



The GO FAIR approach to the practical implementation of data interoperability: The role of machine-actionable metadata

**UKB Meeting
SURF, Utrecht
July 2, 2018**

**Erik Schultes, PhD
International Science Coordinator
GO FAIR International Support and Coordination Office**

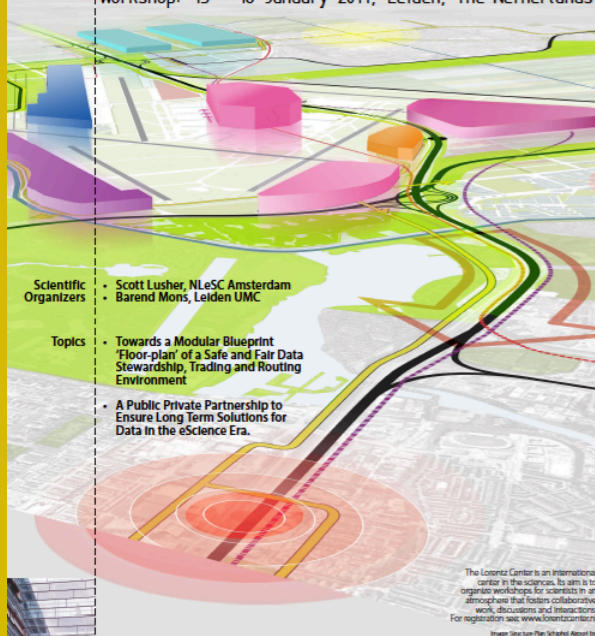
FAIR and GO FAIR

Lorentz

Lorentz center

Jointly Designing a Data FAIRPORT

Workshop: 13 - 16 January 2014, Leiden, the Netherlands



Scientific Organizers

- Scott Lusher, NLeSC Amsterdam
- Barend Mons, Leiden UMC

Topics

- Towards a Modular Blueprint 'Floor-plan' of a Safe and Fair Data Stewardship, Trading and Routing Environment
- A Public Private Partnership to Ensure Long Term Solutions for Data in the eScience Era.

The Lorentz Center is an international center in the sciences. Its aim is to organize workshops for scientists in an atmosphere that fosters collaborative work, discussions and interactions. For registration see www.lorentzcenter.nl

Image: Data Fair - Original Report by ICIP Amsterdam, Leiden, 2014. Please refer to the report for more information.

FOM STW eScience DTL Science NWO Lorentz center

www.lorentzcenter.nl

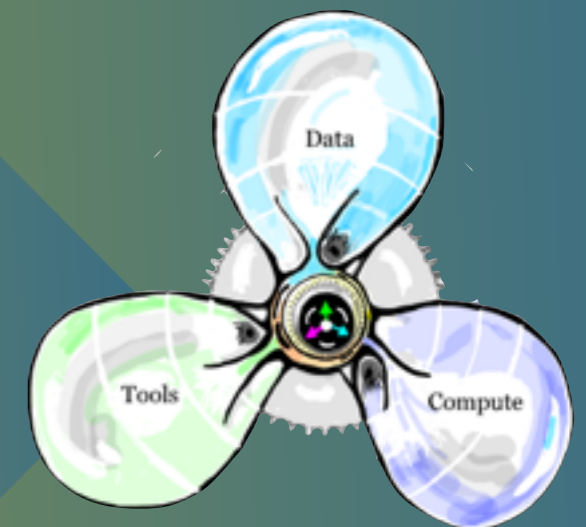
EOSC



FAIRdICT



IFDS



Birth

2014

Infancy

2015

2016

Adolescence

2017

2018...

Maturity

GO BUILD FAIRification technology

- FAIRifier
- FAIR Data Point
- FAIR Metrics
- convergence on standards

GO TRAIN and certification of people

- LERU Summer School
- Lorentz Workshop
- BYODs
- Canonical FAIR processes and curriculum

GO CHANGE organizations (incentives)


- Public: universities, funding organizations
- Private: tech, publishing

GO FAIR offices and Implementati...

Here we keep track of the development of the GO FAIR network

129 views

All changes saved in Drive

 Add layer  Share  Preview

Biodiversity IN


 Individual styles

 Naturalis Biodiversity Center

 Museum für Naturkunde


 Natural History Museum

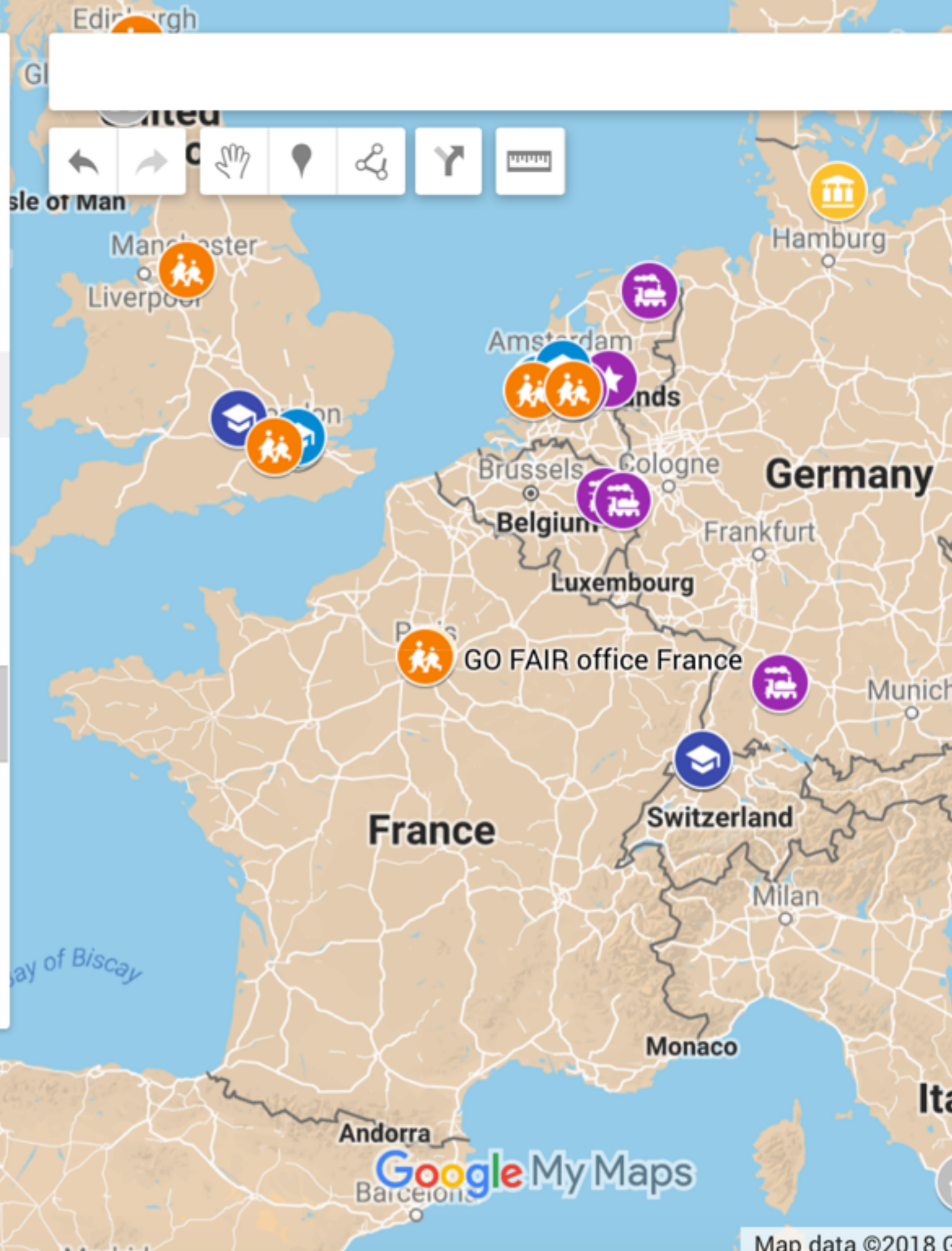
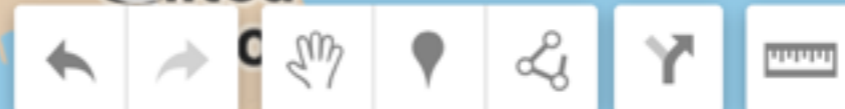
 Natural History Museum

 Tromsø Museum – Universit...

 Royal Botanic Garden Edinbu...

 The Faculty of Natural Scien...

 National Museum of Natural ...



France

Switzerland

Milan

Monaco

Andorra

Google My Maps

What is FAIR Data ?

**“Data and services that are
findable,
accessible,
interoperable,
re-usable
both for machines and for people.”**

The FAIR Guiding Principles for scientific data management and stewardship,
Scientific Data (2016), <https://www.nature.com/articles/sdata201618>

What is FAIR Data ?

Data and services that are
findable,
accessible,
interoperable,
re-usable
for machines (*and sometimes, in rare
circumstances, may be even for people*).

What is FAIR Data ?

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;

What is FAIR Data ?

<https://www.go-fair.org/fair-principles/>



News Contact

[GO FAIR Initiative](#) [Implementation Networks](#) [FAIR Principles](#) [Technology](#) [Training](#) [Certification](#)

FAIR Principles

<https://www.go-fair.org/fair-principles/>

[Home](#) > [FAIR Principles](#)

FAIR Principles

- > **F1: (meta) data are assigned globally unique and persistent identifiers**
- > **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**
- > **A1: (meta)data are retrievable by their identifier using a standardised communication**

On March 15th, 2016, a group has published "**The FAIR Guiding Principles for scientific data management and stewardship**" comment on Nature's Scientific Data. The authors aimed for the principles to act as guidelines for those willing to improve findability, accessibility, interoperability and reuse of their digital assets. With the increase on volume, complexity and creation speed of data, humans are more and more relying on computational support for dealing with data. The principles were, therefore, defined with the focus on machine-actionability, i.e., the capacity of computational systems to find, access, interoperate and reuse data with none or minimal human intervention.

Findable: the first obstacle for someone willing to (re)use data is to find them. Metadata and data should be easy to find by both humans and computers. Machine-readable metadata is essential for automatic discovery of relevant datasets and services, and for this reason are essential to the **FAIRification process**.

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

F2. data are described with rich metadata

What is FAIR Data ?

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;

What is FAIR Data ?

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;

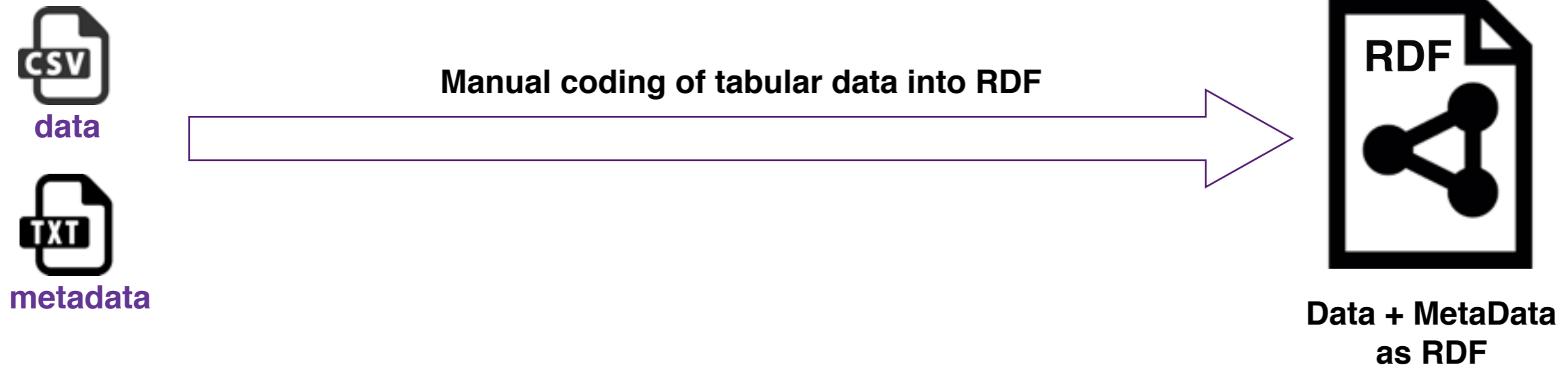
R1.1 (meta)data are released with a clear and accessible data usage license;

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

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

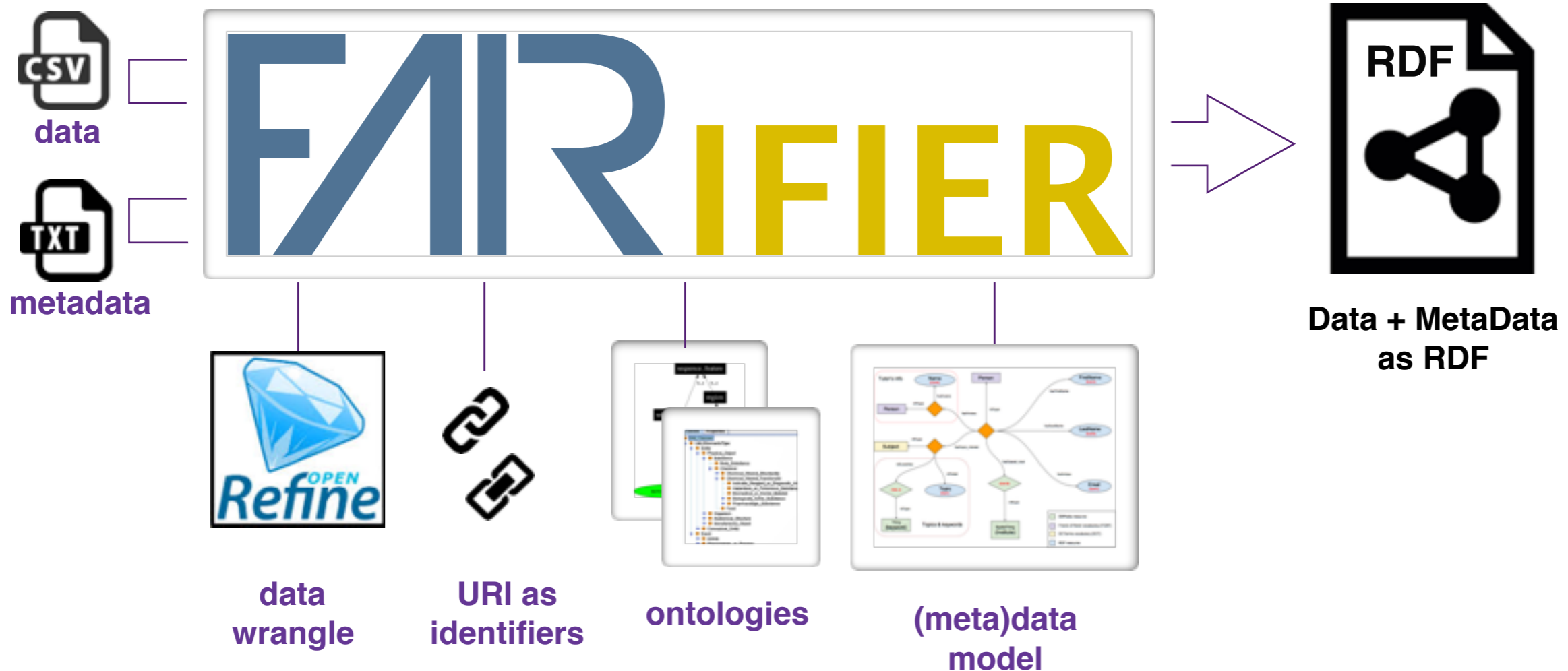
A Reference Implementation of the 15 Principles to show how Data can be made FAIR for machines

(FAIR-dICT project, DTL: <https://www.dtls.nl/fair-data/fair-dict/>)



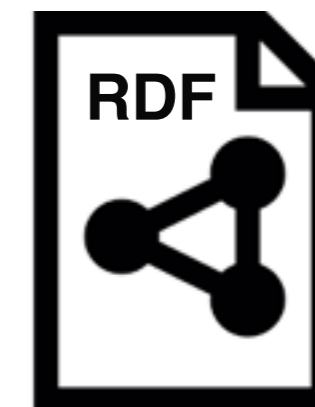
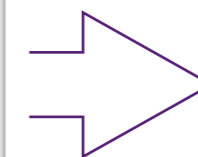
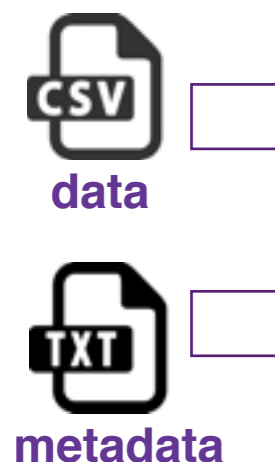
A Reference Implementation of the 15 Principles to show how Data can be made FAIR for machines

(FAIR-dICT project, DTL: <https://www.dtls.nl/fair-data/fair-dict/>)



A Reference Implementation of the 15 Principles to show how Data can be made FAIR for machines

(FAIR-dICT project, DTL: <https://www.dtls.nl/fair-data/fair-dict/>)



Data + MetaData as RDF



data wrangle



URI as identifiers



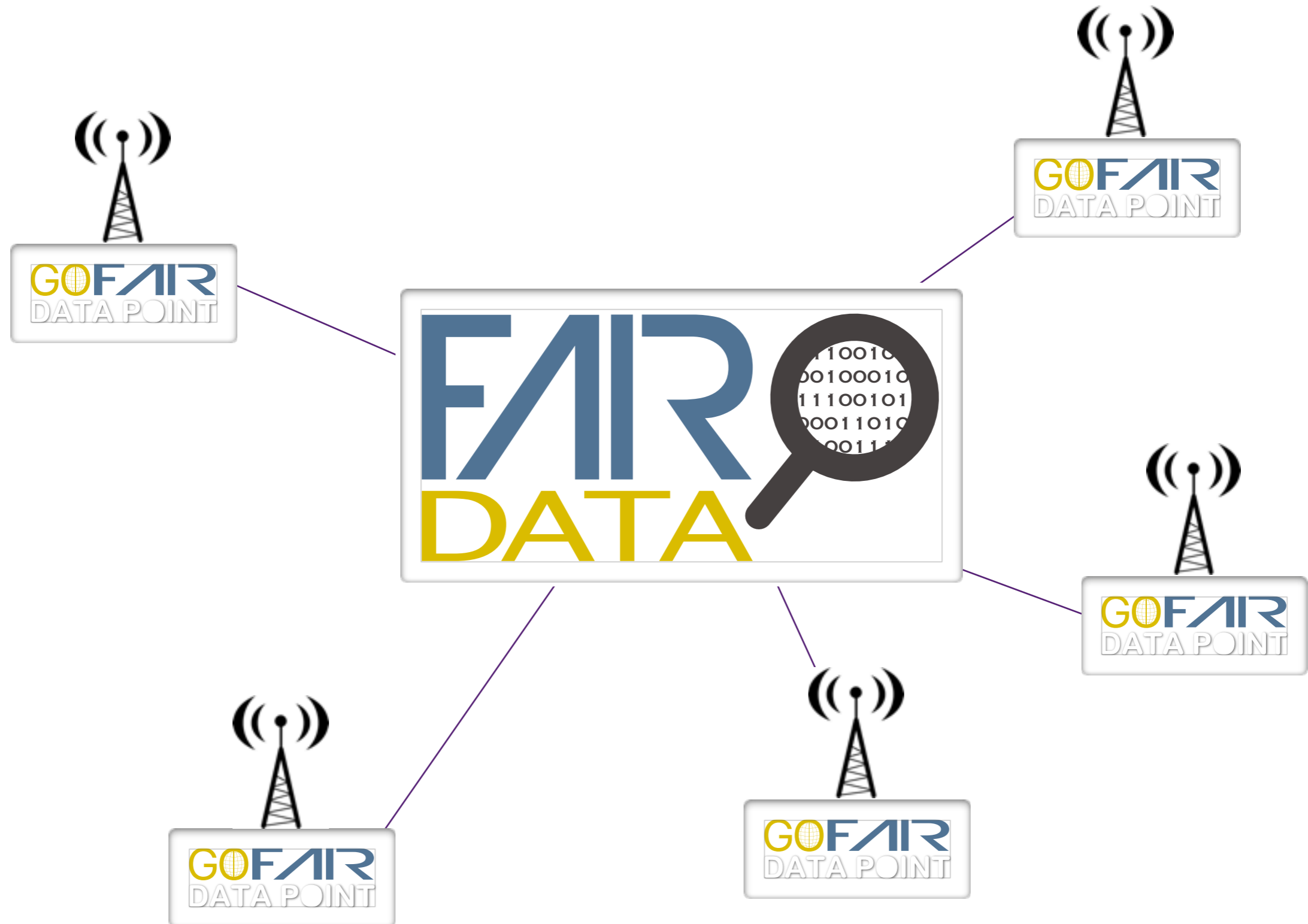
ontologies



(meta)data model

A Reference Implementation of the 15 Principles to show how Data can be made FAIR for machines

(FAIR-dICT project, DTL: <https://www.dtls.nl/fair-data/fair-dict/>)



FAIR Data Point

GET <URL>

<http://www.w3.org/TR/vocab-dcat/>

Catalog 1

DATASET 1

DIST 1

DIST 2

DATASET 2

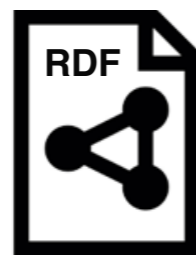
DIST 3

Catalog 2

DATASET 3

DIST 4

DIST 5



FAIR metadata

Title	FDP of lorentz.fair-dtls.surf-hosted.nl
Metadata ID	3e77134d-9338-482a-a4f7-8e6933b47469
Description	FDP of lorentz.fair-dtls.surf-hosted.nl
Issued	2017-05-12T12:11:36.343Z
Modified	2017-05-12T12:11:36.355Z
License	license
Catalogs	https://lorentz.fair-dtls.surf-hosted.nl/fdp/catalog/wur https://lorentz.fair-dtls.surf-hosted.nl/fdp/catalog/DvBC1 https://lorentz.fair-dtls.surf-hosted.nl/fdp/catalog/PKD_2.0 https://lorentz.fair-dtls.surf-hosted.nl/fdp/catalog/AGIS https://lorentz.fair-dtls.surf-hosted.nl/fdp/catalog/PKD_1.0 https://lorentz.fair-dtls.surf-hosted.nl/fdp/catalog/Kenya.2017 https://lorentz.fair-dtls.surf-hosted.nl/fdp/catalog/test_example_number_2_2.0 https://lorentz.fair-dtls.surf-hosted.nl/fdp/catalog/post_example_2.0 https://lorentz.fair-dtls.surf-hosted.nl/fdp/catalog/cds-v1 https://lorentz.fair-dtls.surf-hosted.nl/fdp/catalog/post-test_1.0 https://lorentz.fair-dtls.surf-hosted.nl/fdp/catalog/testcatalog_1.0 https://lorentz.fair-dtls.surf-hosted.nl/fdp/catalog/FAIR_in_Africa_0001



Google Search

I'm Feeling Lucky



| inhibitors p65

Google Search

I'm Feeling Lucky



inhibitors p65



All

Images

Shopping

Videos

News

More

Settings

Tools

About 544.000 results (0,51 seconds)

RelA/NFkB p65 Inhibitors: Novus Biologicals

<https://www.novusbio.com/inhibitors/rela-nfkb-p65> ▼

RelA/NFkB p65 Inhibitors available through Novus Biologicals. Browse our RelA/ NFkB p65 Inhibitor catalog backed by our Guarantee+.

Inhibiting NF-κB Activation by Small Molecules As a Therapeutic ...

<https://www.ncbi.nlm.nih.gov> › NCBI › Literature › PubMed Central (PMC)

by SC Gupta - 2010 - Cited by 345 - Related articles

May 21, 2010 - ... pathway (Fig 3). Table 1. A list of small molecules as inhibitors of NF-κB pathway Blocking NF-κB activation by inhibitors of p65 acetylation.

Suppression of p65 phosphorylation coincides with inhibition of ... - NCBI

<https://www.ncbi.nlm.nih.gov/pubmed/16163708>

by J Hu - 2005 - Cited by 26 - Related articles

Mol Carcinog. 2005 Dec;44(4):274-84. Suppression of p65 phosphorylation coincides with inhibition of IκappaBα polyubiquitination and degradation.

Identification of a p65 peptide that selectively inhibits NF-kappa B ...

<https://www.ncbi.nlm.nih.gov/pubmed/14711835>

by Y Takada - 2004 - Cited by 157 - Related articles

Jan 7, 2004 - Identification of a p65 peptide that selectively inhibits NF-kappa B activation





Find me [all known low molecular weight inhibitors of the Human p65 Protein](#). Separate the list based on those that were [found in curated databases](#), from those that were [found in self-deposited data archives](#). Keep track of the [license and citation information](#) for each one. If data is relevant, but [not public](#), please provide the [contact information for the person I need](#) so I can request the data.

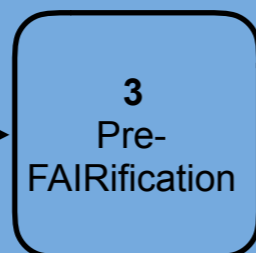
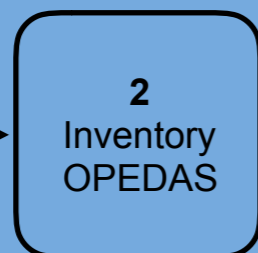
Canonical FAIRification

Institutional Policy for Going FAIR

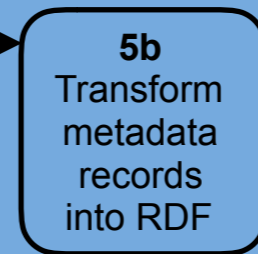
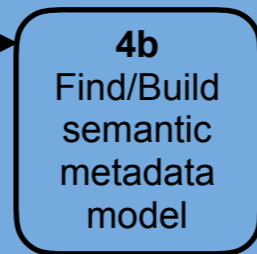
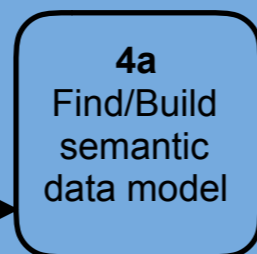
Implementation Strategy



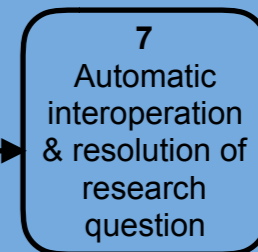
Preparation



FAIRification



Analysis



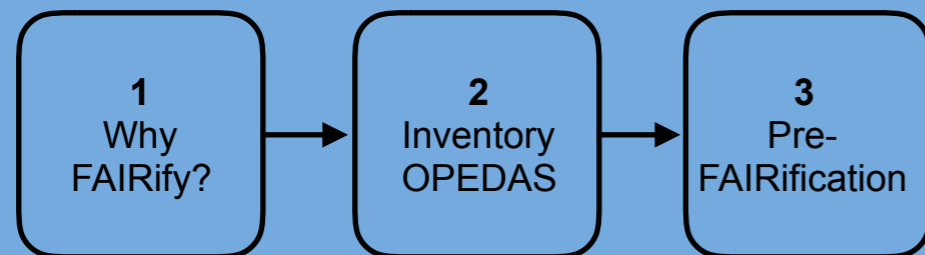
Canonical FAIRification

Institutional Policy for Going FAIR

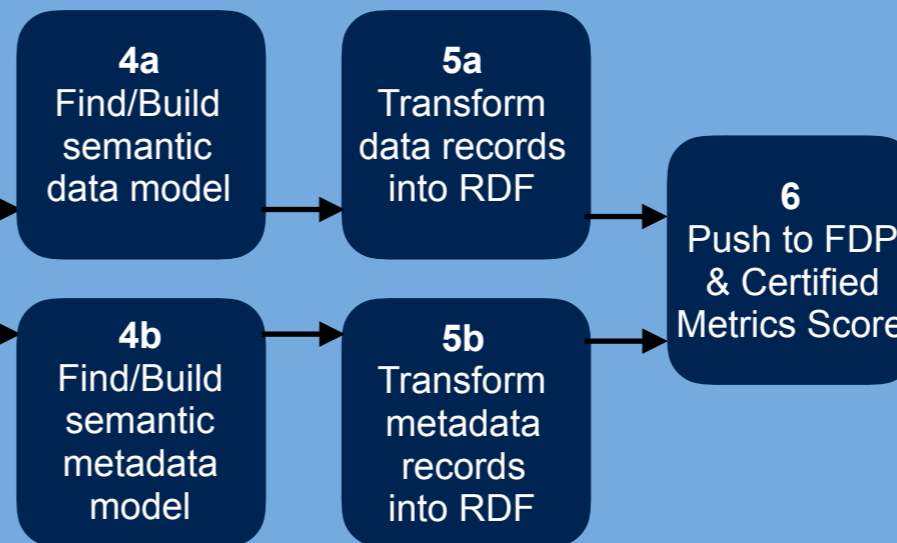
Implementation Strategy



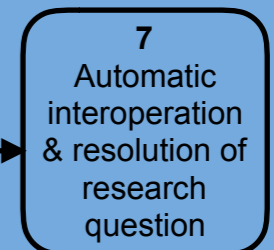
Preparation



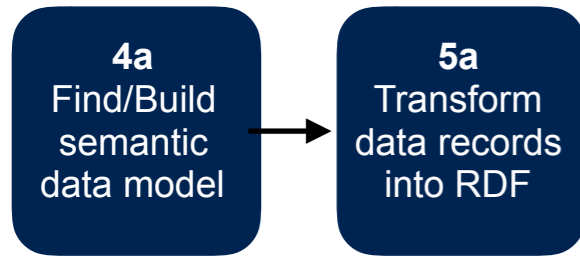
FAIRification



Analysis



Practicing FAIR Data



Distinguish:
Information
Data
Knowledge

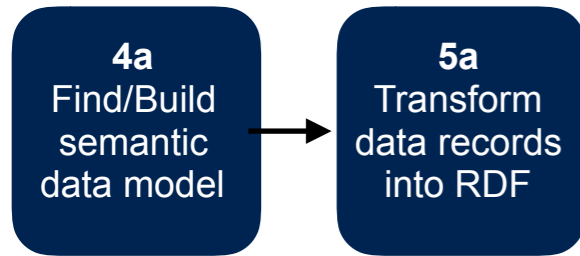
Consider:

Semantics is the **linguistic** and **philosophical** study of **meaning**.

<https://en.wikipedia.org/wiki/Semantics>

Semantic Data is, *data with meaning*.

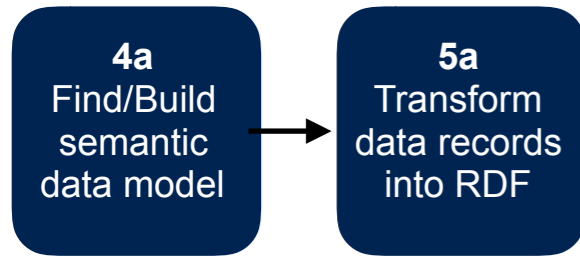
Practicing FAIR Data



Knowledge: associations implied by data

Malaria is transmitted by mosquitoes.

Practicing FAIR Data

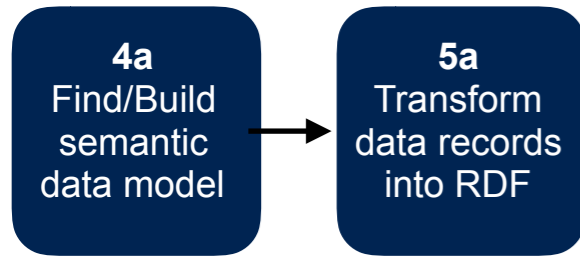


Knowledge: associations implied by data

Malaria is transmitted by mosquitoes.

Subject - Predicate - Object

Practicing FAIR Data



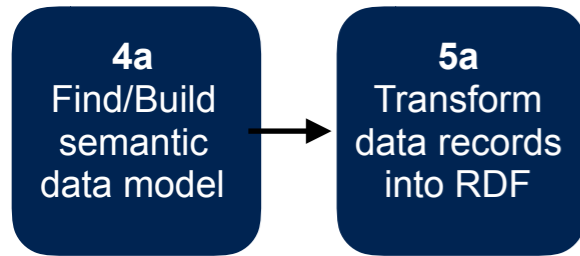
Knowledge: associations implied by data

Malaria is transmitted by mosquitoes.

Subject - Predicate - Object

semantic triple

Practicing FAIR Data



Knowledge: associations implied by data

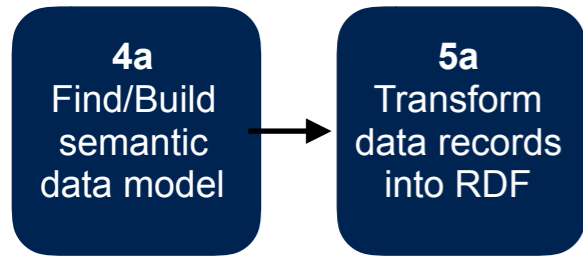
Malaria is transmitted by mosquitoes.

Subject - Predicate - Object

ORPHA673 - D018562 - *Anopheles gambiae*

semantic triple

Practicing FAIR Data



Knowledge: associations implied by data

Malaria is transmitted by mosquitoes.

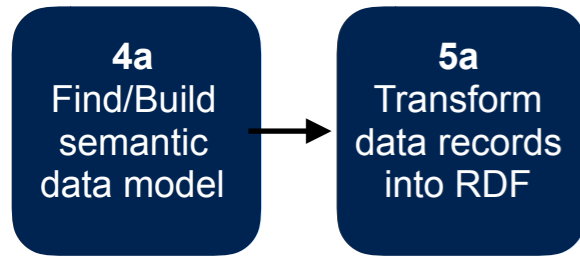
Subject - Predicate - Object

ORPHA673 - D018562 - *Anopheles gambiae*

Ontologies

semantic triple

Practicing FAIR Data



Knowledge: associations implied by data

Malaria is transmitted by mosquitoes.

Subject - Predicate - Object

ORPHA673 - D018562 - *Anopheles gambiae*

http://www.orpha.net/consor/cgi-bin/OC_Exp.php?Lng=GB&Expert=673 -

<http://bioportal.bioontology.org/ontologies/MESH?p=classes&conceptid=D018562> -

<https://www.vectorbase.org/organisms/anopheles-gambiae>

Ontologies

semantic triple

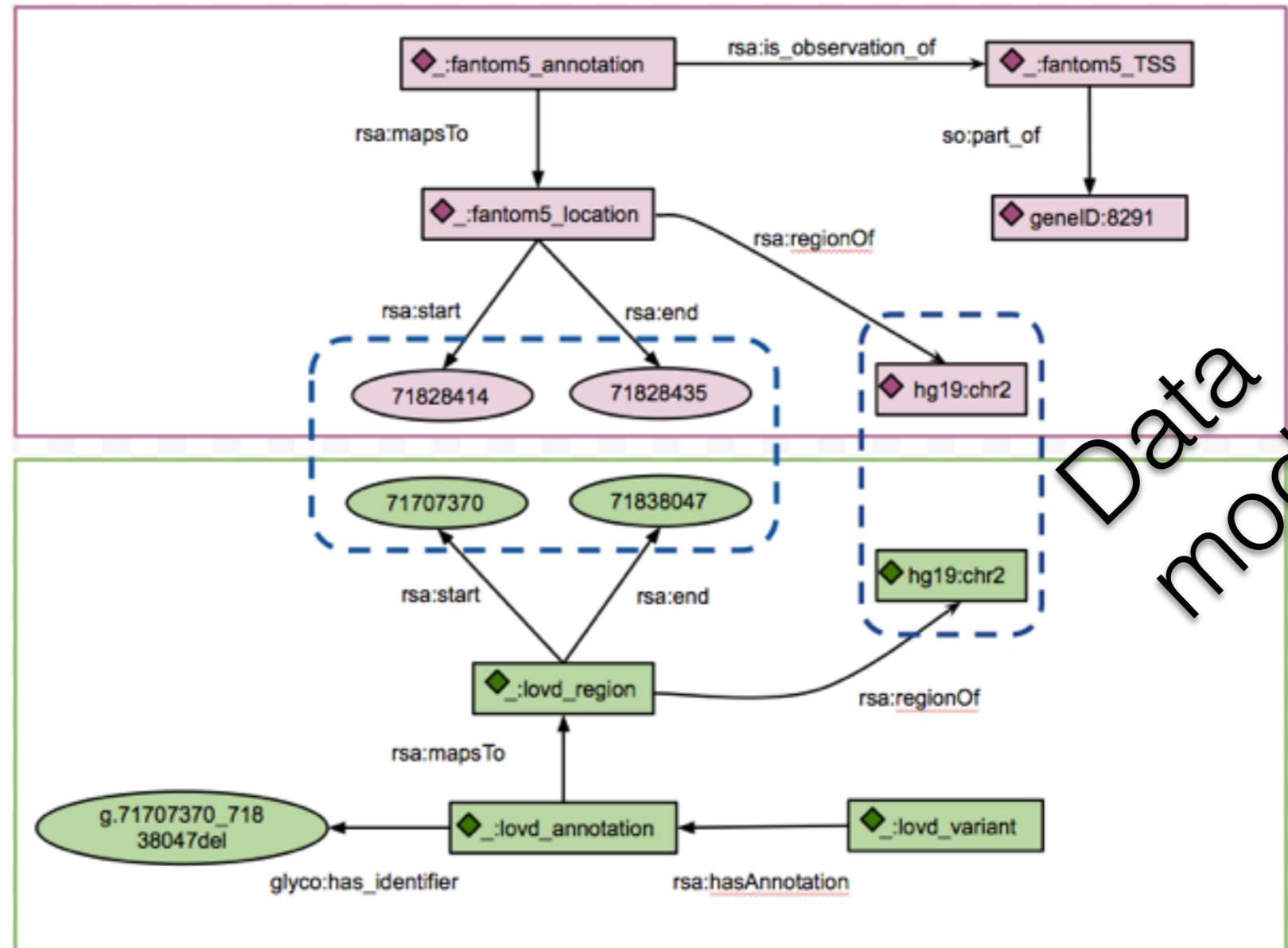
GPRI

Practicing FAIR Data

4a
Find/Build
semantic
data model

5a
Transform
data records
into RDF

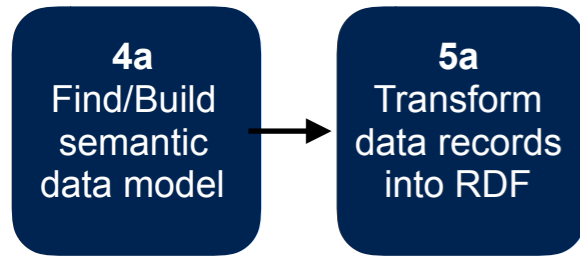
Knowledge: associations implied by data



Data models

20% of known disease causing variants map to TSS

Practicing FAIR Data

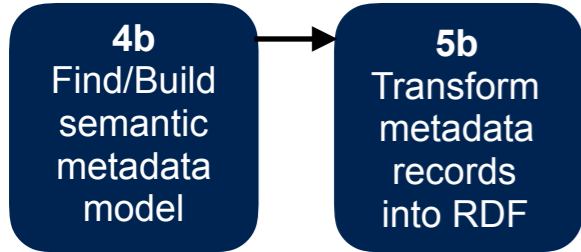


Knowledge: associations implied by data

The Future of Semantic Data Modeling:

- Deep dive course on Data Modeling
- Tools (visual modeling tools)
- Libraries of FAIR Data Models (F1000)
- Common Data Model templates
- Peer review (Myles Axton)

Practicing FAIR MetaData



Metadata: Data about data

www.nature.com/scientificdata

SCIENTIFIC DATA

OPEN

SUBJECT CATEGORIES

- » Research data
- » Publication characteristics

Comment: The FAIR Guiding Principles for scientific data management and stewardship

Mark D. Wilkinson *et al.*[#]

There is an urgent need to improve the infrastructure supporting the reuse of scholarly data. A diverse set of stakeholders—representing academia, industry, funding agencies, and scholarly publishers—have come together to design and jointly endorse a concise and measurable set of principles that we refer to as the FAIR Data Principles. The intent is that these may act as a guideline for those wishing to enhance the reusability of their data holdings. Distinct from peer initiatives that focus on the human scholar, the FAIR Principles put specific emphasis on enhancing the ability of machines to automatically

Received: 10 December 2015

Accepted: 12 February 2016

Published: 15 March 2016

Practicing FAIR MetaData

4b

Find/Build
semantic
metadata
model

5b

Transform
metadata
records
into RDF

Metadata: Data about data

Acknowledgements

The original Lorentz Workshop ‘Jointly Designing a Data FAIRport’ was organized by Barend Mons in collaboration with and co-sponsored by the Lorentz center, The Dutch Techcenter for the Life Sciences and the Netherlands eScience Center. The principles and themes described in this manuscript represent the significant voluntary contributions and participation of the authors at, and/or subsequent to, this workshop and from the wider Force11, BD2K and ELIXIR communities. We also acknowledge and thank the organizers and backers of the NBDC/DBCLS BioHackathon 2015, where several of the authors made significant revisions to the FAIR Principles.

Author Contributions

M.W. was the primary author of the manuscript, and participated extensively in the drafting and editing of the FAIR Principles. M.D. was significantly involved in the drafting of the FAIR Principles. B.M. conceived of the FAIR Data Initiative, contributed extensively to the drafting of the principles, and to this manuscript text. All other authors are listed alphabetically, and contributed to the manuscript either by their participation in the initial workshop and/or by editing or commenting on the manuscript text.

Additional Information

Competing financial interests: M.A. is the *Nature Genetics*’ Editor in Chief; S.A.S. is *Scientific Data*’s Honorary Academic Editor and consultant.

How to cite this article: Wilkinson, M. D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* 3:160018 doi: 10.1038/sdata.2016.18 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0>

Practicing FAIR MetaData

4b

Find/Build
semantic
metadata
model

5b

Transform
metadata
records
into RDF

Metadata: Data about data

Acknowledgements

The original Lorentz Workshop ‘Jointly Designing a Data FAIRport’ was organized by Barend Mons in collaboration with and co-sponsored by the Lorentz center, The Dutch Techcenter for the Life Sciences and the Netherlands eScience Center. The principles and themes described in this manuscript represent the significant voluntary contributions and participation of the authors at, and/or subsequent to, this workshop and from the wider Force11, BD2K and ELIXIR communities. We also acknowledge and thank the organizers and backers of the NBDC/DBCLS BioHackathon 2015, where several of the authors made significant revisions to the FAIR Principles.

The FAIR Guiding Principles for scientific data management and stewardship

MD Wilkinson, M Dumontier, IJJ Aalbersberg, G Appleton, M Axton, ...
Scientific data 3

576

2016

... of the FAIR Data Initiative, contributed extensively to the drafting of the principles, and to this manuscript text. All other authors are listed alphabetically, and contributed to the manuscript either by their participation in the initial workshop and/or by editing or commenting on the manuscript text.

Additional Information

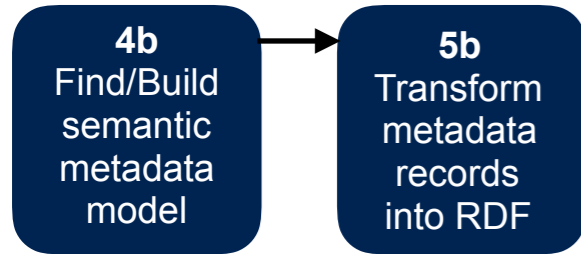
Competing financial interests: M.A. is the *Nature Genetics*' Editor in Chief; S.A.S. is *Scientific Data*'s Honorary Academic Editor and consultant.

How to cite this article: Wilkinson, M. D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* 3:160018 doi: 10.1038/sdata.2016.18 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0>

Practicing FAIR MetaData



53 Authors

47 Affiliations

Metadata: Data about data

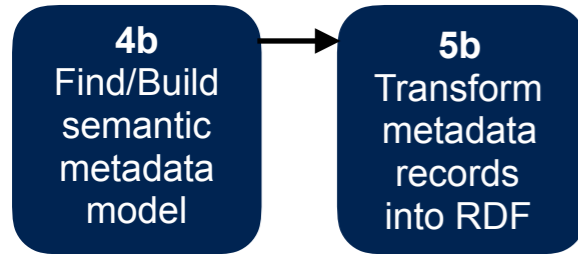
Mark D. Wilkinson¹, Michel Dumontier², IJsbrand Jan Aalbersberg³, Gabrielle Appleton³, Myles Axton⁴, Arie Baak⁵, Niklas Blomberg⁶, Jan-Willem Boiten⁷, Luiz Bonino da Silva Santos⁸, Philip E. Bourne⁹, Jildau Bouwman¹⁰, Anthony J. Brookes¹¹, Tim Clark¹², Mercè Crosas¹³, Ingrid Dillo¹⁴, Olivier Dumon³, Scott Edmunds¹⁵, Chris T. Evelo¹⁶, Richard Finkers¹⁷, Alejandra Gonzalez-Beltran¹⁸, Alasdair J.G. Gray¹⁹, Paul Groth³, Carole Goble²⁰, Jeffrey S. Grethe²¹, Jaap Heringa²², Peter A.C. 't Hoen²³, Rob Hooft²⁴, Tobias Kuhn²⁵, Ruben Kok²², Joost Kok²⁶, Scott J. Lusher²⁷, Maryann E. Martone²⁸, Albert Mons²⁹, Abel L. Packer³⁰, Bengt Persson³¹, Philippe Rocca-Serra¹⁸, Marco Roos³², Rene van Schaik³³, Susanna-Assunta Sansone¹⁸, Erik Schultes³⁴, Thierry Sengstag³⁵, Ted Slater³⁶, George Strawn³⁷, Morris A. Swertz³⁸, Mark Thompson³², Johan van der Lei³⁹, Erik van Mulligen³⁹, Jan Velterop⁴⁰, Andra Waagmeester⁴¹, Peter Wittenburg⁴², Katherine Wolstencroft⁴³, Jun Zhao⁴⁴ & Barend Mons^{45,46,47}

¹Center for Plant Biotechnology and Genomics, Universidad Politécnica de Madrid, Madrid 28223, Spain. ²Stanford University, Stanford 94305-5411, USA. ³Elsevier, Amsterdam 1043 NX, The Netherlands. ⁴Nature Genetics, New York 10004-1562, USA. ⁵Euretos and Phortos Consultants, Rotterdam 2741 CA, The Netherlands. ⁶ELIXIR, Wellcome Genome Campus, Hinxton CB10 1SA, UK. ⁷Lygature, Eindhoven 5656 AG, The Netherlands. ⁸Vrije Universiteit Amsterdam, Dutch Techcenter for Life Sciences, Amsterdam 1081 HV, The Netherlands. ⁹Office of the Director, National Institutes of Health, Rockville 20892, USA. ¹⁰TNO, Zeist 3700 AJ, The Netherlands. ¹¹Department of Genetics, University of Leicester, Leicester LE1 7RH, UK. ¹²Harvard Medical School, Boston, Massachusetts MA 02115, USA. ¹³Harvard University, Cambridge, Massachusetts MA 02138, USA. ¹⁴Data Archiving and Networked Services (DANS), The Hague 2593 HW, The Netherlands. ¹⁵GigaScience, Beijing Genomics Institute, Shenzhen 518083, China. ¹⁶Department of Bioinformatics, Maastricht University, Maastricht 6200 MD, The Netherlands. ¹⁷Wageningen UR Plant Breeding, Wageningen 6708 PB, The Netherlands. ¹⁸Oxford e-Research Center, University of Oxford, Oxford OX1 3QG, UK. ¹⁹Heriot-Watt University, Edinburgh EH14 4AS, UK. ²⁰School of Computer Science, University of Manchester, Manchester M13 9PL, UK. ²¹Center for Research in Biological Systems, School of Medicine, University of California San Diego, La Jolla, California 92093-0446, USA. ²²Dutch Techcenter for the Life Sciences, Utrecht 3501 DE, The Netherlands. ²³Department of Human Genetics, Leiden University Medical Center, Dutch Techcenter for the Life Sciences, Leiden 2300 RC, The Netherlands. ²⁴Dutch TechCenter for Life Sciences and ELIXIR-NL, Utrecht 3501 DE, The Netherlands. ²⁵VU University Amsterdam, Amsterdam 1081 HV, The Netherlands. ²⁶Leiden Center of Data Science, Leiden University, Leiden 2300 RA, The Netherlands. ²⁷Netherlands eScience Center, Amsterdam 1098 XG, The Netherlands. ²⁸National Center for Microscopy and Imaging Research, UCSD, San Diego 92103, USA. ²⁹Phortos Consultants, San Diego 92011, USA. ³⁰SciELO/FAPESP Program, UNIFESP Foundation, São Paulo 05468-901, Brazil. ³¹Bioinformatics Infrastructure for Life Sciences (BILS), Science for Life Laboratory, Dept of Cell and Molecular Biology, Uppsala University, S-751 24, Uppsala, Sweden. ³²Leiden University Medical Center, Leiden 2333 ZA, The Netherlands. ³³Bayer CropScience, Gent Area 1831, Belgium. ³⁴Leiden Institute for Advanced Computer Science, Leiden University Medical Center, Leiden 2300 RA, The Netherlands. ³⁵Swiss Institute of Bioinformatics and University of Basel, Basel 4056, Switzerland. ³⁶Cray, Inc., Seattle 98164, USA. ³⁷Unaffiliated. ³⁸University Medical Center Groningen (UMCG), University of Groningen, Groningen 9713 GZ, The Netherlands. ³⁹Erasmus MC, Rotterdam 3015 CE, The Netherlands. ⁴⁰Independent Open Access and Open Science Advocate, Guildford GU1 3PW, UK. ⁴¹Micelio, Antwerp 2180, Belgium. ⁴²Max Planck Compute and Data Facility, MPS, Garching 85748, Germany. ⁴³Leiden Institute of Advanced Computer Science, Leiden University, Leiden 2333 CA, The Netherlands. ⁴⁴Department of Computer Science, Oxford University, Oxford OX1 3QD, UK. ⁴⁵Leiden University Medical Center, Leiden and Dutch TechCenter for Life Sciences, Utrecht 2333 ZA, The Netherlands. ⁴⁶Netherlands eScience Center, Amsterdam 1098 XG, The Netherlands. ⁴⁷Erasmus MC, Rotterdam 3015 CE, The Netherlands.

● ORCID

- GUPRI
- metadata profile
- machine-actionable

Practicing FAIR MetaData



Metadata: Data about data

The Future of Semantic Data Modeling:

- Metadata Editors:
 - Center for Expanded Data Annotation and Retrieval (CEDAR) <https://metadatacenter.org>
 - DTL <https://fair-course.fair-dtls.surf-hosted.nl/editor/#!/>
- Libraries of FAIR MetaData Models (F1000)
- Common Metadata templates:
 - FAIRSharing <https://fairsharing.org>
 - People: ORCID profiles <https://orcid.org>
 - Institutions:
 - [your institution](https://www.grid.ac) 😊 <https://www.grid.ac>
 - funding organisations <https://www.crossref.org/services/funder-registry/>



dcodings / schemas
forked from shexSpec/schemas

Watch 1 Star 0 Fork 6

Code Pull requests 0 Projects 0 Insights

Join GitHub today

GitHub is home to over 28 million developers working together to host and review code, manage projects, and build software together.

[Sign up](#)

Dismiss

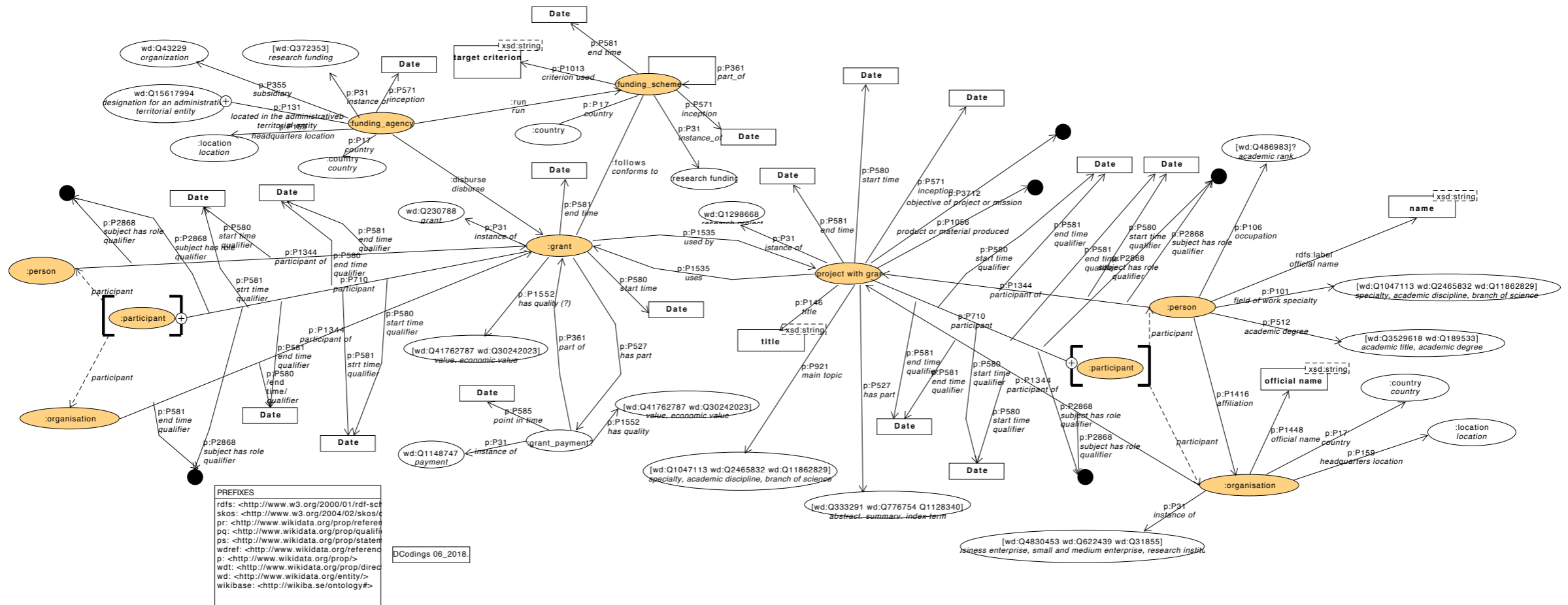
Branch: master schemas / Wikidata / research_funding / doc / GrantModel_RDFGraph.pdf

Find file Copy path

dcodings Revised and more complete version.

156533a 14 days ago

1 contributor



Practicing FAIR Data

6
Push to FDP
& Certified
Metrics Score

FAIR Metrics: <http://fairmetrics.org>

FAIR Metrics

The FAIR Metrics Group took-on the challenge of designing a framework for evaluating "FAIRness".

Discoverability and reusability are not abstract concepts, but imply concrete behaviors and properties that must hold true for the fulfillment of the FAIR objectives. Given this, it must therefore be possible to precisely define a measurable set of properties and behaviors that assess FAIRness. Over the short 1 month lifespan of the FAIR Metrics Working Group, we have created a cogent framework for developing FAIR metrics manifested as a simple form with 8 questions that structures fruitful conversations about proposed metrics.

Our approach recognizes that the diversity in opinion must play a key role in crafting fair and effective FAIR guidelines. Communities must not only understand what is meant by FAIR, but must also be able to monitor the FAIRness of their digital resources, in a realistic, but quantitative manner. We recognize that what is considered FAIR in one community may be quite different from FAIRness in another community - different community norms and practices make this a certainty! As such, our approach focuses on the mechanism by which metrics can be created by community members themselves, rather than attempting to create a set of one-size-fits-all metrics to apply to every resource.

With a mechanism in-place to design metrics, we now open the process of generating metrics to community participation. We have created several exemplar metrics that we think will be broadly applicable; however, additional metrics may be designed and published through **our open submission process**, or simply shared within your community through your normal communication channels.

Our proposed FAIR Metrics can be found [here](#).

We have selected an approach to publishing FAIR Metrics that is, itself, FAIR. This takes the form of a FAIR Accessor (a kind of Linked Data Platform Container), which describes a subset of metrics, the community to which they are applicable, other relevant metadata, and links to each

[FAIR Metrics GitHub](#)

[FAIR Metrics Paper](#)

[Metrics Process](#)

[Metrics Authoring Framework](#)

[Metrics Form](#)

[About Us](#)

Practicing FAIR Data

6
Push to FDP
& Certified
Metrics Score

FAIR Metrics: <http://fairmetrics.org>

FAIR Metrics

The FAIR Metrics Group took-on the challenge of designing a framework for evaluating "FAIRness".

Discoverability and reusability are not abstract concepts, but imply concrete behaviors and

FAIR Metrics GitHub

FAIR Metrics Paper

Process

ing Framework

Framework for authoring FAIR Metrics

broadly applicable, however, additional metrics may be designed and published through our **open submission process**, or simply shared within your community through your normal communication channels.

Our proposed FAIR Metrics can be found [here](#).

We have selected an approach to publishing FAIR Metrics that is, itself, FAIR. This takes the form of a FAIR Accessor (a kind of Linked Data Platform Container), which describes a subset of metrics, the community to which they are applicable, other relevant metadata, and links to each

Practicing FAIR Data

6
Push to FDP
& Certified
Metrics Score

FAIR Metrics: <http://fairmetrics.org>

FAIR Metrics

The FAIR Metrics Group took-on the challenge of designing a framework for evaluating "FAIRness".

Discoverability and reusability are not abstract concepts, but imply concrete behaviors and

FAIR Metrics GitHub

FAIR Metrics Paper

Process

ing Framework

- **Community defined**
- **Objective**
- **Quantifiable**
- **Reproducible**
- **Automatic (scalable)**
- **Certifiable**

broadly applicable, however, additional metrics may be designed and published through our **open submission process**, or simply shared within your community through your normal communication channels.

Our proposed FAIR Metrics can be found [here](#).

We have selected an approach to publishing FAIR Metrics that is, itself, FAIR. This takes the form of a FAIR Accessor (a kind of Linked Data Platform Container), which describes a subset of metrics, the community to which they are applicable, other relevant metadata, and links to each

Practicing FAIR Data

6
Push to FDP
& Certified
Metrics Score

FAIR Metrics: <http://fairmetrics.org>

FAIR Metrics

The FAIR Metrics Group took-on the challenge of designing a framework for evaluating

FAIR Metrics GitHub

www.nature.com/scientificdata

SCIENTIFIC DATA

OPEN Comment: A design framework and exemplar metrics for FAIRness

Mark D. Wilkinson¹, Susanna-Assunta Sansone², Erik Schultes³, Peter Doorn⁴, Luiz Olavo Bonino da Silva Santos^{5,6} & Michel Dumontier⁷

Received: 28 November 2017

Accepted: 9 May 2018

Published: 26 June 2018

The FAIR Principles¹ (<https://doi.org/10.25504/FAIRsharing.WWI10U>) provide guidelines for the publication of digital resources such as datasets, code, workflows, and research objects, in a manner that makes them Findable, Accessible, Interoperable, and Reusable (FAIR). The Principles have rapidly been adopted by publishers, funders, and pan-disciplinary infrastructure programmes and societies. The Principles are aspirational, in that they do not strictly define how to achieve a state of "FAIRness", but rather they describe a continuum of features, attributes, and behaviors that will move a digital resource closer to that goal. This ambiguity has led to a wide range of interpretations of FAIRness, with some resources even claiming to already "be FAIR"! The increasing number of such statements, the emergence

14 Core FAIR Metrics

Findable:

FM-F1A **FM-F1B**

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

FM-F2

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

FM-F3

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

FM-F4

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

Accessible:

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

FM-A1.1

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

FM-A1.2

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

FM-A2

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

Interoperable:

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

FM-I2

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

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

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**; **FM-R1.1**

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

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

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

<http://fairmetrics.org>

<https://github.com/FAIRMetrics/Metrics/blob/master/ALL.pdf>

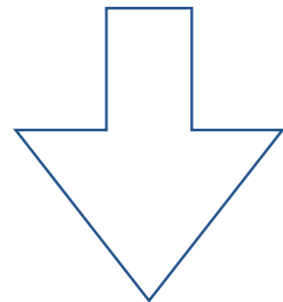
The FAIR Metrics Framework

<i>FIELD</i>	<i>DESCRIPTION</i>
Metric Identifier	FM-F1B: https://purl.org/fair-metrics/FM_F1B
Metric Name	Identifier persistence
To which principle does it apply?	F1
What is being measured?	Whether there is a policy that describes what the provider will do in the event an identifier scheme becomes deprecated.
Why should we measure it?	The change to an identifier scheme will have widespread implications for resource lookup, linking, and data sharing. Providers of digital resources must ensure that they have a policy to manage changes in their identifier scheme, with a specific emphasis on maintaining/redirecting previously generated identifiers.
What must be provided?	A URL that resolves to a document containing the relevant policy.
How do we measure it?	Use an HTTP GET on URL provided.
What is a valid result?	Present (a 200,202,203 or 206 HTTP response after resolving all and any prior redirects. e.g. 301 -> 302 -> 200 OK.) or Absent (any other HTTP code)
For which digital resource(s) is this relevant?	All
Comments	<p>A first version of this metric would focus on just checking a URL that resolves to a document. We can't verify that document.</p> <p>A second version would indicate how to structure the data policy document with a particular section (similar to how the CC license has a section for attribution in PDF)</p>

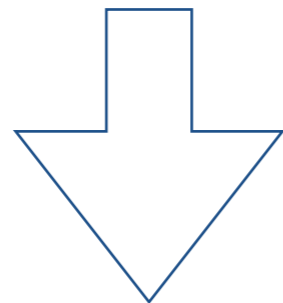
The FAIR Metrics Framework

Example: FM-F1B, Identifier Persistence

v1.0 **checks** for HTTP 200 return



v2.0 **validates** a standard RDF persistence policy

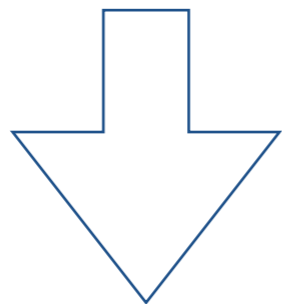


v3.0 **scores** multiple parameters of persistence policy

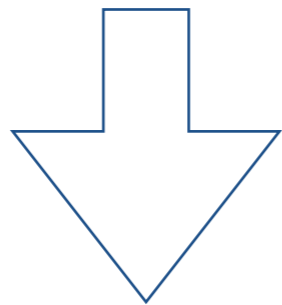
The FAIR Metrics Framework

Example: FM-F1B, Identifier Persistence

v1.0 **checks** for HTTP 200 return



v2.0 **validates** a standard RDF persistence policy



v3.0 **scores** multiple parameters of persistence policy

14 Core FAIR Metrics

21 Questions

22 Community Challenges

FAIR Principle F1: (meta) data are assigned globally unique and persistent identifiers. *Fundamental requirement for accurate and reproducible machine actionability. Examples: Universally unique identifier (UUID): https://en.wikipedia.org/wiki/Universally_unique_identifier ; Digital Object Identifier (DOI): <http://www.doi.org>*

FAIR Metric F1A:

Question 1: *Provide an URL to a registered scheme that defines the globally-unique structure of the identifier(s) for your digital resource.*

FAIR Metric F1B:

Question 2: *Provide an URL to a document that defines the persistence policy of your identifier(s).*

Community challenge:

- (1) What are your required (or preferred) identifier registration services ?
- (2) What is your minimal persistence policy?
- (3) Can you make your persistence policy machine-readable?

14 Core FAIR Metrics

21 Questions

22 Community Challenges

FAIR Principle F2: F2: Data are described with rich metadata

Data should be machine-discoverable by the widest range of stakeholders possible. That is, you should not presume who will want to use your data, or for what purpose. Resource providers should be generous and expansive with their metadata (see also R1). Exemplar rich metadata frameworks: DCAT; ISA framework.

FAIR Metric F2:

Question 3. *Provide the URL to a document that contains machine-readable metadata for the digital resource.*

Question 4. *Provide the URL for the file format of this metadata.*

Community challenge:

(4) Can you define a minimal set of metadata for your community?

(5) Can you make your metadata machine-readable?

14 Core FAIR Metrics

21 Questions

22 Community Challenges

FAIR Principle F3: Metadata clearly & explicitly include the identifier of the data it describes *Metadata and the dataset they describe are often separate (and probably should be, v.v. Principle A2). Because most data formats are not extendible, and therefore cannot always refer to the metadata, the association between metadata and the data can often only be achieved by explicit references to the data's globally unique identifier within the metadata record, thus facilitating ID-based search.*

FAIR Metric F3:

Question 5. *Provide an URL to the metadata document that contains the globally unique and persistent identifier for the digital resource.*

Question 6. *Provide the URL to the data described by in that metadata document.*

Community challenge:

(6) Can you define the metadata model that explicitly links data and metadata?

(7) Can you make this metadata model machine-readable?

FAIR Metrics Community Challenges

Find

- (1) What are your required (or preferred) identifier registration services ?
- (2) What is your minimal persistence policy?
- (3) Can you make your persistence policy machine-readable?
- (4) Can you define a minimal set of metadata for your community?
- (5) Can you make your metadata machine-readable?
- (6) Can you define the metadata model that explicitly links data and metadata?
- (7) Can you make this metadata model machine-readable?
- (8) What is the required (preferred) search engine for your community ?

Access

- (9) What is the required (preferred) communication protocol for your community ?
- (10) What is your required (preferred) protocols for restricting access to data ?
- (11) Can you make this protocol machine-readable?
- (12) What is your minimal persistence policy for metadata?
- (13) Can you make this persistence policy machine-readable?

Interoperate

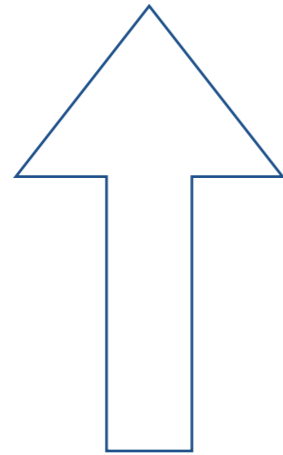
- (14) What is your required (preferred) standards in knowledge representation ?
- (15) What are your required (preferred) vocabularies ?
- (16) What is your required LinkSet ?

Reuse

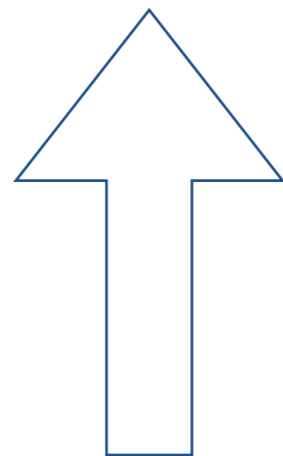
- (17) What is your required (preferred) usage license framework?
- (18) Can you make these usage licenses machine-readable?
- (19) What is your required (preferred) provenance metadata descriptions?
- (20) Can you make this provenance metadata machine-readable?
- (21) What are your certification criteria for data & metadata?
- (22) What is your machine-actionable validation-certification system ?

FAIR Evaluator

FAIR Evaluator (arbitrary collections of FAIR Metrics)



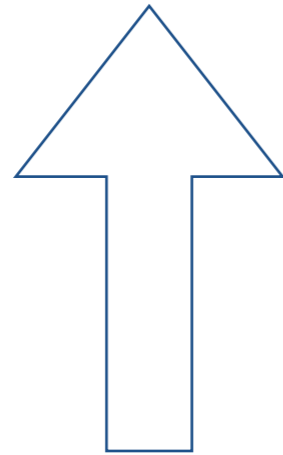
14 Core FAIR Metrics



15 FAIR Principles

FAIR Evaluator

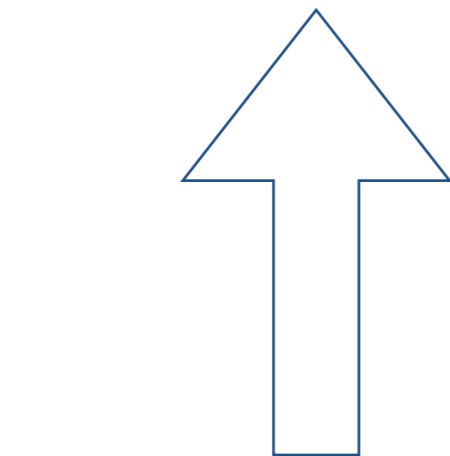
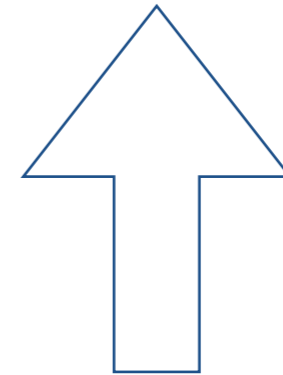
FAIR Evaluator (arbitrary collections of FAIR Metrics)



14 Core FAIR Metrics



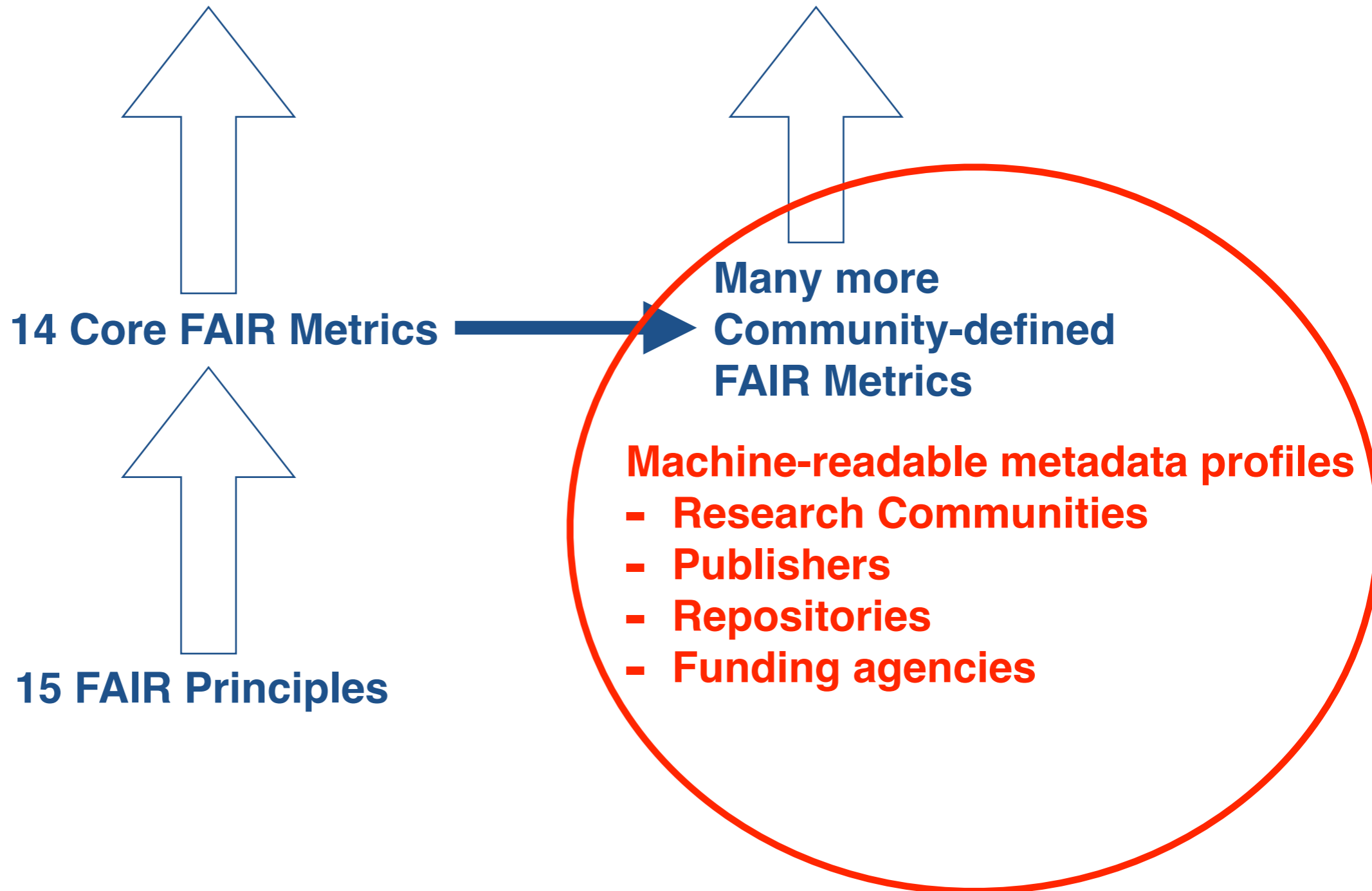
**Many more
Community-defined
FAIR Metrics**



15 FAIR Principles

FAIR Evaluator

FAIR Evaluator (arbitrary collections of FAIR Metrics)



14 Core FAIR Metrics

**Many more
Community-defined
FAIR Metrics**

15 FAIR Principles

Machine-readable metadata profiles

- **Research Communities**
- **Publishers**
- **Repositories**
- **Funding agencies**



Metadata IN

<https://www.go-fair.org/implementation-networks/overview/>