FAIR and Real world

Requirements for being FAIR

Daniele Bailo, EPOS IT Officer

What are the **requirements** to comply with FAIR principles?

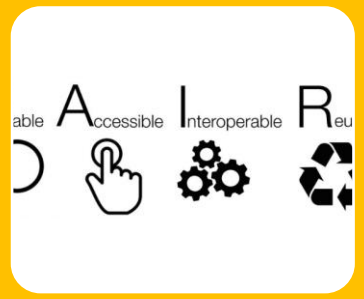
THIS PROJECT HAS RECEIVED FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT N° 676564

What are the **requirements** to comply with FAIR principles?

What **work needs to be done** for a datacenter/RI to be FAIR?

THIS PROJECT HAS RECEIVED FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT N° 676564

1

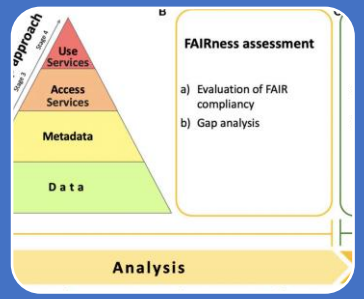


FAIR analysis

2

Reated IT concepts

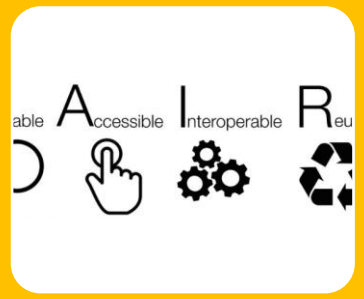
3



Implementation process

THIS PROJECT HAS RECEIVED FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT N° 676564

1 FAIR analysis



THIS PROJECT HAS RECEIVED FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT N° 676564

Fair Guiding 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 Guiding principles

Main concepts:

- 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 Guiding principles

Main concepts:

- 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 Guiding principles

Main concepts:

- 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 Guiding principles

Main concepts:

- Data
- Metadata
- Identifiers
- Protocols

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 Guiding 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

Main concepts:

- Data
- Metadata
- Identifiers
- Protocols
- Authe&Autho

Fair Guiding 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

Main concepts:

- Data
- Metadata
- Identifiers
- Protocols
- Authe&Autho

Standards

Fair Guiding 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

Main concepts:

- Data
- Metadata
- Identifiers
- Protocols
- Authe&Autho

Standards

Semantics

Fair Guiding 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

Main concepts:

- Data
- Metadata
- Identifiers
- Protocols
- Authe&Autho

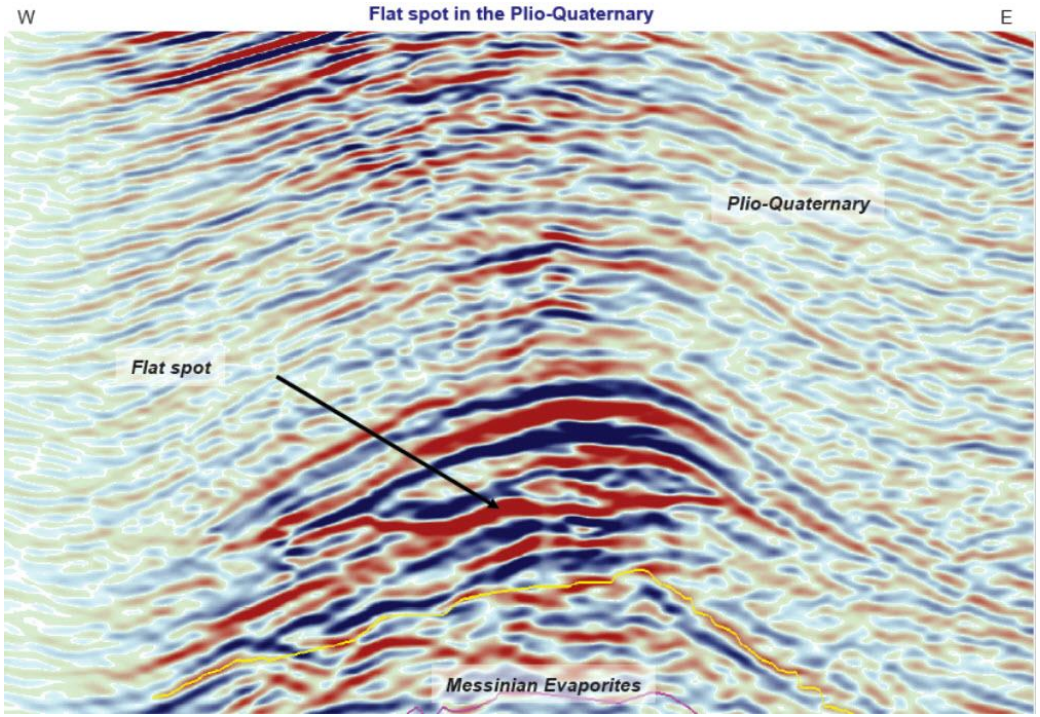
Standards

Semantics

policy & provenance

2  Reated IT concepts

THIS PROJECT HAS RECEIVED FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT N° 676564
Disclaimer: the content of this presentation reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains



Courtesy of Joshua Doubek

Data

Our greatest wealth

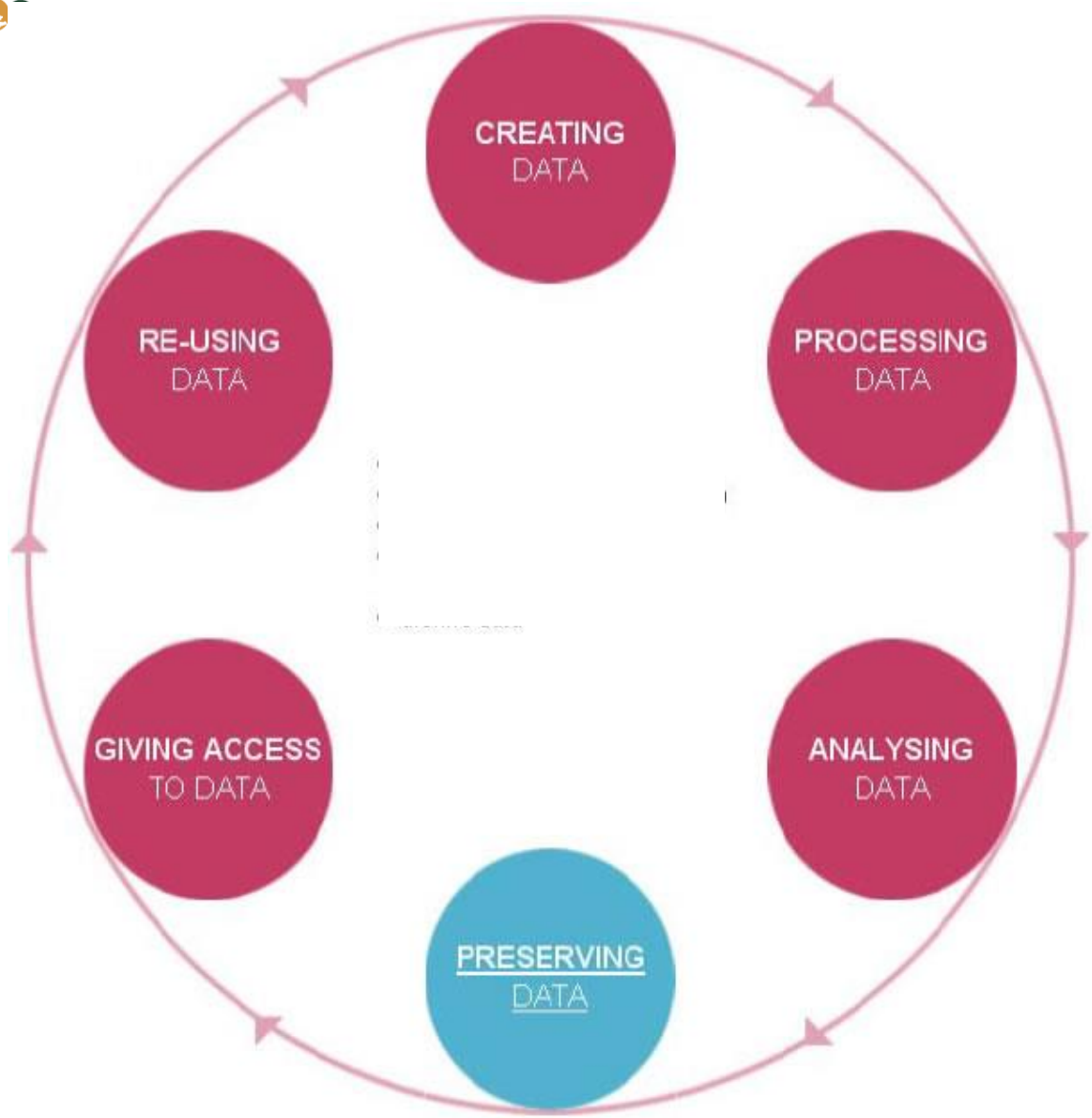
Function

Describe (physical) phenomena

Issues

- *Format & serialization*
- *Harmonization & standards*
- *Proliferation standards*
- *Data storage*

THIS PROJECT HAS RECEIVED FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT N° 676564



[UK data Archive http://www.data-archive.ac.uk/]

Data Lifecycle

What?

Identifies mechanisms, standard, components and interfaces making data science efficient and cost effective

Data Management Plan

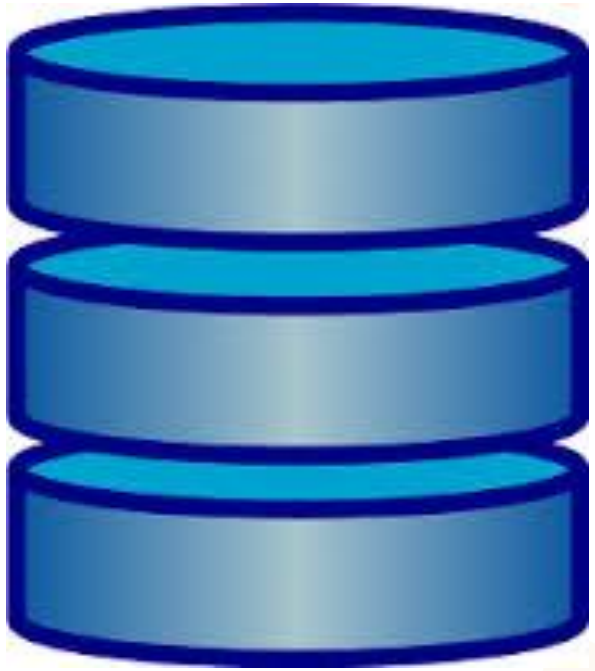
- *Data management*
- *Data analysis*
- *Data preservation*
- *Data publication*
- *Data sharing*

THIS PROJECT HAS RECEIVED FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT N° 676564

CORE TRUST SEAL

TRUST

Transparency | Responsibilty | User focus | Sustainability | Technology



How to store data?

Trusted data repository

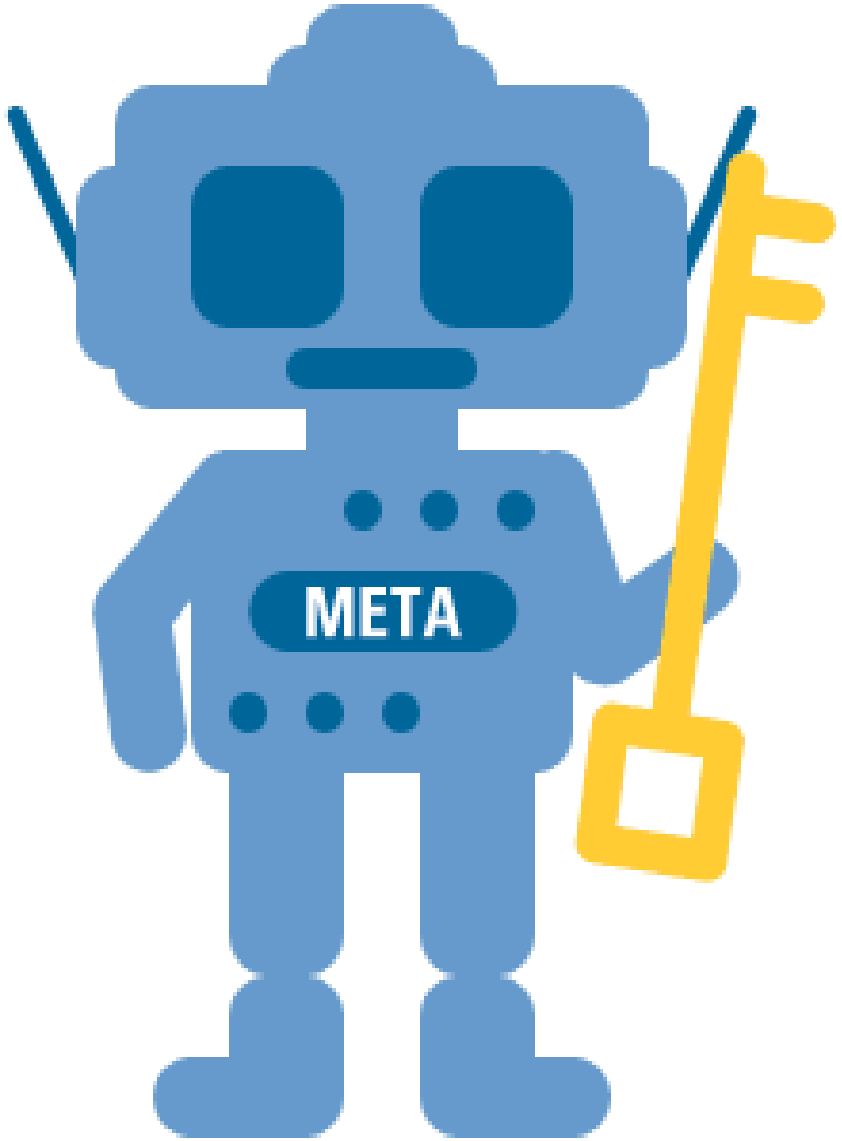
Certifications ensure

- *Storage*
- *Preservation*
- *Sustainability*
- *Access*
- *Integrity*
- *Documentation availability*
- *Couple with PIDs*
- *…*

Question

- *Should we certify our repos?… DEPENDS*

THIS PROJECT HAS RECEIVED FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT N° 676564



Metadata

Data about Data (really?)

Purposes

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

Usually attached to D.O.

Issues (selection of)

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

THIS PROJECT HAS RECEIVED FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT N° 676564

(Metadata) standards in speficic scientific domain

…a real story…

THIS PROJECT HAS RECEIVED FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT N° 676564

(Metadata) standards in speficic scientific domain

…a real story…

THIS PROJECT HAS RECEIVED FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT N° 676564

GeoJSON

DCAT APPLICATION PROFILE FOR EUROPEAN DATA PORTALS

DCAT-AP

ckan

CERIF

Dublin Core Metadata Initiative®
Making it easier to find information.

Metadata

What is the best standard?

DEPENDS ON THE REQUIREMENTS

Fair principles require for

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

Serialization and format:
two different things (almost)

THIS PROJECT HAS RECEIVED FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT N° 676564

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

How to register/cite data or publications?

PID system

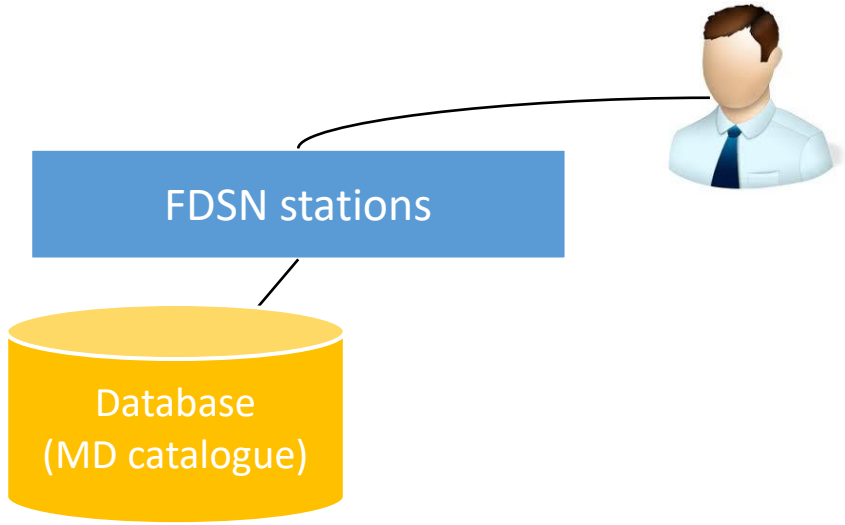
Purpose

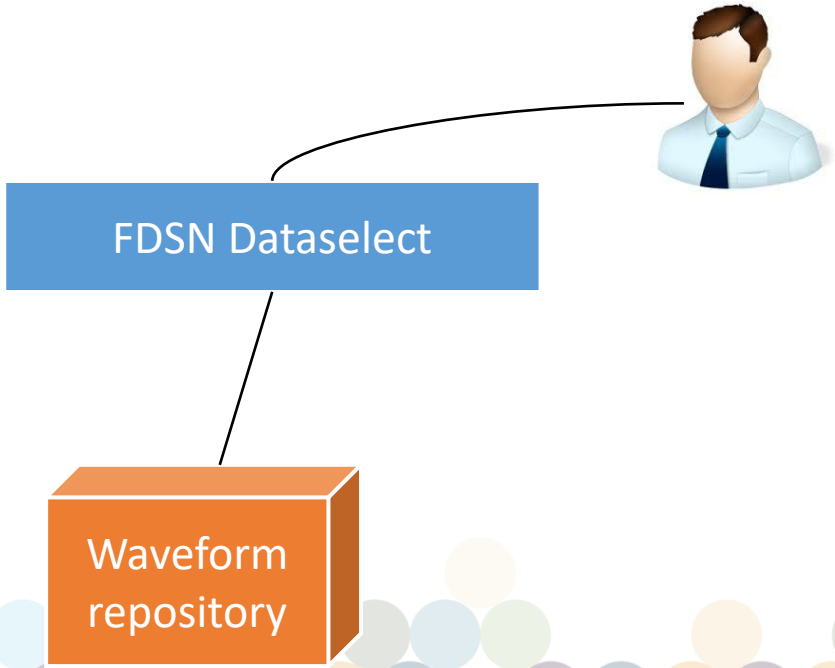
- *DO / publication can be uniquely referenced*
- *Assign a PID at data creation times*

Issues

- *Need for a simple mechanism to implement it*
- *Data and publications*

THIS PROJECT HAS RECEIVED FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT N° 676564





Protocols
web service

What is it?

It does something for the user
*(deliver value to customer)**

A "thin layer"

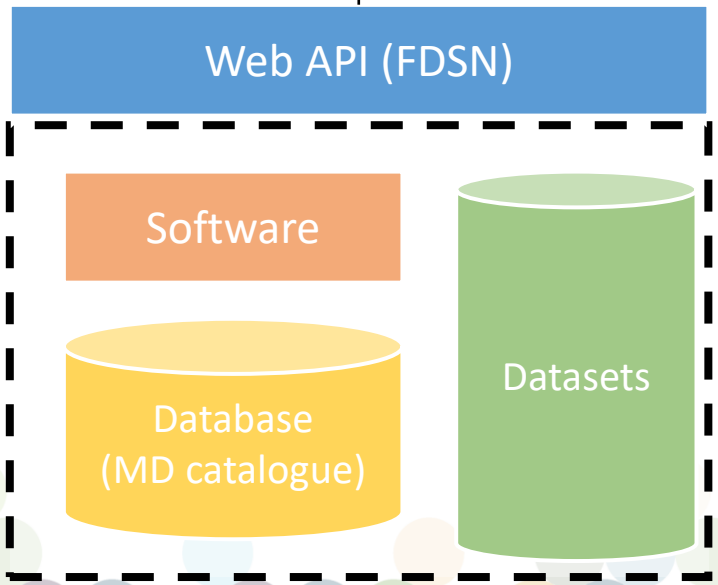
We usually don't know what's under the hood

Examples

- *FDSN stations*
- *FDSN dataselect*

THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT N° 676564

System 1
GUI Client



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*

THIS PROJECT HAS RECEIVED FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT N° 676564

AARC

ORCID

Sign in with EGI CheckIn

Sign in with TCS-AH AAI

Google

facebook

How to access data?

AAI system
(federeated & distributed)

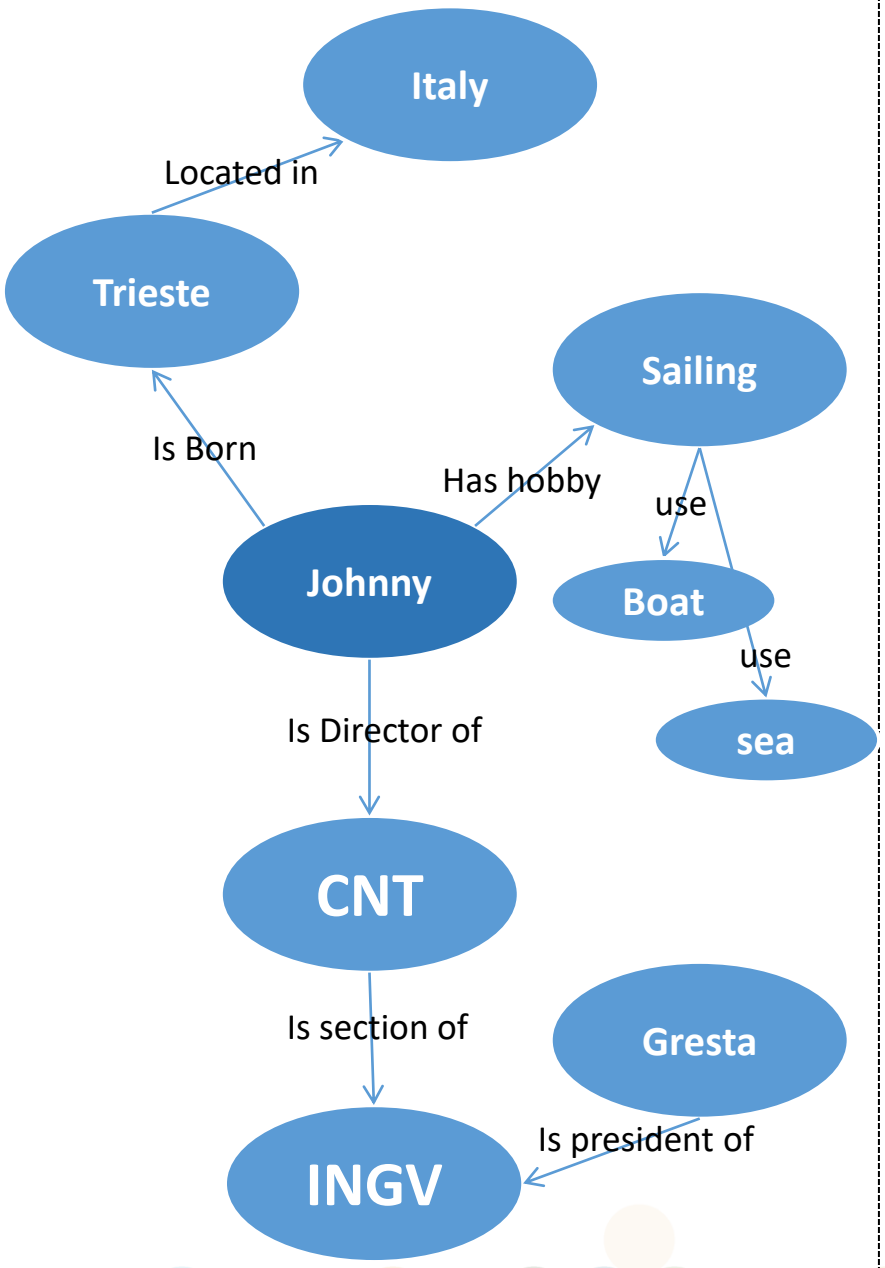
Purpose

- *Authenticate users*
- *Authorize users*

Issues

- *Identity Provider*
- *Federation*
- *Authorization*
- *Unity – integrates AAI*

THIS PROJECT HAS RECEIVED FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT N° 676564



Ontologies & semantics

Why an ontology?

It is the way machines manage “meaning”

How does it work?

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

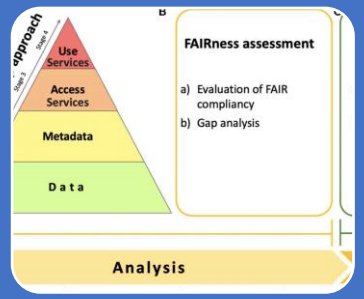
Issues

- Many ontologies exist
- Vocabulary Mapping

THIS PROJECT HAS RECEIVED FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT N° 676564

EPOS approach

3 Implementation process



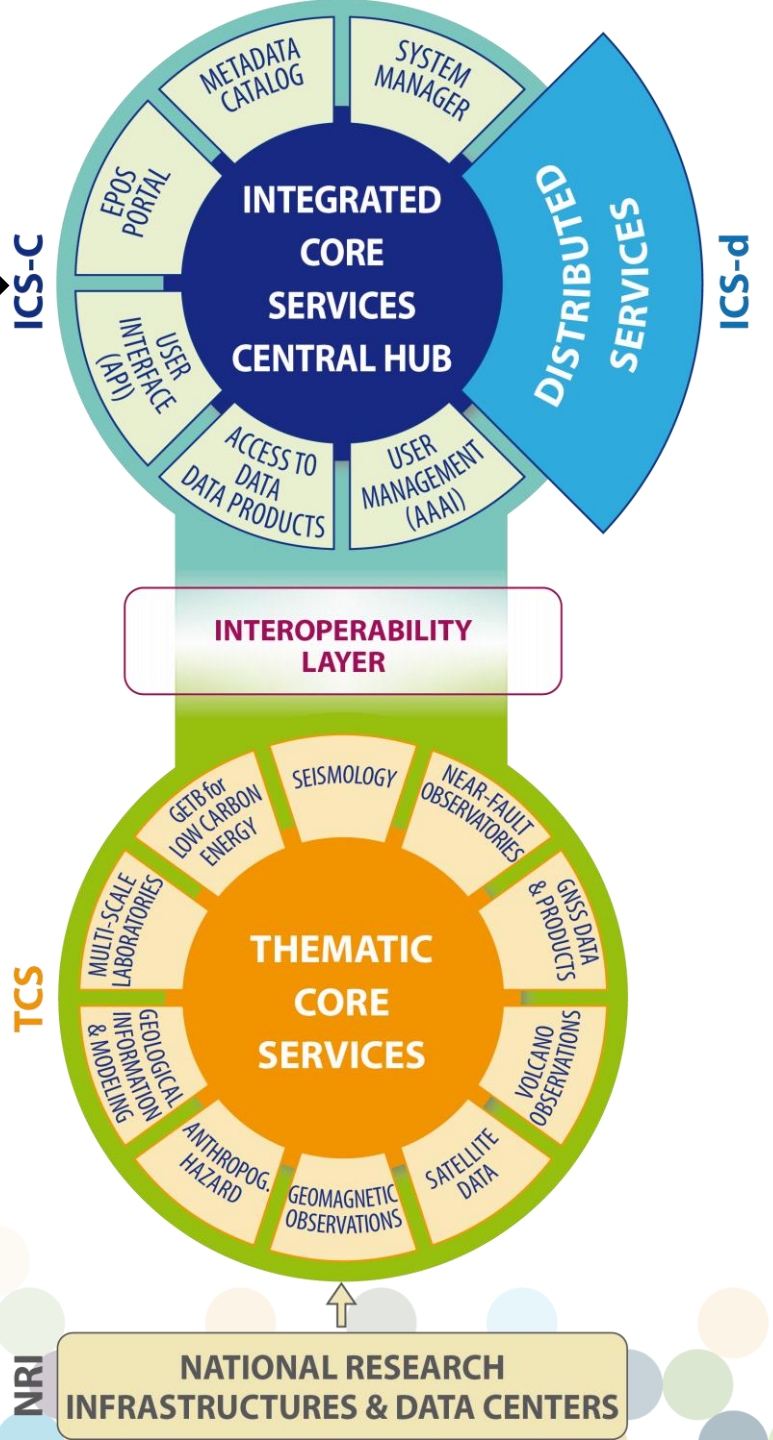
THIS PROJECT HAS RECEIVED FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT N° 676564

EPOS CONTEXT

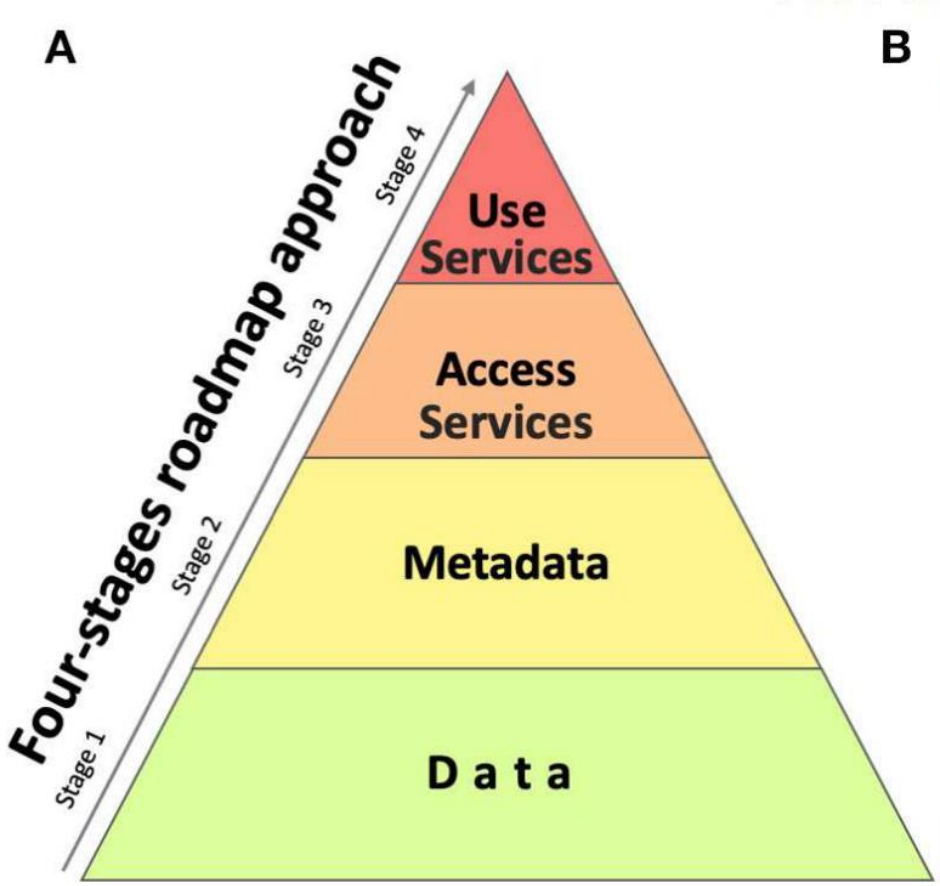
Data & Service integration →

Interoperability →

Communities →



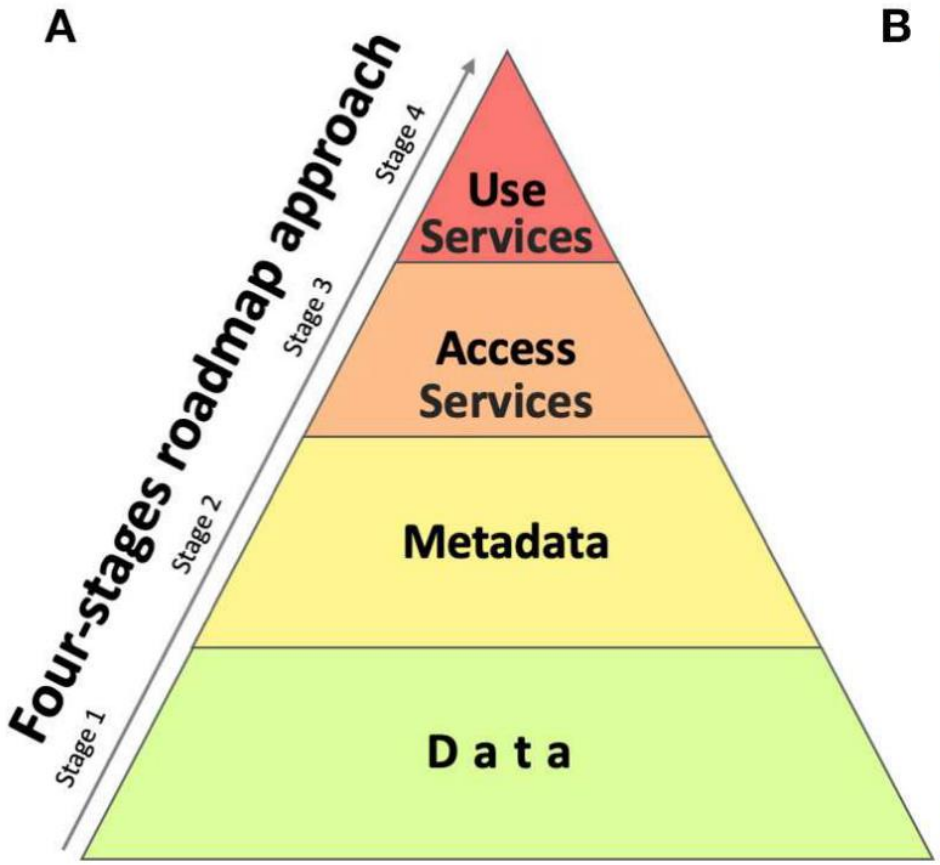
THIS PROJECT HAS RECEIVED FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT N° 676564



Process to achieve fairness

Bailo, D., Paciello, R., Sbarra, M., Rabissoni, R., Vinciarelli, V., & Cocco, M. (2020). Perspectives on the Implementation of FAIR Principles in Solid Earth Research Infrastructures. *Frontiers in Earth Science*, *8*, 3. https://doi.org/10.3389/feart.2020.00003

THIS PROJECT HAS RECEIVED FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT N° 676564



B

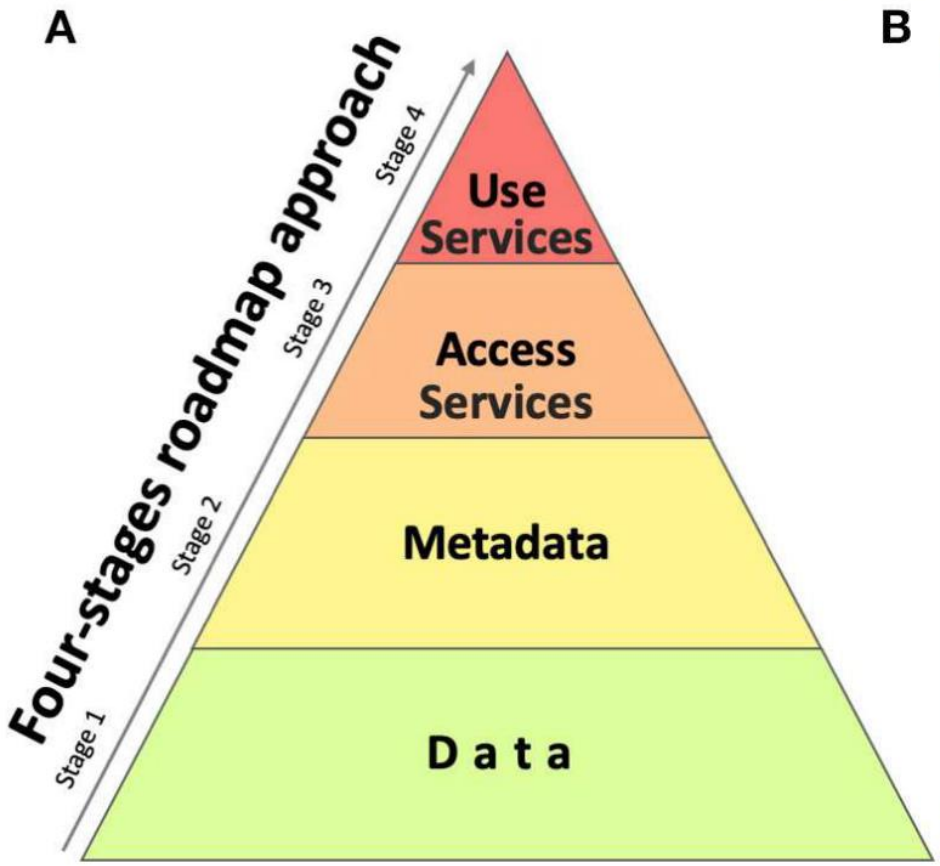
FAIRness assessment

a) Evaluation of FAIR compliancy

b) Gap analysis

Bailo, D., Paciello, R., Sbarra, M., Rabissoni, R., Vinciarelli, V., & Cocco, M. (2020). Perspectives on the Implementation of FAIR Principles in Solid Earth Research Infrastructures. *Frontiers in Earth Science*, *8*, 3. https://doi.org/10.3389/feart.2020.00003

THIS PROJECT HAS RECEIVED FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT N° 676564
Disclaimer: the content of this presentation reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains



B

FAIRness assessment

a) Evaluation of FAIR compliancy

b) Gap analysis

C

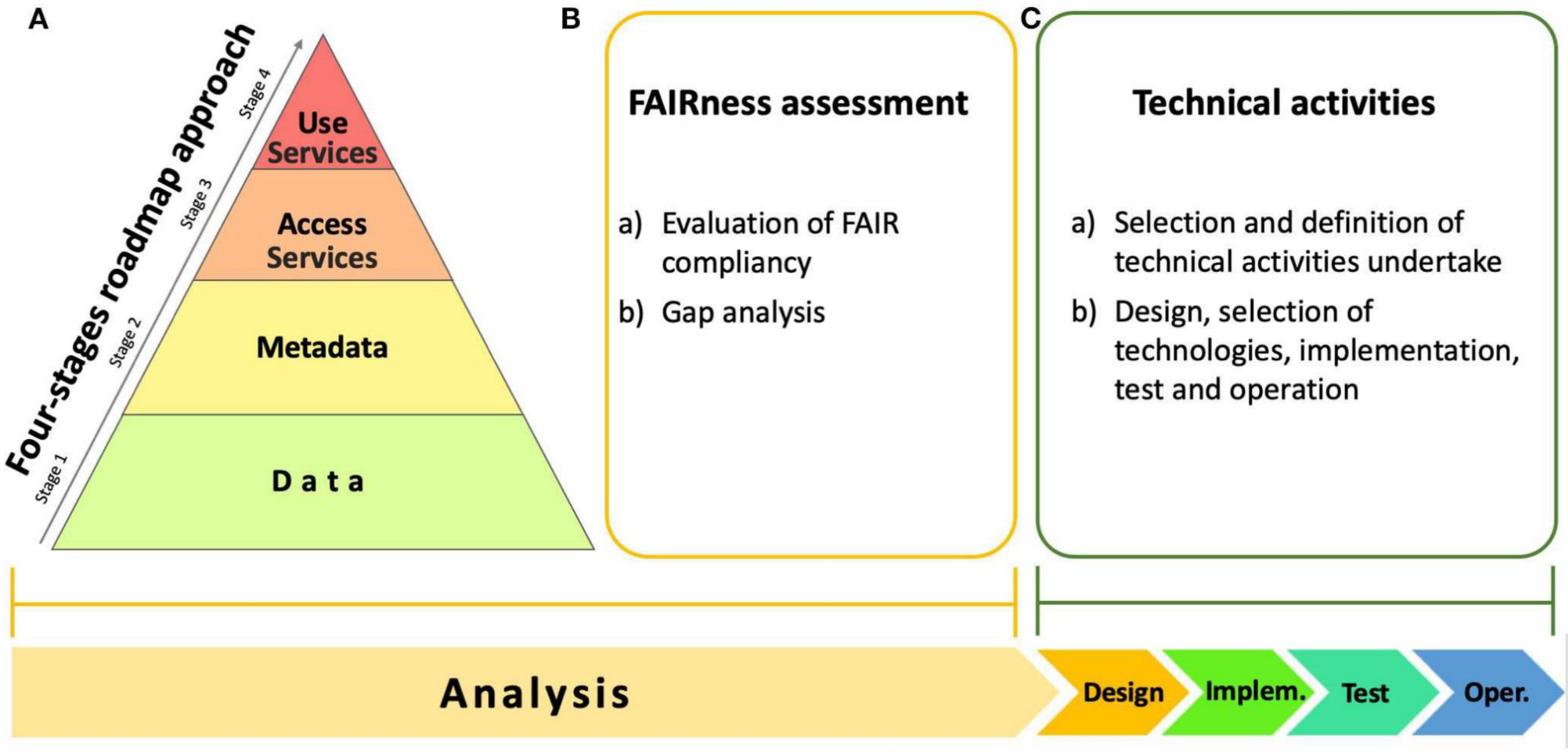
Technical activities

a) Selection and definition of technical activities undertake

b) Design, selection of technologies, implementation, test and operation

Bailo, D., Paciello, R., Sbarra, M., Rabissoni, R., Vinciarelli, V., & Cocco, M. (2020). Perspectives on the Implementation of FAIR Principles in Solid Earth Research Infrastructures. *Frontiers in Earth Science*, *8*, 3. https://doi.org/10.3389/feart.2020.00003

THIS PROJECT HAS RECEIVED FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT N° 676564



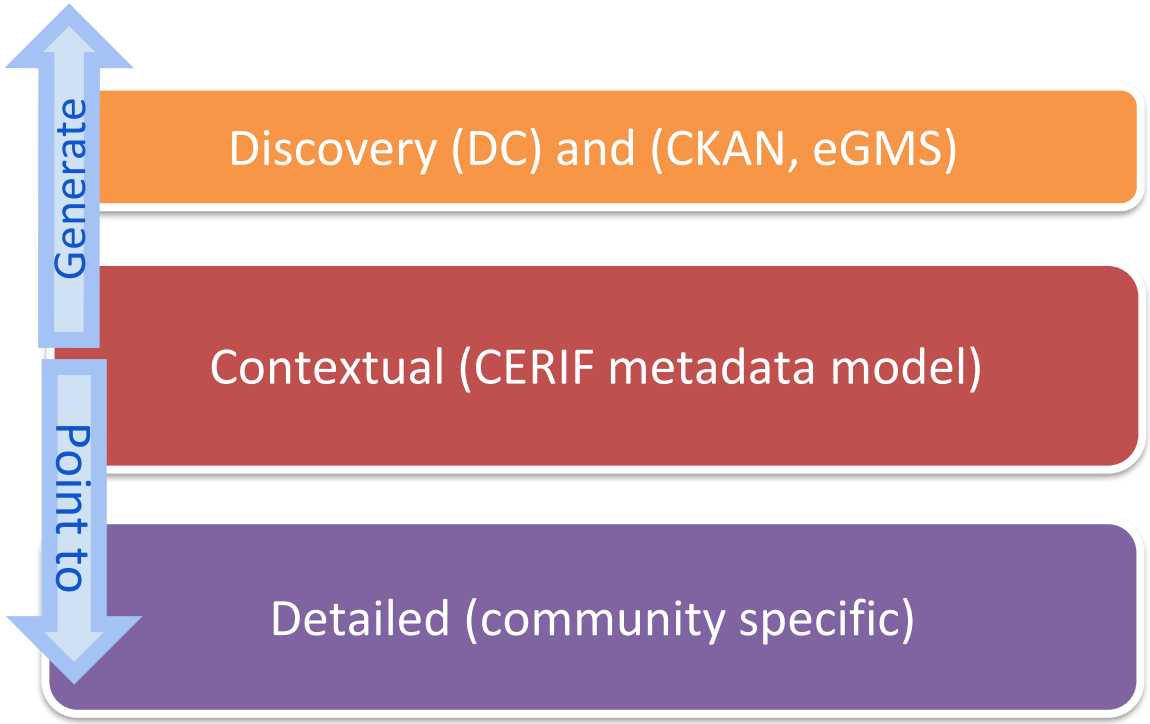
Bailo, D., Paciello, R., Sbarra, M., Rabissoni, R., Vinciarelli, V., & Cocco, M. (2020). Perspectives on the Implementation of FAIR Principles in Solid Earth Research Infrastructures. *Frontiers in Earth Science*, *8*, 3. https://doi.org/10.3389/feart.2020.00003

THIS PROJECT HAS RECEIVED FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT N° 676564

Compliant with A1, I3

EPOS Fair implementation examples (1)

Three layer metadata model



1. *Map & match only contextualization metadata*
2. *Pointers to detailed metadata*
3. *Export metadata in any standard*

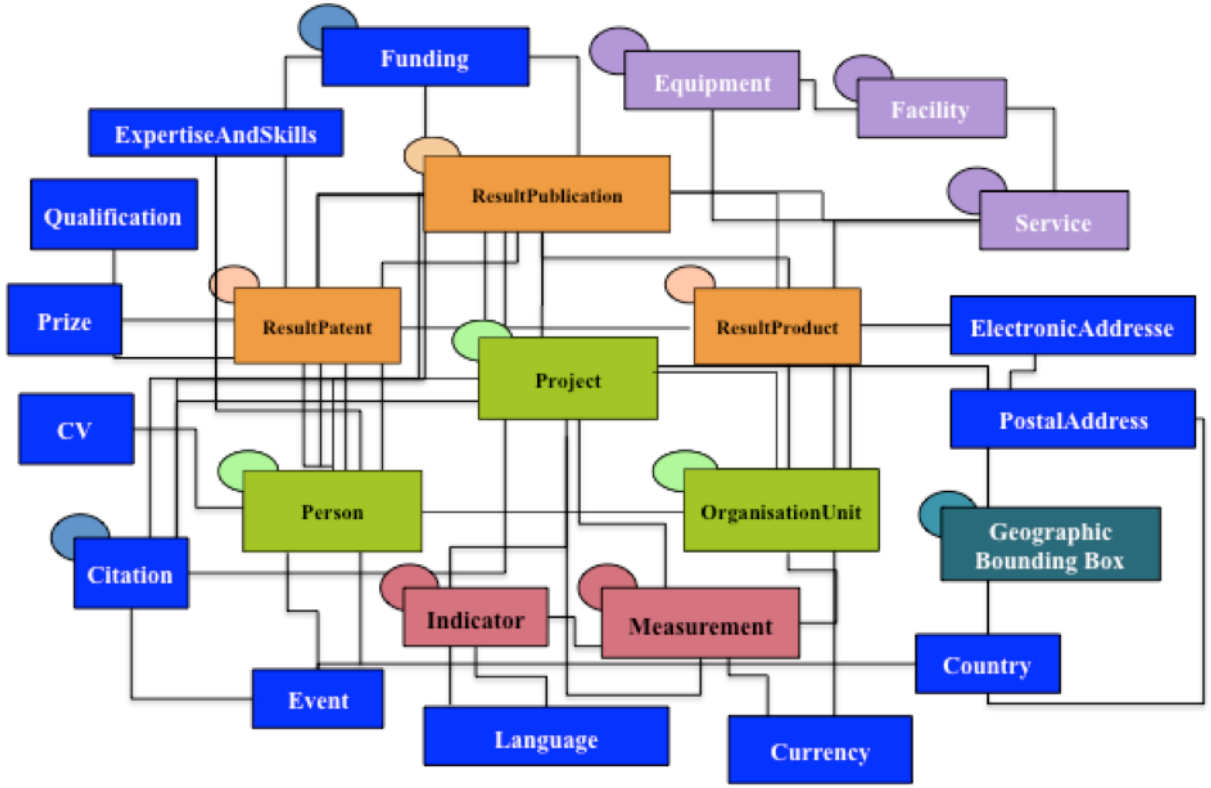
[Keith Jeffery]

THIS PROJECT HAS RECEIVED FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT N° 676564

Compliant with F2, F3, F4, R1.3

EPOS Fair implementation examples (2)

Rich metadata model



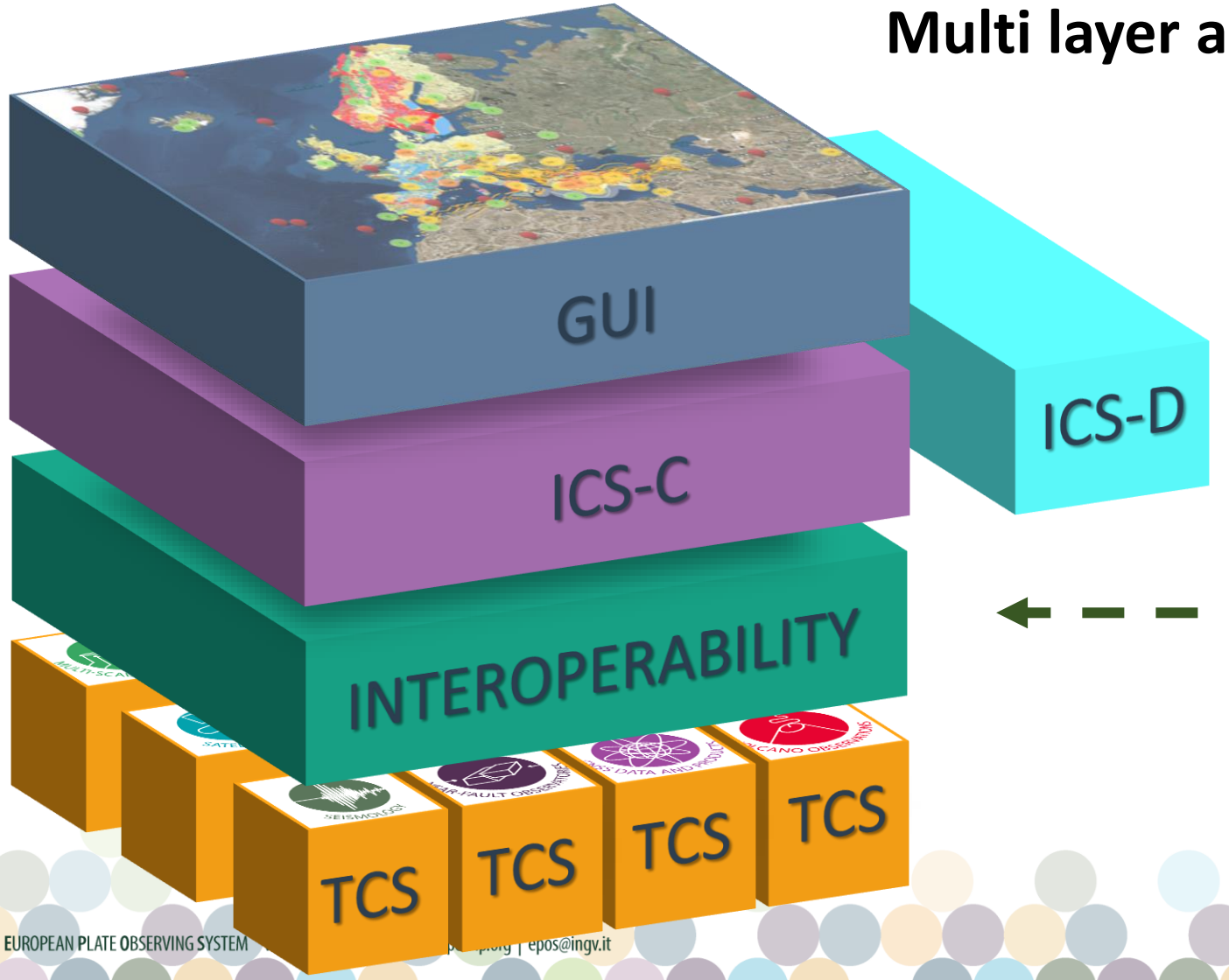
- *Supports several concepts*
- *Superset of many metadata standards*
- *Referencial integrity*
- *Formal syntax, declared semantics*

THIS PROJECT HAS RECEIVED FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT N° 676564

Compliant with A1.1

EPOS Fair implementation examples (3)

Multi layer approach



Web-Api based

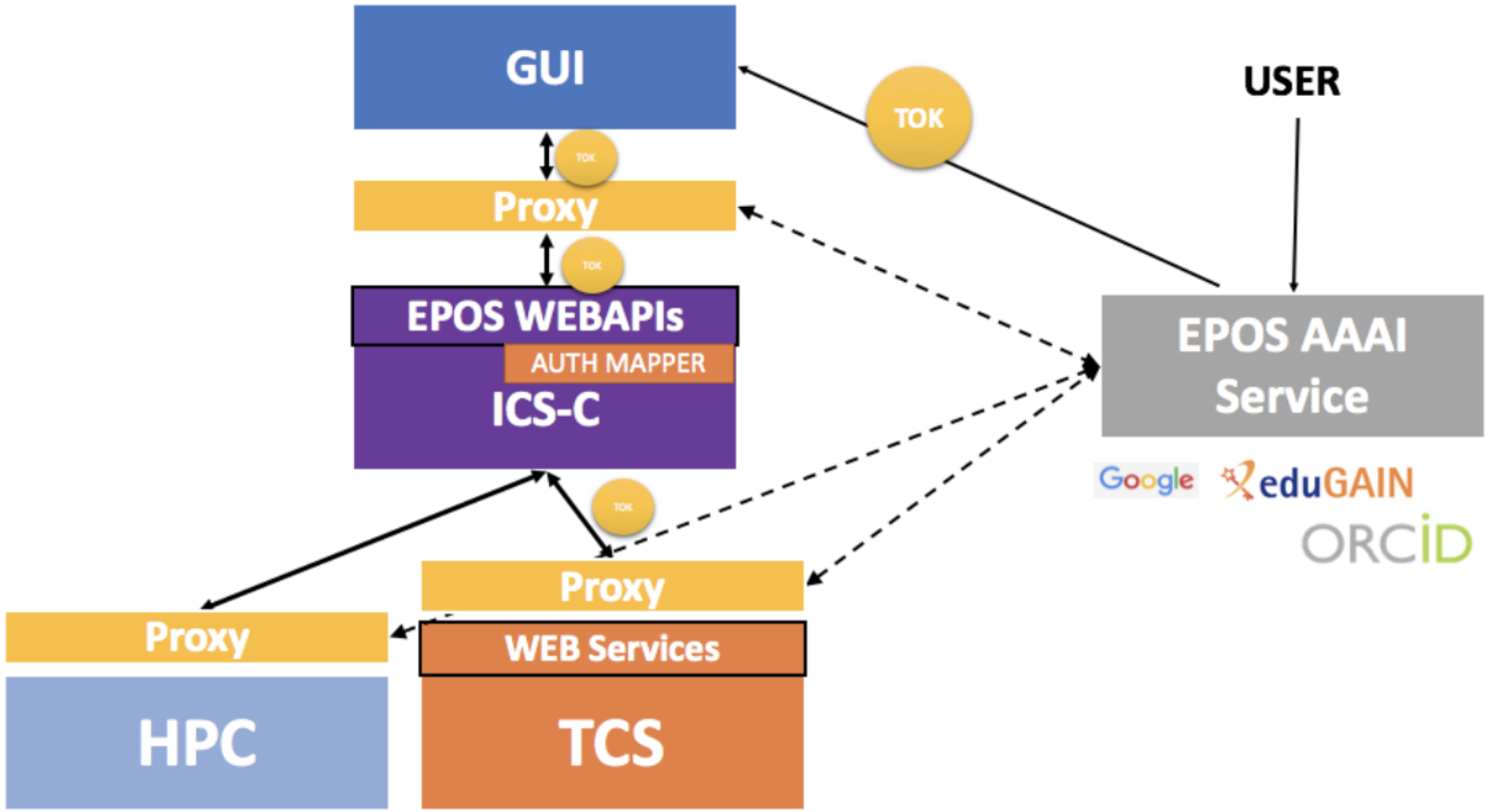
(standard protocols)

THIS PROJECT HAS RECEIVED FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT N° 676564

Compliant with A1.1, A1.2

EPOS Fair implementation examples (4)

AAI integration system



EUROPEAN PL
ROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT N° 676564

EPOS Fair implementation examples (5)

Full list of implementation activities and FAIR compliancy

PERSPECTIVE ARTICLE

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

Check for updates

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

 Daniele Bailo[1*], Rossana Paciello[1], Manuela Sbarra[2], Riccardo Rabissoni[2], Valerio Vinciarelli[2] and Massimo Cocco[1]

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

THIS PROJECT HAS RECEIVED FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT N° 676564

WebSite

www.epos-ip.org

Newsletter

www.epos-ip.org/news-press/epos-ip-newsletter

Social Media

Thank You

THIS PROJECT HAS RECEIVED FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT N° 676564

THIS PROJECT HAS RECEIVED FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT N° 676564

THIS PROJECT HAS RECEIVED FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT N° 676564

A pragmatic approach

- data harmonization
- metadata
- PID
- AAAI systems
- Web-services é communication protocols

THIS PROJECT HAS RECEIVED FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT N° 676564