NOMAD / FAIRmat

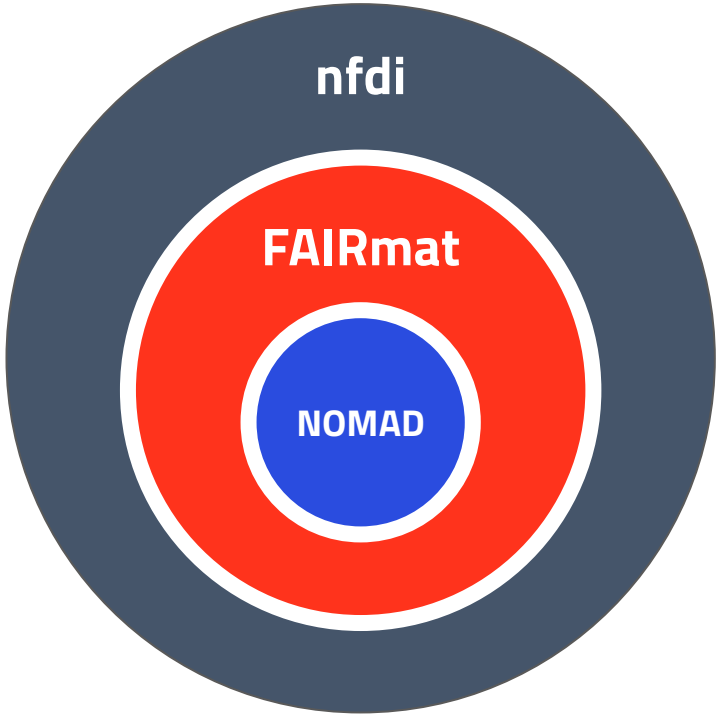
Enabling material scientists to define, share, and use data

Markus Scheidgen
NOMAD, FAIRmat, IRIS, Physik, HU Berlin

Agenda

- What is NFDI, FAIRmat, and NOMAD
- How does NOMAD work?
- Demo
- Comparison to other systems

What are NFDI / FAIRmat / NOMAD



nfdi: Nationale Forschungsdaten Infrastructure, link
(national research data infrastructure)

FAIRmat: NFDI consortium for FAIR materials science data, link
(FAIR: findable, accessible, interoperable, re-usable)

NOMAD: A web-based service and software for managing FAIR materials science data, link
FAIRmat uses NOMAD to build a federated infrastructure of connected NOMAD installations

FAIRmat values

FAIR

Findable, Accessible, Interoperable, Re-usable

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

Open access

Use open processes to support a wide community

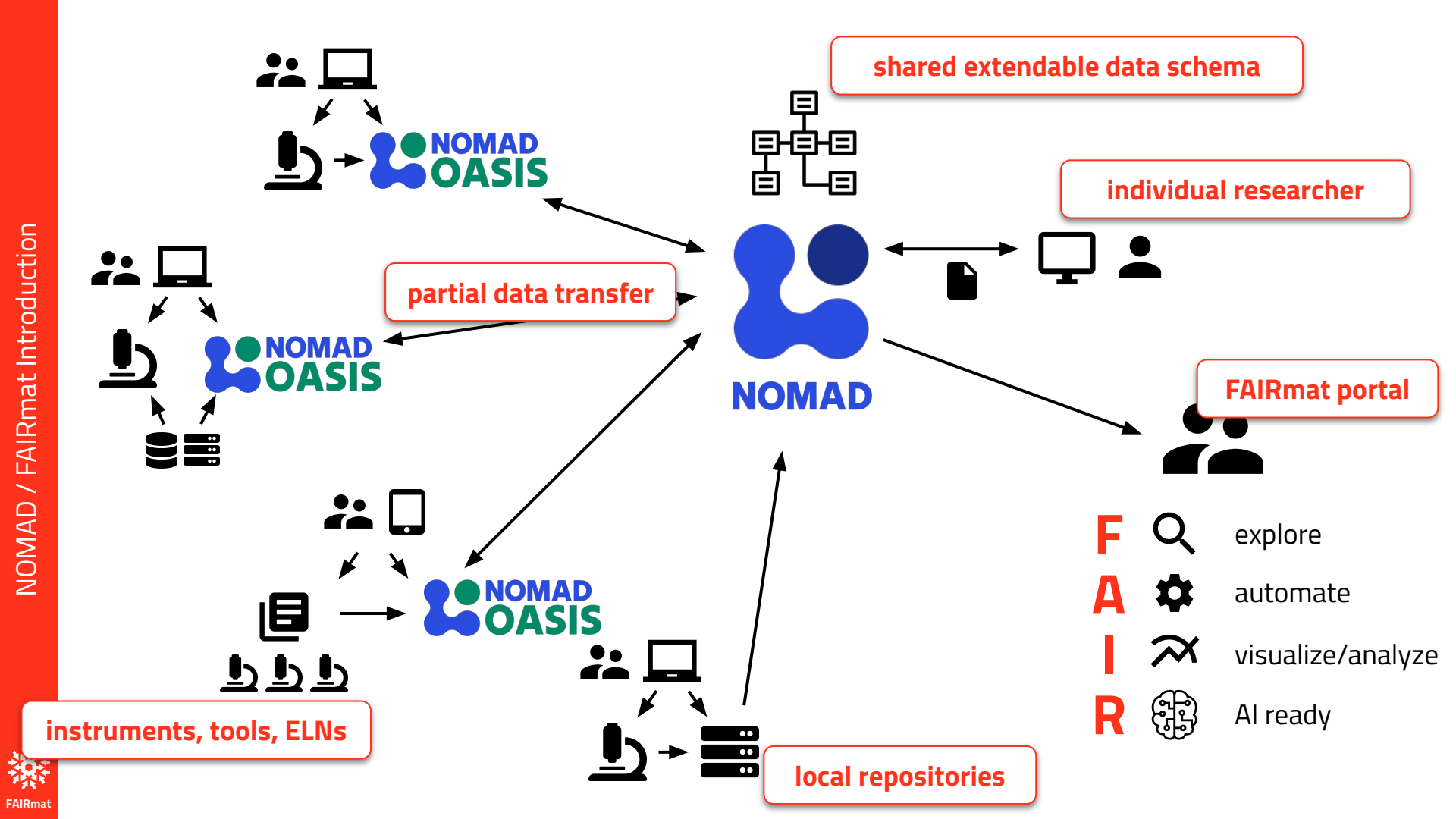
FAIRmat advocates for an urgently needed culture shift towards data sharing, and stands for open access to scientific materials data and tools.

Bottom-up approach

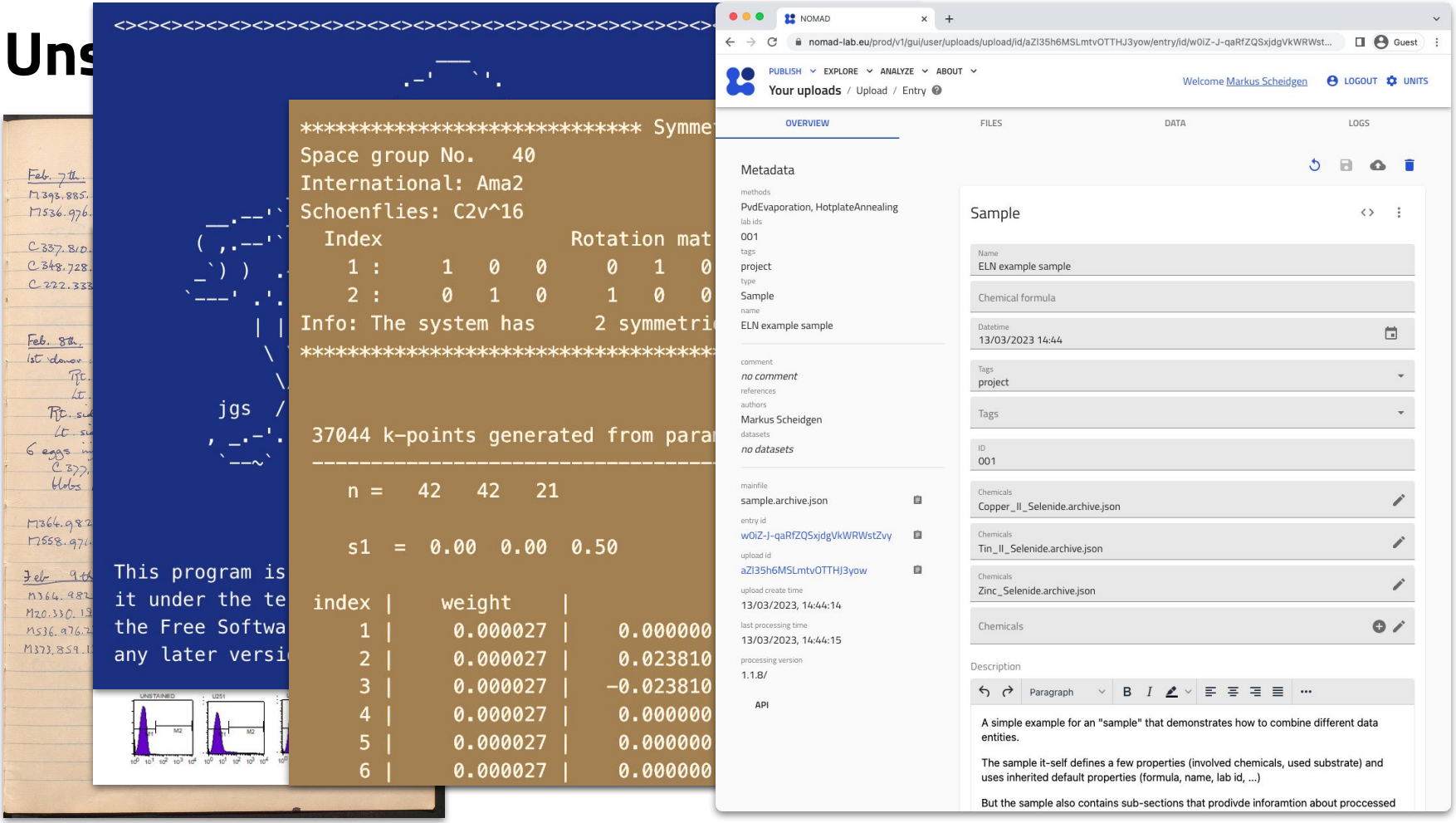
Embracing the community

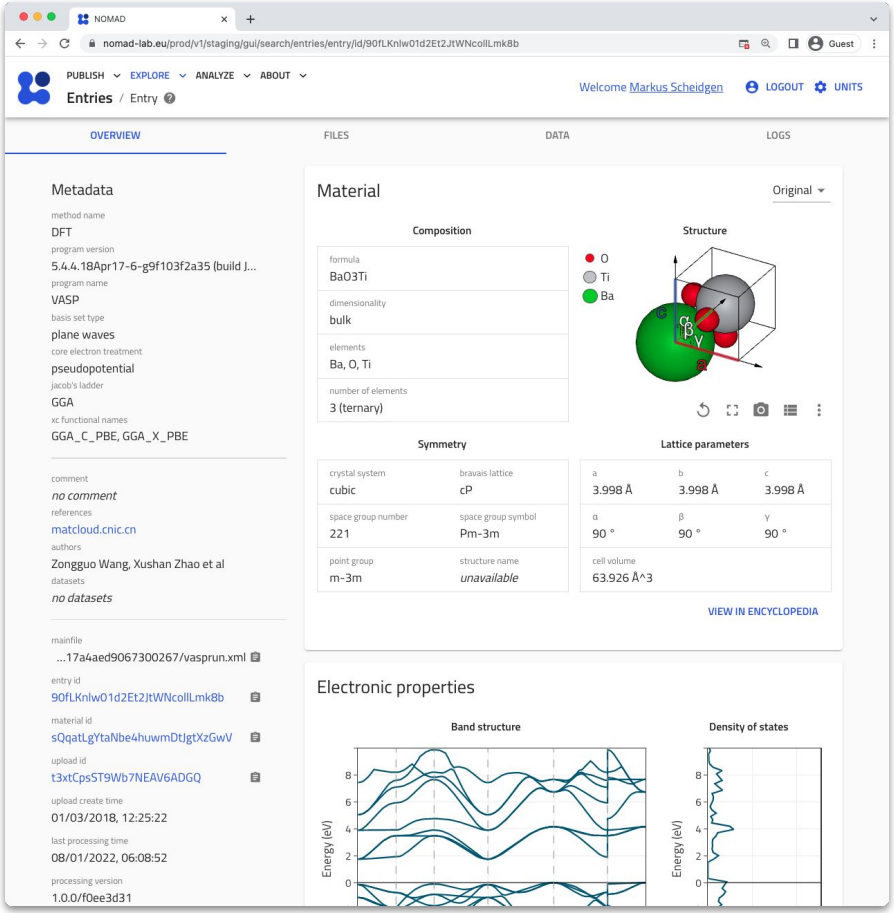
FAIRmat follows an approach that is driven by the needs of scientists and already enjoys strong support from the community.

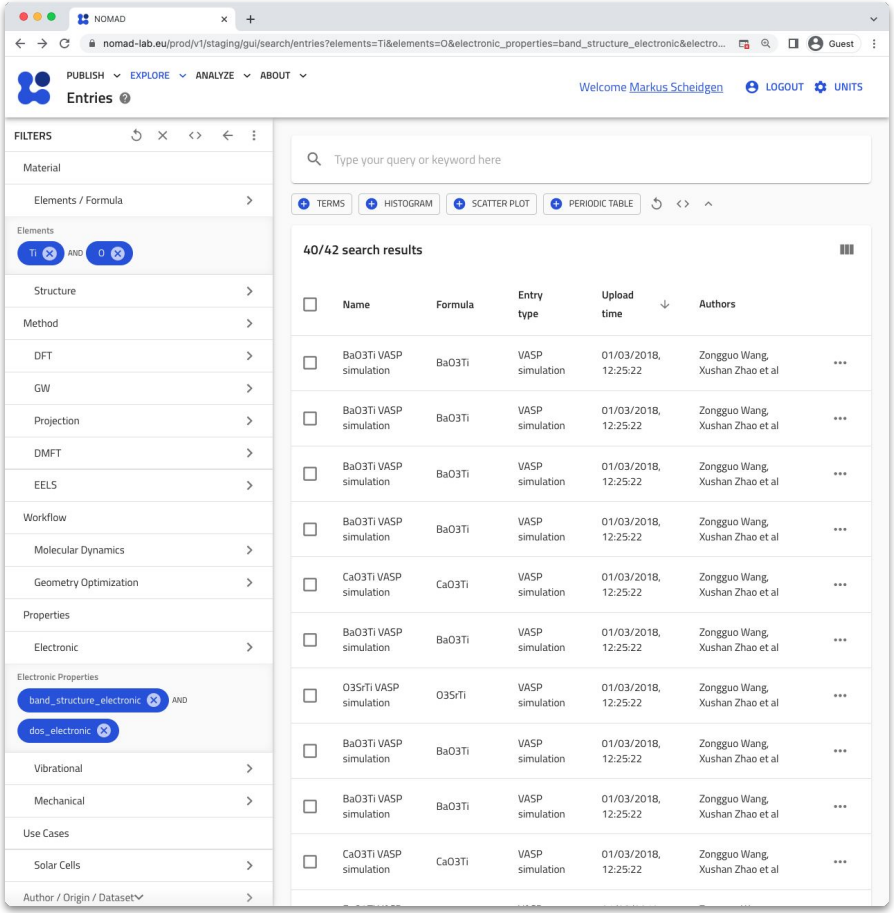
FAIRmat structure


shared extendable data schema
individual researcher
partial data transfer
NOMAD OASIS
NOMAD OASIS
NOMAD OASIS
NOMAD
FAIRmat portal
F explore
A automate
I visualize/analyze
R AI ready
instruments, tools, ELNs
local repositories

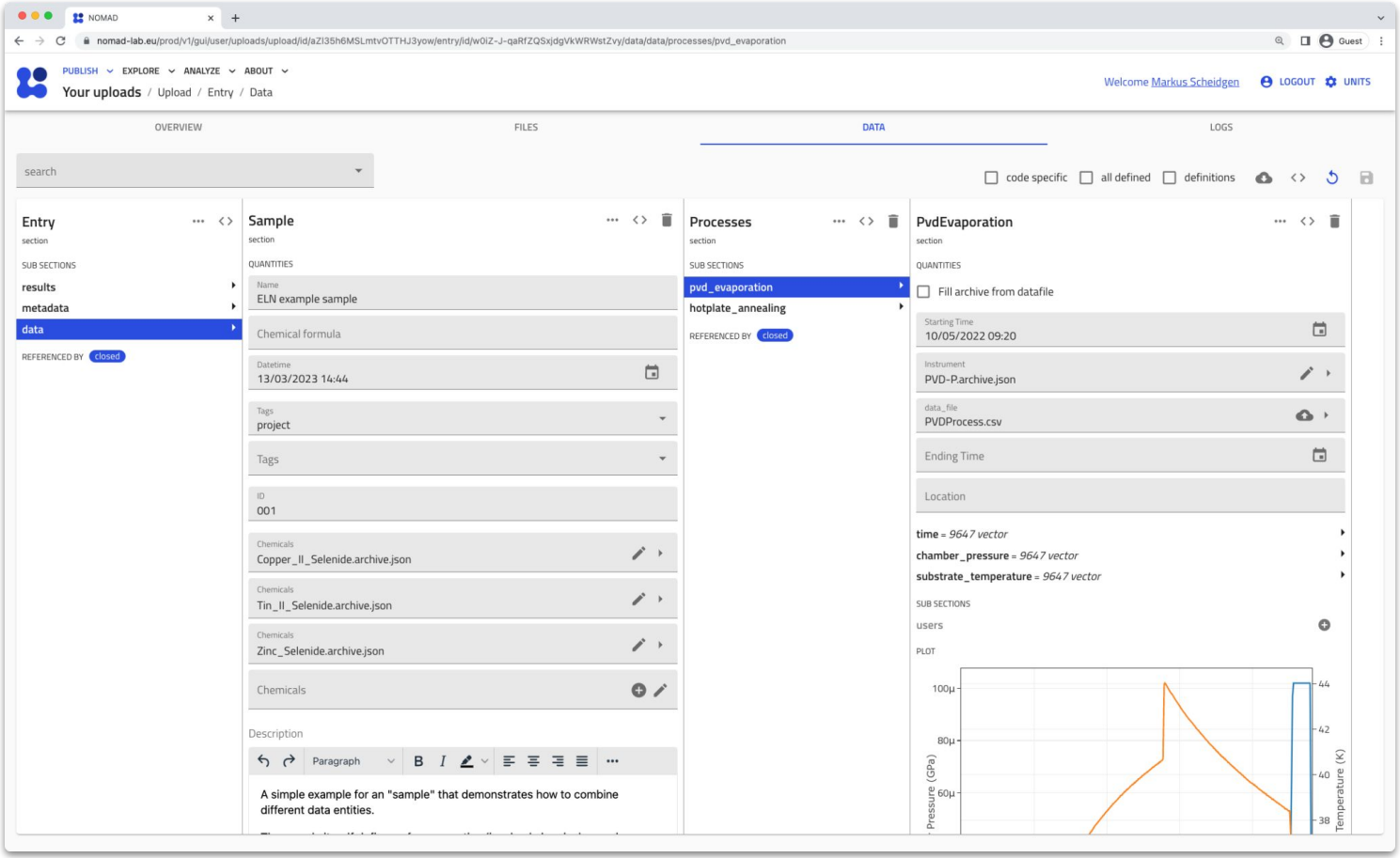
Uns


Space group No. 40
International: Ama2
Schoenflies: C2v^16
Index Rotation mat
1 : 1 0 0 0 1 0
2 : 0 1 0 1 0 0
Info: The system has 2 symmetri
37044 k-points generated from para
n = 42 42 21
s1 = 0.00 0.00 0.50
index | weight |
1 | 0.000027 | 0.000000
2 | 0.000027 | 0.023810
3 | 0.000027 | -0.023810
4 | 0.000027 | 0.000000
5 | 0.000027 | 0.000000
6 | 0.000027 | 0.000000
This program is
it under the te
the Free Softwa
any later versi
jgs
NOMAD
nomad-lab.eu/prod/v1/gui/user/uploads/upload/id/aZI35h6MSLmtvOTTHJ3yow/entry/id/w0iZ-J-qaRfZQSxjdgVkWRWst...
PUBLISH EXPLORE ANALYZE ABOUT
Your uploads / Upload / Entry
Welcome Markus Scheidgen LOGOUT UNITS
OVERVIEW FILES DATA LOGS
Metadata
methods
PvdEvaporation, HotplateAnnealing
lab ids
001
tags
project
type
Sample
name
ELN example sample
comment
no comment
references
authors
Markus Scheidgen
datasets
no datasets
mainfile
sample.archive.json
entry id
w0iZ-J-qaRfZQSxjdgVkWRWstZvy
upload id
aZI35h6MSLmtvOTTHJ3yow
upload create time
13/03/2023, 14:44:14
last processing time
13/03/2023, 14:44:15
processing version
1.1.8/
API
Sample
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
Paragraph
A simple example for an "sample" that demonstrates how to combine different data entities.
The sample it-self defines a few properties (involved chemicals, used substrate) and uses inherited default properties (formula, name, lab id, ...)
But the sample also contains sub-sections that prodivde inforamtion about proccessed

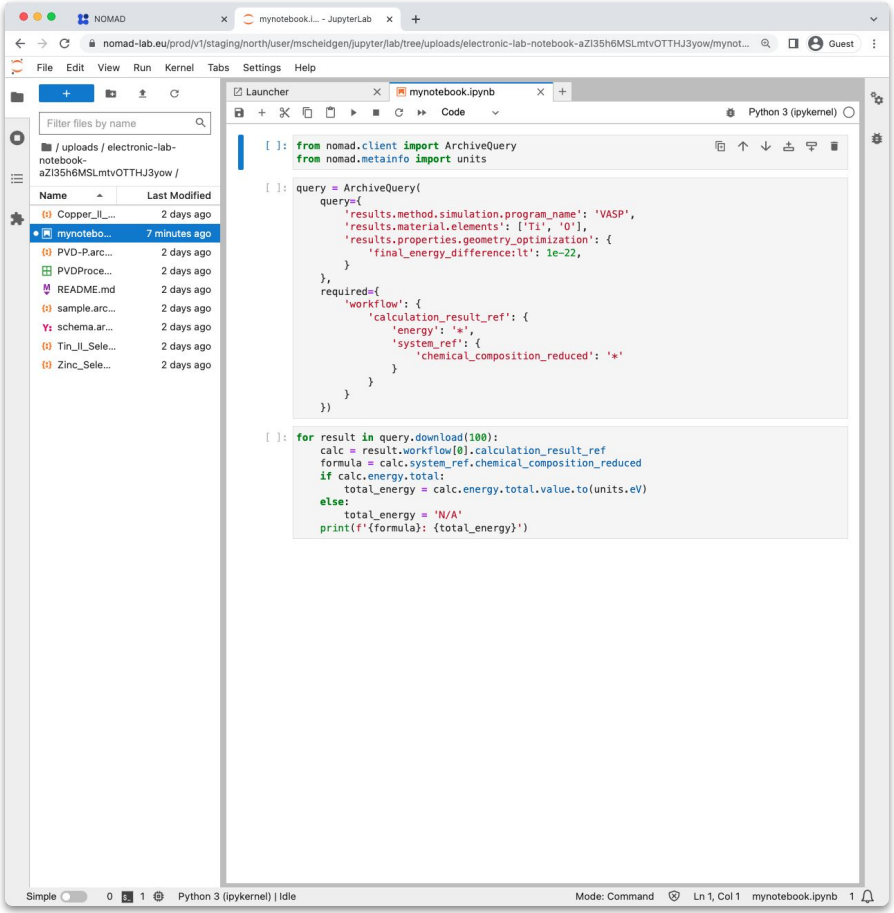

NOMAD
nomad-lab.eu/prod/v1/staging/gui/search/entries/entry/id/90fLKnlw01d2Et2JtWNcollLmk8b
PUBLISH EXPLORE ANALYZE ABOUT
Entries / Entry
Welcome Markus Scheidgen
LOGOUT
UNITS
OVERVIEW
FILES
DATA
LOGS
Metadata
method name
DFT
program version
5.4.4.18Apr17-6-g9f103f2a35 (build J...
program name
VASP
basis set type
plane waves
core electron treatment
pseudopotential
jacob's ladder
GGA
xc functional names
GGA_C_PBE, GGA_X_PBE
comment
no comment
references
matcloud.cnic.cn
authors
Zongguo Wang, Xushan Zhao et al
datasets
no datasets
mainfile
...17a4aed9067300267/vasprun.xml
entry id
90fLKnlw01d2Et2JtWNcollLmk8b
material id
sQqatLgYtaNbe4huwmDtJgtXzGwV
upload id
t3xtCpsST9Wb7NEAV6ADGQ
upload create time
01/03/2018, 12:25:22
last processing time
08/01/2022, 06:08:52
processing version
1.0.0/f0ee3d31
Material
Original
Composition
formula
BaO3Ti
dimensionality
bulk
elements
Ba, O, Ti
number of elements
3 (ternary)
Structure
O
Ti
Ba
Symmetry
crystal system
cubic
bravais lattice
cP
space group number
221
space group symbol
Pm-3m
point group
m-3m
structure name
unavailable
Lattice parameters
a
3.998 Å
b
3.998 Å
c
3.998 Å
α
90 °
β
90 °
γ
90 °
cell volume
63.926 Å^3
VIEW IN ENCYCLOPEDIA
Electronic properties
Band structure
Density of states
Energy (eV)


NOMAD
nomad-lab.eu/prod/v1/staging/gui/search/entries?elements=Ti&elements=O&electronic_properties=band_structure_electronic&electro...
PUBLISH EXPLORE ANALYZE ABOUT
Entries
Welcome Markus Scheidgen
LOGOUT
UNITS
FILTERS
Material
Elements / Formula
Elements
Ti
AND
O
Structure
Method
DFT
GW
Projection
DMFT
EELS
Workflow
Molecular Dynamics
Geometry Optimization
Properties
Electronic
Electronic Properties
band_structure_electronic
AND
dos_electronic
Vibrational
Mechanical
Use Cases
Solar Cells
Author / Origin / Dataset
Type your query or keyword here
TERMS
HISTOGRAM
SCATTER PLOT
PERIODIC TABLE
40/42 search results
Name
Formula
Entry type
Upload time
Authors
BaO3Ti VASP simulation
BaO3Ti
VASP simulation
01/03/2018, 12:25:22
Zongguo Wang, Xushan Zhao et al
CaO3Ti VASP simulation
CaO3Ti
O3SrTi VASP simulation
O3SrTi

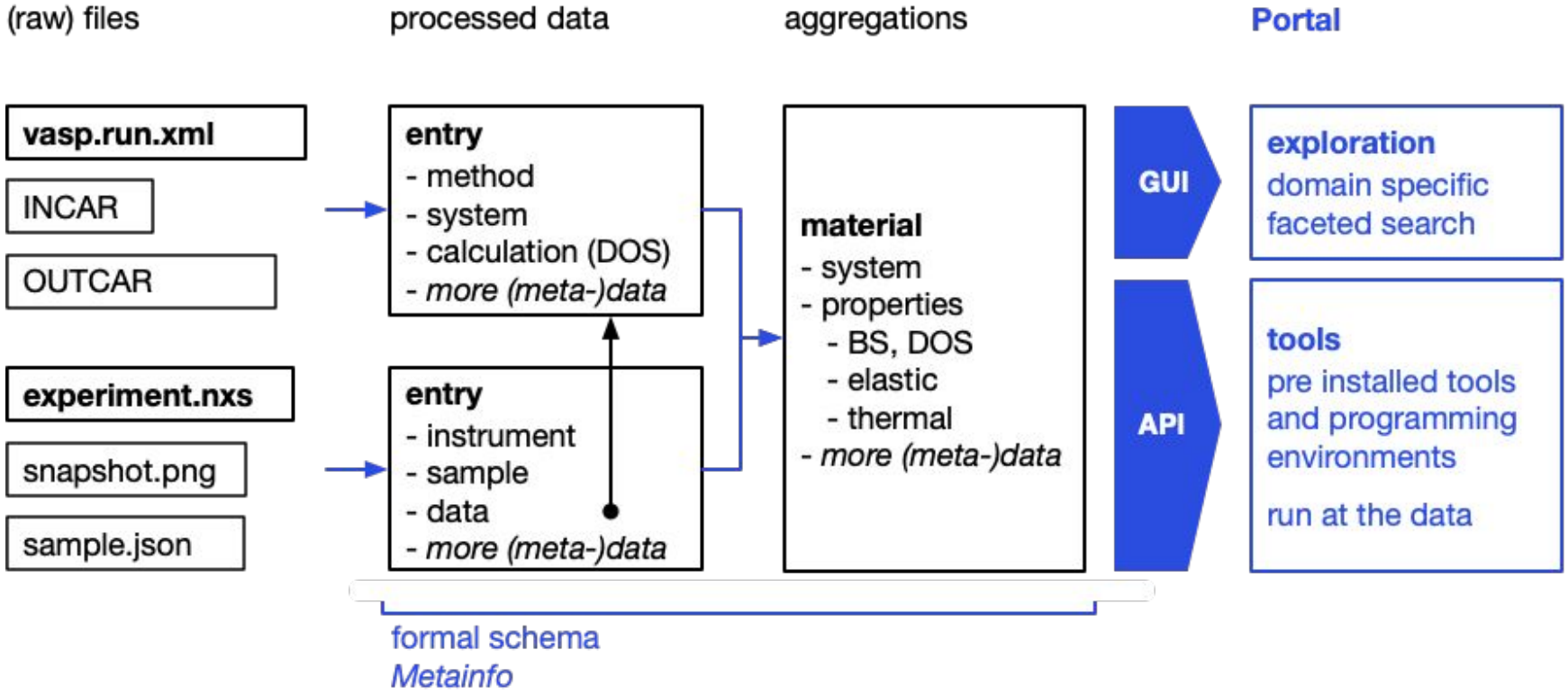
FAIRmat


NOMAD
nomad-lab.eu/prod/v1/gui/user/uploads/upload/id/aZI35h6MSLmtvOTTHJ3yow/entry/id/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
code specific all defined definitions
Entry
section
SUB SECTIONS
results
metadata
data
REFERENCED BY closed
Sample
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
Paragraph
A simple example for an "sample" that demonstrates how to combine different data entities.
Processes
section
SUB SECTIONS
pvd_evaporation
hotplate_annealing
REFERENCED BY closed
PvdEvaporation
section
QUANTITIES
Fill archive from datafile
Starting Time
10/05/2022 09:20
Instrument
PVD-P.archive.json
data_file
PVDProcess.csv
Ending Time
Location
time = 9647 vector
chamber_pressure = 9647 vector
substrate_temperature = 9647 vector
SUB SECTIONS
users
PLOT
100µ
80µ
60µ
Pressure (GPa)
44
42
40
38
Temperature (K)

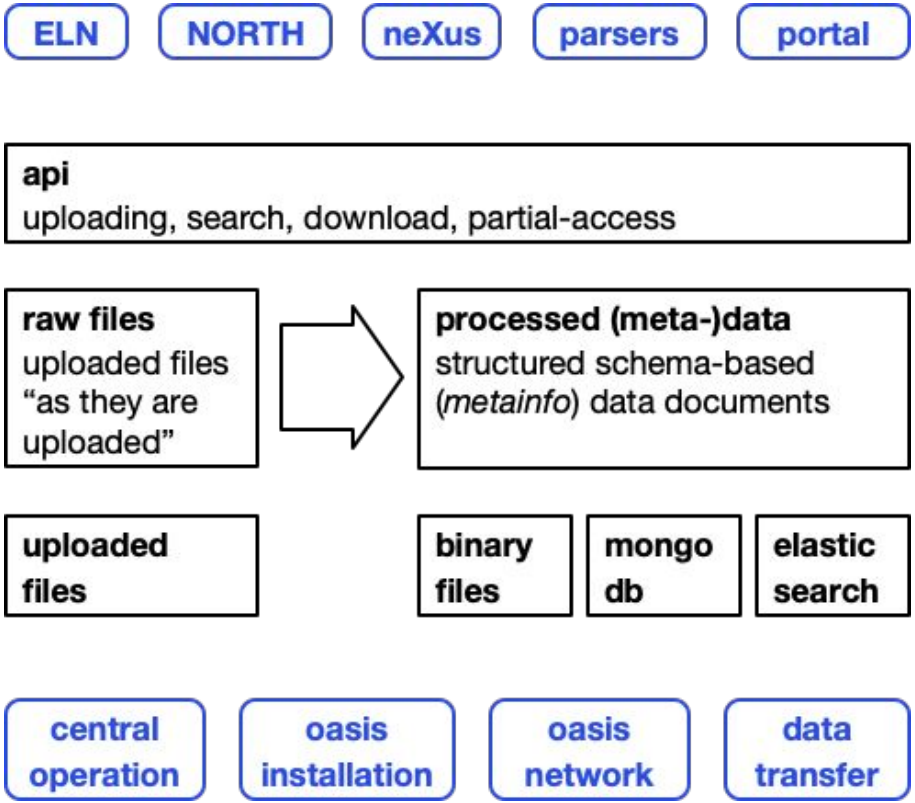
FAIRmat


NOMAD
mynotebook.i... - JupyterLab
nomad-lab.eu/prod/v1/staging/north/user/mscheidgen/jupyter/lab/tree/uploads/electronic-lab-notebook-aZI35h6MSLmtvOTTHJ3yow/mynot...
Guest
File Edit View Run Kernel Tabs Settings Help
Launcher
mynotebook.ipynb
Code
Python 3 (ipykernel)
Filter files by name
/ uploads / electronic-lab-notebook-aZI35h6MSLmtvOTTHJ3yow /
Name
Last Modified
Copper_II_... 2 days ago
mynotebo... 7 minutes ago
PVD-P.arc... 2 days ago
PVDProce... 2 days ago
README.md 2 days ago
sample.arc... 2 days ago
schema.ar... 2 days ago
Tin_II_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
Python 3 (ipykernel) | Idle
Mode: Command
Ln 1, Col 1
mynotebook.ipynb

Extracting structured (meta-)data from files



Extracting structured (meta-)data from files



demo

Pros

- **schema-focused**, produces consistent FAIR data
- supports growing number of file-types, e.g. **neXus**
- **file-based**, easy to import and export
- **API** covers all functionality
- integrated **JupyterHUB** to work with and analyse managed data
- data can be **shared** and **published** through a central NOMAD

Cons

- **schema-focused**, requires you to structure data first
- designed to be **extendable**, adapt NOMAD to your needs
- integration with others ELNs has just begun
- young constantly **evolving project**, not all features are there yet

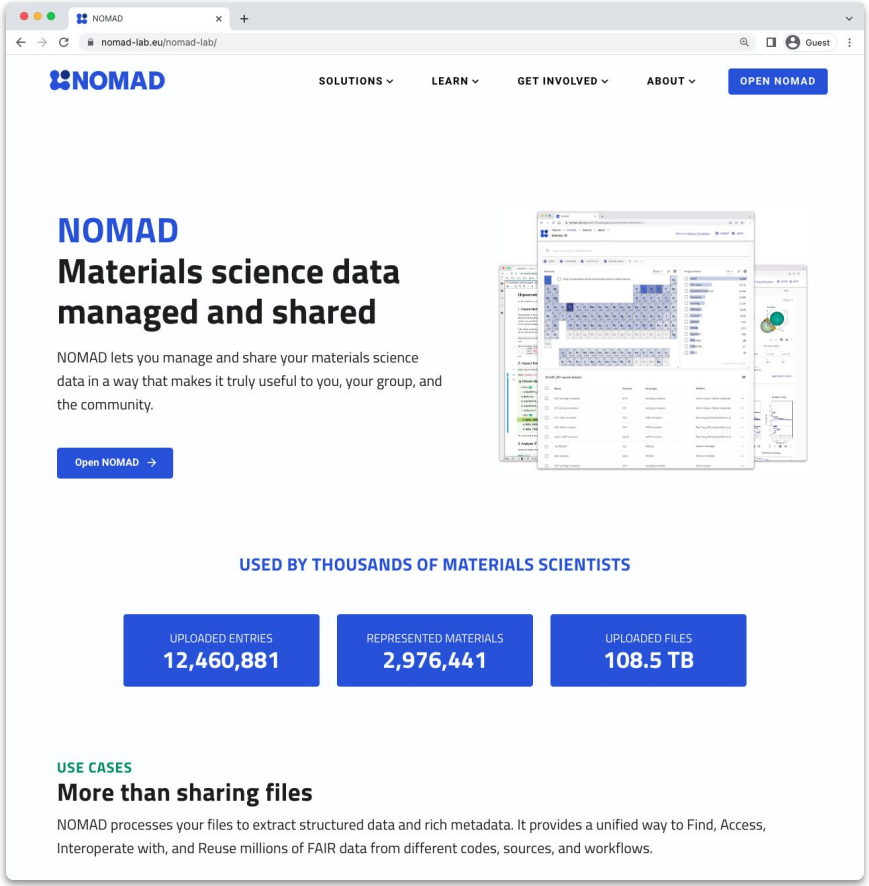
FAIR, extendable, scalable: 

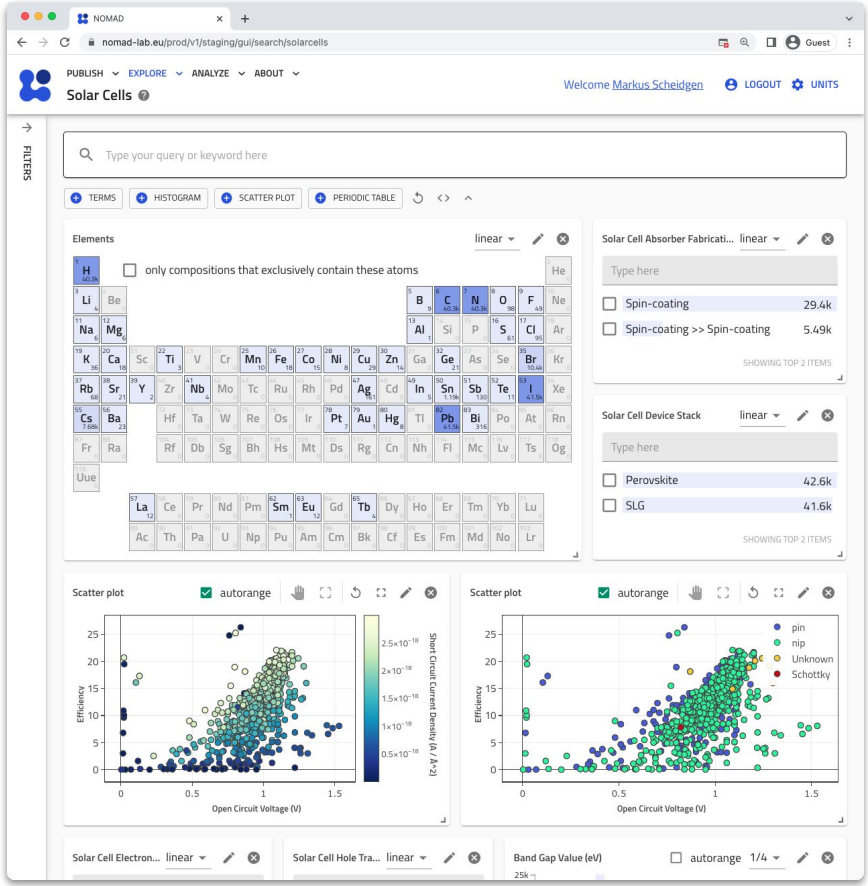
simple, low maintenance: 

NOMAD Links

- homepage – 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)

Questions?


NOMAD
nomad-lab.eu/nomad-lab/
NOMAD
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.


NOMAD
nomad-lab.eu/prod/v1/staging/gui/search/solarcells
PUBLISH
EXPLORE
ANALYZE
ABOUT
Solar Cells
Welcome Markus Scheidgen
LOGOUT
UNITS
FILTERS
Type your query or keyword here
TERMS
HISTOGRAM
SCATTER PLOT
PERIODIC TABLE
Elements
linear
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
Perovskite 42.6k
SLG 41.6k
Scatter plot
autorange
Efficiency
Open Circuit Voltage (V)
Short Circuit Current Density (A / m^2)
pin
nip
Unknown
Schottky
Solar Cell Electron...
Solar Cell Hole Tra...
Band Gap Value (eV)
1/4