

ME/CFS & post-COVID
research consortia: **general introduction**

Collaborating on FAIR data for synergy in research

ZonMw, FAIR data workshop, Zwolle, April 9, 2024

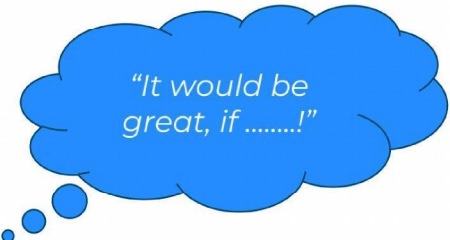
5.1.2e

Implementation lead FAIR data management

5.1.2e

FAIR data & Open Access

Today on the agenda



"It would be great, if!"

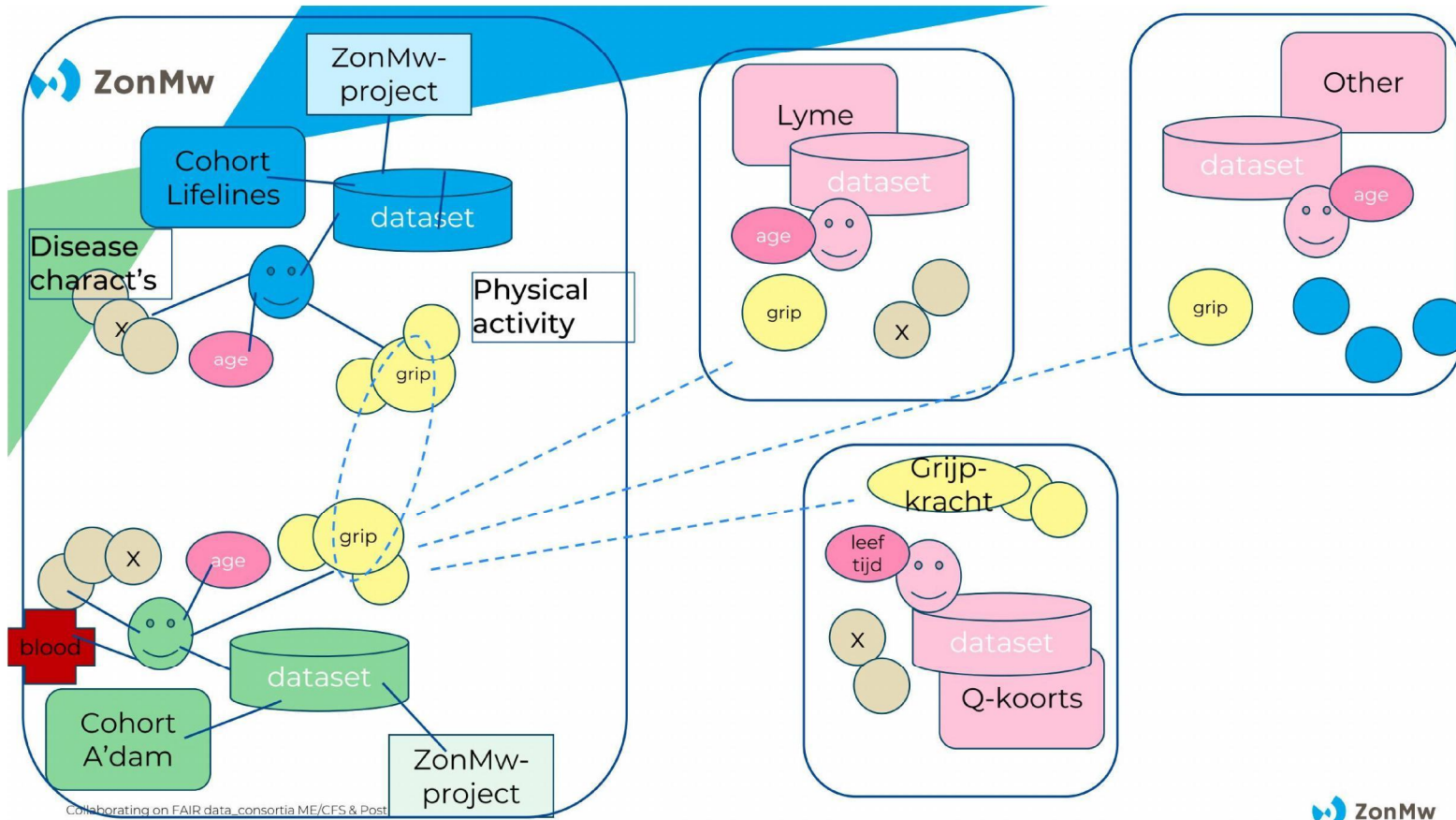
1. Available information > reusable for men and machines (FAIR)
2. The bumpy road to FAIR data
3. It is all about the choices you make as a research community
4. A closer look at data FAIRification
6. From DMPs to a FAIR Implementation Profile
7. Next steps



Session 1:
Explore DMPs to see the standards, tools (etc) that may be useful



Session 2: Compare / review the input from breakout 1, and **choose what fits your consortium best**



Beware of Fake-FAIR!!

FAIR principles

Provide **guidance** for data management:

- F**indable – vindbaar
- A**ccessible – toegankelijk
- I**nteroperable – uitwisselbaar
- R**e-usable – herbruikbaar

*Reusable for both **men** and **machines**
(people and computers)
(**Fully-AI-Ready**)*

FAIR data stay at the source
(data visiting)

FAIR is 90% Metadata

Persistent identifiers
to make metadata
usable for the machine

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

Machine-actionable metadata

Technical infrastructure (accepted generic services)

Social decisions (domain specific agreements)

Blue principles:
researchers & data
stewards

Red principles:
Data stewards,
FAIR data experts

Levels of FAIRness: Examples of what you can do with it?

health RI
enabling data driven health

Projects Dashboard Initiatives

Project overview

Request data Reset

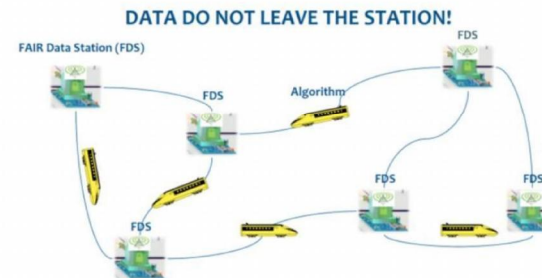
Search

1 to 25 of 193

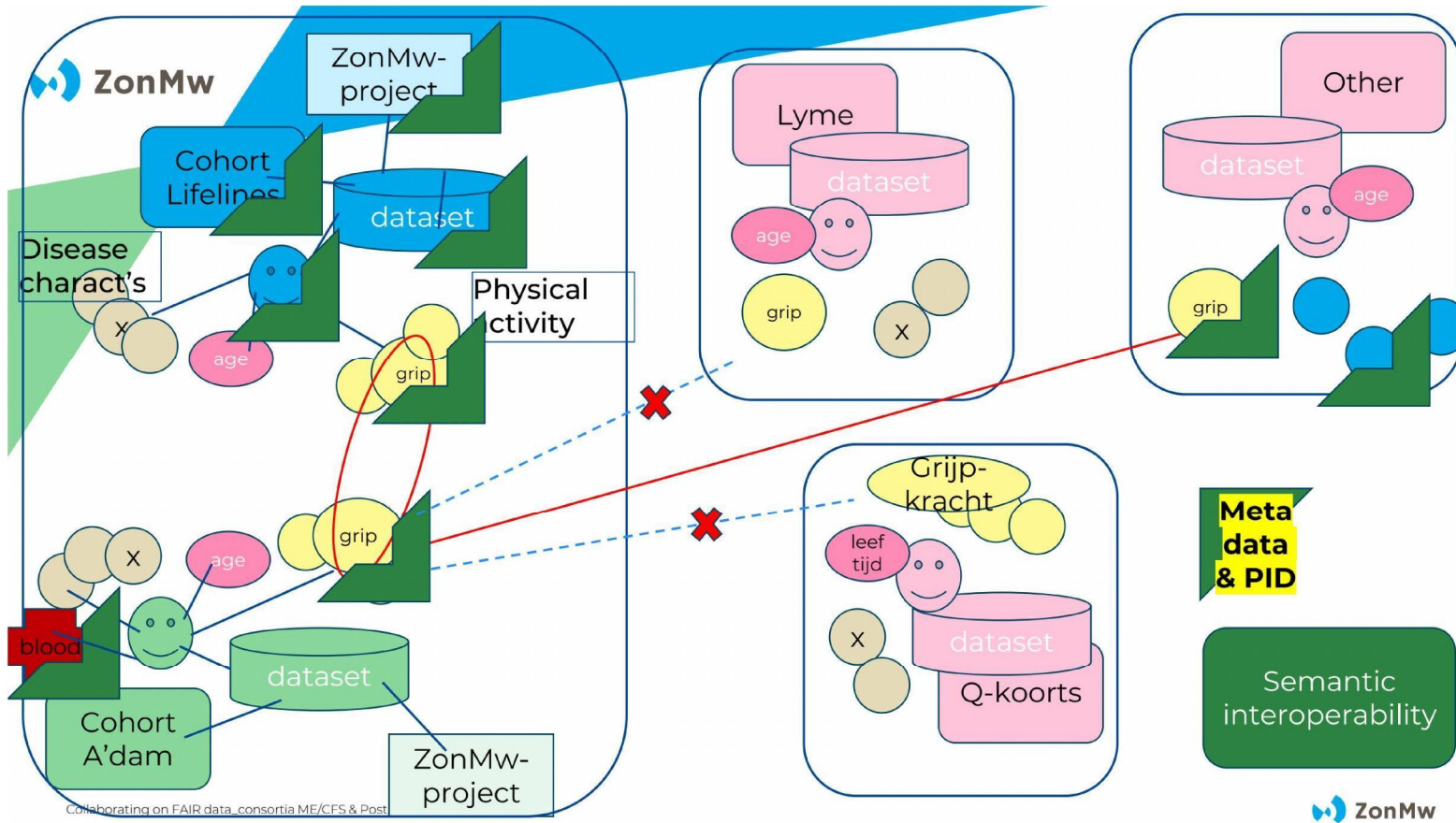
Lead institution	Predicinal development of an antiviral small molecule drug against COVID-19	Last changed 03/04/2024
Data availability	ZonMw, Netherlands Organisation for Health Research and Development, Radboud University Nijmegen	
ZonMw focus area	Model Predictive Pandemic Control	Last changed 02/04/2024
Type of provided assets	ZonMw, Netherlands Organisation for Health Research and Development, Eindhoven University of Technology	
Provided data	Determinants of excess mortality rates in nursing homes	Last changed 20/03/2024
Type of provided biocatal	ZonMw, Netherlands Organisation for Health Research and Development, Eindhoven University of Technology	
Type of provided service	Nasal And Systemic Immunity following COVID Vaccination (NASI-COVAC)	Last changed 22/03/2024

Data portal: data not yet fully FAIR
Project metadata FAIR

FAIR = Fully-AI-Ready



Data visiting > query > personal health train



What steps must we take to improve reusability, interoperability?

What scientific challenges do you want to solve?

- **ME/CFS, post-covid** and its causes
- Disease characteristics, mechanisms
- Diagnosis
- Treatment
- Etc, etc, etc

The research community
must choose!

What choices are relevant for ME/CFS and Post-covid?

- Reusable data? > red/blue principles
- Semantic interoperability? > to align topics, terms

Ad 2 - Consortia meet on the bumpy road to FAIR data



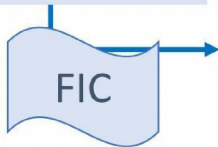
FAIRify all research outputs!
Not only datasets, also e.g.
samples (biobank)

>>

5.1.2e

Ad 4 - A closer look at data FAIRification

Consortia form a
community to
FAIRify

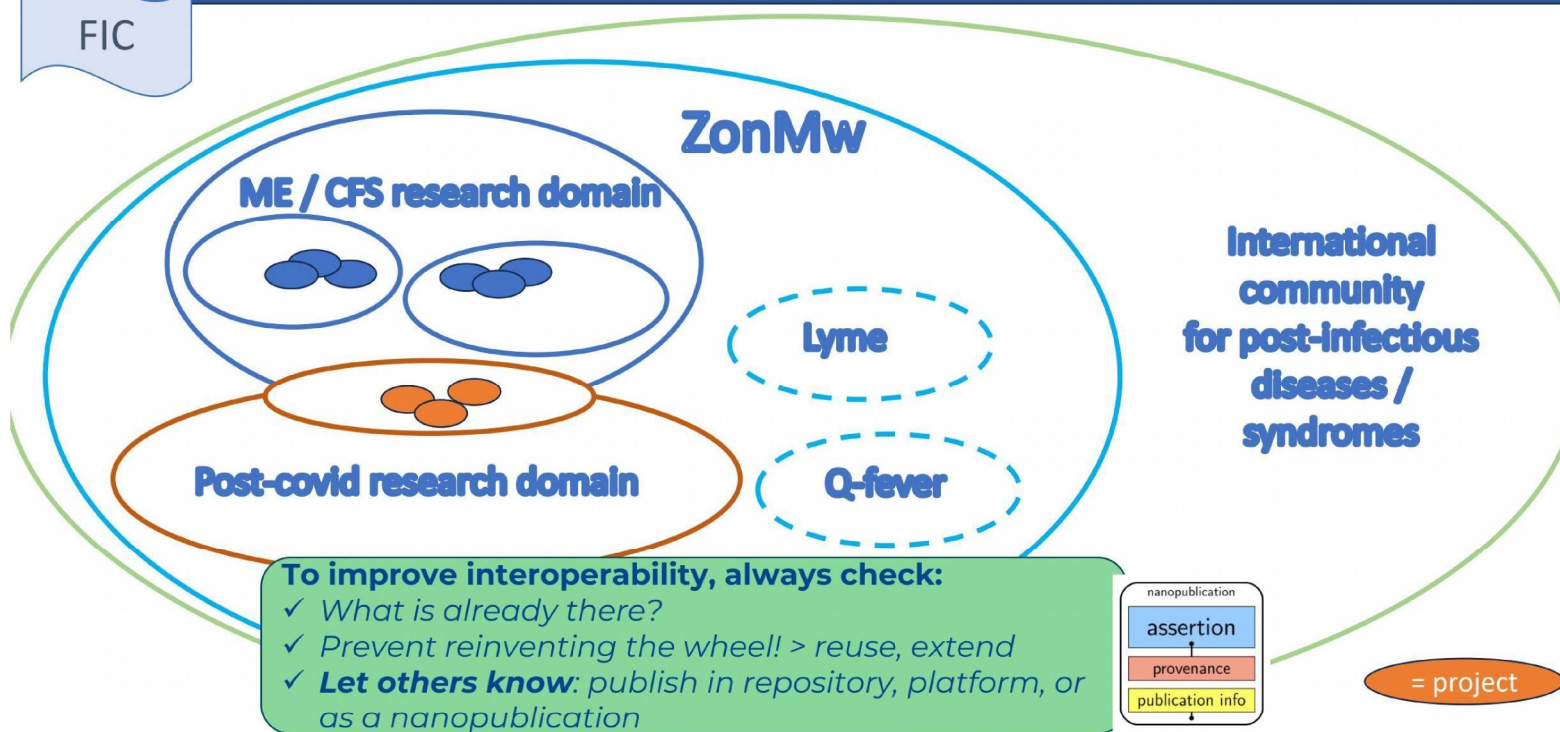


- **What are your goals for reusable / FAIR data?**
- **What do you want to FAIRify?** Datasets, sample collections, cohorts, variables...
- **Who do you need therefore?**
 - Community building, collaborating consortia >
 - **FAIR Implementation Community (FIC)**
 - Expertise from researchers, data stewards and FAIR data experts
- **How will you approach it? Community choices > today:**
 - FAIR Implementation Profile
 - Metadata

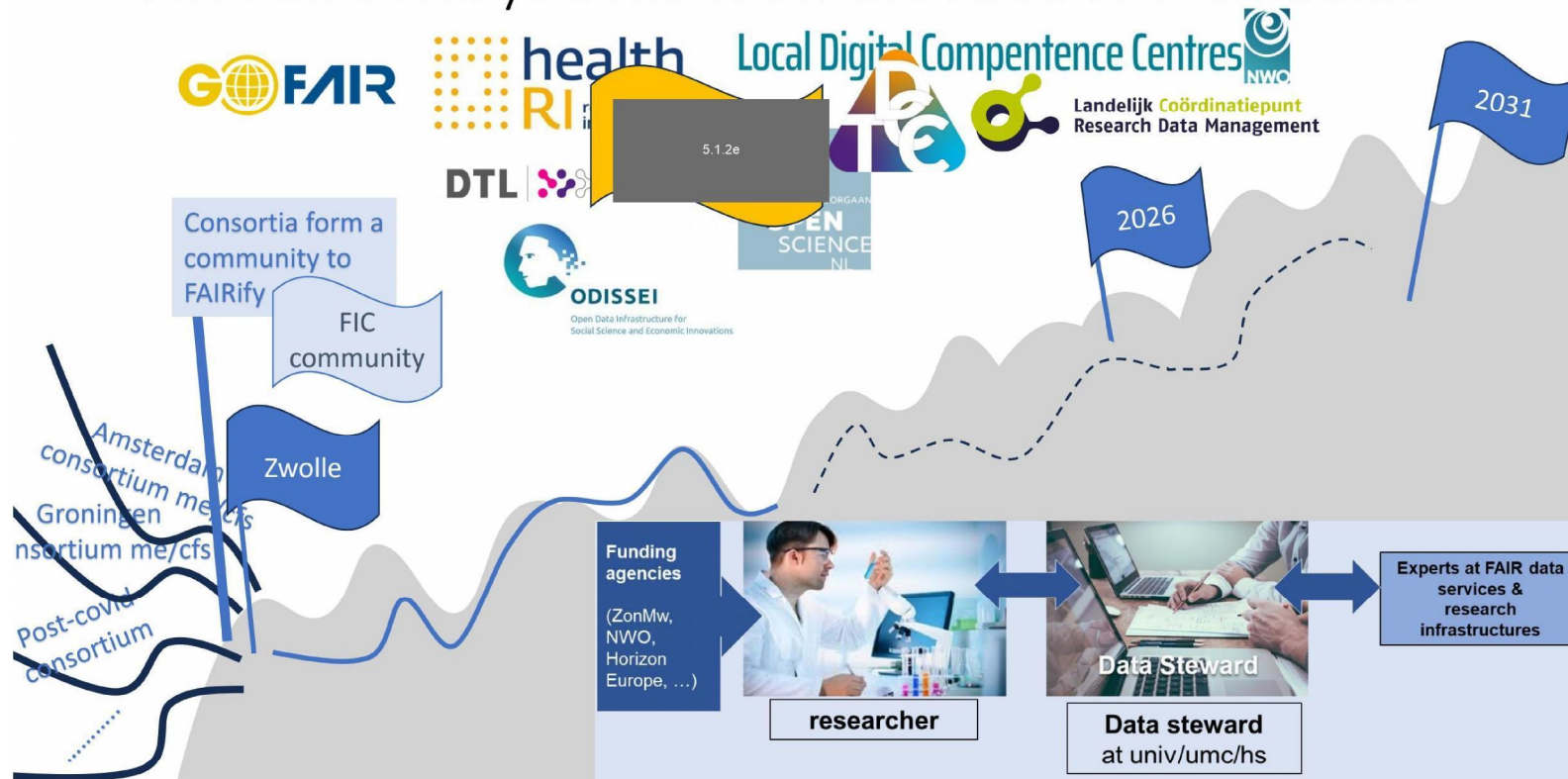
Consortia form a
community to
FAIRify

FIC

The ME/CFS and post-covid FAIR Implementation Community: part of a larger community – including smaller ones



Who else will you meet on the road to FAIR data?



Who will you meet on the road to FAIR data?

GO FAIR Foundation (international)

- [workshops, training, advices](#)
- [FAIR Connect](#) (including nanopublication search engine)

Health-RI (Netherlands)

- [Health-RI nodes: FAIR Data](#)
- > FAIR coordinatoren en regionale knooppunten
- > zie ook presentatie 5.1.2e met contactpersonen
- [Services & tools](#)
- [Data Steward \(DSIG\) Network - TDCC.nl](#)
- [COVID-19 dataportal](#) (will be extended to include all health-research topics)

Background information - guidance:

[FAIR Cookbook and the notion of metadata \(elixir-europe.org\)](#)

[RDMkit \(elixir-europe.org\)](#)

An
consor
Gronin
nsortium
Post-covid
consortium

(ZonMw,
NWO,
Horizon
Europe, ...)



researcher



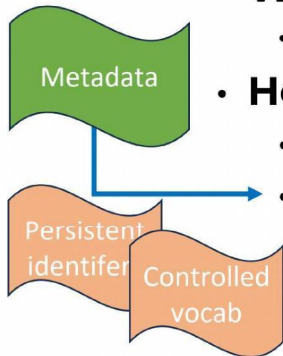
Data steward
at univ/umc/hs

Experts at FAIR data
services &
research
infrastructures

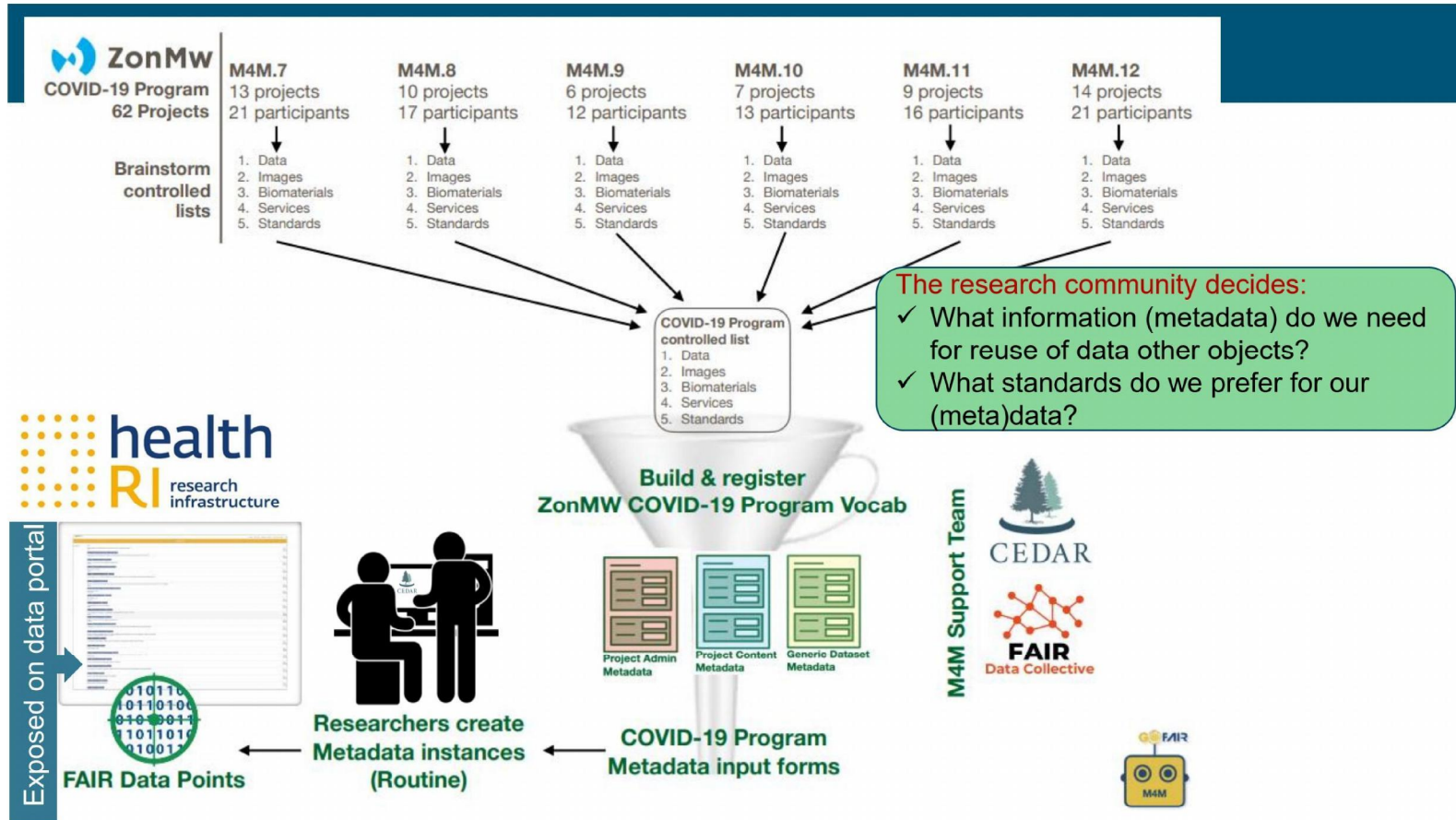
2031

Ad 4 - A closer look at data FAIRification

- **What are your goals for reusable / FAIR data?**
- **What do you want to FAIRify?** Datasets, sample collections, cohorts, variables...
- **Who do you need therefore?**
 - FIC & community, expertise
- **How will you approach it? Community choices > today:**
 - FAIR Implementation Profile > working sessions
 - Metadata



Collaborating on FAIR data_consortia ME/CFS & Post-covid_April 9, 2024





Metadata Forms for ZonMw ID & AMR

Guidelines for completing CEDAR forms

[Link to the guidelines document](#)

CEDAR forms [forms to fill out, CEDAR account required]

Project Admin form
Project Content form
Collection Admin form
Collection Content form

CEDAR forms in OpenView [forms to view/share, no CEDAR account]

Project Admin form in OpenView
Project Content form in OpenView
Collection Admin form in OpenView
Collection Content form in OpenView

BioPortal vocabularies

ID & AMR Vocabulary: AIR data_consortia ME/CFS & Post-covid_April 9, 2024

BioPortal Ontologies Search Annotator Recommender Mappings

Infectious Diseases and Antimicrobial Resistance

Last updated: September 30, 2021

Summary Classes Properties Notes Mappings Widgets

Jump to:

- * access condition
- * animal disease
- * centrifugation information
- * collection status
- * consent restriction
- * data
- * data quality
- * deprecated
- * detailed type of specimen
- * domain
- * environment
- * focus area
- * funding type
- * isolation source
- * macromolecule
- * organisation level
- * organism
 - * animal
 - * microorganism
 - * plant
- * plant disease
- * population group
- * quality control of biomaterials
 - * amplifiability of nucleic acids
 - * concentration of molecular or cellular derivatives
 - * integrity of nucleic acids
 - * purity of nucleic acids
- * quality control of microbiological material
- * quality management of pre-analytical data
- * specimen
 - * biosys sample
 - * blood specimen**
 - * body fluid sample
 - * bronchoalveolar lavage fluid sample
 - * cerebrospinal fluid sample
 - * respiratory sample
 - * skin swab
 - * specimen from genital system
 - * specimen of unknown material
 - * stool specimen
 - * synovial fluid specimen
 - * urine specimen
- * storage information
- * temporal scope
- * time frame
- * viability frequency

Details Visualization Notes (0) Class Mappings (17)

Preferred Name	blood specimen
Synonyms	
ID	http://purl.org/zonmw/id-amr/10069
broader	specimen
exactMatch	http://purl.bioontology.org/ontology/SNOMEDCT/119297000
inScheme	http://purl.org/zonmw/id-amr/
prefLabel	blood specimen

Persistent identifier (PURL)

Persistent identifier (PURL) for the controlled term > the computer can read it now!

Controlled vocab

The thing you measure Or The sample you collected

Class Search

Enter a class, e.g., Melanoma

[Show advanced options >>](#)

Matches in 29 ontologies

CFS-1686 - Medical Subject Headings (MESH)

<http://purl.bioontology.org/ontology/MESH/C000599713>[details](#) - [visualize](#) - [26 more from this ontology](#)

Check

- ✓ *What is already there*
- ✓ *(consider to) reuse, extend, publish*

Reuses in other ontologies

PGP-Me - Robert Hoehndorf Version of MeSH (RH-MESH)

<http://phenomebrowser.net/ontologies/mesh/mesh.owl#PC407124>[details](#) - [visualize](#) - [16 more from this ontology](#)

Chronisches Erschöpfungssyndrom (CFS) - Disease core ontology applied to Rare Diseases (HRDO)

http://www.jimics.org/hrdo/dfns#pat_id_790[details](#) - [visualize](#)

chronic fatigue syndrome (CFS) - Ontology of Consumer Health Vocabulary (OCHV)

<http://sbmi.uth.tmc.edu/ontology/ochv#C0015674>[details](#) - [visualize](#) - [1 more from this ontology](#)

Myalgic encephalomyelitis/chronic fatigue syndrome - International Classification of Diseases, Version 10 - Clinical Modifica

Class Search

Enter a class, e.g. Melanoma

[Show advanced options >>](#)[Search](#)

Matches in 21 ontologies

[Post-COVID Conditions \(Long COVID\) - MedlinePlus Health Topics \(MEDLINEPLUS\)](#)<http://purl.bioontology.org/ontology/MEDLINEPLUS/C5706081>

What are post-COVID conditions (long COVID)? COVID-19 (coronavirus disease 2019) is an illness caused by a virus. Many people infected with the virus. But ...

[details](#) - [visualize](#) - [2 more from this ontology](#)[Post-COVID syndrome - Medical Dictionary for Regulatory Activities Terminology \(MedDRA\) \(MEDDRA\)](#)<http://purl.bioontology.org/ontology/MEDDRA/10085505>[details](#) - [visualize](#) - [24 more from this ontology](#)[Post COVID-19 condition - International Classification of Diseases, Version 10 - Clinical Modification \(ICD10CM\)](#)<http://purl.bioontology.org/ontology/ICD10CM/U09>[details](#) - [visualize](#) - [2 more from this ontology](#)[Post-acute COVID-19 - SNOMED CT \(SNOMEDCT\)](#)<http://purl.bioontology.org/ontology/SNOMEDCT/1119303003>

Symptoms attributed to SARS-CoV-2 infection that persist for more than three weeks following onset. Symptoms attributed to SARS-CoV-2 infection that persist for more than three weeks following onset and with a ...

[details](#) - [visualize](#) - [5 more from this ontology](#)[Post-COVID-19 Symptom - National Cancer Institute Thesaurus \(NCIT\)](#)<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C189195>

A symptom reported following recovery from COVID-19.

Check

- ✓ *What is already there*
- ✓ *(consider to) reuse, extend, publish*

Metadata for variables > lecture

5.1.2e

Check

- ✓ *What is already there > **for example***
- 5.1.2e ***variables!***
- ✓ *(consider to) reuse, extend, publish*

npj | digital medicine www.nature.com/npjdigitalmed

ARTICLE OPEN Check for updates

Harmonization and standardization of data for a pan-European cohort on SARS- CoV-2 pandemic

5.1.2e 5.1.2e 5.1.2e

The European project ORCHESTRA intends to create a new pan-European cohort to rapidly advance the knowledge of the effects and treatment of COVID-19. Establishing processes that facilitate the merging of heterogeneous clusters of retrospective data was an essential challenge. In addition, data from new ORCHESTRA prospective studies have to be compatible with earlier collected information to be efficiently combined. In this article, we describe how we utilized and contributed to existing standard terminologies to create consistent semantic representation of over 2500 COVID-19-related variables taken from three ORCHESTRA studies. The goal is to enable the semantic interoperability of data within the existing project studies and to create a common basis of standardized elements available for the design of new COVID-19 studies. We also identified 743 variables that were commonly used in two of the three prospective ORCHESTRA studies and can therefore be directly combined for analysis purposes. Additionally, we actively contributed to global interoperability by submitting new concept requests to the terminology Standards Development Organizations.

npj Digital Medicine (2022)5:75 ; <https://doi.org/10.1038/s41746-022-00620-x>

What we showed until now:

