

FAIRsharing.org

Promoting the discovery of data standards, policies and
databases across all research domains

Allyson Lister, PhD

Content and Community Coordinator

ORCID: 0000-0002-7702-4495 | Twitter: @allysonlister

Try our new release!
beta.fairsharing.org



datareadiness.eng.ox.ac.uk



contact@fairsharing.org



[@FAIRsharing_org](https://twitter.com/FAIRsharing_org)



[10.25504/FAIRsharing.2abjs5](https://doi.org/10.25504/FAIRsharing.2abjs5)



[FAIRsharing.org](https://fairsharing.org)  Educational

CC BY-SA 4.0 International

FAIRsharing in a nutshell:

scope and mission

The AgBioData RCN

What's an RCN?

Research Coordination Network -
Funded by National Science Foundation

... supporting groups of investigators to **communicate and coordinate** their research, training and educational activities across disciplinary, organizational, geographic and international boundaries.

Overarching RCN Goal:

Make agricultural genetics, genomics and breeding (GGB) data easier to find and use, and available for the long term

When: Sept 1 2021 - Aug 31 2024 (3 years)

Funding: includes a full time Program Coordinator, in-person meeting in year 2, sustainability planning

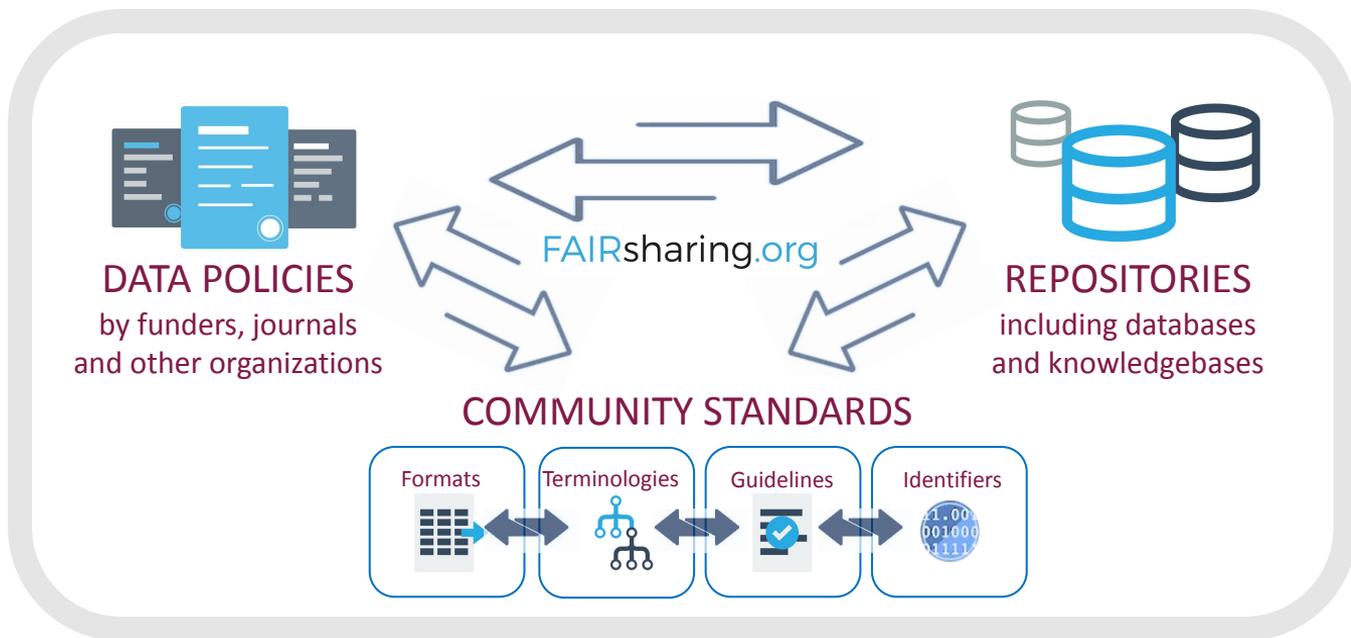


12



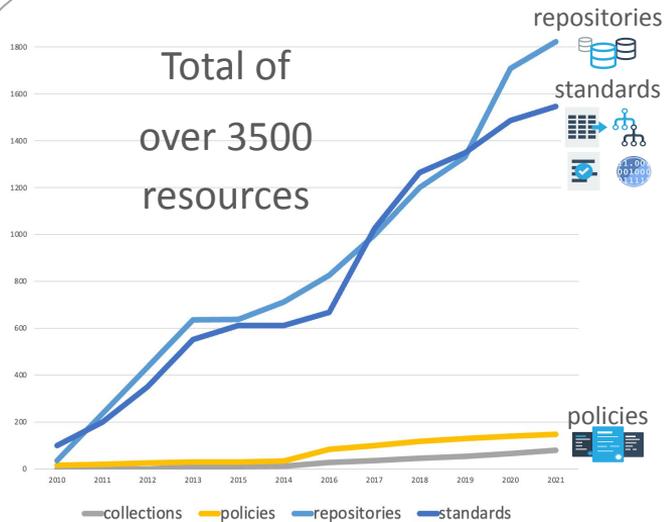
Eva
Huala

A global resource for all disciplines



An informative and educational resource of three *curated* and *interlinked* registries of standards, repositories and policies; three key elements of the FAIR ecosystem

Promoting the value of repositories, standards, policies



Fosters a culture change where the use of these resources for FAIRer data is pervasive and seamless

Researchers in academia, industry, government

Journal publishers or organizations with data policy

Learned societies, unions and associations



Developers and curators of resources

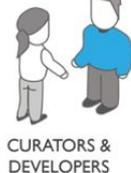
Research data facilitators, librarians, trainers

Funders and data policy makers

Guides **consumers** to *discover, select and use* these resources with confidence
Helps **producers** to make their resources more *visible, more widely adopted and cited*

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 FAIRness of FAIRsharing: increasing your...

Findability

Discovery of your resource via

- standard search
- traversal of our relationship graph

Your stakeholders can use FAIRsharing to find your resource through multiple routes, both directly and via related records

Interoperability

Navigate relationships to see which resources

- utilise the standards you implement
- retrieve data from you

This information can advertise e.g. vocabularies you implement and help identify possible collaborations

Accessibility

- FAIRsharing summarises access methods and protocols
- Users can see your resource metadata, even if your service is unavailable

Reusability

- Showcase your community certifications
- Provide licencing information

A FAIRsharing record: at-a-glance view

Records have >40 descriptors

GENERAL INFORMATION

 **The Arabidopsis Information Resource (TAIR)**

 [10.25504/FAIRsharing.4dqw0](https://doi.org/10.25504/FAIRsharing.4dqw0)

Type Knowledgebase and repository

Registry Database

Description The Arabidopsis Information Resource (TAIR) maintains a database of genetic and molecular biology data for the model higher plant *Arabidopsis thaliana*. Data available from TAIR includes the complete genome sequence along with gene structure, gene product information, gene expression, DNA and seed stocks, genome maps, genetic and physical markers, publications, and information about the Arabidopsis research community. Gene product function data is updated every week from the latest published research literature and community data submissions.

Homepage <http://www.arabidopsis.org/>

Year of Creation 2000

Maintainers [taircurator](#)

Countries developing this resource [United States](#)

Subjects [Metabolomics](#) [Life Science](#)

Domains [Molecular Structure](#) [Genetic Map](#) [Physical Map](#) [Genome Map](#) [Expression Data](#)
[DNA Sequence Data](#) [Publication](#) [Structure](#) [Gene](#)

Taxonomic Range [Arabidopsis Thaliana](#)

User Defined Tags None

[VIEW RELATION GRAPH](#)

 **How to cite this record**

FAIRsharing.org: TAIR; The Arabidopsis Information Resource
Resource DOI: [10.25504/FAIRsharing.4dqw0](https://doi.org/10.25504/FAIRsharing.4dqw0), Last

Including:

- Description
- Country
- Date of creation
- Subject(s) coverage
- Maintainer(s)
- Funder(s)
- Organization(s)
- Support details
- Access
- Licence(s)
- Publication(s)
- Related records (standards and databases, and policies)

DOI: <https://doi.org/10.25504/FAIRsharing.4dqw0>

Records are citable; maintainers identified and credited

GENERAL INFORMATION

 **The Arabidopsis Information Resource (TAIR)**

 10.25504/FAIRsharing.4dqw0

Type Knowledgebase and re...

Registry Database

Description The Arabidopsis Inform... data for the model high... genome sequence alo... seed stocks, genome... Arabidopsis research... published research lite...

Homepage <http://www.arabidopsis.org/>

Year of Creation 2000

Maintainers [taircurator](#)

Countries developing this resource [United States](#)

Subjects [Metabolomics](#) [Life Science](#)

Domains [Molecular Structure](#) [Genetic Map](#) [Physical Map](#) [Genome Map](#)
[DNA Sequence Data](#) [Publication](#) [Structure](#) [Gene](#)

Taxonomic Range [Arabidopsis Thaliana](#)

User Defined Tags None

[VIEW RELATION GRAPH](#)

 **How to cite this record**

FAIRsharing.org: TAIR; The Arabidopsis Information Resource, DOI: 10.25504/FAIRsharing.4dqw0, Last

 **How to cite this record**

FAIRsharing.org: TAIR; The Arabidopsis Information Resource, DOI: 10.25504/FAIRsharing.4dqw0, Last
Accessed: Wednesday, September 29th 2021, 11:37

Homepage <http://www.arabidopsis.org/>

Record created at Tuesday, November 4th 2014, 15:23 |  **Record**

Year of Creation 2000

Homepage <https://www.agbiodata.org/>

Reference URL <https://www.agbiodata.org/databases>

Year of Creation None found

Maintainers [CampbellJD](#)  ID, [bcondon](#)  ID

Relationships among repositories and standards are tracked

GENERAL INFORMATION

 **The Arabidopsis Information Resource (TAIR)**

 10.25504/FAIRsharing.4dqw0

Type Knowledgebase and repository

Registry Database

Description The Arabidopsis Information Resource (TAIR) maintains data for the model higher plant Arabidopsis thaliana, genome sequence along with gene structure, gene expression data, seed stocks, genome maps, genetic and physical maps, and other Arabidopsis research community. Gene product functions, published research literature and community data subsets.

Homepage <http://www.arabidopsis.org/>

Year of Creation 2000

Maintainers [taicurator](#)

Countries developing this resource [United States](#)

Subjects [Metabolomics](#) [Life Science](#)

Domains [Molecular Structure](#) [Genetic Map](#) [Physical Map](#) [DNA Sequence Data](#) [Publication](#) [Structure](#) [Genetics](#)

Taxonomic Range [Arabidopsis Thaliana](#)

User Defined Tags None

[VIEW RELATION GRAPH](#)

How to cite this record

FAIRsharing.org: TAIR; The Arabidopsis Information Resource. DOI: 10.25504/FAIRsharing.4dqw0. Last

RELATED CONTENT

Search through related databases

RELATED STANDARDS (8) **RELATED DATABASES (5)**

 **PlantRNA**
The Arabidopsis Information Resource *related*

 **Planteome**
The Arabidopsis Information Resource *related*

 **ThaleMine**
The Arabidopsis Information Resource *related*

RELATED CONTENT

Search through related standards

RELATED STANDARDS (8) **RELATED DATABASES (5)**

 **Sequence Ontology**
The Arabidopsis Information Resource *implements* Sequence Ontology

 **Evidence and Conclusion Ontology**
The Arabidopsis Information Resource *implements* Evidence and Conclusion Ontology

 **Gene Ontology**
The Arabidopsis Information Resource *implements* Gene Ontology

DOI: <https://doi.org/10.25504/FAIRsharing.4dqw0>

Recommendations by data policies are highlighted

GENERAL INFORMATION

The Arabidopsis Information Resource
DOI: 10.25504/FAIRsharing.4dqw0

Type: Knowledgebase
Registry: Database
Description: The Arabidopsis data for the model genome sequence, seed stocks, gene Arabidopsis research published research
Homepage: <http://www.arabidopsis.org>
Year of Creation: 2000
Maintainers: [taircurator](#)
Countries developing this resource: [United States](#)
Subjects: [Metabolomics](#)
Domains: [Molecular Structure](#), [DNA Sequence Data](#)
Taxonomic Range: [Arabidopsis Thaliana](#)
User Defined Tags: None

[VIEW RELATION GRAPH](#)

How to cite this record
FAIRsharing.org: TAIR; The Arabidopsis Information Resource. DOI: 10.25504/FAIRsharing.4dqw0

COLLECTIONS & RECOMMENDATIONS

Search through in policies

IN POLICIES (6) **IN COLLECTIONS (1)**

-  **PLOS Recommended Data Repositories**
PLOS Recommended Data Repositories **recommends** The Arabidopsis Information Resource
-  **Elsevier's Data Policy**
Elsevier's Data Policy **recommends** The Arabidopsis Information Resource
-  **The Plant Cell - Instructions for Authors**
The Plant Cell - Instructions for Authors **recommends** The Arabidopsis Information Resource

Evolution is monitored and reasons provided

Clear All Ontology Subjects: [agriculture](#) Output status: [deprecated](#)

GENERAL INFORMATION

This record was deprecated on 2021-9-17 for the following reason(s): AgBase became Host Pathogen Interaction Database, and therefore this record has been deprecated.

This record is replaced by: [Host Pathogen Interaction Database](#)

 **AgBase (AgBase)**
[doi](#) [10.25504/FAIRsharing.kn4ycg](https://doi.org/10.25504/FAIRsharing.kn4ycg)

Type Knowledgebase

Registry Database

Description AgBase is a curated, open-source, Web-accessible resource for functional analysis of agricultural plant and animal gene products.

FAIRsharing.kn4ycg

Standards Implemented: 0
Related Databases: 0
Endorsing Policies: 0

Standards Implemented: 0
Related Databases: 0
Endorsing Policies: 0

Standards Implemented: 0
Related Databases: 0
Endorsing Policies: 0

GRO_CPD
Cereal Plant Development

AgMES
Agricultural Metadata Element Set

AgBase
AgBase

FAIRsharing.org  Educational
CC BY-SA 4.0 International

i This record is awaiting review by FAIRsharing curators

ACTIONS ▾

GENERAL INFORMATION

 **Data and Service Center for the Humanities (DaSCH)**

 Awaiting DOI

Type	Repository
Registry	Database

In-house curation and the 'life cycle status' tags

Life cycle tags:

-  Ready for use, implementation, or recommendation
-  In development
-  Status uncertain
-  Deprecated as subsumed or superseded

- Every edit is checked by the in-house curation team
- Maintainers are notified via email when their record is updated

The value of tracking
the **relationships** between
repositories, standards and policies

Emerging data types Working Group

- ▶ Novel data types often results from novel scientific methods
- ▶ The goal of this working group is
 - ▶ To select the most relevant new data types for AgBioData members
 - ▶ Identify the specific challenges to make this data type FAIR
 - ▶ Pass results onto other working groups (eg. metadata, data federation) in following years to develop best practices
 - ▶ We could form a new issue WG after 6 month
- ▶ Example new data types
 - ▶ High-throughput phenotyping data
 - ▶ Metabolomic data
 - ▶ Image data
 - ▶ Microbiome data



27



Sook
Jung

GENERAL INFORMATION



AgBioData
Awaiting DOI

Type Collection

Registry Collection

Description AgBioData is a consortium of agricultural... together to ensure standards and best pr... genetic and breeding data.

Homepage <https://www.agbiodata.org/>

Reference URL <https://www.agbiodata.org/databases>

Year of Creation None found

Maintainers [CampbellJD](#) [bcondon](#)

Countries developing this resource None found

Subjects [Animal Husbandry](#) [Genomics](#) [Plant Ge...](#)

Domains None

Taxonomic Range None

User Defined Tags [Metadata Standardization](#)

[VIEW RELATION GRAPH](#)

URL: <https://beta.fairsharing.org/AgBioData>

MATCH ALL TERMS MATCH ANY TERM

MAINTAINED NOT MAINTAINED

RECOMMENDED NOT RECOMMENDED

Displaying 1 to 30 of 30.

Select all the resources that have been recommended by a data policy from a journal, journal publisher or funder

Variant Call Format

Variant Call Format (VCF) is a text file format (most likely stored in a compressed manner). It contains meta-information lines, a header line, and then data lines ea...

Genomics Sequence ... All

+5 more tags

- Related Standards: 0
- Implementing Databases: 17
- Endorsing Policies: 0

FASTA

FASTA format is a text-based format for representing either nucleotide sequences or peptide sequences, in which nucleotides or amino acids are represented usi...

Life Science Nucleic aci... All

+3 more tags

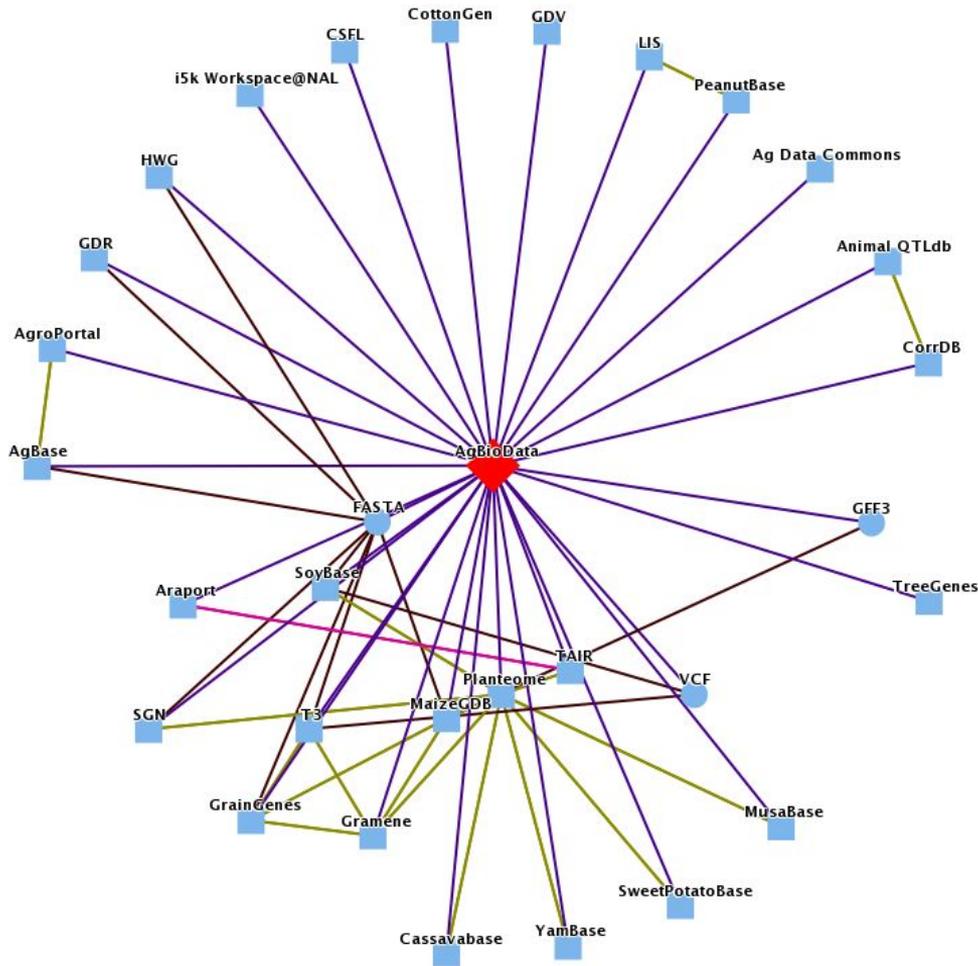
- Related Standards: 4
- Implementing Databases: 266
- Endorsing Policies: 0

GFF3

Generic Feature Format Version 3

TAIR

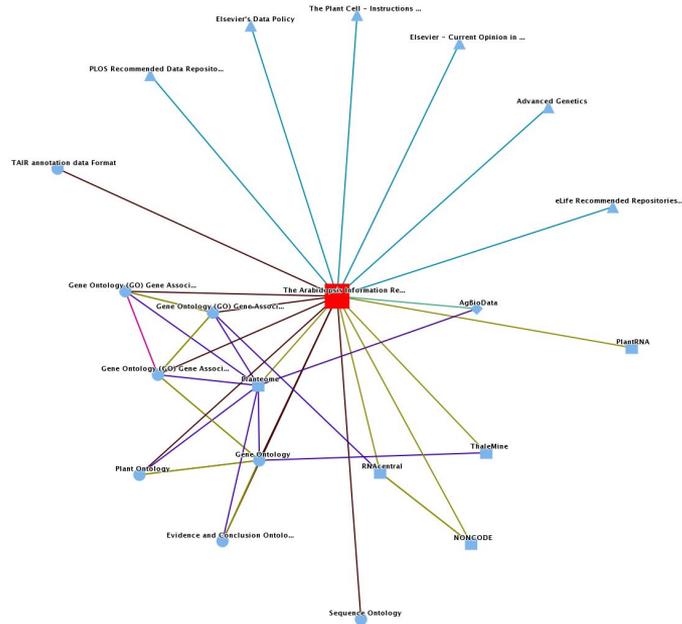
The Arabidopsis Information Resource



AgBioData Graph

- What does it show you?
 - How your resource is related to other resources within the collection
 - Traversal to other, related, graphs
- How can it be improved?
 - Add all connections your resource has to other data standards, databases and policies
 - Add connections to each other (do you share data? Use a common format?)

FAIRsharing Network Graph
The Arabidopsis Information Resource
(Database/Knowledgebase And Repository)

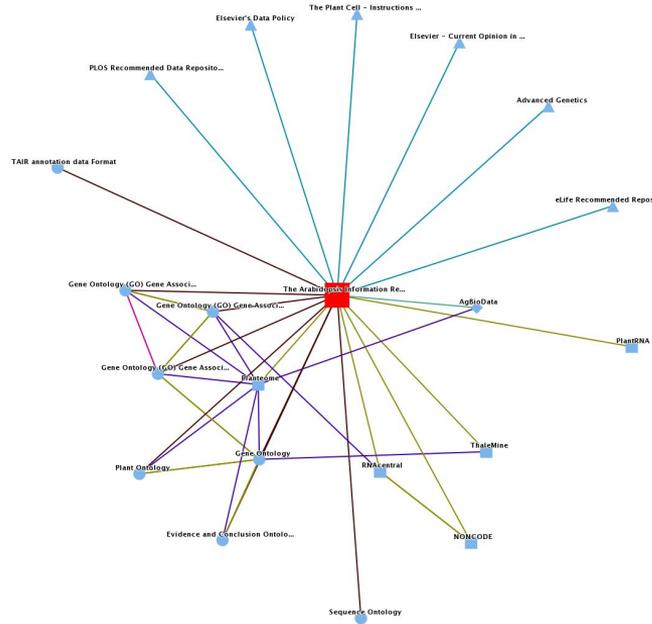


Relationship graphs

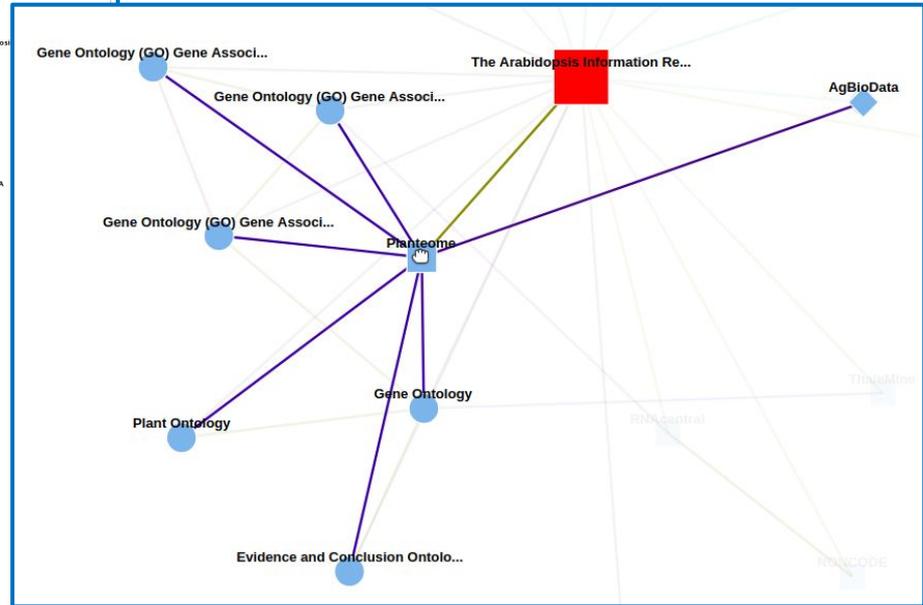
Shows other records related to TAIR:

- standards TAIR implements,
- repositories TAIR has a relationship with (e.g. share data with, import data from)
- policies that recommend TAIR

FAIRsharing Network Graph
The Arabidopsis Information Resource
(Database/Knowledgebase And Repository)



Relationship graphs: how resources in your network are connected



Examples of how *consumers* and *producers*
of standards, repositories and policies
use FAIRsharing and **benefit** from it

Metadata Working Group

- ▶ Long-standing WG that has explored existing metadata standards for a variety of data types.
- ▶ Mission is to improve use of standardized metadata, which is critical especially for finding and relevant datasets and sharing and reuse of data.
- ▶ Proposed activities:
 - ▶ Document recommended metadata standards for member databases.
 - ▶ Develop a metadata assessment tool.
 - ▶ Work with member databases to document metadata practices.



Ethy
Cannon



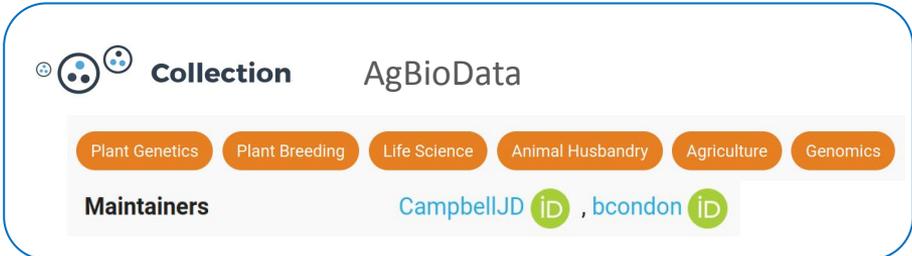
25

Organizations: tailored views for education and promotion

Many groups and projects have created
Collections

These are are branded pages that group
selected **standards** and/or **repositories**
relevant to their community and/or
developed by them

EXAMPLES

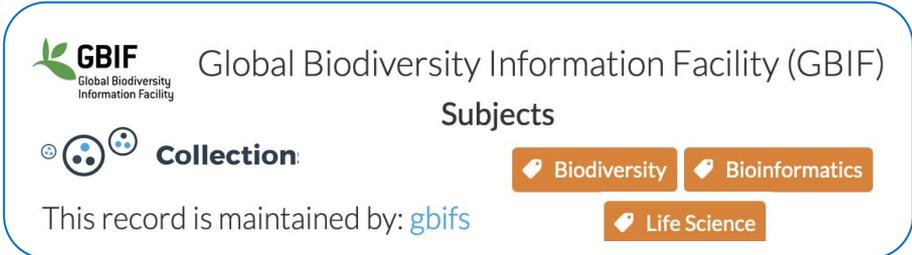


Collection AgBioData

Plant Genetics Plant Breeding Life Science Animal Husbandry Agriculture Genomics

Maintainers CampbellJD , bcondon 

URL: beta.fairsharing.org/AgBioData



 Global Biodiversity Information Facility (GBIF)

Collection:  

Subjects Biodiversity Bioinformatics Life Science

This record is maintained by: [gbifs](#)

URL: beta.fairsharing.org/GBIFPortals

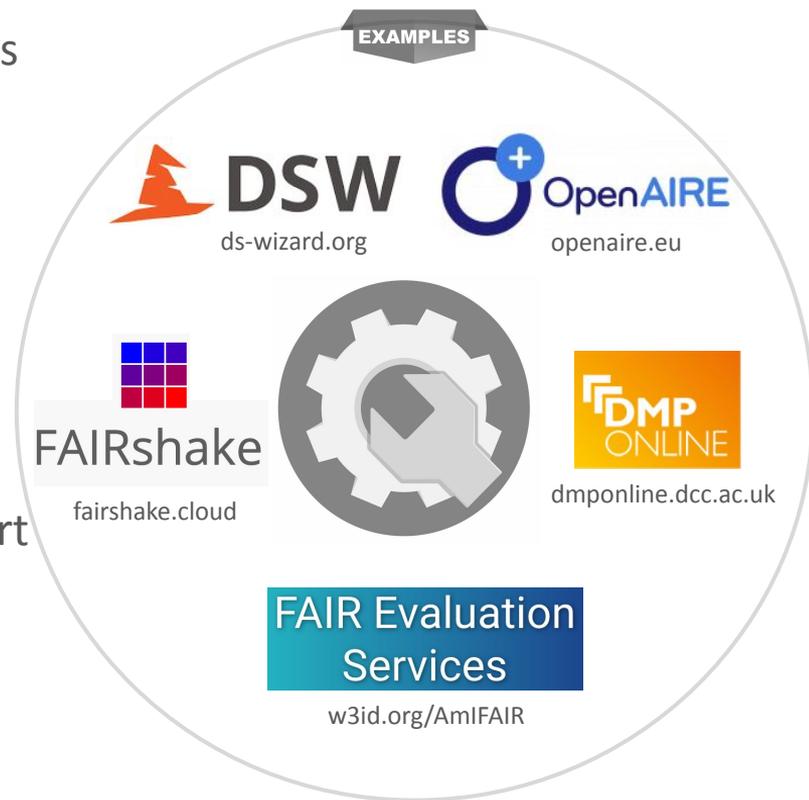
Services: curated content to power 3rd party tools

A growing number of **FAIR-enabling Services** access FAIRsharing API, and use it as **look-up** and **select** service for **standards** and **repositories** for:

- data management plans and guidelines
- FAIR assessment

A new, in-development **Data Discovery Service**, part of ELIXIR-driven EOSC-Life and BY-COVID projects:

- register repositories' access methodologies, e.g. schema.org, bioschemas, OAI-PMH
- enable (meta)data harvesting



FAIR in practice:

FAIRsharing implementation

and collaborations

Implementation Working Group

- ▶ Starts during Q1 of Year 2, with AgBioData members, stakeholders, data generators, agricultural scholarly publishers, major federal & private funding agencies
- ▶ Working group will implement community-wide Aim 1 deliverables and develop effective mechanisms
 - ▶ Develop effective mechanisms and strategies to implement the FAIR data management
 - ▶ Publish an initial set of recommendations/roadmap for funders and journals to support and encourage FAIR data compliance.

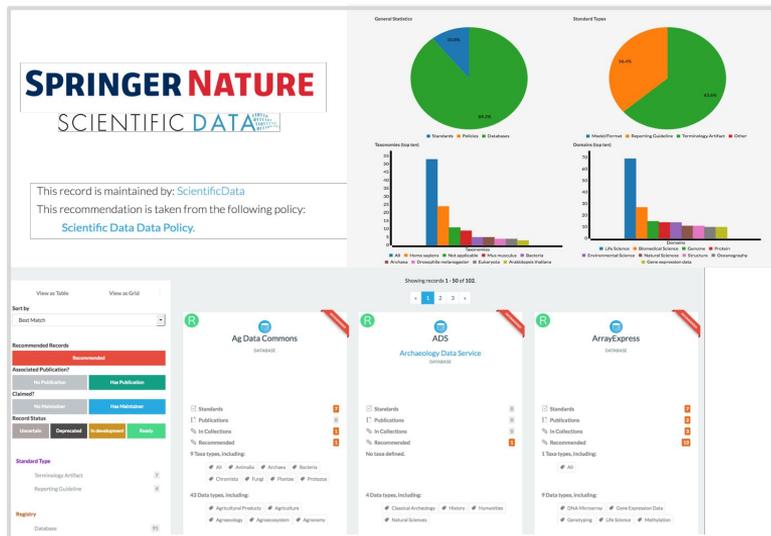


29



Jacqueline
Campbell

Policy makers: policies registration and dashboards



EXAMPLES

Scholarly publishers use FAIRsharing to:

- **list standards** and **repositories** they recommend in their **data policy**,
- **discover** new resources and **monitor** their evolution to **refine** or **update** their policies



The FAIRness of FAIRsharing: nuts and bolts

Findability



Sitemap.xml, JSON

Schema.org

Markup with Schema.org for search indexes



DOI unique persistent identifiers for each record



ORCID for authors

Interoperability



JSON markup



Standardized semantics



Cross-links to records in other registries (*ongoing*)



ROR for organizations (*ongoing*)



FundRef for funders (*ongoing*)

Accessibility



REST Web Services

Reusability



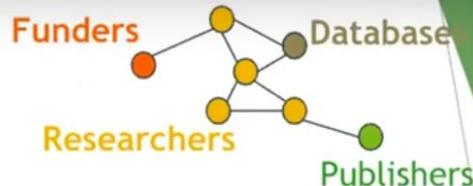
CC BY 4.0 license

Users, adopters, communities and working groups



Help us help you!

Recruitment Working Group



- ▶ **Aim:** Expand the network to include key stakeholders
- ▶ **Goal:** Broadening participation in the AgBioData network
- ▶ Focus efforts to recruit & engage participants
 - ▶ Underrepresented groups (NIFA Tribal Colleges)
 - ▶ Junior researchers (graduate students/postdocs)
- ▶ Identify relevant groups/organizations, establishing contacts, & gathering/reporting on metrics
- ▶ WG will meet virtually & at the beginning of each year
 - ▶ To recruit data generating scientists &/or stakeholders for each working group (Program Coordinator will play a role in this work)
 - ▶ End of year 2, recruit people from agricultural scholarly publishers, major federal & private funding agencies to join the **Implementation Working Group**



Jacqueline Campbell



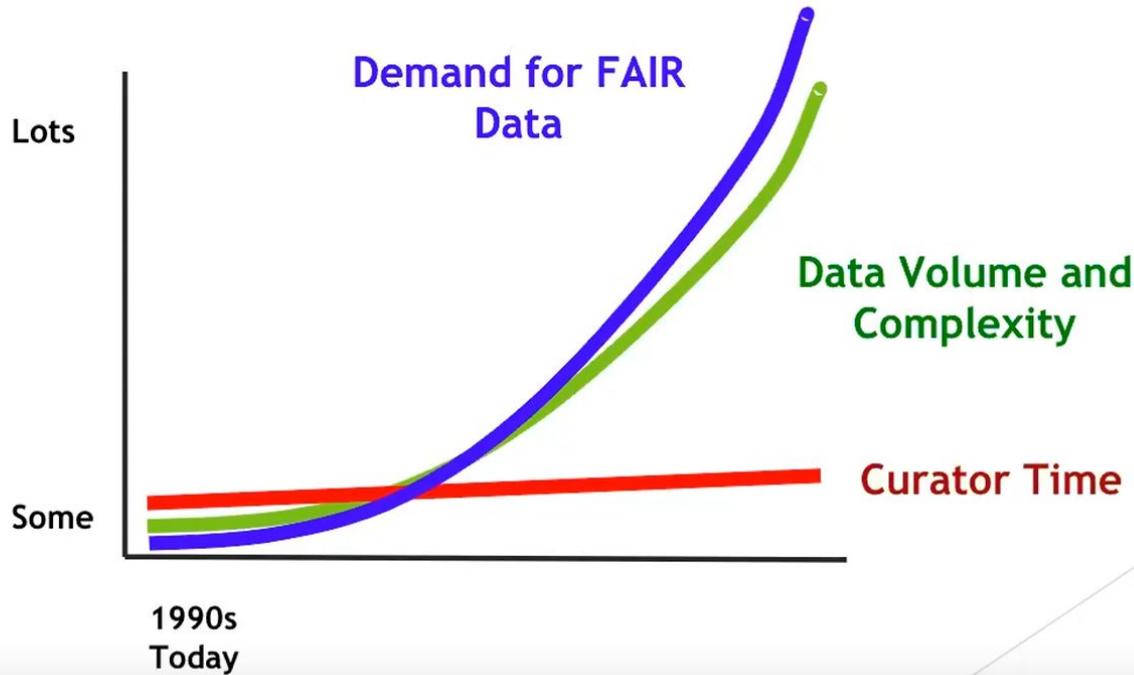
28

Help us to help you promote your repository

- You are the expert, producer, owner or maintainer
 - Use your ORCID to be recognised and credited
- You know more about the key features and descriptors
 - Identify the community standards you use, key to FAIRness
 - Describe the relationships, if any, with other repositories
- Your descriptive record in FAIRsharing informs user decisions
 - Add rich metadata to help with visibility and user selection
 - Indicate if you have any global certifications, or community badges of adoption and recommendation

What is AgBioData?

Genetic, Genomic and Breeding Data



Lisa Harper



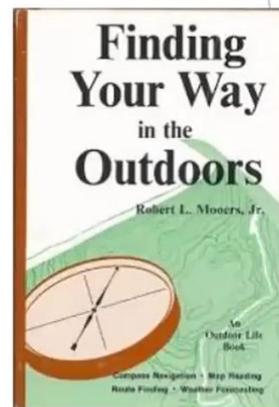
Community Curators: in development

- Do you have a broad and deep understanding of the standards and databases most relevant in your field?
- Be a recognized contributor of the FAIRsharing community
 - How would you like to be recognised?
- Help your community
 - Record maintainers only describe those resources they are directly involved with
 - Community curators can manage the entire ecosystem of records within your domain of interest
- Access all records in your discipline and create new ones
 - Work with existing maintainers
 - Suggest new maintainers for the internal FAIRsharing Curation team to contact
- Gain on-the-job curation expertise
- Engagement and networking with a community of like-minded people
- Become an expert on data and metadata standards, databases, repositories and data policies in your area

Outreach and related projects

Education Working Group

- ▶ FAIR Data Management Guide
 - ▶ Researcher-focused
 - ▶ What GGB databases are out there? Which should I use?
 - ▶ Where and how should I submit my data of type X?
 - ▶ Checklist of FAIR data procedures
 - ▶ Start general, then add specifics by data type as time allows
- ▶ FAIR data curriculum
 - ▶ Foundational knowledge of FAIR and GGBs
 - ▶ Graduate student courses
 - ▶ Workshops at Conferences
 - ▶ Online, at your own pace
- ▶ FAIR data summer workshop



30



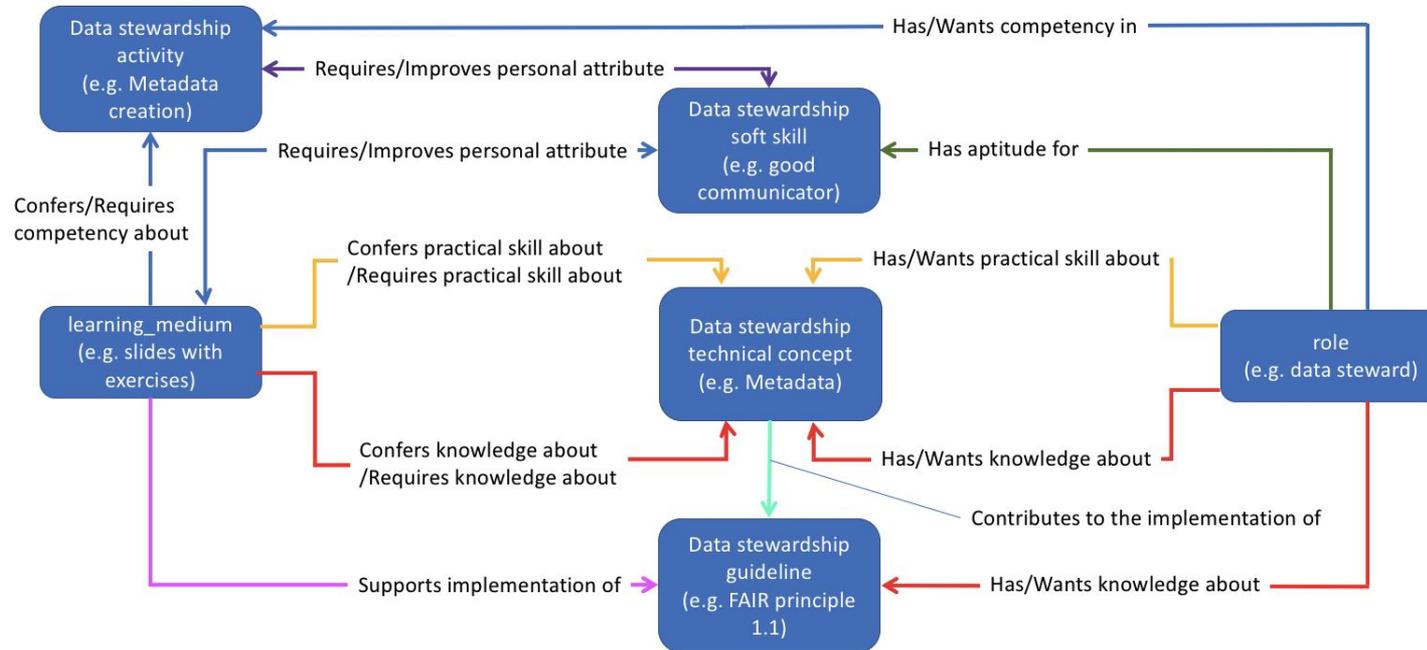
Meg
Staton

Plans for FAIRassist: FAIRsharing educational component

- It will *explain* and *guide* stakeholders to make the best use of functionalities and content offered by FAIRsharing
- It will serve both types of users:
 - **Consumers** (all users) of resources, from one or more **registries**
 - **Producers** (some users) of one or more (types of) **resource**
- It will have two elements:
 - **Educational Guidance** - *visual storytelling via infographics*
 - Targeted advice that explains how and when FAIRsharing helps them
 - **Resources Finder** - *wizard*
 - Decision tree, developed in phases, that guides users to discover, filter, select a subset of resources

terms4FAIRskills

Competencies, skills and knowledge associated with making and keeping data FAIR



FAIR cookbook



Hands-on recipes to make
and keep data FAIR

What is it?

An online, 'live' resource
for the **life sciences**

A collection of **recipes**
that cover the operation
steps of FAIR data
management

Who is it for?

Anyone who needs
practical assistance in
their FAIRification
journey or creates FAIR
guidance and
educational material

Who developed it?

+50 researchers and **data
management**

professionals in the life
sciences, from **academia**
and **industry**

Including ELIXIR members

<https://fairplus.github.io/the-fair-cookbook/content/home.html>

Operational Team

- **Allyson Lister**, Content and Community Lead
- **Milo Thurston**, Technical Lead
- **Ramon Granell**, Data Enrichment & Quality Manager
- **Delphine Dauga**, Data Curator Manager
- **Hossein Mirian**, Web Developer
- **Dominique Batista**, Research Software Engineer
- **Massimiliano Izzo**, Research Software Engineer
- **Philippe Rocca-Serra**, Co-Founder
- **Susanna-Assunta Sansone**, PI and Founder

Executive Advisors

- **Varsha Khodiyar**, HDRUK
- **David Carr**, independent expert
- **Chris Graf**, Springer Nature
- **Marta Tereperek**, Data Stewardship Coordinator, TUDelft
- **Robert Hanisch**, Director, NIST Office of Data & Informatics
- **Peter McQuilton**, GSK

Stakeholder Advisors

- **Emma Ganley**, Director of Strategic Initiatives, Protocols.io
- **Michael Ball**, Biotechnology and Biological Sciences Research Council (BBSRC)
- **Theo Bloom**, British Medical Journal (BMJ)
- **Nick Everitt and Matthew Cannon**, Taylor and Francis
- **Wei-Mun Chan**, eLife
- **Geraldine Clement-Stoneham**, Head of Knowledge Management and Scholarly Communication, Medical Research Council (MRC)
- **Helena Cousijn**, DataCite
- **Scott Edmunds**, GigaScience, Oxford University Press
- **Dominic Fripp**, JISC, UK
- **Simon Hodson**, CODATA
- **Iain Hrynaszkiewicz**, PLoS
- **Mike Huerta**, Coordinator of Data & and Open Science Initiative, Associate Director for Programme Development at the NIH National Library of Medicine
- **Amye Kenall**, VP of Publishing and Product, Research Square
- **Adam Leary**, Oxford University Press
- **Thomas Lemberger**, EMBO Press
- **Luiz Olavo Bonino**, GO-FAIR
- **Kiera McNiece**, Cambridge University Press
- **Dagmar Meyer**, European Research Council (ERC), Executive Agency
- **Marina Soares E Silva, Ilaria Carnevale and Sarah Callaghan**, Elsevier
- **Imma Subirats**, Information Management Officer, FAO of the United Nations
- **Molly Cranston and Guillaume Wright**, F1000Research
- **Kathryn Sharples**, Wiley
- **Catriona MacCallum**, Hindawi

FAIRsharing.org

Thank you!

Homepage: <https://fairsharing.org>

Beta site: <https://beta.fairsharing.org>

Contact us: contact@fairsharing.org

Twitter: [@fairsharing_org](https://twitter.com/fairsharing_org)