

HeFDI Data Talk

Date	Topic	Presenter(s)
06 October 2023	Kadi4Mat – Karlsruhe Data Infrastructure for Materials Science	Dr. Michael Selzer (Institute for Nanotechnology, KIT)



Abstract:

In research today, many digital results are produced from experiments, simulations or measurements. Often, however, only the most important data is stored in the long term, for example if it is part of a scientific publication. The Kadi4Mat software creates technical possibilities to store as much of the accumulating data as possible in such a way that a benefit can be drawn from it. On the one hand, it is particularly important to record the data simply and as automatically as possible, but on the other hand, it is also important to store the information correctly. Kadi4Mat is needed to store information in a meaningful way and to make it available in the long term. Only in this way the accumulated data can be efficiently searched, analysed, visualised and compared with each other in order to maximize the information gain from scientific work.

About the HeFDI Data Talks:

The HeFDI Data Talks are a bi-weekly open information and discussion event focused on data management in the context of science, in which relevant NFDI consortia as well as research data management services present themselves. The series discusses current topics and presents numerous – including local and regional – tools and services. The HeFDI Data Talks are an offer of the HeFDI Initiative (Landesinitiative HeFDI), which is funded by Hesse's Ministry for Science and Arts (HMWK).

DOI-Link: <https://doi.org/10.5281/zenodo.8424794>; Lizenzinformation: Creative Commons Attribution 4.0 International ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/))



Kadi4Mat: General Overview and Sample Applications

Michael Selzer et al.

6 October 2023



Motivation

- In materials science, understanding new materials is becoming increasingly complex
- Without appropriate **data science methods**, it will no longer be possible to manage the ever-growing volumes of heterogeneous data from simulations and experiments
- An important aspect to be able to perform corresponding data analyses smoothly is the structured storage of research (meta)data with the help of a suitable **research data infrastructure**

Research Data Infrastructure

- The development of such an infrastructure is being pursued as part of several research projects, including **FestBatt 2** [1], **POLiS** [2], **AQua** [3], **NFDI4Ing** [4], and others
- Some of the most important objectives of the infrastructure are:
 - Structured **data storage** and web based **data exchange and publication** via different interfaces
 - Integrating heterogeneous and reproducible **workflows** to make use of application-specific tools in an automated and documentable manner
 - Allowing other institutions to set up their **own installations** with the possibility to **connect** multiple of such instances

[1] <https://festbatt.net/>

[2] <https://www.postlithiumstorage.org/>

[3] <https://www.iam.kit.edu/mms/5562.php>

[4] <https://nfdi4ing.de/>

Implementation

- The concept for the research data infrastructure is a modular and generic architecture that combines the two components **electronic lab notebook (ELN)** and **repository**



- There are already several **open source** solutions for both components, but none that combines the desired features of both components
- The goal is to combine **established technologies** and systems with **new concepts**



Karlsruhe Data Infrastructure for Materials Science

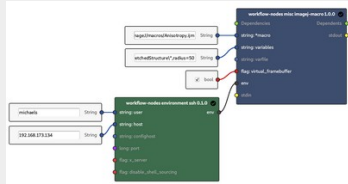
- **Kadi4Mat [1] is the Karlsruhe Data Infrastructure for Materials Science**, a generic open source [2] software under active development for managing research data
- The software as a whole can best be described as a **virtual research environment**

[1] Brandt et al., 2021. Kadi4Mat: A Research Data Infrastructure for Materials Science. Data Science Journal, 20(1), p.8. DOI: <http://doi.org/10.5334/dsj-2021-008>

[2] <https://gitlab.com/iam-cms/kadi>

The Kadi eco-system

KadiStudio Workflows



KadiAPY (Python) API

```

kadiuser@kadi:~$ kadi-apy --help
Usage: kadi-apy [OPTIONS] COMMAND [ARGS]...

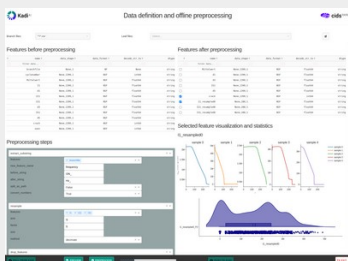
The kadi-apy command line interface.

Options:
  --version  Print the version of this group and exit.
  --commands Print a list of all commands this group contains and exit.
  --help    Show this message and exit.

Commands:
  collections  Commands to manage collections.
  config      Commands to manage configurations.
  group       Commands to manage groups.
  idtool       Commands to manage idtool licenses.
  records     Commands to manage records.
  templates   Commands to manage templates.
  users       Commands to manage users.

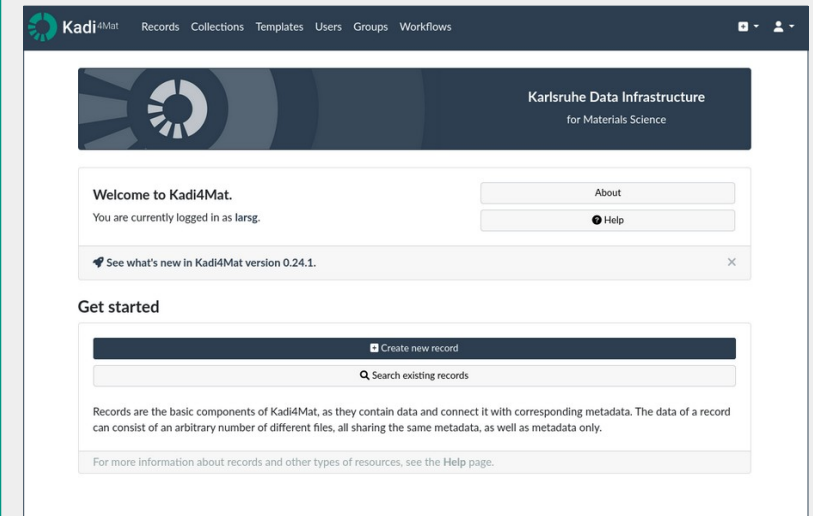
kadiuser@kadi:~$ kadi-apy records create -i kadi-test-record -t Kadi-Test-Record
Successfully created record 'kadi-test-record' (id: '99', identifier: 'kadi-test-record').
kadiuser@kadi:~$ kadi-apy records add-files -r kadi-test-record -s text.txt
Successfully added files to record 'kadi-test-record'.
kadiuser@kadi:~$ kadi-apy records add-metadata -r kadi-test-record -m temperature -f float -u 3.8 -u 8
Successfully added metadata 'temperature' with the value '3.8' and the unit 'K' to record 'kadi-test-record'.
kadiuser@kadi:~$
  
```

KadiAI ML Interface

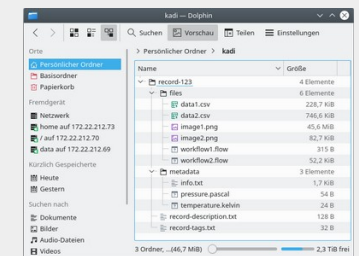


Streamline
your
research

KadiWeb



KadiFS Filesystem



Kadi4Mat in the Research Data Lifecycle

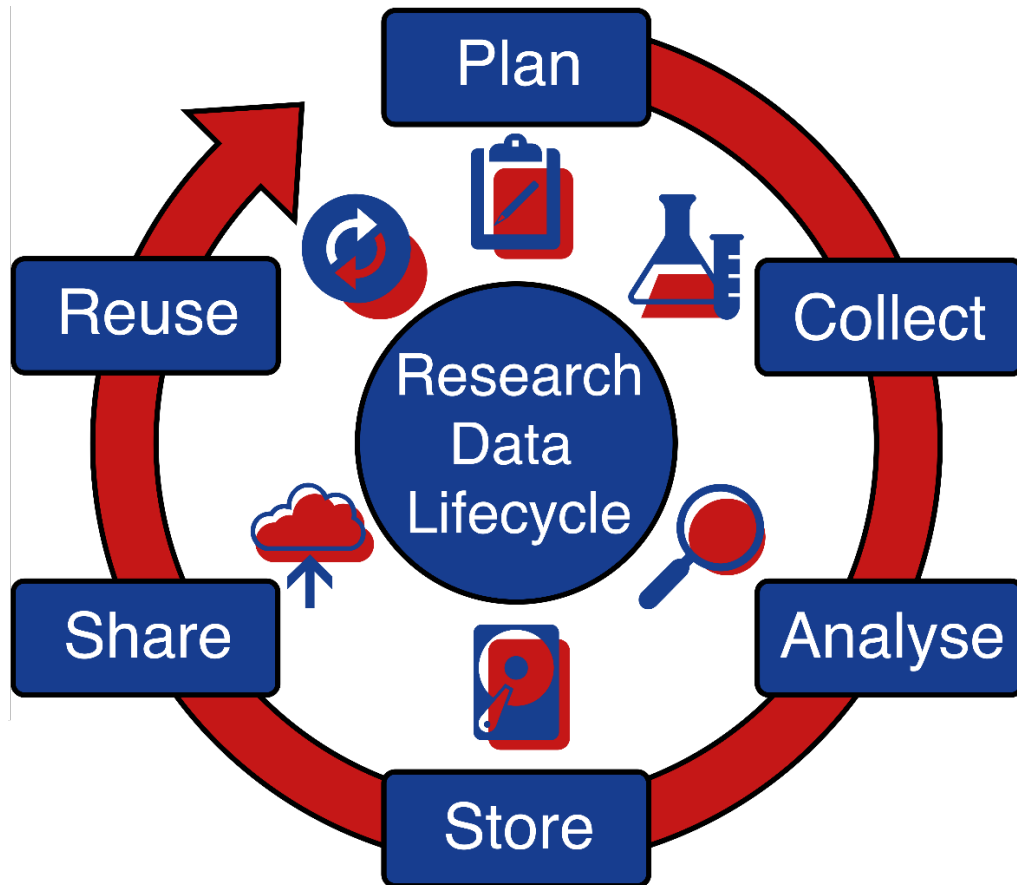


Image: Ephraim Schoof (HIU, Germany)

ELN → “ELN 2.0”

- Focused on automation and reproducibility

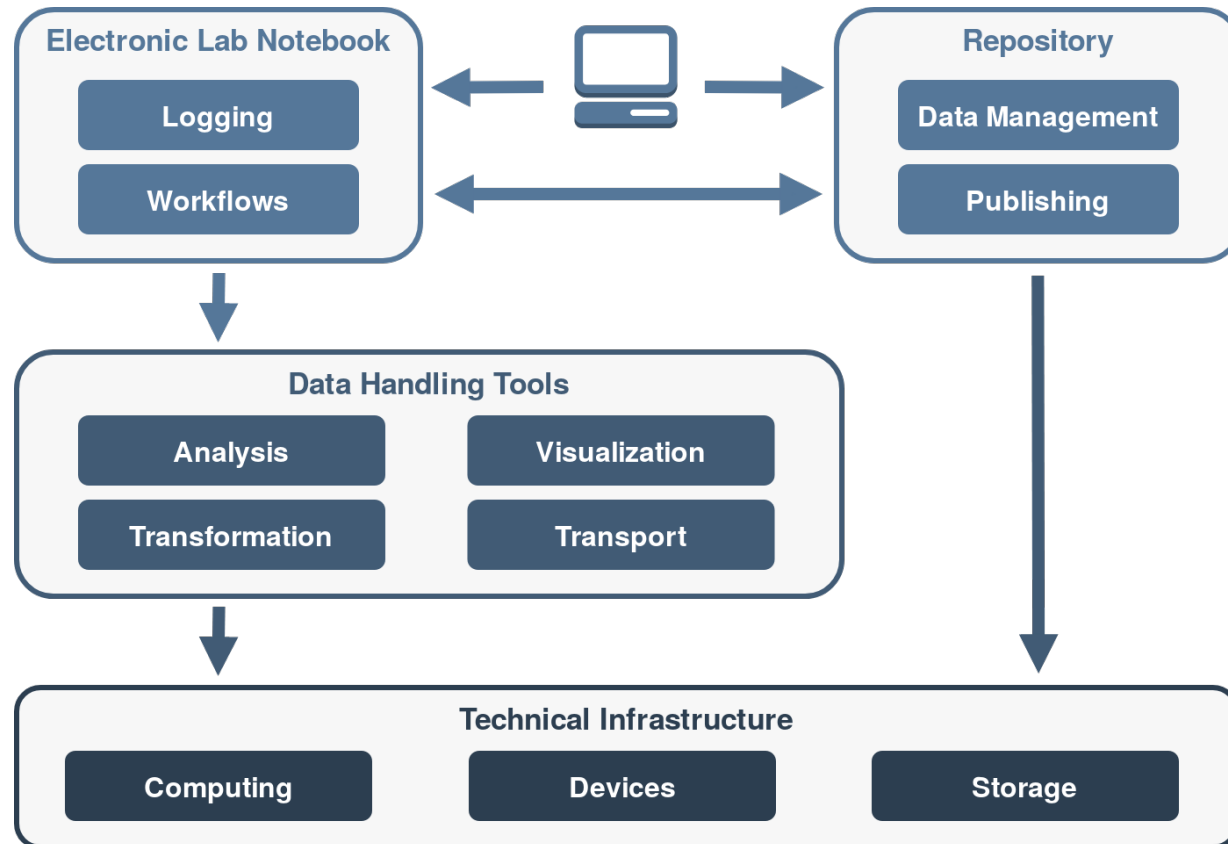
Repository → “Community repository”

- Focused on managing and sharing “warm data” within a community
- Other steps, like **publishing data** (e.g. via Zenodo), can be offered by integrating with existing software as much as possible



Image: <https://zenodo.org>; CC BY 4.0

Conceptual Overview of Kadi4Mat



Electronic Lab Notebook

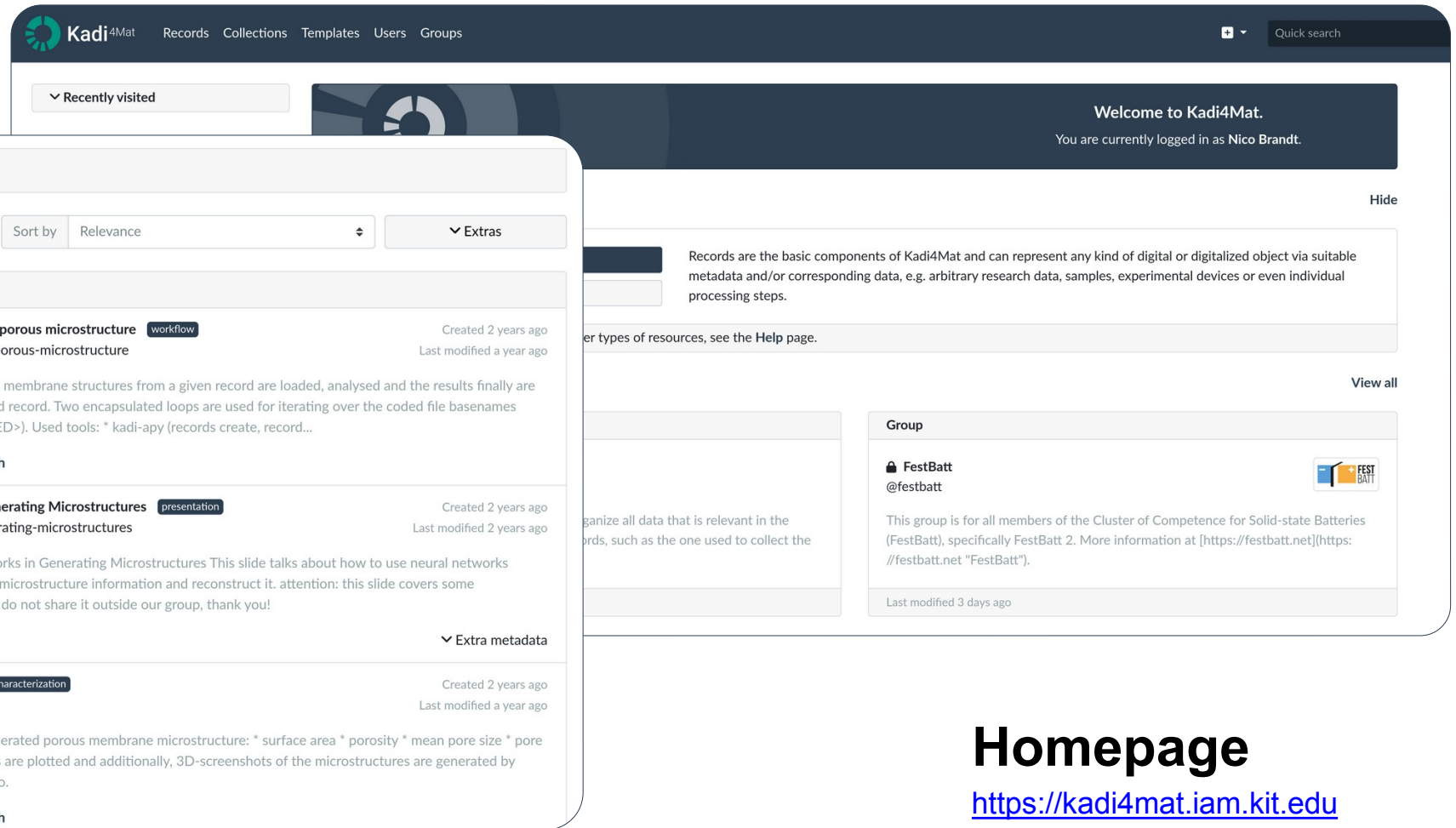
- Documentation and data provenance
- Heterogeneous and reproducible workflows using a compendium of data handling tools

Repository

- Structured data storage and exchange with user-defined access permissions
- Publishing data

Web-Interface of Kadi4Mat

Searching resources



The screenshot displays the Kadi4Mat web interface. At the top, there is a navigation bar with the Kadi4Mat logo and menu items: Records, Collections, Templates, Users, Groups. A search bar on the right contains the text "Quick search". Below the navigation bar, a "Welcome to Kadi4Mat" message is visible, stating "You are currently logged in as Nico Brandt." The main content area shows search results for the query "microstructure". The search bar includes a search icon, a "Search" button, and a "Sort by" dropdown menu set to "Relevance". A "Create new record" button is located on the left side of the search results. The search results list three items:

- Workflow for analysing porous microstructure** (workflow) - Created 2 years ago, Last modified a year ago. Description: "With this workflow, porous membrane structures from a given record are loaded, analysed and the results finally are uploaded in a newly created record. Two encapsulated loops are used for iterating over the coded file basenames (membrane_<PORO>_<SEED>). Used tools: * kadi-apy (records create, record...)"
- Neural Networks in Generating Microstructures** (presentation) - Created 2 years ago, Last modified 2 years ago. Description: "PhD Seminar:Neural Networks in Generating Microstructures This slide talks about how to use neural networks (especially VAE) to capture microstructure information and reconstruct it. attention: this slide covers some unpublished results, please do not share it outside our group, thank you!"
- Analysed-Structures** (characterization) - Created 2 years ago, Last modified a year ago. Description: "Extracted properties of generated porous membrane microstructure: * surface area * porosity * mean pore size * pore size distribution The results are plotted and additionally, 3D-screenshots of the microstructures are generated by executing a paraview-macro."

On the right side of the interface, there is a "Group" section for "FestBatt @festbatt". The description for this group is: "This group is for all members of the Cluster of Competence for Solid-state Batteries (FestBatt), specifically FestBatt 2. More information at https://festbatt.net [https://festbatt.net "FestBatt"]". The group was last modified 3 days ago.

Homepage

<https://kadi4mat.iam.kit.edu>

<https://demo-kadi4mat.iam.kit.edu/>

Basic Structure of Kadi4Mat



Kadi⁴Mat

Records

Collections

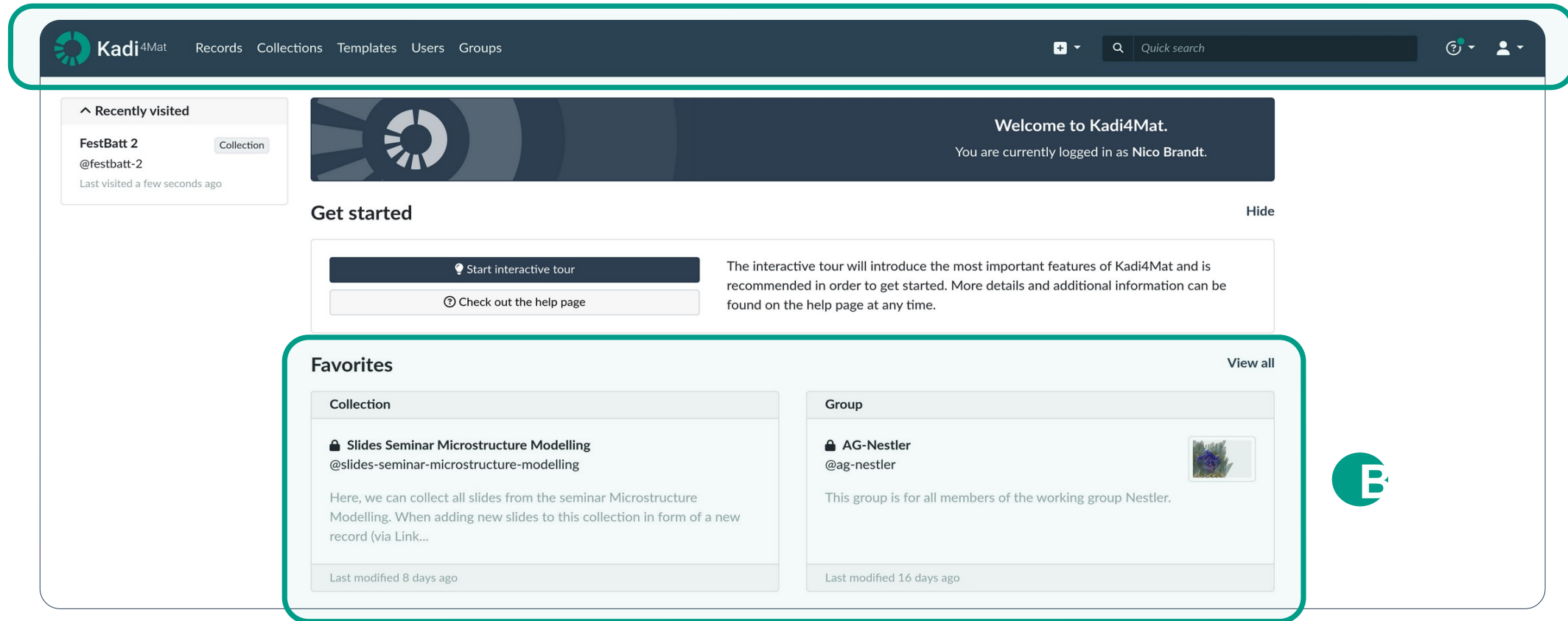
Templates

Users

Groups

- **Records:** Combine data with generic, user-defined metadata
- **Collections:** Group multiple records and/or other collections
- **Templates:** Offer blueprints for records and generic metadata
- **Groups:** Group multiple users to ease access permission management

Basic Structure of Kadi4Mat



Kadi4Mat Records Collections Templates Users Groups

Quick search

Recently visited

FestBatt 2
@festbatt-2
Collection
Last visited a few seconds ago

Welcome to Kadi4Mat.
You are currently logged in as Nico Brandt.

Get started Hide

Start interactive tour

Check out the help page

The interactive tour will introduce the most important features of Kadi4Mat and is recommended in order to get started. More details and additional information can be found on the help page at any time.

Favorites View all

Collection

Slides Seminar Microstructure Modelling
@slides-seminar-microstructure-modelling

Here, we can collect all slides from the seminar Microstructure Modelling. When adding new slides to this collection in form of a new record (via Link...

Last modified 8 days ago

Group

AG-Nestler
@ag-nestler

This group is for all members of the working group Nestler.

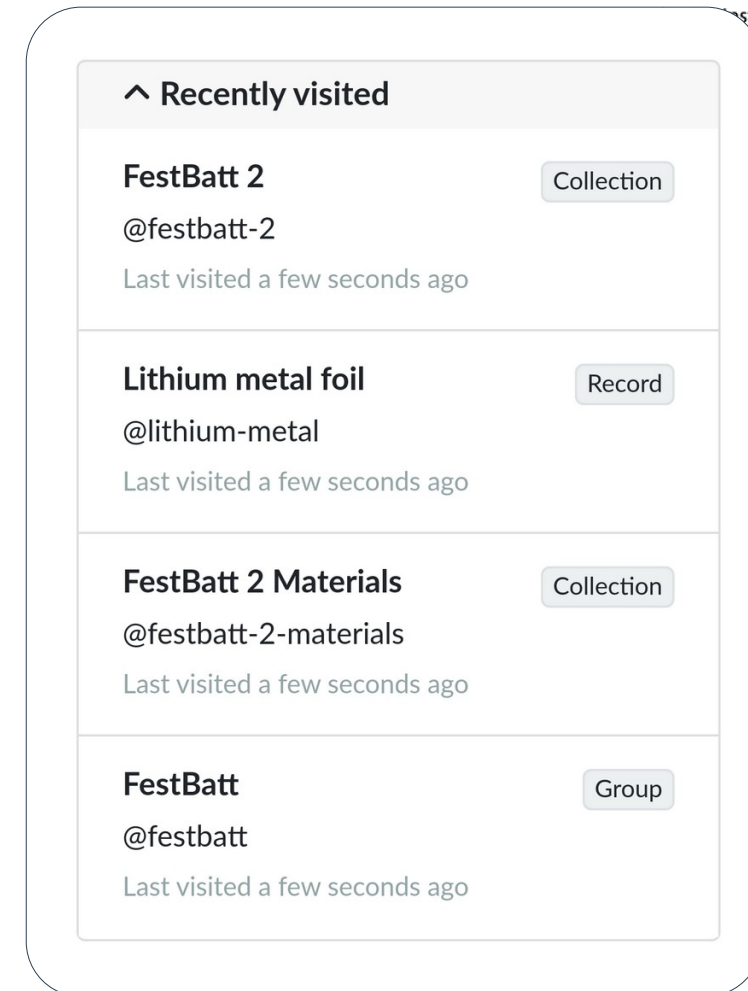
Last modified 16 days ago

(A) Navigation bar with various links and menus

(B) Favorite resources and latest updates (customizable)


New Feature of Kadi4Mat: Recently Visited Resources

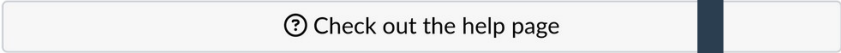
- The five most recently visited resources are shown in a new overview on the left next to the main content
- This should aid with e.g. navigating sub-collection hierarchies



New Feature of Kadi4Mat: Interactive Tour

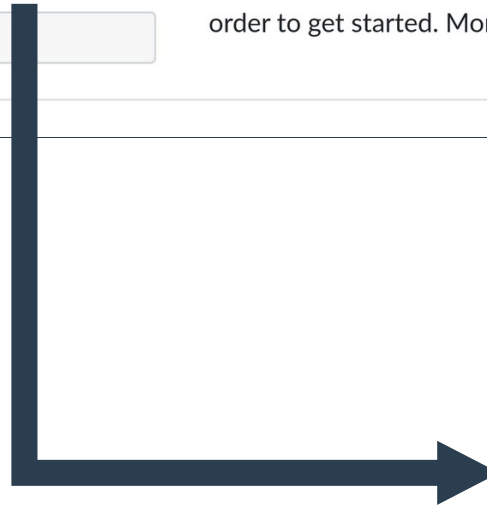
Get started Hide

 Start interactive tour

 Check out the help page

The interactive tour will introduce the most important features of Kadi4Mat and is recommended in order to get started. More details and additional information can be found on the help page at any time.

- The first interactive tour introduces the basic functionality of Kadi4Mat with focus on records



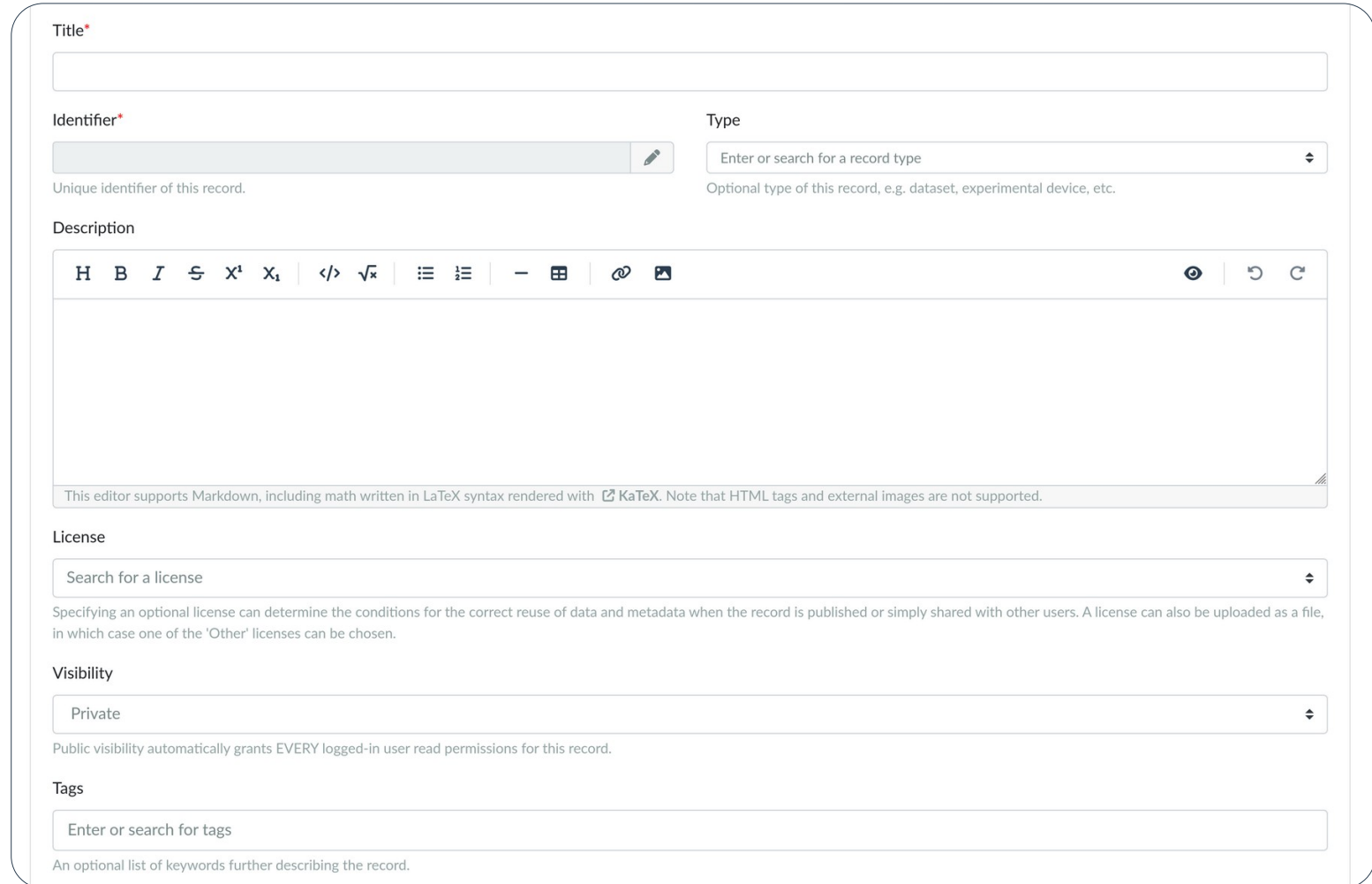
Introduction ✕

This interactive tour will introduce you to the most important features of Kadi4Mat. If you want to pause the tour, you can simply close it and continue it later on this page by clicking on *Continue tour*.

Next →

Basic Feature of Kadi4Mat: Creating new Records

- **Basic metadata schema for each record**
 - Title (required)
 - Unique identifier (required)
 - Type (e.g. *dataset*, *device*, ...)
 - Description (supports Markdown syntax and LaTeX formulas)
 - License
 - Visibility (*private* or *public*)
 - Tags (a.k.a. keywords)

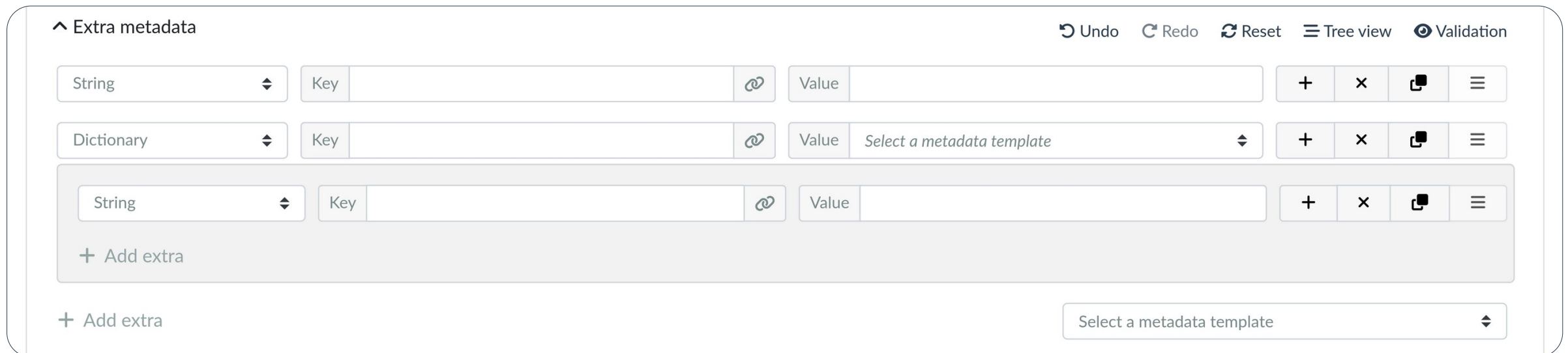


The screenshot shows a web form for creating a new record. It includes the following fields and sections:

- Title***: A text input field.
- Identifier***: A text input field with a unique identifier description below it.
- Type**: A dropdown menu with the placeholder text "Enter or search for a record type" and a description: "Optional type of this record, e.g. dataset, experimental device, etc."
- Description**: A rich text editor with a toolbar containing icons for bold, italic, strikethrough, subscript, superscript, code, link, unlink, list, and image. A note at the bottom states: "This editor supports Markdown, including math written in LaTeX syntax rendered with [KaTeX](#). Note that HTML tags and external images are not supported."
- License**: A dropdown menu with the placeholder text "Search for a license" and a description: "Specifying an optional license can determine the conditions for the correct reuse of data and metadata when the record is published or simply shared with other users. A license can also be uploaded as a file, in which case one of the 'Other' licenses can be chosen."
- Visibility**: A dropdown menu with the selected option "Private" and a description: "Public visibility automatically grants EVERY logged-in user read permissions for this record."
- Tags**: A text input field with the placeholder text "Enter or search for tags" and a description: "An optional list of keywords further describing the record."

Basic Feature of Kadi4Mat: Creating new Records Karlsruhe Institute of Technology

- **Generic metadata (schema free)**
 - To specify user- and application-specific metadata, optionally based on existing schemes
 - Each metadatum consists of a key-value pair of a certain type, allowing for either literals or nested values
 - **Literal values:** String, Integer, Float, Boolean (*true* or *false*), Date
 - **Nested values:** Dictionary, List (uses *indices* instead of *keys*)



^ Extra metadata Undo Redo Reset Tree view Validation

String	Key	Value	+	x	Copy	Menu
Dictionary	Key	Value <i>Select a metadata template</i>	+	x	Copy	Menu
String	Key	Value	+	x	Copy	Menu

+ Add extra

+ Add extra Select a metadata template

New Feature of Kadi4Mat: Permissions Editor

Permissions Roles

Type	User		User	@fhartmann		Role	Editor		+	×
Type	Group		Group	@festbatt		Role	Collaborator		+	×
Type	User		User	<i>Search for users</i>		Role	Member		+	×

Directly add user or group roles to this record.

- Permissions for users or groups can now be specified directly when creating new resources
- The same is possible within record templates, so the editor can be prefilled when creating a corresponding record

New Feature of Kadi4Mat: Record Link Editor

Record links 👁 Link preview

New record created → @test

Direction	Outgoing ▾	Record	@test ✕ ▾	Name	created ✕ ▾	🔗	+	✕
-----------	------------	--------	-----------	------	-------------	---	---	---

New record ← created by @test

Direction	Incoming ▾	Record	@test ✕ ▾	Name	created by ✕ ▾	🔗	+	✕
-----------	------------	--------	-----------	------	----------------	---	---	---

Directly link this record with one or more other records.

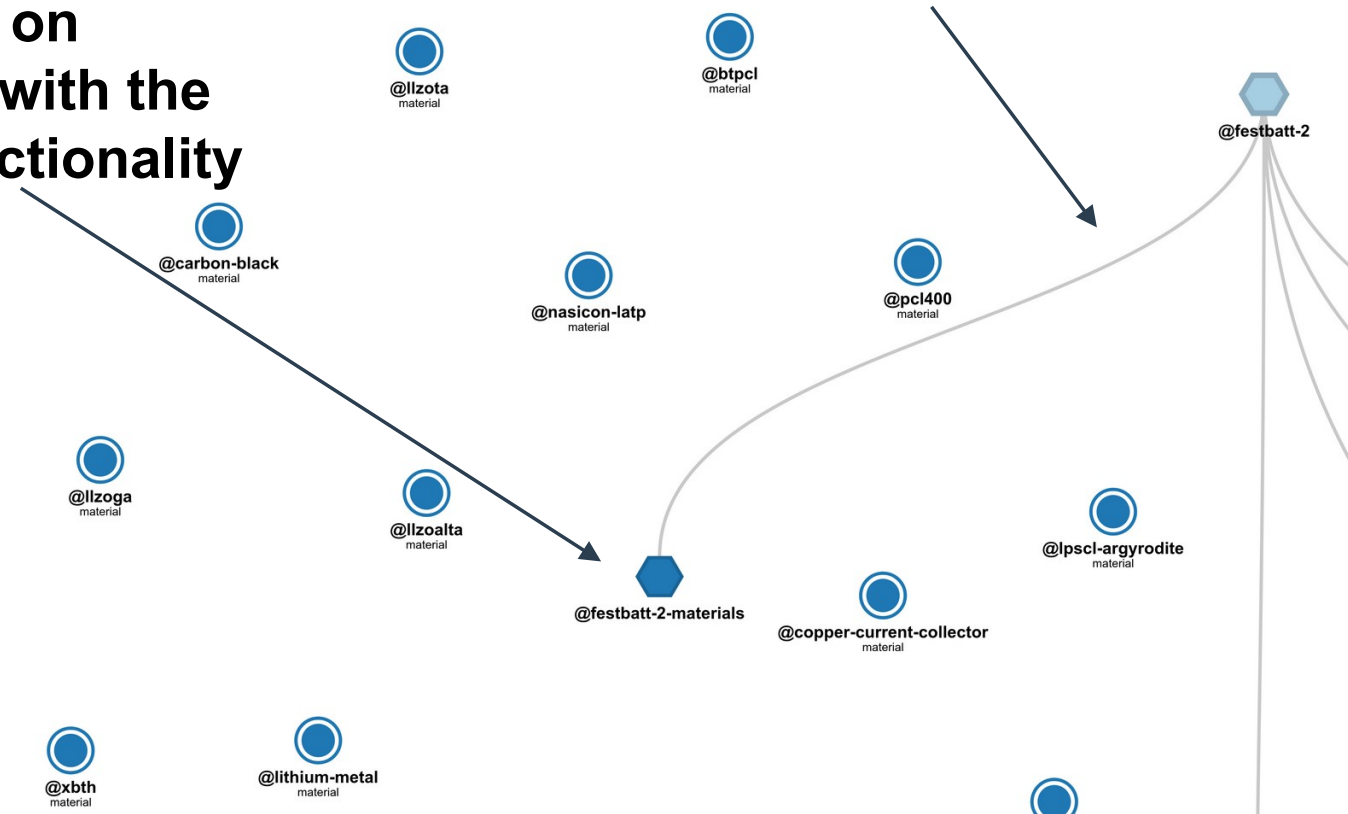
- Similarly, when creating a new record, it is now possible to directly link it with other records using a new editor, which can also show a small preview for each link
- The same is possible within record templates, so the editor can be prefilled when creating a corresponding record

New Feature of Kadi4Mat: Interactive Collection Visualization

Sub/Child collections (loadable on demand) with the same functionality

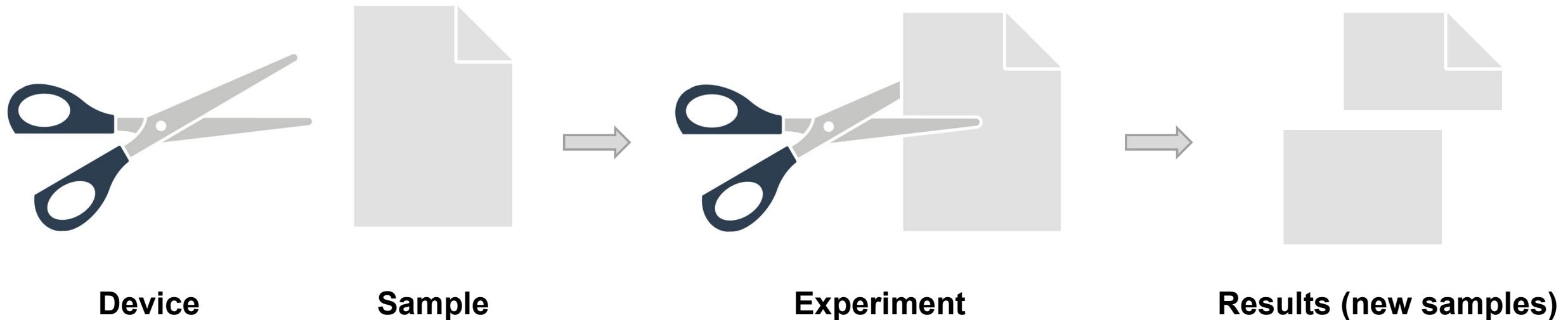
Records (and their links, if applicable) within a specific collection

- An interactive visualization of collections has been added, similar to the existing one for record links
- Next to records, it also allows visualizing the collection hierarchy (and the respective records)

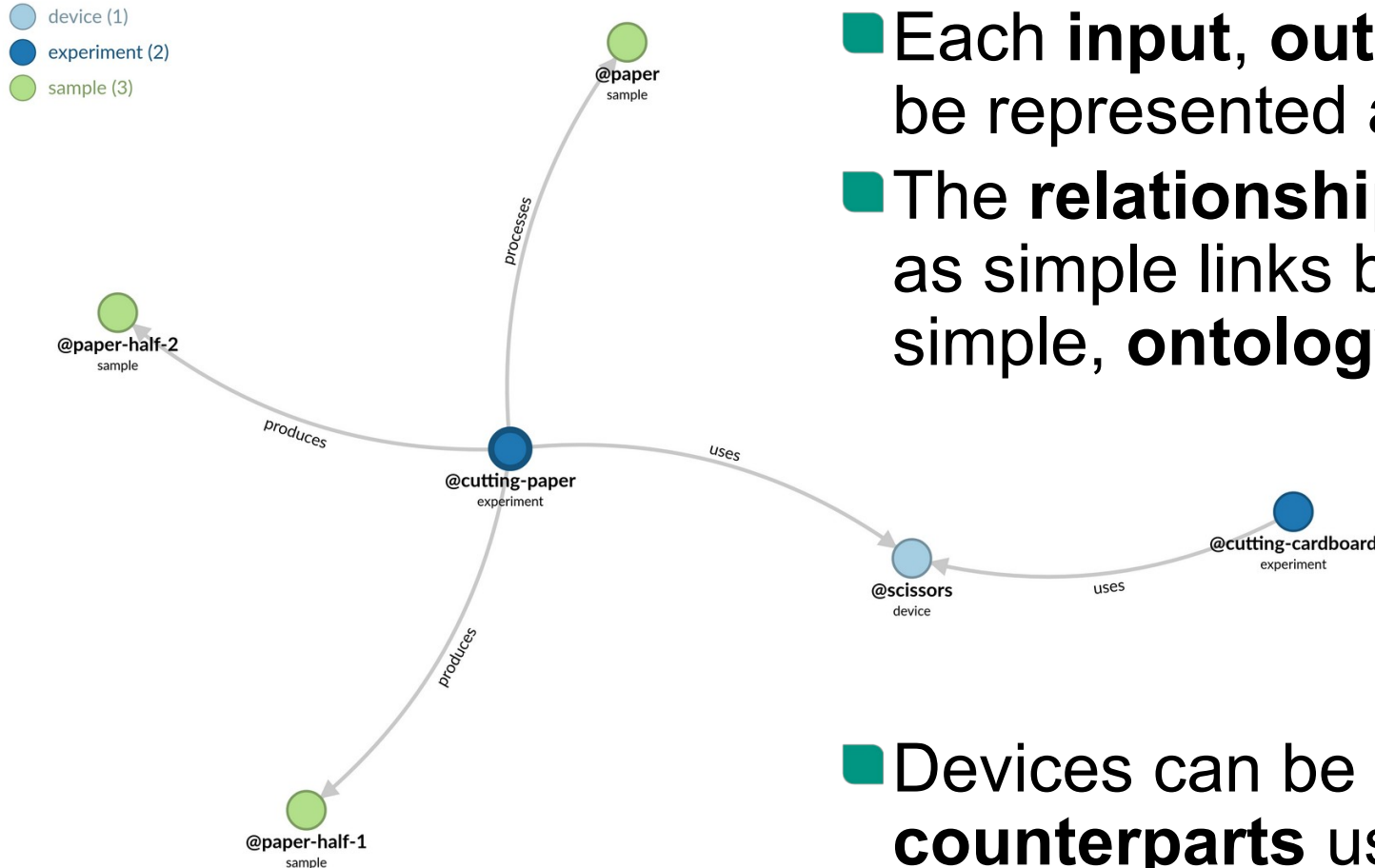


Example: Using Kadi4Mat in Experiments

- A simple **experiment** (cutting a piece of paper) is performed using a **device** (scissors) on some **sample** (paper)



Kadi4Mat in Experiments



- Each **input**, **output** and **processing** step can be represented as a record in Kadi4Mat
- The **relationships** of the steps are represented as simple links between the records, creating a simple, **ontology-like** structure



- Devices can be linked to their **digital counterparts** using e.g. QR codes

Ontologies

*“An ontology is a **formal, explicit** specification of a **shared conceptualization**”*

machine-
readable

consensual knowledge

concepts, properties,
relations, functions,
constrains, axioms are
explicitly defined

abstract model and
simplified view of a
phenomenon

Ontologies

- Important benefits ontologies provide are
 - Semantic interoperability
 - Reasoning and inference – ontologies include logical axioms and rules that capture domain knowledge and relationships
 - Domain understanding and analysis – enables exploration of complex knowledge landscapes
 - Knowledge based systems and applications – enables the creation of intelligent agents and expert systems

Ontologies

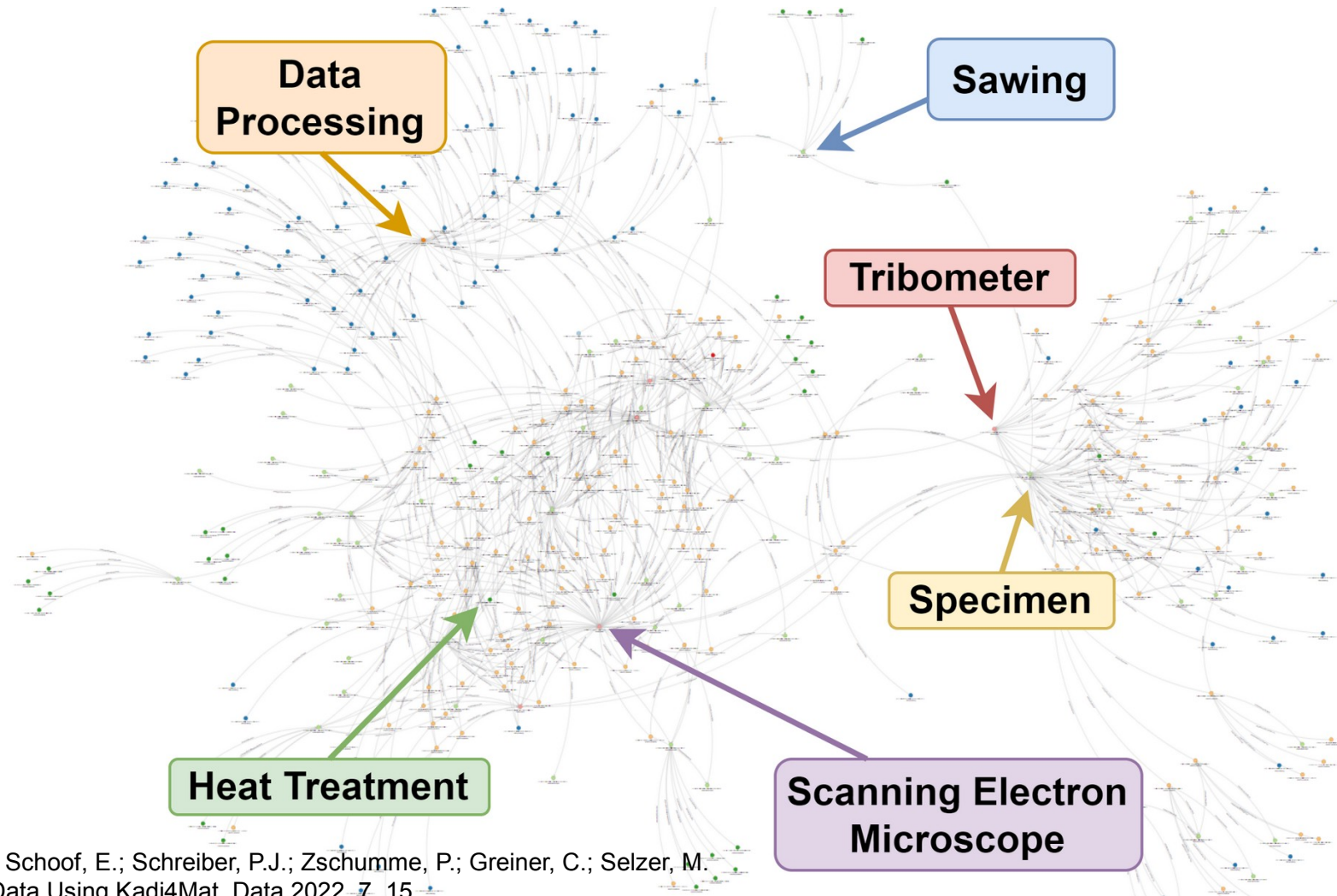
- Crucial for ensuring consistency, interoperability and resuability of data
- Implemented via
 - ~ Metadata standards – metadata templates in Kadi4mat
 - ~ Instrument/device templates

Ontologies



Brandt, N.; Garabedian, N.T.; Schoof, E.; Schreiber, P.J.; Zschumme, P.; Greiner, C.; Selzer, M.
Managing FAIR Tribological Data Using Kadi4Mat. Data 2022, 7, 15.
<https://doi.org/10.3390/data7020015>

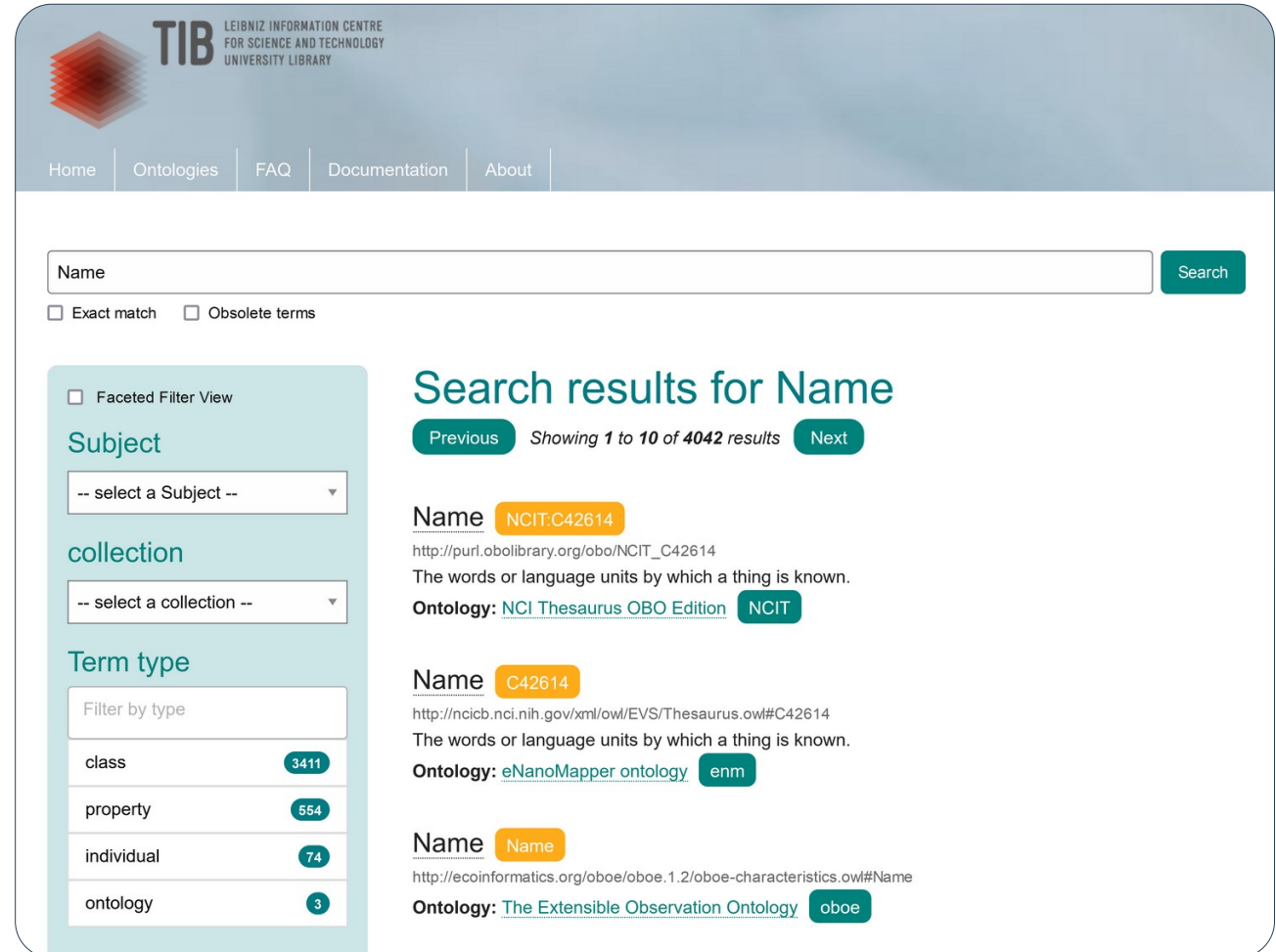
Ontologies



Brandt, N.; Garabedian, N.T.; Schoof, E.; Schreiber, P.J.; Zschumme, P.; Greiner, C.; Selzer, M.
Managing FAIR Tribological Data Using Kadi4Mat. Data 2022, 7, 15.
<https://doi.org/10.3390/data7020015>

New Feature of Kadi4Mat: Metadata Term Search

- Terminology services are one option to reuse existing **ontologies** (which define concepts and their relationship within a specific domain)



The screenshot shows the TIB (Leibniz Information Centre for Science and Technology) Metadata Term Search interface. The header includes the TIB logo and navigation links: Home, Ontologies, FAQ, Documentation, and About. A search bar is located at the top right, with a search button. Below the search bar, there are checkboxes for 'Exact match' and 'Obsolete terms'. A faceted filter view is visible on the left, with sections for 'Subject', 'collection', and 'Term type'. The 'Term type' section shows a list of filters: 'class' (341), 'property' (554), 'individual' (74), and 'ontology' (3). The main search results area displays 'Search results for Name' with navigation buttons for 'Previous' and 'Next', and a message 'Showing 1 to 10 of 4042 results'. Three search results are shown, each with a 'Name' field, a URL, a description, and an 'Ontology' field with a corresponding button.

Search results for Name

Showing 1 to 10 of 4042 results

Name NCIT:C42614
http://purl.obolibrary.org/obo/NCIT_C42614
The words or language units by which a thing is known.
Ontology: NCI Thesaurus OBO Edition NCIT

Name C42614
http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C42614
The words or language units by which a thing is known.
Ontology: eNanoMapper ontology enm

Name Name
http://ecoinformatics.org/oboe/oboe.1.2/oboe-characteristics.owl#Name
Ontology: The Extensible Observation Ontology oboe

<https://service.tib.eu/ts4tib>

New Feature of Kadi4Mat: Metadata Term Search

^ Extra metadata Undo Redo Reset Tree view Show validation

String Key Name Value + x [icon] [icon]

Term IRI **Find term**

An IRI specifying an existing term that the metadatum should represent.

- The term search is available in all places where term IRIs can be specified
- It is implemented using the **plugin infrastructure** of Kadi4Mat, so an integration of other services is easily possible

Find term

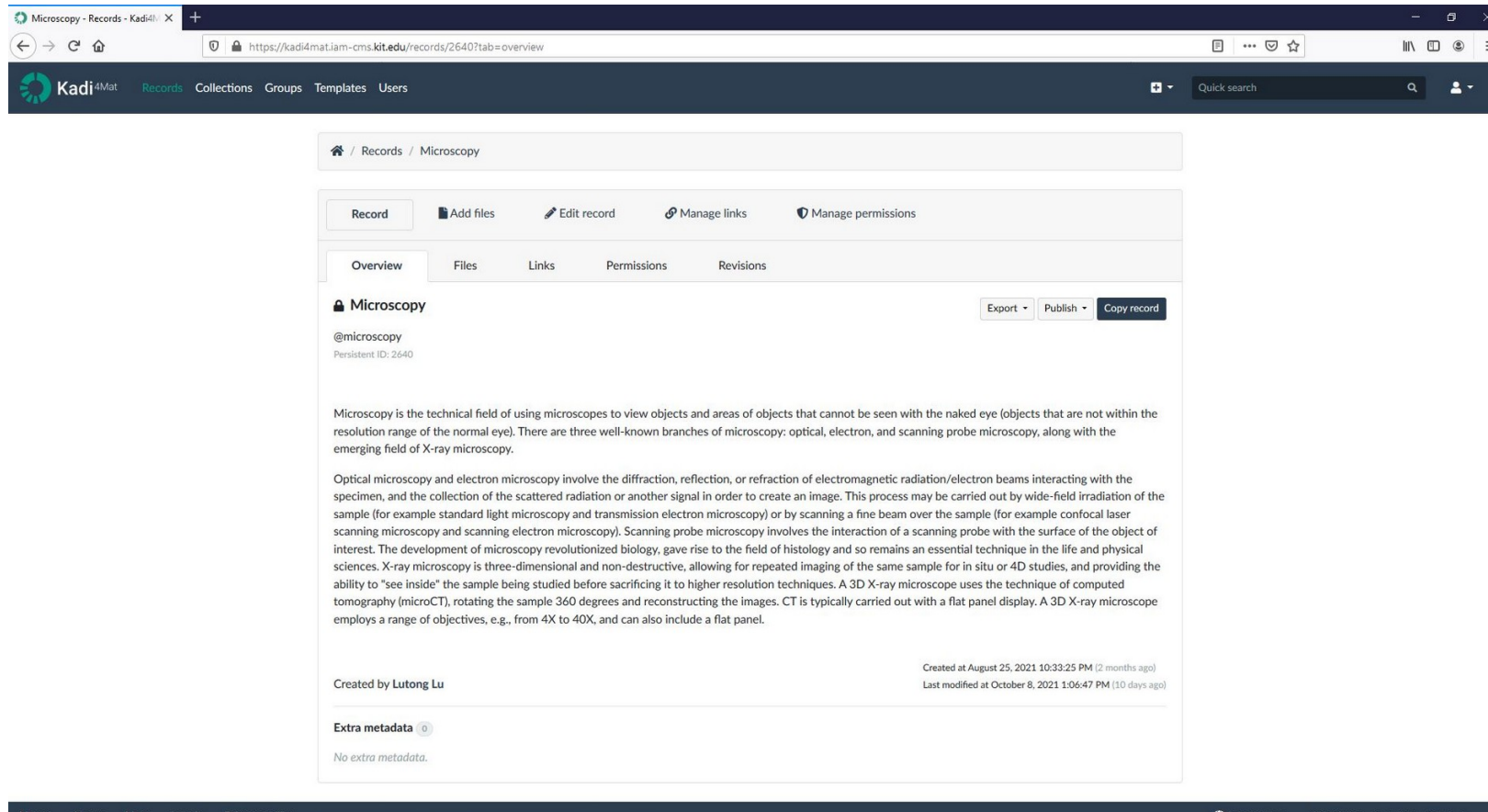
Name

Class Designative Name http://www.ontologyrepository.com/CommonCoreOntologies/DesignativeName A Designative Information Content Entity that consists of a string of characters that designates an entity within a specified cultural or social namespace and which is typically a word or phrase in a natural language that has an accepted cultural or social significance.	Select term
Class Name http://purl.obolibrary.org/obo/NCIT_C42614 The words or language units by which a thing is known.	Select term
Class Name http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C42614 The words or language units by which a thing is known.	Select term
Class Name http://emmo.info/domain-crystallography/cif_top#Name	Select term
Property Name http://rs.tdwg.org/abcd/terms/name The name of a class in a specific language.	Select term

Note that these results are provided by an external terminology service.

<< < Page 1 of 429 > >>

The KIT instrument data base („Gerätepool“) - Inside Kadi4Mat



The screenshot shows a web browser window displaying the Kadi4Mat interface. The address bar shows the URL <https://kadi4mat.iam-cms.kit.edu/records/2640?tab=overview>. The page title is "Microscopy - Records - Kadi4Mat". The interface includes a navigation menu with "Records", "Collections", "Groups", "Templates", and "Users". A search bar is visible. The main content area shows a record for "Microscopy" with the following details:

- Record ID: 2640
- Created by: Lutong Lu
- Created at: August 25, 2021 10:33:25 PM (2 months ago)
- Last modified at: October 8, 2021 1:06:47 PM (10 days ago)

The record description states: "Microscopy is the technical field of using microscopes to view objects and areas of objects that cannot be seen with the naked eye (objects that are not within the resolution range of the normal eye). There are three well-known branches of microscopy: optical, electron, and scanning probe microscopy, along with the emerging field of X-ray microscopy." It further details the techniques of optical, electron, and X-ray microscopy.

Kadi4Mat is the **Karlsruhe Data Infrastructure for Materials Science**, a software for managing research data with the aim of combining new concepts with established technologies and existing solutions.

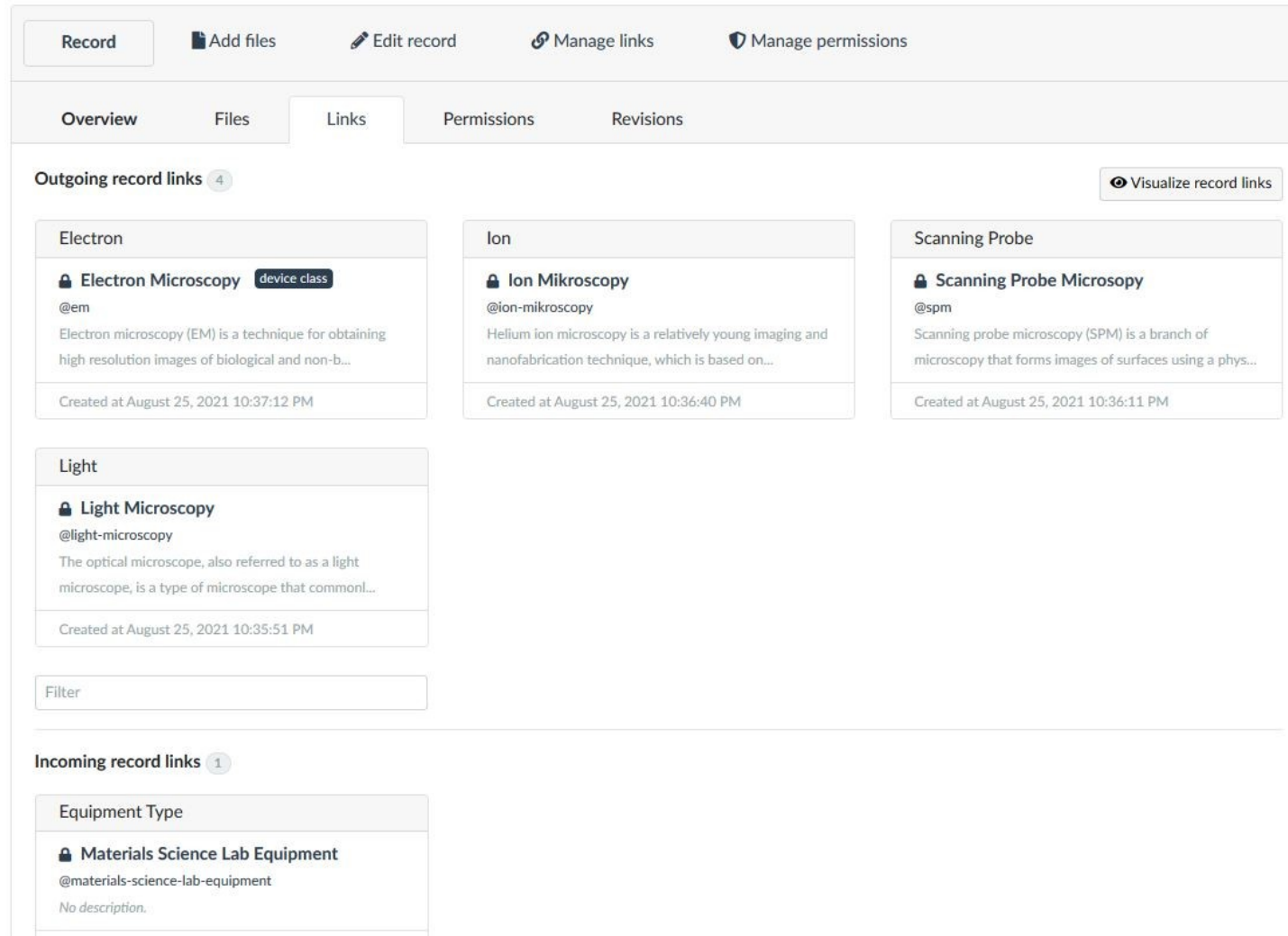
<https://kadi4mat.iam-cms.kit.edu/about>

The KIT instrument data base („Gerätepool“) - Inside Kadi4Mat - Overview



Kadi4Mat is the **Karlsruhe Data Infrastructure for Materials Science**, a software for managing research data with the aim of combining new concepts with established technologies and existing solutions.

<https://kadi4mat.iam-cms.kit.edu/about>



The screenshot displays the Kadi4Mat interface for managing record links. At the top, there are navigation buttons: Record, Add files, Edit record, Manage links, and Manage permissions. Below this is a tabbed interface with 'Overview', 'Files', 'Links', 'Permissions', and 'Revisions'. The 'Links' tab is active, showing 'Outgoing record links' (4) and 'Incoming record links' (1). A 'Visualize record links' button is also present. The outgoing links are categorized into three groups: Electron, Ion, and Scanning Probe. Each group contains a record card with a lock icon, a title, a handle, a description, and a creation timestamp. The incoming link is categorized under 'Equipment Type' and contains a record card for 'Materials Science Lab Equipment' with a handle and no description.

Category	Record Title	Handle	Description	Created At
Outgoing record links	Electron	Electron Microscopy	@em Electron microscopy (EM) is a technique for obtaining high resolution images of biological and non-b...	August 25, 2021 10:37:12 PM
	Ion	Ion Mikroskopy	@ion-mikroskopy Helium ion microscopy is a relatively young imaging and nanofabrication technique, which is based on...	August 25, 2021 10:36:40 PM
	Scanning Probe	Scanning Probe Microscopy	@spm Scanning probe microscopy (SPM) is a branch of microscopy that forms images of surfaces using a phys...	August 25, 2021 10:36:11 PM
Incoming record links	Equipment Type	Materials Science Lab Equipment	@materials-science-lab-equipment No description.	

The KIT instrument data base („Gerätepool“) - Inside Kadi4Mat - Searching



Kadi4Mat is the **Karlsruhe Data Infrastructure for Materials Science**, a software for managing research data with the aim of combining new concepts with established technologies and existing solutions.

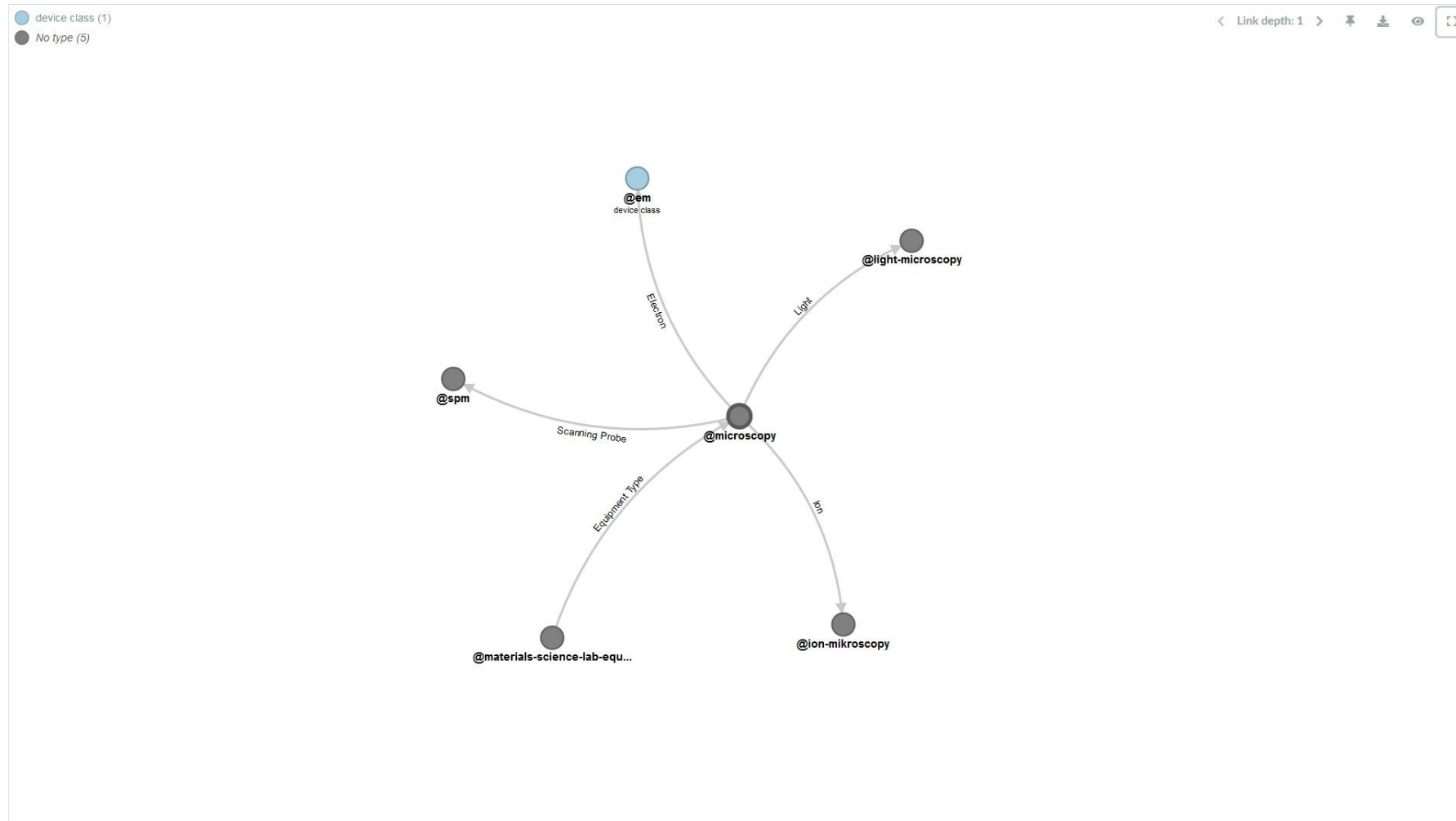
<https://kadi4mat.iam-cms.kit.edu/about>

Search interface showing 49 results found. The interface includes a search bar, a "Sort by" dropdown set to "Relevance", and a "Search extras" button. On the left, there are several filter panels: "Create new record" (with a plus icon), "Select a record template" (set to "microscopy"), "No results found" (with a slider from 10 to 100), "Hide public records" (checkbox), "Filter by collection" (dropdown), "Include child collections" (checkbox), "Filter by type" (dropdown), "Filter by tag" (dropdown), and "Filter by MIME type".

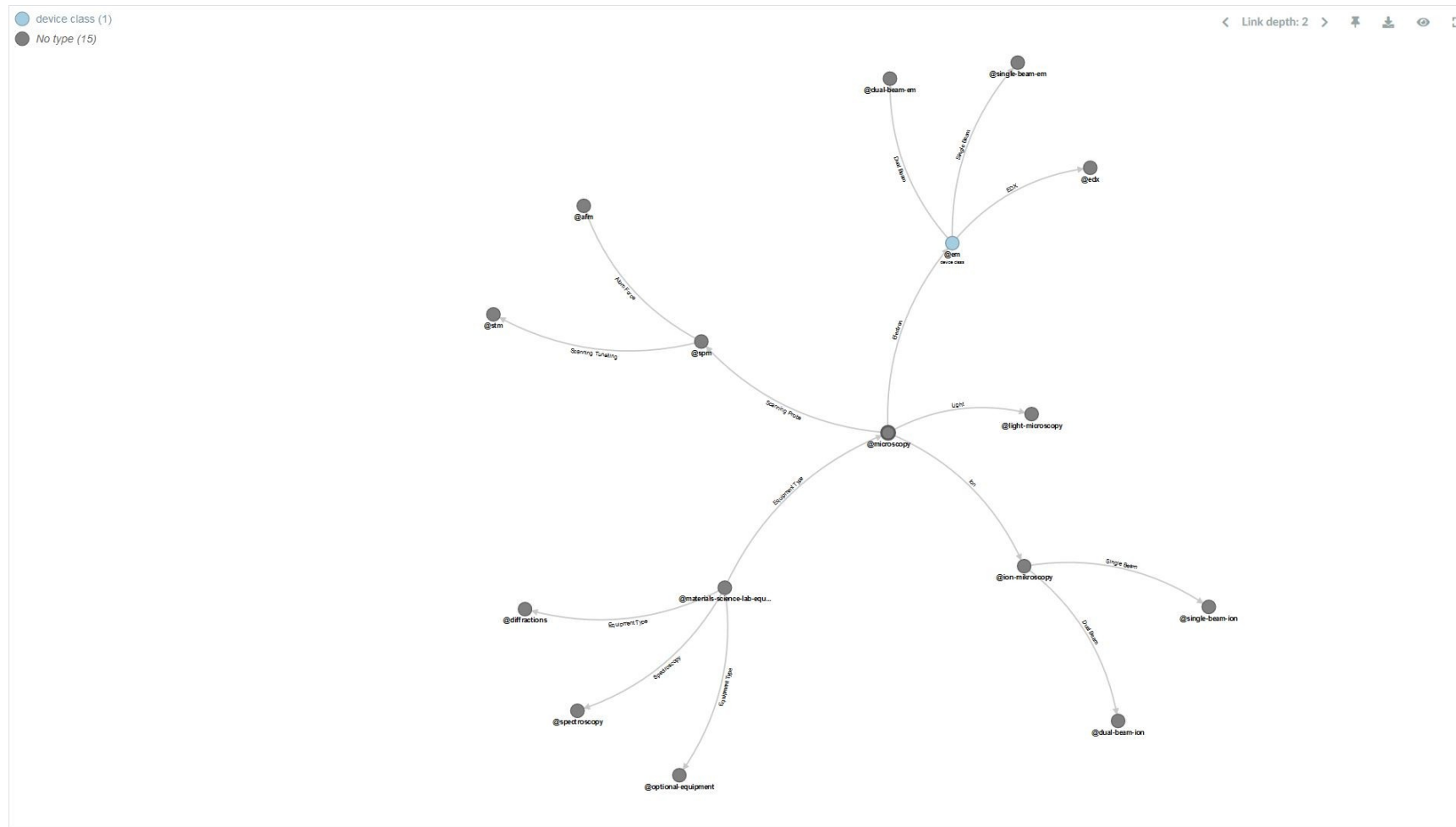
49 results found

- Microscopy** @microscopy
Created 2 months ago
Last modified 10 days ago
Microscopy is the technical field of using microscopes to view objects and areas of objects that cannot be seen with the naked eye (objects that are not within the resolution range of the normal eye). There are three well-known branches of microscopy: optical, electron, and scanning probe microscopy...
Created by Lutong Lu
- Electron Microscopy** @em (device class)
Created 2 months ago
Last modified 10 days ago
Electron microscopy (EM) is a technique for obtaining high resolution images of biological and non-biological specimens. It is used in biomedical research to investigate the detailed structure of tissues, cells, organelles and macromolecular complexes. The high resolution of EM images results from t...
Created by Lutong Lu
- EDX** @edx
Created 14 days ago
Last modified 14 days ago
No description.
Created by Matthias Mail
- Scanning Electron Microscopy** @sem
Created 2 months ago
Last modified 14 days ago
A scanning electron microscope (SEM) is a type of electron microscope that produces images of a sample by scanning the surface with a focused beam of electrons. The electrons interact with atoms in the sample, producing various signals that contain information about the surface topography and compos...
Created by Lutong Lu

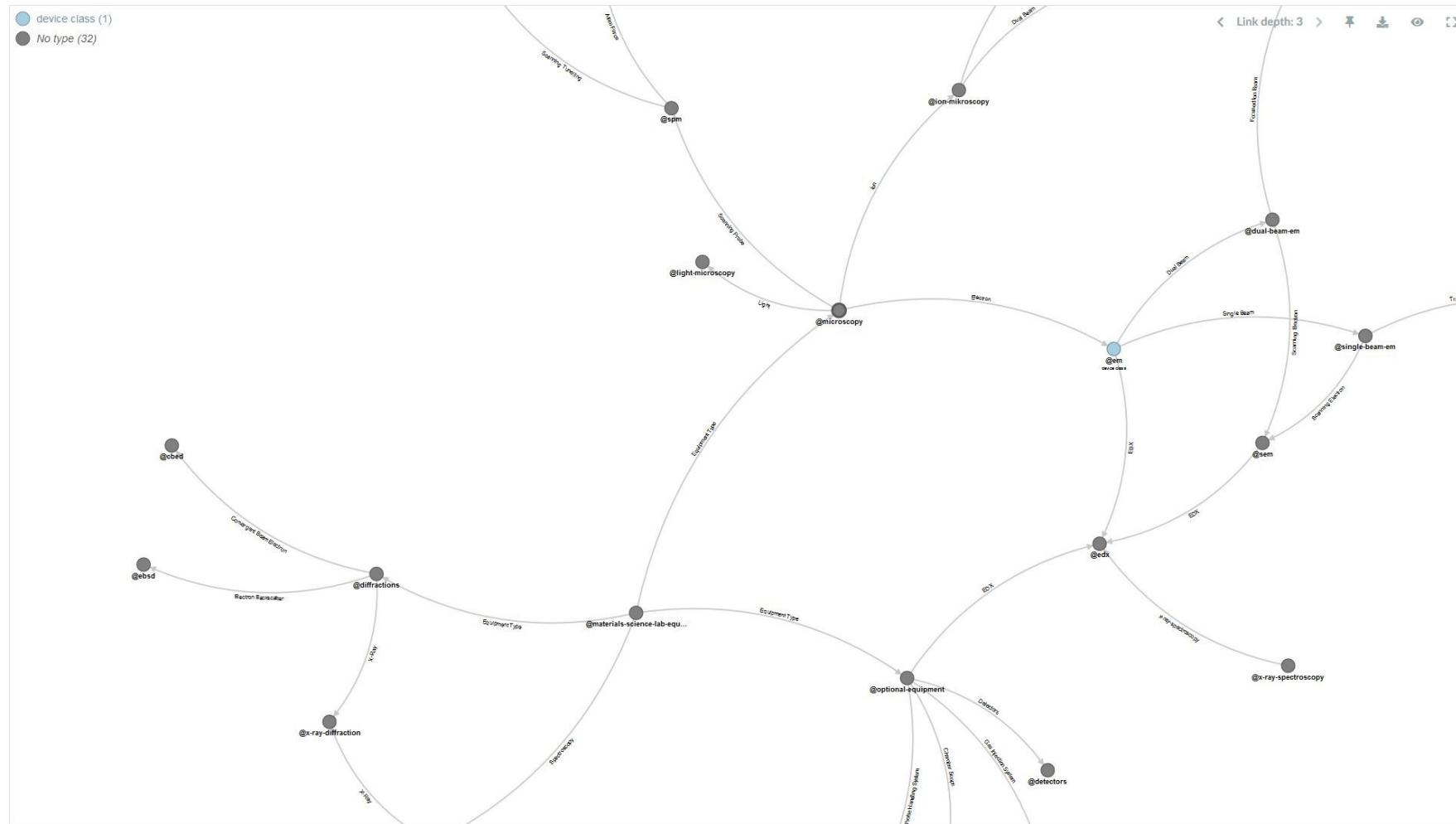
The KIT instrument data base („Gerätepool“) - Inside Kadi4Mat - Searching



The KIT instrument data base („Gerätepool“) - Inside Kadi4Mat - Searching

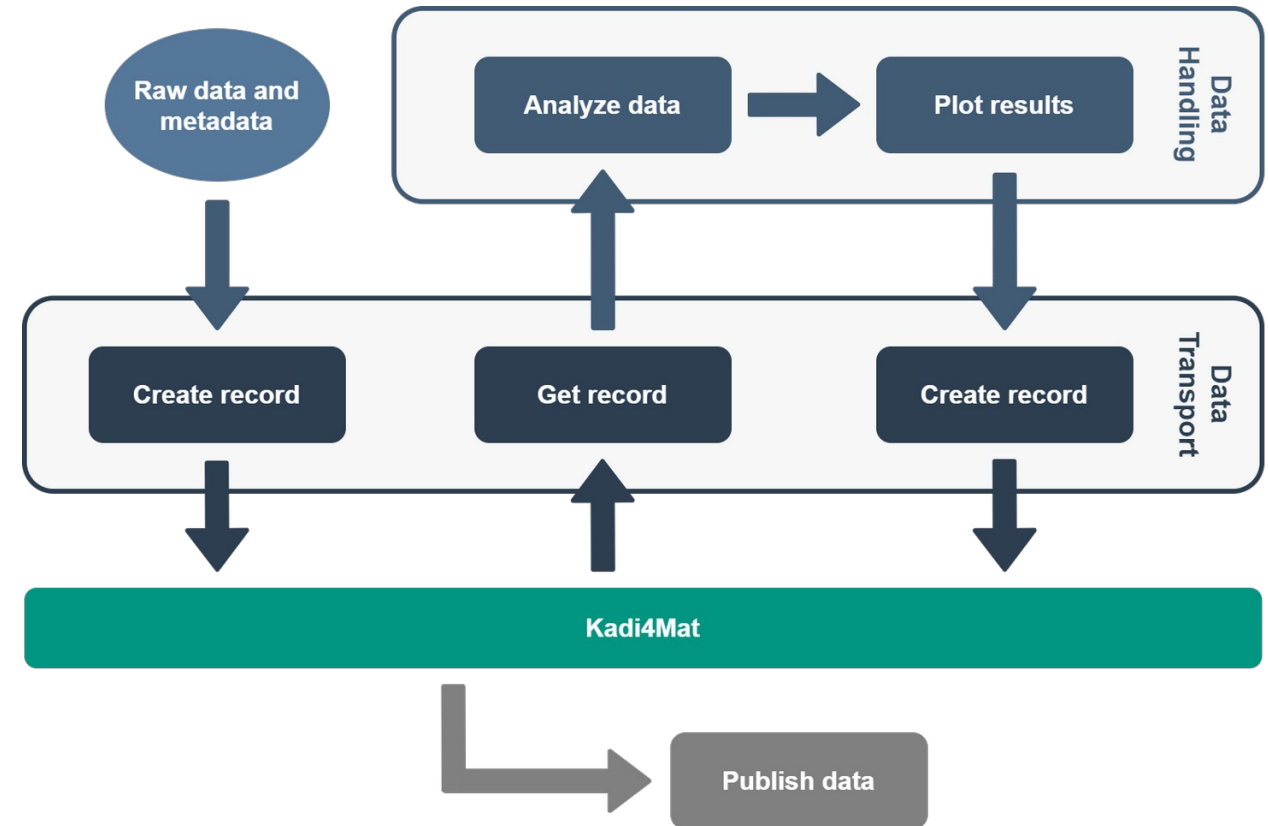


The KIT instrument data base („Gerätepool“) - Inside Kadi4Mat - Searching



Establishing of workflows (ELN)

- A workflow is a well-defined sequence of sequential or parallel steps, which are processed as automatically as possible
- Workflows can be used to integrate existing software (e.g. Origin), devices, and **data science tools**
- A full integration of defining workflows graphically and running them via Kadi4Mat is in development



Establishing of workflows (ELN)

- A workflow is being established as a reference manual for users that will document the sequence of experiments conducted during TEM sample preparation
- The workflows have been designed for two specific applications: preparing TEM samples using FIB Lamella techniques and Ultramicrotomy
- A collection is created at the start of the workflow, to which all the records (all the experiments) are added sequentially.

Using the API via Python

- Sending HTTP request via Python is easiest done with the **requests** library [1]:

```
$ pip install requests
```

- The **GET** request shown previously, done via the browser, can be sent using this library e.g. via the interactive Python shell, using the token created before:

```
$ python
>>> import requests
>>> response = requests.get(
...     "https://kadi4mat.iam-cms.kit.edu/api",
...     headers={"Authorization": "Bearer <token>"},
... )
>>> response.json()
{...}
```

[1] <https://requests.readthedocs.io/en/latest/>

The Kadi-APY Library

- In parallel to the API itself that Kadi4Mat provides, a Python library called **kadi-apy** is available and under development by various people in our group and contributors
- The library is supposed to make using the API via Python as easy as possible, offering higher-level functionality to interact with Kadi4Mat as well as several quality-of-life features
- More information and documentation can be found at <https://gitlab.com/iam-cms/kadi-apy>

The Kadi-APY Library

- The library can be installed as any other library:

```
$ pip install kadi-apy
```

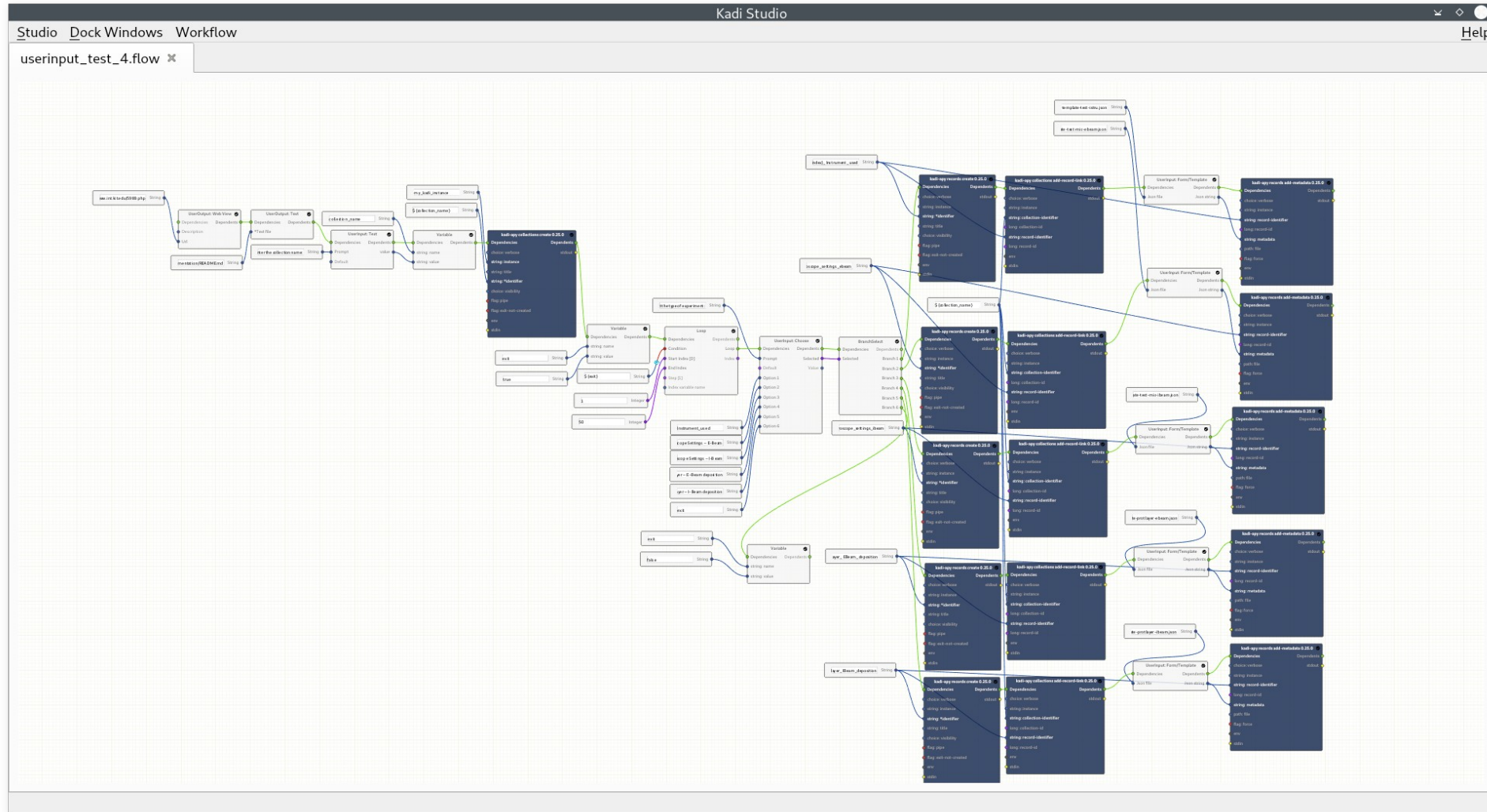
- Besides the **Python module** itself ...

```
$ python  
>>> import kadi_apy  
>>> ...
```

- ... it also includes a **command line interface (CLI)** using the **kadi-apy** command:

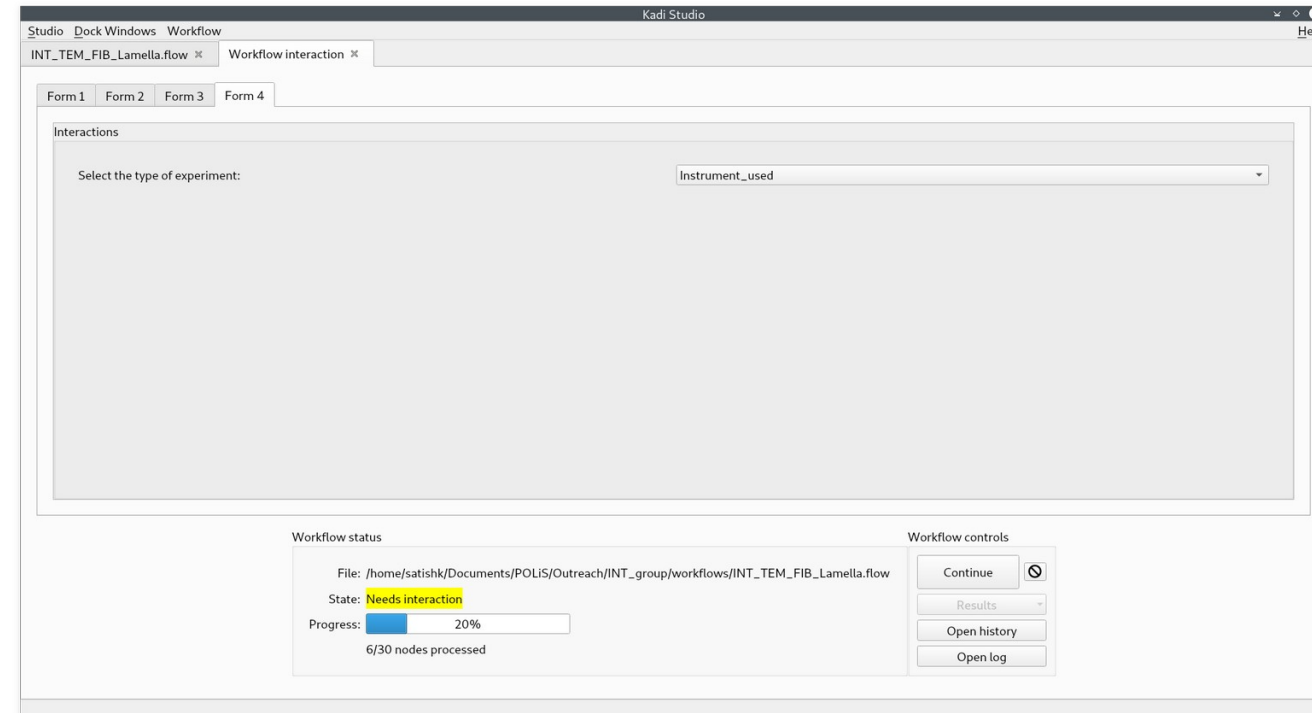
```
$ kadi-apy --help  
...
```

Workflow



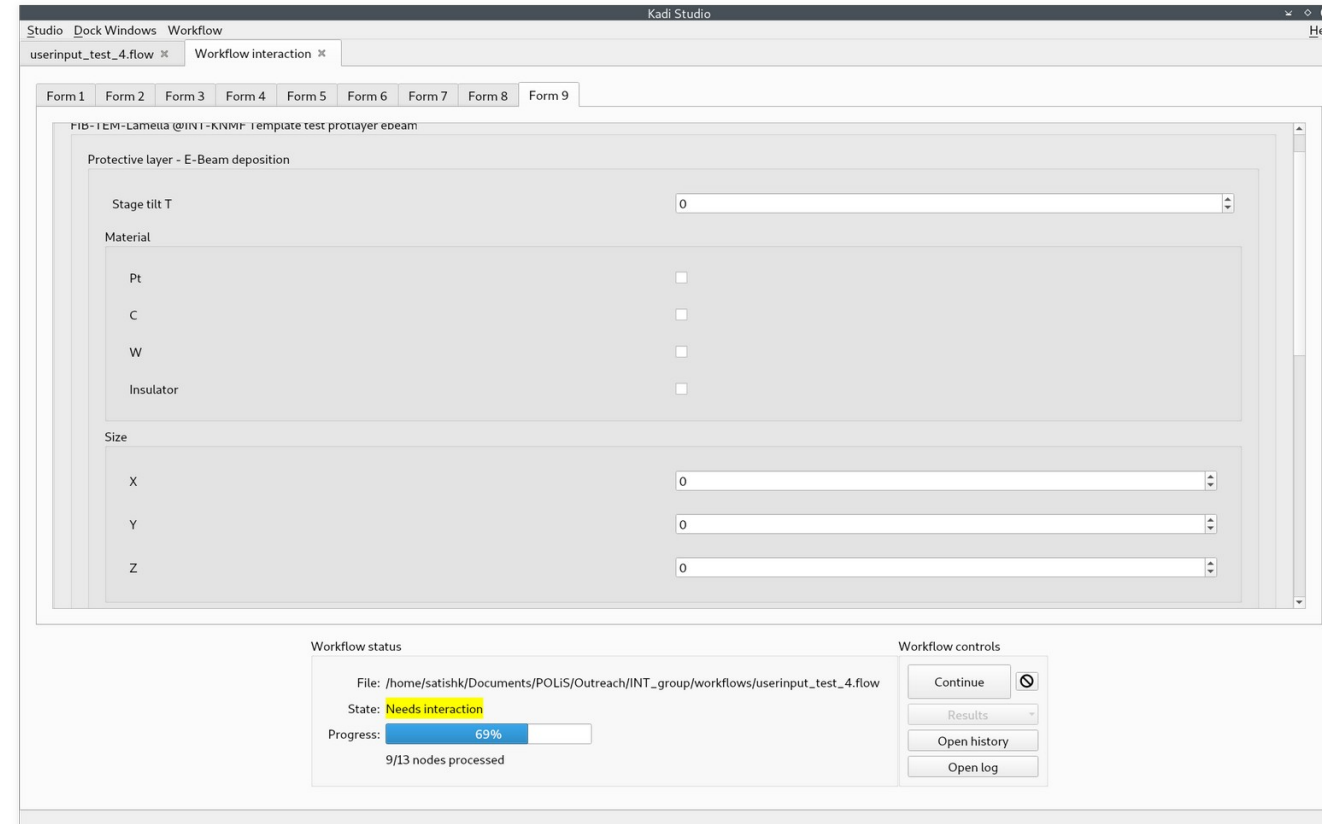
Workflow

- Used instrument can be selected
- User input form is displayed from the metadata templates exported from kadi4mat web application.
- Customized templates in json format
- can be imported into the workflow
- and displayed

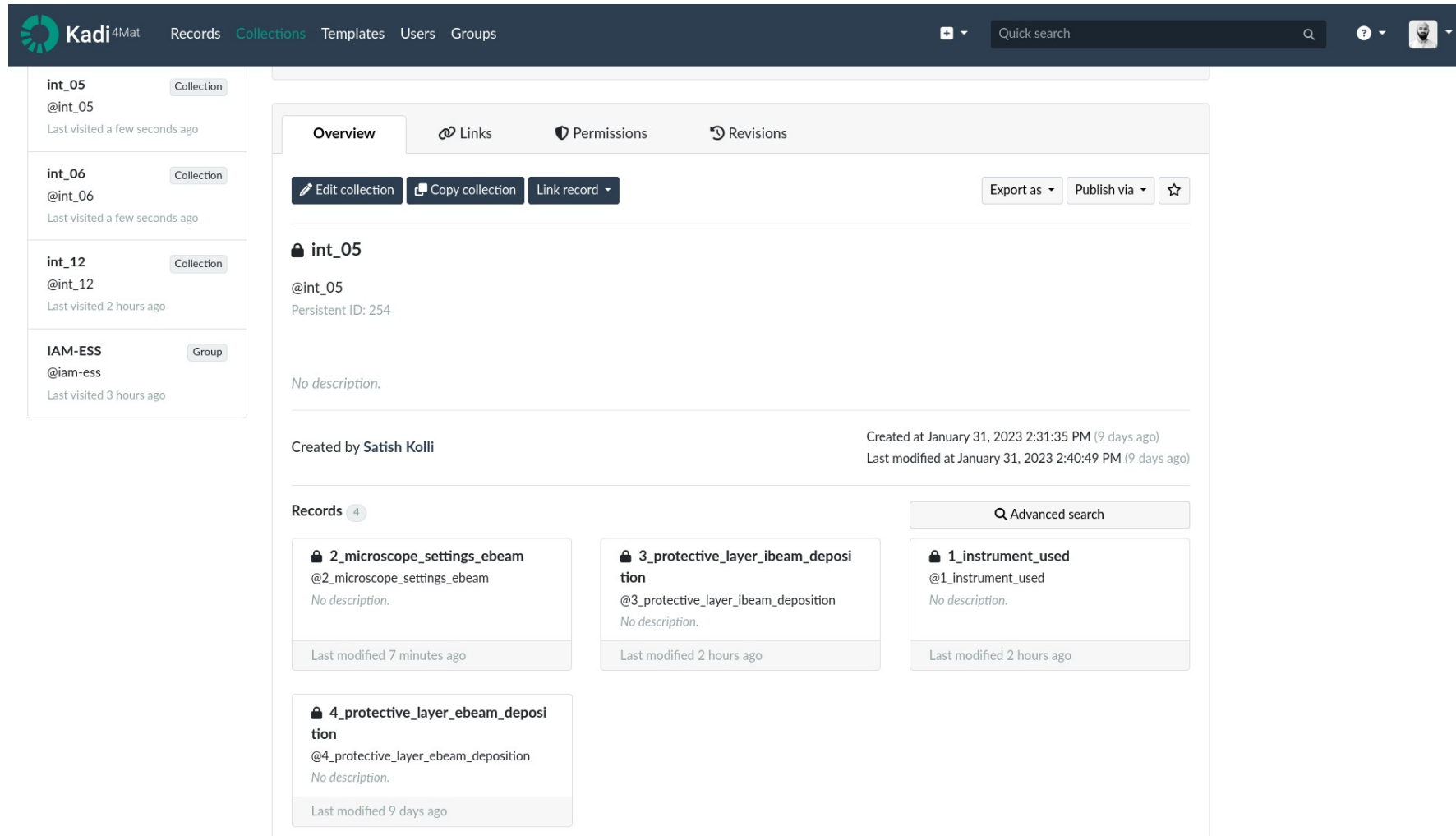


Workflow

- Used instrument can be selected
- User input form is displayed from the metadata templates exported from kadi4mat web application.
- Customized templates in json format can be imported into the workflow and displayed



Collection created in Kadi4mat



The screenshot displays the Kadi4mat web interface. The top navigation bar includes 'Records', 'Collections', 'Templates', 'Users', and 'Groups'. A search bar is present on the right. The left sidebar shows a list of collections and groups, including 'int_05', 'int_06', 'int_12', and 'IAM-ESS'. The main content area shows the details for the 'int_05' collection, including tabs for 'Overview', 'Links', 'Permissions', and 'Revisions'. The 'Overview' tab is active, showing the collection name, user '@int_05', and creation/modification dates. Below this, there are four record cards, each with a lock icon, title, user, description, and last modified time.

int_05 @int_05
Last visited a few seconds ago

int_06 @int_06
Last visited a few seconds ago

int_12 @int_12
Last visited 2 hours ago

IAM-ESS @iam-ess
Last visited 3 hours ago

int_05 @int_05
Persistent ID: 254

No description.

Created by Satish Kolli
Created at January 31, 2023 2:31:35 PM (9 days ago)
Last modified at January 31, 2023 2:40:49 PM (9 days ago)

Records 4

Advanced search

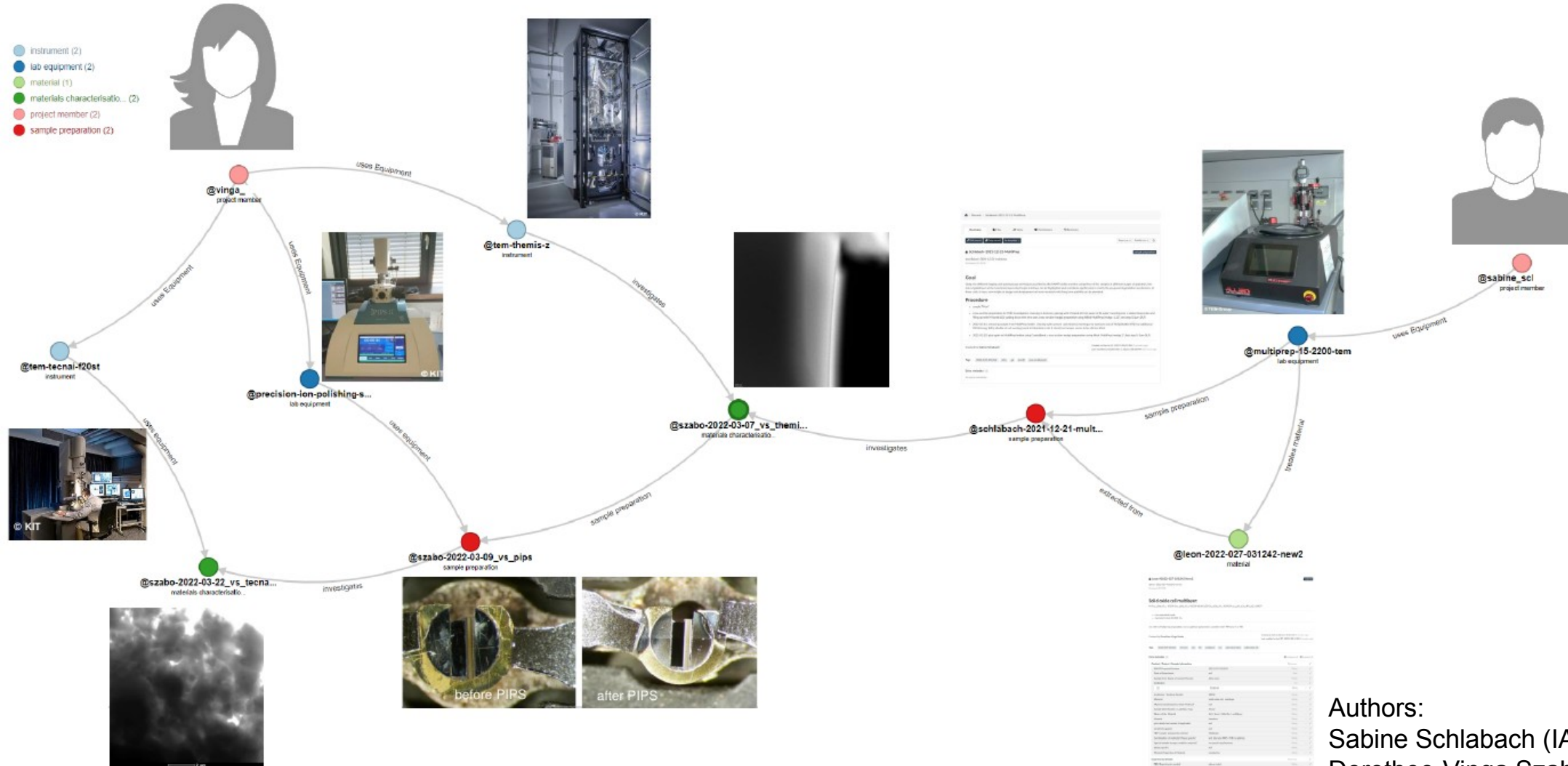
2_microscope_settings_ebeam @2_microscope_settings_ebeam
No description.
Last modified 7 minutes ago

3_protective_layer_ibeam_deposition @3_protective_layer_ibeam_deposition
No description.
Last modified 2 hours ago

1_instrument_used @1_instrument_used
No description.
Last modified 2 hours ago

4_protective_layer_ebeam_deposition @4_protective_layer_ebeam_deposition
No description.
Last modified 9 days ago

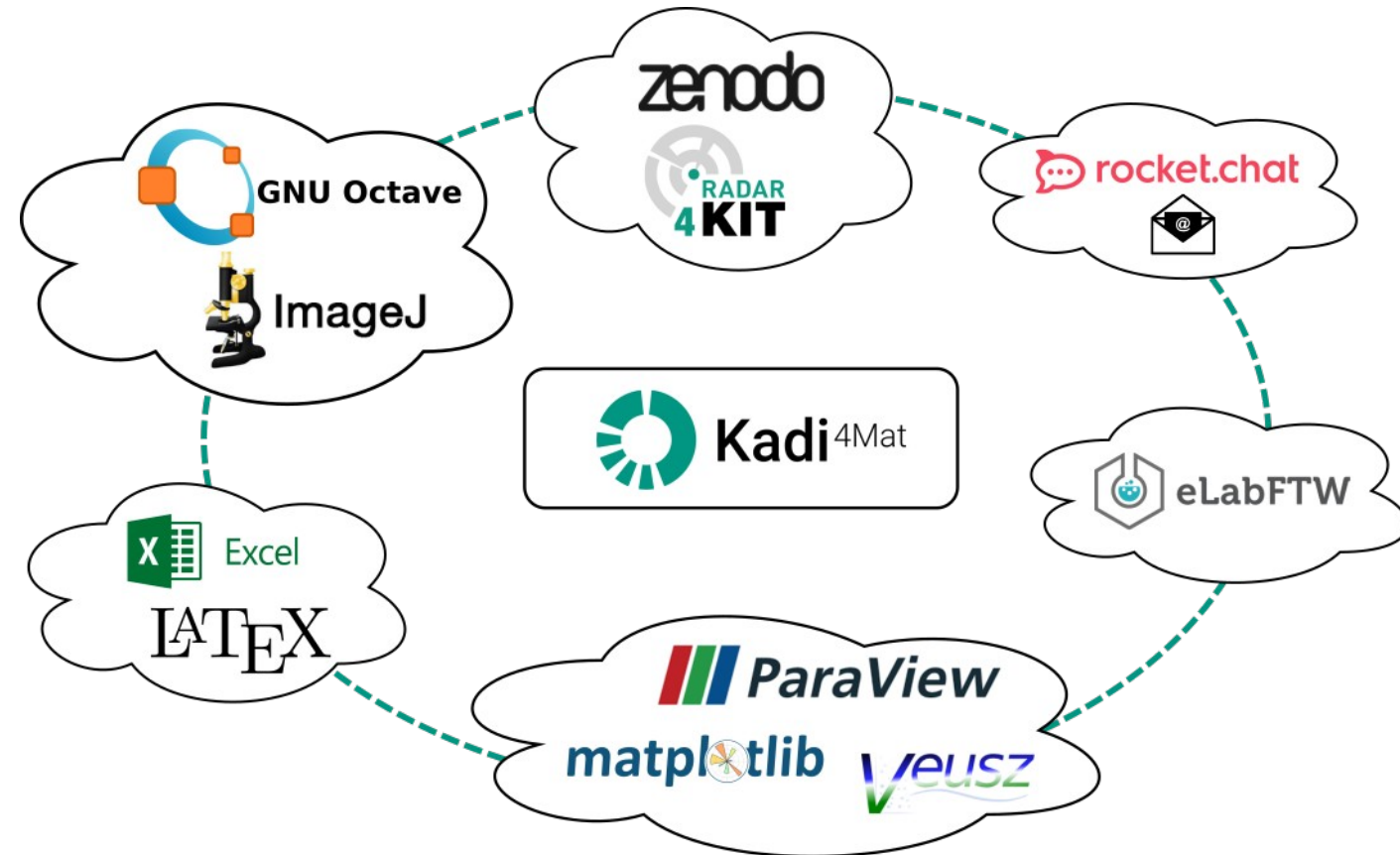
Usage of Instrument descriptions in the research process



Authors:
Sabine Schlabach (IAM)
Dorothee-Vinga Szabo (IAM)

Programming Interface (API) of Kadi4Mat

- A REST-like **HTTP API** enables programmatic and automated usage of most features
- A **Python wrapper library** [1] exists on top of the API, which can also be used to integrate almost any tool via command line interfaces

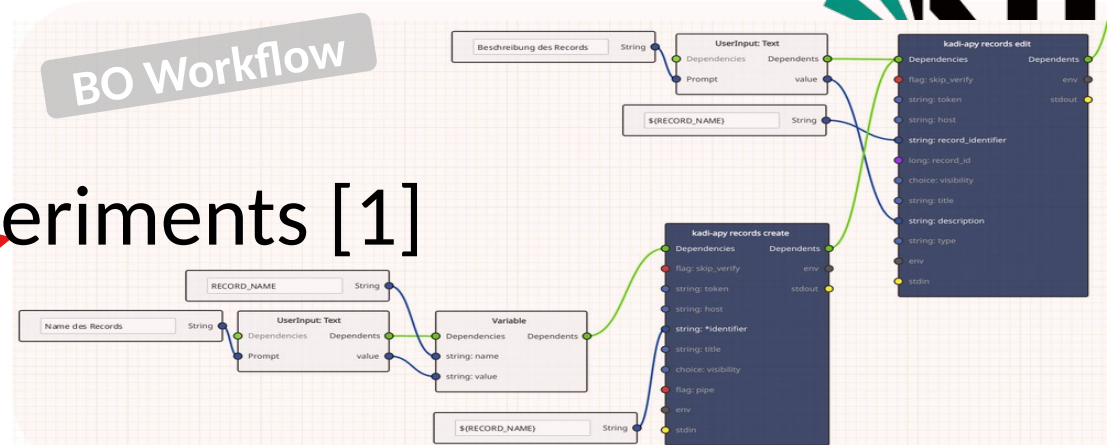
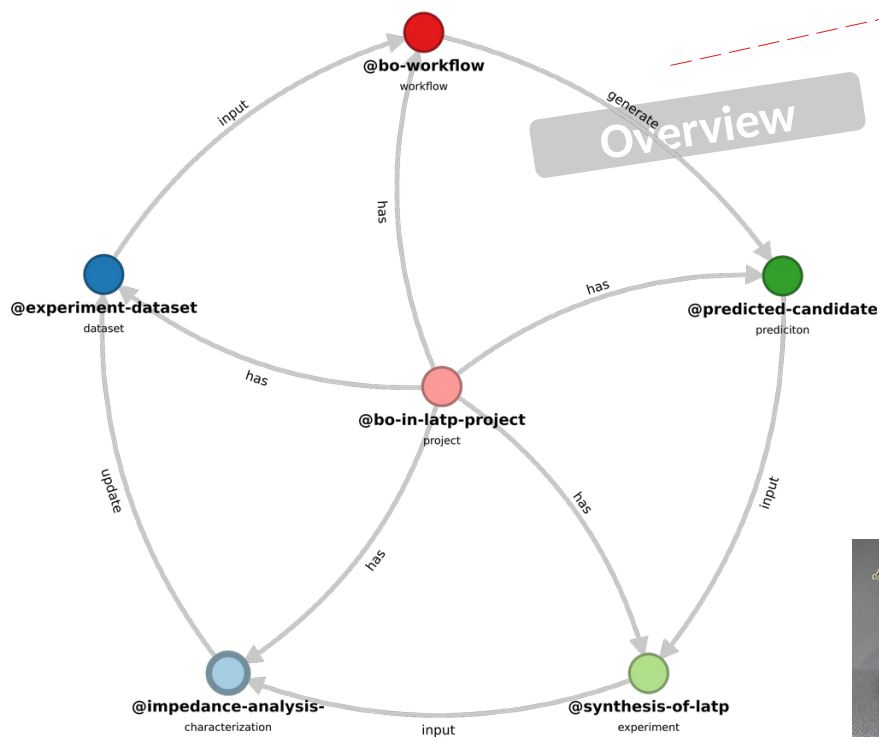


[1] <https://gitlab.com/iam-cms/kadi-apy>

Image: Patrick Altschuh (KIT/HKA, Germany)

Data Science: Application of Bayesian optimization methods

- Reducing required number of experiments [1]



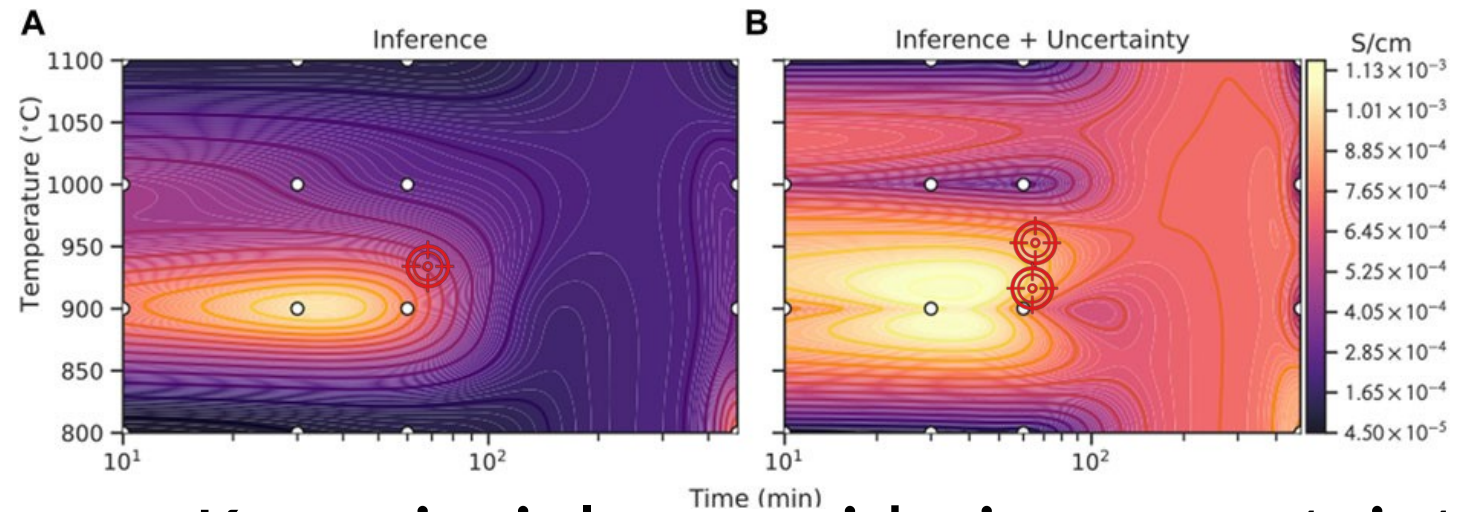
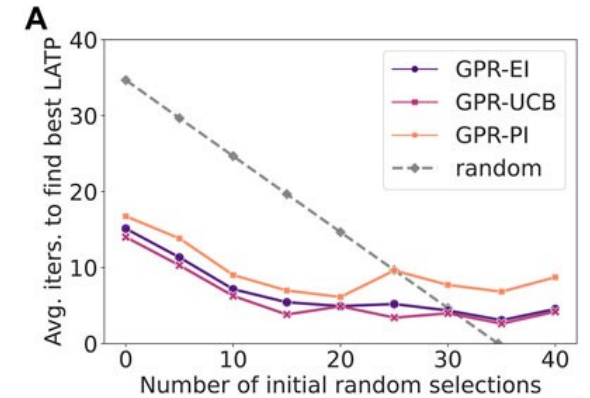
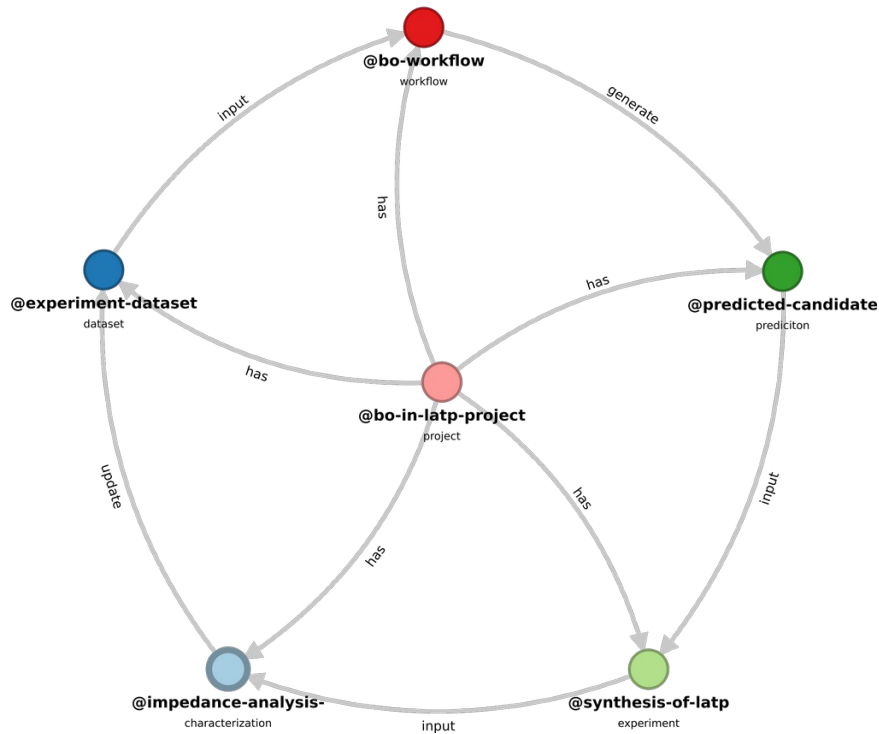
	Input			Output	
	Predicted experiment parameter			Measured properties	
N	rel. H_3PO_4	Sintering	Holding time	Sintered density	Ionic conductivity
o.	wt%	T °C	min	%	S/cm
1	-7.50	1000	40	87.99	5.67E-04
4	-7.50	1000	540	95.81	1.06E-03
1	-11.25	1000	450	97.52	1.09E-03
3					
2	-15.00	1000	450	97.21	7.81E-04
0					

[1] Yinghan *et al.*, 2022, *Frontiers in materials*, w

Data Science: Application of Bayesian optimization methods

- Reducing required number of experiments [1]

~ 1.5 - 2 times faster

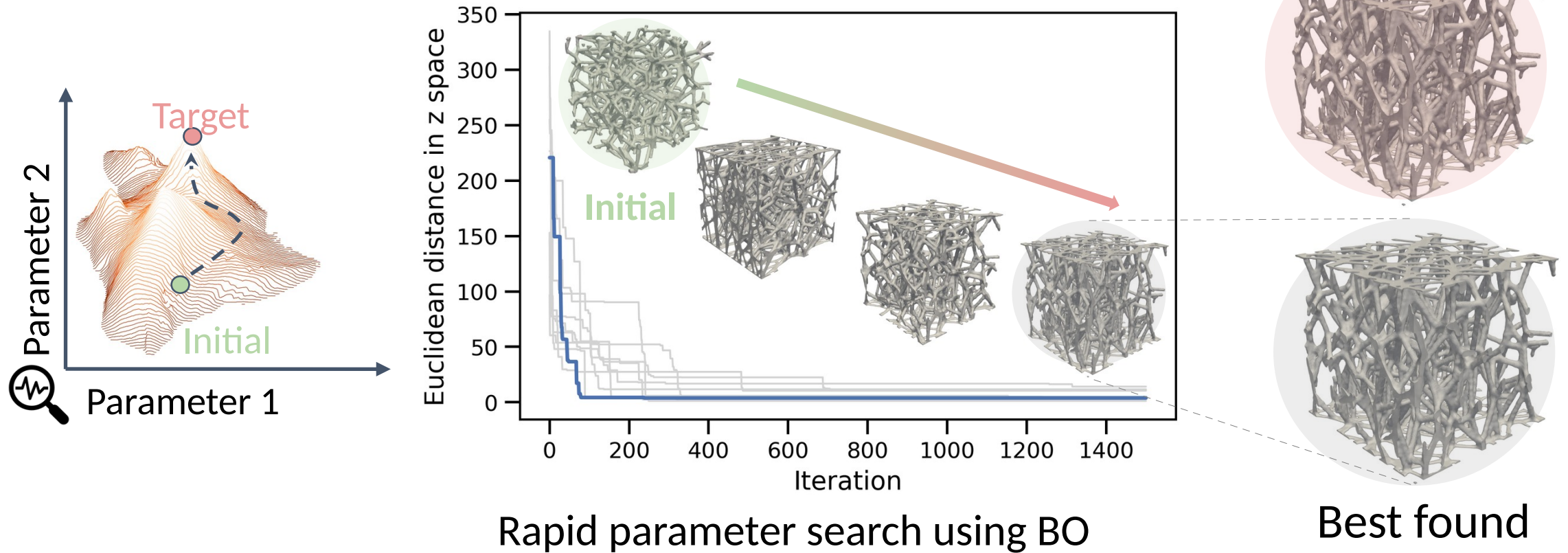


Key principle: considering uncertainty

[1] Yinghan *et al.*, 2022, *Frontiers in materials*, work with Nikolas Schiffmann *et al.*, KIT. BO applied in LATP synthesis

Data Science: Application of Bayesian optimization methods

- Optimizing/searching simulation parameters [1]

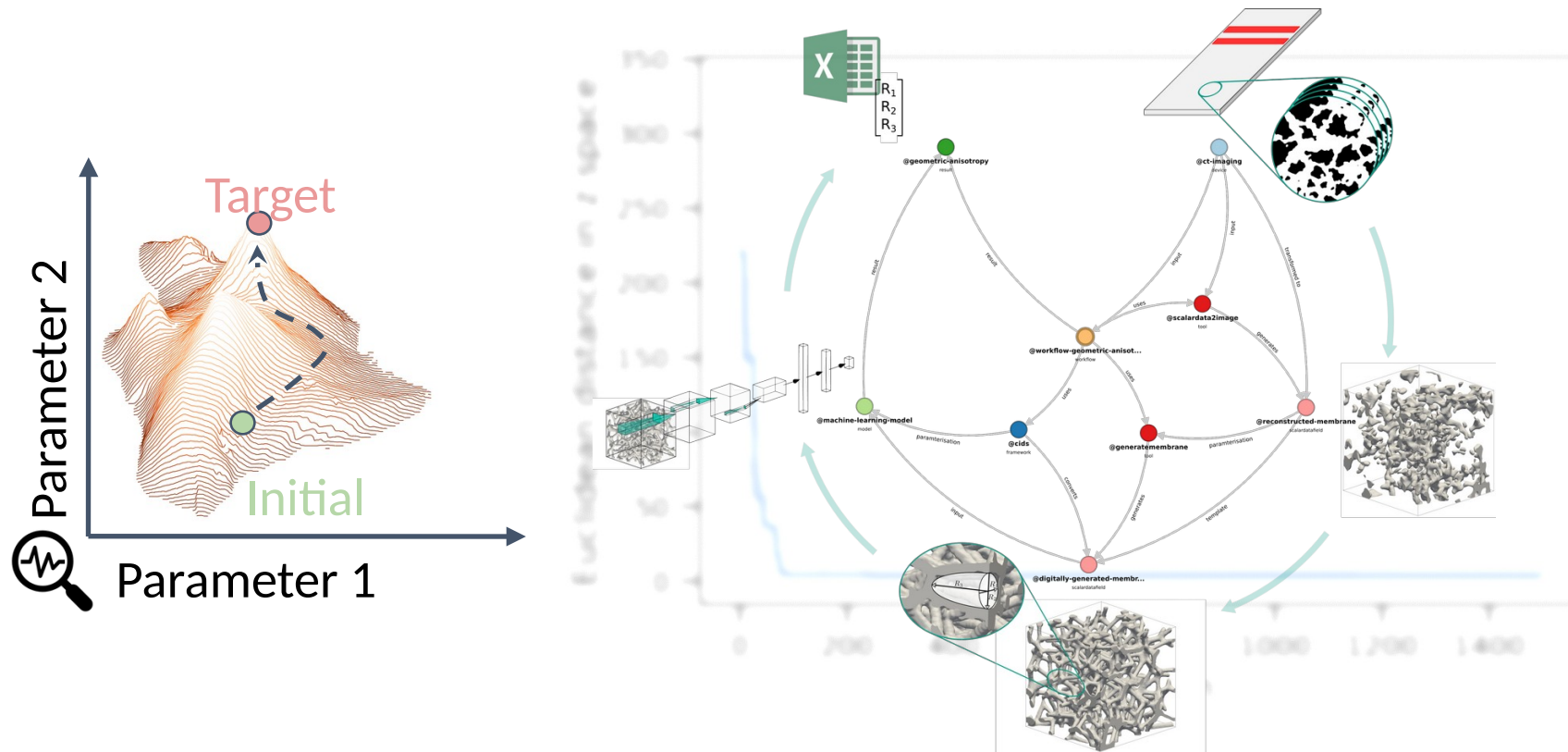


[1] work with Patrick Altschulh *et al.*, KIT. BO applied in simulated porous membranes.

Data Science: Application of Bayesian optimization methods

- Optimizing/searching simulation parameters [1]

Image:
Lars Griem (KIT)

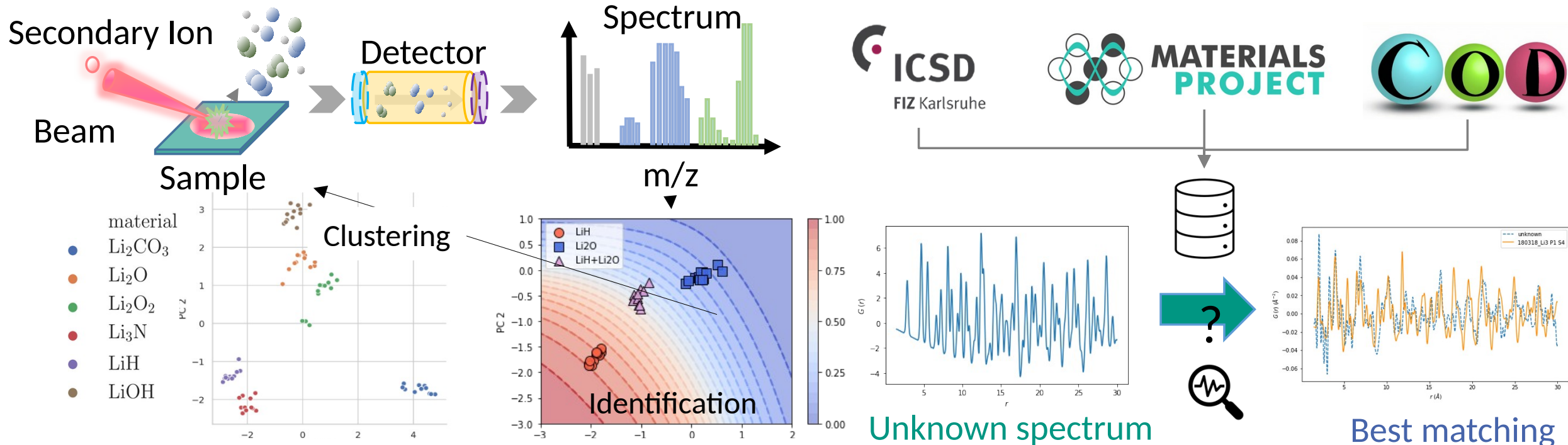


Microstructure-Property
prediction

[1] work with Patrick Altschulh *et al.*, KIT. BO applied in simulated porous membranes.

Data Science: Support in the analysis of complex characterization data

- Enhance data analysis and interpretation
- Collect data from open database and prepare models



PCA of 6 lithium-compounds [1]

Analysis of mixtures

Search PDF candidates in database [2]

[1] Yinghan *et al.*, 2020, Procedia Computer Science, work with S.-K. Otto *et al.*, JLU. Extended draft in preparation
 PCA: principal component analysis, used to extract most important information and reduce data dimension
 [2] work with A.-L. Hansen *et al.*, KIT PDF: pair distribution function

Data Science

Exploration of more application scenarios & integration of data science tools

- Building appropriate models based on the amount/types of data
- Integrating data-science tools (CIDS framework [1]) and providing example use-cases

The image displays three screenshots from a data science tool interface. The left screenshot shows the 'Hyperparameters' section with sliders and dropdowns for 'Num layers encoder', 'Num layers decoder', 'Conv units', 'Conv size', and 'Conv strides'. A purple callout labeled 'Parameter Setting' is overlaid on this section. The middle screenshot shows the 'Model Definition' code editor with Python code for a convolutional VAE model. A purple callout labeled 'User Code' is overlaid on the code. The right screenshot shows a dashboard with an 'AUC.png' plot, a 3D visualization of a protein structure, and a workflow diagram. Two purple callouts labeled 'Online Visualization' and 'Building Workflow' are overlaid on the dashboard.

Interactive data-science tool integration, data visualization and workflow tools

[1] work with Arnd Koeppel et al., KIT

Summary

- The future goal of Kadi4Mat is to support electronically recording the entire **scientific workflow** of daily research work **FAIRly** (*Findable, Accessible, Interoperable, Reusable*)
- Due to the heterogeneous nature of materials science, most features of Kadi4Mat are kept **generic** and are developed **bottom-up**
- All relevant information about Kadi4Mat can also be found at <https://kadi.iam.kit.edu>



Acknowledgments

- This work is funded by the Ministry of Science, Research and the Arts Baden-Württemberg (MWK-BW), in the project MoMaF--Science Data Center (grant number: 34-7547.222).
- We would like to acknowledge the German Federal Ministry of Education and Research (BMBF) for its financial support within the project FB2 TheoDat, under the grant number 03XP0435D.
- This work is funded by the BMBF and MWK-BW as part of the Excellence Strategy of the German Federal and State Governments in the project Kadi4X.
- This work was partly carried out with the support of the Karlsruhe Nano Micro Facility (KNMFi, www.knmf.kit.edu), a Helmholtz Research Infrastructure at Karlsruhe Institute of Technology (KIT, www.kit.edu).
- Part of the model formulation was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), under project ID 390874152, POLiS Cluster of Excellence (grant number: UP 33/1). This work contributes to the research performed at CELEST (Center for Electrochemical Energy Storage Ulm-Karlsruhe).
- The authors would like to thank the Federal Government and the Heads of Government of the Länder, as well as the Joint Science Conference (GWK), for their funding and support within the framework of the NFDI4Ing consortium. Funded by the German Research Foundation (DFG) - project number 442146713."
- This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), under the project number SE 2842/3-1.