Data Specifications for Data Sharing

**Preliminary investigation of data sharing in collaboration with the
Korea Institute of Science and Technology Information, KISTI**

Introduction

Data sharing allows for the provision of sustained and innovative Information services based on the acquisition and processing of data that is published and openly accessible. KISTI currently focusses on research reports, research papers, and images targeted for domestic science and technology information users. KISTI looks to broaden its focus on data sharing by collaborating with recognized organizations beyond Korea for purposes of research. It is further open to cooperate with an established (or to be established) network in response to major science and technology information issues.

1. Contact information

Institution name:	Aquatic Sciences and Fisheries Abstracts; Food and Agriculture Organization of the United Nations, ASFA-FAO
Department name:	ASFA
Person in charge:	Tamsin Vicary
Title of position:	Fishery Officer (Information)
Contact: *(Correspondence)*	Tamsin Vicary
Email:	Tamsin.Vicary@fao.org
Phone Number:	+39 06 570 50289
Website:	https://www.fao.org/asfa/en

2. Key shared data types

Research Reports	Yes
Yearbooks	Yes
Statistics	Yes
Research papers	Yes
Reference databases	No
Scientific and technical terms	Yes
Videos	No
Images	No
Sound recordings	No

3. Description of shared data types

3.1. Data Type: Research reports, Yearbooks, Statistics, Research papers

Origin of the data type: *(Brief description of the data origin and its creators)*
The data type comes from ASFA's network of over 100 institutions who create bibliographic records for their institutions

Source of information: *(Department and contact information)*
ASFA partners, full list available on ASFA website: https://www.fao.org/asfa/partnership/en

Current holdings: *(Amount of data)*
40,000 manually created records; harvested around 100,000 records.

Data update and provision cycle: *(Weekly, monthly, etc.)*
Data entered daily and updated on public search interface weekly

Service homepage: *(URL)*
The application for storing data: https://open-asfa.d4science.org/dashboard
The public search interface: https://www.fao.org/fishery/en/openasfa

Data processing and activities: *(Workflow, procedures in brief)*
ASFA Partners create record; can set up a team in which case a supervisor would review and approve; it is then published and shareable on FAO website.

Metadata fields and elements:
Different for each document type, have sent exports of Monograph, Journal article, Chapter, and Database records. See the 3 files attached as examples.

Data recording period and scope: *(Year of construction)*
Records created from 2021. Data covers materials from past 100 years.

Data Sharing Terms and License Policy: *(Open Access Compliant, CC, etc.)*
CC-BY NC

3.2. Data Type: Scientific and technical terms

Origin of the data type: *(Brief description of the data origin and its creators)*
AGROVOC and ASFA editorial communities

Source of information: *(Department and contact information)*
Ms. Imma Subirats, Partnerships and UN Collaboration Division PSU, FAO

Current holdings: *(Amount of data)*
40000+ concepts

Data update and provision cycle: *(Weekly, monthly, etc.)*

Monthly

Service homepage: *(URL)*

AGROVOC: AGROVOC Multilingual Thesaurus https://agrovoc.fao.org/browse/agrovoc/en/

Data processing and activities: *(Workflow, procedures in brief)*

Editors suggest new concepts; approved by AGROVOC curator

Metadata fields and elements:

See the file attached with examples of S&T terms

Data recording period and scope: *(Year of construction)*

Early 1980s

Data Sharing Terms and License Policy: *(Open Access Compliant, CC, etc.)*

CC-BY-4.0

4. Further interest to collaborate in data sharing for purposes of research and public awareness?

Possibly – schedule allowing.

Data Specifications for Data Sharing

**Preliminary investigation of data sharing in collaboration with the
Korea Institute of Science and Technology Information, KISTI**

Introduction

Data sharing allows for the provision of sustained and innovative Information services based on the acquisition and processing of data that is published and openly accessible. KISTI currently focusses on research reports, research papers, and images targeted for domestic science and technology information users. KISTI looks to broaden its focus on data sharing by collaborating with recognized organizations beyond Korea for purposes of research. It is further open to cooperate with an established (or to be established) network in response to major science and technology information issues.

1. Contact information

Institution name:	GreyNet, Grey Literature Network Service
Department name:	GreyNet International
Person in charge:	Dominic Farace
Title of position:	Data Administrator
Contact: *(Correspondence)*	idem
Email:	info@greynet.org
Phone Number:	31-(0)20 331 2420
Website:	https://greynet.org

2. Key shared data types

Research Reports	No
Yearbooks	No
Statistics	No
Research papers	No
Reference databases	No
Scientific and technical terms	No
Videos	Open Access via TIB AV-Portal
Images	No
Sound recordings	No
Research data	Open Access via DANS Data Station SSH
Conference papers	Open Access via GreyGuide Repository

3. Description of shared data types

3.1 Videos

Origin of the data type: *(Brief description of the data origin and its creators)*

The video recordings are created within the International Conference Series on Grey Literature some of which are recorded onsite and others via MP4 recordings. The collection of conference videos began in 2017 and is housed within the TIB AV-Portal administered by the Leibniz Information Centre for Science and Technology & University Library in Hannover, Germany.

Source of information: (*Department and contact information)*

The metadata assigned the video recordings are defined by the TIB AV-Portal

Current holdings: *(Amount of data)*

As of August 11th 2024, there are 133 videos in the GreyNet Series

Data update and provision cycle: *(Weekly, monthly, etc.)*

Annually, including Incidental retrospective input

Service homepage: *(URL)*

https://av.tib.eu/publisher/GreyNet%20International

Data processing and activities: *(Workflow, procedures in brief)*

TIB AV-Portal Agreement ➢ Data submission ➢ Data entry ➢ Storage & preservation ➢ Open access ➢ Ongoing statistics on video viewings

Metadata fields and elements:

Title	License	Language
Title of Series	Identifiers	Subject Area
Number of Parts	Publisher	Genre
Authors	Release Date	Abstract

Data recording period and scope: *(Year of construction)*

Video recordings date back to 2005; however, the series was constructed and accessible in 2017

Data Sharing Terms and License Policy

Open Access Compliant, http://creativecommons.org/licenses/by/3.0/de/deed.en

3.2 Research Data

Origin of the data type: *(Brief description of the data origin and its creators)*

The research data originates primarily within the International Conference Series on Grey Literature and is linked to the full-text papers submitted by the authors. Some of the research data originates within The Grey Journal and is linked to the journal articles submitted by the authors. Also, some of the research data originates within research projects undertaken in collaboration with GreyNet International. All of the research data is housed in the DANS Data Station for the Social Sciences and Humanities administered by the Royal Netherlands Academy of Arts and Sciences (KNAW) and the Dutch Research Council (NWO) of The Netherlands.

Source of information: (*Department and contact information)*

The metadata assigned the research data are defined by DANS Data Station Social Sciences and Humanities

Current holdings: *(Amount of data)*

As of August 11th 2024, there are 54 datasets accompanied by 122 files in GreyNet's Research Data Collection

Data update and provision cycle: *(Weekly, monthly, etc.)*

Annually, including ad hoc and retrospective input

Service homepage: *(URL)*

https://ssh.datastations.nl/dataverse/root/?q=greynet

Data processing and activities: *(Workflow, procedures in brief)*

DANS Data Agreement ➢ Data submission ➢ Data entry ➢ Storage & preservation ➢ Open access ➢ Ongoing statistics on data downloads

Metadata fields and elements:

Publication Year	Keyword	Spatial Coverage
License	Collection	Language Of Metadata
Subject	Author Name	File Type
Audience	Author Affiliation	Description

Data recording period and scope: *(Year of construction)*

Since 2011

Data Sharing Terms and License Policy

Open Access Compliant, https://creativecommons.org/publicdomain/zero/1.0/

3.3 Conference Papers

Origin of the data type: *(Brief description of the data origin and its creators)*

The conference papers are created within the International Conference Series on Grey Literature and are submitted by authors and researchers worldwide. The collection of conference papers first began in 1994 are now comprehensively housed in the GreyGuide Repository administered by the IT System and Infrastructure within the National Research Council of Italy (CNR) located in Pisa.

Source of information: (*Department and contact information)*

The metadata assigned conference papers are defined by the GreyGuide Repository in collaboration with GreyNet International

Current holdings: *(Amount of data)*

As of August 11th 2024, there are 480 conference papers in the GreyGuide GLP collection

Data update and provision cycle: *(Weekly, monthly, etc.)*

Annually

Service homepage: *(URL)*

http://greyguiderep.isti.cnr.it/listacollection.php

Data processing and activities: *(Workflow, procedures in brief)*

GreyGuide Agreement ➤ Data submission ➤ Data entry ➤ Storage & preservation ➤ Open access ➤ Usage statistics available for repository administrators

Metadata fields and elements:

Authority:	Acronym:	Target Audience
Collection:	Conference topics:	Notes:
Document Type:	Event date:	Subject(s)
Title:	Event place:	Creator of Record
Author(s):	Edited by:	Creator's e-mail:
Corporate Author(s):	Publisher:	Creator's Postal Address
Abstract Language:	Volume:	Creation date:
Abstract (English):	Number:	Date of change:
Abstract (Other language):	Pages:	Document code:
Other Language:	ISSN (print):	File(s):
Year of publication:	ISSN (e-print):	File name title:
Document URL:	ISBN:	File language:
Digital Object Identifier (DOI):	Related Sources or References:	Creative Common license:
Conference Title:	Target Sector	

Data recording period and scope: *(Year of construction)*

Since 2014

Data Sharing Terms and License Policy

Open Access Compliant, https://creativecommons.org/licenses/by-sa/3.0/

4. Further interest to collaborate in data sharing for purposes of research and public awareness?

GreyNet is most interested in collaborating with organizations both within and outside its current network in sharing diverse types of data for purposes of research and public awareness. Such an initiative would be in line with GreyNet's mission dedicated to research, publication, open access, education and public awareness of grey literature. While GreyNet's overall content is small in scale, it provides a paradigm in which full-text, associated research data, and audio-visual material are integrated and comply with FAIR data principles.

Data Specifications for Data Sharing

Preliminary investigation of data sharing in collaboration with the
Korea Institute of Science and Technology Information, KISTI

Introduction

Data sharing allows for the provision of sustained and innovative Information services based on the acquisition and processing of data that is published and openly accessible. KISTI currently focusses on research reports, research papers, and images targeted for domestic science and technology information users. KISTI looks to broaden its focus on data sharing by collaborating with recognized organizations beyond Korea for purposes of research. It is further open to cooperate with an established (or to be established) network in response to major science and technology information issues.

1. Contact information

Institution name:	Korea Institute of Science and Technology Information
Department name:	Digital Curation Center
Person in charge:	Jayhoon Kim
Title of position:	Team Leader
Contact: *(Correspondence)*	Sejong Lim
Email:	seoklim@kisti.re.kr
Phone Number:	82-042-869-1891
Website:	https://scienceon.kisti.re.kr/main/mainForm.do

2. Key shared data types

Research Reports	Accessible by Public Access Policy
Yearbooks	No
Statistics	No
Research papers	Only metadata
Reference databases	No
Scientific and technical terms	No
Videos	No
Images	Limited availability by MOU
Sound recordings	No

3. Description of shared data types

3.1 Research Reports

Origin of the data type: *(Brief description of the data origin and its creators)*

The national R&D research report is created within the performance of national research and development projects and can be viewed in its original form free of charge. Since 2009, the electronic text of reports created within national research projects are provided by specialized research management organizations administered under each ministry or government office.

Source of information: *(Department and contact information)*

Metadata is obtained from the NTIS (National Science & Technology Information Service) Center of KISTI administered under the Ministry of Science and ICT of Korea.

Current holdings: *(Amount of data)*

As of December 2023, there are 240,388 reports

Data update and provision cycle: *(Weekly, monthly, etc.)*

Ongoing, updated twice a week

Service homepage: *(URL)*

https://scienceon.kisti.re.kr/main/mainForm.do

Data processing and activities: *(Workflow, procedures in brief)*

Business plan ➢ Data collection ➢ Database construction ➢ Storage & preservation ➢ Data service/provision ➢ User survey

Metadata fields and elements:

Management number Report name Business name Project Title Host research institution	Research director Participating researcher Country name Language Date/month of publication Project start year	Host government Dept. Research project management agency Unique assignment number Page Date of acquisition

Data recording period and scope: *(Year of construction)*

Since 2009

Data Sharing Terms and License Policy

Free access according to public access policy (Open Access Compliant)

3.2 Research Papers

Origin of the data type: *(Brief description of the data origin and its creators)*

The national R&D research paper means papers published with the support of government funding. For license issue, KISTI cannot collect full-text files. Research papers metadata are sharable.

Source of information: (*Department and contact information)*

Metadata is obtained from the NTIS (National Science & Technology Information Service) Center of KISTI administered under the Ministry of Science and ICT of Korea.

Current holdings: *(Amount of data)*

As of December 2023, there are around 1.2M paper metadata

Data update and provision cycle: *(Weekly, monthly, etc.)*

Updated once a year

Service homepage: *(URL)*

https://scienceon.kisti.re.kr/main/mainForm.do

Data processing and activities: *(Workflow, procedures in brief)*

Business plan ➤ Data collection ➤ Database construction ➤ Storage & preservation ➤ Data service/provision ➤ User survey

Metadata fields and elements:

Management number Project number, ScienceON, NTIS)	Paper title Issue/number Journal title/ISSN Author Published date Full-text URL	Indexing status (SCI, Korea Citation Index)

Data recording period and scope: *(Year of construction)*

Since 2009

Data Sharing Terms and License Policy

Metadata are available

3.3 Images

Origin of the data type: *(Brief description of the data origin and its creators)*

Images are table and figures extracted from national R&D research reports. Image data are for machine learning.

Source of information: (*Department and contact information)*

Image data are extracted from the national R&D research reports.

Current holdings: *(Amount of data)*

As of August 2024, there are 9,142 tables, 10,730 figures and corresponding captions

Data update and provision cycle: *(Weekly, monthly, etc.)*

This data was developed in 2021 by special budget of covering COVID-19

Service homepage: *(URL)*

KISTI Data Sharing & Utilization Service (https://aida.kisti.re.kr)

Data processing and activities: *(Workflow, procedures in brief)*

Business plan ➤ Data collection ➤ Database construction ➤ Storage & preservation ➤ Data service/provision ➤ User survey

Metadata fields and elements:

Image file name	Captions	Image X, Y position information in pages

Data recording period and scope: *(Year of construction)*

Developed in 2021

Data Sharing Terms and License Policy

Non-commercial, Source Citation, No redistribution

Q4. Further interest to collaborate in data sharing for purposes of research and public awareness?

Data Specifications for Data Sharing

Preliminary investigation of data sharing in collaboration with the
Korea Institute of Science and Technology Information, KISTI

Introduction

Data sharing allows for the provision of sustained and innovative Information services based on the acquisition and processing of data that is published and openly accessible. KISTI currently focusses on research reports, research papers, and images targeted for domestic science and technology information users. KISTI looks to broaden its focus on data sharing by collaborating with recognized organizations beyond Korea for purposes of research. It is further open to cooperate with an established (or to be established) network in response to major science and technology information issues.

1. Contact information

Institution name:	National Library of Medicine, National Institutes of Health NLM/NIH
Department name:	NCBI Bookshelf
Person in charge:	Stacy Lathrop
Title of position:	Technical Specialist
Contact: *(Correspondence)*	Stacy Lathrop
Email:	stacy.lathrop@nih.gov
Phone Number:	+1 301-402-2430
Website:	https://www.ncbi.nlm.nih.gov/books

2. Key shared data types

Research Reports	yes
Yearbooks	no
Statistics	yes
Research papers	yes
Reference databases	yes
Scientific and technical terms	yes
Videos	no
Images	no
Sound recordings	no
Systematic reviews	yes
Standards and guidelines	yes

3. Description of shared data types

3.1. Books and Documents

Origin of the data type:

Participating organizations

Source of information:

NCBI Bookshelf; booksauthors@ncbi.nlm.nih.gov

Current holdings:

13,200 books and documents in life science and healthcare.

Data update and provision cycle:

Daily

Service homepage:

https://www.ncbi.nlm.nih.gov/books

Data processing and activities:

Submission of source Word, PDF, XML files to normalized BITS XML for ingest into PubMed Central Book databased for dynamic rendering on Bookshelf site with metadata indexed in PubMed and catalogued in NLM Catalog; subsets of data also available via FTP service as MARC records and/or normalized BITS XML

Metadata fields and elements:

All book metadata BITS supported elements: https://jats.nlm.nih.gov/extensions/bits/tag-library/2.0/element/book-meta.html	All book part metadata BITS supported elements: https://jats.nlm.nih.gov/extensions/bits/tag-library/2.0/element/book-part-meta.html

Data recording period and scope:

Launched 1999

Data Sharing Terms and License Policy:

Per the terms of the source content providers copyrights and licenses

4. Further interest to collaborate in data sharing for purposes of research and public awareness?

I am happy for Bookshelf to collaborate, but any particular details would need to be approved, beyond folks grabbing content in our OA subset for FAIR purposes: https://www.ncbi.nlm.nih.gov/books/about/openaccess/ or reuse of our provided MARC records: https://www.ncbi.nlm.nih.gov/books/about/librarians/#librarians.MARC_Records.

Data Specifications for Data Sharing

Preliminary investigation of data sharing in collaboration with the
Korea Institute of Science and Technology Information, KISTI

Introduction

Data sharing allows for the provision of sustained and innovative Information services based on the acquisition and processing of data that is published and openly accessible. KISTI currently focusses on research reports, research papers, and images targeted for domestic science and technology information users. KISTI looks to broaden its focus on data sharing by collaborating with recognized organizations beyond Korea for purposes of research. It is further open to cooperate with an established (or to be established) network in response to major science and technology information issues.

1. Contact information

Institution name:	German National Library of Science and Technology, TIB
Department name:	Research and Development, Lab for non-textual Materials
Person in charge:	Margret Plank
Title of position:	Head of Lab for non-textual materials
Contact: *(Correspondence)*	Margret Plank
Email:	Margret.plank@tib.eu
Phone Number:	+49 (0)511 762 4884
Website:	https://www.tib.eu/en

2. Key shared data types

Research Reports	No
Yearbooks	No
Statistics	No
Research papers	No
Reference databases	No
Scientific and technical terms	No
Videos	Yes
Images	No
Sound recordings	Yes

3. Description of shared data types

3.1. Scientific Videos and Audio Recordings

Origin of the data type:

Scientific Videos and Audio recordings, such as teaching and learning videos, conference recordings, animations, visualized research data, video abstracts, project recordings, etc.

Source of information:

Coming from scientific institutions such as universities, non-university research institutions, and individual scientists

Current holdings:

Approx. 47.000 Scientific Videos and Audio recordings

Data update and provision cycle:

Daily

Service homepage:

https://av.tib.eu/

Data processing and activities:

https://www.tib.eu/fileadmin/extern/knm/NTM-Metadata-Schema_documentation_v_2.2.pdf

Metadata fields and elements:

https://www.tib.eu/fileadmin/extern/knm/NTM-Metadata-Schema_documentation_v_2.2.pdf

Data recording period and scope:

2014-2024

Data Sharing Terms and License Policy:

95% CC-licensed

4. Further interest to collaborate in data sharing for purposes of research and public awareness?

For Scientific Videos yes, otherwise no