



# Metadata for Machines (M4M): Metadata Templates to Support FAIR Principles for Observational Health Research

EFMI STC 2021 - Workshop

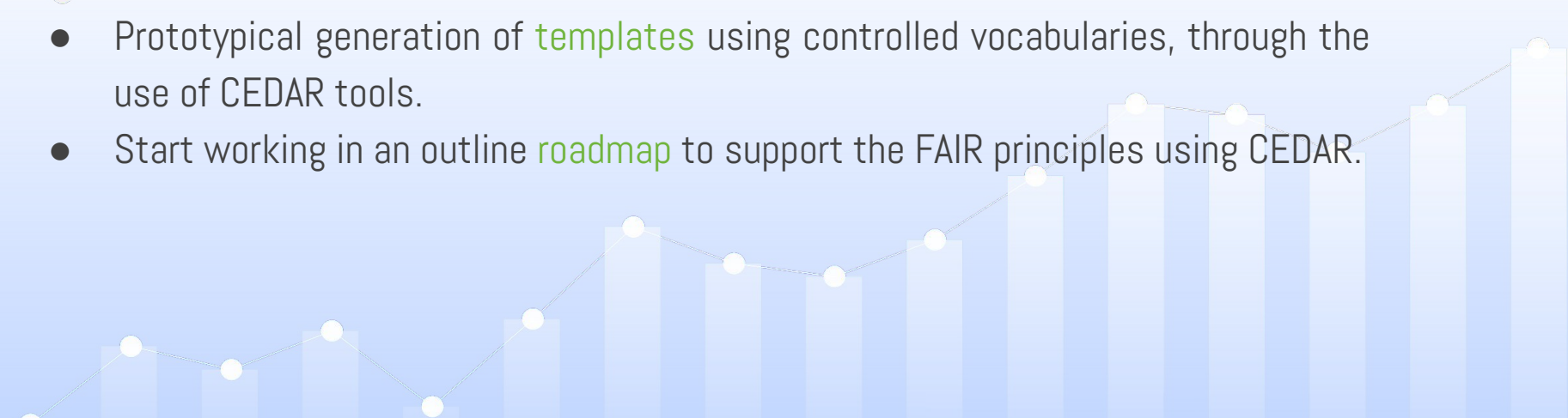
Alicia MARTÍNEZ-GARCÍA, Matthias LÖBE, Ronald CORNET, Mark A. MUSEN, John GRAYBEAL, A. Anil SINACI, Celia ALVAREZ-ROMERO, Erik SCHULTES, Christian DRAEGER, Erika LOGGIN, Carlos L. PARRA-CALDERÓN

23rd November 2021

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# Workshop objectives

- Promote **collaboration** for better metadata to improve the FAIRness of health data sets.
- Prototypical generation of **templates** using controlled vocabularies, through the use of CEDAR tools.
- Start working in an outline **roadmap** to support the FAIR principles using CEDAR.

A decorative bar chart with 12 light blue bars of varying heights, overlaid with a grey line graph and white circular markers. The chart is positioned in the background of the lower half of the slide.


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# Outline

1. Workshop objectives and panellists
2. FAIR4Health project
3. M4M previous workshops (The Rapid M4M)
4. Building FAIR metadata with CEDAR
5. Test the use / requirements of CEDAR tools for creation of metadata templates in specific case studies
  - Metadata generated in FAIR4Health project
  - NFDI4Health C19 metadata schema
  - Health-RI / ZonMw (Covid-19 scenarios)
6. Discussion & AOB

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# Workshop panellists

A decorative bar chart with 12 light blue bars of varying heights. A thin grey line with white circular markers connects the tops of the bars, showing an overall upward trend. The chart is positioned in the background of the slide.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

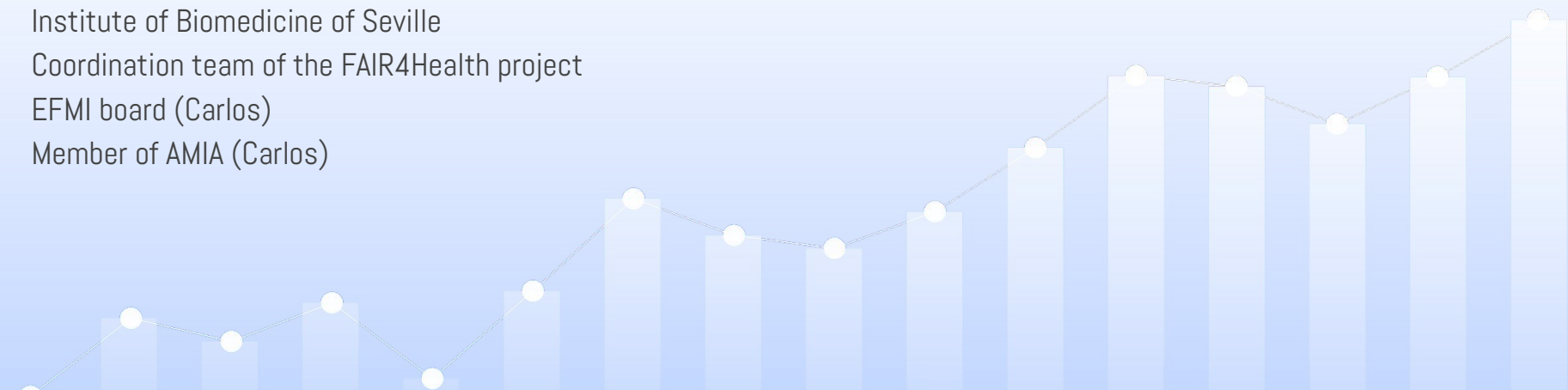


# Panellists

Alicia MARTÍNEZ-GARCÍA  
Celia ALVAREZ-ROMERO  
Carlos L. PARRA-CALDERÓN



Andalusian Health Service (SAS) - Virgen del Rocío University Hospital  
Institute of Biomedicine of Seville  
Coordination team of the FAIR4Health project  
EFMI board (Carlos)  
Member of AMIA (Carlos)

A decorative bar chart with 12 light blue bars of varying heights, overlaid with a white line graph connecting white circular markers.

This project has received funding from the European Union's Horizon 2020  
research and innovation programme under grant agreement No 824666

# Panellists

Matthias LÖBE  
Christian DRAEGER



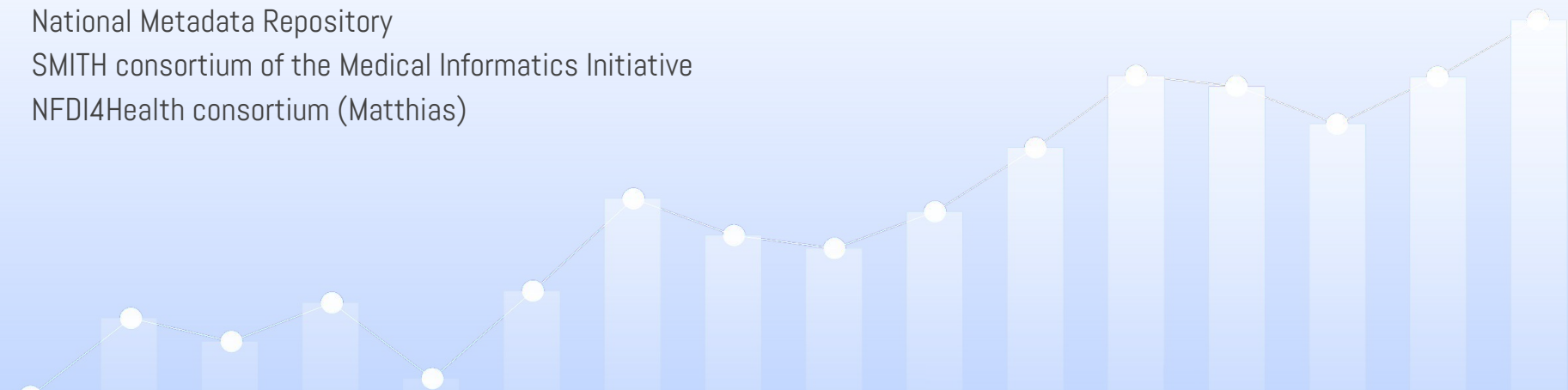
Institute for Medical Informatics, Statistic and Epidemiology (IMISE), University of Leipzig

FAIR4Health consortium

National Metadata Repository

SMITH consortium of the Medical Informatics Initiative

NFDI4Health consortium (Matthias)

A decorative bar chart with 12 light blue bars of varying heights, overlaid with a grey line graph and white circular markers.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# Panellists

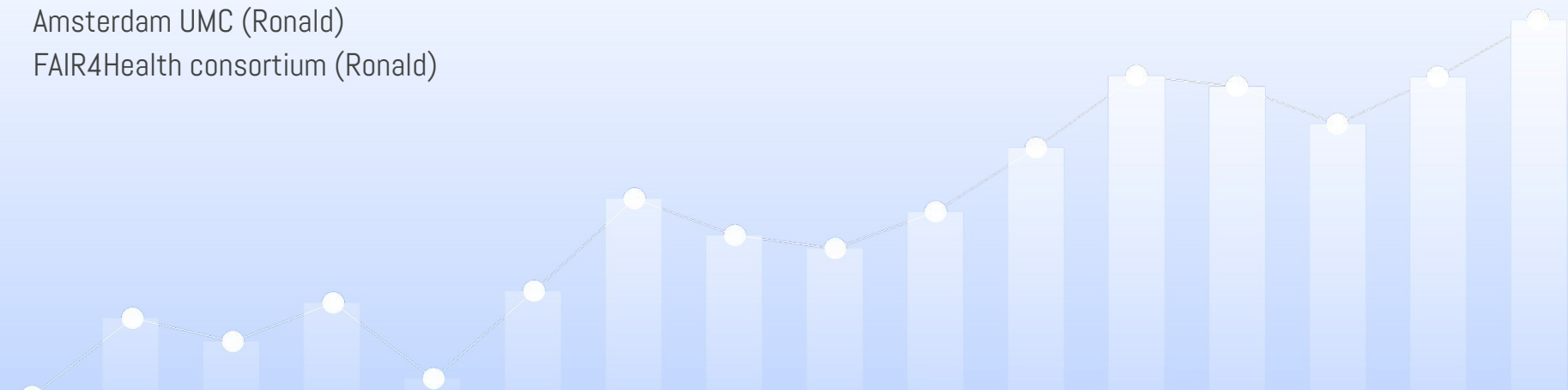
Ronald CORNET  
Erik SCHULTES



GoFAIR

Amsterdam UMC (Ronald)

FAIR4Health consortium (Ronald)

A decorative bar chart with 12 light blue bars of varying heights, overlaid with a grey line graph and white circular markers. The chart is positioned in the background of the bottom section of the slide.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# Panellists

Mark A. MUSEN  
John GRAYBEAL

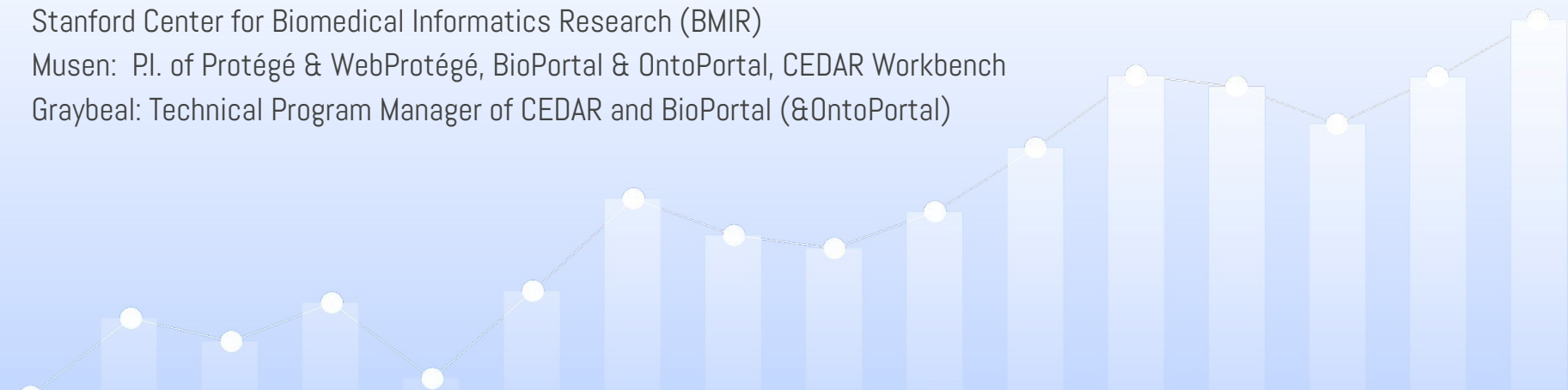


Stanford University School of Medicine

Stanford Center for Biomedical Informatics Research (BMIR)

Musen: P.I. of Protégé & WebProtégé, BioPortal & OntoPortal, CEDAR Workbench

Graybeal: Technical Program Manager of CEDAR and BioPortal (&OntoPortal)

A decorative bar chart with 12 light blue bars of varying heights, overlaid with a line graph showing an upward trend. The chart is positioned in the background of the bottom section of the slide.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

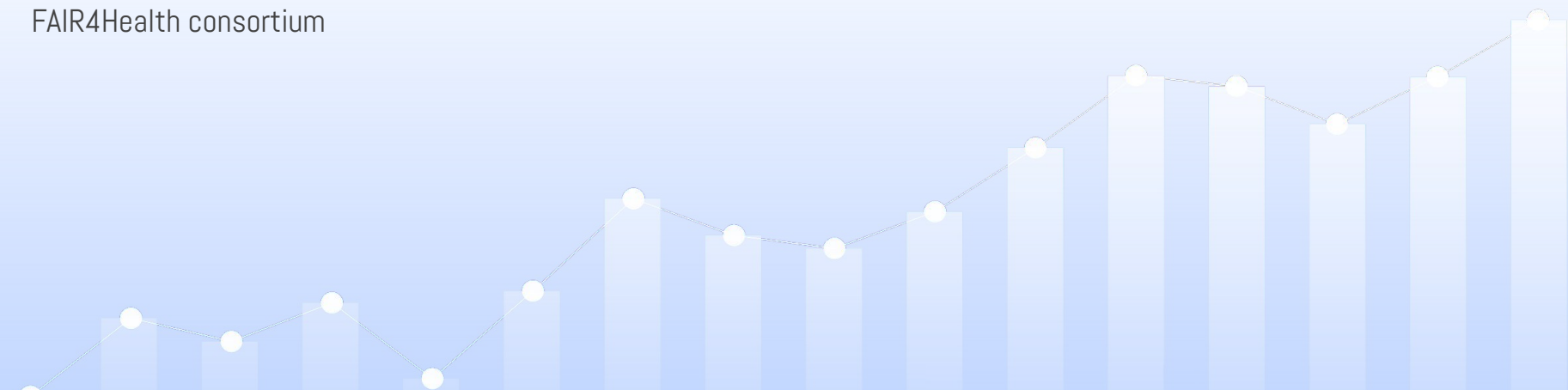
# Panellists

A. Anil SINACI



SRDC Software Research Development and Consultancy Corporation, Turkey

FAIR4Health consortium

A decorative bar chart with 12 light blue bars of varying heights, overlaid with a line graph showing an upward trend. The chart is positioned in the background of the bottom half of the slide.

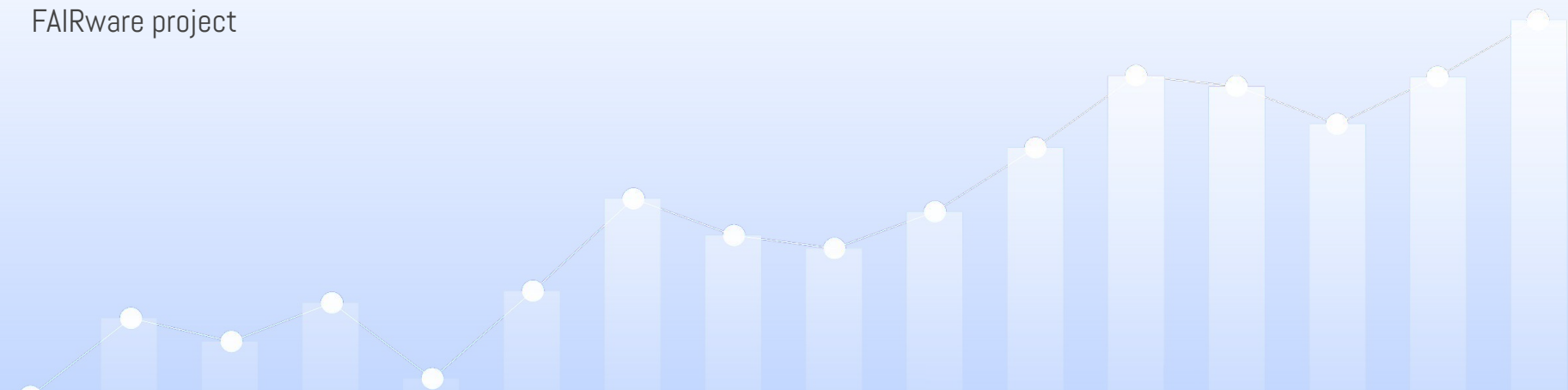
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# Panellists

Erika LOGGIN



Wellcome Trust  
FAIRware project



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# Outline

1. ~~Workshop objectives and panellists~~
2. FAIR4Health project
3. M4M previous workshops (The Rapid M4M)
4. Building FAIR metadata with CEDAR
5. Test the use / requirements of CEDAR tools for creation of metadata templates in specific case studies
  - Metadata generated in FAIR4Health project
  - NFDI4Health C19 metadata schema
  - Health-RI / ZonMw (Covid-19 scenarios)
6. Discussion & AOB

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# FAIR4Health project



[www.fair4health.eu](http://www.fair4health.eu)



<https://www.linkedin.com/company/13977340/>



@FAIR4Health  
#FAIR4Health



<https://www.youtube.com/channel/UCpycUIqaXMAJCZPatqcm4cq>

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

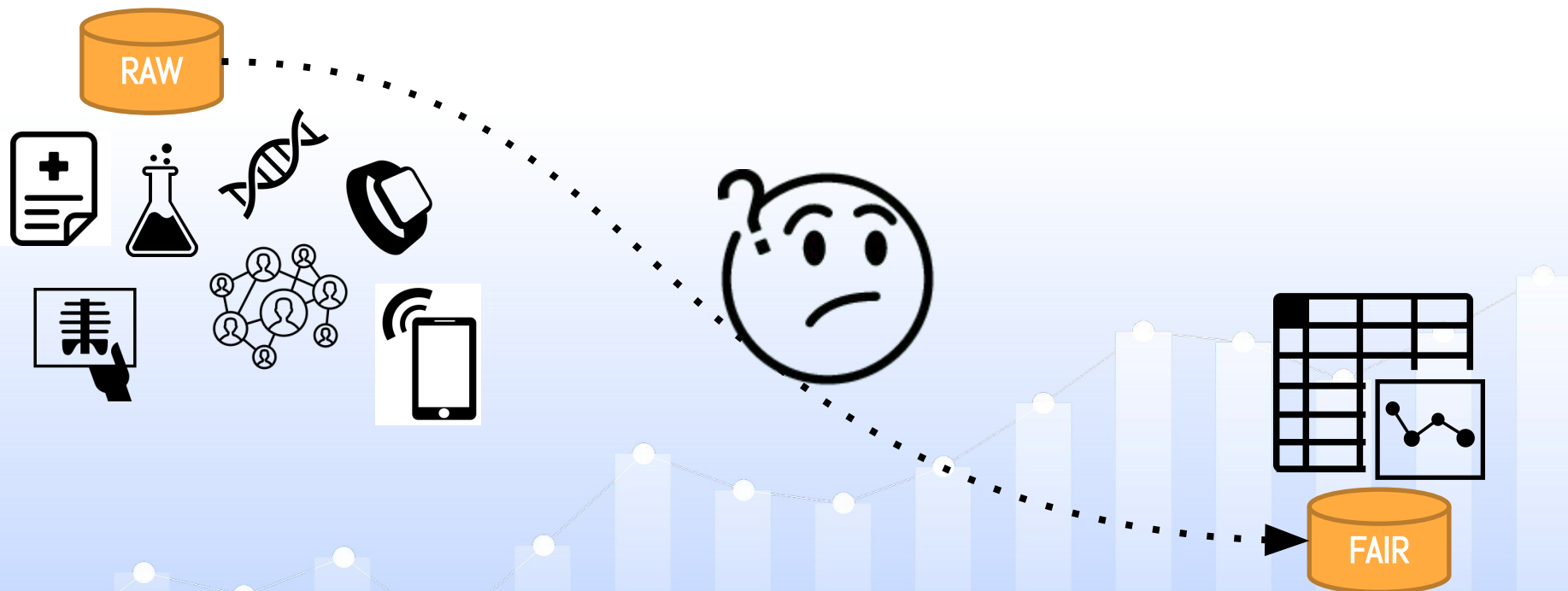


To **facilitate and encourage** the EU Health Research community to **FAIRify, share and reuse** their datasets derived from publicly funded research initiatives through the demonstration of the potential impact that such strategy will have on health outcomes and health research.



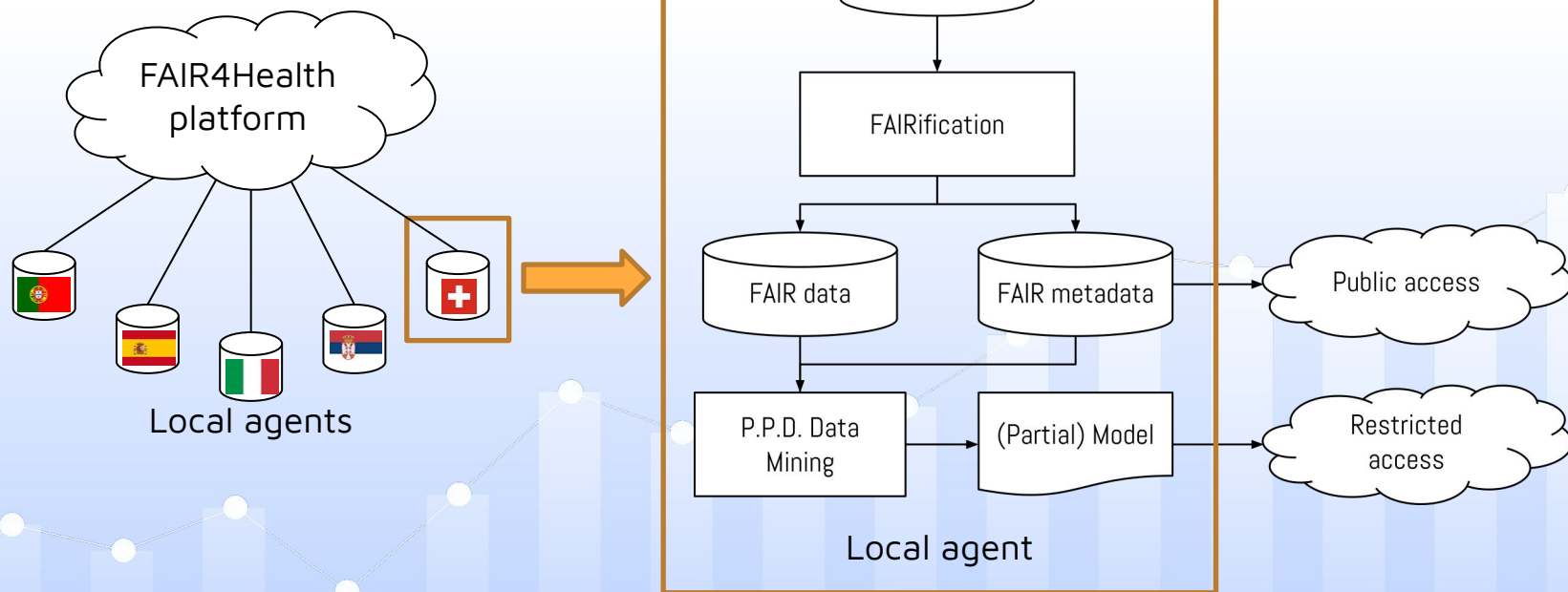
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# FAIRification Workflow



Sinaci, A. A., Núñez-Benjumea, F. J., Gencturk, M., Jauer, M. L., Deserno, T., Chronaki, C., ... & Erturkmen, G. B. L. (2020). From Raw Data to FAIR Data: The FAIRification Workflow for Health Research. *Methods of Information in Medicine*, 59(S 01), e21-e32.

# Technological solution



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# FAIR4Health project: Metadata

## Welcome to FAIR4Health Platform.

Research study	Description	Total number of sites	Total number of patients	AI Algorithm
1	Identification of multimorbidity patterns and polypharmacy correlation on the risk of mortality in elderly, and demonstrate the reproducibility of research	5	11.486	FP Growth
2	Develop and pilot an early prediction service for 30-days readmission risk in COPD (Chronic Obstructive Pulmonary Disease) patients	3	4.944	Support Vector Machine (SVM), Logistic Regression, Decision Trees, Random Forest, Gradient Boosted Trees

**Common Data model:** <https://github.com/fair4health/common-data-model>

**Owner:** [FAIR4Health Consortium](#)

**FAIR4Health FAIR metadata repository:** <https://github.com/fair4health/metadata>

[+ Info](#)
[Contact](#)

© Fair4Health H2020 Project.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# FAIR4Health project: Metadata



Use case: Use Case 1

1 Name & description		2 Feature set selection				3 Eligibility criteria				4 Results & statistics				5 Completed
agent	# of records	Age	Gender	Smoking status	Institutionalized	Nationality	Heart failure	Coronary heart disease	Hypertension	Atrial fibrillation	Pulmonary hypertension	Chronic kidney disease	Prostatic hypertrop	
<div><div><div><div><div></div></div></div><div><div><div></div></div></div></div><div><div><div></div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div>	5812	Min: 65 Max: 104 Null: 0 <a href="#">See all distributions</a> <a href="#">See distributions</a>	female: 2868 male: 2944 <a href="#">See distributions</a>	FALSE: 5788 TRUE: 24 <a href="#">See distributions</a>	FALSE: 5530 TRUE: 282 <a href="#">See distributions</a>	724: 5798 170: 3 380: 2 504: 2 Others: 7 <a href="#">See distributions</a>	FALSE: 3653 TRUE: 2159 <a href="#">See distributions</a>	FALSE: 4456 TRUE: 1356 <a href="#">See distributions</a>	FALSE: 1165 TRUE: 4647 <a href="#">See distributions</a>	FALSE: 3918 TRUE: 1894 <a href="#">See distributions</a>	FALSE: 5341 TRUE: 471 <a href="#">See distributions</a>	FALSE: 4483 TRUE: 1329 <a href="#">See distributions</a>	FALSE: 5200 TRUE: 612 <a href="#">See distributi</a>	
<div><div><div><div><div></div></div></div><div><div><div></div></div></div></div><div><div><div></div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div>	331	Min: 68 Max: 116 Null: 0 <a href="#">See all distributions</a> <a href="#">See distributions</a>	female: 237 male: 94 <a href="#">See distributions</a>	FALSE: 329 TRUE: 2 <a href="#">See distributions</a>	FALSE: 331 <a href="#">See distributions</a>	CS: 183 DE: 63 IT: 62 IL: 20 Others: 3 <a href="#">See distributions</a>	FALSE: 256 TRUE: 75 <a href="#">See distributions</a>	FALSE: 129 TRUE: 202 <a href="#">See distributions</a>	FALSE: 331 <a href="#">See distributions</a>	FALSE: 331 <a href="#">See distributions</a>	FALSE: 331 <a href="#">See distributions</a>	FALSE: 331 <a href="#">See distributions</a>	FALSE: 331 <a href="#">See distributi</a>	
<div><div><div><div><div></div></div></div><div><div><div></div></div></div></div><div><div><div></div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div>	244	Min: 66 Max: 104 Null: 0 <a href="#">See all distributions</a> <a href="#">See distributions</a>	female: 115 male: 129 <a href="#">See distributions</a>	FALSE: 244 <a href="#">See distributions</a>	FALSE: 227 TRUE: 17 <a href="#">See distributions</a>	CH: 183 IT: 13 ES: 8 FR: 7 Others: 33 <a href="#">See distributions</a>	FALSE: 243 TRUE: 1 <a href="#">See distributions</a>	FALSE: 189 TRUE: 55 <a href="#">See distributions</a>	FALSE: 114 TRUE: 130 <a href="#">See distributions</a>	FALSE: 173 TRUE: 71 <a href="#">See distributions</a>	FALSE: 230 TRUE: 14 <a href="#">See distributions</a>	FALSE: 207 TRUE: 37 <a href="#">See distributions</a>	FALSE: 221 TRUE: 23 <a href="#">See distributi</a>	
<div><div><div><div><div></div></div></div><div><div><div></div></div></div></div><div><div><div></div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div> <div><div><div></div></div></div>	3786	Min: 65 Max: 106 Null: 0 <a href="#">See all distributions</a> <a href="#">See distributions</a>	female: 1889 male: 1897 <a href="#">See distributions</a>	FALSE: 3786 <a href="#">See distributions</a>	FALSE: 3786 <a href="#">See distributions</a>	724: 3724 250: 11 642: 7 504: 5 Others: 39 <a href="#">See distributions</a>	FALSE: 2877 TRUE: 909 <a href="#">See distributions</a>	FALSE: 3001 TRUE: 785 <a href="#">See distributions</a>	FALSE: 909 TRUE: 2877 <a href="#">See distributions</a>	FALSE: 2751 TRUE: 1035 <a href="#">See distributions</a>	FALSE: 3502 TRUE: 284 <a href="#">See distributions</a>	FALSE: 2968 TRUE: 818 <a href="#">See distributions</a>	FALSE: 3479 TRUE: 307 <a href="#">See distributions</a>	

# FAIR4Health project: Metadata

agent	# of records	Age	Gender	Smoking status	Institutionalized
 FAIR4Health PPDDM-Agent for 	5812	Min: 65 Max: 104 Null: 0 <a href="#">See all distributions</a> <a href="#">See distributions</a>	female: 2868 male: 2944 <a href="#">See distributions</a>	FALSE: 5788 TRUE: 24 <a href="#">See distributions</a>	FALSE: 5530 TRUE: 282 <a href="#">See distributions</a>
 FAIR4Health PPDDM-Agent for 	331	Min: 68 Max: 116 Null: 0 <a href="#">See all distributions</a> <a href="#">See distributions</a>	female: 237 male: 94 <a href="#">See distributions</a>	FALSE: 329 TRUE: 2 <a href="#">See distributions</a>	FALSE: 331 <a href="#">See distributions</a>

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# FAIR4Health project: Metadata

**Age** ×

Value.	Count	Percentage
88.0	299	5.14 %
87.0	298	5.13 %
85.0	290	4.99 %
86.0	289	4.97 %
Others	4636	79.77 %

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# FAIR4Health project: Metadata

See all distribution

Age		
Value	Count	Percentage
88.0	299	5.14 %
87.0	298	5.13 %
85.0	290	4.99 %
86.0	289	4.97 %
Others	4636	79.77 %

Gender		
Value	Count	Percentage
female	2868	49.346 %
male	2944	50.654 %

Smoking status		
Value	Count	Percentage
0.0	5788	99.587 %
1.0	24	0.413 %

Institutionalized		
Value	Count	Percentage
0.0	5530	95.148 %
1.0	282	4.852 %

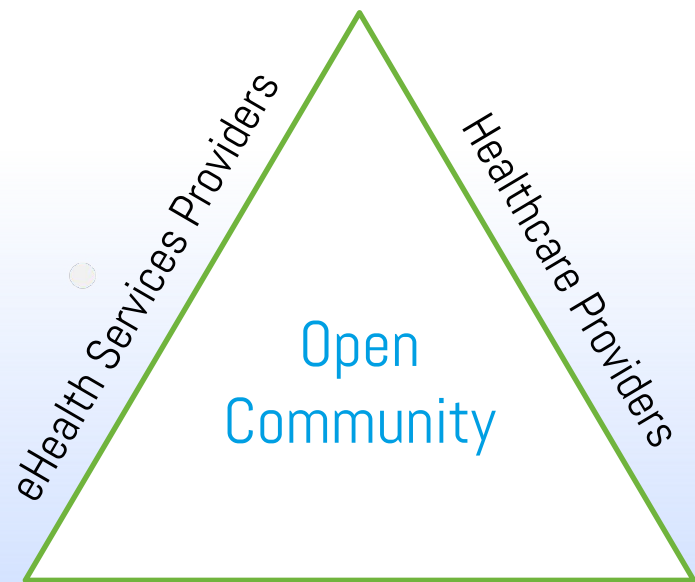
  

Nationality		
Value	Count	Percentage
724	5798	99.759 %
170	3	0.052 %
380	2	0.034 %
504	2	0.034 %
Others	7	0.12 %

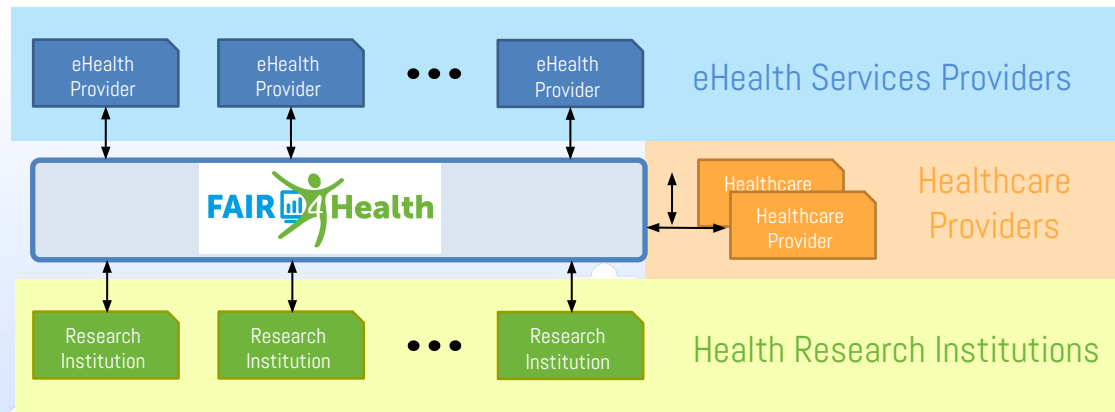
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666



# FAIR4Health Vision

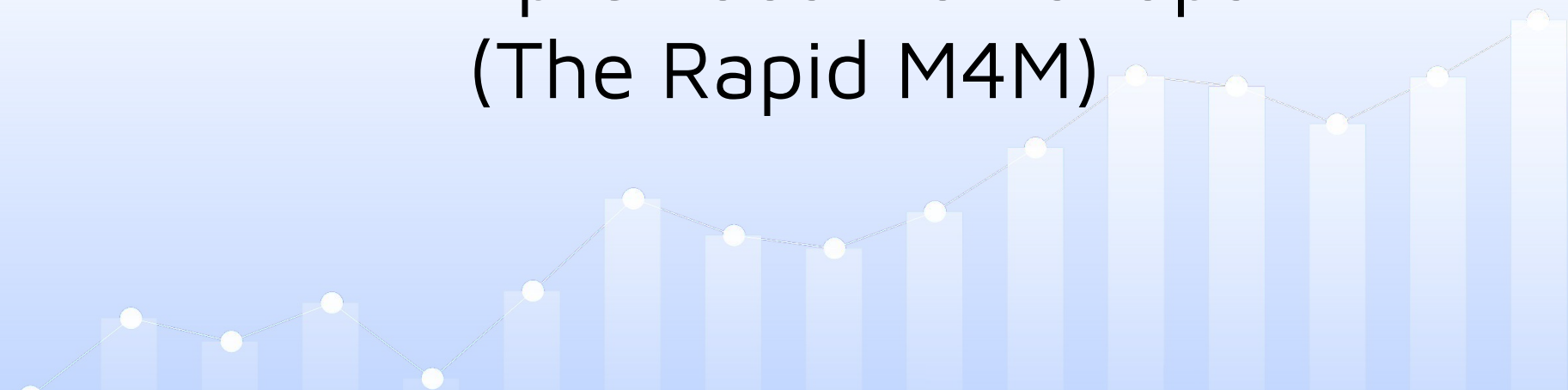


Health Research Institutions



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# M4M previous workshops (The Rapid M4M)

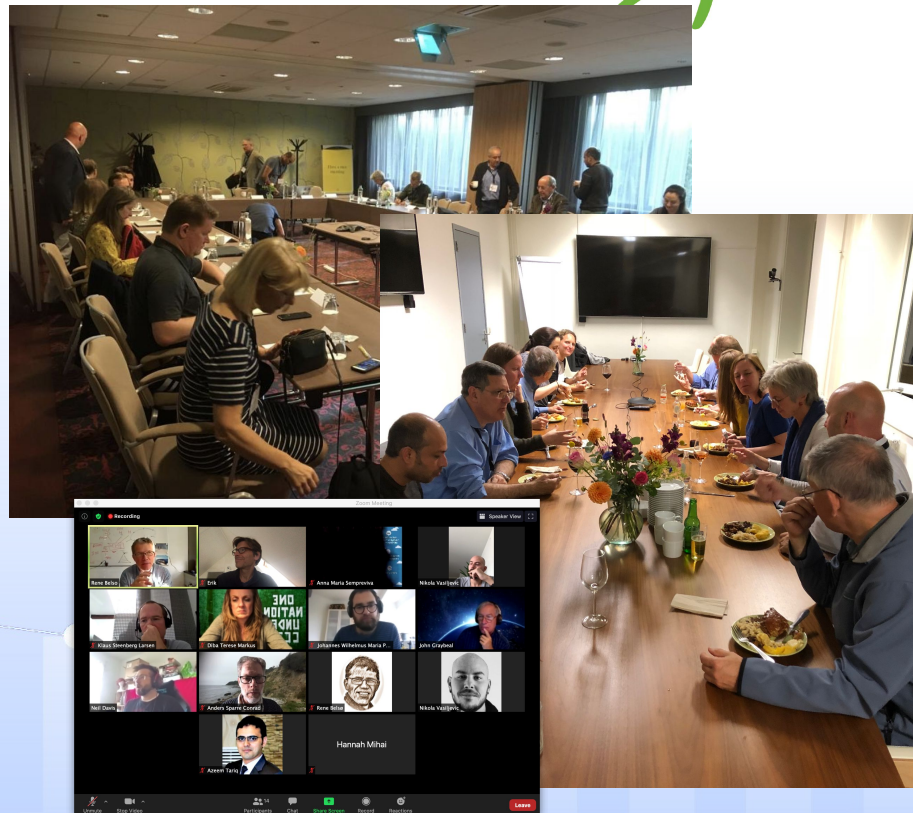
A decorative bar chart with 15 light blue bars of varying heights, overlaid with a grey line graph featuring white circular markers. The chart is positioned at the bottom of the slide, behind the funding text.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# M4M previous workshops

**M4M workshops overview** <https://www.gofairfoundation.org/m4m/>

Workshop Date		Community	Topic	Sponsor
M4M.1	October 2019	Inaugural	Setting up the concept	GO FAIR
M4M.2	January 2020	Funders	ZonMw + HRB	GO FAIR
M4M.3	January 2020	PreClinicalTrails	pre-registration form	GO FAIR
M4M.4	April-Sept 2020	VODAN Africa	Metadata for the FDP	Phillips Foundation
M4M.5	Summer 2020	AnnaEE	Climate data	DeiC
M4M.6	Summer 2020	DTU and others	Wind Energy	DeiC
M4M.7	November 2020	COVID-19 Program	Care (Treatment) / Prevention	ZonMw
M4M.8	November 2020	COVID-19 Program	Diagnostic / Testing – <a href="#">Recordings</a>	ZonMw
M4M.9	November 2020	COVID-19 Program	Prognosis / Risk assessments	ZonMw
M4M.10	November 2020	COVID-19 Program	Virus / Immunology / Molecular – <a href="#">Recordings</a>	ZonMw
M4M.11	November 2020	COVID-19 Program	Organisational / Process related – <a href="#">Recordings</a>	ZonMw
M4M.12	November 2020	COVID-19 Program	Socio-economic / Behavioral – <a href="#">Recordings</a>	ZonMw
M4M.13	February 2021	COVID-19 Program	Vocab	ZonMw
M4M.14	February 2021	COVID-19 Program	Vocab	ZonMw
M4M.15	June 2021	COVID-19 Program	Rapid M4M for datasets	ZonMw
M4M.16	June 2021	COVID-19 Program	I-ADOPT M4M for variables	ZonMw
M4M.17	June 2021	ID & AMR	R4R, COVID—>ID&AMR	ZonMw
M4M.18	Sept 2021	INCENTIVE	Influenza vaccine	EU/Horizon2020

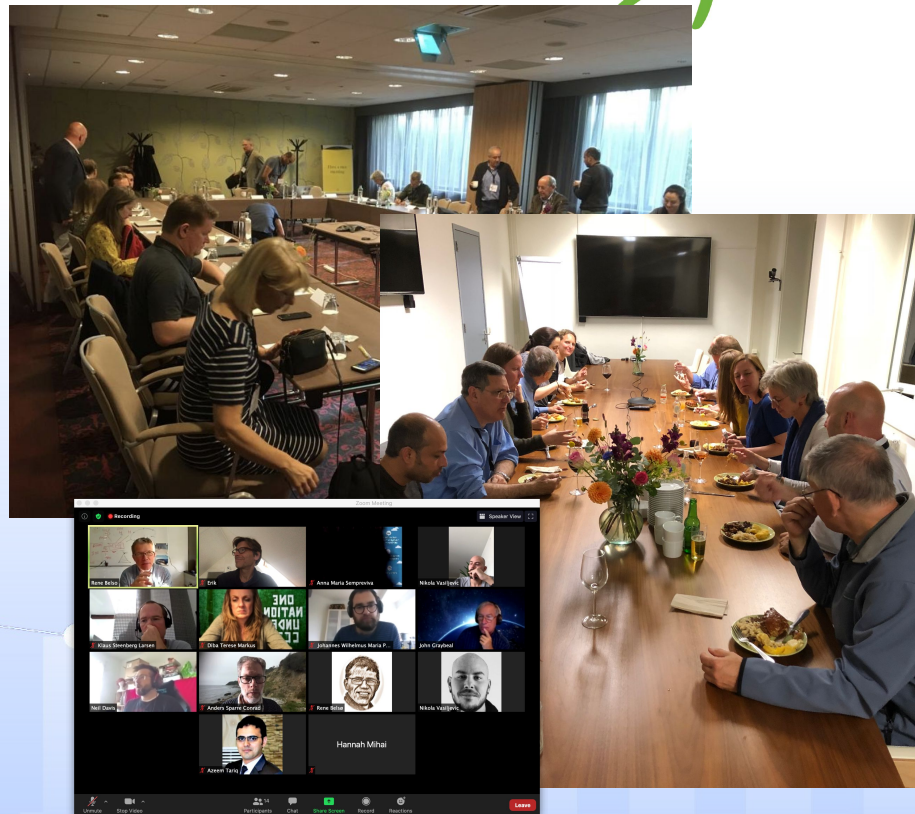


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# M4M previous workshops



**Making it easy for humans  
to make metadata for  
machines**



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666



# M4M previous workshops

## How to GO FAIR

<https://www.go-fair.org/how-to-go-fair/>

Home » How to GO FAIR

### How to GO FAIR

Since its beginning in early 2018, the GO FAIR community has been working towards implementations of the **FAIR Guiding Principles**. This collective effort has resulted in a three-point framework that formulates the essential steps towards the end goal, a global Internet of FAIR Data and Services where data are **F**indable, **A**ccessible, **I**nteroperable and **R**eusable (**FAIR**) for machines.

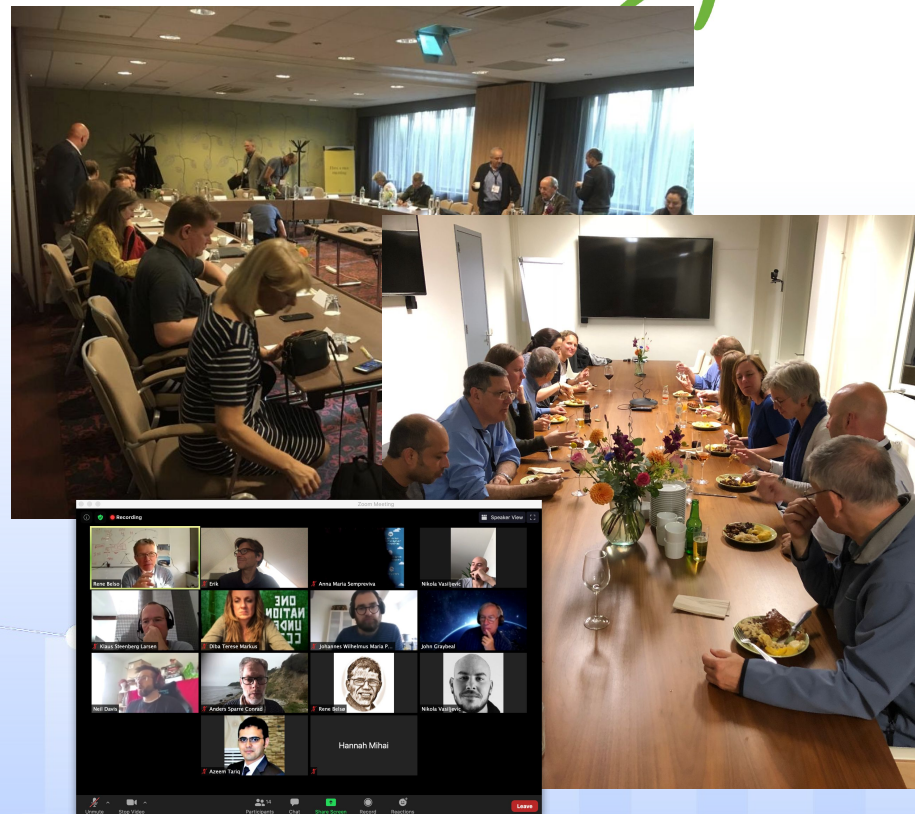


### A framework guiding FAIRification

The Three-point FAIRification Framework provides practical "how to" guidance to stakeholders seeking to go FAIR.

Moreover, by following this framework, stakeholders can rest assured that their efforts toward FAIRification will be optimally coordinated with the efforts of other stakeholders in the GO FAIR community. The three-point framework maximizes reuse of existing resources, maximizes interoperability, and accelerates convergence on standards and technologies supporting FAIR data and services.

- Typically, the FAIRification process begins when a community of practice considers its domain-relevant metadata requirements and other policy considerations, and formulates these considerations as machine-actionable metadata components. These considerations can be guided in **Metadata for Machines (M4M) Workshops**.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# M4M previous workshops



## Box 2 | The FAIR Guiding Principles

<https://www.nature.com/articles/sdata201618>

### 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

## M4M workshop format

Domain  
Experts

+

FAIR  
Metadata  
Experts

=

FAIR  
Metadata

- Domain-relevant
- Community standards

- Solutions
- Tools

- FAIR (schema/vocabs)
- Domain-specific
- Reusable

CEDAR



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# M4M previous workshops

**ZonMw**  
COVID-19 Program  
62 Projects

Brainstorm  
controlled  
lists

November

February

March

COVID  
Program  
Duration

**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

**Feb 9: Consolidation of lists and terms describing the project assets**  
**Feb 17: Building the COVID-19 Program Vocabulary**

**Project Admin  
Metadata**

**Project Content  
Metadata**

**Generic Dataset  
Metadata**

**COVID-19 Program  
metadata input forms**

**Data Stewards  
support researchers  
to routinely create  
metadata instances**

**FAIR  
Data  
Points**

**health RI**  
enabling data driven health

# M4M previous workshops



**BioPortal**

**ZonMW COVID-19**  
Last uploaded: October 20, 2021

Summary **Classes** Properties

**Details**

Acronym	ZONMW-CONTENT
Visibility	Public
Description	Collections of terms for the ZonMW COVID-19 program. This submission was created using SKOS.
Status	Production
Format	SKOS
Contact	Barbara Magagna, barbara.magagna@zonmw.nl

<https://www.gofairfoundations.org/m4m/>

Project Cedar - Metadata Editor

COVID-19 Project Content\_V5

Read & Understood

Read & Understood\*

Project Title (1..N)

Project Title - Multiple answers are allowed.\*

Language\*

Scope

Focus Area

To which focus area of the ZonMW COVID-19 program does your project belong?\*

Most projects have been classified by ZonMW in one of these focus areas. If you can not find your focus area, please contact us.

For all terms provided in the drop-down list it is useful to check the definitions in BioPortal <https://bioportal.bioontology.org/ontologies/ZONMW-CONTENT>.

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)

Area Level - Multiple answers are allowed.

Disease (1..N)

Which category of diseases, if any, is relevant for your project? Multiple answers are allowed.

If you ticked human disease, what specific disease is relevant for your project?

**health RI**  
enabling data driven health

<https://covid19initiatives-test.health-ri.nl/p/ProjectOverview>

Initiatives Dashboard **Projects**

**Project overview**

Request data Reset

Search

1 to 25 of 98

Lead institution	De maatschappelijke impact van COVID-19	Last changed 02/11/2021
Data availability	ZonMw - Netherlands Organisation for Health Research and Development Erasmus University Rotterdam	
ZonMw focus area	Investigating the immune response to COVID-19 Vaccination in Lung Transplantation patients (COVALENT study)	Last changed 01/11/2021
Type of provided assets	Covalent	
Provided data	ZonMw - Netherlands Organisation for Health Research and Development University Medical Center Groningen	
Type of provided biomaterial	Aanhoudende Klachten na COVID-19: perspectief vanuit de populatie, patiënt, en zorg.	Last changed 01/11/2021
Type of provided service	ZonMw - Netherlands Organisation for Health Research and Development University Medical Center Groningen	
	Changes in the use and organization of care in general practices and out-of-hours services: lessons learned from the COVID-19 pandemic	Last changed 28/10/2021
	COVID-GP	
	ZonMw - Netherlands Organisation for Health Research and Development NIVEL - Netherlands Institute for Health Services Research	
	Is COVID-19 a threat to banks and financial stability in Europe?	Last changed 26/10/2021
	Erasmus University Rotterdam	
	Infectiepreventie van COVID-19 in ziekenhuizen - omgevingsstudie	Last changed 26/10/2021
	COCON- environment	

2021 Health-RI DTL GFAIR FAIR NEDERLANDSE FEDERATIE VAN UNIVERSITAIR MEDISCHE CENTRA ZonMw

Terms and Conditions | Privacy Policy | Cookies Policy | Disclaimer

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666



# M4M previous workshops



## The *Rapid* M4M

- 6 hours (1 or 2 day)
- DCAT metadata templates are given:
  - Catalog
  - Dataset
  - Distribution
- Building vocab for (only) two fields:
  - Subject
  - Variable descriptions (advanced I-ADOPT)
- “Introduction to...” or “training with tools”

Next up: M4M.19  
EOSC Nordic NICEST2 (Climate)  
December 7-8

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

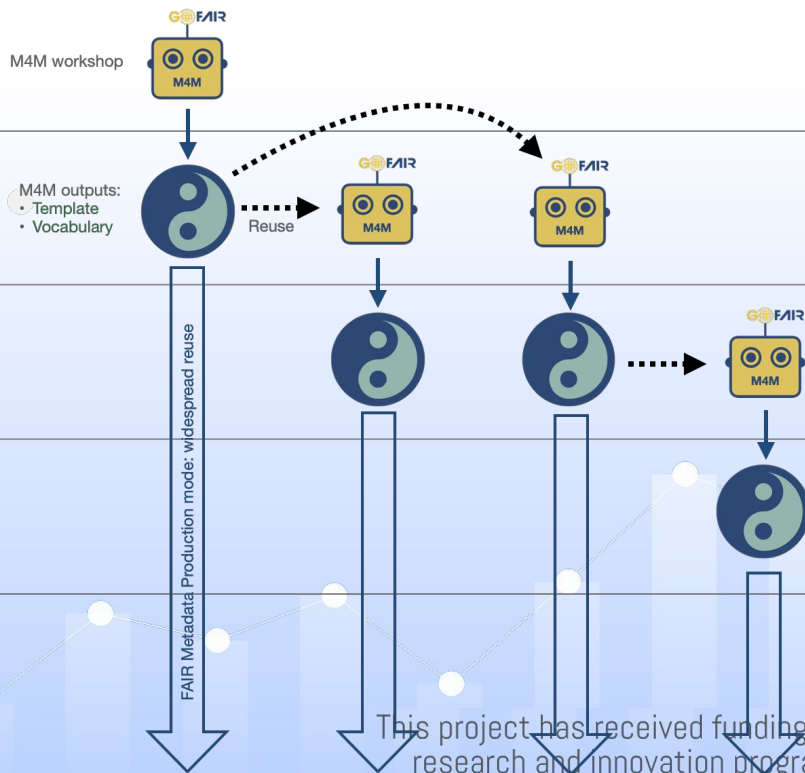
# M4M previous workshops

Community  
A

Community  
B

Community  
C

Community  
D



**M4M Goal:**  
**FAIR metadata at scale**

## INCLUSION CRITERIA

Patients ≥ 65 years old

FHIR QUERY

/Patient?birthdate=le1955

FHIR PATH

Patients with at least two chronic diseases/comorbidities of the following categories

FHIR QUERY

/Condition?code sw=G45,I60,I61,I62,I63,G20,G30,F01,F02,F03,F05,G60,G61,G62,G63,G64,G65,G35,F

FHIR PATH

agg:groupBy(Condition.subject.reference.count()) where(agg >= 2) bucket

## FHIR QUERIES AS TABLE

Table (1..N)

Variable

Age

FHIR QUERY

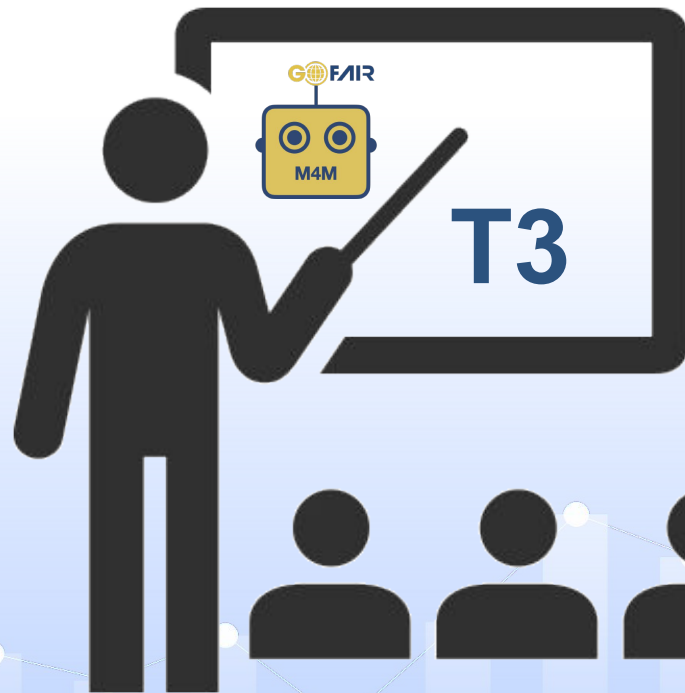
/Patient

FHIR PATH

value:Patient.birthDate getPeriod(\$this, now(), 'years')

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# M4M previous workshops



## M4M Train-the-Trainer Program (2022)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666



# CEDAR

CENTER FOR EXPANDED DATA  
ANNOTATION AND RETRIEVAL

## Building FAIR metadata with CEDAR

<https://cedar.metadatacenter.org>

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666



## Part 1: The Problem

In the biomedical community, we already had minimal information models and excel spreadsheet solutions.

- The former weren't computable—just instructions.
- The latter weren't semantically precise, nor easy to validate.

Both of these solutions were too hard to use.

	A	B	
1	basic_study_design	Schema Version 2.28	Content
2	Please do not delete or edit this column		
3	Column Name		
4			
5	study		
6	User_Def_Study_ID		
7	Brief_Title		
8	Official_Title		
9	Study_Type		
10	Brief_Description		
11	Detailed_Description		
12	Hypothesis		
13	Objectives		
14	Endpoints		
15	Sponsoring_Organization		
16	Target_Enrollment		
17	Condition_Studied		
18	Minimum_Age		
19	Maximum_Age		
20	Age_Limit		
21	Actual_Start_Date		
22	Intervention_Agent		
23			
24	arm_or_cohort		
25	User_Def_Arm_Or_Cohort_ID	Name	Description
26			Population_Selection_Rule
27			Sort_Order
28	study_personnel		
29	User_Def_Study_Personnel_ID		
30	Honoric		
31	Last_Name		
32	First_Name		
33	Suffixes		
34	Organization		
35	Email		
36	Title_In_Study		
37	Role_In_Study		
38	Site_Name		

This project has received funding from the European Union's Horizon research and innovation programme under grant agreement No 101019719.



# Building FAIR metadata with CEDAR

## Part 2: The Vision

We understood how most metadata gets created.

### Project-Based Metadata Ecosystems

- **Investigators** perform experiments in a domain.
- A **Project Authority** makes metadata templates to encourage good experiment data annotation.
- **Investigators** fill out metadata per the templates.
- **Project** stores data (and metadata) in its public repository for data discovery and use.
- **Researchers** use resulting metadata resources to find and reuse the project's resources.





# Building FAIR metadata with CEDAR

## Part 2: The Vision

We wanted metadata to be simple for metadata providers—yet richly meaningful and computable for all its users.

In other words, FAIR.

With funding from NIH's Big Data to Knowledge Project in 2014, we started working on CEDAR.

### The CEDAR Vision

*A simple 'life of metadata' for users*

- Compatibility with best known practices and standards
- Use of existing semantic and data structure resources
- End-to-end improvement in biomedical metadata



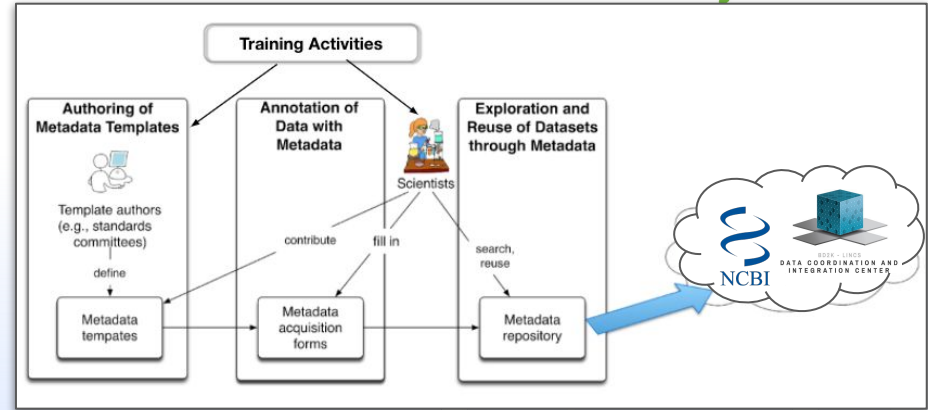
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666



# Building FAIR metadata with CEDAR

## Part 3: The Implementation

We designed and built an open source system that makes it easy to specify metadata forms—we call them 'metadata templates'—and even easier to fill them out.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666





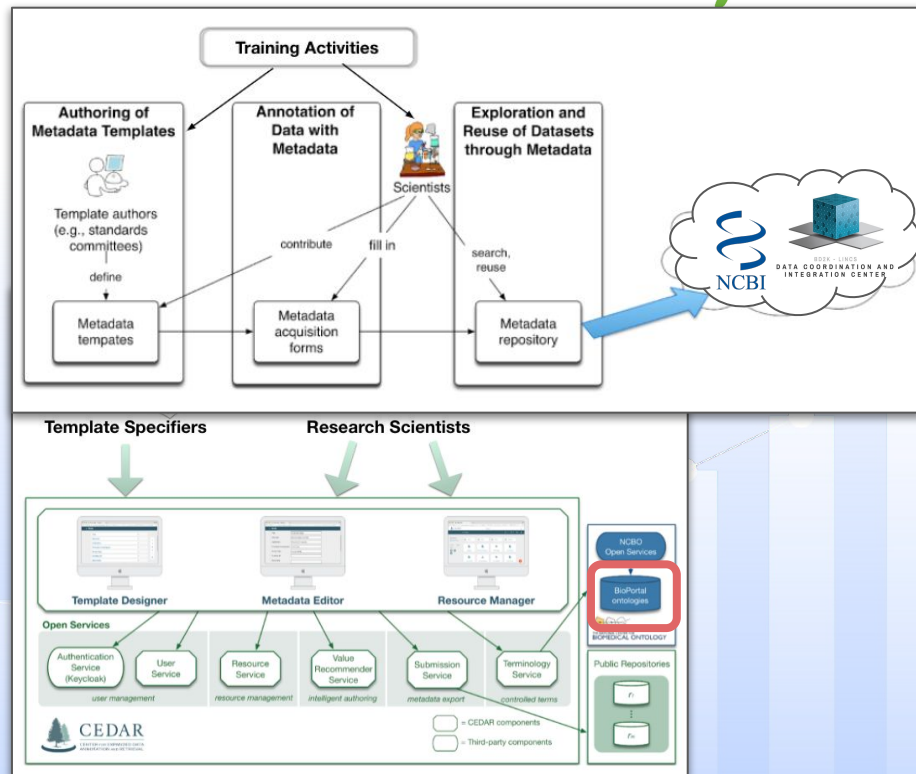
# Building FAIR metadata with CEDAR

## Part 3: The Realization

We designed and built an open source system that makes it easy to specify metadata forms—we call them 'metadata templates'—and even easier to fill them out.

We used the BioPortal ontology repository as a community resource to provide precise terms to fill out the fields.

And we made it easy for authorized users to access the system and see—or even update—their content via the public REST API...



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666



# Building FAIR metadata with CEDAR

## Part 3: The Realization

We designed and built an open source system that makes it easy to specify metadata forms—we call them 'metadata templates'—and even easier to fill them out.

We used the BioPortal ontology repository as a community resource to provide precise terms to fill out the fields.

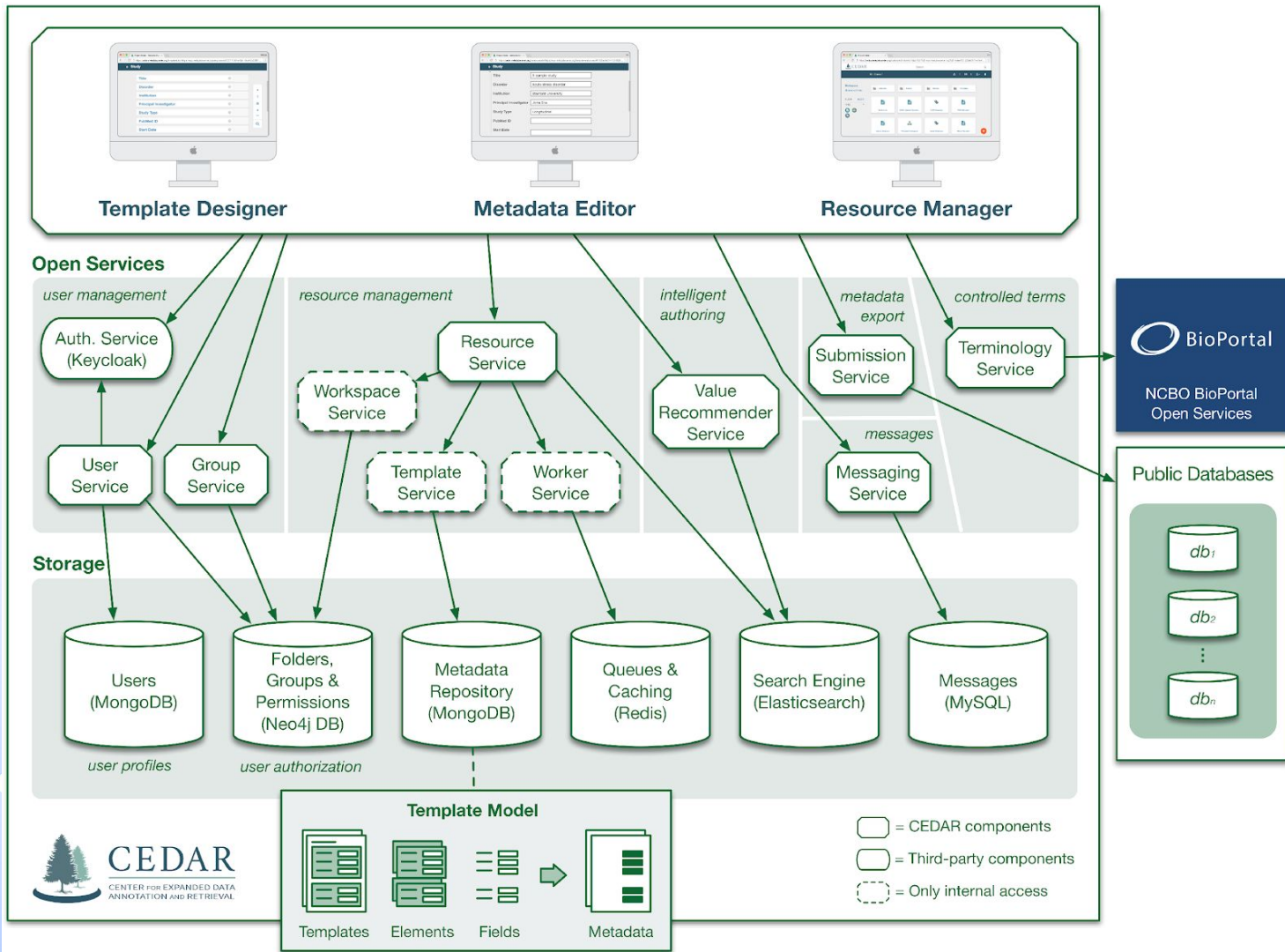
And we made it easy for authorized users to access the system and see—or even update—their content via the public REST API...

And even publish content to repositories or to the open web.

The screenshot shows the CEDAR OpenView interface for the Generic Dataset Metadata Template (GDMT). The page has a header with the CEDAR OpenView logo and the title 'Generic Dataset Metadata Template (GDMT)'. Below the header, the word 'View' is displayed in a large blue font. The main content area is titled 'Generic Dataset Metadata Template (GDMT)' and contains a list of expandable fields, each with a right-pointing chevron and a help icon (a circle with a question mark). The fields are: Resource Type, Dataset Identifier, Title (1 of N), Description (1 of N), Language, Version, Subjects and Keywords, Creator (1 of N), Contributor (1 of N), Rights, and Date (1 of N).

This project has received funding from the European Union's Horizon research and innovation programme under grant agreement No 101019715

[bit.ly/cedar-generic-dataset-openview](https://bit.ly/cedar-generic-dataset-openview)





# Building FAIR metadata with CEDAR

## Part 4: The Demo

The screenshot shows the CEDAR metadatacenter.org interface. On the left, there is a sidebar with a 'New +' button, 'Workspace' section (Shared with Me, Shared with Everybody, Community Folders), and a 'FILTER' section with expandable menus for TYPE, VERSION, and CATEGORIES. The main area displays a table of metadata items for the 'Demo User'.

Title	Created	Modified	
Demonstration Folders	1/4/20	1/4/20	⋮
Demo Generic Metadata Template for Datasets (GDMT)	4/1/21	4/1/21	🔍 ⋮
Resource Type no Details	4/1/21	4/1/21	⋮
Revised case report form for Confirmed Novel Coronavirus C...	6/16/20	6/16/20	📄 ⋮
MIARR V1.1.0 metadata	4/27/20	4/27/20	📄 ⋮
Demonstration Instructions	8/8/19	1/28/20	📄 🔍 ⋮

Displaying 1 - 6 of 6

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666



# Building FAIR metadata with CEDAR

## Part 4: The Demo

← Demo Generic Metadata Template for Datas...

Name	Identifier	Version	Description
Demo Generic Metadata Templa		0.0.1	

+

Dataset Identifier

Dataset Identifier

Dataset Identifier Type

+

Version

Version

CEDAR  
metadatascenter.org

Search

All / Users / Demo User

Title	Created	Modified
Demonstration Folders	1/4/20	1/4/20
Demo Generic Metadata Template for Datasets (GDMT)	4/1/21	4/1/21
Resource Type no Details	4/1/21	4/1/21
Revised case report form for Confirmed Novel Coronavirus C...	6/16/20	6/16/20
MIARR V1.1.0 metadata	4/27/20	4/27/20
Demonstration Instructions	8/8/19	1/28/20

Displaying 1 - 6 of 6

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666



# Building FAIR metadata with CEDAR

## Part 4: The Demo

The screenshot shows the CEDAR metadata center interface. At the top, there is a search bar and a user profile icon. Below the search bar, there is a table of resources. The table has columns for Title, Created, and Modified. The resources listed are:

Title	Created	Modified
Demonstration Folders	1/4/20	1/4/20
Demo Generic Metadata Template for Datasets (GDMT)	4/1/21	4/1/21
Resource Type no Details	4/1/21	4/1/21
Revised case report form for Confirmed Novel Coronavirus C...	6/16/20	6/16/20
MIARR V1.1.0 metadata	4/27/20	4/27/20
Demonstration Instructions	8/8/19	1/28/20

At the bottom of the table, it says "Displaying 1 - 6 of 6".

The screenshot shows the "Demo Generic Metadata Template for Datasets (GDMT)" form. The form has a header with a back arrow, a document icon, and the title. Below the header, there is a table with columns for Name, Identifier, Version, and Description. The table contains one row with the following data:

Name	Identifier	Version	Description
Demo Generic Metadata Templa		0.0.1	

The screenshot shows the "NCBI BioSample - Human Pkg 1.0 - Demo me..." form. The form has a header with a back arrow, a document icon, and the title. Below the header, there is a form with several fields:

- Disease: hepatocellular carcinoma
- Tissue: Hepatic Tissue
- Cell line: A dropdown menu with the following options: HepG2 (80%), SNU398 (10%), Hep3B (10%), C3A (8%), GM08388, GM08390, and GM08398.
- Ethnicity: (empty field)
- Isolate: (empty field)





# Building FAIR metadata with CEDAR

## Part 5: The Resources

### Resources

Web site: <http://metadatacenter.org>

Workbench: <https://cedar.metadatacenter.org>

Technical site: <https://metadatacenter.github.io>

API: <https://resource.metadatacenter.org/api/>

References: <https://metadatacenter.org/references>



# CEDAR

CENTER FOR EXPANDED DATA  
ANNOTATION AND RETRIEVAL

[metadatacenter.org](http://metadatacenter.org)

● Test the use / requirements of CEDAR tools for creation of metadata templates in specific case studies



# Testing CEDAR tools in specific case studies

## Use case 1: Metadata generated in FAIR4Health project

- FAIR4Health Use case 1
  - multimorbidity patterns and polypharmacy on the 6-months mortality rate and cognitive impairment among elderly
- FAIR4Health Use case 2
  - predicting 30-day readmission risk in patients suffering from Chronic Obstructive Pulmonary Disease (COPD)
- Testing for CEDAR was done
  - with FAIR4Health Use Case 1
  - for a handful of variables from one Data Set
  - manually, not automated

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# Testing CEDAR tools in specific case studies

## FHIR Path / Queries shared through Excel

- [Common Data Model](#) (CDM) is implemented as FHIR profiles
  - FHIR Queries search for FHIR Resources
  - FHIR Path filters / selects data
- FHIR Path / Queries create a Feature Set
- Each Local agent creates one Data Set based on
  - the same Feature Set
  - the same Eligibility Criteria
- Feature Sets are created through the [FAIR4Health Portal](#)

	A	B	C
1	<b>Inclusion Criteria</b>	<b>fhir_query</b>	<b>fhir_path</b>
2	Patients ≥65 years old	/Patient?birthdate=le1955	
3	Patients with at least two chronic diseases/comorbidities of the foll	/Condition?code=sw=G45,I60,I61,I62	aggr:groupBy(Condition.subject
4			
5	<b>Variables</b>		
6	Age	See Use case 2	
7	Gender	See Use case 2	
8	Smoking status	See Use case 2	
9	Institutionalized	/Observation?code=160734000&val	value:exists
10	Nationality	/Patient	value:Patient.address.country
11	Heart failure	See Use case 2	
12	Coronary heart disease	See Use case 2	
13	Hypertension	See Use case 2	
14	Atrial fibrillation	See Use case 2	
15	Pulmonary hypertension	See Use case 2	
16	Chronic kidney disease	See Use case 2	
17	Prostatic hypertrophy	See Use case 2	
18	Renal calculi	See Use case 2	
19	Asthma	See Use case 2	
20	Obstructive sleep apnea syndrome	See Use case 2	
21	Pulmonary embolism	See Use case 2	
22	Inflammatory bowel disease	See Use case 2	
23	Chronic liver disease	See Use case 2	
24	Chronic pancreatitis	See Use case 2	
25	Gastroesophageal reflux	See Use case 2	
26	Diverticular disease of colon	See Use case 2	
27	Diabetes mellitus	See Use case 2	
28	Hyperlipidemia	See Use case 2	
29	Obesity	See Use case 2	
30	Hypothyroidism	See Use case 2	
31	Hyperthyroidism	See Use case 2	
32	Stroke TIA	/Condition?code=sw=G45,I60,I61,I62	value:exists
33	Parkinsons disease	/Condition?code=sw=G20	value:exists
34	Alzheimers disease	/Condition?code=sw=G30	value:exists

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# Testing CEDAR tools in specific case studies

## FHIR Path / Queries represented in CEDAR

Template:

<https://bit.ly/3DASuDs>

Exemplary Metadata:

<https://bit.ly/3HADIEz>

**INCLUSION CRITERIA**

Patients ≥ 65 years old

FHIR QUERY

/Patient?birthdate=le1955

FHIR PATH

Patients with at least two chronic diseases/comorbidities of the following categories

FHIR QUERY

/Condition?code:sw=G45,I60,I61,I62,I63,G20,G30,F01,F02,F03,F05,G60,G61,G62,G63,G64,G65,G35,F

FHIR PATH

aggr:groupBy(Condition.subject.reference,count()) where (agg >= 2) bucket

**FHIR QUERIES AS TABLE**

Table (1..N)

Variable

Age

FHIR QUERY

/Patient

FHIR PATH

value:Patient.birthDate.getPeriod(\$this, now(), 'years')

**INCLUSION CRITERIA**

Patients ≥ 65 years old

FHIR QUERY

/Patient?birthdate=le1955

FHIR PATH

Patients with at least two chronic diseases/comorbidities of the following categories

FHIR QUERY

/Condition?code:sw=G45,I60,I61,I62,I63,G20,G30,F01,F02,F03,F05,G60,G61,G62,G63,G64,G65,G35,F

FHIR PATH

aggr:groupBy(Condition.subject.reference,count()) where (agg >= 2) bucket

**FHIR QUERIES AS TABLE**

Table (1..N)

	Variable	FHIR QUERY	FHIR PATH
1	Age	/Patient	value:Patient.birthDate.getPeriod(\$this, now(), 'years')
2	Gender	/Patient	value:Patient.gender
3	Smoking status	/Observation?code=Z7166-2&value-concept=373066001	value:exists
4	Institutionalized	/Observation?code=1607340008&value-concept=373066001	value:exists
5	Nationality	/Patient	value:Patient.address.country
6	Heart Failure	/Condition?code:sw=I46,I50	value:exists
7	Coronary Heart Disease	/Condition?code:sw=I20,I21,I24,I25	value:exists
8	Hypertension	/Condition?code:sw=I10,I11,I12,I13	value:exists
9	Polypharmacy	/MedicationStatement	aggr:groupBy(MedicationStatement.subject.reference,count()) where (agg >= 2) bucket
10	Number of prescribed drugs	/MedicationStatement	aggr:groupBy(MedicationStatement.subject.reference,count()) where (agg >= 2) bucket

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

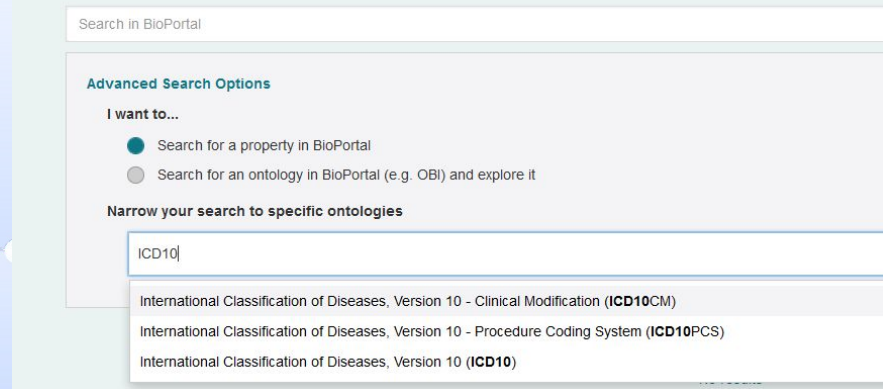
# Testing CEDAR tools in specific case studies

## Adding ICD-10 Codes to CEDAR Elements with BioPortal (1)

- FHIR Queries reference ICD-10 codes
- That information is not represented in previous Table
- CEDAR adds controlled Terms to each Element
- Common Terms can be added from [BioPortal](#)



BioPortal or [Enter Property URI](#)



# Testing CEDAR tools in specific case studies

## Adding ICD-10 Codes to CEDAR Elements with BioPortal (2)

Template:

<https://bit.ly/3cwXJrY>

Exemplary Metadata:

<https://bit.ly/3cw8rik>

- Codes are added on Template generation
- Metadata for a Feature Set could now:
  - contain references to the corresponding ICD-10 Codes
  - be available as JSON-LD and RDF
  - still be easily shared through CEDAR OpenView

### COMORBIDITIES

Heart Failure

FHIR QUERY

/Condition?code:sw=I46,I50

FHIR PATH

value:exists

Coronary Heart Disease

FHIR QUERY

/Condition?code:sw=I20,I21,I24,I25

FHIR PATH

value:exists

Hypertension

FHIR QUERY

/Condition?code:sw=I10,I11,I12,I13,I15,I16

FHIR PATH

value:exists

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# FAIR4Health project: Metadata



Use case: Use Case 1

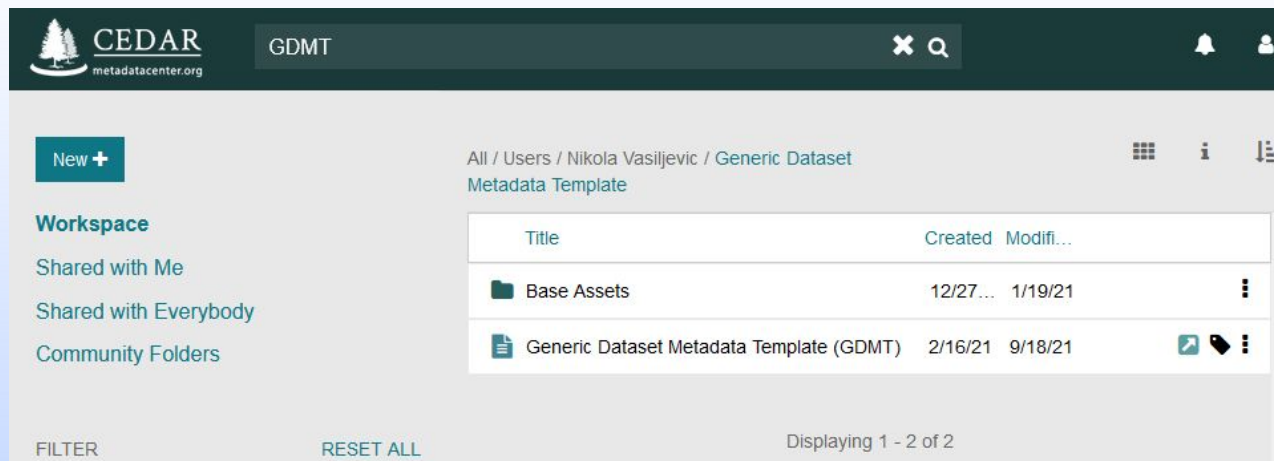
1 Name & description		2 Feature set selection				3 Eligibility criteria				4 Results & statistics				5 Completed
agent	# of records	Age	Gender	Smoking status	Institutionalized	Nationality	Heart failure	Coronary heart disease	Hypertension	Atrial fibrillation	Pulmonary hypertension	Chronic kidney disease	Prostatic hypertrop	
<div><div><div><div><div></div></div></div><div><div><div></div></div></div></div><div>FAIR4Health PPDDM-Agent for <div></div></div></div>	5812 <div><a href="#">See all distributions</a></div>	Min: 65 Max: 104 Null: 0 <div><a href="#">See distributions</a></div>	female: 2868 male: 2944 <div><a href="#">See distributions</a></div>	FALSE: 5788 TRUE: 24 <div><a href="#">See distributions</a></div>	FALSE: 5530 TRUE: 282 <div><a href="#">See distributions</a></div>	724: 5798 170: 3 380: 2 504: 2 Others: 7 <div><a href="#">See distributions</a></div>	FALSE: 3653 TRUE: 2159 <div><a href="#">See distributions</a></div>	FALSE: 4456 TRUE: 1356 <div><a href="#">See distributions</a></div>	FALSE: 1165 TRUE: 4647 <div><a href="#">See distributions</a></div>	FALSE: 3918 TRUE: 1894 <div><a href="#">See distributions</a></div>	FALSE: 5341 TRUE: 471 <div><a href="#">See distributions</a></div>	FALSE: 4483 TRUE: 1329 <div><a href="#">See distributions</a></div>	FALSE: 5200 TRUE: 612 <div><a href="#">See distributions</a></div>	
<div><div><div><div><div></div></div></div><div><div><div></div></div></div></div><div>FAIR4Health PPDDM-Agent for <div></div></div></div>	331 <div><a href="#">See all distributions</a></div>	Min: 68 Max: 116 Null: 0 <div><a href="#">See distributions</a></div>	female: 237 male: 94 <div><a href="#">See distributions</a></div>	FALSE: 329 TRUE: 2 <div><a href="#">See distributions</a></div>	FALSE: 331 <div><a href="#">See distributions</a></div>	CS: 183 DE: 63 IT: 62 IL: 20 Others: 3 <div><a href="#">See distributions</a></div>	FALSE: 256 TRUE: 75 <div><a href="#">See distributions</a></div>	FALSE: 129 TRUE: 202 <div><a href="#">See distributions</a></div>	FALSE: 331 <div><a href="#">See distributions</a></div>	FALSE: 331 <div><a href="#">See distributions</a></div>	FALSE: 331 <div><a href="#">See distributions</a></div>	FALSE: 331 <div><a href="#">See distributions</a></div>	FALSE: 331 <div><a href="#">See distributions</a></div>	
<div><div><div><div><div></div></div></div><div><div><div></div></div></div></div><div>FAIR4Health PPDDM-Agent for <div></div></div></div>	244 <div><a href="#">See all distributions</a></div>	Min: 66 Max: 104 Null: 0 <div><a href="#">See distributions</a></div>	female: 115 male: 129 <div><a href="#">See distributions</a></div>	FALSE: 244 <div><a href="#">See distributions</a></div>	FALSE: 227 TRUE: 17 <div><a href="#">See distributions</a></div>	CH: 183 IT: 13 ES: 8 FR: 7 Others: 33 <div><a href="#">See distributions</a></div>	FALSE: 243 TRUE: 1 <div><a href="#">See distributions</a></div>	FALSE: 189 TRUE: 55 <div><a href="#">See distributions</a></div>	FALSE: 114 TRUE: 130 <div><a href="#">See distributions</a></div>	FALSE: 173 TRUE: 71 <div><a href="#">See distributions</a></div>	FALSE: 230 TRUE: 14 <div><a href="#">See distributions</a></div>	FALSE: 207 TRUE: 37 <div><a href="#">See distributions</a></div>	FALSE: 221 TRUE: 23 <div><a href="#">See distributions</a></div>	
<div><div><div><div><div></div></div></div><div><div><div></div></div></div></div><div>FAIR4Health PPDDM-Agent for <div></div></div></div>	3786 <div><a href="#">See all distributions</a></div>	Min: 65 Max: 106 Null: 0 <div><a href="#">See distributions</a></div>	female: 1889 male: 1897 <div><a href="#">See distributions</a></div>	FALSE: 3786 <div><a href="#">See distributions</a></div>	FALSE: 3786 <div><a href="#">See distributions</a></div>	724: 3724 250: 11 642: 7 504: 5 Others: 39 <div><a href="#">See distributions</a></div>	FALSE: 2877 TRUE: 909 <div><a href="#">See distributions</a></div>	FALSE: 3001 TRUE: 785 <div><a href="#">See distributions</a></div>	FALSE: 909 TRUE: 2877 <div><a href="#">See distributions</a></div>	FALSE: 2751 TRUE: 1035 <div><a href="#">See distributions</a></div>	FALSE: 3502 TRUE: 284 <div><a href="#">See distributions</a></div>	FALSE: 2968 TRUE: 818 <div><a href="#">See distributions</a></div>	FALSE: 3479 TRUE: 307 <div><a href="#">See distributions</a></div>	



# Testing CEDAR tools in specific case studies

## Reusing existing Templates in CEDAR

Generic Dataset Metadata Template (GDMT): <https://bit.ly/3qTYwv9>



The screenshot shows the CEDAR metadata center interface. The header includes the CEDAR logo and the text "metadatascenter.org". A search bar contains the text "GDMT". The main content area displays a list of templates under the heading "Generic Dataset Metadata Template (GDMT)". The list includes two entries: "Base Assets" and "Generic Dataset Metadata Template (GDMT)". The "Generic Dataset Metadata Template (GDMT)" entry is highlighted, showing its creation and modification dates.

Title	Created	Modified
Base Assets	12/27...	1/19/21
Generic Dataset Metadata Template (GDMT)	2/16/21	9/18/21

Displaying 1 - 2 of 2

View

Generic Dataset Metadata Template (GDMT) ⌵

- > Resource Type ⌵
- > Dataset Identifier ⌵
- > Title (1 of N) ⌵
- > Description (1 of N) ⌵
- > Language ⌵
- > Version ⌵
- > Subjects and Keywords ⌵
- > Creator (1 of N) ⌵
- > Contributor (1 of N) ⌵
- > Rights ⌵
- > Date (1 of N) ⌵
- > Funding (1 of N) ⌵
- > Related Resources (1 of N) ⌵
- > Variable (1 of N) ⌵
- > Data Source (1 of N) ⌵
- > Data Stream (1 of N) ⌵
- > Spatial Coverage ⌵
- > Vertical Coverage ⌵
- > Temporal Coverage ⌵
- > Distribution (1 of N) ⌵




# Testing CEDAR tools in specific case studies

## Metadata Schema for Use Case 1 in CEDAR (1)

- Reuse elements from GDMT
- Version / Date could capture how distributions change over time
- Eligibility criteria contains FHIR Query / Path expressions
- Similarly Metadata from Feature Sets or a reference to a CEDAR Metadata Instance could be included
- Distributions replaces GDMT Elements based on specific needs of FAIR4Health Use Case 1

[View](#)

FAIR4Health Use Case 1 Dataset Template (based on GDMT) 

- > Dataset Identifier 
- > Title (1 of N) 
- > Description (1 of N) 
- > Version 
- > Date (1 of N) 
- > Creator (1 of N) 
- > Data Source (1 of N) 

### Eligibility criteria

- > Eligibility criteria (1 of N)

### Distributions

- > Number of Records
- > Age
- > Gender
- > Smoking status

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# Testing CEDAR tools in specific case studies

## Metadata Schema for Use Case 1 in CEDAR (2)

Template:

<https://bit.ly/3FxXBow>

Exemplary Metadata:

<https://bit.ly/3HBVMZw>

agent	# of records	Age	Gender	Smoking status	Institutionalized
<input checked="" type="checkbox"/> FAIR4Health PPDDM-Agent for <input type="text"/>	5812	Min: 65 Max: 104 Null: 0 <a href="#">See all distributions</a> <a href="#">See distributions</a>	female: 2868 male: 2944 <a href="#">See distributions</a>	FALSE: 5788 TRUE: 24 <a href="#">See distributions</a>	FALSE: 5530 TRUE: 282 <a href="#">See distributions</a>
<input checked="" type="checkbox"/> FAIR4Health PPDDM-Agent for <input type="text"/>	331	Min: 68 Max: 116 Null: 0 <a href="#">See all distributions</a> <a href="#">See distributions</a>	female: 237 male: 94 <a href="#">See distributions</a>	FALSE: 329 TRUE: 2 <a href="#">See distributions</a>	FALSE: 331 <a href="#">See distributions</a>

### DISTRIBUTIONS

Number of Records

Number of Records

5812

Age

Min

65

Max

104

Null

0

Gender

female

2868

male

2944

Smoking status

False

5788

True

24

Institutionalized

False

5530

True

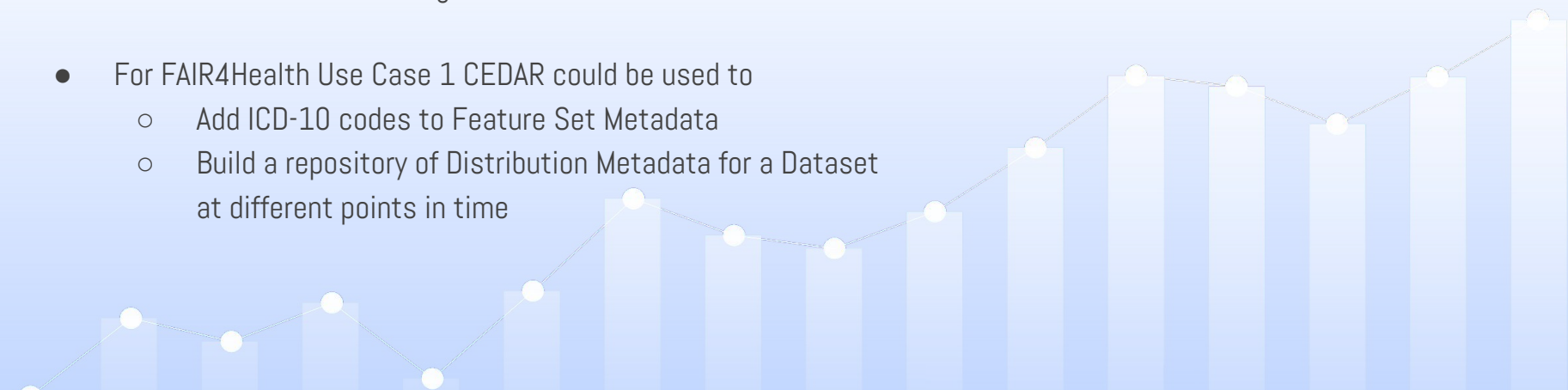
282

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# Testing CEDAR tools in specific case studies

## Conclusion

- CEDAR allows easy manual creation of Templates
- Templates can be published or shared with CEDAR OpenView
- Templates and Metadata Instances are available as JSON-LD and RDF
- Possible automation through CEDAR API
  
- For FAIR4Health Use Case 1 CEDAR could be used to
  - Add ICD-10 codes to Feature Set Metadata
  - Build a repository of Distribution Metadata for a Dataset at different points in time



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# Testing CEDAR tools in specific case studies

## Use case 2: NFDI4Health Covid-19 metadata schema

- [NFDI4Health](#) is the National Research Data Infrastructure for Personal Health Data in Germany
  - Part of a network of NFDI\* consortia covering different scientific disciplines
  - 5-year grant (2020-2025) from the German Research Foundation (DFG)
  - Main goal is to build a networked infrastructure to connect existing datasets
  - Full coverage of epidemiological cohorts, clinical trials, public health surveys and administrative databases
  - Strongly committed to FAIR
- [NFDI4Health Task Force COVID-19](#)
  - 3-year grant (2020-2023) from the German Research Foundation
  - Mission is to support clinical research by providing links (ongoing studies, publications, datasets)



# Testing CEDAR tools in specific case studies

## Use case 2: NFDI4Health Covid-19 metadata schema

### NFDI4Health Task Force COVID-19:

- Creating an inventory of German COVID-19 studies
  - Covering structured health data from clinical trials incl. vaccination studies, epidemiological studies, and public health surveillance
- Result: [Study Hub NFDI4Health COVID-19](#), consisting of three main components
  - Central search portal
  - Document portal
  - Instrument portal
- Development is based on a consented Metadata Schema influenced by standards like Dublin Core, DataCite, CDISC BRIDG and HL7 FHIR as well as de-facto standard dataset conventions from international trial registers

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# Testing CEDAR tools in specific case studies

## German Central Health Study Hub COVID-19: Central search portal

- Metadata Schema is used for faceted search, browsing and filtering



Studies Questionnaires

Refine your search. Currently 893 studies are matched

<https://covid19.studyhub.nfdi4health.de/study/>

Quick filters

- Available Documents 893
  - no
- Available Instruments 893
  - no
- Primary Design 536
  - Non-interventional
  - Interventional
  - N/A
- Study Status 425
  - Recruiting

29 columns hidden Sort fields

View Details	Acronyms	Title	Webpage...	Source Identifiers
Open		Rationale and investigational study for the treatment of COVID-19 with severe viral pneumonia with isolated, placental, mesenchymal stem cell exosomes	Visit	ISRCTN33578935
Open	SAPRIS	The SAPRIS Survey - Health, practices, relations, and social inequalities in France's general population during the COVID-19 crisis		
Open	Essen Corona	Survey for a study to fight corona pandemic in Essen.	Visit	
Open	Covid Mental Health	The German COVID-19 Survey on Mental Health		
Open	FoodCOVID-19	Food systems in times of COVID-19	Visit	
Open	CoPE with Corona	Study of general population coping with COVID-19/SARS-CoV-2	Visit	
Open	CRISIS - Pilot	The CoRonaviruS Health Impact Survey - Pilot Survey	Visit	
Open	CoVerlauf	Studie zum Erkrankungsverlauf bei Personen mit einer COVID-19-Erkrankung bzw. einem positiven Test auf SARS-CoV-2	Visit	
Open	CoCCS	Covid-19 Case-Cluster-Study (Heinsberg Study)	Visit	
Open	CoMolo	CORONA-MONITORING lokal	Visit	

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# Testing CEDAR tools in specific case studies

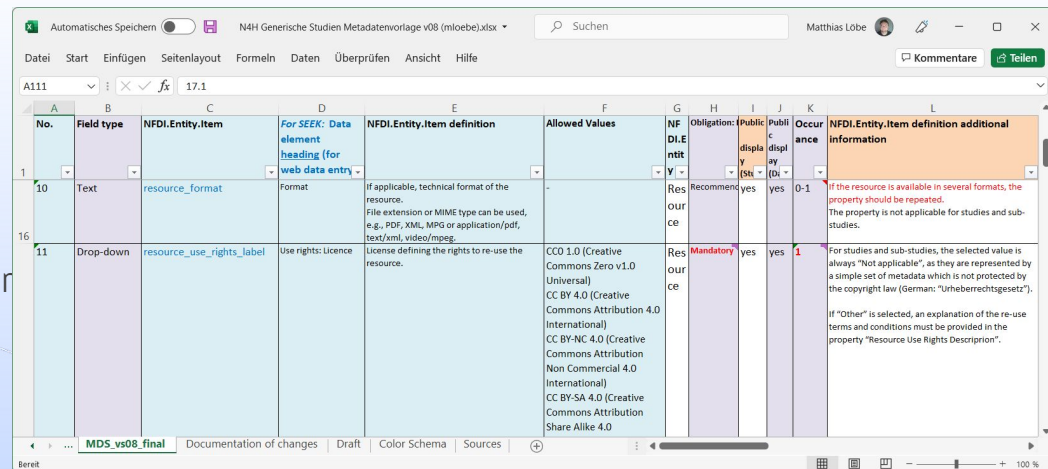
## Metadata Schema in Microsoft Excel

- As with the development of any good metadata schema, Excel was used
  - Easy collaborative creation
  - Easy to comment and change

- Now it would be version 1.0 in a more strictly formalized way

- Changes must be tracked
- Data annotations
- API access to read definitions

Element



No.	Field type	NFDI.Entity.Item	For SEEK: Data element heading (for web data entry)	NFDI.Entity.Item definition	Allowed Values	NFDI.Entity.Item	Obligation	Public display	Occurrence	NFDI.Entity.Item definition additional information
10	Text	resource_format	Format	If applicable, technical format of the resource. File extension or MIME type can be used, e.g., PDF, XML, MPG or application/pdf, text/xml, video/mpeg.	-	Recommendation	yes	yes	0-1	If the resource is available in several formats, the property should be repeated. The property is not applicable for studies and sub-studies.
11	Drop-down	resource_use_rights_label	Use rights: Licence	License defining the rights to re-use the resource.	CC0 1.0 (Creative Commons Zero v1.0 Universal) CC BY 4.0 (Creative Commons Attribution 4.0 International) CC BY-NC 4.0 (Creative Commons Attribution Non Commercial 4.0 International) CC BY-SA 4.0 (Creative Commons Attribution Share Alike 4.0)	Mandatory	yes	yes	1	For studies and sub-studies, the selected value is always "Not applicable", as they are represented by a simple set of metadata which is not protected by the copyright law (German: "Urheberrechtsgesetz").  If "Other" is selected, an explanation of the re-use terms and conditions must be provided in the property "Resource Use Rights Description".

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666



# Testing CEDAR tools in specific case studies



## Metadata Schema in CEDAR (1)

- CEDAR allowed us to create forms consisting that correspond to the data elements in the Excel sheet
- Using a easy graphical editor
- Supporting collaborative editing

A screenshot of a web-based metadata schema editor. The title bar reads "NFDI4Health C19 StudyHub Metadata Schema". The interface shows a list of four schema elements, each with a plus icon, a list icon, a red star icon, a pencil icon, and a close icon. The first element is "ID of the resource" (1), the second is "Type of the resource" (2), the third is "General type of the resource" (2.a), and the fourth is "Title" (1..4). The "Title" element is expanded, showing a text input field, a dropdown menu for "Language" with options "EN (English)", "DE (German)", and "ES (Spanish)", and a question mark icon. A sidebar on the right contains icons for a list, a calendar, an envelope, a hash, and a magnifying glass.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

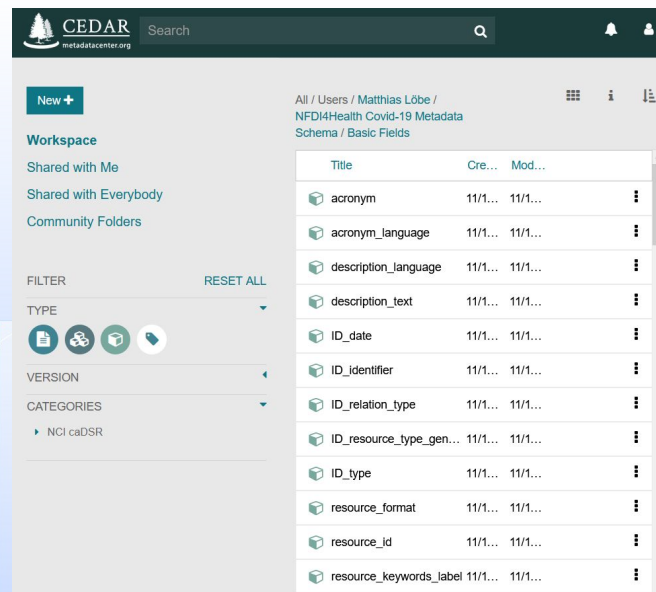
# Testing CEDAR tools in specific case studies



## Metadata Schema in CEDAR (2)



- To enable re-use, we first defined all data elements as CEDAR fields
- Currently, the Metadata Schema consists of 109 fields
- All data elements from the NFDI4Health could be realized

A screenshot of the CEDAR metadata center web application. The interface shows a sidebar with navigation options like 'New', 'Workspace', 'Shared with Me', 'Shared with Everybody', and 'Community Folders'. It also has filter sections for 'TYPE', 'VERSION', and 'CATEGORIES'. The main area displays a table of metadata fields for 'NFDI4Health Covid-19 Metadata Schema / Basic Fields'.

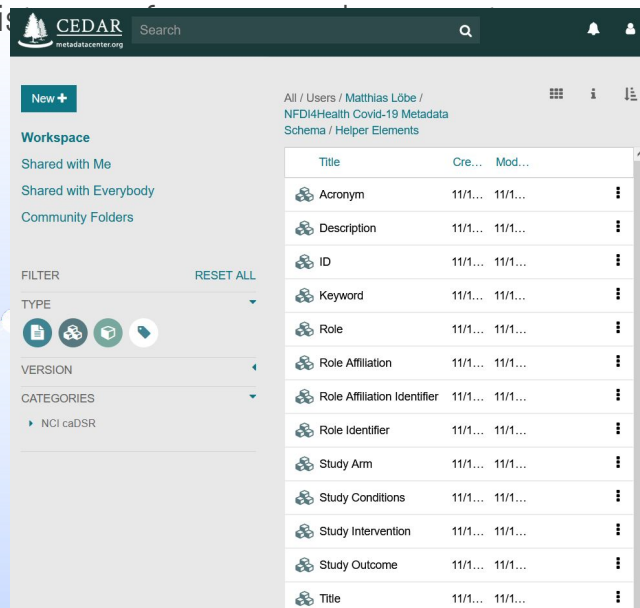
Title	Cre...	Mod...
acronym	11/1...	11/1...
acronym_language	11/1...	11/1...
description_language	11/1...	11/1...
description_text	11/1...	11/1...
ID_date	11/1...	11/1...
ID_identifier	11/1...	11/1...
ID_relation_type	11/1...	11/1...
ID_resource_type_gen...	11/1...	11/1...
ID_type	11/1...	11/1...
resource_format	11/1...	11/1...
resource_id	11/1...	11/1...
resource_keywords_label	11/1...	11/1...

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# Testing CEDAR tools in specific case studies

## Metadata Schema in CEDAR (3)

- Complex data elements, which logically consist of several elements, can occur several times, were implemented as CEDAR elements

A screenshot of the CEDAR metadata schema interface. The interface shows a sidebar with navigation options like "Workspace", "Shared with Me", and "Community Folders". The main area displays a table of metadata elements. The table has columns for "Title", "Cre...", and "Mod...". The elements listed include Acronym, Description, ID, Keyword, Role, Role Affiliation, Role Affiliation Identifier, Role Identifier, Study Arm, Study Conditions, Study Intervention, Study Outcome, and Title. Each element has a version number "11/1..." and a "Mod..." column with a dropdown arrow.

Title	Cre...	Mod...
Acronym	11/1...	11/1...
Description	11/1...	11/1...
ID	11/1...	11/1...
Keyword	11/1...	11/1...
Role	11/1...	11/1...
Role Affiliation	11/1...	11/1...
Role Affiliation Identifier	11/1...	11/1...
Role Identifier	11/1...	11/1...
Study Arm	11/1...	11/1...
Study Conditions	11/1...	11/1...
Study Intervention	11/1...	11/1...
Study Outcome	11/1...	11/1...
Title	11/1...	11/1...

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# Testing CEDAR tools in specific case studies



## Metadata Schema in CEDAR (4)



- CEDAR templates can be shared as OpenView: <https://bit.ly/3HxCcxy> (no account needed)
- CEDAR templates can be populated (e.g. for rapid prototyping)

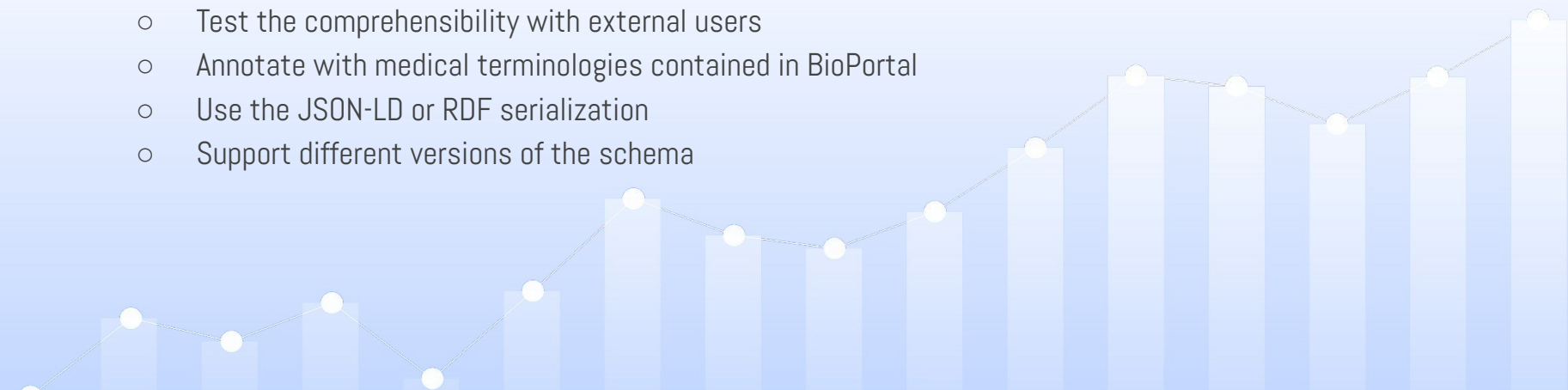
A screenshot of a web browser window displaying the "CEDAR OpenView" interface. The browser's address bar shows the URL "https://openview.metadatascenter.org/templates/https%2F%2F...". The page title is "NFDI4Health Covid-19 Metadata Schema: Basic Information". The form contains several fields: "Type of the resource" with a dropdown arrow, "General type of the resource" with a dropdown arrow, "Title (1 of 4)", "Acronym (1 of 4)", "Description (1 of 2)", and "Keyword (1 of 5)". Below these is a section for "Primary language of the resource" with radio buttons for "EN (English)", "DE (German)", "ES (Spanish)", "FR (French)", and "Other". There are also fields for "Web page" and "Version".

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# Testing CEDAR tools in specific case studies

## Conclusion

- CEDAR can make valuable contributions for developing metadata schemas
  - Ambiguities are particularly apparent in an implementation!
- Additionally, in the future we could:
  - Test the comprehensibility with external users
  - Annotate with medical terminologies contained in BioPortal
  - Use the JSON-LD or RDF serialization
  - Support different versions of the schema



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# Testing CEDAR tools in specific case studies

## Use case 3: Health-RI / ZonMw (Covid-19 scenarios)

- ❖ Dutch Covid-19 funding program (ZonMw)
- ❖ Total budget: € 115.000.000
- ❖ 3 focus areas
  - Predictive diagnostics and treatment
  - Care and prevention
  - Societal dynamics
- ❖ Cooperations
  - The Viral Outbreak Data Network (VODAN) of GO-FAIR
  - WHO Coordinated Global Research Roadmap
  - Health-RI (Dutch Health Research Infrastructure)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# Testing CEDAR tools in specific case studies

Health-RI / ZonMw (Covid-19 scenarios)

Quote from the website:

## Open Science, Open Access and FAIR data

In line with ZonMw and NWO signing the [international statement](#) to make all data and publications publicly available as quickly as possible in the fight against COVID-19, **all data and results must be published without embargo immediately** in Open Access form and in accordance with [research data management and stewardship \(RDM\)](#) so that all data are FAIR.

Open Science in COVID-19 research: [All you need to know about open science, open access publications and FAIR data stewardship in COVID-19 research projects.](#)

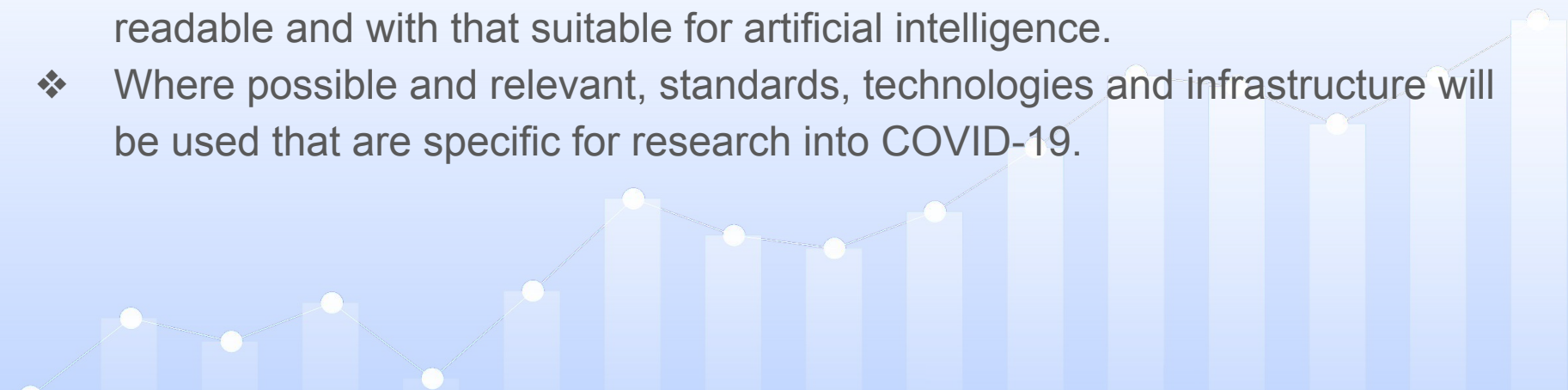
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666



# Testing CEDAR tools in specific case studies

## ... so that all data are FAIR

- ❖ Application of **FAIR data stewardship / management** to make data in the project **reusable** for the entire research community, and preferably computer-readable and with that suitable for artificial intelligence.
- ❖ Where possible and relevant, standards, technologies and infrastructure will be used that are specific for research into COVID-19.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# Testing CEDAR tools in specific case studies

## Practical approach

- ❖ **130 projects**, all needed to make their data FAIR
- ❖ **Centralized process** to capture ZonMw COVID-19 Program metadata
- ❖ Three core templates:
  - Project Admin [in OpenView](#)
  - Project Content [in OpenView](#)
  - GDMT - Generic Dataset Metadata Template (now the DCAT templates)
- ❖ Required:
  - Vocab creation → BioPortal [ZonMw COVID-19 Vocabulary](#)
  - Collection & deposition of metadata in CEDAR

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# Testing CEDAR tools in specific case studies

❖ Year-long M4M-program, a.k.a. “Slow M4M” 😊

Monday, November 1st ▾



**Mijke Jetten** 11:55 AM

Reminder: very welcome to join us in a few minutes to ask your Covid CEDAR form questions. 📌 (edited)

*Members of the M4M support team will be available on Zoom ([https://zoom.us/j/\[REDACTED\]](https://zoom.us/j/[REDACTED])) to answer your questions on the CEDAR forms. Today meetings are scheduled for:*

- 1 November 12-13 o'clock
- 1 November 15-16 o'clock

(edited)

Tuesday, November 16th ▾



**Mijke Jetten** 12:23 AM

☀️ Reminder: very welcome to join us today to ask your Covid CEDAR form questions. ☀️ (edited)

*Members of the M4M support team will be available on Zoom ([https://zoom.us/j/\[REDACTED\]](https://zoom.us/j/[REDACTED])) to answer your questions on the CEDAR forms.*

- 16 November 11-12 o'clock
- 18 November 16-17 o'clock

*You can also contact us via [servicedesk@health-ri.nl](mailto:servicedesk@health-ri.nl)*

Today ▾



**Mijke Jetten** 12:08 PM

!🔔 Last help session for Covid-19 forms, today Nov 18 16-17 CET: Members of the M4M support team will be available on Zoom ([https://zoom.us/j/\[REDACTED\]](https://zoom.us/j/[REDACTED])) to answer your questions

## Project overview



Request data

Reset

1 to 25 of 105

Lead institution



- ☐ Academic Medical Center - Amsterdam UMC
- ☐ Eindhoven University of Technology
- ☐ Erasmus Centrum voor Zorgbestuur
- ☐ Erasmus MC
- ☐ Erasmus University Rotterdam
- ☐ GGD Amsterdam - Public Health Service of Amsterdam
- ☐ GGD Zuid Holland Zuid
- ☐ Hogeschool van Amsterdam - Amsterdam University of Applied Sciences
- ☐ IKNL - Netherlands Comprehensive Cancer Organisation
- ☐ Leiden University Medical Center
- ☐ Maastricht University
- ☐ Maastricht University Medical Centre
- ☐ NIVEL - Netherlands Institute for Health Services Research
- ☐ NSCR - Netherlands Institute for the Study of Crime and Law Enforcement
- ☐ Radboud University Nijmegen
- ☐ Radboud University Nijmegen Medical Centre
- ☐ Sanquin
- ☐ Saxion University of Applied Sciences
- ☐ SEO Economisch Onderzoek
- ☐ Spaarne Gasthuis - Spaarne Ziekenhuis
- ☐ Tilburg University
- ☐ University Medical Center Groningen
- ☐ University Medical Center Utrecht
- ☐ University of Amsterdam
- ☐ University of Groningen
- ☐ University of Leiden

De impact van COVID-19 onder laaggeletterden en mensen met een LVB: maatregelen, (mentale) gezondheid en handelingsperspectieven t.a.v. de zorg- en ondersteuningsbehoefte en beleid

Last changed  
18/11/2021

COVID-19: LVB en laaggeletterden

ZonMw - Netherlands Organisation for Health Research and Development

Radboud University Nijmegen Medical Centre

Measuring understanding & reducing respiratory droplet spreading

Last changed  
16/11/2021

10430012010022

ZonMw - Netherlands Organisation for Health Research and Development

University of Twente

COVID-NL cohort ErasmusMC

Last changed  
16/11/2021

ErasmusMC COVID-NL cohort

other

Erasmus MC

Prospective cohort study of non-hospitalised COVID-19 patients: determining length of isolation and patient clinical development at home (COVID-HOME study)

Last changed  
16/11/2021

COVID-HOME

ZonMw - Netherlands Organisation for Health Research and Development

University Medical Center Groningen

VAN PREVENTIE NAAR VEERKRACHT: een design framework voor de 1,5 meter-samenleving

Last changed  
15/11/2021

ZonMw - Netherlands Organisation for Health Research and Development

Hogeschool van Amsterdam - Amsterdam University of Applied Sciences

Kansen in Crisis: Hoe geven scholen vorm aan educatief partnerschap?

Last changed  
12/11/2021

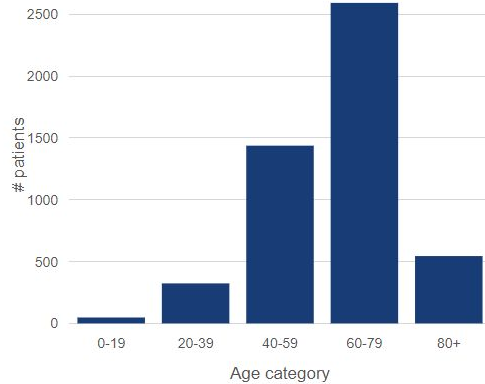
ZonMw - Netherlands Organisation for Health Research and Development

University of Amsterdam

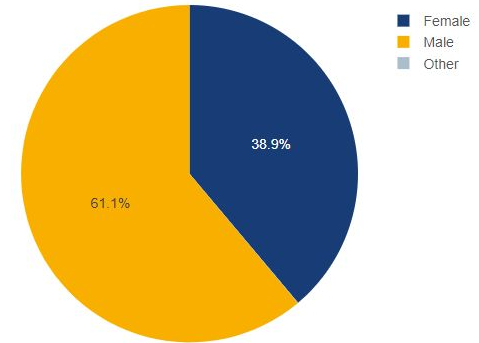
## COVID-NL clinical data dashboard



# patients included •	# COVID positive patients included •	# patients in IC •	# deceased patients •
5871	4527	1595	826



# centers included •
6



[Search deeper in the data](#)


[Request data](#)

Last updated: 18/11/2021

# Testing CEDAR tools in specific case studies

## Conclusion

- ❖ CEDAR and DCAT templates contribute to deposition of harmonized descriptive metadata
- ❖ CEDAR API enables processing and visualization of metadata
- ❖ Metadata made FAIR
- ❖ (all) data made Findable
  - Accessible, Interoperable, Reusable are work in progress



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

# Discussion & AOB

Join us!

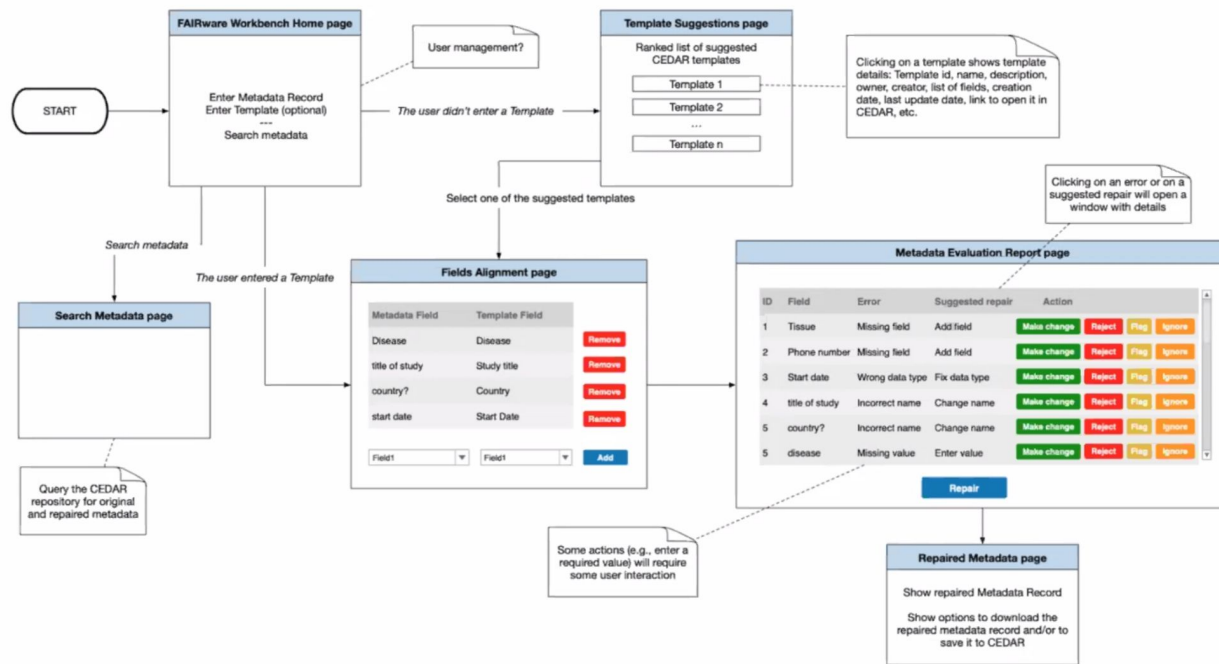
<https://bit.ly/M4M-FAIR>

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666



# FAIRware workbench

## FAIRWare Workbench Workflow



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

спасибо  
danke 謝謝  
ngiyabonga  
tesekkür ederim  
tapadh leat  
dank je  
gracias  
moichakkeram  
go raibh maith agat  
arigato  
dankjem  
merci  
ευχαριστώ  
감사합니다  
obrigado  
sagolun  
sukriya  
kop khun krap  
grazie  
terima kasih  
dziękuję  
hvala  
mauruuru  
bedankt



[www.fair4health.eu](https://www.fair4health.eu)



@FAIR4Health  
#FAIR4Health



<https://www.linkedin.com/company/13977340/>



<https://www.youtube.com/channel/UCpycUIqaXMAJCZPatqcm4cg>

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666