EOSC Symposium 2022
14-17 November 2022

Making data policies FAIR

Joy Davidson, DCC

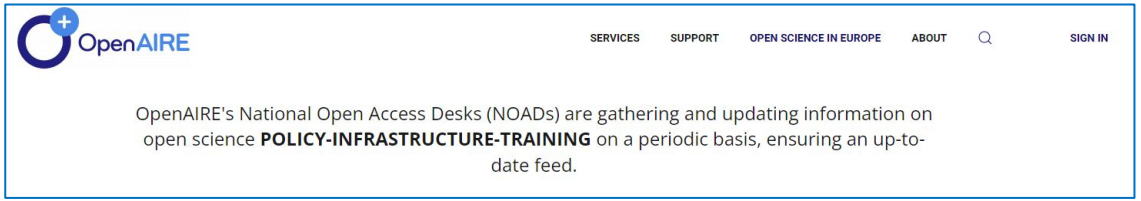
Allyson Lister
FAIRsharing Content & Community Lead

@digitalcuration
@FAIRsFAIR_EU
@jd162a

RDA Europe / EOSC Future Domain Ambassador for standards, repositories and policies

@FAIRsharing_org
@allysonlister
@allysonlister@fediscience.org

Many stakeholders want to monitor the OS policy landscape



https://www.openaire.eu/os-eu-countries

https://www.eosc.eu/monitoring-and-reporting

https://eosc-portal.eu/policy/EU-Countries

Currently, a manual process

- Mainly through surveys at national level
- Time consuming to collect and analyse
- Hard to compare the actual *content* and coverage of the policies across Europe

AA9 Landscape Monitoring
*Monitor **standardised** national Open Science and FAIR data strategies, including the description of these policies.*

SRIA: https://www.eosc.eu/sites/default/files/EOSC-SRIA-V1.0_15Feb2021.pdf

FAIRsFAIR Policy Resources

FAIR-enabling data policy checklist

helps policy makers review whether their data policies are FAIR-enabling and provides practical recommendations

Policy element	Good practice recommendation	Tick the statement that best reflects your policy
Title of the policy	To support findability, policies should have a title that makes clear whose policy it is and what the policy relates to.	• The policy has a clear and appropriate title.
		• The title of the policy would benefit from being made more explicit.
		• The policy does not have a title.
Year the policy was introduced	To support both human interpretation and machine actionability, the policy should make clear the period of validity, differentiating between the date it was written and the date it was implemented where necessary.	• The policy clearly states when it came into effect and provides a scheduled review date.
		• The policy clearly states when it came into effect but does not provide a scheduled review date.
		• The policy does not make clear when it came into effect or provide a scheduled review date.
Persistent Identifier (PIDs) of the policy document	PIDs should be assigned to clearly versioned policies to ensure that the right version can be found and fed into machine actionable pipelines.	• The policy has a persistent identifier such as a DOI.
		• The policy does not have a persistent identifier such as a DOI.
Policy is registered	Policies should be registered with services such as FAIRsharing or similar to improve their findability for various end users.	• The policy is registered with one or more registry services
		• The policy is not registered
Availability of the policy	Policies should be openly available and described consistently using a structured data markup schema to support both human and machine	• The policy is openly available online in a structured format.
		• The policy is openly available online (e.g., HTML, PDF).

Structured policy description template

Includes policy checklist elements and enables policy makers create and share structured versions of their data policies

FAIRsFAIR Policy Characterisation Form			
Number	Section	Policy Element	Options
1	Policy Context	Title of policy	Free text
2		Availability of the policy	
3		Country/Countries	Free text
4		Year policy was introduced	Free text
5		Status of policy	
6		Policy is scheduled for review	
7		If scheduled for review, when will this take place?	Free text
8		Persistent Identifier (PID) for policy (i.e., not a question about whether policy requires PIDs but rather is there a PID for the policy document itself)	
9		Policy is registered	
10		Subject area/discipline covered	
11		Sub-discipline if applicable	Free text
12		Type of policy	
13		Scope of the policy	
14		Definition of data provided	
15		Sharing research software	Suggested / Required / Not covered / Other
16		Data sharing	
17		Alignment with FAIR principles	
18		Metadata sharing	
19		Exceptions to data sharing are allowed	
20		Policy includes expectation on data management planning (DMP) development	
21		If a DMP should be developed, please indicate the stage at which it should be produced	
22		Is there an expectation that the DMP will be updated and clearly versioned?	

Resources are available from https://fairsfair.eu/policy-recommendations-and-support-programme

RDA Journal Policy Features

Policy feature definitions.

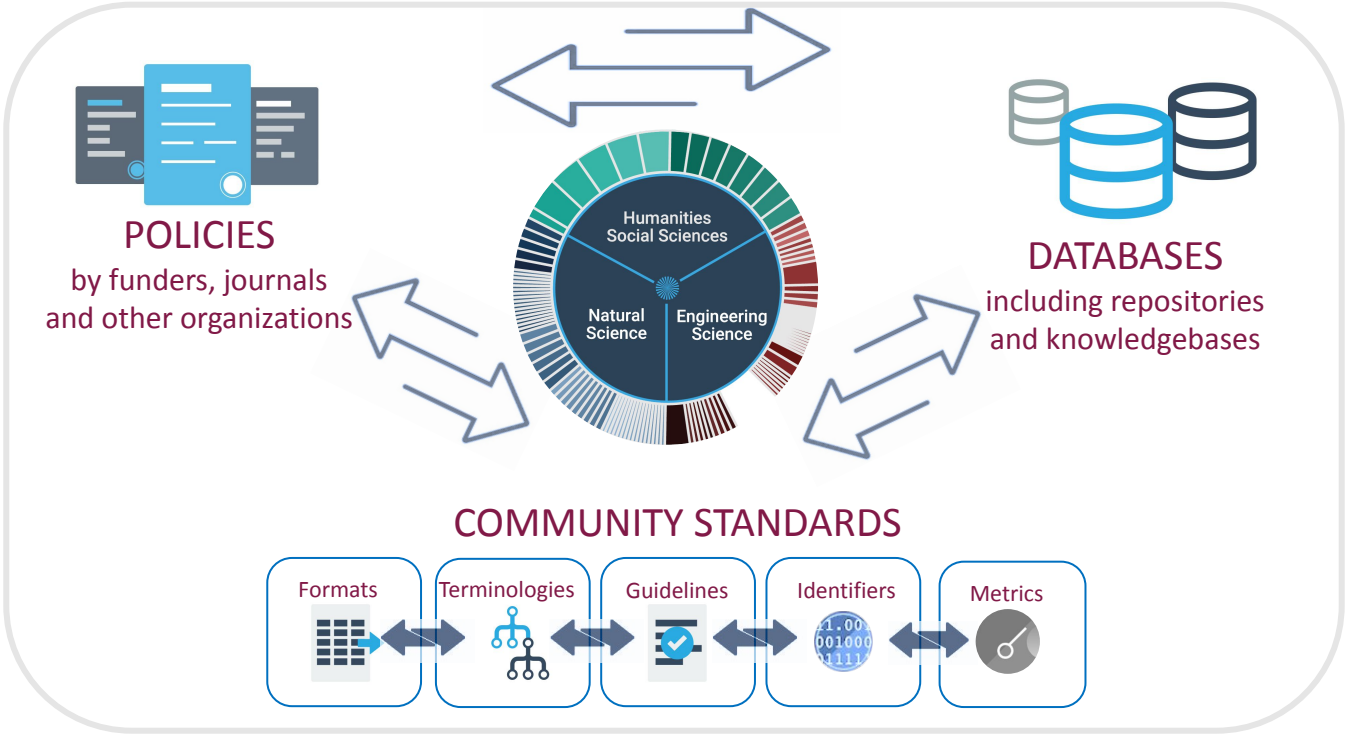
Feature	What	Why
Definition of research data	Define which research data the policy applies to, and the types of research data covered by the policy.	This enables the policy to define its scope and, where appropriate, provide general or discipline specific information on research data and file and format types (**Hodson & Molloy, 2015**). Specifying non-numeric data types (images, video, text etc) helps ensure relevance and applicability across research disciplines.
Definition of exceptions	Define what data do not need to be, or should not be made publicly available, under the policy and the alternative options for describing the availability of these data.	Ensures data policy is applicable to all research publications, but acknowledges legitimate exceptions and makes clear the policy does not create new legal or ethical precedents.
Embargoes	Define if and what embargoes on data release are permitted.	Researchers' reasonable right of first use of data generated during their research is a widely accepted principle of data sharing (**Anon, 2016b**), but reasonable lengths of embargo may vary by discipline, data type and study.
Supplementary materials	Define the journal/publisher's position on data sharing via supplementary materials, and if and when sharing data as supplementary materials is permitted under the policy.	While many policies preference sharing data in repositories (**McCarthy, 2009**; **Santos, Blake & States, 2005**; **Evangelou, Trikalinos & Ioannidis, 2005**), sharing data as supplementary materials remains very common. Some journals have strong data sharing policies and specify supplementary materials as the mechanism for data sharing. Supplementary materials are often a solution for researchers without discipline specific repositories and the definitions of supplementary material and research data often overlap.
Data repositories	State position on the use of data repositories. Data repositories	Lack of an appropriate repository or lack of awareness of repositories are common

http://doi.org/10.5334/dsj-2020-005

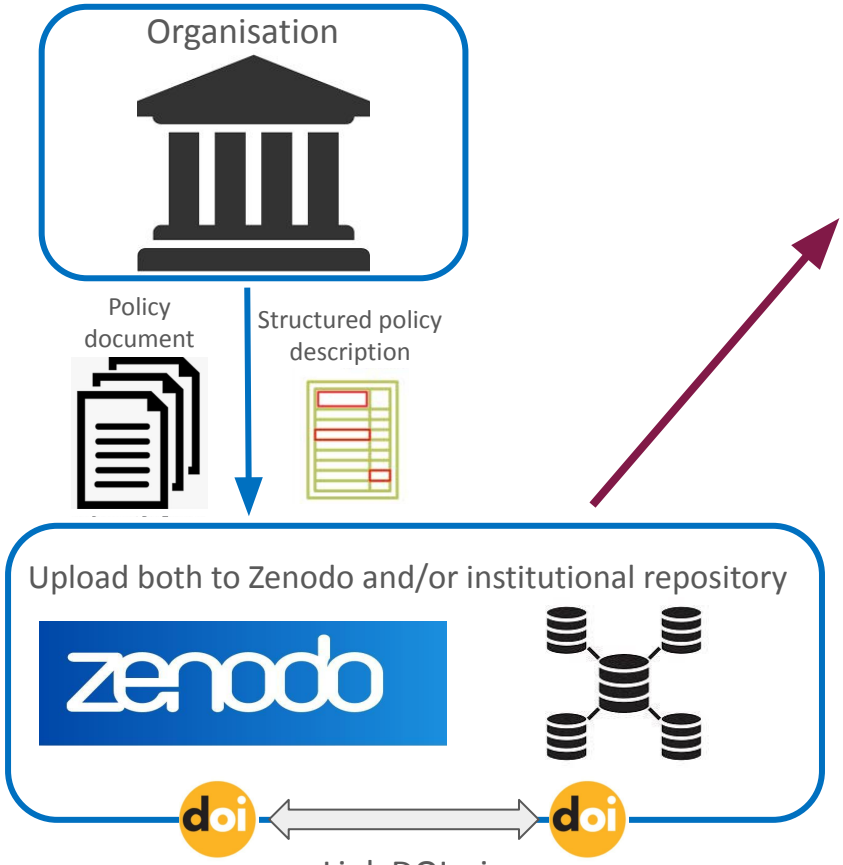
The **RDA Data Policy Standardisation and Implementation Interest Group** has defined and described **14 features** of journal research data policies and arranged these into a set of six standard policy types or tiers.

These can be adopted by journals and publishers to promote data sharing in a way that encourages good practice and is appropriate for their audience's needs.

FAIRsharing registry



FAIRsharing provides **curated descriptions** and **relationship graphs** of standards, databases and policies in **all disciplines**



FAIRsharing.org
standards, databases, policies

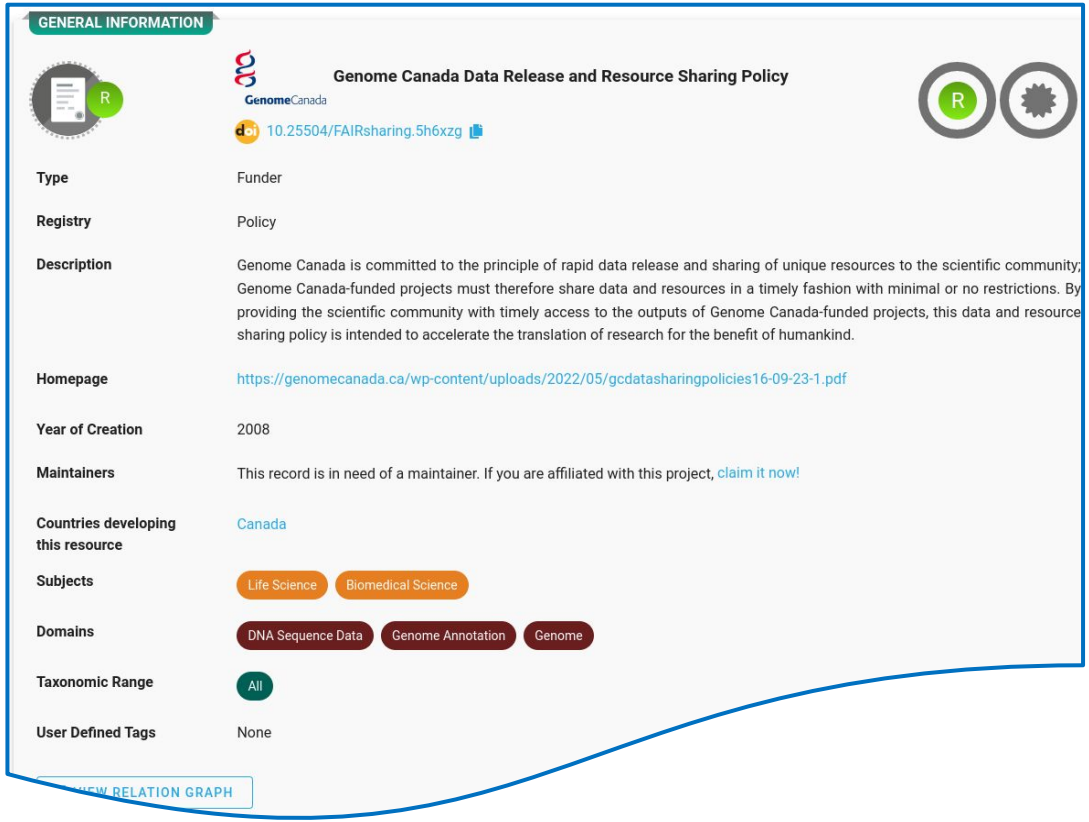
Register policy and include DOI to deposited versions and vice versa

- FAIRsharing recently updated their metadata to include most of the fields recommended by FAIRsFAIR so descriptions of content are more comprehensive
- Registering provides an openly available structured description of your data policy that is **machine readable**
- Very useful for monitoring the landscape with minimal effort!

Curating FAIRsharing Policy Metadata

*I found that adding [the new policy] metadata was **extremely helpful** for me to understand concretely what should go in such a Policy, ideally. And found out [a] particular policy is lacking many of these important elements ... Which I can now bring to their attention!*

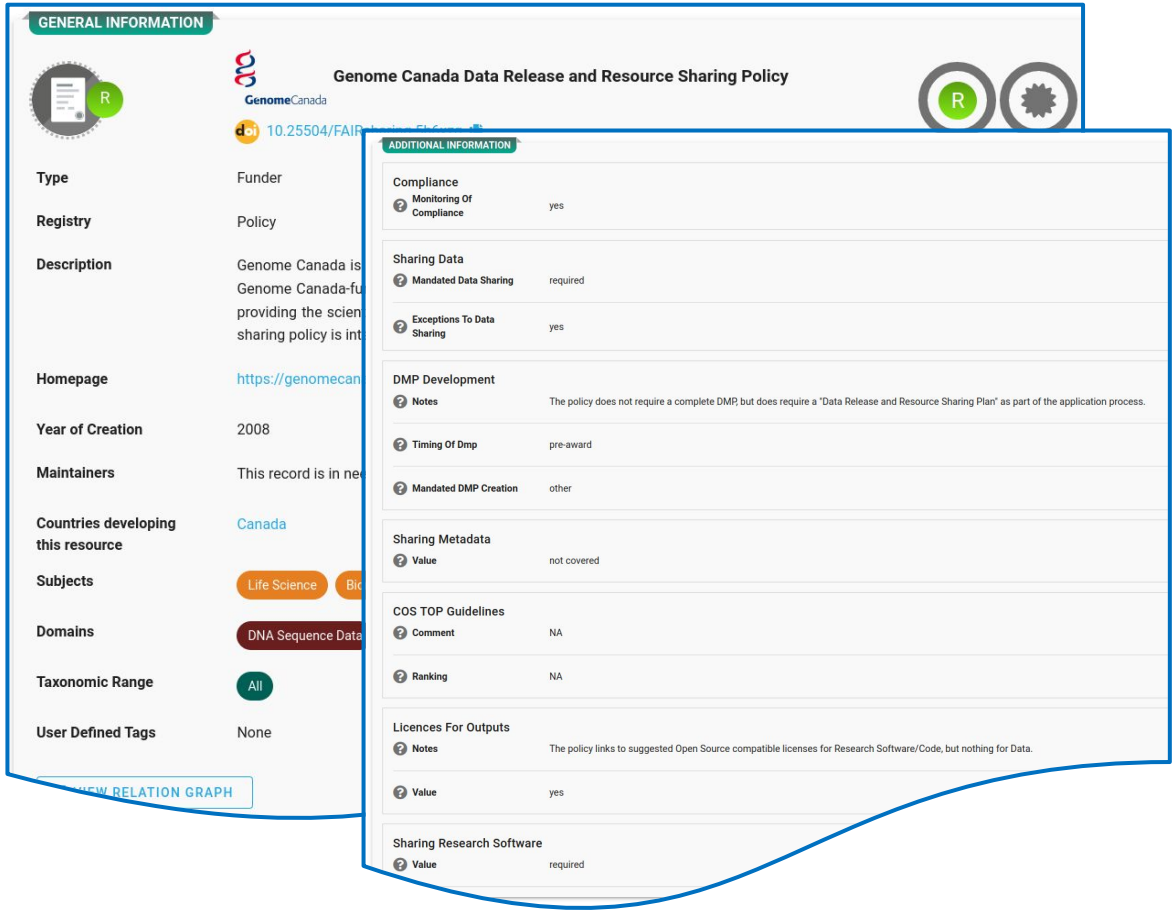
Gabriel Pelletier, FAIRsharing Community Curator for Neuroscience



Curating FAIRsharing Policy Metadata

*I found that adding [the new policy] metadata was **extremely helpful** for me to understand concretely what should go in such a Policy, ideally. And found out [a] particular policy is lacking many of these important elements ... Which I can now bring to their attention!*

Gabriel Pelletier, FAIRsharing Community Curator for Neuroscience



Benefits of this approach

- Ensures policies align with **FAIR**
- Ensures policymakers are **in control of updates**
- Utilises **freely-available repositories** and **registries**
- Policies and a structured description of their content are available to **multiple stakeholders** for a variety of uses (EOSC Association, OpenAIRE, EOSC Future, etc).

Thank you!

More information at
https://dcc.ac.uk/blog/fairsharing-and-dcc-collaborate-align-policy-metadata

https://blog.fairsharing.org/?p=451

@digitalcuration
@FAIRsFAIR_EU
@jd162a

RDA Europe / EOSC Future Domain Ambassador for standards, repositories and policies

@FAIRsharing_org
@allysonlister
@allysonlister@fediscience.org

COST – Promoting and spreading excellence

16 November 2022, Prague

cost
EUROPEAN COOPERATION
IN SCIENCE & TECHNOLOGY

Funded by
the European Union

European Cooperation in Science & Technology

+300
Running Actions

250
Researchers per Actions

30
Countries per Action

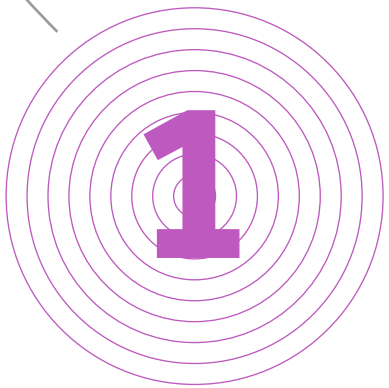
45,000
Researchers involved in COST Actions

Horizon Europe =
Widening participation and strengthening the European Research Area pillar

cost
EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

Role of COST in the ERA

3 STRATEGIC PRIORITIES



Promoting and spreading **excellence**

Fostering **interdisciplinary** research for breakthrough science

Empowering and retaining **young researchers**

cost
EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

For many of these 300 COST networks, data sharing lies at the heart of their activities

- Leukaemia gene discovery by data sharing, mining and collaboration
- Performance and reliability of photovoltaic systems: evaluations of large-scale monitoring data
- Mobilising Data, Policies and Experts in Scientific Collections

Over more than 50 years, COST Actions have shown innovative ways in how to collect, manage, use and visualize data

COST Rules include a clear commitment by COST to promote Open Access to research outputs resulting from COST Actions.

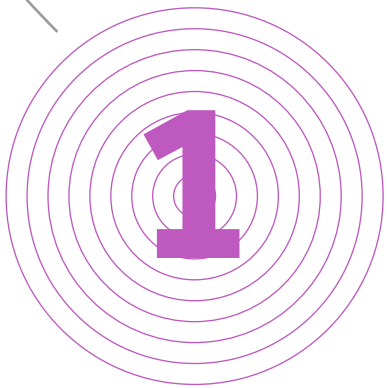
Since September 2022 COST Actions have the opportunity to submit their scientific publications directly to Open Research Europe (ORE), the European Commission's Open Access publishing platform for research.

Example of COST Success Stories on data related activities

- COST Action MP1304
- Running from 16/05/2013 - 24/11/2017
- NewCompStar - Exploring fundamental physics with compact stars
- On 10 April 2019, the world discovered the first-ever picture of a black hole in stunning detail

"Neutron stars like to become black holes," says Action Chair Prof. Luciano Rezzolla of Goethe University Frankfurt, Germany, who also played a prominent role in the collaboration that had presented the first-ever image of a black hole. "The two have a lot in common, and so the expertise that was developed and nurtured in NewCompStar also provided fertile ground for this imaging work."

How to engage with COST?



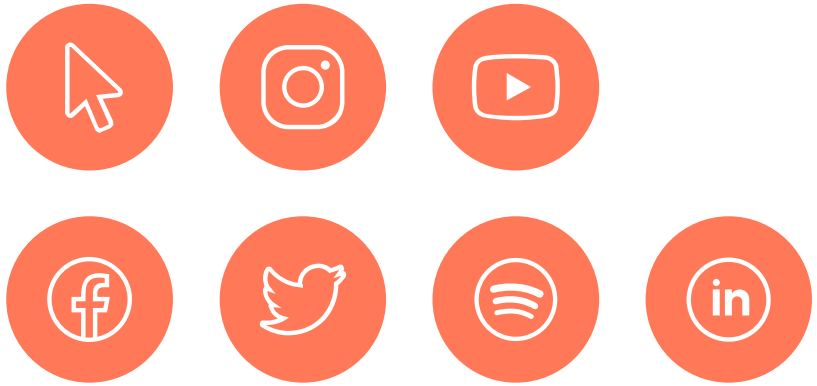
As a **participant**
of a running Action

As a **proposer**
of a new COST Action

As **expert** to evaluate
submitted proposals

cost
EUROPEAN COOPERATION
IN SCIENCE & TECHNOLOGY

THANK YOU!



Contact and engage with us!

cost
EUROPEAN COOPERATION
IN SCIENCE & TECHNOLOGY

Funded by
the European Union

WorkflowHub

FAIR Workflow Registry

Justin Clark-Casey
EOSC Programme Manager, EMBL-EBI
ELIXIR Europe

On behalf of:
Professor Carole Goble
The University of Manchester, UK
ELIXIR-UK

WorkflowHub Club

FAIR enabling practices, EOSC Symposium, Prague, CZ 2022-11-16

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824087

Computational Workflows make EOSC data analysis FAIR

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824087

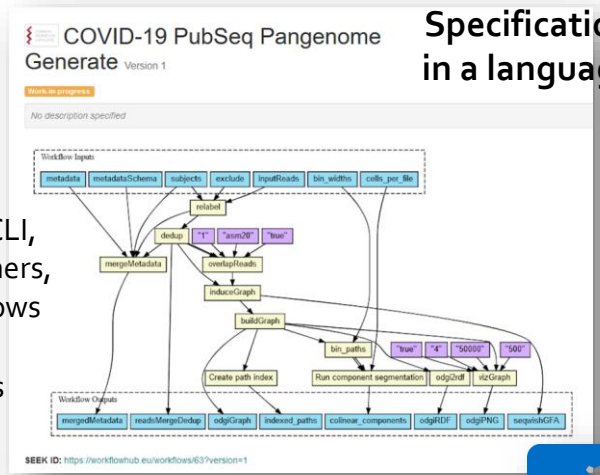
What is a workflow …

Precise Specification in a language

inputs

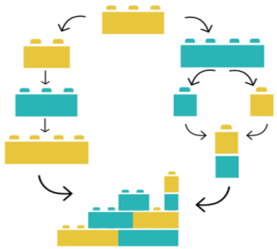
tools, CLI, containers, workflows

outputs



WfMS Engine

Computational data pipelines and analytics, chain codes, handle data flow, manage execution on computational platforms.



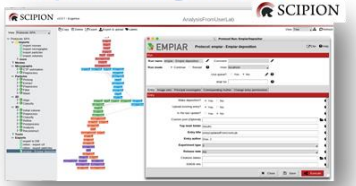
Access EOSC computation and data infrastructure, tool **interoperability**, processing **portability** and optimisation

Software Execution on EOSC computational platforms / clouds

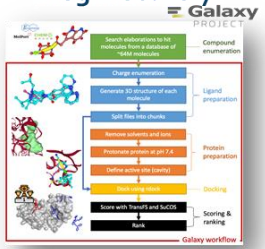
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824087

Increasing widespread tracking rise in data-driven reproducible analytics, data processing at scale & professionalisation of data pipelines.

CryoEM Image Analysis

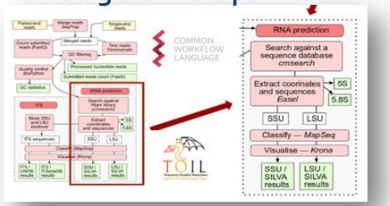


Drug Discovery

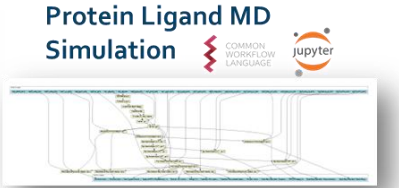


Genome Annotation

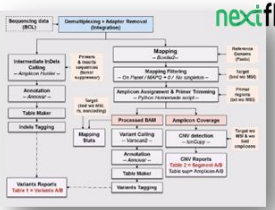
Metagenomic Pipelines



Protein Ligand MD Simulation



High Throughput Sequencing



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824087

Explosion of Workflow Management Systems

Interactive Electronic Research Notebooks
Scripting environments
Workflow management systems and execution platforms

This project has received funding from the European Union's Horizon

Peter Amstutz, Maxim Mikheev, Michael R. Crusoe, Nebojša Tijanić, Samuel Lampa, et al. (2022): **Existing Workflow systems.** *Common Workflow Language wiki*, GitHub. https://s.apache.org/existing-workflow-systems updated 2022-08-30, accessed 2022-08-30.

Are workflows
Findable, Accessible, Interoperable, Reusable?

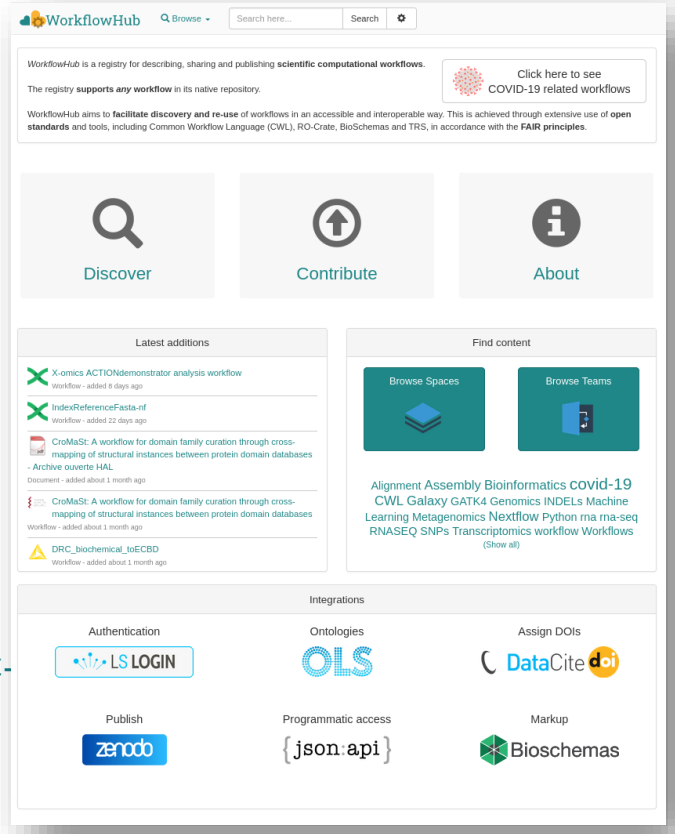
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824087

https://workflowhub.eu

WorkflowHub

FAIR Workflow Registry

Launched Sept 2020
EOSC service provided by ELIXIR, EOSC-Life, The University of Manchester
Open Development



283 workflows

11 system types

110 teams

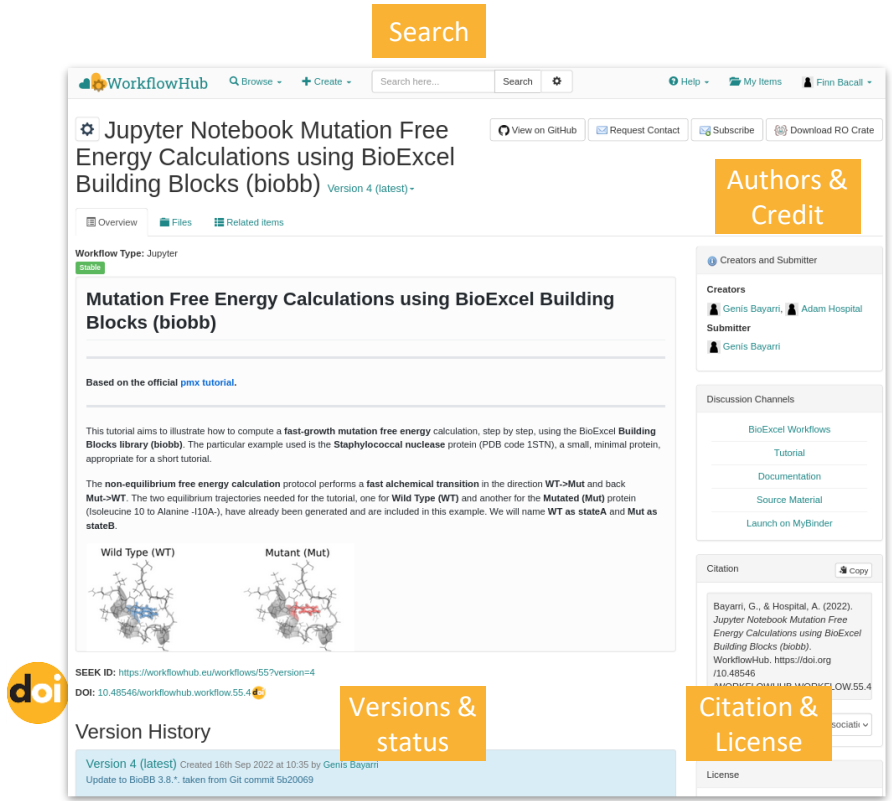
103 organisations

360 people

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824087

FAIR Workflow registry for EOSC

- Workflow-system **agnostic**
- Search for and **discover** workflows
- **Metadata** standardization
  (CWL, schema.org, custom tags, RO-Crate)
- **DOI** publication, citation & credit
- **Collections,** Teams, Organizations and **Communities**
- **Programmatic** access: GA4GH TRS API, RO-Crate
- **Registry**, not repository
  Workflows can live elsewhere, e.g. GitHub
- Integration with EOSC **execution** platforms



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824087

WorkflowHub

Low barrier to entry for publishing workflows

- Workflow-system agnostic
- Import from native repositories
- Git integration
- Versions, any stage of development
- Automated metadata extraction
- Include documents, test data

tl;dr: Workflows can remain in existing repositories

GitHub

git

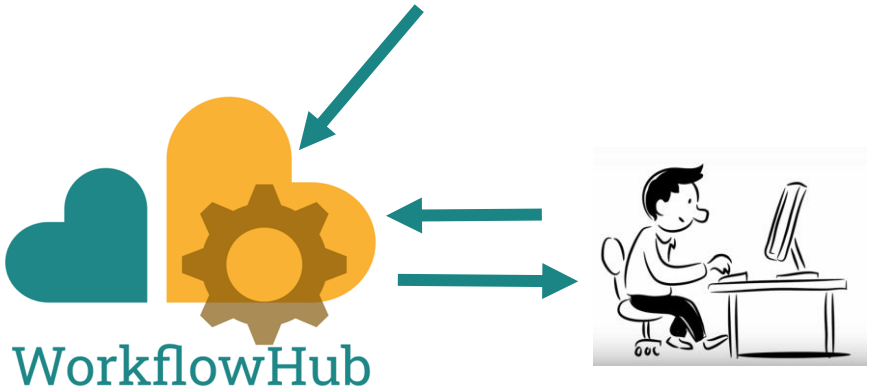
Snakemake workflow catalog

nf-core

A community effort to collect a curated set of analysis pipelines built using Nextflow.

IWC - Intergalactic Workflow Commission

Galaxy PROJECT

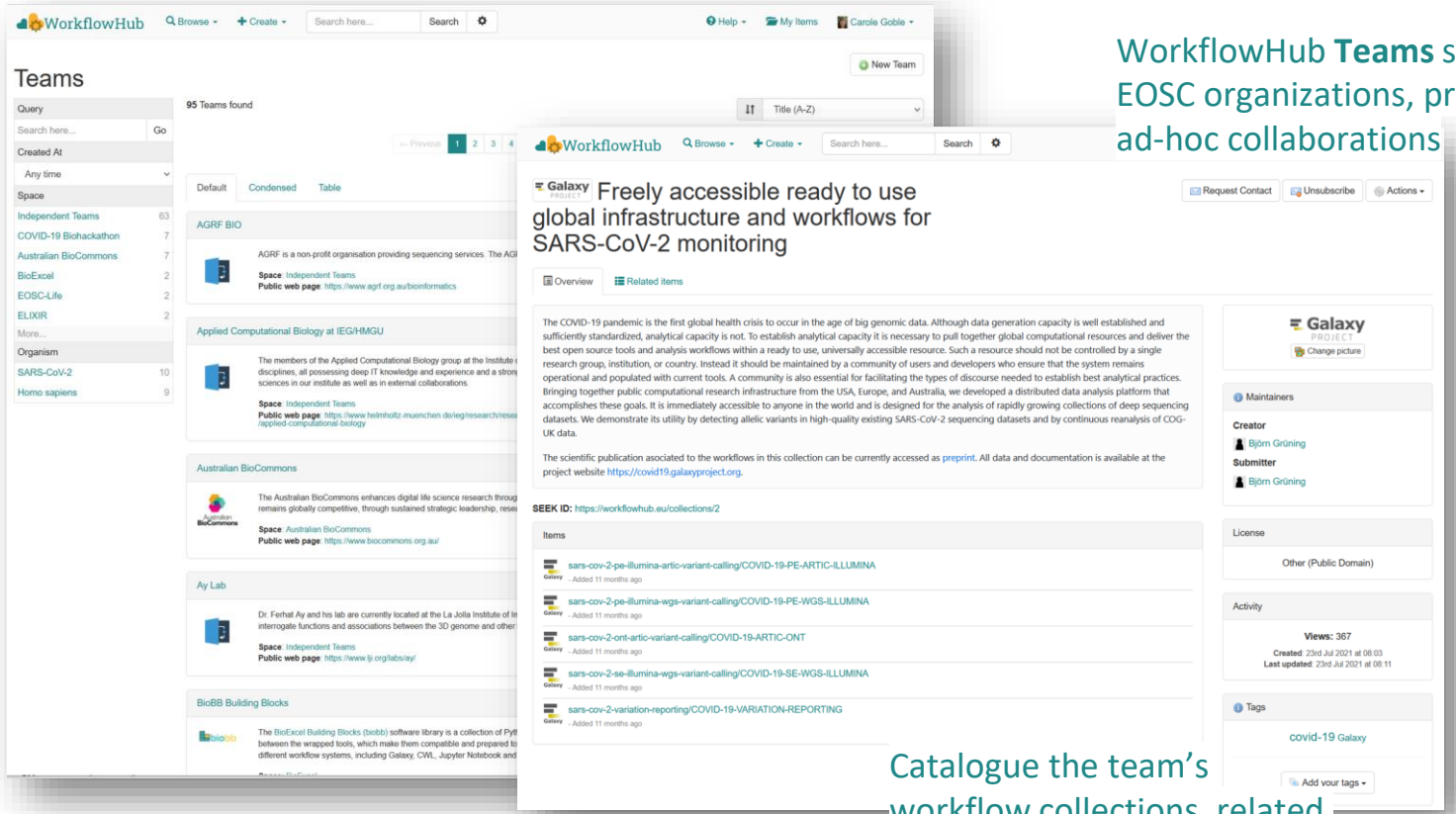


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824087

Helping EOSC communities make FAIR workflows

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824087

Supporting workflow communities in EOSC



WorkflowHub **Teams** serve numerous EOSC organizations, projects and ad-hoc collaborations

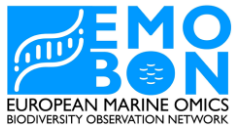
Catalogue the team's workflow collections, related resources, people and institutions

Request Contact

Subscribe

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824087

Communities at WorkflowHub



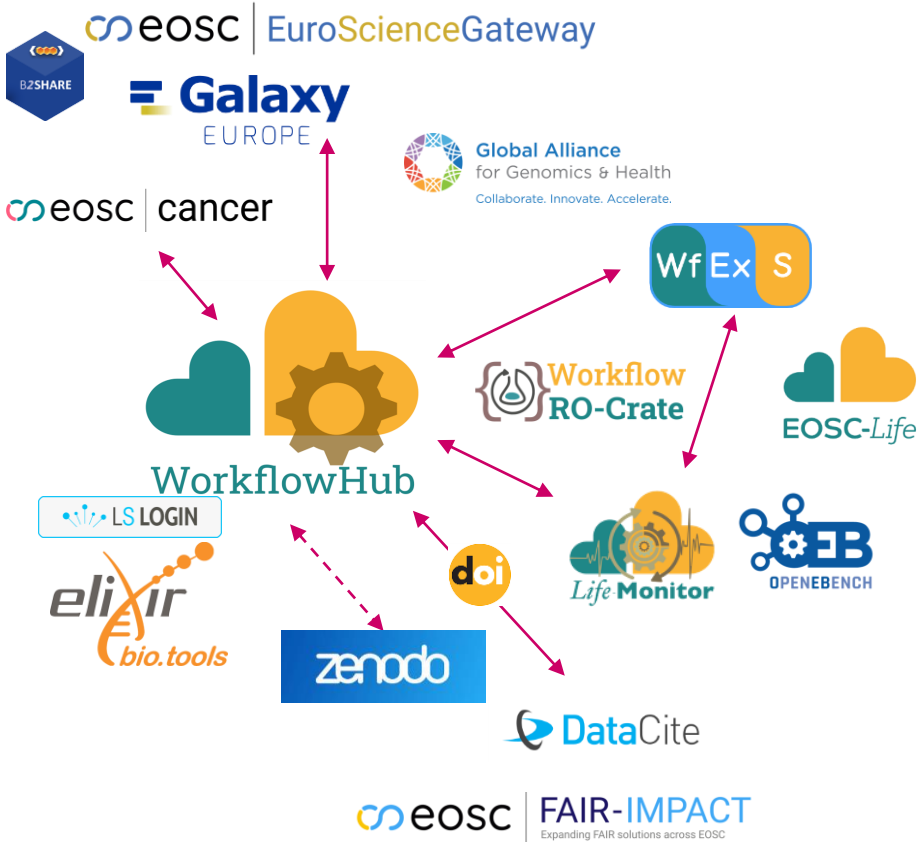
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824087

https://workflowhub.eu/projects

WorkflowHub integrates with EOSC services

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824087

WorkflowHub in the EOSC ecosystem



Services in the

Workflow Collaboratory exchange workflows as

FAIR Digital Objects using **RO-Crate**

- Packaging workflow files & companion objects
- Exchange between services & systems
- Reproducibility & Testing
- Citation
- Workflows in the EOSC PID graph
- Standards-based exchange

https://doi.org/10.5281/zenodo.4605654

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824087

Acknowledgements

WorkflowHub Club

https://about.workflowhub.eu/

Open Development

Bi-weekly community calls

Workflow Community Initiative https://workflows.community/about
EOSC-Life https://www.eosc-life.eu/
ELIXIR http://elixir-europe.org
RO-Crate https://www.researchobject.org/ro-crate/
WorkflowHub https://workflowhub.eu/
Galaxy Europe https://galaxyproject.eu/
Bioschemas https://bioschemas.org
Common Workflow Language https://www.commonwl.org/
Life Monitor https://crs4.github.io/life_monitor/

FAIR Computational Workflows at WCI

Finn Bacall, The University of Manchester
Stuart Owen, The University of Manchester
Stian Soiland-Reyes, The University of Manchester

National Initiatives for Open Science in Europe

RoLECT: a self-assessment tool against EOSC RoP Legal & Ethics Compliance

FAIR Enabling Practices EOSC Symposium 2022, Prague, 14-17 November 2022

Eleni Toli
ATHENA RC
NI4OS-Europe Project Director

What is RoLECT?

- Self-assessment tool against EOSC RoP focusing on Legal & Ethics aspects.
- Designed to promote compliance and identify non-compliance with EOSC RoP.
- Prospective resource providers have a hands-on opportunity to verify if the resources meet the legal & ethics standards of EOSC Rules of Participation.
- Targeted audience: any resource providers familiar / concerned with legal & ethics conditions related to EOSC RoP.
- Version released: June 2022

Ethics & Legal compliance as part of FAIR

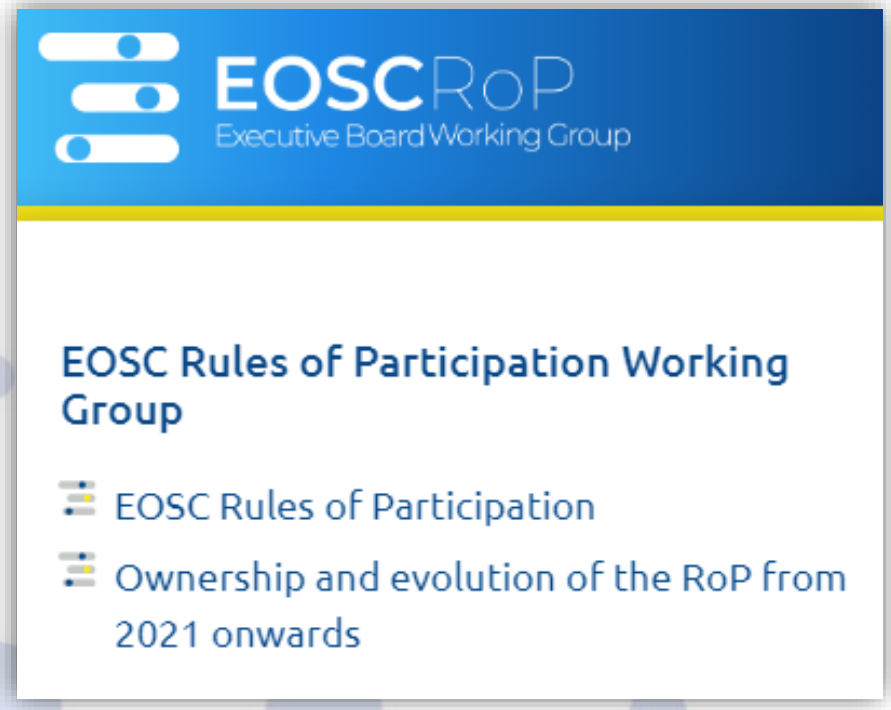
- Legal and ethics conditions are important elements of FAIR
  - have a great impact on how we grant access to resource, use it and allow its reuse
  - are important parts in policies of publishing platforms and funding bodies → and EOSC!

More than half of respondents say that the policies of funding bodies and publishers are the most influential when it comes to their RDM and data sharing

European Research Data Landscape Study, Executive Summary – Draft

The EOSC context

EUROPEAN OPEN SCIENCE CLOUD



eosc

Advisory Group	**Implementation of EOSC**
Task Force	**Rules of Participation Compliance Monitoring**

Aims

Provide criteria based on the principles expressed in the high-level EOSC RoP.

Define monitoring process of the different RoPs by setting up a dedicated framework

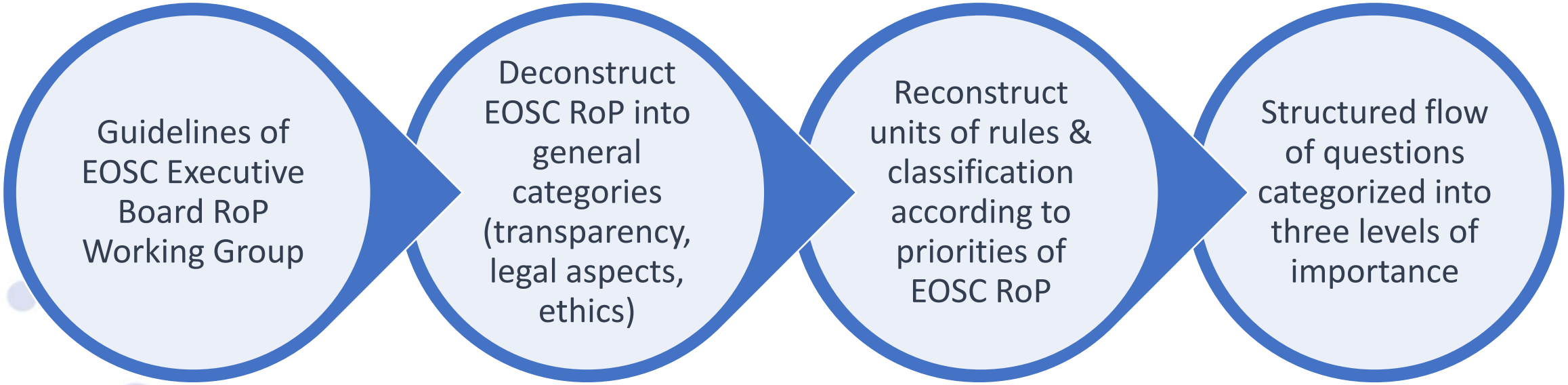
Flashback: Key principles of EOSC RoP v0.5

- EOSC is based on the principle of openness
- EOSC data align with FAIR principles
- EOSC services align with EOSC architecture & interoperability guidelines
- EOSC is based on principles of ethical behaviour and integrity
- EOSC users are expected to contribute to a successful EOSC and active EOSC community
- EOSC users adhere to the terms of use of the resources they use
- EOSC users reference the resources they use in their work
- Participation in EOSC is subject to the policy and legislation of the EU

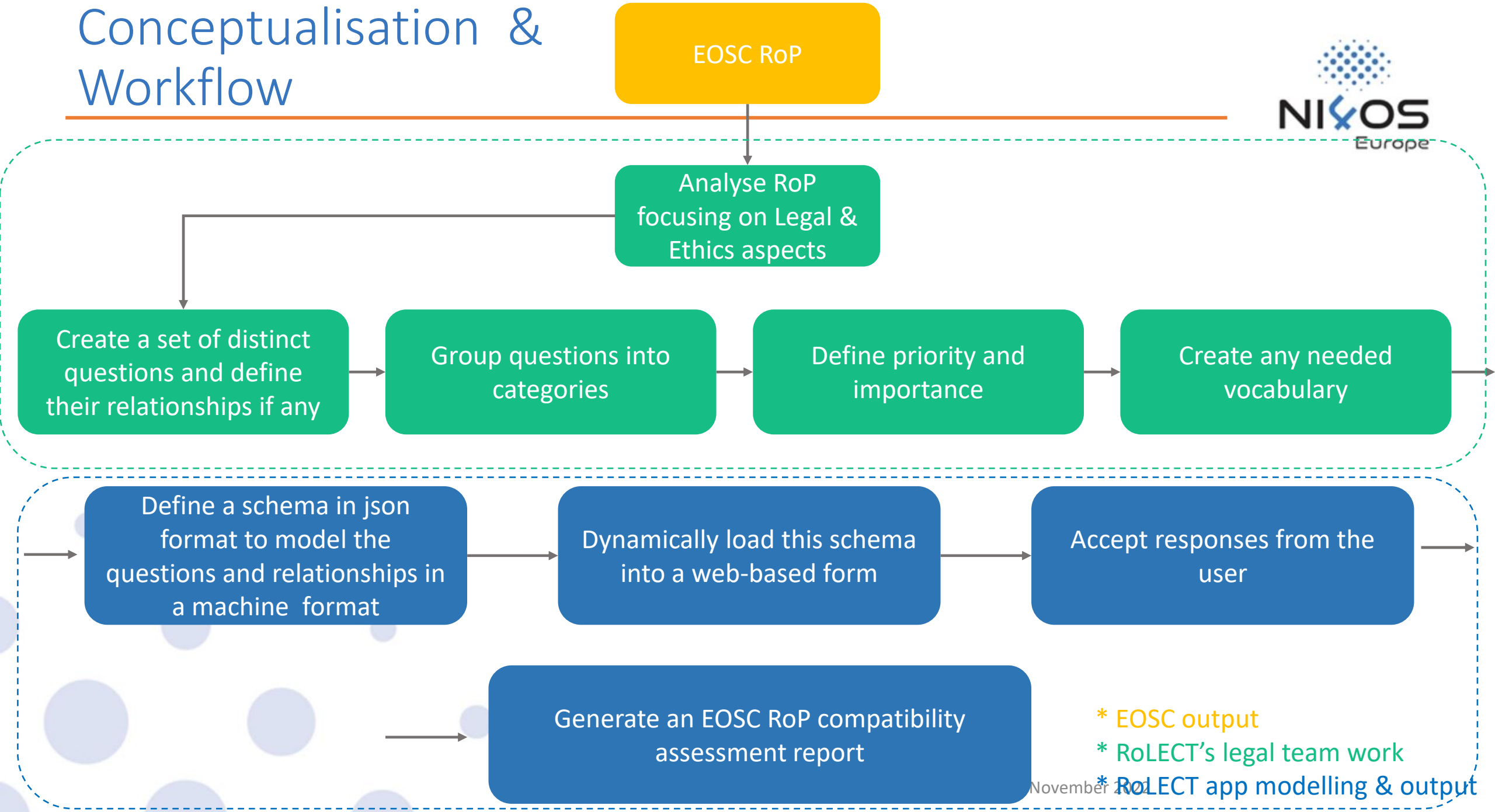
Legal background

- Personal Data: minimum level of compliance with GDPR → lawfulness of processing (Art 6§4 GDPR), information obligation (Art. 13, 14 GDPR), and data subjects' rights (Art. 15-22 GDPR).
- Definitions of PSI and trade secrets according to Directives (EU) 2019/1024 and 2016/943 respectively.

Development approach



Conceptualisation & Workflow



How does it work? The process

Step 1
General info about resource (URL, type etc.)

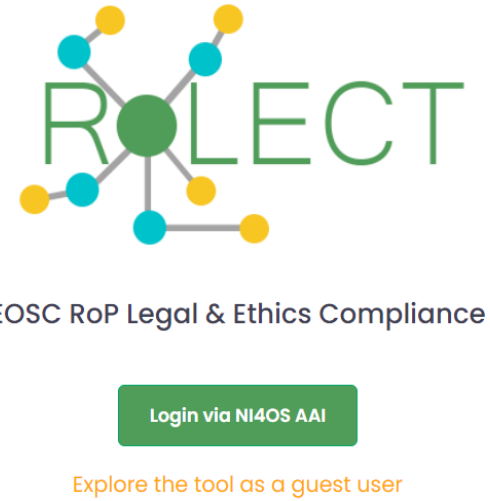
Step 2
If IP, select the types of IP and licensing conditions

Step 3
Check for PSI, cultural info, personal data, ethics.

Step 4
Overview & assessment report

→ Explore the tool as guest or authenticated user

https://rolect.ni4os.eu/rolect/auth/login

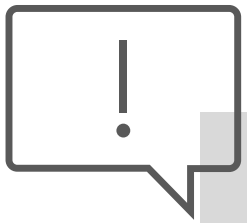


Levels of Importance

High Importance: Critical information that reflects the basic principles of RoP

Medium Importance: Information that is consistent with the EOSC RoP

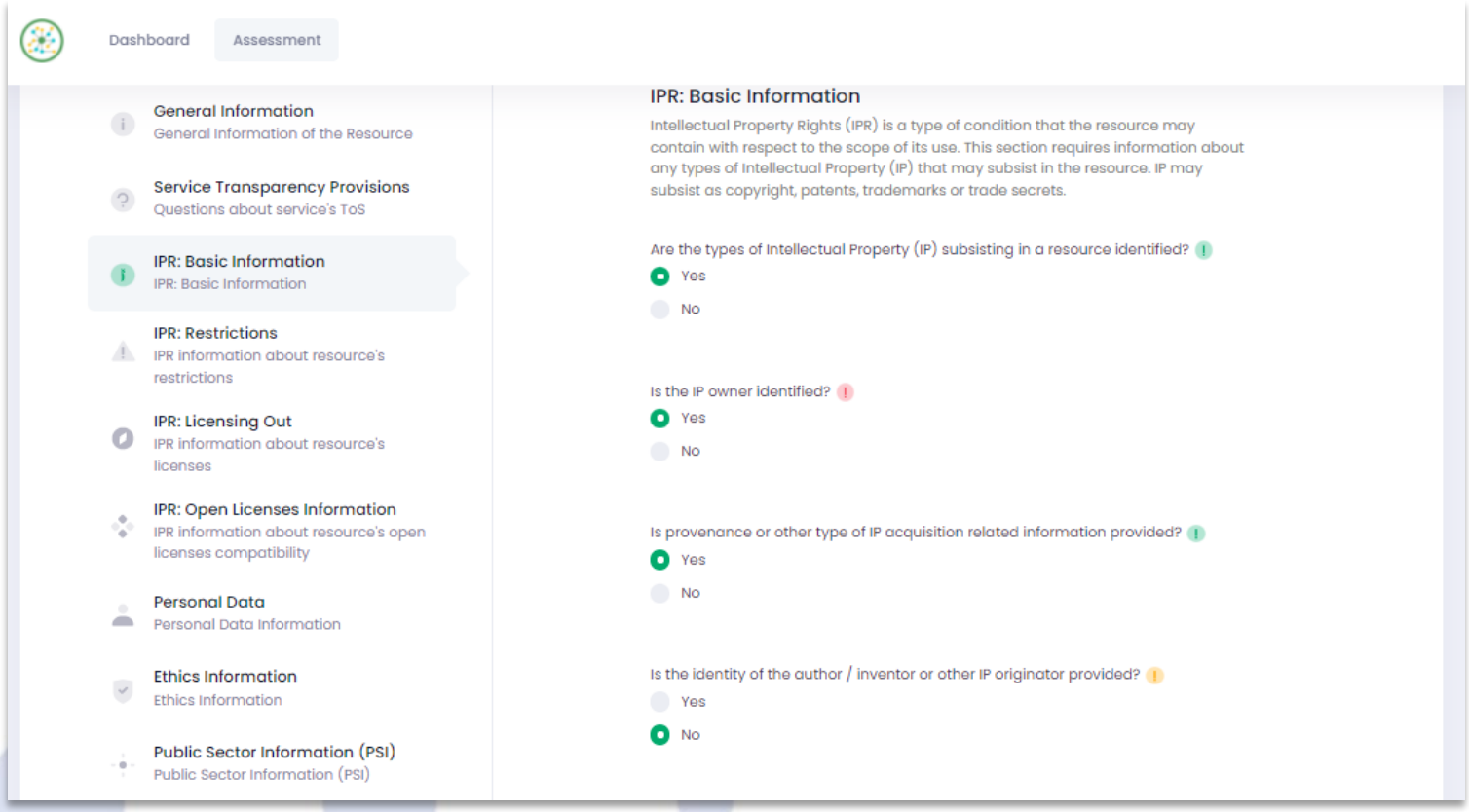
Low Importance: Information with a relatively low impact on the assessment result



help resource providers to assess what information or omission of information is most important

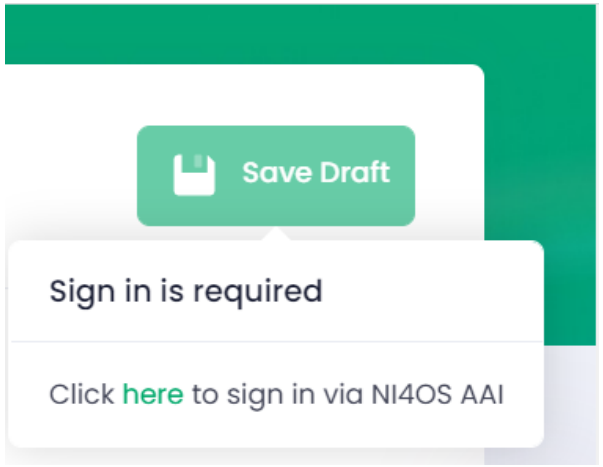
FAIR, legal & ethics aspects, boundaries and restrictions

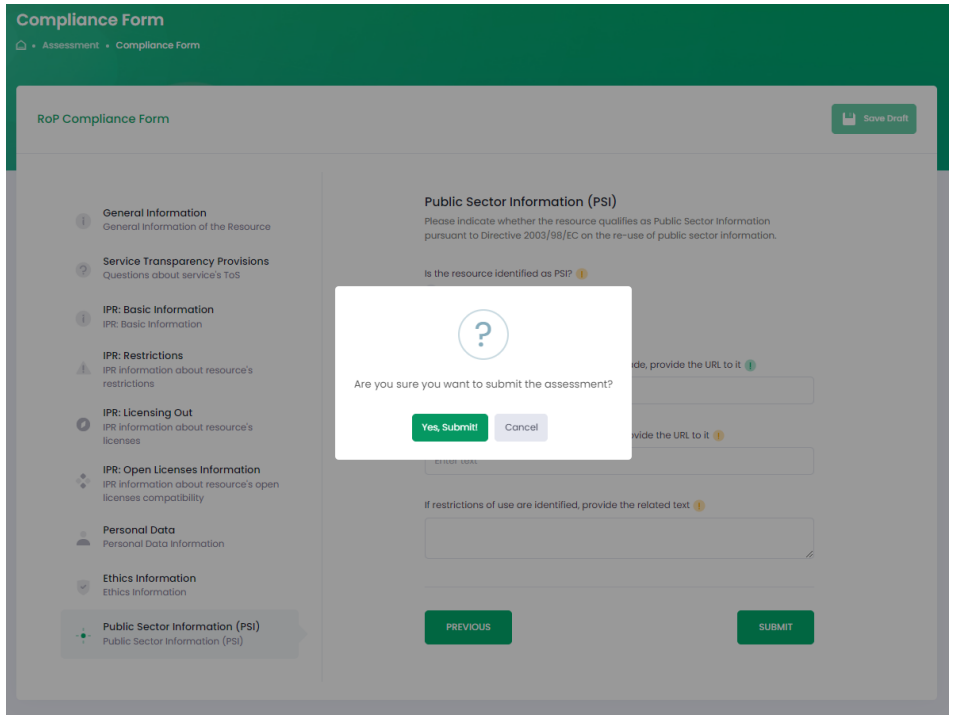
"As open as possible, as closed as necessary"



- Intellectual Property Rights
- Licensing conditions
- License compatibility assessment
- Public domain
- Cultural heritage law
- Code of Conduct
- Public Sector Information
- Trade secrets, confidential information

Assessment report





Thank you!

Eleni Toli, elto@athenarc.gr

NI4OS

NI4OS_eu

NI4OS.eu

Join NI4OS-Europe Community:
https://ni4os.eu/contact-us

ATHENA Research & Innovation Information Technologies

@openaire_eu

DMPs, FAIR principles & the European Landscape

Elli Papadopoulou

ATHENA RC / OpenAIRE

DATA MANAGEMENT PLANS - TODAY

Policy

- Public DMPs
- Rich content
- Support

Technical

- Search & retrieve
- Interoperability
- Semantics

Adoption

- FAIR practices
- Disciplines / Communities
- Data sharing

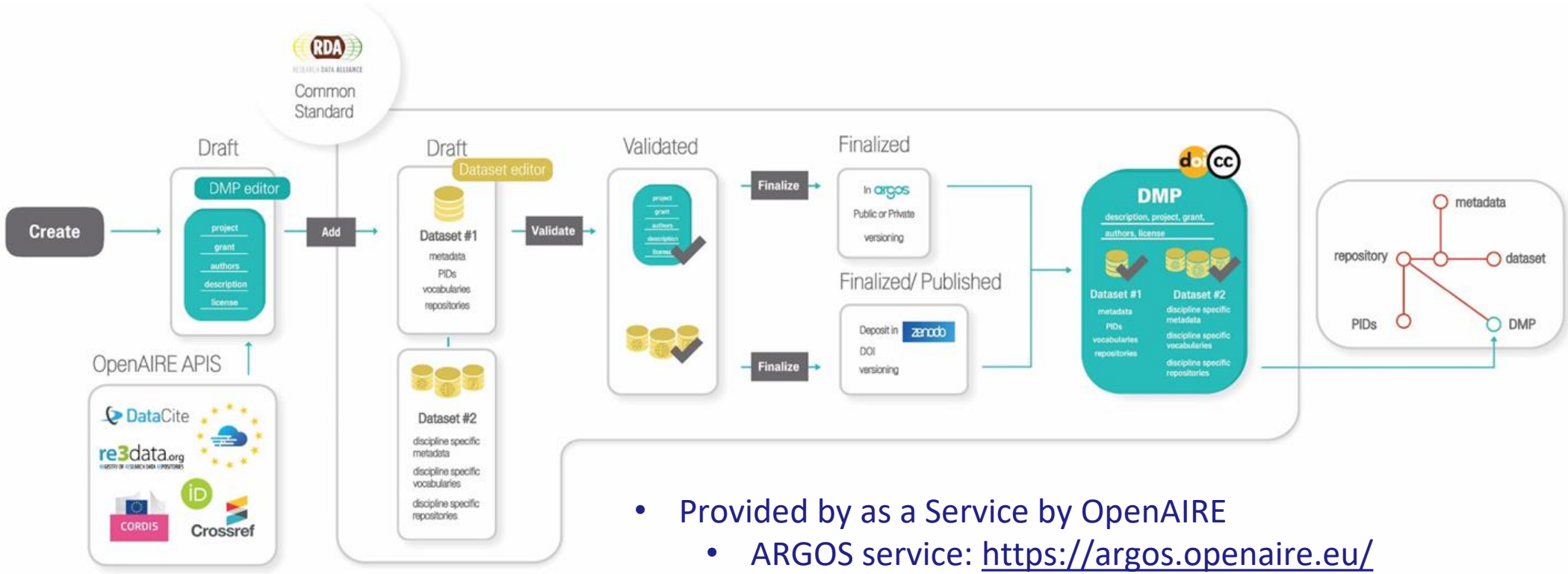
- Not Open Access
- Granularity of content
- Unclear about the process, e.g. where to start, how to get input

- Resource_type
- Reused datasets
- Qualified references

- Reproducibility
- Domain Data Protocols
- Input & Output Data

ARGOS – A GLIMPSE

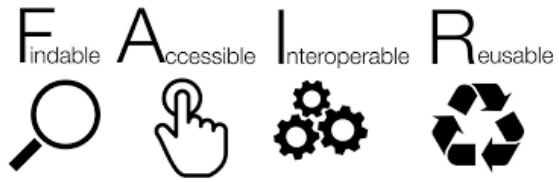
ARGOS is an open source, configurable and extensible **tool** for **planning Research Data Management** (RDM) activities according to **Open Access & FAIR data policies.**



- Provided by as a Service by OpenAIRE
  - ARGOS service: https://argos.openaire.eu/

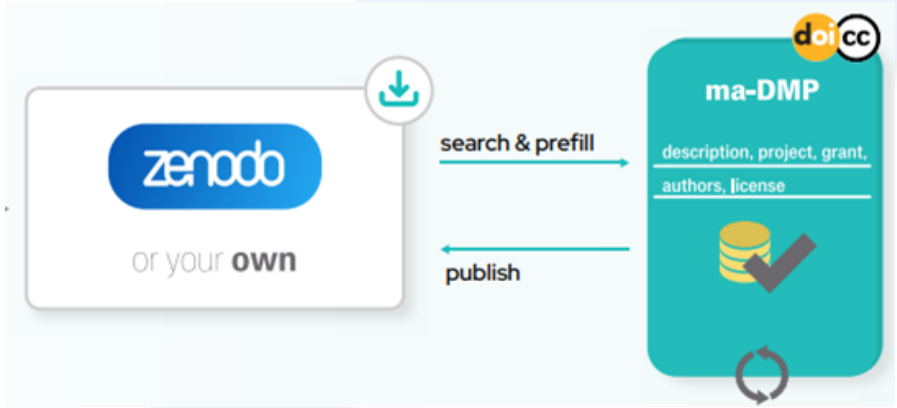
EOSC Resource (eosc-portal.eu)

ARGOS – OUTPUTS & PROCESSES



- **Findable**: OpenAIRE EXPLORE
- **Accessible**: PIDs (ORCIDs & DOIs)
- **Interoperable**: RDA DMP Common Standard
- **Reusable**: Licenses
- **Versioned** (history/provenance)
- **Published** and **preserved** in Zenodo

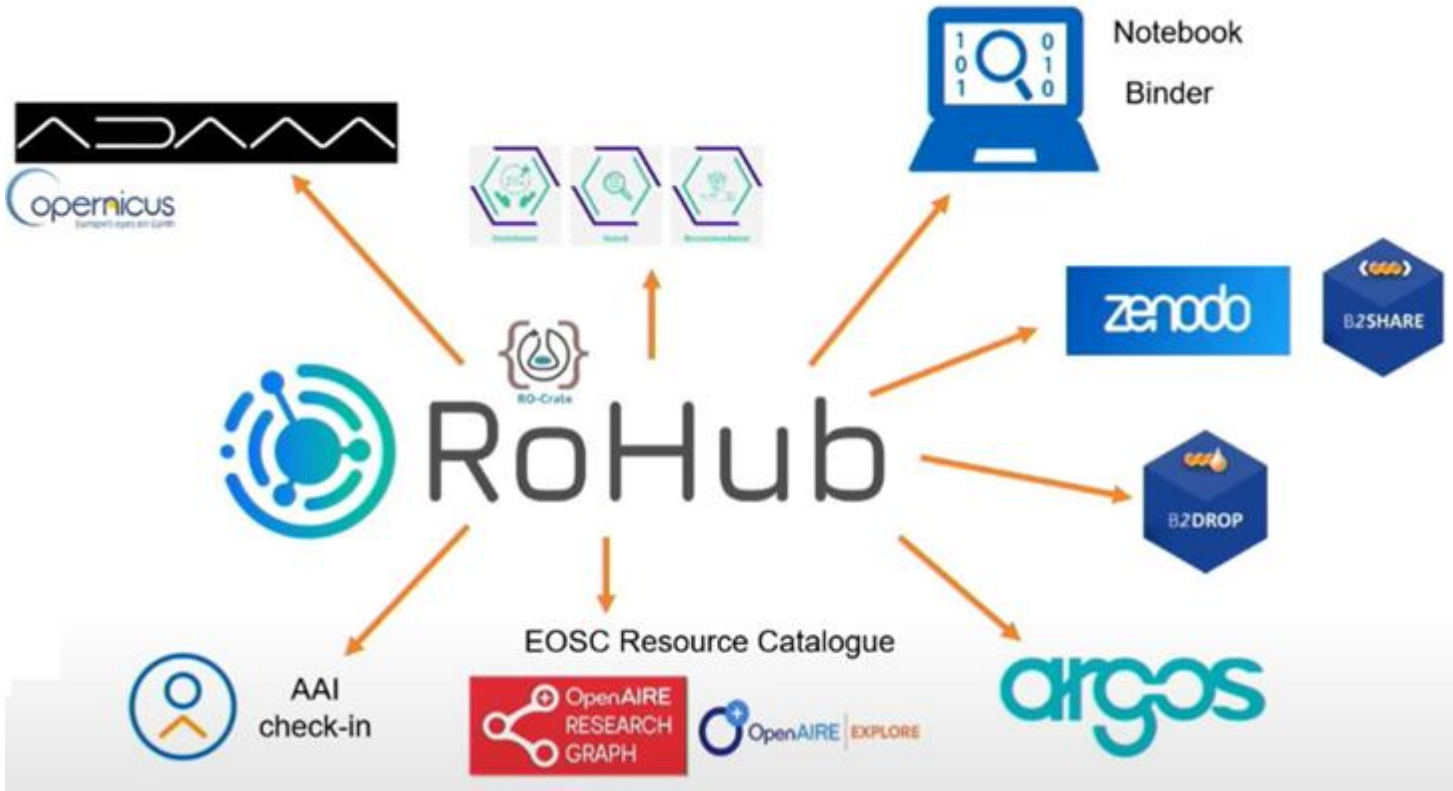
- **Automated**
  - **Writing:** inferred content and **allocation** in Templates as answers
  - **Searching:** Published Dataset / Software metadata
  - **Publication:** DMP metadata and files in Zenodo



ARGOS – NEW

Reliance

- **Collaborations**
  - Funders and Institutions
    - Deployments and Integrations with local research services
    - Migrating mechanism to move DMPs from other software to ARGOs
- **Templates**
- **Features**
  - Extend publishing mechanism
  - Extend semantics
  - Machine actionable table for input
- **Feedback** from the community



ARGOS - COMMUNITY

Argos

Community Calls

Are you a researcher or administrator of Argos? Got questions on how to write your Data Management Plan (DMP) or how to create your Template and connect DMPs with other data services and outputs? Join us and learn more!

OpenAIRE is running a series of **community calls for Argos** to **support all researchers** in meeting their **Horizon Europe requirements** by creating **FAIR** (**Findable, Accessible, Interoperable, Reusable**) **DMPs**. Similarly, it supports **all research performing and funding organisations** to orchestrate their data services around Argos and **connect data workflows** contributing to **interconnected Research Data Management ecosystems**.

These calls offer the opportunity to discover Argos novelties and learn how to benefit from them in your practice, share feedback and discuss the future of DMPs as FAIR and machine actionable outputs, i.e. as complete outputs that bring validated information, qualified references and automations to the table to assist the processes of collecting, documenting and publishing your data.

The **Argos Community Calls** will run **every last Wednesday of the month at 14.00 CEST**, starting from June 29th!

https://www.openaire.eu/argos-community-calls

THANKS

Email
elli.p@athenarc.gr

Twitter
@elli_lib

ORCID
0000-0002-0893-8509

RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

recherche.data.gouv.*fr*

Research data support at Recherche Data Gouv

@RechercheDataGv

RÉPUBLIQUE FRANÇAISE
Liberté Égalité Fraternité

recherche.data.gouv.*fr*

What is Recherche Data Gouv?

An ecosystem providing...

A sovereign solution for publishing and reporting data

- A federated national research data platform...
- … complementary to the thematic repositories...
- ... planned to integrate EOSC services, offering access to shared and open research data

Research Data Management Support

- Train, support and raise the awareness of research teams throughout the data life cycle

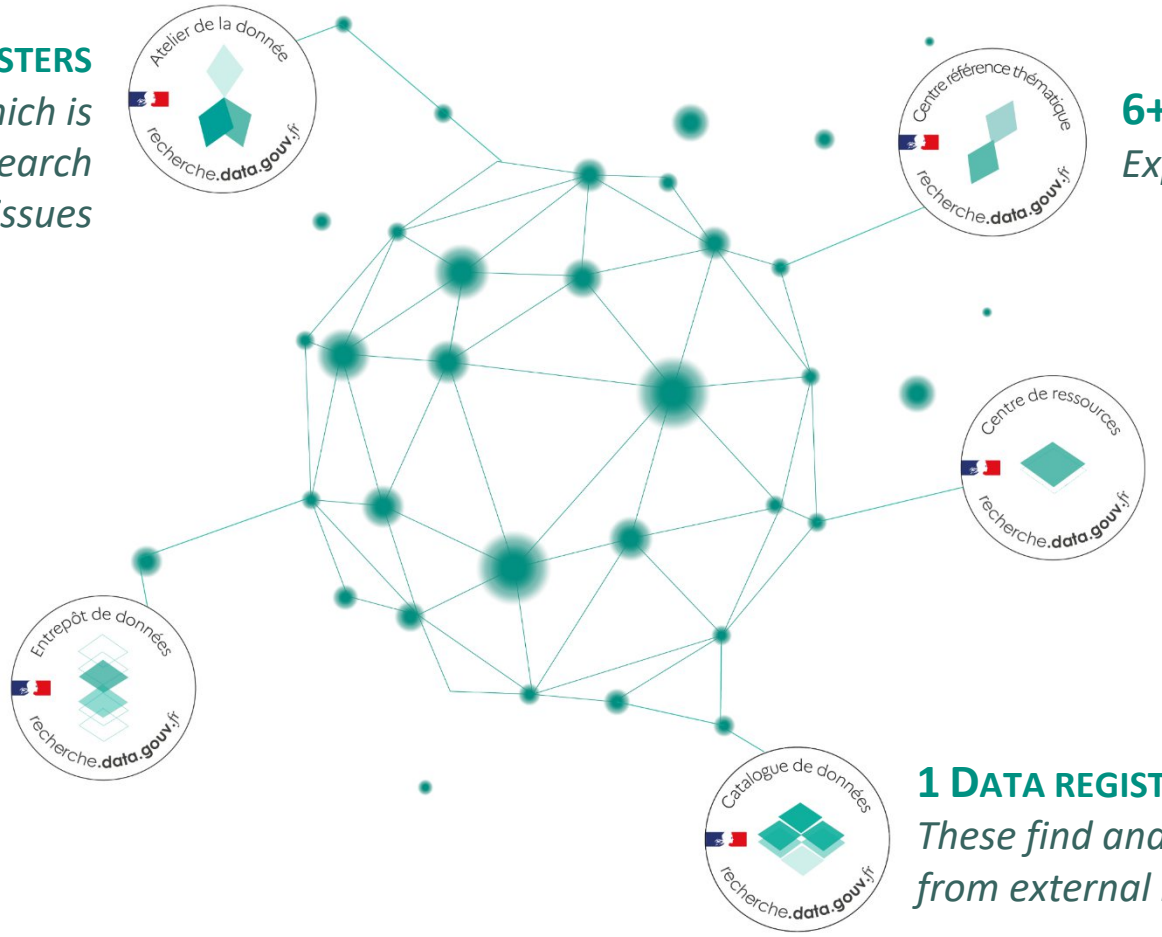
RÉPUBLIQUE FRANÇAISE
Liberté Égalité Fraternité

recherche.data.gouv.*fr*

An ecosystem of five types of stakeholders

13+ DATA MANAGEMENT CLUSTERS

Generalist expertise which is geographically close to the research teams it supports on all data issues



6+ THEMATIC REFERENCE CENTRES

Expertise in a given scientific field

4 RESOURCE CENTRES

These support the data management clusters and capitalise on their practices

1 DATA REPOSITORY

A mutualised repository for all institutions to deposit and publish data

1 DATA REGISTRY

These find and harvest data from from external repositories

@RechercheDataGv

RÉPUBLIQUE FRANÇAISE
Liberté Égalité Fraternité

recherche.data.gouv.fr

Data management clusters

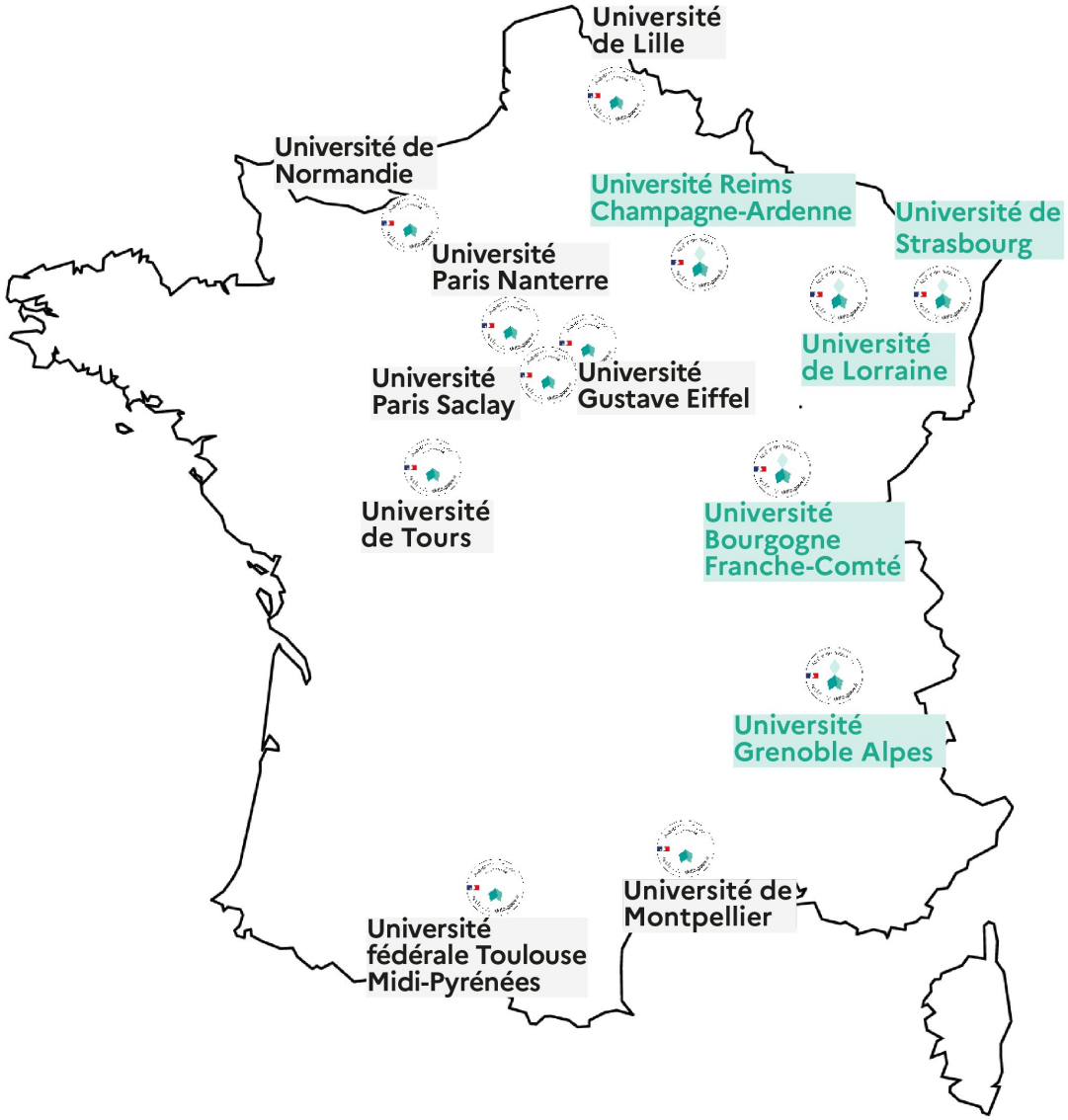
Atelier de la donnée
recherche.data.gouv.fr

Generalist expertise, close to the research teams

Shared services and skills for organisations from a given territory

These clusters are progressively developed

- *As they are designed by the universities involved*
- *Following successive calls for expressions of interest (three a year as of the end of 2023)*



@RechercheDataGv

RÉPUBLIQUE FRANÇAISE
Liberté Égalité Fraternité

recherche.data.gouv.*fr*

Thematic reference centres

Centre référence thématique
recherche.data.gouv.fr

IFB
INSTITUT FRANÇAIS DE BIOINFORMATIQUE

PROGEDO
DATA INFRASTRUCTURE

DATA TERRA

CDS

PNDB Pôle National de Données de Biodiversité

HN Huma-Num IR*

Expertise in a given scientific field

Definition and dissemination of **best practices and international standards for data management, processing and dissemination** by scientific field

The first landscape put forward:

Research infrastructures that structure data management and dissemination for their scientific community

@RechercheDataGv

RÉPUBLIQUE FRANÇAISE
Liberté Égalité Fraternité

recherche.data.gouv.*fr*

Resource centres

Centre de ressources recherche.data.gouv.fr

Training and e-training resource centre

Recherche Data Gouv

*Sharing and certification of **teaching tools** along with the development and provision of **generalist e-training***

Skills resource centre

Recherche Data Gouv

*A system aimed at **increasing the skills of the data management clusters***

Shared tools for the clusters resource centre

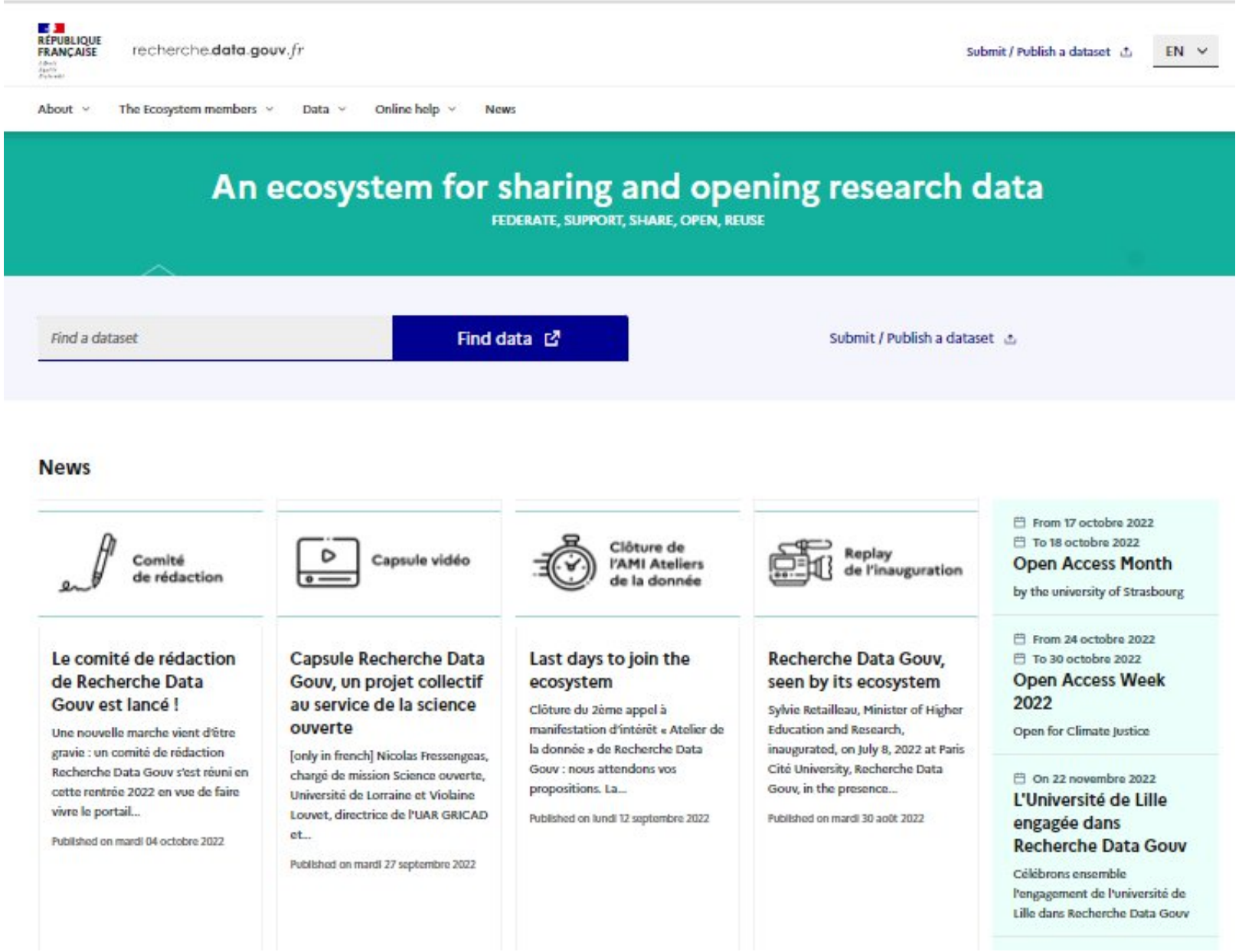
Recherche Data Gouv

*Development and maintenance of **shared tools for the data management clusters***

Repository-registry resource centre

Recherche Data Gouv

Development and maintenance of the repository-registory along with user support



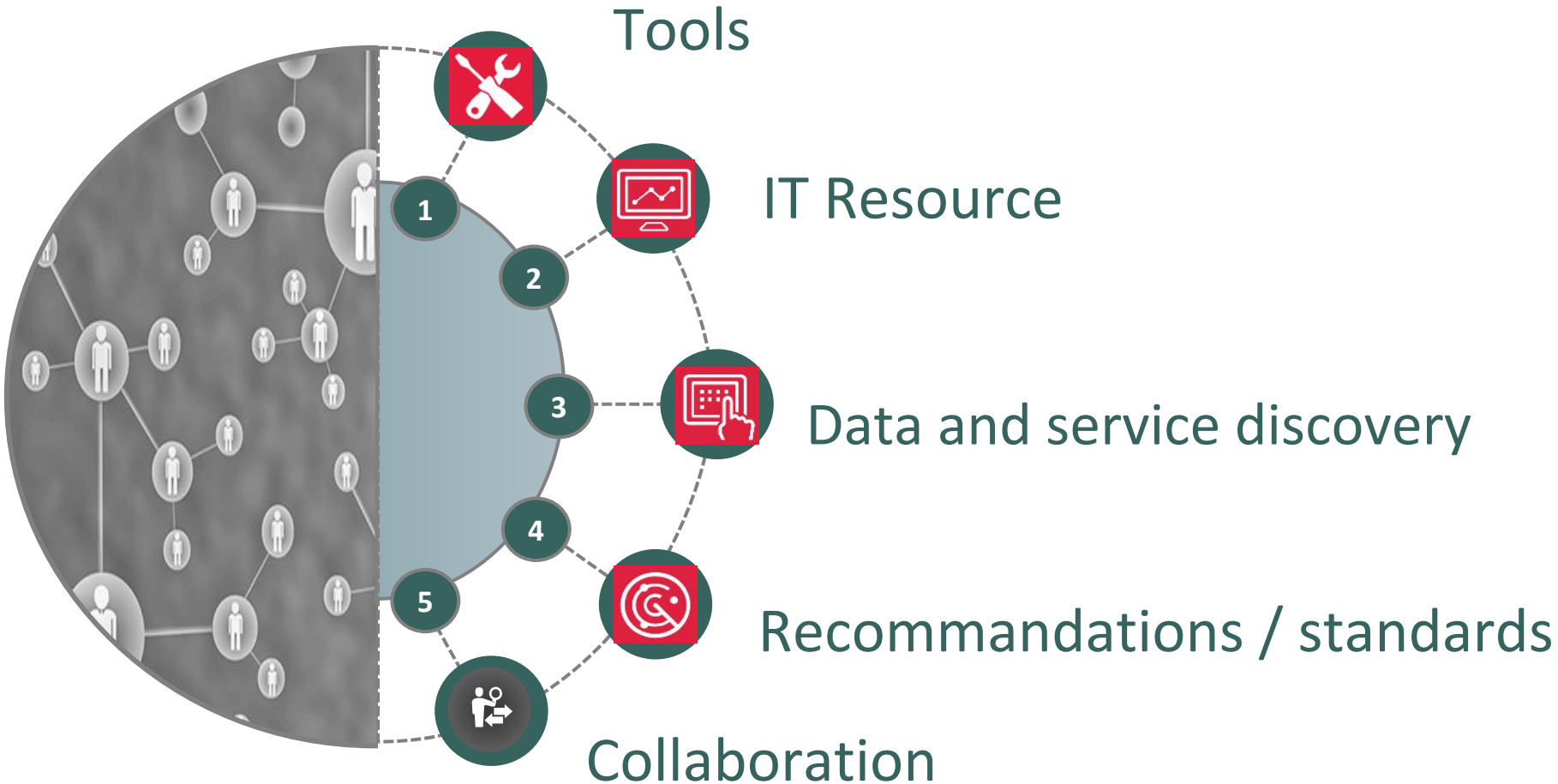
https://recherche.data.gouv.fr/en

@RechercheDataGv

RÉPUBLIQUE FRANÇAISE
Liberté Égalité Fraternité

recherche.data.gouv.fr

EOSC – Recherche Data Gouv: several points of convergence



Tools

IT Resource

Data and service discovery

Recommandations / standards

Collaboration

@RechercheDataGv

RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

recherche.data.gouv.*fr*

Thank you!

https://recherche.data.gouv.fr/en

@RechercheDataGv

https://www.linkedin.com/company/recherche-data-gouv/

Making EOSC Research Objects FAIR with RO-Crate

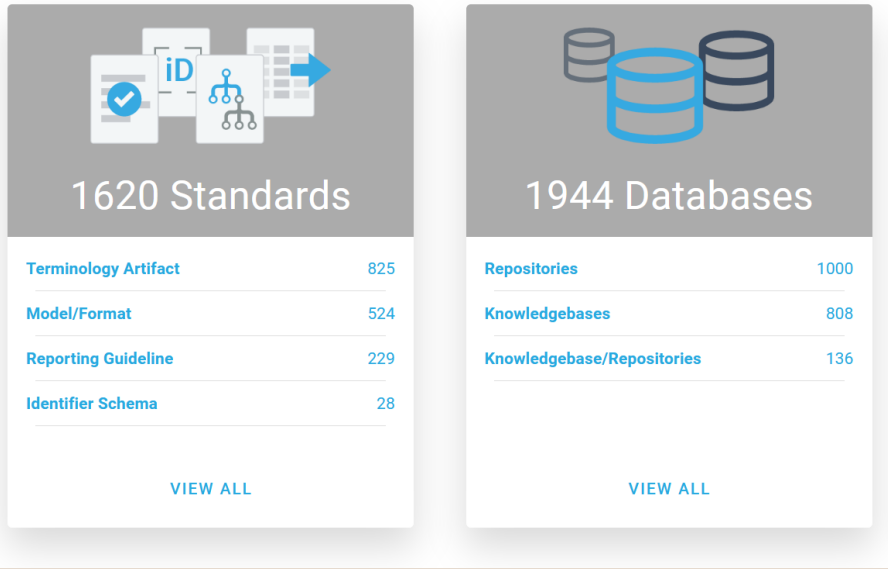
Justin Clark-Casey
RO-Crate Community
EMBL-EBI
ELIXIR Europe

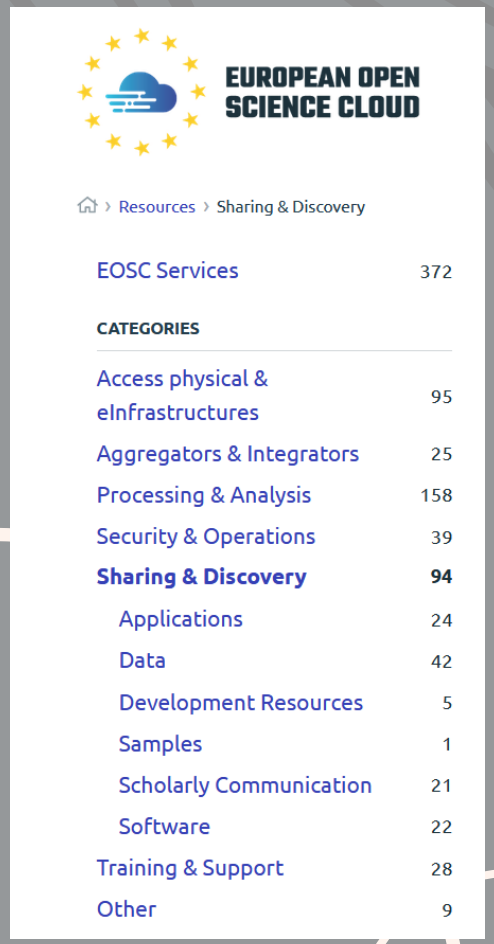
Stian Soiland-Reyes
RO-Crate Community
The University of Manchester
ELIXIR Europe

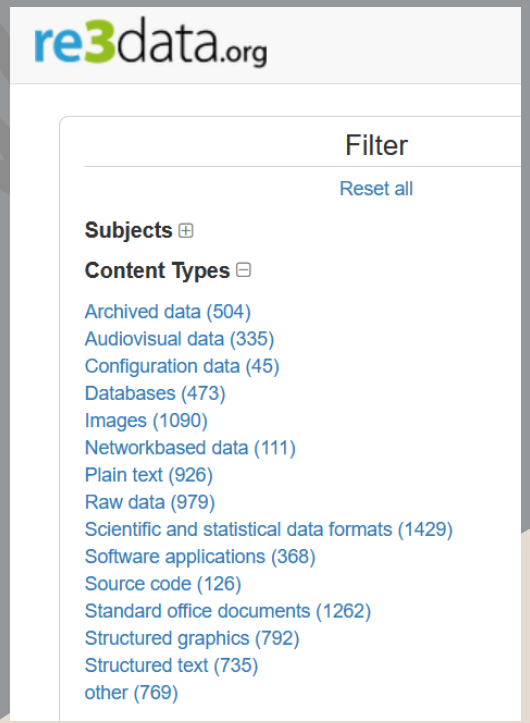
EOSC Symposium
Prague, CZ, 2022-11-16

Is it FAIR to use these repositories?

FAIRsharing.org
standards, databases, policies







eosc

Researchers are asked to make their research outputs FAIR – where to publish?

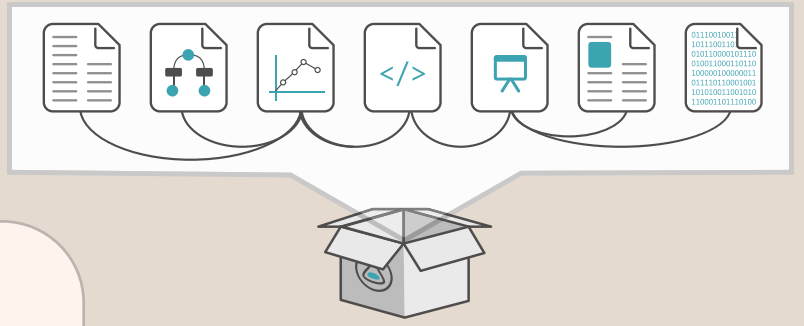
Thousands of public, institutional and domain-specific repositories

Help from guidance and catalogues (FAIRsharing, re3data, EOSC Catalogue)

..but how to gather and reference outputs across multiple repositories?

What about contextual information?

Aims of FAIR Research Objects



Describe and package data collections, datasets, software etc. with their metadata

Platform-independent object exchange between repositories and services

Support reproducibility and analysis: link data with codes and workflows

Transfer of secure/large distributed datasets with persistent identifiers

Aggregate citations and persistent identifiers

Propagate provenance and existing metadata

Publish and archive mixed objects and references

Reuse existing standards, but hide their complexity

research object.org

Practical and general purpose packaging of Research Objects

Infrastructure independent – avoiding repository/service silos
Practical, lightweight, robust

Familiar, developer friendly, web native, machine- and human-readable, search-engine accessible
Adoptable Linked Data JSON and PIDs

Embrace diversity, legacy, unknowns, open-ended, multi-interpretation, self-describing, interlingua
Adaptable Metadata Profiles

RO-Crate

https://www.researchobject.org/ro-crate/

Realizing FAIR Digital Objects with RO-Crate

Reference existing repositories
Re-use Web standards (JSON-LD, schema.org)
Persistent identifiers w/FAIR Signposting
Add context: people, projects, etc.

PID

RO-Crate Metadata file

id
type
description
datePublished
...

author

organisation

license

Structured metadata about the RO-Crate and content

JSON-LD
schema.org

RO-Crate Content

PID

files

id
type
description
datePublished
creator
size
format
...

PID

directories

By reference (PID, URL)

PDF

https://doi.org/10.5281/zenodo.5841615

https://github.com/o/script

https://www.researchobject.org/ro-crate/specification.html

RO-Crate in practice

RO-Crate is used by multiple international projects

Applied across research domains – from **life sciences** to **cultural heritage**

https://www.researchobject.org/ro-crate/in-use/

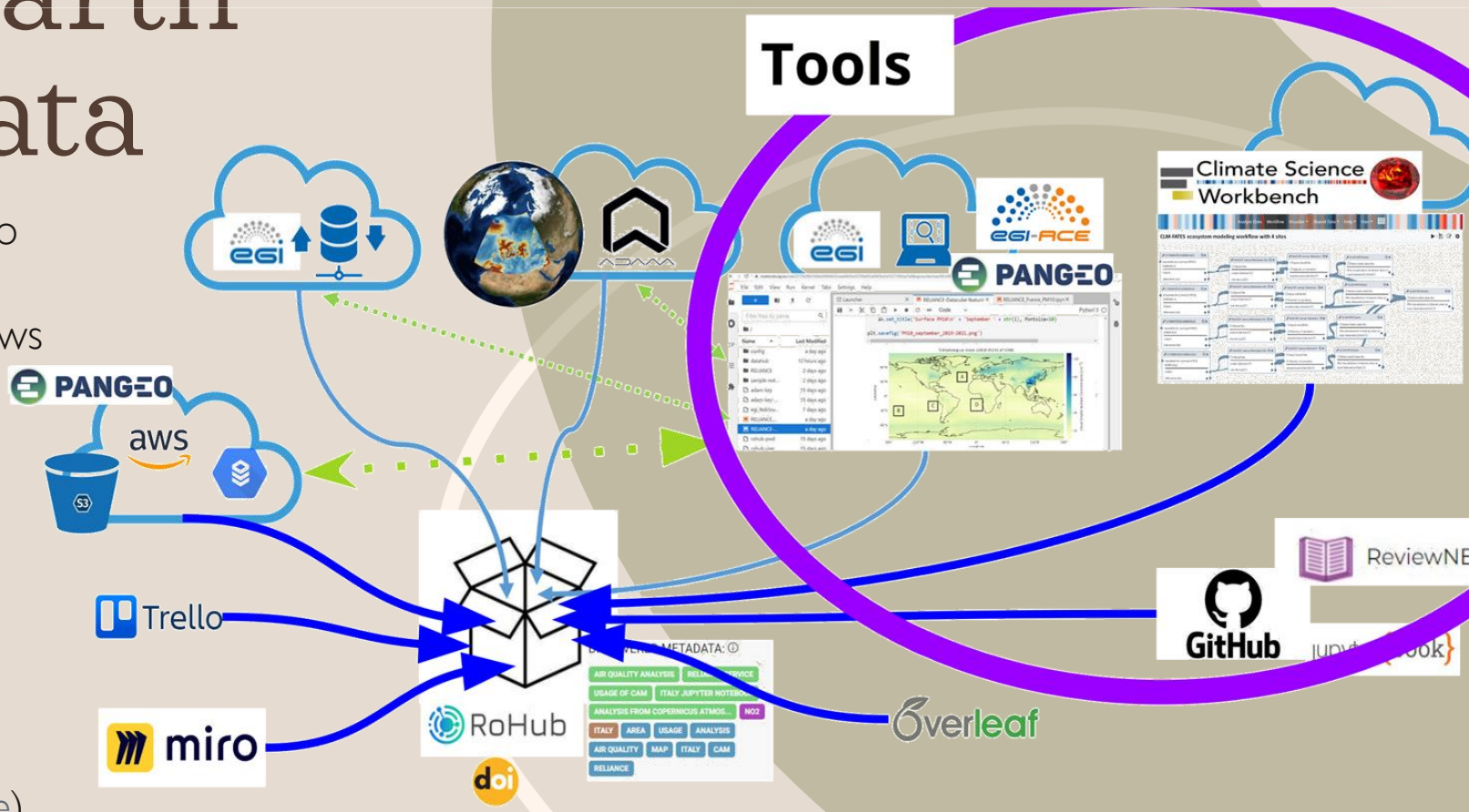
Data Cubes – earth observation data

The EOSC project RELIANCE use RO-Crate to package data cubes of earth observation data, along with documentation, images and workflows

Connects to related infrastructures for execution/analysis.

Metadata includes temporal coverage, spatial coverage and vertical coverage.

ROHub publishes the archived RO-Crates to general-purpose repositories (Zenodo, B2Share) for longevity and PIDs.

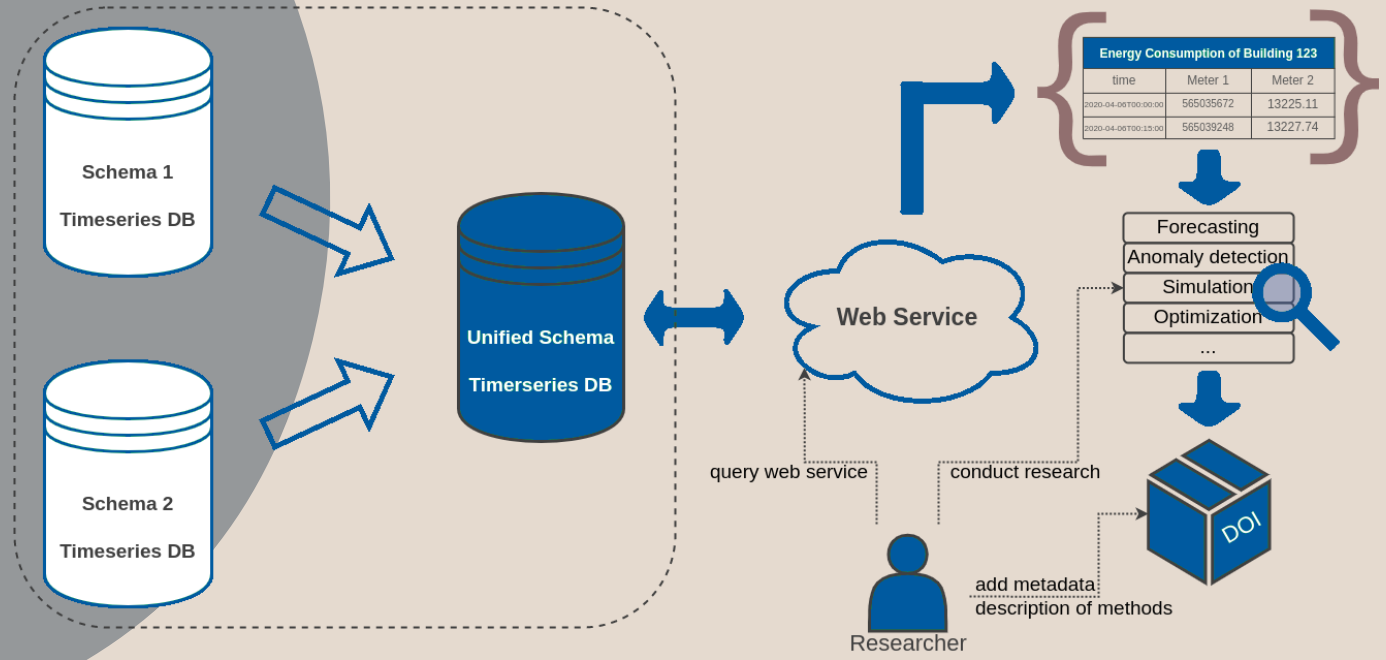


Reliance

RoHub

Fouilloux et al (2022):
International conference on FAIR Digital Objects
https://doi.org/10.24424/nz65-v565
https://doi.org/10.3897/rio.8.e93940

HMC Hub Energy:
FAIR Time Series of energy consumption measurements



Capturing & describing Time series:

- Electricity
- Gas
- Heat
- Drinking Water
- Compressed air

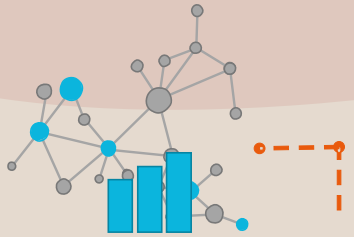
Günther et al (2022): Use Cases in HMC – from Generation to Reuse of Data
Helmholtz Metadata Collaboration
https://doi.org/10.5281/zenodo.7157694
https://helmholtz-metadaten.de/en/fair-data-commons/overview

Annotating plant research data

EXPERIMENTAL DATA

ANNOTATION

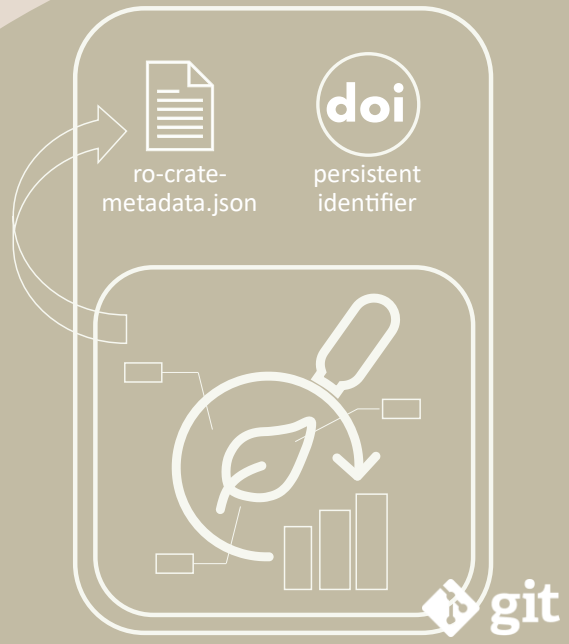
COMPUTATION



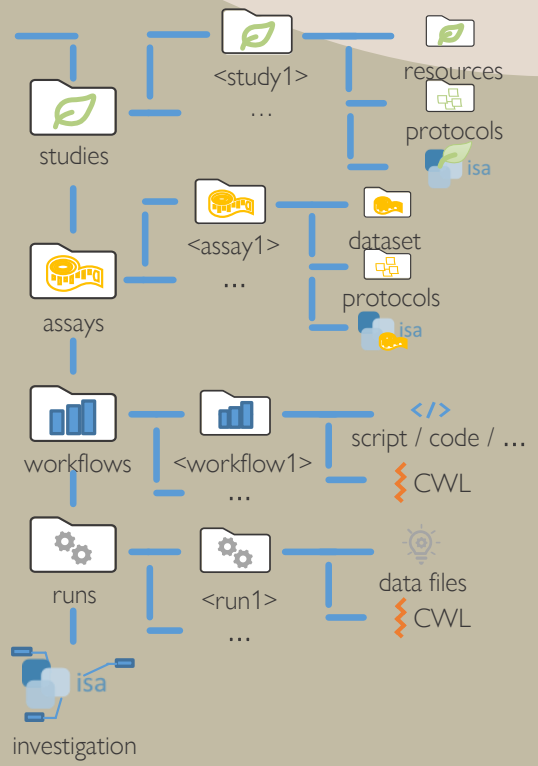
nfdi Nationale Forschungsdaten Infrastruktur

Data PLANT

https://nfdi4plants.de/



ANNOTATED RESEARCH CONTEXT (ARC)



RO metadata file is automatically generated, converting from ISA folder structure and annotations

Stored using Git LFS to support large data

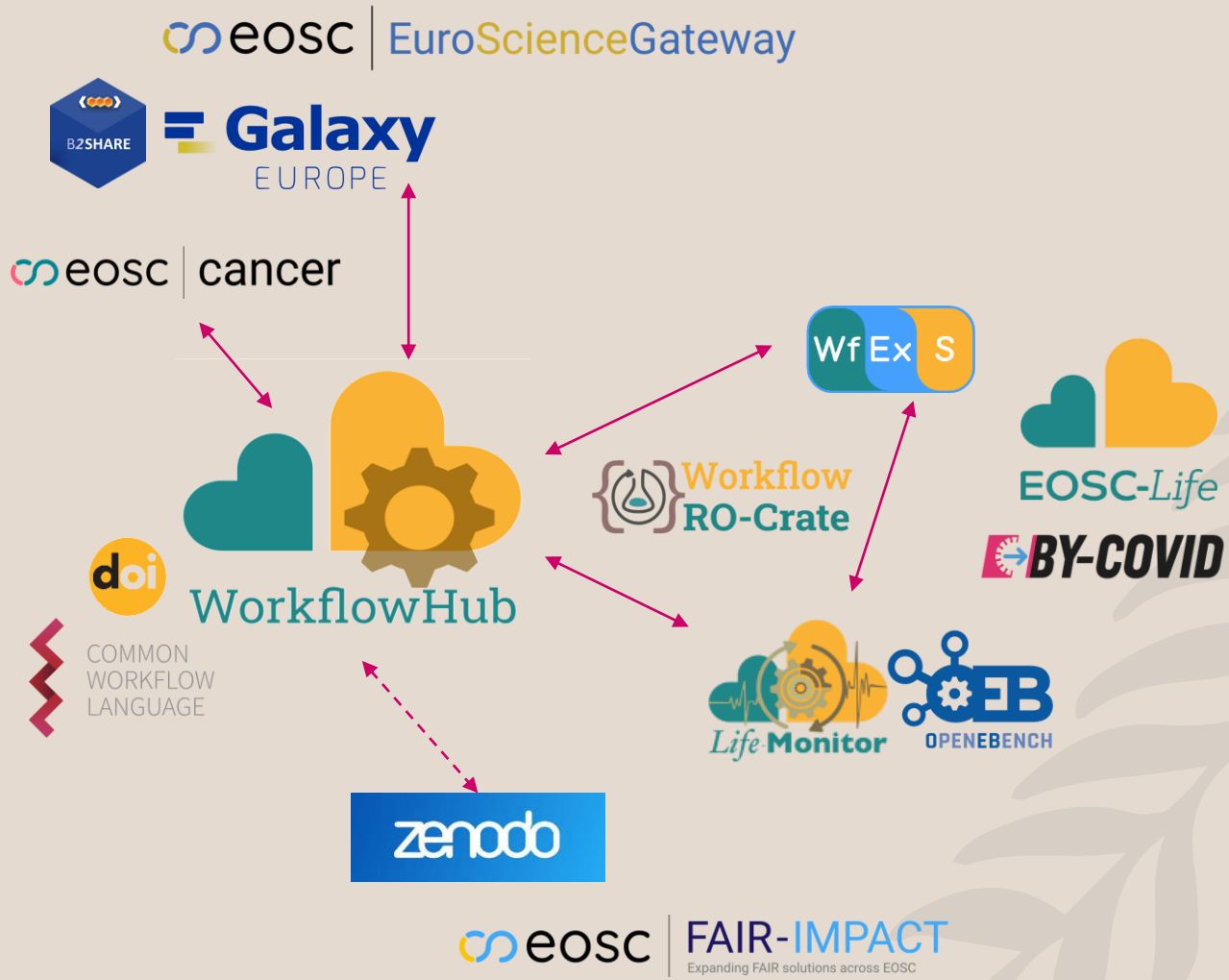
Investigation
Study
Assay
https://isa-tools.org/

Reuse established standards including ISA Model, CWL

https://nfdi4plants.de/content/learn-more/annotated-research-context.html

Building an EOSC ecosystem of FAIR Workflows

- EOSC projects BY-COVID, EOSC-Life, EuroScienceGateway exchange rich Workflow RO-Crate within an emerging EOSC ecosystem
- Workflow Crates capture
  - executable workflows in their native format (e.g. Galaxy)
  - interoperable CWL description of the workflow
  - software citations (e.g. tools used)
  - required data sources
  - test suites
  - workflow execution provenance



https://workflowhub.eu/
https://www.researchobject.org/workflow-run-crate/

https://www.researchobject.org/ro-crate/community.html

Join RO-Crate

The *RO-Crate* team is:

- Peter Sefton https://orcid.org/0000-0002-3545-944X
- Stian Soiland-Reyes https://orcid.org/0000-0001-9842-9718
- Eoghan Ó Carragáin https://orcid.org/0000-0001-8131-2150
- Oscar Corcho https://orcid.org/0000-0002-9260-0753
- Daniel Garijo https://orcid.org/0000-0003-0454-7145
- Raul Palma https://orcid.org/0000-0003-4289-4922
- Frederik Coppens https://orcid.org/0000-0001-6565-5145
- Carole Goble https://orcid.org/0000-0003-1219-2137
- José María Fernández https://orcid.org/0000-0002-4806-5140
- Kyle Chard https://orcid.org/0000-0002-7370-4805
- Jose Manuel Gomez-Perez https://orcid.org/0000-0002-5491-6431
- Michael R Crusoe https://orcid.org/0000-0002-2961-9670
- Ignacio Eguinoa https://orcid.org/0000-0002-6190-122X
- Nick Juty https://orcid.org/0000-0002-2036-8350
- Kristi Holmes https://orcid.org/0000-0001-8420-5254
- Jason A. Clark https://orcid.org/0000-0002-3588-6257
- Salvador Capella-Gutierrez https://orcid.org/0000-0002-0309-604X
- Alasdair J. G. Gray https://orcid.org/0000-0002-5711-4872
- Stuart Owen https://orcid.org/0000-0003-2130-0865
- Alan R Williams https://orcid.org/0000-0003-3156-2105
- Giacomo Tartari https://orcid.org/0000-0003-1130-2154
- Finn Bacall https://orcid.org/0000-0002-0048-3300
- Thomas Thelen https://orcid.org/0000-0002-1756-2128
- Hervé Ménager https://orcid.org/0000-0002-7552-1009
- Laura Rodríguez-Navas https://orcid.org/0000-0003-4929-1219
- Paul Walk https://orcid.org/0000-0003-1541-5631
- brandon whitehead https://orcid.org/0000-0002-0337-8610
- Mark Wilkinson https://orcid.org/0000-0001-6960-357X
- Paul Groth https://orcid.org/0000-0003-0183-6910
- Erich Bremer https://orcid.org/0000-0003-0223-1059
- LJ Garcia Castro https://orcid.org/0000-0003-3986-0510
- Karl Sebby https://orcid.org/0000-0001-6022-9825
- Alexander Kanitz https://orcid.org/0000-0002-3468-0652
- Ana Trisovic https://orcid.org/0000-0003-1991-0533
- Gavin Kennedy https://orcid.org/0000-0003-3910-0474
- Mark Graves https://orcid.org/0000-0003-3486-8193
- Jasper Koehorst https://orcid.org/0000-0001-8172-8981
- Simone Leo https://orcid.org/0000-0001-8271-5429
- Marc Portier https://orcid.org/0000-0002-9648-6484
- Paul Brack https://orcid.org/0000-0002-5432-2748
- Milan Ojsteršek https://orcid.org/0000-0003-1743-8300
- Bert Droesbeke https://orcid.org/0000-0003-0522-5674
- Chenxu Niu https://orcid.org/0000-0002-2142-1731
- Kosuke Tanabe https://orcid.org/0000-0002-9986-7223
- Tomasz Miksa https://orcid.org/0000-0002-4929-7875
- Marco La Rosa https://orcid.org/0000-0001-5383-6993
- Cedric Decruw https://orcid.org/0000-0001-6387-5988
- Andreas Czerniak https://orcid.org/0000-0003-3883-4169
- Jeremy Jay https://orcid.org/0000-0002-5761-7533
- Sergio Serra https://orcid.org/0000-0002-0792-8157
- Ronald Siebes https://orcid.org/0000-0001-8772-7904
- Shaun de Witt https://orcid.org/0000-0003-4196-3658
- Shady El Damaty https://orcid.org/0000-0002-2318-4477
- Douglas Lowe https://orcid.org/0000-0002-1248-3594
- Xuanqi Li https://orcid.org/0000-0003-1498-6205
- Sveinung Gundersen https://orcid.org/0000-0001-9888-7954
- Muhammad Radifar https://orcid.org/0000-0001-9156-9478
- Rudolf Wittner https://orcid.org/0000-0002-0003-2024
- Oliver Woolland https://orcid.org/0000-0002-4565-9760
- Paul De Geest https://orcid.org/0000-0002-8940-4946
- Douglas Fils https://orcid.org/0000-0002-2257-9127
- Florian Wetzels https://orcid.org/0000-0002-5526-7138
- Raül Sirvent https://orcid.org/0000-0003-0606-2512
- Abigail Miller https://orcid.org/0000-0001-9228-2882
- Jake Emerson https://orcid.org/0000-0003-0617-9219
- Davide Fucci https://orcid.org/0000-0002-0679-4361
- …you?

Engaging stakeholder communities in Skills4EOSC and FAIR-IMPACT

Loek Brinkman, DANS / TUDelft
Joy Davidson, DCC

Skills 4 eosc

eosc | FAIR-IMPACT
Expanding FAIR solutions across EOSC

H2020 EOSC ecosystem

EUROPEAN OPEN SCIENCE CLOUD

(5a)

EOSC secretariat.eu

Setup and management of the EOSC Secretariat supporting the EOSC Governance

Regional Nodes / Thematic Projects (5b)

NI4OS

EOSC SYNERGY

EOSC-Pillar

EOSC NORDIC

ExPaNDS

FAIR (5c)

Synchronisation Force

FAIRSFAIR

EGFC

European Group of FAIR Champions

EOSC Governance Board

EOSC Executive Board

EOSC Working Groups

- Landscape WG
- Sustainability WG
- Rules of Participation WG
- Architecture WG
- FAIR WG
- Skills and Training WG

FAIR WG Task Groups

- FAIR practice
- Interoperability
- PIDs
- Metrics and certification

INFRAEOSC-5 Cross Project Collaboration Board (CPCB)

INFRAEOSC-5 Task Forces:

- Landscaping
- FAIR data and Infrastructures
- Services onboarding
- National policies and governance
- Training and skills
- Dissemination and events

EOSC Interest Groups

- Researcher engagement and use cases
- Service and research product catalogue
- Federating core
- Glossary
- Cluster Collaboration

"Horizontals"

ARCHIVER

eInfraCentral

TRIPLE-A

RDA RESEARCH DATA ALLIANCE

OpenAIRE

OCRE Open Clouds for Research Environments

EOSC-hub

ESFRI Clusters

ENVRI FAIR

panosc

SSHOC

ESCAPE

EOSC-Life

Other FAIR Initiatives

FAIR4Health

STM Research Data

GO FAIR

FAIRsharing.org standards, databases, policies

FAIRplus

eosc | FAIR-IMPACT
Expanding FAIR solutions across EOSC

Call HORIZON-INFRA-2021-EOSC-01-05

Enabling discovery and interoperability of federated research objects across scientific communities

Expanding FAIR solutions in Europe

Partly following up on FAIRsFAIR

EU funded project

Coordination and Support Action

10 million euro

36 months, starting 1 June 2022

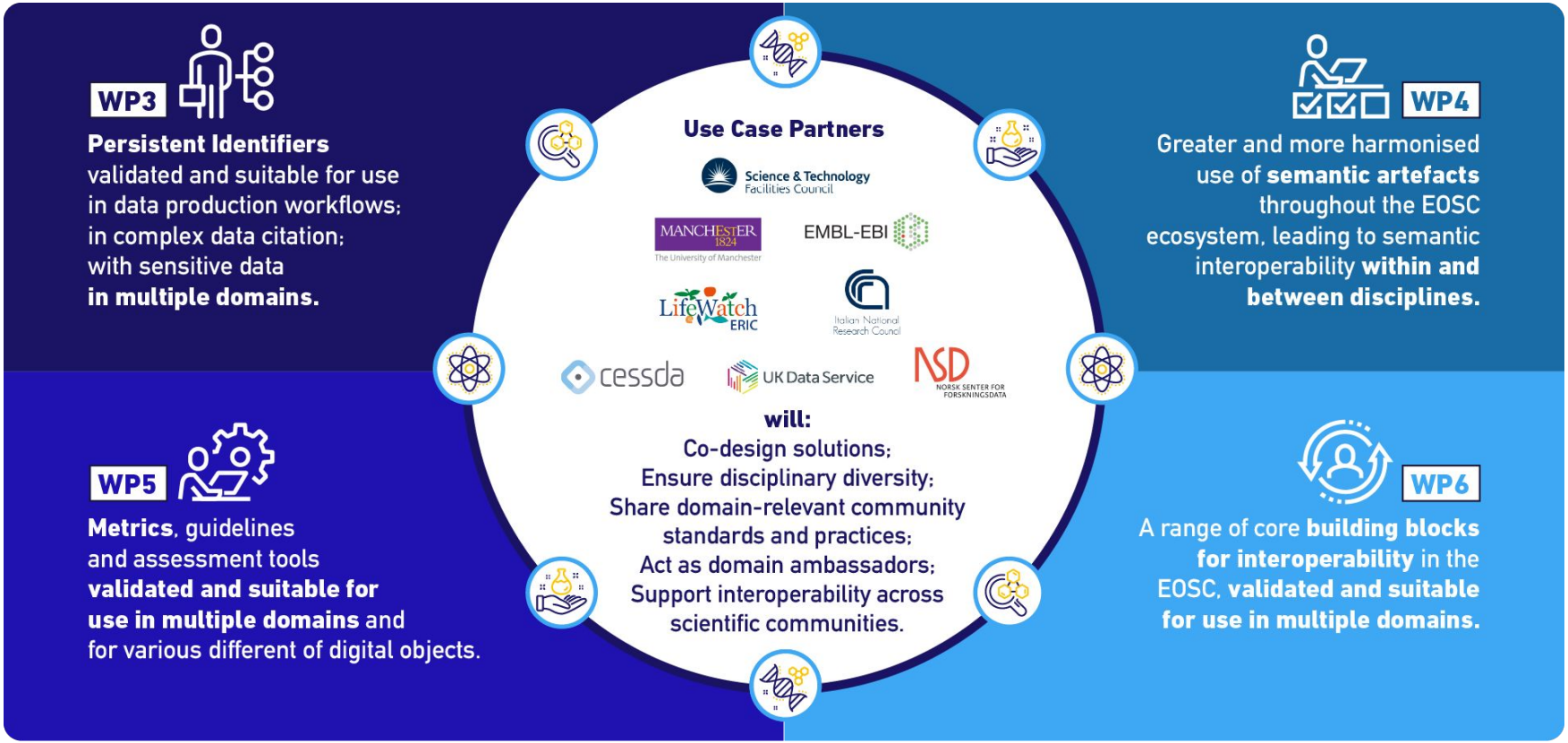
28 partners and affiliate entities

From 10 EU member states: NL, FI, FR, DK, IT, DE, ES, NO, BE, RO

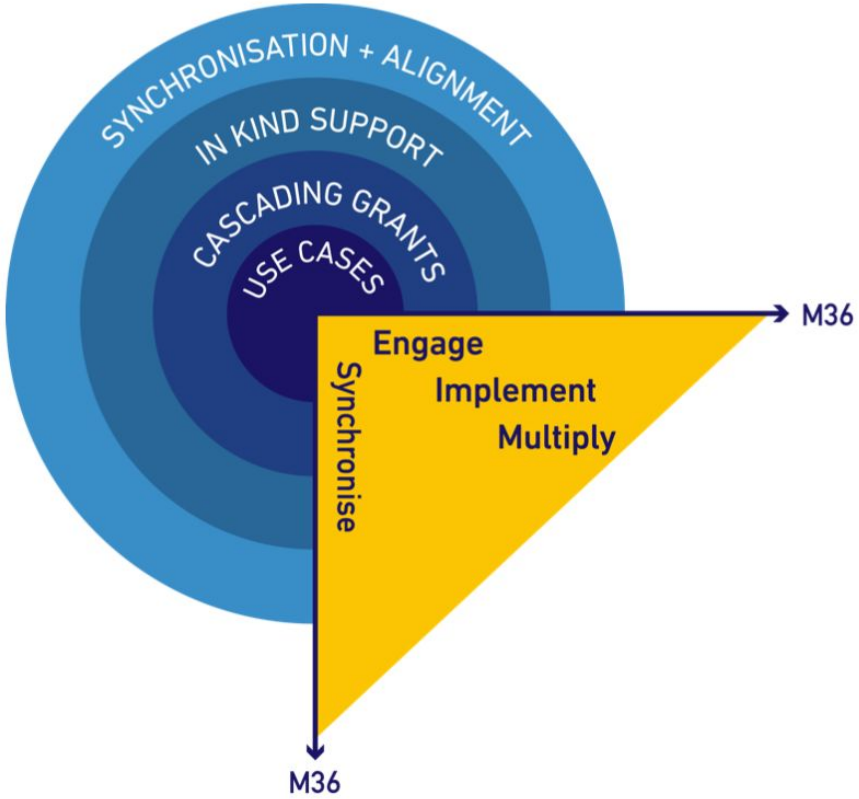
and the UK

Goal: Supporting the implementation of FAIR-enabling practices

Engaging stakeholder communities

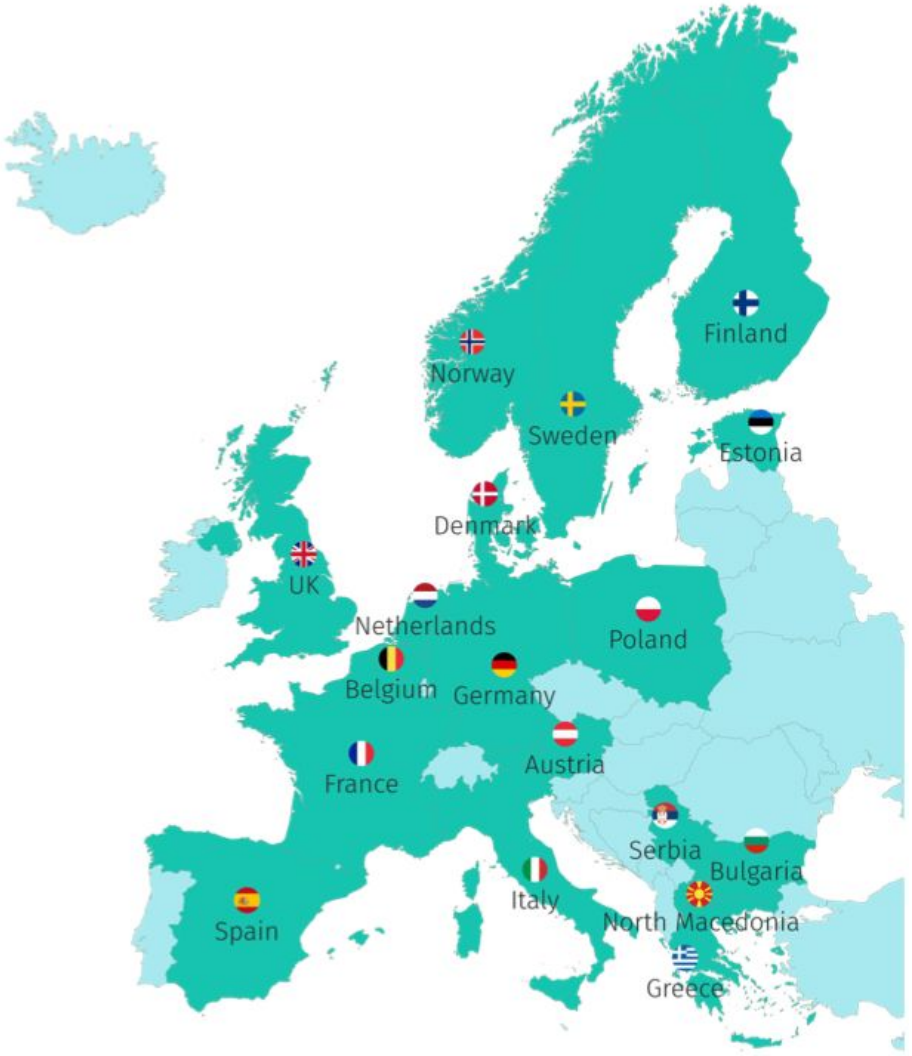


Engaging stakeholder communities



Skills 4 eosc

- 44 **Participants**, 18 **Countries**
- "**Key doers**" in Open Science in their Country/Region/Domain
- 2 ESFRI Research Infrastructures
- 7 millions €



Skills 4 eosc

Goal: Supporting an EOSC-ready digitally skilled workforce

Institutions

Professional networks and communities

Science for Policy

Engaging stakeholder communities

Network of Competence Centres

- Institutional
- Regional
- Thematic

Skills 4 eosc

Peer networks: OSC Incubator Program

Next run: March 2023

Register at www.osc-international.com

INOSC
International Network of
Open Science & Scholarship Communities

Ways you can engage with us

- Apply to take part in one of the support programmes
  - FAIR-IMPACT open calls (March 2023)
  - Skill4EOSC Open Science Incubator Programme (March 2023)
- Review and comment on our draft deliverables
  - Skills4EOSC Zenodo Community
  - FAIR-IMPACT Zenodo Community
- Sign up to the project newsletters to be kept informed about our outputs and open calls

Related EOSC Symposium sessions

Talkers, Thinkers, Doers: Stakeholder Engagement and EOSC synergies through Skills4EOSC

Towards a shared value proposition for Persistent Identifiers in EOSC

Semantic Interoperability in EOSC

EOSC PID Policy and FAIRCORE4EOSC: Measuring compliance

FAIR enabling practices

- Engaging stakeholder communities in Skills4EOSC and FAIR-IMPACT

Training & Skills for EOSC: Lightning talks

- The Skills4EOSC project contribution
- The Open Science Communities Incubator Program

Thank you!

@Skills4Eosc

Skills 4 eosc

https://www.fair-impact.eu

@fairimpact_eu

/company/fair-impact-eu-project

eosc | FAIR-IMPACT

Expanding FAIR solutions across EOSC

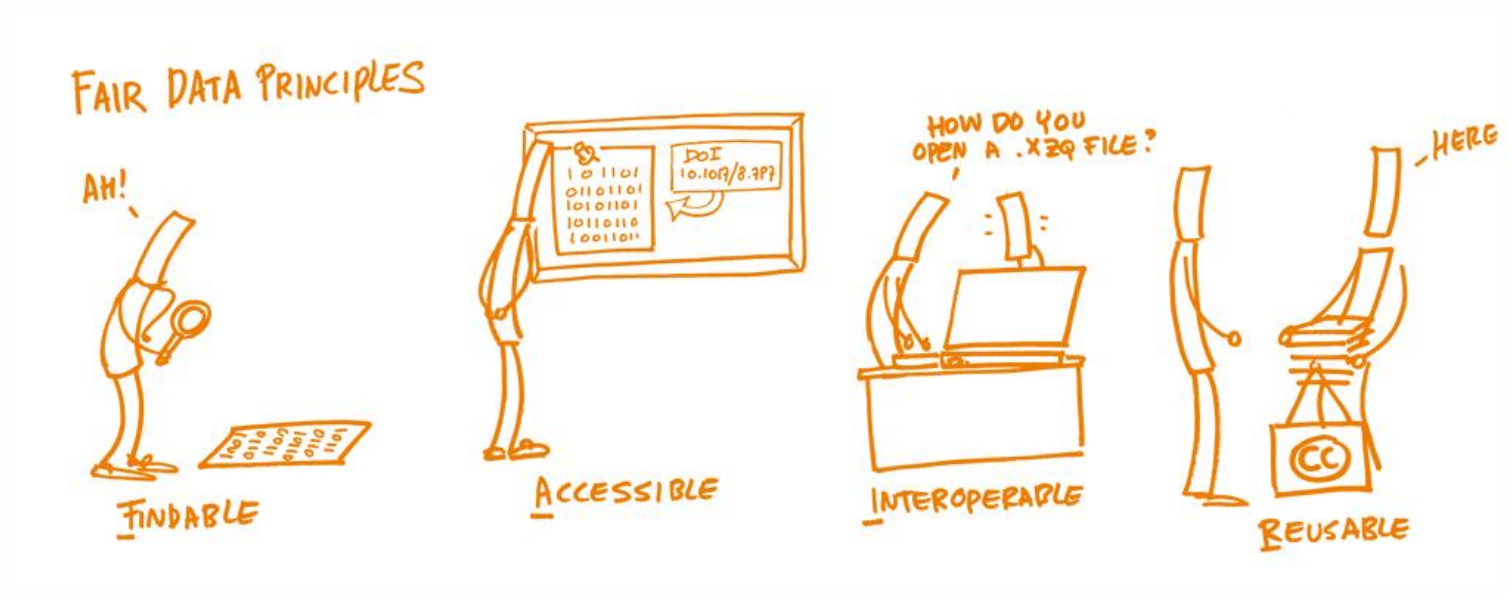
Tackling Research Data Management challenges with FAIRness

Giulia Caldoni giulia.caldoni2@unibo.it
Bianca Gualandi bianca.gualandi4@unibo.it
Mario Marino **mario.marino6@unibo.it**

Research Services Coordination Unit Research Services Division (ARIC)
University of Bologna

RDM Challanges with FAIRness

- Data need to be Findable, Accessible, Interoperable and Reusable (FAIR): need acknowledged by RDM communities
- FAIR principles can guide and be beneficial to researchers: it is not trivial to make them understand how
- RDM take place in a context of specific and restrictive ethical and legal requirements
- FAIR principles and DMP are mandatory in many funded programmes, including Horizon Europe



(2018). Open Science Training Handbook (1.0) [Computer software]. Zenodo. https://doi.org/10.5281/zenodo.1212496 (CC0)

Researchers should be aware of how all these topics intersect the different phases of data lifecycle, and support for researchers should be tailored on this need for a global and integrated overview

Support @ FAIRness

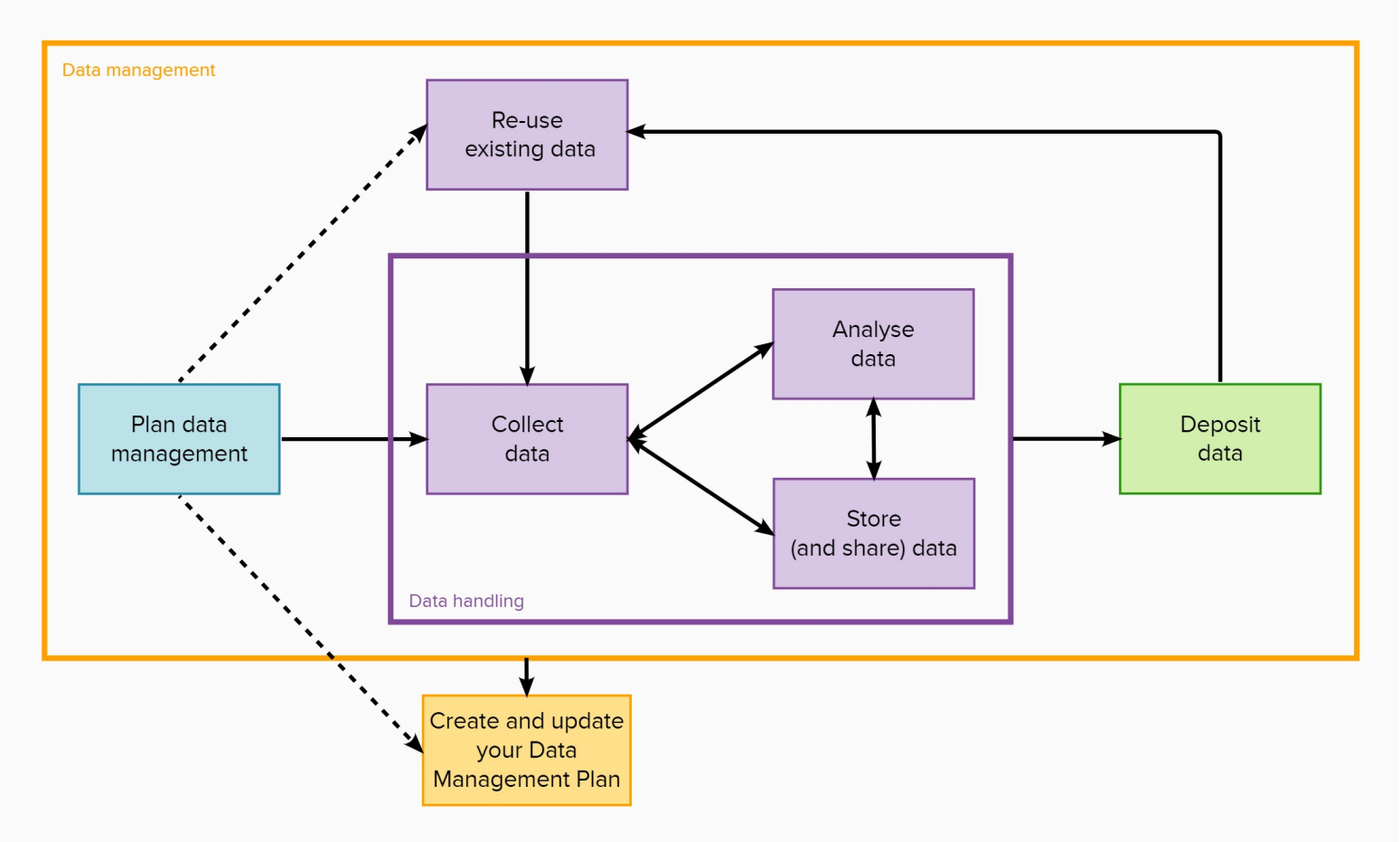
In our experience across different research domains, researchers often face the same stumbling blocks.

To support them, we have developed a **RDM Decisional Tree** starting from the fundamental bricks of the data lifecycle and posing a series of questions to help researchers navigate:

1) the domain specific nature and origin of the data they are handling;
2) Privacy/Ethics requirements (e.g. GDPR);
3) Intellectual Property Rights;
4) active data storage;
5) long-term deposit and preservation.

With all these questions in mind, we support researchers in finding answers by supplying a **Data Lifecycle Model**, including a series of possible actions and resources to protect/organize/preserve their data.

Data Life-cycle



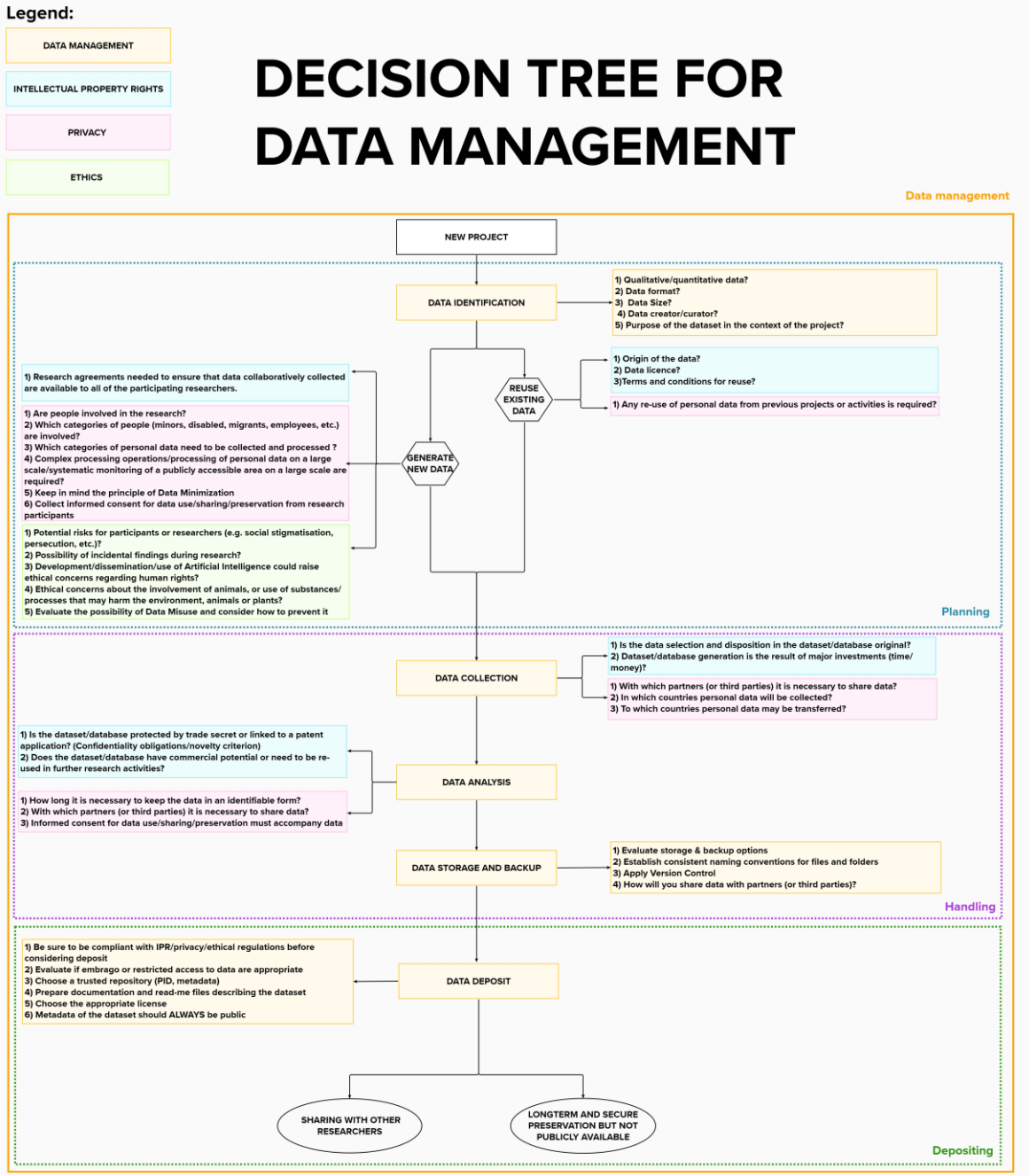
Source: https://zenodo.org/record/7249051#.Y1pAi3ZBxPY

This diagram proposes a data lifecycle model inspired by the University of Virginia Library's model (https://guides.lib.virginia.edu/c.php?g=515290&p=3522215)

Research Data Management Decision Tree

Source: https://zenodo.org/record/7190005#.Y1ptGXZBxPY

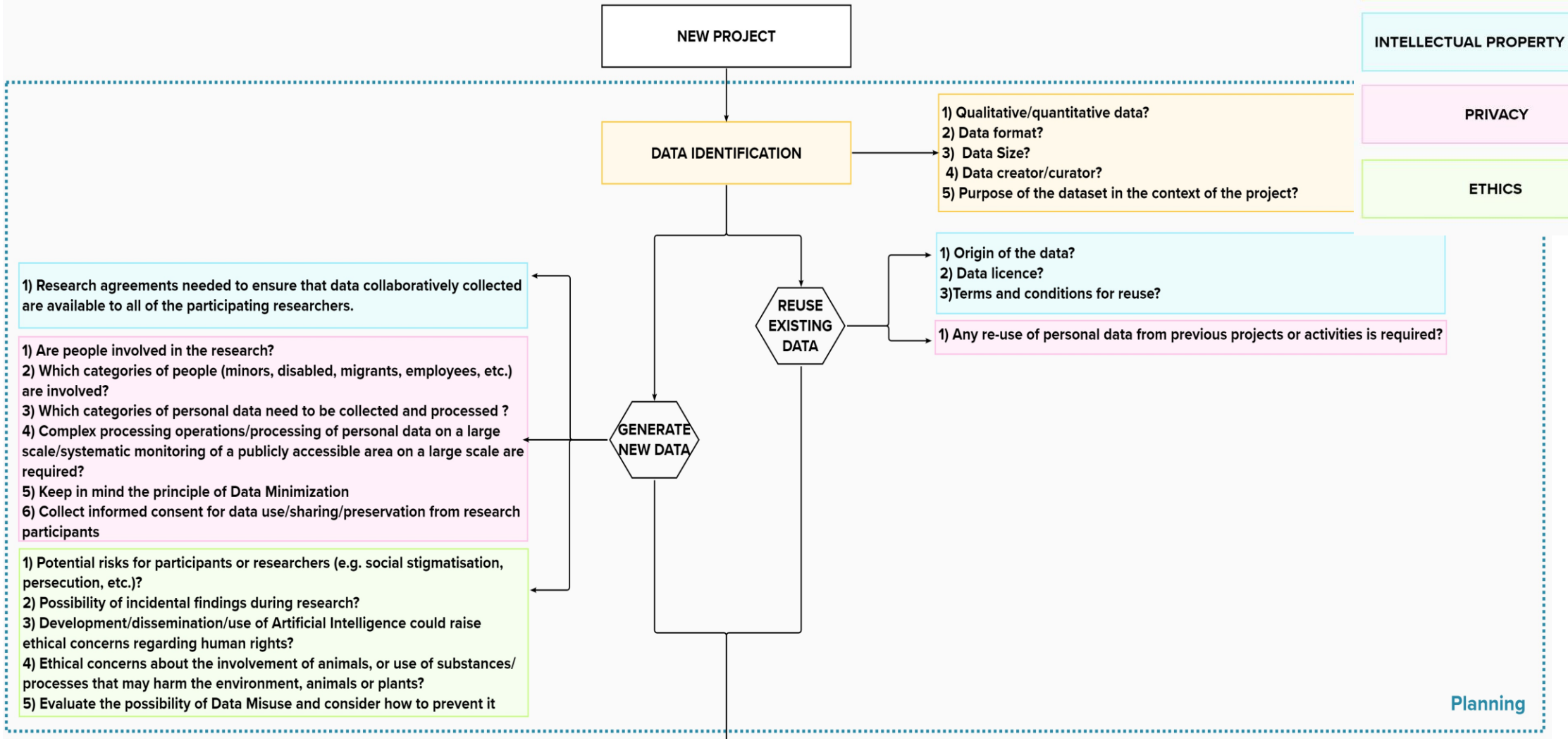
Decisional tree to help researchers ask, and answer, questions they did not know they had. Starting from the fundamental bricks of the data lifecycle, integrates a series of questions aiming to encourage the researchers address some major attention points, encompassing issues that range from privacy/ethics requirements to IPR legislation and FAIR principles.



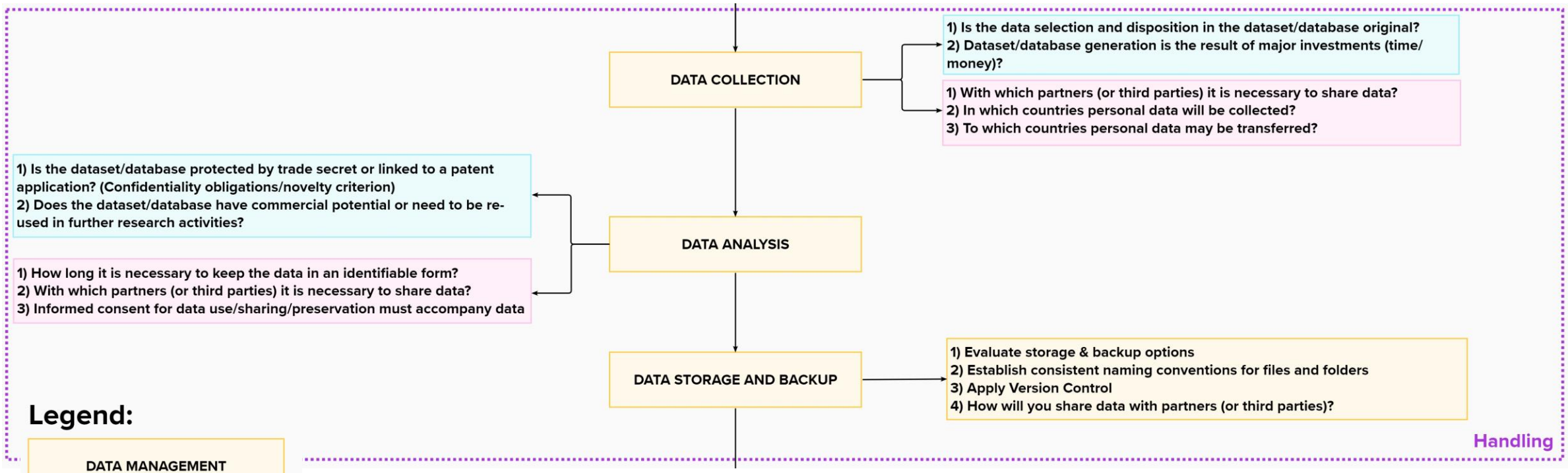
EOSC Symposium, November 2022, Prague

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

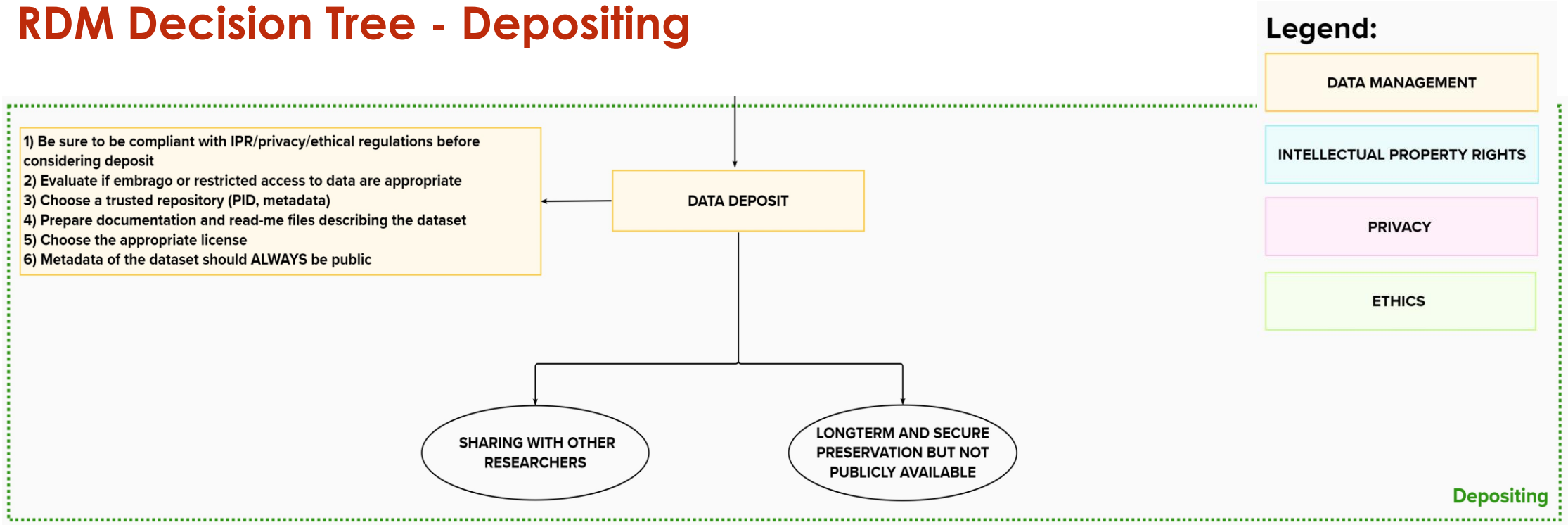
RDM Decision Tree - Planning



RDM Decision Tree - Handling



RDM Decision Tree - Depositing



Take Home Messages:

- RDM Decision Tree is useful also as a self-standing tool to support researchers. It improves researchers' awareness on the importance and value of research data
- It helps the researcher to think about both their inbound (IPR, privacy, ethics) and outbound data, and to make upstream decisions with a view to long-term preservation.
- It highlights that FAIR RDM implies a plurality of supporting expertise and a multitude of services, resources, tools
- RDM DT is a tool that can be used in any context, it might also need to be integrated depending on the situation (e.g. funded project, different funders requirements, etc.)
- We are collecting feedback from researchers to identify further critical points and integrate the Tree
- The feedback gathered so far from researchers has shown considerable appreciation for the tool provided

Thank you!

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA