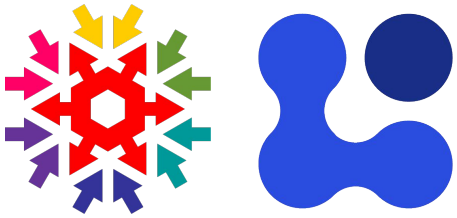


FAIR research data with NOMAD

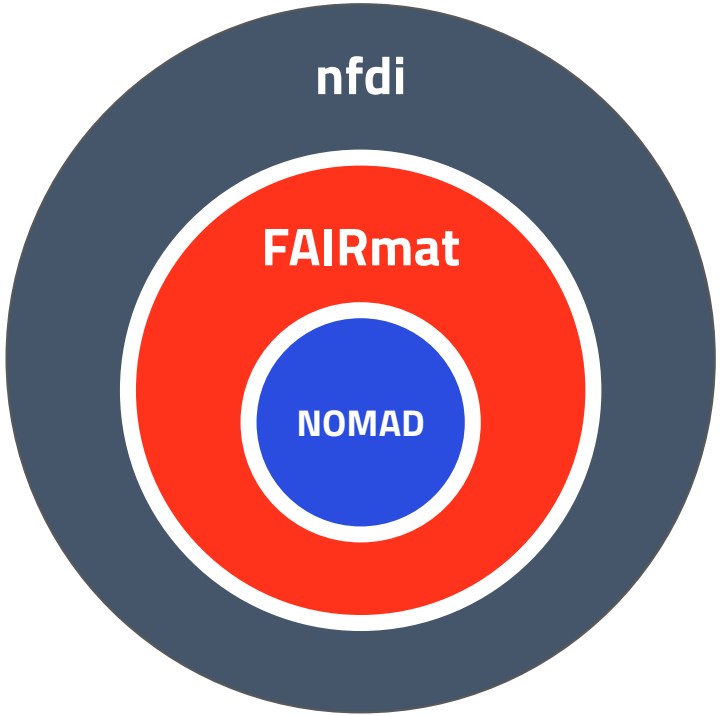
Operating NOMAD: A FAIR data service on the MPCDF HPC Cloud with Kubernetes

Markus Scheidgen

Agenda

- ndfi, FAIRmat, and NOMAD
- Why kubernetes
- NOMAD at the HPC Cloud
- From commit to deployment
- Application monitoring

What are nfdi / FAIRmat / NOMAD

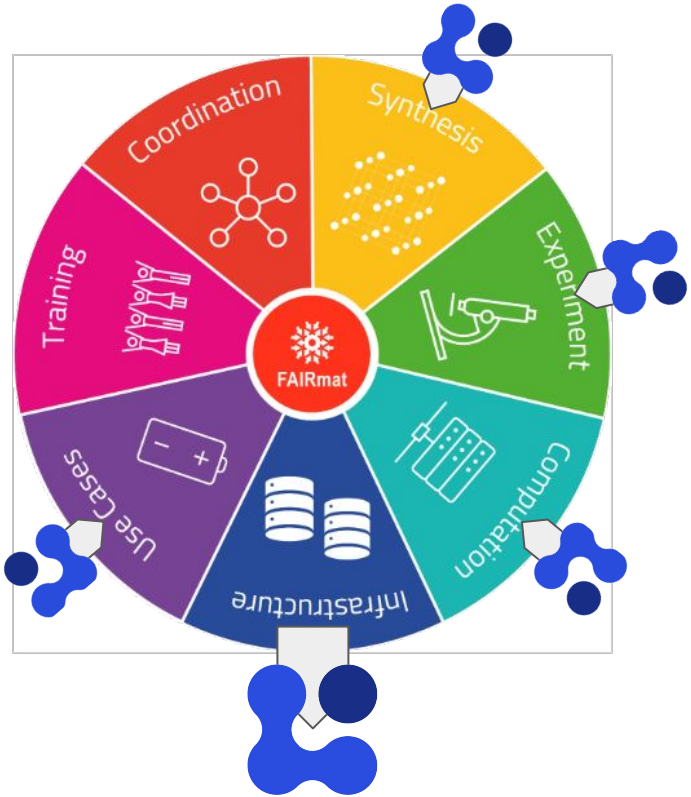


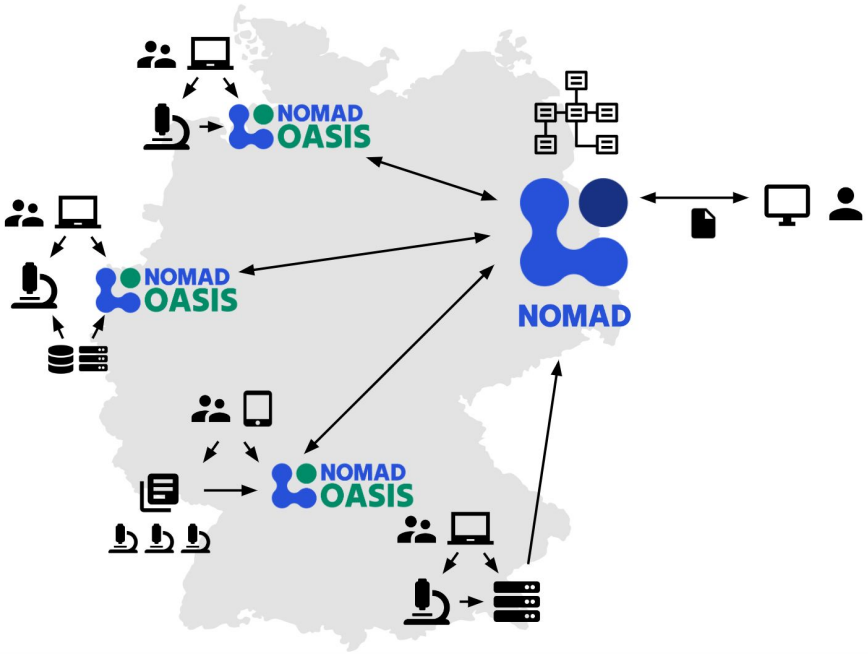
nfdi: Nationale Forschungsdaten Infrastructure, link
(national research data infrastructure)

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

What are nfdi / FAIRmat / NOMAD





MPCDF HPC Cloud Workshop

FAIRmat

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/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... linear

Type here

Spin-coating 29.4k

Spin-coating >> Spin-coating 5.49k

SHOWING TOP 2 ITEMS

Solar Cell Device Stack linear

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

pin
nip
Unknown
Schottky

Efficiency

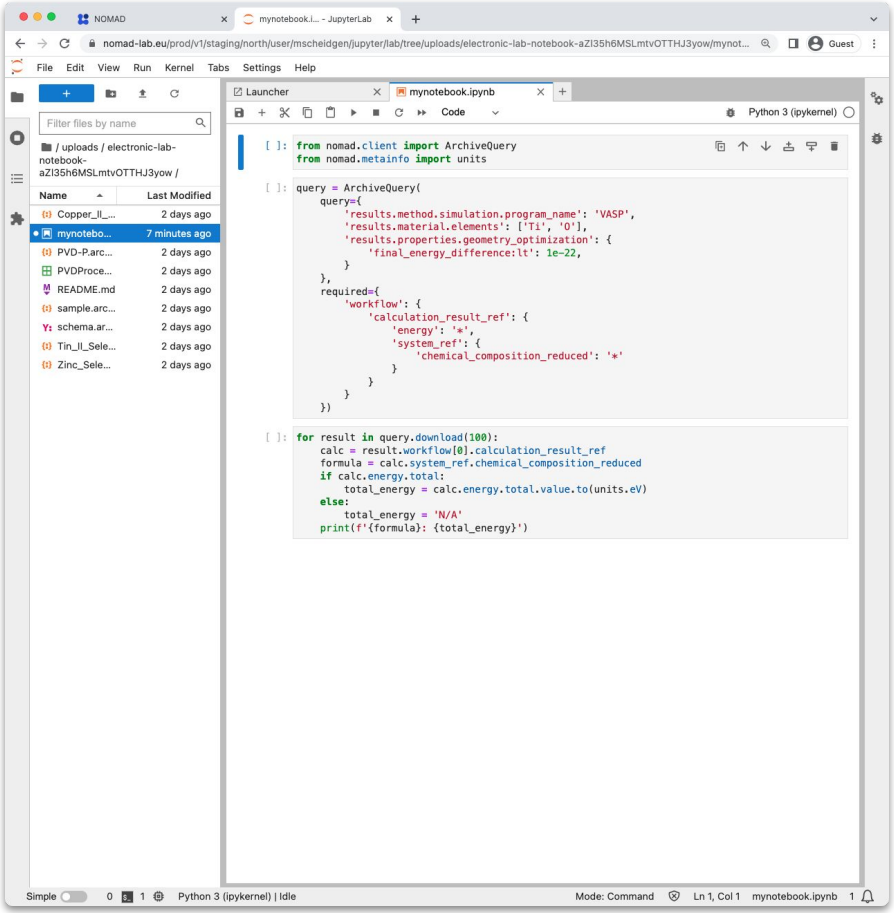
Open Circuit Voltage (V)

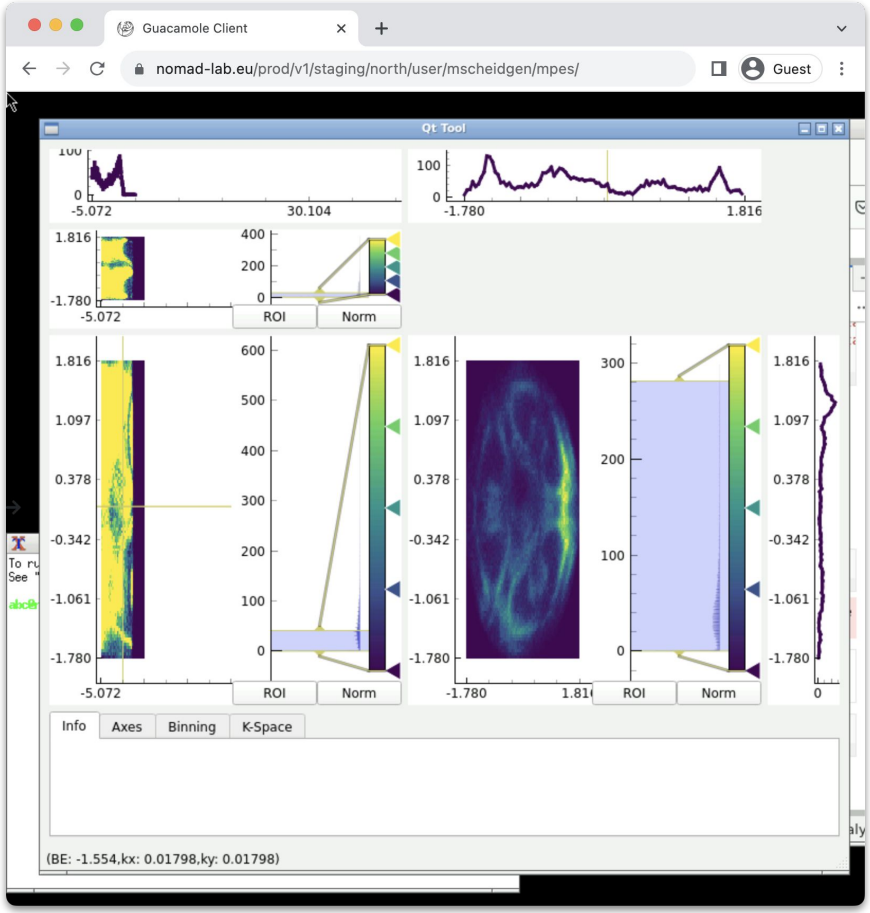
Solar Cell Electron... linear

Solar Cell Hole Tra... linear

Band Gap Value (eV) autorange 1/4

MPCDF HPC Cloud Workshop





FAIRmat

NOMAD

nomad-lab.eu/prod/v1/gui/search/entries/entry/id/sCltpLugL04XS2EVLsv6Us6Eznjl/data/run/system/atoms/positions

PUBLISH EXPLORE ANALYZE ABOUT

Entries / Entry / Data

Welcome Markus Scheidgen LOGOUT UNITS

OVERVIEW FILES DATA LOGS

search

code specific all defined definitions

Entry
section

SUB SECTIONS
- results
- metadata
- workflow
- run

REFERENCED BY closed

Run
section

SUB SECTIONS
- program
- method
- system
- calculation

REFERENCED BY closed

System
section

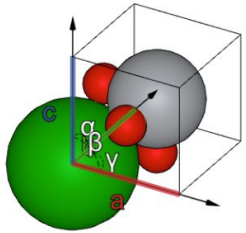
QUANTITIES
- type = bulk
- configuration_raw_gid = gSWzJ-w8k8S3xs3i_x34kF6VPcJU
- is_representative = true
- chemical_composition = BaO0OTi
- chemical_composition_hill = BaO3Ti
- chemical_composition_reduced = BaO3Ti

SUB SECTIONS
- atoms
- springer_material
- symmetry

REFERENCED BY closed

Atoms
section

O
Ti
Ba



QUANTITIES
- species = *5 vector*
- labels = *5 list*
- positions = *5 × 3 matrix*
- lattice_vectors = *3 × 3 matrix*
- periodic = *3 list*

REFERENCED BY closed

positions
quantity

VALUE

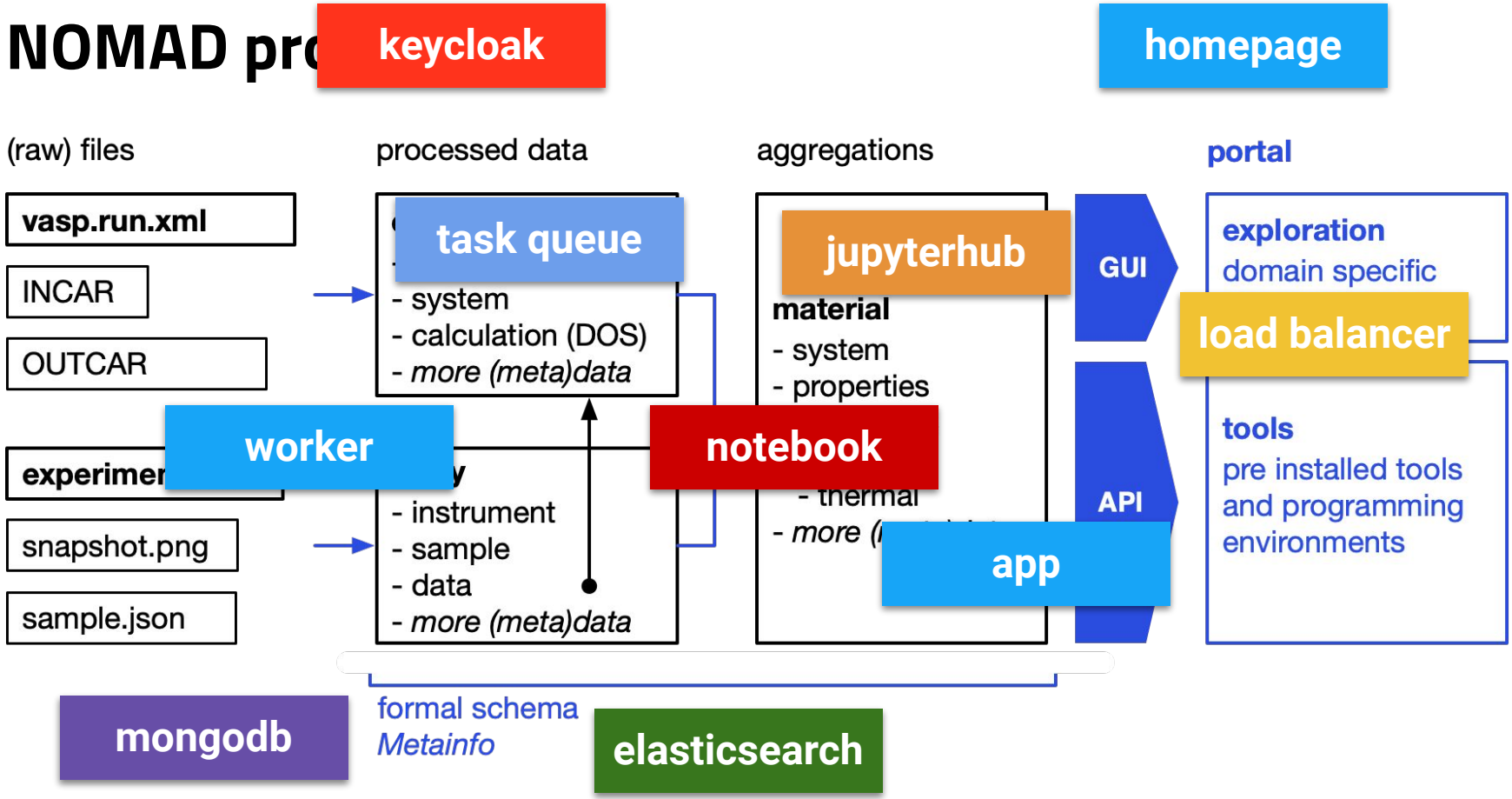
0	0	2
0	2	0
0	2	2

(n_atoms × 3)

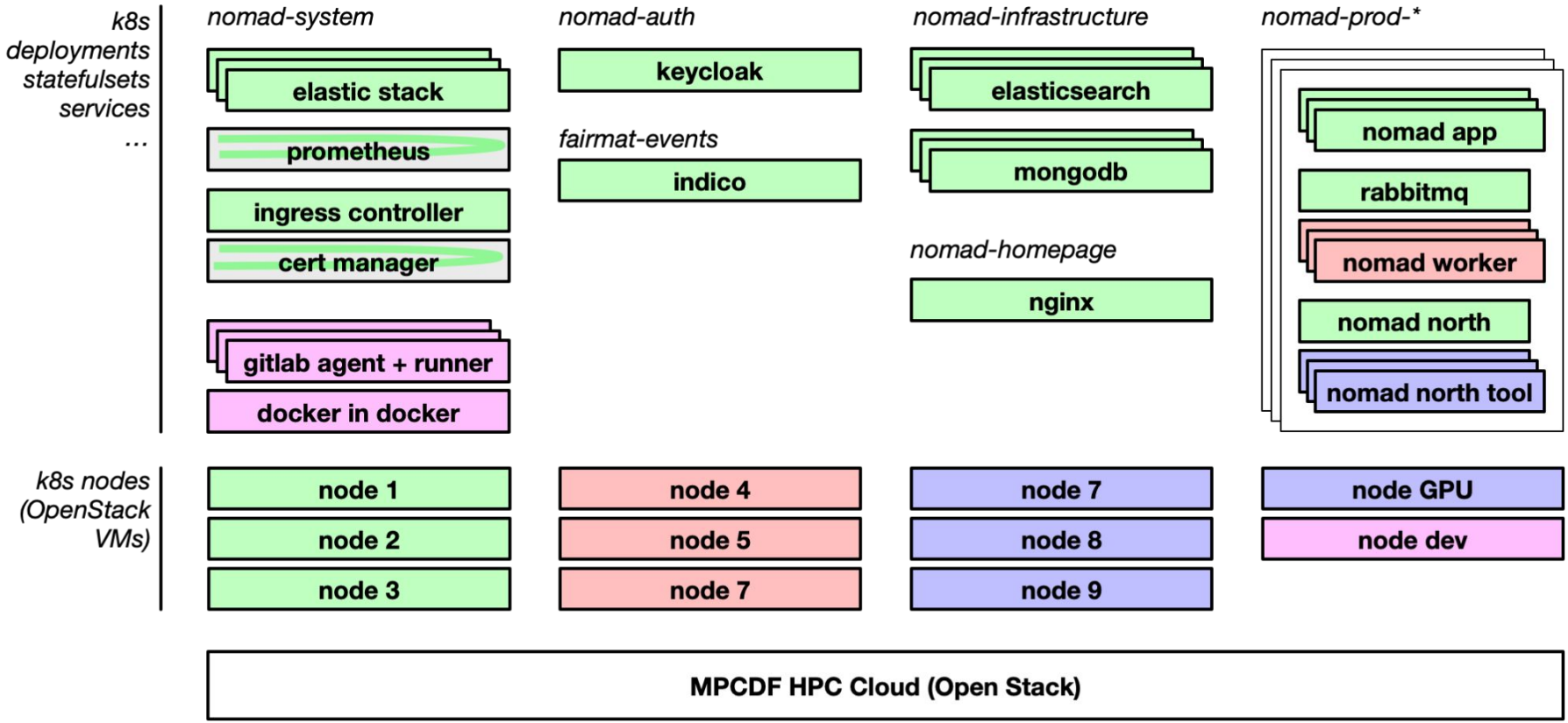
Å

REFERENCED BY closed

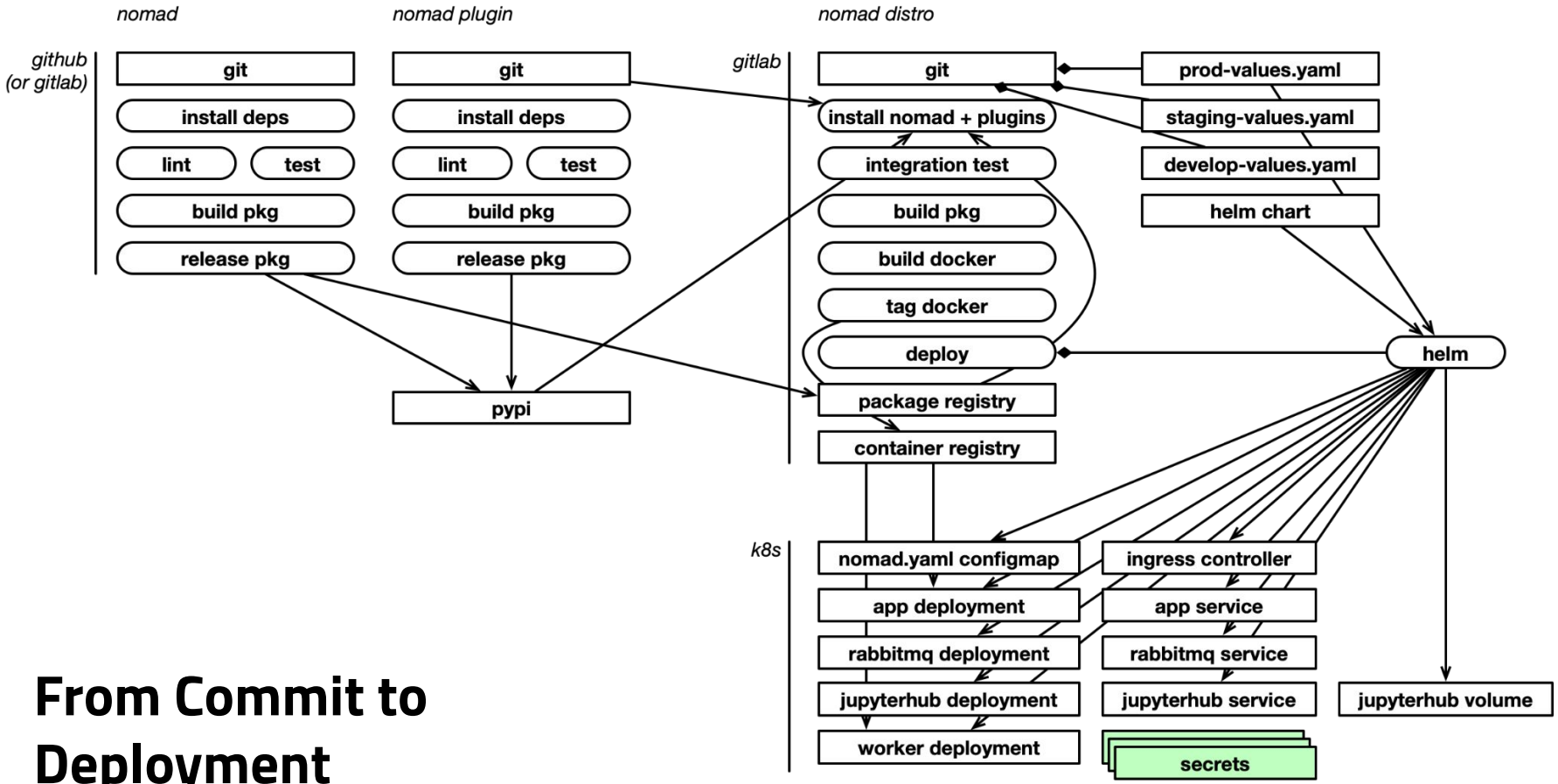
NOMAD pro


keycloak
homepage
(raw) files
processed data
aggregations
portal
vasp.run.xml
INCAR
OUTCAR
experimen
snapshot.png
sample.json
task queue
- system
- calculation (DOS)
- more (meta)data
worker
- instrument
- sample
- data
- more (meta)data
jupyterhub
material
- system
- properties
- thermal
- more (
notebook
GUI
API
app
exploration
domain specific
load balancer
tools
pre installed tools
and programming
environments
mongodb
formal schema
Metainfo
elasticsearch

NOMAD on the HPC Cloud

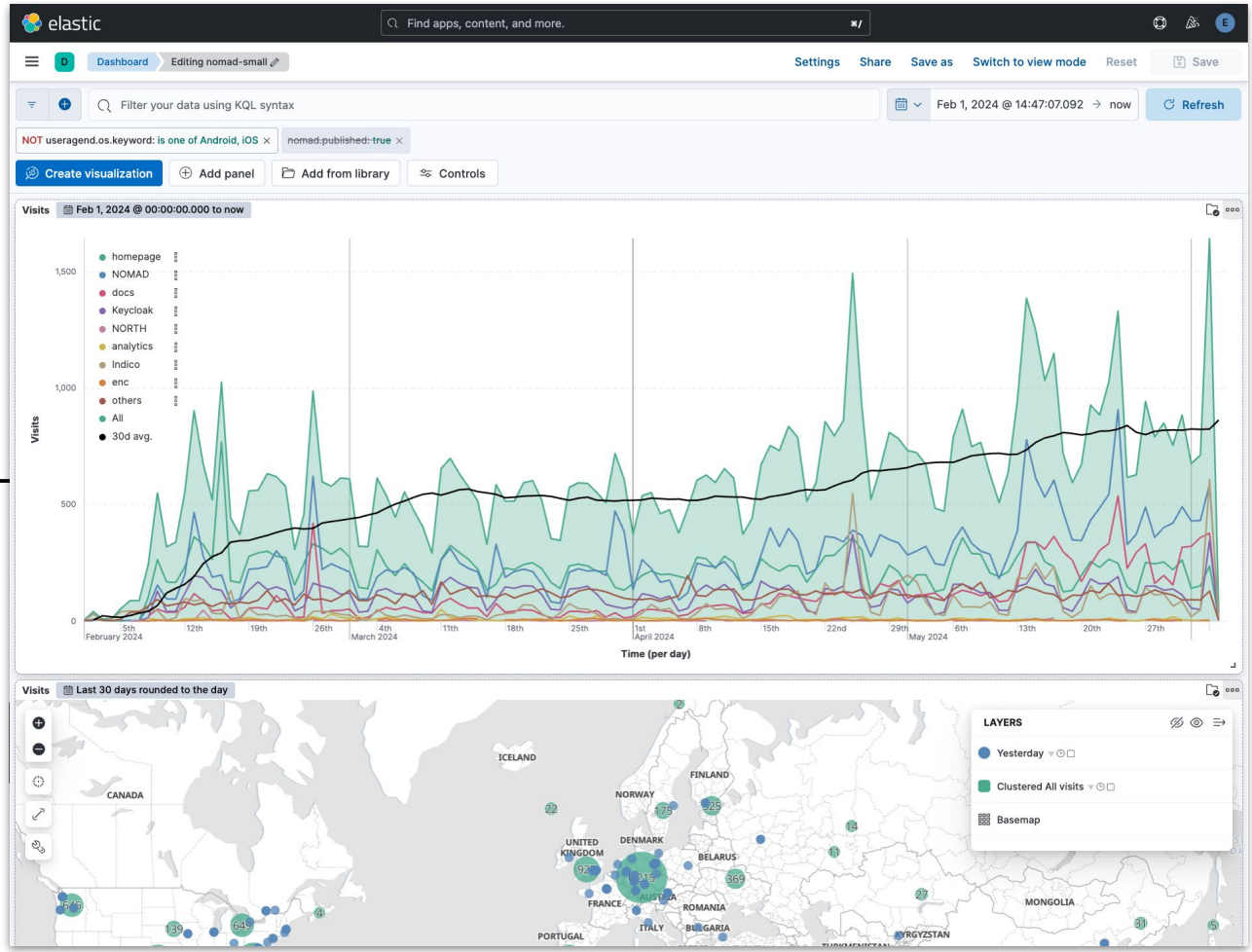


From Commit to Deployment



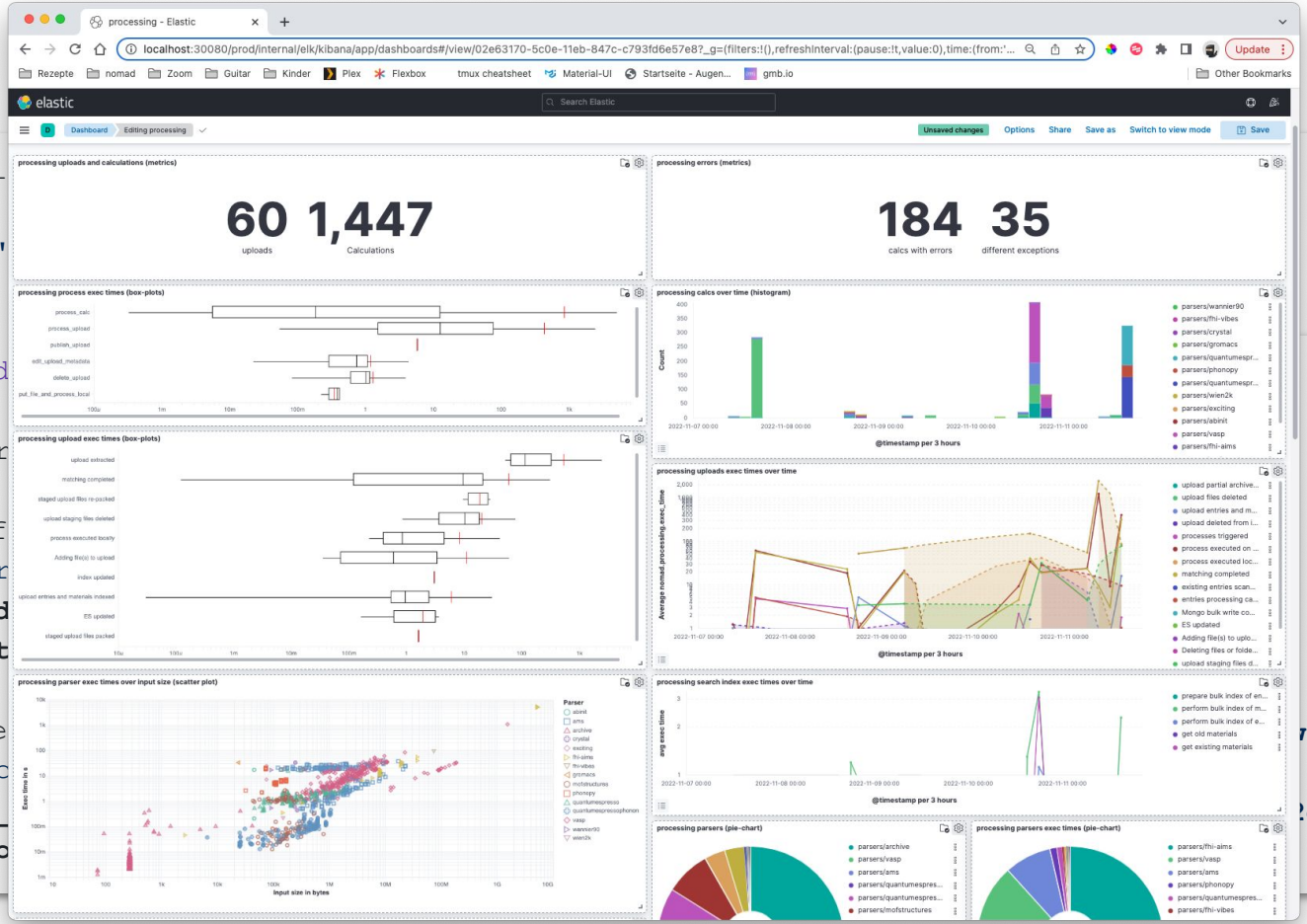
Elasticstack

keycloak



Structlog

```
1.   logger = get_l
2.       __name__,
3.       api="upload"
4.
5.   @app.get
6.   def get_upload
7.       try:
8.           with timer
9.               ...
10.          logger.inf
11.              'retur
12.              upload
13.              exec_t
14.      except e:
15.          logger.e
16.              'exc
17.              exc_
18.              uplo
```



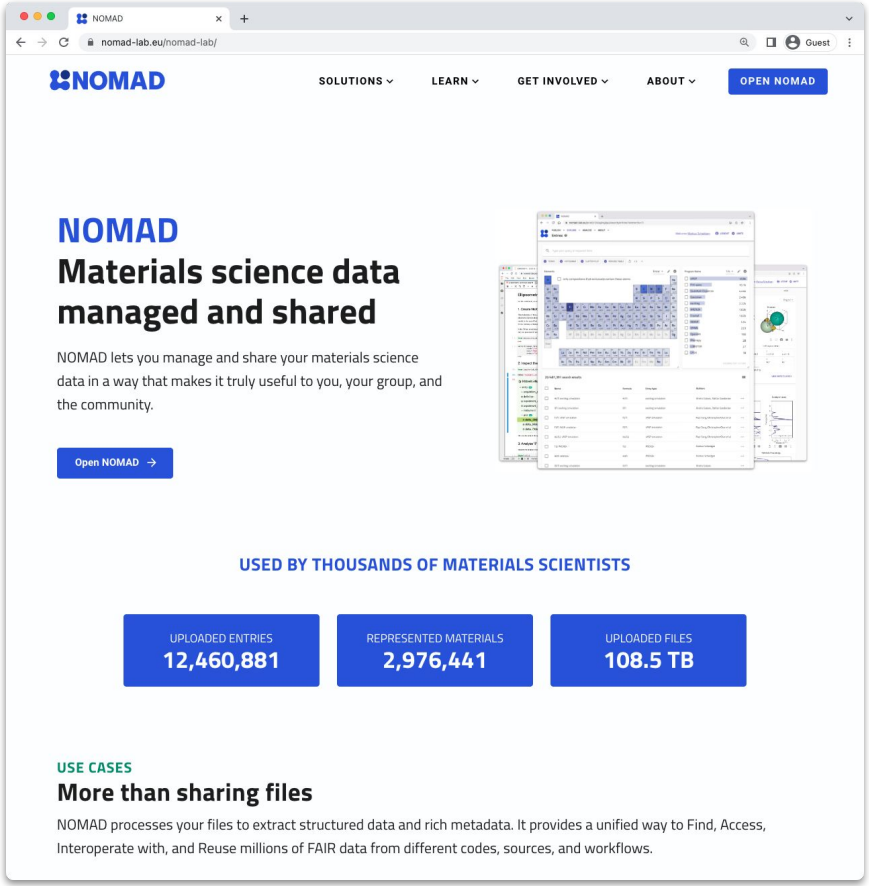
MPCDF HPC Cloud Workshop

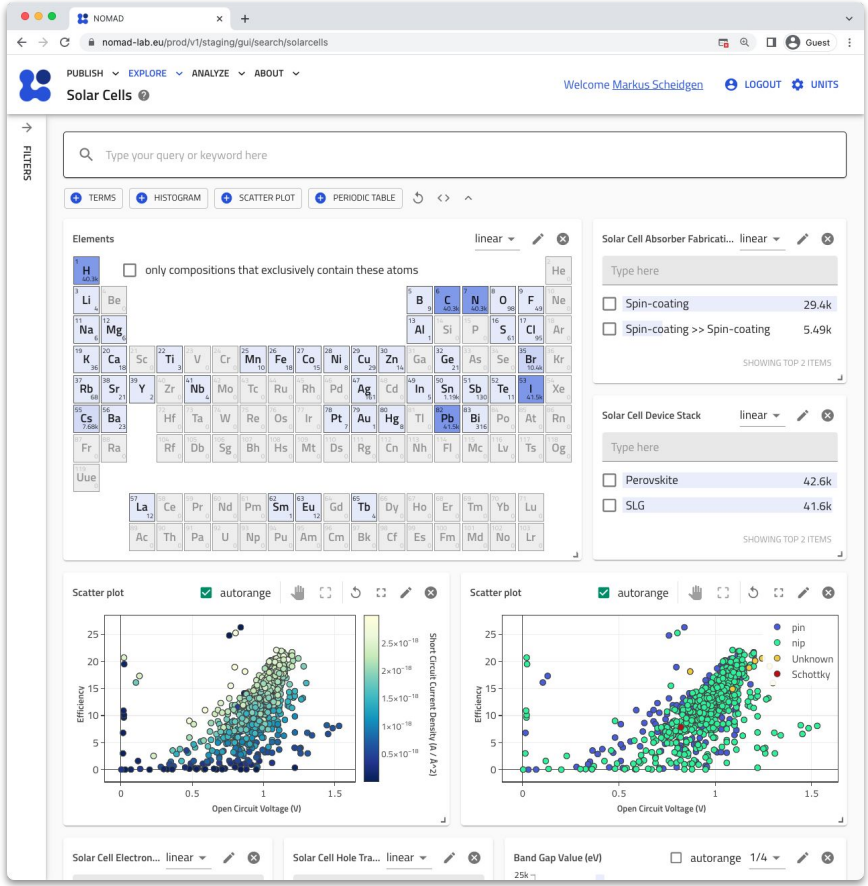
FAIRmat

Summary

- “easy” orchestration and scaling to create a complex service
- automated pipelines from commit to deployment
- monitoring an application on millions of logs

Questions?





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)