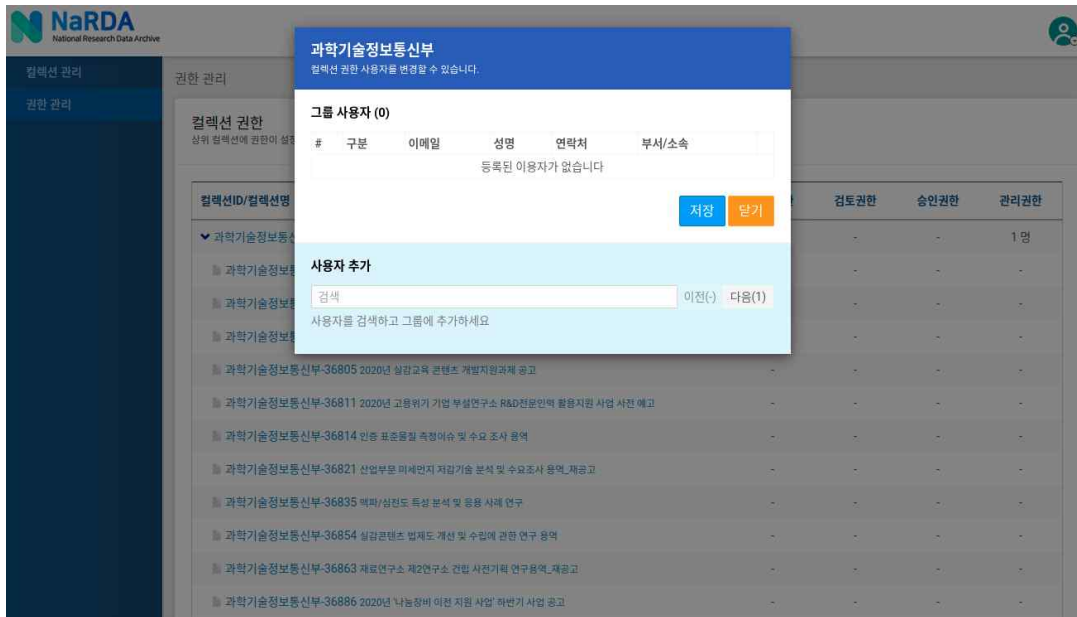


그림 37. Adding submission users to the specific collection

- 일괄 연구데이터 등록 기능은 계층 구조를 지니는 메타데이터의 일괄 입력을 위해 XML 형식의 메타데이터를 일괄 반입 가능하다. 또한, 파일명만 작성하면 해당 파일명을 지니는 데이터 파일을 “내 저장소”에서 자동으로 탐색하여 XML에 기록된 파일명과 자동으로 매칭해 준다. 일괄 등록 기능의 이용을 위해서는 아래 그림에서 “연구데이터” => “제출” 메뉴를 선택하거나 첫 화면의 상단에 나와 있는 “연구데이터 데이터 제출” 메뉴를 클릭한다.

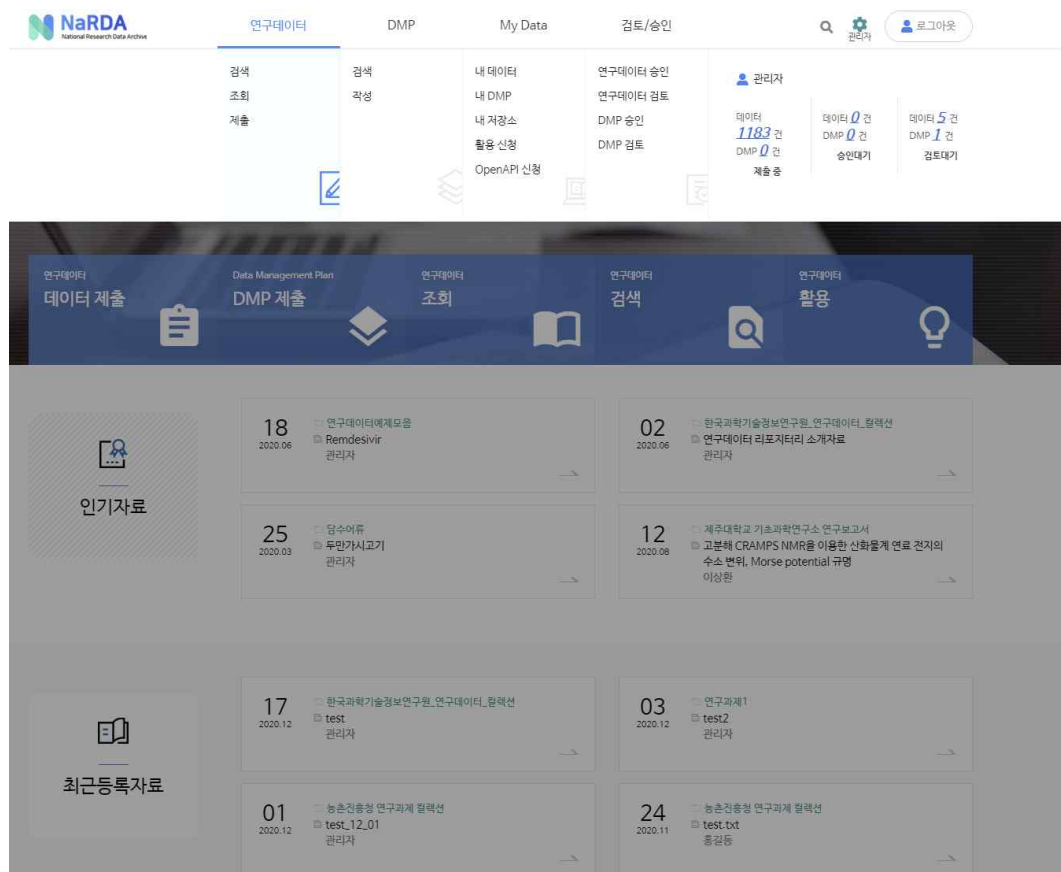


그림 38. Research data submission menu in data repository

- 연구데이터 제출 메뉴를 클릭하면, 아래 그림과 같이 작성 중인 연구데이터 목록이 표시되고, 신규 제출을 위해 목록 우상단에 있는 “제출” 버튼을 클릭한다.

NaRDA National Research Data Archive

연구데이터 DMP My Data 검토/승인

연구데이터 제출

작성중/반려/승인 대기 상태 목록을 조회할 수 있습니다.
데이터의 제출 작업을 계속하려면 제목을 클릭하세요.

작성중 데이터 : 총 1183건의 데이터가 있습니다.

Search for... [제출] [제출완료 데이터 보기]

번호	컬렉션	제목	작성자	상태	수정일	도구
1183	연구데이터_제출_시연_컬렉션	Untitled	관리자	임시저장	2021-01-19	삭제
1182	연구데이터_제출_시연_컬렉션	Untitled	관리자	임시저장	2021-01-19	삭제
1181	담수어류	Untitled	관리자	임시저장	2021-01-15	삭제
1180	연구데이터_제출_시연_컬렉션	Untitled	관리자	임시저장	2021-01-14	삭제
1179	곤충, 나비목 DB	nitens	관리자	임시저장	2020-01-17	삭제
1178	곤충, 나비목 DB	ophiogramma	관리자	임시저장	2020-01-17	삭제
1177	곤충, 나비목 DB	scolopacina	관리자	임시저장	2020-01-17	삭제
1176	곤충, 나비목 DB	rubrirena	관리자	임시저장	2020-01-17	삭제
1175	곤충, 나비목 DB	hampsoni	관리자	임시저장	2020-01-17	삭제
1174	곤충, 나비목 DB	brunnescens	관리자	임시저장	2020-01-17	삭제

« < 1 2 3 4 5 6 7 8 9 10 > »

그림 39. Display of Research data application List

- 연구데이터 제출을 위해서는 컬렉션을 선택해야 한다. 컬렉션이란, 연구데이터셋을 묶은 논리적인 단위로 연구 기관의 데이터셋 분류기준에 따라, 과학분야 또는 과제 단위, 또는 부서 단위로 분류할 수 있다.

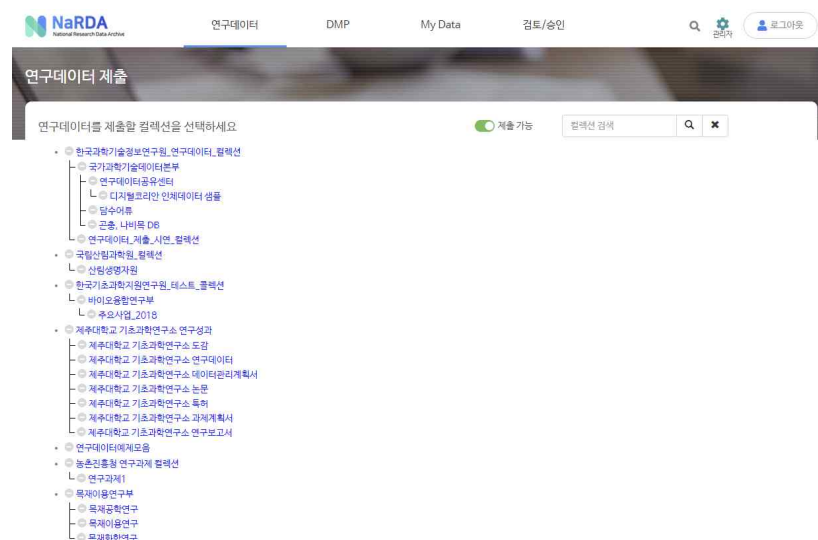


그림 40. Display of collection list to submit research data

- 연구데이터의 컬렉션 선택이 완료되면, 제출해야 되는 연구데이터의 메타데이터 형식(스키마)을 선택한다. 관리자 화면에서, 연구데이터 컬렉션과 연구자에게 보여주는 메타데이터 형식을

선택할 수 있다. 아래 그림과 같이 메타데이터 형식은 미리보기 버튼을 통해 해당 메타데이터의 항목을 확인할 수 있다. 일괄 데이터 제출을 위해서는 “일괄 제출” 버튼을 클릭한다.

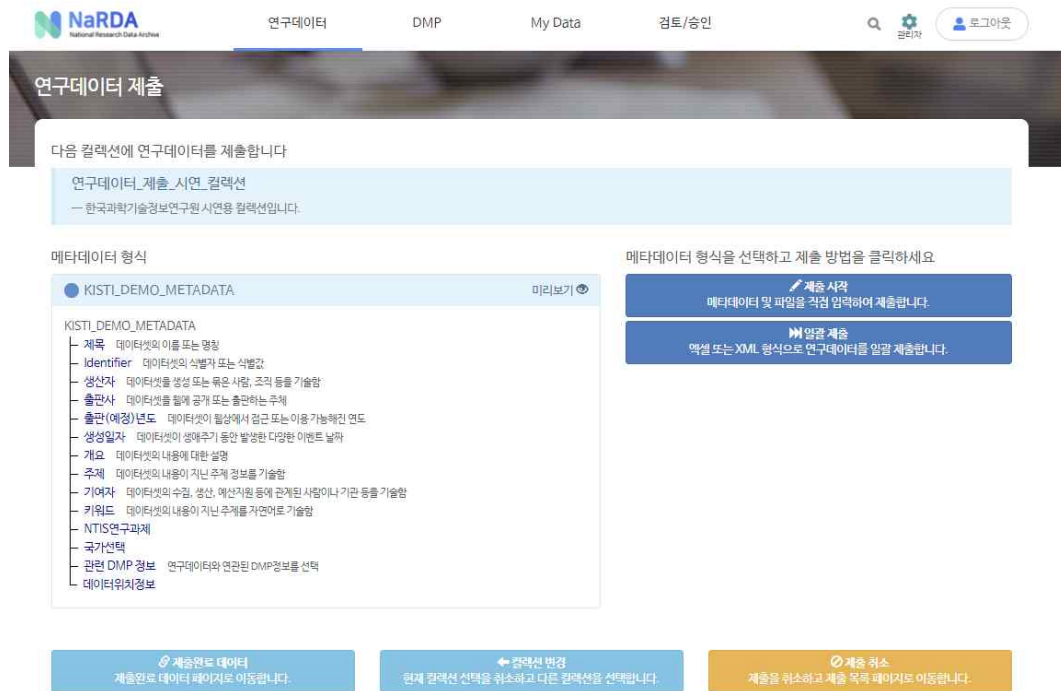


그림 41. Display of metadata form(scheme) to submit research data

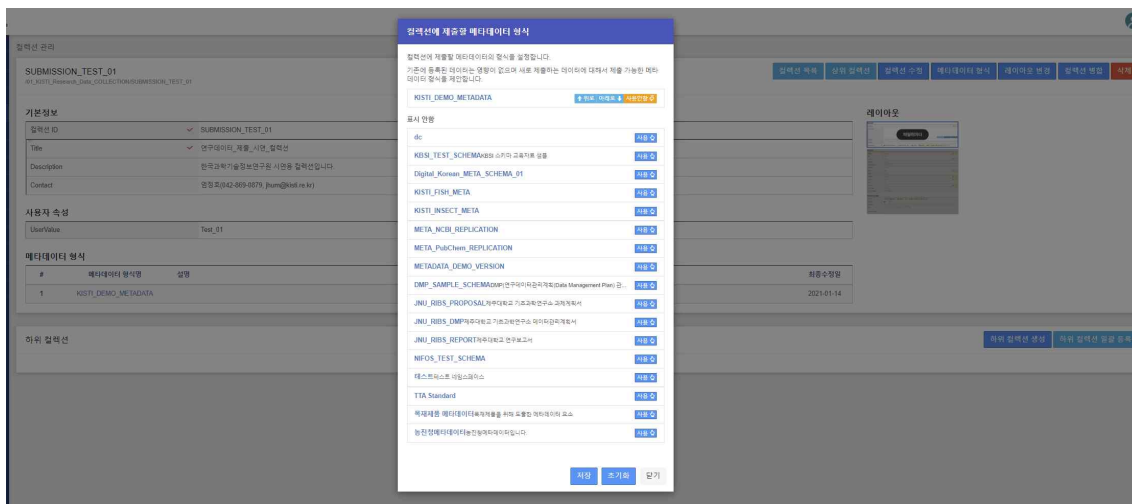


그림 42. Metadata scheme selection menu in collection management menu

- 연구데이터의 일괄 제출 메뉴는 아래 그림과 같다. 기존 EXCEL 뿐만 아니라 XML 형식의 메타데이터 일괄 반입 또한 가능하다. 먼저 해당 메타데이터 형식에 따른 XML 양식을 “XML” 버튼을 클릭하여 다운로드 받을 수 있다. XML 문서를 작성한 후, “데이터목록 업로드” 버튼을 클릭하여 해당 목록을 반입한다. 파일 데이터가 “내 저장소”에 없을 경우에는 “업로드 필요”라는 문구가 뜨고, 있을 경우에는 “내 저장소” 문구가 파일 옆에 붙는다. 해당 메타데이터 목록 및 파일 업로드 여부를 확인한 후, “반입 시작” 버튼을 클릭하여 연구데이터 제출 목록을 일괄 반입한다. 반입이 완료되면, 제출중인 연구데이터 목록에서 반입된 연구데이터를 확인할 수 있다. 연구데이터를 확인 및 더 필요한 수정을 완료하면, DMP와 마찬가지로 연구데이터 검토/승인 절차를 거쳐 공개여부에 따라 최종적으로 리포지터리에 게시된다.

일괄 제출

- 일괄 반입한 데이터는 작성중인 데이터로 등록됩니다. 개별 데이터를 확인하여 승인 요청을 실행해야 합니다.
- 한번에 많은 데이터를 올리는 것보다 100건 이하로 나누어 올리는 것이 좋습니다.
- 엑셀 파일로 제출하는 경우 엑셀 파일의 첫번째 시트에서만 데이터 목록을 읽으며 하위 필드를 사용할 수 없으며 반복 필드는 한번만 입력할 수 있습니다.

다음 컬렉션에 연구데이터를 제출합니다

연구데이터_제출_시연_컬렉션
— 한국과학기술정보연구원 시연용 컬렉션입니다.

메타데이터 형식
KISTI_DEMO_METADATA

일괄 반입할 데이터를 다음의 템플릿에 작성하십시오

엑셀 XML

작성한 엑셀 파일 또는 XML 파일을 업로드하십시오

데이터목록 업로드

연구데이터 항목에 연결된 파일데이터를 업로드 하십시오

파일데이터 업로드

업로드할 데이터가 올바른지 확인하고 반입을 시작하십시오

반입 시작

파일을 업로드하면 아래에 데이터가 표시됩니다.

제출완료 데이터
제출완료 데이터 페이지로 이동합니다.

← 컬렉션 변경
현재 컬렉션 선택을 취소하고 다른 컬렉션을 선택합니다.

제출 취소
제출을 취소하고 제출 목록 페이지로 이동합니다.

그림 43. batch research data submission menu

```

1 <?xml version="1.0" encoding="UTF-8"?>
2 <record>
3   <!-- 파일명, 데이터 간수명을 record 를 반복사용하십시오 -->
4   <!-- 필수/선택 속성을 확인하십시오 -->
5   <metadata>
6     <title 필드명="제목" 필수="예">
7       <!-- 데이터셋의 제목을 쓰는 명칭 -->
8       일괄 반입 테스트 데이터
9     </title>
10    <identifier 필드명="Identifier" 필수="예" 반복="예">
11      <!-- 데이터셋의 구분자, 또는 식별자 -->
12      KISTI-TEST-01
13    </identifier>
14    <creator 필드명="성상자" 필수="예" 반복="예">
15      <!-- 데이터셋을 생성 또는 복본 사람, 조직 등을 기술함 -->
16      윤정호
17    </creator>
18    <publisher 필드명="출판사">
19      <!-- 데이터셋을 통해 공개 또는 출판하는 주체 -->
20      KISTI
21    </publisher>
22    <publicationYear 필드명="출판(연장)년도">
23      <!-- 데이터셋이 출판해져 있는 혹은 이용 가능한 연도 -->
24      2020
25    </publicationYear>
26    <date 필드명="상장일자" 반복="예">
27      <!-- 데이터셋이 형제주기 동안 발생한 다양한 이벤트 날짜 -->
28      2020-12-31
29    </date>
30    <description 필드명="개요">
31      <!-- 데이터셋의 내용에 대한 설명 -->
32      <!-- 데이터셋의 내용에 대한 설명 -->
33    </description>
34    <subject 필드명="주제">
35      <!-- 데이터셋의 내용이 지닌 주제 정보를 기술함 -->
36      연구데이터
37    </subject>
38    <contributor 필드명="기여자" 반복="예">
39      <!-- 데이터셋의 수집, 생산, 예산지원 등에 관계된 사람이나 기관 등을 기술함 -->
40      최명석
41    </contributor>
42    <contributor 필드명="기여자" 반복="예">
43      <!-- 데이터셋의 수집, 생산, 예산지원 등에 관계된 사람이나 기관 등을 기술함 -->
44      이상환
45    </contributor>
46    <keyWord 필드명="키워드" 반복="예">
47      <!-- 데이터셋의 내용이 지닌 주제를 자연어 기술함 -->
48      연구데이터
49    </keyWord>
50    <funder 필드명="NTIS연구과제" 형식="ntis.project"><!-- NTIS연구과제 --></funder>
51    <nation 필드명="국가선택" 선택형속="country" 형식="select"><!-- 국가선택 --></nation>
52    <owp 필드명="연구과제 DWP 정보" 형식="dwp">
53      <!-- 연구데이터와 연관된 DWP정보를 선택 -->
54      </owp>
55    <location 필드명="데이터위치정보" 형식="ukt-geo"><!-- 데이터위치정보 --></location>
56    </metadata>
57    <bundle>
58      <!-- 파일 필드명="데이터 파일" 반복="예"><!-- 파일명을 입력합니다 --> 연구데이터리포지터리_설명자료_KIST.pptx </file>
59      <!-- description 필드명="파일설명"><!-- 파일에 대한 설명 --></description>
60    </bundle>
61  </record>
62 </record>

```

그림 44. Display of XML metadata scheme

NaRDA National Research Data Archive

연구데이터 DMP My Data 검토/승인

로그아웃

일괄 제출

다음 컬렉션에 연구데이터를 제출합니다

연구데이터_제출_시연 컬렉션
— 한국과학기술정보연구원 시연용 컬렉션입니다.

메타데이터 형식

KISTI_DEMO_METADATA

일괄 반입 데이터를 다음의 템플릿에 작성하십시오

엑셀 XML

작성한 엑셀 파일 또는 XML 파일을 업로드 하십시오

데이터목록 업로드

연구데이터 항목에 연결된 파일데이터를 업로드 하십시오

파일데이터 업로드

업로드할 데이터가 올바른지 확인하고 반입을 시작하십시오

반입 시작

KISTI_DEMO_METADATA.xml 총 1건 (2.00 Kb)

- 1 일괄 반입 테스트 데이터
 - 연구데이터리포지터리_설명자료_KIST.pptx **데이터파일**
 - 일괄 반입 테스트 데이터
 - Identifier KISTI-TEST-01
 - Creator 윤정호
 - Publisher KISTI
 - PublicationYear 2020
 - Date 2020-12-31
 - Subject 연구데이터
 - Contributor #1 최명석
 - Contributor #2 이상환
 - Keyword 연구데이터
 - File #1 연구데이터리포지터리_설명자료_KIST.pptx

파일데이터
연구데이터리포지터리_설명자료_KIST.pptx **데이터파일**

업로드된 모든 파일을 삭제하고 제출을 다시 시작합니다.

제출 다시 시작
컬렉션/메타데이터 형식을 선택하여 제출을 다시 시작합니다.

제출 취소
제출을 취소하고 계층 목록 페이지로 이동합니다.

그림 45. Display of metadata attribute list for batch submission data



연구데이터

DMP

My Data

검토/승인

관리자

로그아웃

연구데이터 제출

이력

관계 설정

검토 요청

수정

삭제

컬렉션 & 메타데이터 형식

컬렉션

연구데이터_제출_시연_컬렉션 한국과학기술정보연구원 시연용 컬렉션입니다.

메타데이터 형식

KISTI_DEMO_METADATA

파일 데이터

파일 데이터

#1

데이터파일

✓

연구데이터포지터리_설명자료_KIST.pptx

25.00 Mb

파일설명

자료유형

파일공개 여부

파일명 공개 (파일 요청/승인 후 다운로드)

메타데이터

제목

일괄 반입 테스트 데이터

Identifier

KISTI-TEST-01

생산자

엄정호

출판사

KISTI

출판(예정)년도

2020

생성일자

2020.12.31

주제

연구데이터

기여자

최명석

기여자

이상환

키워드

연구데이터

공개 및 라이선스 설정

공개 구분

✓

원래공개

DOI

✓

사용안함

라이선스

저작권

자료 사용시 주의사항

목록

이력

관계 설정

검토 요청

수정

삭제

그림 46. submission state of batch research data application

- 개발한 연구데이터 리포지터리에서는 OpenAPI를 통해 연구데이터의 메타데이터 및 데이터 파일을 공유할 수 있다. “My Data” => “OpenAPI 신청” 메뉴를 통해 OpenAPI 사용을 관리자에게 요청할 수 있다. OpenAPI 신청 시에는 원의 데이터에 한정하여, 활용할 도메인 또는 IP를 작성하면 해당하는 IP에 대해서는 OpenAPI를 개방한다.

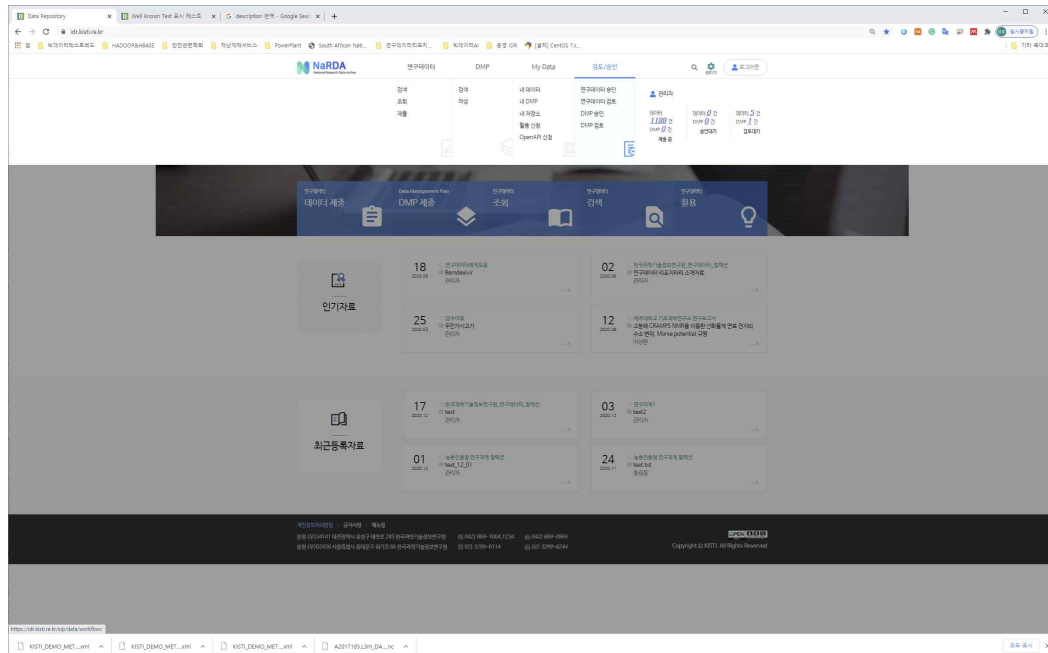


그림 47. OpenAPI request menu

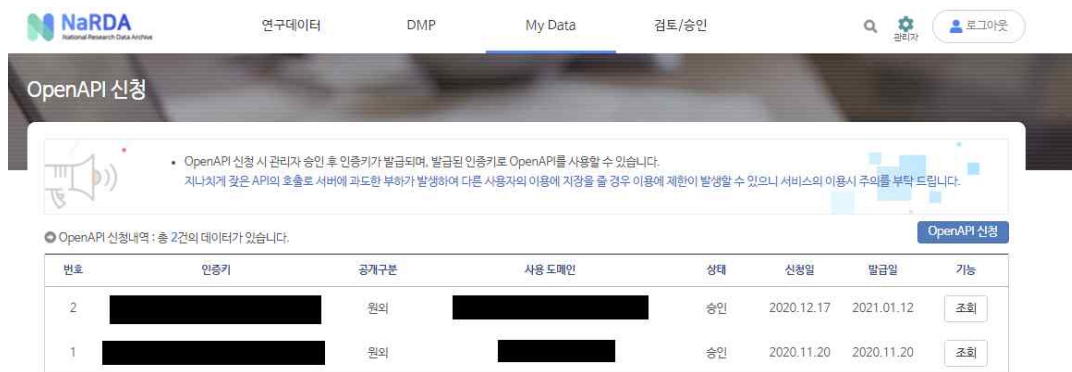


그림 48. Display of OpenAPI application list


```

localhost:19080/openapi/data/a416bdb1-87e0-4fc4-ad21-5a
{
  id: "a416bdb1-87e0-4fc4-ad21-5ab3b49a50ec",
  version: "1",
  title: "An analysis on satisfaction level of clinicians on implant surgical guidance system",
  submitter: "최남민",
  collection: {
    id: "aec8c966-fd2e-4dec-9dae-7d3db02e0c01",
    name: "geo",
    title: "디지털매핑에 의한 통합 지질정보 제공기술 개발",
    description: "2018년도 디지털매핑에 의한 통합 지질정보 제공기술 개발.",
    collectionPath: "/aec8c966-fd2e-4dec-9dae-7d3db02e0c01",
    lastModified: 1594367982246
  },
  lastModified: 1594367999002,
  accessFrom: 1594306800000,
  license: null,
  copyright: null,
  notes: null,
  doi: "10.22686/ldr/9",
  date_created: 1594367998600,
  date_submitted: 1594367998600,
  date_approved: 1594367998968,
  date_archived: 1594367998968,
  metadata: {
    저널정보: {
      저널명: "구강회복융합과학기술",
      기관명: "대한학관철교합학회",
      출판년도: "2015",
      출판월: "09",
      권: "31",
      호: "3"
    }
  }
}

```

그림 51. OpenAPI search result for the specific research data

```

localhost:19080/openapi/collection?key=s927ldnWSn9c
{
  - {
    id: "6a337a17-e629-43f8-b953-eb8c51d8f4c9",
    name: "drone3d",
    title: "드론 3D 매핑 기반 건설 현장관리 실용화기술 개발",
    description: "드론 3D 매핑 기반 건설 현장관리 실용화기술 개발.",
    collectionPath: "/6a337a17-e629-43f8-b953-eb8c51d8f4c9",
    lastModified: 1594367982422,
    metadata: { },
    attributes: { },
    subCollections: [ ]
  },
  - {
    id: "aec8c966-fd2e-4dec-9dae-7d3db02e0c01",
    name: "geo",
    title: "디지털매핑에 의한 통합 지질정보 제공기술 개발",
    description: "2018년도 디지털매핑에 의한 통합 지질정보 제공기술 개발.",
    collectionPath: "/aec8c966-fd2e-4dec-9dae-7d3db02e0c01",
    lastModified: 1594367982246,
    metadata: { },
    attributes: { },
    subCollections: [ ]
  },
  - {
    id: "3a9ac5da-aa47-42b1-a066-f6403fa551a5",
    name: "puriton",
  }
}

```

그림 52. OpenAPI search result for collection metadata

```
{
  id: "653e8494-ad57-44e8-8a01-5dc9ff1911d4",
  name: "경찰청",
  title: "경찰청",
  description: null,
  collectionPath: "/653e8494-ad57-44e8-8a01-5dc9ff1911d4",
  lastModified: 1598499858061,
  metadata: {
    Title: "경찰청",
    Subject: [
      "경찰",
      "Police"
    ],
    Contact: "112",
    Rights: "저작권",
    Keyword: [
      "형사",
      "수사",
      "수배"
    ]
  },
  attributes: {
    만듦이: "김진섭",
    그림: "김길남"
  },
  subCollections: [
    {
      id: "132986a-d925-424d-9cc8-acfe40e9c572",
      name: "경찰청-37210",
    }
  ]
}
```

그림 53. OpenAPI search result for the specific collection

- 연구데이터 제출 시 선택하였던 컬렉션을 변경할 경우가 발생한다는 연구기관의 요구사항을 반영하여, 연구데이터와 맵핑된 컬렉션을 변경하는 기능을 연구데이터 리포지터리에서 제공한다. 해당하는 기능은 관리자 메뉴에서 “제출 완료된 데이터”를 선택하여 “컬렉션 변경” 메뉴를 통해 맵핑된 컬렉션을 수정할 수 있다.

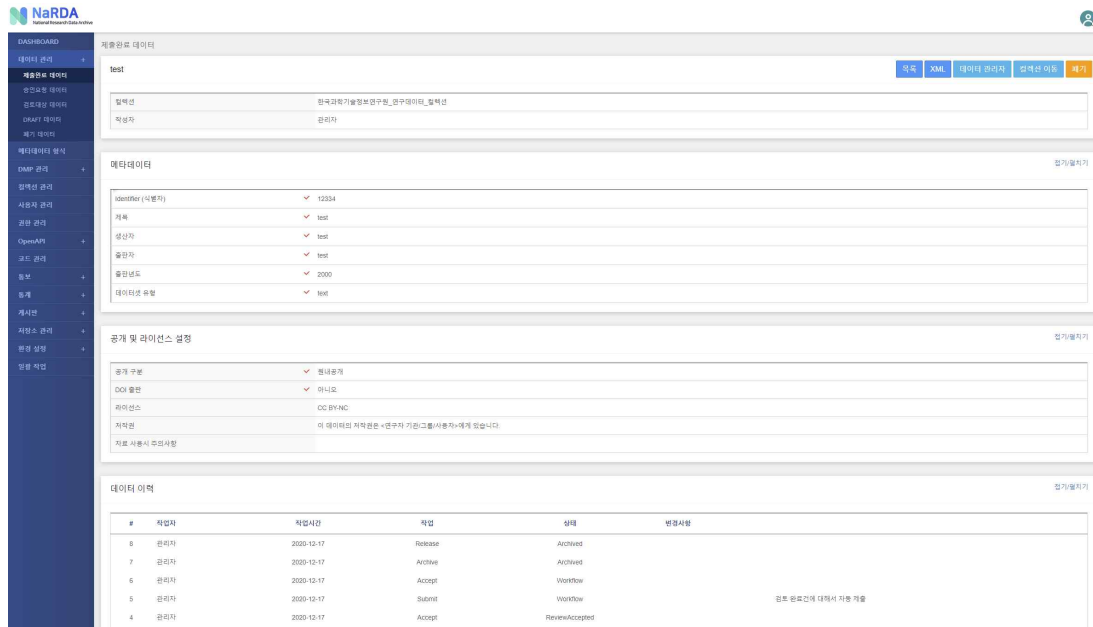


그림 54. Collection movement menu

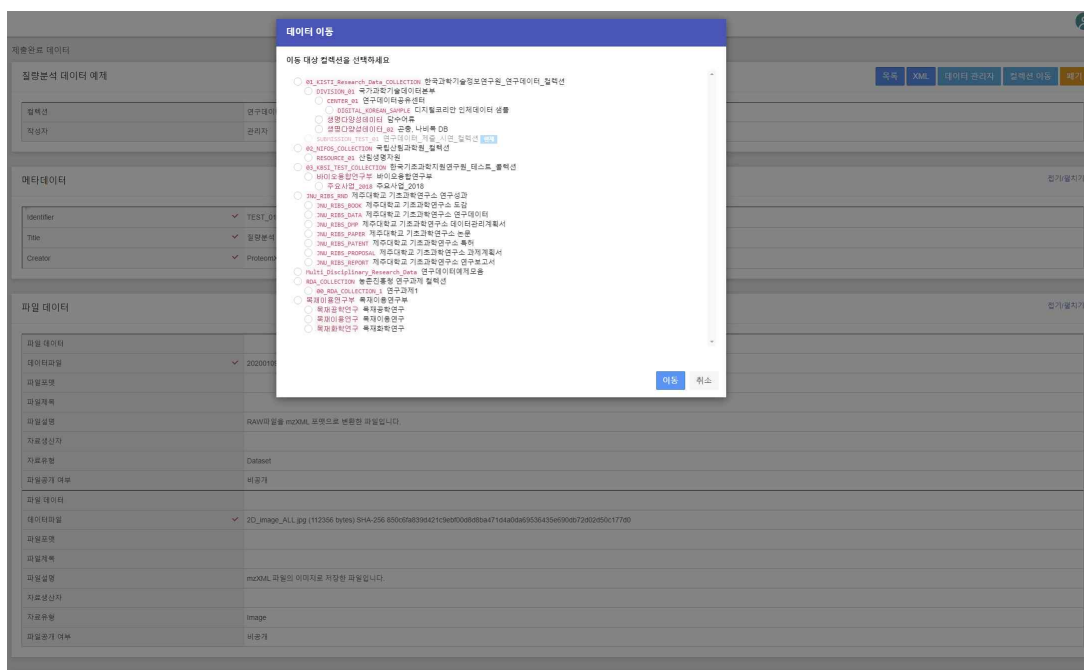


그림 55. Change collection of research data to the one of other collections

- 연구데이터 컬렉션에는 기본적으로 리포지터리에서 정의한 메타데이터를 사용한다. 정의된 메타데이터는 Title(제목), Description(설명), Subject(주제), Contact(담당자), Rights(저작권 정

보), Keyword(키워드) 이다. 그 이외의 메타데이터가 필요할 경우에는 아래의 그림과 같이 사용자 속성으로 정의할 수 있다.

The screenshot shows the NaRDA submission interface. On the left is a sidebar with navigation links. The main area is titled '입력한 컬렉션' (Submitted Collection) and shows details for 'SUBMISSION_TEST_01'. The '기본정보' (Basic Information) section includes fields for Collection ID, Title, Description, Subject, Contact, Rights, and Keyword. Below this is the '사용자 속성' (User Attributes) section, which contains a table with one row: 'UserValue' with the value 'Test_01'.

그림 56. Typing user defined attributes to a collection

The screenshot shows the NaRDA submission interface with additional tabs at the top: '입력한 컬렉션', '입력한 필드', '입력한 속성', '메타데이터 형식', '메타데이터 링크', '입력한 정보', and '입력'. The '기본정보' section is visible. To the right, there is a '제어번호' (Control Number) field. Below the '사용자 속성' section, there is a '메타데이터 형식' (Metadata Format) table with one row: 'HIST_DEMO_METADATA' with the value '2021-01-14'.

그림 57. Display of collection metadata with the user defined attribute

- 한국기초과학지원연구원의 바이오융합분석본부에 배포되는 연구데이터 리포지터리에서, 장비에서 생산되는 단백질 연구데이터에 대한 검토를 위해 가시화 기능에 대한 필요성이 제기되었으며, 한국기초과학지원연구원과의 협업을 통해 데이터 가시화 기능을 개발하였다. 한국기초과학지원연구원에서는 단백질 연구데이터에 대한 특성과 가시화에 필요한 요구사항을 정의하였으며, 본 과제에서는 이러한 요구사항을 수렴하고, 데이터에 대한 이해를 바탕으로 웹 브라우저 상에서 가시화하는 데이터 가시화 기능을 개발하였다. 해당 기능은 내부적으로 플러그인 형태로 개발되어 다른 종류의 데이터 가시화에도 적용 및 확장 할 수 있는 형태로 리포지터리의 게시 기능을 수정, 고도화 하였다.

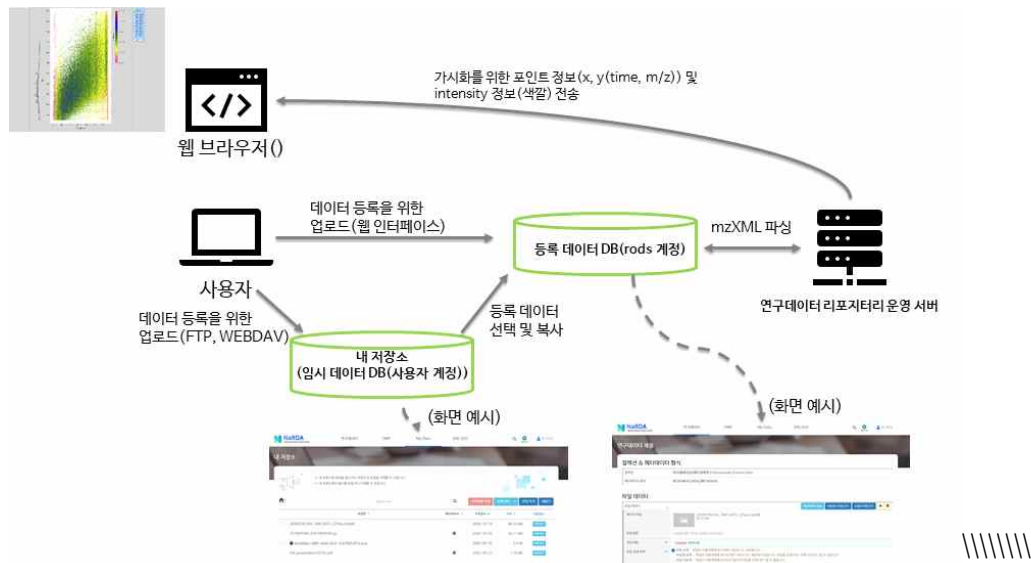


그림 58. System process for visualizing mzXML formatted data in data repository

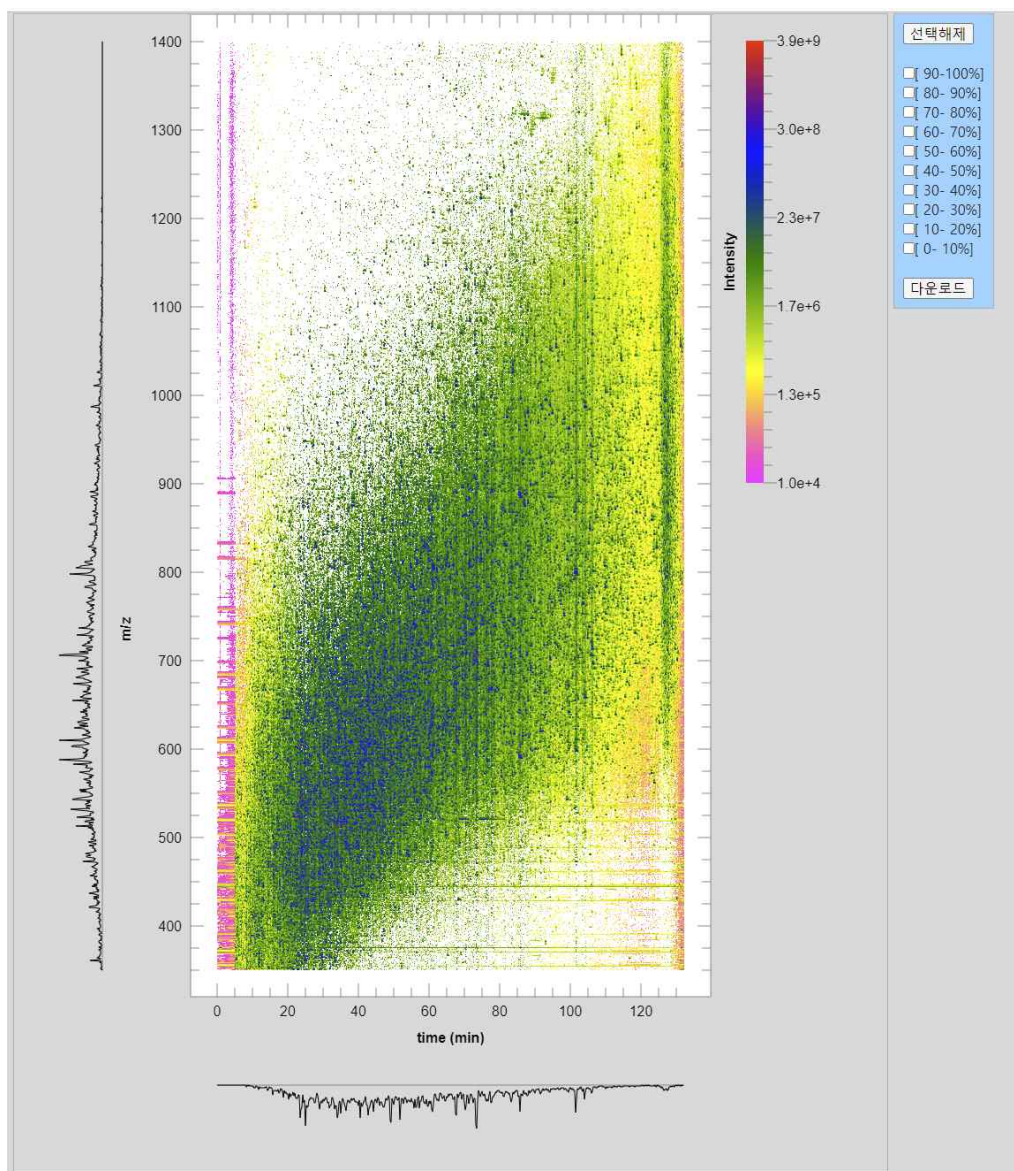


그림 59. Example of plotting mzXML formatted proteomics data

- 연구데이터 중에서 지리 정보를 표현하는 데이터가 많기 때문에, WKT(Well Known Text) 형태로 시공간 좌표를 입력하면, 이를 구글 지도 등으로 맵핑하여 게시 화면에 보여주는 가시화 기능을 개발하였다.

The screenshot displays the NaRDA web interface. At the top, there's a header with the NaRDA logo and navigation links: '연구데이터', 'DMP', 'My Data', and '검토/승인'. A search bar and a '로그아웃' button are also present.

The main content area is titled 'Well Known Text 표시 테스트'. It features a '메타 데이터' (Metadata) section on the left with links for 'Identifier', 'Title', and '조사지역'. The central part shows a map of the world with a black polygon highlighting a region in the Middle East. The map is labeled 'WKT-01' and 'Well Known Text 표시 테스트'. Below the map, there's a '파일 데이터' (File Data) section showing a file icon and the text '[신형] 44.4 bin' and '45,465,600 bytes'.

On the right side, there's a sidebar with statistics: '19 views' and '0 downloads'. Below this, there's a '나의 별점' (My Rating) section with a star rating of 2.5. The '컬렉션' (Collection) section lists '연구데이터공유센터 컬렉션'. The '제출자' (Submitter) is '관리자'. The 'Publication date' is '2020-10-19'. The 'DOI' is '10.22686/idr/2'. The 'Versions' section shows 'Version 1' with the DOI '10.22686/idr/2' and the date '2020-10-19'. There's a 'Share' section with icons for various social media platforms. The 'Cite as' section provides the citation information: '관리자 (2020-10-19) Well Known Text 표시 테스트' and the DOI link 'https://doi.org/10.22686/idr/2'. The 'Export' section has buttons for 'JSON' and 'XML'.

그림 60. Display of geographical data

□ 연구데이터 리포지터리의 안정적인 운영을 위해 메타데이터를 관리하는 데이터베이스를 이중화하고, 각종 컨테이너로 구성되는 모듈들을 분산 설치하여 운영 테스트를 진행해 보았다. 이중화에 활용된 데이터베이스는 PostgreSQL이며, PostgreSQL은 이중화를 위해 총 3대의 머신을 필요로 하고, 컨테이너 모듈에 분산 설치를 위해 가상머신 3대를 활용하였다. 향후에는 쿠버네티스와 같은 엔터프라이즈 도커 운영 소프트웨어를 활용하여 효율적인 운영 서비스 구축을 위해 개발할 것이다.

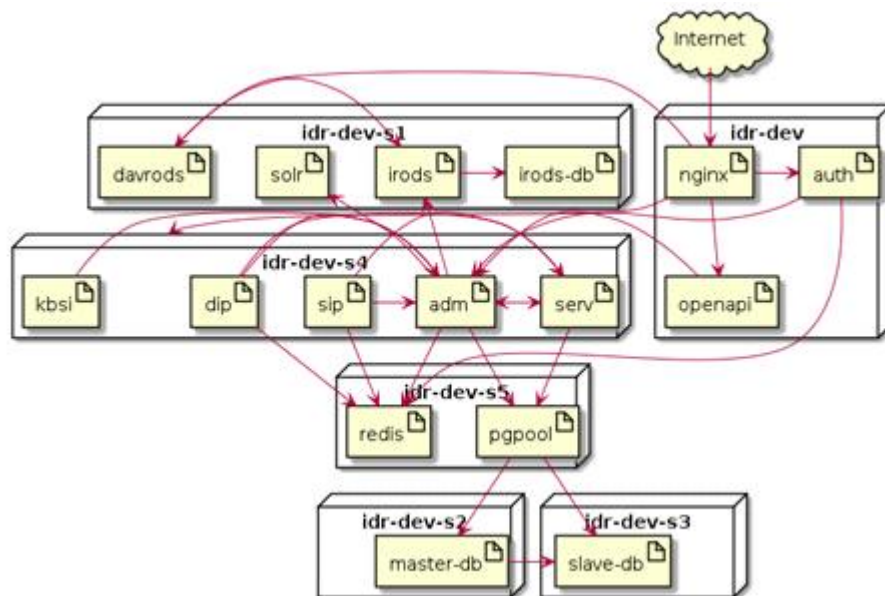


그림 61. System architecture of replication database server and distributing docker containers for research data repository