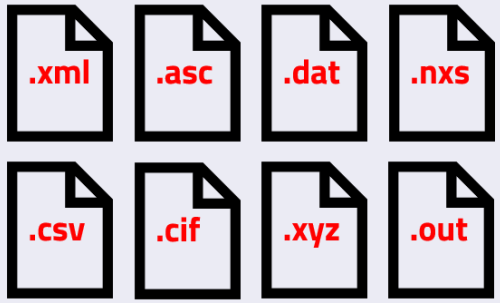
From raw files to published datasets

FAIRmat tutorial 16
February 26, 2025

Publishing your data with NOMAD

To begin the journey from uploading raw files to publishing datasets with a Digital Object Identifier (DOI), you need to understand the key NOMAD elements involved in the process.



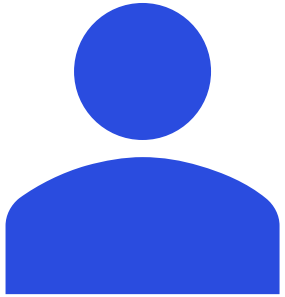
raw data files

NOMAD

published datasets

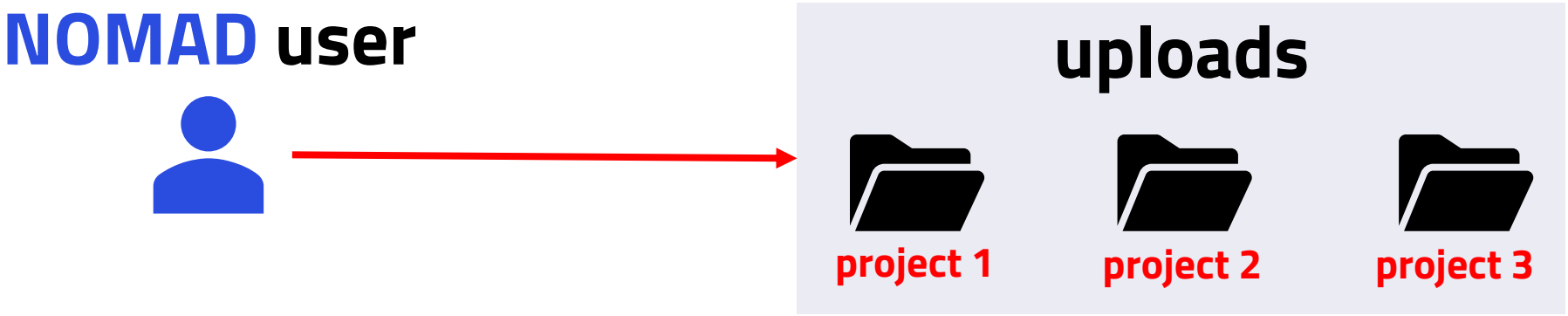
NOMAD user

To create a user account, follow the steps on this link.



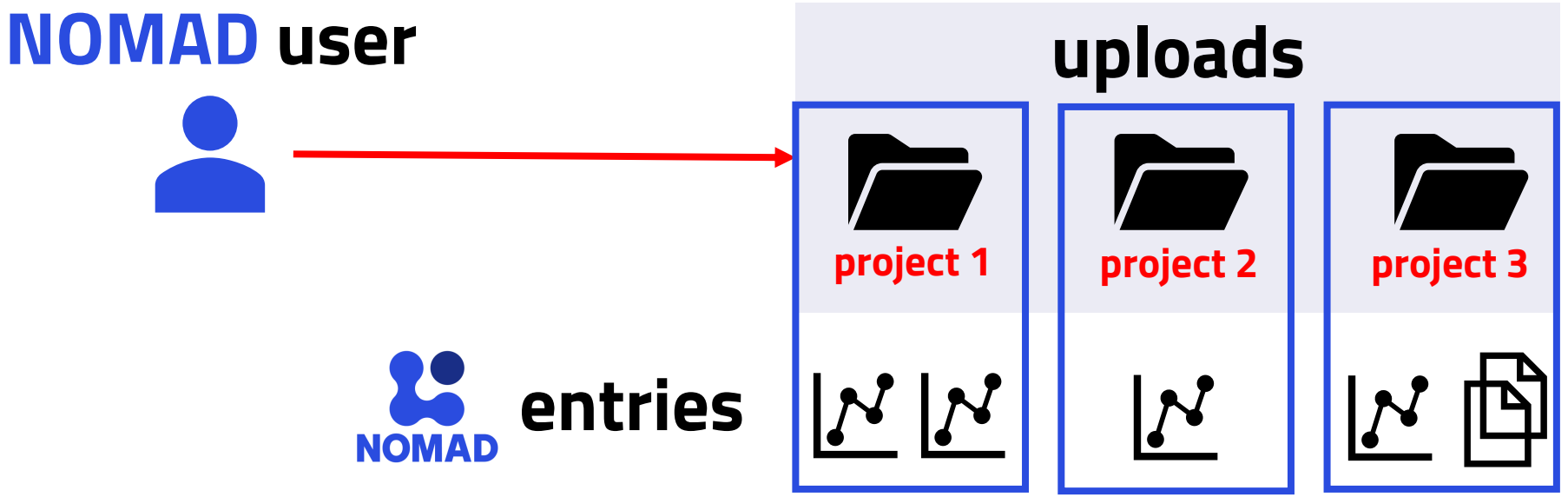
steps to create an account

NOMAD uploads



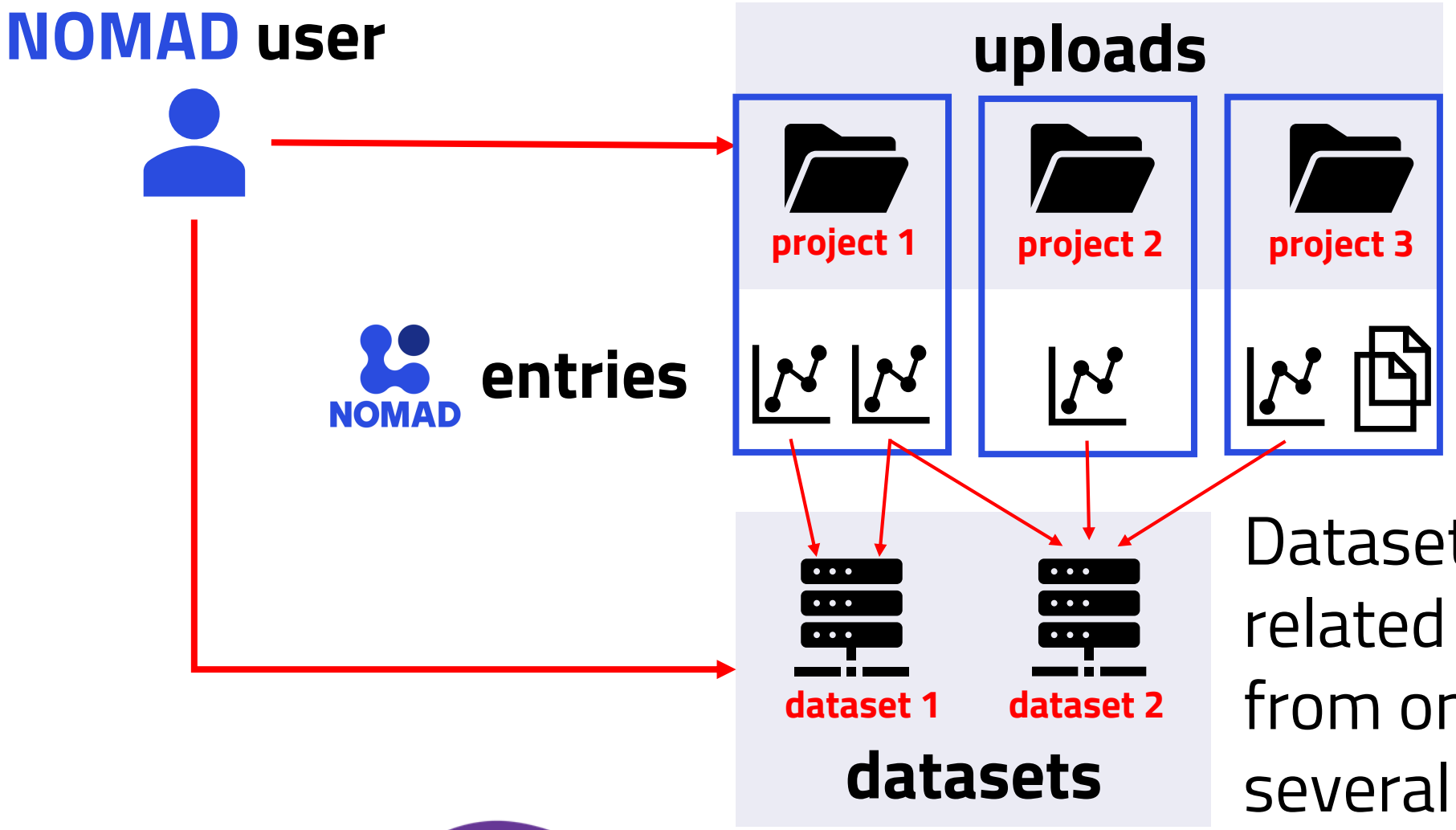
- Research data in NOMAD are organized inside uploads.
- Upload can be created for each project and be organized in nested folder directories.
- Each upload can contain multiple files and entries.

NOMAD entries



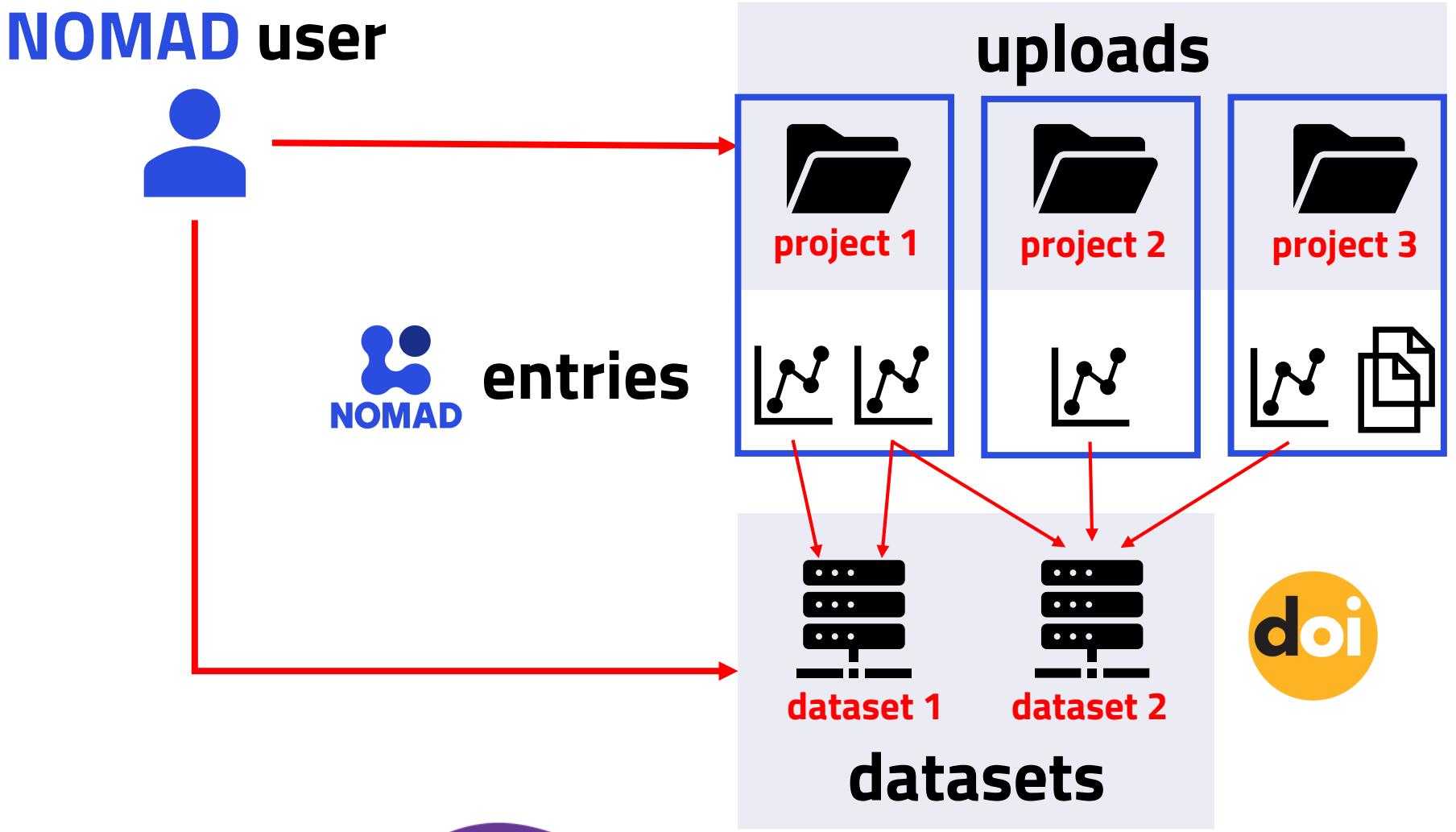
- Entries are generated automatically by NOMAD.
- Entries can be searched and viewed through dedicated pages.
- Files are processed into entries if a built-in parser is available.

Datasets in NOMAD



Datasets group related entries either from one upload or several uploads.

The key elements in NOMAD



Creating and managing uploads

Hands-on tasks:

- ❑ Create your first upload in NOMAD.
- ❑ Rename your upload.
- ❑ Understand the elements and views of the upload page.
- ❑ Share your upload with a collaborator.
- ❑ Publish your upload.

Lets go to NOMAD, click here!

Uploading raw files to NOMAD

Hands-on tasks:

- ❑ Add files to your upload.
- ❑ Distinguish between generated entries and stored files.
- ❑ Preview and visualize your uploaded files/entries.

Uploaded raw files: entries vs. files

NOMAD Entries

- Originate from files that have a **built-in parser in NOMAD exists.**
- NOMAD extracts the data based on the data schema.
- This allows for searchability, analysis and generating visualization.

Other Files

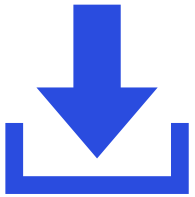
- Originate from files that **do not have a built-in parser in NOMAD**.
- NOMAD cannot automatically extract and structure their data.
- They are stored as part of your upload, and can be downloaded or shared.

Exercise 1: uploading miscellaneous files

In this exercise, we will use files that are not processed into entries in NOMAD, such as, .pdf, .jpg, .txt, and .csv.

Hands-on tasks:

- ❑ Add files to your upload by drag-and-drop and by file browser.
- ❑ View and organize your files in your upload.
- ❑ Preview the files contents in the GUI.



Download the files!
click here!

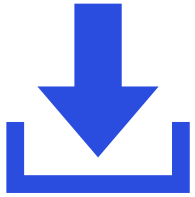
Lets go to NOMAD,
click here!

Exercise 2: uploading computations data files

In this exercise, we will use input and output files of a DFT calculation for Iron(III) Oxide.

These calculations were preformed using a code that is supported by NOMAD, i.e. the *FHI-aims* code.

NOMAD has a parser for the *FHI-aims* code. This means it will create an **entry** for these data.



Download the files!
click here!

Exercise 2: uploading computations data files

Hands-on tasks:

- ❑ Add files to your upload by drag-and-drop.
- ❑ Explore the generated entry page from the files.
- ❑ Provide additional metadata to the entry.

Lets go to NOMAD, click here!

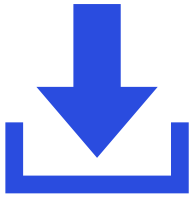
Exercise 3: uploading experimental data files

In this exercise, we will upload files of an x-ray photoelectron spectroscopy (XPS) measurement.

NOMAD supports *.nxs* experimental data files.

Scientific instruments generate raw files in other formats, such as *.xml* (from SPECS spectrometer).

NexusDataConverter converts common formats to the .nxs format.



Download the files!

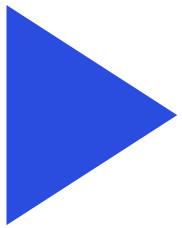
click here!

Exercise 3: uploading experimental data files

In this exercise, we will upload files of x-ray photoelectron spectroscopy (XPS) measurement on polymers.

Hands-on tasks:

- ❑ Add .nxs files to your upload by drag-and-drop.
- ❑ Explore the generated entry page from the files.
- ❑ Use the *NexusDataConverter* to upload files with other formats, such as .xml.



Lets go to NOMAD, click here!

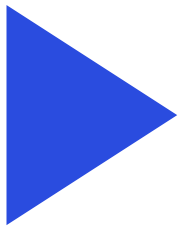
Creating and publishing datasets

- Datasets are used to group entries, making it easier to manage related data.
  - Datasets can include entries from several uploads.
  - The same entry can be in several datasets.
- Datasets have no influence on the processing of data.
- Users can get a DOI for their datasets.

Creating and publishing datasets

Hands-on tasks:

- ❑ Create a new dataset, and include entries into the dataset.
- ❑ Explore and manage datasets.
- ❑ Assign a DOI to your datasets.



Lets go to NOMAD, click here!