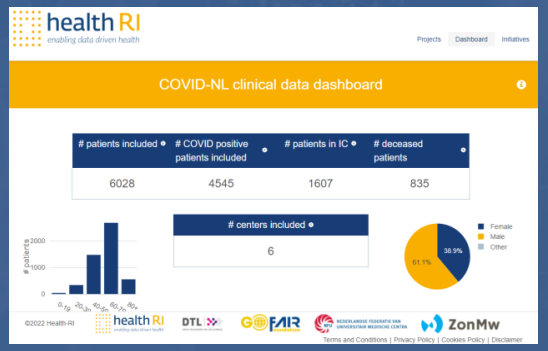
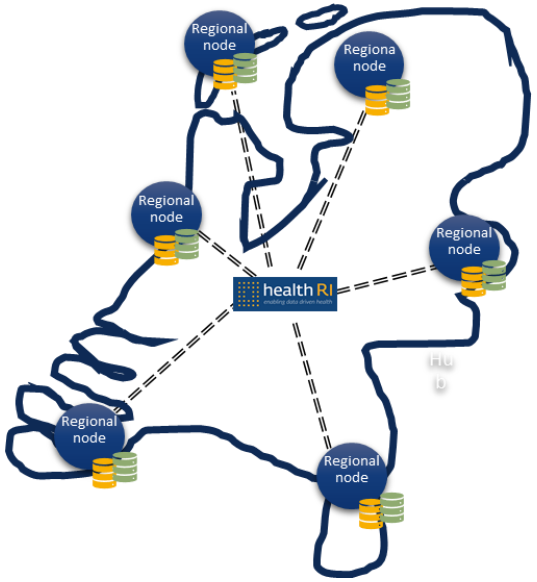
health RI

enabling data driven health



Covid-19 Health-RI Portal

Think big – start small – act now

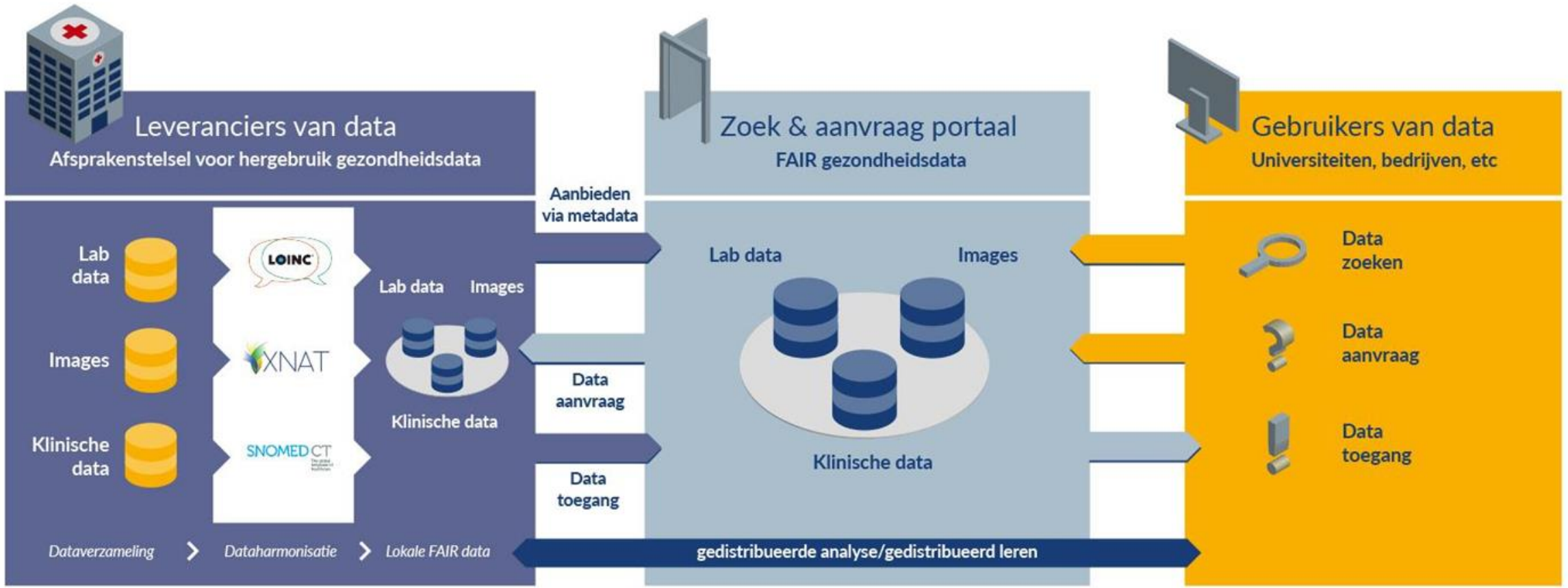


***Think big**: national health data infrastructure*

1. Connecting communities
2. Expertise, tools & services
3. Best practices

***Start small**: [national COVID-19 portal]*

Towards a National COVID-19 observational data portal
(see also https://www.health-ri.nl/covid-19-data-portal)

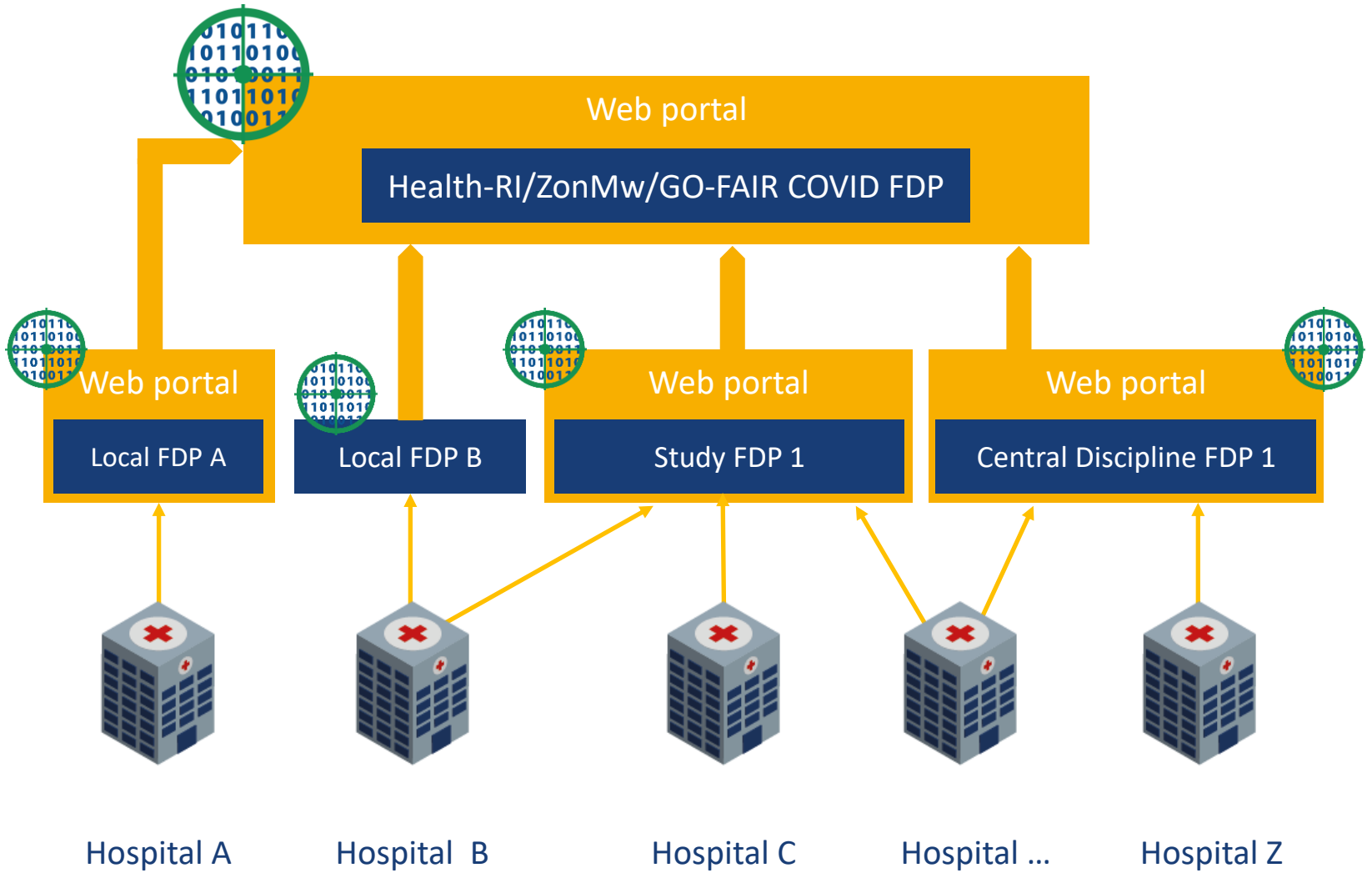


The hospitals - data contributors: Legal basis, data harmonization

The national portal - data discovery & accessibility

Users – controlled access & data reuse

Towards a national COVID-19 observational data portal



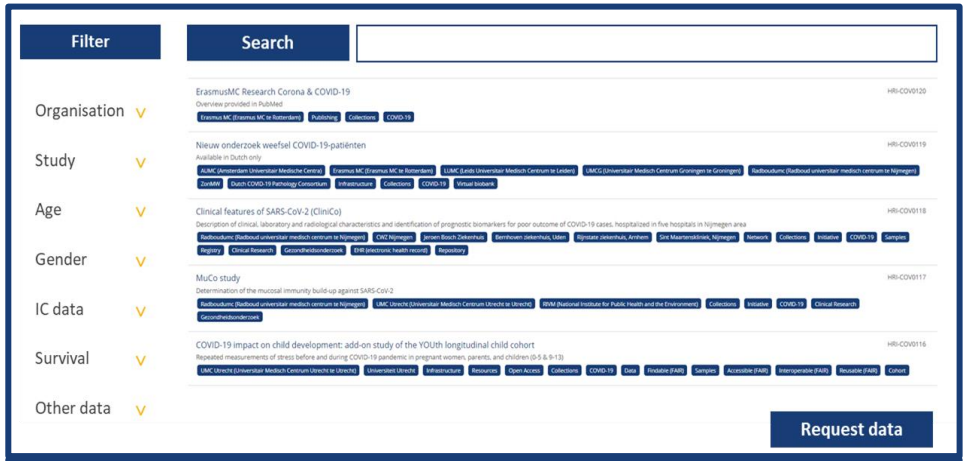
Administrative and descriptive metadata to facilitate data discovery & accessibility

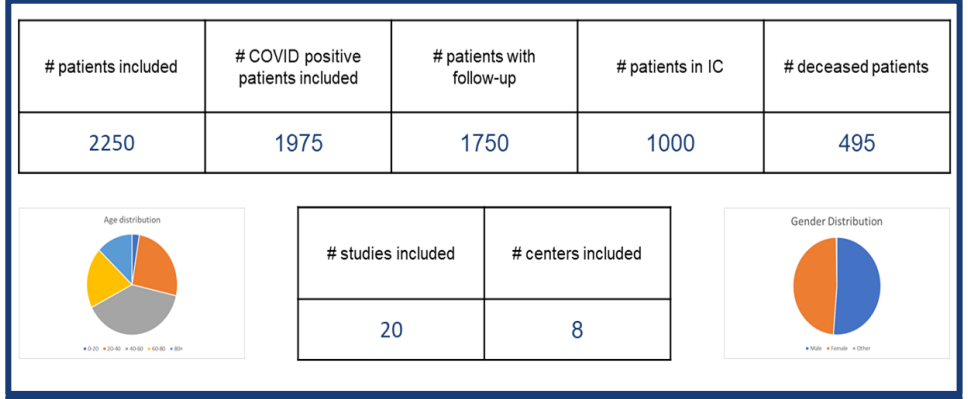
Data at patient/record level

Pseudonymisation keys

FDP: FAIR Data Point

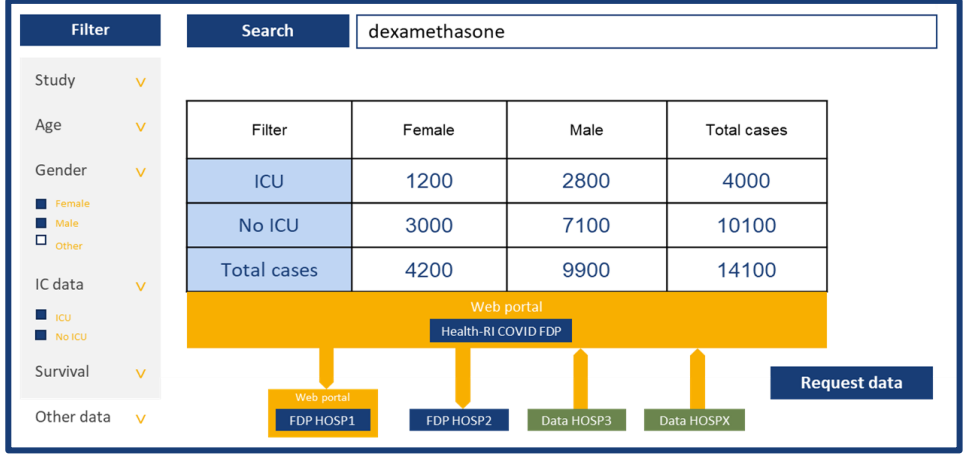
Portal - Different data presentation levels





# patients included	# COVID positive patients included	# patients with follow-up	# patients in IC	# deceased patients
2250	1975	1750	1000	495

# studies included	# centers included
20	8



Filter	Female	Male	Total cases
ICU	1200	2800	4000
No ICU	3000	7100	10100
Total cases	4200	9900	14100

Catalogus with metadata: **ZonMw studies**

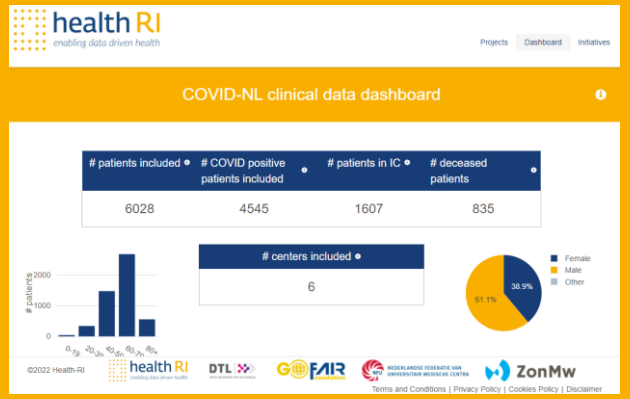
- NFU COVID cohort
- COVID cohort per UMC
- Other studies, COVIDpredict, CAPACITY, etc

High level aggregates (Dashboard)

Search on allowed items and present aggregates

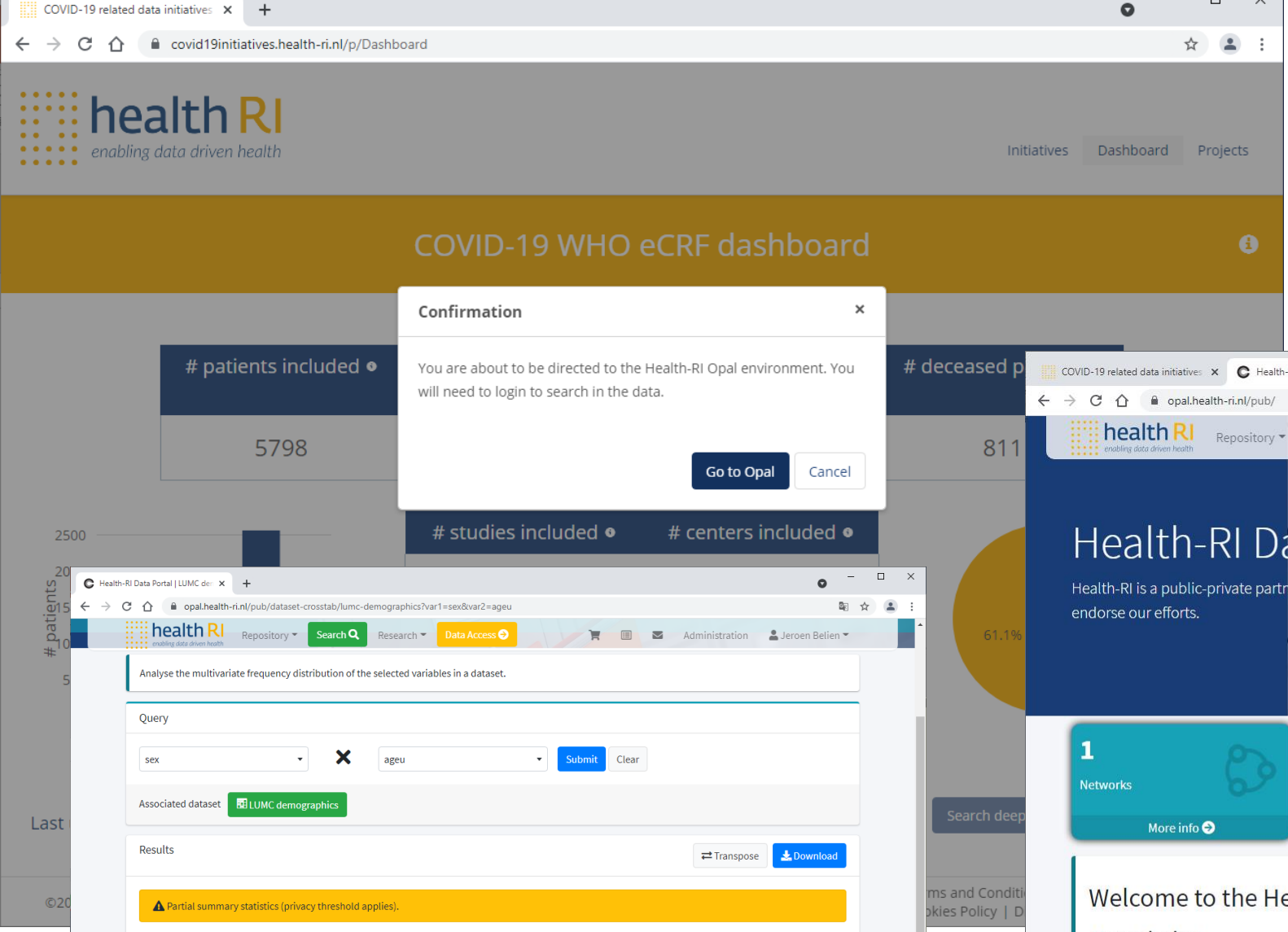
Lowest level - raw data (incl ETL)

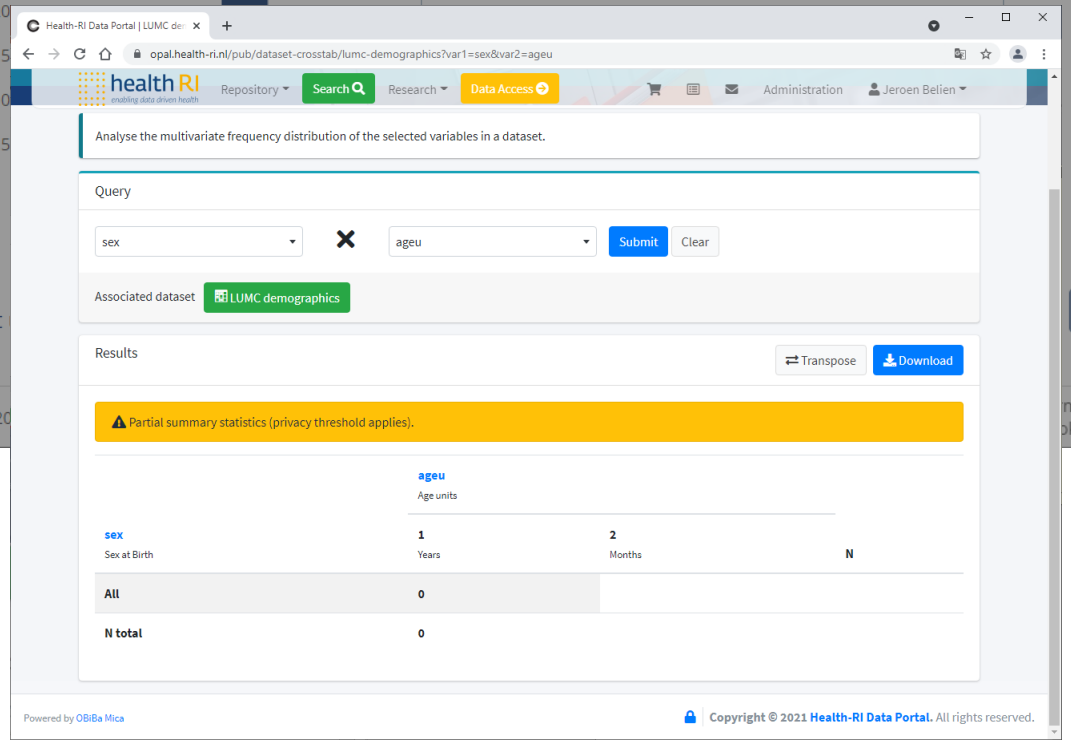
NEDERLANDSE FEDERATIE VAN UNIVERSITAIR MEDISCHE CENTRA

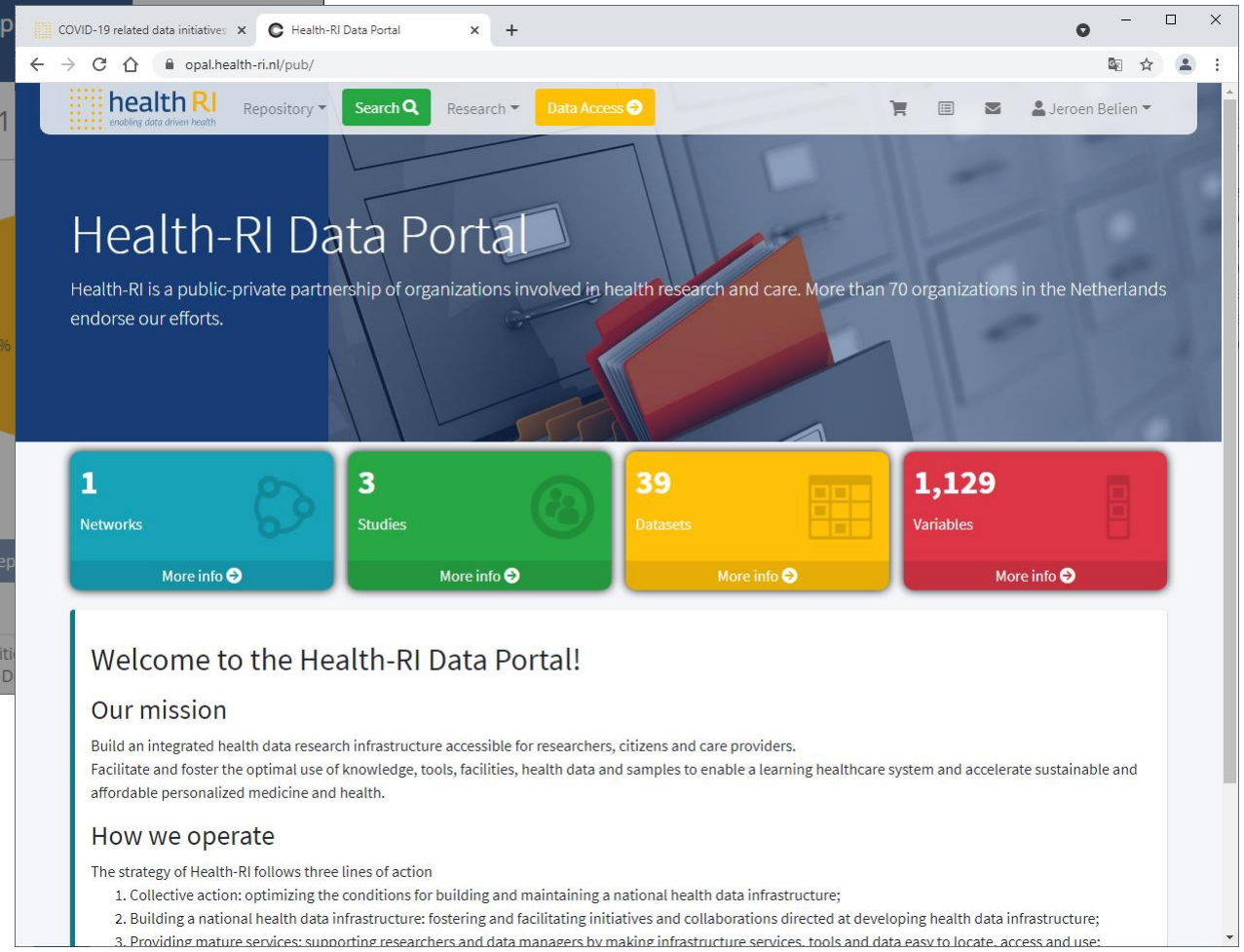


What a potential researcher can see and do:

Live demo

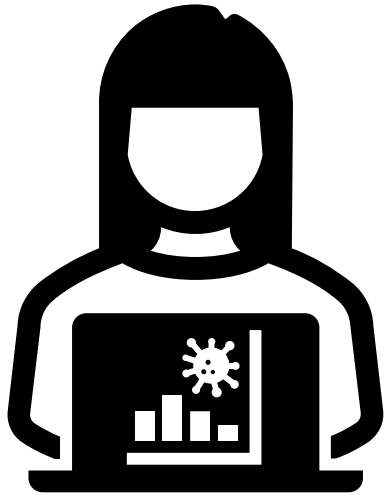

COVID-19 related data initiatives
covid19initiatives.health-ri.nl/p/Dashboard
health RI
enabling data driven health
Initiatives
Dashboard
Projects
COVID-19 WHO eCRF dashboard
Confirmation
You are about to be directed to the Health-RI Opal environment. You will need to login to search in the data.
Go to Opal
Cancel
patients included
5798
deceased p
811
studies included
centers included
2500
patients
61.1%
Search deep
Last
©20
ms and Conditi
okies Policy | D


Health-RI Data Portal | LUMC de
opal.health-ri.nl/pub/dataset-crosstab/lumc-demographics?var1=sex&var2=ageu
health RI
Repository
Search
Research
Data Access
Administration
Jeroen Belien
Analyse the multivariate frequency distribution of the selected variables in a dataset.
Query
sex
ageu
Submit
Clear
Associated dataset
LUMC demographics
Results
Transpose
Download
Partial summary statistics (privacy threshold applies).
ageu
Age units
sex
Sex at Birth
1
Years
2
Months
N
All
0
N total
0
Powered by OBiBa Mica
Copyright © 2021 Health-RI Data Portal. All rights reserved.


COVID-19 related data initiatives
Health-RI Data Portal
opal.health-ri.nl/pub/
health RI
Repository
Search
Research
Data Access
Jeroen Belien
Health-RI Data Portal
Health-RI is a public-private partnership of organizations involved in health research and care. More than 70 organizations in the Netherlands endorse our efforts.
1
Networks
More info
3
Studies
More info
39
Datasets
More info
1,129
Variables
More info
Welcome to the Health-RI Data Portal!
Our mission
Build an integrated health data research infrastructure accessible for researchers, citizens and care providers.
Facilitate and foster the optimal use of knowledge, tools, facilities, health data and samples to enable a learning healthcare system and accelerate sustainable and affordable personalized medicine and health.
How we operate
The strategy of Health-RI follows three lines of action
1. Collective action: optimizing the conditions for building and maintaining a national health data infrastructure;
2. Building a national health data infrastructure: fostering and facilitating initiatives and collaborations directed at developing health data infrastructure;
3. Providing mature services: supporting researchers and data managers by making infrastructure services, tools and data easy to locate, access and use;

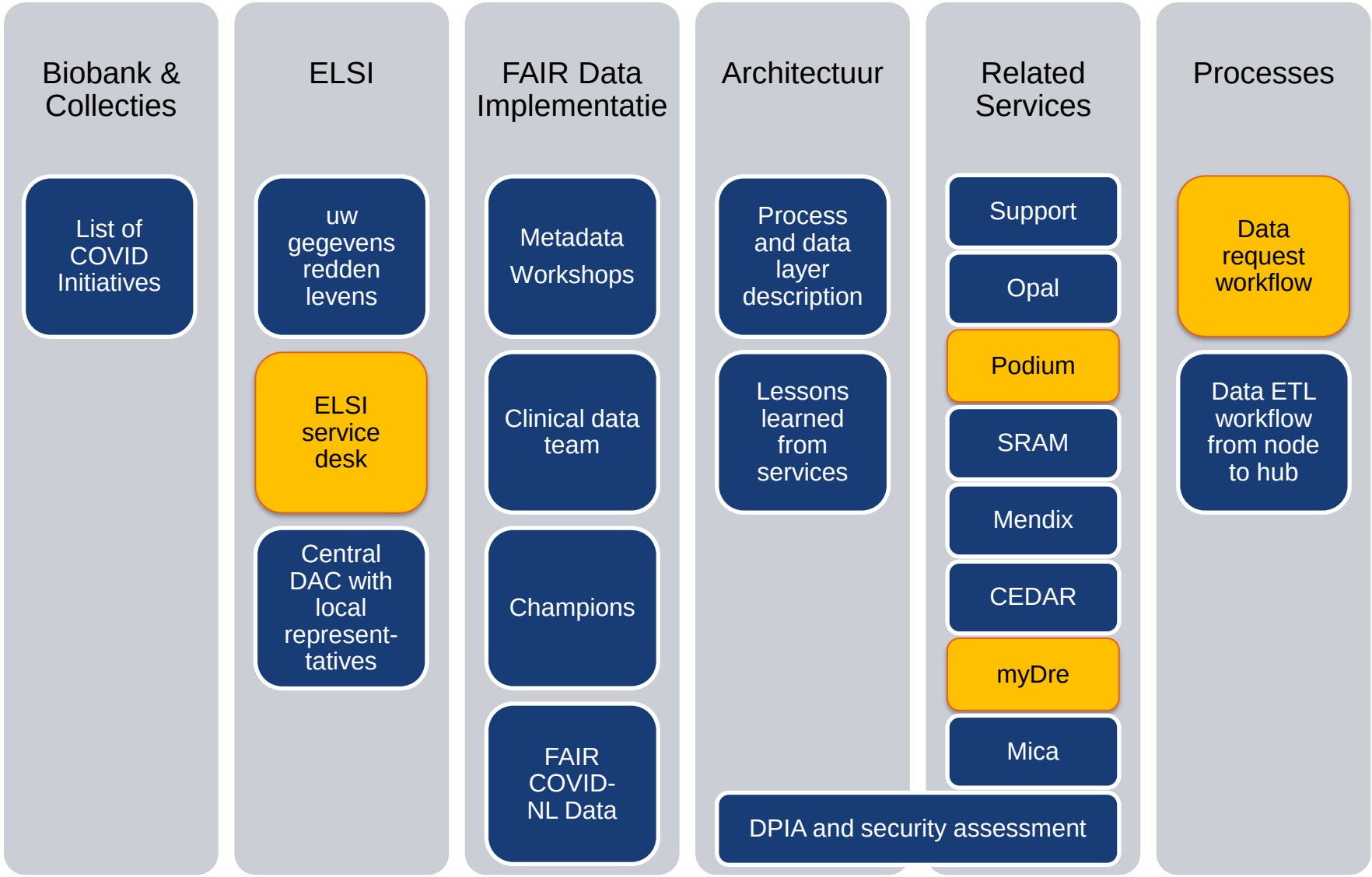
Data source

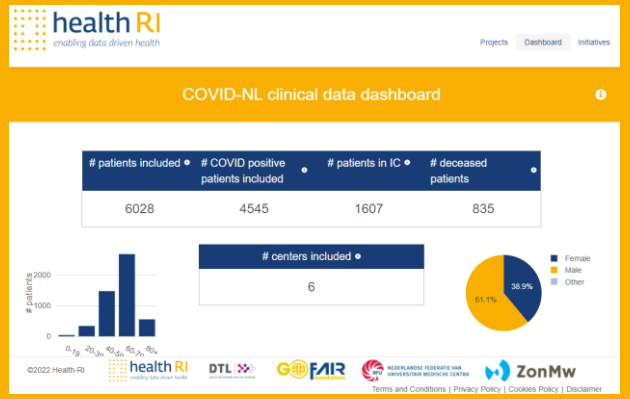
COVID PORTAL



Secondary/ tertiary Data Use

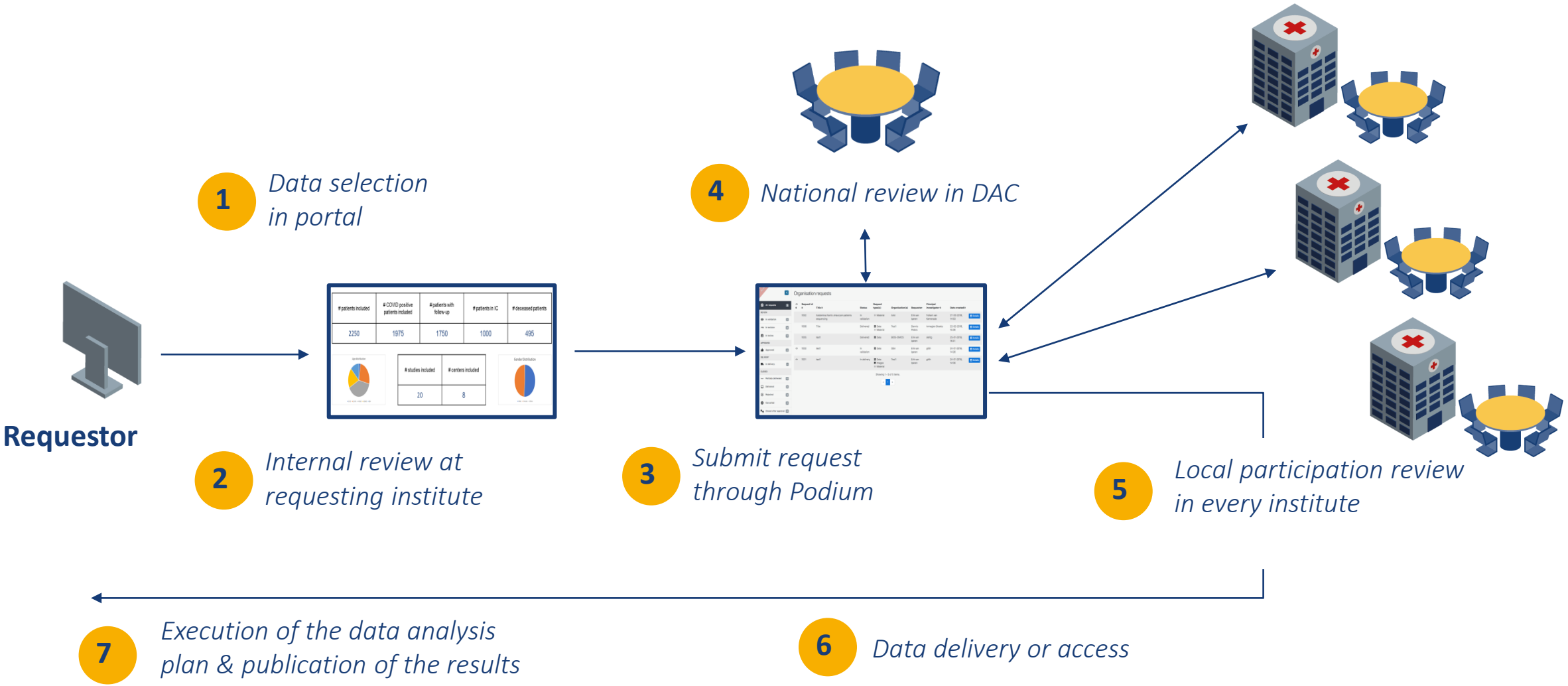
Covid Portal in Context – Achievements of Pilot





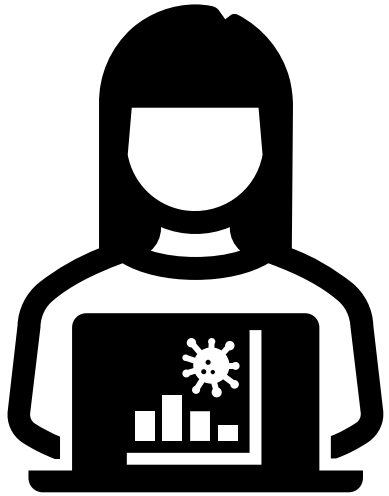
How to request and get data:

Data governance: COVID-NL data request workflow



Data source

COVID PORTAL



Secondary/ tertiary Data Use

Secondary Data Use

COVID PORTAL

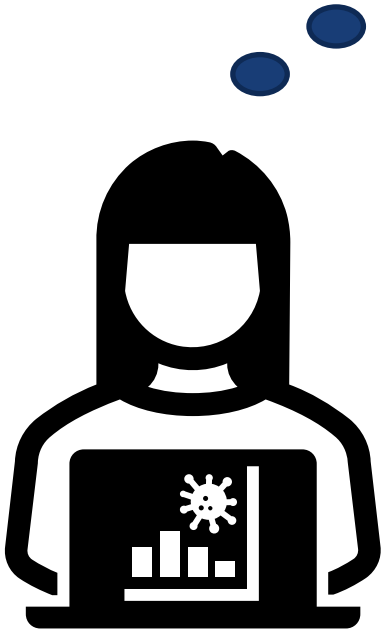
Query by researcher

Check if Consent is still applicable

Transfer to safe environment/ known and trusted

API from anDREa

Workspace in myDRE

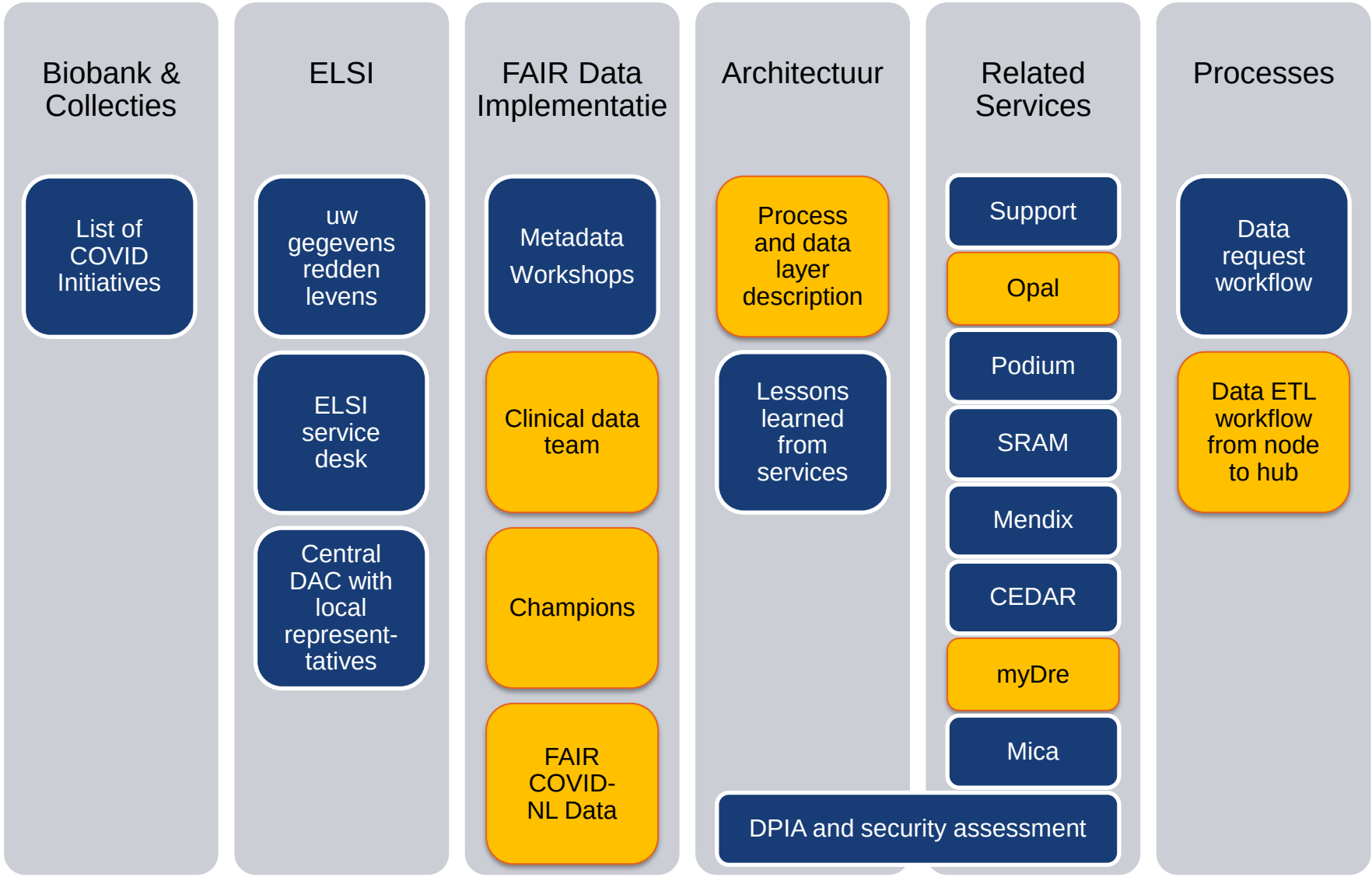


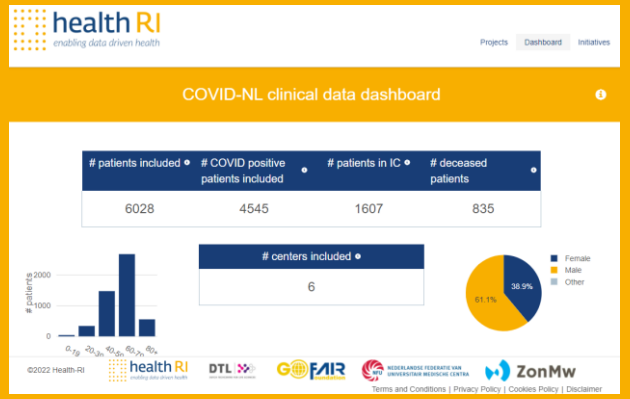
DMP (i.a. long-term storage)

Archiving (with DOI)

Yoda, Figshare…

Covid Portal in Context – Achievements of Pilot





How the data get into the portal:

- Process and data layer description
- Data ETL workflow from node to hub
- Clinical Data team and data champions

COVID-19 two major data flows

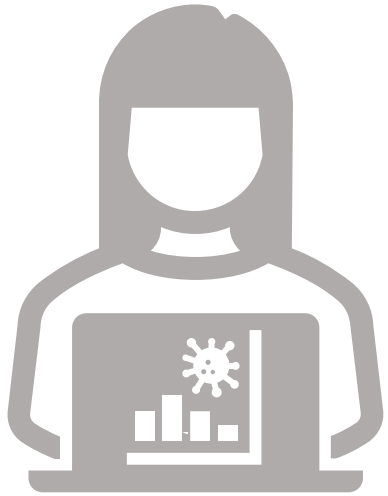
Two major (funded) COVID-19 research programs initiated by:
1) ZonMw: Dutch funder of health research and innovation projects.
2) NFU: Dutch Federation of University Medical Centers

Objective: Build national observational COVID-19 data portal that will help researchers to find and reuse COVID-19 related observational data

Ad 1)
- Involves over 60 funded projects covering topics from health care incl. diagnostics, prognosis and immunology, to socio-economical
- a series of training events for principal investigators and/or data stewards in applying FAIRification approaches to improve interoperability of study related (meta)data

Ad 2)
- Data from (UMC) hospitalized COVID-19 patients
- captured in various systems like EHRs and EDCs (Castor & Redcap)
- Mappings and transformations to WHO/ISARIC Core or Rapid CRF codebooks
- Re-use templates developed by 1) for aggregated (meta)data on portal

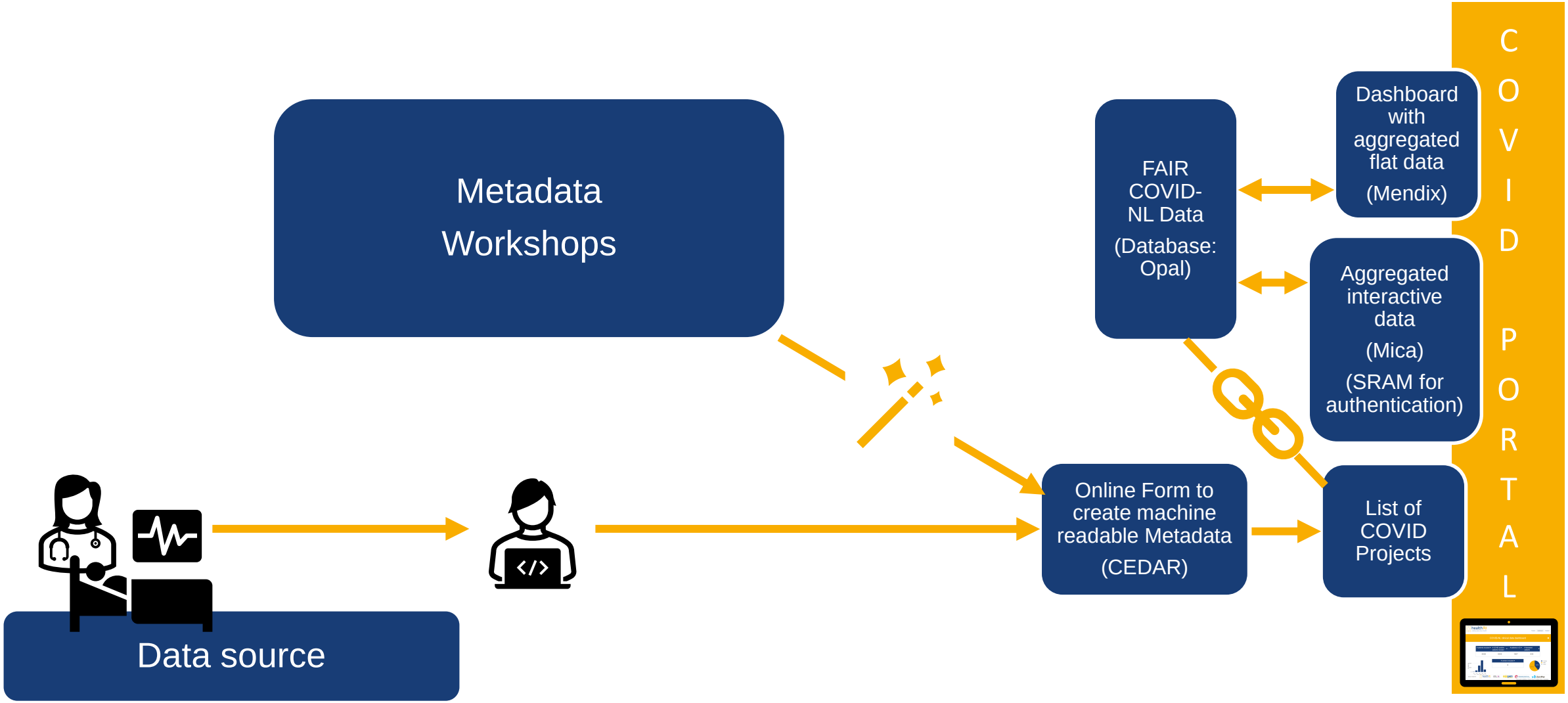
COVID PORTAL



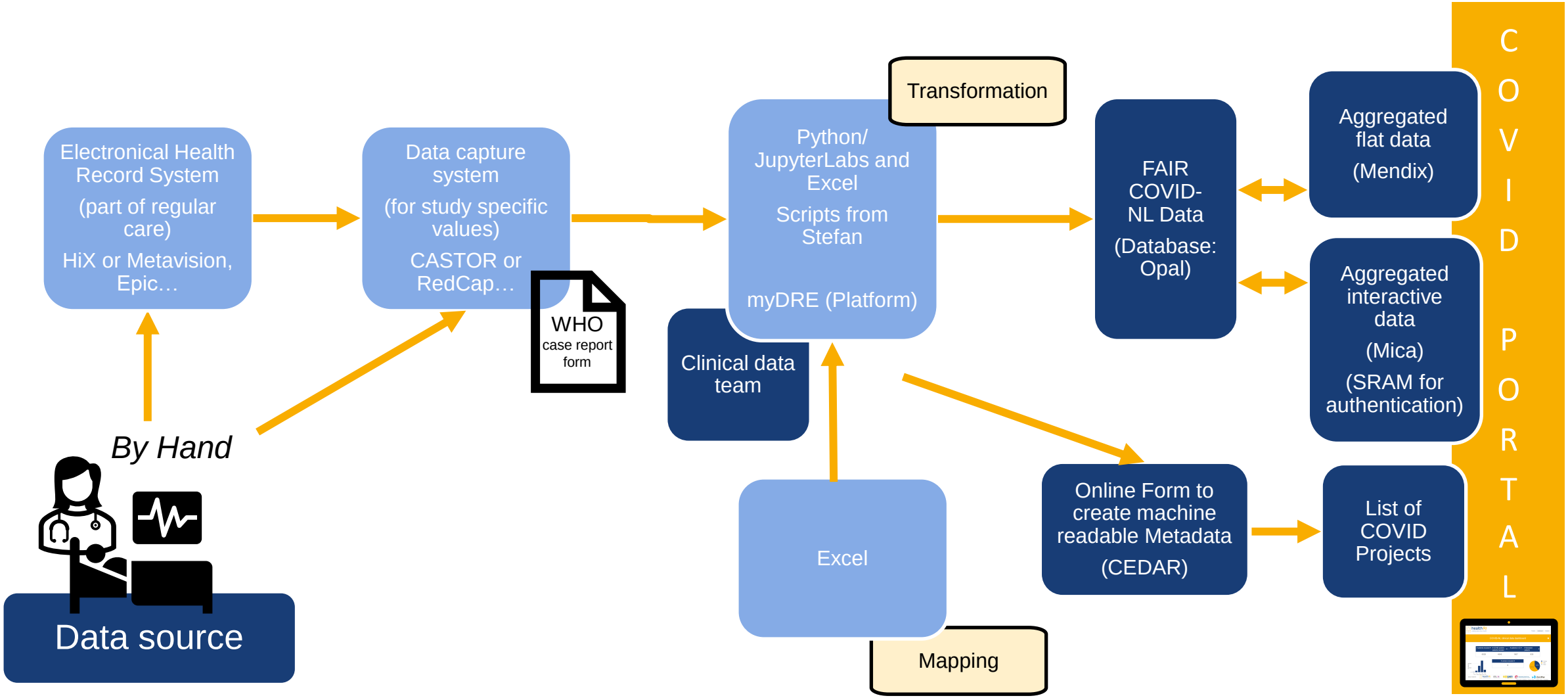
Data source

Secondary/ tertiary Data Use

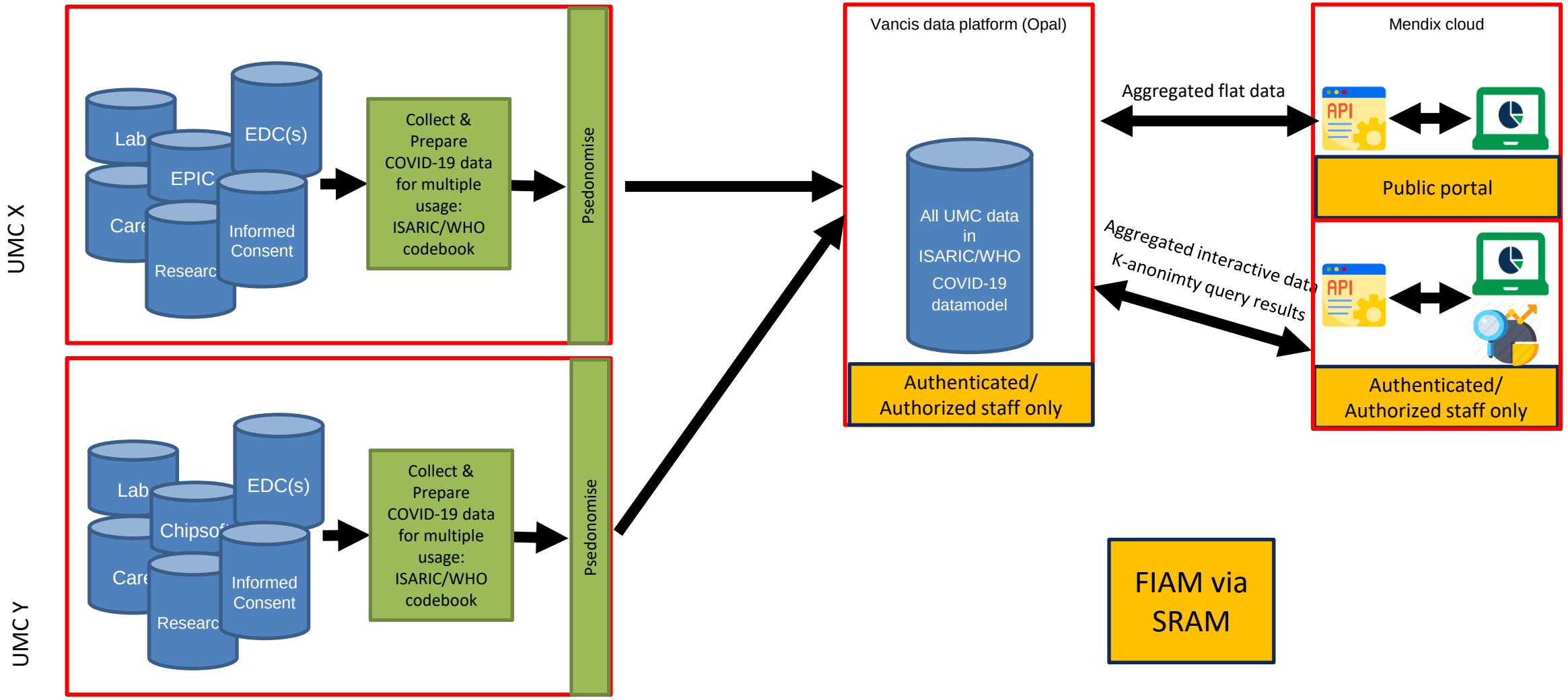
Metadata


Metadata Workshops
FAIR COVID-NL Data
(Database: Opal)
Dashboard with aggregated flat data
(Mendix)
Aggregated interactive data
(Mica)
(SRAM for authentication)
COVID PORTAL
Online Form to create machine readable Metadata
(CEDAR)
List of COVID Projects
Data source

Data ETL workflow from node to hub

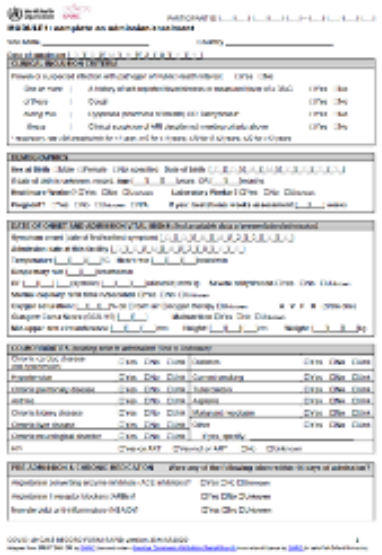


Q1-2022: COVID-19 NL dataflow UMC's via central hosted Opal @Vancis

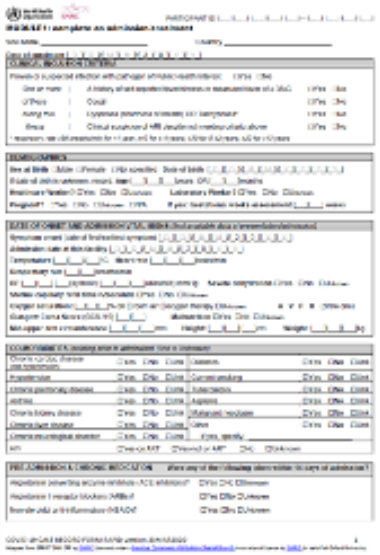


Disclaimer: deduplication of patients is not (yet) possible

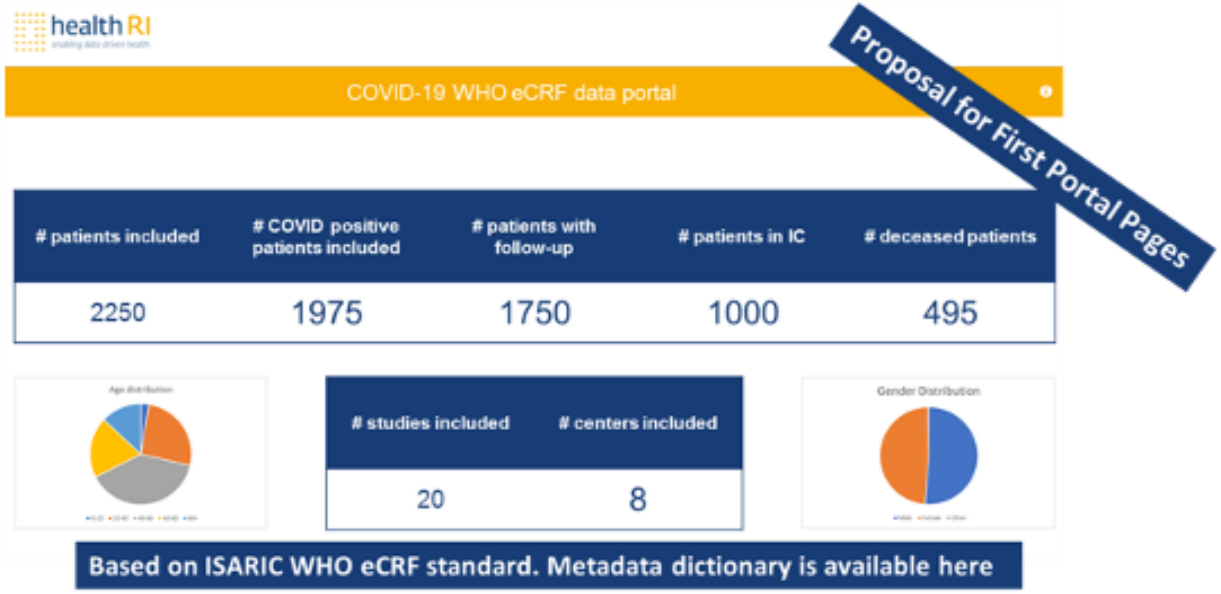
The road (initial steps) towards a national COVID-19 observational data portal



WHO-CRF codebook per hospital/study/..



WHO-CRF codebook harmonised to ISARIC



Simple website; central only metadata and aggregated numbers of limited set of data items. No choice of underlying tool (s) yet

Global timeline

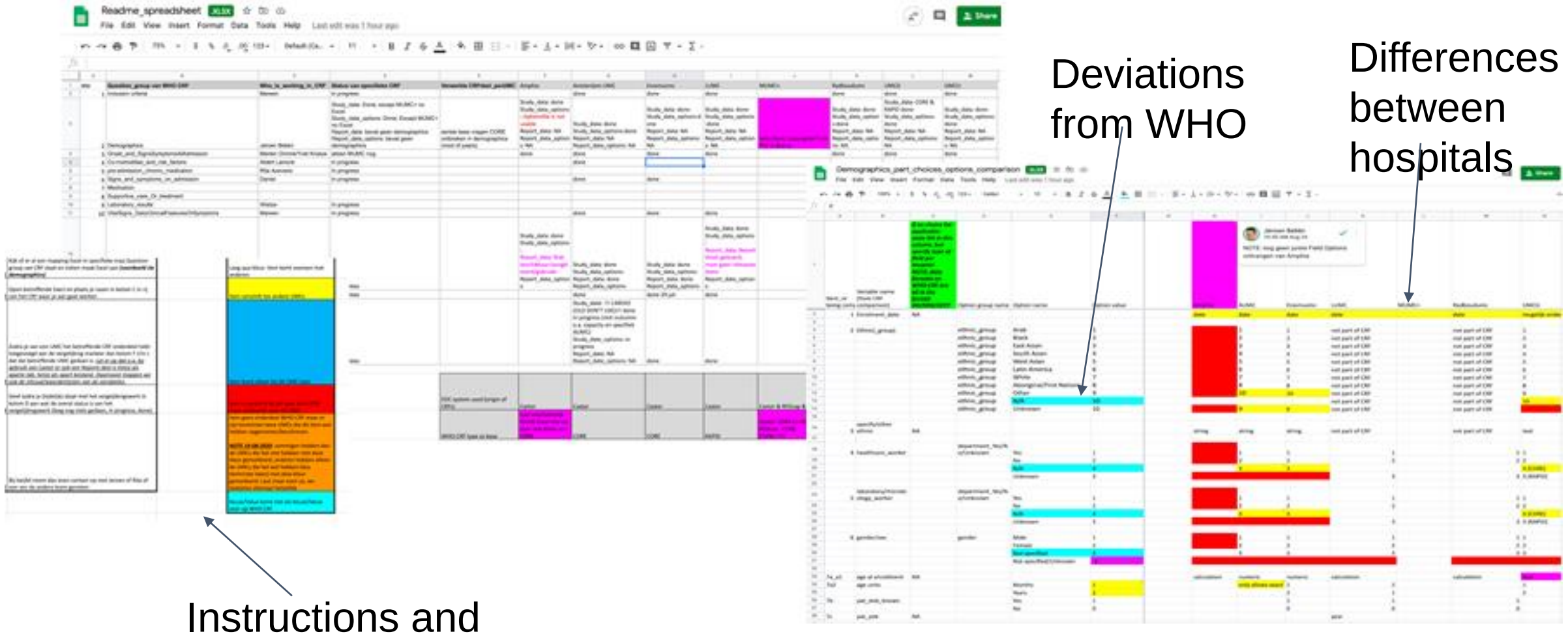
First week July 2020

October 2020

October 2020

NFU & STZ COVID Clinical data working group

Snapshots Mapping (impression): each local codebook -> ISARIC



COVID-NL data – status

UMC	Data in dashboard	CRF mapping	Data transformatie	Data beschikbaar voor OPAL	Scrambled data beschikbaar voor OPAL	Scrambled data in OPAL *	Data in OPAL *
AUMC	Ja	Ja	Ja	Ja	Ja	Nee	Nee
ErasmusMC	Ja	Ja	Ja	Ja	Ja	Ja	Nee
LUMC	Ja	Ja	Ja	Ja	Ja	Ja	Nee
MUMC	Ja	Ja	Ongoing	Nee	Nee	Nee	Nee
RadboudUMC	Ja	Ja	Ja	Ja	Ja	Ja	Nee
UMCG	Ja	Ja	Ja	Ja	Ja	Nee	Nee
UMCU	Nee	Ja	Ja	Ja	Ja	Nee	Nee

* Pending on update of OPAL environment after security fixes

NFU Clinical data team/werkgroep : collaboration of Health-RI nodes in practice

Health-RI	Wiro Niessen (ErasmusMC)
	Gerrit Meijer (NKI)
	Mijke Jetten (DTL)
	Jeroen Beliën (AUMC)
	Rita Azevedo (Lygature)
	Julia Lischke (Lygature)
Amphia	Henriette Cuijpers
	Wilco Jacobs
AUMC	Meriem Manaï
	Sjoerd Manger
	Rudy Scholte
	Mareen Datema
ErasmusMC	Heleen Pruijn
	Stefan van Aalst
	Raoul Boomsma
	Anne-Lotte Masson
LUMC	Karin van der Pal
	Merijn Driessen
	Anke Wegman
MUMC+	Olav Palmen
	Pascal Suppers
	Wendy Neyens
OLVG	Ise Butter
RadboudUMC	Kit Roes
	Marten Onnink
	Mirjam Brullemans
	Yvet Kroeze
UMCG	Rene Oostergo
	Marloes Cardol
	Bas de Reuver
UMCU	Nicole Boekema-Bakker
	Marieke Vianen
	Sam Keukens
	Joost Schotsman
	Paula Jansen
	Wietze Pasma

ZonMw

COVID-19 Program aims to make data FAIR

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

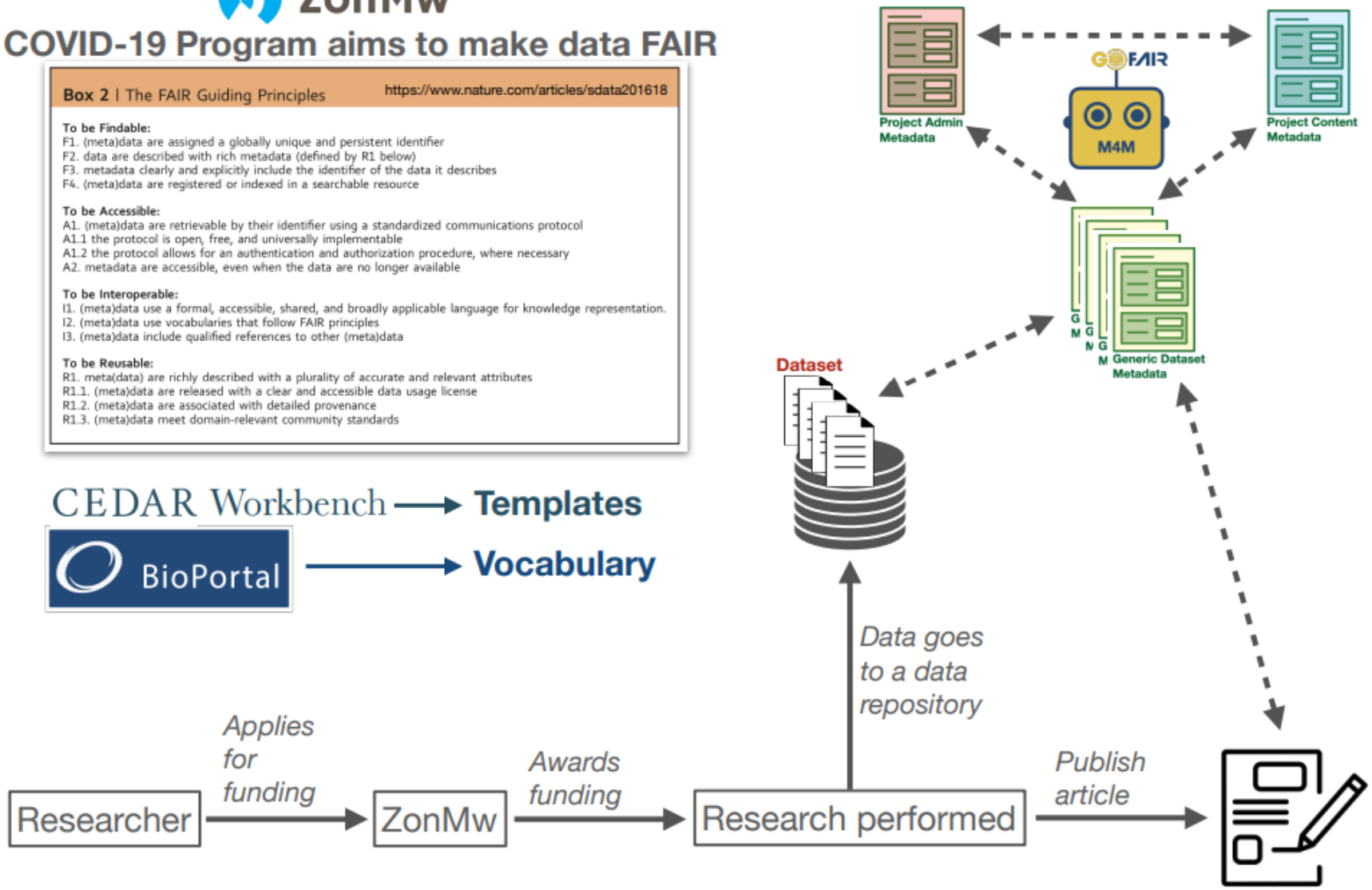
To be Accessible:
A1. (meta)data are retrievable by their identifier using a standardized communications protocol
A1.1 the protocol is open, free, and universally implementable
A1.2 the protocol allows for an authentication and authorization procedure, where necessary
A2. metadata are accessible, even 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

CEDAR Workbench → **Templates**

BioPortal → **Vocabulary**



FAIR Data Collective

Road (so far) to machine actionable (meta)data

Domain experts

+

FAIR metadata experts

=

Metadata schema

+

Vocabulary

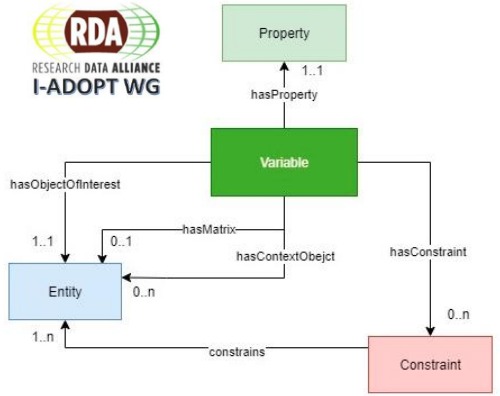
=

Variable

+

Vocabulary

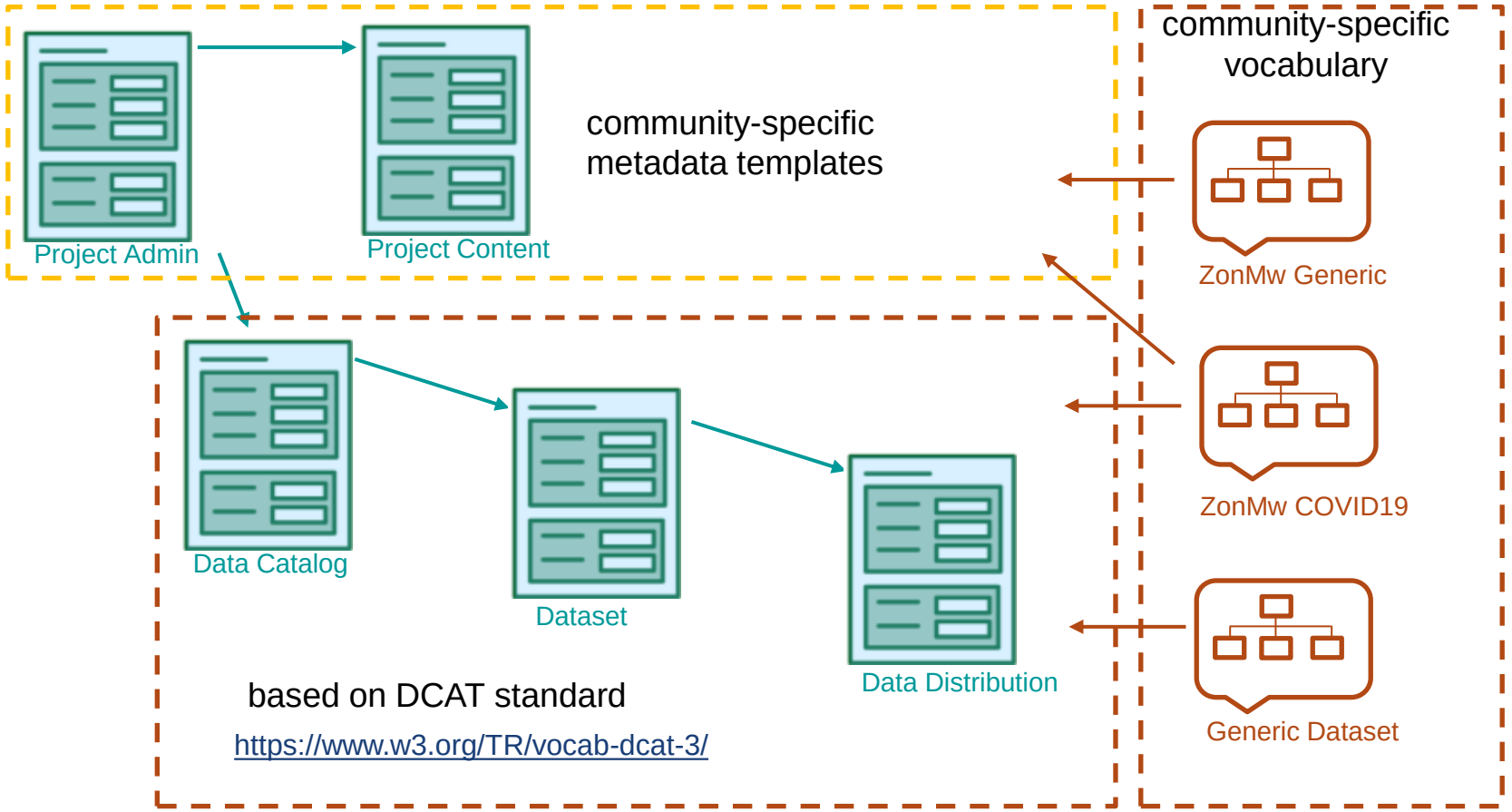
=



Data + FAIR metadata + FAIR Variable = FAIR Data

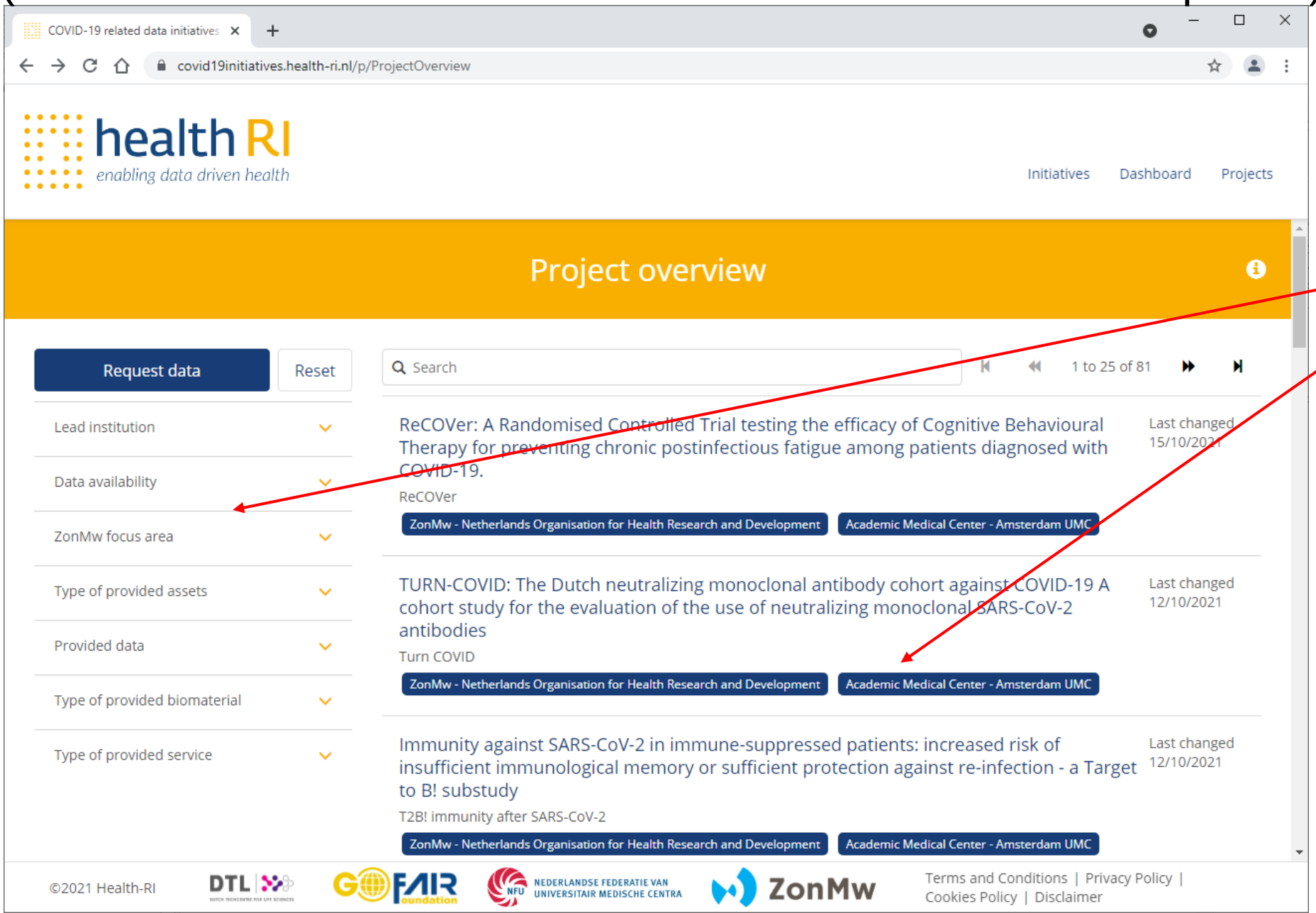
CC BY NC ND

Metadata templates and vocabulary for projects and datasets

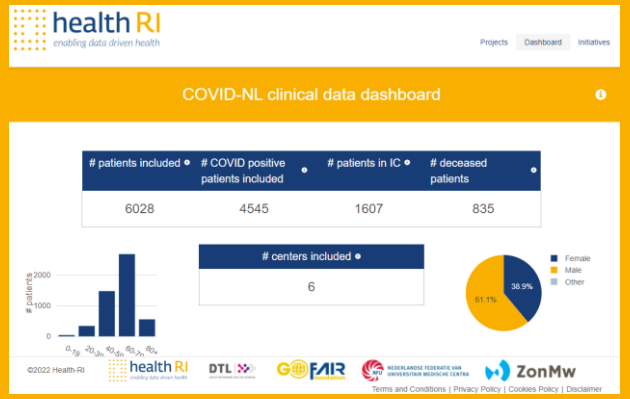


CC BY NC ND

ZonMw metadata portal
(machine readable metadata → easier → human readable portal)

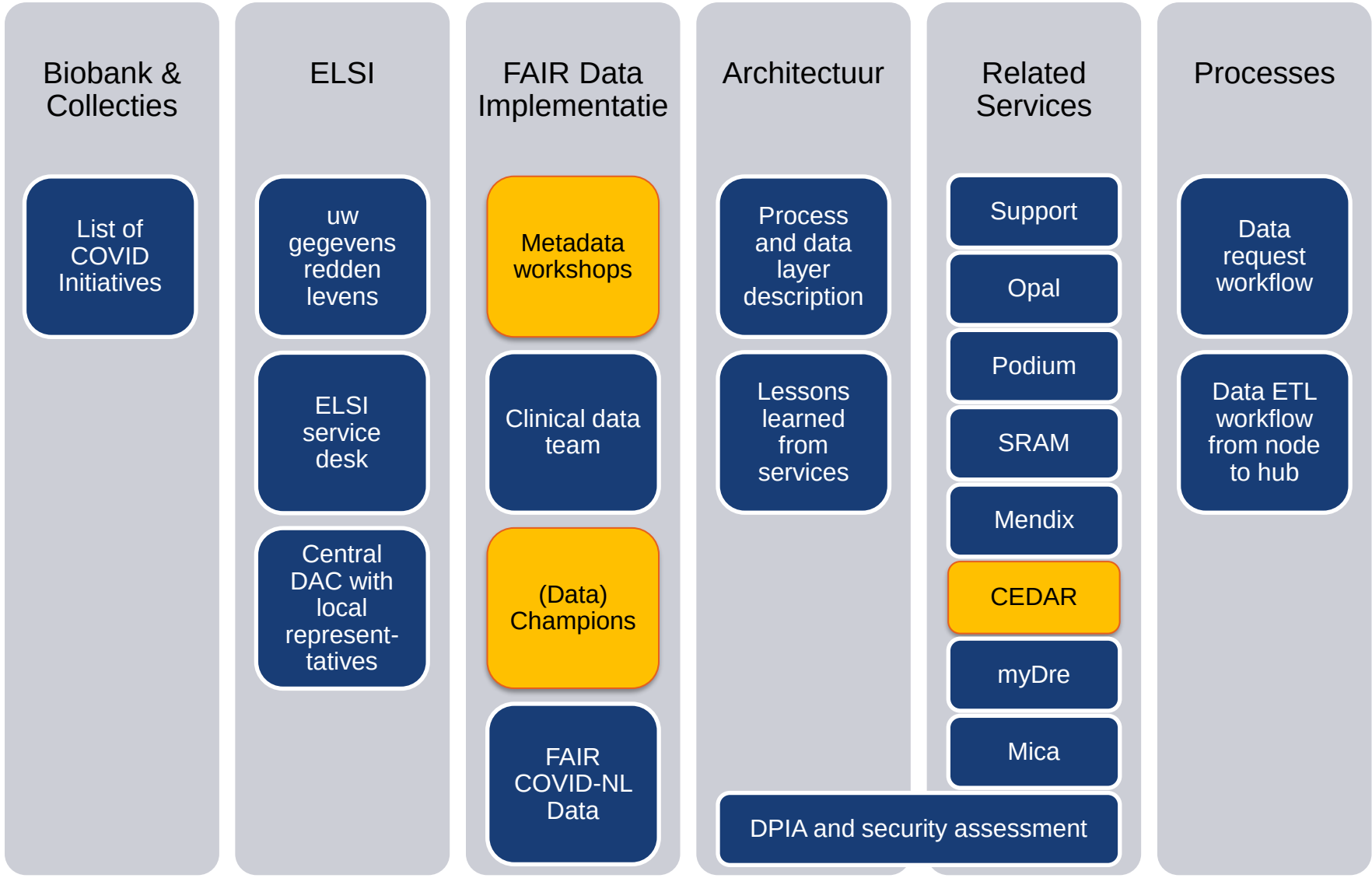


semantic faceted search using the provided metadata



How the standards were defined + the Metadata work

Covid Portal in Context - Achievements



A data steward perspective - How it all started: joint webinars

ZonMw

Why these webinars (incl. audiences)

- About the ZonMw requirements for Covid-19 grants
- For data stewards (experts) that (will) support ZonMw grants, and for those interested in further developments, as we expect Covid-19 research to be a catalysator for FAIR data in general sense
- Community of data stewards (experts) as a vital backbone for making appro developing new elements for FAIR data management in COVID-19 research
- Share your questions, experiences, challenges and good practices with us, s we can develop further services and translate gained knowledge into a share base/portal

Why?

Transparency and rapid communication about current research activities is needed to synergize the many efforts in the current hectic of COVID-19 research, to benefit from new developments, to find options to collaborate, and to prevent double work. ZonMw and NWO have therefore signed the Statement on sharing research findings and data in public health emergencies that are relevant to the novel coronavirus (COVID-19) outbreak. Moreover, ZonMw aligns actions with the GLOPID-R collaboration of research funders.

DTL | health RI

ZonMw

Activities

Based on the requirements & recommendations for the activities of open science and fair data. **Goals**: raise awareness, create/share good practices in research data management and data stewardship

Webinars for support and community building

>> Introduction to FAIR data management (May 7)
>> Procedures, requirements and community building (June 8)
>> Requirement 1: Alignment and reuse (today's session, June 18)
>> Requirement 2: Pre registration (June 22)
>> Requirement 3: FAIR data within COVID-19 research community (June 25)
>> Requirement 4: Budget for FAIR data and Open Access Publications (June 30)
>> Requirement 5: DMP and key items (July 6)

Community: invitation to join the DTL Data Stewards Interest Group (DSIG), which facilitates a Special Interest Group (SIG) for Covid-19 research. Planning: monthly virtual 'walk in sessions' to ask ZonMw Covid-19 grant related questions, exchange experiences and challenges and create/share good practices. More information and announcements via DSIG mailing list

DTL | health RI | GO FAIR

Workshops - Creating FAIR (meta)data together

3. Creating FAIR data, tailored to the COVID-19 research community

Here one finds the actions that enable researchers in ZonMw's COVID-19 projects to create FAIR data or - as a minimum - FAIR metadata. These actions were initiated by ZonMw in summer 2020 together with GO FAIR Foundation, DTL, and Health-RI. Below one finds the resources that resulted from ZonMw's actions, and those that are provided by other organisations, services and infrastructures.

Update on ZonMw's actions to optimize reuse of COVID-19 related data (summer 2021)

As a follow up on the early projects to FAIRify COVID-19 research data, GO FAIR Foundation, DTL, and Health-RI developed the three-point FAIRification Framework (3-PFF). In a series of M4M Workshops (M4M), metadata experts from (among others) GO FAIR and domain experts (COVID-19 researchers and their data stewards) participated to get training and to jointly create FAIR metadata: i.e., metadata that are both machine-actionable and follow domain-relevant community standards. You can read more about this approach in FAIRification. The concrete outputs of the workshops are domain specific metadata forms. Information about the COVID-19 projects derived from the metadata will be exposed on the Dutch Health Data portal at Health-RI.

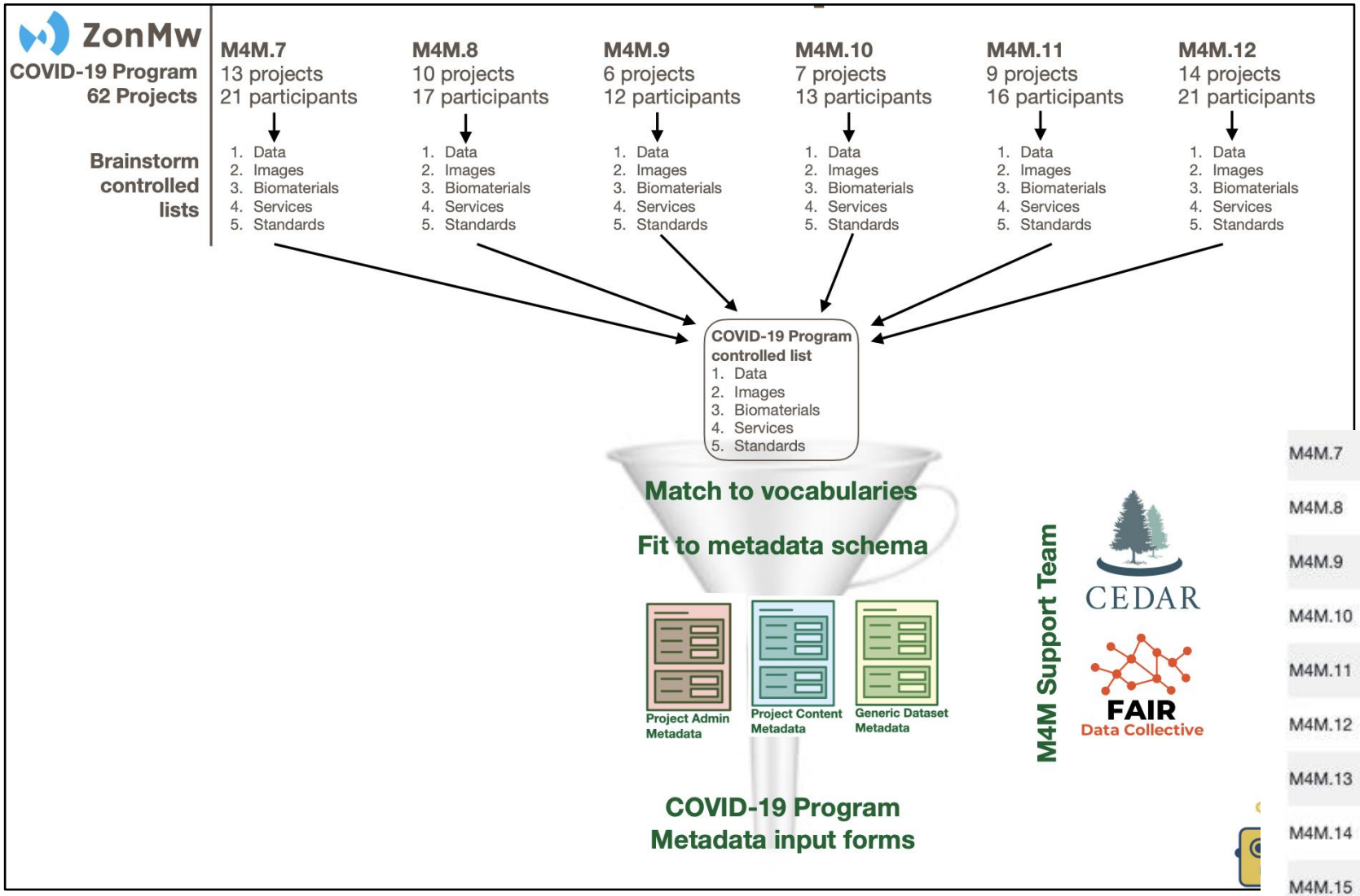
Find here the Guidelines with instructions and links to fill in the metadata forms
Find the metadata templates on the M4M resource page

The M4M Workshop concept

Making it easy for humans to make metadata for machines

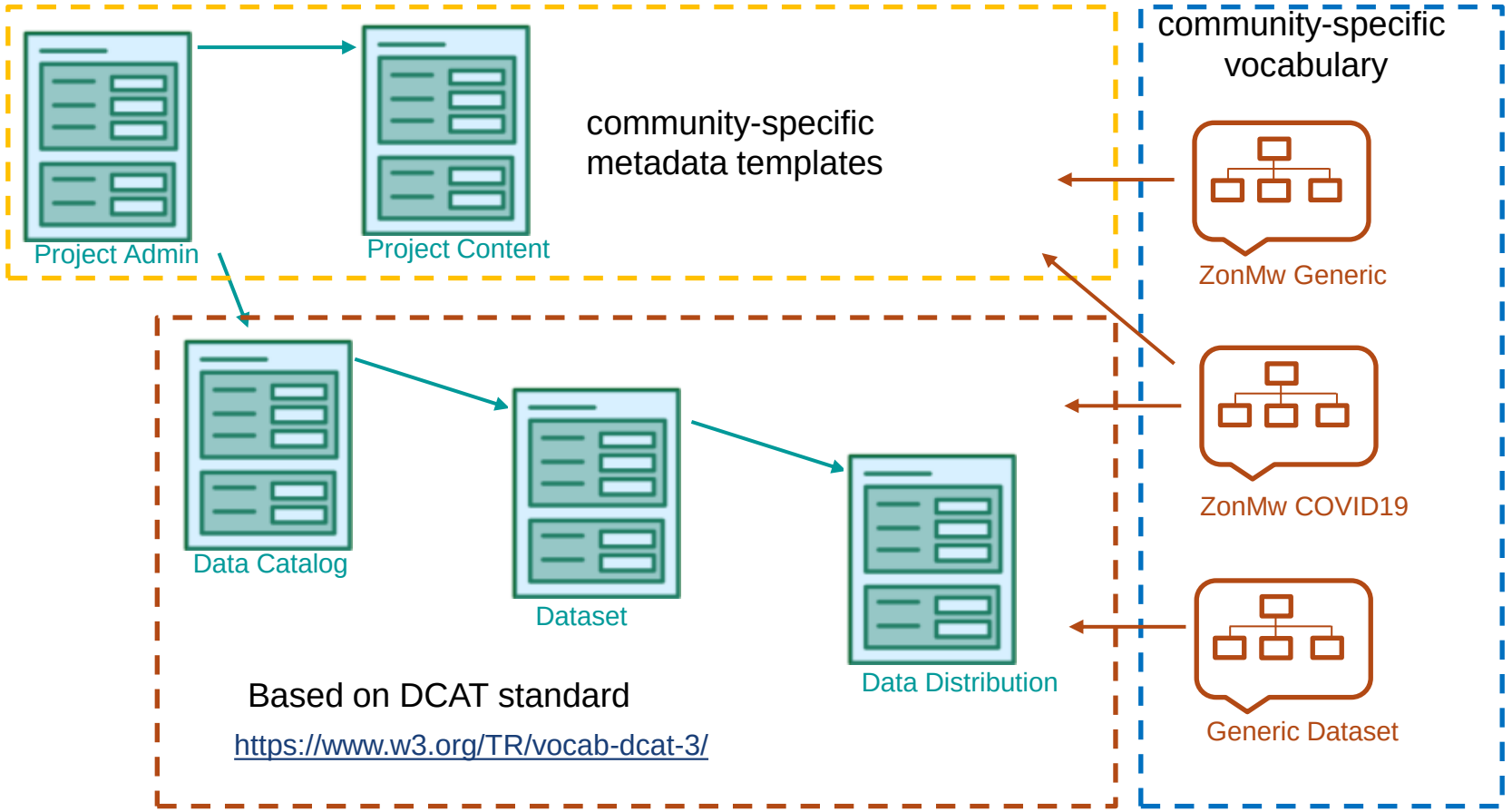
Metadata for Machines (M4M) workshops are agile, hackathon-style events that bring together domain experts (who are able and willing to represent a domain community) with FAIR metadata experts (data stewards) who guide a discussion leading to the metadata requirements that meet the FAIR data needs of that domain community. M4M Workshops are lightweight, fast-track (often 1-day) events where policy and domain experts can build new, or make informed choices regarding the reuse of already existing metadata schema. Although M4M Workshops can serve many purposes, they are usually intended to kick-start FAIRification efforts with minimally viable metadata components that are modular, and can be later extended as needed.

M4M in theory - Creating the CEDAR templates



M4M.7	November 2020	COVID-19 Program	Care (Treatment) / Prevention	ZonMw
M4M.8	November 2020	COVID-19 Program	Diagnostic / Testing – Recordings	ZonMw
M4M.9	November 2020	COVID-19 Program	Prognosis / Risk assessments	ZonMw
M4M.10	November 2020	COVID-19 Program	Virus / Immunology / Molecular – Recordings	ZonMw
M4M.11	November 2020	COVID-19 Program	Organisational / Process related – Recordigs	ZonMw
M4M.12	November 2020	COVID-19 Program	Socio-economic / Behavioral – Recordings	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

Metadata templates and vocabulary for projects and datasets



CC BY NC ND

M4M in practice - Creating the CEDAR templates

Format:

- We have an initial list of project assets (from vairous sources)
- Today we brainstorm collectively in short time-intervals using Padlet
- From Padlet, we export spreadsheets
- From spreadhsheets, we create semantic artefacts, CEDAR templates, and metadata forms

Chaos - Freedom without FAIR

:Padlet

Controlled lists

CSV

Controlled vocabularies

BioPortal

Metdata schema

Project Content Metadata

Order - COVID Program metadata instances

"Filling in" the form

M4M

M4M in practice - Creating the CEDAR templates

M4M in practice - Outcome: CEDAR templates

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]

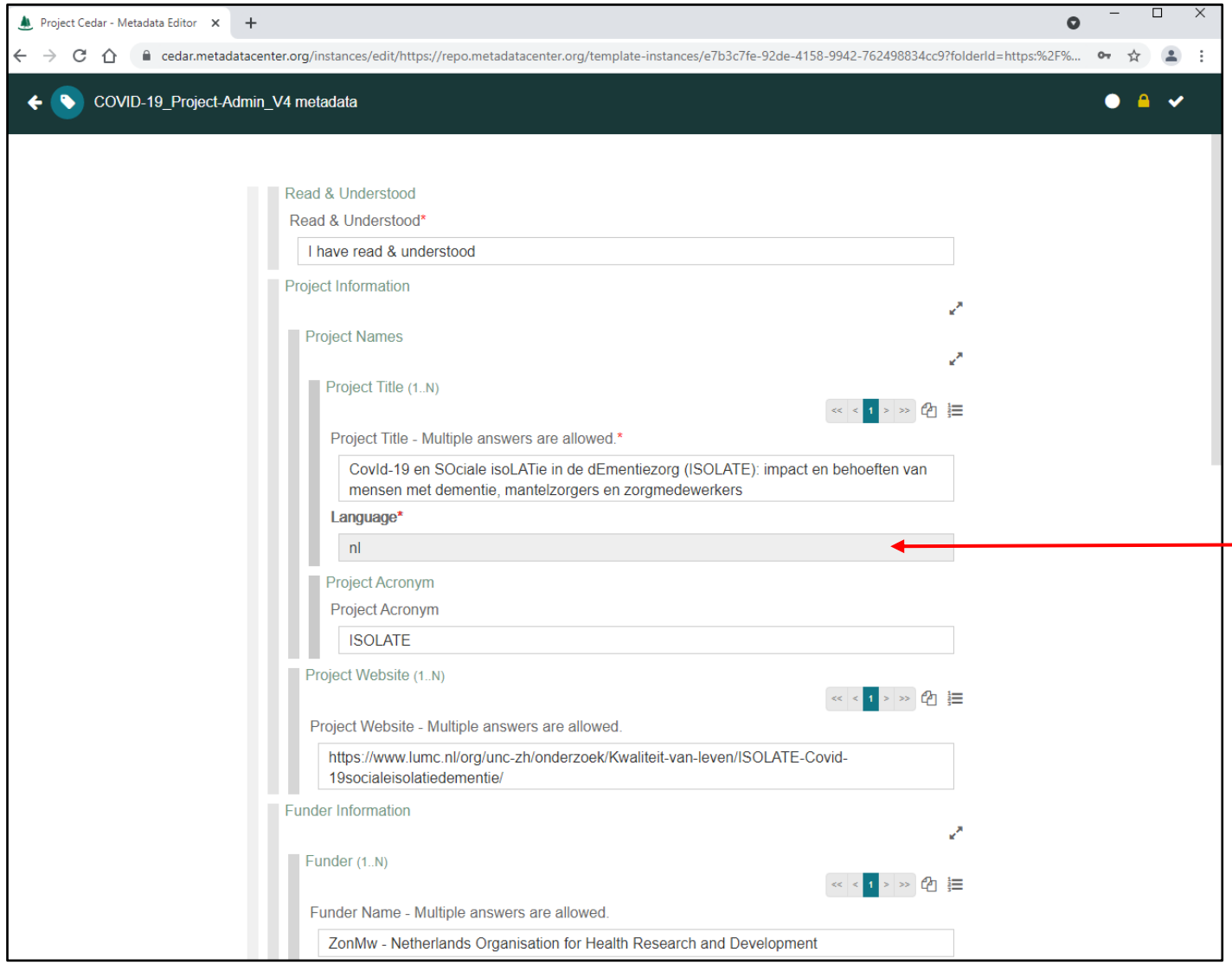
CEDAR forms in OpenView – easy way to view/share forms

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

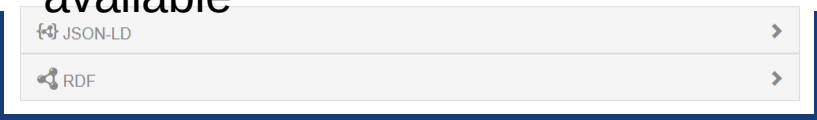
ZonMw Generic Terms
ZonMw COVID-19 Vocabulary

M4M in practice - Outcome: CEDAR templates



Controlled vocabularies from bioportal.bioontology.org

Machine readable metadata available



Supporting the community - Filling out the templates

Community Q&A sessions

When

Mondays 15:00-16:00

Join Zoom Meeting: https://zoom.us/j/97415399025

Meeting ID: 974 1539 9025

Details

This Q&A session is aimed at anyone from ZonMw COVID funded projects and will offer answers to questions on ZonMw COVID FAIR data services project:

- Experts will be available weekly in freely walk-in remote sessions
- Any question is welcome, please take advantage of these sessions
- If experts can't answer, your question will be forwarded

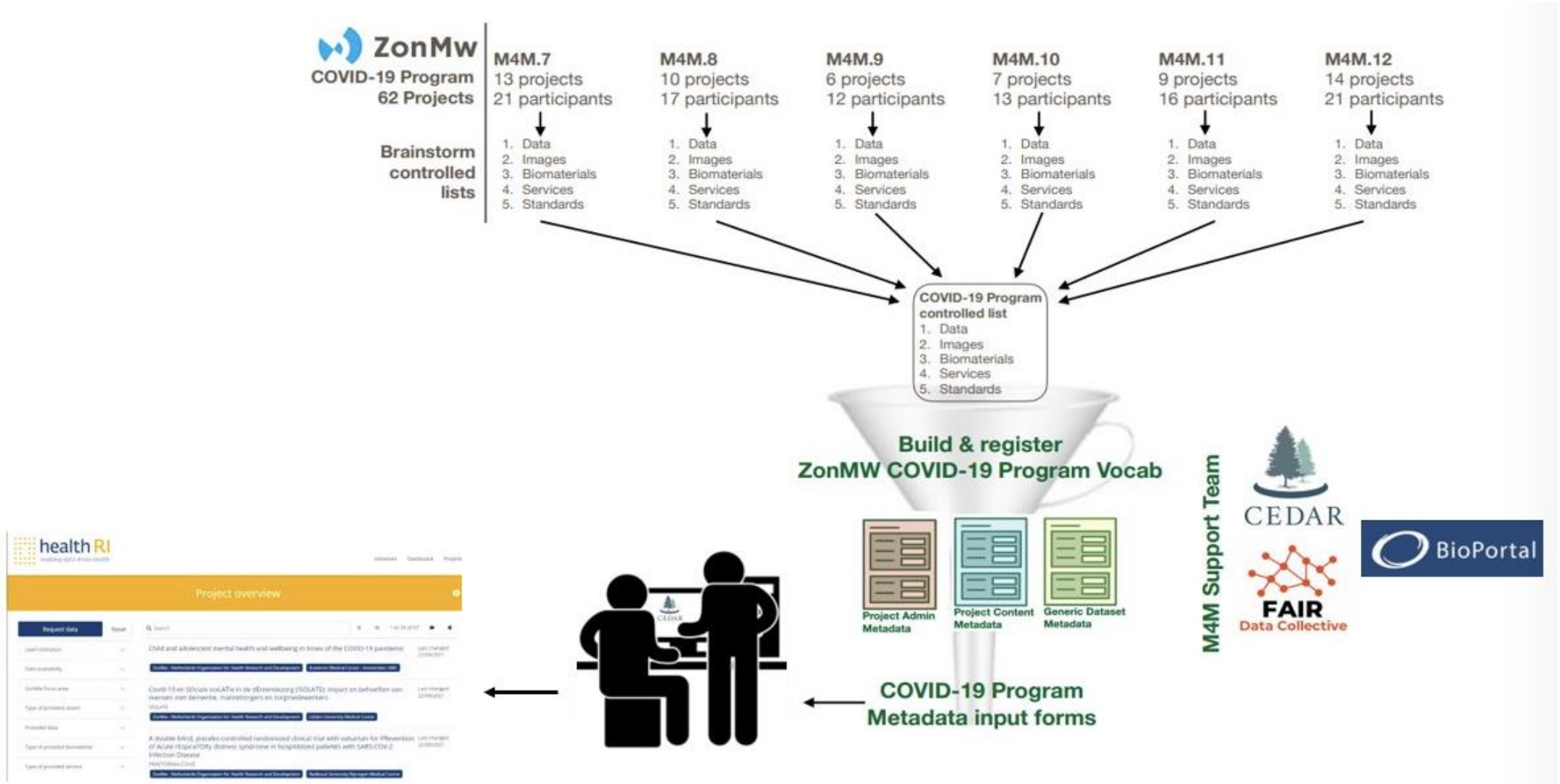
The Questions and Answers from these sessions are publicly available. You can also add your questions to the document so that they can be discussed in the next session.

← Home

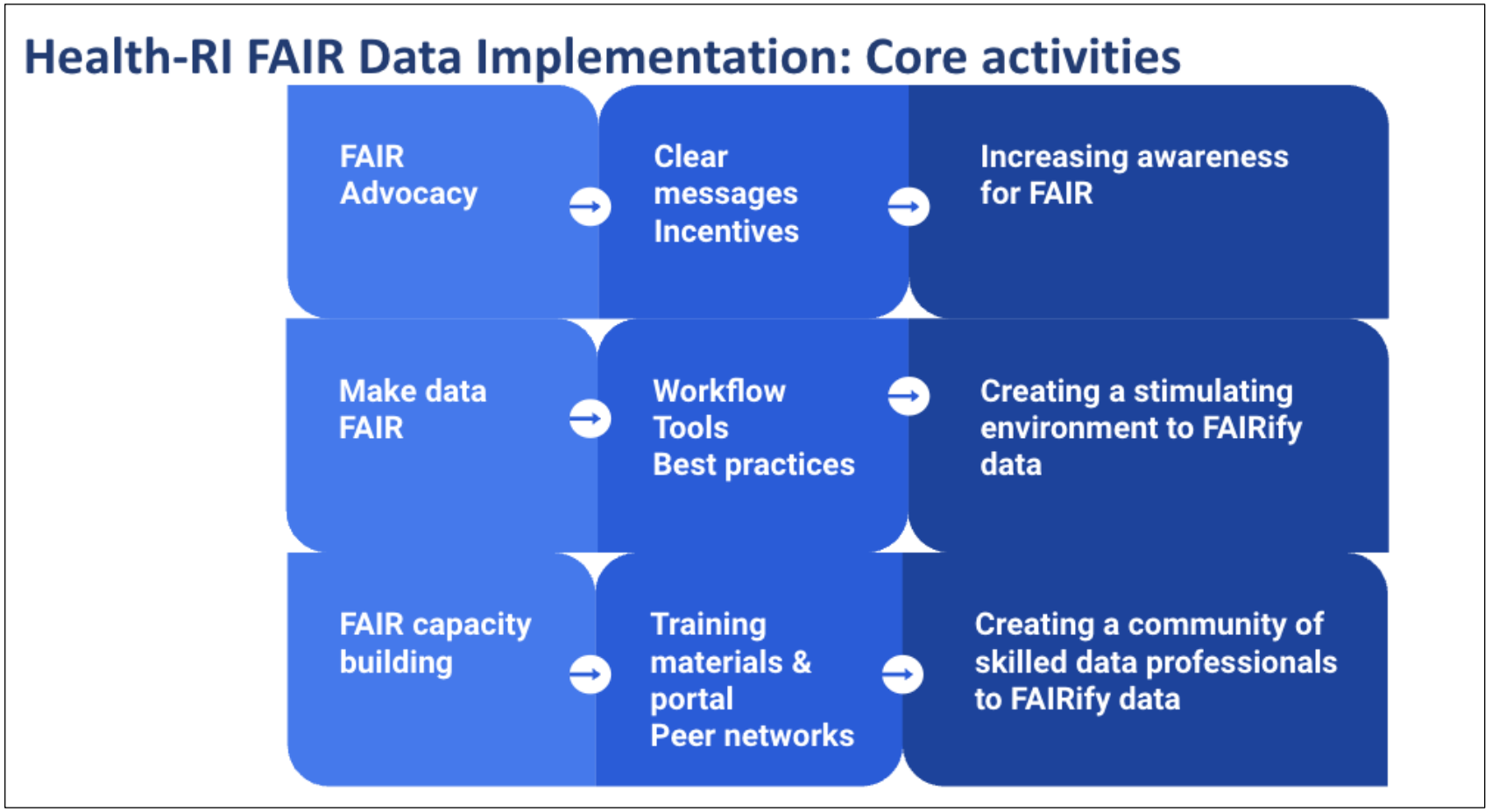
Health-RI Service Desk

You can contact the Health-RI service desk for all your inquiries regarding Health-RI and associated initiatives and services.

From Community to Data Portal

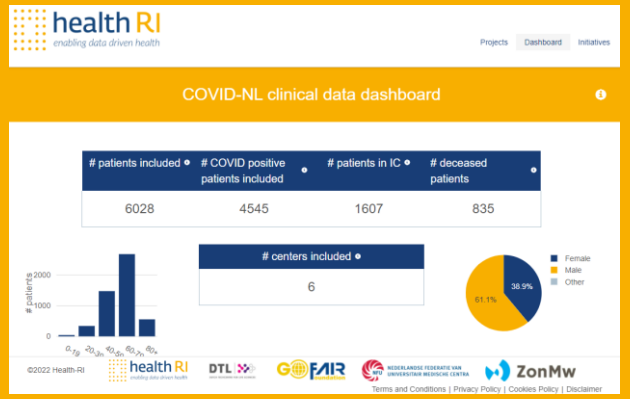


Making Supporting the Community scalable



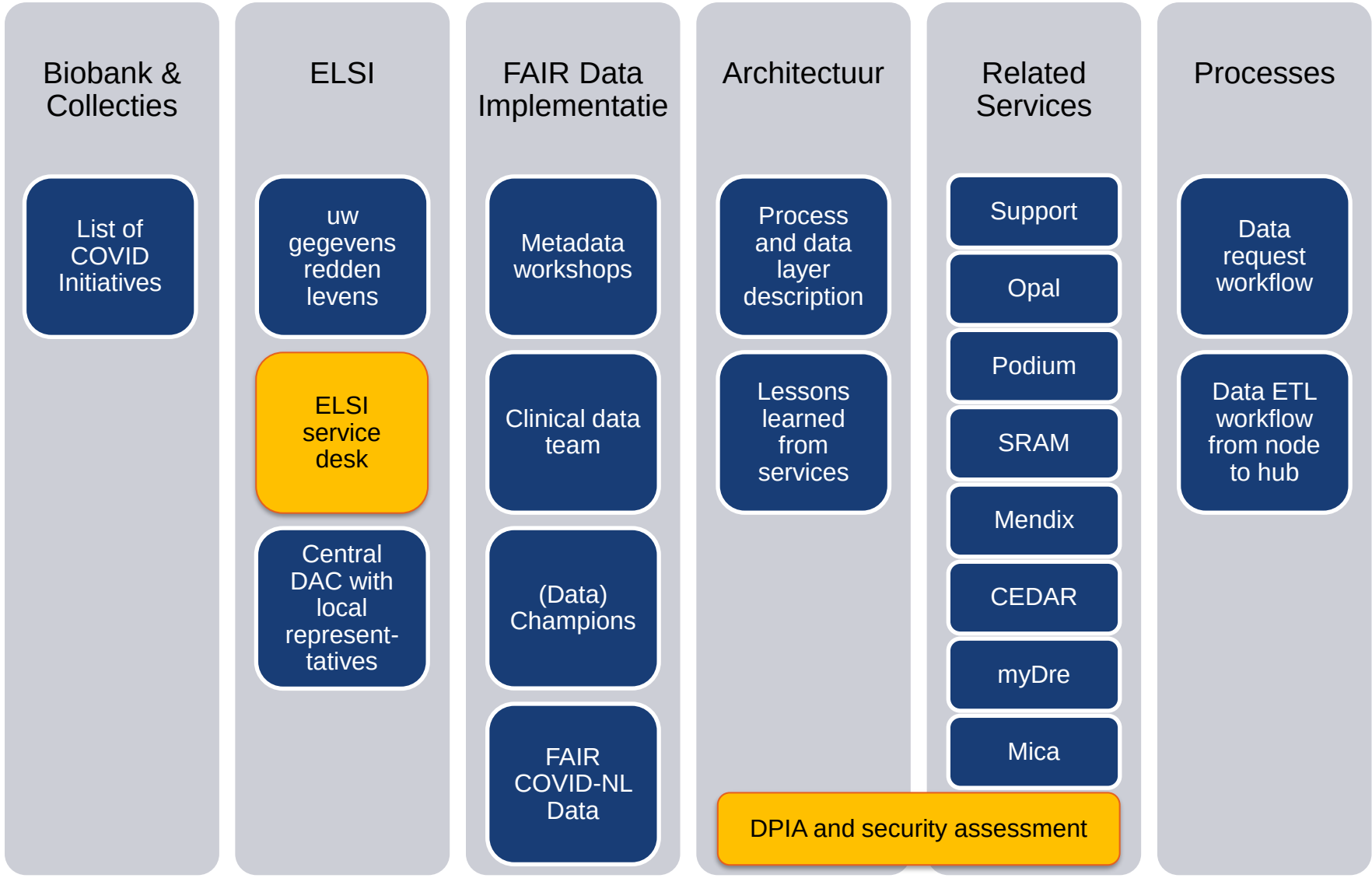
Data reuse obstacles removal traject

Collaborative commitment to remove obstacles that block the (re)use of health data for research and innovation



What needed to be done in terms of legal and societal groundwork:

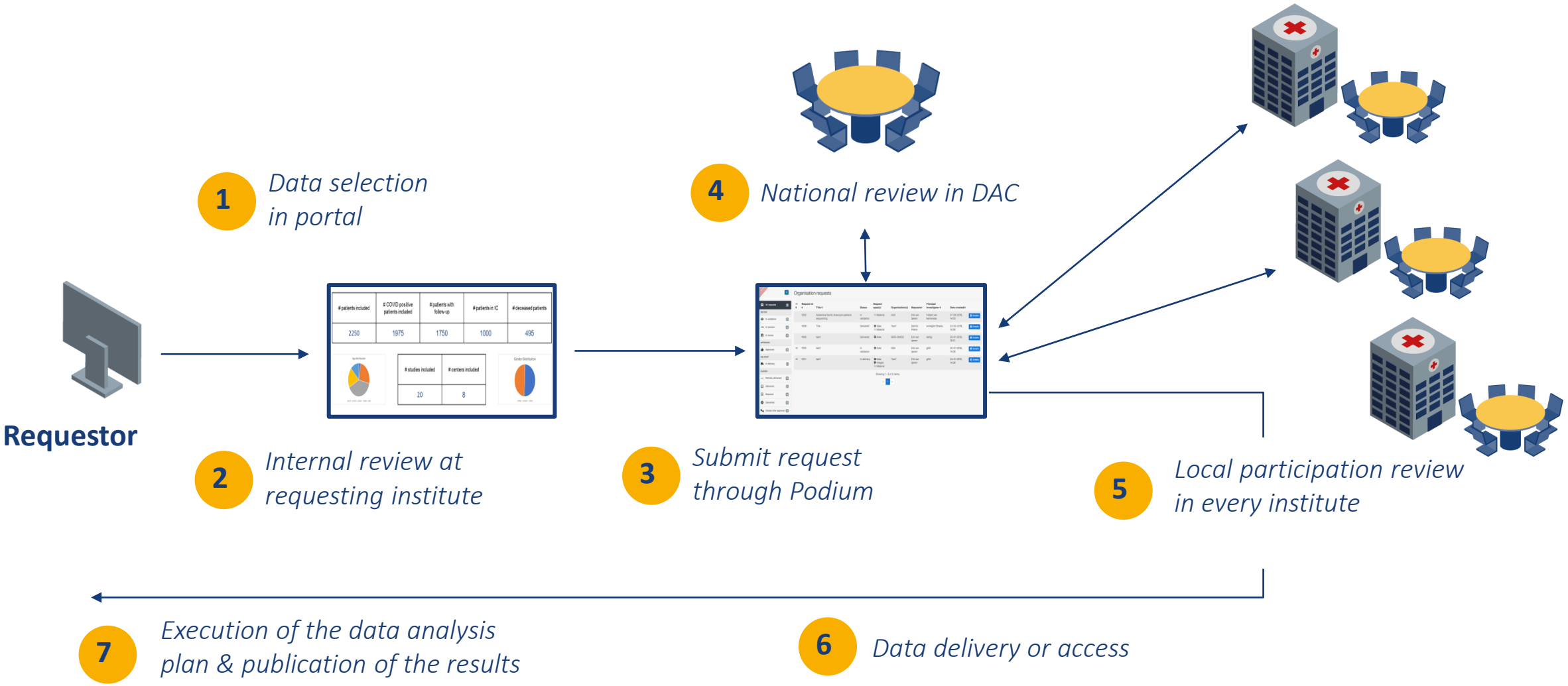
Covid Portal in Context - Achievements



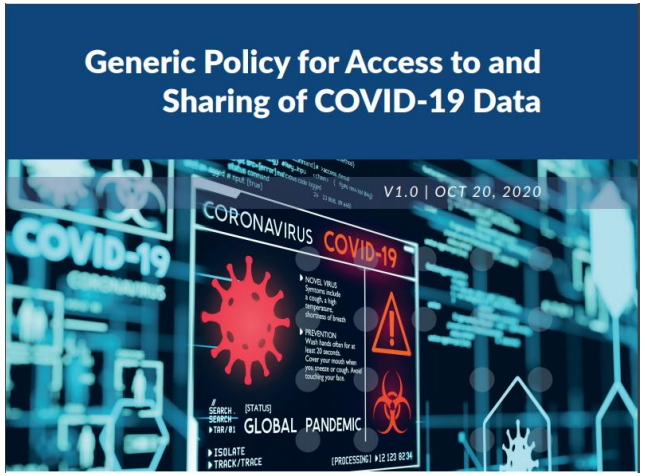
COVID-NL data governance – the challenges

- COVID-NL is not only technology: data governance & regulatory aspects
- Already some procedures in place in most institutes
- Often based on a somewhat different interpretation of the legislation
  - In particular for the GDPR legal basis
- Responsibilities for secondary use of data are not always clearly defined
- Key stakeholders (legal support (often outsourced), DPOs, data managers, security officers, etc.) had to be identified while working on the procedures

Data governance workflow – the result

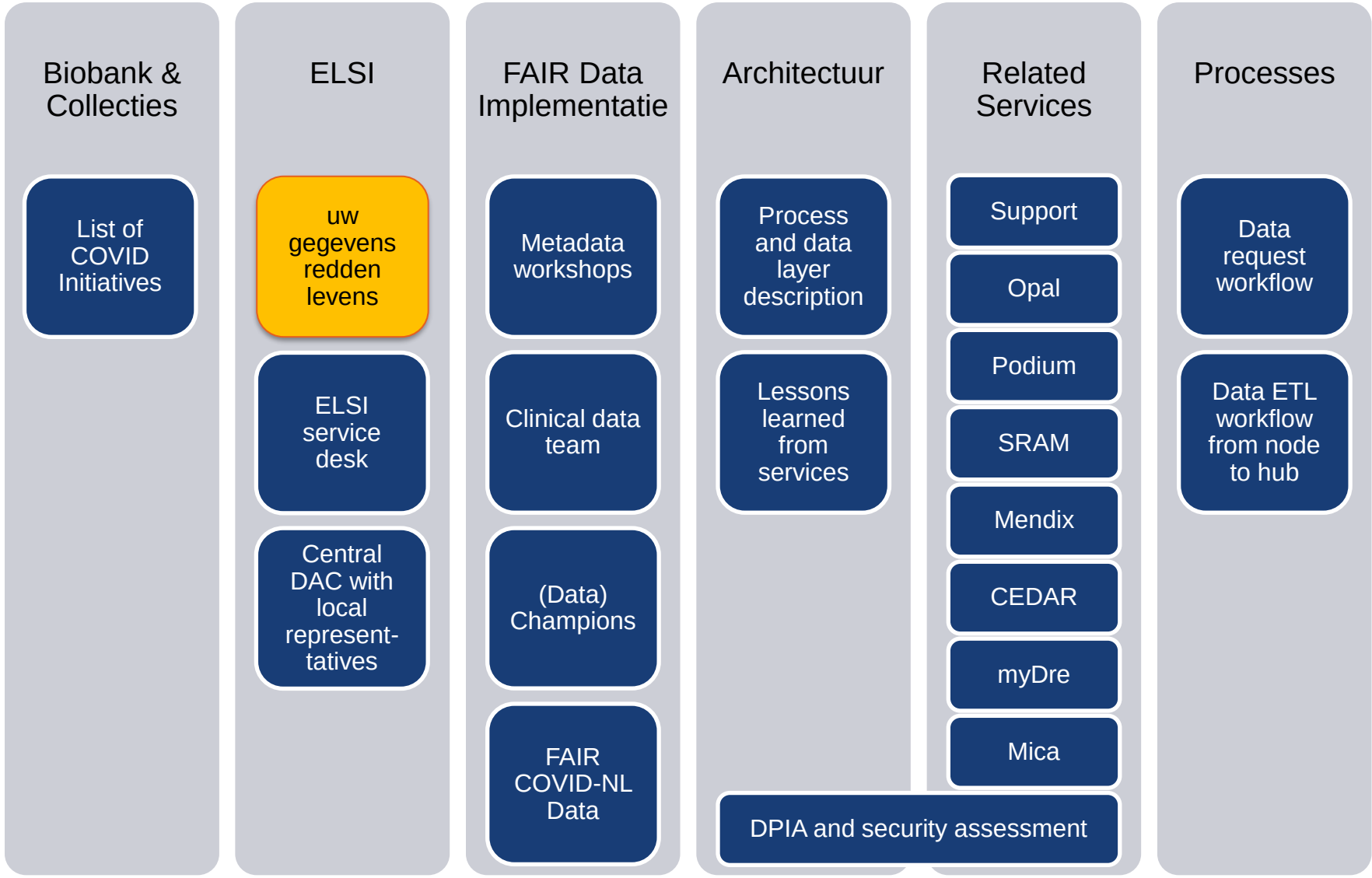


What is available?



- Documents publicly accessible via ELSI Servicedesk:
  - Feb 2022: Contract agreed between partners (still to be signed off)
  - Oct 2020 (update Nov 2021): Policy for Access to and Sharing of COVID-19 Data
  - June 2020: Guide data governance procedures
  - June 2020: Standard Data Transfer Agreement

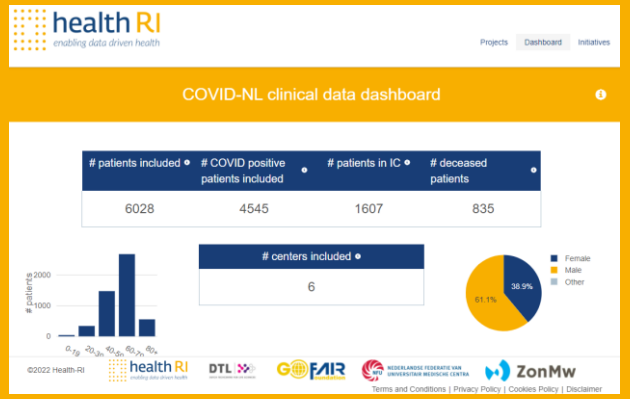
Covid Portal in Context - Achievements



UW GEGEVENS REDDEN LEVENS

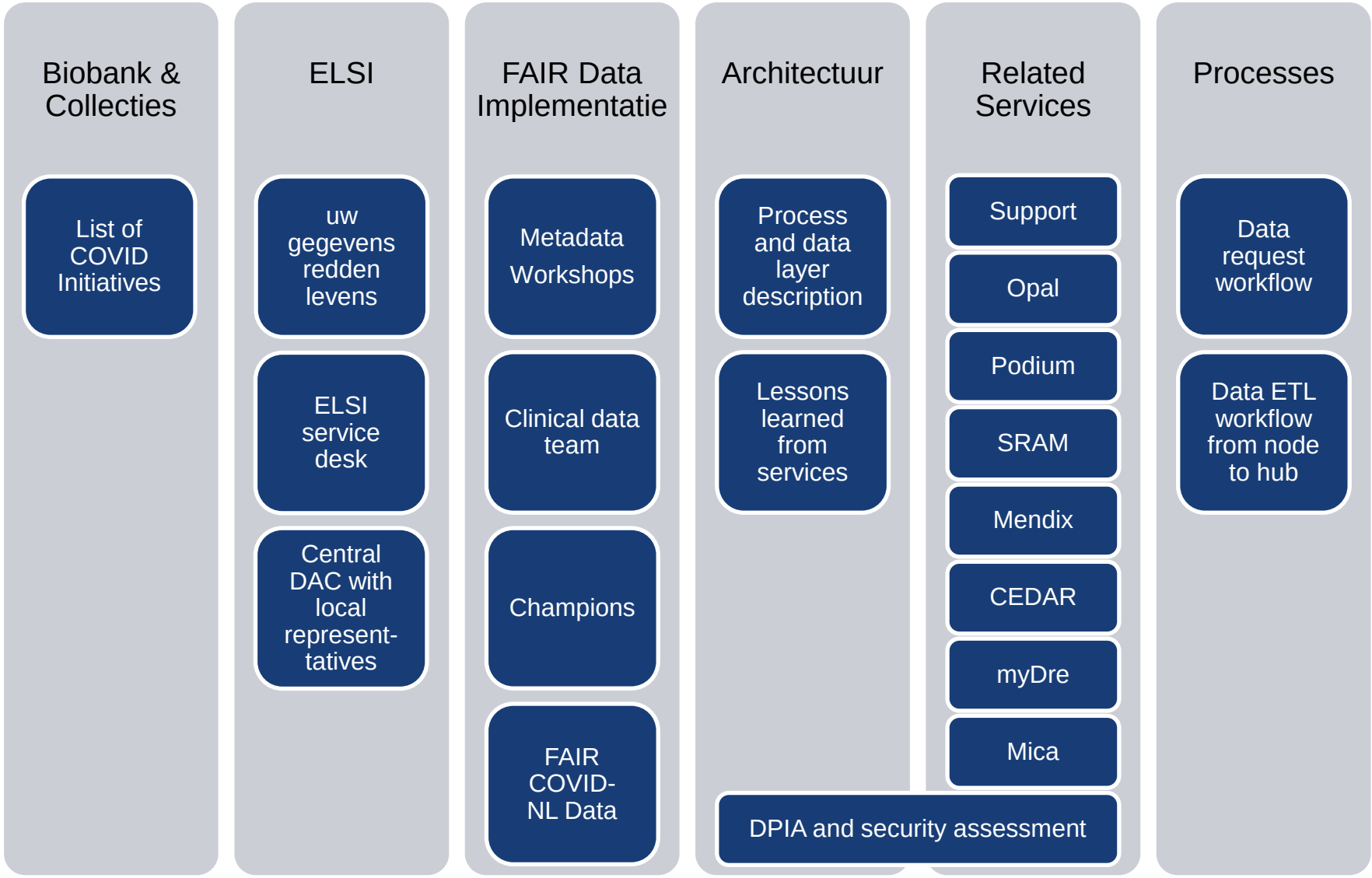
Deliverables van project

- Toestemmingsverklaring goedgekeurd door VWS voor nationaal gebruik door ziekenhuizen bij opgenomen Covid patiënten
- Informatie tekst goedgekeurd door VWS voor nationaal gebruik door ziekenhuizen bij opgenomen Covid patiënten
- Animatie, talking heads film en beeldmateriaal
- Website www.uwgegevensreddenlevens.nl
- Eindverslag lessons learned

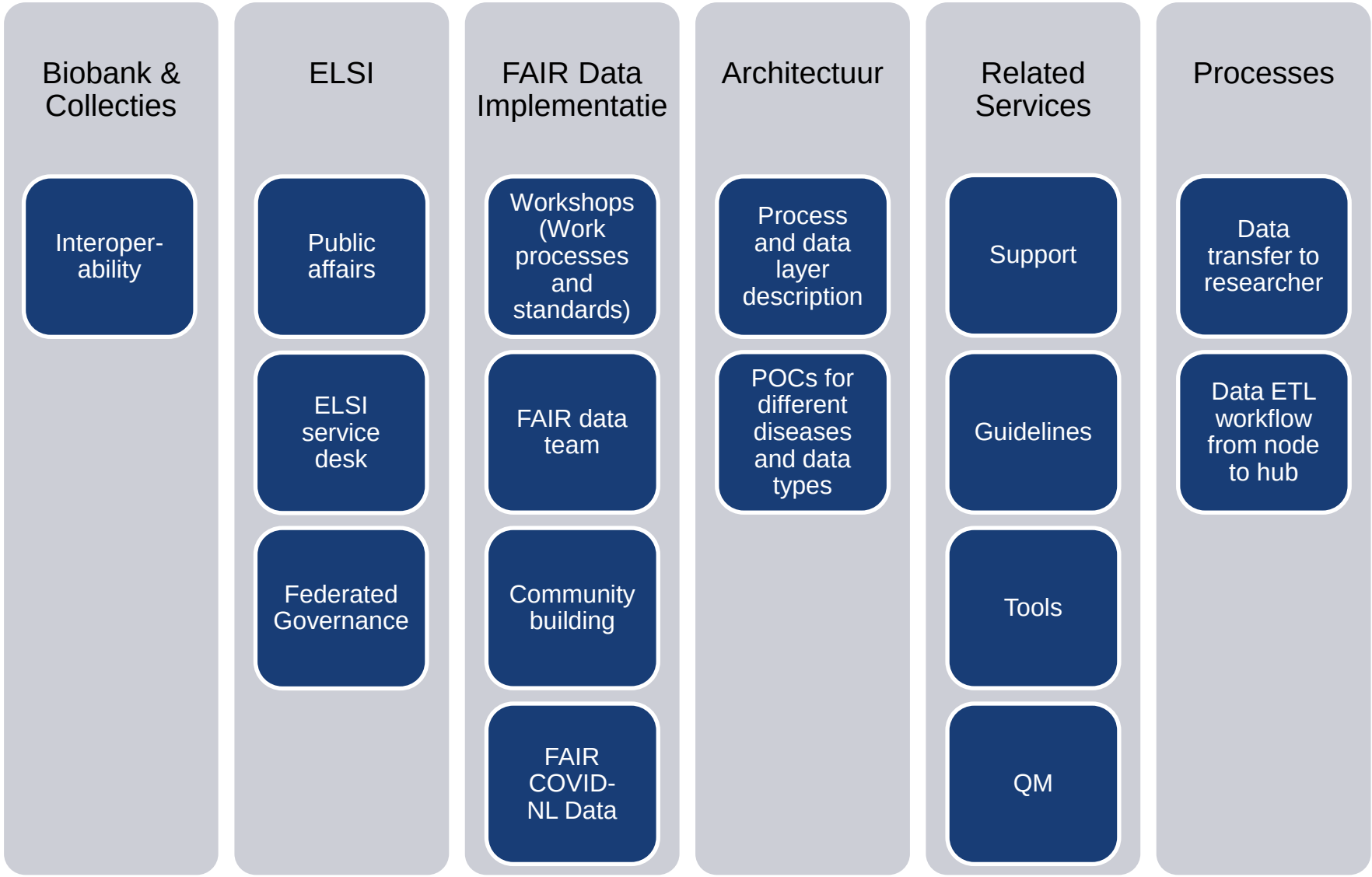


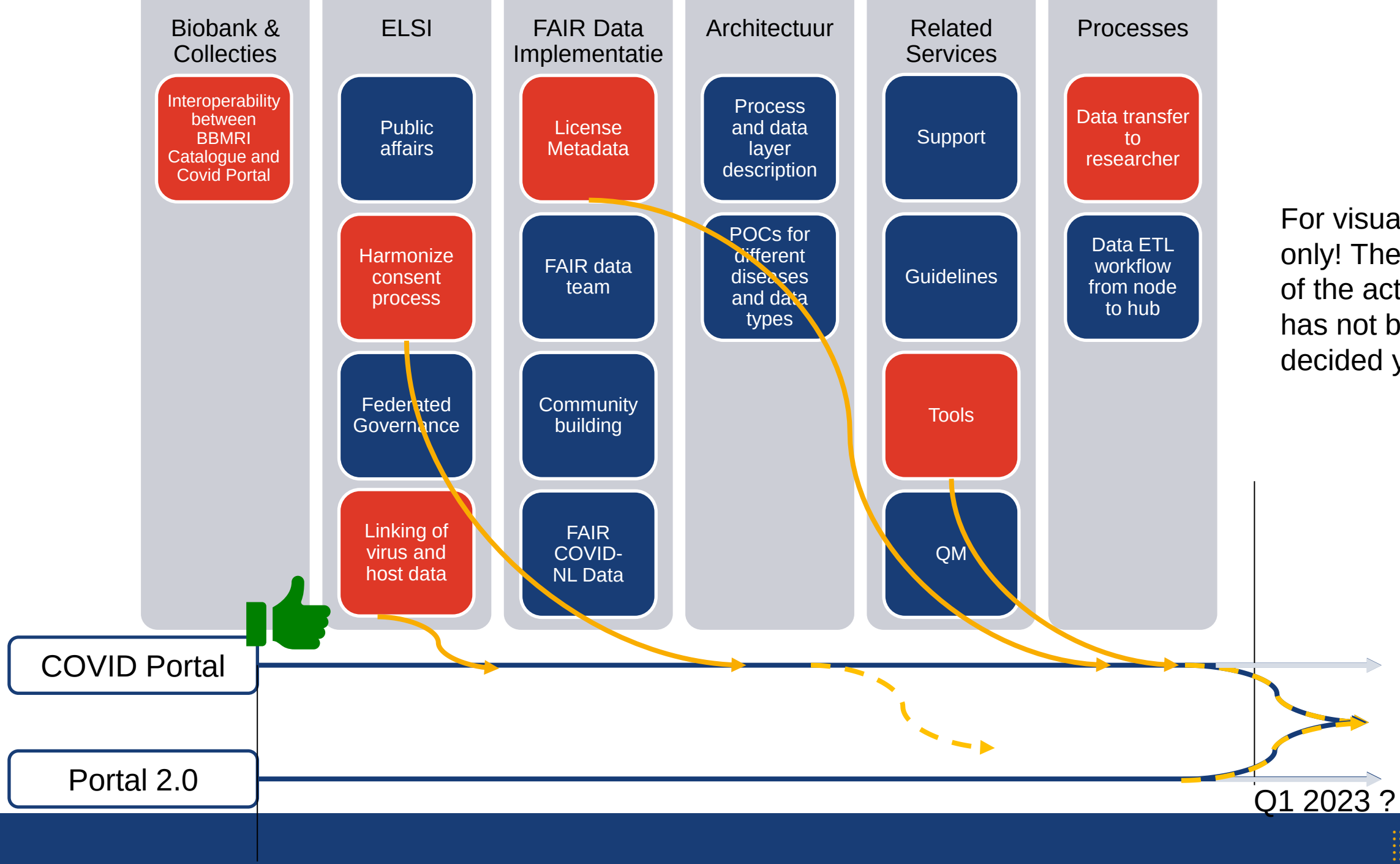
What needs to be done next:

Covid Portal in Context – Achievements of Pilot



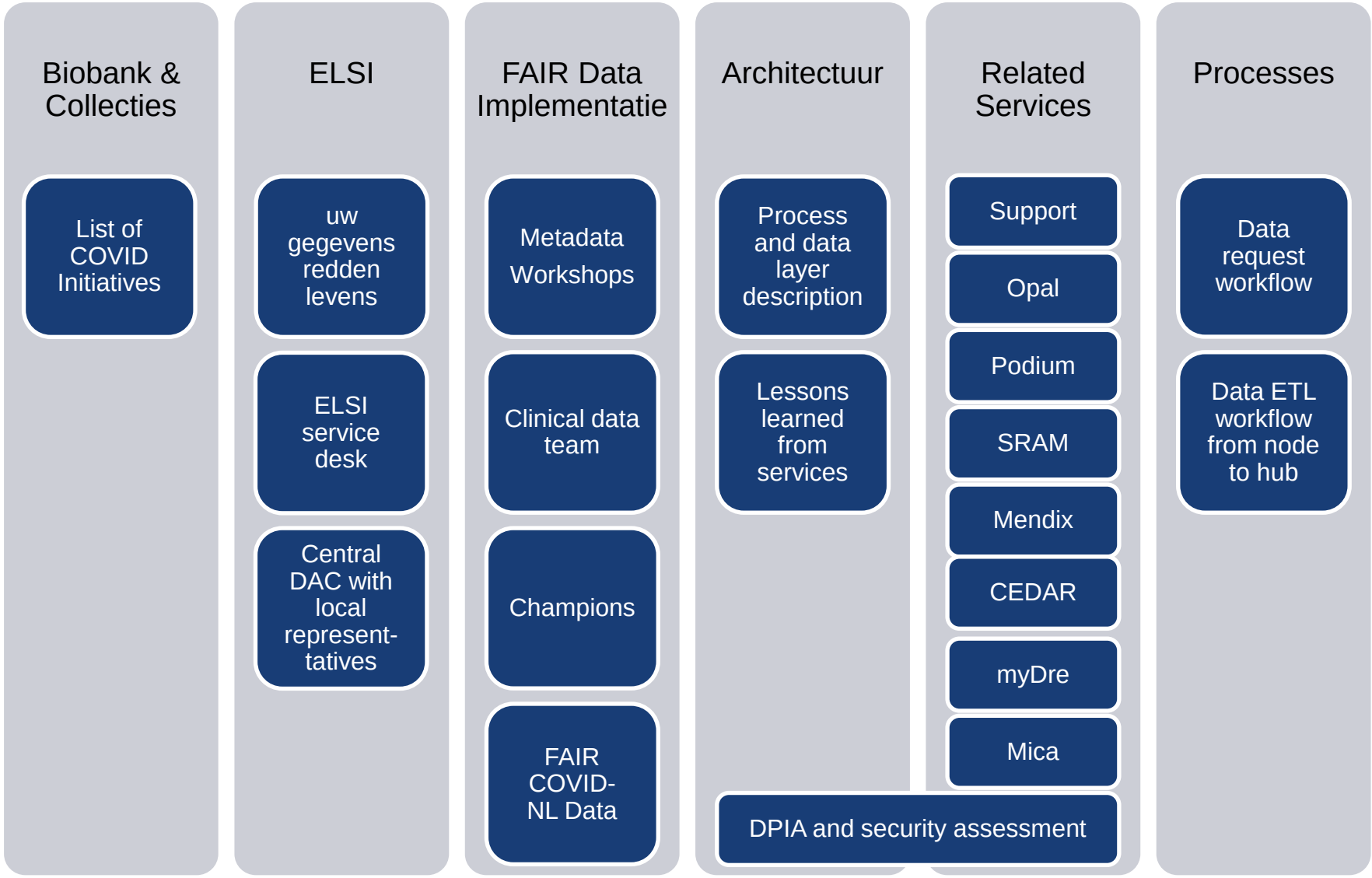
Covid Portal in Context – Relation to Health-RI future activities





For visualization only! The order of the activities has not be decided yet.

Covid Portal in Context – Achievements of Pilot



Metadata Workshops

Data ETL workflow from node to hub

FAIR COVID-NL Data

Clinical data team

Support

Process and data layer description

Champions

DPIA and security assessment

Data source

uw gegevens redden levens

COVID PORTAL

Opal

Podium

SRAM

Mendix

CEDAR

Mica

myDre

myDre

Lessons learned from services

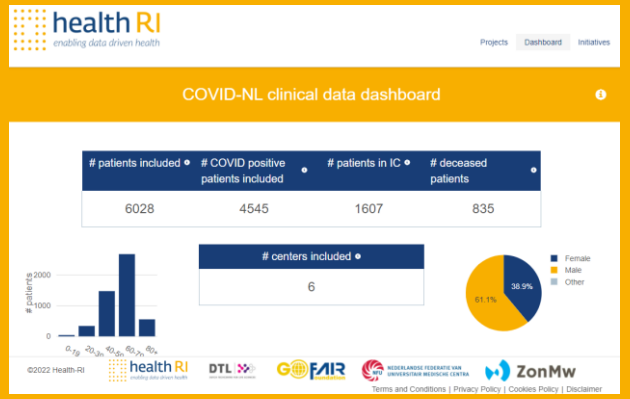
ELSI service desk

List of COVID Initiatives

Central DAC with local represent-tatives

Data request workflow

Secondary Data Use


health RI
enabling data driven health
Projects
Dashboard
Initiatives
COVID-NL clinical data dashboard
patients included
COVID positive patients included
patients in IC
deceased patients
6028
4545
1607
835
centers included
6
patients
2000
1000
0
0-19
20-39
40-59
60-79
80+
Female
Male
Other
38.9%
61.1%
©2022 Health-RI
health RI
DTL
GO FAIR
ZonMw
Terms and Conditions | Privacy Policy | Cookies Policy | Disclaimer

Thank you!