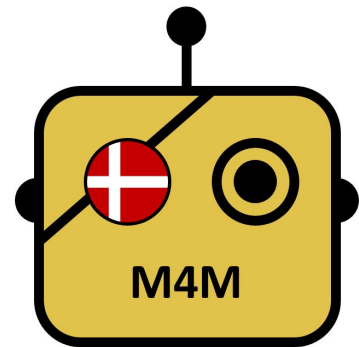


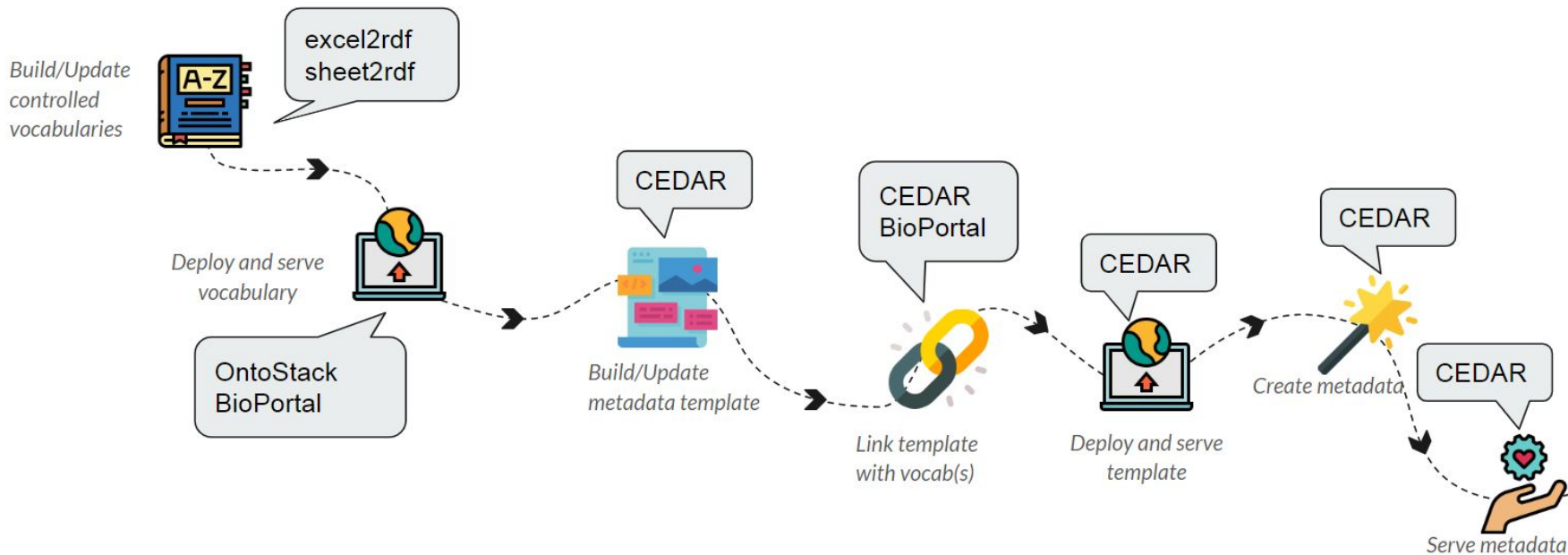
# Metadata for Machines

DeiC Konferencen 2021, 4/11-2021

Hannah Mihai, DeiC

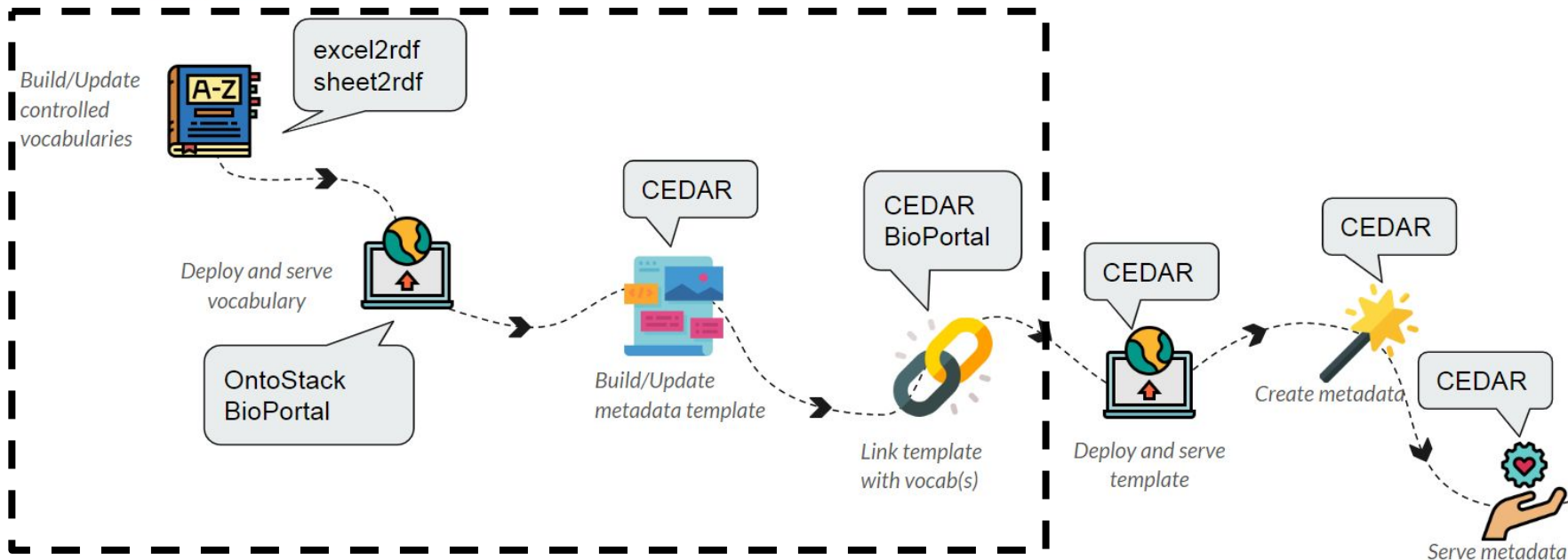


## FAIRification roadmap



Slide from: [https://zenodo.org/record/4705970#.YX\\_g7p6ZM2w](https://zenodo.org/record/4705970#.YX_g7p6ZM2w)

## FAIRification roadmap

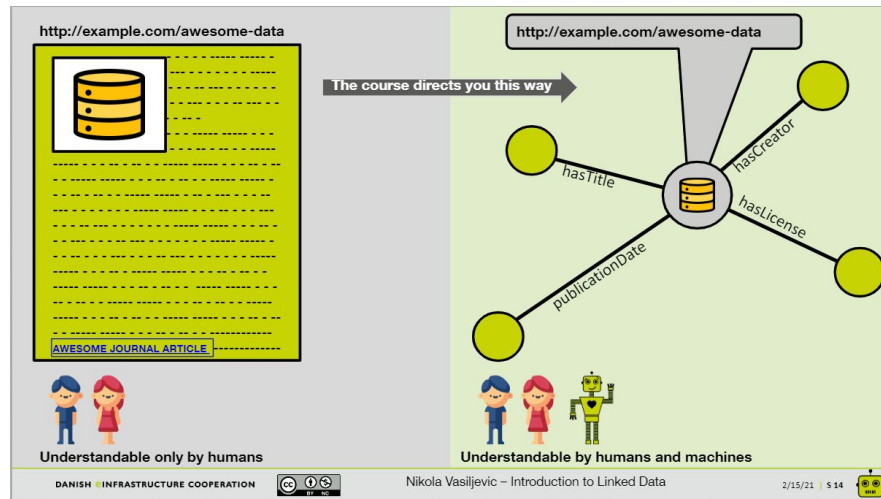


Slide from: [https://zenodo.org/record/4705970#.YX\\_g7p6ZM2w](https://zenodo.org/record/4705970#.YX_g7p6ZM2w)

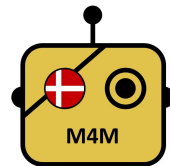
## The reason for M4M workshops

Machine actionable metadata is important, because:

- You learn good data handling practices, which strengthens your data handling for yourself individually as well as with your collaborators
- It increases the impact of your research group's data by making more of your data discoverable and increase the number of citable outputs
- It can help you receive funding.



Slide from: <https://zenodo.org/record/4462058>

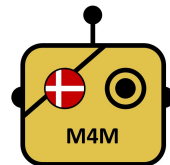


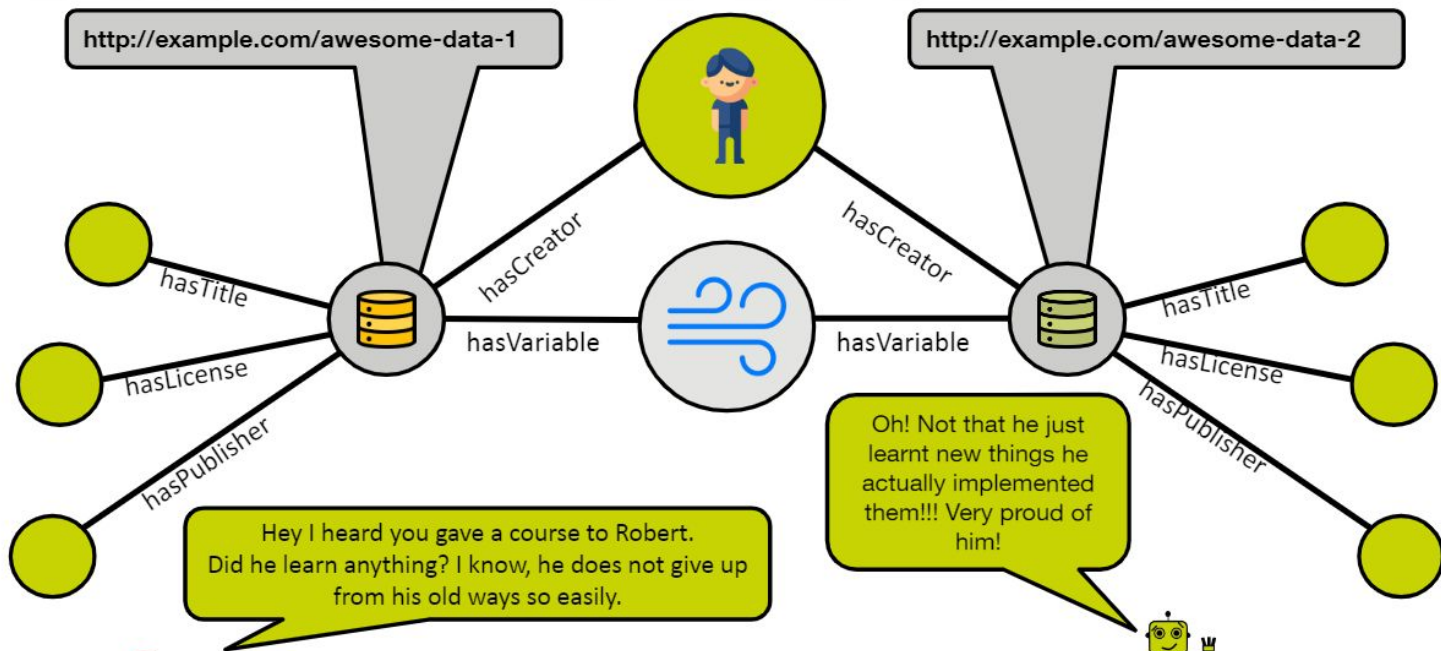
## The 'rapid' M4M concept - a collaboration with GO FAIR

- Homework before the workshop
  - Mainly consisting of getting the right accounts, finding a suitable dataset, familiarizing yourself with the concept of FAIR
- 1 hour pre-WS meeting
  - To avoid technical issues during the workshop and to assist with structuring the dataset
- 2 x ½ days online workshop or one full day physical workshop
- All information can be found here:  
<https://m4m-dk.readthedocs.io/en/latest/>

This collaboration was made possible by:

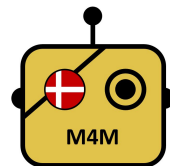
Nikola Vasiljevic [0000-0002-9381-9693](tel:0000-0002-9381-9693) (DTU, DK)  
Diba Terese Markus [0000-0001-8025-5131](tel:0000-0001-8025-5131) (AAU, DK)  
John Graybeal [0000-0001-6875-5360](tel:0000-0001-6875-5360) (Stanford University, USA)  
Erik Schultes [0000-0001-8888-635X](tel:0000-0001-8888-635X) (GFF)  
Hannah Mihai [0000-0002-0454-4289](tel:0000-0002-0454-4289) (DeiC, DK)  
Anders Sparre Conrad [0000-0002-5283-2074](tel:0000-0002-5283-2074) (DeiC, DK)  
Rene Belsø [0000-0001-6912-0498](tel:0000-0001-6912-0498) (DeiC, DK)



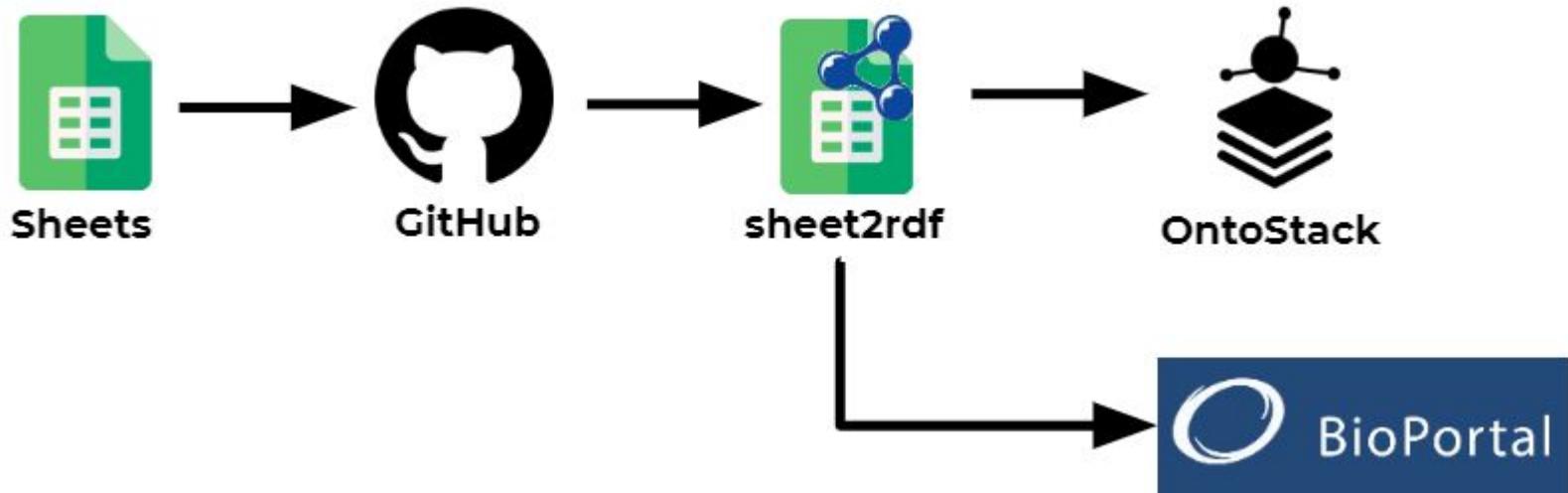


Slide from:

<https://zenodo.org/record/4705676>



## The workflow



<b>ConceptScheme URI</b>	http://ontology.deic.org/cv/vocab-name/	
<b>PREFIX</b>	vocab-name	http://ontology.deic.org/cv/vocab-name/
<b>PREFIX</b>	pav	http://purl.org/pav/
<b>PREFIX</b>	dct	http://purl.org/dc/terms/
<b>PREFIX</b>	owl	http://www.w3.org/2002/07/owl#
<b>PREFIX</b>	xsd	http://www.w3.org/2001/XMLSchema#
<b>PREFIX</b>	skos	http://www.w3.org/2004/02/skos/core#
<b>dct:title</b>		
<b>dct:description</b>		
<b>dct:creator</b>		
<b>dct:rights</b>		
<b>pav:version</b>		
<b>pav:createdOn</b>		
<b>pav:lastUpdatedOn</b>		

Identifier	skos:prefLabel@en	skos:altLabel(separator=" ", "#")
vocab-name:		
vocab-name:		
vocab-name:		
vocab-name:		
vocab-name:		
vocab-name:		
vocab-name:		
vocab-name:		
vocab-name:		
vocab-name:		

### ENVS variables

Content language English

Search

Alphabetical Hierarchy Groups

1  
1-[[2-(2,4-dichlorophenyl)-4-propyl-1,3-dioxolan-2-yl)methyl]-1,2,4-triazole → Propiconazole

6  
60207-90-1 → Propiconazole

A  
Atmospheric Terms  
atmospheric\_pressure

G  
Generic Terms

L  
Latitude  
Longitude

P  
Propiconazole

R  
relative\_humidity

W  
Water Terms  
water\_temperature  
wind\_direction  
wind\_speed

#### Vocabulary information

TITLE	ENVS variables
DESCRIPTION	Quantitative and qualitative and quantitative variables describing ...
CREATOR	https://orcid.org/0000-0002-0329-0192 https://orcid.org/0000-0002-9381-9693 https://orcid.org/0000-0002-9823-589X https://orcid.org/0000-0003-2971-3402
RIGHTS	https://spdx.org/licenses/CC0-1.0
TYPE	http://www.w3.org/2004/02/skos/core#Concept http://www.w3.org/2004/02/skos/core#ConceptScheme
CONCEPT SCHEME	ENVS variables
BELONGS TO VOCABULARY	ENVS variables
URI	http://purl.org/m4m-dk-2/variables/

#### Resource counts by type

Type	Count
------	-------



## ENVS variables

Last updated: September 27, 2021

Summary Classes Properties Notes Mappings Widgets

## Details

Acronym	ENVS_VARIABLES
Visibility	Public
Description	Quantitative and qualitative variables describing environmental state
Status	Alpha
Format	SKOS
Contact	Pedro Carvalho, pedro.carvalho@envs.au.dk

## Metrics

Classes	2
Individuals	12
Properties	0
Maximum depth	0
Maximum number of children	2
Average number of children	2
Classes with a single child	0
Classes with more than 25 children	0
Classes with no definition	2

## Submissions

Version	Released	Uploaded	Downloads
0.1.1 (Parsed, Indexed, Metrics, Annotator, Error Diffs)	09/27/2021	09/27/2021	SKOS   CSV   RDF/XML
0.1.1 (Archived)	09/27/2021	09/27/2021	SKOS

## Views of ENVS\_VARIABLES

No views of ENVS\_VARIABLES available

## Visits

We are still collecting data for ENVS\_VARIABLES

## Projects using ENVS\_VARIABLES

No projects using ENVS\_VARIABLES

Slide taken from:

<https://zenodo.org/record/4705676>

The ontology can be found on our local instance of OntoStack:

<http://ontology.deic.dk/envs-vars/en/>

And publicly on BioPortal:

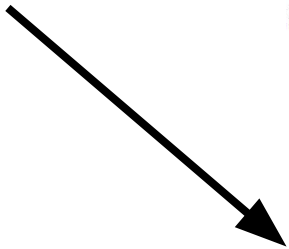
[https://bioportal.bioontology.org/ontologies/ENVS\\_VARIABLES](https://bioportal.bioontology.org/ontologies/ENVS_VARIABLES)





### Definition of terms (optionally properties)

Identifier	skos:prefLabel@en	qudt:unit(separator="," )	skos:altLabel(separator=";")	skos:definition@en	
vars:AtmosphericTerms	Atmospheric Terms			This is a parent concept containing controlled terms representing a	
vars:WaterTerms	Water Terms			This is a parent concept containing controlled terms representing w	
vars:GenericTerms	Generic Terms			This is an overarching concept containing variables that are not fiel	
vars:wind_speed	wind_speed	unit:M-PER-SEC		Speed is the magnitude of velocity. Wind is defined as a two-dimer	
vars:wind_direction	wind_direction	unit:Degree			
vars:Latitude	Latitude			Latitude is positive northward; its units of degree_north (or equivale	
vars:Longitude	Longitude			Longitude is positive eastward; its units of degree_east (or equivale	
vars:relative_humidity	relative_humidity				
vars:atmospheric_pressure	atmospheric_pressure	unit:MilliBAR,unit:HectoPA		Air pressure is the force per unit area which would be exerted when	
vars:water_temperature	water_temperature				
vars:Propiconazole	Propiconazole		60207-90-1;1-[[2-(2,4-dichlorophenyl)-4-propyl-1,3-dioxolan-2-yl]methyl]-1,2,4-triazole		
vars:					
vars:					



```
<?xml version="1.0" encoding="utf-8" ?>  
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:skos="http://www.w3.org/2004/02/skos/core#">  
  
  <skos:Concept rdf:about="http://purl.org/m4m-dk-2/variables/relative_humidity">  
    <skos:prefLabel xml:lang="en">relative_humidity</skos:prefLabel>  
    <skos:broader rdf:resource="http://purl.org/m4m-dk-2/variables/AtmosphericTerms"/>  
  </skos:Concept>  
  
  <skos:Concept rdf:about="http://purl.org/m4m-dk-2/variables/AtmosphericTerms">  
    <skos:topConceptOf rdf:resource="http://purl.org/m4m-dk-2/variables/" />  
    <skos:narrower>  
      <skos:Concept rdf:about="http://purl.org/m4m-dk-2/variables/atmospheric_pressure">  
        <skos:prefLabel xml:lang="en">atmospheric_pressure</skos:prefLabel>  
        <skos:broader rdf:resource="http://purl.org/m4m-dk-2/variables/AtmosphericTerms"/>  
      </skos:Concept>  
    </skos:narrower>  
  
    <skos:narrower rdf:resource="http://purl.org/m4m-dk-2/variables/relative_humidity"/>  
  </skos:narrower>  
  
    <skos:Concept rdf:about="http://purl.org/m4m-dk-2/variables/wind_direction">  
      <skos:prefLabel xml:lang="en">wind_direction</skos:prefLabel>  
      <skos:broader rdf:resource="http://purl.org/m4m-dk-2/variables/AtmosphericTerms"/>  
    </skos:Concept>  
  </skos:narrower>  
</rdf:RDF>
```

## Cedar



CEDAR

Email

|

Password

Remember me

[Forgot Password?](#)

LOG IN

Cedar is a service that enables:

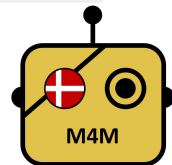
- Creation of and collaboration on metadata templates
- Define metadata for single datasets
- Access and search other's metadata to find useful datasets

“The CEDAR platform provides an easy-to-use solution for creating and reusing FAIR metadata. CEDAR’s metadata modeling, flexible but rigorous semantics, and ability to quickly produce structured metadata makes it perfect for ongoing Metadata 4 Machines workshops and emerging FAIR training courses.”

**Erik Schultes**

International Science Coordinator, GO FAIR International Support and Coordination Office (GFISCO) – GO FAIR

<https://metadatacenter.org/>



Slide from:

[https://zenodo.org/record/4621141#.YUBd\\_Z37SUK](https://zenodo.org/record/4621141#.YUBd_Z37SUK)

## CEDAR artifacts



Metadata template **fields** (e.g., *title, description, start date*)



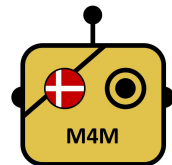
Metadata template **elements** (groups of fields and/or elements) (e.g., *Creator*)



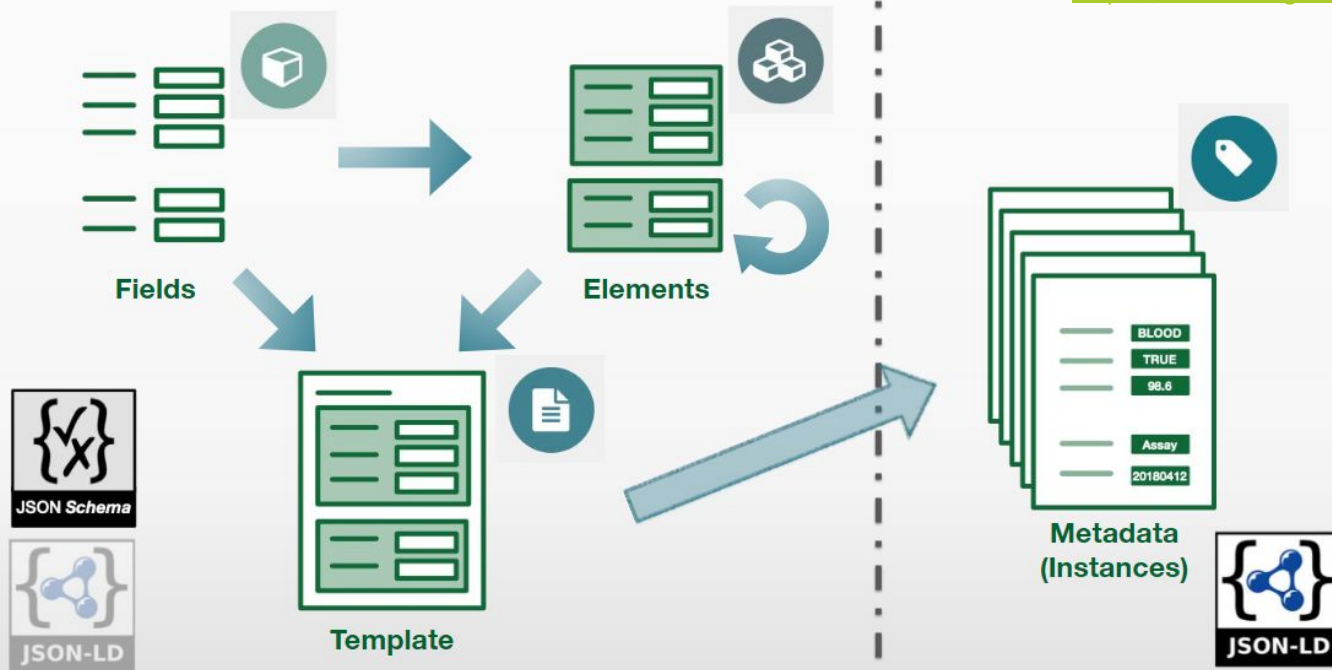
Metadata **templates** (e.g., *GDMT*)



Metadata (aka template **instances**)



## CEDAR's Template Model



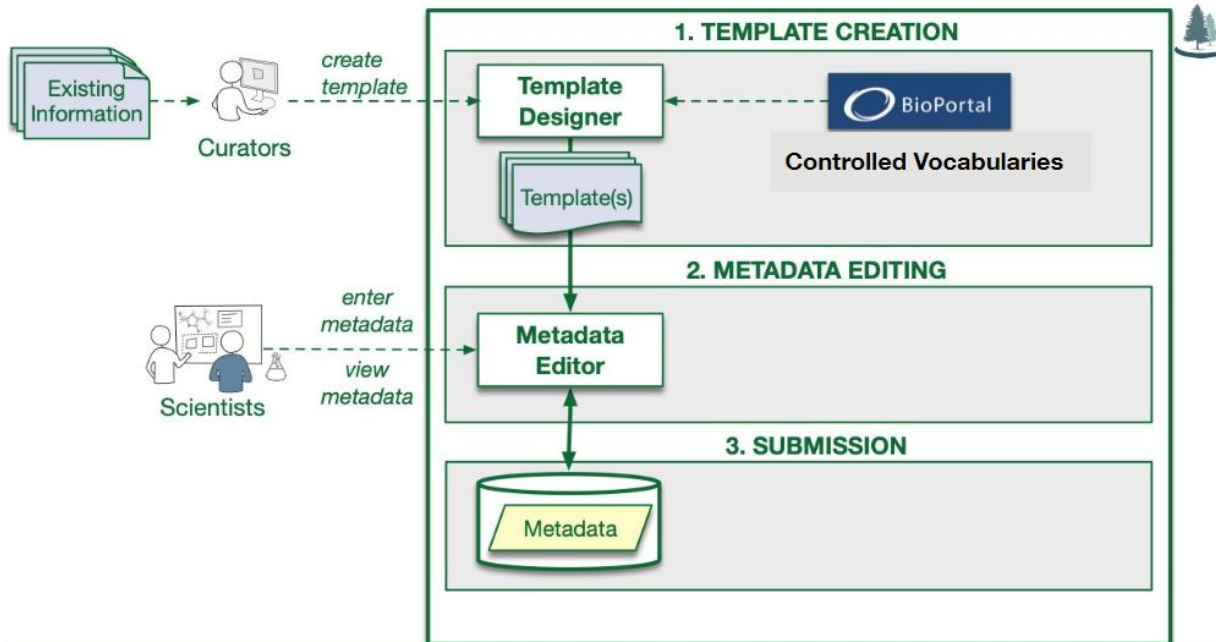
Slide from:

[https://zenodo.org/record/4621141#.YUBd\\_Z37SUk](https://zenodo.org/record/4621141#.YUBd_Z37SUk)

## CEDAR-Only Workflow

Slide from:

[https://zenodo.org/record/4621141#.YUBd\\_Z37SUk](https://zenodo.org/record/4621141#.YUBd_Z37SUk)





1..N

Creator

Creator Type

Creator Name

Creator Given Name

Creator Family Name

Creator Identifier

Creator Identifier Scheme

Creator Affiliation

Creator Affiliation Identifier

Creator Affiliation Identifier Scheme

- A
- 📅
- ✉
- #
- ...
- 🔍



bestrock

Creator (1..N)

Creator Type\*

Person

Creator Name

Creator Name\*

Hannah Mihai

Creator Given Name

Hannah

Creator Family Name

Mihai

Creator Identifier

Creator Identifier\*

https://orcid.org/0000-0002-0454-4289

Creator Identifier Scheme\*

ORCID

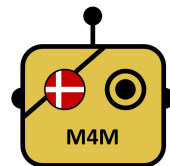
## What happened so far...

- Workshops with 2 research groups have been conducted in 2021
  - One online event with NeiC - PaRI (Nordic Pandemic Research Infrastructure / Bioinformatics)
  - One F2F event with a Danish research group from Aarhus University (Department of Environmental Science)



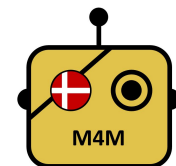
**Pedro Carvalho** 4:05 PM

Thank you [@Nikola Vasiljevic](#) for guiding us through an excellent workshop yesterday, and thank you for all the participants - AU and external - all discussions elevated the work being done. I think this short and focused 5h were great to show what needs to be done, but that it can also be done. There are several steps that need to be taken to fully implement all these tools and a big work to be done in terms of culture in our workplaces on how data and metadata is handled. I know that I will not look at data and metadata the same way as before.



## The way forward

- Workshop with 1 more research group is planned for December 2021
- Concept planning for a Train the Trainers Event
  - Training people (research support staff, data stewards) within the Danish Universities
  - Expanding the reach and flexibility to the benefit of researchers
  - Collaboration with FAIR Data Collective







**Any questions or comments?**

