

## *ELIXIR Training Programme in Management*

### **Module 2 - RDM strategy**

Best practices from the ELIXIR portfolio of data management tools and services. A quick guide to adopting ELIXIR resources and optimising an ELIXIR Node's data management strategy

Facilitators: Mijke Jetten & Celia van Gelder  
Finland, 27-28-29 June, 2023



ELIXIR-CONVERGE has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 871075.

# Time schedule

Day 2	Time (CEST/EEST)	Module 2: Best practices from the ELIXIR portfolio of data management tools and services. A quick guide to adopting ELIXIR resources and optimising an ELIXIR Node's data management strategy <i>Contacts: Celia van Gelder &amp; Mijke Jetten</i>
June 28	9:00/10:00	<a href="#">Welcome &amp; introduction</a>
June 28	9:30/10:30	<a href="#">Data management: examples of ELIXIR tools and services from a Node strategy perspective - part 1</a>
June 28	11:00/12:00	Lunch
June 28	12:00/13:00	<a href="#">Data management: examples of ELIXIR tools and services from a Node strategy perspective - part 2</a>
June 28	13:00/14:00	<a href="#">Funder requirements to FAIR data</a>
June 28	14:00/15:00	Break
June 28	14:30/15:30	<a href="#">Data stewardship roles and competences</a>
June 28	15:30/16:30	<a href="#">Learn from &amp; share with</a>
June 28	16:00/17:00	Finland: dinner & social event

Day 3	Time (CEST/EEST)	Module 2: Best practices from the ELIXIR portfolio of data management tools and services. A quick guide to adopting ELIXIR resources and optimising an ELIXIR Node's data management strategy <i>Contacts: Celia van Gelder &amp; Mijke Jetten</i>
June 29	9:00/10:00	<a href="#">Maturity/services model, Node strategy and the ideal Node team</a>
June 29	10:30/11:30	Break
June 29	11:00/12:00	<a href="#">Maturity/services model, Node strategy and the ideal Node team</a>
June 29	12:00/13:00	Finland: feedback lunch



# Day 2/28 June

## Module 2 - RDM strategy



# Welcome & introduction

(9-9.30 CEST / 10-10.30 EEST)



# Zoom housekeeping



We invite you to contribute actively, but to mute yourself if not speaking



Welcome to use the chat for further comments or discussions, however speaking up is preferred



Please use the “hand-raising function” to indicate you would like to contribute directly



Speak up if you feel unheard or unseen, and share suggestions to engage more from the F2F room



As a hybrid training with a big group is challenging, bare with us if we deviate from the time schedule

# ELIXIR Code of Conduct principles

## ELIXIR Code of Conduct principles



### We value

each other's perspectives

### We adopt

a zero-tolerance approach to harassment and discrimination

### We maintain

high ethical standards

### We'll apply

honesty and integrity in the dealing of any transgressions against the Code.

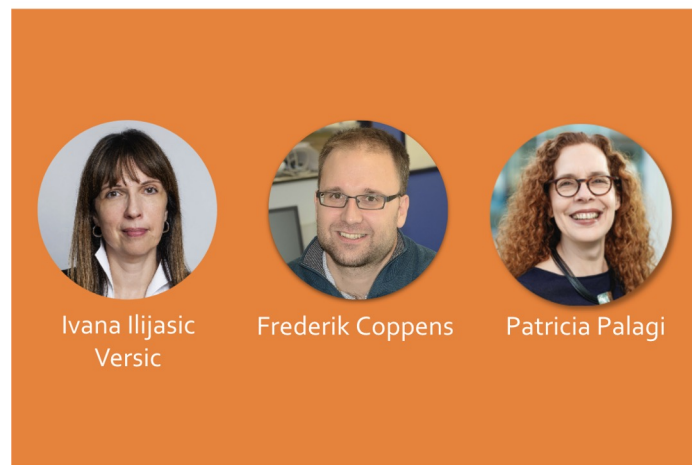
### We're committed

to making ELIXIR events collaborative, supportive and enjoyable

### We'll ensure

a respectful and inclusive environment

## ELIXIR Code of Conduct oversight group



Ivana Ilijasic  
Versic

Frederik Coppens

Patricia Palagi

Reach out to the Oversight Group to enforce the Code of Conduct

[coc.oversight@elixir-europe.org](mailto:coc.oversight@elixir-europe.org)



# ELITMa - ELIXIR training program in Management

- History
  - ✓ Building on survey Training Platform ([link](#))
  - ✓ ELITMa Curriculum designed in ELIXIR-CONVERGE WP6 and WP2 collaboration ([link to Task 6.4 report](#))
  - ✓ CONVERGE WP2 Uplift allowed to build the material and deliver pilot edition of two (tbc) Modules
- Benefit for ELIXIR
  - ✓ Node staff trained for more effective operations, including the ability to navigate ELIXIR complexities. Strengthening of ELIXIR people network.
- Lead (Ana, PT)
  - ✓ Committed Nodes: NL, NO, FI, UK, and Hub





# ELITMa Modules' description



## ELIXIR Governance and Organizational Behaviour

Short element on European Research Infrastructures (RIs) focusing on ELIXIR, from mission to operations, after which trainees should be able to navigate ELIXIR features and entwine these with national nodes'.

## Data Management

Strengthening DM in Nodes by using best practices from the ELIXIR portfolio of data management tools and services.

## Communication

Educate on how to communicate properly inside the organization.

## Financial Management

Theoretical tools and practical techniques used in financial management of ELIXIR Nodes; role of internal control and risk management in Nodes.

## Legal and ethics issues

Tools for a correct approach to stewardship of common legal issues within an RI in general, and to understand ELIXIR's legal framework in particular.

## Leadership

Support ELIXIR Node-, WP- and task leaders in developing their leadership- and team management skills.

## Strategic Management

Frameworks and tools for designing and implementing a successful strategy for a national node integrated in a European research infrastructure (includes impact evaluation).

## Project Management

Theoretical tools and practical techniques used in financial management of an ELIXIR Node, and role of internal control and risk management in the organisation.





# Building your Node DM portfolio



- Focus of the module: **Node data management strategy**
  - ✓ We will define together what we precisely mean along the way
  - ✓ Potentially: your Node's strategy on data management **services**?
  - ✓ Roles Data Managers & Node Coordinators
- Thank you for the preparation work. We will jointly further populate your Node DM portfolio
  - ✓ It is a working session and we want to define practical next steps for all involved
  - ✓ Share with & learn from other Nodes
- Please build your [Node's DM portfolio](#) in the shared notes document, as they are for your use
  - ✓ Formulate your own lessons learned
  - ✓ Formulate your take home messages & next steps
- It is the **start of a journey**, that will continue after today
  - ✓ Get inspired by what your peers in other Nodes have written down



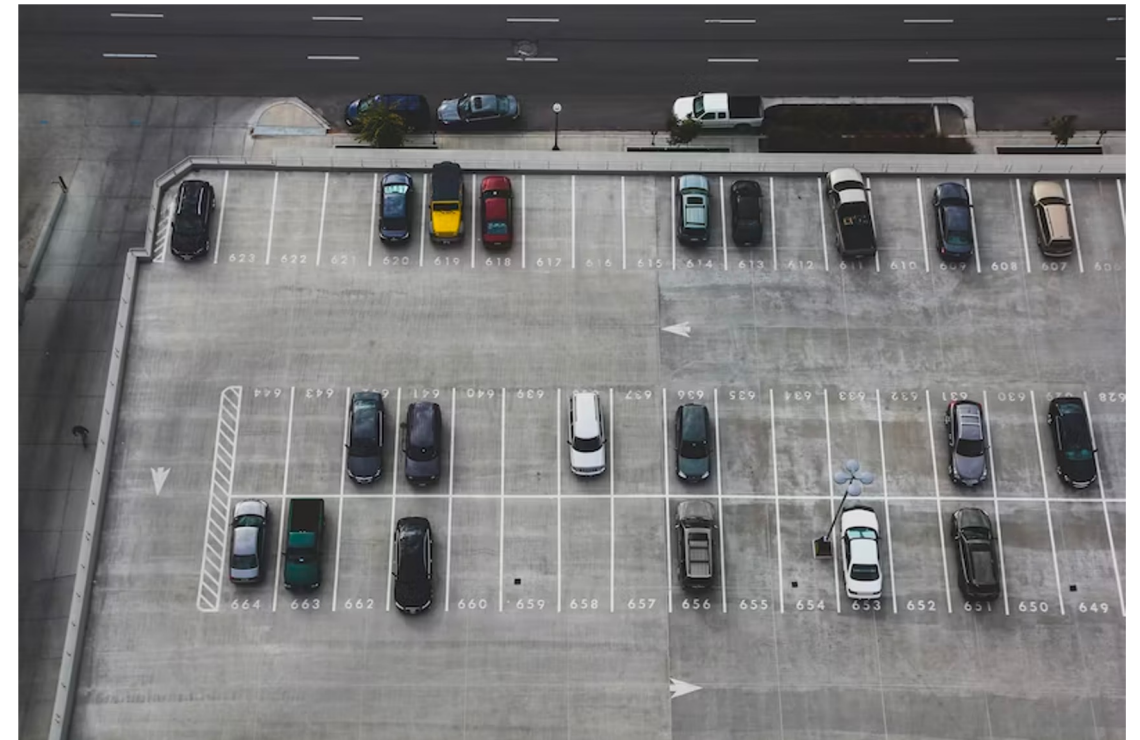
# How we will work

- You are the **first cohort** of the module in this form. Please share your **experiences** so that we can improve
- We have a challenging mix of **online** and **onsite** participants, so bear with us; **speak up** if you don't feel heard
- Participants come from **different backgrounds** and have **different levels of expertise**
- Focus is **knowledge exchange**, i.e., sharing and disseminating of knowledge and providing inputs to problem solving, so we are **all key to the success** of this workshop
- We try to make the module as **interactive** as possible. We invite you to participate actively, and relate the information to your daily Node practice



# Parking Lot

- We have only 2 days to start reflecting and discussing
- We will not have time to go into detail on all interesting topics
- Let us collect together topics that we should not forget and/or that you miss in the current programme
- We will try to come back to them as much as possible during the module and will also make sure that these topics are not forgotten going forward
- Please add your topics in the **Parking Lot section of the Shared Notes!**



# Introduction round

- Visit the [Shared notes document](#)
- Add your Node, name, role, F2F/online, email address
- And a **fun fact** about yourself



# ELIXIR Hub

- Supports executive management and administrative staff.
- Develops and delivers ELIXIR scientific strategy and manages Commissioned Services by Nodes.
- Coordinates services from ELIXIR Nodes.
- Coordinates and supports ELIXIR governance bodies and technical committees.
- Leads ELIXIR communications and external relations, including industry support.
- Supports institutions within the Nodes.
- Collaborates with national and European funders and policy-makers.



**Fabio Liberante**  
Programme Manager  
Data & Interoperability



# ELIXIR Hub - Technical Team

- Oversee operation/development of ELIXIR Services at European level
- Interface between Nodes and International activities
- Technical coordination of Nodes
- Co-located with EMBL-EBI
- Do *not* run ELIXIR Services directly



**Peter Maccallum**  
Chief Technical Officer



**Fabio Liberante**  
Programme Manager  
Data & Interoperability



**Jonathan Tedds**  
Programme Manager  
Compute & Tools



**Katharina Heil**  
Programme Manager  
Communities & Training



**Manthos Pitoulis**  
Senior Science Officer  
Tools & Training



**David Lloyd**  
Technical Support Officer



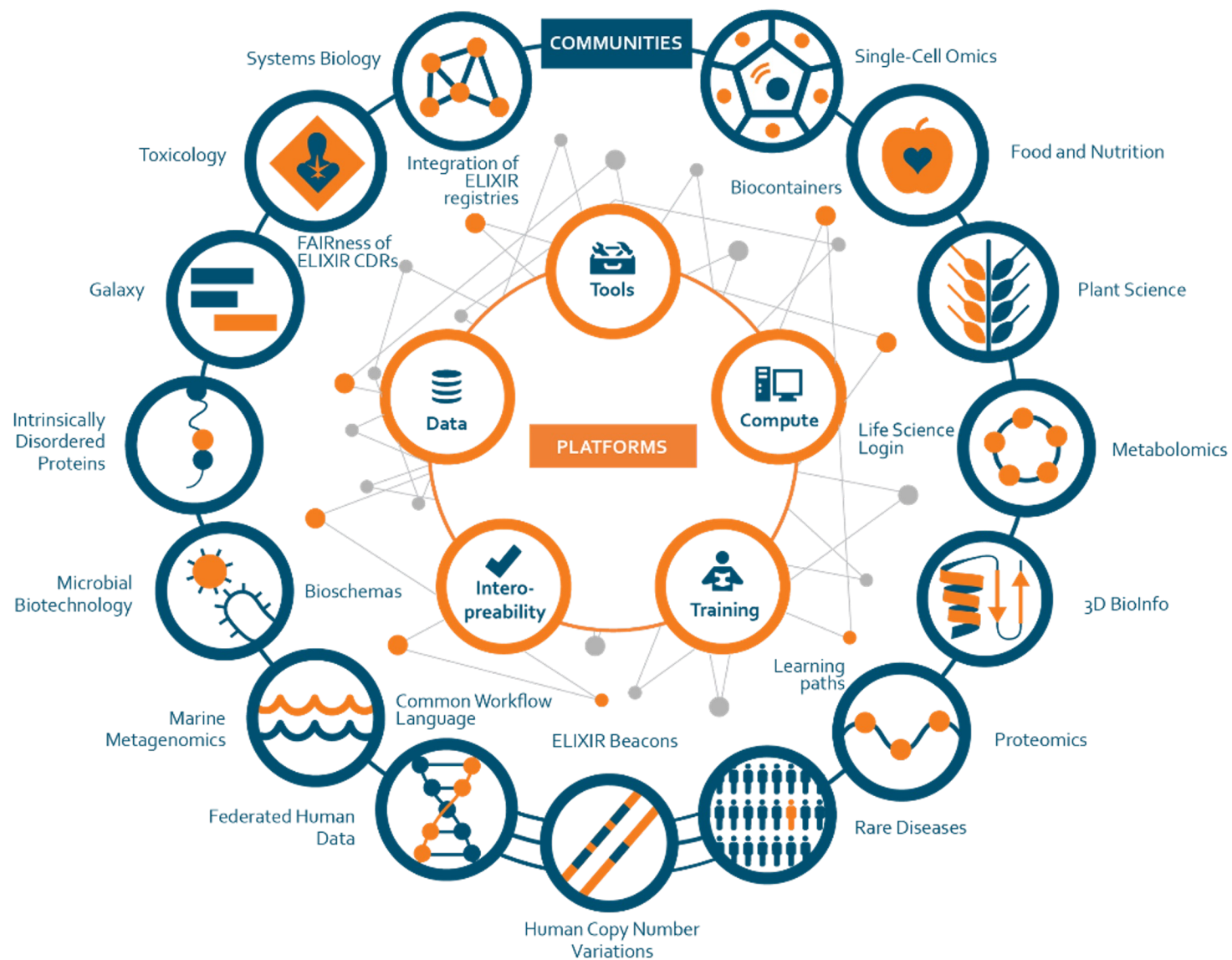
**Gavin Farrell**  
Technical Officer  
Compute & Data



**Clare Garrard**  
Science Officer  
Communities & Interoperability



# Platforms & Communities



# Data Platform Overview

## Mission

01

Sustainably Funded  
Core Data Resources

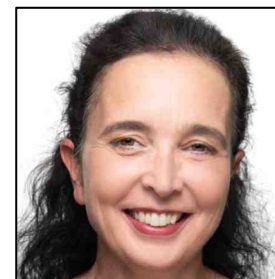
02

Coordinated Ecosystem of  
Node Data Resources

03

Support European  
Research Community

## Leadership



**Ulrike Wittig**  
ELIXIR Germany  
Platform Lead



**Patrick Ruch**  
ELIXIR Switzerland  
Platform Lead



**Silvio Tosatto**  
ELIXIR Italy  
Platform Lead



**Fabio Liberante**  
ELIXIR Hub  
Programme Manager



**Gavin Farrell**  
ELIXIR Hub  
Platform Officer





# Interoperability Platform Overview

## Mission

01

Help people and machines to discover, access, integrate and analyse biological data

02

Encourage adoption of standardised file formats, metadata, vocabularies and identifiers

## Leadership



**Susanna-Assunta Sansone**  
ELIXIR UK  
Platform Lead



**Chris Evelo**  
ELIXIR Netherlands  
Platform Lead



**Tony Burdett**  
EMBL-EBI  
Platform Lead

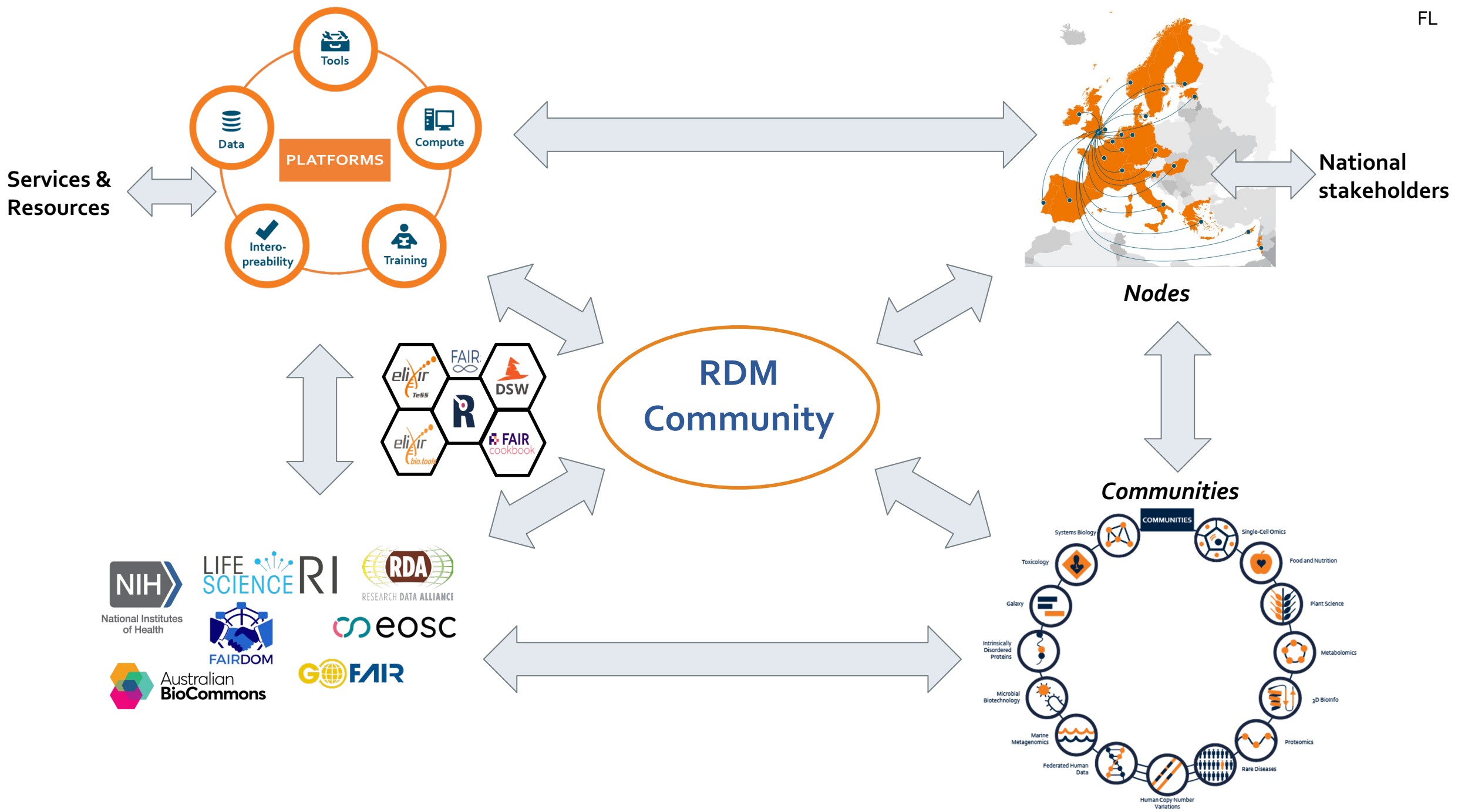


**Fabio Liberante**  
ELIXIR Hub  
Programme Manager



**Clare Garrard**  
ELIXIR Hub  
Science Officer  
Communities &  
Interoperability





# ELIXIR Services - some definitions

An **ELIXIR Service** is a set of technology, software, expertise and processes which together provide a defined capability to deliver data, analysis or a supporting function for ELIXIR's objectives.

An ELIXIR Service can be one of the following service types:

- **ELIXIR Node Service** - identified as part of an ELIXIR Node's contribution to ELIXIR and is listed on a Node's Service Delivery Plan.
- **ELIXIR Multi-Node Service** - provided by two or more ELIXIR Nodes working together on one or more of the activities required to deliver the Service. The ELIXIR Hub does not count as a Node in this definition.
- **ELIXIR Commissioned Service** - a type of *contract* - for Service Delivery or Development.
- **ELIXIR Infrastructure Service** - ELIXIR Node or Multi-Node Service - identified as a critical dependency or core capability for ELIXIR such that its continued operation (or, where not possible, its replacement) is a strategic, funded, priority.



# Multi-Node Services and Federation

We are increasingly seeing ELIXIR Services delivered as Multi-Node Services, ranging from pairs of Nodes working together to fully federated Europe-wide networks

Some **Benefits** of being identified as a Multi-Node Service:

- Public statement of multi-source support
- Clarity on leadership & contributions to help funding decisions
- Support for candidate Infrastructure Service status; Some costs (operations) eligible for Inf. Service funding
- Some costs (development) eligible for Implementation Study funding
- Status with National funders, attractive to EU funders

Some **Responsibilities**:

- Good governance
- A transparent process between partners
- Quality
- Oversight/interest by Technical Coordinators' Group



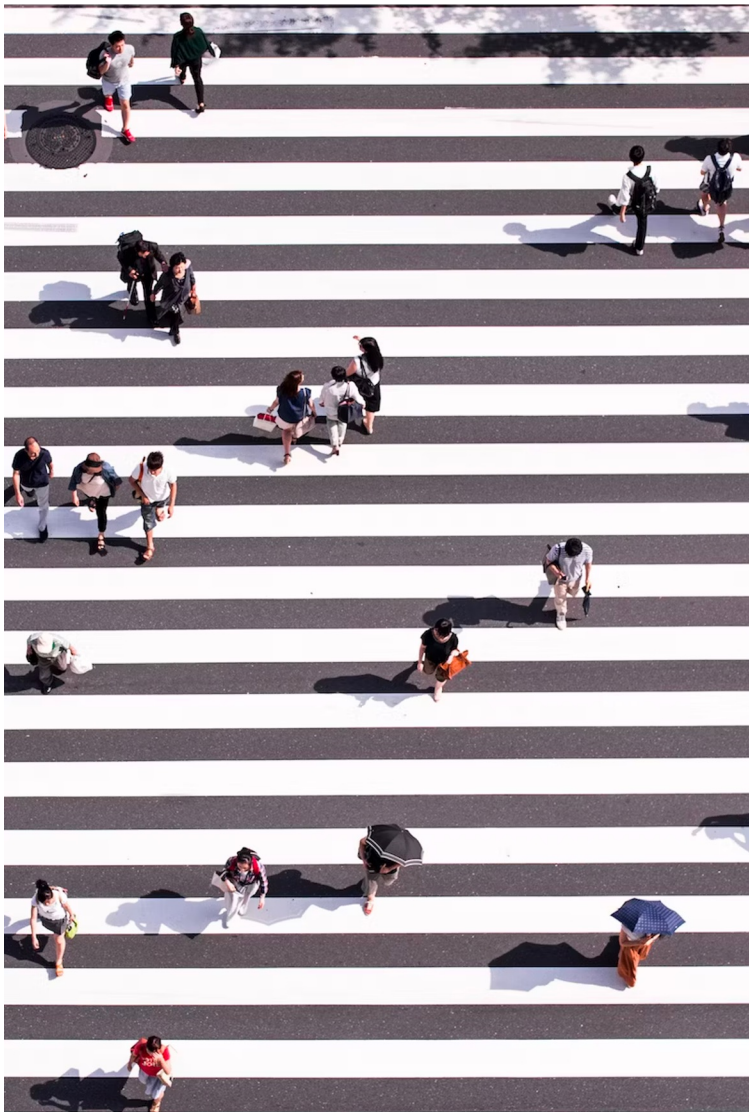
# Data management: examples of ELIXIR tools and services from a Node strategy perspective - **Part 1**

(9.30-11 CEST / 10.30-12 EEST)





# Nodes' good practices



- Presentations [10 min]
  - ✓ [RDMkit](#):
    - [LU \[Pinar Alper\]](#)
    - [NO \[Korbinian Bösl\]](#)
  - ✓ [DSW](#):
    - [CZ \[Anna Strachotova\]](#)
    - [NO \[Korbinian Bösl\]](#)
  - ✓ [Federated EGA \(CDR/EDD\)](#):
    - [FI \[Riku Riski\]](#)
    - [NO \[Korbinian Bösl\]](#)
- Discussion
  - ✓ What services are relevant to your Node?
  - ✓ What/who do you need (from others) to implement it?

\* CDR = ELIXIR Core Data Resource (CDR); EDD = ELIXIR Deposition Database



# RDMkit, ELIXIR-LU, Pinar Alper





# RDMkit as a Node Service at ELIXIR LU [LU, Pinar]

## Why this service

- RDM has become prominent in our Node agenda
  - Research Data Management and Stewardship
  - FAIRification
- LU is represented in the FCB and RDMkit Editorial Boards
- LU has the FCB in its service portfolio.
- Luxembourg National Data Service has joined the Node.
  - RDM Support at a national level
  - Coordination of a national RDM Community.



# RDMkit as a Node Service at ELIXIR LU [LU, Pinar]

## What are the challenges

- RDMkit is helpful as a consolidating overview for RDM and RDM resources. But,
  - There are national guidelines and grey literature, which does not fit in RDMkit.
  - We're kickstarting a national RDM Community and need a place to capture the knowledge of this community.
- Current RDMkit Node services focus on the central RDMkit (UK, BE)



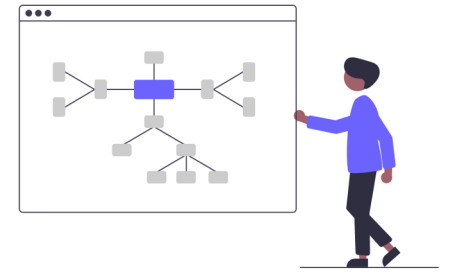
## And RDMkit-based Node Service at ELIXIR LU [LU, Pinar]

### What could work well with RDMkit

- RDMkit is not just a resource, it offers the means to spawn other resources:
  - Site theme.
  - Approach: open, documented, reusable processes.
  - Content.
- RDMkit Alliance is in the works
  - We know we will not be alone.



# RDMkit as a Node Service at ELIXIR LU [LU, Pinar]



## How will we implement it?

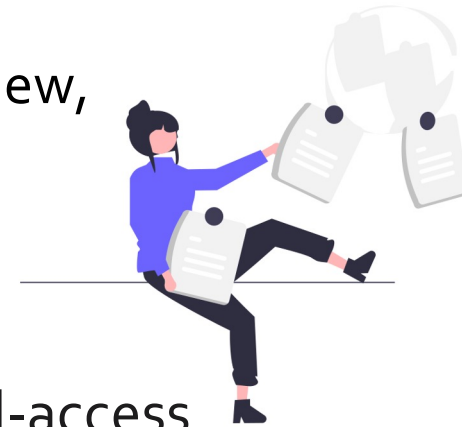
- Outline possible implementation options
- Estimate the implementation/maintenance cost for those options
- Perform a Decision Analysis and Resolution (DAR).
- If an RDMkit-based service is the chosen option, then apply [our formal decision process](#).



# End2End Support Service for controlled-access data sharing [LU, Pinar]

## Background

- We provide a Data Hosting service for Translational Biomedicine projects that are nearing their end.
- We host data with the condition that it is shared.
- As part of hosting we facilitate controlled-access
  - collect and publish ELSI Metadata,
  - handle data access requests and coordinate the DAC review,
  - provision the data for approved requests.



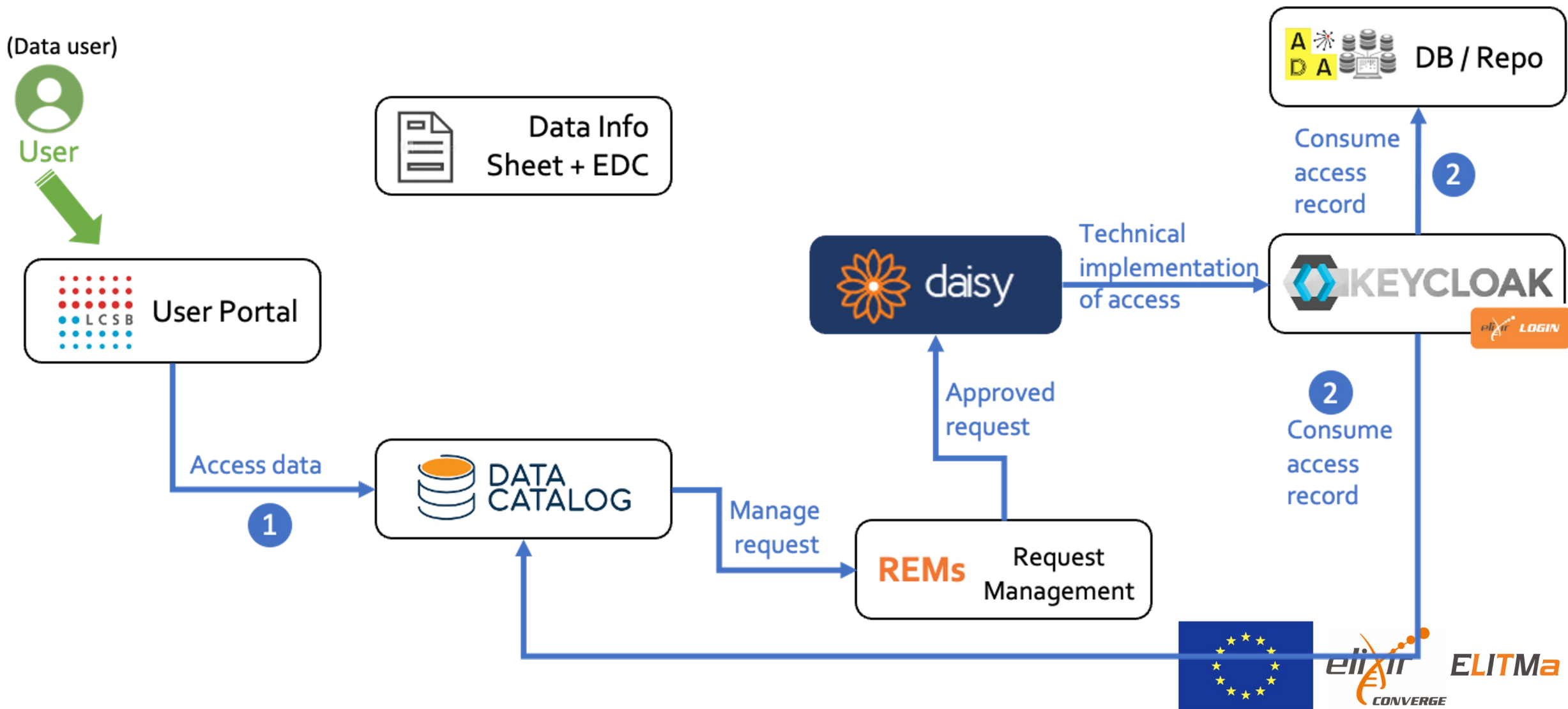
## Why this service

- As the hosting service increased in popularity, the workload for controlled-access facilitation became rather significant,
- GDPR documentation requirements were pervasive,
- Data provision is transitioning from file download to cloud access.



# End2End Support Service for controlled-access data sharing [LU, Pinar]

## Access request handling, access management.



# End2End Support Service for controlled-access data sharing [LU, Pinar]

## How we implemented it



- We gradually increased automation in our process.
- Development spanned several years.
- All Node teams collaborated; ELSI, Data Stewardship, Tech/Infra.
- Data Protection By Design and By Default (DPbDD) was applied.
- Where possible:
  - we re-used RDM tools from other Nodes (REMS)
  - we adopted standards (OIDC)



# End2End Support Service for controlled-access data sharing [LU, Pinar]

## What worked well

- Focusing on the process and info models helped.
  - First we optimised those, then we introduced automation.
- Our Node will join one or more data sharing federations.
  - End2End was our prep work and we went live just in time.
- Our ELSI team is working on a DPbDD checklist for future assemblies/services.





# End2End Support Service for controlled-access data sharing [LU, Pinar]

## What were the challenges

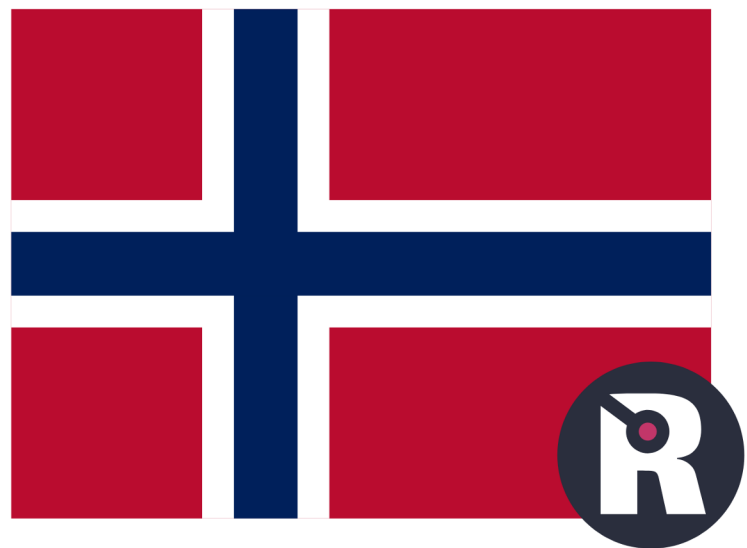
- We underestimated the assembly development effort.
  - Assembling tools takes time (connectors/parsers/mappers).
  - Testing assemblies is non-trivial.



# RDMkit, ELIXIR Norway - Korbinian Bösl



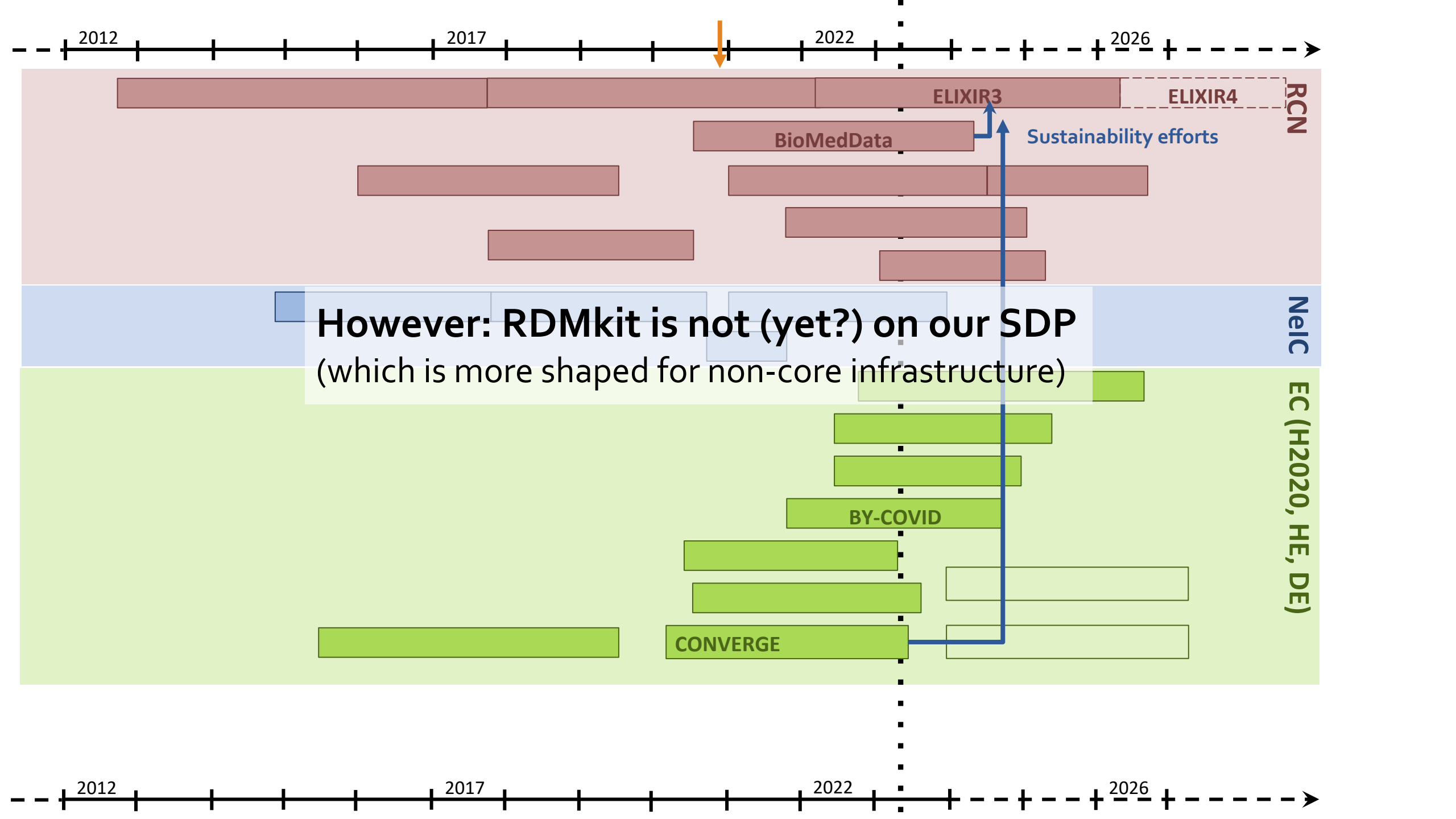
# RDMkit at ELIXIR Norway, Korbinian Bösl & Nazeefa Fatima



## Why this service:

- RDM is central for us as a node
- Need for  $\geq 1$  place for RDM knowledge sharing
- Synergistic effect with national project
- Involved from the very beginning





# RDMkit at ELIXIR-Norway, Korbinian Bösl & Nazeefa Fatima

## What works well

- Good uptake from institutions and users
- Recognition from funders
- Usage of RDMkit in training
- Reuse of theme

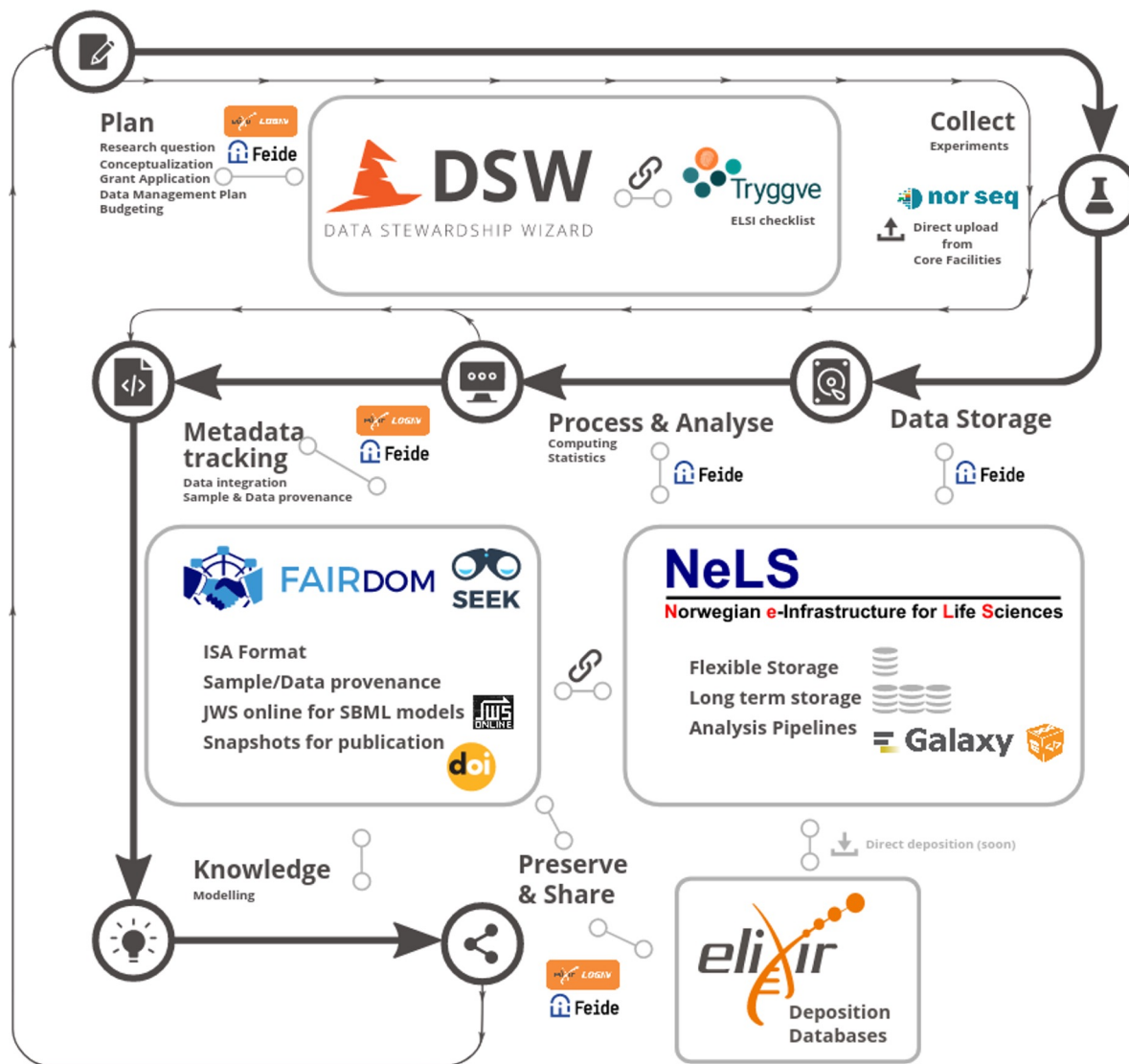
# RDMkit at ELIXIR-Norway, Korbinian Bösl & Nazeefa Fatima

## Challenges

- Guiding users to our DS-Wizard instance
- Multiple layers of information in RDMkit - overwhelming for passing by user
- Cross linking with our additional RDM information resource (RDM lookup)
- Convincing national RIs/community to contribute (often referring to their ESFRI instead)
- Space to capture knowledge out of the RDMkit scope
- Maintenance debts



# NeLS tool assembly, Korbinian Bösl & Nazeefa Fatima



## How we implemented it:

- Historically evolved from our core storage infrastructure
- Complementary functionality added through other tools over time
- Based on national HPC/big data infrastructure



# NeLS tool assembly, Korbinian Bösl & Nazeefa Fatima

## What worked well

- Covering classic sequencing/bioinformatic use cases
- Focus on single processes
- End-to-end PoC





# NeLS tool assembly, Korbinian Bösl & Nazeefa Fatima

## Challenges

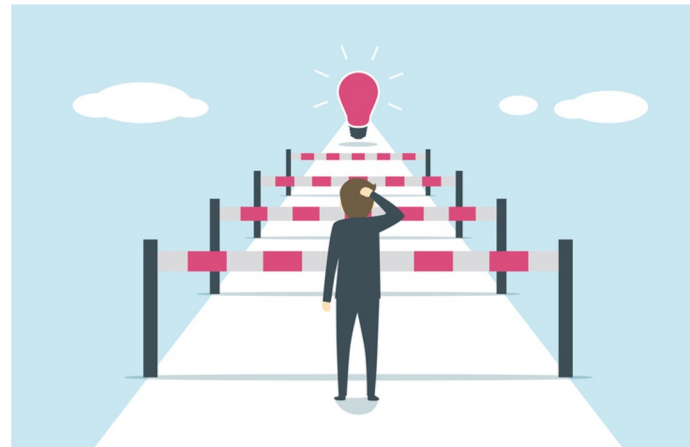
- Other use cases
- Synergy with sensitive data tool assemblies
- Long term effort
- Automatisation of support processes
- User experience
- Many different technologies
- Technical debts/History to carry
- Projects w/o Data Stewards
- Brokering



**DSW, ELIXIR-CZ, Anna Strachotova**



# Data Stewardship Wizard <https://ds-wizard.org/>



# Open science -> Open Science & Data management in CZ

EC Directive 2023/98/EC, on the reuse of public sector information

Open Research Data Pilot in H2020 for all disciplines

Implementation of EU Directive 2019/1024 - Amendment published in the collection of laws under number 241/2022 Coll.

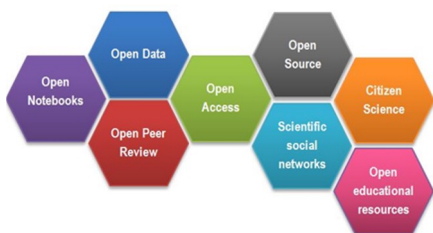
Open Research Data Pilot in H2020 for selected disciplines



Czech funding programmes start to require DMPs (usually based on Horizon EU)



OECD Principles and Guidelines for Access to Research Data from Public Funding



Codification of FAIR principles: WILKINSON? Mark D. et al. (53 authors, Nature)

[EU Directive 2019/1024, on open data and the reuse of public sector information](#)



# Why DSW



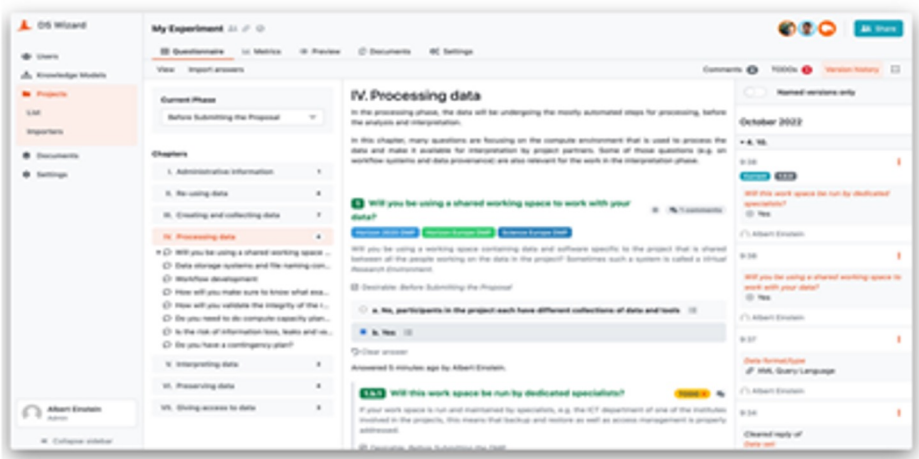
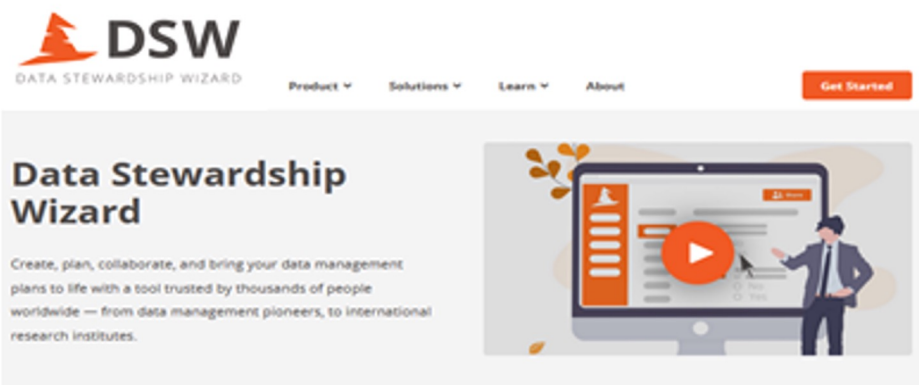
Guidelines for open access to publications, data and other research outputs

## Data Management Plans: tools

The following resources support the generation of DMPs:

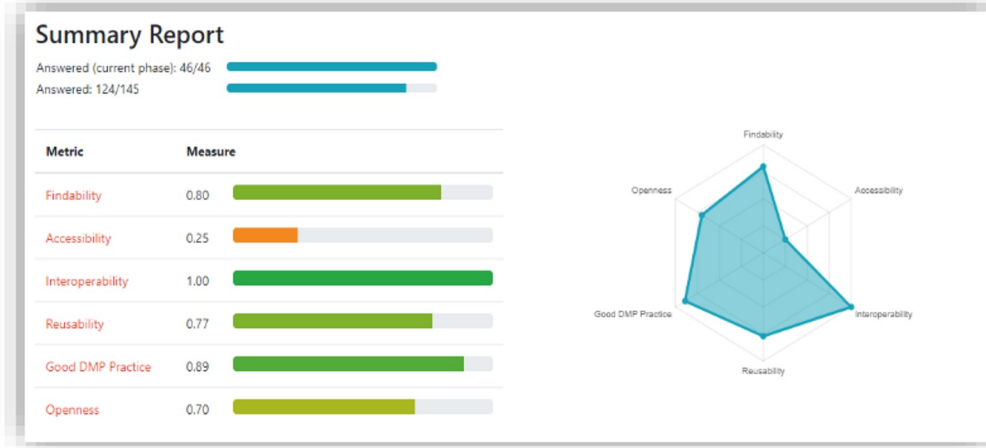
A template for the DMP can be found in the online H2020 guide for applicants, here.

- The DMPONLINE tool, specifically supports the development of project data management plans: <https://dmponline.dcc.ac.uk>
- ARGOS is an online tool to automate the creation, management, sharing and linking DMPs with the research artefacts to which they correspond. It is the joint effort of OpenAIRE and EUDAT CDI to deliver an open platform for Data Management Planning that addresses FAIR and Open best practices and assumes no barriers for its use and adoption (<https://argos.openaire.eu>).
- **The Data Stewardship Wizard is a joint ELIXIR CZ and ELIXIR NL tool, bringing a simple but powerful solution for researchers to help them understand what is needed for good, FAIR-oriented data stewardship, and build their own Data Management Plans (<https://ds-wizard.org/>).**



# Why DSW

- Living DMPs as Inverative Questionnaires
- Guidance and linded (ELIXIR) resources
  - Collaboration in Interactive Questionnaires (history, sharing project, comments)
  - Knowledge Models
  - Document Templates
  - Multiple Documents for Single Questionnaire



**Question** → 1.a.4.b.1.a.1 What repository will this data be stored in?

**Guidance text** → Domain repositories often have the best functionality to make the data findable and reusable: even though it may look like a database that could be reused in a completely different field would be better findable in a generic repository, the limited availability of domain-specific metadata make that less valuable.

Many repositories are listed in <https://fairsharing.org/>

If a repository offers to give your data set a DOI or alternative persistent identifier it is a good idea to use that option.

External links: [FAIRSharing](#), [Registry of Research data Repositories](#)

**Possible answers (input based on question type)**

- a. A domain-specific repository
- b. Our national repository
- c. Our institutional repository
- d. A special-purpose repository for the project

**Who answered and when** → Answered in less than 5 seconds by Albert Einstein.

**Advice for answer** → Disadvantage of a general purpose repository is the lack of data-specific features (e.g. 'play' instead of 'download' for an audio file) and limited findability

References

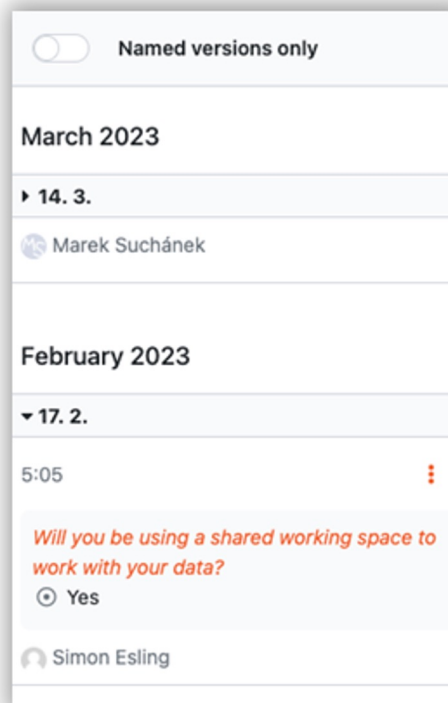
(FAIR) Metrics

Follow-up questions indication

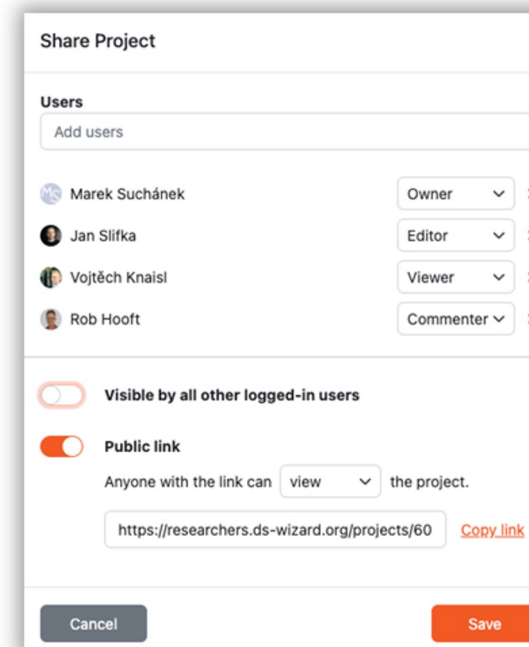


# Why DSW

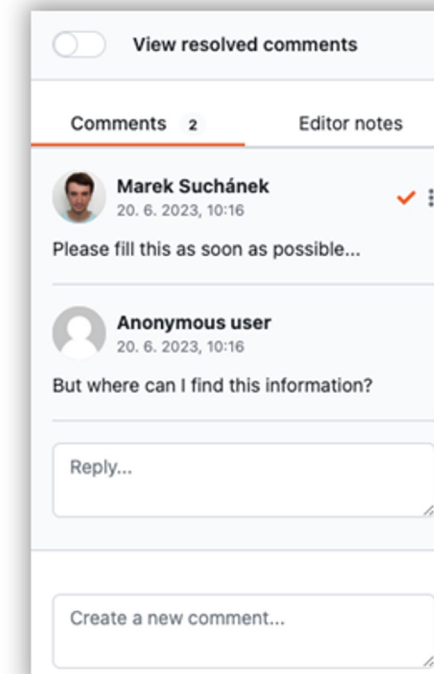
- Living DMPs as Inverative Questionnaires
- Guidance and linded (ELIXIR) resources
- **Collaboration in Interactive Questionnaires (history, sharing project, comments)**
- Knowledge Models
- Document Templates
- Multiple Documents for Single Questionnaire



Version History



Sharing Projects  
(permissions, public links)



Comments

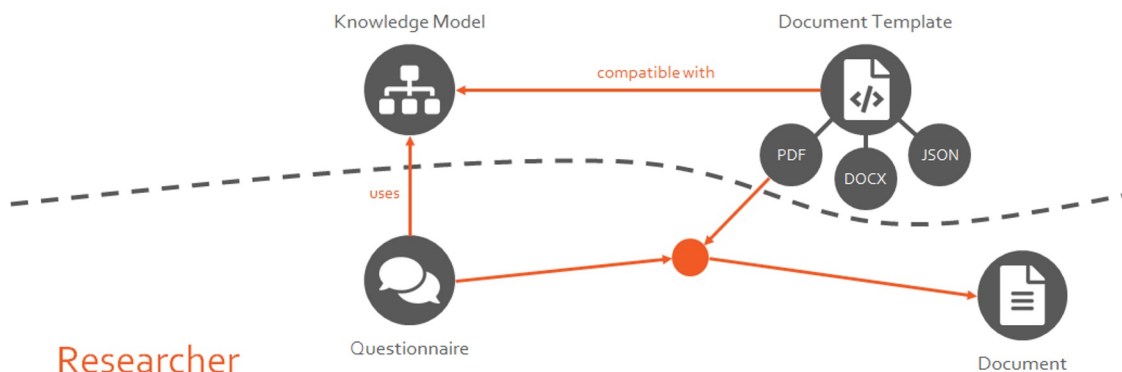


# Why DSW

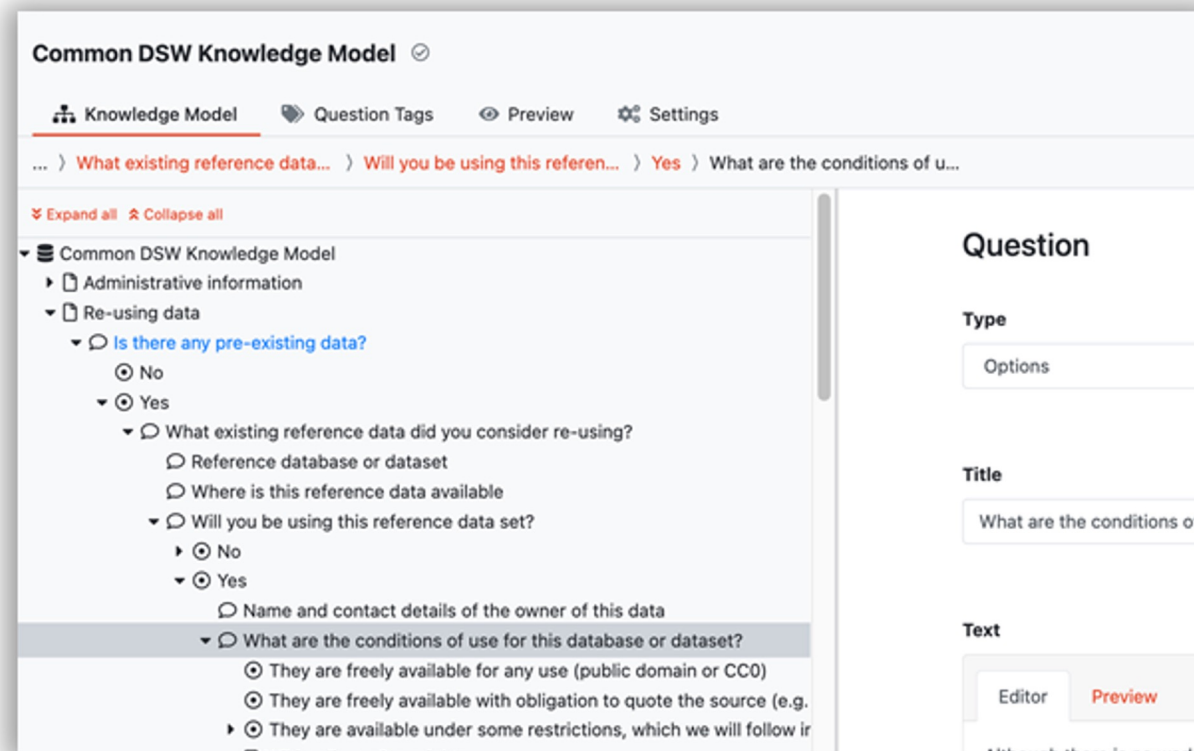
- Living DMPs as Inverative Questionnaires
- Guidance and linded (ELIXIR) resources
- Collaboration in Interactive Questionnaires (history, sharing project, comments)
- **Knowledge Models**
- Document Templates
- Multiple Documents for Single Questionnaire



## Data Steward



## Researcher



A screenshot of the 'Common DSW Knowledge Model' interface. The top navigation bar includes 'Knowledge Model', 'Question Tags', 'Preview', and 'Settings'. The breadcrumb trail shows: '... > What existing reference data... > Will you be using this referen... > Yes > What are the conditions of u...'. Below the breadcrumb, there are 'Expand all' and 'Collapse all' options. The main content area is a tree view under 'Common DSW Knowledge Model', with 'Re-using data' expanded to show a question: 'What existing reference data did you consider re-using?'. The question has two main options: 'No' and 'Yes'. Under 'Yes', there are three sub-questions: 'Reference database or dataset', 'Where is this reference data available', and 'Will you be using this reference data set?'. The 'Will you be using this reference data set?' question has 'No' and 'Yes' options. Under 'Yes', there are two sub-questions: 'Name and contact details of the owner of this data' and 'What are the conditions of use for this database or dataset?'. The 'What are the conditions of use for this database or dataset?' question has three radio button options: 'They are freely available for any use (public domain or CC0)', 'They are freely available with obligation to quote the source (e.g. ...)', and 'They are available under some restrictions, which we will follow in ...'. On the right side, there is a 'Question' panel with fields for 'Type' (Options), 'Title' (What are the conditions of ...), and 'Text' (Editor and Preview buttons).





# Why DSW

- Living DMPs as Inverative Questionnaires
- Guidance and linded (ELIXIR) resources
- Collaboration in Interactive Questionnaires (history, sharing project, comments)
- Knowledge Models
- **Document Templates**
- **Multiple Documents for Single Questionnaire**

## New Document

**Name**

Answered (current phase): 25/28   
 Answered: 109/126 

**Document Template**

**H** **Horizon Europe DMP** 1.3.0  
 Data Management Plan according to the Horizon Europe template



**Format**

HTML  PDF Document

MS Word Document

## New Document

**Name**

Answered (current phase): 25/28   
 Answered: 109/126 

**Document Template**

**m** **maDMP (RDA DMP Common Standard)** 1.14.0  
 Machine-actionable DMP according to RDA Common Standard

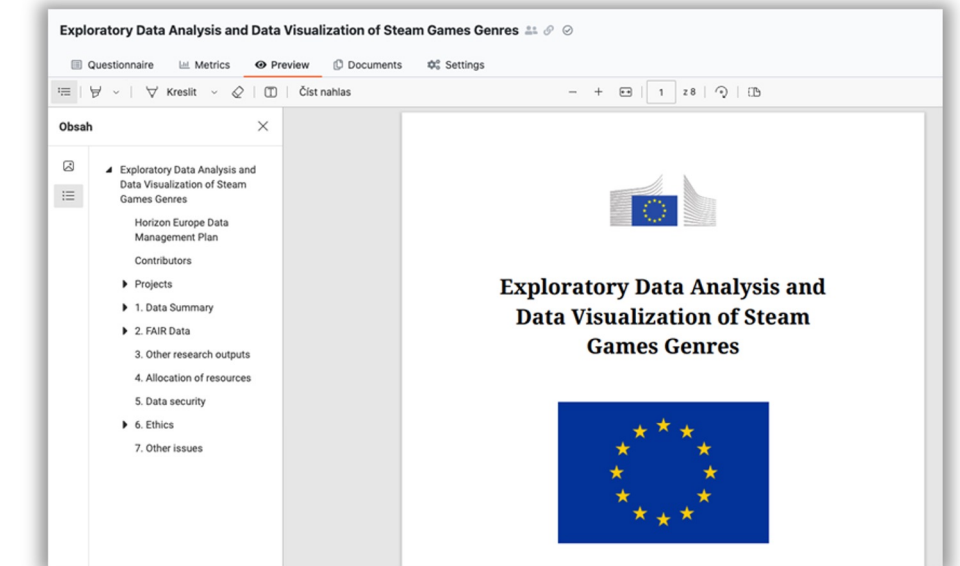
**Format**

JSON  Turtle (no blank nodes)

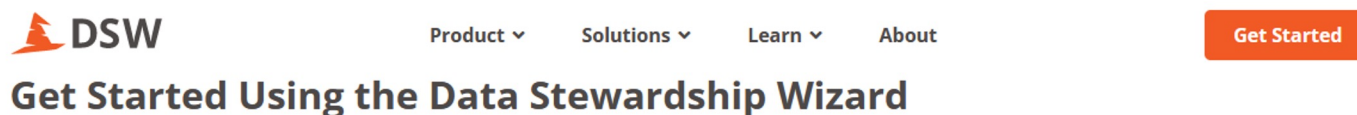
Turtle  N3

RDF/XML  JSON-LD

N-Triples  Trig



# Implementation of DSW - how to start



DSW Cloud for ELIXIR	FAIR Wizard <small>COMMERCIAL</small>	Self-Managed
For researchers in ELIXIR	For groups and institutions	For open-source adventurers
<ul style="list-style-type: none"><li>• Easy to sign up and use</li><li>• Ready to use Knowledge Models</li><li>• Privacy and stability</li></ul>	<ul style="list-style-type: none"><li>• All the DSW features available</li><li>• Your own instance with ease</li><li>• Enterprise-ready ecosystem</li></ul>	<ul style="list-style-type: none"><li>• Requires own IT infrastructure</li><li>• Absolute freedom in setup</li><li>• Keeping everything in-house</li></ul>
<ul style="list-style-type: none"><li>✓ Projects &amp; DMPs</li><li>✓ Online Collaboration</li><li>✓ Version History</li><li>✓ Life Science Login</li></ul>	<ul style="list-style-type: none"><li>✓ Projects &amp; DMPs</li><li>✓ Online Collaboration</li><li>✓ Version History</li><li>✓ Knowledge Models Management</li><li>✓ Template Management</li><li>✓ User Management</li><li>✓ Settings &amp; Configuration</li><li>✓ Enhanced SSO and Integrations</li><li>✓ Deployment Customisations</li><li>✓ Support and Consultancy</li><li>✓ Additional Services</li></ul>	<ul style="list-style-type: none"><li>✓ Projects &amp; DMPs</li><li>✓ Online Collaboration</li><li>✓ Version History</li><li>✓ Knowledge Models Management</li><li>✓ Template Management</li><li>✓ User Management</li><li>✓ Settings &amp; Configuration</li></ul>
<a href="#">Learn more</a>	<a href="#">Start now</a>	<a href="#">Explore guide</a>

Several institutions (mainly universities) even **beyond ELIXIR** started or are **starting to adopt DSW** (with our support) for planning RDM...



# Implementation of DSW - DSW Cloud for ELIXIR

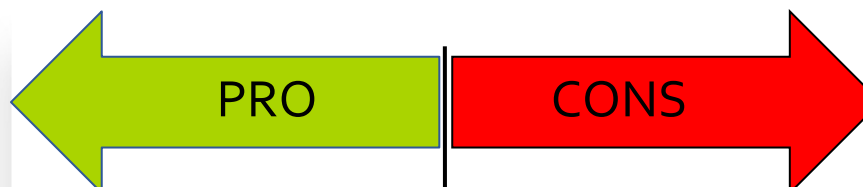
## DSW Cloud for ELIXIR

For researchers in ELIXIR

- Easy to sign up and use
- Ready to use Knowledge Models
- Privacy and stability

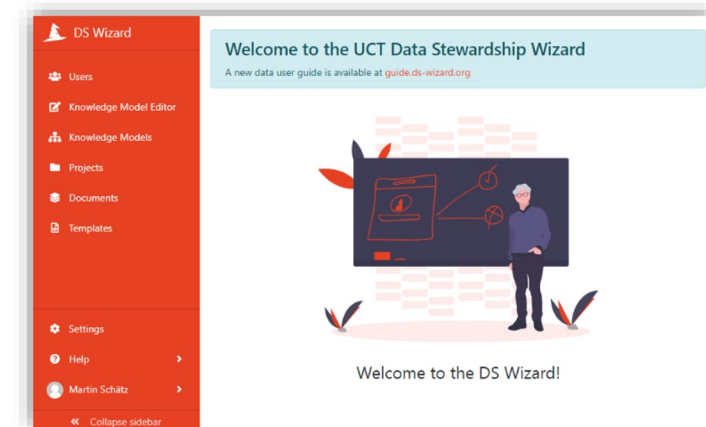
- ✓ Projects & DMPs
- ✓ Online Collaboration
- ✓ Version History
- ✓ Life Science Login

Learn more



- minimization of claims on DSW users
- **DSW team operates the system** -> installing and updating, hardware, storage, etc.
- the institution **uses DSW as a service** and only takes care of configuring end-user settings

- **full dependence** on the operators of the DSW Cloud for ELIXIR infrastructure
- **the end user cannot influence** any maintenance windows, service outages, but neither can they prevent them transitions to new versions of software, etc.,
- **the user doesn't have control over the physical location** of its data storage, backup, etc.



The ELIXIR operates a cloud infrastructure in which it is possible to host institutional DSW installations that are used by users as a service. The user has access on its own subdomain (e.g. <https://vscht.ds-wizard.org/>). Institutional instances in the ELIXIR infrastructure are available to ELIXIR members or projects with ELIXIR associates.

A special shared is available for end-researchers-individuals instance <https://researchers.ds-wizard.org/> that individuals can directly use for preparation of their DMPs.



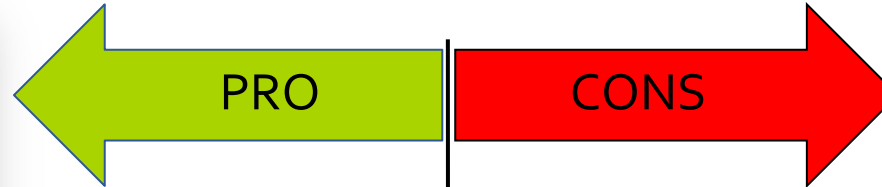
# Implementation of DSW - Self-Managed

## Self-Managed

For open-source adventurers

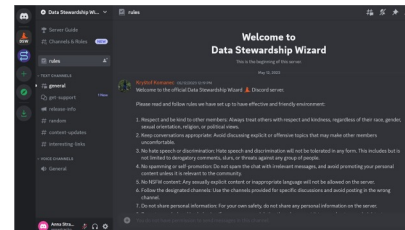
- Requires own IT infrastructure
- Absolute freedom in setup
- Keeping everything in-house
- ✓ Projects & DMPs
- ✓ Online Collaboration
- ✓ Version History
- ✓ Knowledge Models Management
- ✓ Template Management
- ✓ User Management
- ✓ Settings & Configuration

Explore guide

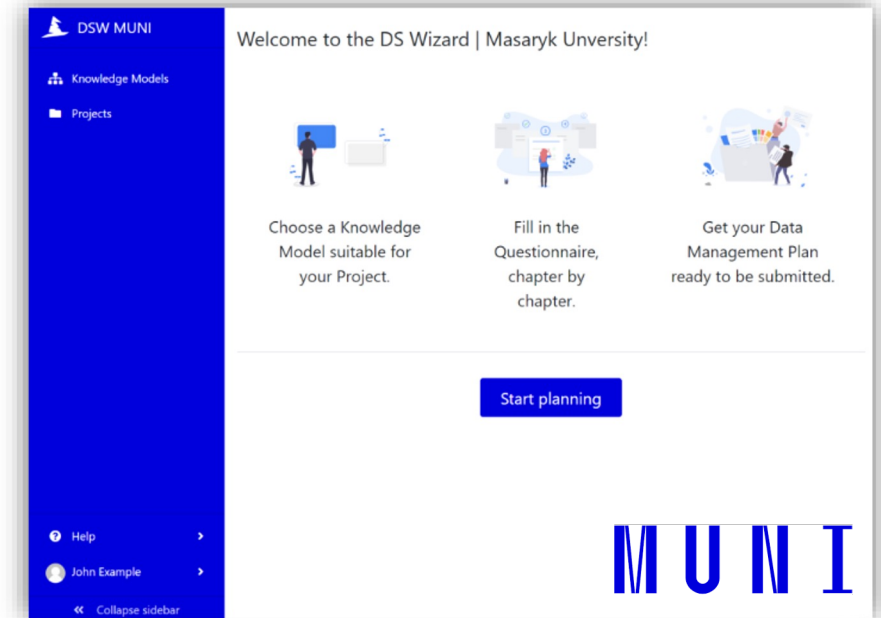


- ➔ **full control** over installation, upgrade dates, configuration changes, etc.
- ➔ freely available source codes allows **multiple DSW instances** -> easily run test or
- ➔ development through its **RESTful API** can facilitate integration with local systems

- ➔ **full responsibility** for the complete management, including security, backup, updating, etc.
- ➔ **actively monitor** the development of DSW



<https://discord.com/channels/693010138302513174/1102900416783597620>



Installation of the application on own hardware/infrastructure from downloaded source codes or Docker images. About the operation, backup, updates, settings, etc. of the entire instance of DSW then it must be taken care of by its user-operator.



# Challenges

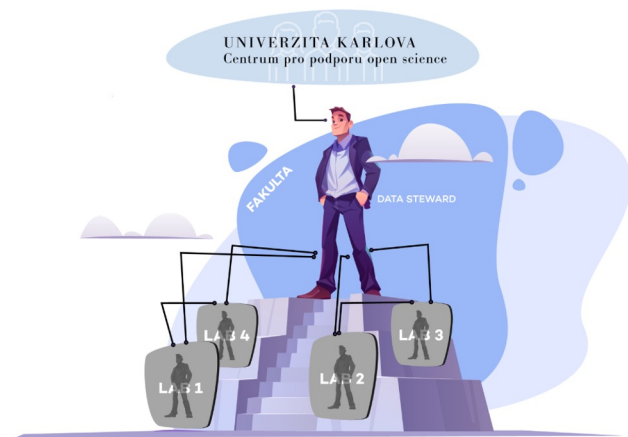


# Challenges



Czech funding programmes also start to **require DMPs** (usually based on Horizon EU)

**National Data Steward community** with more than 120 people



Several institutions (mainly universities) even **beyond ELIXIR** started or are **starting to adopt DSW** (with our support) for planning RDM...



There is still relatively early phase of "proper" RDM in Czechia; however, **National Repository Platform** project is being prepared... and DSW with **customized content** is included!



# DSW, ELIXIR-NO, Korbinian Bösl



DMP Landscape analysis

- Funder requirements (specific tool?)
- Features

Self-hosting vs **SaaS?**

KM from scratch or **adapt life science KM**

First workshops & Feedback

Dedicated Feature developments (financed by Centre for Digital Life Norway)



Exemplar DMPs from other RIs

Recommended tool by major funder

Integration with knowledge resource

Project for national pilot from major universities



2019

2023



		DMPonline	DMPTool	Argos	DataWiz	easyDMP
<a href="#">Template Selection</a>	✓	✓	✓	✓	Limited	✓
<a href="#">Detailed Export Settings</a>	✓	✓	✓	✓	Limited	Limited
<a href="#">Template Creation</a>	✓	✗	✗	✗	✗	Limited
<a href="#">Template Styling</a>	✓	✗	✗	✗	✗	✗
<a href="#">Project Phases</a>	✓	✗	✗	Unknown	Limited	✗
<a href="#">Suitable For Worldwide Use</a>	✓	Limited	Limited	Limited	Limited	Limited
<a href="#">Various Sharing Options</a>	✓	Limited	Limited	Limited	✓	Limited
<a href="#">Comments</a>	✓	✓	✓	✗	✗	✗
<a href="#">Machine-Actionability</a>	✓	Limited	Limited	✓	Limited	Limited
<a href="#">FAIR Metrics Information</a>	✓	✗	✗	Limited	✗	Limited
<a href="#">TODOs</a>	✓	✗	✗	✗	✗	✗
<a href="#">Live Collaboration</a>	✓	✗	✗	✗	✗	✗
<a href="#">Migrations</a>	✓	✗	Unknown	Unknown	✗	✗
<a href="#">Version History</a>	✓	Unknown	Unknown	Limited	Limited	Unknown
<a href="#">Languages (other than English)</a>	✓	✓	✓	Limited	✓	✗
<a href="#">Funder Templates</a>	✓	✓	✓	✓	Limited	Limited



## The Data Stewardship Wizard is provided by:



### Demo

For exploring the DSW features

- Easy to sign up and use
- A shared instance with other users
- **Not for serious usage**

### Researchers

For individual researchers

- Easy to sign up and use
- Ready to use Knowledge Models
- Privacy and stability

### Self-Managed

For organizations

- All the DSW features available
- Your own instance
- You need to host and run the instance by yourself

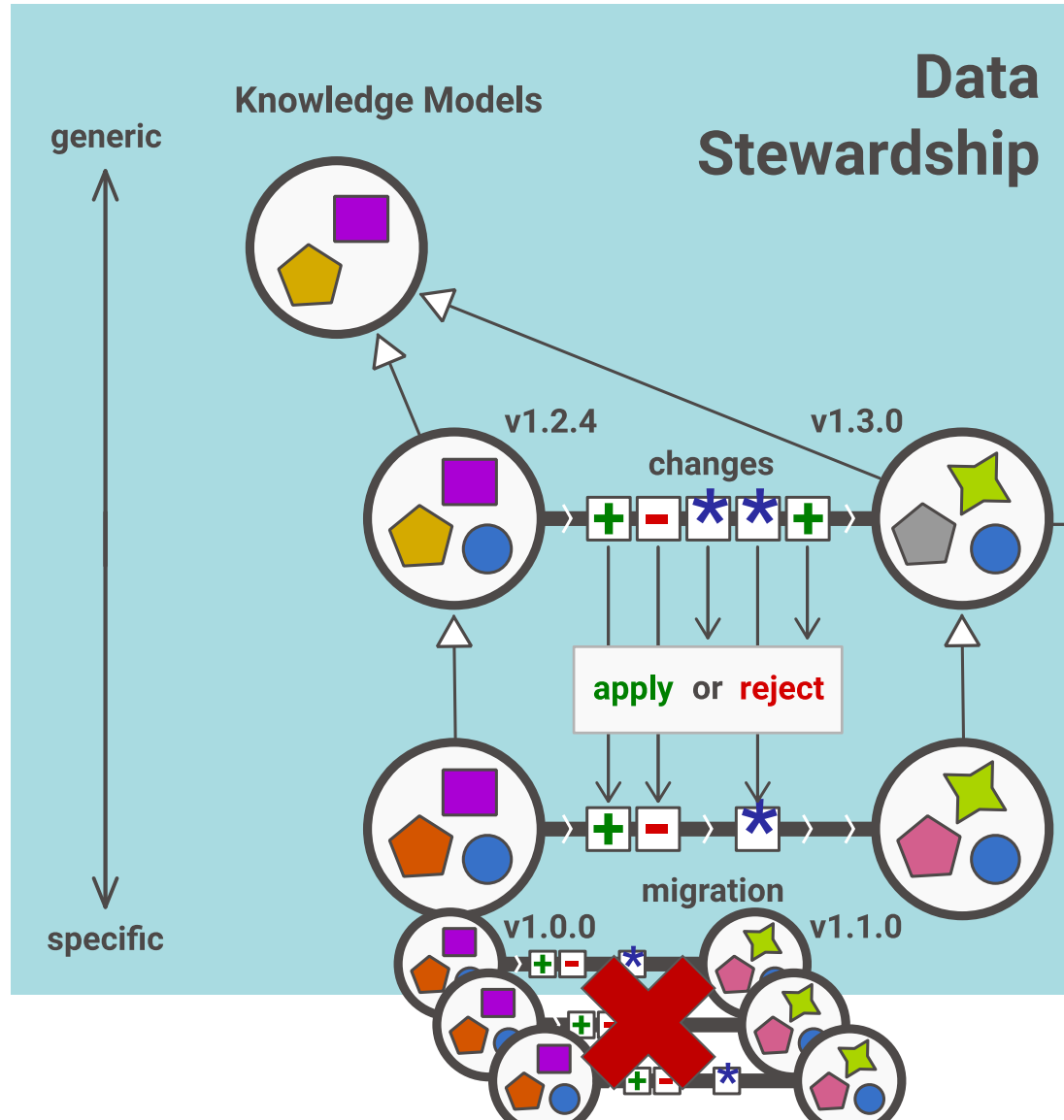
### DSW Cloud

For organizations

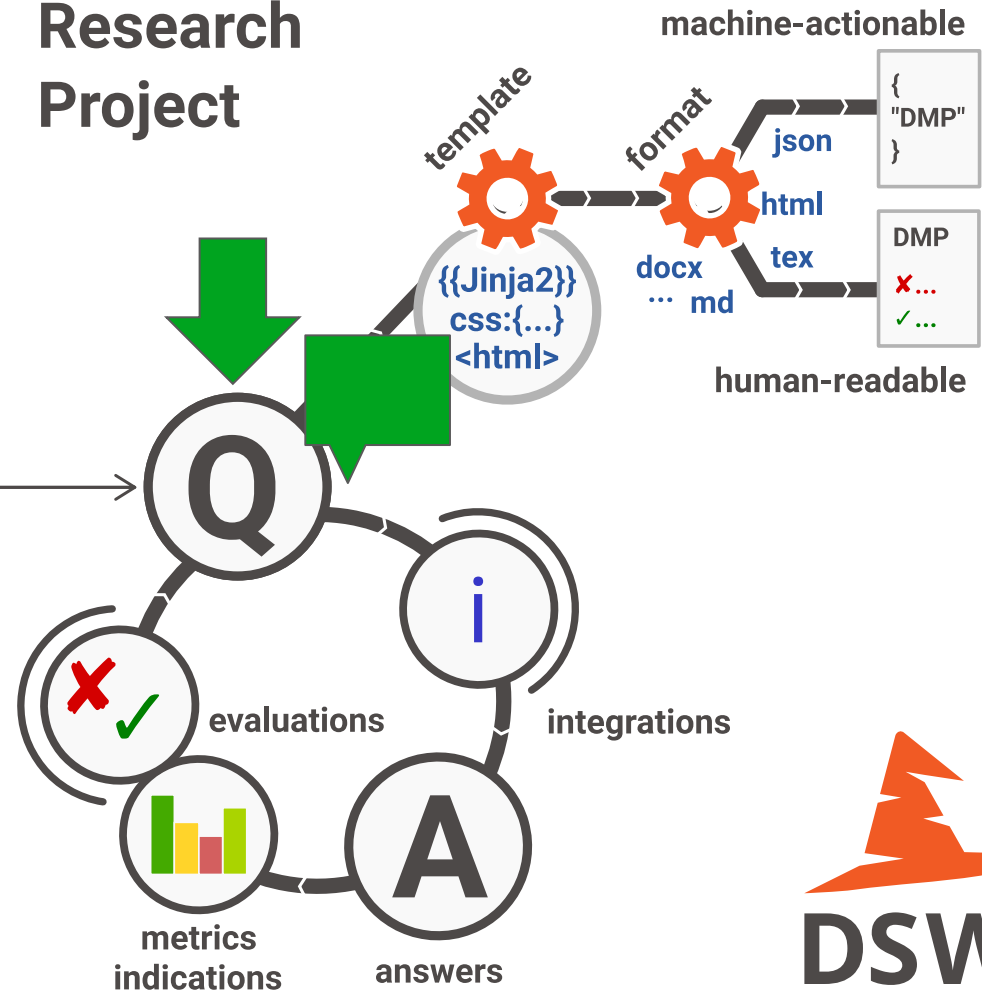
We offer managing the DS Wizard instance for interesting projects that want to use it seriously but don't want to run it by themselves. \*



# Considerations



## Research Project



# Recommendations



## Create Project

From Template  Custom

### Name

### Template

--

- M **Mass spectrometry-based proteomics**  
current best practice recommendations from NAPI
- N** **NGS of non-human samples**  
best practise recommendations from NorSeq
- D **Preclinical imaging**

### Chapters

- I. Administrative information 1
- II. Re-using data 3
- III. Creating and collecting data 6**
- IV. Data sensitivity 1
- V. Processing data 3
- VI. Interpreting data 3
- VII. Preserving data 5
- VIII. Giving access to data 3



#### 4.a.2.b.1.a.1 Minimal Information Standard

Horizon 2020 DMP **Sigma2**

Darwin Core

[FAIRsharing](https://fairsharing.org/bsg-s000195) <https://fairsharing.org/bsg-s000195>

Answered about 1 month ago by Ola Nordmann.

#### 4.a.2.b.1.b.1 Minimal Information Standard

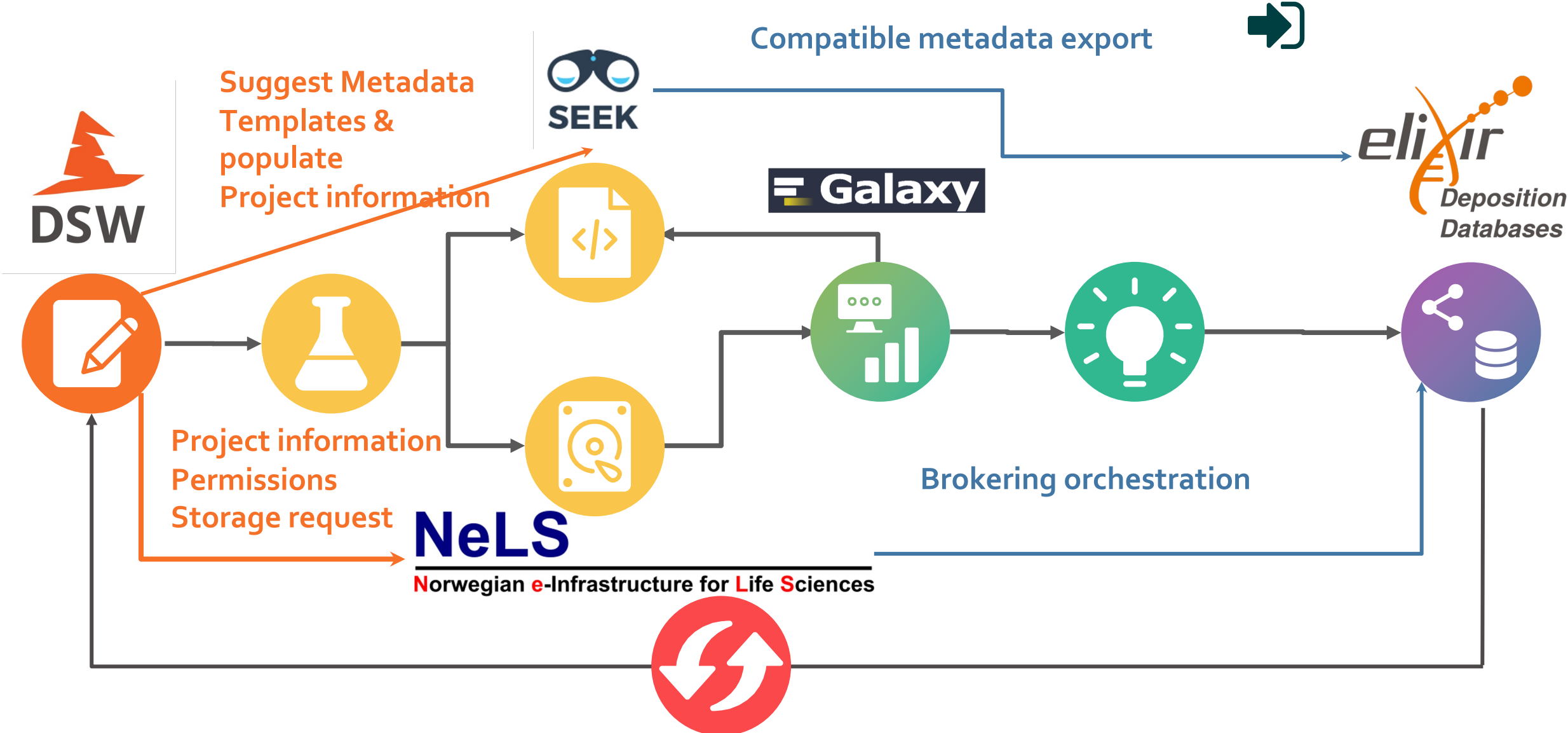
Horizon 2020 DMP **Sigma2**

Minimum Information about any (x) Sequence

[FAIRsharing](https://fairsharing.org/bsg-s000518) <https://fairsharing.org/bsg-s000518>

Answered about 1 month ago by Ola Nordmann.

# Ambition



# What works well

- Using DSW in workshops, ...
- Convincing stakeholder about advantages of DSW
- 'Smooth sailing' in DSW Cloud (updates, ...)
- Linking 'exemplar DMPs' from our RDMlookup (=knowledge platform)

# Challenges

- Default KM can be overwhelming for 1<sup>st</sup> time users
- Getting users to our instance (instead of generic DSW instance)
- Difficulty for some stakeholders to understand tool (=DSW) vs. KM (e.g. LifeScience KM)
- Text export from DSW w/o questionnaire - challenging for institutional support units
- OpenID challenges with national edu ID service

fEGA, ELIXIR-FI, Riku Riski





# Federated European Genome-phenome Archive (FEGA) Finland



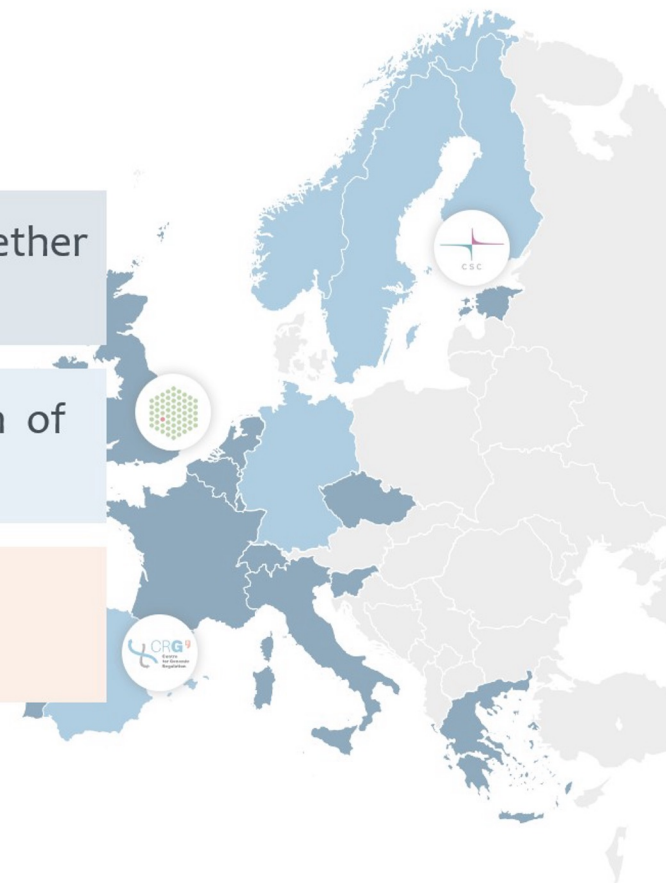
Federation consists of national nodes working together with **European Genome-phenome Archive (EGA)**



FEGA has been developed with the contribution of central EGA and the **ELIXIR** network



The Finnish FEGA Node is operated by **CSC**



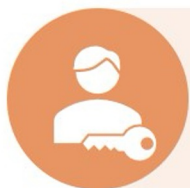
# Federated European Genome-phenome Archive (FEGA) Finland



Complies with national laws and the General Data Protection Regulation (GDPR)



All sensitive data are stored within Finland at CSC data centers



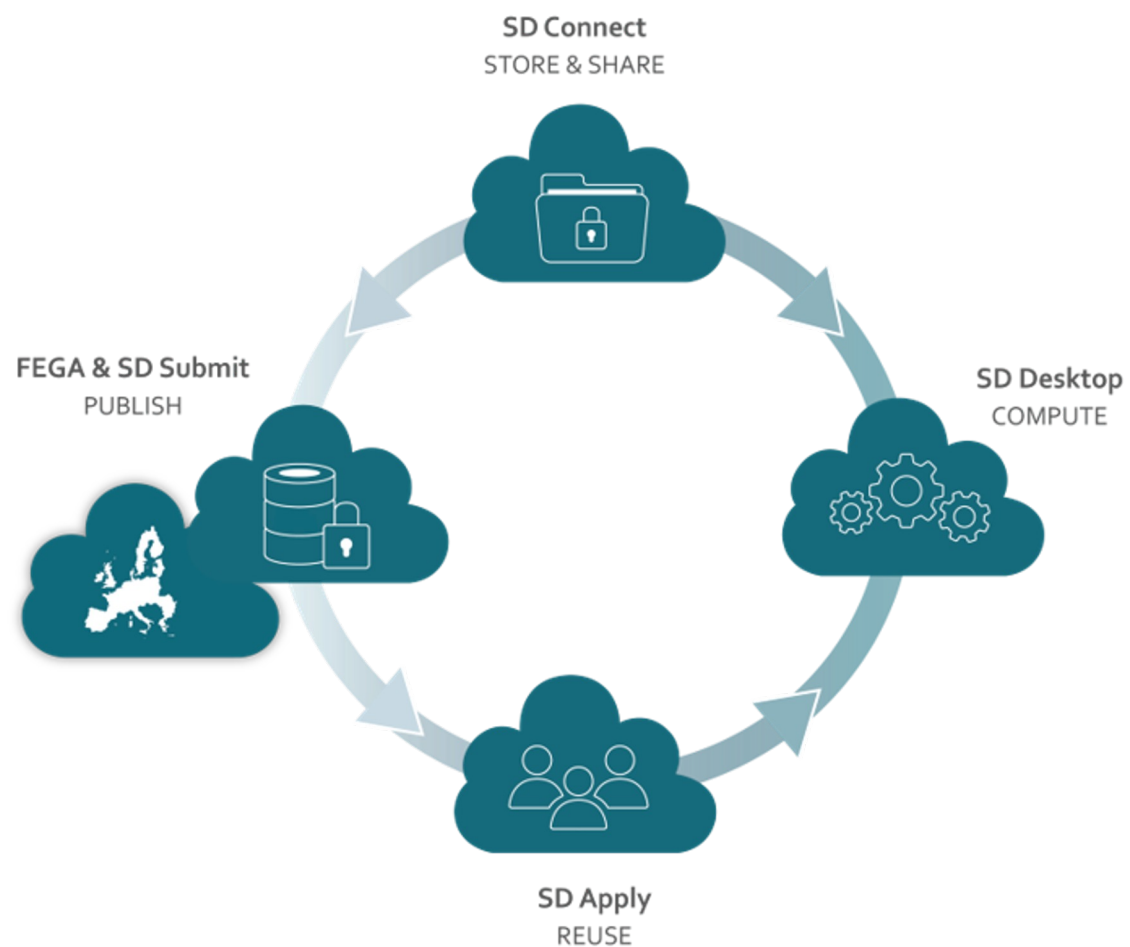
The data owner acts as a data controller and CSC as a data processor



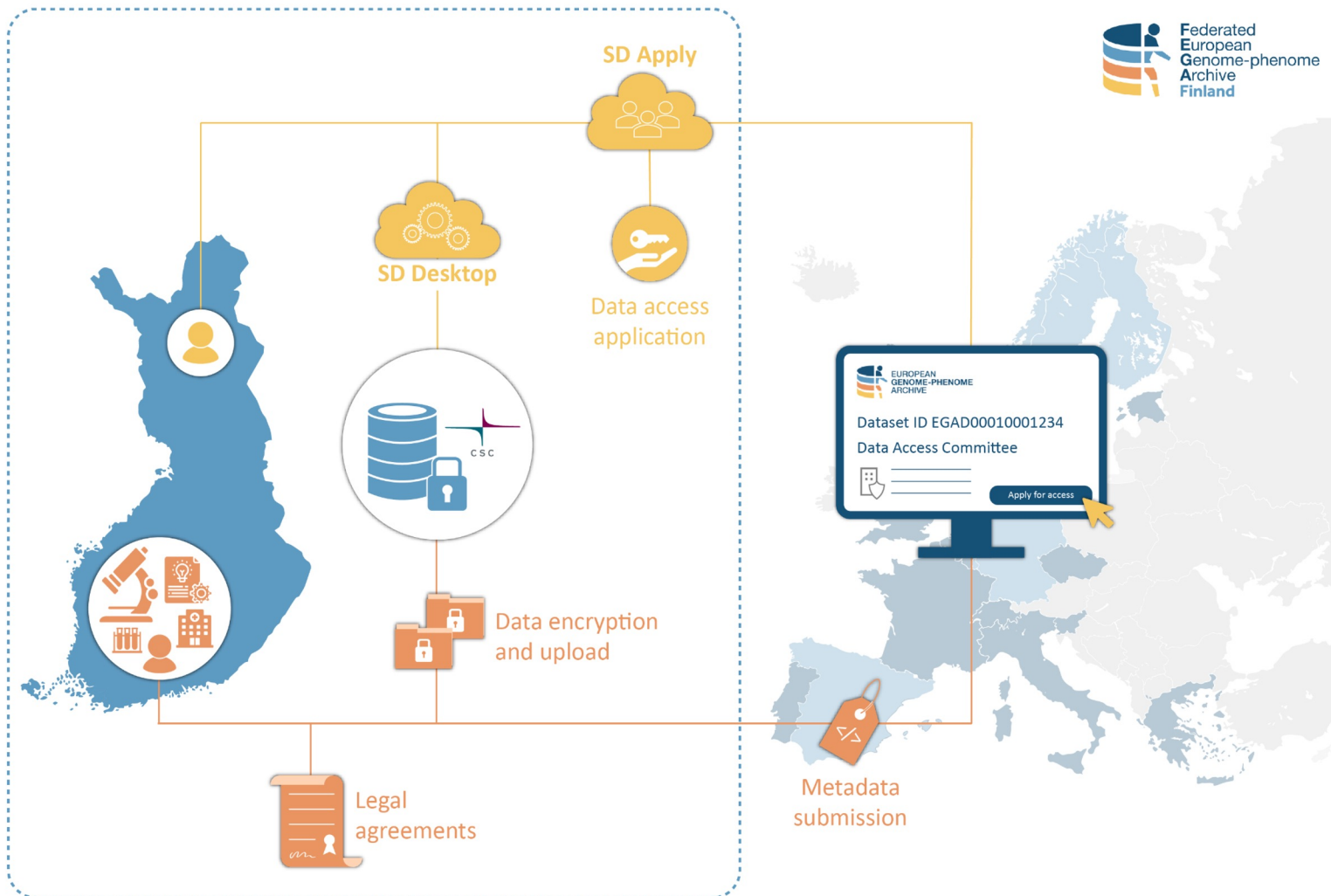
Public metadata is discoverable in the international discovery service



# Sensitive Data Services for Research



# FEGA Finland



# Summary

## Why FEGA?

- National customers/data
- FHD community, centrepiece in IS 22-23
- Co-lead the community, founding member

## How we implemented it?

- Reference, using global standards
- Suite of audited sensitive data services
- Growing a team systematically (e.g. PO, UX, tech support)
- In collaboration with EGA & other nodes, Nordic development team

## What works well

- Cross-node collaboration
- Scaling of existing components
- Company policy SOPs/SLAs

## Challenges

- Contracts, contracts... (DPAs, CA slow)
- Software dependencies / release management
- Production stability



# fEGA, ELIXIR-NO, Korbinian Bösl





[Docs and Info](#) [Privacy Policy](#) [Terms of Service](#) [Contact Us](#) [About](#)

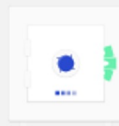
# Federated EGA Norway node



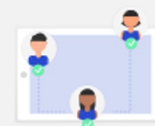
The Federated EGA (European Genome-phenome Archive) is a distributed solution for sharing and exchange of human -omics data across national borders. Federated instance collects metadata of -omics data collections stored in national or regional archives and make them available for search through the main EGA portal.



Searchable

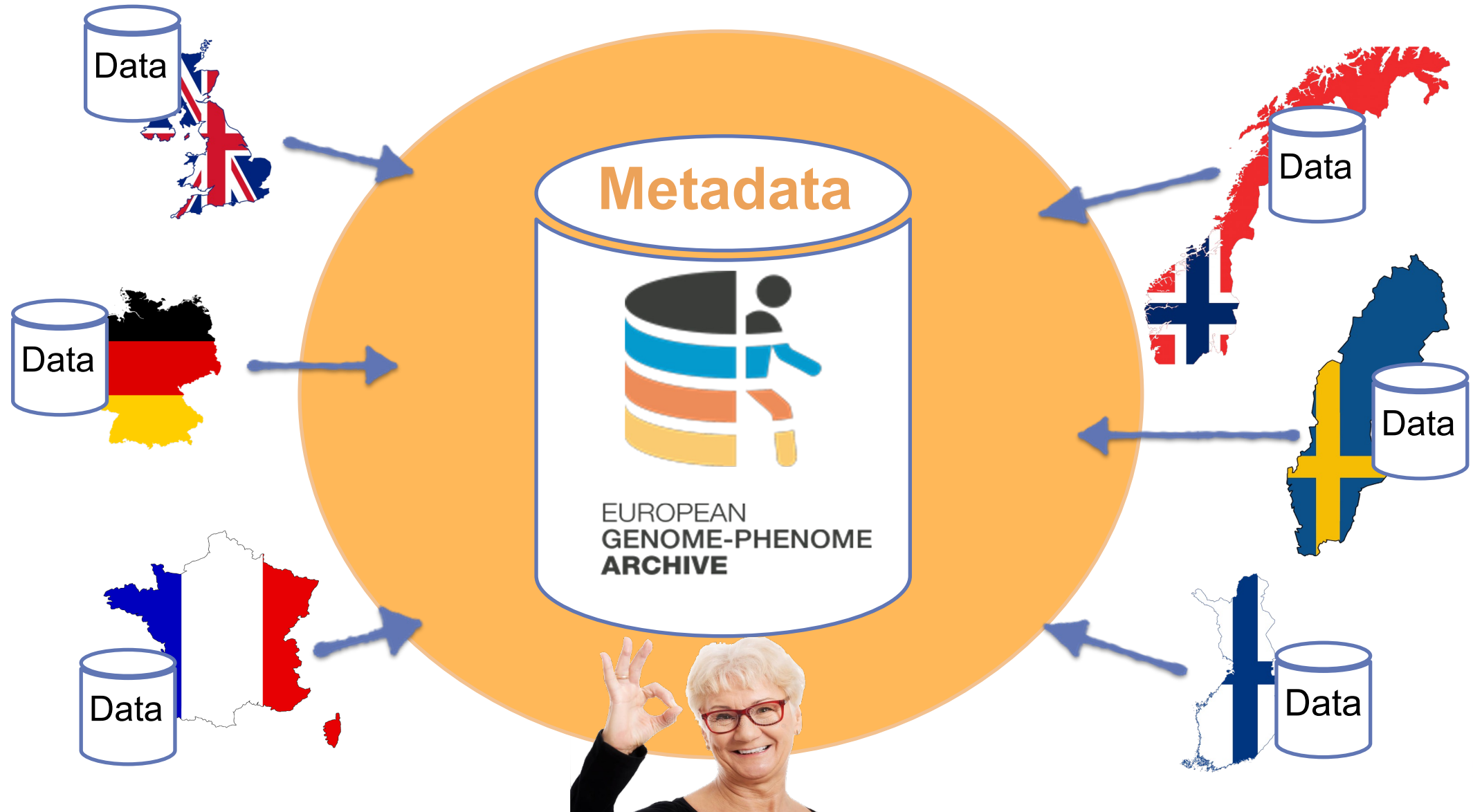


Secure



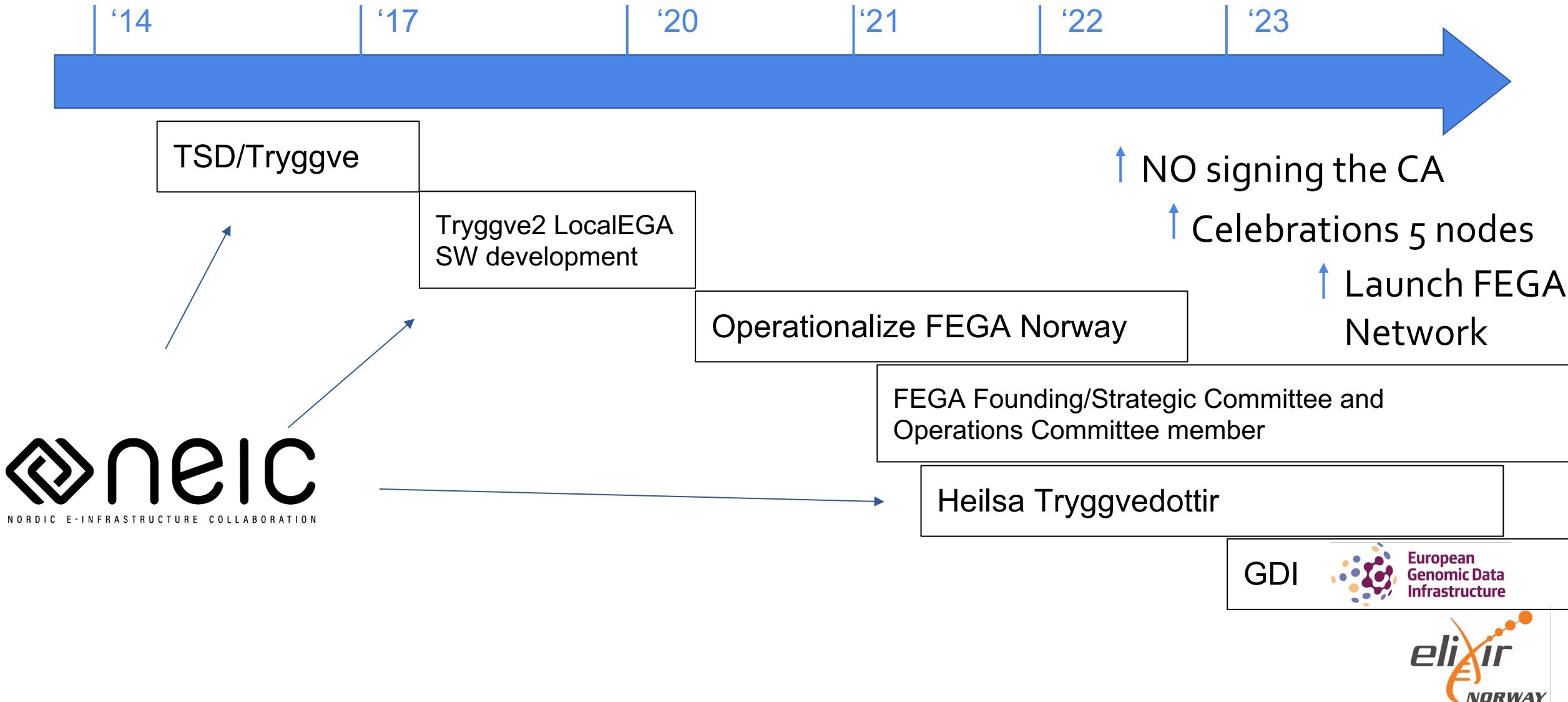
Shareable

# Federated EGA makes EGA scalable and GDPR-compliant

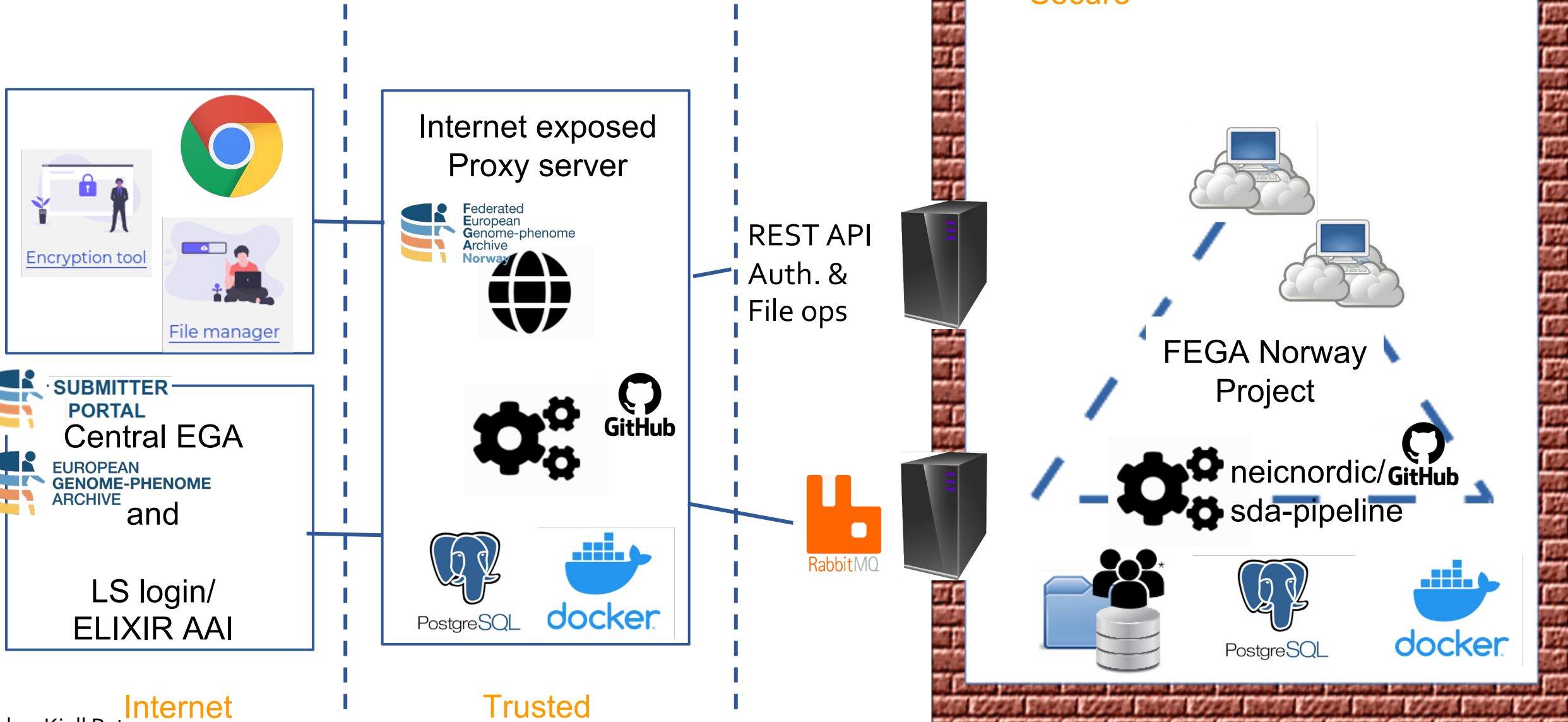




# Timeline FEGA Norway (and GDI)



# FEGA Norway overview





## Protecting

The sensitive data stored at Federated EGA instance is encrypted with Crypt4GH - a new standard file container format from the Global Alliance for Genomics and Health (GA4GH), allowing genomic data to remain secure throughout their lifetime, from initial sequencing to sharing with professionals at external organizations.

*Strong encryption both in transit and at rest are critical requirements for Genomics England and other genomics in healthcare initiatives.*

**Augusto Rendon, Crypt4GH implementation leader at Genomics England**

Hosted by



UiO : Universitetet i Oslo



### Deposit data

Please follow the instructions on how to perform a [submission](#).



### Search data

Please follow the instructions on how to perform a [search](#).



### Retrieve data

Please follow the instructions on how to perform a [retrieval](#).



### Approve access

DAC, please follow the instructions of data [approval](#).

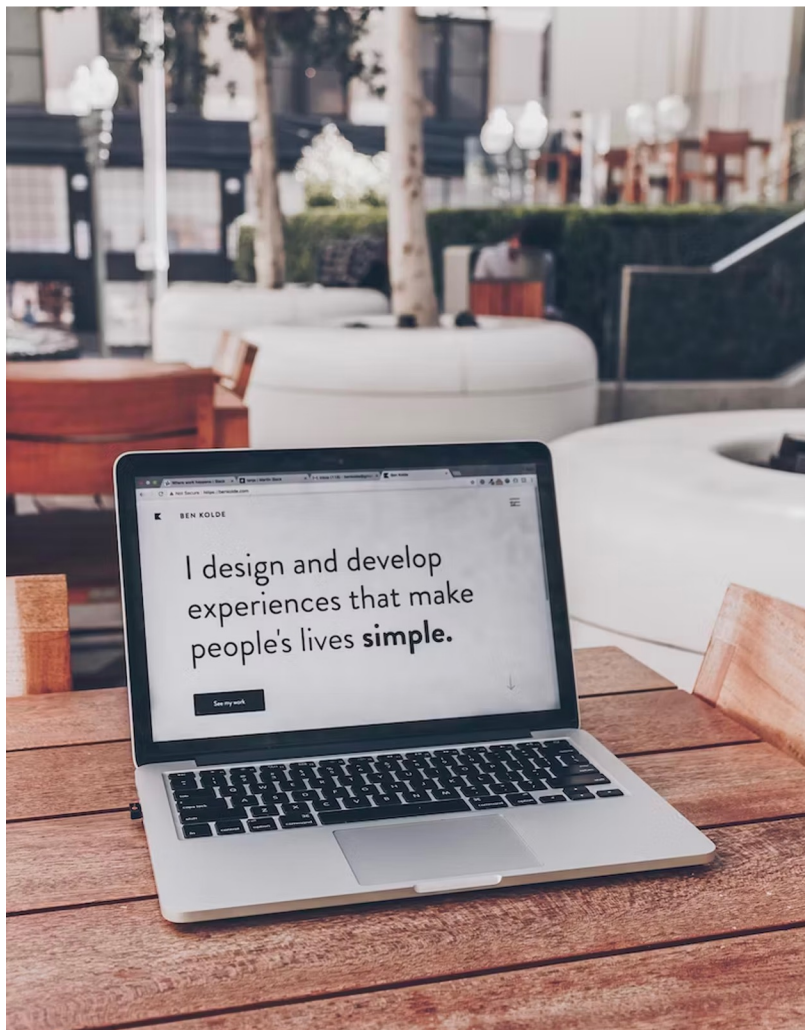
# What works well

- Small team of generalists -> agile development
- Using existing TRE legal and technical framework
- Nordic collaboration and funding; synergy & exchange
- Great recognition by national authorities

# Challenges

- Very long process
- Vulnerable to staff turnover
- Started work on legal framework late(r) than technical development
- Legal status as consortium (!= single legal entity) complicates situation
- Strong dependency on hosting TRE (vulnerable to changes and higher technical complexity)

# Discussion



- Based on the presentations (RDMkit, DSW, Federated EGA), plenary discussion:
  - ✓ What services are relevant to your Node?
  - ✓ What/who do you need (from others) to implement it?
- Please build your [Node's DM portfolio](#) in the shared notes document, as they are for your use
  - ✓ Formulate your own lessons learned
  - ✓ Formulate your take home messages & next steps



# Lunch

(11-12 CEST / 12-13 EEST)

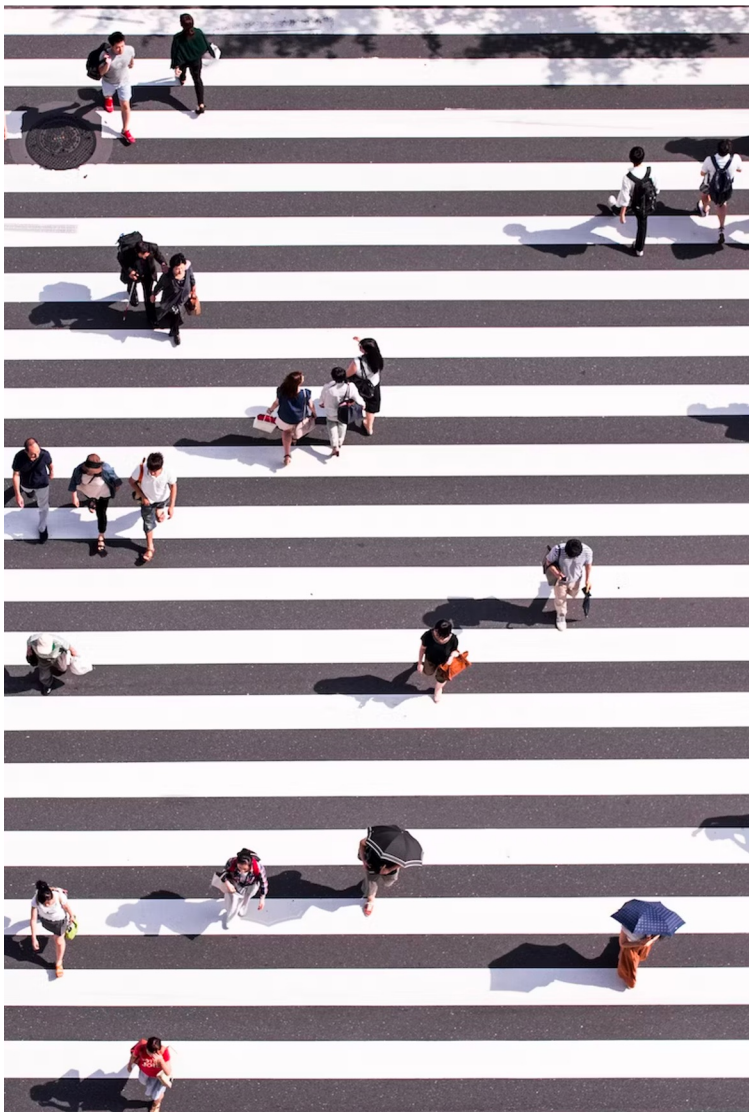


# Data management: examples of ELIXIR tools and services from a Node strategy perspective - **Part 2** (12-13 CEST / 13-14 EEST)





# Nodes' good practices



- Presentations [10 min]
  - ✓ [MOLGENIS](#) (RIR)
    - [NL \[Joeri van der Velde\]](#)
  - ✓ [EDAM](#)
    - [FR \[Hervé Ménager\], with help of NO \[Korbinian Bösl\]](#)
  - ✓ [FAIRsharing](#) (RIR)
    - [UK \[Allyson Lister\]](#)
  - ✓ [FAIR Cookbook](#)
    - [LU \[Danielle\] Welter, with help of CH \[Vassilios Ioannidis\]](#)
- Discussion
  - ✓ What services are relevant to your Node?
  - ✓ What/who do you need (from others) to implement it?

\* RIR = ELIXIR Recommended Interoperability Resource



Molgenis, ELIXIR-NL, Joeri van der Velde



# MOLGENIS [NL, Joeri van der Velde]

The screenshot shows the Molgenis website homepage. At the top left is the Molgenis logo (a blue 'M' in a circle) and the text 'MOLGENIS'. To its right is a green rounded rectangle containing the text 'Web-based scientific data platform'. Further right is a navigation menu with links: 'GET', 'FEATURES', 'COMMUNITIES', 'LEARN', and 'TEAM'. Below the navigation is a horizontal line. On the left side, the text 'FOR SCIENTIFIC DATA' is displayed in large white letters. Below it, the text 'Free. Open source. Simple to use.' is shown. Further down, a paragraph describes Molgenis: 'MOLGENIS: a data platform for researchers to accelerate scientific collaborations and for bioinformaticians who want to make researchers happy.' In the center, a blue rounded rectangle contains the text 'Used for 15+ years' and 'Currently 100+ servers', with 'Free & open source (GNU LGPL 3.0)' below it. To the right of this, an orange rounded rectangle contains the text 'NEW: 'EMX2''. At the bottom left, two URLs are listed: 'https://www.molgenis.org' and 'https://github.com/molgenis'. At the bottom right, the Elixir logo is displayed, featuring a stylized orange and blue 'E' and the text 'elixir Recommended Interoperability Resources'.

- Run it yourself (Java + Postgres), *it's all yours, for free, forever*
- SaaS (we host it for you, €2500 per year or via a partnership)
- Data management services available!

# MOLGENIS: research data platform

## Simple upload formats

**Up/Download for pet store**

**Upload**  
Import and export data (tables) and metadata (schema, settings) in bulk.  
Import data by uploading files in excel, zip, json or yaml format.

**Download**  
Export data by downloading various file formats:  
Export schema as [csv](#) / [json](#) / [yaml](#)  
Export all data as [excel](#) / [csv.zip](#) / [tst](#) / [jsonid](#)

**Data model**

```

tableName,tableExtends,columnName,columnType,key,required,ref
Analyses,,,,,,,,,http://edamontology.org/operation_2945,Bes
Analyses,,id,string,1,TRUE,,,,,,,,,http://purl.allotrope.org/ont
Analyses,,aligner,string,,,,,,,,,http://edamontology.org/oper
Analyses,,analysisDate,date,,,,,,,,,http://purl.obolibrary.org
Analyses,,biosampleId,string,,,,,,,,,http://purl.obolibrary.org
Analyses,,individualId,string,,,,,,,,,http://purl.obolibrary.org
Analyses,,pipelineName,string,,,,,,,,,http://edamontology.org/
Analyses,,pipelineRef,string,,,,,,,,,http://purl.obolibrary.org
Analyses,,runId,string,,,,,,,,,http://purl.obolibrary.org/obo
Analyses,,variantCaller,string,,,,,,,,,http://edamontology.org
    
```

**Data**

```

id,aligner,analysisDate,biosampleId,individualId,pipelineName
A01,bwa-0.7.8,2021-01-21,Sample001,Ind001,Pipeline-panel-000
A02,bwa-0.7.8,2021-02-12,Sample001,Ind001,Pipeline-panel-000
A03,bwa-0.8.0,2022-01-10,Sample001,Ind001,Pipeline-panel-000
A04,bwa-0.9.0,2022-02-11,Sample002,Ind002,Pipeline-panel-000
    
```

## 100% flexible datamodel

**Schema editor: fairdatahub**

columnName	columnType	required	refTable	refLink	refBack	semantics	validation	visibility	description
id	STRING	1				http://purl.allotrope.org/ontologies/result#FAFR_0001979			E.g. Analysis012345
aligner	STRING					http://edamontology.org/operation_0292			E.g. bwa-0.7.8
analysisDate	DATE					http://purl.obolibrary.org/obo/date			E.g. 2021-10-17
biosampleId	STRING					http://purl.obolibrary.org/obo/NCIT_C93400			E.g. S0001
individualId	STRING					http://purl.obolibrary.org/obo/NCIT_C164337			E.g. P0001
pipelineName	STRING					http://edamontology.org/data_2531			E.g. Pipeline-panel-0001
pipelineRef	STRING					http://purl.obolibrary.org/obo/NCIT_C17564			E.g. ...
runId	STRING					https://doi.org/10.48511/workflowhub/workflow/111			E.g. SRR10903401
variantCaller	STRING					http://edamontology.org/operation_3227			E.g. GATK4.0

Add, explore, filter, publish

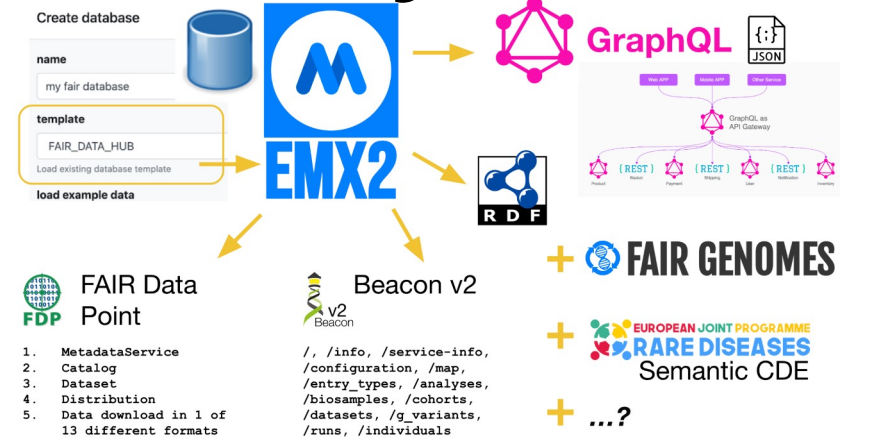
it's semantic

#	id	aligner	analysisDate	biosampleId	individualId	pipelineName	pipelineRef	runId	variantCaller
A01	bwa-0.7.8	2021-01-21	Sample001	Ind001	Pipeline-panel-0001-v1	https://doi.org/10.48511/workflowhub/workflow/111	SRR10903401	GATK4.0	
A02	bwa-0.7.8	2021-02-12	Sample001	Ind001	Pipeline-panel-0001-v2	https://doi.org/10.48511/workflowhub/workflow/111	SRR10903402	GATK3.0	
A03	bwa-0.8.0	2022-01-10	Sample001	Ind001	Pipeline-panel-0001-v1	https://doi.org/10.48511/workflowhub/workflow/111	SRR10903403	GATK3.0	
A04	bwa-0.9.0	2022-02-11	Sample002	Ind002	Pipeline-panel-0001-v2	https://doi.org/10.48511/workflowhub/workflow/111	SRR10903404	GATK4.0	

## Many import/export APIs

- CSV
- Excel
- ZIP
- RDF (7 formats)
- GraphQL
- JSON-LD
- Beacon v2
- FAIR Data Point
- Files
- Tasks
- Python
- YAML

## FAIR-out-of-the-box



# MOLGENIS: What works well?

- Data management made easy and accessible
  - Flexible data model, no mistakes are final
  - Import/export using Excel/CSV/JSON if you want
- Can be extended with rich user interfaces
- New generation of MOLGENIS every ~5 years since 2002
  - Keep the best bits, fresh start with latest tech
  - Migrations are usually painless, import-export, done
- Open source development
  - The real value is not in the code, but in the people!
  - We welcome contributions, extensions, feedback, etc.
- Run MOLGENIS servers on separate virtual machines or k8s
  - Allow customization and fine-tuning of each instance
  - Zero chance of accidental cross-app data leaks
- Blur the border between relational database and triple store
  - Data can be represented differently, no problem
  - A little semantics goes a long way - benefits everyone



<https://fairgenomes.org>



# MOLGENIS: Communities & Best practice templates



- **Examples:** cohort/biobank catalogue federations, patient registries, large genomics projects
- **Challenge:** allow for *heterogeneity* in a complex data landscape *while maintaining interoperability*
- **Our solution:** offer best-practice templates, but you can *always* extend or deviate when needed

MOLGENIS

The image features a world map with various logos of genomics and healthcare organizations overlaid. The logos include:

- EUCAN connect**
- ATHLETE** (The Exposome from: Evidence to Translation)
- LifeCycle**
- longITools** (HEALTH & ENVIRONMENT DYNAMICS)
- VAC4EU**
- conCEPTION** (SAFETY EVIDENCE ECOSYSTEM)
- Immune Mage**
- EUROPEAN MEDICINES AGENCY** (SCIENCE MEDICINES HEALTH)
- MINERVA**
- DCAT** (Bioschemas)
- Global Alliance for Genomics & Health**
- BBMRI-ERIC**
- 1+MG**
- FAIR GENOMES**
- European Genomic Data Infrastructure**
- eosc | cancer**
- CINECA**
- X-omics.nl**
- SolveRD** (Solving the Unsolved Rare Diseases)
- European Reference Networks**
- EOSC-Life**
- health RI** (enabling data driven health)
- VKGL** (vereniging klinisch genetische LABORATORIUMDIAGNOSTIEK)
- isatools**
- Nictiz**
- EUROPEAN OPEN SCIENCE CLOUD**
- EUROPEAN JOINT PROGRAMME RARE DISEASES**

Only together we can define how to find, share and reuse \*omics data in research and healthcare

Contact: [k.j.van.der.velde@umcg.nl](mailto:k.j.van.der.velde@umcg.nl)  
Support: [molgenis-support@umcg.nl](mailto:molgenis-support@umcg.nl)



umcg



university of  
 groningen



ELITMa


# EDAM, ELIXIR-NO and ELIXIR-FR, Hervé Ménager



# EDAM: a standard to improve findability of resources

The screenshot shows a web browser window displaying the BioCatalogue website. The address bar shows the URL: <https://www.biocatalogue.org/search?utf8=✓&q=sequence+similarity+search+protein>. The BioCatalogue logo is at the top left, with the tagline "The Life Science Web Services Registry". Navigation links include "Getting Started", "About Us", "Contact Us", and "API Docs". A search bar contains the text "sequence similarity search prote" and a "Go!" button. Below the search bar are links for "Home", "Services", "Register a Service", and "Service Providers". A "Search by Data" link is highlighted, and a "Latest" link is also visible. A "SHARE" button with social media icons and an "EMAIL" button are on the right. The main content area displays "Search Results" and a message: "Try our new Search by Data feature - to search using your input/output data". Below this, it states: "Search query 'sequence similarity search protein domains' returned 0 items". The footer contains copyright information: "Copyright (c) 2008 - 2014 The University of Manchester and the European Bioinformatics Institute (EMBL-EBI)".

biocatalogue.org <https://www.biocatalogue.org/search?utf8=✓&q=sequence+similarity+search+protein> W - Wikipedia (en)

**BioCatalogue**   
"The Life Science Web Services Registry"

[Getting Started](#) | [About Us](#) | [Contact Us](#) | [API Docs](#)

[Sign up](#) | [Sign in](#)

Search:  [Go!](#) | [Home](#) | [Services](#) | [Register a Service](#) | [Service Providers](#) | [RSS](#) | [Twitter](#)

[Search by Data](#) | [Latest](#)

[Home](#) » [Search](#) » All » Results for 'sequence similarity search protein domains'

[SHARE](#) [EMAIL](#)

## Search Results

Try our new [Search by Data](#) feature - to search using your input/output data

Search query "sequence similarity search protein domains" returned 0 items

Copyright (c) 2008 - 2014 [The University of Manchester](#) and the [European Bioinformatics Institute \(EMBL-EBI\)](#)



# What is EDAM

## Topic

Paleogenomics  
Electron cryotomography

## Operation

Molecular docking  
Image segmentation

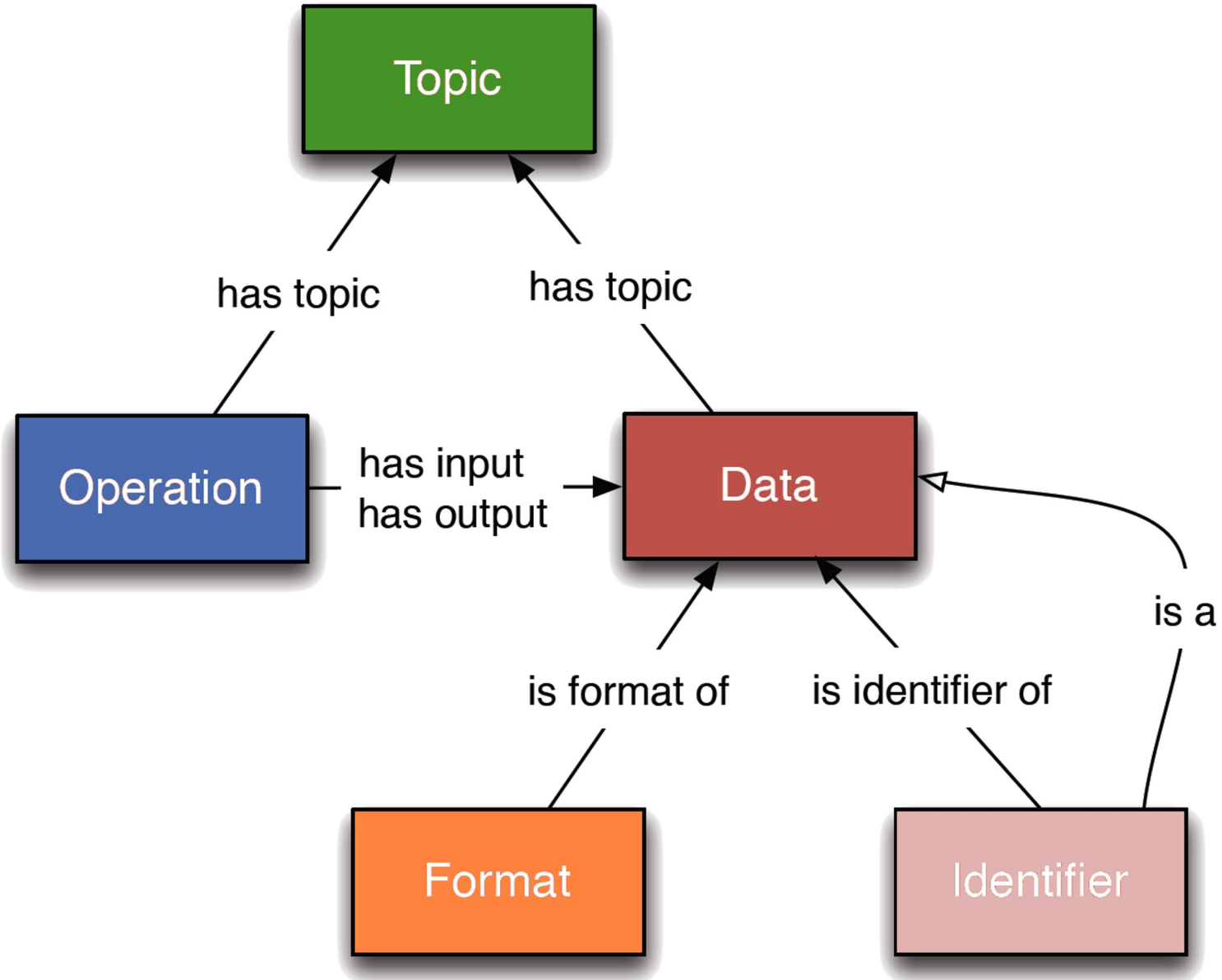
## Data

Sequence assembly  
Gene ID

## Format

FASTQ  
SBML

# What is EDAM



# When and how can you use EDAM



Search bio.tools

25402 tools

Explore ▾

Menu ▾

llamothe ▾

## MacSyFinder (biotools:macsyfinder)

👍 ID Verified

<https://github.com/gem-pasteur/macsyfinder>

Available versions

1.0.2

Functional genomics >

Mature

GPL-3.0

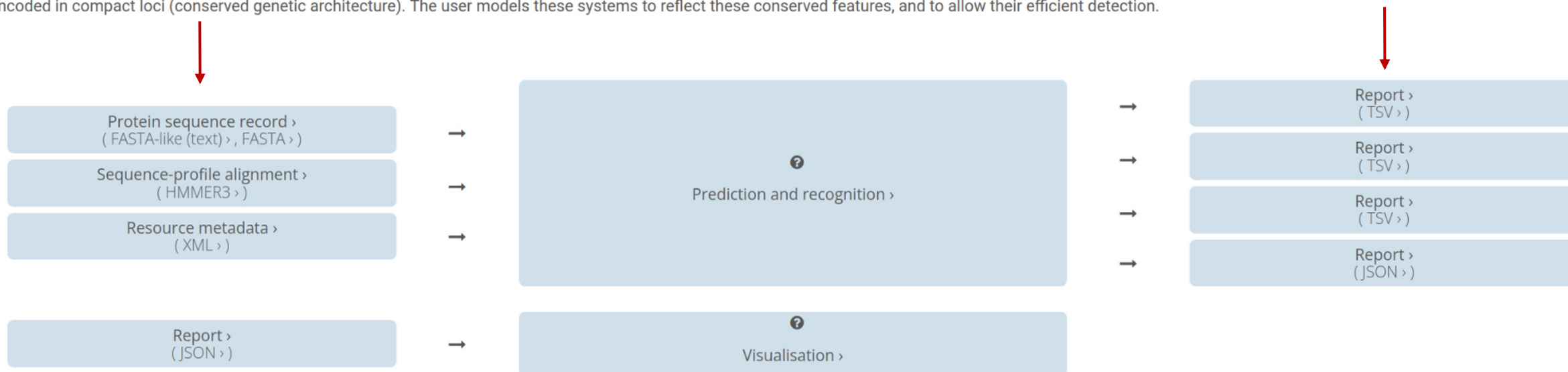
Command-line tool

Web application

Python



Program to model and detect macromolecular systems, genetic pathways... in protein datasets. In prokaryotes, these systems have often evolutionarily conserved properties: they are made of conserved components, and are encoded in compact loci (conserved genetic architecture). The user models these systems to reflect these conserved features, and to allow their efficient detection.



# When and how can you use EDAM

[Events](#)[Materials](#)[e-Learning](#)[Workflows](#)[Providers](#)[About](#)[Home](#) / [Materials](#)

## Date Added

Anytime ▼

## Scientific topic

RNA-Seq 53

Data management 42

Rare diseases 36

Statistics and probability 20

Bioinformatics 17

Genomics 16


Proteomics 16

Phylogenetics 13


Pathway or network 11

**Show more scientific topics** ▼

## Operation

 Register training material

Subscribe ▼

 What are materials in TeSS?



 Most recent ▼

**1513 materials found**

[← Previous](#) [1](#) [2](#) [...](#) [8](#) [9](#) [10](#) [11](#) **[12](#)** [13](#) [14](#) [15](#) [16](#) [...](#) [50](#) [51](#) [Next →](#)

### Analysis of RNA-Seq Data

Bioconductor provides tools for the analysis and comprehension of high-throughput genomic data. Bioconductor uses the R statistical programming language, and is open source and open development. It has two releases each year, 1560 software packages, and an...

**Scientific topics:** RNA-Seq




### eQTL / molecular-QTL analyses

Bioconductor provides tools for the analysis and comprehension of high-throughput genomic data. Bioconductor uses the R statistical programming language, and is open source and open development. It has two releases each year, 1560 software packages, and an...

# When and how can you use EDAM

HP



## Human Phenotype Ontology


The Human Phenotype Ontology has been developed to provide a structured and controlled vocabulary for the phenotypic features encountered in human hereditary and other disease. The goal is to provide resource for t...

Genetics Phenomics Phenotype Disease Homo sapi... +2 more tags

---

- ▶ Related Standards
- ▶ Implementing Databases
- ▶ Endorsing Policies

15	<b>Type:</b> subjects
	<b>Name:</b> Phenomics
19	<b>URL:</b> <a href="http://edamontology.org/topic_3298">http://edamontology.org/topic_3298</a>
1	<b>Definition:</b> Phenomes, or the study of the change in phenotype (the physical and biochemical traits of organisms) in response to genetic and environmental factors.



<https://fairsharing.org>

# When and how can you use EDAM

**Galaxy Europe**

[EU infrastructure] 30.05.2022 - There is a performance issue

**Tools**

search to  Full Tool Panel

... by Ontology

- EDAM Operations
- EDAM Topics

**BIOLOGY**

Microbiology

Transcriptomics

Gene transcripts

Structural variation

Genetic variation

Epigenetics

Metagenomics

Microbial ecology

Biodiversity

Ecology

**Expasy**  
Swiss Bioinformatics Resource Portal

Home About SIB News Contact

SIB Swiss Institute of Bioinformatics

e.g. BLAST, UniProt, MSH6, Albumin...

**SIB Resources**

- Genes & Genomes**
  - Genomics
  - Metagenomics
  - Transcriptomics
- Proteins & Proteomes**
- Evolution & Phylogeny**
  - Evolution biology
  - Population genetics
- Structural Biology**
  - Drug design
  - Medicinal chemistry
  - Structural analysis
- Systems Biology**
  - Glycomics
  - Lipidomics
  - Metabolomics

**SwissRegulon Portal**  
Tools and data for regulatory genomics

**SWISS-MODEL**  
Protein structure homology-modelling

**V-pipe**  
Viral genomics pipeline

**SwissDrugDesign**  
Widening access to computer-aided drug design

**EPD**  
Eukaryotic Promoter Database

**Bgee**  
Gene expression expertise

**UniProtKB/Swiss-Prot**  
Protein knowledgebase

**STRING**  
Protein-protein interaction networks and enrichment analysis

<https://expasy.org>

**Biii**  
Biolmage Informatics Index

Advanced search (under construction)

**All Content**

Displaying 1 - 20 of 41. Total number of pages: 3

Content type: Software For which operation? ---Filament tracing X

for which topic? Choose some values (one or more) Supported image dimension Choose some values (one or more) Type -Any-

Interaction level Choose some values (one or more) License/Openness Choose some values (one or more) Programming language Choose some values (one or more)

specific biological term Choose some values (one or more)

Apply

listing content

Name	Category	Image	requires	Type	Interaction level	License/Openness	has function	has topic
Alivia - Alivia Cloud - Alivia Web	Software			Collection, Workflow	Automated, Manual, Semi-automated	Commercial	Analysis, Frequency analysis, Image analysis, Colocalisation analysis, Object-based colocalisation, Pixel-based colocalisation, Filament tracing, Image feature detection, Image segmentation, Cell segmentation, Pixel classification, Neuron image analysis, Object classification, Object counting, Texture extraction, Object tracking, Cell tracking, Particle tracking, Annotation, Classification, Editing, Image deconvolution, Visualisation, Image visualisation, Volume rendering, Plotting	Machine learning
AnalyzeSkeleton	Software		ImageJ, Analyze 12.0	Component	Automated, Semi-automated		Filament tracing, Object detection, Pixel classification	
AnaMorf	Software		ImageJ	Component	Automated	Free and open source	Filament tracing	
Angiogenesis / Sprout Analyzer (ImageJ)	Software		ImageJ	Component	Semi-automated	Free and open source	Filament tracing, Neuron image analysis	Biolmage informatics, Fluorescence microscopy
APP (All-path pruning)	Software		Vaa3D	Component	Automated	Free and open source	Filament tracing, Image thresholding, Neuron image analysis	Confocal microscopy

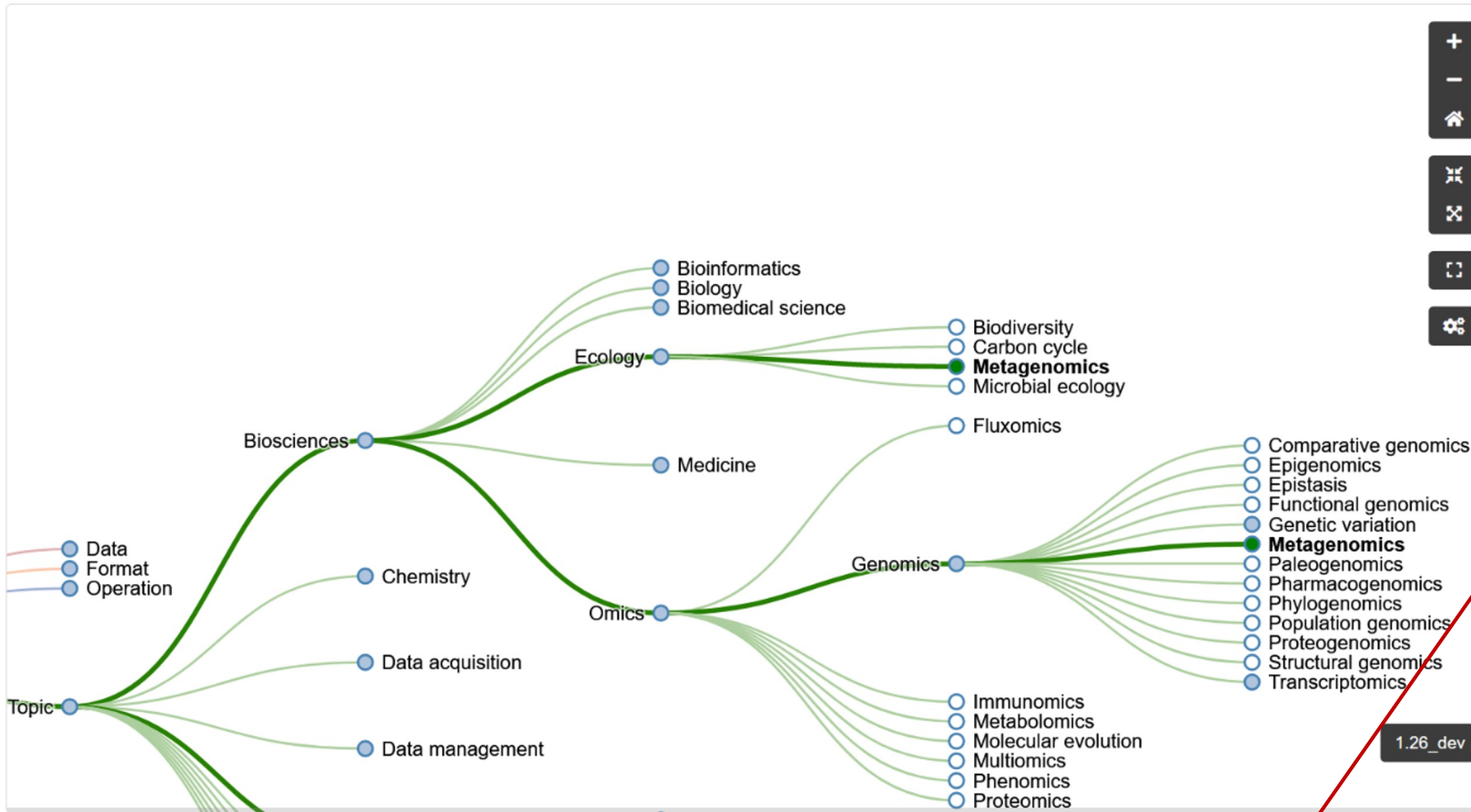
<https://biii.eu>

<https://usegalaxy.eu>

# EDAM Browser: visualize the ontology

Custom Version ▾ All EDAM ▾

Metagenomics 🔍



Details of term "Metagenomics" + ✎

<b>Term</b>	Metagenomics
<b>Definition</b>	The study of genetic material recovered from environmental samples, and associated environmental data.
<b>Comment</b>	
<b>Exact synonyms</b>	
<b>Narrow synonyms</b>	Shotgun metagenomics
<b>URI</b>	<a href="#">topic_3174</a> 📄
<b>Parents</b>	<a href="#">Ecology</a> 📄 <a href="#">Genomics</a> 📄
<b>Community usage</b>	
<b>bio.tools</b>	715 times
<b>Biosphere</b>	2 times by appliances, 7 times by tools.
<b>TeSS</b>	9 times
<b>Links</b>	Open in <a href="#">AberOWL</a> , <a href="#">BioPortal</a> , <a href="#">OLS</a> or <a href="#">WebProtégé</a> .

- Previously seen:
- Details of term "Ecology" + ✎
  - Details of term "Environmental sciences" + ✎
  - Details of term "Data" + ✎

# Contribute to EDAM!

Metagenomics

### Details of term "Metagenomics"

<b>Term</b>	Metagenomics
<b>Definition</b>	The study of genetic material recovered from environmental samples, and associated environmental data.
<b>Comment</b>	
<b>Exact synonyms</b>	
<b>Narrow synonyms</b>	Biome sequencing Shotgun metagenomics
<b>URI</b>	<a href="#">topic_3174</a>
<b>Parents</b>	<a href="#">Genomics</a> <a href="#">Ecology</a>
<b>Community usage</b>	
<b>bio.tools</b>	714 times
<b>Biosphere</b>	2 times by appliances, 7 times by tools.
<b>TeSS</b>	9 times
<b>Links</b>	Open in <a href="#">AberOWL</a> , <a href="#">BioPortal</a> , <a href="#">OLS</a> or <a href="#">WebProtégé</a> .

Previously seen:

- Details of term "Operation"
- Details of term "Sequence clustering"
- Details of term "Clustering"
- Details of term "Sequence annotation"

## Propose a new term in "topic" of EDAM ontology

Label:

Definition:

Parent:

Exact synonyms:

Narrow synonyms:

Issue comments:

[Continue to GitHub](#)

Search or jump to...

edamontology / edamontology

### New operation "Disease transmission analysis" #760

**Operation name:** < Disease transmission analysis >  
**Short description:** < The analysis of the transmission of disease in a population >  
**Parent URI(s):** <http://edamontology.org/operation\_2945 >  
**Has Input:** <>  
**Has Output:** <http://edamontology.org/data\_2048 >  
**Has Topic:** <http://edamontology.org/topic\_3324 | http://edamontology.org/topic\_3305 >  
**exactSynonym(s):** <>  
**Wikipedia:** <>  
**Comment:** <>

**Assignees:** No one—assign yourself  
**Labels:** [conceptterm addition](#)  
**Projects:** None yet  
**Milestone:** No milestone  
**Development:** Successfully merging a pull request may close this issue.  
**Add terms (December 2021)** Kigaard/edamontology  
**Notifications:** Unsubscribe  
**Participants:** 2 participants



# Contribute to EDAM!

The screenshot shows the EDAM live dev version web interface. The top navigation bar includes "EDAM live dev version (currently 1.26\_dev, 04.08.2021 21:02 UTC, 4f7cbbfd main)", "Home", "Display", "Project", "Herve Manager", and "Help". Below the navigation bar, there are tabs for "Classes", "Properties", "Individuals", "Comments", "Changes by Entity", and "History". The main content area is divided into two panels. The left panel, titled "Class Hierarchy", shows a tree view of classes under "Owl:Thing". The "ACE" class is highlighted in blue. The right panel, titled "Class: ACE", shows the details for the "ACE" class, including its IRI, annotations, parents, and relationships. The "Annotations" section lists several annotations, including "rdf:type", "hasDbXref", "Created in", "Documentation", "hasDefinition", "isSubset", and "isSubset". The "Parents" section lists "textual format" and "Sequence assembly format". The "Relationships" section is currently empty.

The screenshot shows the GitHub repository page for "edamontology/edamontology". The top navigation bar includes "Code", "Issues" (195), "Pull requests" (33), "Discussions", "Actions", "Projects", "Wiki", "Security", "Insights", and "Settings". The "Issues" section is active, showing a list of issues. The issues are filtered by "is:issue is:open". The list of issues includes:

- #864: New format: MSP (concept/term addition) - opened 5 days ago by xtrojak
- #862: Requesting assistance on behalf of Human Cell Atlas Ontology - opened on May 24 by bvamer-ebi
- #861: Dereferencing EDAM classes is not working properly (access | publication | release | bug | partially fixed) - opened on May 10 by albangainard
- #860: Improve citation info (access | publication | release | documentation) - opened on Apr 19 by matuskalas
- #859: Consider a GitHub Action for testing integration into TeSS (access | publication | release | validation) - opened on Mar 9 by matuskalas
- #854: [BUG] (a priori) Erroneous hierarchy for "qualsolid" [format:3610] (bug | high priority) - opened on Oct 6, 2022 by M-casado
- #851: Obsolete Infectious tropical disease has wrong 'replacedBy' (bug | good first issue | polishing) - opened on Sep 21, 2022 by matuskalas
- #850: Add metadata spreadsheet under resource metadata - opened on Sep 12, 2022 by FSapenaVentura

# FAIRsharing, ELIXIR-UK, Allyson Lister





# FAIRsharing.org

## Integrating FAIRsharing into ELIXIR Node strategies for data management

Allyson Lister, Content & Community Lead

*with thanks to*

Susanna    one, PI

[datareadiness.eng.ox.ac.uk](http://datareadiness.eng.ox.ac.uk)

Allyson's RDA/EOSC Future Domain Ambassadorship supported by

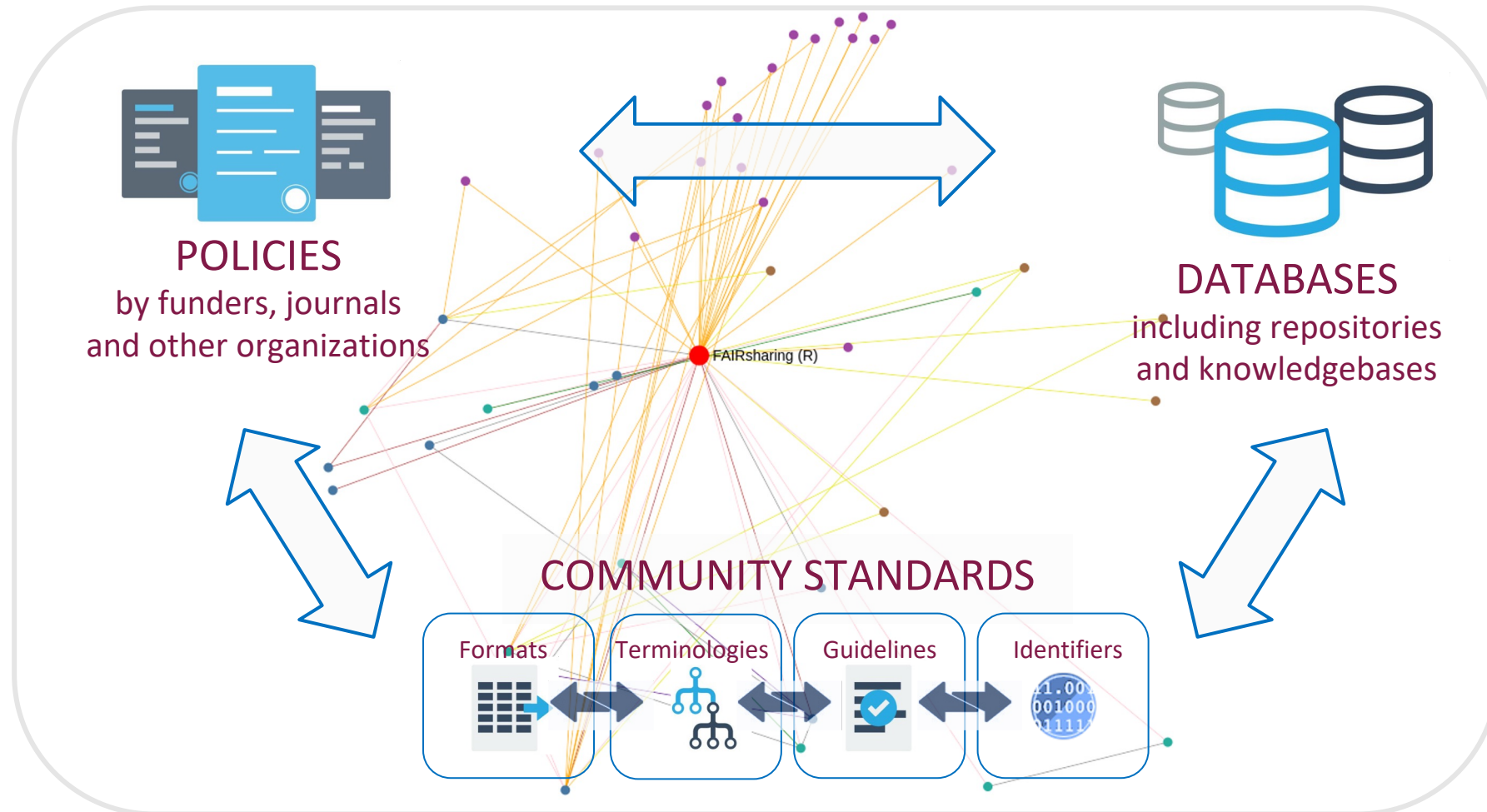


Supported by the Research Data Alliance through the RDA Open Calls as part of the EOSC Future project.

The project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 101017536.



# An informative and educational resource, and a service



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

# FAIRsharing in numbers: content, contributions, coverage

**1282**

**Contributors**

Our community of record maintainers and curators

**3983**

**Records**

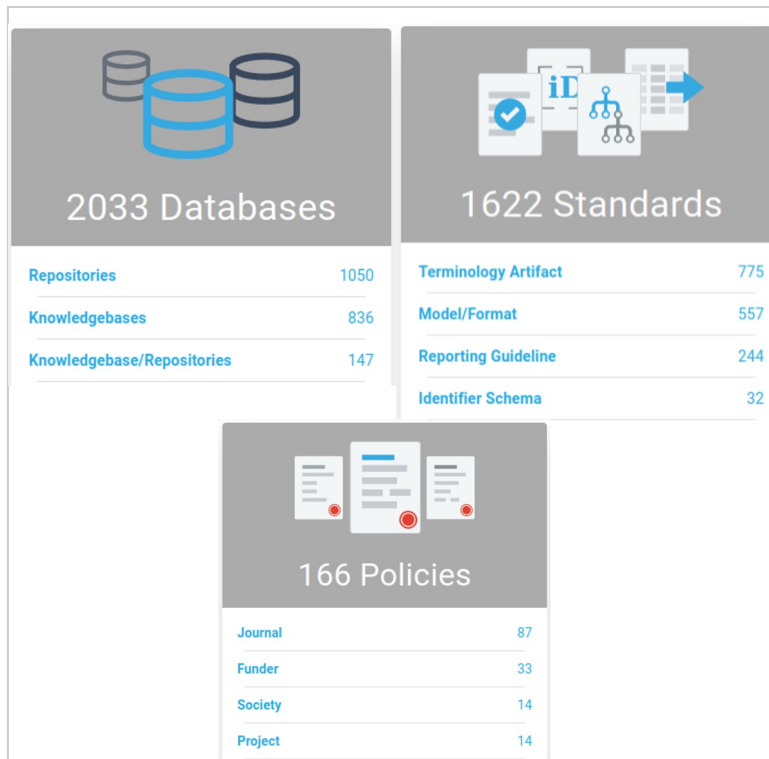
Descriptions of standards, databases and policies

**198K**

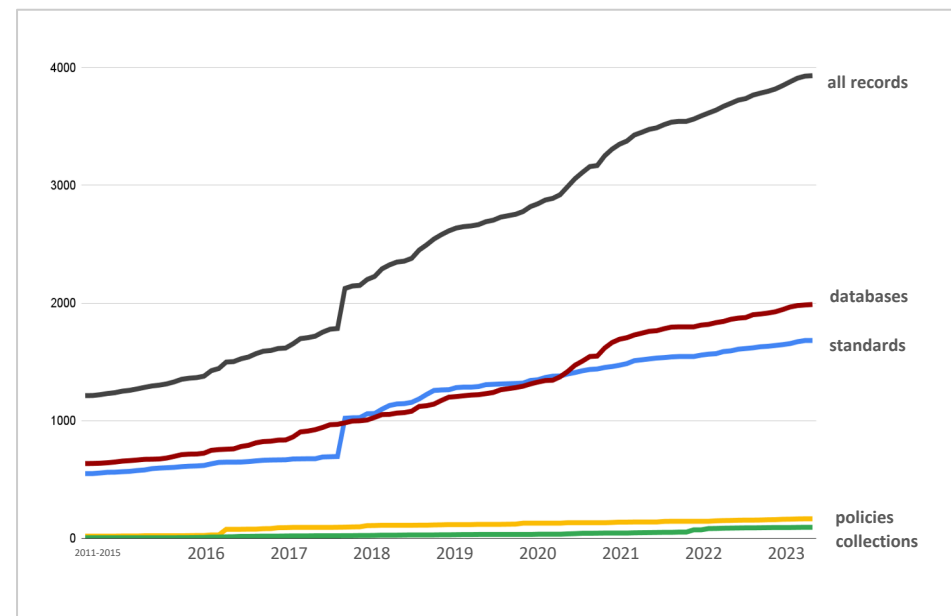
**Views**

Page visits since 2015

## Types and subtypes



## Growth since launch



as of June 2023

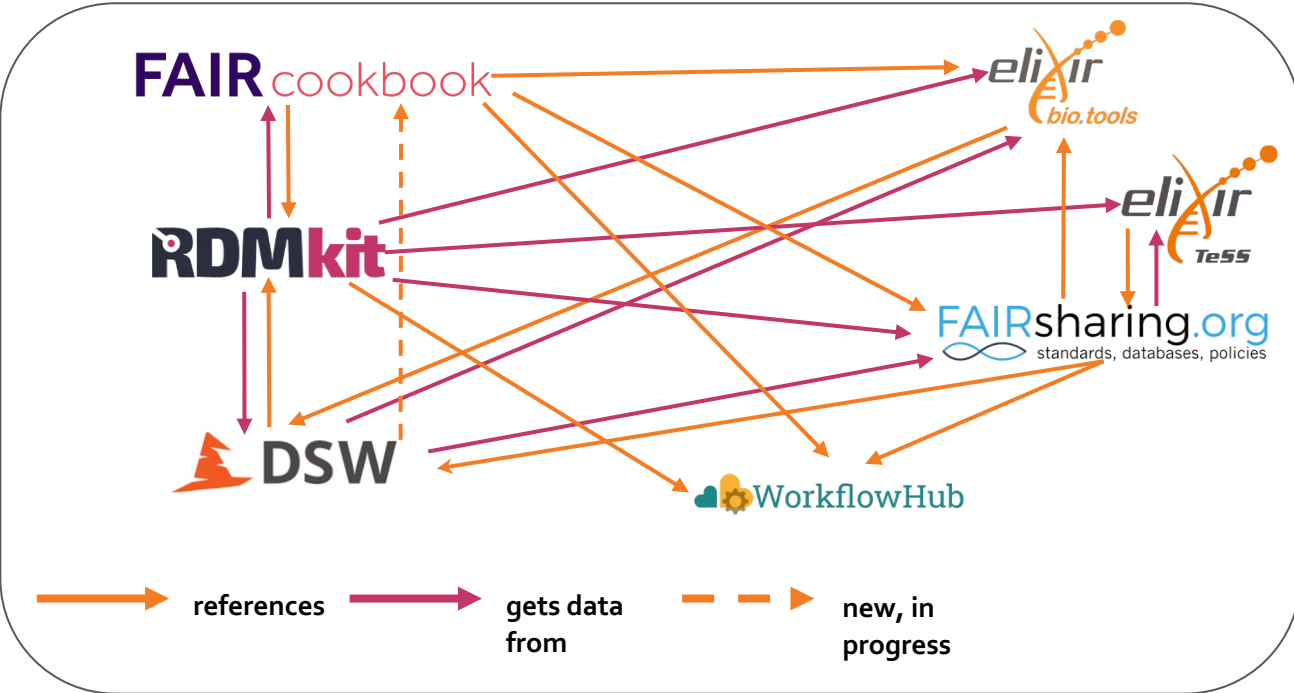
## Disciplines



# FAIRsharing connects your resource to dozens of services within ELIXIR and beyond

A growing number of tools and services access the FAIRsharing API, and use it for **look-up, selection and content retrieval for standards and repositories** in:

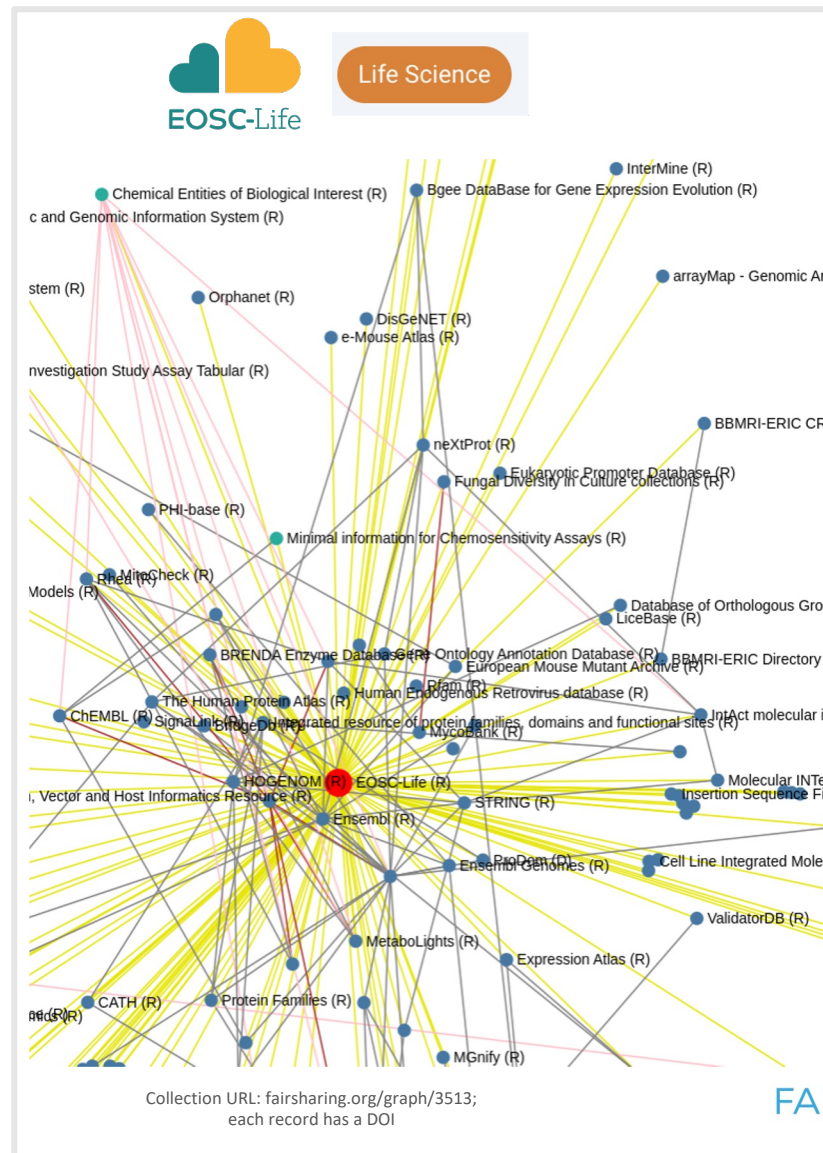
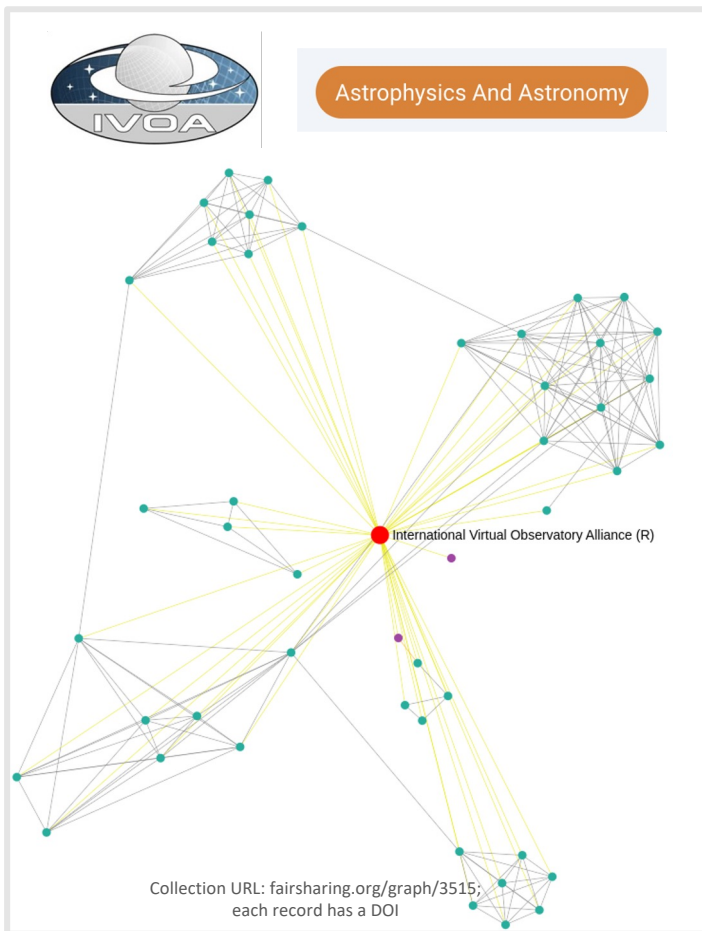
- creation of data management plans
- enrichment of guidance and training material
- assessment of FAIRness



	FAIR Metrics and Data Quality Task Force	<a href="#">ORCID</a> Wilkinson, Mark D; <a href="#">ORCID</a> Sansone, Susanna-Assunta; <a href="#">ORCID</a> Grootveld Marjan; <a href="#">ORCID</a> Nordling, Josefine; <a href="#">ORCID</a> Dennis, Richard; <a href="#">ORCID</a> Hecker, David
	<b>FAIR Assessment Tools: Towards an “Apples to Apples” Comparisons</b>	Hosting institution(s) <a href="#">ORCID</a> Schultes, Erik
		Project member(s) <a href="#">ORCID</a> Goble, Carole; <a href="#">ORCID</a> Lyster, Allyson; <a href="#">ORCID</a> Thurston, Milo; <a href="#">ORCID</a> Maayan, Avi

# FAIRsharing makes your standard, database or policy discoverable to an entire ecosystem of resources

- Deprecates
- Recommends
- Collects
- Profiles
- Accepts
- Outputs
- Shares code with
- Shares data with
- Implements
- Extends
- Related to



The **standards**, **repositories** and **policies** each EOSC Cluster uses or endorses

## Registry

- COLLECTION
- DATABASE
- STANDARD
- POLICY

Status (shown on mouseover)

- READY
- IN DEVELOPMENT
- UNCERTAIN
- DEPRECATED

Distance from centre

- ONE HOP
- TWO HOPS
- THREE HOPS

# FAIRsharing benefits the databases, standards and policies registered within it

The **registration** of **databases** and/or **standards**




- increases their **discoverability**, and
- provides them (the maintainers of the records) with visibility and credit.

The more **connected** the resource, the greater its **visibility**

- if a **standard** is implemented by a **database**, these two records will be interlinked
- these links can be navigated, allowing new relationships to be explored and collaborations to be formed

<https://doi.org/10.25504/FAIRsharing.zv11j3>  
<https://www.gbif.org/news/6P2o4iu4iF3AXqQvUUusuzZ/gbif-staff-join-the-fairsharing-community-champions-programme>


This record replaces or incorporates the following deprecated resources:  
[Canadian Biodiversity Information Facility](#)



  GBIF.org  
 [10.25504/FAIRsharing.zv11j3](https://doi.org/10.25504/FAIRsharing.zv11j3)

**Type** Repository  
**Registry** Database  
**Description** GBIF—the Global Biodiversity Information Facility—is funded by the world's governments and is the world's largest biodiversity database. It provides access to all types of life on Earth datasets through a discoverable, accessible, interoperable and citable, assigning each record a unique identifier and linking it to peer-reviewed research. GBIF's mission is to increase the visibility and impacts of climate change research and to share data. To deposit or host data, contact your local node or publishers with a host.


**Homepage** <https://www.gbif.org>

**Year of Creation** 2001

**Maintainers** [gbifs](mailto:gbifs), [kcopas@gbif.org](mailto:kcopas@gbif.org) 

The following community curators have contributed to this record: [kcopas@gbif.org](mailto:kcopas@gbif.org) , [jmiller@gbif.org](mailto:jmiller@gbif.org) 

NEWS | 20 MARCH 2023  
**GBIF staff join the FAIRsharing Community Champions Programme**  
*BiCIKL guides development of FAIRsharing-powered web app enabling EU-fund partners to share their collective knowledge on biodiversity-related standards, data and policies*



The BiCIKL web app reveals links between organizations that implement FAIR data standards, databases and policies in FAIRsharing.

This news item also appears on the [FAIRsharing blog](#)

Members of the GBIF Secretariat joined staff from the Pacific Northwest National Laboratory last year as early adopters of the [FAIRsharing Community Champions Programme](#).



# FAIRsharing for all

## FAIRsharing in a nutshell

FAIRsharing.org

help in selecting the right resources to fulfil data management

### FAIRsharing for all: user groups

#### JOURNAL PUBLISHERS

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



#### LIBRARIANS/TRAINERS

Use FAIRsharing to provide a foundation on which to create or enrich educational lectures, training and teaching material, and to plug into data management planning tools

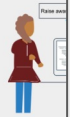


#### FUNDERS

Recommend FAIRsharing to your awardees to inform the development of their data management plan, and select the appropriate resources to recommend in your data policies

## FAIR

JOURNAL  
Create and  
citable s  
repositories  
users or th  
recommen



SOCIETIE  
Raise awa  
databases, r  
as well as r  
take action  
use and cita

\*datasets, soft

FAIR  
CC BY-SA



#### SOCIETIES AND ALLIANCES

Raise awareness around standards, databases, repositories and data policies, as well as mobilise your community to take action to promote the registration, use and citation of key resources

#### DEVELOPERS AND CURATORS

Increase the discoverability and adoption of your resource by adding it in FAIRsharing; if you are a tool developer, access FAIRsharing content via its API to power your system



#### RESEARCHERS

Identify and cite the standards, databases or repositories that exist for your discipline when creating a data management plan, releasing data or submitting a manuscript to a journal

# ELIXIR in FAIRsharing

DATABASES   POLICIES   COLLECTIONS   **ORGANISATIONS**   ADD CONTENT

---

### Organisations List

Name	Types	Homepage	ROR
ELIXIR Czech Republic	Government body	<a href="https://www.elixir-czech.cz/">https://www.elixir-czech.cz/</a>	
ELIXIR Europe	Government body	<a href="https://elixir-europe.org">https://elixir-europe.org</a>	<a href="https://ror.org/044rwnt51">https://ror.org/044rwnt51</a>
ELIXIR-EXCELERATE programme, ELIXIR-HUB, UK	Government body	<a href="https://www.elixir-europe.org/about-us/how-funded/eu-projects/excelerate">https://www.elixir-europe.org/about-us/how-funded/eu-projects/excelerate</a>	
ELIXIR-UK	Government body	<a href="https://elixiruknode.org/">https://elixiruknode.org/</a>	
ELIXIR Italy	Government body	<a href="http://elixir-italy.org">http://elixir-italy.org</a>	
ELIXIR Machine Learning Focus Group	Consortium	<a href="https://elixir-europe.org/focus-groups/machine-learning">https://elixir-europe.org/focus-groups/machine-learning</a>	
ELIXIR Norway	Government body	<a href="https://elixir.no/">https://elixir.no/</a>	
Horizon 2020 Programme of the European Union - EXCELERATE	Government body	<a href="https://elixir-europe.org/about-us/how-funded/eu-projects/excelerate">https://elixir-europe.org/about-us/how-funded/eu-projects/excelerate</a>	
Spanish Instituto de Salud Carlos III Instituto Nacional de Bioinformatica	Government body	<a href="https://inb-elixir.es/">https://inb-elixir.es/</a>	
ELIXIR CONVERGE	Consortium	<a href="https://elixir-europe.org/about-us/how-funded/eu-projects/converge">https://elixir-europe.org/about-us/how-funded/eu-projects/converge</a>	

Rows per page: 10 1-10 of 11 < >

# ELIXIR in FAIRsharing

## General Information

**Homepage:** <https://elixiruknode.org/>

**Alternative Names:**

**Types:** Government body































**This organisation is a member of:**



**This organisation has the following members:**

**Associated countries:**  
United Kingdom

## Records related to this organisation

[FILTER RECORDS](#)

Name	Status	Relation	Grant	Actions
 Integrated resource of protein families, domains and functional sites		Funds		
 The European Genome-phenome Archive		Maintains		
 Bioschemas Dataset Profile		Maintains		
 Investigation Study Assay JSON		Maintains		
 Bioschemas Protein Profile		Maintains		
 Bioschemas FormalParameter Profile		Maintains		
 Bioschemas Gene Profile		Maintains		
 Bioschemas Taxon Profile		Maintains		
 Identifiers.org Central Registry		Maintains		
 Genetic and Genomic Information System		Maintains		

1-10 of 19  

## Users affiliated with this organisation

Username	Email address	ORCID ID	Twitter
No data available			

<https://fairsharing.org/organisations/843>

# ELIXIR in FAIRsharing

## ELIXIRUK



### ELIXIR UK node resources

The UK Node supports the bioinformatics and broader biological research community by providing training and resources that help researchers to find and share data, exchange expertise, and agree ...

Bioinforma... All +2 more tags

- ▶ Related Standards 1
- ▶ Related Databases 10
- ▶ Related Policies 0

## ELIXIR RIRs



### ELIXIR Recommended Interoperability Resources

The ELIXIR Recommended Interoperability Resources (RIRs) are tools and registries that have been selected by a panel of external reviewers. These resources build a network of foundational componen...

Data Integr... FAIR All +2 more tags

- ▶ Related Standards 3
- ▶ Related Databases 9
- ▶ Related Policies 0

## ELIXIR Core Data Resources and Deposition Databa...



### ELIXIR Core Data Resources and Deposition Databases


ELIXIR unites Europe's leading life science organisations in managing and safeguarding the increasing volume of data being generated by publicly funded research. It coordinates, integrates and sustai...

Environme... Literature ... All +5 more tags

- ▶ Related Standards 1
- ▶ Related Databases 25
- ▶ Related Policies 0

# ELIXIR in FAIRsharing

**ELIXIRUK**



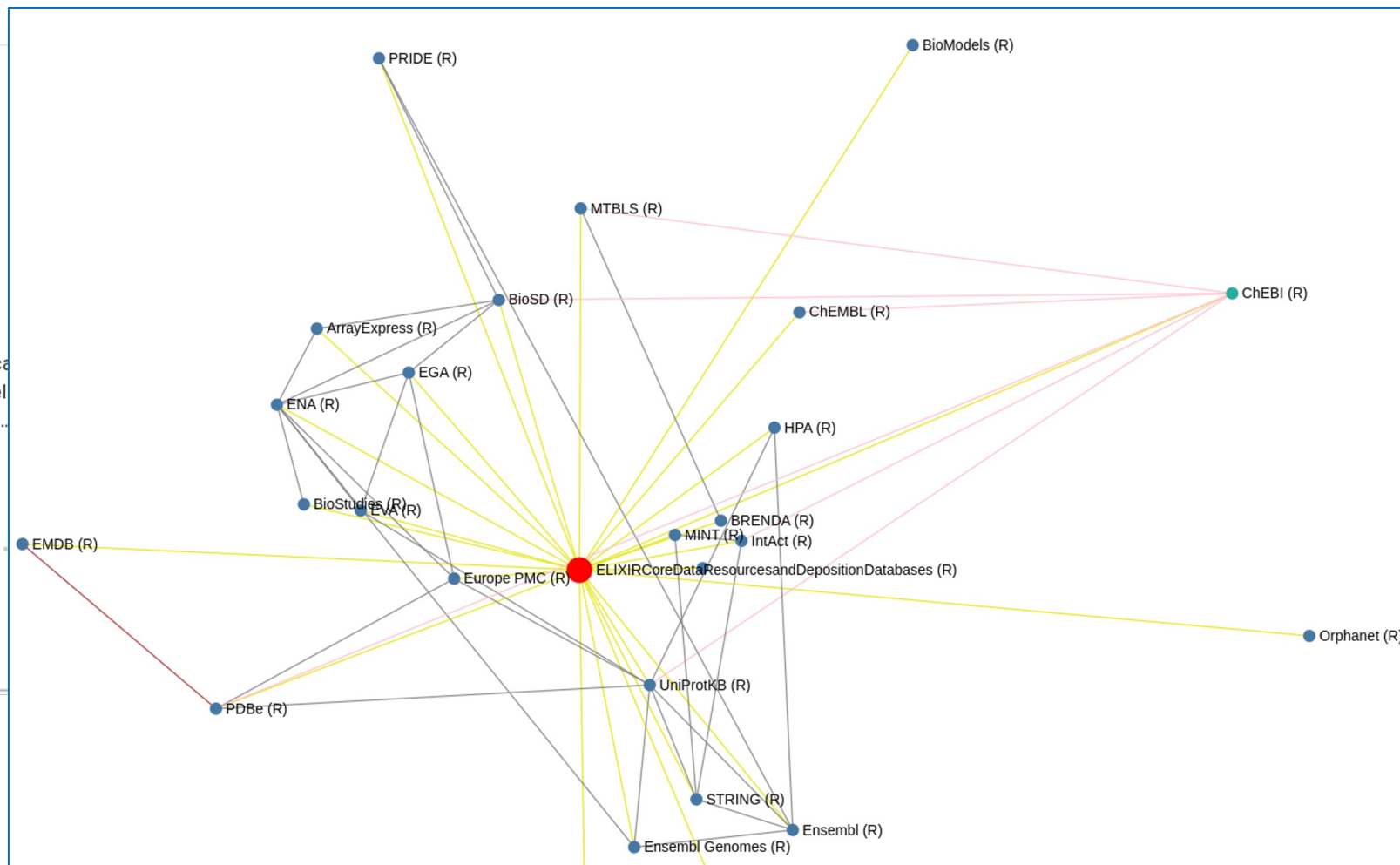
ELIXIR UK node resources

The UK Node supports the bioinformatics and broader biological research community by providing training and resources that help researchers to find and share data, exchange expertise, and agree ...

Bioinforma... All +2 more tags

---

▶ Related Standards 1  
▶ Related Databases 10  
▶ Related Policies 0



ELIXIR CDRs: <https://fairsharing.org/graph/3527>

# How can your Node integrate with FAIRsharing?

1. **Policy and Strategy:** Development, implementation and monitoring of the research data management policy and strategy of the institute
2. **Compliance:** Compliance with relevant codes of conduct, legislation and field specific standards
3. **FAIR data:** Alignment to the FAIR data principles and the principles of Open Science
4. **Services:** Availability of adequate support on research data management, in staff or services
5. **Infrastructure:** Availability of adequate infrastructure for research data management
6. **Knowledge management:** Adequate level of knowledge and skills on research data management in the institute
7. **Network:** Obtaining and maintaining a network of aligned expertise areas and relevant organisations by the institute
8. **Data archiving:** Adequate support and infrastructure for FAIR and long-term archiving of the data of the institute

# How can your Node integrate with FAIRsharing?

## Nodes require...

## FAIRsharing.org helps you via...

Visibility for the <b>resources</b> they produce and endorse	<b>Collections</b> of Node resources to provide <b>live representations</b> of these resources
Visibility for the <b>organisations</b> responsible for them (incl the Nodes themselves)	<b>Organisation-specific FAIRsharing pages</b> that list all <b>relationships</b> e.g. all resources associated with your Node
Integration of Node <b>resources</b> with the ELIXIR ecosystem of tools	<b>API</b> access used by many <b>tools</b> including those for FAIR assessment and data management planning
Ways to discover <b>resources</b> of use to Node members	<b>Search</b> functionality to discover resources for collaboration and use

# How can your Node integrate with FAIRsharing?

## Nodes require...

## FAIRsharing.org helps you via...

Visibility for the <b>resources</b> they produce and endorse	<b>Collections</b> of Node resources to provide <b>live representations</b> of these resources
Visibility for the <b>organisations</b> responsible for them (incl the Nodes themselves)	<b>Organisation-specific FAIRsharing pages</b> that list all <b>relationships</b> e.g. all resources associated with your Node
Integration of Node <b>resources</b> with the ELIXIR ecosystem of tools	<b>API</b> access used by many <b>tools</b> including those for FAIR assessment and data management planning
Ways to discover <b>resources</b> of use to Node members	<b>Search</b> functionality to discover resources for collaboration and use



# How can your Node integrate with FAIRsharing?

## Nodes require...

## FAIRsharing.org helps you via...

Visibility for the <b>resources</b> they produce and endorse	<b>Collections</b> of Node resources to provide <b>live representations</b> of these resources
Visibility for the <b>organisations</b> responsible for them (incl the Nodes themselves)	<b>Organisation-specific FAIRsharing pages</b> that list all <b>relationships</b> e.g. all resources associated with your Node
Integration of Node <b>resources</b> with the ELIXIR ecosystem of tools	<b>API</b> access used by many <b>tools</b> including those for FAIR assessment and data management planning
Ways to discover <b>resources</b> of use to Node members	<b>Search</b> functionality to discover resources for collaboration and use

# How can your Node integrate with FAIRsharing?

## Nodes require...

## FAIRsharing.org helps you via...

Visibility for the <b>resources</b> they produce and endorse	<b>Collections</b> of Node resources to provide <b>live representations</b> of these resources
Visibility for the <b>organisations</b> responsible for them (incl the Nodes themselves)	<b>Organisation-specific FAIRsharing pages</b> that list all <b>relationships</b> e.g. all resources associated with your Node
Integration of Node <b>resources</b> with the ELIXIR ecosystem of tools	<b>API</b> access used by many <b>tools</b> including those for FAIR assessment and data management planning
Ways to discover <b>resources</b> of use to Node members	<b>Search</b> functionality to discover <b>resources</b> for collaboration and use

# How can your Node integrate with FAIRsharing?

To enhance...

**Nodes** should integrate FAIRsharing by...

**citability** of Node resources

Using FAIRsharing record links in your guidance to members, **providing stable PIDs** for resources you want to recommend

discoverability of Node resources

Recommending that members **register** any newly-developed **resources** in FAIRsharing to improve their **visibility** to **humans** and **machines**

exploration of the existing ecosystem

Suggesting that members use FAIRsharing to **discover standards** and **databases**

**Representation** and **visibility** of your Node in FAIRsharing

Nominating Node representatives to **join** the **FAIRsharing Community Champions** program to aid delivery of all of these features

# How can your Node integrate with FAIRsharing?

## To enhance...

## Nodes should integrate FAIRsharing by...

**citability** of Node resources

Using FAIRsharing record links in your guidance to members, **providing stable PIDs** for resources you want to recommend

**discoverability** of Node resources

Recommending that members **register** any newly-developed **resources** in FAIRsharing to improve their **visibility** to **humans** and **machines**

exploration of the existing ecosystem

Suggesting that members use FAIRsharing to **discover standards** and **databases**

Representation and visibility of your Node in FAIRsharing

Nominating Node representatives to **join the FAIRsharing Community Champions** program to aid delivery of all of these features

# How can your Node integrate with FAIRsharing?

To enhance...

**Nodes** should integrate FAIRsharing by...

**citability** of Node resources

Using FAIRsharing record links in your guidance to members, **providing stable PIDs** for resources you want to recommend

**discoverability** of Node resources

Recommending that members **register** any newly-developed **resources** in FAIRsharing to improve their **visibility** to **humans** and **machines**

**exploration** of the existing ecosystem

Suggesting that members use FAIRsharing to **discover standards** and **databases**

**Representation and visibility** of your Node in FAIRsharing

Nominating Node representatives to **join the FAIRsharing Community Champions** program to aid delivery of all of these features

# How can your Node integrate with FAIRsharing?

To enhance...

**Nodes** should integrate FAIRsharing by...

**citability** of Node resources

Using FAIRsharing record links in your guidance to members, **providing stable PIDs** for resources you want to recommend

**discoverability** of Node resources

Recommending that members **register** any newly-developed **resources** in FAIRsharing to improve their **visibility** to **humans** and **machines**

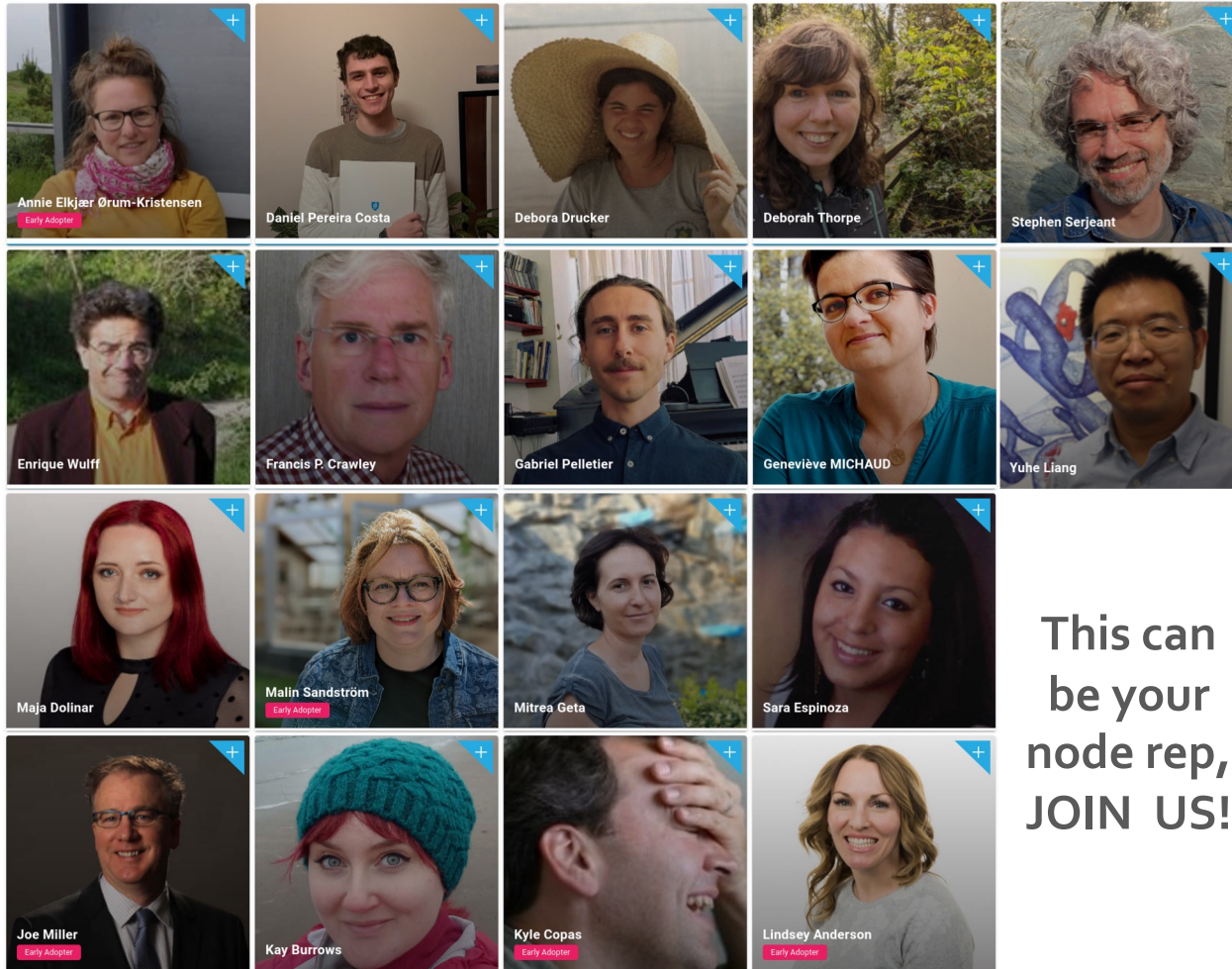
**exploration** of the existing ecosystem

Suggesting that members use FAIRsharing to **discover standards** and **databases**

**Representation** and **visibility** of your Node in FAIRsharing

Nominating Node representatives to **join** the **FAIRsharing Community Champions** program to aid delivery of all of these features

# FAIRsharing Community Champion Programme





This can  
be your  
node rep,  
JOIN US!

A thriving community of domain and discipline experts who:

- 1 . **act as advocates** to promote the value of standards, databases and policies for digital objects (incl. data, software).
- 2 . **create educational material** describing these resources helping researchers and other stakeholders to find, use and adopt them.
- 3 . **enrich the content** of FAIRsharing, adding and enhancing the description and discoverability of these resources.

ORCID

Membership and service (1)	The following community curators have contributed to this record:
FAIRsharing.org: Oxford, GB	<a href="mailto:kcopas@gbif.org">kcopas@gbif.org</a>  <a href="mailto:jmiller@gbif.org">jmiller@gbif.org</a> 
2022-01-01 to 2022-12-31   FAIRsharing Community Curator (Service)	
URI: <a href="#">FAIRsharing user page for kcopas@gbif.org</a>	
URI: <a href="#">FAIRsharing Community Curation Programme</a>	
Source: FAIRsharing.org	

[fairsharing.org/community\\_champions/our\\_champions](https://fairsharing.org/community_champions/our_champions)

# Thank you!



## Stakeholder Advisors

- **Amye Kenall**, Research Square
- **Adam Leary**, Oxford University Press
- **Catriona MacCallum**, Hindawi
- **Dagmar Meyer**, ERC, Executive Agency
- **Dominic Fripp**, JISC
- **Emma Ganley**, Protocols.io
- **Geraldine Clement-Stoneham**, MRC
- **Helena Cousijn**, DataCite
- **Lauren Cadwallader**, PLoS
- **Imma Subirats**, FAO of the United Nations
- **Kiera McNiece**, Cambridge University Press
- **Lauren Cadwallader**, PLoS
- **Lorenzo Feri**, Elsevier
- **Luiz Olavo Bonino**, GO-FAIR
- **Mark Leggott**, Digital Research Alliance of Canada
- **Marta Teperek**, TUDelft
- **Michael Ball**, Medical Research Council
- **Mike Huerta**, NIH National Library of Medicine
- **Nick Everitt** and **Matthew Cannon**, Taylor and Francis
- **Rebecca Grant**, F1000
- **Sarah Callaghan**, Strategy & Policy Unit, University of Oxford
- **Sarah Stewart**, Bodleian Libraries, University of Oxford
- **Scott Edmunds**, GigaScience, Oxford University Press
- **Simon Hodson**, CODATA
- **Sowmya Swaminathan**, Springer Nature
- **Theo Bloom**, British Medical Journal
- **Thomas Lemberger**, EMBO Press
- **Wei-Mun Chan**, eLife

## Executive Advisors

- **David Carr**, Independent expert
- **Peter McQuilton**, (FAIRsharing Founding Member), GSK
- **Robert Hanisch**, NIST
- **Varsha Khodiyar**, Independent Expert

## RDA FAIRsharing WG Chairs

- **Graham Smith**, Springer Nature
- **Holly Murray**, Health Data Research UK
- **Peter McQuilton**, GSK
- **Rebecca Grant**, F1000
- **Simon Hodson**, CODATA
- **Allyson Lister**, University of Oxford
- **Susanna-Assunta Sansone**, University of Oxford

## Community Champions

[fairsharing.org/community\\_champions/our\\_champions](https://fairsharing.org/community_champions/our_champions)

[fairsharing.org/communities#governance](https://fairsharing.org/communities#governance)

WG

**FAIRsharing Registry: Connecting data policies, standards and databases RDA WG**

[www.rd-alliance.org/group/fairsharing-registry-connecting-data-policies-standards-databases.html](http://www.rd-alliance.org/group/fairsharing-registry-connecting-data-policies-standards-databases.html)

## Operational Team



**Allyson Lister**

*Content and Community Coordinator*



**Milo Thurston**

*Technical Coordinator*



**Susanna-Assunta Sansone**

*Principal Investigator and Founder*



**Philippe Rocca-Serra**

*Co-Founder*



**Ramon Granell**

*Data Manager and Research Software Engineer*



**Dominique Batista**

*Front End Manager*



**Delphine Dauga**

*Curation Manager*



**Prakhyat Gailani**

*Web Developer*



# FAIR Cookbook, ELIXIR-LU, Danielle Welter



# The FAIR Cookbook as a Node Service

## Background

<https://faircookbook.elixir-europe.org/>

- Developed by the FAIRplus project
- Open, online resource of hands-on recipes for “FAIR doers” in the Life Sciences
- Practical guidance on specific FAIR & RDM challenges



ELITMa

# The FAIR Cookbook as a Node Service

## Value of the FAIR Cookbook to ELIXIR nodes

- Quality research data management practices & FAIR data are at the core of ELIXIR's mission
- ELIXIR nodes build up expertise in different areas of RDM and FAIR
- The FAIR Cookbook is one way of sharing expertise with the wider community, both at node level and across ELIXIR

# The FAIR Cookbook as a Node Service

## Why we integrated the FAIR Cookbook into our Service Delivery Plan

- **Integration into Service Delivery Plan = sustainability!**
  - The Cookbook is deeply interconnected with other ELIXIR resources, such as the RDMkit
  - ELIXIR is an active network of experts in a wide range of domains
  - ELIXIR-LU & ELIXIR-CH represented on the Cookbook Editorial Board
- **Special case:** no dedicated infrastructure or deployment requirement
  - FAIR Cookbook runs on Github with open access tools and technologies, under the ELIXIR banner



# The FAIR Cookbook as a Node Service

## Integration into the Service Delivery Plan for ELIXIR-LU - practical steps

- [General ELIXIR-LU decision process](#) - based on ELIXIR guidelines
  - Review of Cookbook against inclusion criteria
    - Including but not limited to quality of documentation, usability, licensing, scientific excellence, alignment with mission of ELIXIR/ELIXIR-LU
  - Approval of inclusion proposal by ELIXIR-LU HoN
  - Referral of inclusion proposal to ELIXIR



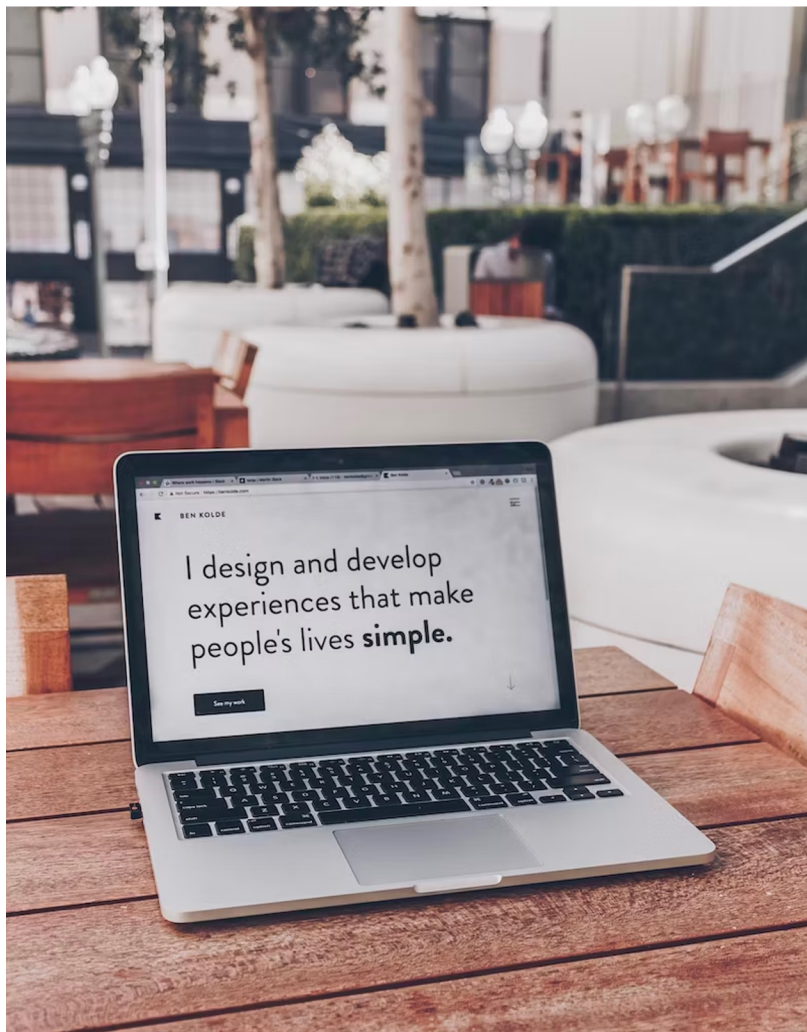
# The FAIR Cookbook as a Node Service

## Integration into the Service Delivery Plan for ELIXIR-CH - practical steps

- Similar process to ELIXIR-LU, again based on ELIXIR guidelines
  - Review of Cookbook against inclusion criteria such quality, purpose and sustainability of the resource
  - Facilitated by precedent of similar assessment in other Nodes
  - Inclusion proposal presented to ELIXIR-CH HoN for approval
  - Referral of inclusion proposal to ELIXIR



# Discussion



- Based on the presentations (Molgenis, EDAM, FAIRsharing, FAIRcookbook), plenary discussion:
  - ✓ What services are relevant to your Node?
  - ✓ What/who do you need (from others) to implement it?
- Please build your [Node's DM portfolio](#) in the shared notes document, as they are for your use
  - ✓ Formulate your own lessons learned
  - ✓ Formulate your take home messages & next steps



# Funder requirements for FAIR data

(13-14 CEST / 14-15 EEST)





# Funder requirements for FAIR data

- **Topic:** Funder requirements for FAIR data
- **Description:** Examples of FAIR requirements by funders (ZonMw: Covid & Dementia data) to showcase what services are needed to help researchers meet those requirements
- **Learning outcome:** Reflect on how funder requirements to FAIR data may affect the Nodes' DM strategy and identify potential implications
- **Form:** Showcasing and discussion
- **Assignment:** Add lessons learned to Node portfolio



# Funder requirements

for research data management and FAIR data:

*From burden to benefit!*

Alles over FAIR-datamanagement bij ZonMw  
[Toegangtotdata@zonmw.nl](mailto:Toegangtotdata@zonmw.nl)



**ZonMw**

## Learning outcomes RDM strategy

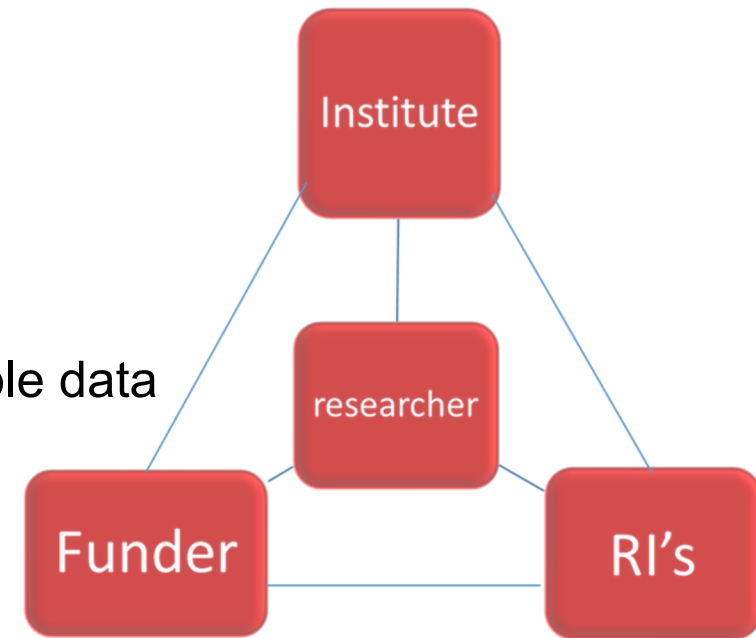
[Elixir](#): How funder requirements may affect the Nodes' DM strategy

>> potential implications

[ZonMw's perspective](#) (and of other funders):

- Creating FAIR data is our mutual interest: value from research > reusable data
- The funder is not a bureaucrat, but a collaborator!
- Distinguish RDM - data FAIRification
- From RDM requirements for single projects

towards organising research communities and domain specific FAIR implementations

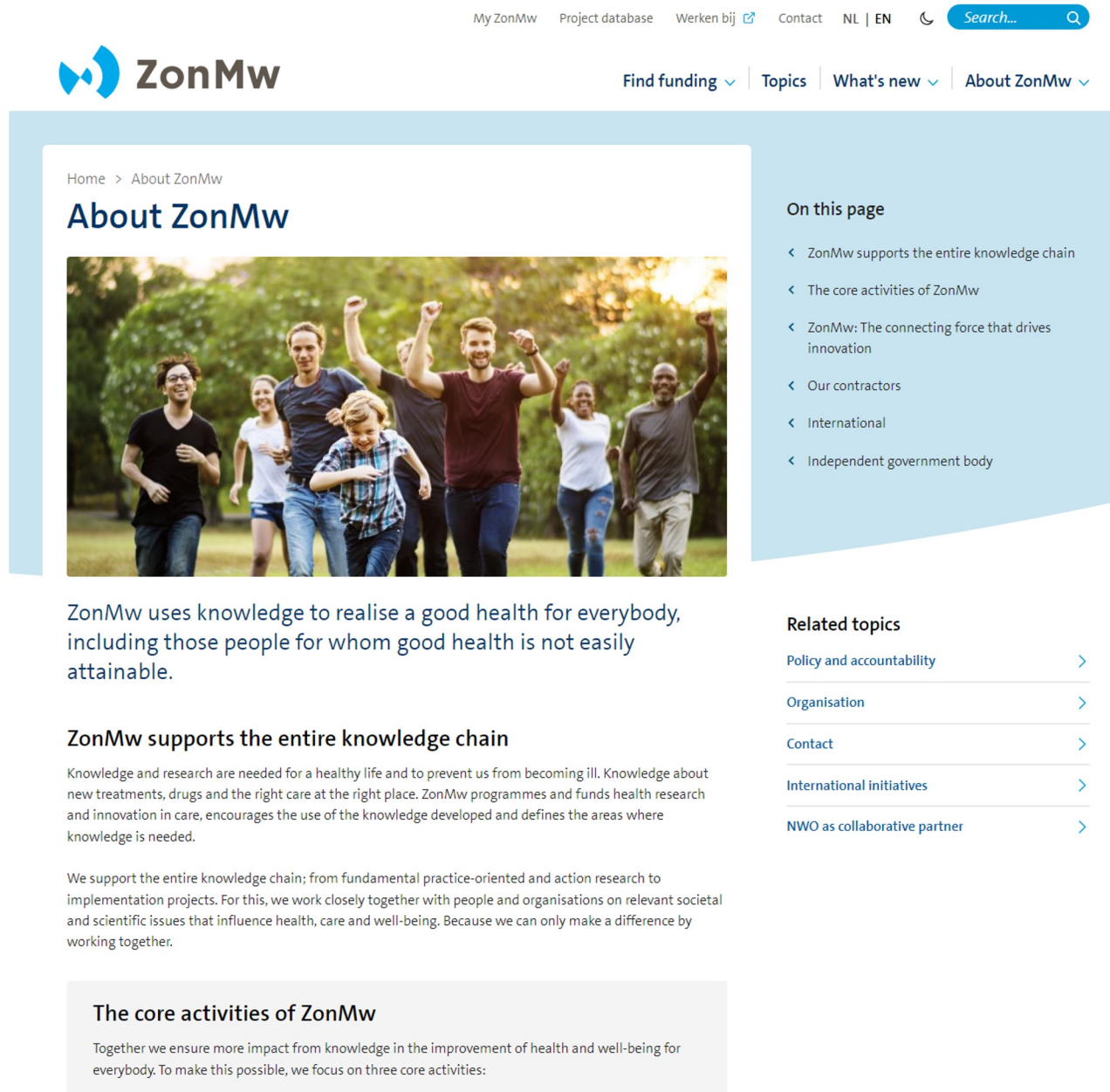


## Outline

1. About ZonMw and the role of a funding organisation in FAIR RDM
2. WHY does a funder bother about reusable data?
3. What & how? Collaboration!
4. RDM requirements in ZonMw's grant procedures
5. Discussion about your experiences (perspective from an ELIXIR node)
6. Steps towards FAIR data – we started with COVID-19
7. Developing approaches and services for data FAIRification
8. Implementing data FAIRification in RDM requirements
9. The funder's role is evolving .....
10. Discussion about your opportunities and collaborators

# What is ZonMw?

- Funding organisation for health research
  - around 400 million a year
- Signaling, research, implementation, impact
- Collaborates (inter)nationally
- **Extra focus on Open Science**
  - Recognition & reward
  - Education
  - Citizen science
  - Open Access and FAIR data



The screenshot shows the ZonMw website's 'About ZonMw' page. At the top, there is a navigation bar with links for 'My ZonMw', 'Project database', 'Werken bij', 'Contact', 'NL | EN', and a search bar. Below the navigation bar, the ZonMw logo is displayed on the left, and a menu with 'Find funding', 'Topics', 'What's new', and 'About ZonMw' is on the right. The main content area features a breadcrumb trail 'Home > About ZonMw' and the title 'About ZonMw'. A large photograph of a diverse group of people running in a park is shown. Below the photo, the text reads: 'ZonMw uses knowledge to realise a good health for everybody, including those people for whom good health is not easily attainable.' This is followed by a section titled 'ZonMw supports the entire knowledge chain' with a sub-header 'ZonMw supports the entire knowledge chain' and a paragraph: 'Knowledge and research are needed for a healthy life and to prevent us from becoming ill. Knowledge about new treatments, drugs and the right care at the right place. ZonMw programmes and funds health research and innovation in care, encourages the use of the knowledge developed and defines the areas where knowledge is needed.' Below this is another paragraph: 'We support the entire knowledge chain; from fundamental practice-oriented and action research to implementation projects. For this, we work closely together with people and organisations on relevant societal and scientific issues that influence health, care and well-being. Because we can only make a difference by working together.' At the bottom of the main content area, there is a section titled 'The core activities of ZonMw' with the text: 'Together we ensure more impact from knowledge in the improvement of health and well-being for everybody. To make this possible, we focus on three core activities:'. On the right side of the page, there is a sidebar with 'On this page' containing a list of links: 'ZonMw supports the entire knowledge chain', 'The core activities of ZonMw', 'ZonMw: The connecting force that drives innovation', 'Our contractors', 'International', and 'Independent government body'. Below this is a 'Related topics' section with links: 'Policy and accountability', 'Organisation', 'Contact', 'International initiatives', and 'NWO as collaborative partner'.

## Why does ZonMw require RDM & FAIR data?

> improve impact of research outputs and accelerate innovation!

### □ Enable data reuse and data sharing:

- Less duplication: money, time, laboratory animals, subjects
- Enrich research (combining datasets)
- Pattern recognition in large datasets
- Accelerate knowledge and innovation for improving health



Data and other research outputs, e.g.:

Biological materials, recordings, images, etc.

### □ Enable verification of research results:

- Research integrity, responsible science
- Reproduction research



### □ Ensure added value for the researcher!

- Improve efficiency; spend less time on 'ordering data'
- Quality improvement

# What is good datamanagement and FAIR data for ZonMw?

- Good **data stewardship** is essential (has been mandatory since 2013)
- **FAIR data**: Scientific input and output is available for re-use **for humans** and **for computers**
- First FAIR implementations with COVID-studies (2020-2022),
- Now spreading across other domains

## General Terms and Conditions Governing Grants of ZonMw

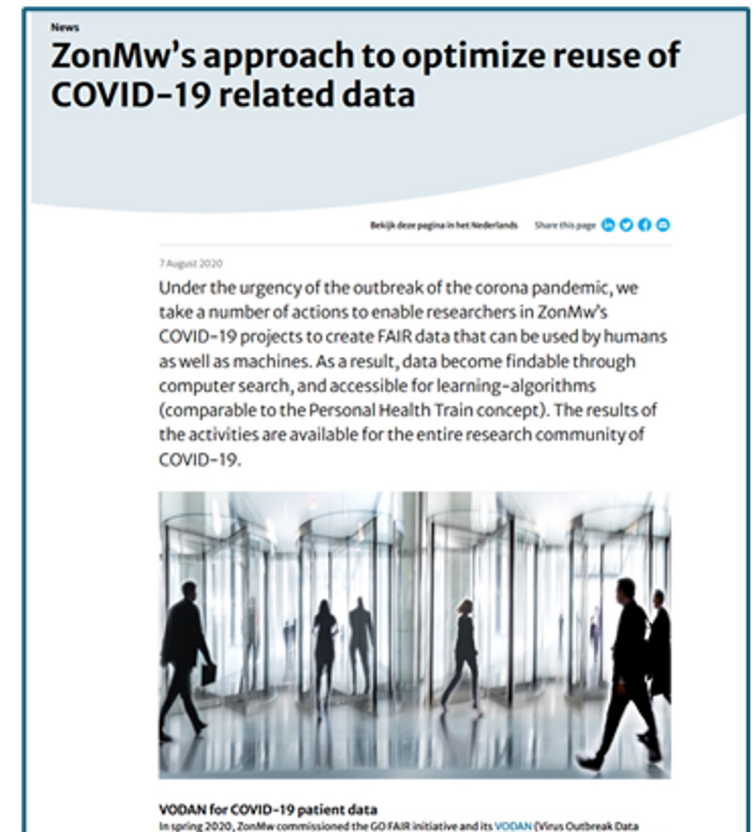
Grant conditions in short:

1 • **Data management plan**

2 • Databases and accompanying explanatory notes will **in any event be made available** in the manner indicated by ZonMw for use for the benefit of further scientific and/or academic research. The data that are collected within a project must therefore be documented and stored in an accessible manner in accordance with the applicable standards.

3 in an accessible manner that is in accordance with the applicable standards.

[Grant Terms and Conditions from 1st July 2013.pdf \(zonmw.nl\)](#)




News  
**ZonMw's approach to optimize reuse of COVID-19 related data**

Bekijk deze pagina in het Nederlands Share this page

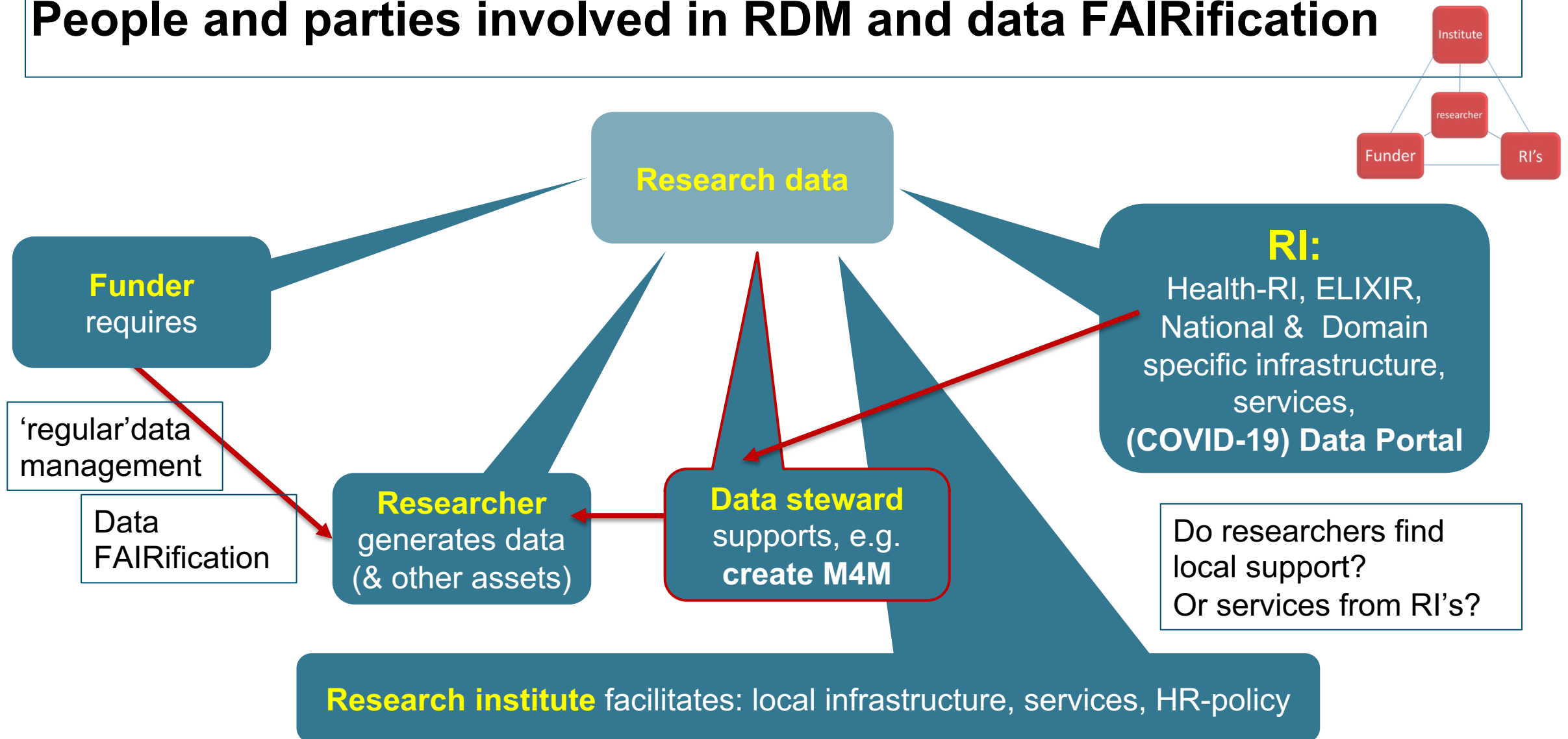
7 August 2020

Under the urgency of the outbreak of the corona pandemic, we take a number of actions to enable researchers in ZonMw's COVID-19 projects to create FAIR data that can be used by humans as well as machines. As a result, data become findable through computer search, and accessible for learning-algorithms (comparable to the Personal Health Train concept). The results of the activities are available for the entire research community of COVID-19.



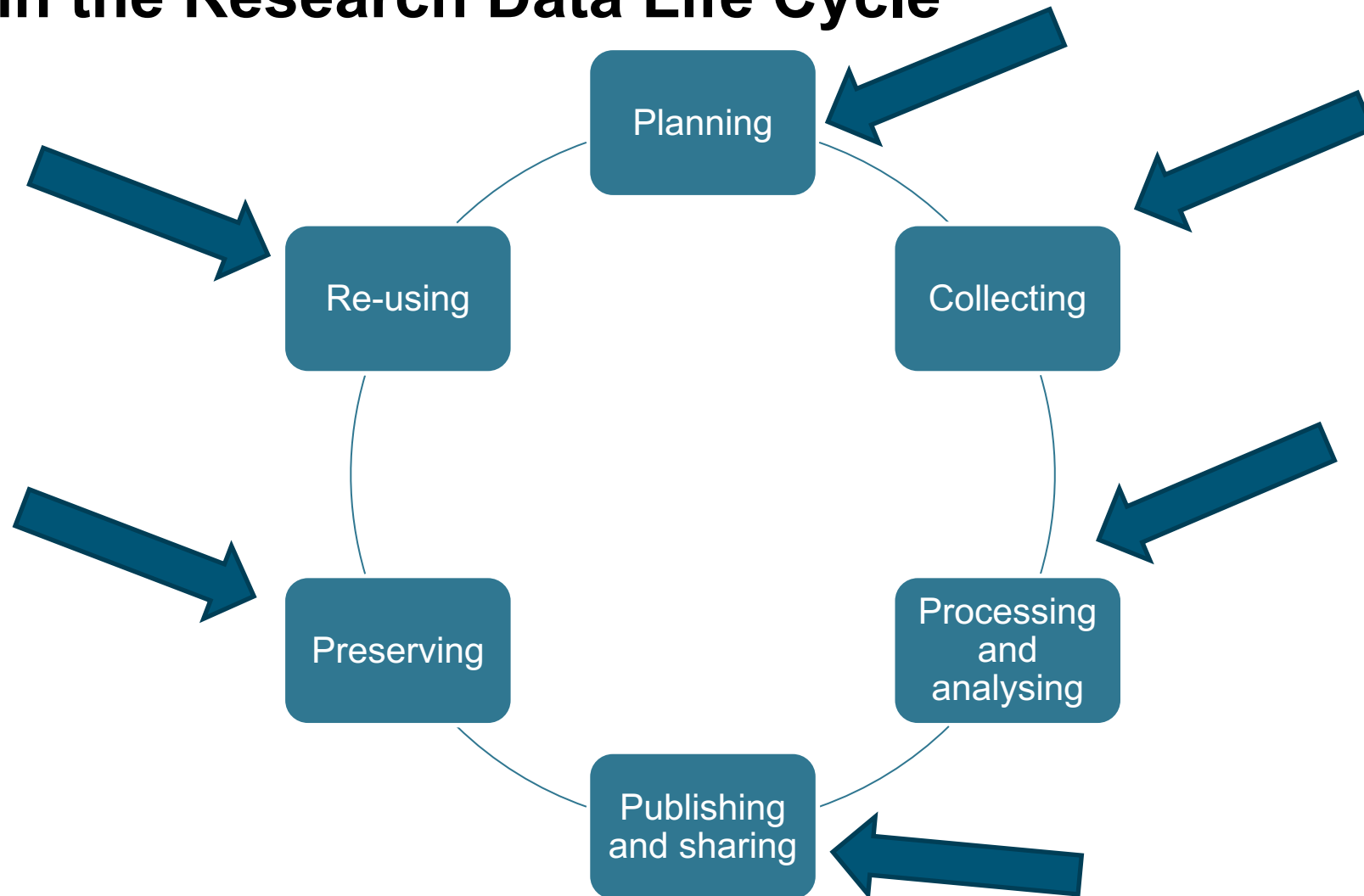
**VODAN for COVID-19 patient data**  
In spring 2020, ZonMw commissioned the GO FAIR initiative and its **VODAN** (Virus Outbreak Data

# People and parties involved in RDM and data FAIRification



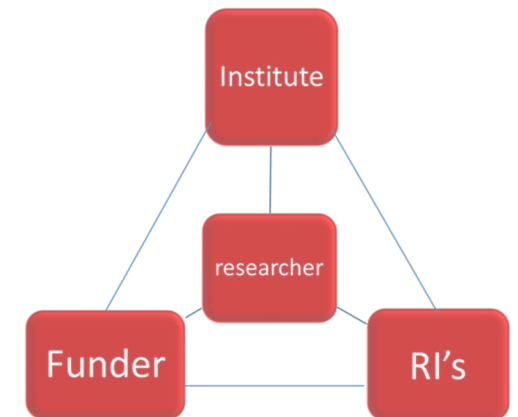


## Phase in the Research Data Life Cycle

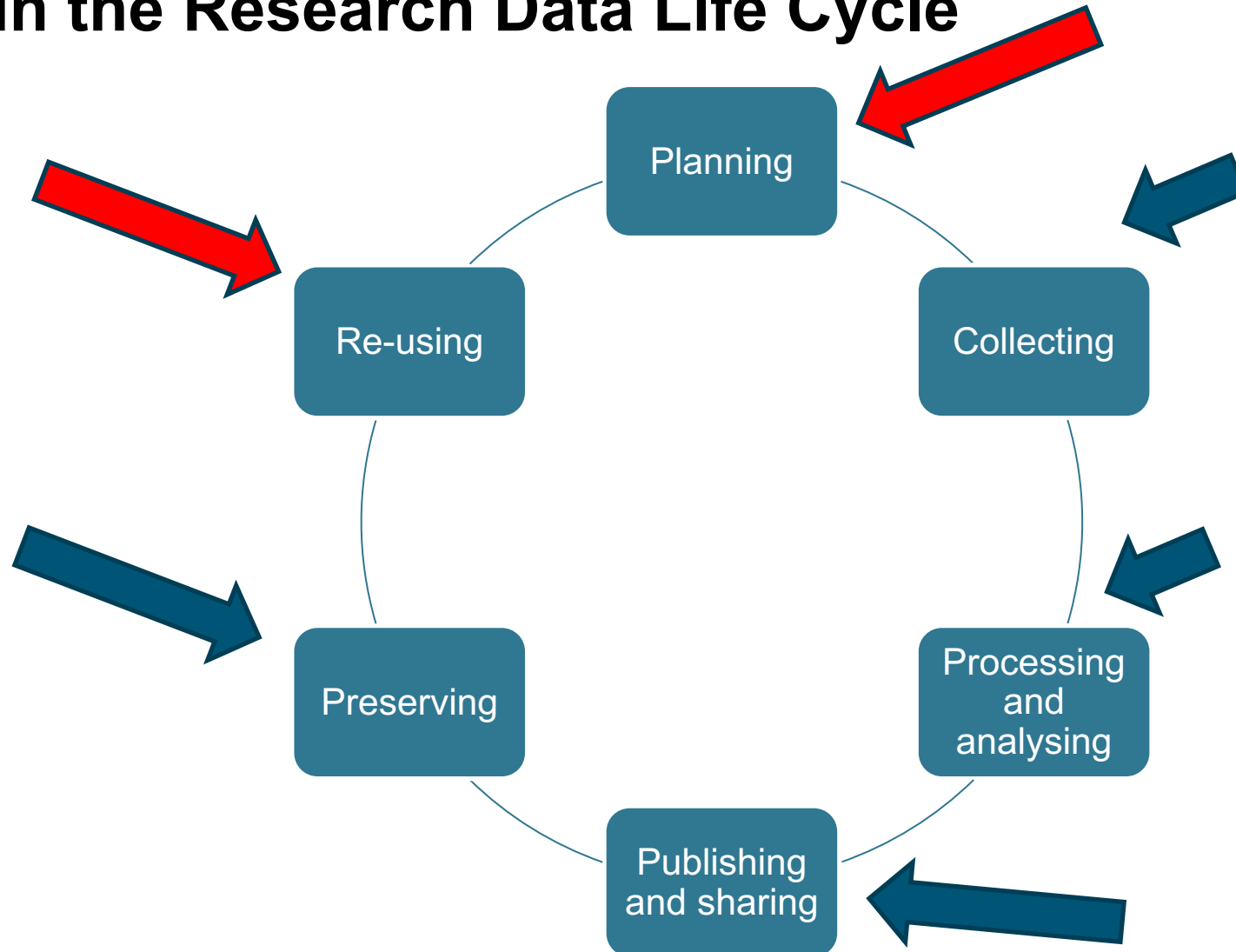


Planning and executing RDM in your research project:

1. **Funder requirements**
2. Support & services & tools



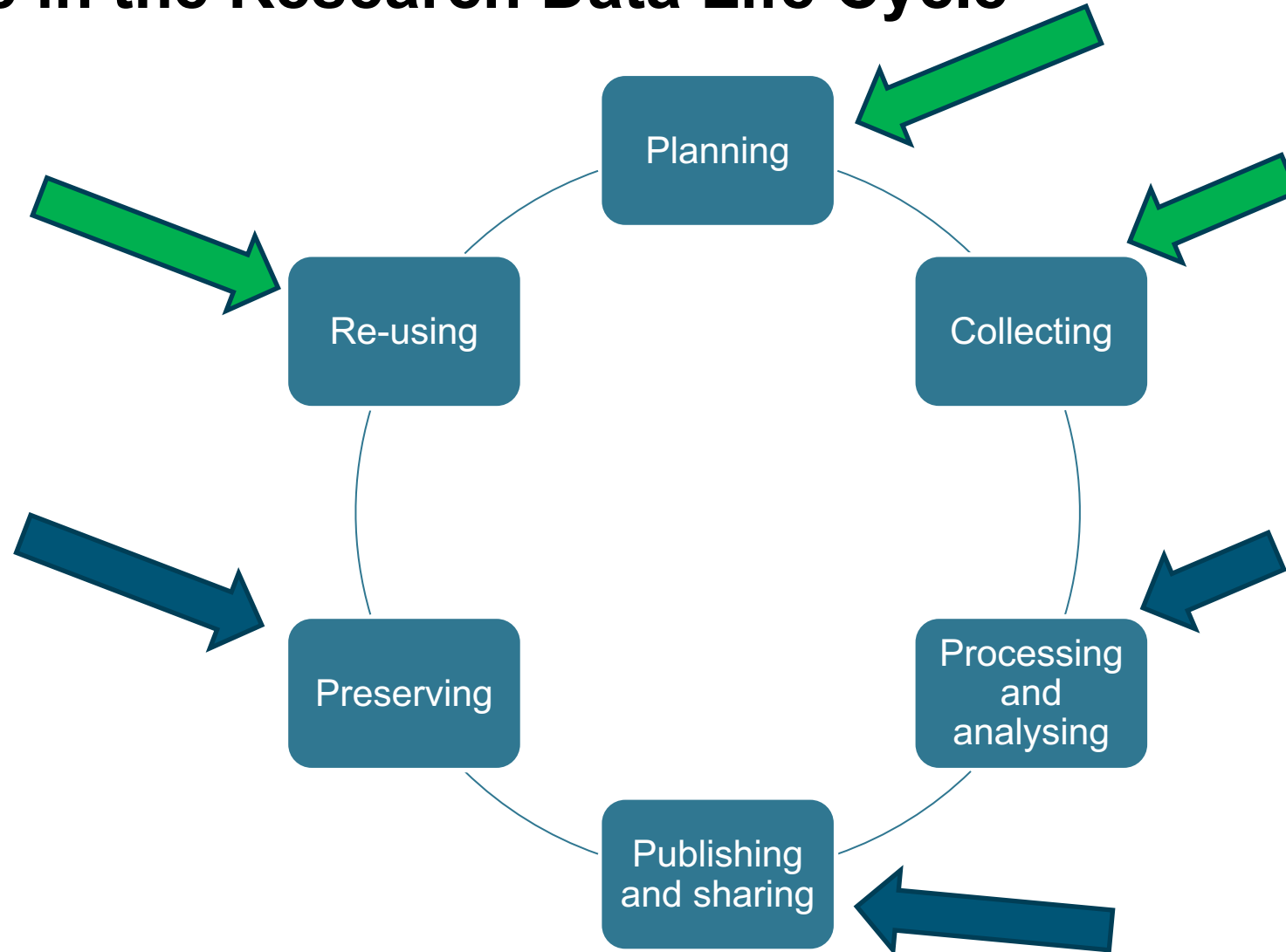
## Phase in the Research Data Life Cycle



### Step 1- Planning RDM in application phase:

1. Check call for proposal.
2. Prepare for requirements in granted project.
3. Look for RDM services & tools\*
4. Fill in the RDM-section in the grant application:
  - a. Reuse existing data
  - b. Explain the need to collect new data
  - c. Who is your data steward
  - d. If possible: use standards of your research domain
  - e. Reserve sufficient budget
  - f. FAIR data requirements?\*

# Phase in the Research Data Life Cycle

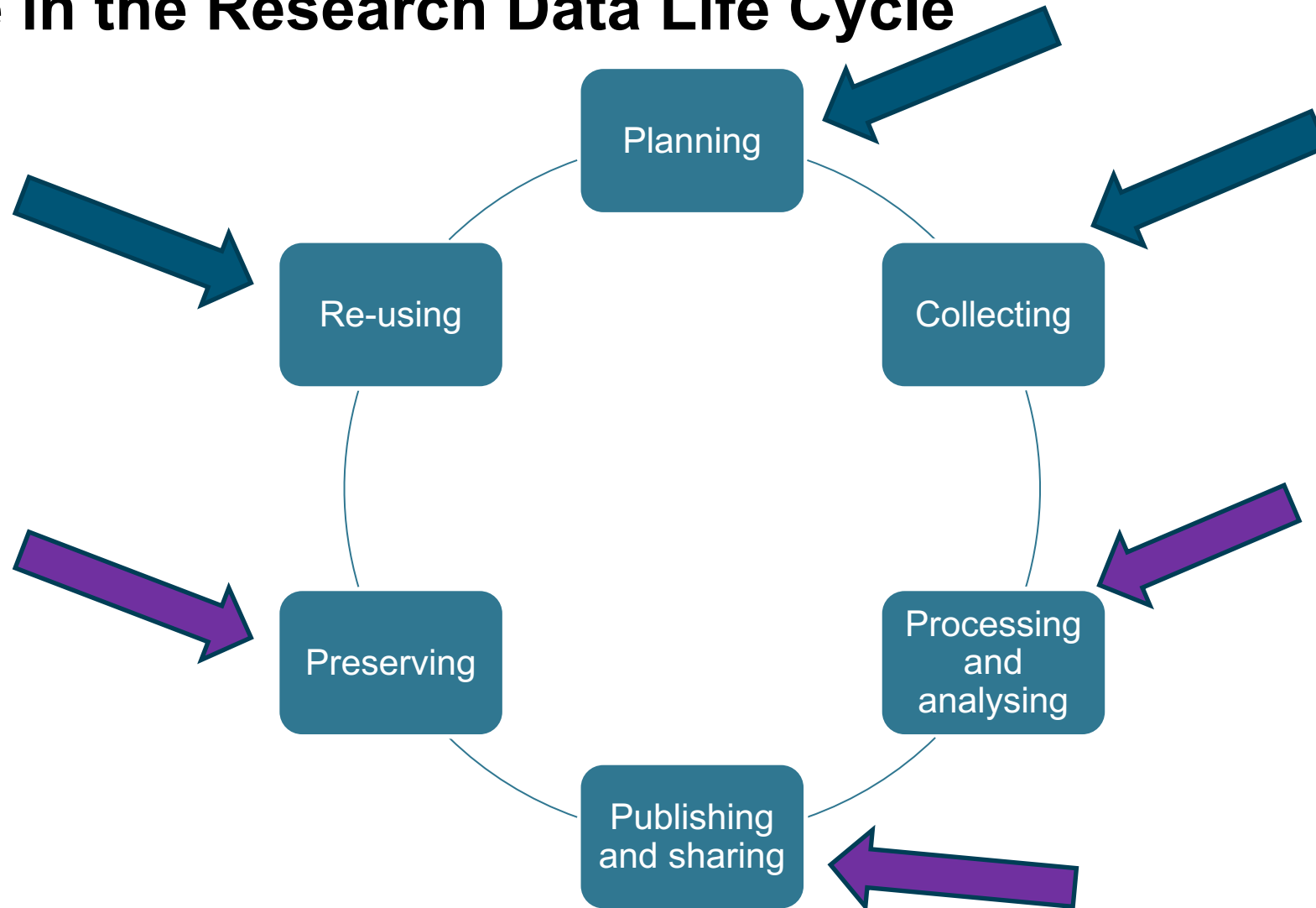


## Step 2 – Planning RDM early in granted project:

1. Involve your data steward (DS)
2. Select a DMP template\*\*
3. Look for RDM services & tools
4. Write a DMP
5. Check quality with DS
6. Deliver DMP @ZonMw (check on completeness)
7. Prepare for ZonMw's monitoring of key items\*\*\*
8. FAIR data activities?

Voorbereiding van FAIR-datamanagement in onderzoeksprojecten | ZonMw

## Phase in the Research Data Life Cycle



### Step 3 – Reporting on RDM output in monitoring phase

1. Involve your datasteward
2. Update your DMP
3. Deliver updated DMP @ZonMw
4. Report on ZonMw's **key items**
5. Produce FAIR (meta)data?

Instructies over rapportage over FAIR-datamanagement met kerngegevens | ZonMw

## How does ZonMw monitor the outcome of a researcher's DM efforts?

### Monitoring KEY ITEMS

1. Persistent Identifier referring to the dataset
  2. Link to the repository or digital archive
  3. Search engine of repository/ online catalogue/ web portal
  4. Terms of Use of dataset
  5. Machine readable data format
  6. Standardized terminology for data
  7. Metadata scheme
- Diagram illustrating the mapping of key items to FAIR principles:
- Items 1, 2, and 3 are grouped by a bracket and labeled **Findable**.
  - Item 4 is linked by an arrow to **Accesible**.
  - Items 5 and 6 are grouped by a bracket and linked by an arrow to **Interoperable**.
  - Item 7 is linked by an arrow to **Reusable**.

### How FAIR are the data?

>> key items give some indication of 'FAIRness'



## Select a DMP template

### Options:

- DMP template of your research institute, provided it is acknowledged by ZonMw (procedure, DS support, key items)

Erkende DMP-templates

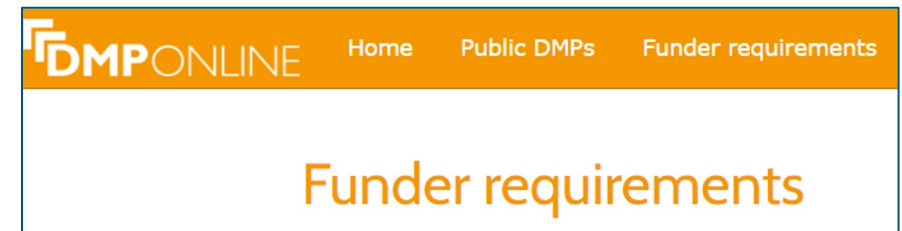
Hogescholen +

Universiteiten +

Academische medische centra -

Instituție	Template	Info	Contact
AmsterdamUMC	<a href="#">Link ↗</a>	<a href="#">Link ↗</a>	<a href="#">Link ↗</a>
UMCU	<a href="#">Link ↗ (DMPonline)</a>	<a href="#">Link ↗ (UMCU)</a>	<a href="mailto:researchIT@umcutrecht.nl">researchIT@umcutrecht.nl</a>

- DMP template of ZonMw on DMPonline



- DS-wizard



## Discussion:

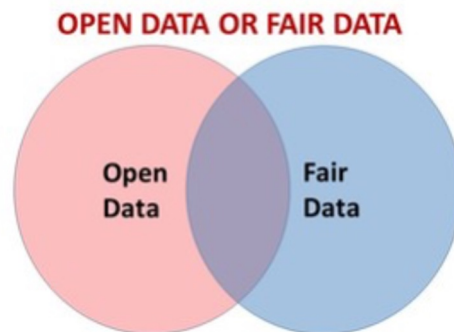
1. What kind of funder's requirements do you know of? Or have you experienced?
2. 'Regular' research data management? *(we will later talk about FAIR data)*
3. What services do you need?
4. Are they available for your RDM? What kind of opportunities or obstacles do you experience?



Please write in your portfolio your most important observations.

List the issues that you want to find out more about.

F<sub>indable</sub> A<sub>ccessible</sub> I<sub>nteroperable</sub> R<sub>eusable</sub>



FAIR is not the same as Open.

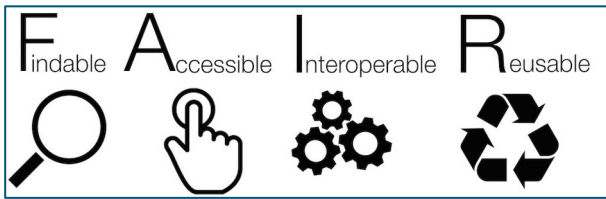
*“as open as possible, as closed as necessary<sup>1</sup>”*



# FAIR principles are a guide

towards

reusable data



FAIR data can be **found, understood, and reused** by humans and by **machines** (computers) (**Fully-AI-Ready**)

FAIRification is for **90% about metadata** that tell us where to find, what they are about and how to use them

- FAIR is not all or nothing
- >> FAIRification is a process
- >> Levels of FAIRness

## Technical infrastructure (generic operations) Data/metadata (domain-specific content)

### Box 2 | The 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:

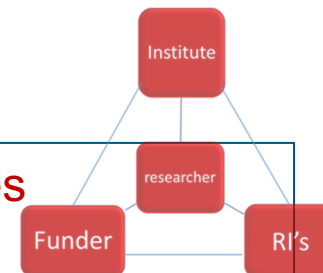
- I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (meta)data use vocabularies that follow FAIR principles
- I3. (meta)data include qualified references to other (meta)data

#### To be Reusable:

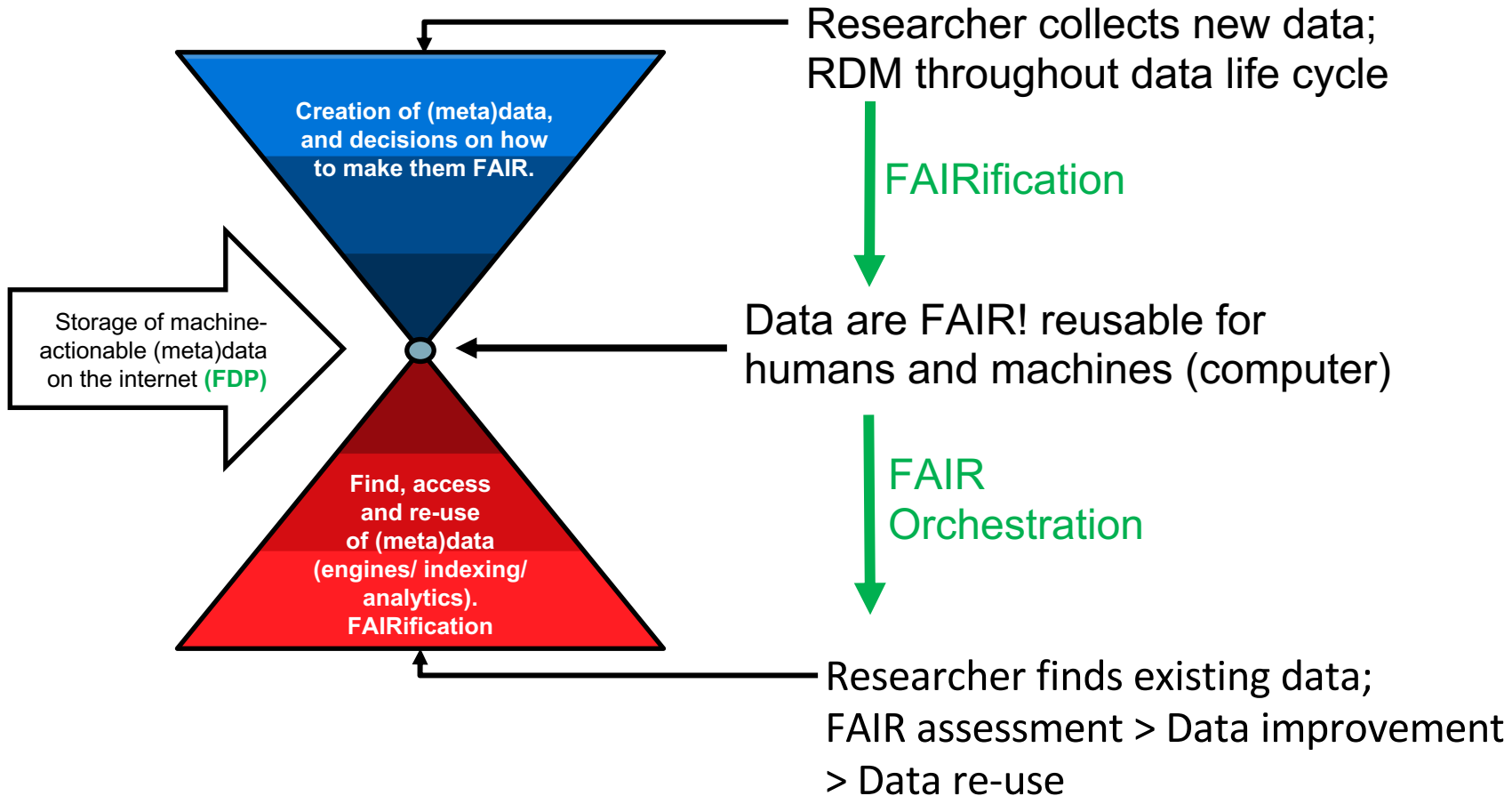
- R1. meta(data) are richly described with a plurality of accurate and relevant attributes
  - R1.1. (meta)data are released with a clear and accessible data usage license
  - R1.2. (meta)data are associated with detailed provenance
  - R1.3. (meta)data meet domain-relevant community standards

FAIRification is about **community specific choices**

- Researchers are 'domain experts'
- A domain agrees about the terms (**controlled vocabularies**) to describe projects; about variables within datasets;
- **FAIR data experts** (e.g. trained data stewards) build the **metadata** schemes (metadata-for-machines, M4M)



# Steps in data FAIRification Hourglass model (GO FAIR)



## 3-Point FAIRification Framework 3-PFF



### Typical resources for data FAIRification:

**FIP** – FAIR Implementation Profile

**M4M** – Metadata-for-Machines

**FDP** - FAIR data point

### Controlled vocabularies


**GUPRI's** (persistent identifiers)

# Collaboration!

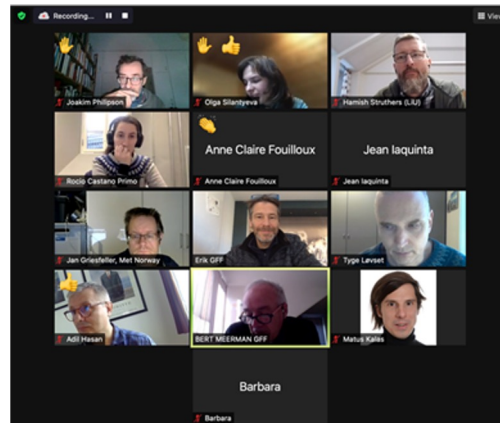
**ZonMw's approach to optimize reuse of COVID-19 related data**

7 August 2020

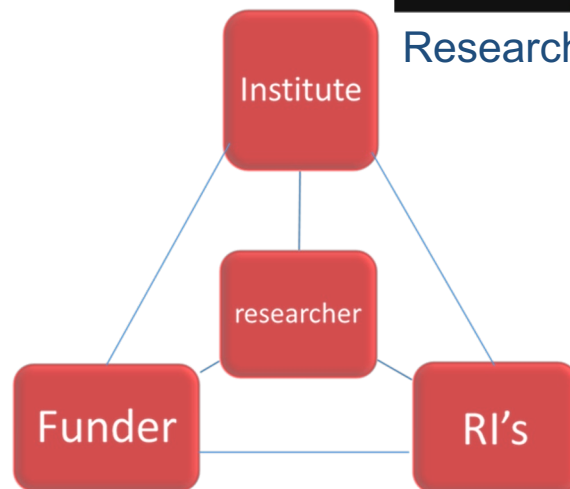
Under the urgency of the outbreak of the corona pandemic, we take a number of actions to enable researchers in ZonMw's COVID-19 projects to create FAIR data that can be used by humans as well as machines. As a result, data become findable through computer search, and accessible for learning-algorithms (comparable to the Personal Health Train concept). The results of the activities are available for the entire research community of COVID-19.



**VODAN for COVID-19 patient data**  
In spring 2020, ZonMw commissioned the GO FAIR initiative and its VODAN (Virus Outbreak Data



Researchers, Domain experts



**ZonMw**  
**COVID-19 Program**  
**62 Projects**

**Brainstorm**  
**controlled**  
**lists**

**M4M.7**  
 13 projects  
 21 participants

1. Data
2. Images
3. Biomaterials
4. Services
5. Standards

**M4M.8**  
 10 projects  
 17 participants

1. Data
2. Images
3. Biomaterials
4. Services
5. Standards

**M4M.9**  
 6 projects  
 12 participants

1. Data
2. Images
3. Biomaterials
4. Services
5. Standards

**M4M.10**  
 7 projects  
 13 participants

1. Data
2. Images
3. Biomaterials
4. Services
5. Standards

**M4M.11**  
 9 projects  
 16 participants

1. Data
2. Images
3. Biomaterials
4. Services
5. Standards

**M4M.12**  
 14 projects  
 21 participants

1. Data
2. Images
3. Biomaterials
4. Services
5. Standards

**COVID-19 Program**  
**controlled list**  
 1. Data  
 2. Images  
 3. Biomaterials  
 4. Services  
 5. Standards

**Build & register**  
**ZonMW COVID-19 Program Vocab**

**Project Admin**  
**Metadata**

**Project Content**  
**Metadata**

**Generic Dataset**  
**Metadata**

**M4M Support Team**



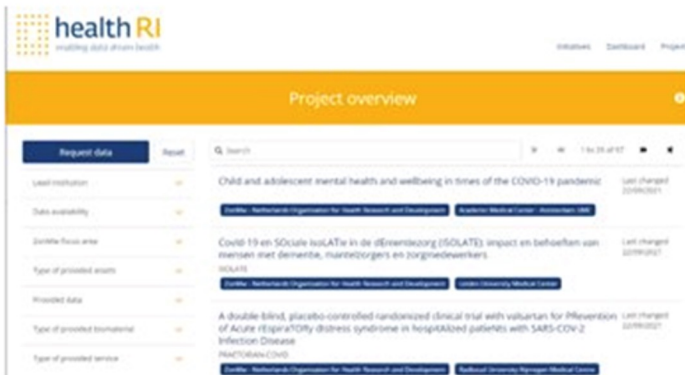
**CEDAR**

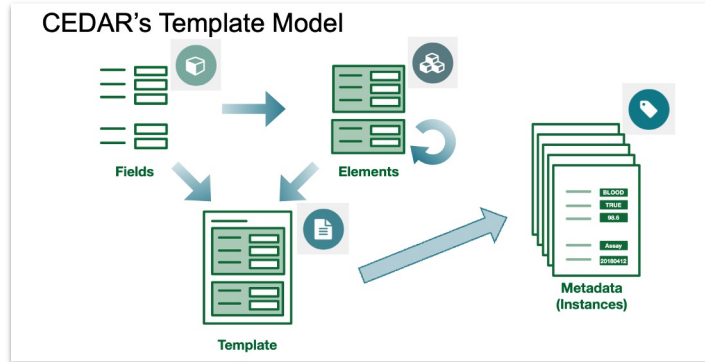


**FAIR**  
**Data Collective**



**COVID-19 Program**  
**Metadata input forms**



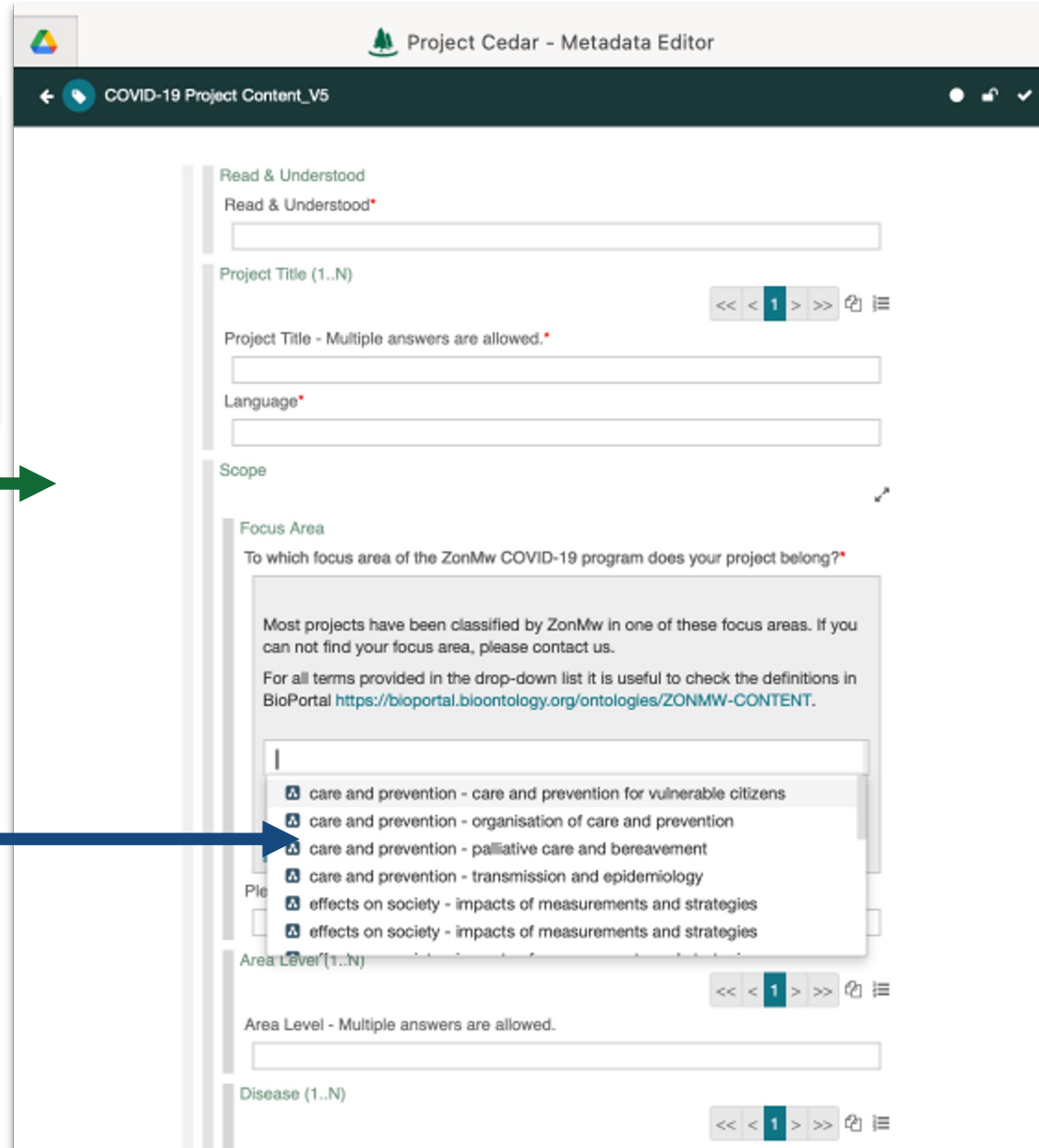


## CEDAR Workbench

- Metadata schema > template > form to be "filled in"
- The "questions"
- Metadata Center



- Controlled vocabularies
- Drop-downs and autocomplete
- The "answers"
- NCBO BioPortal (bioontology.org)



The screenshot shows the 'Project Cedar - Metadata Editor' interface. The form is titled 'COVID-19 Project Content\_V5' and contains several sections:

- Read & Understood:** A text input field with a red asterisk.
- Project Title (1..N):** A text input field with a red asterisk and a note: "Project Title - Multiple answers are allowed." Below the input is a pagination control showing page 1 of 1.
- Language:** A text input field with a red asterisk.
- Scope:** A section with a red asterisk.
- Focus Area:** A section with a red asterisk. It contains a text input field with a dropdown menu open, showing several options:
  - care and prevention - care and prevention for vulnerable citizens
  - care and prevention - organisation of care and prevention
  - care and prevention - palliative care and bereavement
  - care and prevention - transmission and epidemiology
  - effects on society - impacts of measurements and strategies
  - effects on society - impacts of measurements and strategies
- Area Level (1..N):** A text input field with a red asterisk and a note: "Area Level - Multiple answers are allowed." Below the input is a pagination control showing page 1 of 1.
- Disease (1..N):** A text input field with a red asterisk and a pagination control showing page 1 of 1.

# COVID-19 Program aims to make data FAIR

## ZonMw COVID-19 Community

### Guidelines for completing CEDAR forms

[Link to the guidelines document](#)

### CEDAR forms – to be filled out

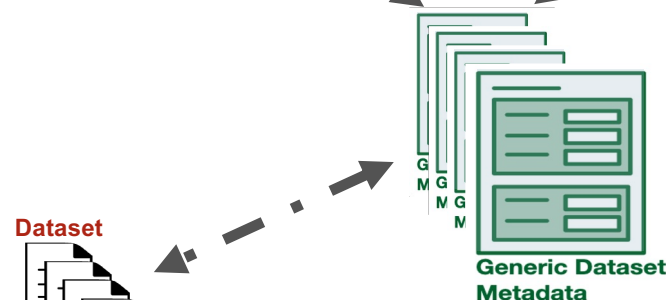
- [Project Admin form](#) [CEDAR account required]
- [Project Content form](#) [CEDAR account required]
- [Data Catalogue form](#) [CEDAR account required]
- [Dataset form](#) [CEDAR account required]
- [Data Distribution form](#) [CEDAR account required]

### CEDAR forms in OpenView – easy way to view/share forms

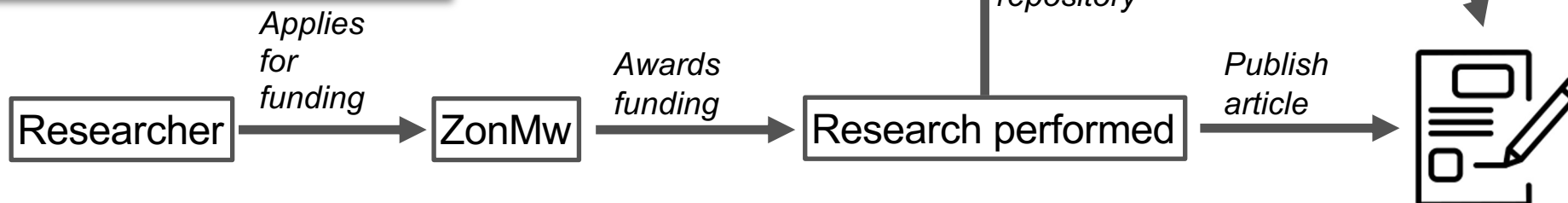
- [Project Admin form in OpenView](#) [no CEDAR account required]
- [Project Content form in OpenView](#) [no CEDAR account required]
- [Data Catalogue form in OpenView](#) [no CEDAR account required]
- [Dataset form in OpenView](#) [no CEDAR account required]
- [Data Distribution form in OpenView](#) [no CEDAR account required]

### BioPortal vocabularies

- [ZonMw Generic Terms](#)
- [ZonMw COVID-19 Vocabulary](#)



*Data goes to a data repository*

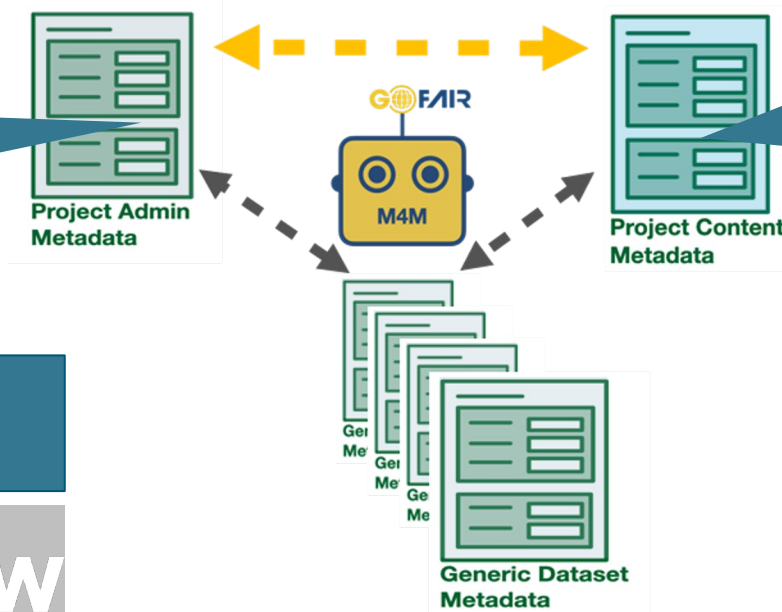


# FIP overview

FAIR principle	Question	FAIR enabling resource types	Your answers
F1	What globally unique, persistent, resolvable identifiers do you use for metadata records?	Identifier type	<i>e.g. PURL, DOI</i>
F1	What globally unique, persistent, resolvable identifiers do you use for datasets?	Identifier type	
F2	Which metadata schemas do you use for findability?	Metadata schema	
F3	What is the technology that links the persistent identifiers of your data to the metadata description?	Metadata-Data linking mechanism	
F4	In which search engines are your metadata records indexed?	Search engines	
F4	In which search engines are your datasets indexed?	Search engines	
A1.1	Which standardized communication protocol do you use for metadata records?	Communication protocol	
A1.1	Which standardized communication protocol do you use for datasets?	Communication protocol	
A1.2	Which authentication & authorisation technique do you use for metadata records?	Authentication & authorisation technique	
A1.2	Which authentication & authorisation technique do you use for datasets?	Authentication & authorisation technique	
A2	Which metadata longevity plan do you use?	Metadata longevity	
I1	Which knowledge representation languages (allowing machine interoperation) do you use for metadata records?	Knowledge representation language	
I1	Which knowledge representation languages (allowing machine interoperation) do you use for datasets?	Knowledge representation language	
I2	Which structured vocabularies do you use to annotate your metadata records?	Structured vocabularies	
I2	Which structured vocabularies do you use to encode your datasets?	Structured vocabularies	
I3	Which models, schema(s) do you use for your metadata records?	Metadata schema	
I3	Which models, schema(s) do you use for your datasets?	Data schema	
R1.1	Which usage license do you use for your metadata records?	Data usage license	
R1.1	Which usage license do you use for your datasets?	Data usage license	
R1.2	Which metadata schemas do you use for describing the provenance of your metadata records?	Provenance model	
R1.2	Which metadata schemas do you use for describing the provenance of your datasets?	Provenance model	

# Extending the approach for FAIRification to other domains

**M4M:**  
Other funders plan to use M4M Project Admin



**M4M:**  
Other research domains plan to develop specific M4M Project Content:

- COVID, ID & AMR,
- Pandemic Preparedness
- Dementia
- ME/CFS

Collaboration – workflow - services:



**Other 'FAIR supporting resosurces':**

- Services & tools > catalogue @Health-RI
- FIP
- Controlled vocabularies
- COVID-19 Data portal > health data portal



# Discussion

## 1. Funders evolve from RDM requirements for single projects towards organising research communities, and domain specific FAIR implementations

- a. What examples do you know about?
- b. What is your opinion on this role?
- c. What do you expect from a funder?

## 2. Collaboration for FAIRification

- a. What opportunities for collaboration do you know? Are / Were they fruitful?
- b. Who are / were collaborating?
- c. What opportunities do you see for FAIRification within your Node



**Reflect on how funder requirements to FAIR data may affect your Node DM strategy and identify potential implications**

Reminder: Please build your [Node's DM portfolio](#) in the shared notes document, as they are for your use

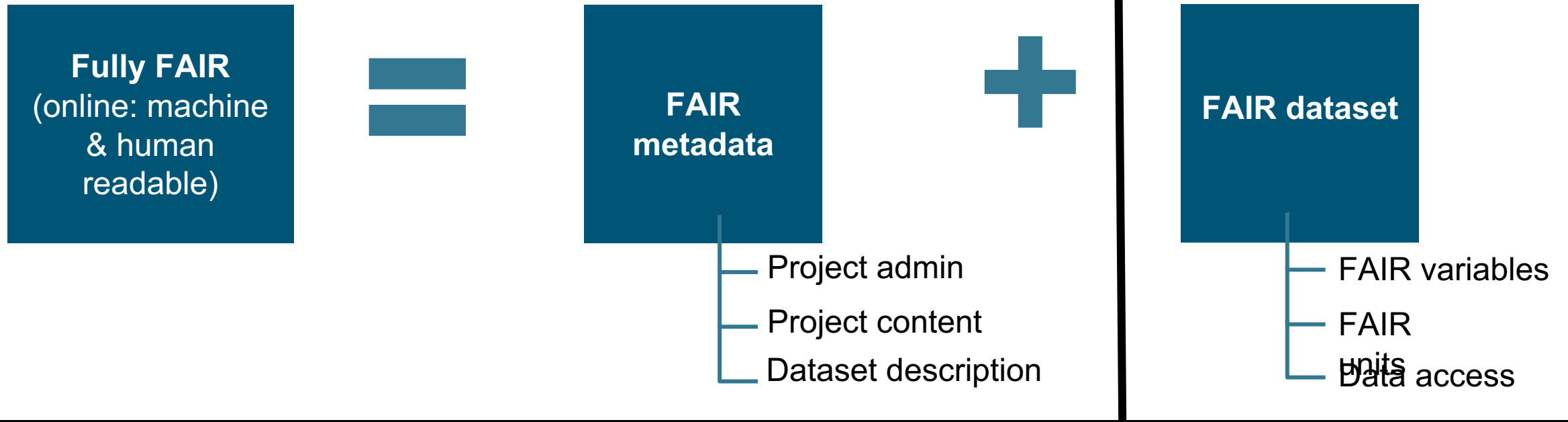
- ✓ Formulate your own lessons learned
- ✓ Formulate your take home messages & next steps

**Extra slides:**

## Types of information – access – reuse

Type	Information	Access - Reuse																									
General	I (dr. Smith) do research with funding X: Do people with many UTI's have a higher change of Alzheimers?	Everyone																									
Article	Written report (publication) on the research question.	Everyone if Open Access																									
Metadata	Information about the context of the sources, the project, information about the data, the conditions of how it was collected, and the conditions for (re)use.	Everyone (no sensitive information)																									
	<table border="1"> <thead> <tr> <th></th> <th>Name</th> <th>Age</th> <th>Tau (i)</th> <th>Fr UTI.</th> </tr> </thead> <tbody> <tr> <td></td> <td>(#/y)</td> <td></td> <td></td> <td></td> </tr> <tr> <td>a.</td> <td>J. de Boer</td> <td>60</td> <td>40</td> <td>3</td> </tr> <tr> <td>b.</td> <td>A. Vlecht</td> <td>51</td> <td>20</td> <td>6</td> </tr> <tr> <td>c.</td> <td>L. Smid</td> <td>62</td> <td>24</td> <td>0</td> </tr> </tbody> </table>		Name	Age	Tau (i)	Fr UTI.		(#/y)				a.	J. de Boer	60	40	3	b.	A. Vlecht	51	20	6	c.	L. Smid	62	24	0	
	Name	Age	Tau (i)	Fr UTI.																							
	(#/y)																										
a.	J. de Boer	60	40	3																							
b.	A. Vlecht	51	20	6																							
c.	L. Smid	62	24	0																							
Data		Depending on conditions (access control)																									

# How do you set steps to FAIRify data?



## Examples:

- *Zonmw nr: 461000000*
- *Onderwerp: Alzheimer*
- *Methode: PET & CSF*
- *Onderzoekers: Dr. Piet van UMC afd. VUmc*
- *Beschikbaar: voor universiteits onderzoekers*
- *Dataset:*
  - *mensen van boven de 50, n = 40*
  - *Aβ42 meting in CSF*
  - *Distribution*
- *Geslacht: man, vrouw of anders*
- *Leeftijd: jaren*
- *Aβ42: picogramen/milliliter*
- *License*

# Break

(14-14.30 CEST / 15-15.30 EEST)



# Data stewardship roles and competences

(14.30-15.30 CEST / 15.30-16.30 EEST)



# Data stewardship roles and competences



- **Topic:** Data stewardship roles and competences
- **Description:** Brief overview of roles and competences and stock taking by the Nodes
- **Learning outcome:** Reflect on the roles and the competences you have & need in your current Node team **to implement the Node DM strategy**
- **Form:** Landscape overview and reflection from a Node perspective
- **Assignment:** Add lessons learned to your Node portfolio



Role	Description
Data steward	A person responsible for keeping the quality, integrity, and access arrangements of data and metadata in a manner that is consistent with applicable law, institutional policy, and individual permissions. Data stewardship implies professional and careful treatment of data throughout all stages of a research process. A data steward aims at guaranteeing that data is appropriately treated at all stages of the research cycle (i.e., design, collection, processing, analysis, preservation, data sharing and reuse).
Data analyst	This is someone who knows statistics. They may know programming, or they may be an Excel wizard. Either way, they can build models based on low-level data. Most importantly, they know which questions to ask of the data.
Data engineer	Operating at a low level close to the data, they are people who write the code that handles data and moves it around. They may have some machine learning background.
Data manager	A data manager is a person responsible for the management of data objects including metadata. These people think about managing and preserving data. They are information specialists, archivists, librarians and compliance officers.
Data scientist	A practitioner of data science. It is a generic term that encompasses many fields of specialised expertise. In the current report, data analysts, data stewards and research software engineers are considered as sub-groups of data scientists. In certain contexts, data scientist is also sometimes used in a more limited way that makes it equivalent to either the data analyst or software engineer roles.
Research software engineer	A growing number of people in academia combine expertise in programming with an intricate understanding of research. These Research Software Engineers may start off as researchers who spend time developing software to progress their research or they may start off from a more conventional software-development background and be drawn to research by the challenge of using software to further research.
Research support professional	In the context of digitalisation, these are the people who support scientific researchers conducting data-intensive science. They are not necessarily part of a research team and might be considered as service providers. This is a broad category that can include data stewards, RSEs, data managers, librarians and archivists.

Table 2.1 Roles in the data professional landscape<sup>43</sup>

## Data-related roles

OECD publishing

### BUILDING DIGITAL WORKFORCE CAPACITY AND SKILLS FOR DATA-INTENSIVE SCIENCE

OECD SCIENCE, TECHNOLOGY AND INNOVATION  
POLICY PAPERS  
July 2020 No. 90



OECD report [Building digital workforce capacity and skills for data-intensive science](#) (2020)

Table from [our NPOS report](#) (page 20-21)



nationaal  
programma  
open  
science



# Data stewardship

**Data Stewardship** is the responsible planning and executing of all actions on digital data before, during and after a research project, with the aim of optimizing the usability, reusability and reproducibility of the resulting data

Data stewardship and data management skills are **essential in research**



Role	Task	FTEs needed per 1000 researchers
Data Steward	Assisting researchers with effective management of research data	26
Trainer on data stewardship	Training researchers on data management skills	4

<https://doi.org/10.5281/zenodo.4769514>

WORLD VIEW · 25 FEBRUARY 2020

## Invest 5% of research funds in ensuring data are reusable



It is irresponsible to support research but not data stewardship, says Barend Mons.

Barend Mons 

<https://doi.org/10.1038/d41586-020-00505-7>



# DM Tasks identified by ELIXIR NC

[Spreadsheet Tasks Node coordinators](#)

This inventory by the NC contributed to the formulation of the ELITMa DM Module

	A	B	C	D	E
1		<b>ELIXIR Belgium (Kim)</b>	<b>ELIXIR Netherlands (roles split over Rob, Merlijn, Celia)</b>	<b>ELIXIR-ES/INB (Salva &amp; Eva)</b>	<b>ELIXIR Germany (Vera, Rabeaa)</b>
56					
57	Data management related	Owncloud and Google Drive management	Participate in ELIXIR Converge (Celia/Rob/Mijke)	Work closely together with the Spanish Network for Research Data	Own Cloud and Google Drive management
58		Our ELIXIR-BE Data steward and Bioinformatics trainer are mainly responsible for DM related tasks: identify and organise trainings related to RDM, connecting with institutes and researchers for RDM community, the ELIXIR-Converge project	Run the national Data Stewards interest group (Mijke Jetten)	Ensure ELIXIR-ES participation in projects like ELIXIR-Converge	Ensure ELIXIR-DE participation in projects like ELIXIR-Converge
59			Participate in the National Platform Open Science (Celia/Mijke)	Identify training opportunities related to data management	Communication and Support to data management team involved in Converge
60			Participate in the definition of EOSC (Celia / Rob)	Act as a reference Hub for connecting researchers with data management resources	
61			Definition of Data Management Competency Framework (Mijke/Celia)		
62			<i>Note: We are fulfilling the roles mentioned here, but we do not think that these are roles for the Node Coordinator</i>		
63					

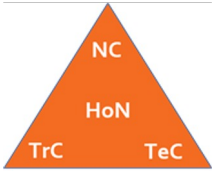
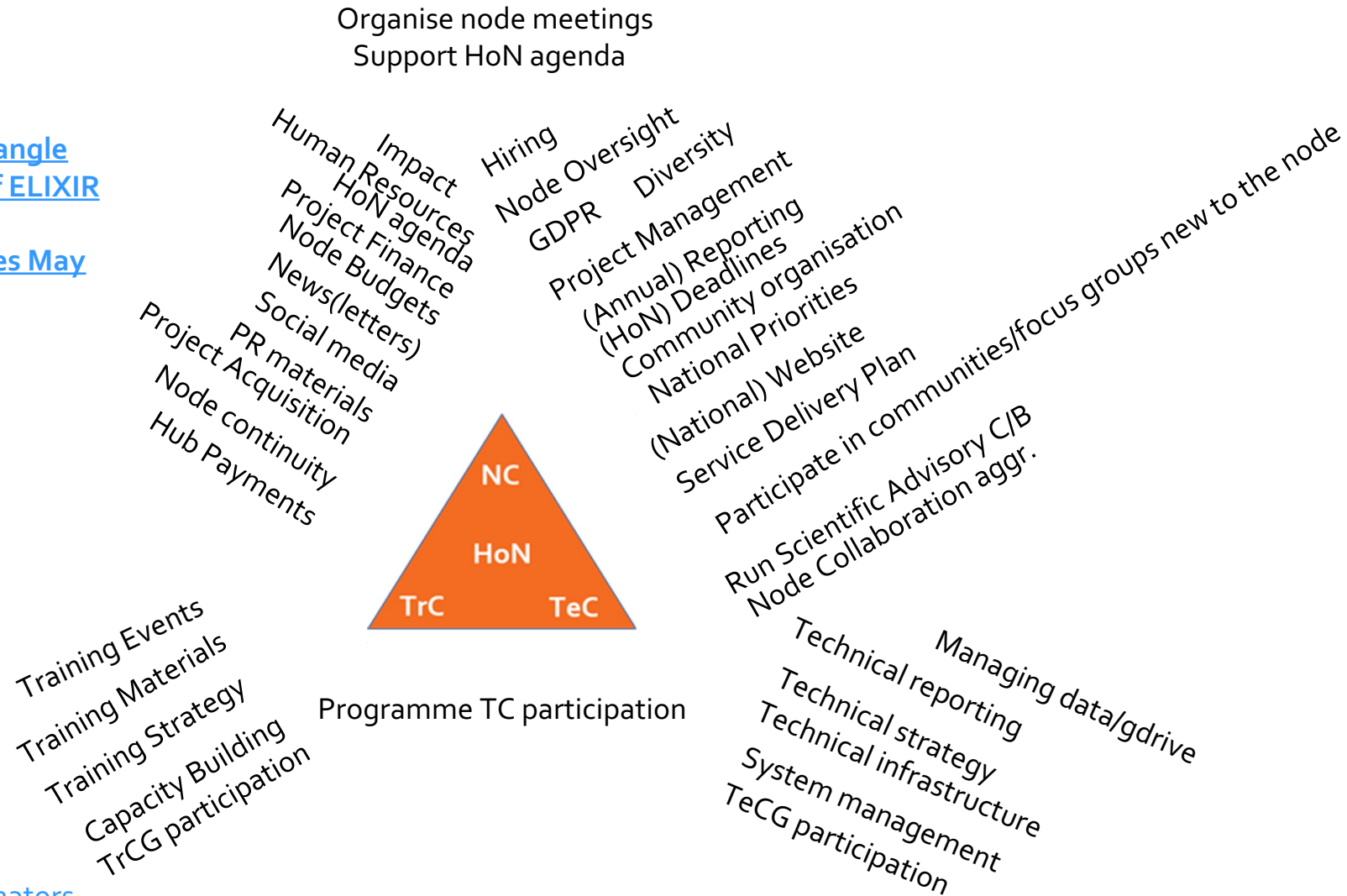
	ELIXIR Germany (Vera, Rabeaa)	ELIXIR PT (Sofia)
1		
56		
57	Data management related	Own Cloud and Google Drive management
58		Google Drive management - Organization of several types of data pertaining to ELIXIR-PT
59		Support to the data management training programme of ELIXIR-PT - Ready for BioData Management?
60		Identify training opportunities related to data management
61		Act as a reference Hub for connecting researchers with data management resources
62		Further training would be useful to support grant writing and project technical support

	A	J	K	L
1		ELIXIR EE (Hedi) / there is no such position as Node Coordinator in EE	ELIXIR ITA (we do not a position, position is on volunteer basis)	<b>ELIXIR Norway (Christine)</b>
55				
56				
57	Data management related	Starting the DM local community and actions		We have a data manager that coordinates all DM work f ELIXIR-NO
58		Overseeing Estonian action in Converge		Keep an overview of national activities in BioMedData project, engage and interact with leaders of partner RIs
59				Keep an overview of node involvements and activities in CONVERGE
60				
61				

# The Node Triangle - Many tasks and responsibilities

Source

[Embedding the Coordinators' Triangle \(TeC-TrC-NC\) in the operations of ELIXIR Nodes](#) (AHM2021 Symposium)  
[Node Triangle Status & Challenges May 2021](#)

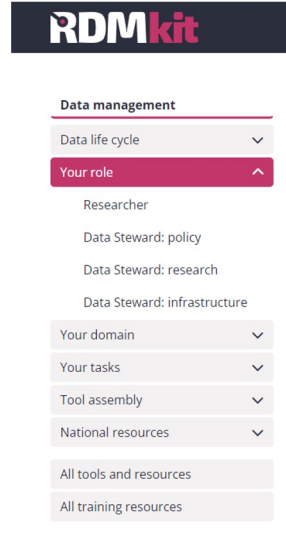
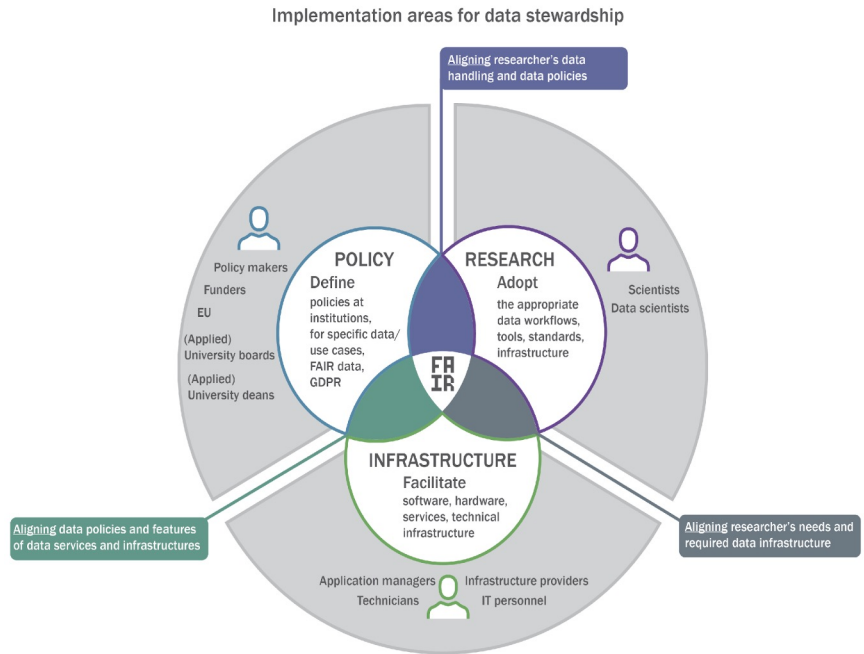


[Spreadsheet Tasks Node coordinators](#)



# Data stewardship roles and competences - the landscape

## ELIXIR/NPOS data stewardship competency framework



- **Data Steward Policy** - institute and policy focused
- **Data Steward Research** - project and research focused
- **Data Steward Infrastructure** - data and e-infrastructure focussed
- Responsibilities, Activities, KSAs (knowledge, skills & abilities), Learning Objectives
- <https://competency.ebi.ac.uk/framework/datasteward/1.0>

Professionalising data stewardship in the Netherlands. Competences, training and education. *Dutch roadmap towards national implementation of FAIR data stewardship*  
<https://doi.org/10.5281/zenodo.4320504>;



# 8 Competence Areas

- **Policy and Strategy:**  
Development, implementation and monitoring of the research data management policy and strategy of the institute
- **Compliance:**  
Compliance with relevant codes of conduct, legislation and field specific standards
- **FAIR data:**  
Alignment to the FAIR data principles and the principles of Open Science
- **Services:**  
Availability of adequate support on research data management, in staff or services
- **Infrastructure:**  
Availability of adequate infrastructure for research data management
- **Knowledge management:**  
Adequate level of knowledge and skills on research data management in the institute
- **Network:**  
Obtaining and maintaining a network of aligned expertise areas and relevant organisations by the institute
- **Data archiving:**  
Adequate support and infrastructure for FAIR and long-term archiving of the data of the institute

## Your task for your Node Portfolio:

Analyse the competences of your current (and future) Node team, based on the identified competence areas for data stewardship (see Portfolio template)



# Roles and competences - the landscape, cntd

Check out other competency profiles in the Competency hub for inspiration too!

<https://competency.ebi.ac.uk/>

<b>Students and professionals in computational biology</b> → Explore the ISCB competencies and: <ul style="list-style-type: none"><li>• Career profiles</li><li>• Training resources</li></ul>	<b>Professionals in computational biomolecular research</b> → Explore the BioExcel competencies and: <ul style="list-style-type: none"><li>• Career profiles</li><li>• Training resources</li></ul>	<b>Professionals in computational personalised medicine</b> → Explore the PerMedCoE competencies and: <ul style="list-style-type: none"><li>• Career profiles</li><li>• Training resources</li></ul>
<b>Research infrastructure scientists</b> → Explore the ARISE competencies	<b>Managers of research infrastructures</b> → Explore the RITrain competencies	<b>Technical operators of research infrastructures</b> → Explore the CORBEL competencies and: <ul style="list-style-type: none"><li>• Training resources</li></ul>
<b>Data steward: policy, research and infrastructure</b> → Explore the NPOS/ELIXIR competencies	<b>Professionals working with clinical human data</b> → Explore the CINECA competencies	<b>Clinical practitioners using genomics</b> → Explore the NHS competencies



# Roles and competences - the landscape, cntd

- Netherlands eScience Center has [released its role description and job profiles](#) for Research Software Engineers (RSEs)!
- Role description: [10.5281/zenodo.7994285](https://doi.org/10.5281/zenodo.7994285)
- Job profiles: [10.5281/zenodo.7805869](https://doi.org/10.5281/zenodo.7805869)



About What we do Resources Insights

## The Netherlands eScience Center Releases Job Profile and Role Description for Research Software Engineers

22 Jun 2023 - 3 min

### What is a Research Software Engineer?

The term Research Software Engineer (RSE) refers to a professional who combines expertise in software development and methodology with deep knowledge of one or more research fields. RSEs focus on developing and/or maintaining research software, with the aim to answer research questions and solve problems within their research disciplines.

The term RSE does not refer to a specific job profile, but rather to a role that can vary significantly in practice. This implies that the educational level and daily work of RSEs can vary widely, and that their roles can be fulfilled in academic institutions both under the umbrella of research support staff and academic staff, depending on their activities<sup>1</sup>.

Within the academic staff, the label RSE can be adopted by any level, from PhD candidates to professors. The emphasis of these RSEs is on high-quality and innovative methodological research, analogous to scientists who focus primarily on instrumentation, such as research groups in methods and statistics in the social sciences or in radio telescopes in astronomy.

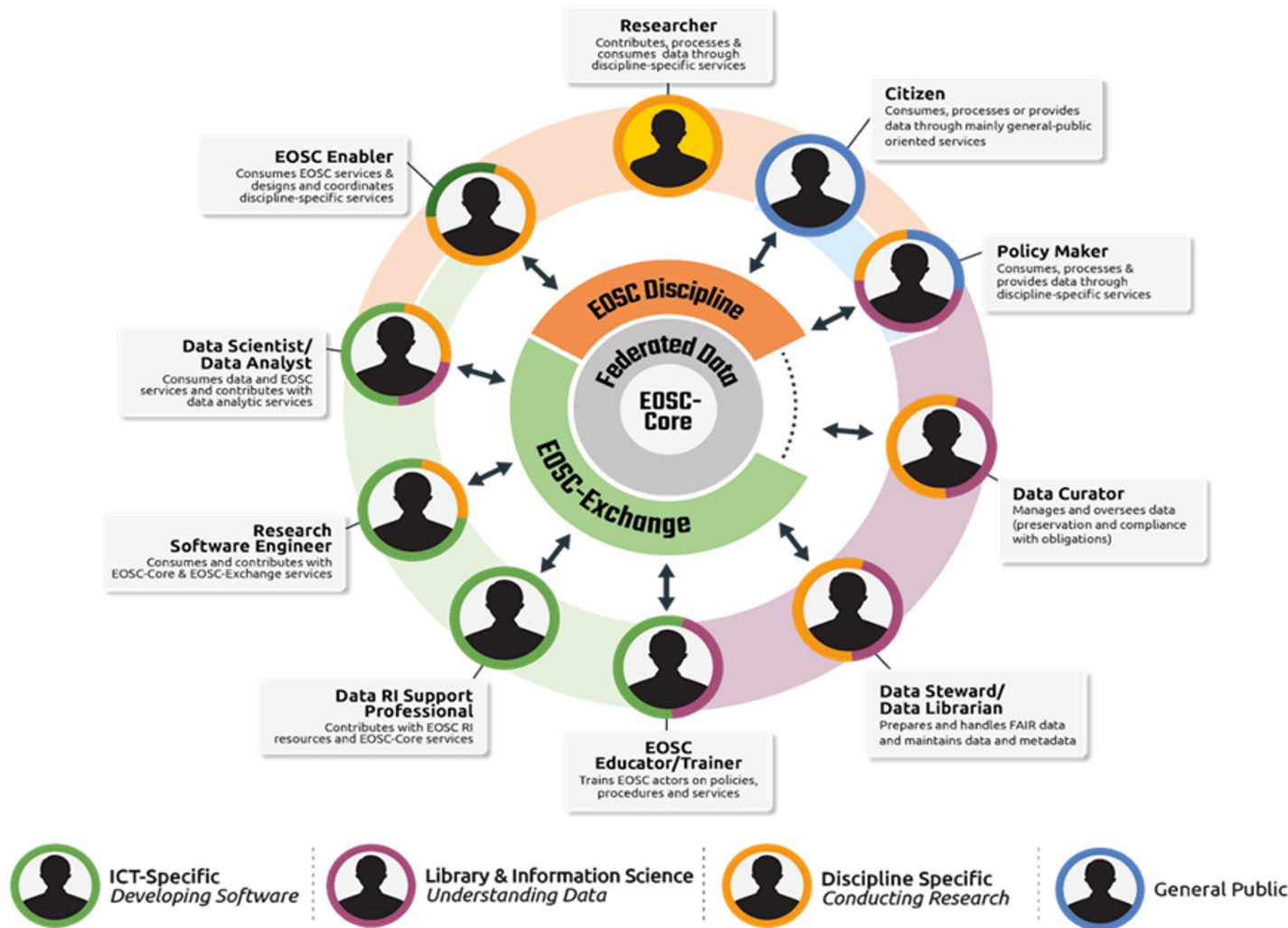
On the research support side of this spectrum, it is appropriate to use the term "software stewardship" as an analogy to "data stewardship." In this case, RSEs primarily focus on advising on best practices, including documentation, licensing, archiving, version control and testing of research software, as well as (supporting) the development and maintenance of existing software and workflows.

In both cases, it is assumed that an RSE has knowledge of designing, writing, implementing and documenting reliable, reusable and sustainable research software that meets the necessary technological, ethical and legal requirements. Many RSEs also provide training to researchers or support staff.



# EOSC Actors

Description of 10 roles within the EOSC ecosystem, one situational example for each role and a list of required skills



Researcher	
<b>Overview</b>	The researcher is the main target of the EOSC ecosystem and interacts with it to obtain, process, produce, deposit and share research data, using mainly high-level services provided by the ecosystem.
<b>Examples</b>	A researcher would browse and identify data related to different genetic variants of the flu available in a FAIR data repository and perform a phylogenetic study on samples through a service from the EOSC marketplace, creating a graphical representation as a phylogenetic tree to include in an article, referencing the data sources and processing pipelines.
<b>Required skills</b>	<ul style="list-style-type: none"> <li>General knowledge on the EOSC ecosystem, covering the EOSC-Core and EOSC-Exchange interdisciplinary services for data access, sharing, reuse and processing, and relevant discipline-specific services.</li> <li>Knowledge of the added value that EOSC services provide to research and publication workflows.</li> <li>Understanding of how to assess the FAIRness of services when searching for and producing research data.</li> <li>Skills to apply EOSC services wherever they support the research and innovation lifecycle and contribute to their development.</li> <li>Training and communication skills to teach and educate other researchers and students on how to conduct research in the frame provided by EOSC.</li> </ul>

Data Steward/Data Librarian	
<b>Overview</b>	A Data Steward is an expert on the preparation and treatment of data including data selection, storage, preservation, annotation provenance and other metadata maintenance, and dissemination. Data librarians are professional library staff who are experts on RDM, using research data as a resource or supporting researchers dealing with data (description, archiving and dissemination). Other closely related roles will also be considered under this category.
<b>Examples</b>	A Data Steward could be an expert who validates, recodes, trims or applies any other action on each source dataset of genomic samples related to influenza to guarantee that they can be properly used and integrated according to domain-specific standard formats.
<b>Required skills</b>	<ul style="list-style-type: none"> <li>Deep understanding of FAIR principles to ensure that research data from various domains is aligned with FAIR and CARE (Collective benefit, Authority to control, Responsibility, Ethics) principles.</li> <li>Ability to use EOSC-Core and EOSC-Exchange services for data publication and preservation and to facilitate the continued development of an infrastructure and library services to support data discovery, curation, preservation and sharing according to those principles.</li> <li>Ability to validate the fulfillment of open science principles in EOSC-Core and EOSC-Exchange services related to data.</li> <li>Ability to advise faculty and students on RDM according to the FAIR and CARE principles, including the discovery and reuse of existing datasets, through the EOSC services and ecosystem.</li> </ul>

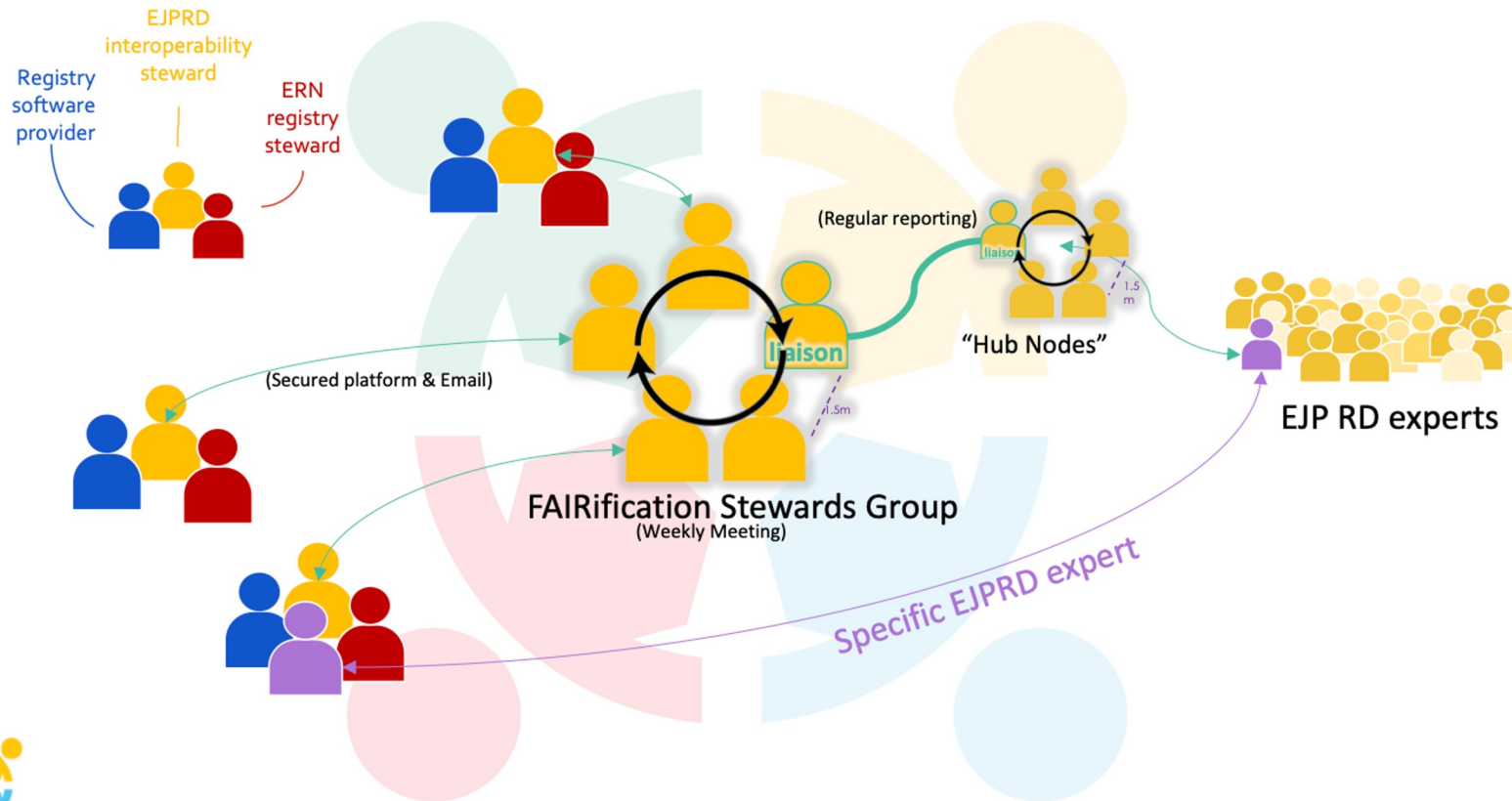
EOSC Working Group Report Digital skills ofr FAIR and open science

<https://op.europa.eu/en/publication-detail/-/publication/af7f7807-6ce1-11eb-aeb5-01aa75ed71a1/language-en/format-PDF/source-190694287%20%20>





# The European Joint Programme on Rare Diseases (EJPRD) FAIRification steward



**FAIRification steward:** tasked to work with the other two parties to help manage FAIRification process of ERN registry towards a converged, EJPRD [Virtual Platform \(VP\)](#) compliant implementation of the ERN registries that the steward is assigned to

**Registry software provider:** tasked with implementing/configuring the software for the registry

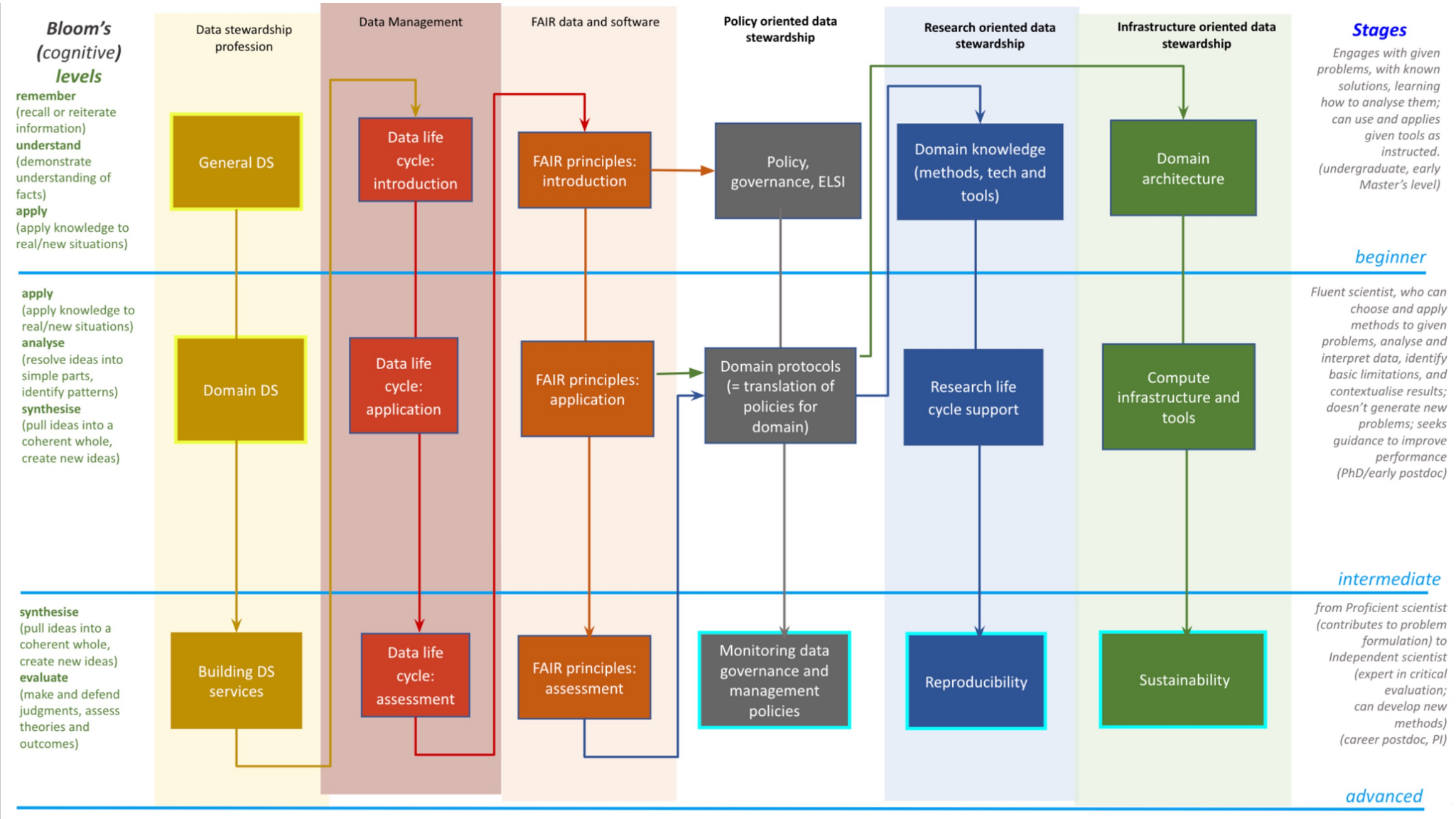
**ERN registry steward:** tasked with one or more aspects of data or project management within the Registry project

[EJPRD FAIRification website](#)

[ERN](#) = European Reference Networks

# Learning Paths for data steward (1)

[BioHackathon 2022 project](#)



# Learning Paths for data steward (2)

[BioHackathon 2022 project](#)

<p><b>Bloom's (cognitive) levels</b></p> <p><b>remember</b> (recall or reiterate information)</p> <p><b>understand</b> (demonstrate understanding of facts)</p> <p><b>apply</b> (apply knowledge to real/new situations)</p>	<p>Data stewardship profession</p> <p>Landscape (FAIR, OS, RI)</p> <p>Roles &amp; Responsibilities</p>	<p>FAIR data and software</p> <p>FAIR principles</p> <p>F -Discovery</p> <p>F - PIDs</p>	<p>Policy oriented data stewardship</p> <p>Understanding Governance landscape</p> <p>Data management policies (RPO, RFO &amp; publishers)</p> <p>Data management planning</p> <p>Legislation, ethical principles, codes of conduct (ELSI)</p>	<p>Data Management</p> <p>Planning for Data Management</p>	<p>Research oriented data stewardship</p> <p>Domain knowledge (methods, tech and tools)</p>	<p>Infrastructure oriented data stewardship</p> <p>Knowledge of and advice on compute infrastructure and tools available</p> <p>Knowledge of and <del>astorage</del> storage options available</p>	<p><b>Stages</b></p> <p>Engages with given problems, with known solutions, learning how to analyse them; can use and applies given tools as instructed. (undergraduate, early Master's level)</p> <p><b>beginner</b></p>
<p><b>apply</b> (apply knowledge to real/new situations)</p> <p><b>analyse</b> (resolve ideas into simple parts, identify patterns)</p> <p><b>synthesise</b> (pull ideas into a coherent whole, create new ideas)</p>	<p>Communication and outreach</p> <p>Community building and networking</p> <p>(Domain) research process</p>	<p>F - Metadata</p> <p>A - Access protocols (formats)</p> <p>A - Access permissions (roles)</p> <p>A - Access control</p> <p>R - Licences</p> <p>R - Metadata</p> <p>R - Provenance</p>	<p>DMP templates, requirements and tools</p> <p>Domain protocols (incl. legal, financial, technical implications)?</p> <p>Governance implementation</p> <p>Policies around storage and compute</p>	<p>Versioning</p> <p>Data wrangling</p> <p>Data preservation</p> <p>Storing data</p> <p>Data sharing</p>		<p>Workflows</p> <p>Pipelines (the process and tools used to gather raw data from multiple sources, analyze it, and present the results in an understandable format)</p> <p>software sustainability??</p>	<p>Fluent scientist, who can choose and apply methods to given problems, analyse and interpret data, identify basic limitations, and contextualise results; doesn't generate new problems; seeks guidance to improve performance (PhD/early postdoc)</p> <p><b>intermediate</b></p>
<p><b>synthesise</b> (pull ideas into a coherent whole, create new ideas)</p> <p><b>evaluate</b> (make and defend judgments, assess theories and outcomes)</p>	<p>Building and delivering support services</p> <p>Developing and Delivering Training</p>	<p>I - Linked data</p> <p>I - Vocabularies &amp; Ontologies</p> <p>I - Standards &amp; Formats</p> <p>R - Reproducible</p>	<p>Impact tracking (metrics)</p> <p>Monitoring data governance and management policies</p>	<p>Handling sensitive data</p> <p>Data quality</p> <p>Data integration &amp; interoperability</p> <p>Data architecture</p> <p>Data security</p>	<p>Pipelines design</p> <p>Workflows design</p>	<p>scripting/coding</p>	<p>from Proficient scientist (contributes to problem formulation) to Independent scientist (expert in critical evaluation; can develop new methods) (career postdoc, PI)</p> <p><b>advanced</b></p>



# Roles and competences - your reflections

- Analyse the competences of your current (and future) Node team, based on the identified competence areas for data stewardship
  - ✓ Fill the section in your Node Portfolio
- Identify a challenge in/for your Node in relation to roles and competences
  - ✓ For plenary discussion and to take further tomorrow

Competence area	- How is this competence area tackled in your Node? - How does your Node want to tackle this going forward?	By whom? Name and/or role (NC, DM, other,...)
Policy and Strategy		
Compliance		
FAIR Data		
Services		
Infrastructure		
Knowledge Management		
Network		
Data archiving		

# Learn from & Share with

(15.30-16 CEST / 16.30-17 EEST)



# Wrap up

- 3 minutes per Node, **for example on:**
  - ✓ Identified challenge in/for your Node in relation to roles and competences
  - ✓ Take-home messages for the day, eye openers, plans for the next steps for your Node DM strategy
  - ✓ Parking Lot topics

Please help make notes

# Tomorrow



## Maturity/services model, Node strategy and the ideal Node team

### With Ivan and Niclas

- Presentation: current Node services maturity model
- Template discussion: what should be in a model?
- Node examples (vision, mission & Node team)
- Breakout discussions: positioning your Node regarding selected topics from the model + ideal Node team
- Discuss next steps for your Node's strategy on data management services
- Role of the newly established RDM Community and its network

- ✓ *Maturity model: a framework for measuring an organization's maturity*
- ✓ *Measuring what? RDM services, activities, readiness, preparedness, awareness*
- ✓ *Which organizations? Research institutions and national or Node level*



**Day 3/29 June**  
**Module 2 - RDM strategy**

**Maturity/services model, Node strategy  
and the ideal Node team**





# Planning [updated version]



## Maturity/services model, Node strategy and the ideal Node team

### With Ivan and Niclas

- **Part 1: 9-10:30 CEST / 10-11:30 EEST**
  - Example ELIXIR-SE: what is a Node strategy for DM [10 min]  
Reflection: has anyone else done a similar exercise yet? [15 min]
  - Node services maturity model [5 min]  
Template discussion: what should be in a model? [15 min]
  - Breakout discussions: positioning your Node regarding selected topics from the model [40 min]
  - Example ELIXIR-EE: ideal Node team to cater for a Node DM strategy [5 min]
- **Break: 10:30-11 CEST / 11:30-12 EEST**
- **Part 2: 11-12 CEST / 12-13 EEST**
  - Discuss next steps for your Node's strategy on data management services [20 min]
  - Where does this effort sit? Role of the newly established RDM Community and its network [20 min]
  - Where do the Parking Lot items sit? Any Other Business? [20 min]



# Example ELIXIR-SE / what is a node strategy for DM

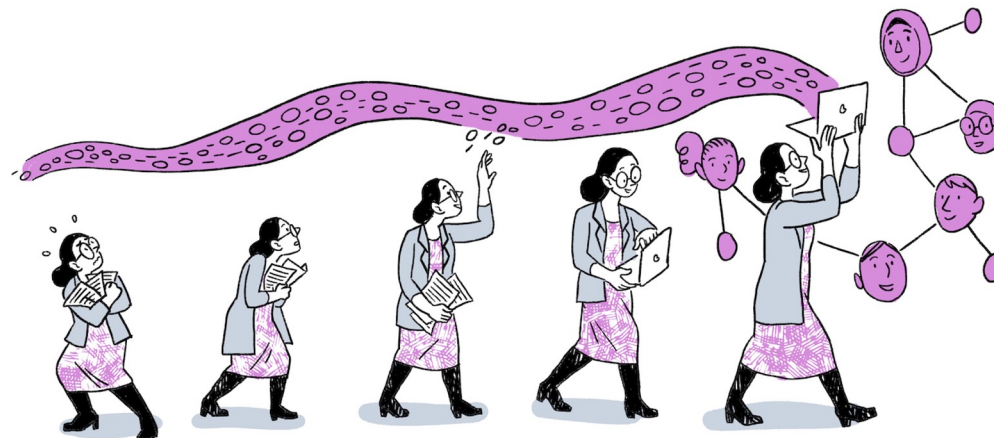
- DM team of eight
  - “Data manager” and seven Data Stewards
- Not yet a full-fledged strategy
- Vision & Mission statements as a starting point



# ELIXIR-SE What's our vision?

*The desired future position*

“*Swedish Life Science researchers apply good Data Management practices so that the research outputs produced are available to the global research community, and to society at large, according to the principles of Open Science, Reproducible Research, and FAIR*”



EVOLVING TOWARDS AN ERA OF  
**OPEN SCIENCE**

# What's our mission?

*The overall objectives*

## **Support**

- *Swedish Life Science researchers shall have access to Data Management support to be able to manage their research data according to international best practice*

## **Training**

- *Swedish Life Science researchers shall have access to training in efficient Data Management practices that can be applied in their daily work*

## **Collaboration**

- *Data Management practices, solutions and training for Swedish Life Science shall be established in collaboration with NBIS internally, and relevant national and international stakeholders*

# Support

## *The overall objective*

*Swedish Life Science researchers shall have access to Data Management support to be able to manage their research data according to international best practice*

## *The approaches*

- ❑ Develop **procedures** for providing efficient Data Management support to research projects throughout the research data lifecycle
- ❑ Provide **data management planning** knowledge and resources/tools for research projects
- ❑ Provide support and resources/tools for **data submission** to recommended community deposition databases
- ❑ Provide support and resources/tools to enable **reproducible research**

# Training

## *The overall objective*

*Swedish Life Science researchers shall have access to training in efficient Data Management practices that can be applied in their daily work*

## *The approaches*

- ❑ Provide Data Management **training** for researchers
- ❑ Provide **guidelines and documentation** for Data Management best practises
- ❑ Promote awareness of the importance of Open Science, Reproducible Research, and FAIR in the Swedish Life Science community

[NBIS DM Team Training Strategy 2023 -- 2026](#)

# DM training strategy

- Target audience
  - Roles & Skill levels
- Topics
- Delivery formats
- Roadmap 2023-26
- Work Plan 2023

## DM team training Strategy 2023 – 2026

Version 1.0

Last updated 2022-12-19 by Niclas Jareborg

Approval date: 2022-12-19

Next time for revision: 2023-10

### Introduction

This document describes the general Data Management Team Training Strategy, with the aims and explicit goals for the team training efforts to be realized during 2023. By the end of 2023 the current training strategy will be evaluated and updated as a continued strategy for the following year. The revised strategy document will be versioned and published along the previous one(s).

*Mission statement for training* - Swedish Life Science researchers shall have access to training in efficient Data Management practices, concepts and practicals, that can be applied in their daily work. The Data Management achievements and efforts at NBIS should be shared with interested collaborators within the ELIXIR network.

Approaches, activities and resources defined as necessary for fulfilling delivery:

- Provide Data Management training for researchers
  - Courses
  - Workshops
- Provide guidelines and documentation for Data Management "best practises"
  - SciLifeLab RDM-guidelines (collaboration with DC)
  - ELIXIR RDMkit (collaboration with DC)
- Promote awareness of the importance of Open Science, Reproducible Research, and FAIR in the Swedish Life Science community
  - Outreach
  - Community building
  - Networking

### Target audience

The primary target audience for the Data Management training strategy are primarily researchers and staff active in Swedish Life Science research. However, several actors on both national and international level who work in close proximity to, or in collaborations with, the primary target audience can become secondary parts in the NBIS DM training efforts.

There is no aim to target activities solely to third party actors (e.g. private companies and industry), unless there is an explicit on-demand request.

### Definitions



# Collaboration

## *The overall objective*

*Data Management practices, solutions and training for Swedish Life Science shall be established in collaboration with NBIS internally, and relevant national and international stakeholders*

## *The approaches*

- ❑ Collaborate with **NBIS staff** to develop Data Management aspects of support procedures and policies
- ❑ Contribute to Data Management activities and capacity building within **ELIXIR**
- ❑ Collaborate with **SciLifeLab Data Centre** to develop Data Management policies, procedures and solutions to support researchers
- ❑ Interact with **other national and international stakeholders** to harmonise Data Management policies, procedures and solutions to support researchers



# RDM Services strategy

- Scope of your target audience? Who do we serve?
- What is your mandate? Expectations from, and dependencies of, external funders/infrastructures/organizations?
- What are the needs of your target audience?
- What services do you want provide to fulfill those needs, given your staffing ? Which services to you want to prioritise? Are some needs solved by other infrastructures / organizations?
- What's your funding model(s)?
- Roadmap for developing RDM services over time
  
- RDM Services maturity model as a tool for strategy planning

# Plenary reflection

- Has anyone done a similar exercise for their Node?
- Would you be able to do so for your Node based on the example of SE?
- How to share our experiences, challenges, solutions when drafting a Node DM strategy?

## Node services maturity model

- **Maturity model:** a framework for measuring an organization's maturity
- **Measuring what?** RDM services, activities, readiness, preparedness, awareness
- **Which organizations?** Research institutions and national or Node level

Strategy and sustainability	Data and metadata management
Legal	Support

## Strategy and sustainability

- Organisational DM strategy
- cost recovery model
- short-to-mid term funding
- long-term sustainability

## Legal

- Data protection
- GDPR assessment
- Data Protection Impact Analysis
- Policies for data management
- Policies for sensitive data infrastructure

## Support

- Training
- Capacity building
- Expert network
- Communications
- Onboarding and guides for DM personas
- Guidelines for researchers (example: RDMkit)
- DM support infrastructure

## Data and metadata management

- Data organisation
  - standards, versioning, licensing
  - file types, folder structure, naming conventions
- Metadata standards (& harmonisation)
  - identifiers, provenance, ontologies and CVs
  - software tools (ELNs, LIMS, EDC systems...)
- Data ingestion
- Data storage
- Data archiving
- Reproducible data analysis
  - Best practices, SOP, Policies
- Data sharing
  - machine actionability, publication
  - brokering
- DMP (support in developing) (example: DSW)
- Risk assessment
- Security breach / incident response plan



# Template discussion

- Are the proposed aspects relevant and useful?
- If gaps, what should we further add?

# Breakout discussions: Node positioning

- Breakout session per 2 Nodes
- Discuss the orange aspects from the model
- Use the table in the Portfolio template to answer the following questions:
  - How do you cater for the aspects of **strategy** and **policies** in your Node?
  - On **training, guidelines, metadata standards, data storage, data sharing** and **DMP**:
    - Are you offering services? Explain
    - If not yet, do you plan to? Explain
    - If out of scope, explain

## Strategy and sustainability

- **Organisational DM strategy**
- cost recovery model
- short-to-mid term funding
- long-term sustainability

## Legal

- Data protection
- GDPR assessment
- Data Protection Impact Analysis
- **Policies for data management**
- Policies for sensitive data infrastructure

## Support

- **Training**
- Capacity building
- Expert network
- Communications
- Onboarding and guides for DM personas
- **Guidelines for researchers** (example: [RDMkit](#))
- DM support infrastructure

## Data and metadata management

- Data organisation
  - standards, versioning, licensing
  - file types, folder structure, naming conventions
- **Metadata standards** (& harmonisation)
  - identifiers, provenance, ontologies and CVs
  - software tools (ELNs, LIMS, EDC systems...)
- Data ingestion
- **Data storage**
- Data archiving
- Reproducible data analysis
  - Best practices, SOP, Policies
- **Data sharing**
  - machine actionability, publication
  - brokering
- **DMP (support in developing)** (example: [DSW](#))
- Risk assessment
- Security breach / incident response plan

IM

# Break

(10.30-11 CEST / 11.30-12 EEST)





# ELIXIR Estonia node

M2 module ELITMa





# A small node, we fit on a unstable pier

We are

- 17ish people (more to join in autumn) in head node Tartu
- 10ish people in partnering universities

# Data management by University of Tartu



**UNIVERSITY OF  
TARTU LIBRARY:**

- Credited course on DMPs for PhDs
- Online material for DMP and Open Science
- Every spring DMP courses (for grant applicants)



**UNIVERSITY OF  
TARTU:**

- Credited course "Data Stewards and Data management Principles";  
Lecturer: Priit Adler

# Data management by ELIXIR Estonia

## ON A UNIVERSITY LEVEL

- 10 different courses in data management
- DMP evaluation consulting for Estonian Research Council
- Data management consultation
- Currently working on a webpage for data management needs tailored for Estonia

# Data management by ELIXIR Estonia

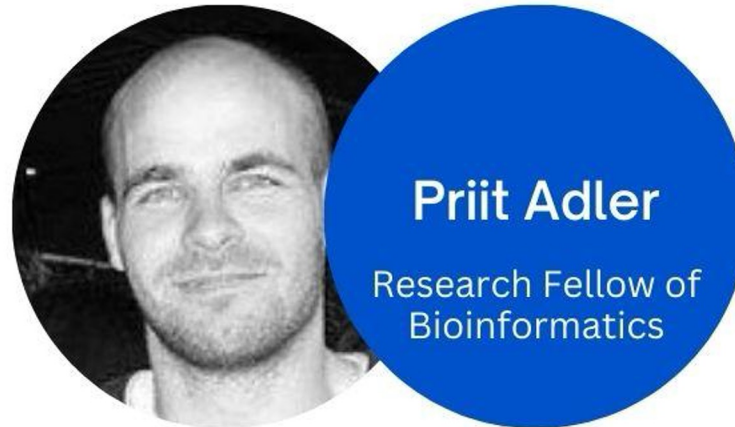
## ON A INSTITUTE LEVEL

- Helping researchers to write DMP
- General DM consultations
- Data managers for different projects

## ON A NODE LEVEL

- Making sure GDPR is being followed
- DM consultations
- In-house training

# Data management team



# Currently

## 3 = 1

One person needs to know how to do it all, both data manager and data steward tasks

- Heleri is on maternity leave
- Priit does the Uni lecture and only carpentry teaching (Git)
- Everything else -> Diana



ONE PERSON FOR EACH TASK

# Ideal ELIXIR-EE DM team



Training



Creating RDM  
policy



RDM counseling



DM community  
building



DM infrastructure



**Thank You**



# Discussing next steps

- What would you need to start/continue working with the model from your Node's perspective to build your Node strategy for DM services?

# RDM Community and wrap up

- Where does this effort sit? Role of the newly established RDM Community and its network
- Where do the Parking Lot items sit?
- Any Other Business?



## RDM Community: Co-leads



Munazah Andrabi  
ELIXIR-UK



Flora D'Anna  
ELIXIR-BE



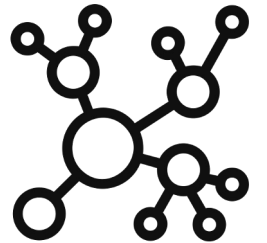
Mijke Jetten  
ELIXIR-NL



Niclas Jareborg  
ELIXIR-SE

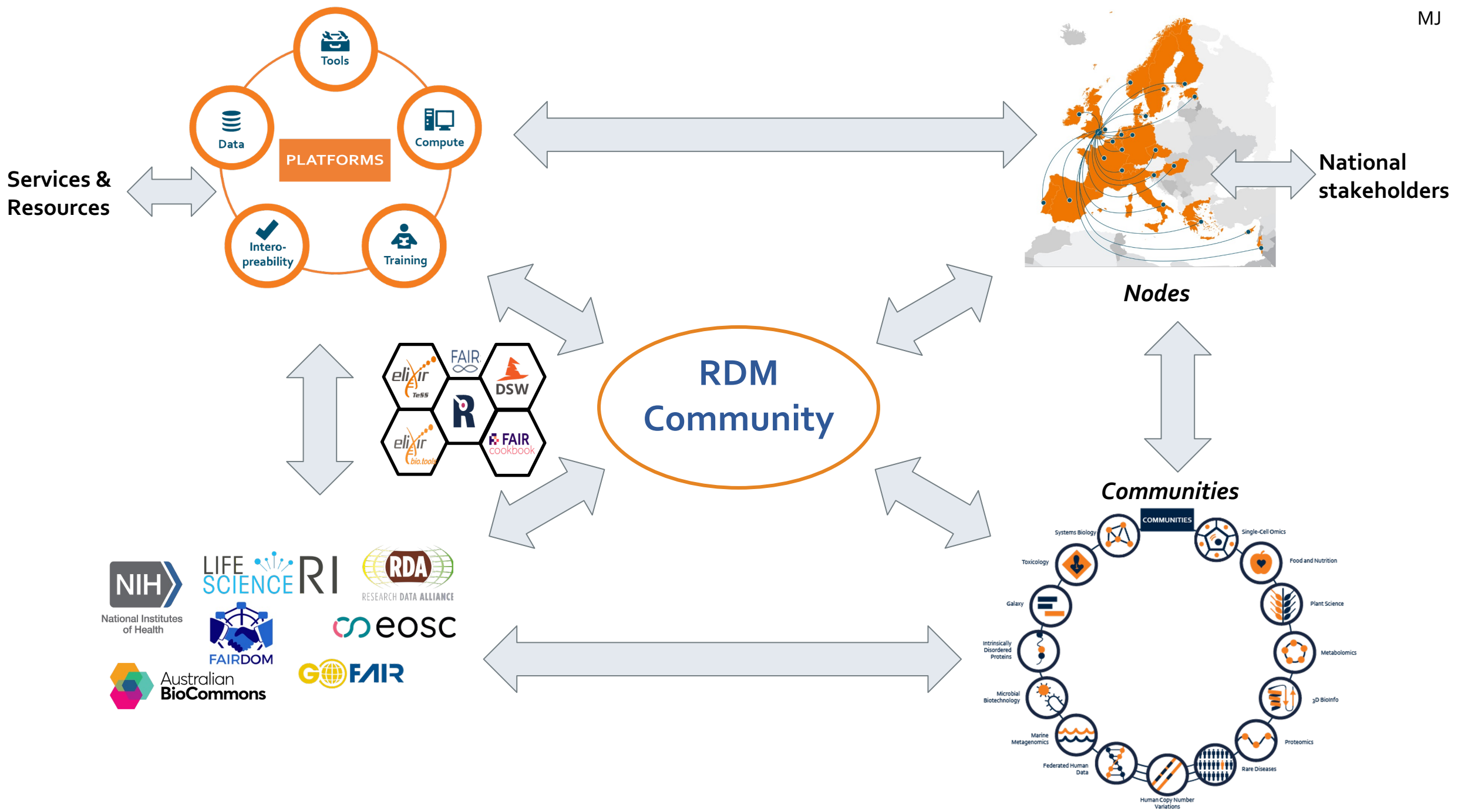
# An ELIXIR RDM Community

- Members identity
  - Life science RDM professionals in ELIXIR and beyond.
  - Inclusive to anyone that has a professional interest in RDM for the life sciences.
- Scope of the Community
  - The network of RDM professionals
    - RDM knowledge exchange and capacity building
  - Management of RDM knowledge
    - Coordination of content management and editorial functions for “The RDM Knowledge Ecosystem” - RDMkit, FAIR Cookbook, DSW
  - Management of RDM training resources and expertise
    - The network of RDM Trainers, RDM learning paths and training resources
  - External stakeholders
    - Coordinate the interactions with external infrastructures, organizations and communities active in RDM



# Out of scope for RDM Community

- Technical service provisioning
  - RDMkit
  - Data Stewardship Wizard
  - FAIR cookbook
- RDM helpdesk support for researchers and consortia
  - DMP writing / reviewing
  - Data depositions support (consultations / hands-on)
- Delivering training programmes
- Curation of data deposited in data resources



## Objectives of the RDM Community- Short term (~2 years) - I

Area	Objective
People network	Provide a forum that enables regular inter-personal knowledge exchange between life science RDM professionals.
	Facilitate ELIXIR Node capacity building in deploying RDM services, e.g. by establishing a knowledge handbook for RDM service providers, <b>and developing an RDM services maturity model for ELIXIR Nodes.</b>
RDM knowledge	Coordinate and contribute to the RDM ecosystem content by collaborating with the editorial boards of RDMkit and FAIR Cookbook, to ensure that the content is continuously fit for purpose.
	Collect requirements that would facilitate flow of FAIR data from researchers to repositories (such as the ELIXIR Deposition Databases) via RDM services, in order to improve the FAIRness of research data made available for life science and society at large.

## Objectives of the RDM Community - Short term (~2 years) - II

Area	Objective
RDM training	Provide a forum that enables knowledge and teaching best practices exchanges among RDM trainers.
	Ensure that RDM training materials are being developed and FAIRified, enabling reuse and adoption by the ELIXIR Nodes.
	Develop RDM learning paths for researchers, data stewards, tool developers and trainers.
External stakeholders	Create and maintain an overview of external stakeholder engagements with RDM aspects across ELIXIR Platforms, Communities, Focus Groups and projects.



# RDM Community: Next steps

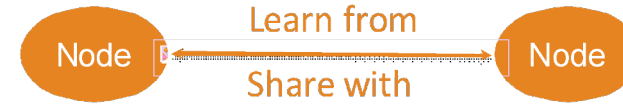
- Implementation study
  - January 2024 – December 2025
  - Draft project plan
    - Organisational Framework
    - Knowledge management and dissemination
    - Capacity building
- F2F meeting during the fall
- Join the Community!
  - **#community\_rdm** slack channel
  - [rdm@elixir-europe.org](mailto:rdm@elixir-europe.org) mailing list ([sign-up link](#))
  - Web page to follow

# RDM Community

- Where does this DM strategy effort sit?
  - ✓ Explicit part of the RDM Community short term goals and Implementation Study.  
*People network: Facilitate ELIXIR Node capacity building in deploying RDM services, e.g. by establishing a knowledge handbook for RDM service providers, and developing an RDM services maturity model for ELIXIR Nodes.*
  - ✓ We will use input today as guidance for RDM Community to further develop the model aspects (template, descriptions etc.)
  - ✓ Next steps: from January 2024 onwards formally via the RDM Community Implementation Study
  - ✓ Don't wait, as you are the RDM Community and it is already active: exchange with your peers (NC, DM) for your Node's DM portfolio next steps
  - ✓ Any need for reminders/follow-up activities, to coordinate your DM portfolio next steps?
- Where do the Parking Lot items sit?
  - ✓ Work that needs to be done anyway by your Nodes, so use the RDM Community to work together/prevent siloes
- Any Other Business?



# Topics missing



## Are there important [topics missing](#)?

- DM Community building (EE, UK, LU, CH, EE, CZ)
  - Create engagement
  - Share experiences
  - Role of emerging RDM community (o.a. financial model, relation to Platforms)
- Policies related to the data management strategies and best practices in different nodes (FI)
- Formal education & to hear from other nodes if there is stories to share to establish formal programs on RDM (NO)
- Monitoring routines for RDM (NO)
- Best practice and emerging patterns in core tool deployments (LU)
- DSW deployments and KM customisations in nodes (LU)
- RDMkit instances, plans, if any, for further instances (LU)
- Guidance documents for our application purposes (DE)
- Show-and-tell sessions to exchange experiences (DE)
- DMP and FAIRification as tightly bound issues (CZ)
- Formal DS curriculum (ECTS awarded with a certificate or degree) (NO, CZ)
  - Role and position Node data manager
  - Training for Node data manager



# Best practices

What are the current [best practices](#) in your Node on the topics above, to bring into the course to inspire others and share experiences?

- Training & webinars around useful resources/knowledge dissemination via communities and ambassadors (UK, FI, BE, IT, NO, CH, EE, CZ)
  - Self-study online course on RDM basics + [National guidelines](#) incl. declarations, policies and recommendations (FI)
  - Fellowship (UK)
  - Interactions with various stakeholders, incl. work regarding financial sustainability (NO)
  - Training course in collaboration with Portugal (CY)
  - Preparation to DMP via training (CH)
  - National resources on RDMkit (CH, EE, CZ)
  - DMPs (DE)
  - DM at institutional level/as part of ELIXIR-CZ consortium (CZ)
  - DS training (NL)
- (Inter)national collaboration:
  - We have a lot of national collaboration on RDM issues under national open science coordination (FI)
  - We collaborate with the Flemish Research Data Network (BE)
  - Currently Italy does not have a national program on data management although several local initiatives are building up (IT)
  - Formal project with data producing RIs on RDM (NO)
  - Currently, the national program on data management is being implemented (CY)
  - DMPs training in national spring/summer schools + international bioinformatics conferences such as ECCB and GCB (DE)
  - Training courses (ELIXIR-CZ consortium partners level so far) (CZ)
  - Collaboration with funders (NL)
  - Very well organisational structures (local and [thematic DCCs](#); [LCRDM](#); [DSIG](#)) (NL)
- Tools:
  - Endzend solution for the stewardship of sensitive data (LU)
  - DSW implementation (SE)
- DM services strategy of the Node (SE)





Thanks a lot for your  
active participation &  
safe travels home

