



# Metadata for Machines Workshop: VODAN Africa

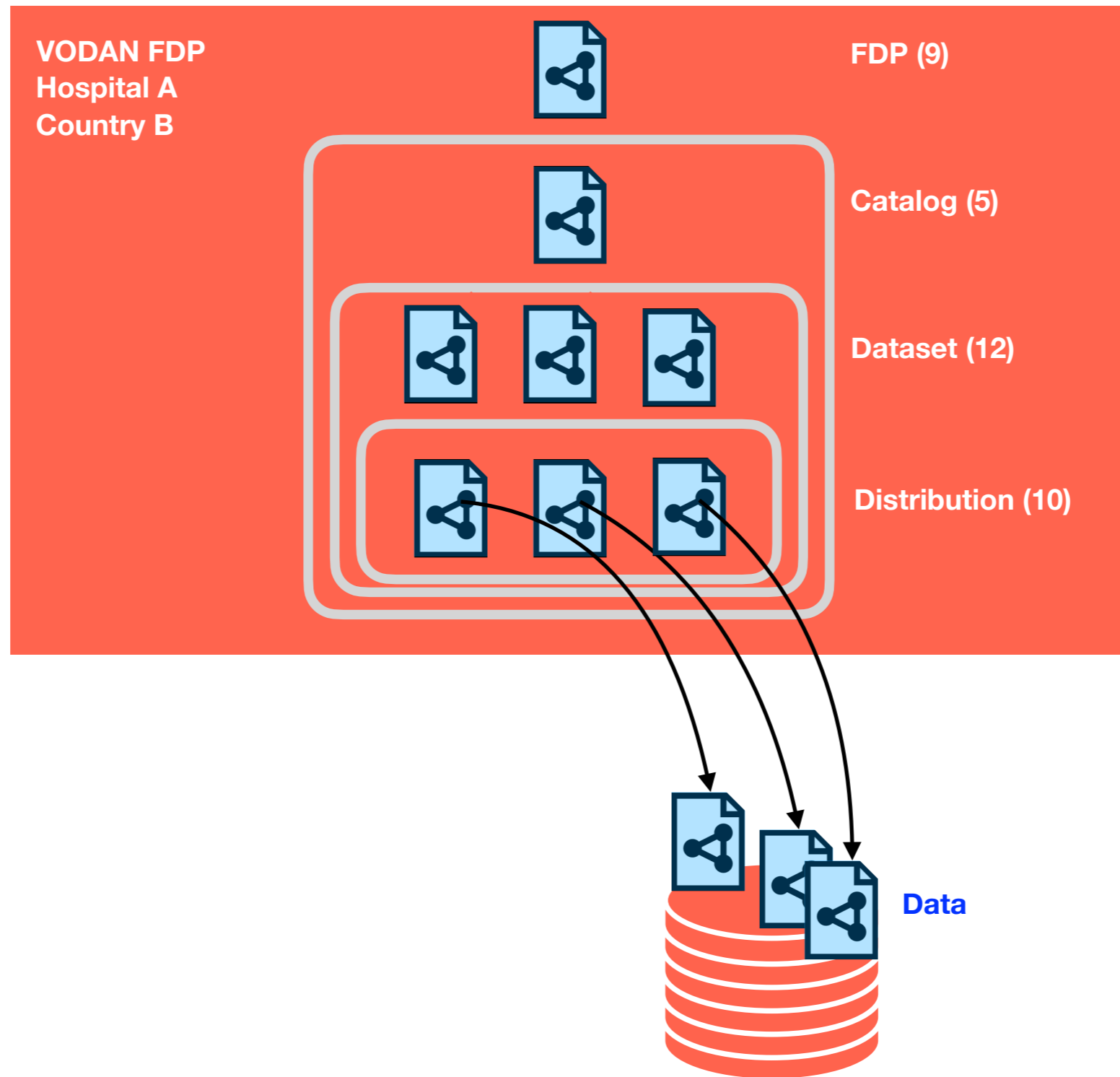
Dr. Erik Schultes, PhD  
International Science Coordinator  
GO FAIR International Support and Coordination Office  
<http://orcid.org/0000-0001-8888-635X>  
[erik.schultes@go-fair.org](mailto:erik.schultes@go-fair.org)  
[go-fair.org](http://go-fair.org)  
<https://bit.ly/M4Mhandbook>

M4M.4.12, July 13, 2020



- There are 15 people on the call potentially, and 1 hour reserved, so I want to keep it focused on KIU.
- I will moderate the discussion. I'll have a few slides to kick off the webinar.
- Primary goal: Configure the KIU FDP.
- For this, we share screens, and we go through each of the 32 questions to complete the configuration.
- Bonus goal:
  - connect KIU FDP to datasets (perhaps those from Prof Mariem, Tunisia CEDAR instances?)
- Next steps are to demonstrate some form of query between the KIU and a Dutch FDP. Options:
  - 1) scrape metadata, pool, and analyse
  - 2) using Uganda cloud environment, load data into a triple store and run SPARQL query

# KIU FDP Configuration



# Configuring the FAIR Data Point

## VIDAN FAIR Data Point ES

### Erik's FAIR Data Point Edit

This is a demonstration FDP, especially for M4M Workshops

**Catalogs** + Create

**Catalog 1**  
Literal  
Issued 29-04-2020 Modified 05-05-2020

**Catalog 2**  
Example data catalog number 2  
Issued 05-05-2020 Modified 05-05-2020

**My Catalog**  
Issued 29-04-2020 Modified 29-04-2020

**Vojta**  
Issued 29-04-2020 Modified 29-04-2020

Metadata Issued 24-04-2020 Metadata Modified 05-05-2020

Version 1.0

License [cc-by-nc-nd3.0](#)

Specification [fdpMetadata](#)

Language [en](#)

Download RDF [ttl](#) [rdf+xml](#) [json-ld](#)

# Configuring the FAIR Data Point

## VIDAN FAIR Data Point ES

### Erik's FAIR Data Point Edit

This is a demonstration FDP, especially for M4M Workshops

**Catalogs** + Create

- Catalog 1**  
Literal  
Issued 29-04-2020 Modified 05-05-2020
- Catalog 2**  
Example data catalog number 2  
Issued 05-05-2020 Modified 05-05-2020
- My Catalog**  
Issued 29-04-2020 Modified 29-04-2020
- Vojta**  
Issued 29-04-2020 Modified 29-04-2020

Metadata Issued: 24-04-2020    Metadata Modified: 05-05-2020

Version: 1.0

License: [cc-by-nc-nd3.0](#)

Specification: [fdpMetadata](#)

Language: [en](#)

Download RDF: [ttl](#) [rdf+xml](#) [json-ld](#)

ES

Edit

### Erik's FAIR Data Point

This is a demonstration FDP, especially for M4M Workshops

#### Catalogs

+ Create

##### Catalog 1

Literal

Issued 29-04-2020 Modified 05-05-2020

##### Catalog 2

Example data catalog number 2

Issued 05-05-2020 Modified 05-05-2020

##### My Catalog

Issued 29-04-2020 Modified 29-04-2020

##### Vojta

Issued 29-04-2020 Modified 29-04-2020

Metadata Issued  
24-04-2020

Metadata Modified  
05-05-2020

Version  
1.0

License  
[cc-by-nc-nd3.0](#)

Specification  
[fdpMetadata](#)

Language  
[en](#)

Download RDF  
[ttl](#) [rdf+xml](#) [json-ld](#)

ES

Edit



# KIU FAIR Data Point

FDP (9 components)		
1	Title	KIU FAIR Data Point
2	Description	This FAIR Data Point contains the metadata of datasets and other artefacts
3	Publisher	<a href="https://www.kiu.ac.ug">https://www.kiu.ac.ug</a>
4	Version	1.0
5	Language	<a href="http://id.loc.gov/vocabulary/iso639-1/en">http://id.loc.gov/vocabulary/iso639-1/en</a>
6	License	<a href="http://rdjlicense.appspot.com/rdjlicense/cc-by-nc-nd4.0">http://rdjlicense.appspot.com/rdjlicense/cc-by-nc-nd4.0</a>
7	Start date	3rd June, 2020
8	Last update	13 July, 2020
9	Institution country	<a href="https://www.wikidata.org/wiki/Q1036">https://www.wikidata.org/wiki/Q1036</a>

# KIU FAIR Data Point

Catalog metadata (5 components)		
1	Title	COVID Case Report Forms
2	Description	Description of dataset
3	Publisher	<a href="https://www.kiu.ac.ug">https://www.kiu.ac.ug</a>
4	Version	1.0
5	Language	<a href="http://id.loc.gov/vocabulary/iso639-1/en">http://id.loc.gov/vocabulary/iso639-1/en</a>

# KIU FAIR Data Point

Dataset metadata (12 components)		
1	Title	Name of dataset
2	Description	Description of data set
3	Publisher	<a href="https://www.kiu.ac.ug">https://www.kiu.ac.ug</a>
4	Version	1.0
5	Language	<a href="http://id.loc.gov/vocabulary/iso639-1/en">http://id.loc.gov/vocabulary/iso639-1/en</a>
6	License	<a href="https://creativecommons.org/licenses/by-nc">https://creativecommons.org/licenses/by-nc</a>
7	Issued	13 July, 2020
8	Modified	13 July, 2020
9	Keywords	COVID-19; Case Report Form; KIU; Uganda
10	Theme	<a href="http://bioportal.bioontology.org/ontologies/C">http://bioportal.bioontology.org/ontologies/C</a>
11	Contact point	-
12	Landing page	-





# KIU FAIR Data Point

Distribution Metadata (10 components)		
1	Title	Name of dataset distribution
2	Description	Description of dataset and distribution type
3	License	<a href="https://creativecommons.org/licenses/by-nc/">https://creativecommons.org/licenses/by-nc/</a>
4	Issued	13 July, 2020
5	Modified	13 July, 2020
6	Download URL	Link to download file
7	Access URL	Link to access point (SPARQL end point)
8	Media Type	<a href="https://w3id.org/mediatype/text/turtle">https://w3id.org/mediatype/text/turtle</a>
9	Format	Human readable
10	Byte Size	Estiamted file size

Repository metadata (9 components)	
1	<b>Title:</b> Give a title, or name, of the FAIR Data Point. Ex.: VODAN FAIR Data Point, KIU FAIR Data Point, LUMC FAIR Data Point
2	<b>Description:</b> A textual description of the FAIR Data Point explaining its purpose, the context, etc. E.g., "This FAIR Data Point contains the metadata of datasets and other artefacts related to the COVID-19 pandemic in hospital ACME".
3	<b>Publisher:</b> The person or organization responsible for the FAIR Data Point. E.g., Hospital A, University B, etc.
4	<b>Version:</b> The version of this deployment of the FAIR Data Point. E.g., 1.0
5	<b>Language</b> The URI of the main language of this FAIR Data Point according to ISO 639-1. E.g: <a href="http://id.loc.gov/vocabulary/iso639-1/en">http://id.loc.gov/vocabulary/iso639-1/en</a> , <a href="http://id.loc.gov/vocabulary/iso639-1/fr">http://id.loc.gov/vocabulary/iso639-1/fr</a>
6	<b>License</b> The URI of the usage license for the FAIR Data Point. Notice that this license does not need to be the same of the datasets' licenses. This is for the usage of the FAIR Data Point as a metadata provisioning service. E.g., <a href="http://rd]license.appspot.com/rd]license/cc-by-nc-nd4.0">http://rd]license.appspot.com/rd]license/cc-by-nc-nd4.0</a>
7	<b>Start date</b> The date when the FDP started operating. MM DD YYYY X/Y/ 2020
8	<b>Last update</b> The date when the FDP was last updated operating. MM DD YYYY X/Y/ 2020
9	<b>Institution country:</b> The URI of the country of origin of the institution responsible for the FDP. E.g., <a href="https://www.wikidata.org/entity/Q55">https://www.wikidata.org/entity/Q55</a> (URI for The Netherlands in WikiData), <a href="https://www.omg.org/spec/LCC/Countries/ISO3166-1-CountryCodes/Netherlands">https://www.omg.org/spec/LCC/Countries/ISO3166-1-CountryCodes/Netherlands</a> (URI for The Netherlands in ISO 3166-1 Country Codes Ontology).

- FDP
- VODAN
- Country

- Catalog
- Hospital

- Dataset
- eCRF

- Distribution
- RDF
- Txt
- XML

Catalog metadata (5 components)	
1	<b>Title:</b> Give a title, or name, of the Catalog. Ex.: COVID-19 CRFs; Virus Sequence Data; Refugee Interviews
2	<b>Description:</b> A textual description of the catalog. E.g., "This catalog contains patient CRFs following the WHO standard".
3	<b>Publisher:</b> The publisher (person or organization) responsible for this catalog. E.g., Hospital A, University B, etc.
4	<b>Version:</b> The version of this catalog. E.g., 1.0
5	<b>Language:</b> The URI of the main language used in this catalog according to ISO 639-1. E.g: <a href="http://id.loc.gov/vocabulary/iso639-1/en">http://id.loc.gov/vocabulary/iso639-1/en</a> , <a href="http://id.loc.gov/vocabulary/iso639-1/fr">http://id.loc.gov/vocabulary/iso639-1/fr</a>

- FDP
- VODAN
- Country

- **Catalog**
- Hospital

- Dataset
- eCRF

- Distribution
- RDF
- Txt
- XML

Dataset metadata (12 components)	
1	<b>Title:</b> Title, or name, of the dataset. Ex.: Clinical trial dataset, COVID-19 WHO CRF dataset, etc
2	<b>Description:</b> A textual description of the dataset. E.g., "This dataset is the COVID-19 WHO case report form data for hospital X."
3	<b>Publisher:</b> The publisher (person or organization) responsible for this dataset. E.g., Hospital A, University B, etc.
4	<b>Version:</b> The version of this dataset. E.g., 1.0
5	<b>Language:</b> The URI of the main language used in this dataset according to ISO 639-1. E.g: <a href="http://id.loc.gov/vocabulary/iso639-1/en">http://id.loc.gov/vocabulary/iso639-1/en</a> , <a href="http://id.loc.gov/vocabulary/iso639-1/fr">http://id.loc.gov/vocabulary/iso639-1/fr</a>
6	<b>License:</b> The URI of the usage license for the dataset. E.g., <a href="http://rdlicense.appspot.com/rdlicense/cc-by-nc-nd4.0">http://rdlicense.appspot.com/rdlicense/cc-by-nc-nd4.0</a>
7	<b>Issued</b> The date when the dataset was first issued. MM DD YYYY X/Y/ 202
8	<b>Modified</b> The date when the dataset was last modified. MM DD YYYY X/Y/ 2020
9	<b>Keywords:</b> List of terms related to the dataset. These keywords/terms give hints for the human user on what the dataset is about. The machine equivalent of keyword is theme.
10	<b>Theme:</b> URIs of concepts related to the dataset. These URIs give hints for the machine user - the artificial agents - on what the dataset is about. The human equivalent of theme is keyword.
11	<b>Contact point:</b> The URI representing a contact point to get more information about the dataset
12	<b>Landing page:</b> The URI of the landing page (if other than the FDP Client) of the dataset

- FDP
- VODAN
- Country

- Catalog
- Hospital

- Dataset
- eCRF

- Distribution
- RDF
- Txt
- XML

**Distribution Metadata (10 components, technical description of the digital representation of the dataset)**

1	<b>Title:</b> Title, or name, of the distribution. Ex.: COVID-19 WHO CRF dataset in RDF, Interview document in txt, temperature in XML, etc
2	<b>Description:</b> A textual description of the distribution. E.g., "The RDF distribution of the COVID-19 WHO case report form data for hospital X."
3	<b>License:</b> The URI of the usage license for the distribution. E.g., <a href="http://rdVicense.appspot.com/rdVicense/cc-by-nc-nd4.0">http://rdVicense.appspot.com/rdVicense/cc-by-nc-nd4.0</a>
4	<b>Issued</b> The date when the distribution was first issued. MM DD YYYY X/Y/ 2020
5	<b>Modified</b> The date when the distribution was last modified. MM DD YYYY X/Y/ 2020
6	<b>Download URL:</b> The URL to download the file. Only use this if there is a file to be downloaded.
7	<b>Access URL:</b> A URL of the resource that gives access to a distribution of the dataset. E.g. landing page, feed, SPARQL endpoint.
8	<b>Media Type:</b> The media type of the distribution as defined by IANA [ <a href="https://www.w3.org/TR/vocab-dcat-2/#bib-iana-media-types">https://www.w3.org/TR/vocab-dcat-2/#bib-iana-media-types</a> ].
9	<b>Format:</b> The file format of the distribution
10	<b>Byte Size:</b> The size of a distribution in bytes

- **FDP**
  - VODAN
  - Country
- **Catalog**
  - Hospital
- **Dataset**
  - eCRF
- **Distribution**
  - RDF
  - Txt
  - XML