



30
JANUARY
2018

RDA EU
Data Innovation
Forum

Brussels, Belgium

Join the discussion on role of research data in the
Data Economy context and the RDA contribution to the
Data Economy building blocks

rd-alliance.org/rdae-data-innov-forum-2018



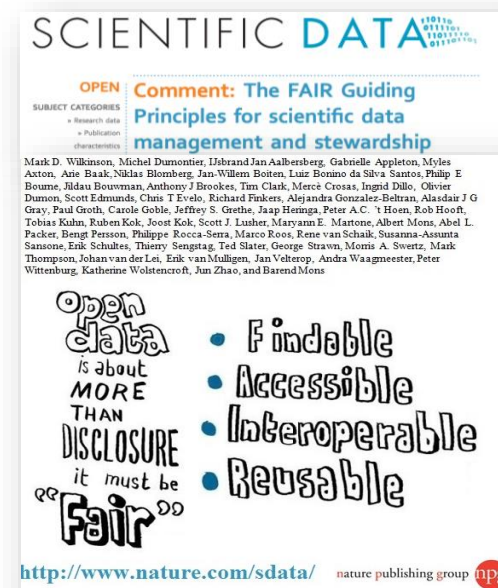
The FAIRsharing Registry – an output of the BioSharing WG
Peter McQuilton, PhD
Co-chairs: Susanna-Assunta Sansone,
Rebecca Lawrence and Simon Hodson

Summary of the Problem

The WG was set up to construct a set of guidelines for linking information on Standards, Databases and Data Policies

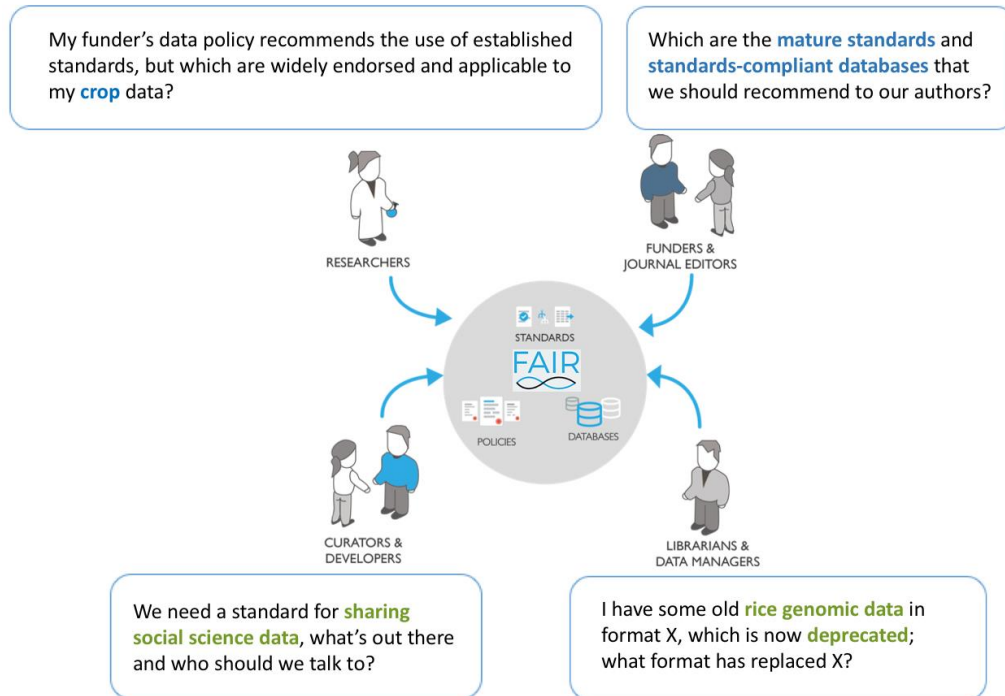
- Standards are reinvented
- Databases use the wrong standards
- Data policies fail to endorse appropriate resources
- Increasingly, data is lost as it is not
 - Findable
 - Accessible
 - Interoperable
 - Reusable

(The FAIR principles – [DOI:10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18))



Stakeholders affected by this WG

- Data creators and (re)users
- People who recommend or endorse the use of data standards and databases
- Focused on the life, biomedical and environmental sciences

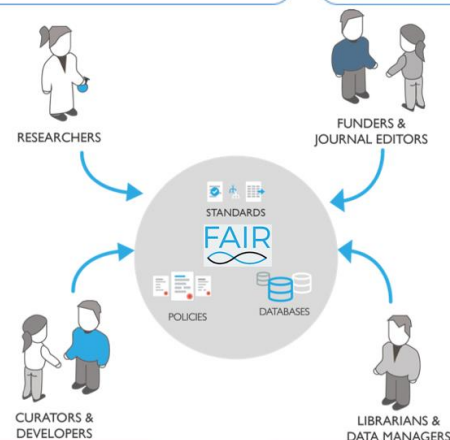


Stakeholders affected by this WG

- Data creators and (re)users
- People who recommend or endorse the use of data standards and databases
- ~~■ Focused on the life, biomedical and environmental sciences~~
- Domain agnostic

My funder's data policy recommends the use of established standards, but which are widely endorsed and applicable to my **crop** data?

Which are the **mature standards** and **standards-compliant databases** that we should recommend to our authors?



We need a standard for **sharing social science data**, what's out there and who should we talk to?

I have some old **rice genomic data** in format X, which is now **deprecated**; what format has replaced X?

The Working Group structure and composition

BioSharing Registry: connecting data policies, standards & databases in life sciences WG

Status: Recognised & Endorsed

Chair(s): Susanna-Assunta Sansone, Rebecca Lawrence, Simon Hodson & Peter McQuilton

Group Email:

Secretariat Liaison: Herman Stehouwer

TAB Liaison: Peter Wittenberg and Simon Cox

BioSharing WG

Recommendation Title: BioSharing Recommendations: data repositories, standards and policies in the life, biomedical and environmental sciences

Recommendation package DOI: <http://dx.doi.org/10.15497/RDA00017>

Group Co-Chairs:

Peter McQuilton, University of Oxford, UK

Simon Hodson, CODATA

Rebecca Lawrence, F1000

Susanna-Assunta Sansone, University of Oxford, UK

- >80 members from around the world
 - Standard/database developers
 - Funders
 - Journal data editors
 - Researchers
 - Data managers/librarians

Highlights of the Recommendation

- Guidelines for each stakeholder to describe and link standards, databases and data policies
 - Standard and database creators and maintainers
 - Researchers
 - Data Managers
 - Journal publishers and funders who have a data policy
- A manually curated registry describing and interlinking standards, databases (and repositories), and data policies.

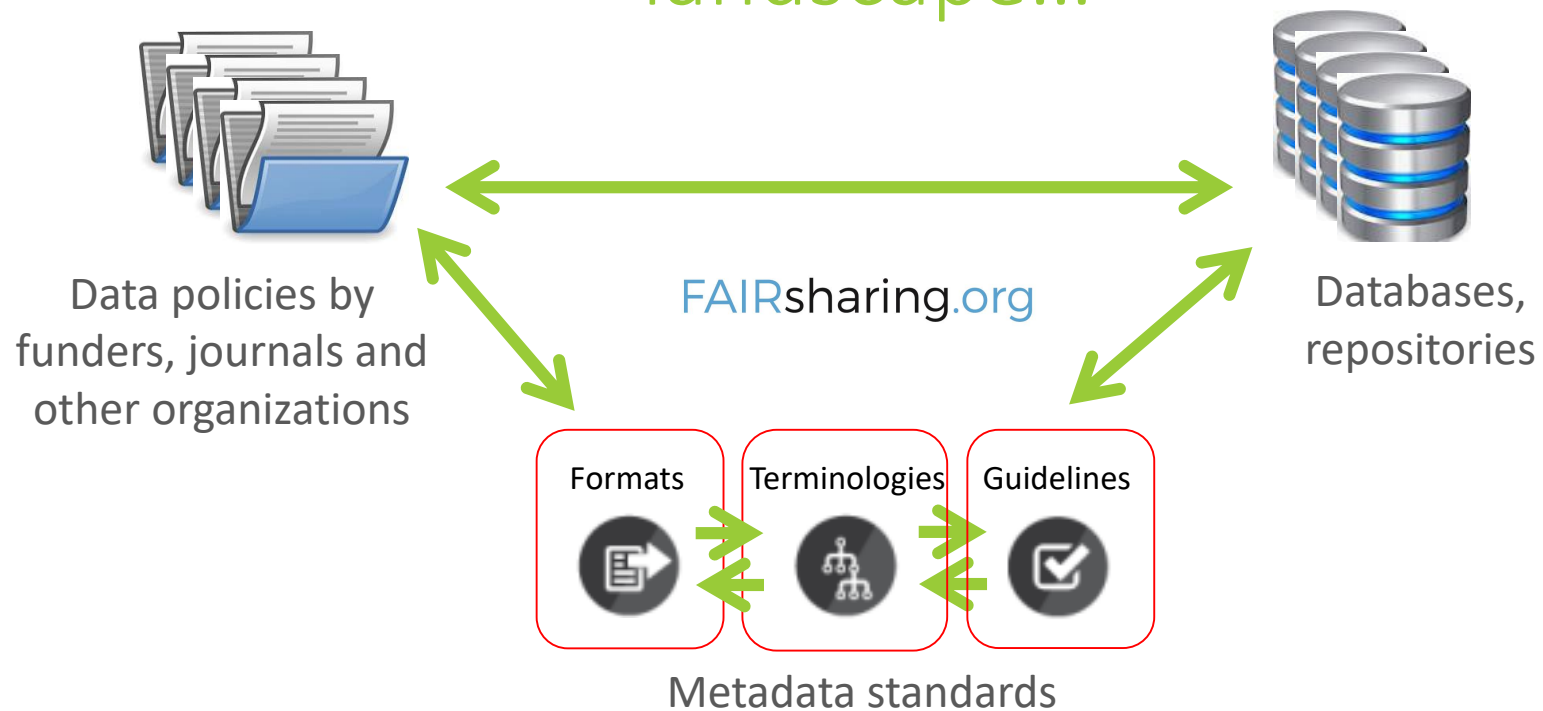
FAIRsharing.org

FAIRsharing.org

A web-based, curated, and searchable portal that monitors the development and evolution of standards, across all disciplines, inter-related to databases/repositories and data policies

Making Standards, Databases and Policies FAIR

Mapping a complex and evolving landscape...



Assigning status indicators

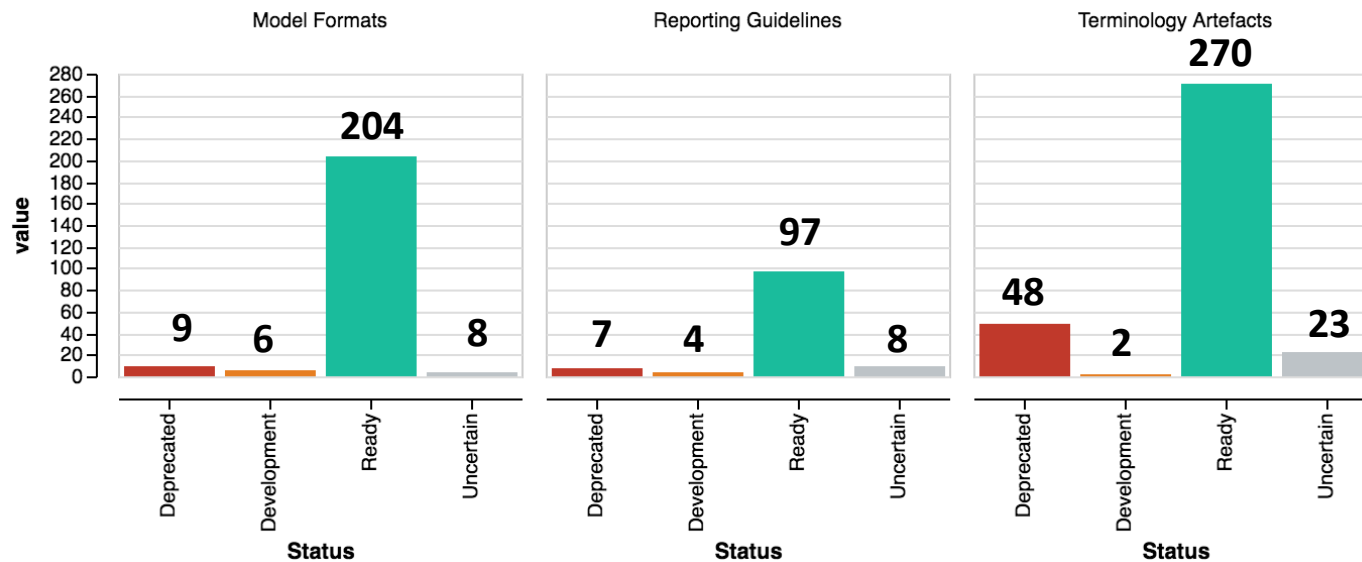
R Ready for use, implementation, or recommendation

Dev In development

U Status uncertain

D Deprecated as subsumed or superseded

All records are manually curated
in-house and verified by the
community behind each resource



*Paper submitted,
preliminary information as of July 2017*

A curated, informative and educational resource on data and metadata *standards*, inter-related to *databases* and data *policies*.

[Find](#)


Recommendations

Standards and/or databases recommended by journal or funder data policies.

[Discover](#)


Collections

Standards and/or databases grouped by domain, species or organization.

[Learn](#)


Educational

About standards, their use in databases and policies, and how we can help you.



[Search](#)

[Standards](#)

[Databases](#)

[Policies](#)

[Collections/Recommendations](#)

Advanced Search



Fine grained control over your search.

Search Wizard



ask FAIRsharing

Let us guide you to your results.



1098 Standards

Terminology Artifact
Model/Format
Reporting Guideline

736
242
120

[View all](#)


1012 Databases

Life Science
Biomedical Science
General Purpose

752
185
11

[View all](#)


105 Policies

Funder
Journal
Society

22
74
5

[View all](#)



Environmental Data Initiative Data Repository

Abbreviation: EDI

General Information

The Environmental Data Initiative is an NSF-funded project meant to accelerate curation and archive of environmental data, emphasizing data from projects funded by the NSF DEB. Programs served include Long Term Research in Environmental Biology (LTREB), Organization for Biological Field Stations (OBFS), Macrosystems Biology (MSB), and Long Term Ecological Research (LTER).

Homepage <https://portal.edirepository.org/>

Developed in [United States of America](#)

Created in 2009

Taxonomic range

All

Scope and data types

- Agriculture
- Atmospheric Science
- Biology
- Earth Science
- Ecology
- Environmental Science
- Geophysics
- Life Science
- Natural Sciences
- Oceanography
- Plant Science

Collected/Recommended By



This record is maintained by [jbrunt](#)

Record updated: Aug. 4, 2017, 10:24 a.m. by [The FAIRsharing Team](#).



Associated recommendations

[PLOS](#)

[Scientific Data](#)

Support

General

support email info@environmentaldatainitiative.org

help [https://environmentaldatainitiative.org/ ...](https://environmentaldatainitiative.org/)

online documentation [http://pastaplus-core.readthedocs.io/en/ ...](http://pastaplus-core.readthedocs.io/en/...)

Additional Information

Contact [Mark Servilla](#) [ORCID](#)

Tools

No tools defined

Access / Retrieve Data

Conditions of Use

Applies to: Data use

[Creative Commons Attribution Zero \(CC0\)](#)

[Creative Commons Attribution \(CC-BY\) 3.0](#)

Data Access

Advanced Search <https://portal.edirepository.org/nis/advancedSearch.jsp>

Related Standards

Reporting Guidelines

No guidelines defined

Terminology Artifacts

No semantic standards defined

Models and Formats

[Ecological Metadata Language](#)

Related Databases

No related databases defined

Implementing Policies

[Scientific Data's Recommended Data Repositories](#)

[PLOS Recommended Data Repositories](#)

Credit

Record Maintainer

This record is maintained by [jbrunt](#) [ORCID](#)

Funds

[US National Science Foundation \(NSF\)](#) (Government body)

Undefined

[University of New Mexico \(University\)](#)

[University of Wisconsin-Madison](#)

Grant Number(s)

[DBI-1565103](#) (US National Science Foundation (NSF))

[DEB-1629233](#) (US National Science Foundation (NSF))

Grouping the data

Collections group together one or more types of resource by domain, project or organization.

Recommendations are a core-set of resources that are selected and recommended by a funder or journal data policy.

The image displays six panels, each representing a different organization or journal's data collection interface. Each panel includes a header with the organization's name and logo, followed by a list of filters (Standards, Policies, Databases) with associated counts. Below these are sections for 'Taxa types' and 'Data types' with specific selection options.

- Clinical Research (CDISC) ORGANIZATION**: 12 Standards, 0 Policies, 0 Databases. 1 Taxa type: Homo Sapiens. 3 Data types: Biomedical Science, Clinical Data, Metadata Standards.
- Computational Modeling (COMBINE) ORGANIZATION**: 7 Standards, 0 Policies, 0 Databases. 1 Taxa type: All. 3 Data types: Computational Modeling, Life Science, Metadata Standards.
- EMBL-ABR ORGANIZATION**: 3 Standards, 0 Policies, 1 Databases. 5 Taxa types: Homo Sapiens, Invertebrata, Plantae, Protozoa, Vertebrata. 5 Data types: Bioinformatics Tools And Formats, Biological And Functional Annotation.
- EMBO Press JOURNAL**: 9 Standards, 1 Policies, 21 Databases. 1 Taxa type: All. 5 Data types: Biological Databases, Biomedical Science, Data Sharing, Life Science.
- F1000Research JOURNAL**: 5 Standards, 1 Policies, 47 Databases. 1 Taxa type: All. 5 Data types: Biological Databases, Biomedical Science, Data Sharing, Life Science.
- PLOS JOURNAL**: 1 Standards, 1 Policies, 59 Databases. 1 Taxa type: All. 5 Data types: Biological Databases, Biomedical Science, Data Sharing, Life Science.

PLOS has identified a set of established repositories and data standards, which are recognized and trusted within their respective communities, for their journals PLOS Biology, PLOS Medicine, PLOS ONE, PLOS Genetics, PLOS Computational Biology, PLOS Pathogens, and PLOS Neglected Tropical Diseases. PLOS requires that authors comply with field-specific standards for preparation and recording of data and select repositories appropriate to their field, for example deposition of microarray data in ArrayExpress or GEO; deposition of gene sequences in GenBank, EMBL or DDBJ; and deposition of ecological data in Dryad. Authors are encouraged to select repositories that meet accepted criteria as trustworthy digital repositories.

This record is maintained by: [eganley](#) , [PLOSData](#)

This recommendation is taken from the following policy:

[PLOS Recommended Data Repositories.](#)

Record updated: April 10, 2017, 11:41 a.m. by [The FAIRsharing Team](#).

[Homepage](#) [Reference](#)

Taxonomic range

[All](#)

Scope and data types

[Biological Databases](#) [Biomedical Science](#) [Data Sharing](#) [Life Science](#) [Metadata Standards](#)

[View as Graph on Preview Server \(BETA\)](#)

[View as Grid](#) [View as Table](#)

Sort by
Best Match

Recommended Records
[Recommended](#)

Associated Publication?
[No Publication](#) [Has Publication](#)


Claimed?
[No Maintainer](#) [Has Maintainer](#)

Record Status
[Uncertain](#) [Deprecated](#) [In developme](#) [Ready](#)

Standard Type

Showing records 1 - 50 of 62.

« 1 2 »



ArrayExpress
DATABASE


[Standards](#) 6

[Publications](#) 2

[In Collections](#) 1

[Recommended](#) 6

1 Taxa types, including:



BioGRID
Biological General Repository for Interaction Datasets
DATABASE


[Standards](#) 3

[Publications](#) 1

[In Collections](#) 3

[Recommended](#) 2

3 Taxa types, including:



BMRB
Biological Magnetic Resonance Databank
DATABASE

[Standards](#) 3

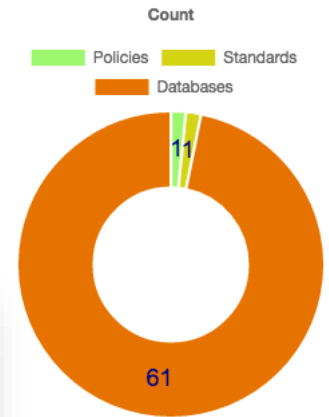
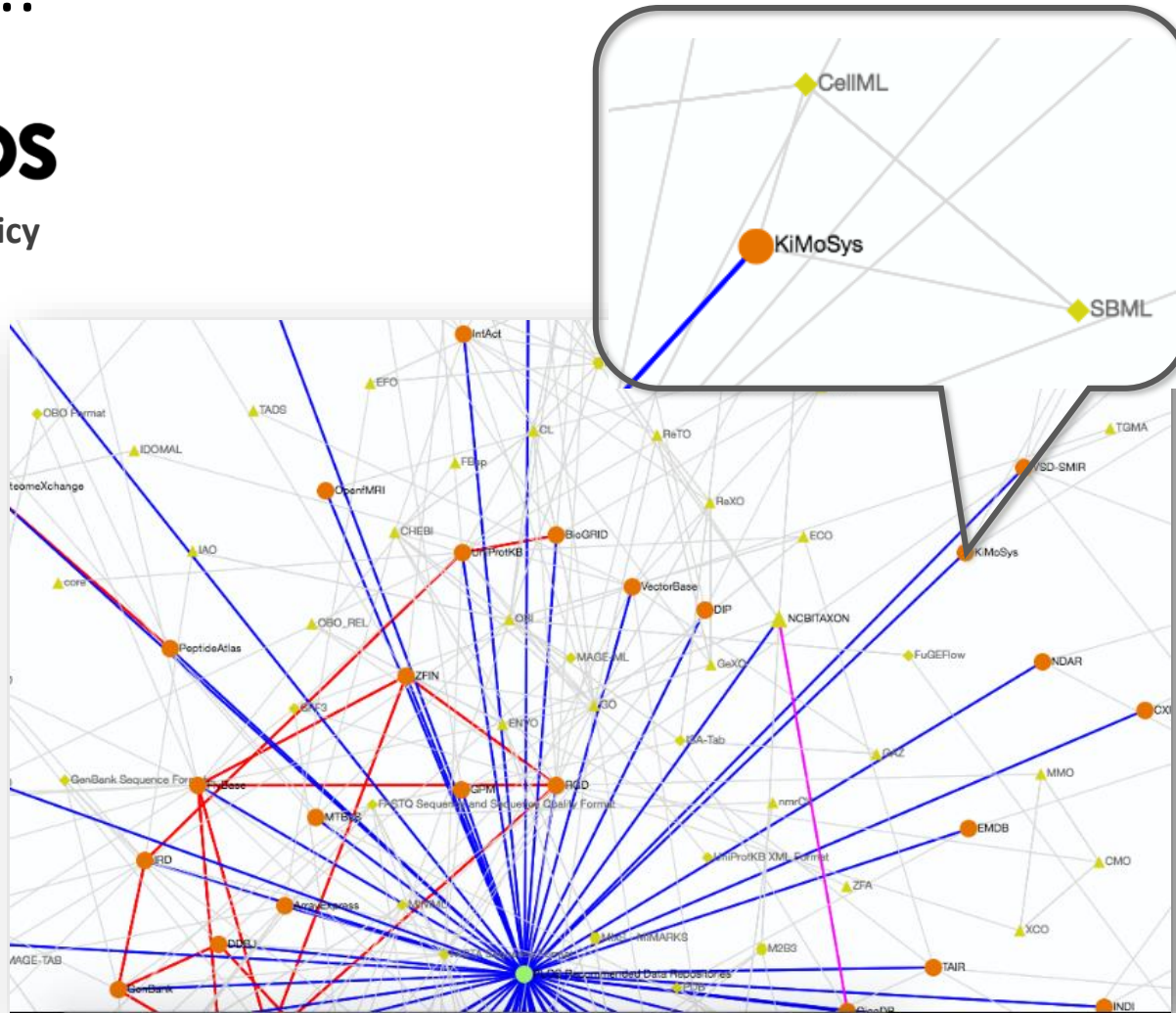
[Publications](#) 2

[In Collections](#) 1

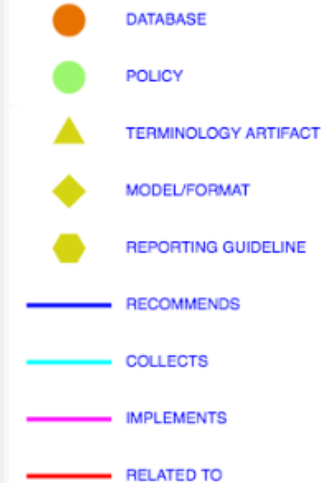
[Recommended](#) 2

1 Taxa types, including:

Visualizing the relationships between data...



Legend



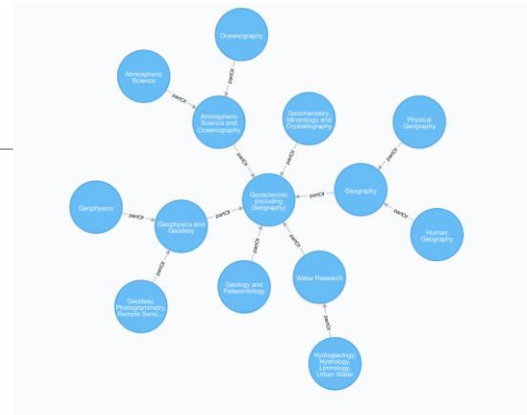
Making data discoverable

FAIRsharing widget

FAIRsharing API

Open ontologies

Bioschemas.org markup



```

1 {
2   "@context": "http://schema.org",
3   "@type": "DataCatalog",
4   "identifier": "https://identifiers.org/MIR:00100800",
5   "name": "FAIRsharing.org",
6   "description": "A manually curated, informative and educational resource
on data and metadata standards, inter-related to databases/data repositories
and funder and journal publisher data policies from across disciplines.
FAIRsharing is an ELIXIR-UK node resource and has an active role in the RDA
and Forcell data initiatives.",
7   "url": "https://fairsharing.org/",
8   "dataset": [
9     {
10      "@type": "Dataset",
11      "name": "Metadata Standard",
12      "description": "A manually curated registry of standards, split
into three types - Terminology Artifacts (ontologies, e.g. Gene Ontology),
Models and Formats (conceptual schema, formats, data models, e.g. FASTA),
and Reporting Guidelines (e.g. the ARRIVE guidelines for in vivo animal
testing). These are linked to the databases that implement them and the
funder and journal publisher data policies that recommend or endorse their
use.",
13      "identifier": "https://www.fairsharing.org/bsq-sXXXXXX",
14      "url": "https://fairsharing.org/standards",
15      "keywords": "Standards, Metadata, Formats, Ontologies,
Terminology Artifacts, Reporting Guidelines",
16      "includedInDataCatalog": "https://www.fairsharing.org"
17    },
18    {

```

Endorsements/ Adopters

Societies and organisations, incl:



Cross-links, d



Institutional RDM services, incl:



mes:



Journal, publishers, incl:



Standard developing groups, incl:



Impact of the Deliverable

Before

- Researchers didn't know which databases to use to source and deposit data
- Standard developers struggled to know what is already out there
- Database developers reinvented standards and procedures
- Journal publishers didn't know which databases and standards to recommend or endorse to their authors/fundees

After

- FAIRsharing maps the landscape of standards, repositories/databases and data policies across domain
- Manually curated, with the community – information is accurate and timely
- FAIRsharing raises the visibility of standards and databases so reducing reinvention and contributing to FAIR data and digital objects
- FAIRsharing works directly with journal publishers to inform and direct their data policies
- FAIRsharing helps make data FAIR

The FAIRsharing team

Our Advisory Board



Susanna-Assunta Sansone
Principal Investigator and Associate Director



Philippe Rocca-Serra
Senior Research Lecturer



Alejandra Gonzalez-Beltran
Research Lecturer



Emma Ganley (PLOS) Co-chair
Varsha Khodiyar (NPG) Co-chair
Michael Ball (BBSRC)



Rafael Jimenez (ELIXIR)
Amye Kenall (BMC)
Rebecca Lawrence (F1000), Co-Chair of the Force11 WG

Database, standard maintainer,
policy maker? Work with us!

contact@fairsharing.org

peter.mcquilton@oerc.ox.ac.uk

Is Lemberger (EMBO Press)
er Lin (CrossRef)
lla Rustici (ELIXIR UK Node Industry
g Sector Lead, University of Cambridge,

Subirats (Information Management Officer,
the United Nations, Italy)

Teperek (Data Stewardship Coordinator,

TUDEft, The Netherlands)

Michael Witt (re3data; Purdue University Libraries)



Massimiliano Izzo
Research Software Engineer



Eamonn Maguire
Web Design Contractor



Allyson Lister
Knowledge Engineer



Delphine Dauga
Curator contractor



Melanie Adekale
Curator contractor

Science Initiative, Associate Director for
Programme Development at the NIH National
Library of Medicine)