

UNIVERSITY  
OF TWENTE.



## A SEMANTIC DATA MODEL FOR THE WHO COVID-19 CASE REPORT FORM

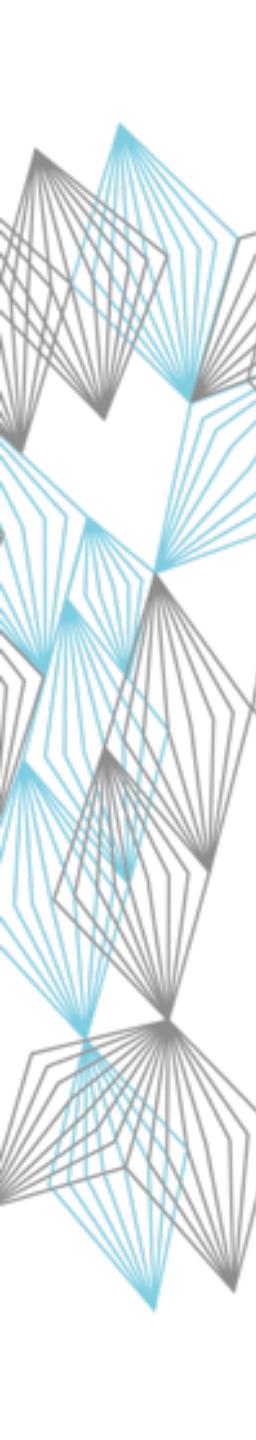
Luiz Bonino, Rajaram Kaliyaperumal, Marc Nieuwland, Jeroen Beliën, Bruna dos Santos Vieira and Lana Meiqari.



19 OCTOBER 2020



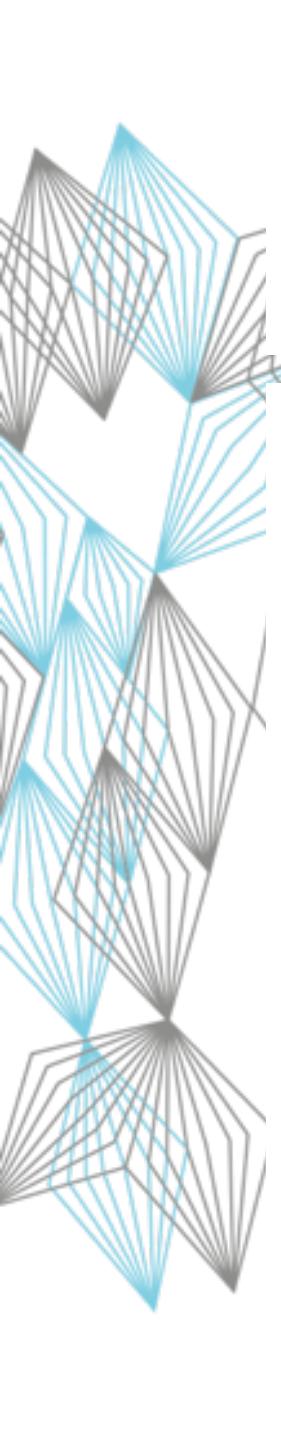
Leiden University  
Medical Center



# VIRUS OUTBREAK DATA NETWORK (VODAN)

A GO FAIR implementation network (<https://www.go-fair.org/implementation-networks/overview/vodan/>):

- Support, create, develop, demonstrate and deliver elements for improving aspects related to reuse of virus outbreaks data;



# VODAN IN A BOX

- Goals:
  - Create an easy-to-deploy solution for capturing, storing and exposing data and metadata related to virus outbreaks.
  - Support for local deployment of the complete solution → data and metadata stay on wherever their owners/creators decide;
- Composed of:
  - Configurable and deployable data entry tool → CRF Wizard;
  - Configurable and deployable FAIR metadata publication tool → FAIR Data Point;
  - Semantic data model of the WHO COVID-19 CRF for reporting COVID-19 cases.

# VODAN IN A BOX



COVID-19 case

# VODAN IN A BOX

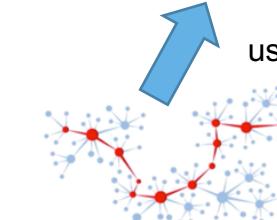


COVID-19 case



WHO CRF  
Data

enters data



WHO CRF  
Semantic Data Model

uses model

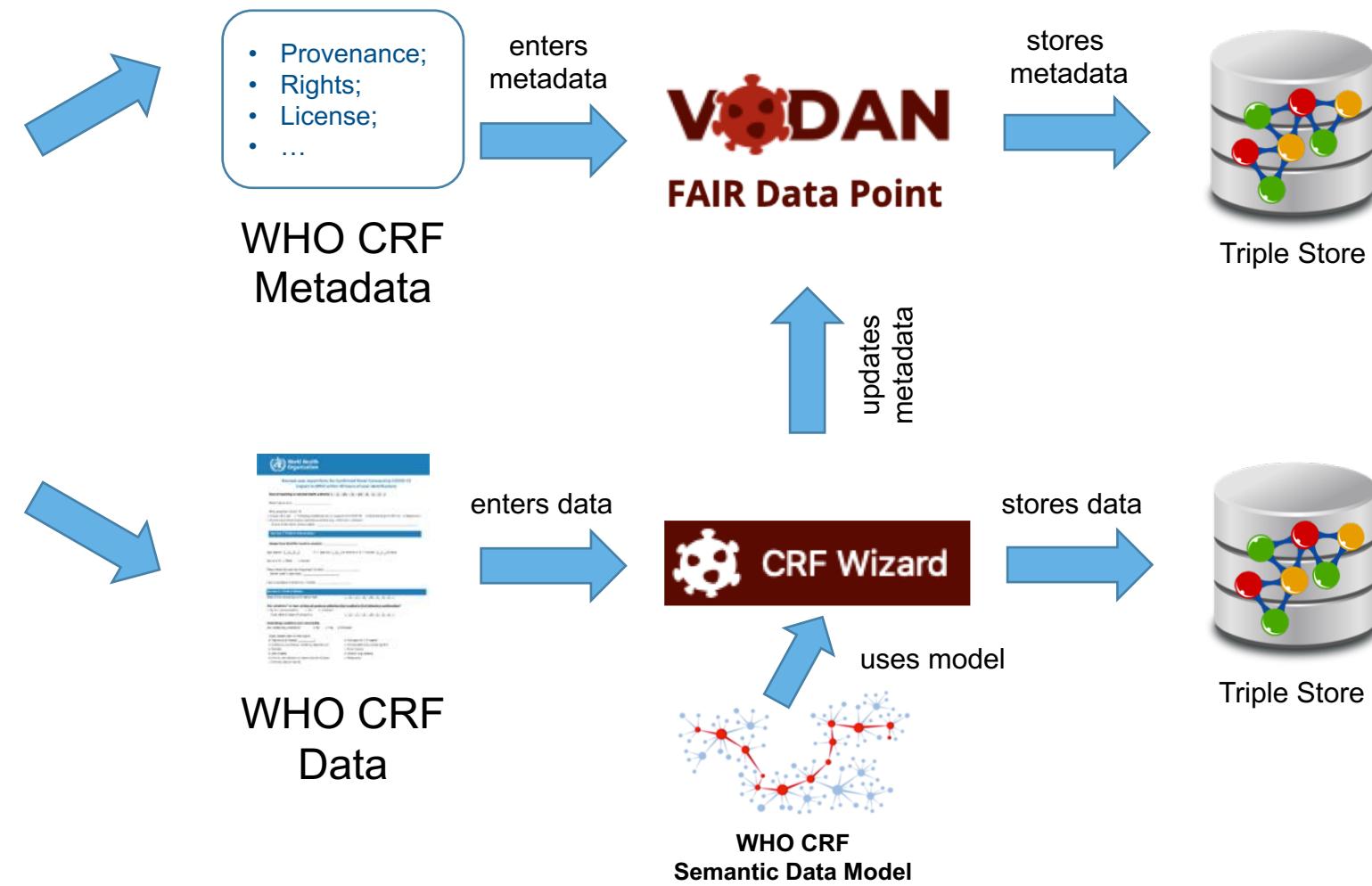
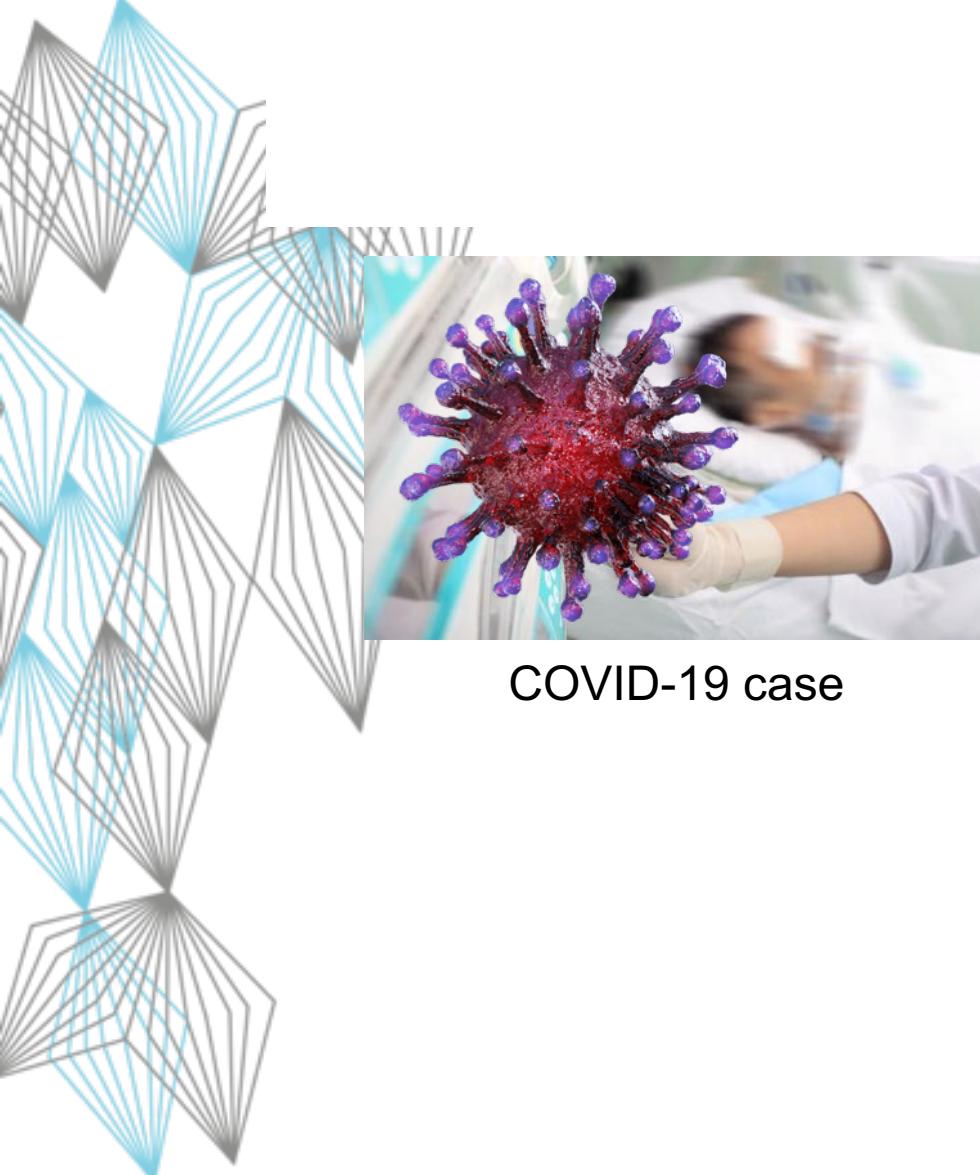


stores data



Triple Store

# VODAN IN A BOX



# WORLD HEALTH ORGANIZATION ELECTRONIC RECORD FORM FOR COVID-19



**Global COVID-19 Clinical Platform**

**NOVEL CORONAVIRUS (COVID-19) - RAPID VERSION**

## DESIGN OF THIS CASE RECORD FORM (CRF)

This CRF has 3 modules:

**Module 1** to be completed on the first day of admission to the health centre.

**Module 2** to be completed on first day of admission to ICU or high dependency unit. Module 2 should also be completed daily for as many days as resources allow. Continue to follow-up patients who transfer between wards.

**Module 3** to be completed at discharge or death.

[https://www.who.int/docs/default-source/coronaviruse/who-ncov-crf.pdf?sfvrsn=84766e69\\_2](https://www.who.int/docs/default-source/coronaviruse/who-ncov-crf.pdf?sfvrsn=84766e69_2)

# WORLD HEALTH ORGANIZATION ELECTRONIC RECORD FORM FOR COVID-19



PARTICIPANT ID 1\_1\_2\_1\_3\_1\_4\_1\_5\_1-1\_6\_1\_7\_1\_8\_1\_9\_1

## MODULE1: complete on admission/enrolment

Site name ACME hospital

Country Wakanda

Date of enrolment [D][D][M][M][Y][Y]

### CLINICAL INCLUSION CRITERIA

Proven or suspected infection with pathogen of Public Health Interest  Yes  No

One or more | A history of self-reported feverishness or measured fever of  $\geq 38^{\circ}\text{C}$   Yes  No

of these | Cough  Yes  No

during this | Dyspnoea (shortness of breath) OR Tachypnoea\*  Yes  No

illness | Clinical suspicion of ARI despite not meeting criteria above  Yes  No

\* respiratory rate  $\geq 50$  breaths/min for <1 year;  $\geq 40$  for 1-4 years;  $\geq 30$  for 5-12 years;  $\geq 20$  for  $\geq 13$  years

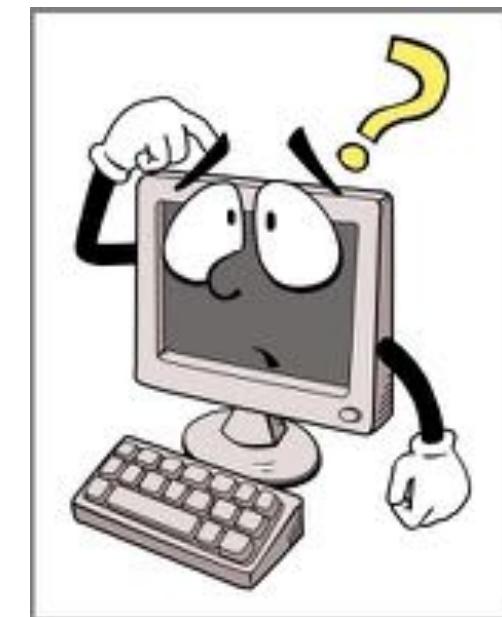
### DEMOGRAPHICS

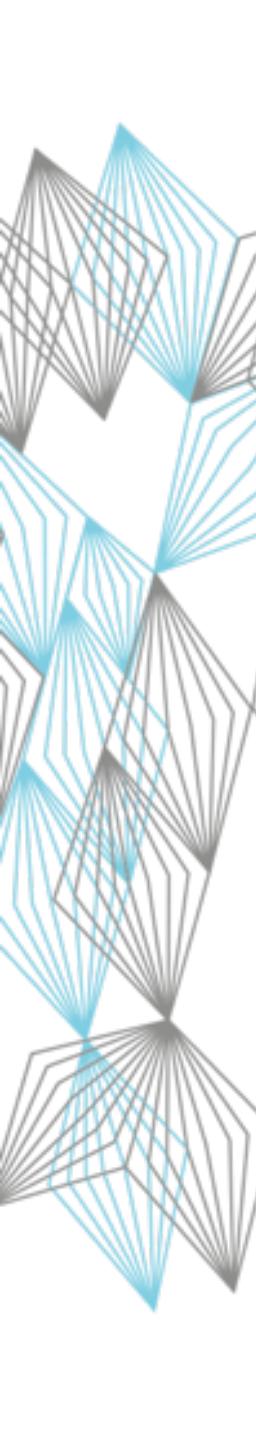
Sex at Birth  Male  Female  Not specified Date of birth [D][M][Y][Y][D][M][Y][Y]

If date of birth is unknown, record: Age [D][M][Y]years OR [D][M]months

Healthcare Worker?  Yes  No  Unknown Laboratory Worker?  Yes  No  Unknown

Pregnant?  Yes  No  Unknown  N/A If yes: Gestational weeks assessment [D][M] weeks

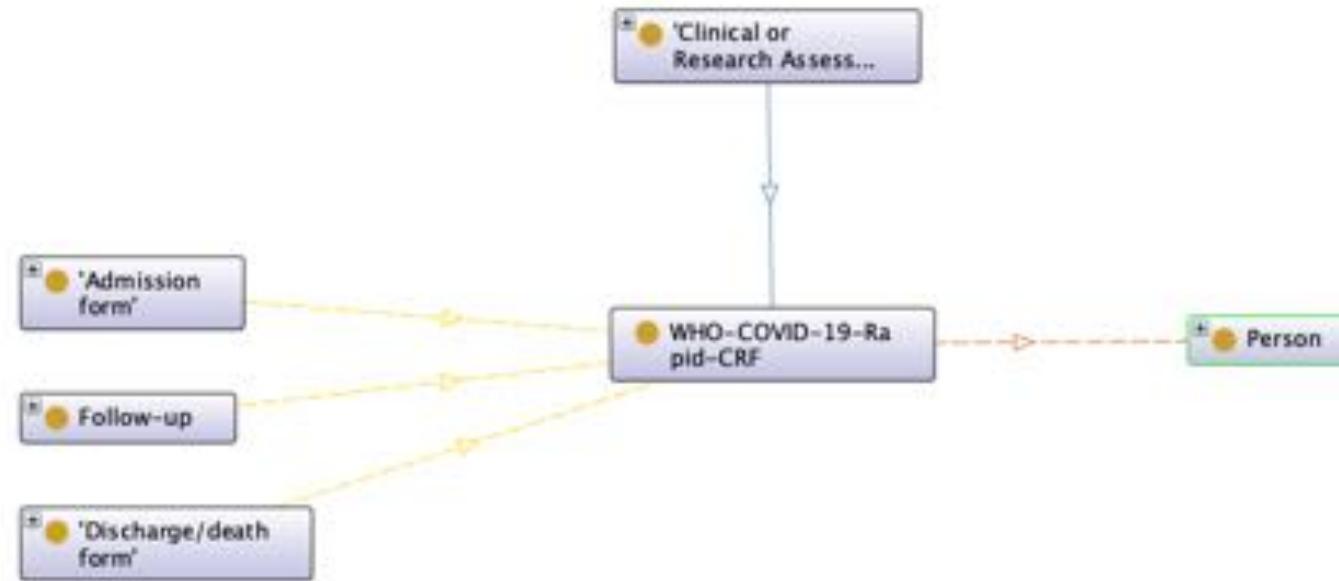




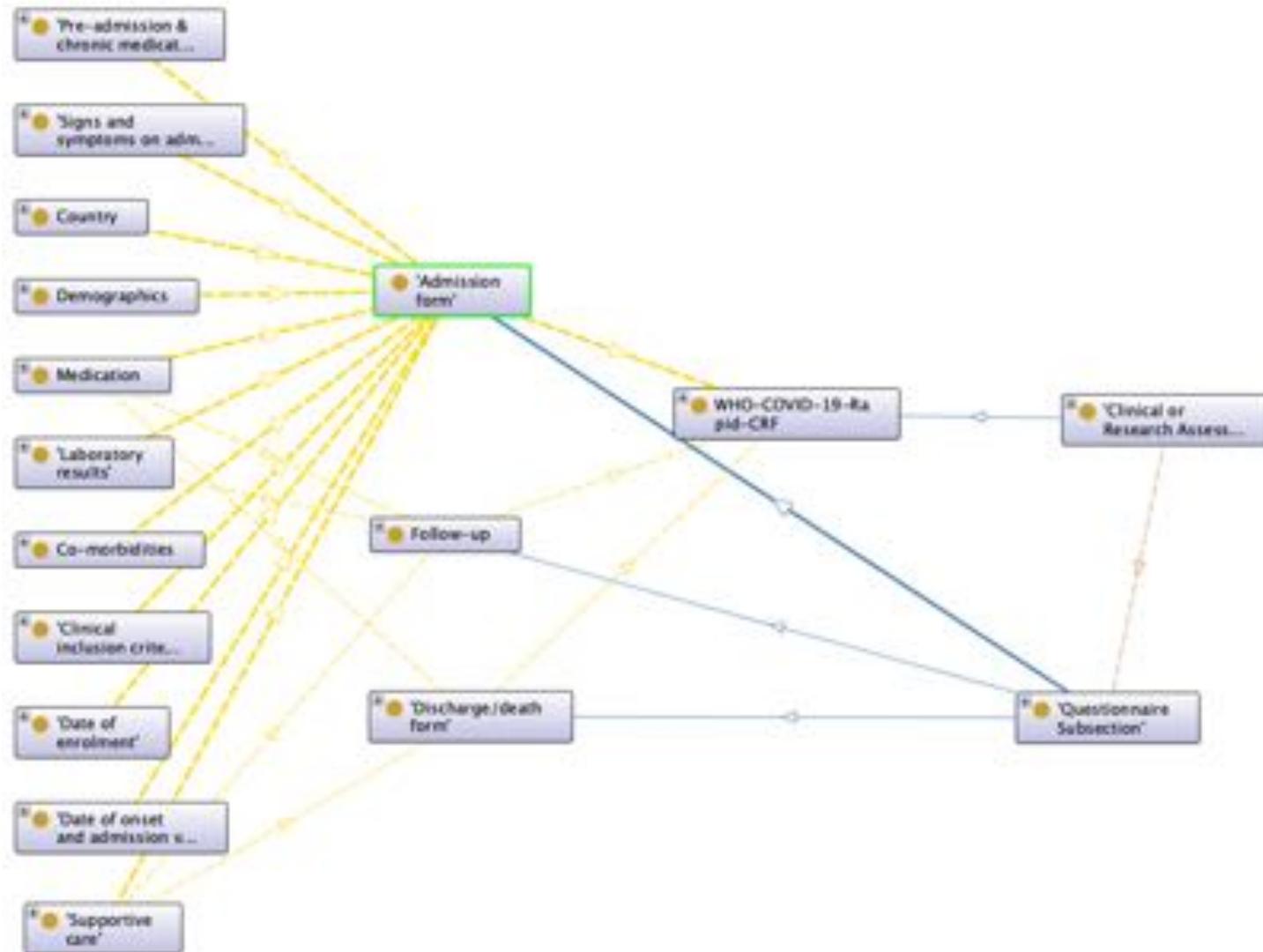
# WHO'S COVID-19 CRF SEMANTIC DATA MODEL

- Model to support semantic annotation of data created based on WHO's COVID-19 CRF rapid version;
- Focus on machine-actionable semantic references for the form's questions and answers, not biological phenomena;
- Open access: CC BY 4.0
- Persistent identifier:  
<http://purl.org/vodan/whocovid19crfsemdatamodel/>;
- BioPortal:  
<http://bioportal.bioontology.org/ontologies/COVIDCRFRAPID>
- Documentation: <https://vodan-ontology.github.io>

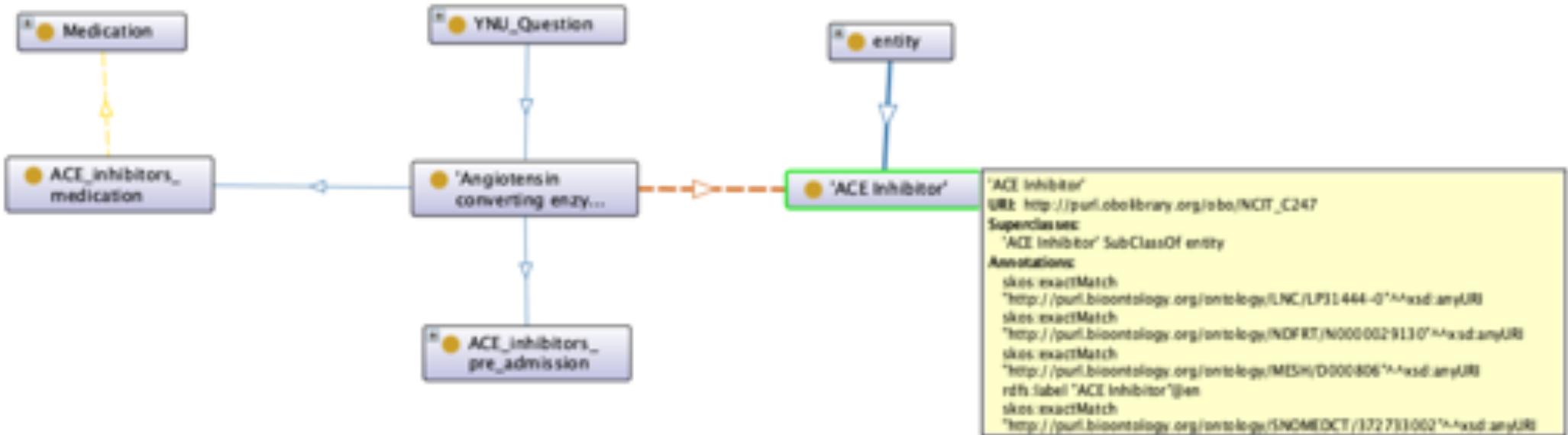
# WHO'S COVID-19 CRF SEMANTIC DATA MODEL



# WHO'S COVID-19 CRF SEMANTIC DATA MODEL



# WHO'S COVID-19 CRF SEMANTIC DATA MODEL



# WHO'S COVID-19 CRF SEMANTIC DATA MODEL

**MEDICATION** Is the patient CURRENTLY receiving any of the following?

Oral/orogastric fluids?  Yes  No  Unknown Intravenous fluids?  Yes  No  Unknown

Antiviral?  Yes  No  Unknown If yes:  Ribavirin  Lopinavir/Ritonavir  Neuraminidase inhibitor  
 Interferon alpha  Interferon beta  Other, specify: \_\_\_\_\_

Corticosteroid?  Yes  No  Unknown If yes, route:  Oral  Intravenous  Inhaled  
If yes, please provide agent and maximum daily dose: \_\_\_\_\_

Antibiotic?  Yes  No  Unknown Antifungal agent?  Yes  No  Unknown

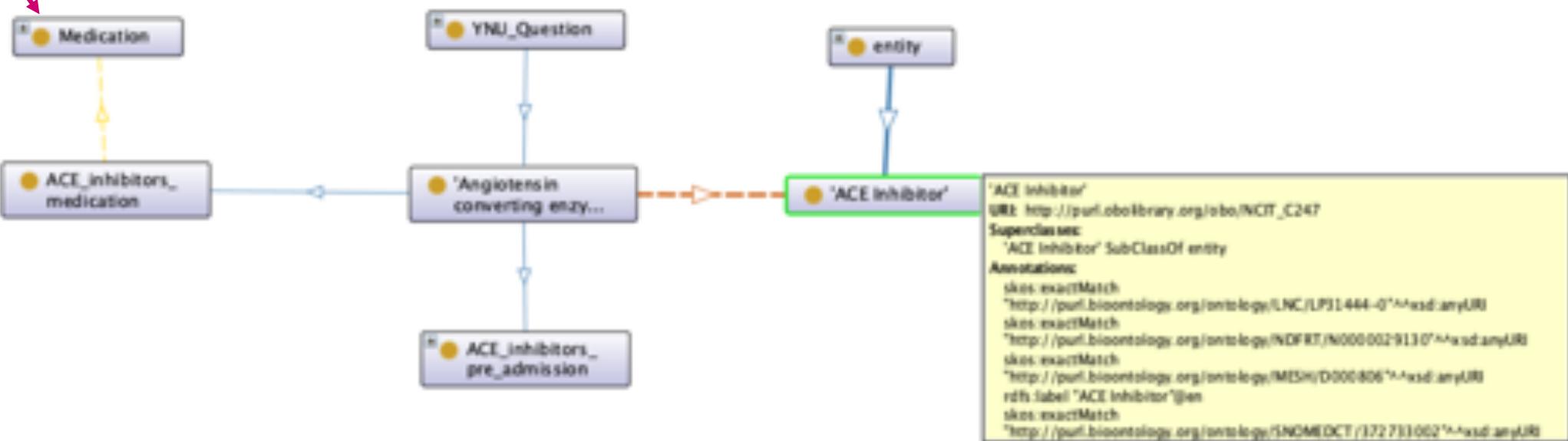
Antimalarial agent?  Yes  No  Unknown If yes, specify: \_\_\_\_\_

Experimental agent?  Yes  No  Unknown If yes, specify: \_\_\_\_\_

Non-steroidal anti-inflammatory (NSAID)  Yes  No  Unknown

Angiotensin converting enzyme inhibitors (ACE inhibitors)  Yes  No  Unknown

Angiotensin II receptor blockers (ARBs)  Yes  No  Unknown



# WHO'S COVID-19 CRF SEMANTIC DATA MODEL

**MEDICATION** Is the patient CURRENTLY receiving any of the following?

Oral/orogastric fluids?  Yes  No  Unknown Intravenous fluids?  Yes  No  Unknown

Antiviral?  Yes  No  Unknown If yes:  Ribavirin  Lopinavir/Ritonavir  Neuraminidase inhibitor  
 Interferon alpha  Interferon beta  Other, specify: \_\_\_\_\_

Corticosteroid?  Yes  No  Unknown If yes, route:  Oral  Intravenous  Inhaled  
If yes, please provide agent and maximum daily dose: \_\_\_\_\_

Antibiotic?  Yes  No  Unknown Antifungal agent?  Yes  No  Unknown

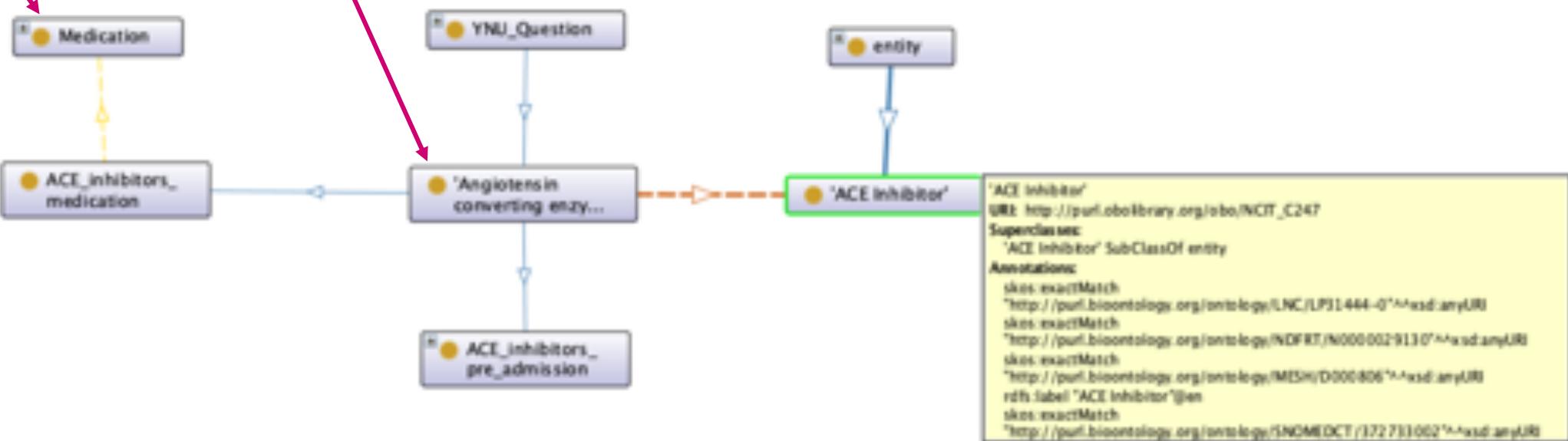
Antimalarial agent?  Yes  No  Unknown If yes, specify: \_\_\_\_\_

Experimental agent?  Yes  No  Unknown If yes, specify: \_\_\_\_\_

Non-steroidal anti-inflammatory (NSAID)  Yes  No  Unknown

Angiotensin converting enzyme inhibitors (ACE inhibitors)  Yes  No  Unknown

Angiotensin II receptor blockers (ARBs)  Yes  No  Unknown



# WHO'S COVID-19 CRF SEMANTIC DATA MODEL

**MEDICATION** Is the patient CURRENTLY receiving any of the following?

Oral/orogastric fluids?  Yes  No  Unknown Intravenous fluids?  Yes  No  Unknown

Antiviral?  Yes  No  Unknown If yes:  Ribavirin  Lopinavir/Ritonavir  Neuraminidase inhibitor  
 Interferon alpha  Interferon beta  Other, specify: \_\_\_\_\_

Corticosteroid?  Yes  No  Unknown If yes, route:  Oral  Intravenous  Inhaled  
If yes, please provide agent and maximum daily dose: \_\_\_\_\_

Antibiotic?  Yes  No  Unknown Antifungal agent?  Yes  No  Unknown

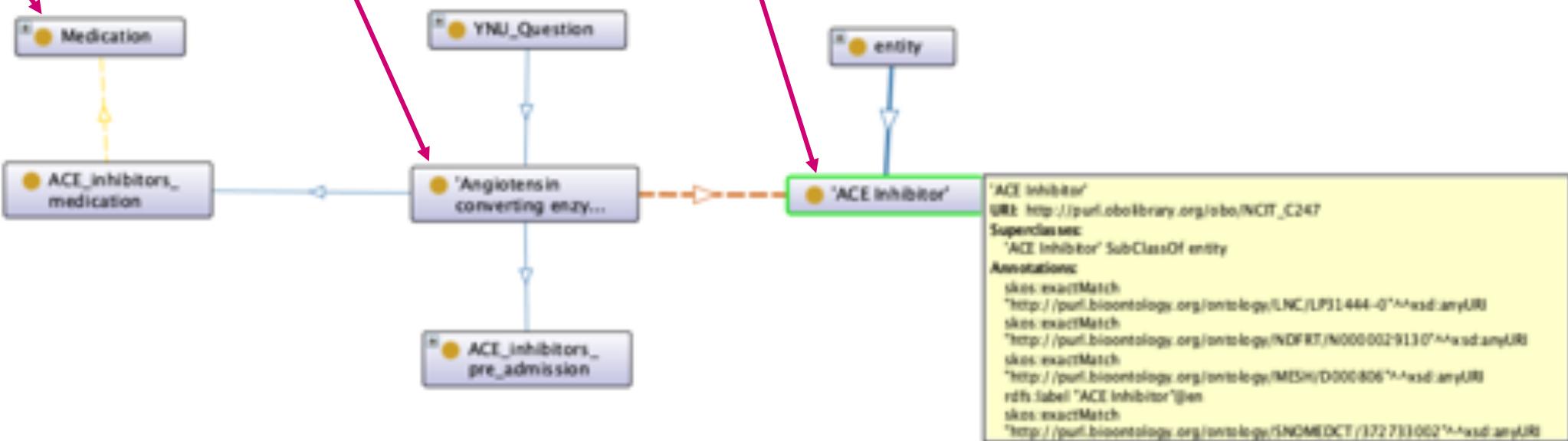
Antimalarial agent?  Yes  No  Unknown If yes, specify: \_\_\_\_\_

Experimental agent?  Yes  No  Unknown If yes, specify: \_\_\_\_\_

Non-steroidal anti-inflammatory (NSAID)  Yes  No  Unknown

Angiotensin converting enzyme inhibitors (ACE inhibitors)  Yes  No  Unknown

Angiotensin II receptor blockers (ARBs)  Yes  No  Unknown



# WHO's COVID-19 CRF semantic data model

Antiviral → SNOMEDCT:372701006

Antiviral Agent → NCIT:C281

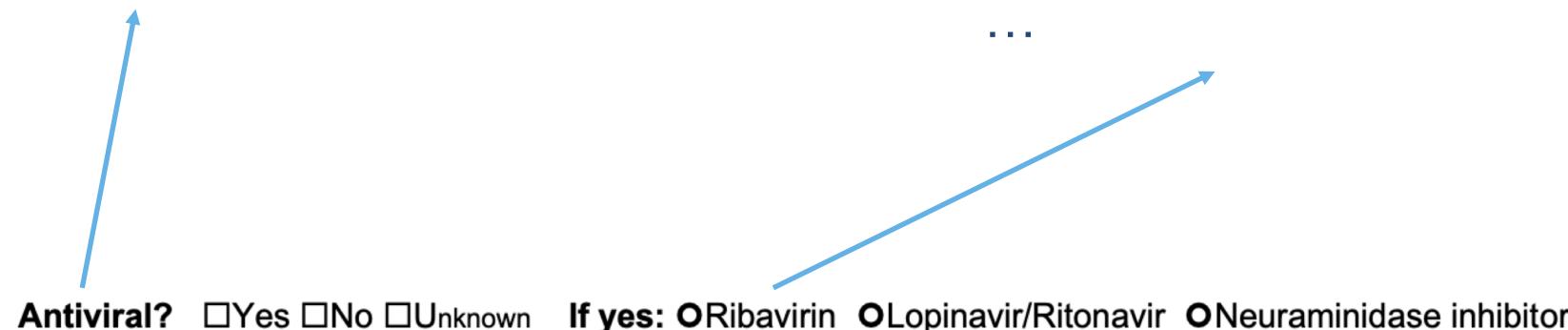
Antiviral Agents → MESH:D000998

Ribavirin → SNOMEDCT:387188005

Ribavirin → NCIT:C807

Ribavirin → MESH:D012254

...



...

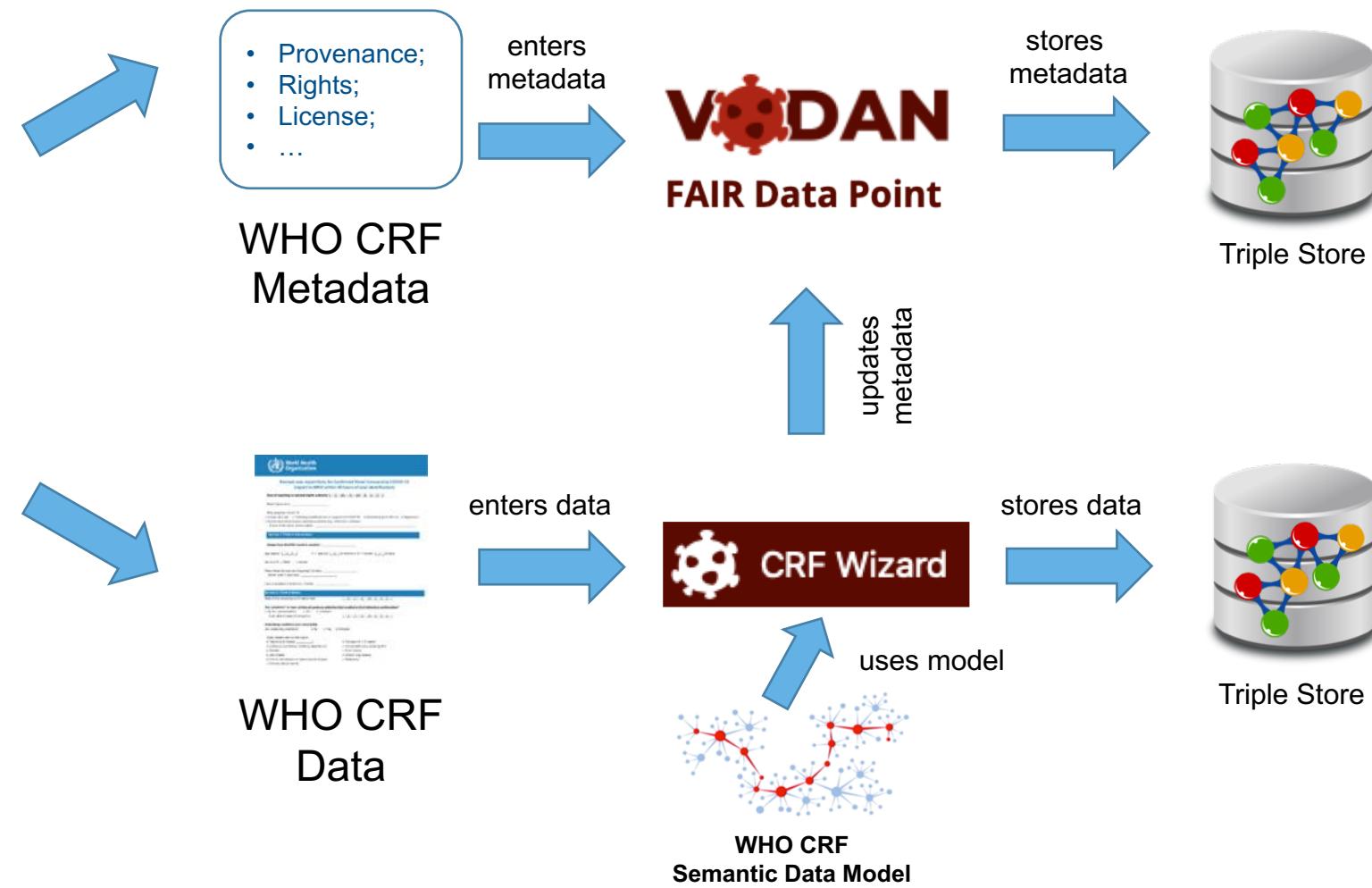
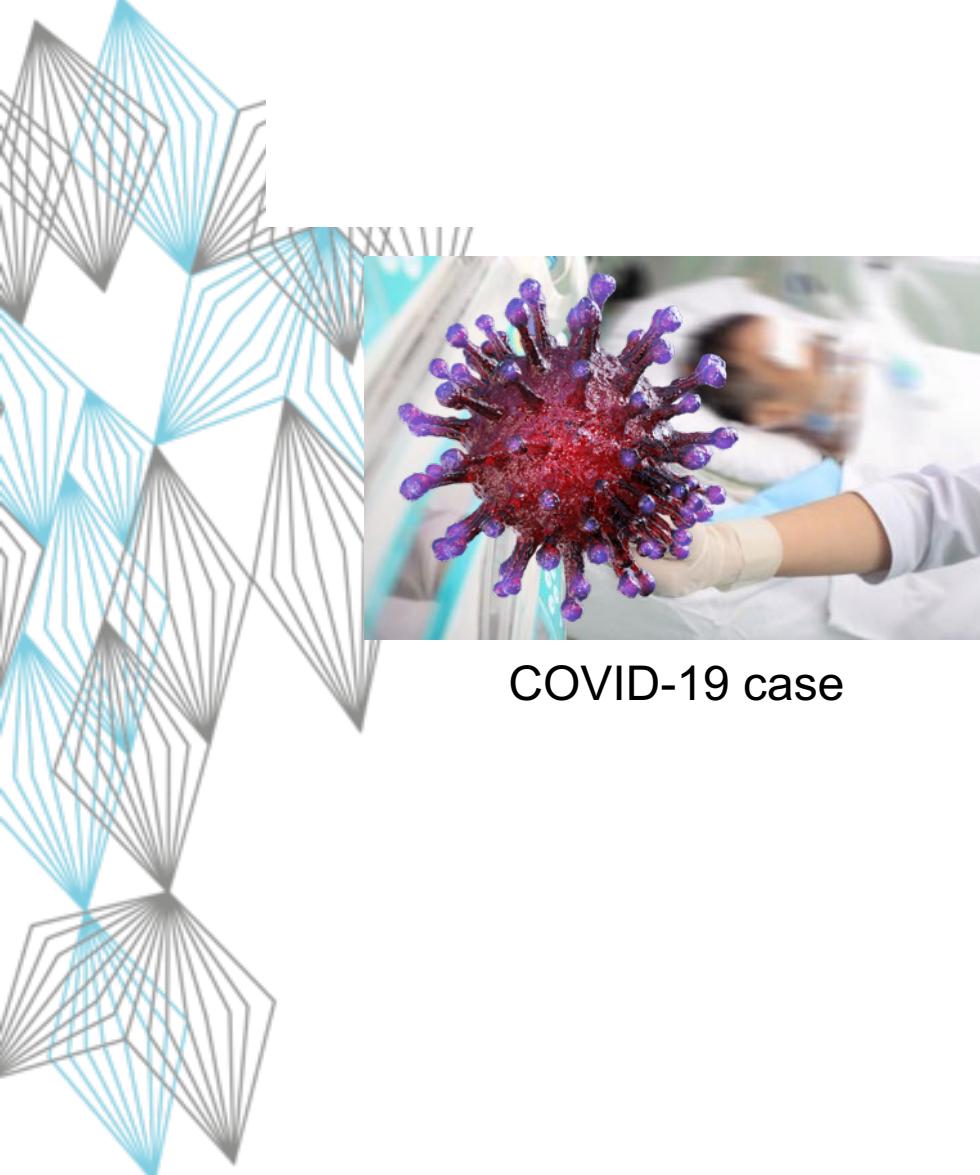
Lopinavir- and ritonavir-containing product → SNOMEDCT:387067003

Lopinavir/Ritonavir → NCIT:C2096

Lopinavir-ritonavir drug combination → MESH:D558899

...

# VODAN IN A BOX



# DATA ENTRY

## Existing cases (CRFs)

The screenshot shows the 'CRFs' section of the CRF Wizard interface. On the left, a sidebar lists navigation options: CRF Wizard, Users, CRF's Template Editor, CRF's Templates, CRFs, Reports, Templates, Settings, Help, and Logout. The main area displays a list of existing forms:

- test [submitted] WHO COVID-19 Case Report Form, 0.1.2 (voden:who-covid19-crft:0..1..2) - Answered 21/01 Updated about 15 hours ago
- test WHO COVID-19 Case Report Form, 0.1.3 (voden:who-covid19-crft:0..1..3) - Answered 21/01 Updated about 15 hours ago
- VODAN-patient-000003 [submitted] (f) Albert Einstein - WHO COVID-19 Case Report Form, 0.1.2 (voden:who-covid19-crft:0..1..2) - Answered 06/05 Updated 3 days ago
- V VODAN-patient-000004 WHO COVID-19 Case Report Form, 0.1.3 (voden:who-covid19-crft:0..1..3) - Answered 58/07 Updated 4 minutes ago
- V VODAN-patient-000005 WHO COVID-19 Case Report Form, 0.1.3 (voden:who-covid19-crft:0..1..3) - Answered 288/292 Updated 6 minutes ago
- V VODAN-patient-000006 WHO COVID-19 Case Report Form, 0.1.3 (voden:who-covid19-crft:0..1..3) - Answered 288/292 Updated 1 minute ago
- V VODAN-patient-0001 [submitted] (f) Albert Einstein - WHO COVID-19 Case Report Form, 0.1.2 (voden:who-covid19-crft:0..1..2) - Answered 128/135 Updated 2 days ago
- V VODAN-test-000002 [submitted] (f) Albert Einstein - WHO COVID-19 Case Report Form, 0.1.2 (voden:who-covid19-crft:0..1..2) - Answered 47/62 Updated 2 days ago

A 'Create' button is located at the top right of the list.

## New case (CRFs)

The screenshot shows the 'VODAN-patient-000006 (WHO COVID-19 Case Report Form, 0.1.3)' form in the CRF Wizard. The left sidebar includes 'Users', 'CRF's Template Editor', 'CRF's Templates', 'CRFs', 'Reports', 'Templates', 'Settings', 'Help', and 'Logout'. The right panel is titled 'I. Introduction' and contains the following text:

This is adjusted version of COVID-19 CASE RECORD FORM ISAPD version 6APR2020. In response to the coronavirus disease 2019 (COVID-19) epidemic, the World Health Organization (WHO) is launching a global COVID-19 Anonymized Clinical Data Platform (the "COVID-19 Data Platform") to enable State Parties to the International Health Regulations (IHR) (2005) to share with WHO anonymized clinical data and information related to patients who suspected or confirmed infections with the 2019-nCoV (collectively "Anonymized COVID-19 Data"). The Anonymized COVID-19 data received from State Parties through the COVID-19 Data Platform will remain property of the constituting State Party and will be used by WHO for purposes of verification, assessment and assistance pursuant to the IHR (2005), including to inform the public health and clinical operation response in connection with the COVID-19 outbreak. To help achieve such purposes, WHO will establish an independent Clinical Advisory Group to advise WHO on global reporting and analysis of the Anonymized COVID-19 Data. State Parties are invited to contribute Anonymized COVID-19 Data to the COVID-19 Data Platform. State Parties should please contact WHO at to obtain more information about, including log-in credentials for, the COVID-19 Platform. To preserve the security and confidentiality of the Anonymized COVID-19 Data, State Parties are respectfully requested to take all necessary measures to protect their respective log-in credentials and passwords to the COVID-19 Data Platform.

**Design of this Case Report Form**

The CRF is designed to collect data obtained through examination, interview and review of hospital notes. Data may be collected retrospectively if the patient is enrolled after the admission date. The data collection period is defined as the period from hospital admission to discharge, transfer, death, or continued hospitalization without possibility of continued data collection.

This CRF has 3 modules:

- Module 1 to be completed on the first day of admission to the health centre.
- Module 2 to be completed on first day of admission to ICU or high dependency unit. Module 2 should also be completed daily for as many days as resources allow. Continue to follow-up patients who transfer between wards.
- Module 3 to be completed at discharge or death.

**General Guidance**

- Participant identification Numbers consist of a site code and a participant number. You can register on the data management system by contacting [COVID\\_CRFPlatform@who.int](mailto:COVID_CRFPlatform@who.int), and our data management team will contact you with instructions for data entry and will assign you a 5-digit site code at that time.
- Please contact us at [COVID\\_CRFPlatform@who.int](mailto:COVID_CRFPlatform@who.int) if we can help with databases, if you have comments and to let us know that you are using the forms.

**1. Do you have site code assigned?**

a. Yes  
 b. No  
 Clear answer

Internal UUID of CRF will be used for both reference in triple store but there will be the official entered Participant ID to identify the participant.

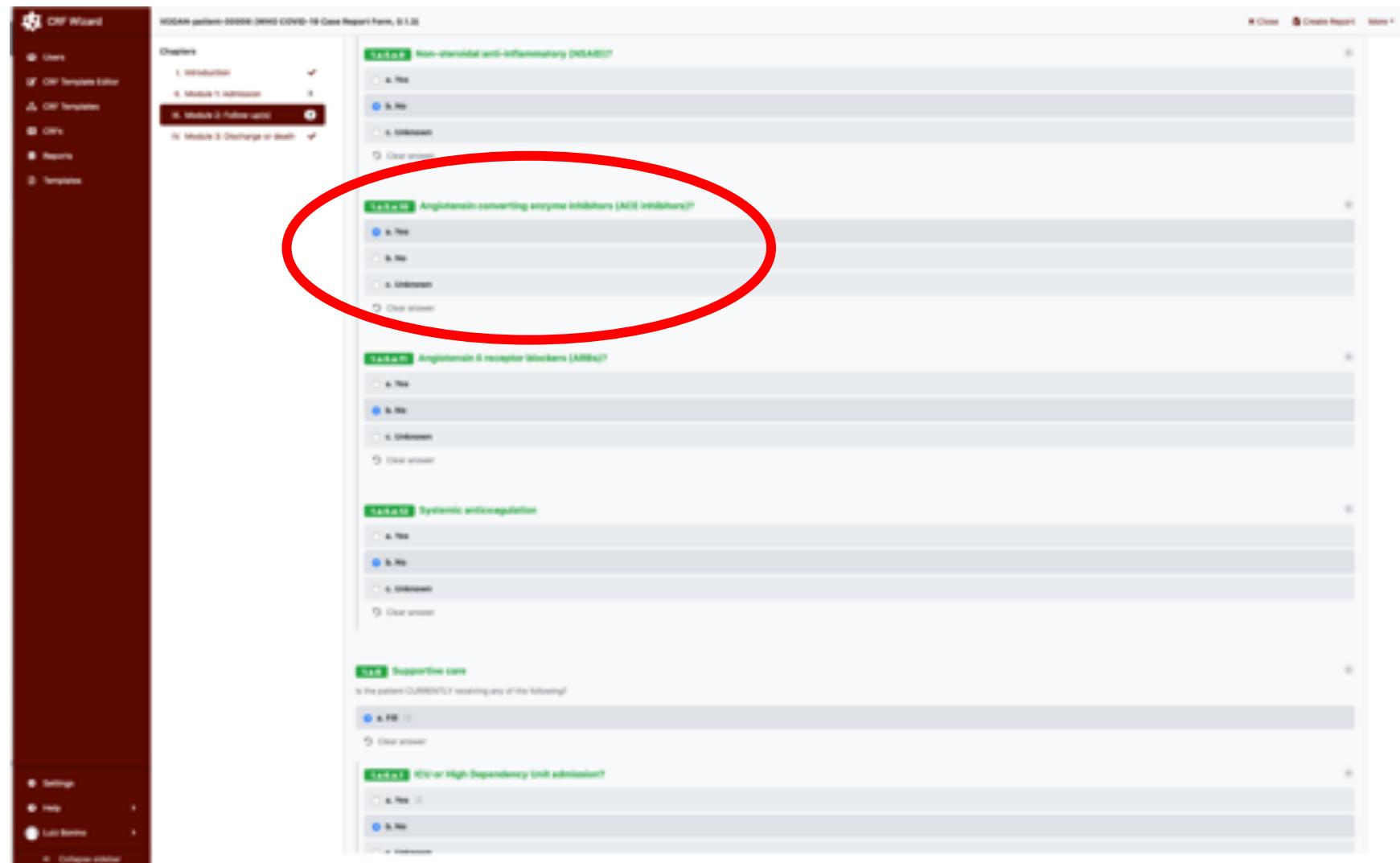
**1.a.1 Participant ID**

VODAN-patient-000007

Participant identification Numbers consist of a site code and a participant number. You can obtain a site code and register on the data management system by contacting [EDCARN@who.int](mailto:EDCARN@who.int). Participant numbers should be assigned sequentially for each site beginning with [NNNN1](#). In the case of a single site recruiting participants on different wards, or where it is otherwise difficult to assign sequential numbers, you can assign numbers in blocks or incorporate alpha-characters. E.g. Ward X will assign numbers from [NNNN1](#) onwards, and Ward Y will assign numbers from [NNNN1](#) onwards.

# DATA ENTRY

Follow-up module → Medication



The screenshot shows the CRF Wizard software interface for a COVID-19 Case Report Form (version 0.1.0). The left sidebar contains navigation links: User, CRF Template Editor, CRF Templates, CRFs, Reports, Templates, Settings, Help, Last Review, and Support center. The main content area displays the 'Follow-up' module. The 'Medication' section includes the following questions:

- Q.1.a.i.e.i Non-steroidal anti-inflammatory (NSAID)?
  - a. Yes
  - b. No
  - c. Unknown
  -
- Q.1.a.i.e.ii Angiotensin-converting enzyme inhibitors (ACE inhibitors)?
  - a. Yes
  - b. No
  - c. Unknown
  -
- Q.1.a.i.e.iii Angiotensin II receptor blockers (ARBs)?
  - a. Yes
  - b. No
  - c. Unknown
  -
- Q.1.a.i.e.iv Systemic anticoagulation
  - a. Yes
  - b. No
  - c. Unknown
  -
- Q.1.a.i.e.v Supportive care
  - a. Yes
  - b. No
  - c. Unknown
  -
- Q.1.a.i.e.vi ICU or High Dependency Unit admission?
  - a. Yes
  - b. No
  - c. Unknown
  -

Reports

Filter by name... Name: 4

1	test ■ Turtle (TTL) - test	Updated 7 minutes ago
V	VODAN-patient-00003 ■ Turtle (TTL) - VODAN-patient-00003	Updated 2 days ago
V	VODAN-patient-00004 ■ Turtle (TTL)	Updated 2 days ago
V	VODAN-patient-00004 ■ Turtle (TTL) - VODAN-patient-00004	Updated 8 minutes ago
V	VODAN-patient-00005 ■ Turtle (TTL) - VODAN-patient-00005	Updated 10 minutes ago
V	VODAN-patient-00006 ■ Turtle (TTL) - VODAN-patient-00006	Updated 5 minutes ago
V	VODAN-test ■ Turtle (TTL) - VODAN-patient-0001	Updated 2 days ago
V	VODAN-test-00002 ■ turtle (TTL) - VODAN-test-00002	Updated 2 days ago

Download Submit Delete

Download the CRF RDF

Submit the CRF RDF  
to the triple store

Submit VODAN-patient-00006

FAIR Data Storage  
Store data in configured triple store and update related metadata in FAIR Data Point.

Cancel Submit

Submit VODAN-patient-00006

✓ The report was successfully submitted.

Cancel Done

# CRF METADATA – FAIR DATA POINT

Metadata repository/FAIR Data Point metadata

A complex, abstract graphic on the left side of the slide consists of numerous thin, intersecting lines forming a three-dimensional, crystalline-like structure. The lines are primarily grey and blue, with some highlights in white and yellow.

**VODAN FAIR Data Point**

**VODAN FAIR Data Point - Demonstration**

This is the demonstration FAIR Data Point of the Virus Outbreak DATA Network (VODAN) initiative. The content in this FDP is for demonstration purposes only and does not represent real cases. The VODAN FDP is part of the VODAN in a Box, a set of tools and models supporting quick deployment of solutions to allow recording reports of COVID-19 cases based on WHO's COVID-19 CRF (with the VODAN DSML exposing the metadata of these gathered data (with the VODAN FDP) and improve the semantic interoperability by having the case reports semantically enriched (with the embedded WHO COVID-19 CRF Semantic Data Model).

**Catalogs**

**Demonstration WHO COVID-19 CRF catalog**

A catalog containing a number of datasets created using the VODAN in a Box toolset. These datasets have mocked-up metadata for demonstration purposes only.

Issued 08-07-2020 Modified 25-07-2020

Log in

Metadata Issued 08-07-2020	Metadata Modified 25-07-2020
Version 1.6.0	
Language en	
License cc-by-nc-nd4.0	
Institution gofairfoundation.org	
Start date 01-07-2020	
Institution country Q29999	
Download RDF <a href="#">rdf+xml</a> <a href="#">json-ld</a>	

FAIR Data Point · About

# CRF METADATA – FAIR DATA POINT

## Dataset metadata



**VODAN FAIR Data Point** [Log in](#)

VODAN.FAIR Data Point · Demonstration / Demonstration WHO COVID-19 CRF.xlsx / ACME Hospital COVID-19 CRF synthetic dataset

### ACME Hospital COVID-19 CRF synthetic dataset

Dataset containing the case reports of COVID-19 for the ACME Hospital

Distributions

**ACME Hospital COVID-19 CRF distribution**  
Issued 13-07-2020 Modified 25-07-2020 Media Type text/txt

Metadata Issued **08-07-2020** Metadata Modified **25-07-2020**

Version **0.1**

Language **en**

License **cc-by-nc-nd3.0**

Theme

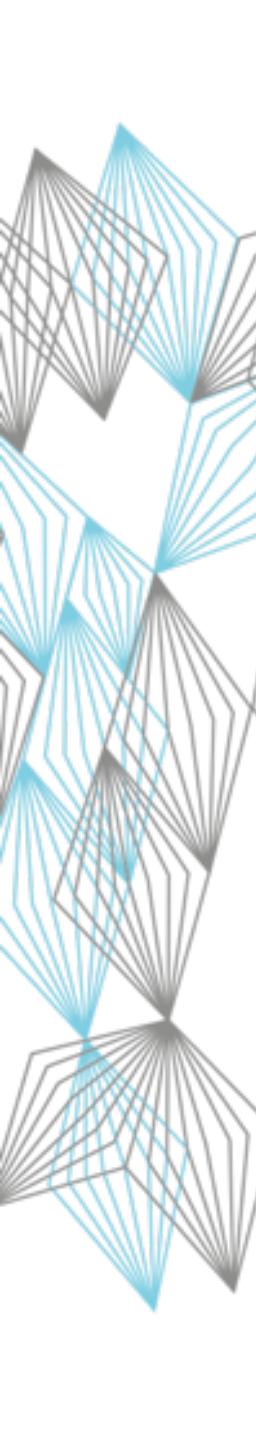
- DOID\_0080609
- MONDO\_0100096
- [RA6DKJHGPI\\_2mE9ojtB3YQXlwGGdCC8Z](#)
- [WpkxElloHsxOjE](#)

[▼ Show more](#)

Keyword

- COVID-19
- CRF
- VODAN IN

[Download RDP](#)  
[ttl](#) [rdf+xml](#) [json-id](#)



# EXAMPLE DATA ANALYSIS

- With a number of VODAN FDPs deployed, we can start asking questions to the data (if licenses permit).
- Example question:
  - Present the names of the health care facilities that have had male patients that were admitted with fever (temperature => 38)

# EXAMPLE SPARQL QUERY

```
# View triples

PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX crfsm: <http://purl.org/vodan/whocovid19crfsemdatamodel/>
PREFIX crfsmi: <http://purl.org/vodan/whocovid19crfsemdatamodel/instances/>
PREFIX obo: <http://purl.obolibrary.org/obo/>

SELECT ?facilityName ?participantID
WHERE
{
  # Selects entities that are of type WHO COVID-19 RAPID CRF
  ?case rdf:type crfsm:who-covid-19-rapid-crf .

  # Retrieves the participant ID (patientID) of the case report
  ?case crfsm:participant_id ?participantID .

  # Selects the case reports having a module1 as one of its parts (BFO_0000051 = hasPart)
  ?case obo:BFO_0000051 ?module1 .

  # View triples
  ?module1 rdf:type crfsm:Module_1 .

  # Selects the case reports having as one of its parts an element of type Facility_name
  ?module1 obo:BFO_0000051 ?facility .
  ?facility rdf:type crfsm:Facility_name .

  # Retrieves the facility name
  ?facility crfsm:has_literal_value ?facilityName .

  # Selects the case reports with a module1 subsection of type Demographics
  ?module1 obo:BFO_0000051 ?demographics .
  ?demographics rdf:type crfsm:Demographics .

  # Selects the case reports whose patients are males (C46109). For females, use C46110
  ?demographics obo:BFO_0000051 ?genderQuestion .
  ?genderQuestion rdf:type crfsm:Sex .

  # Selects the case reports with a Vital Signs subsections with the question temperature on admission filled
  ?genderQuestion crfsm:has_value crfsmi:C46109 .

  # Selects the case reports with a module1 subsection of type Vital Signs
  ?module1 obo:BFO_0000051 ?vitalSigns .
  ?vitalSigns rdf:type crfsm:Vital_signs .

  # Selects the case reports with a Vital Signs subsections with the question temperature on admission filled
  ?vitalSigns obo:BFO_0000051 ?temperatureAtAdmission .
  ?temperatureAtAdmission rdf:type crfsm:Temperature_admission .

  # Retrieves the temperature at admission
  ?temperatureAtAdmission crfsm:has_literal_value ?temperatureValue .

  # Filters the case reports whose patients had fever (temperature >= 38) at admission
  FILTER (?temperatureValue >= 38) .
}
```

# EXAMPLE SPARQL QUERY RESULTS

## Edit query

```
20  
21 # Selects the case reports whose patients are males (046109). For females, use 046110  
22 Demographic obo:RPO_0000051 genderQuestion .  
23 genderQuestion rdfs:type owl:Class .  
24 genderQuestion rdfs:label_value owl:xml:046109 .  
25  
26 # Selects the case reports with a modular subsection of type Vital Signs  
27 Modular obo:RPO_0000051 vitalSigns .  
28 vitalSigns rdfs:type owl:Class .  
29  
30 # Selects the case reports with a Vital Signs subsection with the question temperature on admission filled  
31 vitalSigns obo:RPO_0000051 TemperatureAtAdmission .  
32 TemperatureAtAdmission rdfs:type owl:Temperature_admission .  
33  
34 # Retrieves the temperature at admission  
35 TemperatureAtAdmission rdfs:label_value TemperatureValue .  
36  
37 # Filters the case reports whose patients had fever (temperature > 38) at admission  
38 FILTER (TemperatureValue > 38) .  
39
```

Language: SPARQL   
 Limit to 1000 results  
 Reasoning  
 Long paths  
 Cancel on warnings  
 Show namespaces  
 Add a namespace  
 Edit source  
[Permalink to query](#)

Execute Log Query Show Plan Save as Add to repository

3 results in 55.062 ms		Information
facilityName	participantID	
"Neverland Hospital"	"VODAN-patient-00006"	
"ACME Hospital"	"VODAN-patient-00004"	
"ACME Hospital"	"VODAN-patient-00003"	



# NEXT STEPS

- Extend the model to other types of data relevant to virus outbreaks;
- Define/reuse a domain ontology describing the concepts and relations related to virus infections;
- Develop client applications to facilitate analysis of the data
- Collect feedback from the real-world deployments;



Leiden University  
Medical Center

### Luiz Bonino

International Technology Coordinator – GO FAIR  
Associate Professor – SCS – U. Twente  
Associate Professor – BioSemantics – LUMC

E-mail: [luiz.bonino@go-fair.org](mailto:luiz.bonino@go-fair.org)  
Skype: [luizolavobonino](#)  
Web: [www.go-fair.org](http://www.go-fair.org)