

# Implementing FAIR

## Consideration from Solid Earth Sciences

*Daniele Bailo, EPOS IT Officer, INGV*

*Rossana Paciello, EPOS-ERIC, INGV*

A photograph of a severely damaged building, likely a government palace, with rubble and debris in the foreground. The text 'PALAZZO DEL GOVERNO' is visible on the facade.

EARTHQUAKES

A photograph of a volcanic eruption, showing a large plume of white smoke or ash rising from a dark, rocky landscape. A red line is visible on the right side of the image.

VOLCANIC ERUPTIONS

A photograph of a tsunami, showing a large wave crashing over a coastal area with buildings and infrastructure.

TSUNAMIS

A photograph of a tectonic landscape, showing a large, rocky mountain range with a deep, narrow valley or rift.

SURFACE DYNAMICS  
& TECTONICS

A photograph of a geothermal or hydrothermal system, showing a large, rocky structure with water splashing or bubbling around it.

GEORESOURCES

# Topics

- a) FAIR principles
- b) FAIR «Topics»:
  - a) Data, metadata, semantics, storage, PID, AAAI...
- c) How to implement FAIR? Existing initiatives and tools
- d) «The FAIR implementation process»:  
EPOS experience on how to implement FAIR in the real world



# The principles precede implementation.

They act as a **guide to data publishers** and **stewards to assist them in evaluating** whether their particular implementation choices are rendering their digital research artefacts Findable, Accessible, Interoperable, and Reusable.

*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).*

# FAIR: evaluation & best practices



<https://www.rd-alliance.org/groups/fair-data-maturity-model-wg>

FAIR Implementation Profile



From FAIR Leading Practices to FAIR Implementation and Back: An Inclusive Approach to FAIR at Leiden University Libraries

**Authors:** [Kristina Maria Hettne](#) ✉, [Peter Verhaar](#), [Erik Schultes](#), [Laurents Sesink](#)

*Papers: dicussion on Practices, training, consultancy*



<https://fairsharing.github.io/>

# Europe Embraces FAIR principles

2018



## 1.7.2 Step 2: Implement – culture, technology and skills for FAIR

- » Rec. 4: Develop interoperability frameworks for FAIR sharing within disciplines and research
- » Rec. 5: Ensure Data Management via DMPs
- » Rec. 6: Recognise and reward FAIR data and data stewardship
- » Rec. 7: Support semantic technologies
- » Rec. 8: Facilitate automated processing
- » Rec. 9: Develop assessment frameworks to certify FAIR services
- » Rec. 10: Professionalise data science and data stewardship roles and train researchers
- » Rec. 11: Implement curriculum frameworks and training

# RI Implementer perceptions

FAIR implementation  
For system architect  
and Engineers

DIY (Do It Yourself ) activity





# Whah are FAIR principles really about?

# Fair Data principles

## To be Findable:

- F1. (meta)data are assigned a globally unique and persistent identifier
- F2. data are described with rich metadata (defined by R1 below)
- F3. metadata clearly and explicitly include the identifier of the data it describes
- F4. (meta)data are registered or indexed in a searchable resource

## To be 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

## To be 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

## To be 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 Data principles

Topics:

- Data

## To be Findable:

- F1. (meta)data are assigned a globally unique and persistent identifier
- F2. data are described with rich metadata (defined by R1 below)
- F3. metadata clearly and explicitly include the identifier of the data it describes
- F4. (meta)data are registered or indexed in a searchable resource

## To be 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

## To be 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

## To be 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 Data principles

## Topics:

- Data
- Metadata

### To be Findable:

- F1. (meta)data are assigned a globally unique and persistent identifier
- F2. data are described with rich metadata (defined by R1 below)
- F3. metadata clearly and explicitly include the identifier of the data it describes
- F4. (meta)data are registered or indexed in a searchable resource

### To be 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

### To be 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

### To be 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 Data principles

## Topics:

- Data
- Metadata
- Identifiers

### To be Findable:

- F1. (meta)data are assigned a globally unique and persistent identifier
- F2. data are described with rich metadata (defined by R1 below)
- F3. metadata clearly and explicitly include the identifier of the data it describes
- F4. (meta)data are registered or indexed in a searchable resource

### To be 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

### To be 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

### To be 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 Data principles

## To be Findable:

- F1. (meta)data are assigned a globally unique and persistent identifier
- F2. data are described with rich metadata (defined by R1 below)
- F3. metadata clearly and explicitly include the identifier of the data it describes
- F4. (meta)data are registered or indexed in a searchable resource

## To be 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

## To be 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

## To be 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

## Topics:

- Data
- Metadata
- Identifiers
- Protocols



# Fair Data principles

## To be Findable:

- F1. (meta)data are assigned a globally unique and persistent identifier
- F2. data are described with rich metadata (defined by R1 below)
- F3. metadata clearly and explicitly include the identifier of the data it describes
- F4. (meta)data are registered or indexed in a searchable resource

## To be 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

## To be 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

## To be 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

## Topics:

- Data
- Metadata
- Identifiers
- Protocols
- Authentication & Authorization



# Fair Data principles

## To be Findable:

- F1. (meta)data are assigned a globally unique and persistent identifier
- F2. data are described with rich metadata (defined by R1 below)
- F3. metadata clearly and explicitly include the identifier of the data it describes
- F4. (meta)data are registered or indexed in a searchable resource

## To be 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

## To be 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

## To be 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

## Topics:

- Data
- Metadata
- Identifiers
- Protocols
- Auth&Autho

Standards





# Fair Data principles

## To be Findable:

- F1. (meta)data are assigned a globally unique and persistent identifier
- F2. data are described with rich metadata (defined by R1 below)
- F3. metadata clearly and explicitly include the identifier of the data it describes
- F4. (meta)data are registered or indexed in a searchable resource

## To be 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

## To be 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

## To be 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

## Topics:

- Data
- Metadata
- Identifiers
- Protocols
- Auth&Autho

Standards  
Semantics



# Fair Data principles

## To be Findable:

- F1. (meta)data are assigned a globally unique and persistent identifier
- F2. data are described with rich metadata (defined by R1 below)
- F3. metadata clearly and explicitly include the identifier of the data it describes
- F4. (meta)data are registered or indexed in a searchable resource

## To be 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

## To be 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

## To be 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

## Topics:

- Data
- Metadata
- Identifiers
- Protocols
- Auth&Autho

Standards  
Semantics

policy &  
provenance



# Fair Data principles

## To be Findable:

- F1. (meta)data are assigned a globally unique and persistent identifier
- F2. data are described with rich metadata (defined by R1 below)
- F3. metadata clearly and explicitly include the identifier of the data it describes
- F4. (meta)data are registered or indexed in a searchable resource

## To be 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

## To be 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

## To be 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

## Hidden Topics:

- services+databases

## Topics:

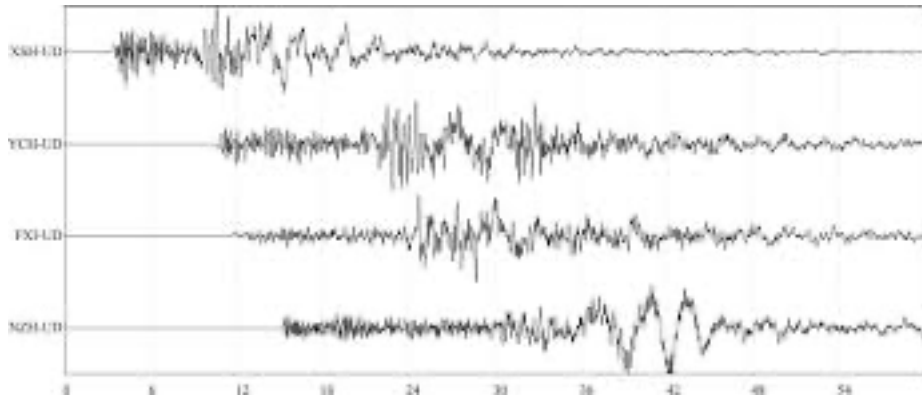
- Data
- Metadata
- Identifiers
- Protocols
- Auth&Autho

Standards  
Semantics

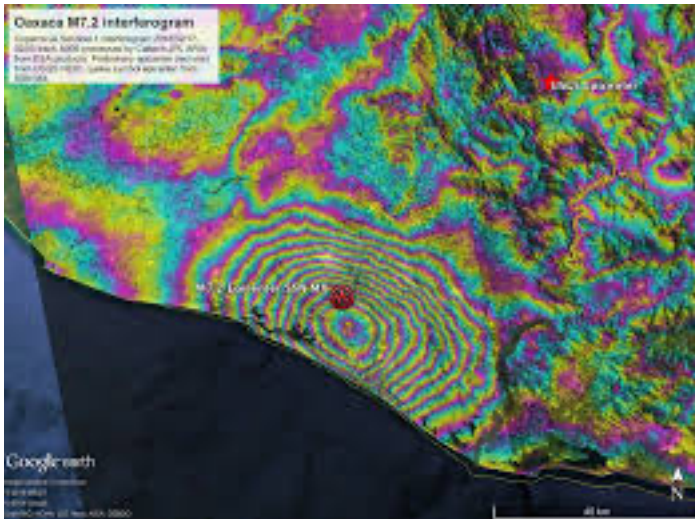
policy &  
provenance

# FAIR Topics





*Seismic Waveform*



*InSAR image*

## Data

*Our greatest wealth*

## Function

*Describe (physical) phenomena*

## Related topics & Issues

- *Format, standards, harmonization*
- *Proliferation standards*
- *Continuous or atomic data?*
- *How to store data?*

# How to store data?



## Lev. 1: Researcher Hard Disk



## Lev. 2: Institutional Services



## Lev. 3a: Certified Repository

- Lin, D., Crabtree, J., Dillo, I. *et al.* The TRUST Principles for digital repositories. *Sci Data* **7**, 144 (2020). <https://doi.org/10.1038/s41597-020-0486-7>
- <https://www.coretrustseal.org/>



## Lev. 3b: 3rd party Repository

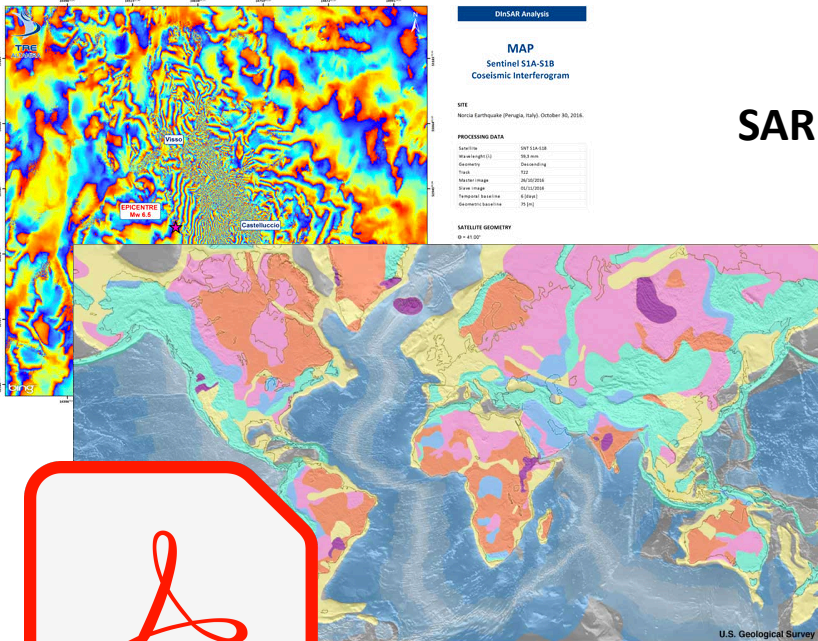
- <https://zenodo.org/>



# Static Data

vs

# Dynamic Data

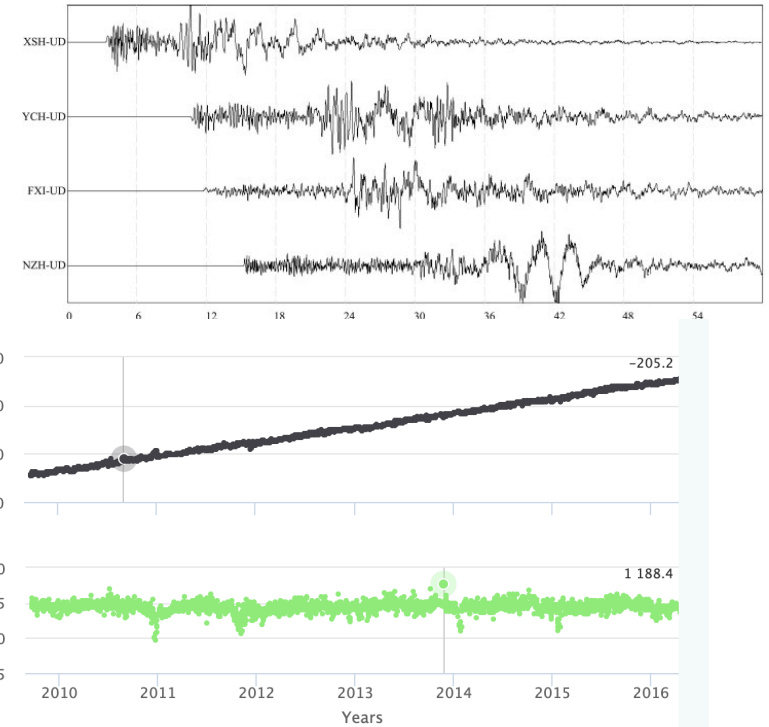


SAR images

Geological maps

Seismic Waveforms produced by stations

GNSS displacement time-series

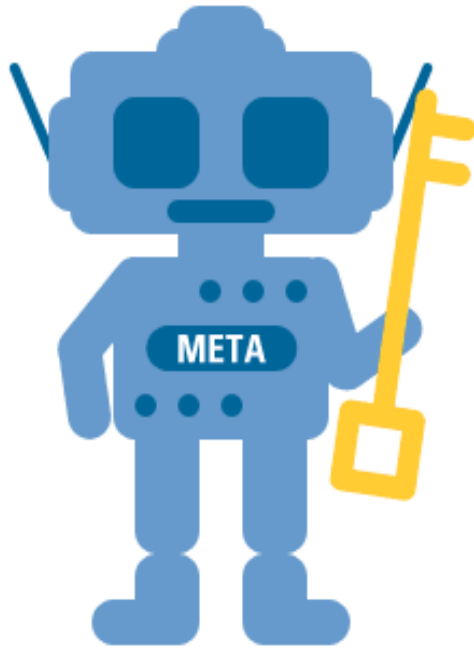


Different solutions for access and identification

## FAIR related Topics:

- ~~Data~~
- Metadata
- Identifiers
- Protocols
- Authentication & Authorisation
- Standards
- Semantics
- policy & provenance
- Services and databases





Digital Object:

**DO:** <data, metadata, handle>

Kahn, R., & Wilensky, R. (2006). A framework for distributed digital object services. *International Journal on Digital Libraries*, 6(2), 115–123.

<https://doi.org/10.1007/s00799-005-0128-x>

## Metadata

*Data about Data*

### **Purposes**

1. *Discovery (humans & machines)*
2. *Contextualization*
3. *Use it for processing or other advanced tasks*

*Usually attached to D.O.*

### **Related topics & Issues**

- *Data or metadata?*
- *Many standards*
- *Catalogue*
- *Ontologies*

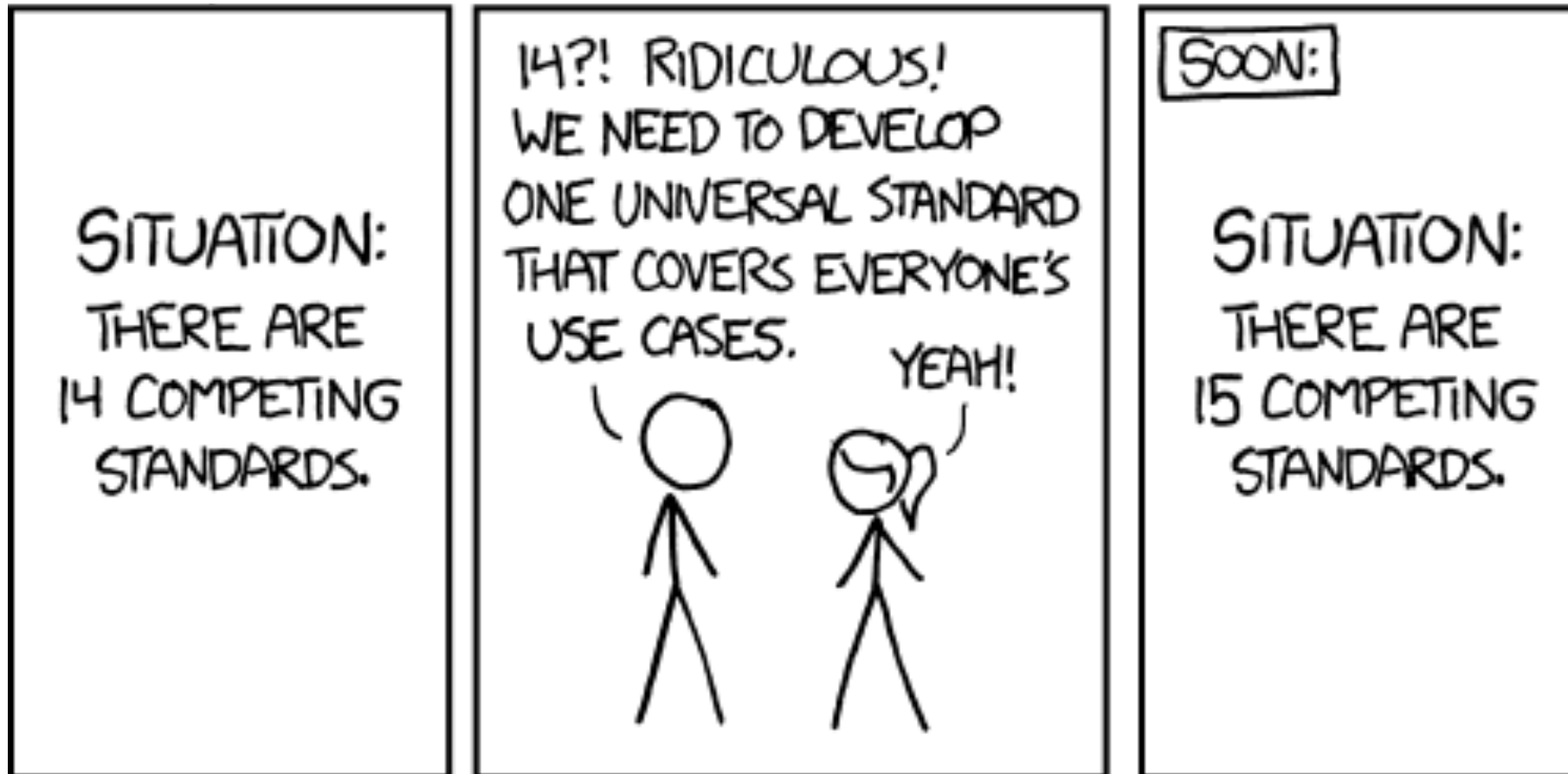
# (Metadata) standards in specific scientific domain

*...a real story...*



# (Metadata) standards in specific scientific domain

*...a real story...*



# GEOJSON



DCAT-AP



## Metadata

*What is the best standard?*

**DEPENDS ON THE REQUIREMENTS**

*FAIR principles require*

- *Rich metadata standard*
- *Usage of formal, accessible, shared, and broadly applicable language for knowledge representation.*

*Serialization and format:  
two different things (almost)*



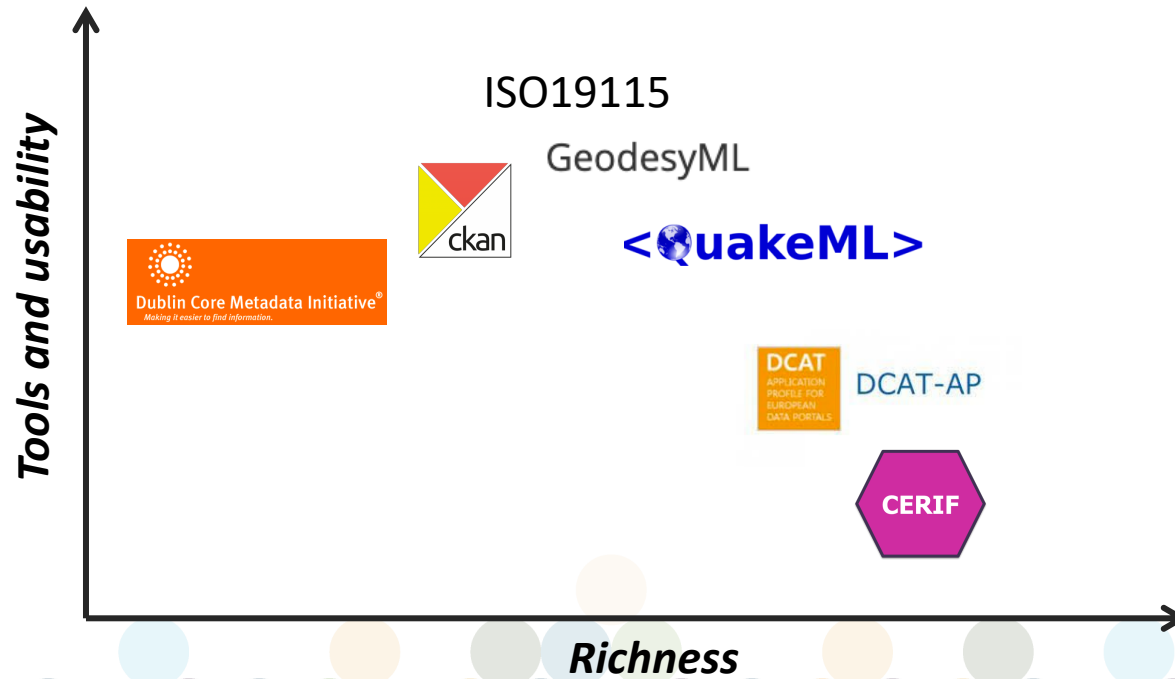
Dublin Core Metadata Initiative®

*Making it easier to find information.*



# Rich Metadata

1. including descriptive information about the **context, quality and condition**, or characteristics of the data.
2. someone should be able to **find data based on the information provided by their metadata**, even without the data's identifier



## Book – Dublin Core example

```
<rdf:Description>
  <dc:creator>Peter Noeller</dc:creator>
  <dc:title>Algebra</dc:title>
  <dc:subject>mathematics</dc:subject>
  <dc:date>2008-04-23</dc:date>
  <dc:language>EN</dc:language>
  <dc:description>
    An Introduction to Algebra
  </dc:description>
</rdf:Description>
```

## Person – DCAT-AP Example

```
<http://orcid.org/0000-0001-7750-7254> a schema:Person ;
  schema:identifier [ a schema:PropertyValue ;
    schema:propertyID "orcid" ;
    schema:value "0000-0001-7750-7254" ;
  ];
  schema:identifier [ a schema:PropertyValue ;
    schema:propertyID "ScopusAuthorID" ;
    schema:value "55861456000" ;
  ];
  schema:familyName "Trendl" ;
  schema:givenName "Lucas" ;
  schema:address [ a schema:PostalAddress ;
    schema:streetAddress "Wolrechtsweg, 207" ;
    schema:addressLocality "De Bilt" ;
    schema:postalCode "3720CA" ;
    schema:addressCountry "The Netherlands" ;
  ];
  schema:email "lucastrendl@private.com" ;
  schema:telephone "+310302200207" ;
  schema:url "http://orcid.org/0000-0001-7750-7254"^^xsd:anyURI ;
  schema:qualifications "Senior Advisor, Researcher" ;
  schema:affiliation <PIC:999518944> ;
  schema:contactPoint <http://orcid.org/0000-0001-7750-7254/legalContact> ;
  schema:contactPoint <http://orcid.org/0000-0001-7750-7254/scientificContact> ;
  schema:contactPoint <http://orcid.org/0000-0001-7750-7254/financialContact> ;
```

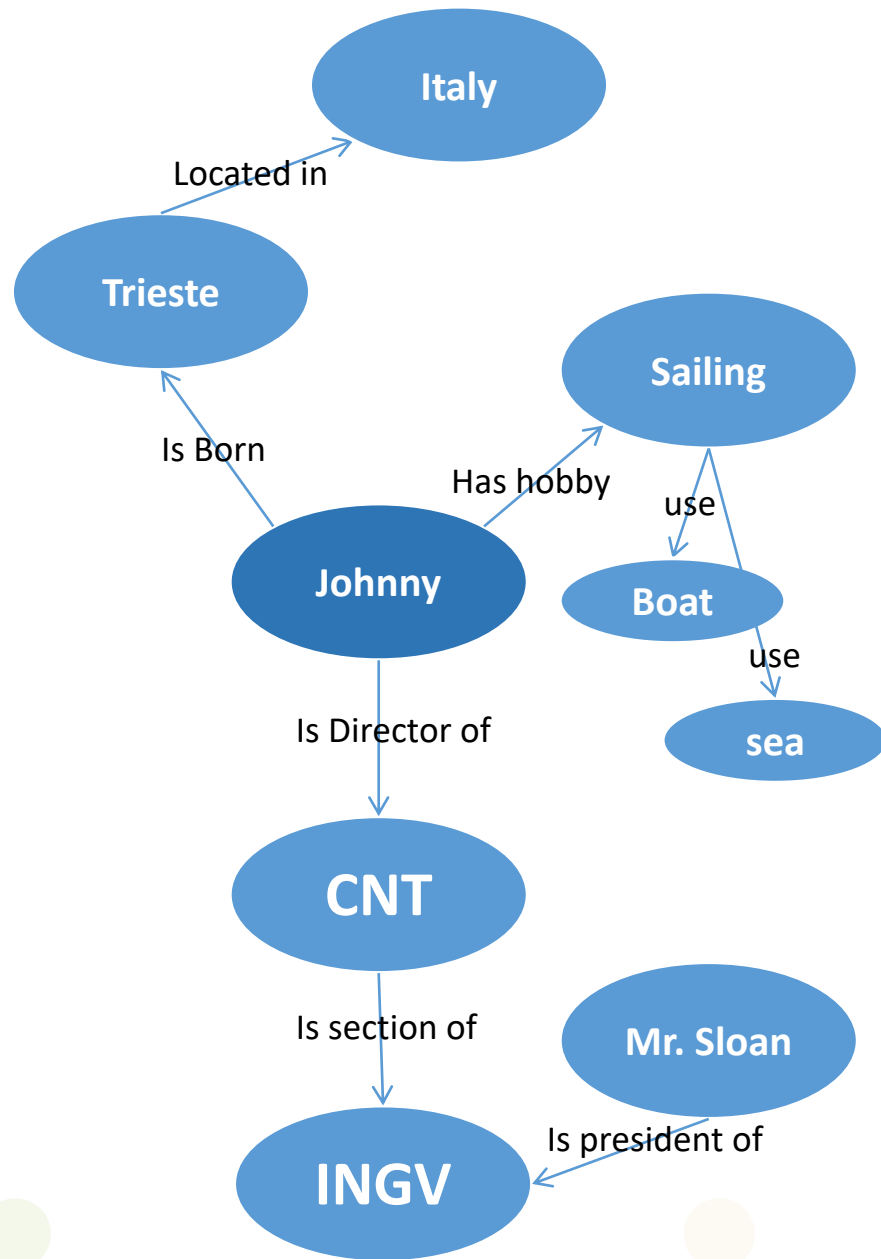
## FAIR related Topics:

- ~~— Data~~
- ~~— Metadata~~
- Identifiers
- Protocols
- Authentication & Authorisation
- Standards
- Semantics
- policy & provenance
- Services and databases

*Metadata:*  
*applicable language for  
knowledge representation*







# Ontologies & semantics

## *Why an ontology?*

It is the way machines manage “meaning”

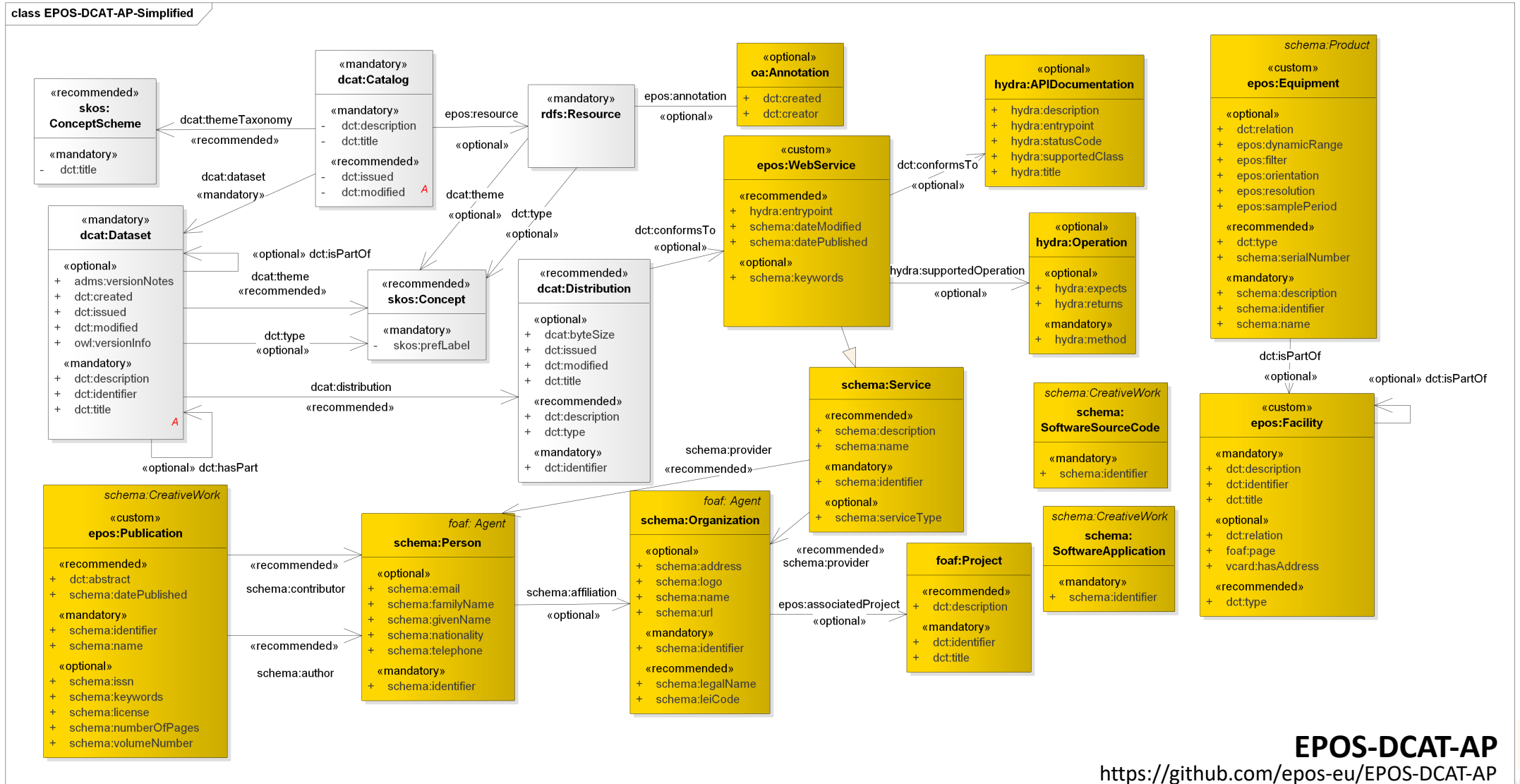
## *How does it work?*

1. *Connects concepts*
2. *Needs vocabulary*

## *Issues*

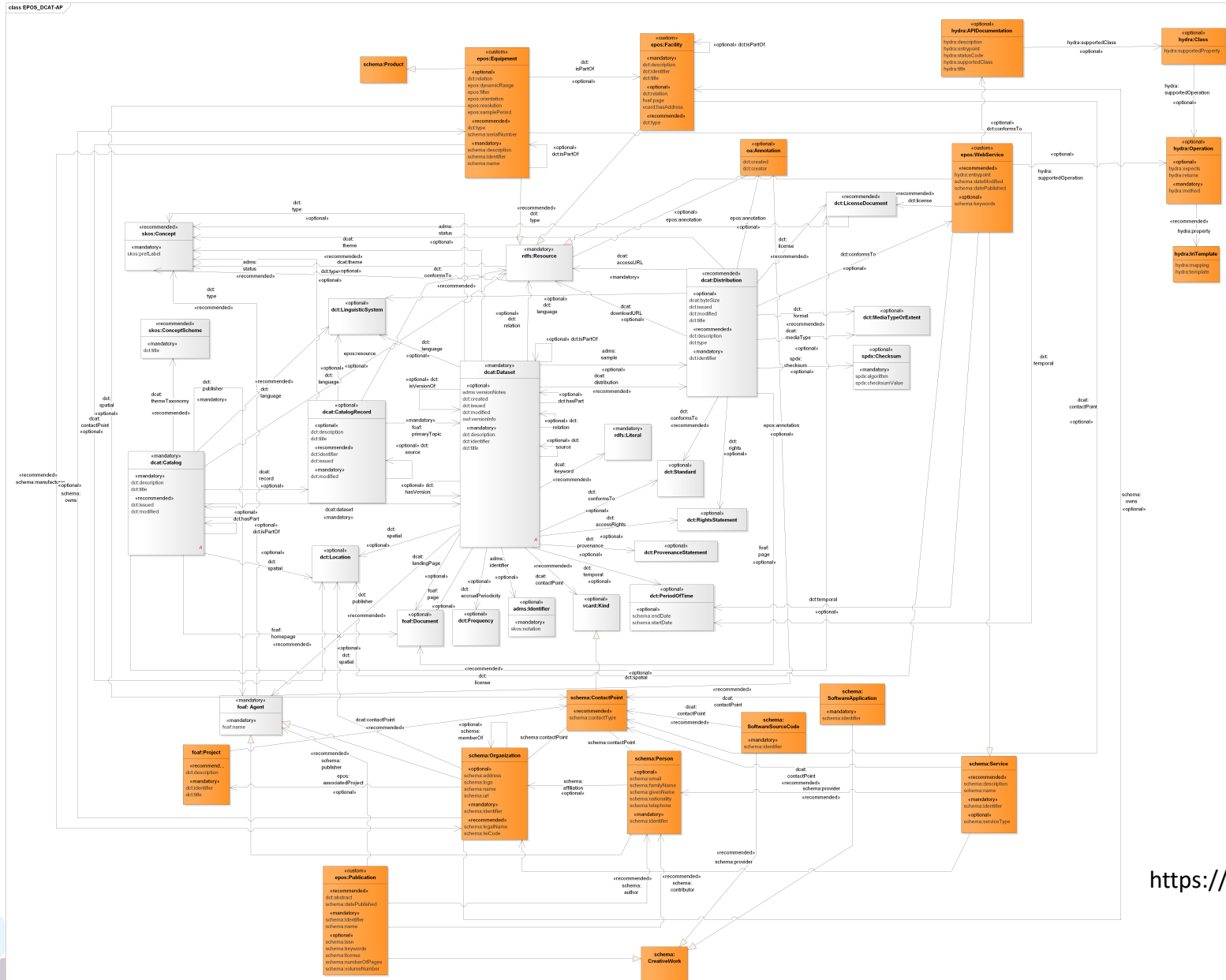
- Many ontologies exist
- Vocabulary Mapping

# Ontologies in real world (human - readable)



Metadata and ontologies

# Ontologies in real world (machine - readable)



**EPOS-DCAT-AP**  
<https://github.com/epos-eu/EPOS-DCAT-AP>





## FAIR related Topics:

- ~~— Data~~
- ~~— Metadata~~
- Identifiers
- Protocols
- Authentication & Authorisation
- Standards
- ~~— Semantics~~
- policy & provenance
- Services and databases



ORCID

<http://dx.doi.org/doi:10.30/tqk>  
<http://hdl.handle.net/hdl:13030/tqk>  
<http://purl.org/tqk>  
 ... <urn:13030:tqk>  
<http://n2t.net/ark:/13030/tqk>  
<http://OwlBike.example.org/ark:/13030/tqk>

## How to register/cite data or publications?

### **PID system**

#### *Purpose*

- *DO / publication can be uniquely referenced*
- *Assign a PID at data creation times*

#### *Issues*

- *Dataset and publications are two different things*
- *Contracts with PID issuer*
- *How to identify data streams?*
- *Landing page problem*

## FAIR related Topics:

- ~~— Data~~
- ~~— Metadata~~
- ~~— Identifiers~~
- Protocols
- Authentication & Authorisation
- Standards
- ~~— Semantics~~
- policy & provenance
- Services and databases



<definition> «Protocol»

a defined set of rules and regulations that determine how data is transmitted in telecommunications and computer networking

# Protocols

## *web service example*

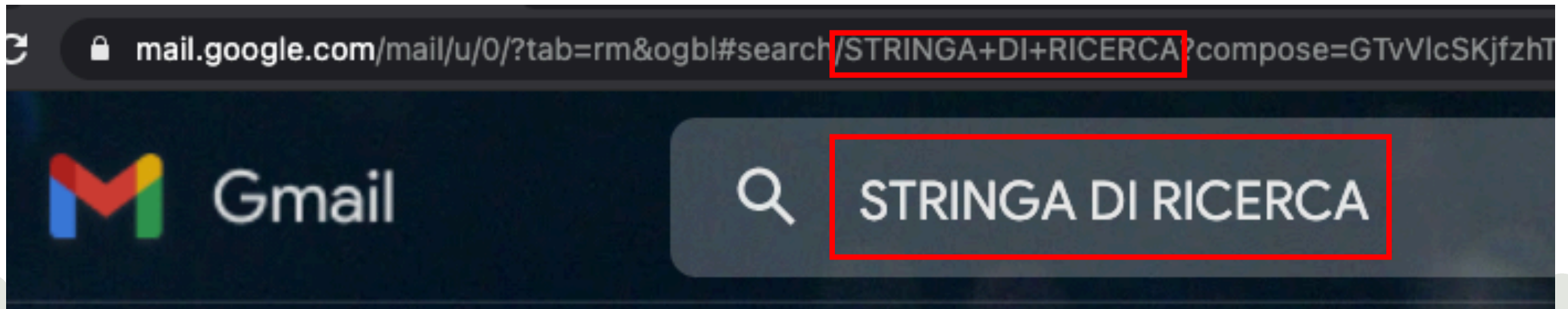
### *What is it?*

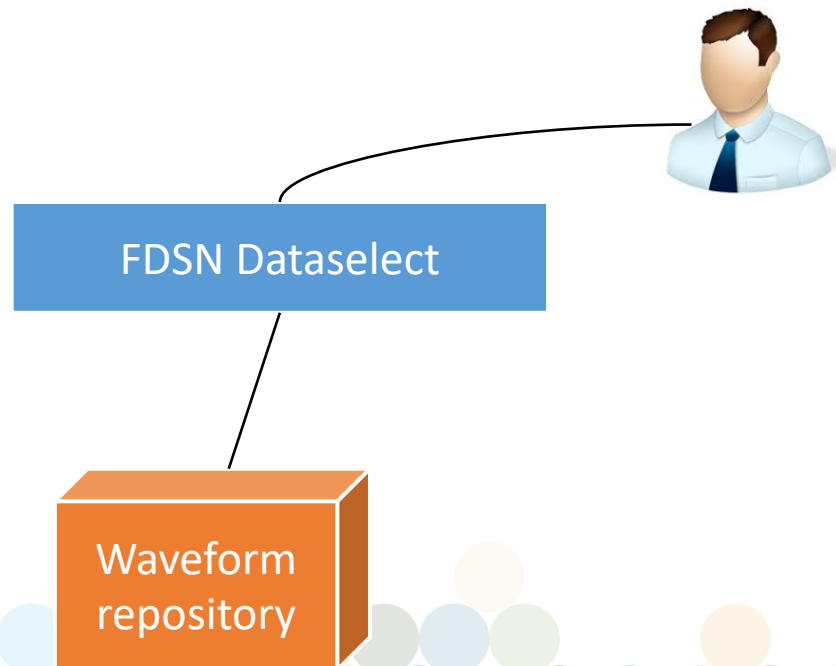
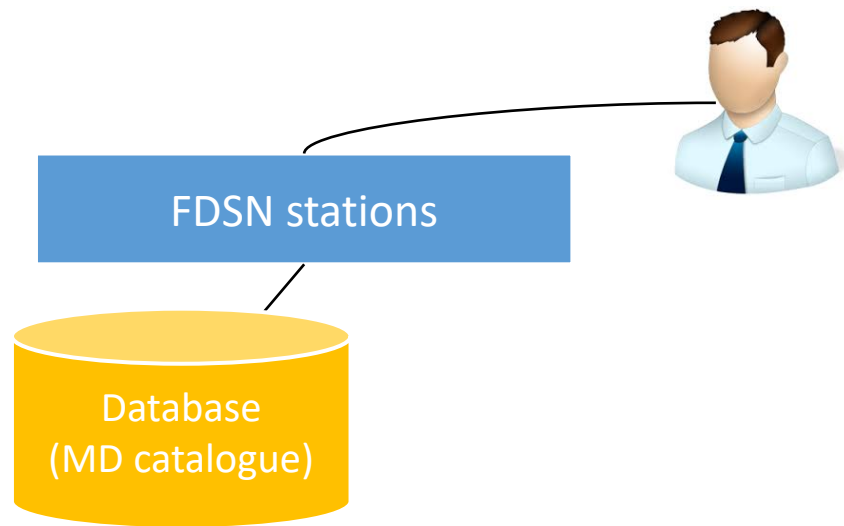
*server software running on a computer device, listening for requests and serving web documents (HTML, JSON, XML, images)*

Base URL  
service provider

Methods / collections / etc.

Parameters





# Protocols

## *web service example*

### *What is it?*

*server running on a computer device, listening for requests and serving web documents (HTML, JSON, XML, images)*

### *A "thin layer"*

*We usually don't know what's under the hood*

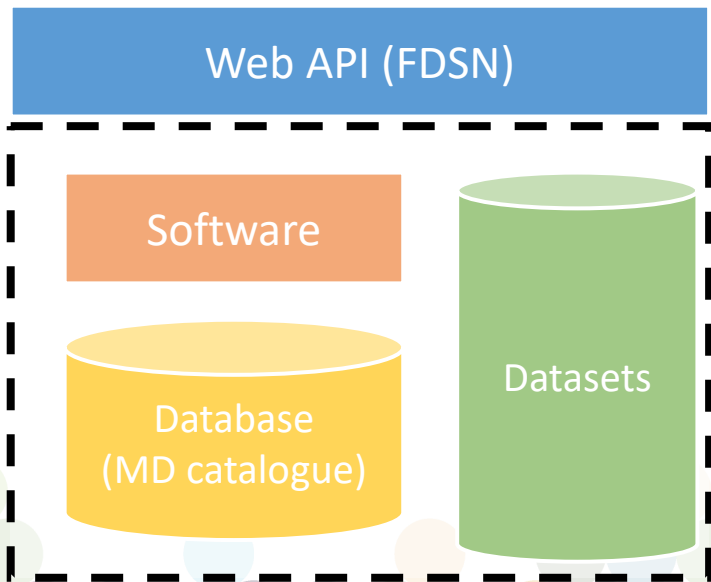
### *Examples*

- *FDSN stations*
- *FDSN dataselct*



System 1  
GUI Client

Agreed protocol



System 2  
(Black Box model)

# Protocols Interoperability

## What & Why

Enables 2 system to

1. Exchange information
2. Understand information

## Usually achieved through:

- Agreed language (protocols)
- Software “translators” interfaces → thin layers → webapis
- Standards (http, POP3, IMAP, ftp)



## FAIR related Topics:

- ~~Data~~
- ~~Metadata~~
- ~~Identifiers~~
- ~~Protocols~~
- Authentication & Authorisation
- ~~Standards~~
- ~~Semantics~~
- policy & provenance
- Services and databases



ORCID

Google



## How to access data?

# AAI system (federated & distributed)

### *Purpose*

- *Authenticate users*
- *Authorize users*

### *Topics*

- *Identity Provider*
- *Federation*
- *Authorization*
- *AAI integration*

# FAIR AAI - real life problems



I'm a Data Infrastructure Manager. How can I implement an AAI system?

Building in-house software can be resource consuming. Do off-the-shelf solutions exist?

my Data  
infrastructure



# FAIR AAAI - real life problems



Hi. I'm eduGAIN Identity Provider, I can help you.

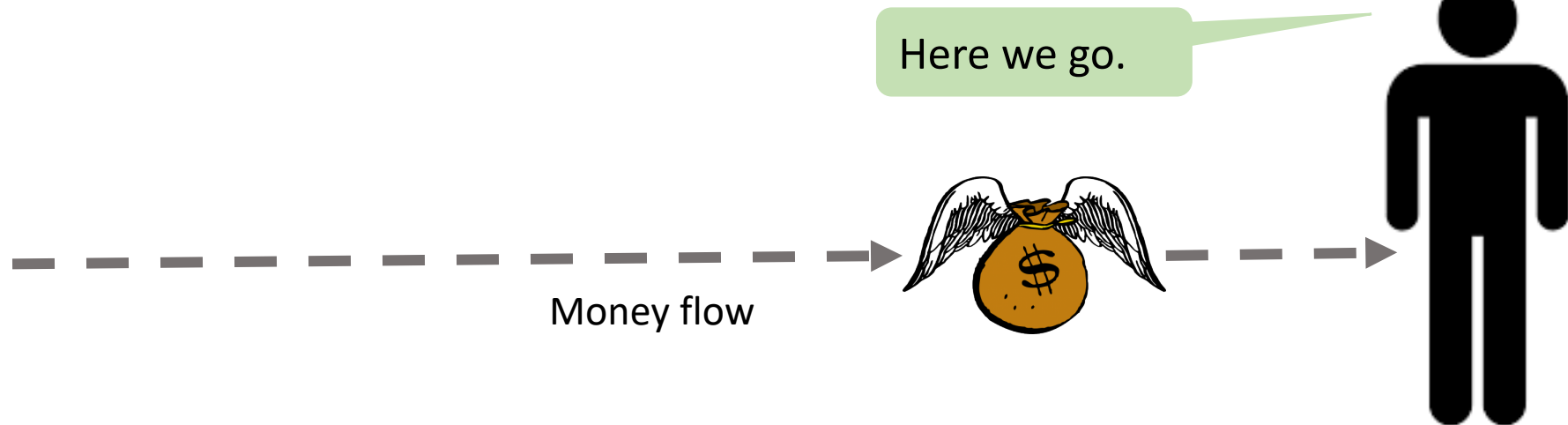
**You, as Service Providers** may delegate the authentication to **Identity Providers** in order to control access to the provided resources.



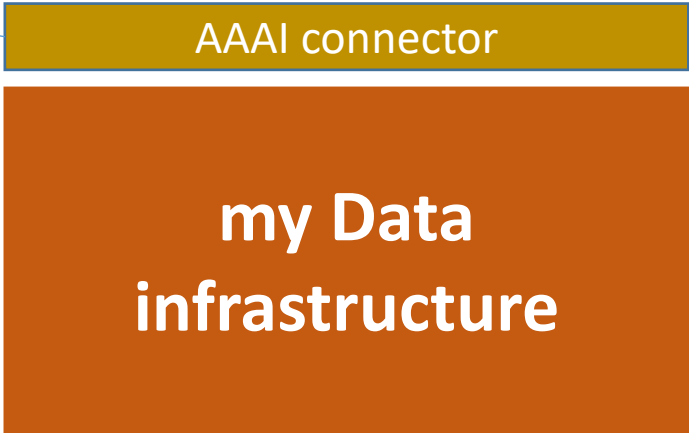
**NB: eduGAIN is here used merely as an example**

my Data  
infrastructure

# FAIR AAI - real life problems



Identity Provider  
Infrastructure  
(e.g. eduGAIN)



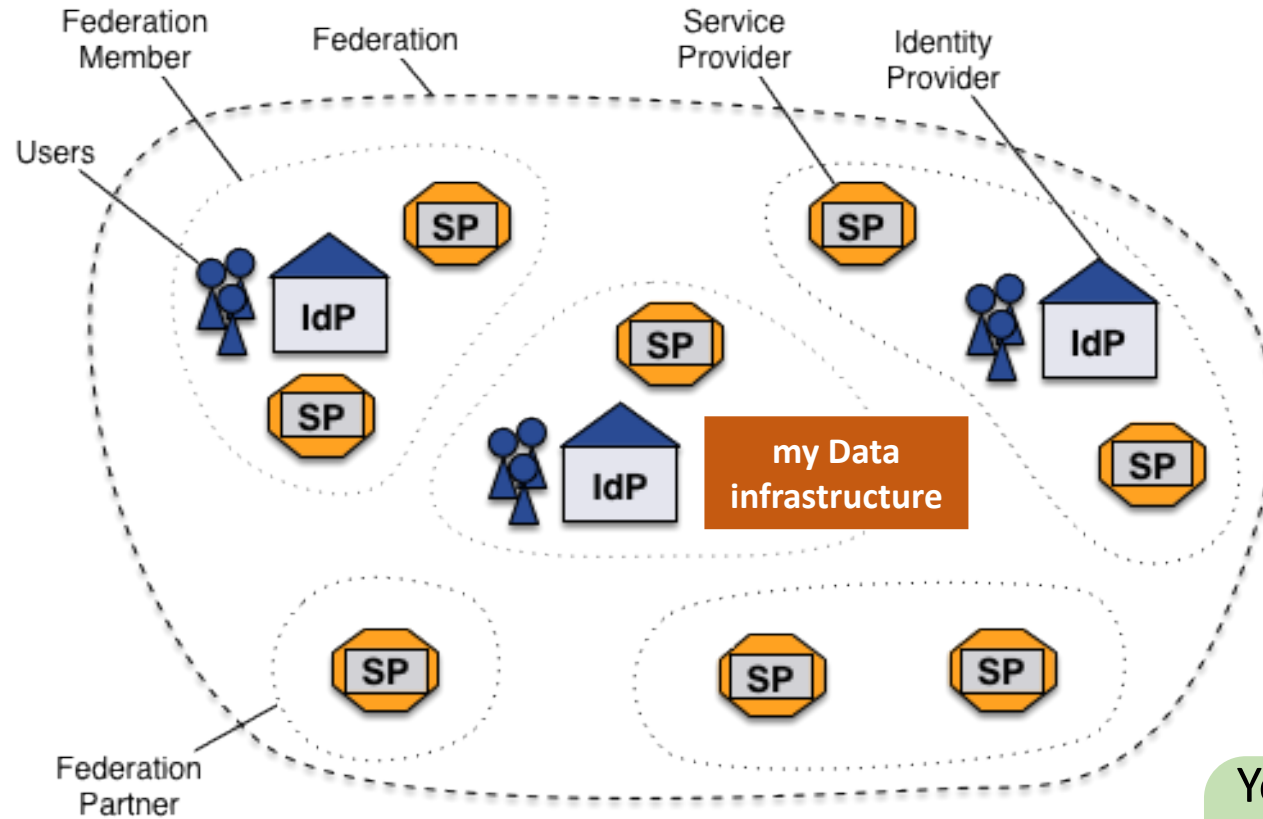
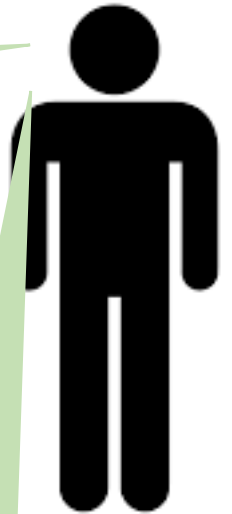
NB: eduGAIN is here used merely as an example

Photo by [Christina @ wocintechchat.com](https://www.wocintechchat.com) on [Unsplash](https://www.unsplash.com)

# FAIR AAAI - real life problems



We can do more.



You can share resources (data) with other Organizations in the federation

NB: eduGAIN is here used merely as an example

Photo by [Christina @ wocintechchat.com](#) on [Unsplash](#)

# FAIR AAAI - real life problems

Sharing resources (data) with other Organizations  
in the federation

=

Defining common **authorization** schema, e.g.



#	Level	Required attributes	Requirements for user	Rights
0	untrusted (anonymous)	(none)	no email confirmation	just access metadata
1	normal user	user	email confirmation (not confirmed name)	access open resources (e.g services)
2	resource access (including workflows and sensor networks, TNA?)	user & access	trusted affiliation and name	access open resources and resources that require trusted users
3	access to embargoed resources	users & specific restricted access attribute (e.g. EPOS IP member)	trusted affiliation, name and user group	access to embargoed resources
4	processing	user & processing	trusted affiliation and user role	access to computational resources
5	admin	admin	manually added to admin group	

Requires discussion and definition of:

- Data policy
- GDPR
- ...others



# FAIR related Topics:

- ~~— Data~~
- ~~— Metadata~~
- ~~— Identifiers~~
- ~~— Protocols~~
- ~~— Authentication & Authorisation~~
- ~~— Standards~~
- ~~— Semantics~~
- ~~— policy & provenance~~
- Services and databases



Leave it open for now...

# How to implement FAIR?

State of the art

# FAIR Implementation in EU Initiatives

## Projects and initiatives



FAIRsFAIR - Fostering Fair Data Practices in Europe - aims to supply practical solutions for the use of the FAIR data principles throughout the research data life cycle.



- Metadata 4 Machines: metadata standard
- FAIR Implementation Profile working group: FAIR evaluation
- FAIR Data Point: metadata repository that provides access to metadata in a FAIR way

**Other initiatives exist. In general most initiatives are about:**  
*Evaluations, metadata discussions, training, adoption of 3<sup>o</sup> party certification, best practices*

## Community specific initiatives



Promotes building of FAIR Ris in the environmental domain. This will enhance the efficiency and productivity of researchers, support innovation and connect the [ENVRI](#) Cluster to the [EOSC](#).

**FAIR implementation is delegated to the communities, with support in terms of best practices, tools, training**

# Standards and Certification initiatives



<https://www.coretrustseal.org/>

<https://doi.org/10.5281/zenodo.3638211>

Certification based on the [DSA–WDS Core Trustworthy Data Repositories Requirements](#):

Access Rights Information, Appraisal  
 Archive, Authenticity, Consumer,  
 Curation, Data  
 Database, Dataset, Designated  
 Community,  
 Digital Object, Digital Preservation ....

German Standards Committee (DIN) adopted **DIN 31644 Information and documentation - Criteria for trustworthy digital archives.**

Consultative Committee for Space Data Systems released recommended practice on "**Audit and certification of trustworthy digital repositories**", published as *ISO 16363 2012* ([ISO,2012b](#)).



<https://doi.org/10.1038/s41597-020-0486-7>

Transparency  
 Responsibility  
 User Focus  
 Sustainability  
 Technology



Digital Preservation Coalition

<https://www.dpconline.org/handbook>



# Implementation in Literature (excerpt)

## FAIR Principles: Interpretations and Implementation Considerations

Annika Jacobsen et al. *Data Intelligence* 2020 2:1-2, 10-29

## The FAIR Principles: First Generation Implementation Choices and Challenges

Barend Mons et al. *Data Intelligence* 2020 2:1-2, 1-9

[https://doi.org/10.1162/dint\\_e\\_00023](https://doi.org/10.1162/dint_e_00023)

## OSSE Goes FAIR - Implementation of the FAIR Data Principles for an Open-Source Registry for Rare Diseases.

Schaaf, J. et al. *Studies in Health Technology and Informatics*, 253, 209–213. <https://doi.org/10.3233/978-1-61499-896-9-209>

## Perspectives on the implementation of FAIR principles in solid Earth Research Infrastructures.,

Bailo, Daniele, et al.



*FrEaS* 8 (2020): 3. <https://doi.org/10.3389/feart.2020.00003>

**Inputs for RI managers/implementers:  
Best practices, discussion specific topics, experiences**



## Volume 2, No. 1-2 - Special Issue on Emergent FAIR Practices.

Issue Editors: Barend Mons, Erik Schultes & Annika Jacobsen



### Implementation Articles

-  **Unique, Persistent, Resolvable: Identifiers as the Foundation of FAIR**  
 Nick Juty , Sarala M. Wimalaratne, Stian Soiland-Reyes, John Kui and Tim Clark  
*Data Intelligence* Winter-Spring 2020, Vol. 2, No. 1-2, pp. 30–39



[Abstract](#) [Full Text](#) [PDF \(225 KB\)](#) [PDF Plus \(268 KB\)](#)

---
-  **Making Data and Workflows Findable for Machines**  
 Tobias Weigel , Ulrich Schwardmann, Jens Klump, Sofiane Bendt Quick  
*Data Intelligence* Winter-Spring 2020, Vol. 2, No. 1-2, pp. 40–46

[Abstract](#) [Full Text](#) [PDF \(233 KB\)](#) [PDF Plus \(283 KB\)](#)

---
-  **The “A” of FAIR – As Open as Possible, as Closed as Necessary**  
 Annalisa Landi , Mark Thompson, Viviana Giannuzzi, Fedele Boni Labastida, Luiz Olavo Bonino da Silva Santos, and Marco Roos  
*Data Intelligence* Winter-Spring 2020, Vol. 2, No. 1-2, pp. 47–55

[Abstract](#) [Full Text](#) [PDF \(244 KB\)](#) [PDF Plus \(293 KB\)](#)

---
-  **A Generic Workflow for the Data FAIRification Process**  
 Annika Jacobsen , Rajaram Kaliyaperumal, Luiz Olavo Bonino da Barend Mons, Erik Schultes, Marco Roos, and Mark Thompson  
*Data Intelligence* Winter-Spring 2020, Vol. 2, No. 1-2, pp. 56–65

[Abstract](#) [Full Text](#) [PDF \(1154 KB\)](#) [PDF Plus \(325 KB\)](#)

---

# How to technically implement FAIR?



# The FAIR implementation process

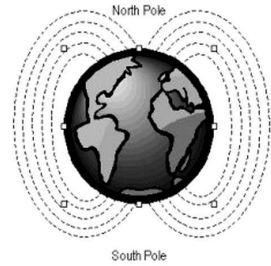
## a scalable approach from Solid Earth Domain

# EPOS Data Integration

Community Specific  
Data & Services



laboratory



Geomagnetic



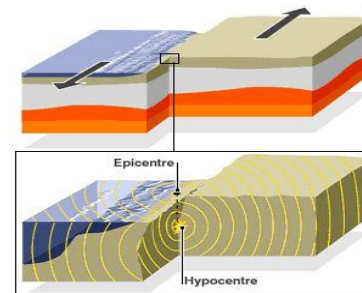
GPS



Volcanos



satellite

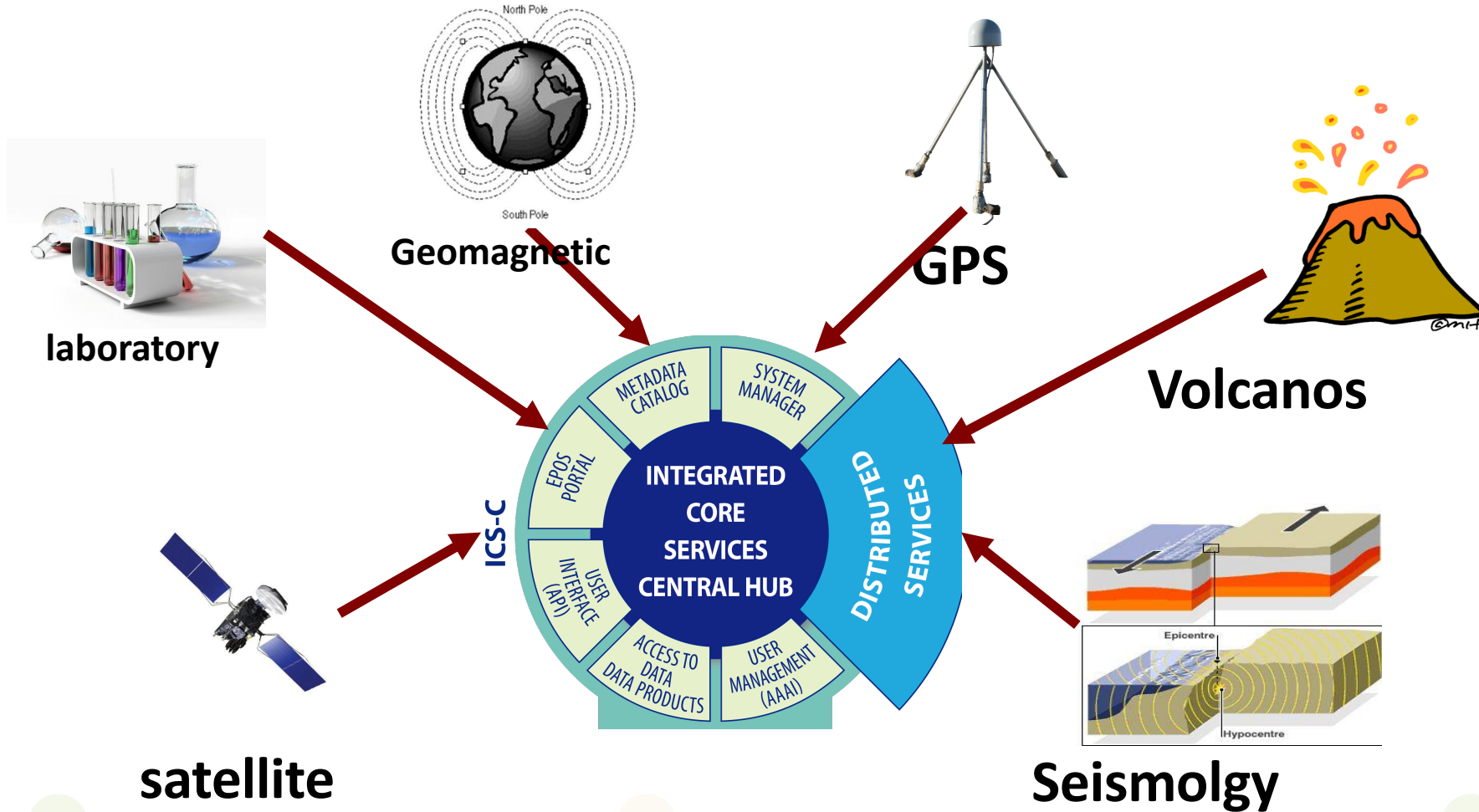


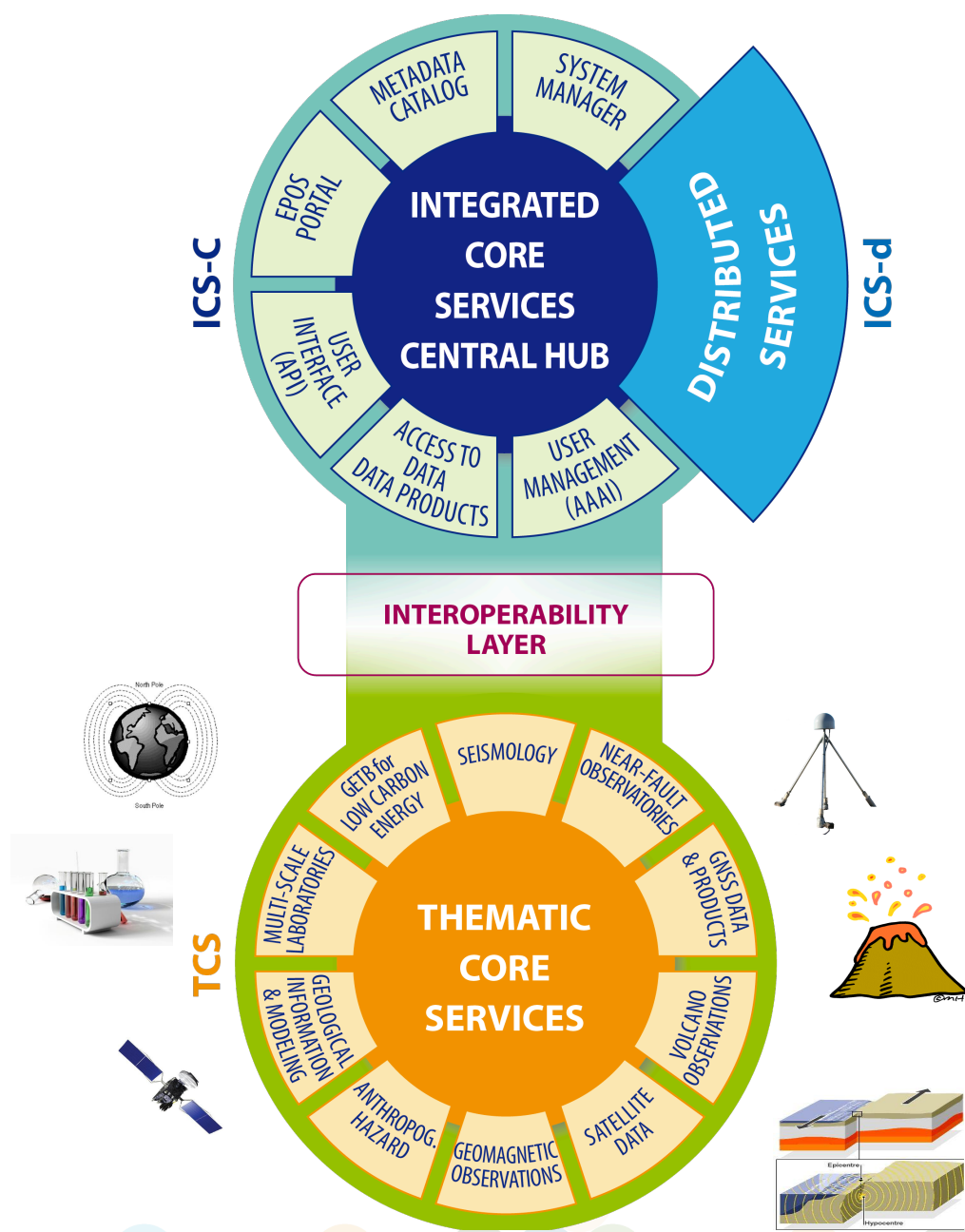
Seismolgy



# EPOS Data Integration

*Data & Services  
Integration in a central hub*

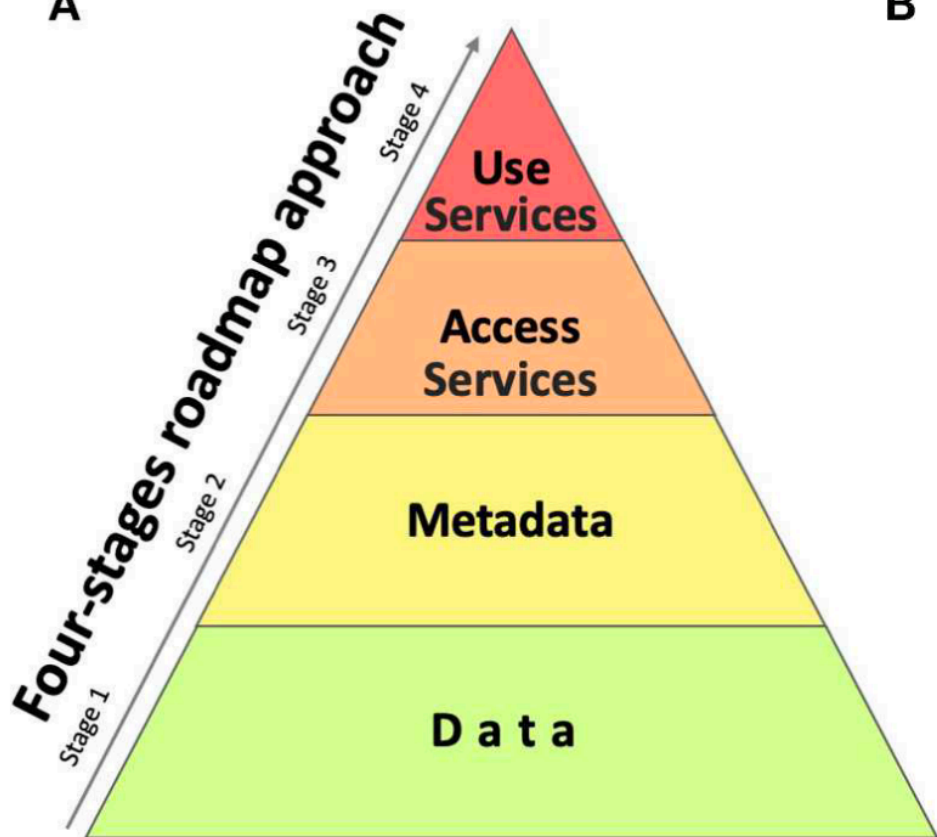




# FAIRness at Integration level

# FAIRness at Community level

A



B

# Process to achieve fairness

## FAIR PRINCIPLES

Use  
Services

Access  
Services

Metadata

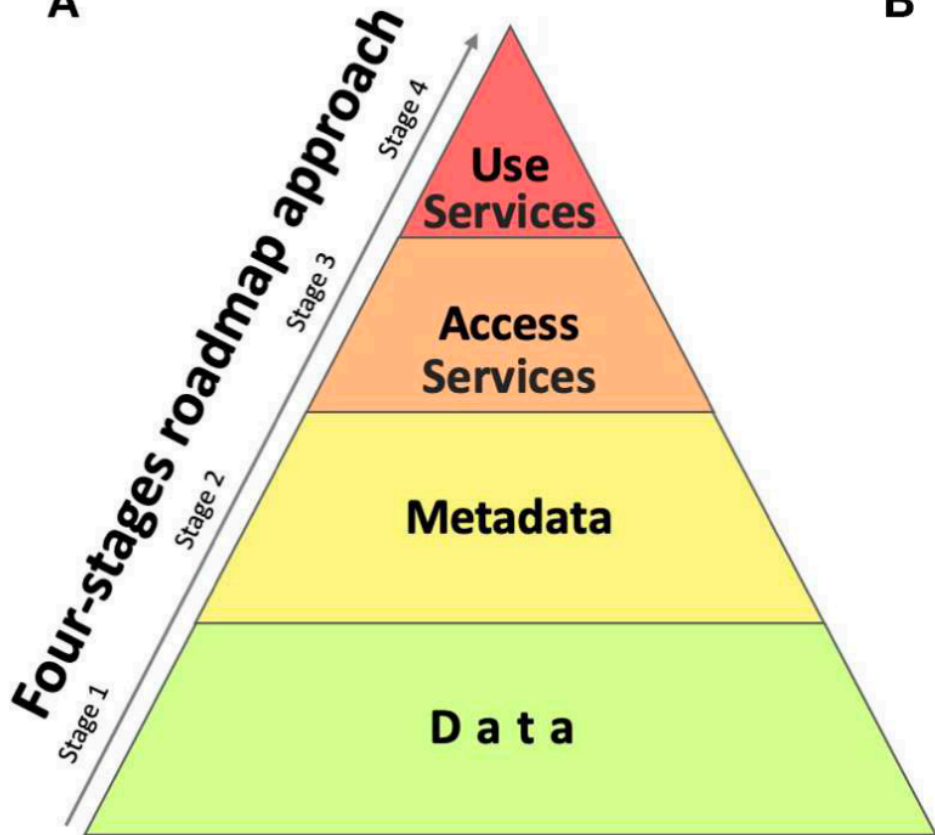
Data

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

- F1.** Metadata are assigned a globally unique and eternally persistent identifier.
- F2.** data are described with rich metadata.
- F3.** metadata specify the data identifier.
- F4.** metadata are registered or indexed in a searchable resource.
- A2.** metadata are accessible, even when the data are no longer available.
  - I1.** metadata use a formal, accessible, shared, and broadly applicable.
  - I2.** metadata use vocabularies that follow FAIR principles.
  - I3.** metadata include qualified references to other metadata.
- R1 (R1.1 - R1.2 - R1.3)** Metadata are richly described

- F1.** Data are assigned a globally unique persistent identifier.
- F4.** Data are registered or indexed in a searchable resource.
- I1.** Data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2.** Data use vocabularies that follow FAIR principles.
- I3.** Data include qualified references to other (meta)data.
- R1.1.** Data are released with a clear license.
- R1.3.** Data meet domain-relevant community standards

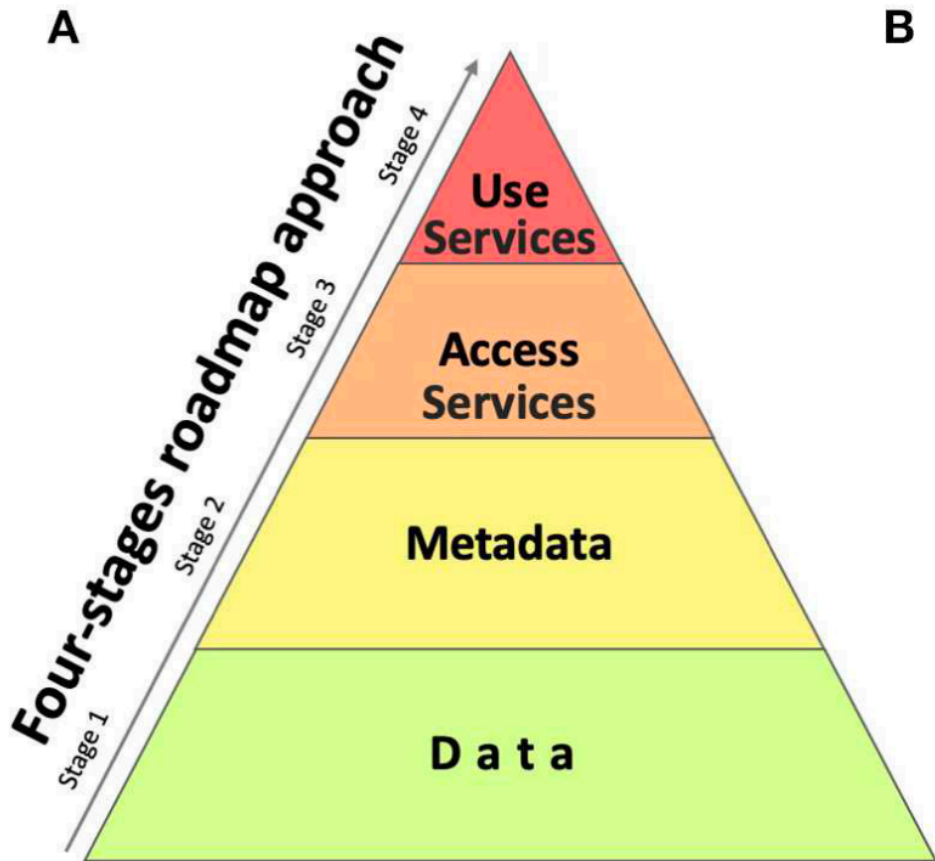


**A****B**

## **FAIRness assessment**

- a) Evaluation of FAIR compliancy
- b) Gap analysis

**Existing evaluation frameworks can be reused, e.g. FIP**



**B**

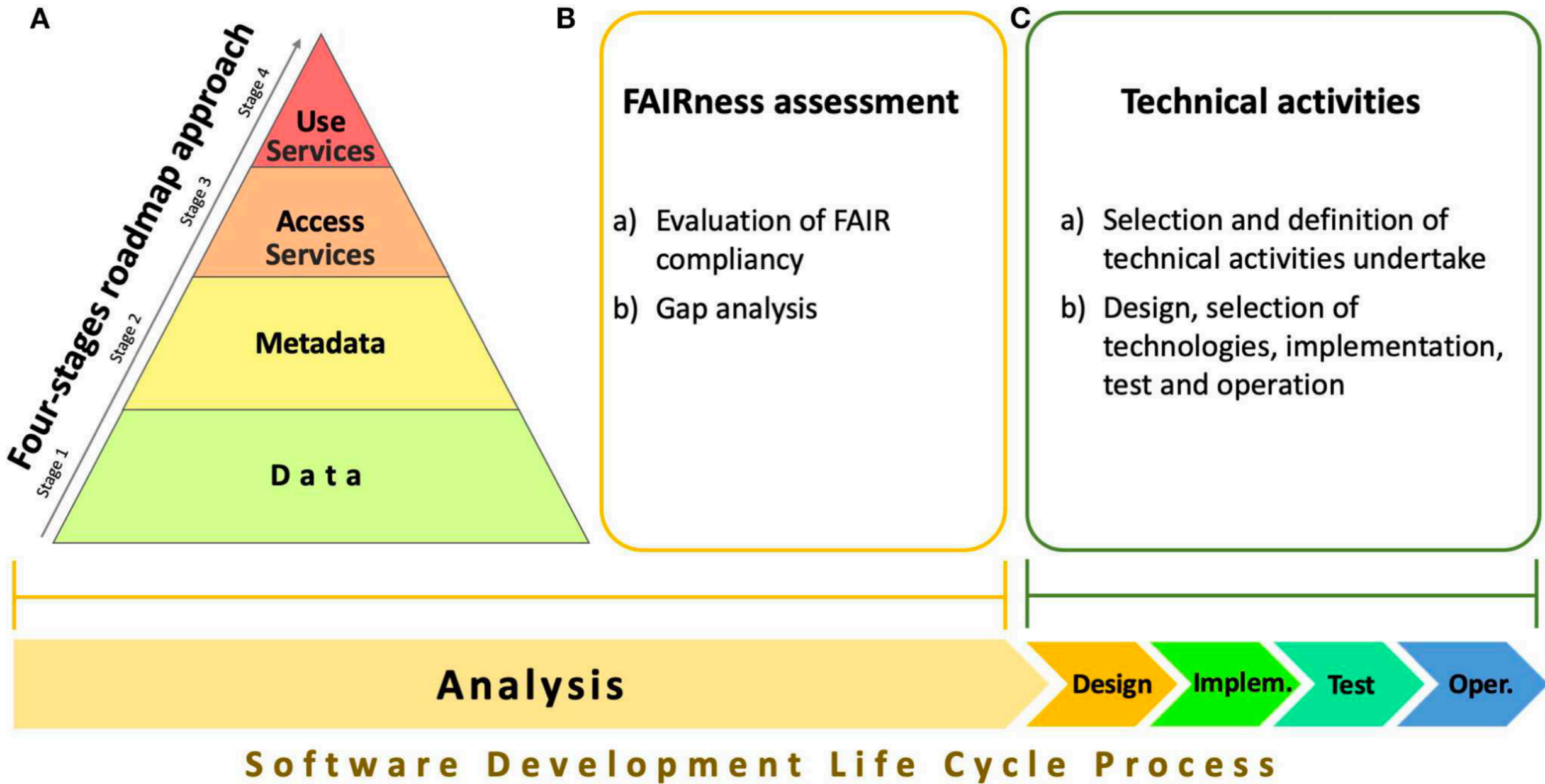
## **FAIRness assessment**

- a) Evaluation of FAIR compliancy
- b) Gap analysis

**C**

## **Technical activities**

- a) Selection and definition of technical activities undertake
- b) Design, selection of technologies, implementation, test and operation



*[Perspectives on the Implementation of FAIR Principles in Solid Earth Research Infrastructures.*

Bailo, D. et al. *Frontiers in Earth Science*, 8, 3. <https://doi.org/10.3389/feart.2020.00003>

**FAIR PRINCIPLES**

**Researcher  
Mindset**

**TYPES of TECHNOLOGIES**

Processing facilities  
Workflows  
Analysis or visualization tools

**Use  
Services**

*Search/discovery service  
SOA, web services  
Standard communication Protocols  
Authentication, Authorisation, Accounting*

**Access  
Services**

*Ontologies and Vocabularies  
Serialisation  
Persistent Identifier  
License handling  
Metadata Catalogue  
Provenance*

**Metadatas**

**Data**

*Standards Formats  
Data catalogue  
Persistent Identifiers  
License handling*

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

- F1.** Metadata are assigned a globally unique and eternally persistent identifier.
- F2.** data are described with rich metadata.
- F3.** metadata specify the data identifier.
- F4.** metadata are registered or indexed in a searchable resource.
- A2.** metadata are accessible, even when the data are no longer available.
- I1.** metadata use a formal, accessible, shared, and broadly applicable.
- I2.** metadata use vocabularies that follow FAIR principles.
- I3.** metadata include qualified references to other metadata.
- R1 (R1.1 - R1.2 - R1.3)** Metadata are richly described

- F1.** Data are assigned a globally unique persistent identifier.
- F4.** Data are registered or indexed in a searchable resource.
- I1.** Data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2.** Data use vocabularies that follow FAIR principles.
- I3.** Data include qualified references to other (meta)data.
- R1.1.** Data are released with a clear license.
- R1.3.** Data meet domain-relevant community standards



DATA (S1)

### FAIR principle

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

**R1.3.** data meet domain-relevant community standards

**I2.** data use vocabularies that follow FAIR principles.

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

**F4.** data are registered or indexed in a searchable resource.

**F1.** data are assigned a globally unique and persistent identifier.

**R1.1.** data are released with a clear and accessible data usage license.

### Related technical activities

***Selection or creation of a data model which provides a standard format for information resource description, supporting an ontology and providing standard vocabulary of terms.***

A main step for a FAIR data provision is its harmonization by adopting a standard format within a domain-specific community (e.g., shapefiles ESRI, 1998); need for harmonization frequently occurs typical “long-tail” contexts where data are produced by several different instruments on a per-sample basis, for instance in laboratories. Some formats embed in their serialization also an ontology (e.g. stationXML<sup>f</sup>); in such cases, appropriate FAIR, community-agreed ontologies need to be selected and adopted, possibly reusing or inheriting from existing ones (e.g., EnvO Buttigieg et al., 2016).

***Selection of a Data catalog which supports discovery of datasets, and adoption of appropriate storage and preservation strategies.***

A system for data storage should be selected on the basis of data type, e.g., georeferenced layered data might need a database with GIS support, while other types of data might require filesystem with hierarchical folders or noSQL databases. Furthermore, Certification of Data repository (Dillo and de Leeuw, 2018) might be also considered.

***Selection of a PID system which guarantees technical reliability, authority, and ensures a long-term viability.***

PIDs require dedicated machinery and software to be issued, resolved and managed. A common solution is to rely on organizations—e.g., Datacite (Jan, 2009), ePIC<sup>g</sup>—implementing the above, but other options might be considered (Sicilia et al., 2019).

***Development and adoption a Data policy which ensures that data collected or created by the communities, once quality controlled, are made available under clear usage conditions licenses.***

From a technical point of view, licenses require rich enough information to be machine-readable and properly cited. A relevant example is creative commons which provide an easy way to choose a license and to share it in a machine-readable way<sup>h</sup>.

Besides technical aspects, for establishing data policies community-wide agreements are needed and, according to the size of the community, they may require implementation of appropriate communication, community building, and organizational strategies.

# Final Considerations

**Need for a pragmatic approach to help and guide Data Providers**

**FAIR is not only about (examples):**

- Defining metadata schemas
- Choosing a good ontology
- Select the serialization format
- Defining PID issuer

**It has pragmatic aspects (examples):**

- select an existing metadata standards with associated tools for metadata manipulation and management
- integrate PID services in the existing e-Infrastructure
- Integrate AAAI system in the existing infrastructure

# Final Considerations

## **Generalistic repositories: do they serve the purpose?**

### **Yes, when:**

- You have few datasets
- Metadata is simply a list of tags
- Your data is linked to a publication

### **No, when:**

- You have many datasets, or dynamic data
- You want to perform parameter-specific search (e.g CO2)
- You need an institutional-based approach

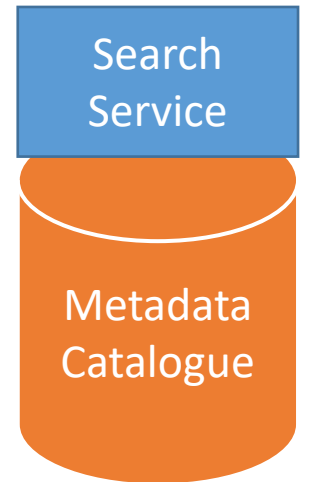
# Final Considerations

**If we want FAIR to be implemented more clear technical directions are needed**

What tools or technical activities to satisfy specific principles?

Example: “**F4**. Data are registered or indexed in a searchable resource”.

- *Need metadata catalogue (database)*
- *Selection of a standard metadata format (need list of available standards)*
- *Need a service to search into the catalogue (web service)*





# Final Considerations

**(web) Services are the «missing guest» in FAIR principles**

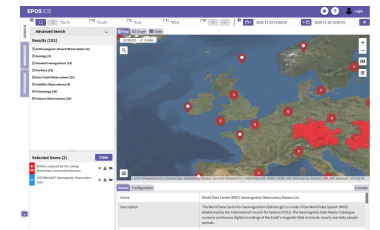
Data is usually made available through a «service»,

Example:

- RESTful web service
- SOAP web Services

**Machines or GUI Client use such services**

GUI client



Data/metadata  
Service

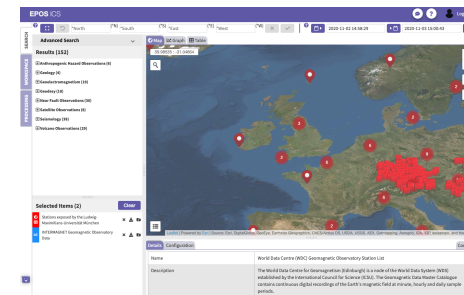
Metadata  
Catalogue

# Final Considerations

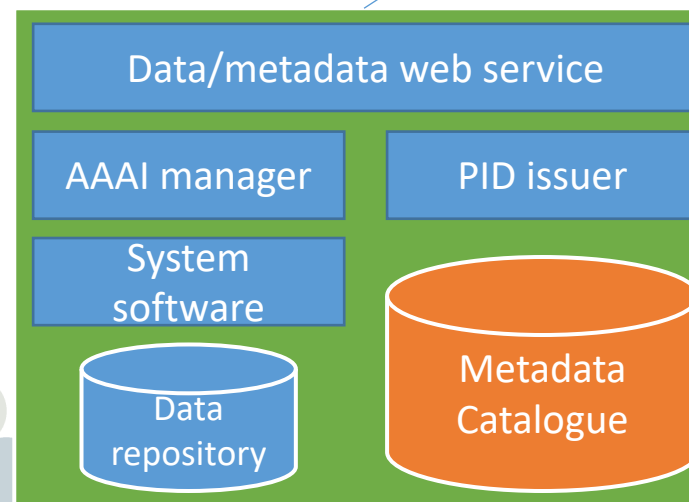
## Freedom of implementation vs. design guidelines

**FAIR community is still too shy**

- reference architectures and technology checklist might be provided
- Otherwise, risk of getting «ingested» by problem-solving companies or certification authorities



GUI Client



# Final Considerations

## Technical Sustainability

Implementing a FAIR system for data stewardship requires **long term sustainability plan** which includes (at least):

- Technical aspects (e.g. maintainability of the software)
- Financial aspects (e.g. funding for the long term maintenance)
- Governance aspects (e.g. involvement of appropriate individuals or stakeholders from the organization)

It is sometimes (often?) preferable to «cluster» with existing initiatives that **tackle FAIR data stewardship at community level**

# Thank You

## WebSite



[www.epos-ip.org](http://www.epos-ip.org)

## Newsletter



[www.epos-ip.org/news-press/epos-ip-newsletter](http://www.epos-ip.org/news-press/epos-ip-newsletter)

## Social Media



### PERSPECTIVE ARTICLE

Front. Earth Sci., 31 January 2020 | <https://doi.org/10.3389/feart.2020.00003>



## Perspectives on the Implementation of FAIR Principles in Solid Earth Research Infrastructures

Daniele Bailo<sup>1\*</sup>, Rossana Paciello<sup>1</sup>, Manuela Sbarra<sup>2</sup>, Riccardo Rabisoni<sup>2</sup>, Valerio Vinciarelli<sup>2</sup> and Massimo Cocco<sup>1</sup>

[<https://www.frontiersin.org/articles/10.3389/feart.2020.00003/full>]