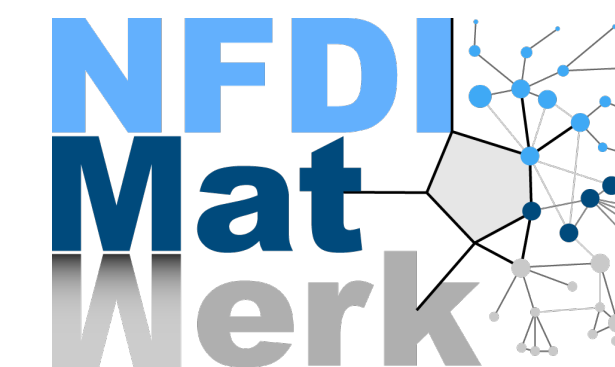




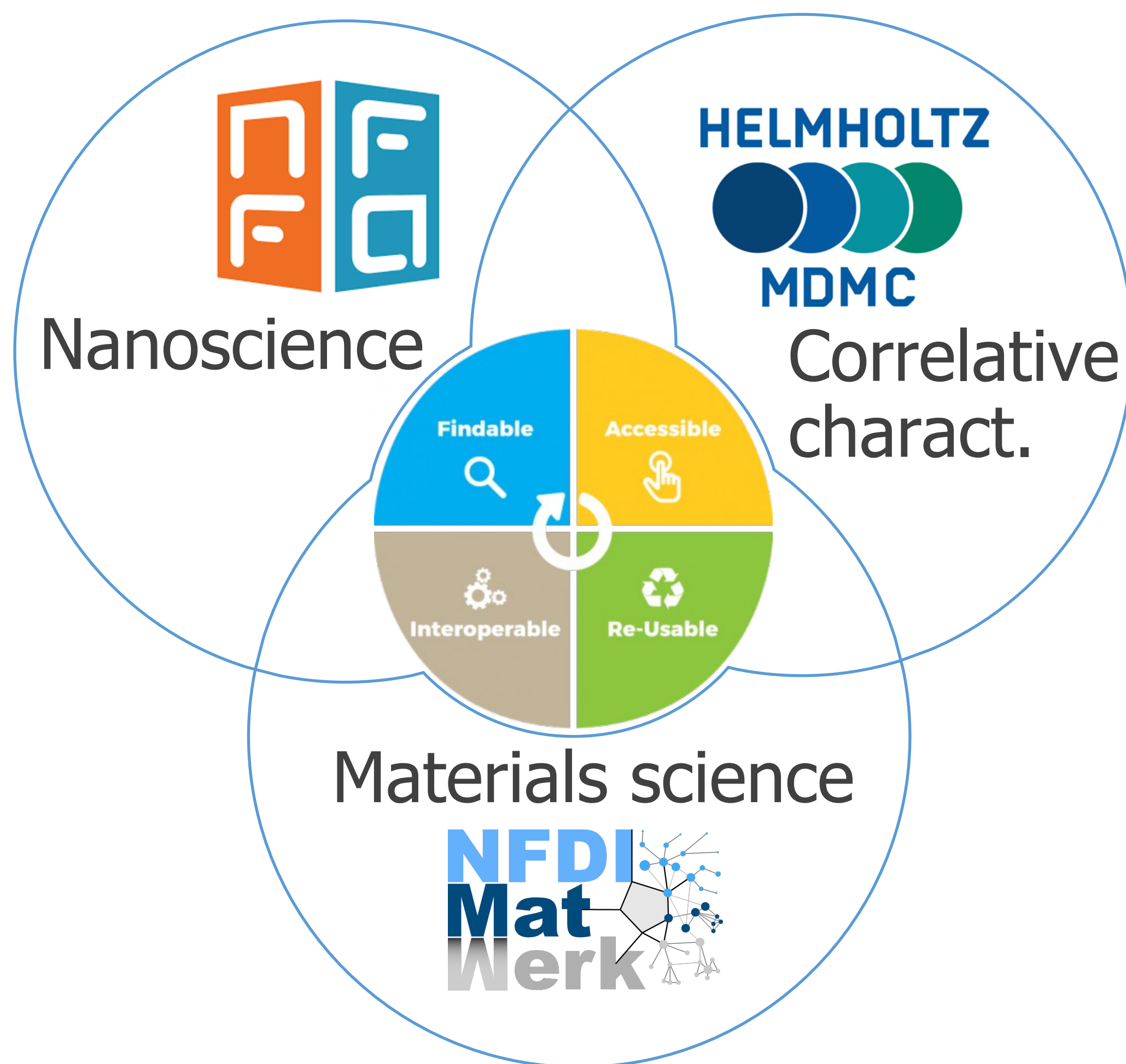
Successful stories in NFFA-Europe Pilot and beyond

26.09.2023 – Heinz Maier-Leibnitz Zentrum (MLZ), Garching bei München, Germany

Dr. Rossella Aversa
Karlsruhe Institute of Technology
NFFA-Europe Pilot – WP9 Leader



It's all about Data Management



Common aim:

- Data management practices following the FAIR principles
- Tools and infrastructure solutions

Outline:

- Ongoing activities and achievements
- Reached in collaboration
- Development guided by requirements

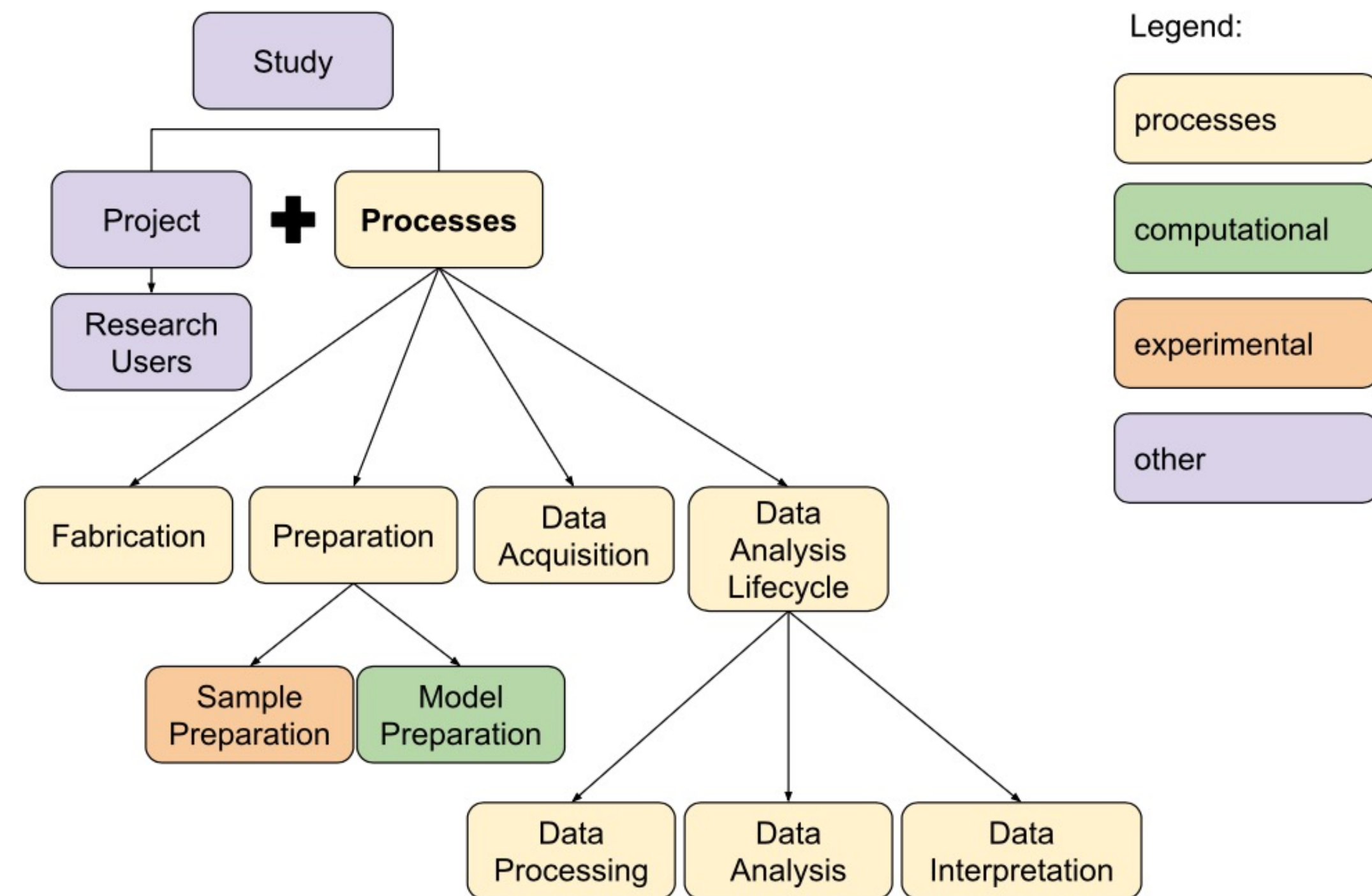
Are we talking about the same thing?

- Glossary

Research User

Person, usually member of a **Project**, who conducts any part of the **Study**, in order to collect and/or analyse **Research Data**, or is interested in reusing **Research Data** by a third party (e.g. **Reference Data**) with the final aim to extract insights that support the answer to some specific research question (i.e., **Conclusions**).

Research Users may be assigned with a role (data curator, instrument scientist, team leader, team member).



<https://jl-mdmc-helmholtz.de/mdmc-activities/metadata-working-group/metadata-wg-topics/semantics/glossary/>

Are we talking about the same thing?

- Glossary
- Vocabulary service

PREFERRED TERM

User Role 

NARROWER CONCEPTS

Data Curator
Instrument Scientist
Team Leader
Team Member

URI

<http://matwerk.datamanager.kit.edu:8001/DemoTerms-1/en/page/userrole> 

DOWNLOAD THIS CONCEPT:

[RDF/XML](#) [TURTLE](#) [JSON-LD](#)

Research User

User Name

Rossella Aversa

User Role

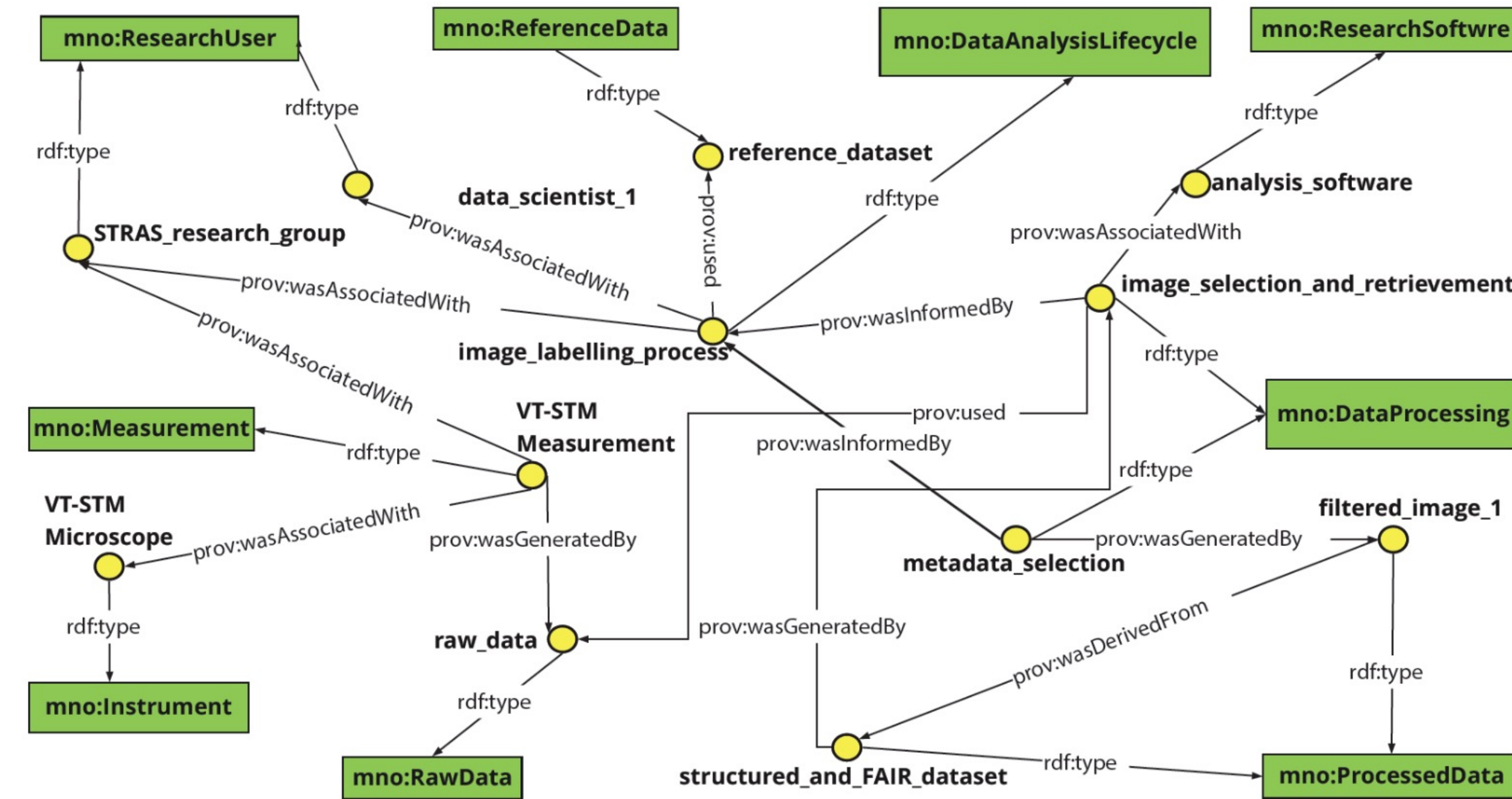
- Data Curator
- Instrument Scientist
- Team Leader
- Team Member

- Terms available online
- Websites, automatic processes, ...
- Centrally maintained

<https://github.com/kat-data-manager/EVOKS>

Are we talking about the same thing?

- Glossary
- Vocabulary service
- Ontology

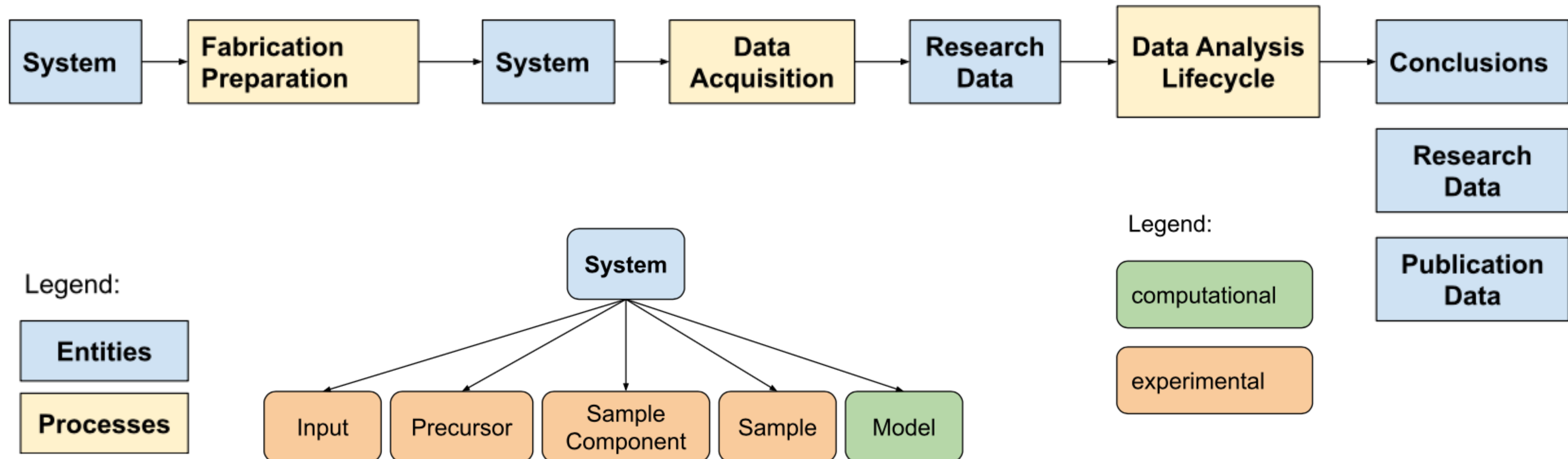


<https://doi.org/10.5281/zenodo.7432171>

<https://doi.org/10.1039/D1NR06485A>

What can I describe?

Every Entity of the Study:



How do I describe my data?

Metadata: descriptive information about data

Metadata Document: structured information about a data resource

Metadata Schema: outline of the overall structure of the metadata (elements, value types, rules, ...)

```

"properties": {
  "sampleName": {
    "type": "string"
  },
  "sampleComponents": {
    "type": "array",
    "items": {
      "type": "object",
      "properties": {
        "sampleComponentName": {
          "type": "string"
        },
        "sampleComponentChemicalFormula": {
          "type": "string"
        }
      }
    }
  },
  "required": ["sampleComponentName"]
},
"samplePreparation": { ...
  }

```

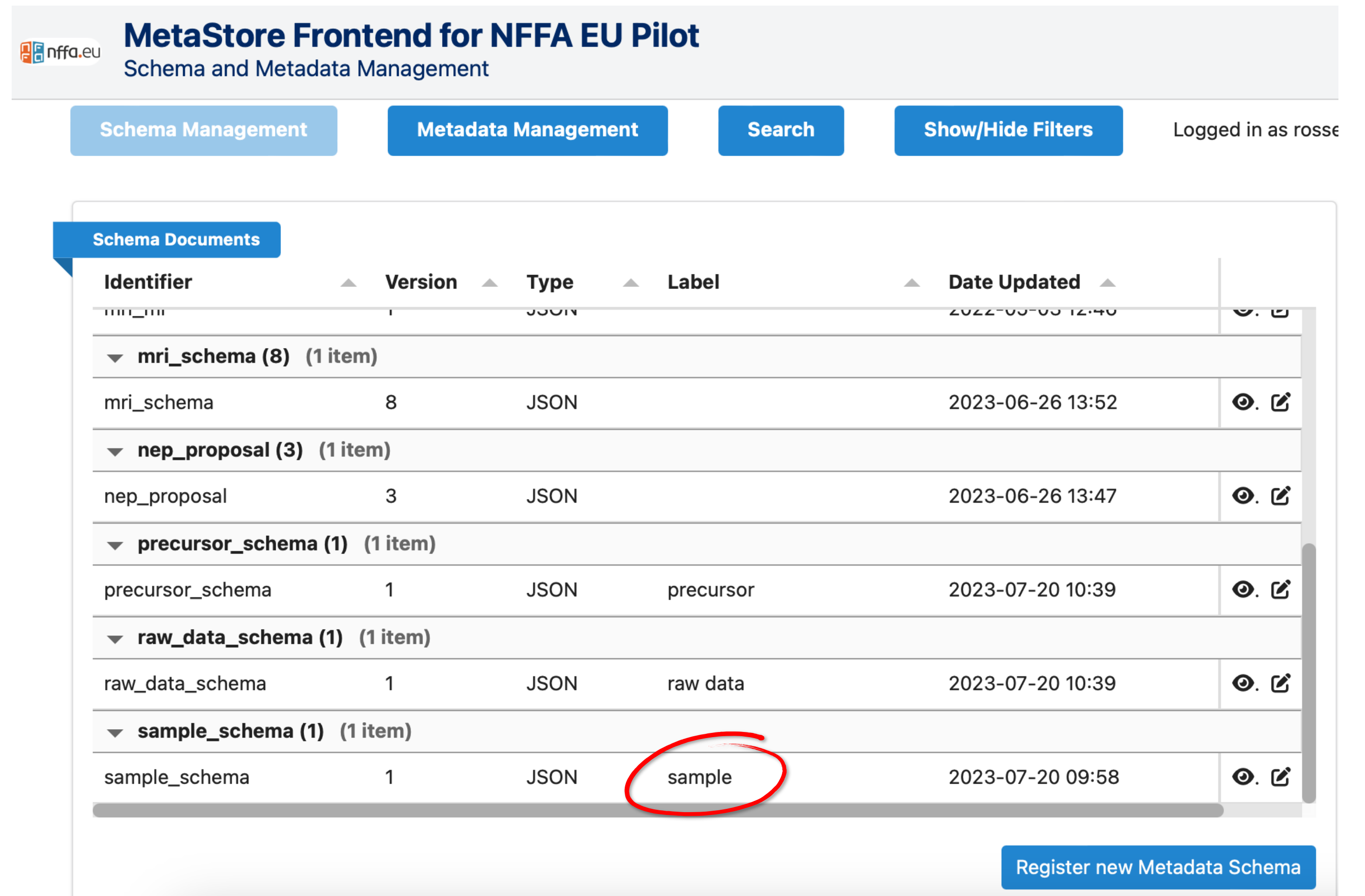
Example: JSON metadata schema for Sample

Where do I find the metadata schemas?

MetaRepo:

- Metadata schema registry
- Metadata repository
- API and GUI
- Virtual Access service

<https://metarepo.nffa.eu>



MetaStore Frontend for NFFA EU Pilot
Schema and Metadata Management

Schema Management Metadata Management Search Show/Hide Filters Logged in as rosse

Schema Documents

Identifier	Version	Type	Label	Date Updated
mri_schema	8	JSON		2023-06-26 13:52
▼ mri_schema (8) (1 item)				
nep_proposal	3	JSON		2023-06-26 13:47
▼ nep_proposal (3) (1 item)				
precursor_schema	1	JSON	precursor	2023-07-20 10:39
▼ precursor_schema (1) (1 item)				
raw_data_schema	1	JSON	raw data	2023-07-20 10:39
▼ raw_data_schema (1) (1 item)				
sample_schema	1	JSON	sample	2023-07-20 09:58
▼ sample_schema (1) (1 item)				

Register new Metadata Schema

Which metadata schemas are available?

- proposal
- (minimal) input
- (minimal) precursor
- (minimal) sample
- (minimal) raw data
- (minimal) analysed data

- SEM
- SEM/FIB tomography
- Nano CT/micro CT
- TEM
- MRI
- STM → see talk by Mirco

How do I write my metadata?

MetaRepo:

- Register metadata documents according to existing schemas
- Validate them

Schema Identifier

The identifier of the validation schema for the provided metadata document

- ✓ sample_schema
- raw_data_schema
- precursor_schema
- nep_proposal
- mri_schema
- mri_ml
- mldata_basic_schema
- input_schema
- datacite
- analysed_data_schema

Metadata Document File

A local file containing the metadata document.

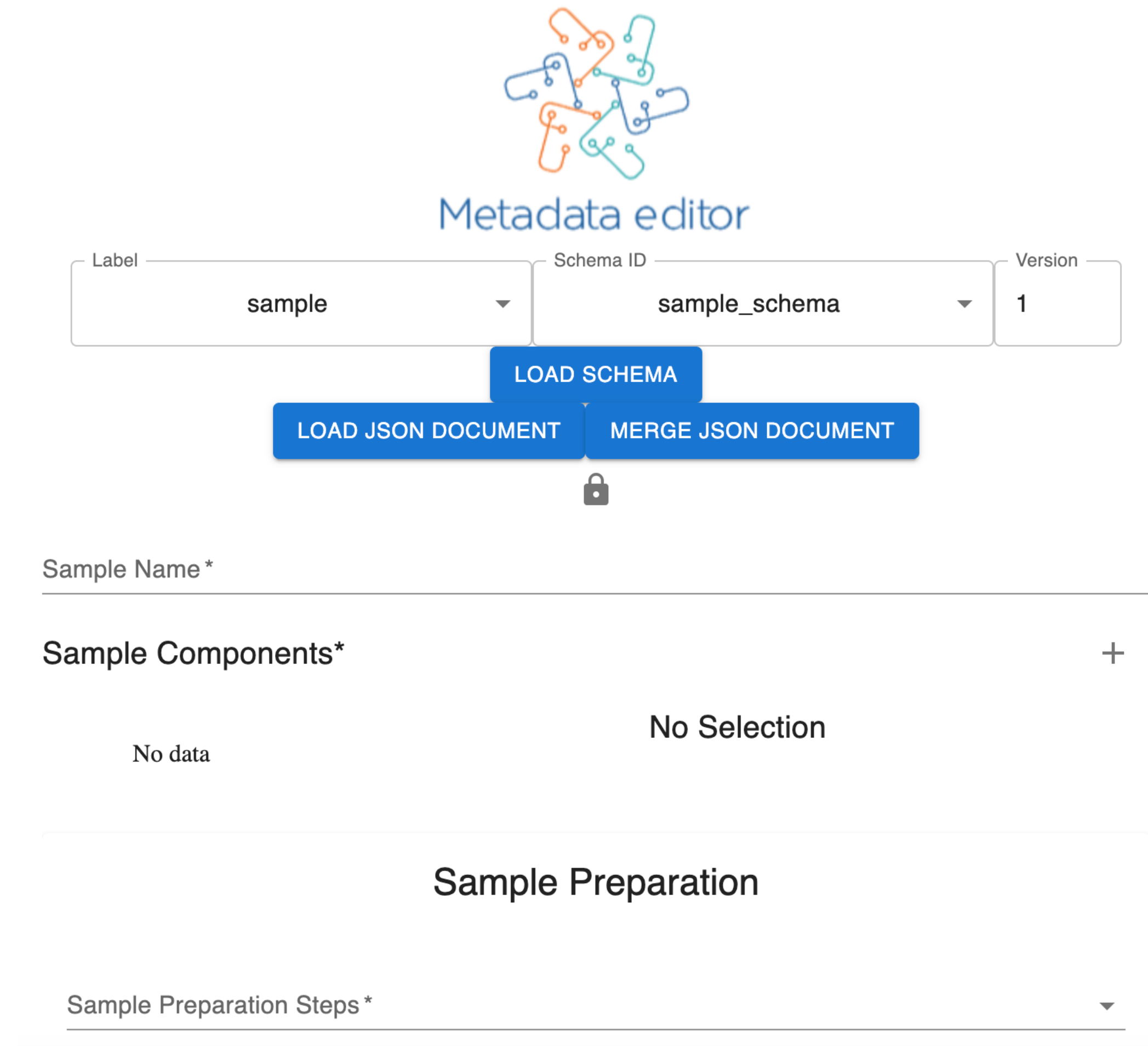
Choose File no file selected

Validate

What if I'm not familiar with JSON?

Metadata editor:

- Edit metadata documents according to schemas
- Register documents via the GUI
- Local application (exposes an API for instruments/ELNs)



The screenshot shows the 'Metadata editor' interface. At the top, there is a logo consisting of several interlocking loops in orange and blue. Below the logo, the title 'Metadata editor' is displayed. The main form contains three dropdown menus: 'Label' with the value 'sample', 'Schema ID' with the value 'sample_schema', and 'Version' with the value '1'. Below these are three blue buttons: 'LOAD SCHEMA', 'LOAD JSON DOCUMENT', and 'MERGE JSON DOCUMENT'. A small lock icon is positioned below the buttons. The form also includes a 'Sample Name*' field, a 'Sample Components*' field with a plus sign and the text 'No Selection', and a 'Sample Preparation Steps*' field. The text 'No data' is visible below the 'Sample Components*' field.

<https://metadata-editor.gitlab.io/documentation/>

Should I fill all metadata by hand?



NFDI-MatWerk Mapping Service

Automatic extraction and mapping of metadata.

Metadata editor (local)

Automatic Metadata Extraction and Mapping tool

- Input: image, file, ...
- Output: metadata document
- New plugins continuously under development

Mapping Component

Choose a suitable mapping from available options

SEM/FIB Tomography Acquisition to TXT

Creates a summary of all metadata extracted from images in a comma delimited txt file. LU: 23.08.2023

Select

SEM/FIB Tomography Acquisition to JSON

Extracts metadata from a SEM/FIB Tomography Acquisition zip file and maps it to the appropriate schema to create a JSON metadata document. LU: 24.08.2023

Select

SEM Zip to txt

Takes a zip folder of arbitrary SEM tiff images and returns a comma separated txt file with a summary of all of the embedded metadata. LU: 05.09.2023

Select

SEM to JSON

This plugin is able to handle a variety of SEM images and processes them using the Hyperspy library. A resulting metadata document in .JSON format is then

Select

Drag & Drop your files or [Browse](#)

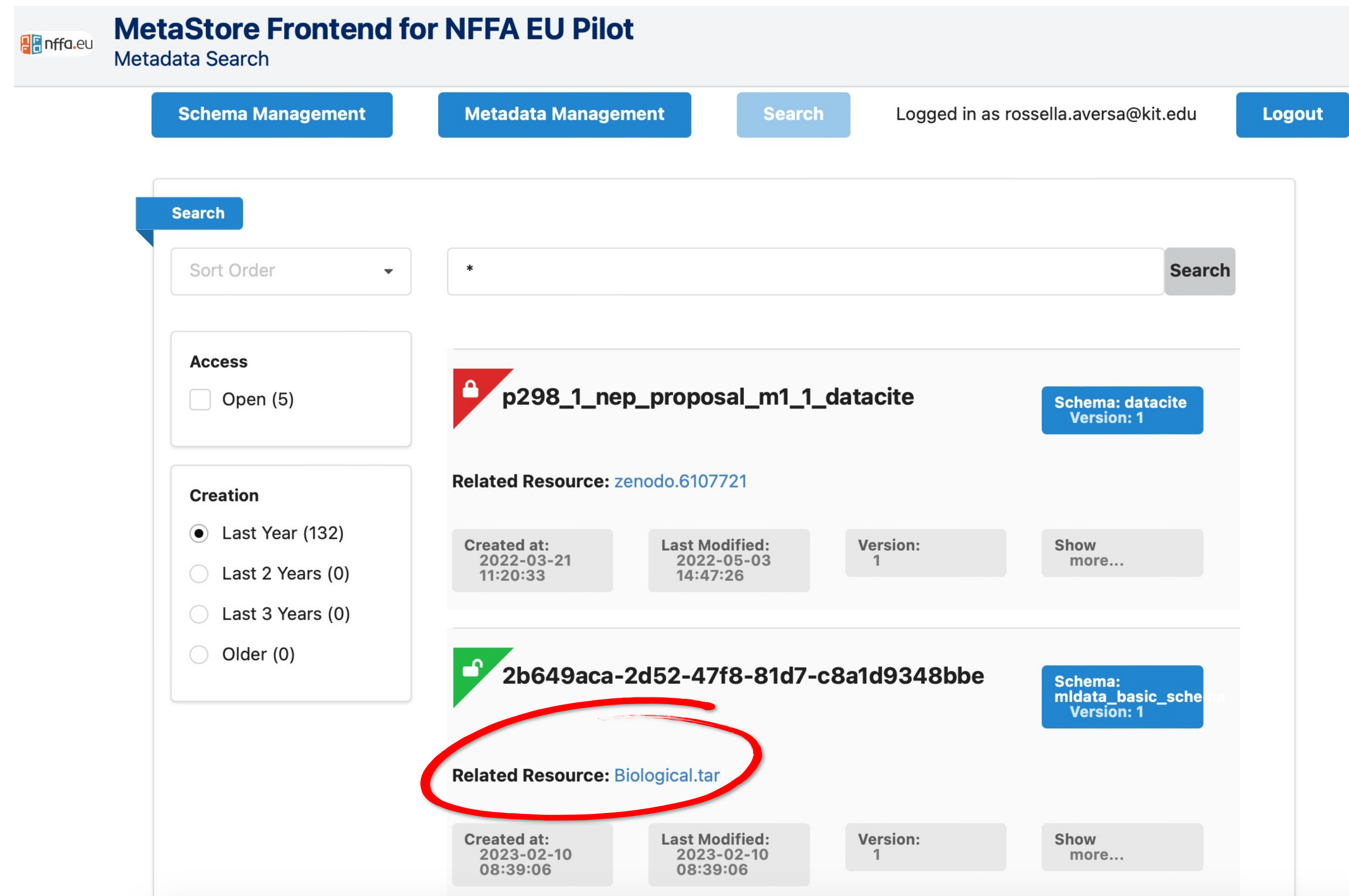
<https://matwerk.datamanager.kit.edu/mapping-service-ui.html>

Map document

How do I search for data?

MetaRepo:

- Use content of metadata documents to search for relevant data
- What is the data about?
Is it useful for my needs?
- Full-text search
- (basic) faceted search
- Private vs Public resources



MetaStore Frontend for NFFA EU Pilot
Metadata Search

Schema Management | Metadata Management | Search | Logged in as rossella.aversa@kit.edu | Logout

Search

Sort Order: [v] | * | Search

Access
 Open (5)

Creation
 Last Year (132)
 Last 2 Years (0)
 Last 3 Years (0)
 Older (0)

Search Results:

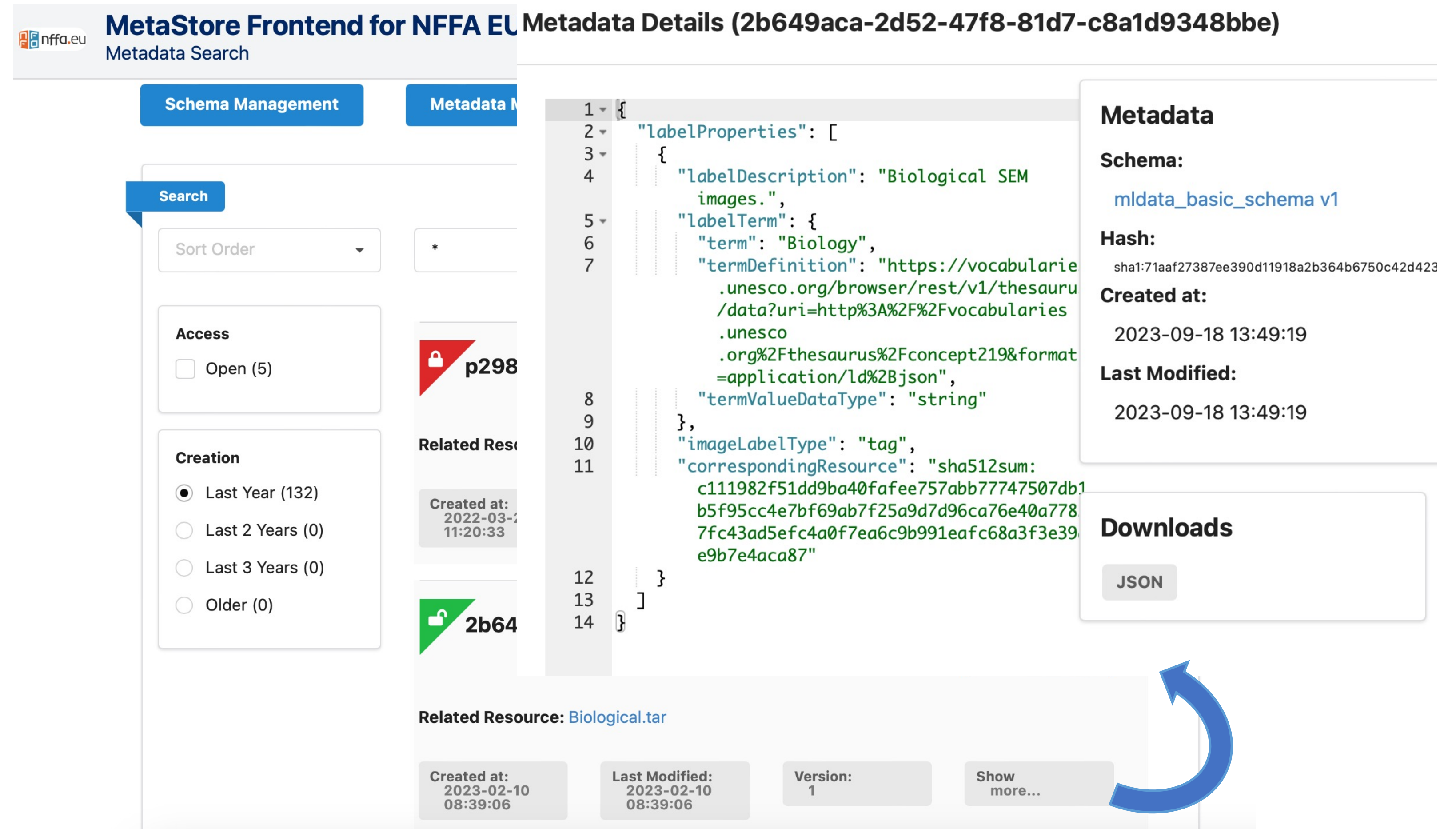
1. p298_1_nep_proposal_m1_1_datacite (Schema: datacite, Version: 1)
 Related Resource: [zenodo.6107721](#)
 Created at: 2022-03-21 11:20:33 | Last Modified: 2022-05-03 14:47:26 | Version: 1 | Show more...

2. 2b649aca-2d52-47f8-81d7-c8a1d9348bbe (Schema: mldata_basic_sche, Version: 1)
 Related Resource: [Biological.tar](#) (circled in red)
 Created at: 2023-02-10 08:39:06 | Last Modified: 2023-02-10 08:39:06 | Version: 1 | Show more...

How do I search for data?

MetaRepo:

- Use content of metadata documents to search for relevant data
- What is the data about?
Is it useful for my needs?
- Full-text search
- (basic) faceted search
- Private vs Public resources



MetaStore Frontend for NFFA EU Metadata Search

Schema Management **Metadata**

Search

Sort Order: [v] * []

Access

Open (5)

Creation

Last Year (132)

Last 2 Years (0)

Last 3 Years (0)

Older (0)

Related Resources

p298

Created at: 2022-03-11 11:20:33

2b64

Related Resource: Biological.tar

Created at: 2023-02-10 08:39:06

Last Modified: 2023-02-10 08:39:06

Version: 1

Show more...

```

1 {
2   "labelProperties": [
3     {
4       "labelDescription": "Biological SEM
5         images.",
6       "labelTerm": {
7         "term": "Biology",
8         "termDefinition": "https://vocabularie
9           .unesco.org/browser/rest/v1/thesauru
10            /data?uri=http%3A%2F%2Fvocabularies
11              .unesco
12                .org%2Fthesaurus%2Fconcept219&format
13                =application/ld%2Bjson",
14         "termValueType": "string"
15       },
16       "imageLabelType": "tag",
17       "correspondingResource": "sha512sum:
18         c111982f51dd9ba40fafee757abb77747507db1
19         b5f95cc4e7bf69ab7f25a9d7d96ca76e40a778
20         7fc43ad5efc4a0f7ea6c9b991eafc68a3f3e39
21         e9b7e4aca87"
22     }
23   ]
24 }
  
```

Metadata

Schema:

[mldata_basic_schema v1](#)

Hash:

sha1:71aaf27387ee390d11918a2b364b6750c42d423

Created at:

2023-09-18 13:49:19

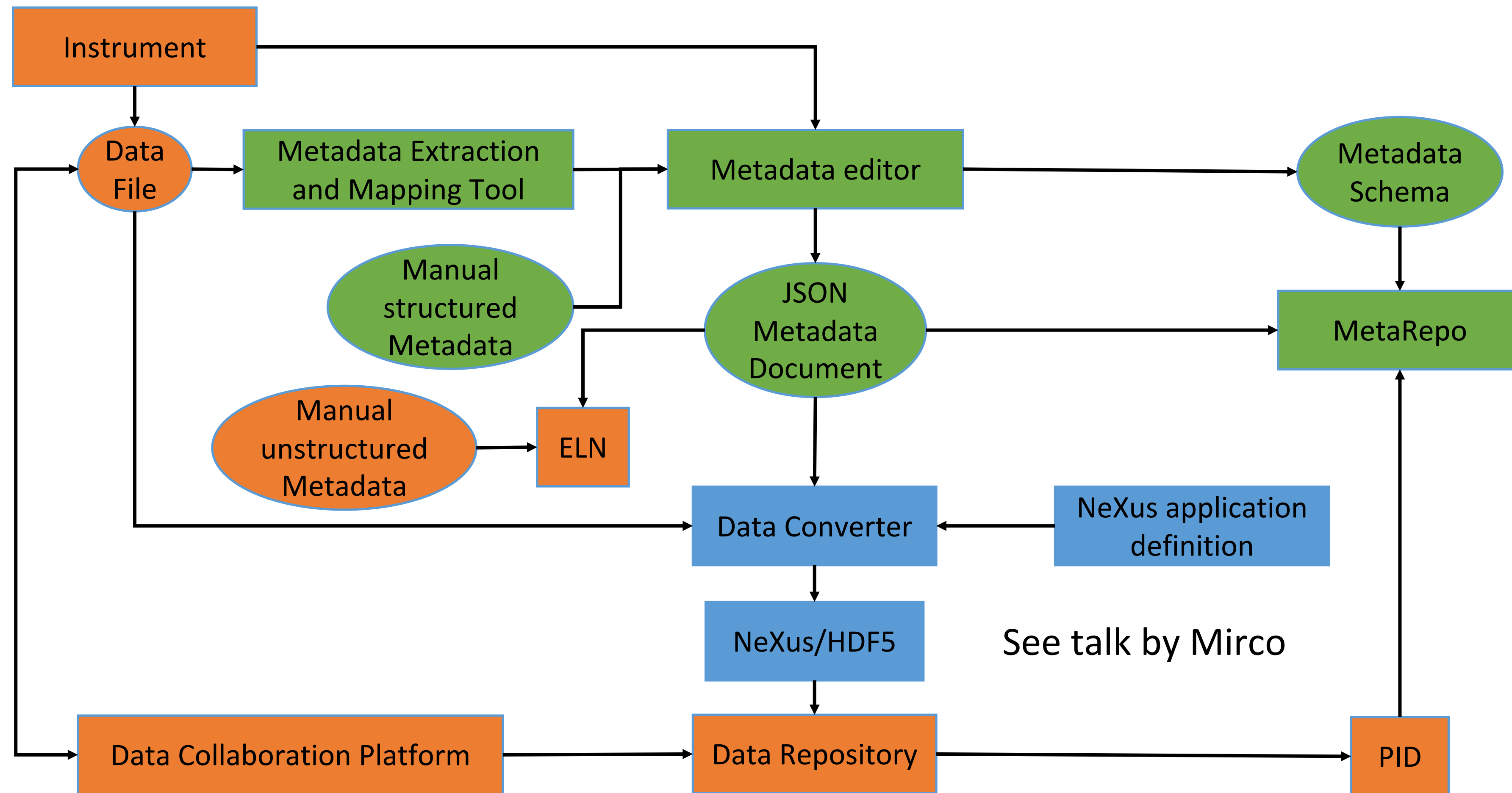
Last Modified:

2023-09-18 13:49:19

Downloads

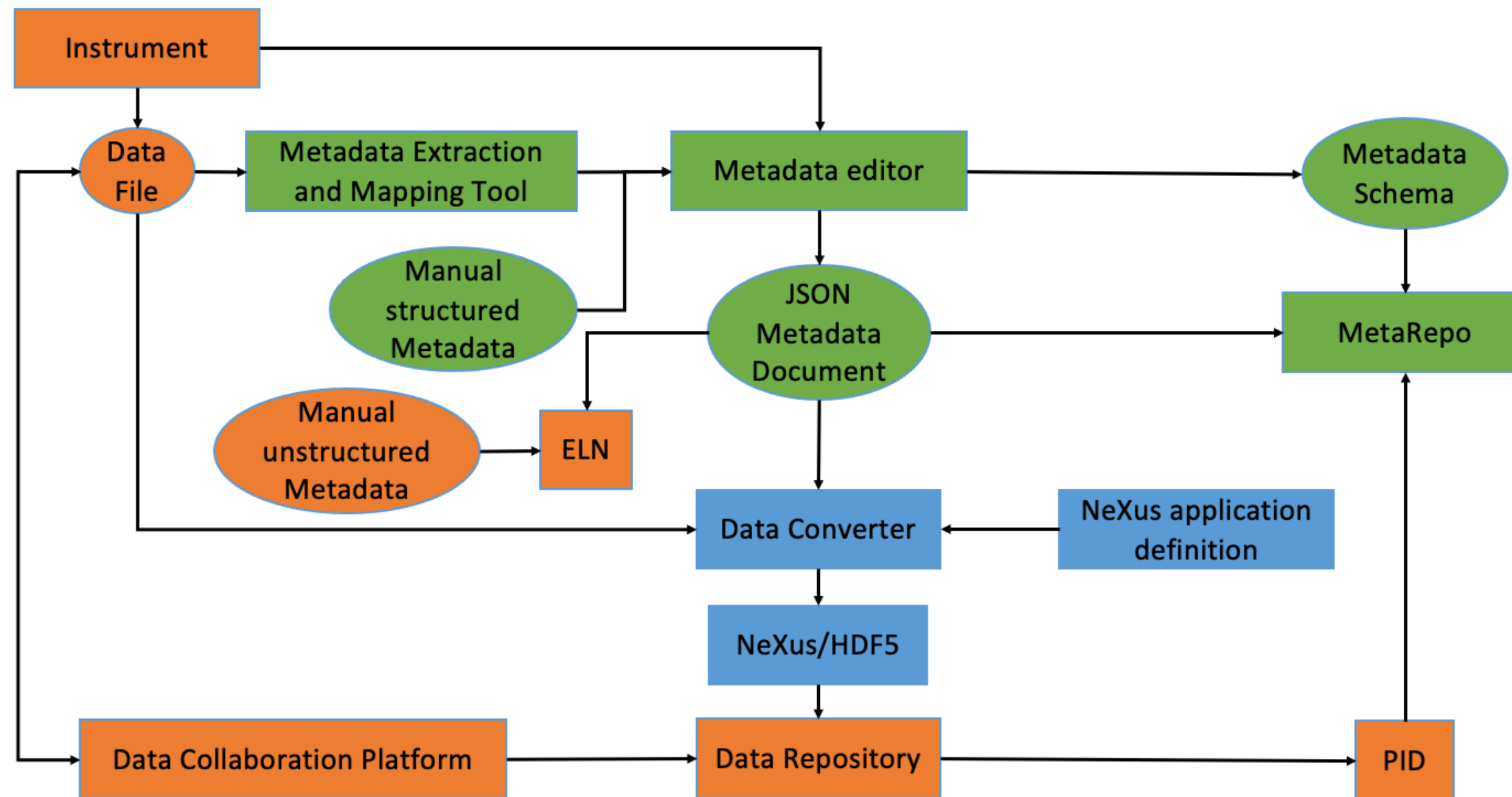
[JSON](#)

Let's put it all together



See talk by Mirco

Closing Remarks



- Tools and services to facilitate (meta)data management
- Framed in a (meta)data management workflow
- Targeted to be useful to researchers
- Applied in daily research life



nffa.eu

PILOT 2021
2026

Contacts

ROSSELLA AVERSA

rossella.aversa@kit.edu

Acknowledgements to: JA6 members, JL-MDMC Metadata WG members, G. Abdildina, N. Blumenröhr, V. Hartmann, M. Inkmann, T. Jejkal, A. Kirar, E. Vitali, and all the colleagues who contributed with tests and consultancy

