

Joint Session IG Metadata, WG Metadata Standards Catalog, IG Data in Context - RDA 13th Plenary Meeting

<https://rd-alliance.org/joint-session-ig-metadata-wg-metadata-standards-catalog-ig-data-context-rda-13th-plenary-meeting>

Collaborative session notes: <https://docs.google.com/document/d/1L9SSdtug7JDBFJZF4Bsgeuv40E8wvQCxMqWm0yE58YI/edit?usp=sharing>

April 4 2019

M4M Workshops

For (faster) FAIRification

Erik Schultes, PhD
International Science Coordinator
GO FAIR International Support and Coordination Office
erik.schultes@go-fair.org
go-fair.org



We (co-chairs of MIG) are inviting various groups with an **interest in metadata** (and particularly **criteria for good metadata and metadata standards**) to give a short presentation so that we can focus **all** the metadata activity in RDA.

References for M4M Workshops

- **M4M Workshops** <https://www.go-fair.org/resources/go-fair-workshop-series/metadata-for-machines-workshops/>
- **On OSF** <https://osf.io/qe9fa/>
- **On Arxiv** <https://arxiv.org/abs/1902.11162>
- **Guidelines document is in preparation**
 - Wittenburg, Koskela, Schultes

Acknowledgements

- Peter Wittenburg
- Rebecca Koskela
- Keith Jeffery

FAIR Principles: for machine F, A, I, and R

Sci. Data 3:160018 doi: 10.1038/sdata.2016.18 (2016)

Findable:

F1 (meta)data are assigned a globally unique and persistent identifier;

F2 data are described with rich metadata;

F3 metadata clearly and explicitly include the identifier of the data it describes;

F4 (meta)data are registered or indexed in a searchable resource;

Interoperable:

I1 (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.

I2 (meta)data use vocabularies that follow FAIR principles;

I3 (meta)data include qualified references to other (meta)data;

Accessible:

A1 (meta)data are retrievable by their identifier using a standardized communications protocol;

A1.1 the protocol is open, free, and universally implementable;

A1.2 the protocol allows for an authentication and authorization procedure, where necessary;

A2 metadata are accessible, even when the data are no longer available;

Reusable:

R1 meta(data) are richly described with a plurality of accurate and relevant attributes;

R1.1 (meta)data are released with a clear and accessible data usage license;

R1.2 (meta)data are associated with detailed provenance;

R1.3 (meta)data meet domain-relevant community standards;

FAIR Principles: for machine F, A, I, and R

Sci. Data 3:160018 doi: 10.1038/sdata.2016.18 (2016)

Findable:

F1 (**meta**)data are assigned a globally unique and persistent identifier;

F2 data are described with rich **metadata**;

F3 **metadata** clearly and explicitly include the identifier of the data it describes;

F4 (**meta**)data are registered or indexed in a searchable resource;

Interoperable:

I1 (**meta**)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.

I2 (**meta**)data use vocabularies that follow FAIR principles;

I3 (**meta**)data include qualified references to other (**meta**)data;

Accessible:

A1 (**meta**)data are retrievable by their identifier using a standardized communications protocol;

A1.1 the protocol is open, free, and universally implementable;

A1.2 the protocol allows for an authentication and authorization procedure, where necessary;

A2 **metadata** are accessible, even when the data are no longer available;

Reusable:

R1 **meta**(data) are richly described with a plurality of accurate and relevant attributes;

R1.1 (**meta**)data are released with a clear and accessible data usage license;

R1.2 (**meta**)data are associated with detailed provenance;

R1.3 (**meta**)data meet domain-relevant community standards;

FAIR Principles: for machine F, A, I, and R

Sci. Data 3:160018 doi: 10.1038/sdata.2016.18 (2016)

Findable:

F1 (**meta**)data are assigned a globally unique and persistent identifier;

F2 data are described with rich metadata;

F3 **metadata** clearly and explicitly include the identifier of the data it describes;

F4 (**meta**)data are registered or indexed in a searchable resource;

Interoperable:

I1 (**meta**)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.

I2 (**meta**)data use vocabularies that follow FAIR principles;

I3 (**meta**)data include qualified references to other (**meta**)data;

Accessible:

A1 (**meta**)data are retrievable by their identifier using a standardized communications protocol;

A1.1 the protocol is open, free, and universally implementable;

A1.2 the protocol allows for an authentication and authorization procedure, where necessary;

A2 **metadata** are accessible, even when the data are no longer available;

Reusable:

R1 **meta**(data) are richly described with a plurality of accurate and relevant attributes;

R1.1 (meta)data are released with a clear and accessible data usage license;

R1.2 (meta)data are associated with detailed provenance;

R1.3 (meta)data meet domain-relevant community standards;

There is no FAIR data without machine-actionable **metadata**

- Researcher / Domain community
- University / Research organization
- Funder
- Publisher
- Repository
- FAIR tools and services
- Instrumentation and other 3rd parties

**There is no FAIR data without
machine-actionable metadata**

FAIR metadata Challenges:

1. Its **not always** easy for domain experts to make machine-ready metadata.
2. Its **never** easy to get widespread reuse of machine-ready metadata.



Making it easy for humans to make metadata for machines



Making it easy for humans to make metadata for machines *and reuse*



and reuse

Making it easy for humans to make metadata for machines

**Domain
Experts**

+

**Metadata
Experts**

- Registries
- Tools

=

FAIR-ready metadata templates

- target FAIR Principles
- compliant to community standards
- machine-actionable
- registered/reusable



Making it easy for humans to make metadata for machines *and reuse*

**Domain
Experts**

+

**Metadata
Experts**

- Registries
- Tools

=

FAIR-ready metadata templates

- target FAIR Principles
- compliant to community standards
- machine-actionable
- registered/reusable

Rough consensus, running code.

15 FAIR Guiding Principles Digital Object Framework

Define metadata functions

Domain Communities

aiming to become more FAIR

Self-Identified

- EOSC
- NIH Data Commons
- Preclinical Trials
- Funders
- American Geophysical Union
- Bayer
- Journalists
- Financial industry

Implementation

Choice

Community chooses to re-use existing metadata elements as needed to implement FAIR

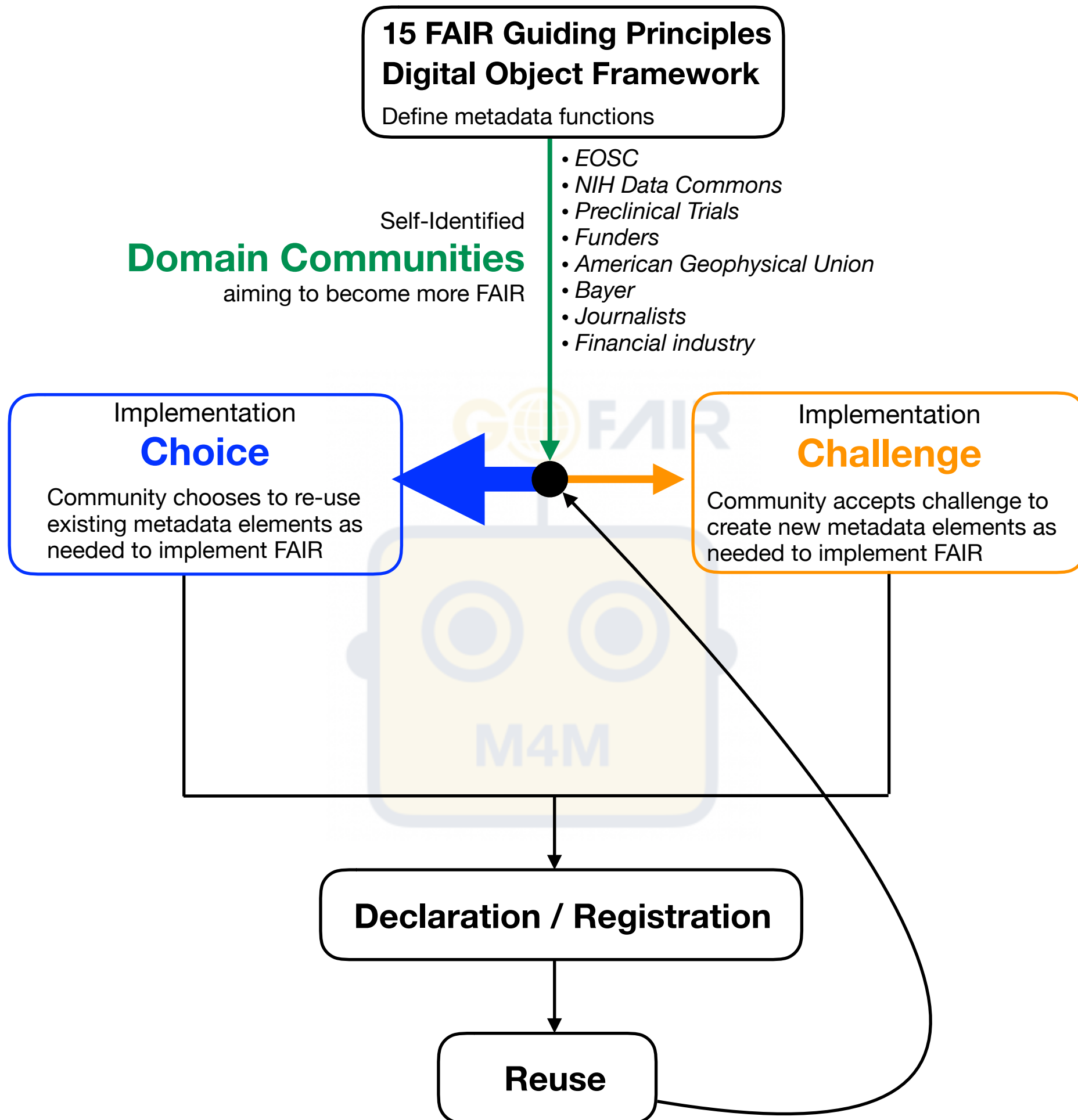
Implementation

Challenge

Community accepts challenge to create new metadata elements as needed to implement FAIR

Declaration / Registration

Reuse





Join

to create a user account

Preclinicaltrials aims to provide a comprehensive listing of preclinical animal study protocols.

Preferably registered at inception in order to **increase transparency**, help **avoid duplication**, and **reduce the risk of reporting bias** by enabling comparison of the completed study with what was planned in the protocol.

Registration of your study requires you to create an account that is

- Anonymous
- Free of charge
- Has an optional embargo period

This register is web-based, open to all types of animal studies and freely accessible and searchable to all with a preclinicaltrials.eu account.

The [registration form](#) is designed by experts on preclinical animal studies and preclinical evidence synthesis.

Please [join](#) us and create an user account, this will provide access to the database and enables you to register your preclinical trial.

Contact us
at info@preclinicaltrials.eu.



Join

to create a user account

Preclinicaltrials aims to provide a comprehensive listing of preclinical animal study protocols.

Preferably registered at inception in order to **increase transparency**, help **avoid duplication**, and **reduce the risk of reporting bias** by enabling comparison of the completed study with what was planned in the protocol.

Registration of your study requires you to create an account that is

- Anonymous
- Free of charge
- Has an optional embargo period

This register is web-based, open to all types of animal studies and freely accessible and searchable to all with a preclinicaltrials.eu account.

The [registration form](#) is designed by experts on preclinical animal studies and preclinical evidence synthesis.

Please [join](#) us and create an user account, this will provide access to the database and enables you to register your preclinical trial.

Contact us
at info@preclinicaltrials.eu.

Section 1. General information

1. * Title of the study

Enter the full title of the study

2. Acronym/short title

Enter optional acronym/short title for the study

3. * Contact details

Give the name of the main administrative contact for the study

Name

Role

What is the role the main contact in the study (e.g. executive researcher, research group supervisor)?

Email address

Provide the email address of the main contact

4. * Study centre details

Give the details of the institutions where the experiments will be undertaken. Add additional lines if there is more



Not logged in

Twitter

Designed by
for studies
synthesis.

an user
access to the
to register



CEDAR

Better metadata means better science

[WATCH VIDEO](#)

[LEARN MORE](#)



SECTION 1: GENERAL INFORMATION

1. Title of the study*

2. Acronym/short title

3. Contact Details

Name*

Role*

Email address*

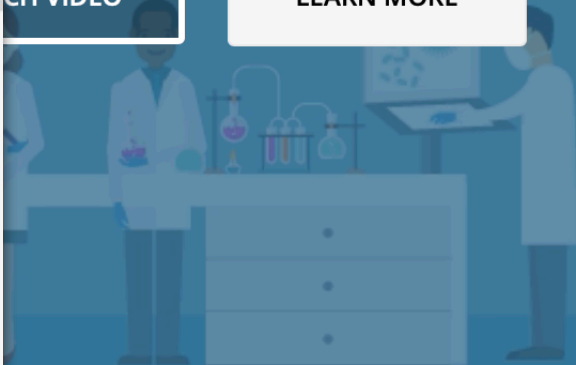
4. Study centre details*

CEDAR

means better science

CH VIDEO

LEARN MORE



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

HOW CAN WE HELP?

We guide consumers to discover, select and use these resources with confidence, and producers to make their resource more discoverable, more widely adopted and cited.

Journal editors & publishers

Create and maintain an interrelated list of citable standards, databases and repositories to recommend to your authors, users or their community, and revise this recommendation over time...

[\[read more\]](#)





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

HOW CAN WE HELP?

We guide consumers to discover, select and use these resources with confidence, and producers to make their resource more discoverable, more widely adopted and cited.

Journal editors & publishers

Create and maintain an interrelated list of citable standards, databases and repositories to recommend to your authors, users or their community, and revise this recommendation over time...

[\[read more\]](#)

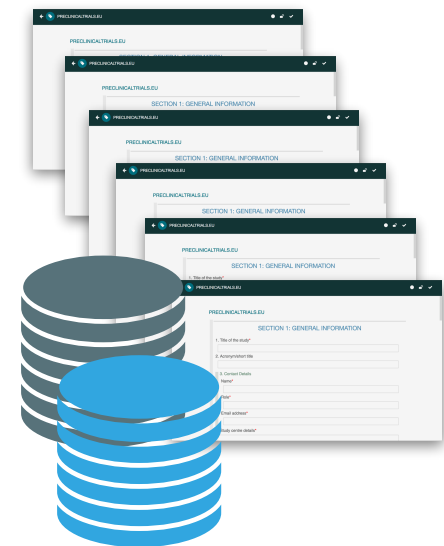




PRECLINICALTRIALS.EU



Metadata templates



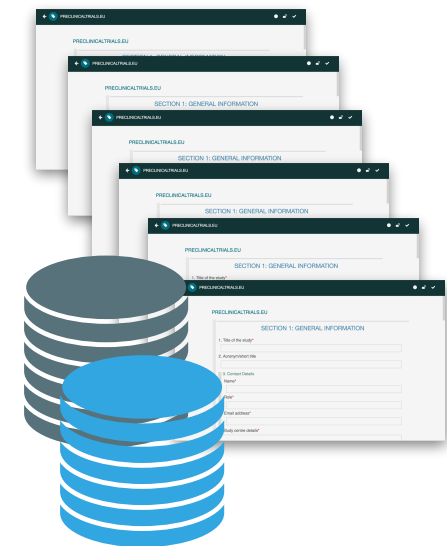


+



=

Metadata templates



[Home](#) › [Events](#) › [M4M #2: Preclinical Trials + M4M #3: Funders](#)

M4M #2: Preclinical Trials + M4M #3: Funders

Jan 14 2019



PRECLINICALTRIALS.EU



Making it easy for humans to make metadata for machines

Following the **Inaugural M4M Workshop** (October 15-16 2018, Leiden), members from GO FAIR and RDA are supporting back-2-back M4M Workshops to create machine-actionable metadata for the Preclinical Trials Research Community (#2) and for Scientific Funders (#3). Participants include: members from the Preclinical Trials and Funders communities and metadata experts from the Centre for Extended Data and Annotation and Retrieval.

At 12.00 there will be a social lunch hosted by GFISCO

Metadata for Machines Workshops

October 15-16 2018

[Home](#) › [Resources](#) › [GO FAIR Workshop Series](#) › Metadata for Machines Workshops

Resources

- › [RDM Starter Kit](#)
- › [GO FAIR Materials](#)
- › [GO FAIR Workshop Series](#)
 - › [Metadata for Machines Workshops](#)
- › [Germany goes FAIR Workshops](#)
- › [Papers & Publications](#)
- › [Videos](#)
- › [Certification](#)
- › [Glossary](#)
- › [FAQ](#)

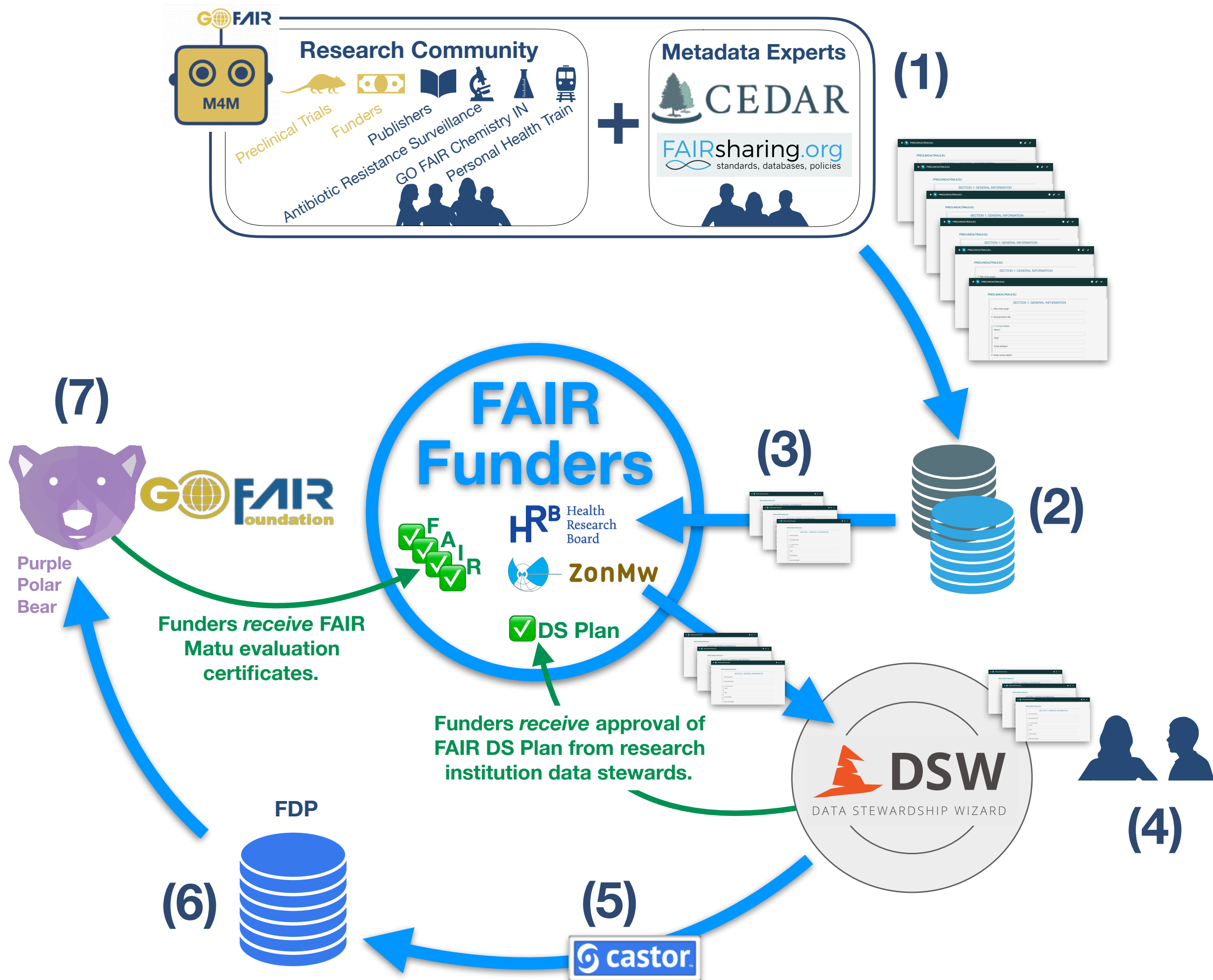
Making it easy for humans to make metadata for machines



Machine-actionable metadata are core to the FAIR Principles. GO FAIR and RDA members have launched the “Metadata for Machines” workshop series (M4M) to assess the state of metadata practices in data-related communities and stimulate the creation and re-use of FAIR metadata standards and machine-ready metadata templates (definitions of metadata categories).

The M4M workshops are agile, hackathon-style events that bring together domain experts with metadata and technical specialists to accomplish 5 objectives:

1. Assess the state of metadata practices in the various scientific communities, look for improvements of the current fragmentation and promote good FAIR compliant practices.
2. Using the FAIR principles as a guide, define essential metadata elements and standards to support F, A, I, and R by machines, drawing on the deep domain knowledge of existing communities.





arXiv.org > cs > arXiv:1902.11162

Search or Article ID

All fields



([Help](#) | [Advanced search](#))

Computer Science > Digital Libraries

The FAIR Funder pilot programme to make it easy for funders to require and for grantees to produce FAIR Data

P. Wittenburg, H. Pergl Sustkova, A. Montesanti, S. M. Bloemers, S. H. de Waard, M. A. Musen, J. B. Graybeal, K. M. Hettne, A. Jacobsen, R. Pergl, R. W. W. Hooft, C. Staiger, C. W. G. van Gelder, S. L. Knijnenburg, A.C. van Arkel, B. Meerman, M. D. Wilkinson, S-A Sansone, P. Rocca-Serra, P. McQuilton, A. N. Gonzalez-Beltran, G. J. C. Aben, P. Henning, S. Alencar, C. Ribeiro, C. R. L. Silva, L. Sayao, L. Sales, V. Veiga, J. Lima, S. Dib, P. Xavier, R. Murtinho, J. Tendel, B. F. Schaap, P. M. Brouwer, A. K. Gavai, Y. Bouzembrak, H. J. P. Marvin, A. Mons, T. Kuhn, A. A. Gambardella, R. de Miranda Azevedo, V. Muhonen, M. van der Naald, N. W. Smit, M. J. Buys, T. F. de Bruin, F. Schoots, H. J. E. Goodson, H. S. Rzepa, K. G. Jeffery, H. P. Shanahan, M. Axton, V. Tkachenko, A. D. Maya, N. K. Meyers, M. Conlon, L. L. Haak, E. A. Schultes

(Submitted on 26 Feb 2019 ([v1](#)), last revised 6 Mar 2019 (this version, v2))

Download:

- [PDF only](#)



Current browse context:

cs.DL

[< prev](#) | [next >](#)

[new](#) | [recent](#) | [1902](#)

Change to browse by:

[cs](#)

References & Citations

























- [NASA ADS](#)

Google Scholar

Bookmark ([what is this?](#))



M4M Workshop requested

  #4 Publishers 	Schultes, Wittenburg, Mons & 7 more Initiated by Myles Axton (now at Wiley). The idea is that a library of metadata statements will enable a publisher to (easily) compose metadata profil...
  #5 Antimicrobial Resistance Surveillance 	Schultes, Wittenburg, Mons & 7 more Community tying to get real-time interoperation of clinical findings in Antimicrobial Resistance.
  #6 Integrating the Healthcare Enterprise (IHE) 	Schultes, Wittenburg, Mons & 7 more IHE is an initiative by healthcare professionals and industry to improve the way computer systems in healthcare share information. IHE promotes the co...
  #7 GO FAIR Chemistry IN 	Schultes, Wittenburg, Mons & 7 more Community (IUPAC and others) that is setting standards for FAIR Chemistry.
  #8 GO FAIR Personal Health Train IN 	Schultes, Wittenburg, Mons & 7 more PHT Community needs to decide on "red box" / "blue box" issues.
  #9 FAIR Data Point 	Schultes, Wittenburg, Mons & 9 more
  #10 Rare Diseases 	Schultes & Roos
  #11 EOSC-Nordic project 	Schultes, Mons, Sustkova & 9 more

<https://osf.io/qe9fa/>

M4M Workshop Focus Areas

Beginners M4M

- a simple 1 or 2 day event
- minimal self-descriptive metadata profile (ORCID, Affiliation)
- licensing issue (R1.1)
- link to funding agency
- automated service

Advanced M4M

- 2-day (or longer) events
- designed to crack a complex F2 or R1.2 metadata problem
- could involve creation of new definitions

Institutional M4M

- Organisations that promote FAIR need to have FAIR metadata themselves
- highly effective as convergers in global interoperability of administrative metadata
- likely to be highly repreative
- modelling internal structure of organisations, and the roles played by different personnel

M4M Workshop Focus Areas

Core M4M Workshops

focus on foundational metadata that are generic and used throughout numerous application areas. Core M4M workshop leverage metadata experts from RDA and in various research infrastructure initiatives.

- Terminology Workshop
- Simple Ontology Supported Metadata
- Registries for Schemas
- Registries for Semantic Categories and Vocabularies
- Component Infrastructure
- Metadata Extraction Library
- Metadata in Workflows
- Flexible Semantic Mapping Framework
- Metadata Tool Registry

M4M Workshop Focus Areas

Content M4M Workshops

workshops will focus on domain-relevant metadata standards to facilitate scientific and scholarly functions in specific research disciplines and industrial applications. Content M4M workshops leverage practicing domain experts who are the only authorities that can craft real-world, semantically rich metadata in specialized application areas.

- **help in metadata terminology**
- **help to start with simple metadata solutions**
- **help in choosing schemas & semantic vocabularies already in use**
- **help with respect to finding suitable tools to create and curate metadata**
- **help finding suitable components augmenting schema**
- **help registering schemas, components, and defined categories**
- **help in easy semantic mapping facilities**
- **help exporting metadata to meet the requirements for discoverability**

Harnessing RDA

Catalog of metadata schemas organised by MSCWG

<https://www.rd-alliance.org/groups/metadata-standards-catalog-working-group.html>

The original directory

<http://rd-alliance.github.io/metadata-directory/>

More recently

<https://rdamsc.dcc.ac.uk/>

The ‘umbrella’ MIG has a set of metadata principles and is working on a canonical set of metadata elements (with the idea that common metadata formats can use it as the interconversion medium)

<https://rd-alliance.org/groups/metadata-ig.html>

M4M Workshops are a vehicle for community decision making around metadata standards.



For (faster) FAIRification