



# FAIR research data with NOMAD

FAIRmat's distributed, schema-based research-data infrastructure to harmonize RDM in materials science

Markus Scheidgen

NOMAD, FAIRmat, IRIS, Physik, HU Berlin

# Agenda

- Research data management and FAIR data
- Structured data and data schemas
- A distributed scalable and extendable data infrastructure

# NFDI goals

## FAIR

Findable, Accessible, Interoperable,  
Re-usable

FAIR principles can transform the  
field of condensed-matter physics  
and the chemical physics of solids.

## Open

Use open processes to support a  
wide community

FAIRmat advocates for an  
urgently needed culture shift  
towards data sharing, and stands  
for open access to scientific  
materials data and tools.

## Bottom-up approach

Embracing existing communities  
and data formats

FAIRmat follows an approach  
that is driven by the needs of  
scientists and already enjoys  
strong support from the  
community.

Feb. 7th.  
M393.885.  
M536.976.

C337.810.  
C348.728.  
C222.333

Feb. 8th.  
1st cloner.  
Tf.  
LT.  
Tf. sid  
LT. sid  
6 eggs in  
C377.  
Hobs.

M366.982  
M558.976.

Feb. 9th  
M364.982  
M20.330.13  
M536.976.2  
M373.859.1

— . — . ' .  
( , . — . ' .  
' ) ) .  
' — . ' . ' .  
| |  
/ \  
/ \  
/ \  
' . — . ' .  
' — . ' .

jgs /




\*\*\*\*\*

Space group  
Internationa  
Schoenflies:  
Index  
1 :  
2 :  
Info: The sy  
\*\*\*\*\*

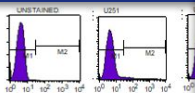
37044 k-poi  
-----  
n = 42  
s1 = 0

index |  
1 |  
2 |  
3 |  
4 |  
5 |  
6 |

This program is  
it under the te  
the Free Softwa  
any later versi

This program is  
it under the te  
the Free Softwa  
any later versi



4



NOMAD

mynotebook.L... - JupyterLab

nomad-lab.eu/prod/v1/staging/north/user/mscheidgen/jupyter/lab/tree/uploads/electronic-lab-notebook-az135h6MSLmtvOTTHJ3yow/mynot...

File Edit View Run Kernel Tabs Settings Help

Filter files by name

uploads / electronic-lab-notebook-az135h6MSLmtvOTTHJ3yow /

Name	Last Modified
mynotebo...	7 minutes ago
PVD-Parc...	2 days ago
PVDProce...	2 days ago
README.md	2 days ago
sample.arc...	2 days ago
schema.ar...	2 days ago
Tin_I_Sele...	2 days ago
Zinc_Sele...	2 days ago

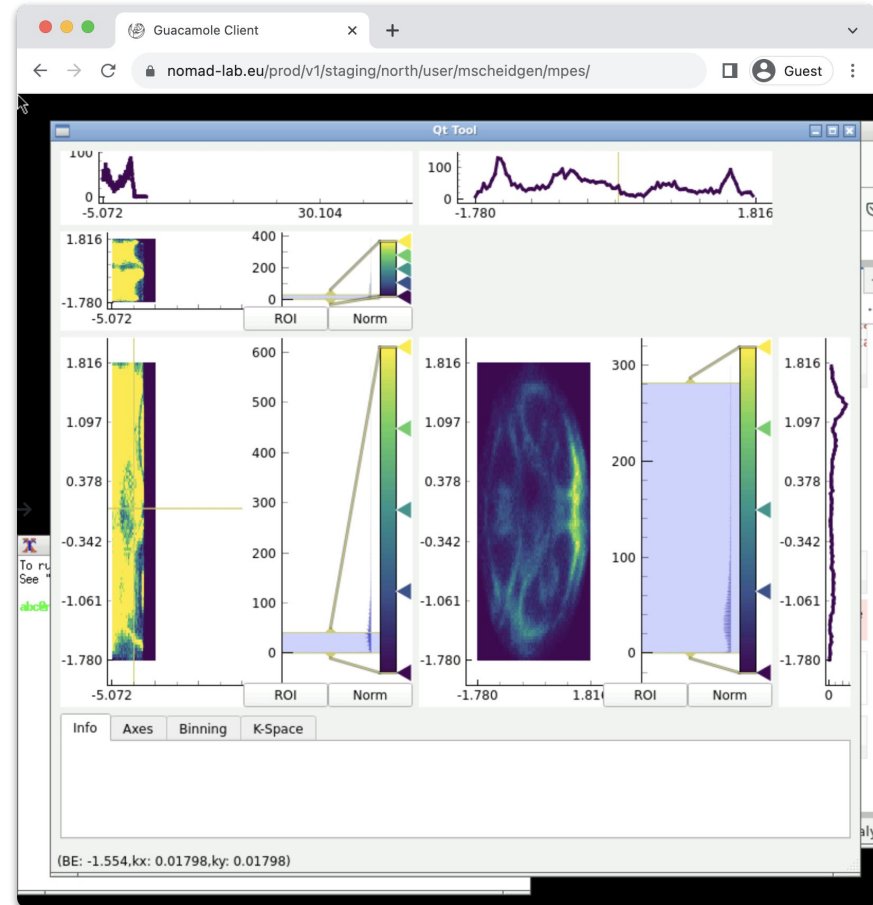
```

[ ]: from nomad.client import ArchiveQuery
    from nomad.metainfo import units

[ ]: query = ArchiveQuery(
    query={
        'results.method.simulation.program_name': 'VASP',
        'results.material.elements': ['Ti', 'O'],
        'results.properties.geometry_optimization': {
            'final_energy_difference:lt': 1e-22,
        }
    },
    required={
        'workflow': {
            'calculation_result_ref': {
                'energy': '*',
                'system_ref': {
                    'chemical_composition_reduced': '*'
                }
            }
        }
    }
)

[ ]: for result in query.download(100):
    calc = result.workflow[0].calculation_result_ref
    formula = calc.system_ref.chemical_composition_reduced
    if calc.energy.total:
        total_energy = calc.energy.total.value.to(units.eV)
    else:
        total_energy = 'N/A'
    print(f'{formula}: {total_energy}')
  
```

Simple 0 1 Python 3 (ipykernel) | Idle Mode: Command Ln 1, Col 1 mynotebook.ipynb 1



NOMAD

nomad-lab.eu/prod/v1/gui/user/uploads/upload/fd/az135h6MSLmtvOTTHJ3yow/entry/fid/w0iZ-J-qaRfZQSxjdgVkwRWstZvy/data/data/processes/pvd\_evaporation

[PUBLISH](#)
[EXPLORE](#)
[ANALYZE](#)
[ABOUT](#)

Your uploads / Upload / Entry / Data

Welcome [Markus Scheidgen](#) [LOGOUT](#) [UNITS](#)

OVERVIEW FILES **DATA** LOGS

search

**Entry** **Sample** **Processes** **PvdEvaporation**

section

SUB SECTIONS

results

metadata

**data**

REFERENCED BY **closed**

section

QUANTITIES

Name

ELN example sample

Chemical formula

Datetime

13/03/2023 14:44

Tags

project

Tags

ID

001

Chemicals

Copper\_II\_Selenide.archive.json

Chemicals

Tin\_II\_Selenide.archive.json

Chemicals

Zinc\_Selenide.archive.json

Chemicals

Description

A simple example for an "sample" that demonstrates how to combine different data entities.

section

SUB SECTIONS

**pvd\_evaporation**

hotplate\_annealing

REFERENCED BY **closed**

section

QUANTITIES

☐ Fill archive from datafile

Starting Time

10/05/2022 09:20

Instrument

PVD-Parchive.json

data\_file

PVDProcess.csv

Ending Time

Location

time = 9647 vector


chamber\_pressure = 9647 vector

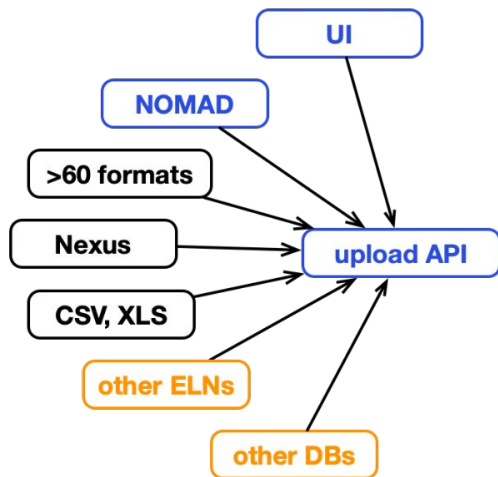
substrate\_temperature = 9647 vector

SUB SECTIONS

users

PLOT

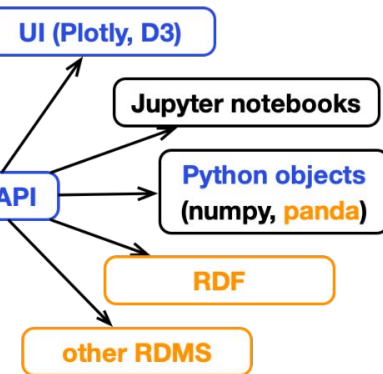


*data upload**data specification***structured data**

NOMAD specific

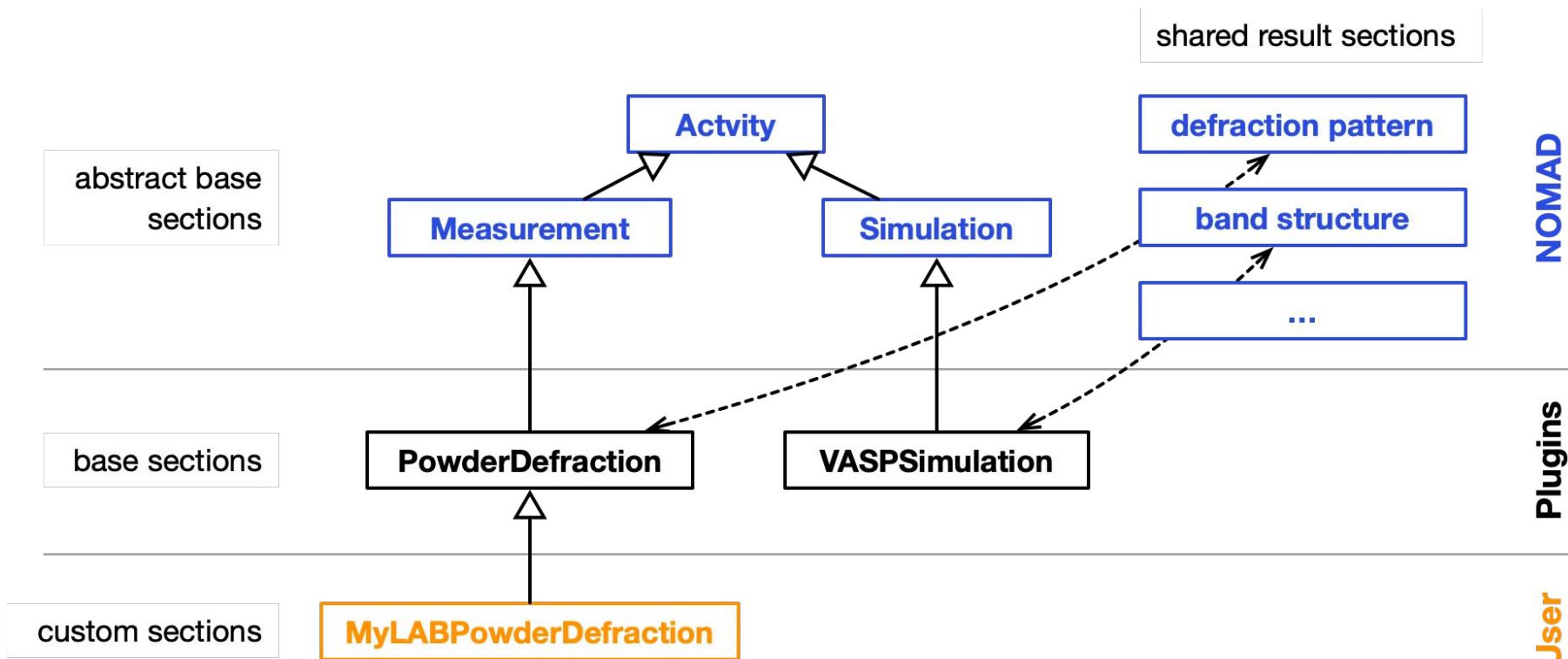
already supported

support planned

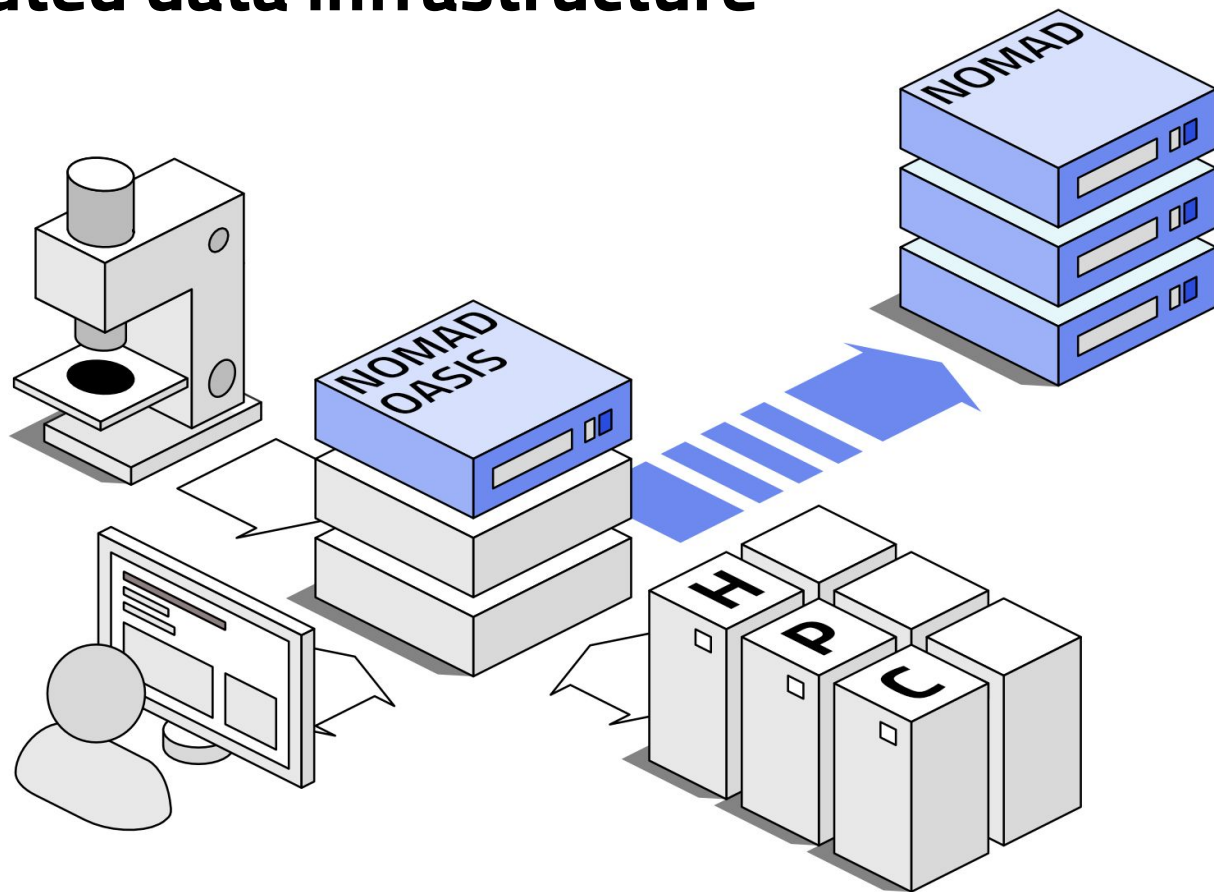
**data API***data access**data storage*

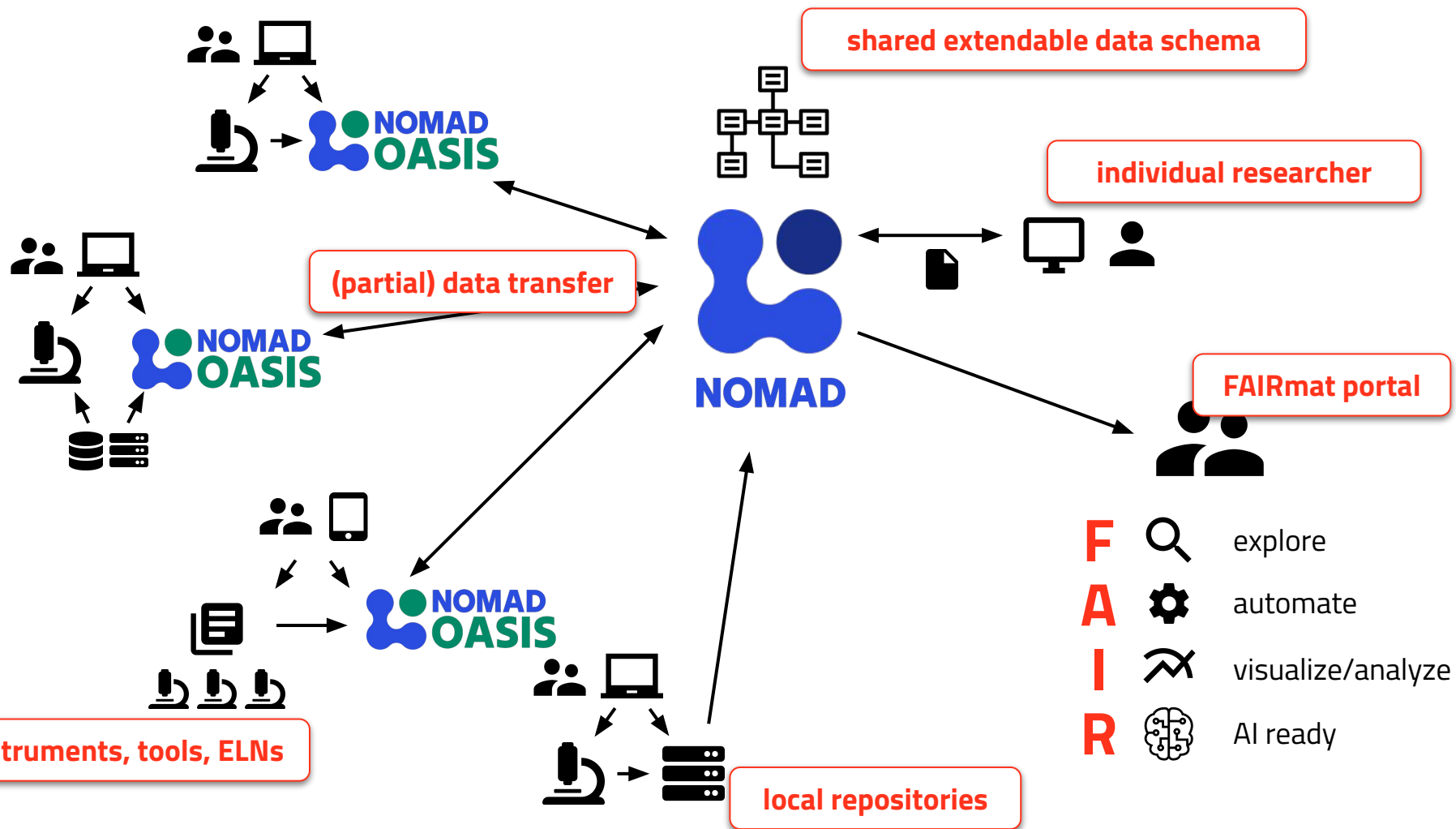


# Harmonization through abstraction



# A distributed data infrastructure






# Summary

- FAIR, open, bottom-up
- Structured data, harmonization through abstraction
- A distributed scalable and extendable data infrastructure

NOMAD software development is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation); project 460197019 (FAIRmat NFDI consortium).

# Questions?



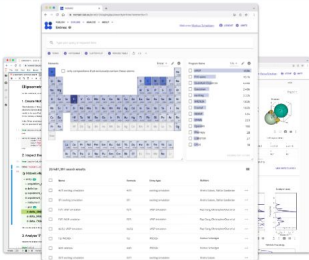
SOLUTIONS ▾ LEARN ▾ GET INVOLVED ▾ ABOUT ▾ [OPEN NOMAD](#)

## NOMAD

### Materials science data managed and shared

NOMAD lets you manage and share your materials science data in a way that makes it truly useful to you, your group, and the community.

[Open NOMAD →](#)



USED BY THOUSANDS OF MATERIALS SCIENTISTS

UPLOADED ENTRIES  
**12,460,881**


REPRESENTED MATERIALS  
**2,976,441**

UPLOADED FILES  
**108.5 TB**

USE CASES

### More than sharing files

NOMAD processes your files to extract structured data and rich metadata. It provides a unified way to Find, Access, Interoperate with, and Reuse millions of FAIR data from different codes, sources, and workflows.

PUBLISH ▾ EXPLORE ▾ ANALYZE ▾ ABOUT ▾Welcome [Markus Scheidgen](#) [LOGOUT](#) [UNITS](#)

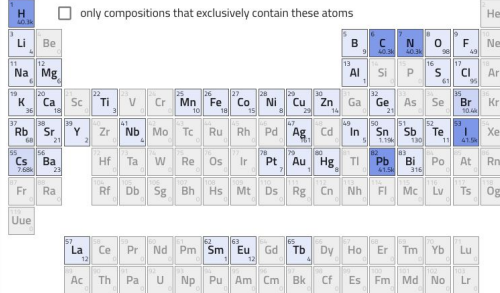
### Solar Cells

Type your query or keyword here

[TERMS](#) [HISTOGRAM](#) [SCATTER PLOT](#) [PERIODIC TABLE](#)

Filters

only compositions that exclusively contain these atoms



#### Solar Cell Absorber Fabricati...

Type here

- ☐ Spin-coating 29.4k
- ☐ Spin-coating >> Spin-coating 5.49k

SHOWING TOP 2 ITEMS

#### Solar Cell Device Stack

Type here

- ☐ Perovskite 42.6k
- ☐ SLG 41.6k

SHOWING TOP 2 ITEMS

#### Scatter plot

☒ autorange

Efficiency

Open Circuit Voltage (V)

Short Circuit Current Density ( $A/A^2$ )

#### Scatter plot

☒ autorange

Efficiency

Open Circuit Voltage (V)

pin  
nip  
Unknown  
Schottky

#### Solar Cell Electron...

linear

#### Solar Cell Hole Tra...

linear

#### Band Gap Value (eV)

☐ autorange 1/4

# NOMAD Links

- homepage – [nomad-lab.eu](https://nomad-lab.eu)
- central NOMAD service – <https://nomad-lab.eu/prod/v1>
- [NOMAD Dokumentation](#)
  - [installing Oasis](#)
  - [schemas and ELNs](#)
- Tutorials and videos
  - [NOMAD Oasis tutorial](#) video playlist
  - [Publishing, Exploring, API](#) video playlist
  - All FAIRmat [tutorials](#)
- [NOMAD forum](#)
- NOMAD's [main gitlab](#) (and [github](#))