
연구데이터 수집 지침

2023. 5.

KIGAM 한국지질자원연구원
지질자원데이터센터

목 차

| | |
|-------------------|---|
| 1. 개요 | 1 |
| 2. 연구데이터 정의 | 2 |
| 3. 연구데이터 수집 | 4 |
| 4. 연구데이터 기탁 | 7 |
| 5. DMP 작성 | 9 |



표 목차

| | |
|----------------------------------|----|
| 〈표 1〉 Dublin Core 메타데이터 요소 | 5 |
| 〈표 2〉 DMP 작성 예시(1) | 9 |
| 〈표 3〉 DMP 작성 예시(2) | 10 |



그림 목차

| | |
|-----------------------------------|---|
| 〈그림 1〉 연구데이터 기탁 시 데이터 등록 절차 | 8 |
|-----------------------------------|---|



【 목 적 】

- ☐ 한국지질자원연구원 지질자원데이터센터에서 연구데이터 수집 시 적용하는 연구데이터 수집 지침(안) 제시

【 대 상 】

- ☐ 원내 연구자 및 지오빅데이터 오픈플랫폼에 연구데이터를 기탁하려는 이용자

【 적용범위 】

- ☐ 원내에서 연구 활동을 통해 생산된 연구데이터와 외부 기관 및 개인으로부터 기탁된 연구데이터에 적용함

【 준 용 】

- ☐ 본 지침에서 정하지 아니한 사항은 국가과학기술연구회의 연구데이터 관리 가이드라인 및 한국지질자원연구원의 제·규정 준용 가능

【 연구데이터 정의 】¹⁾

- ☐ “연구데이터”라 함은 연구사업의 각종 실험, 관찰, 조사 및 분석 등을 통하여 산출된 사실 자료로 연구성과의 검증에 필수적인 데이터를 의미함

【 연구데이터 분류 】²⁾

- ☐ 연구데이터는 1차 데이터, 2차 데이터, 디지털 변환 데이터, 메타데이터, 외부 데이터, 컬렉션으로 분류함
- ☐ 단, 연구노트, 과학 논문 초안, 장래 연구 계획, 동료와의 커뮤니케이션 및 실험 자료는 제외하되, 시추코어, 암석, 화석 등 실험 자료에 대한 메타데이터는 연구 데이터에 포함함
- 1차 데이터: 연구사업의 각종 조사, 실험, 관찰, 관측, 측정, 분석 등을 통하여 산출된 자료로 연구 성과의 재현에 필수적이고 객관적인 사실 데이터
 - 2차 데이터: 1차 데이터를 기반으로 가공, 분석 처리과정을 거쳐 생산한 데이터로서 표, 이미지, 그래프, 모델 및 각종 지질도, 지구물리이상도, 지화학도 등의 데이터
 - 디지털 변환 데이터: 기존의 아날로그 데이터를 디지털 형태로 변환한 데이터
 - 메타데이터: 데이터의 설명에 필요한 데이터의 제목, 데이터의 생산자, 데이터의 생산 장비 및 방법, 데이터의 내용, 획득지역(위치좌표) 및 시기, 데이터의 포맷, 데이터의 품질 등으로 구성되어 데이터를 설명하기 위해 사용되는 데이터
 - 외부 데이터: 외부로부터 유·무상의 형태로 제공받은 데이터
 - 컬렉션: 물리적 또는 디지털 리소스의 모든 집합체

【 데이터 생산 방법 기준에 따른 분류 】

- ☐ 관측 및 관찰 데이터(observational data): 관측이나 관찰을 통해 생산되는 데이터
- ☐ 실험 데이터(experimental data): 실험을 통해 생산되는 데이터
- ☐ 시뮬레이션 데이터(simulation data): 시뮬레이션을 통해 생산되는 데이터
- ☐ 추출 및 컴파일 데이터(derived or compiled data): 컴파일이나 추출을 통해 생성

1) 과학기술정보통신부(2020). 국가연구개발정보처리기준 제2조 6항.
<https://www.msit.go.kr/bbs/view.do?sCode=user&mId=108&mPid=103&bbsSeqNo=83&nttSeqNo=3175298>
 2) STAR Library. Research Data Management. <https://starlibrary.org/research-data-management>

된 데이터

【 연구데이터 공유의 중요성 및 필요성 】³⁾

- ☐ 다양한 측정 및 실험 장비로부터 생산되는 방대한 연구데이터를 통해 과학적 발견을 시도하는 데이터 집종형 연구로 전환함에 따라, 연구데이터의 재활용과 관리의 중요성이 증대됨
- ☐ 오픈 사이언스, 오픈 액세스 운동 등과 더불어, 공공 자금에서 출자한 연구과제를 통해 생산된 데이터와 출판물은 공용 리포지터리에 게시하는 형태로 전환 중임
- ☐ 연구데이터 공유는 다른 연구자 또는 기관에 데이터 활용도를 높임
- ☐ 데이터 인용(Data Citation)등을 통해 연구자 개인 또는 연구 기관의 평판을 높이고, 데이터 검증 등을 통한 더 나은 연구로 발전할 수 있는 기회가 됨
- ☐ 연구데이터의 공유는 데이터의 중복 생산·게재 비용을 줄일 수 있으며, 이를 통해 미래의 연구에 더욱 집중이 가능함
- ☐ 사생활 보호, 연구 윤리 등의 민감한 사항들을 잘 고려하여 전 처리된 대부분의 연구데이터 공유가 가능함
- ☐ 연구과제 관리 기관이 데이터관리계획(DMP: Data Management Plan, 이하 DMP) 등을 통해 연구데이터 공개를 추진하면, 데이터 획득, 저장, 메타데이터 작성 표준화는 연구데이터 공유를 실현하기 위한 연구데이터 관리의 최선의 방안임

3) 국가과학기술연구회. 연구데이터 관리·활용 지침.

<https://www.nst.re.kr/www/selectBbsNttView.do;JSESSIONID=0C1DA6F31EBF881198E95FF8105B74AB?key=64&bbsNo=34&nttNo=2205&searchCtgry=&searchCnd=all&searchKrwd=&integrDeptCode=&pageIndex=5>

【 연구데이터 수집 범위 】

- ☐ 한국지질자원연구원은 국토지질, 광물자원, 석유해저, 지질환경 등 지질자원 데이터셋을 수집 및 관리하며 이는 연구원에서 연구 활동을 통해 획득한 연구데이터(연구원에서 생산한 데이터 및 외부로부터 제공받은 데이터)를 포함함
- ☐ 상기 명시된 연구데이터 수집 범위(연구 활동을 통해 획득한 연구데이터)에는 DMP가 포함됨
 - 한국지질자원연구원의 연구데이터관리규정에 따라 연구사업 진행 시 DMP를 제출해야 하므로 이는 연구자산으로 관리·보존하는 수집 대상에 해당됨
- ☐ DMP에 명시된 연구데이터의 유형은 다음과 같음
 - 야외조사 데이터
 - 야외탐사/측정 데이터
 - 시료 및 시료분석 데이터
 - 지도자료(지질도 및 지질주제도 등 도면형태의 최종결과물)
 - 강의자료(국제지질자원교육센터)
 - 문헌/설문조사 데이터
 - 외부데이터
 - 기타-분류데이터(상기 대분류에 포함되지 않은 데이터)

【 연구데이터 선정 및 평가 】

- ☐ 연구데이터 수집 기준
 - 연구 및 교육에 대한 실질적인 가치, 지속적인 보존가치 중요성을 입증한 데이터
 - 기관이 정의한 핵심 영역에서 생산된 데이터
 - 기관이 자금을 지원하는 지질·자원분야 프로젝트에서 생산된 데이터
- ☐ 보안, 개인정보 보호 및 기밀 유지 고려사항
 - 기타 내용은 연구데이터 윤리·저작권·라이선스 지침을 따름
 - 수집 및 보관되는 민감 데이터의 경우, 프라이버시 및 기밀성에 대해 공인된 표준을 충족해야 함
 - 특정 데이터 수집에 적용할 수 있는 라이선스는 리포지터리의 자원, 목표 및 임무를 고려할 때 법률에 따라 관리되어야 함
- ☐ 저작권 및 라이선스

- 기타 내용은 연구데이터 윤리 · 저작권 · 라이선스 지침을 따름
- 연구데이터 수집 시 지적재산권의 소유자를 식별할 수 있어야 함
- 제출되는 연구데이터에 대한 지적재산권을 가진 개인 또는 기관은 리포지터리에서 정한 기탁 조건에 동의해야 함
- 연구데이터의 저작권 소유자는 데이터 이용 라이선스 설정 시 적절한 라이선스 유형을 선택 및 적용함
- 연구데이터 공유 · 활용 시 CC 라이선스 적용이 우선적으로 검토됨

☐ 연구데이터 품질

- 연구데이터 수집 시 이용자가 데이터 품질과 신뢰성을 평가할 수 있는 정보를 제공하는 포괄적 기술문서를 가진 연구데이터의 수집을 권장함
- 원본 형태의 연구데이터를 선호함

☐ 메타데이터

- 수집 및 기탁 시 사용되는 메타데이터는 <표 1>의 Dublin Core의 형식을 사용함⁴⁾

<표 1> Dublin Core 메타데이터 요소

| 요소 | | 내용 |
|-------------|-------|--|
| Title | 표제 | • 자원에 부여된 제목 |
| Creator | 창작자 | • 자원 내용에 주된 책임을 가진 개체 |
| Type | 유형 | • 자원의 내용이 가지는 성격 또는 장르로, 내용의 일반적인 범주, 기능, 장르 등을 표현함 |
| Contributor | 기여자 | • 제작자 요소에 명시된 개체 이외에 자원의 내용에 기여한 책임이 있는 기타 개체 |
| Publisher | 발행자 | • 자원을 현재의 형태로 이용가능하게 만든 실체 |
| Date | 날짜 | • 자원의 존재 기간 동안 어떠한 사건이 발생한 날짜, 자원의 제작일 또는 자원이 현재의 형태로 이용가능하게 된 시점과 관련됨 |
| Language | 언어 | • 자원의 지적인 내용을 기술하고 있는 언어 |
| Format | 형식 | • 자원의 물리적 표현형식 및 디지털 표현 형식 |
| Description | 설명 | • 자원의 내용에 대한 설명 |
| Subject | 주제 | • 자원의 내용적 주제(topic) |
| Relation | 관련 자원 | • 관련 자원에 대한 참조 |
| Identifier | 식별자 | • 자원을 식별하기 위한 식별기호 |
| Rights | 저작권 | • 자원이 가지고 있는 권리나 자원에 대한 권리에 관한 정보 |
| Source | 출처 | • 현재 자원의 출처가 되는 원 정보자원으로의 참조 |
| Coverage | 수록범위 | • 자원의 내용이 다루는 범위로 공간적 위치나 시간적 범위(시대)가 해당됨 |

4) DCMI(1999) 'Dublin Core Metadata Element Set, Version 1.1'

국립중앙도서관. DCMES 특징. <https://librarian.nl.go.kr/LI/contents/L10202000000.do>

【 선호되는 연구데이터 형식 】

- ☐ 다양한 컴퓨팅 및 기술 환경에서 사용 가능한 데이터
- ☐ 이용자가 쉽게 접근 가능한 형식의 데이터
- ☐ 연구 가치를 훼손하지 않으면서 접근 및 이용의 용이성을 갖춘 데이터
- ☐ 다양한 통계 또는 분석 소프트웨어에서 사용 가능한 형식으로 변환할 수 있는 데이터
- ☐ 데이터 해석을 위해 부가적으로 필요한 소프트웨어가 없는 데이터

【 연구데이터 윤리 】

- ☐ 기타 내용은 연구데이터 윤리 · 저작권 · 라이선스 지침을 따름
- ☐ 연구데이터 생산 · 수집 시 발생할 수 있는 윤리적 문제를 방지하기 위한 항목으로 한국지질자원연구원은 연구데이터 윤리를 준수하여 관련 데이터를 수집함

【 연구데이터 기탁의 의미 】

- ☐ 연구데이터의 향후 재사용 및 지속적 접근을 목적으로 생산한 연구데이터를 시스템(리포지터리)에 등록하는 것

【 연구데이터 기탁의 필요성 】⁵⁾

- ☐ 연구데이터의 지속적인 액세스 보장
- ☐ 전문적인 연구데이터 관리·보존·접근의 제공
- ☐ 연구데이터 기탁을 통한 스토리지 비용 절감 및 대용량 데이터의 안정적 저장 가능
- ☐ 연구데이터를 안전한 환경에서 장기적으로 관리·보존 가능
- ☐ 향후 재사용할 수 있는 잠재적 가치를 가진 자원으로 활용 가능

【 연구데이터 기탁 시 연구자 제출 자료 】

- ☐ 연구데이터 파일
 - 소프트웨어로 연구데이터의 개별 내용이나 내부구조를 가공(확인, 수정, 변환, 추출 등)할 수 있는 상태의 데이터임
 - 기탁하고자 하는 연구 데이터셋(Research Dataset) 또는 연구데이터 파일은 향후 재사용을 고려하여 범용적인 포맷으로 제공하거나 해당 도메인 분야 커뮤니티에서 보편적으로 사용되는 특정 포맷으로 제공할 것을 권장함
- ☐ 도큐멘테이션
 - 연구데이터 파일을 해석하는데 필요하거나 설명할 수 있는 도큐멘테이션 파일을 연구데이터 파일과 함께 제공함
 - 도큐멘테이션 파일에는 코드북, 데이터 수집 도구, 요약 통계, 프로젝트 요약 및 데이터 관련 출판물 목록이 포함될 수 있음
- ☐ 메타데이터
 - 연구데이터 파일과 함께 연구데이터 파일의 내용을 기술하고 있는 메타데이터를

5) 국가과학기술연구회. 출연(연) 연구데이터 관리·활용 방안 연구. 출처::
<https://www.nst.re.kr/www/selectBbsNttView.do?SESSIONID=0C1DA6F31EBF881198E95FF8105B74AB?key=64&bbsNo=34&nttNo=2205&searchCtgry=&searchCnd=all&searchKrwd=&integrDeptCode=&pageIndex=5>

제공해야 함

- 사용하는 메타데이터 형식(DC): 표제, 생성자, 유형, 기여자, 형식, 설명, 주제, 이용조건 등(연구데이터 수집 지침<표 1> 참조)

【 연구데이터 등록 절차 】

□ 등록

- 연구자는 연구데이터의 품질 및 이상 여부를 확인하고 데이터 리포지터리에 연구데이터를 제출<그림 1>

□ 승인 및 검토

- 연구책임자 또는 기관·부서의 연구데이터 관리 담당자는 제출된 연구데이터에 대한 검토 및 승인<그림 1>
- 주요 검토사항: 데이터 형식, 메타데이터 기술 형식 준용 여부, 연구데이터 공개 및 라이선스 적합 여부, 민감 정보 포함 여부 등

□ 기탁 시 검토사항

- checklist
 - 명명규칙 준수
 - 오탈자 확인
 - 파일 크기(최대 단일파일 크기 제한: 100GB)
 - DMP 분류코드 체크



<그림 1> 연구데이터 기탁 시 데이터 등록 절차

5

DMP 작성

【의미】

- ☐ DMP는 연구데이터의 생산·보존·관리 및 공동 활용 등에 관한 계획임⁶⁾
- ☐ 연구데이터의 수집, 관리, 보존, 개방, 활용을 위하여 연구계획수립 시 연구계획서와 함께 작성하여 제출하는 문서를 의미함

【작성 목적】

- ☐ 연구과정에서 생산될 데이터의 종류와 취득방법을 사전에 계획하고, 생산한 데이터를 효율적으로 관리하기 위하여 작성함

【작성 및 이행, 변경】

- ☐ DMP 작성 및 이행에 대한 사항은 한국지질자원연구원 연구데이터관리규정 제13조(데이터관리계획 작성 및 이행)에 명시되어 있으며, 작성 대상, 작성 후 검토사항 등이 포함되어 있음
- ☐ DMP 변경에 대한 사항은 한국지질자원연구원 연구데이터관리규정 제14조(데이터관리계획의 변경)에 명시되어 있음

【구성】

- ☐ DMP 작성은 <표 2,3>과 같이 작성할 수 있음

<표 2> DMP 작성 예시(1)

| | | | | |
|---------------|--------|----|--------------------------------|-----|
| 과제명 | | | | |
| 연구기간 | | | 연구책임자 | |
| 개방여부 | 개방/비개방 | 사유 | (보안과제/대외비/비공개/협약기준 등 비개방사유 기재) | |
| 1. 야외조사 데이터 | | | | |
| ID | 분류코드 | 제목 | 조사방법 및 내용 | 획득자 |
| 1-1 | | | | |
| 1-2 | | | | |
| | | | | |
| 2. 야외탐사/측정데이터 | | | | |
| ID | | 제목 | 탐사방법 및 내용 | 획득자 |
| 2-1 | | | | |

6) 국가연구개발사업의 관리 등에 관한 규정 제2조(정의)

| | | | | | |
|--|------|-------|------|------|-----|
| 2-2 | | | | | |
| 3. 시료 및 시료분석 데이터 | | | | | |
| ID | 분류코드 | 시료명 | 분석방법 | 시료 수 | 획득자 |
| 3-1 | | | | | |
| 3-2 | | | | | |
| 4. 지도자료(지질도 및 지질주제도 등 도면형태의 최종결과물) | | | | | |
| ID | 분류코드 | 제목 | 축척 | | 획득자 |
| 4-1 | | | | | |
| 4-2 | | | | | |
| 5. 강의자료(국제지질자원교육센터) | | | | | |
| ID | 분류코드 | 교육과정명 | 제목 | | 획득자 |
| 5-1 | | | | | |
| 5-2 | | | | | |
| 6. 문헌/설문조사 데이터 | | | | | |
| ID | 분류코드 | 제목 | 조사목적 | | 획득자 |
| 6-1 | | | | | |
| 6-2 | | | | | |
| 7. 외부데이터 | | | | | |
| ID | 분류코드 | 제목 | 생산기관 | | 획득자 |
| 7-1 | | | | | |
| 7-2 | | | | | |
| 8. 기타 - 분류되지 않은 데이터 (상기 대분류에 포함되지 않은 데이터) | | | | | |
| ID | 분류코드 | 제목 | 내용 | | 획득자 |
| 8-1 | ZZ00 | | | | |
| | | | | | |
| | | | | | |

☐ 메타데이터 표준은 <표 3>과 같이 작성할 수 있음

<표 3> DMP 작성 예시(2)

| 구분 | 내용 및 예시 |
|----------|--|
| 메타데이터 표준 | <ul style="list-style-type: none"> • TTAS.IS-19115: 지리정보 관리용 메타데이터 표준 • Dublin Core: 인터넷의 다양한 디지털 자원을 효율적으로 검색 및 관리하기 위한 메타데이터의 집합으로, ISO에서 표준화한 메타데이터 요소 집합 • Darwin Core: 생물 다양성 정보를 위한 Dublin Core의 확장 • ABCD(Access to Biological Collection Data): 표본 및 관찰을 포함한 1차 생물다양성 데이터의 액세스 및 교환을 위한 표준 |

| 구분 | 내용 및 예시 |
|----|--|
| | <ul style="list-style-type: none"> • ABCDEFG(Access to Biological Collection Databases Extended for Geosciences): Geoscience 데이터를 위한 ABCD의 확장 |

【 DMP 작성 시 기대 효과 】

- ☐ 과제를 통해 생성된 연구데이터의 생성, 관리, 공유 현황 파악에 용이함
- ☐ 연구데이터 재현성 확보가 가능함
- ☐ 연구데이터 재사용으로 인한 연구의 진실성 확보할 수 있음
- ☐ 중복 연구 여부를 확인하여 연구자의 연구 효율성을 제고할 수 있음



| 버전 | 일자 | 내용 |
|-----|---------------|---------------|
| 0.1 | 2023. 03. 20. | 문서 Outline 작성 |
| 0.6 | 2023. 04. 28. | 분야별 초안 작성 |
| 0.8 | 2023. 05. 08. | 지침 기관 검토 |
| 1.0 | 2023. 05. 19. | 검토 의견 반영 및 보완 |

