

# 생태 분야 데이터 리포지터리 운영 현황 분석

2024. 06. 27.

전북대학교 연구데이터융복합연구소  
김주섭 / kimjuseop@jbnu.ac.kr

# Table of Contents

- I. 데이터 리포지터리 현황
- II. 도메인별 데이터 리포지터리 운영 현황 분석
- III. 생태 분야 데이터 리포지터리
- IV. 마치며...
- V. 덧붙여...

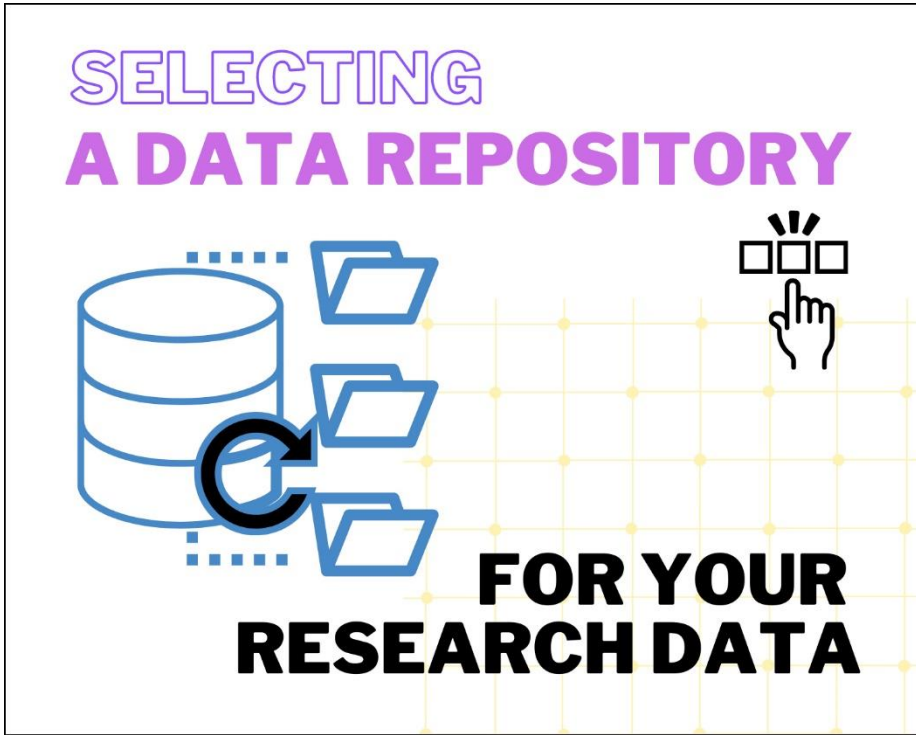
# I . 데이터 리포지터리 현황

# 데이터 리포지터리 현황

데이터 리포지터리는 연구데이터를 수집, 저장 및 보존을 통해 재사용할 수 있는 수단

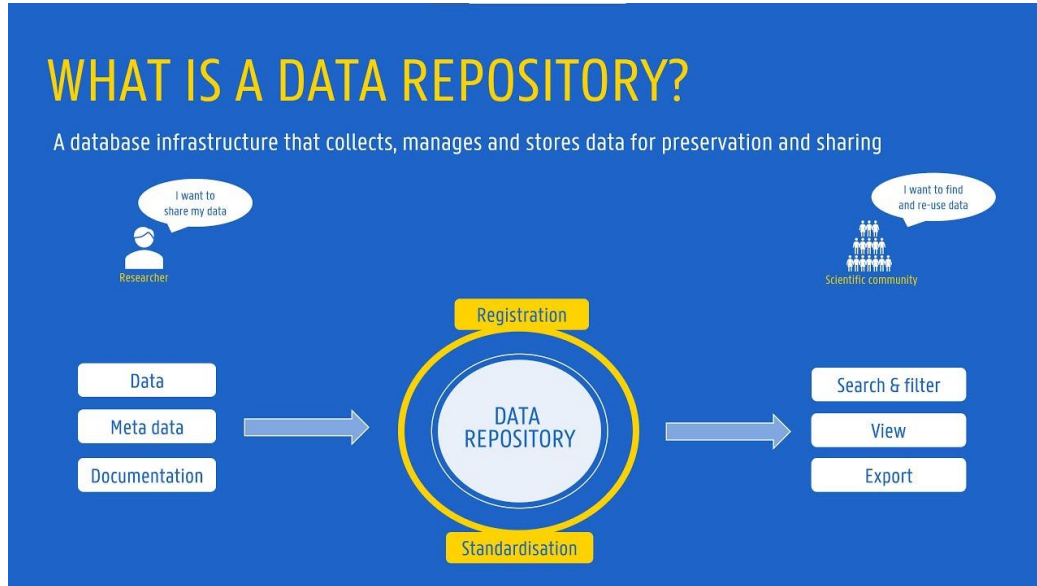
- 연구의 재현성 및 검증 차원에서 재사용하고 공유할 수 있는 데이터를 어떻게 확보할 수 있는지가 오픈사이언스 그리고 데이터 집중형 연구의 성공으로 이끌 수 있는 원동력
- 낮은 데이터의 품질은 연구 생산성을 떨어뜨릴 수 있으며 연구개발 비용과 연구 중복성을 증가시키는 등 결국에는 데이터 공유 및 재사용에 방해 원인이 될 수 있음
- 개별 데이터가 아닌 데이터를 수집, 저장 그리고 보존하는 리포지터리 자체를 평가하는 것이 결국 데이터의 신뢰성을 보장할 수 있는 방법
- 데이터 리포지터리 평가는 장기적으로 조직의 지속 가능성을 보장하고 업무의 관리 및 기술적 측면과 조직의 작업 방식에 대한 깊은 이해 나아가 향후에 데이터 품질 관리에 활용할 수 있다는 장점이 있음

Gualo, Fernando & Rodríguez, Moisés & Verdugo, Javier & Caballero, Ismael & Piattini, Mario. (2021). Data quality certification using ISO/IEC 25012: Industrial experiences. Journal of Systems and Software. 176. 110938. 10.1016/j.jss.2021.110938.

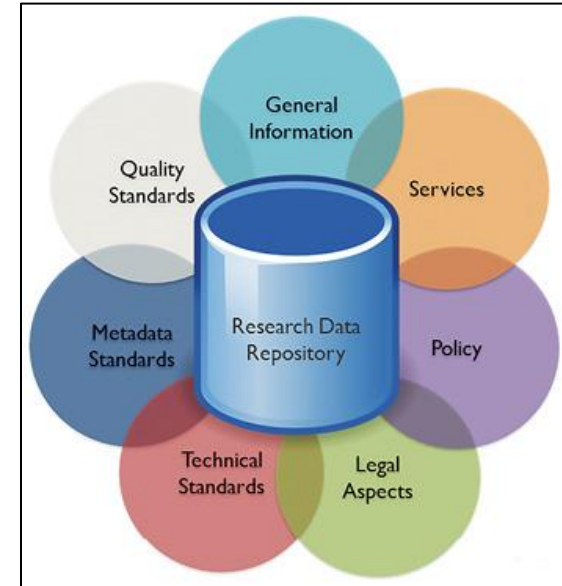


<<https://www.lib.polyu.edu.hk/publications/2021/09/data-repository/>>

# 데이터 리포지터리 현황



[https://i.ytimg.com/vi/pm\\_COU8ByYE/maxresdefault.jpg](https://i.ytimg.com/vi/pm_COU8ByYE/maxresdefault.jpg)



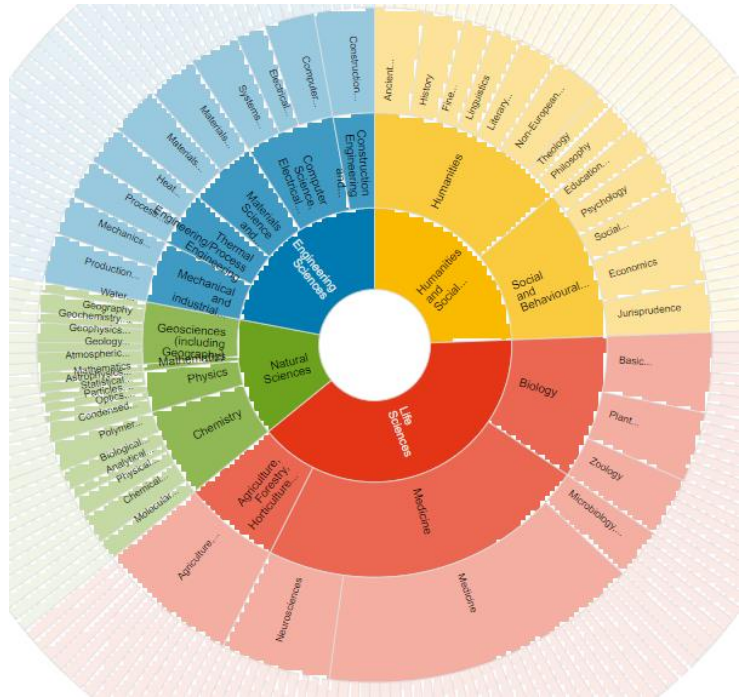
<https://www.infotoday.com/cilmag/apr16/Uzwysbyn--Research-Data-Repositories.shtml>

Where?

**re3data.org**  
REGISTRY OF RESEARCH DATA REPOSITORIES

 Search

# 데이터 리포지터리 현황



인문학 및 사회과학 / 생명과학 / 자연과학 / 공학

No,	국가	개수	No,	국가	개수
1	미국	1,181	15	오스트리아	47
2	독일	517	16	벨기에	40
3	캐나다	398	17	노르웨이	32
4	영국	323	18	스웨덴	29
5	유럽 연합	290	19	폴란드	30
6	프랑스	132	20	덴마크	27
7	오스트레일리아	105	21	러시아 연방	23
8	스위스	84	22	브라질	21
9	중국	83	23	남아프리카 공화국	20
10	네덜란드	79	24	멕시코	18
11	일본	64	25	포르투갈	15
12	스페인	55	26	뉴질랜드	15
13	인도	51	27	핀란드	13
14	이탈리아	51	28	대한민국	13

<국가별 데이터 리포지터리 현황 (2024년 6월 현재)>

## 데이터 리포지터리 현황

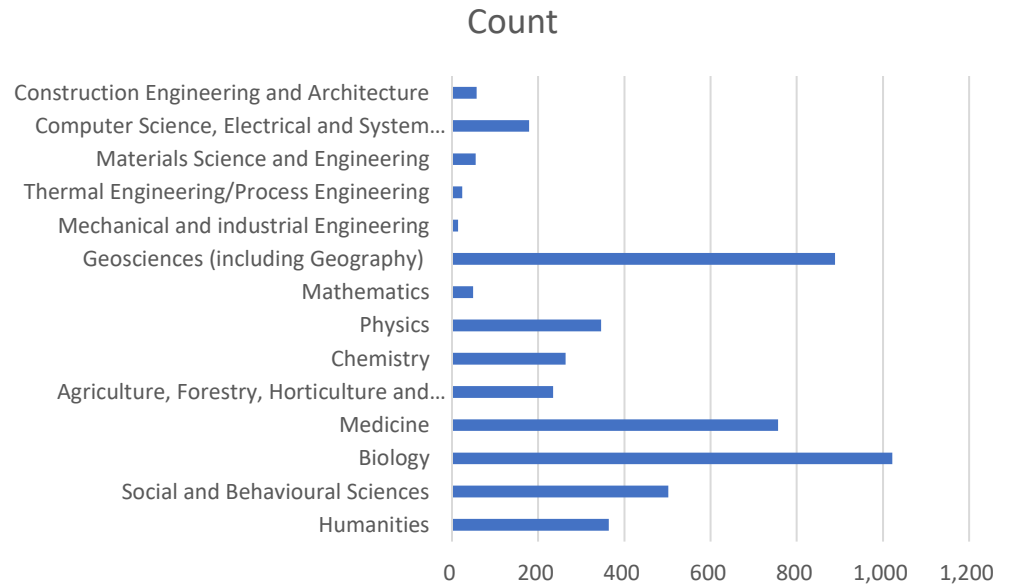
### 13개 리포지터리 중 국제 컨소시엄 제외한 7개 데이터 리포지터리

- Korean Labor & Income Panel Study (한국노동연구회)
- Jurisdictional Ocean Information Sharing System (해양수산부 연구개발사업인 '관할해역 해양 정보 공동활용 구축 연구')
- DataON Repository (한국과학기술정보연구원) ※ CTS 인증 획득 (2024년 3월 28일자)
- International Space Environment Service (국립전파연구원 우주전파센터)
- Conservation Genome Resource Bank for Korean Wildlife (서울대학교 수의과대학)
- Geo Big Data Open Platform (한국지질자원연구원)
- Korea Polar Data Center (극지연구소)

# 데이터 리포지터리 현황

## 주제별 데이터 리포지터리 현황

	Domain	Count
Humanities and Social Sciences	Humanities	364
	Social and Behavioural Sciences	502
Life Sciences	Biology	1,022
	Medicine	757
	Agriculture, Forestry, Horticulture and Veterinary Medicine	235
Natural Sciences	Chemistry	264
	Physics	346
	Mathematics	49
	Geosciences (including Geography)	889
Engineering Sciences	Mechanical and industrial Engineering	14
	Thermal Engineering/Process Engineering	24
	Materials Science and Engineering	55
	Computer Science, Electrical and System Engineering	179
	Construction Engineering and Architecture	57

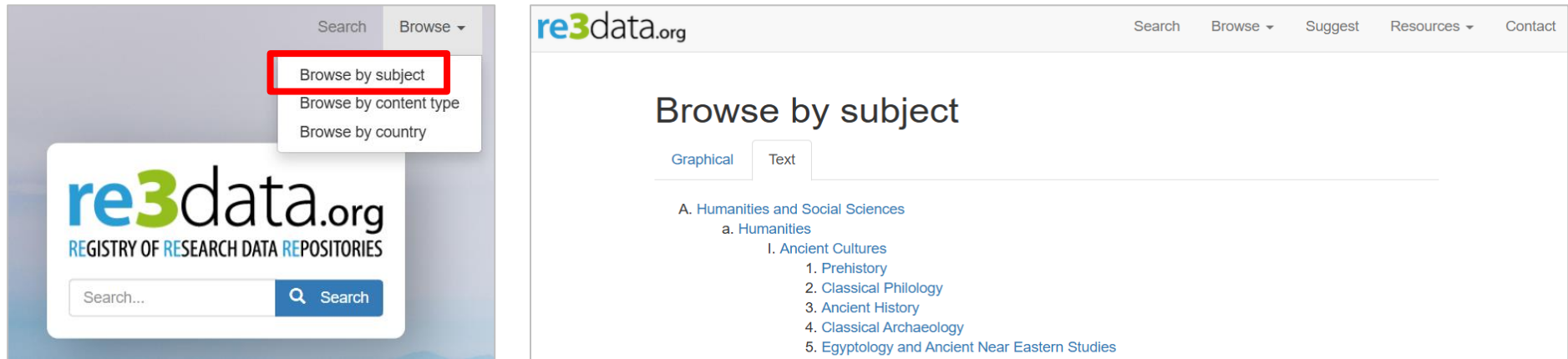




## II. 도메인별 데이터 리포지터리 운영 현황 분석

# 도메인별 데이터 리포지터리 운영 현황 분석

## 분석대상 및 방법



- Re3data.org > Browse by subject > Text 에서 각 주제분야명을 클릭했을 때 나열되는 리포지터리 목록을 참고하여 분석 진행
- **Biology, Chemistry, Geoscience, Mathematics, Physics** 5가지의 주제 분야를 대상으로 분석을 진행하였음
- 단, Biology의 경우 'II-2. Plant Ecology and Ecosystem Analysis'와 'III-3. Animal Ecology, Biodiversity and Ecosystem Research'에 한하여 분석 진행

# 도메인별 데이터 리포지터리 운영 현황 분석

## 분석대상 및 방법

	A	B	C	D	E	F	G	H	I
1	re3data_id	of_repos	itional_na	pository_u	subject	description	contact	content_type	keyword
2	r3d100011	Claremont	CCDL	https://ccc	1 Human	The Clare	digitalcoll	Audiovisu	anthropol
3	r3d100013	Kikapu	UWC Data	https://kik	1 Human	The Unive	research-	Audiovisu	FAIR
4	r3d100013	HepSim		https://atl	3 Natural	HepSim is	chekanov	Images	Monte-Ca
5	r3d100012	Dalhousie	University	https://bo	1 Human	The Dalhc	data.mana	Plain text	Canada
6	r3d100014	bonndata		https://bo	1 Human	bonndata	forschung	Audiovisu	FAIR
7	r3d100012	ETH Zürich	Research	https://ww	1 Human	The Resea	research-c	Audiovisu	ETH Zurich
8	r3d100010	DOE Data	DDE	https://ww	2 Life Scie	The DOE	ddedcomm	Audiovisu	Energy
9	r3d100010	The Netlib		https://ww	3 Natural	The Netlib	netlib@ne	Plain text	mathemat
10	r3d100012	CU Schola	University	https://sch	1 Human	CU Schola	cuscholar	Archived	c
11	r3d100013	Bilkent Ur	BUIR	https://rep	1 Human	This site	p automatio	Audiovisu	IEEE
12	r3d100010	National	SNSDL	https://oe	1 Human	The Natio	https://hel	Audiovisu	space scie
13	r3d100013	TU Wien	Research D	https://res	3 Natural	TU Wien F	https://res	Raw data	Code
14	r3d100014	datastore	by Univers	https://dat	1 Human	datastore	forschung	Archived	cmultidiscip
15	r3d100012	melbourn	University	https://me	1 Human	melbourn	research-d	Archived	ccreative ar



A	B	C	D	E	F	G	H	I	J	K	L	M
만족지표	유형	name_of_repository	오픈액세스	라이선스	인증	정책	PID	저자식별자	quality_management	버저닝	메타데이터	
9	우수	Apollo	O	O	O	Apollo Rep	DOI	ORCID	yes	yes	DataCite Metadata Schema	
6	보편	Bilkent University Institutional Rep	O	O		Content Po	hdl		yes		Dublin Core	
7	우수	bonndata	O	O		Privacy Poli	DOI		yes	yes	DataCite Metadata Schema	
7	우수	Brunel figshare	O	O		Brunel Rese	DOI	ORCID		yes	DataCite Metadata Schema	
2	열악	CEPDB		O					yes			

- Re3data에서 제공하는 API(Beta, api/beta/repositories)를 활용하여 각 주제분야에 해당하는 리포지터리 정보 추출
- [API | re3data.org](https://re3data.org)
- 리포지터리 평가 지표로 선정된 9개 요소 충족 여부를 확인, 몇 개의 평가 지표를 만족하였는지에 따라 리포지터리의 연구데이터 관리 운영수준 평가

# 도메인별 데이터 리포지터리 운영 현황 분석

## 평가 지표 (9개)

오픈액세스

라이선스

인증

정책

PID  
(Persistent Identifier Definition)

저자식별자

품질 관리  
(quality management)

버저닝  
(versioning)

메타데이터



A	B	C	D	E	F	G	H	I	J	K	L	M
만족지표	유형	name_of_repository	오픈액세스	라이선스	인증	정책	PID	저자식별자	quality_management	버저닝	메타데이터	
9	우수	Apollo	O	O	O	Apollo Repo	DOI	ORCID	yes	yes	DataCite Metadata Schema	
6	보편	Bilkent University Institutional Repoc	O	O		Content Po	hdl		yes		Dublin Core	
7	우수	bonndata	O	O		Privacy Poli	DOI		yes	yes	DataCite Metadata Schema	
7	우수	Brunel figshare	O	O		Brunel Rese	DOI	ORCID		yes	DataCite Metadata Schema	
2	열악	CEPDB		O					yes			

# 도메인별 데이터 리포지터리 운영 현황 분석

## 평가 지표 (9개)

오픈액세스

라이선스

인증

- Re3data 리포지터리 목록의 리포지터리명 오른쪽에서 확인할 수 있는 6개 아이콘으로 확인 가능
- 두 번째 아이콘 : 오픈액세스
- 세 번째 아이콘 : 라이선스
- 다섯 번째 아이콘 : 인증

**Apollo**  
DSpace@Cambridge

Subject(s)

Natural Sciences Life Sciences Humanities and Social Sciences Engineering Sciences  
Geosciences (including Geography) Mathematics Physics Chemistry  
Agriculture, Forestry, Horticulture and Veterinary Medicine Medicine Biology  
Social and Behavioural Sciences Humanities

Content type(s)

other Software applications Raw data Audiovisual data Images  
Standard office documents Scientific and statistical data formats Source code  
Plain text

Country

United Kingdom

Apollo (previously DSpace@Cambridge) is the University of Cambridge's Institutional Repository (IR), preserving and providing access to content created by members of the University. The repository stores a range of content and provides different levels of access, but its primary focus is on providing open access to the University's research publications.

### ❖ 오픈액세스 유형

- open access
- restricted access
- closed access

# 도메인별 데이터 리포지터리 운영 현황 분석

## Biology

B. Life Sciences

a. Biology

I. Basic Biological and Medical Research

1. Biochemistry
2. Biophysics
3. Cell Biology
4. Structural Biology
5. General Genetics
6. Developmental Biology
7. Bioinformatics and Theoretical Biology
8. Anatomy

II. Plant Sciences

1. Plant Systematics and Evolution
2. Plant Ecology and Ecosystem Analysis
3. Inter-organismic Interactions of Plants
4. Plant Physiology
5. Plant Biochemistry and Biophysics
6. Plant Cell and Developmental Biology
7. Plant Genetics

III. Zoology

1. Systematics and Morphology
2. Evolution, Anthropology
3. Animal Ecology, Biodiversity and Ecosystem Research
4. Sensory and Behavioural Biology
5. Biochemistry and Animal Physiology
6. Animal Genetics, Cell and Developmental Biology

- Life Science에 해당하는 **Biology** 분야
- Biology 내에서도 'II-2. Plant Ecology and Ecosystem Analysis'와 'III-3. Animal Ecology, Biodiversity and Ecosystem Research'에 해당하는 리포지터리를 대상으로 분석 진행
- II-2. Plant Ecology and Ecosystem Analysis : 72개 리포지터리
- III-3. Animal Ecology, Biodiversity and Ecosystem Research : 94개 리포지터리
- 중복되는 리포지터리를 제외하여 **총 123개의 리포지터리**를 대상으로 지표 분석·평가

# 도메인별 데이터 리포지터리 운영 현황 분석

## Biology

만족지표	유형	name of repository	오픈액세스	라이선스	인증	정책	PID	저자식별자	quality management	버저닝	메타데이터
6	보편	AEKOS Data Portal	○	○		Data Licensing Guidelines	DOI		unknown	yes	
8	우수	Ag Data Commons	○	○		Project Open Data - Open Data Policy	DOI	ORCID	yes	yes	DCAT - Data Catalog Vocabulary
6	보편	Alaska Ocean Observing System	○	○		AOOS Data Policy	DOI		yes		Darwin Core FGDC/CSDGM - Federal Geographic Data Committee Content Standard for Digital Geospatial Metadata
6	보편	Alaska Science Center	○	○		Accessibility and the U.S. Geological Survey AlgaeBase content	DOI		unknown		
4	보편	AlgaeBase	○	○			none		yes		
5	보편	AmphibiaWeb	○	○		AmphibiaWeb Data Use Policy	none	none	unknown	yes	Repository-Developed Metadata Schemas ABCD - Access to Biological Collection Data
5	보편	ARAMOB - Mobilizing Spider Data for Research	○	○		Directions for use / Presentation of results			yes		

# 도메인별 데이터 리포지터리 운영 현황 분석

## Chemistry

C. Natural Sciences  
a. Chemistry  
I. Molecular Chemistry  
1. Inorganic Molecular Chemistry  
2. Organic Molecular Chemistry  
II. Chemical Solid State and Surface Research  
1. Solid State and Surface Chemistry, Material Synthesis  
2. Physical Chemistry of Solids and Surfaces, Material Characterisation  
3. Theory and Modelling  
III. Physical and Theoretical Chemistry  
1. Physical Chemistry of Molecules, Interfaces and Liquids - Spectroscopy, Kinetics  
2. General Theoretical Chemistry  
IV. Analytical Chemistry, Method Development (Chemistry)  
1. Analytical Chemistry, Method Development (Chemistry)  
V. Biological Chemistry and Food Chemistry  
1. Biological and Biomimetic Chemistry  
2. Food Chemistry  
VI. Polymer Research  
1. Preparatory and Physical Chemistry of Polymers  
2. Experimental and Theoretical Physics of Polymers  
3. Polymer Materials

- Natural Science에 해당하는 **Chemistry** 분야
- a. **Chemistry**에 해당하는 리포지터리를 대상으로 분석 진행
- 총 263개의 리포지터리를 대상으로 지표 분석·평가



# 도메인별 데이터 리포지터리 운영 현황 분석

## Chemistry

만족지표	유형	name of repository	오픈액세스	라이선스	인증	정책	PID	저자식별자	quality management	버저닝	메타데이터
4	보편	Adult Blood Lead Epidemiology and Surveillance Interactive Database	○	○		Policies and Regulations			yes		
2	열악	Alberta Food Composition Database	○	○							
3	열악	Alcohol and Gaming Commission of Ontario Data Inventory	○	○		Digital and Data Directive					
4	보편	American Mineralogist Crystal Structure Database	○	○					yes	yes	
2	열악	AMODS	○	○							
9	우수	Apollo	○	○	○	Apollo Repository Preservation Policy	DOI	ORCID	yes	yes	DataCite Metadata Schema
6	보편	ASTM International	○	○		ASTM Policies	DOI		yes	yes	
7	우수	ASU Library Research Data Repository	○	○		ASU Library Privacy Statement	DOI	ORCID		yes	DDI - Data Documentation Initiative
3	열악	Atlantic Oceanographic and Meteorological Laboratory	○	○		Privacy Policy					
5	보편	Atmospheric Chemistry Experiment	○	○		ACE Mission Information for Public Data Release			yes	yes	

# 도메인별 데이터 리포지터리 운영 현황 분석

## Geoscience

### d. Geosciences (including Geography)

#### I. Atmospheric Science and Oceanography

1. Atmospheric Science
2. Oceanography

#### II. Geology and Palaeontology

1. Geology and Palaeontology

#### III. Geophysics and Geodesy

1. Geophysics
2. Geodesy, Photogrammetry, Remote Sensing, Geoinformatics, Cartography

#### IV. Geochemistry, Mineralogy and Crystallography

1. Geochemistry, Mineralogy and Crystallography

#### V. Geography

1. Physical Geography
2. Human Geography

#### VI. Water Research

1. Hydrogeology, Hydrology, Limnology, Urban Water Management, Water Chemistry, Integrated Water Resources Management

- Natural Science에 해당하는 **Geoscience** 분야
- **d. Geoscience (including Geography)**에 해당하는 리포지터리를 대상으로 분석 진행
- **총 889개의 리포지터리** 대상으로 지표 분석·평가

# 도메인별 데이터 리포지터리 운영 현황 분석

## Geoscience

만족지표	유형	name_of_repository	오픈액세스	라이선스	인증	정책	PDI	저자식별자	quality_management	버저닝	메타데이터
9	우수	4TU.ResearchData   science.engineering.design	○	○	○	Core Trust Seal OCE General Data Policy	DOI	ORCID	yes	yes	CF (Climate and Forecast) Metadata Conventions
5	보편	Academic Seismic Portal at LDEO	○	○					yes		ISO 19115
3	열악	Academic Seismic Portal at UTIG	○	○		Terms and Conditions					
6	보편	ACTRIS Data Centre	○	○		ACTRIS Data Centre Data Policy			yes	yes	CF (Climate and Forecast) Metadata Conventions
3	열악	Adjusted and Homogenized Canadian Climate Data	○	○		Terms and conditions					
6	보편	ADS	○	○		ADS Privacy Policy	DOI			yes	DIF - Directory Interchange Format
6	보편	Advanced Cooperative Arctic Data and Information Service	○	○		NSF Directorate for Geosciences Data Policies Legals, Licensing and Attributions	DOI		yes	yes	
5	보편	AEKOS Data Portal	○	○			DOI			yes	
5	보편	AERONET	○	○		DATA - Access and Dissemination Tools			yes		FGDC/CSDGM - Federal Geographic Data Committee Content Standard for Digital Geospatial Metadata
4	보편	AfDB Statistical Data Portal	○	○		Knoema Terms of Use			yes		

# 도메인별 데이터 리포지터리 운영 현황 분석

## Mathematics / Physics

### c. Mathematics

#### I. Mathematics

##### 1. Mathematics

- Natural Science에 해당하는 **Mathematics** 분야
- **c. Mathematics** 에 해당하는 리포지터리를 대상으로 분석 진행
- 총 49개의 리포지터리 대상으로 지표 분석·평가

### b. Physics

#### I. Condensed Matter Physics

1. Experimental Condensed Matter Physics
2. Theoretical Condensed Matter Physics

#### II. Optics, Quantum Optics and Physics of Atoms, Molecules and Plasmas

1. Optics, Quantum Optics, Atoms, Molecules, Plasmas

#### III. Particles, Nuclei and Fields

1. Particles, Nuclei and Fields

#### IV. Statistical Physics, Soft Matter, Biological Physics, Nonlinear Dynamics

1. Statistical Physics, Soft Matter, Biological Physics, Nonlinear Dynamics

#### V. Astrophysics and Astronomy

1. Astrophysics and Astronomy

- Natural Science에 해당하는 **Physics** 분야
- **b. Physics** 에 해당하는 리포지터리를 대상으로 분석 진행
- 총 346개의 리포지터리 대상으로 지표 분석·평가

# 도메인별 데이터 리포지터리 운영 현황 분석

## 도메인별 데이터 리포지터리 현황 비교

주제 분야	Biology		Chemistry		Geoscience		Mathematics		Physics	
	빈도	비율(%)	빈도	비율(%)	빈도	비율(%)	빈도	비율(%)	빈도	비율(%)
우수	18	14.6	58	22.1	164	18.4	21	42.9	58	16.8
보편	79	64.2	167	63.5	496	55.8	24	49	205	59.2
열악	25	20.3	38	14.4	223	25.1	4	8.2	82	23.7
N/P	1	0.8	0	0	6	0.67	0	0	1	0.3
계	123	100	263	100	889	100	49	100	346	100

- 우수: 지표 7개 이상 만족
- 보편: 지표 4개 이상 6개 이하 만족
- 열악: 지표 3개 이하 만족
- N/P: 오픈액세스 closed access 유형

### III. 생태 분야 데이터 리포지터리

## 생태 분야 데이터 리포지터리

### 생태 분야 데이터 리포지터리

리포지터리명	국가	편당기관	비고
GBIF	덴마크		글로벌 네트워크 / WDS 인증
BExIS	독일	DFG	
EDI Data Portal	미국	NSF	CTS 인증
KNB Data Repository	미국	NCEAS / NSF	
VertNet	미국	NSF	
EIDC	영국	NERC	WDS/CTS 인증
EcoDB	중국	CAS	
Canadensys repository	캐나다, 덴마크	CFI	
SiB Colombia Portal de Datos	콜롬비아		글로벌 네트워크
FinBIF	핀란드	Partners of LAJI.FI	

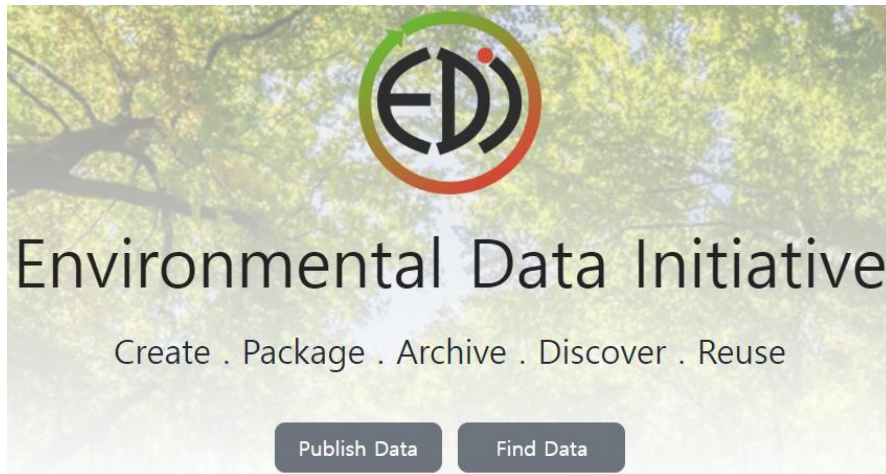
## 생태 분야 데이터 리포지터리

### 생태 분야 데이터 리포지터리 예시

구분 리포지터리	국가	편입기관	메타데이터 표준	영구 식별자	운영현황		기타
					서비스 현황	데이터 현황	
GBIF	덴마크	글로벌 네트워크	Dublin Core EML ABCD	DOI	<ul style="list-style-type: none"> <li>데이터 paper 및 웹이나 제공</li> <li>GBIF API</li> <li>데이터 정책: 개인정보보호 정책, 인용지침 등</li> <li>도구: 출판 툴킷(IPT 등) / 데이터 접근 및 사용/ GBIF LABS</li> <li>데이터셋(품질) 지원 서비스</li> <li>데이터 출판 서비스</li> <li>킬리조 노트</li> </ul>	<ul style="list-style-type: none"> <li>- 레코드: 1,900,000,000</li> <li>- 데이터셋: 60,000</li> </ul>	WDS 인증
BExIS	독일	DFG	Repository-Developed Metadata Schemas EML Darwin Core	DOI	<ul style="list-style-type: none"> <li>시스템 지원</li> <li>- 자원 예약, 문서 업로드, 이벤트 등록 시스템</li> <li>- 포토 갤러리/그래픽 연구 영역 관리</li> <li>- 기후 데이터 임베딩 시스템/ 토지 이용의 중요 매개 변수 계산</li> <li>공개 자료: 출판물, 용어집, 공공데이터</li> <li>도구: 기후 도구(ClimateTool) / LUI 계산 도구</li> </ul>	<ul style="list-style-type: none"> <li>- 데이터셋: 1,500 이상</li> <li>- 출판물: 900개 이상</li> </ul>	
EDI Data Portal	미국	NSF	Dublin Core OAI-OR- EML	DOI	<ul style="list-style-type: none"> <li>리소스: 리포지터리를 이용하는 이용자 특성에 따른 정책 안내</li> <li>데이터 정책</li> <li>- 용어 정의/ 데이터 패키지 접근성/ 민감 데이터</li> <li>- 오프라인 데이터/데이터 기여자의 지적 권리</li> <li>언어 국제화 정책</li> <li>환경 및 생태 연구 커뮤니티를 위해 설계된 도메인별 데이터 리포지터리 운영</li> <li>데이터 관리 및 출판 서비스</li> <li>데이터 패키지 DOI 발급</li> <li>데이터 큐레이터, FAIR 기준의 준수 여부 평가</li> </ul> <p>교육: 웨비나, 워크숍, EDI 정보 관리자의 개별 지원을 통해 데이터 보관 프로세스와 데이터 관리 모범 사례에 대한 교육 제공</p> <p>도구 지원: ezEML(메타데이터 편집기) 등</p>	<ul style="list-style-type: none"> <li>- 디지털 데이터 파일: 약 405,000개</li> </ul>	CTS 인증
KNB Data Repository	미국	NCEAS NSF	Dublin Core EML	PURL URN ARK DOI ORCID iD	<p>데이터 정책</p> <ul style="list-style-type: none"> <li>- 데이터 기탁 및 라이선스, 데이터 보존 정책</li> </ul> <p>도구</p> <ul style="list-style-type: none"> <li>- MetacatUI: 웹 인터페이스/ rDataONE: KNB 및 기타</li> <li>- Metacat: 메타데이터 DB</li> <li>- EML: 생태학 메타데이터 표준</li> <li>- KNB 개발자 도구: shell, R, Matlab, Python 등</li> </ul> <p>메타데이터 품질 확인: MetaDig(품질연진) 활용</p>	<ul style="list-style-type: none"> <li>- 데이터셋: 28,983</li> <li>- 메타데이터: 28,983</li> <li>- 데이터: 65,393</li> <li>- 복제본(replicas): 15,779</li> </ul>	



## 생태 분야 데이터 리포지터리



### Environmental Information Data Centre

The Environmental Information Data Centre (EIDC) is part of the Natural Environment Research Council's (NERC) Environmental Data Service and is hosted by the UK Centre for Ecology & Hydrology (UKCEH). We manage nationally-important datasets concerned with the terrestrial and freshwater sciences.



<b>Find &amp; use data</b> Find and access the data & services we provide in our data catalogue	<b>Deposit data</b> We'll guide you through the deposit process with a dedicated team of Data Centre Operatives	<b>Impact</b> Open science maximises the impact of your science, increases the visibility and transparency of your research and inspires confidence in your work	<b>Support</b> We provide support to NERC-funded researchers in data management planning and deposit to an appropriate data centre
<b>About</b> Who we are and what we do	<b>Contact us</b> Get in touch for help and advice		

- EDI는 FAIR 데이터를 만들기 위해 최선을 다하며, 고품질 데이터와 메타데이터를 보관하고 출판하는 데 도움이 되는 지원, 교육 및 리소스를 제공
- 안전한 데이터 리포지터리를 운영하고 LTER Network Communications Office 및 DataONE과 같은 정보 관리 분야의 다른 리더들과 긴밀히 협력하여 데이터 관리 모범 사례 및 관리를 장려

- NERC의 펀딩을 통해 생성된 데이터 등 장기 저장, 공개 배포, 큐레이션 을 위해 재사용 가능한 형식의 환경 데이터를 수집하며, 연구자들이 생성 하는 데이터가 장기적으로 적절하게 선별되어 이용자 커뮤니티에 대한 보안, 접근성 및 재사용성을 보장
- 프로젝트 기간에 생성된 데이터 관리에 대한 모범사례 조언 및 장기적 큐레이션을 위해 EIDC에 데이터 저장 지원을 목적

## 생태 분야 데이터 리포지터리

### 생태 분야 데이터 리포지터리 비교 분석

리포지터리		GBIF	BExIS	EDI	KNB	VertNet	EIDC	EcoDB (ScienceDB)	Canadensys repository	SIB Colombia	FinBIF	EcoBank	Total
서비스 현황	데이터 정책	0		0	0	0	0	0		0	0		8
	데이터 품질 체크	0		0	0				0	0	0		6
	데이터 교육 및 워크숍	0		0		0	0			0			5
	용어집		0										1
	출판	0	0	0		0		0	0	0	0	0	9
	도구 및 API	0	0	0	0	0		0	0	0	0	0	10
	지도 서비스	0		0	0	0			0	0	0	0	8
	리포지터리 인증	0 (WDS)		0 (CTS)			0 (CTS/WDS)						

## 생태 분야 데이터 리포지터리

### 데이터 리포지터리 서비스 제안

- 1) 데이터 정책:** 연구데이터 관리를 위한 데이터 정책 및 가이드라인이 해외 데이터 리포지터리의 경우 해당 사이트에 게시하여 이용자들에게 서비스하고 있다. 이러한 데이터 정책은 연구자와 일반 이용자를 구분하여 구성될 필요가 있으며 또한 데이터 생명 주기에 맞춘 정책 및 가이드라인으로 구현
- 2) 데이터 품질 검토:** 데이터 형식, 품질 및 적합성 확보를 위한 최소 표준 준수 여부 체크, 메타데이터 품질, 데이터 정리 및 구조화를 위한 도구 제공 그리고 데이터를 필터링하기 위한 도구 지원 등
- 3) 데이터 관리 교육 및 워크숍:** 학문 커뮤니티별로 데이터 관리가 차별화되어 이용자에게 제공되어야 하며, 서비스 제안 사항으로 제시된 품질 검토도 오프라인 및 온라인 교육 과정을 개설하여 연구자 또는 일반 이용자를 대상으로 제공하여야 함

## IV. 나가며...

나가며...

## 1. 데이터 정책 예시

대분류	소분류	세부내용
개요	1.1 연구데이터 가이드라인 개요	가. 연구데이터 정의 .....
수집	2.1 데이터 분류	가. 연구데이터 분류 .....
	2.2 데이터 기탁	가. 데이터 기탁의 의미 .....
관리	3.1 개요	가. 목적, 나. 대상 .....
	3.2 DMP 작성	가. DMP 개요, 나. 작성 목적 .....
	3.3 데이터 저장	가. 데이터의 저장 .....
	3.4 데이터 기술(data description)	가. 메타데이터의 의미 .....
	3.5 데이터 품질보장(data quality assurance)	가. 품질 관리의 필요성 .....
	3.6 FAIR 데이터	가. FAIR 데이터 정의 .....
보존	4.1 개요	가. 목적 나. 적용범위 .....
	4.2 데이터 관리 계획과 보존	
	4.3 데이터 파일 정리 및 기술	
	4.4 장기 보존 데이터	가. 장기 보존 데이터 선택의 필요성 .....
	4.5 기술 및 절차 안정성을 위한 보존 전략	
	4.6 연구데이터 백업	
공유 및 활용	5.1 데이터 접근(data access)	가. 연구데이터 활용 나. 공유 기본 원칙 .....
	5.2 인용 및 출판(citation and publish)	가. 목적, 나. 대상, 다. 적용범위 .....
	5.3 라이선스 및 저작권 (license and copyright)	가. 개요, 나. 필요성 .....
	5.4 엠바고(embargo)	가. 엠바고의 정의, 나. 엠바고의 적용 .....
	5.5 인센티브	가. 연구데이터 인센티브 정의 및 목적 .....

나가며...

## 2. 데이터 품질 검토 (1)

### 연구데이터 자체 평가 도구 : FAIR 데이터 체크리스트

- 예시 : 개인 체크리스트 - Reusable(재사용가능성)

구분	점수	체크리스트	해당 여부
다음 중 데이터에 첨부된 라이선스/사용 권한을 가장 잘 설명한 것은 무엇입니까?	0	라이선스 없음	
	1	비표준 텍스트 기반 라이선스	
	2	비표준 머신러닝 라이선스 (데이터를 어떤 조건에서 재사용할 수 있는지를 나타내는 라이선스)	
	2	표준 텍스트 기반 라이선스	
	3	표준 기계판독형 라이선스 (예 : Creative Commons)	
데이터 재사용을 촉진하기 위해 얼마나 많은 프로벤스 정보가 수집되었습니까?	0	프로벤스 정보가 기록되지 않음	
	1	일부(부분적) 기록	
	2	텍스트 형식으로 전체 기록	
	3	기계 판독 가능한 형식으로 완전 기록	

나가며...

## 2. 데이터 품질 검토 (2)

### 예시 : 조직 체크리스트 - 지표, 우선순위, 체크리스트

ID	Useful 지표
RDA-A1.2-02D	인증 및 허가를 지원하는 액세스 프로토콜을 통해 데이터에 액세스할 수 있습니다.
RDA-I2-01D	데이터에서 공정 준수 어휘 사용
RDA-I3-01D	데이터에는 다른 데이터에 대한 참조가 포함됩니다.
RDA-I3-02M	메타데이터는 다른 데이터에 대한 참조를 포함합니다.
RDA-I3-04MRD	데이터에 다른 데이터에 대한 정규화된 참조가 포함됩니다.

Essential	대부분의 상황에서 공정성을 달성하기 위해 가장 중요한 측면. 본 지표가 충족되지 않는다면 실질적인 공정성 달성이 불가능함
Important	특정 상황에서 가장 중요하지는 않을 수 있지만, 만족도 충족 시 공정성을 상당히 높일 수 있음
Useful	충족되면 좋은 측면을 다루지만 반드시 필수적이지는 않음

FAIR	ID	지표 설명	우선순위	해당 여부
F1	RDA-F1-01M	메타데이터는 영구 식별자로 식별됩니다.	●●●	
	RDA-F1-01D	데이터는 영구 식별자로 식별됩니다.	●●●	
	RDA-F1-02M	메타데이터는 전체적으로 고유한 식별자로 식별됩니다.	●●●	
	RDA-F1-02D	데이터는 전체적으로 고유한 식별자로 식별됩니다.	●●●	
A1	RDA-A1-01M	메타데이터에는 사용자가 데이터에 액세스할 수 있도록 하는 정보가 포함되어 있습니다.	●●○	
I2	RDA-I2-01D	데이터는 Fair 호환 어휘를 사용합니다.	●○○	
R1	RDA-R1-01M	재사용이 가능하도록 정확하고 관련성이 높은 여러 속성이 제공됩니다.	●●●	

나가며...

### 3. 데이터 관리 교육

## RDM 교육 훈련 프로그램

구분	교육 프로그램	내용
1. 오리엔테이션 및 배경	RDM 교육 소개	<ul style="list-style-type: none"> <li>RDM 과정 개요</li> <li>학습목표</li> <li>코스형식</li> <li>온라인 과정, 튜토리얼 목록</li> <li>날짜 및 장소</li> </ul>
2. 연구환경의 이해와 DMP	RDM 정의 및 중요성	<ul style="list-style-type: none"> <li>연구 성격 및 RDM 필요성</li> <li>데이터 라이프사이클</li> <li>연구데이터 및 RDM의 기초</li> </ul>
	오픈 사이언스와 데이터 사이언스	<ul style="list-style-type: none"> <li>오픈 사이언스 개념</li> <li>오픈 사이언스와 데이터 사이언스 및 연구데이터 관리와 차이점 설명</li> <li>오픈 사이언스의 연구 무결성 및 재현성 지원 방법 설명</li> </ul>
	DMP 작성 지원	<ul style="list-style-type: none"> <li>DMP 안내</li> <li>DMP 작성 도구(솔루션 등) 안내</li> <li>펀더 요구사항 안내</li> <li>DMP 템플릿 제시</li> <li>DMP 템플릿에 따른 작성 교육</li> </ul>
	DMP 컨설팅	<ul style="list-style-type: none"> <li>DMP Q&amp;A (질의응답)</li> <li>개인 또는 연구실 단위로 DMP 작성 지도와 첨삭 서비스</li> </ul>



## V. 덧붙여...

## 연구데이터 관리 체계 수립

### Mission

연구데이터 관리체계 안정화와 데이터 공유 및 연계 활성화

### Vision

연구데이터의 계획, 제출, 관리 등을 통한 공유 활용 생태계 구축



### Key Strategies



조직 (인력)



교육 및 훈련



인프라 구축

### Implementation Objectives

- 서비스 이용자의 요구사항을 수준 높게 해결하는 **연구데이터 관리 전담 부서**
- DMP 작성 컨설팅, 연구데이터 제출 및 검증, **연구데이터 가이드라인** 재개정, DB 백업 및 장기보존, 대외 서비스 지원

- 학문분야, 교육 대상의 특성에 맞는 교육프로그램 개발, 직접적인 **연구데이터 관리 훈련 시행**
- 연구자들의 **연구데이터 공개와 공유에 대한 부정적 인식 쇠퇴**, DMP 관련 안내 및 DMP 작성 교육, 연구데이터 관리 교육

- 데이터 공유 물리적 한계 극복, 데이터 저장 및 공유 솔루션, 연구자의 데이터 공유 동기요인 강화, **연구데이터의 저장, 접근, 공유를 위한 통합적 시스템 설계**
- 조직 차원에서 연구데이터 관리 서비스에 대한 역할과 책임 배분

## CTS 인증 및 데이터 품질 관리

### CoreTrustSeal



### CoreTrustSeal 데이터 리포지터리 인증

- RDA(Research Data Alliance) 산하에 있는 WDS(International Science Council)와 DSA(Data Seal of Approval)에서 출범한 인증 기관
- WDS-DSA 데이터 리포지터리 요구조건의 카탈로그와 절차를 기반으로 이해 관계가 있는 모든 데이터 리포지터리에 핵심 수준 인증을 제공
- 기존의 DSA 인증과 WDS 일반 회원 인증을 대체하며, DSA와 WDS의 인증을 갱신하거나 재 인증하려는 경우 CoretrustSeal 인증으로 전환하여야 함

REPOSITORY	INSTITUTION	WEBSITE	ADDRESS	VALID UNTIL	PDF
ICSD	FIZ Karlsruhe - Leibniz Institut für Informationsinfrastruktur GmbH	<a href="https://icsd.fiz-karlsruhe.de/">https://icsd.fiz-karlsruhe.de/</a>	Hermann-von-Helmholtz-Platz 1, Eggenstein-Leopoldshafen, DE	2027-04-15	
CSIRO Data Access Portal	Commonwealth Scientific and Industrial Research Organisation (CSIRO)	<a href="https://data.csiro.au/">https://data.csiro.au/</a>	GPO Box 1700 Canberra ACT 2601, Canberra, AU	2027-04-09	
DataON	Korea Institute of Science and Technology(KISTI)	<a href="https://dataon.kisti.re.kr/index.jsp">https://dataon.kisti.re.kr/index.jsp</a>	245 Daehak-ro, Yuseong-gu, Daejeon, 34141, Korea, Daejeon, KR	2027-03-28	

Certificates	Status
Coretrustseal	149
WDS	45
RatsSWD	41
CLARIN certificate B	27
DSA	15
DINI Certificate	11
DIN 31644	1
ISO 16363	1
Trusted Digital Repository	3
other	28
Total	321

# References

김우중, 김형준, 박희진, 최상옥 (2021). 자연과학 분야 연구데이터 리포지터리 속성 분석: 운영수준, 주제, 유형, 활용도를 중심으로. 기술혁신학회지, 24(4), 577-595.  
<https://doi.org/10.35978/jktis.2021.2.24.1.577>

김주섭, 강효숙, 김선태. (2023). 생태 분야 데이터 리포지터리 운영 현황 분석 및 EcoBank 서비스 제안. 한국문헌정보학회지, 57(4), 289-310.  
<https://doi.org/10.4275/KSLIS.2023.57.4.289>

EDI Data Portal. Available: <https://portal.edirepository.org/nis/home.jsp>

Environmental Information Data Centre. Available: <https://eidc.ac.uk/>.

Re3data. Available: <https://re3data.org>