



FAIR Data Management

Sandor Brockhauser^{*} and the FAIRmat team

^{*}Center for Materials Science Data,
Humboldt-Universität zu Berlin, Germany

Kepler's Laws

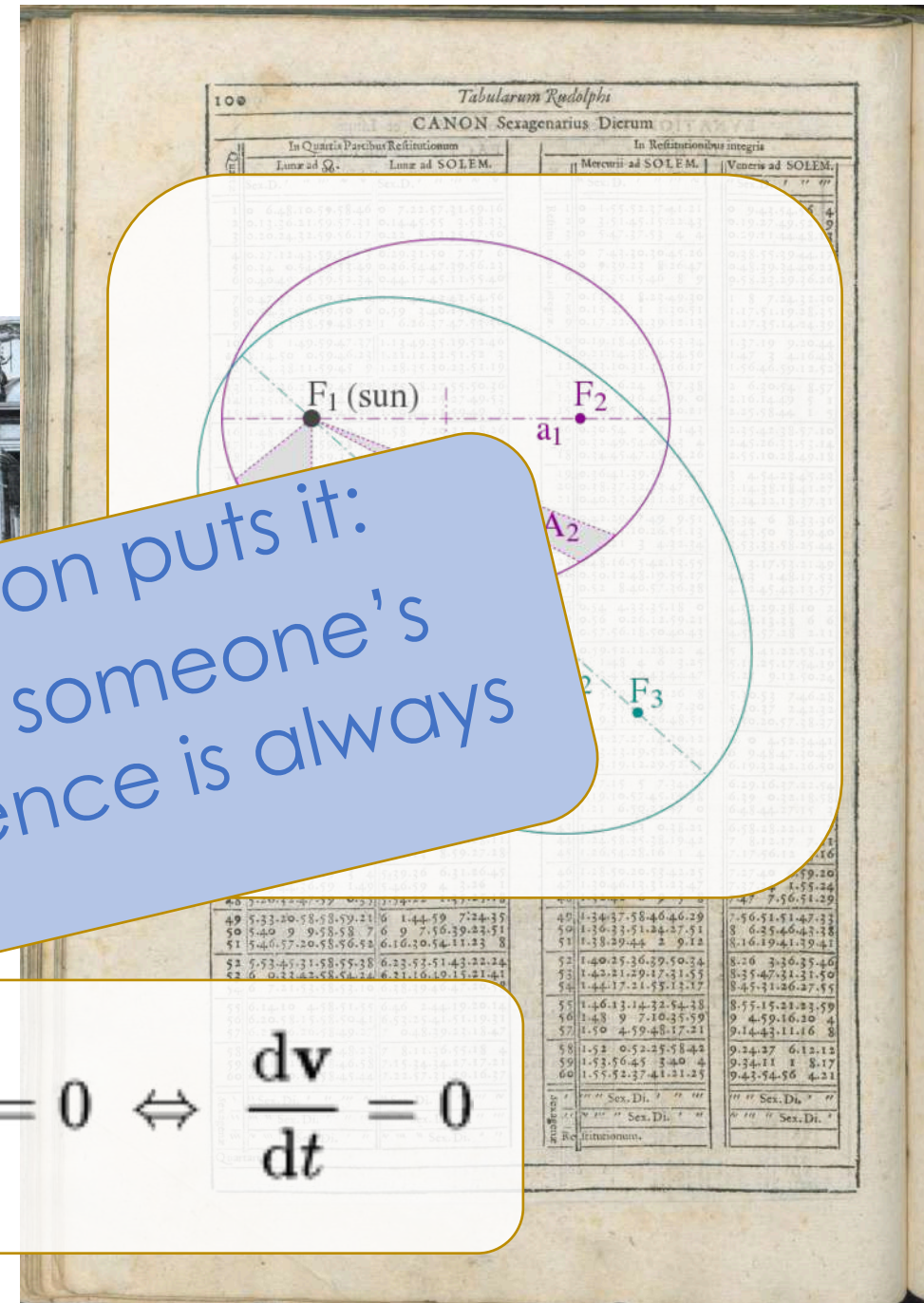
Tycho de Brahe (1546 – 1601)

Johannes Kepler (1571 – 1630)

Astronomia Nova, 1609

Harmonices Mundi, 1619

Tabulae Rudolphinae



As the Royal Society of London puts it:
 „Science is about not taking someone's
 word and so, instead, the science is always
 in the data”

$$\mathbf{F}_{\text{Net}} = 0 \Leftrightarrow \frac{d\mathbf{v}}{dt} = 0$$



European Open Science Cloud

EOSC is the **Science, Research and Innovation data space** in the European strategy for data. (*European Commission*)

- seamless access
- FAIR (Findability, Accessibility, Interoperability and Reusability) management
- reliable reuse of research data and all other digital objects produced along the research life cycle (e.g. methods, software and publications)

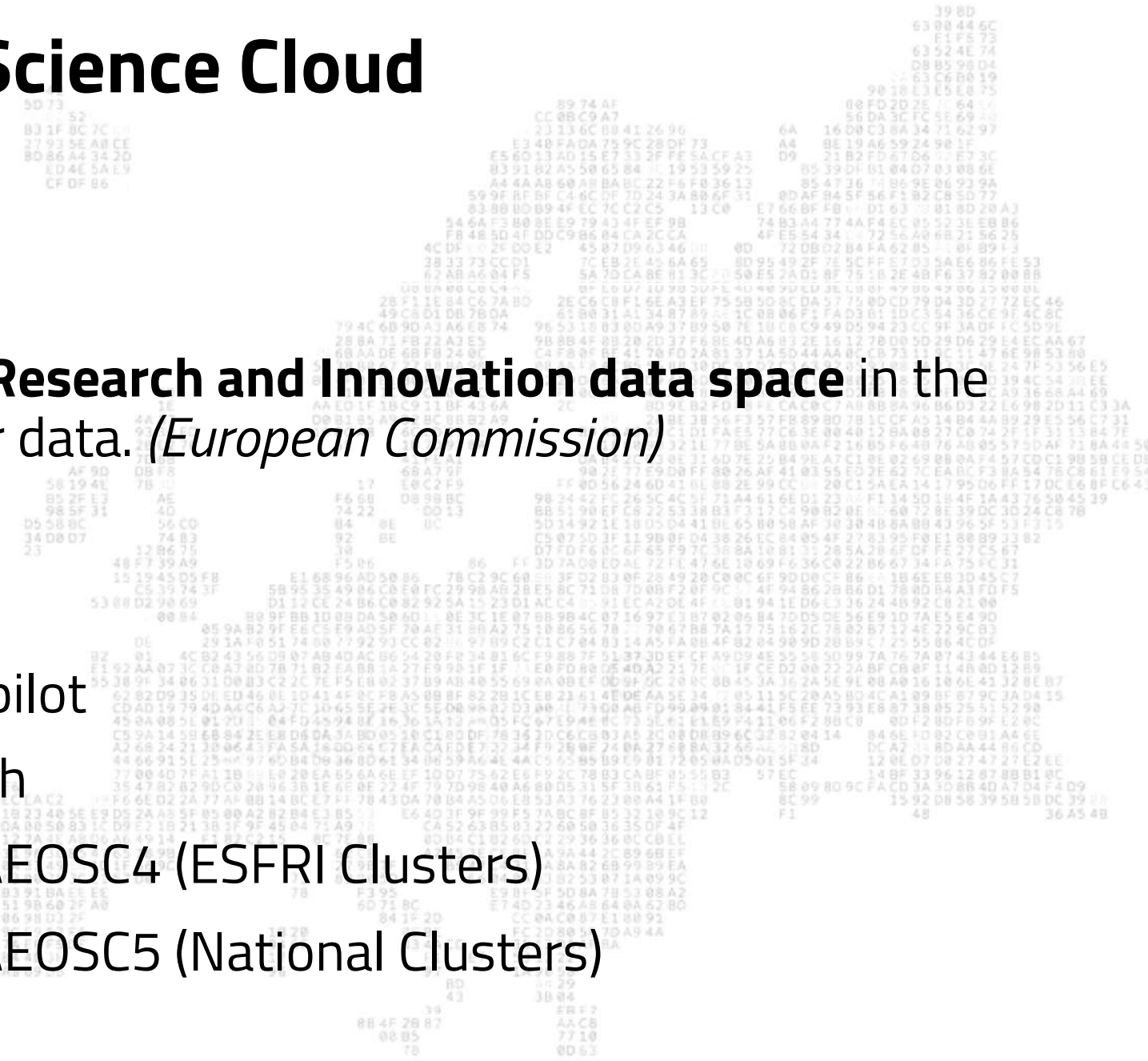


European Open Science Cloud

EOSC is the **Science, Research and Innovation data space** in the European strategy for data. *(European Commission)*

320 million €

- 2016-2018 EOSCpilot
- 2018: EOSC Launch
- 2018-2022 INFRAEOSC4 (ESFRI Clusters)
- 2019-2022 INFRAEOSC5 (National Clusters)





The EU's open science policy

https://ec.europa.eu/info/research-and-innovation/strategy/strategy-2020-2024/our-digital-future/open-science_en

- Open Data
FAIR and open data sharing should become the default for the results of EU-funded scientific research.
- European Open Science Cloud
store, share, process and reuse research digital objects (like **publications, data, and software**)
- New generation metrics
New indicators must be developed to complement the conventional indicators **for research quality and impact**, so as to do **justice to open science** practices.
- Future of scholarly communication
All peer-reviewed scientific **publications should be freely accessible**
- Research integrity & reproducibility of scientific results
All publicly funded research in the EU should adhere to **commonly agreed standards** of research integrity.



wwPDB / IUCr (2008)

- Macromolecular coordinates: mandatory deposition in PDB (created 1971)
- Structure factors: mandatory to accompany coordinates since 2008

11/29/2007

Announcement: Experimental Data Will Be Required for Depositions Starting February 1, 2008

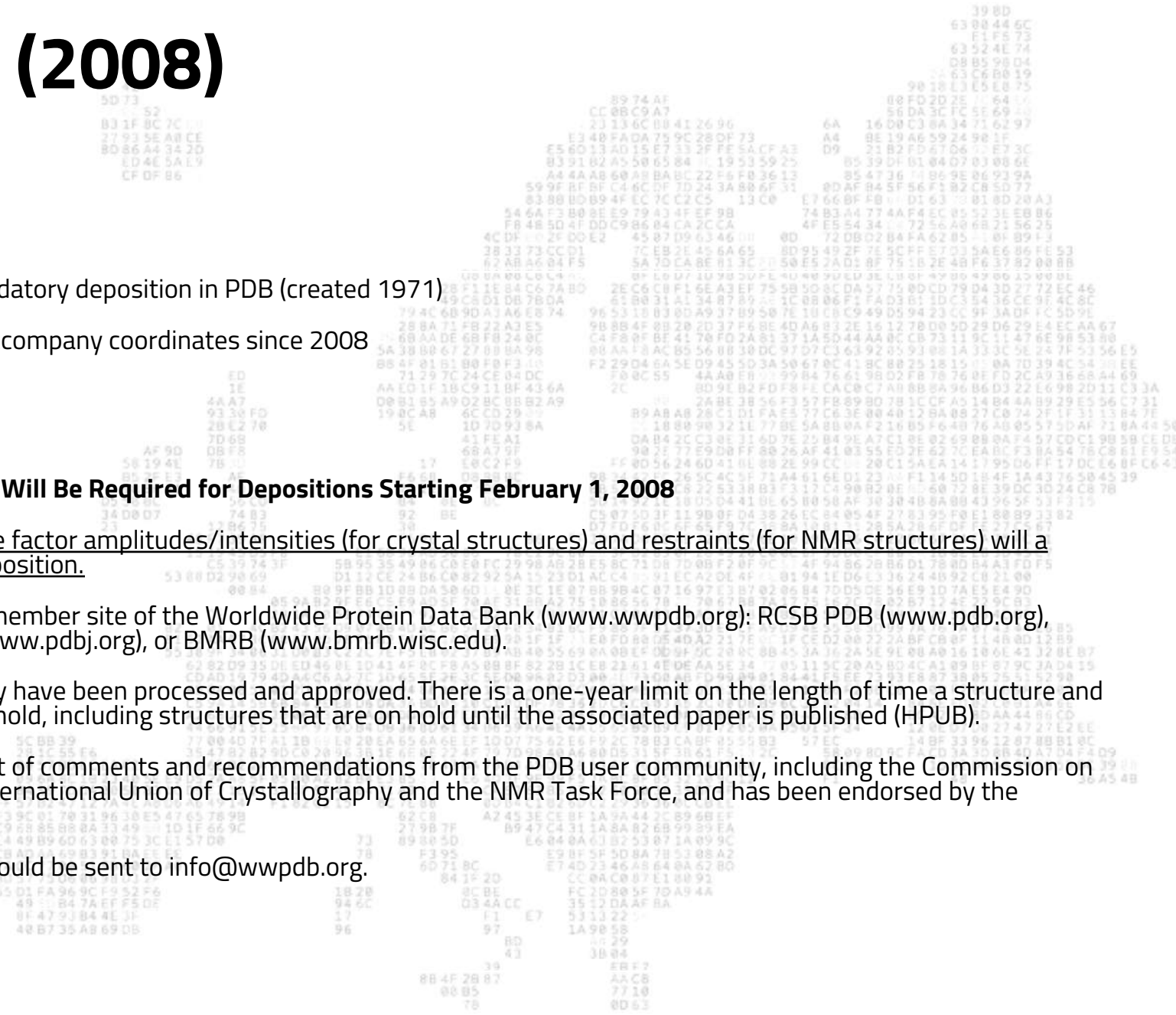
Effective February 1, 2008, structure factor amplitudes/intensities (for crystal structures) and restraints (for NMR structures) will a mandatory requirement for PDB deposition.

These data must be deposited at a member site of the Worldwide Protein Data Bank (www.wwpdb.org): RCSB PDB (www.pdb.org), PDBe (www.ebi.ac.uk/msd), PDBj (www.pdbj.org), or BMRB (www.bmrwisc.edu).

Data can be released as soon as they have been processed and approved. There is a one-year limit on the length of time a structure and its experimental data can be put on hold, including structures that are on hold until the associated paper is published (HPUB).

This policy was developed as a result of comments and recommendations from the PDB user community, including the Commission on Biological Macromolecules of the International Union of Crystallography and the NMR Task Force, and has been endorsed by the wwPDB Advisory Committee.

Questions relating to depositions should be sent to info@wwpdb.org.





Nature Editorials (2009)

Data's shameful neglect. Nature 461, 145 (2009). <https://doi.org/10.1038/461145a>

Research cannot flourish if data are not preserved and made accessible. All concerned must act accordingly.

More and more often these days, a research project's success is measured not just by the publications it produces, but also by the data it makes available to the wider community....

Research funding agencies need to recognize that preservation of and access to digital data are central to their mission, and need to be supported accordingly. ...

"Data management should be woven into every course in science."...

What is more, funding agencies and researchers alike must ensure that they support not only the hardware needed to store the data, but also the software that will help investigators to do this. One important facet is metadata management software: tools that streamline the tedious process of annotating data with a description of what the bits mean, which instrument collected them, which algorithms have been used to process them and so on — information that is essential if other scientists are to reuse the data effectively.



Naure Editorials & Science Editorial (2014)

Journals unite for reproducibility. Nature 515, 7 (2014). <https://doi.org/10.1038/515007a>

a group of editors representing more than 30 major journals; representatives from funding agencies; and scientific leaders assembled at the American Association for the Advancement of Science's headquarters in June 2014 to discuss principles and guidelines for preclinical biomedical research. The gathering was convened by the US National Institutes of Health, Nature and Science (see Science 346, 679; 2014)....

Reproducibility, rigour, transparency and independent verification are cornerstones of the scientific method...

Journals should recommend deposition of data in public repositories, where available, and link data bidirectionally when the paper is published. Journals should strongly encourage, as appropriate, that all materials used in the experiment be shared with those who wish to replicate the experiment....

The hope is that these guidelines will be viewed not as onerous, but as part of the quality control that justifies the public trust in science.



Acta Crystallographica D, IUCr (2019)

Findable Accessible Interoperable Re-usable (FAIR) diffraction data are coming to protein crystallography, Acta Cryst D, Struct. Biology

<https://doi.org/10.1107/S2059798319004844>

(i) Authors should provide a permanent and prominent link from their article to the raw data sets which underpin their journal publication and associated database deposition of processed diffraction data (e.g. structure factor amplitudes and intensities) and coordinates, and which should obey the 'FAIR' principles that their raw diffraction data sets should be Findable, Accessible, Interoperable and Re-usable (<https://www.force11.org/group/fairgroup/fairprinciples>).

(ii) A registered Digital Object Identifier (doi) should be the persistent identifier of choice (rather than a Uniform Resource Locator, url) as the most sustainable way to identify and locate a raw diffraction data set.

In 2018, the IUCr Commission on Biological Macromolecules (CBM) and the IUCr Committee on Data submitted a memorandum to the IUCr Executive Committee and proposed a mechanism for making diffraction experiments publicly available. The goal of ensuring better reproducibility of scientific discoveries in structural biology would be achieved, in part, by:

- (1) Allowing the scientific community to identify and re-use the original diffraction image data from a diffraction experiment, which is the primary source of information used to determine a particular macromolecular structure.
- (2) Facilitating structure re-determination using those original diffraction image data.
- (3) Providing researchers with a straightforward mechanism that will permit assessing the correctness of the structure determination process.
- (4) Providing a mechanism to ensure that the structures in the PDB and the publications derived from them are of the highest possible quality.

IUCr Journals are now taking the lead by encouraging authors to provide a doi for their deposited original raw diffraction data when they submit an article describing a new structure or a new method tested on unpublished diffraction data



Nature Chemistry (April 2022)

Making the collective knowledge of chemistry open and machine actionable

Jablonka, K.M., Patiny, L. & Smit, B. Making the collective knowledge of chemistry open and machine actionable. *Nat. Chem.* 14, 365–376 (2022).

<https://doi.org/10.1038/s41557-022-00910-7>

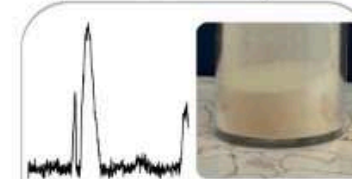
3D 73 52
83 1F 8C 7C
27 93 5E AB CE
80 86 A4 34 2D
ED 4E 5A 19
CF DF 86

MOF that can be made in one step in water



About 3'430 results (0.97 seconds)

Synthesis Recipes



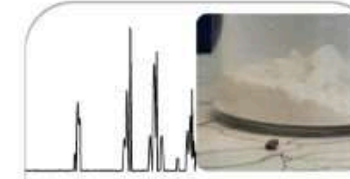
**Nanocrystalline MIL-53
at room temperature**

Díaz's MOF notebook

★★★★★ 56 ratings

4 h at room temperature
Na₂BDC in H₂O added to
Al(NO₃)₃·9H₂O in water

breathing abundant metal



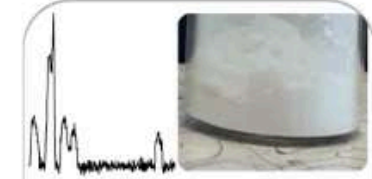
**Nanocrystalline MIL-53
at room temperature**

Díaz's MOF notebook

★★★★★ 89 ratings

72 h at room temperature
Na₂BDC in H₂O added to
Al(NO₃)₃·9H₂O in water

breathing abundant metal



**MIL-808(Hf)
for methane storage**

Dan's MOF blog

★★★★★ 19 ratings

12 h at 100 °C
HfCl₄ in water/formic acid
then BTC ligand

CH₄ storage

show more results

German National Research Data Infrastructure, NFDI

FAIRmat @ NFDI

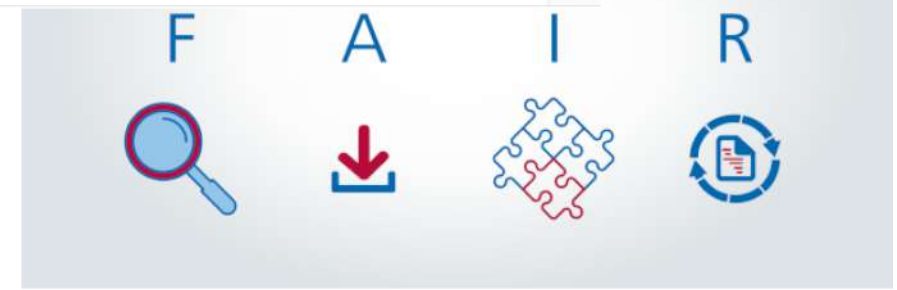
9 projects started in 2020
10 projects started in 2021

FAIRmat project
5 years
9 institutes
52 research groups
14 M€

The screenshot shows the NOMAD web interface. At the top, there is a search bar with 'NOMAD' and 'Search > Entry' buttons. Below the search bar are navigation tabs: PUBLISH, EXPLORE, ANALYZE, and ABOUT. The main content area is divided into several sections: 'OVERVIEW', 'RAW DATA', 'ARCHIVE', and 'LOGS'. The 'OVERVIEW' section is active and displays the following information:

- Method**
code name: VASP 5.4.4 DFT
code version: 5.4.4
electronic structure method: DFT
xc functional family: xc functional names
GGA: GGA_C_PBE, GGA_X_PBE
basis set type: plane waves
- Author metadata**
comment: Materials Project Upload at 2021-06-23 06:00:32.111642
references: <https://materialsproject.org/tasks/mp-14406...>
authors: Materials Project datasets: no datasets
- Material**
formula: Ca4O12Ti4
material type: bulk
material name: unavailable
crystal system: orthorhombic
spacegroup: Pnma (62)

Below the material information, there is a 3D visualization of the crystal structure (green and red spheres) and a 'Band structure' plot. The 'Electronic properties' section is also visible.



FAIRmat for Condensed-matter and chemical-physics

FAIRmat @ NFDI



https://www.fairmat-nfdi.eu/fairmat/fairmat_/fairmatteam

FAIRmat for Condensed-matter and chemical-physics

FAIRmat @ NFDI



FAIRmat for Condensed-matter and Chemical-physics

FAIRmat @ NFDI

Data Center;
Digitalisation;
ELNs; LIMS

Handwritten lab notebook page with the following content:

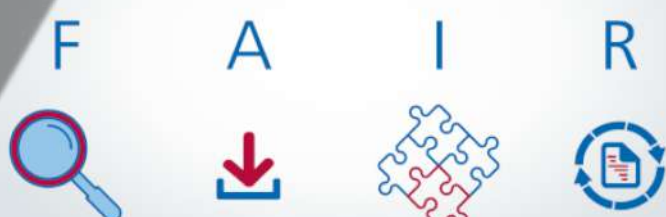
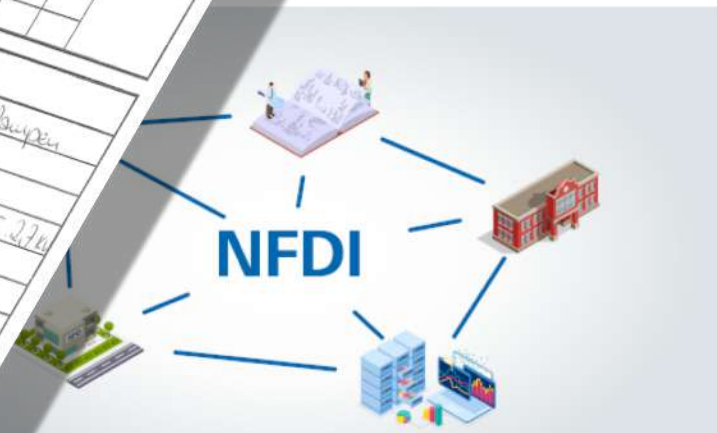
Schreibweise: 3H min
 Bearbeiter/in: Sus. LIP
 Datum: 8.10.19
 Versuch Nr.: 984
 Anlage: 1505/1 1505/2 1510 1520

Material: Wadras
 Ausgangspol (mm): Soll (mm) Ist (mm)
 103
 Induktor: 42 DIFF-2-19
 Versuchsziel: Induktortest $<100>$ vH (PTP)

Wickler: ASM sonstige
 Dotierung: P% B%
 HV: $5 \cdot 10^4$ (V) (min) (max) | LR (V) (technisch) | P (V) (max)
 SG Ar: 15 (min) + N_2 35 (V) + P_2 +

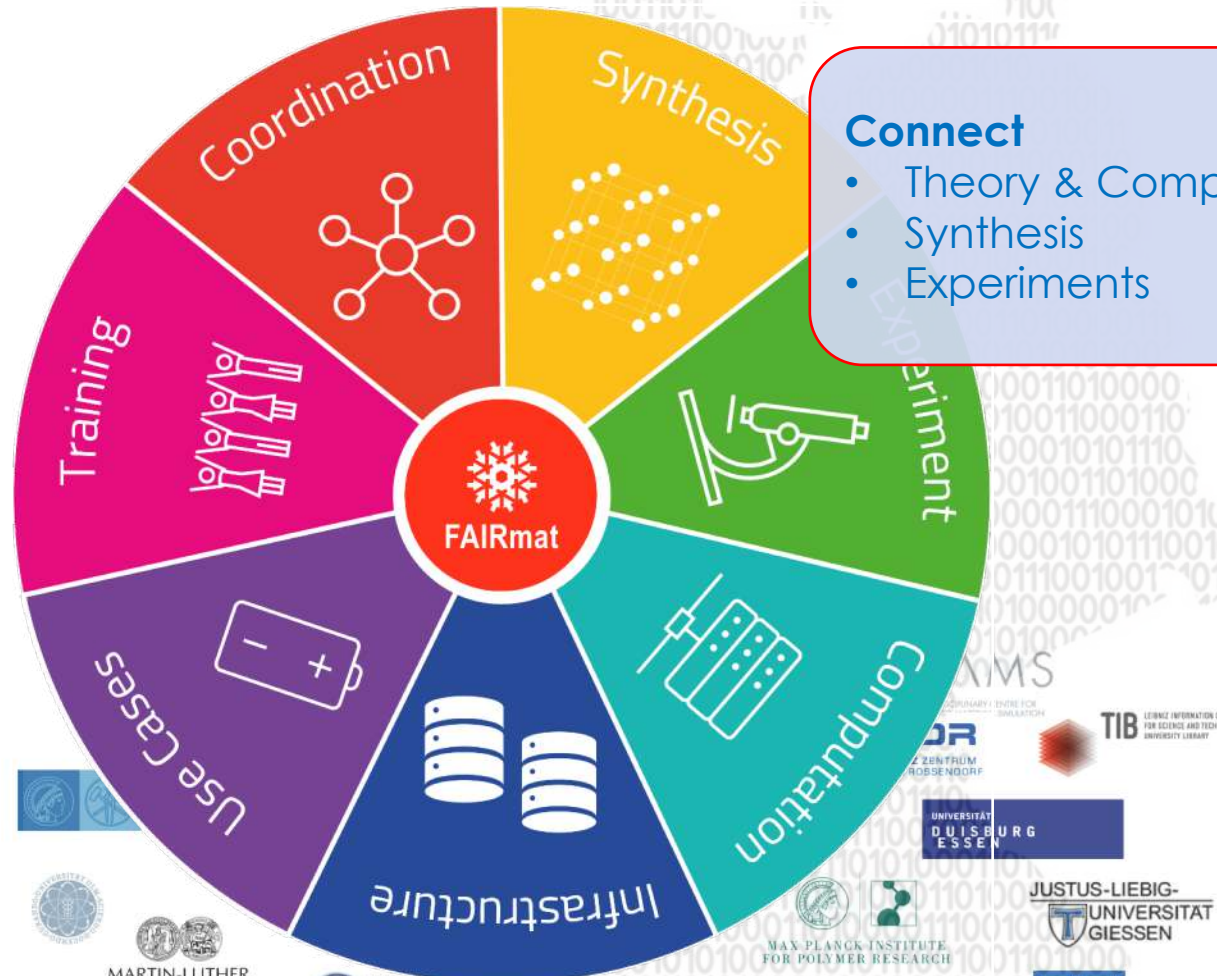
Pos. (unten)	Hub (mm)		Rotation (min)		Soll	U _h (V)	I _h (A)	I _g (A)	I _{max} (A)	Bemerkung
	oben	unten	oben	unten						
259	3,62	3,50	0,8	6,0	5,14	8,09	2,69	1,84	1,99	P. 39, 1,3 kW

Versuchsdurchführung:
 Einbau, Pumpen: 2F Pumpen, Saugen
 Luft Zünden: P₁ für 36V, 45° stabil 2,5-2,7V
 Doppelt ein
 Tropfen: 3,6V Primfien - 3,15 kV
 Rückstromen (Kam. bewegt wenig aus
 DH vH: (2,4-2,4-3,4) kV
 Kugel vH Beschreibung - 2,5mm
 Aufbau aufgangs Steuerung wegen
 Ringbildung
 Minimum 2,86 kV
 103 Übernahme Regelung



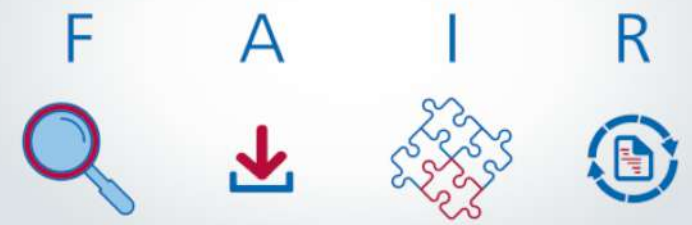
FAIRmat for Condensed-matter and chemical-physics

FAIRmat @ NFDI

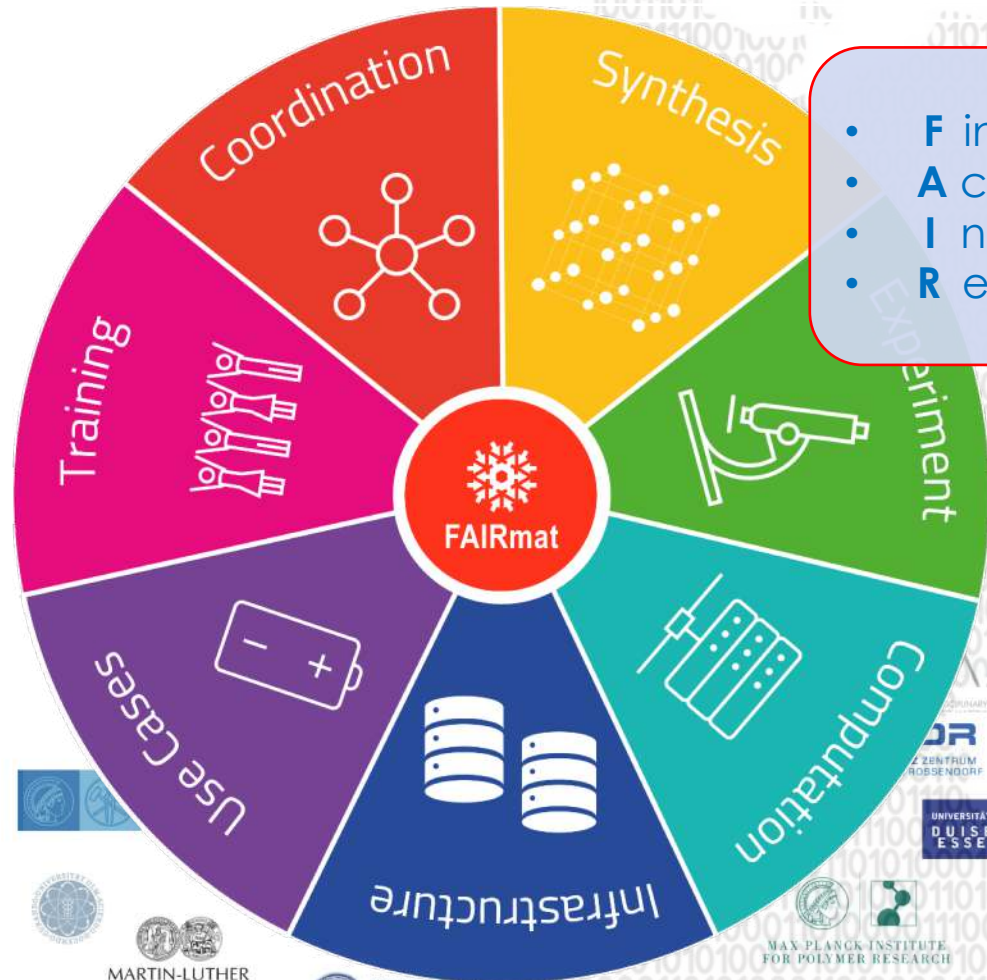


Connect

- Theory & Computation
- Synthesis
- Experiments



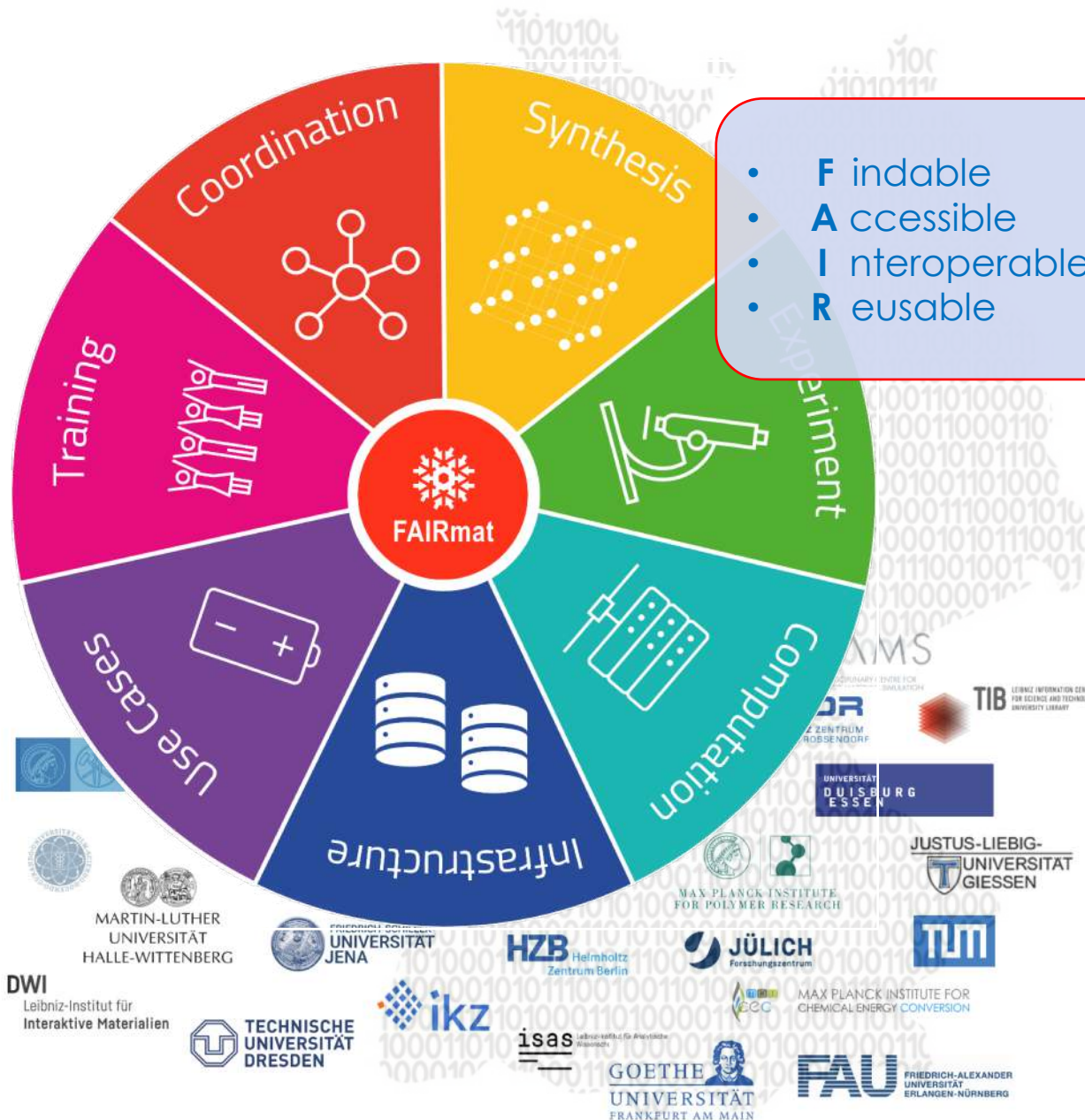
FAIRmat



- **F** indable
- **A** ccessible
- **I** nteroperable
- **R** eusable



Findable

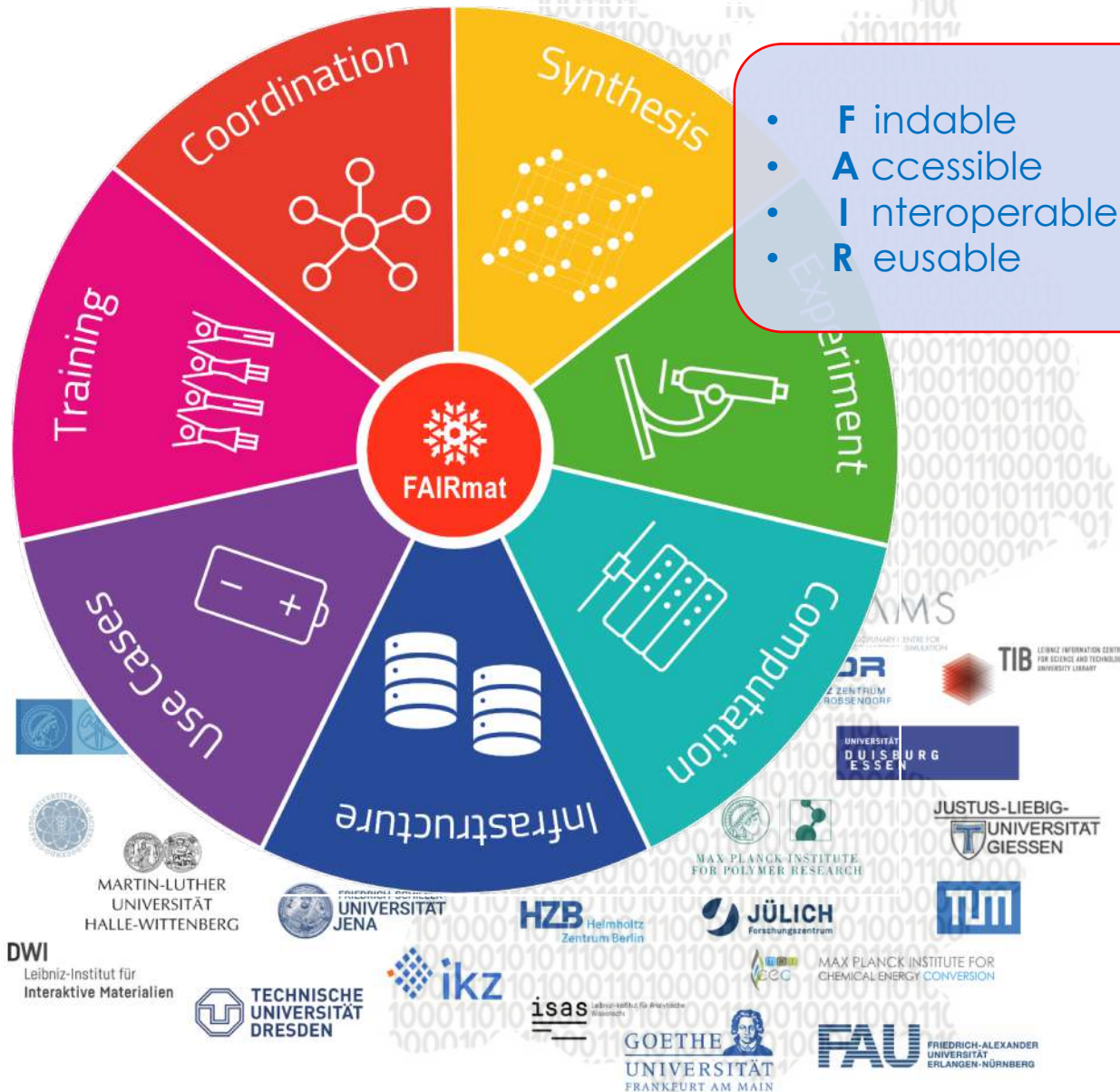


- **F** indable
- **A** ccessible
- **I** nteroperable
- **R** eusable

- Findable**
- Search platform
 - who / what / which project / ...
 - Computation metadata:
 - Input parameters / Output
 - Experiment/Synthesis parameters
 - ELN integration
 - Materials properties
 - Encyclopedia



Accessible



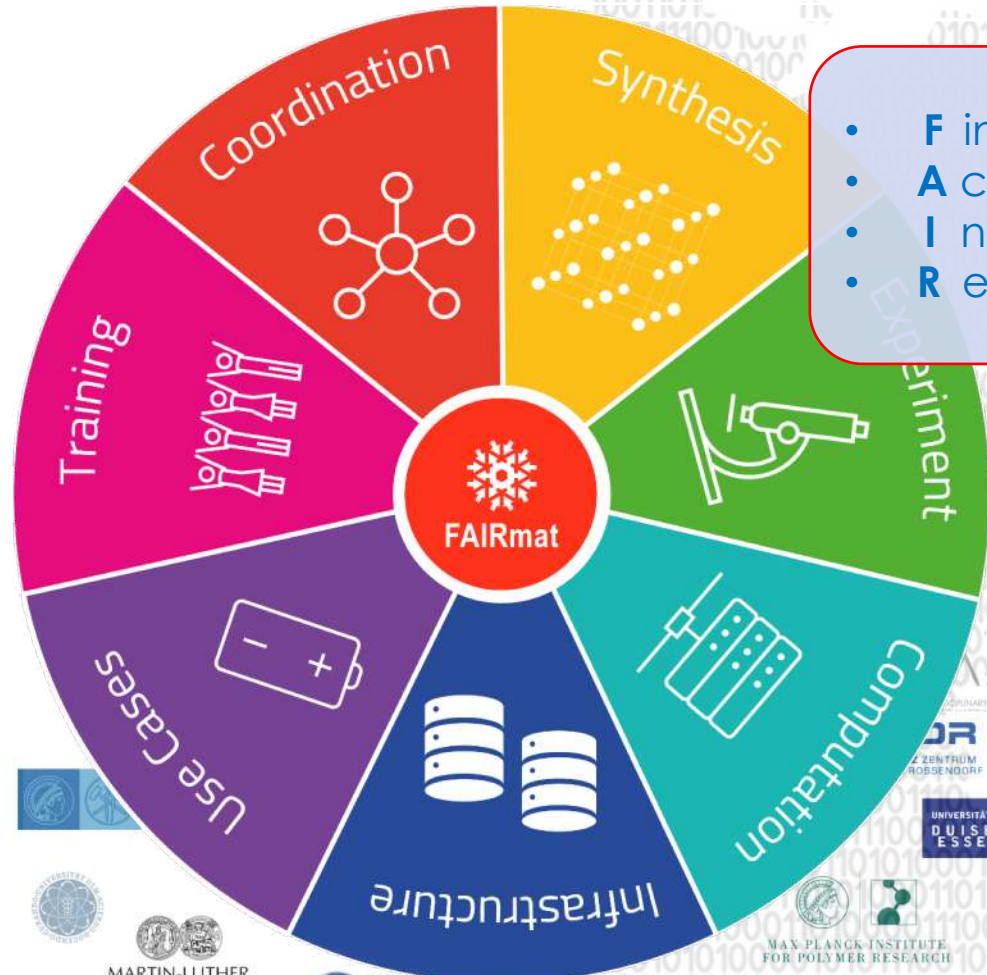
- **F** indable
- **A** ccessible
- **I** nteroperable
- **R** eusable

Accessible

- Access & Download portal (UX)
- Data Exploration & Visualisation
- DOI
- Version control
- Machine readable (API)
- Cloud Computing (big data)

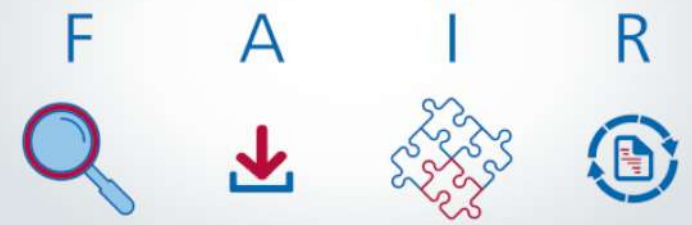


Interoperable

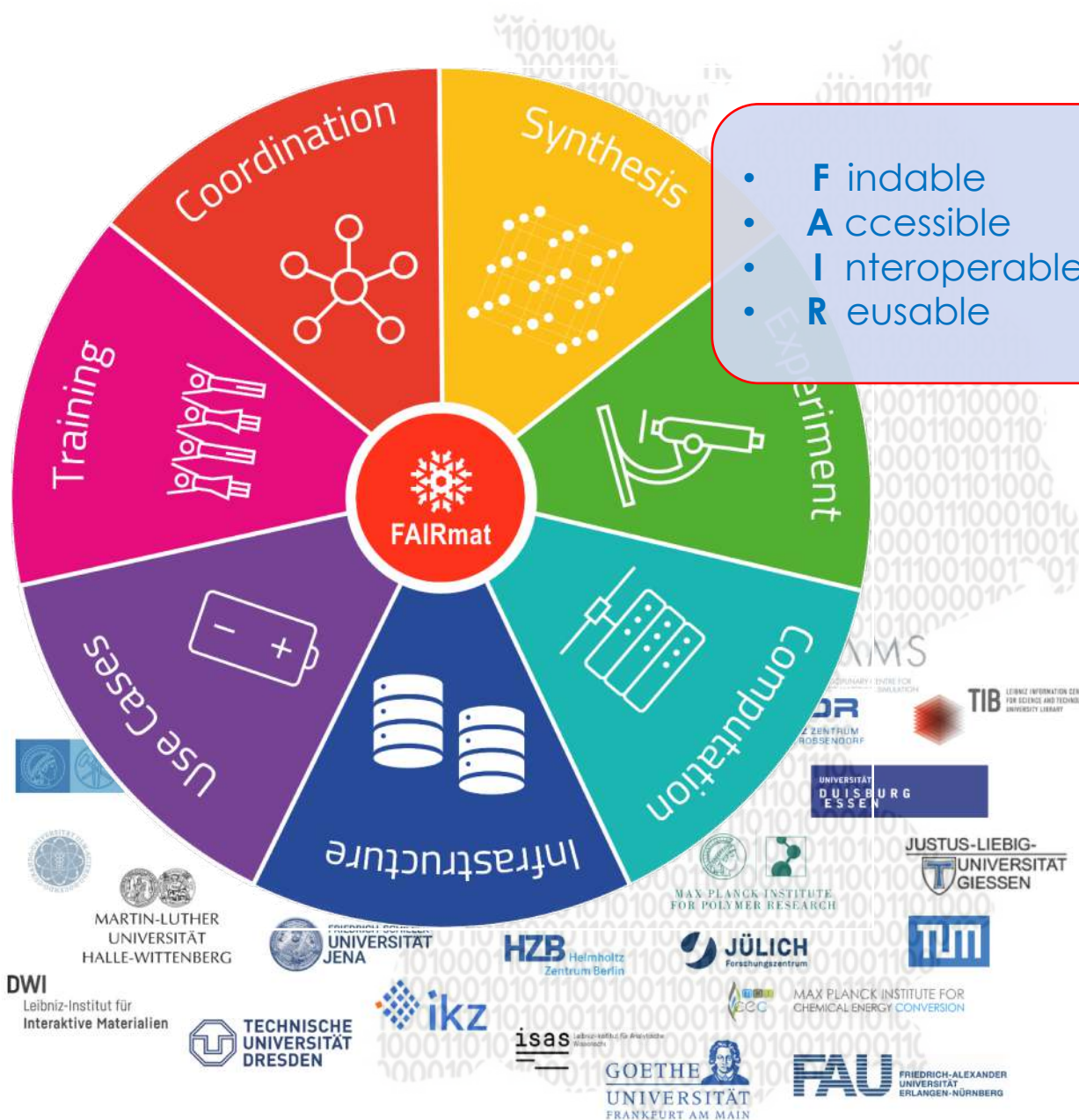


- **F** indable
- **A** ccessible
- **I** nteroperable
- **R** eusable

- Interoperable**
- Glossary with defined vocabulary
 - Machine interpretable
 - Ontology relationships
 - Connection to external ontologies



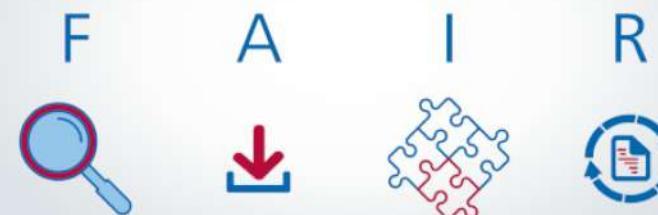
Reusable



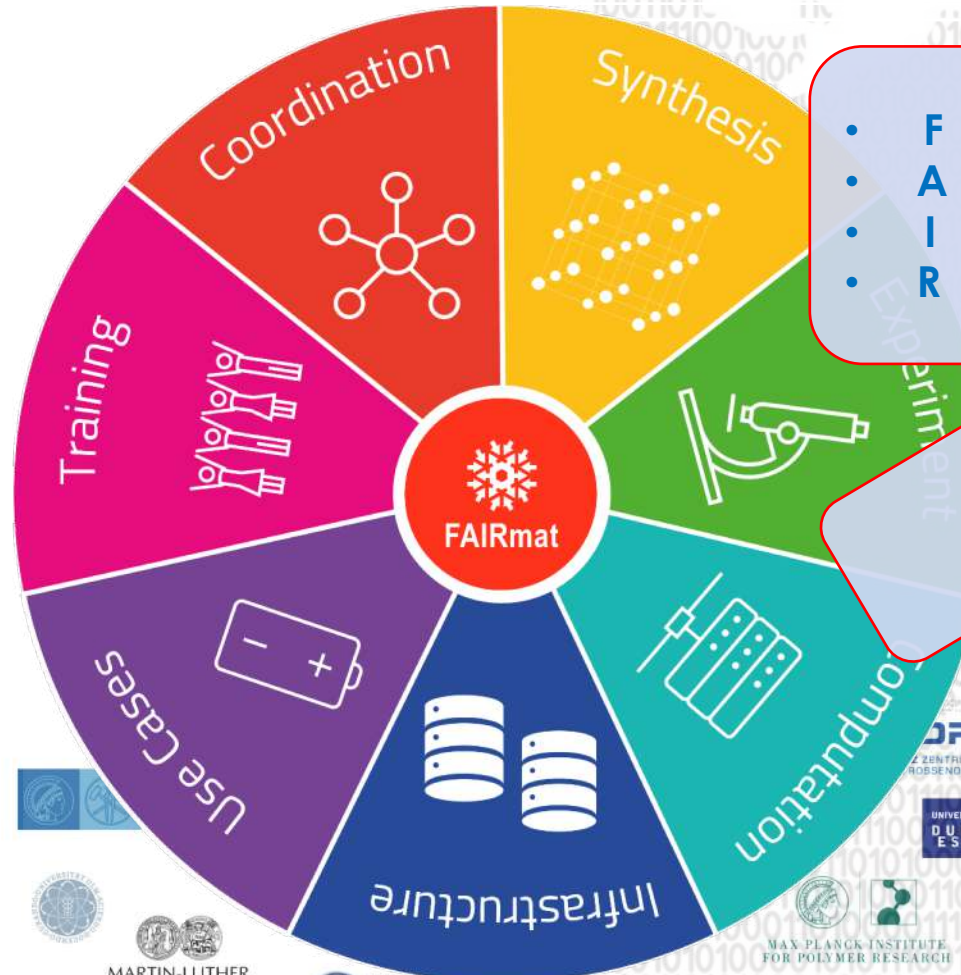
- **F** indable
- **A** ccessible
- **I** nteroperable
- **R** eusable

Reusable

- Data manipulation
- Version control
- Remote Data Analysis Services
- Machine Readable
- Machine Interpretable

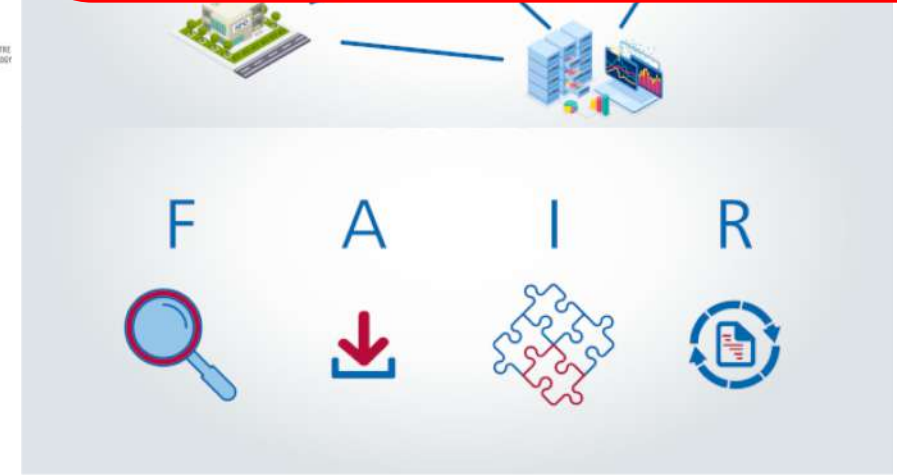


AI Ready



- F indable
- A ccessible
- I nteroperable
- R eusable

- F**indable and **A**I Ready
- Data manipulation
 - Version control
 - Remote Data Analysis Services
 - Machine Readable
 - Machine Interpretable



Data Management and Reproducibility (in Computational Science)

- DATA (information representation)
- Metadata (contextual information)
- Data + Metadata => Input data (for use)
- **Application (uses** Data for search, analysis, visualisation, etc.)
- Data + Metadata + Application => Output (new information)
- OUTPUT DATA + NEW METADATA (new information representation)

```
010010010110011011000010101000000101010011000010101
0000001101001011001100100000011000010010000001101
1000110100101011001100101010000101100100010000000
110100011001000110000101101100110001010010010111
0111001001101010100001011010000101001010110110111
000100000011001100110010011011011010100100000011
0000100
01110010
01011100
011101100
0000011
10110000
```

“If T is a linear transformation from a vector space V over a field F into itself and \mathbf{v} is a nonzero vector in V , then \mathbf{v} is an eigenvector of T if $T(\mathbf{v})$ is a scalar multiple of \mathbf{v} .”





Reproducibility in Experiments and Synthesis

- **Experiment**

- *Instrumentation*
- *Sample* Preparation
- Sample Environment
- Monitors and Detectors
- Data Processing
- *Notes*

- **Synthesis**

- *Sample* History
- *Processes*
- *Characterisation*
- *Notes*

Reproducibility in Experiments and Synthesis

• Experiment

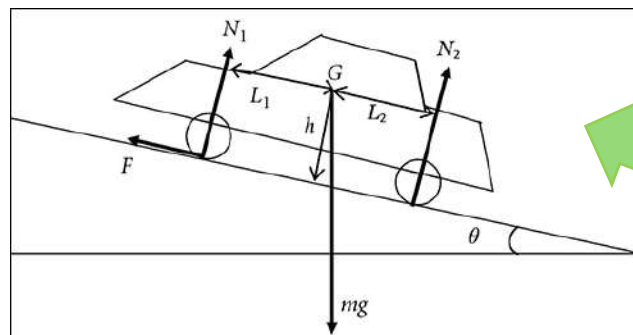
- *Instrumentation*
- *Sample* Preparation
- Sample Environment
- Monitors and Detectors
- Data Processing
- *Notes*

• Synthesis

- *Sample* History
- *Processes*
- *Characterisation*
- *Notes*

Experiment Data Models

Physics Model



Experiment
<ul style="list-style-type: none"> • start time, t_0 • duration, t • distance, d

Slope
<ul style="list-style-type: none"> • angle, θ • length, l

Car
<ul style="list-style-type: none"> • mass, m • center of mass, $[G, h]$ • wheel positions, $[L_1, L_2]$ • force on wheels, $[N_1, N_2]$ • resistive force, F

Experiment Parameters



Reproducibility in Experiments and Synthesis

• Experiment

- *Instrumentation*
- *Sample* Preparation
- Sample Environment
- Monitors and Detectors
- Data Processing
- *Notes*

• Synthesis

- *Sample* History
- *Processes*
- *Characterisation*
- *Notes*

Conceptual Design Model

- *Implementation independent Physics Model*
- *Does not tell how exactly it is performed*

Samples:

- Composition
- Geometry
- History
- ...

Sample Environment.:

- Temperature
- Pressure
- ...

Measurements:

- ARPES experiment
- XRD
- XRF measurement
- Hyperspectral PL imaging
- THz spectroscopy
- I-V measurement
- ...

Processes:

- Spin coating
- PVD deposition
- Solution preparation
- PLD deposition
- Hot plate annealing
- ...



Reproducibility in Experiments and Synthesis

• Experiment

- *Instrumentation*
- *Sample* Preparation
- Sample Environment
- Monitors and Detectors
- Data Processing
- *Notes*

• Synthesis

- *Sample* History
- *Processes*
- *Characterisation*
- *Notes*

Technical Design Model

- *Implementation specific*
- *Not only Physics Model, But also Technical Details*

E.g.

what was the pressure in a chamber



how it has been produced and maintained

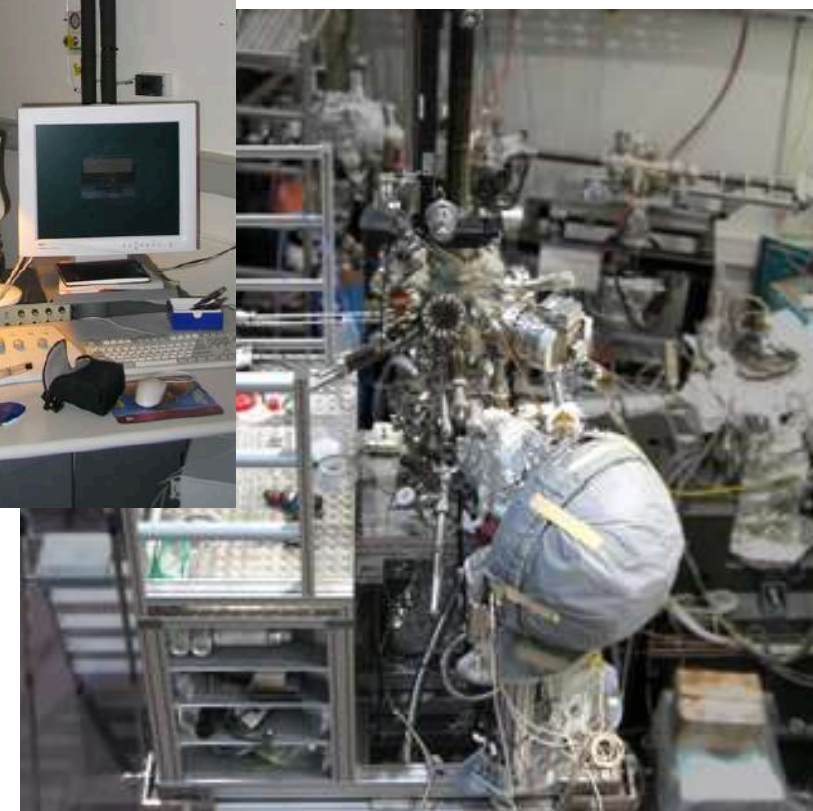
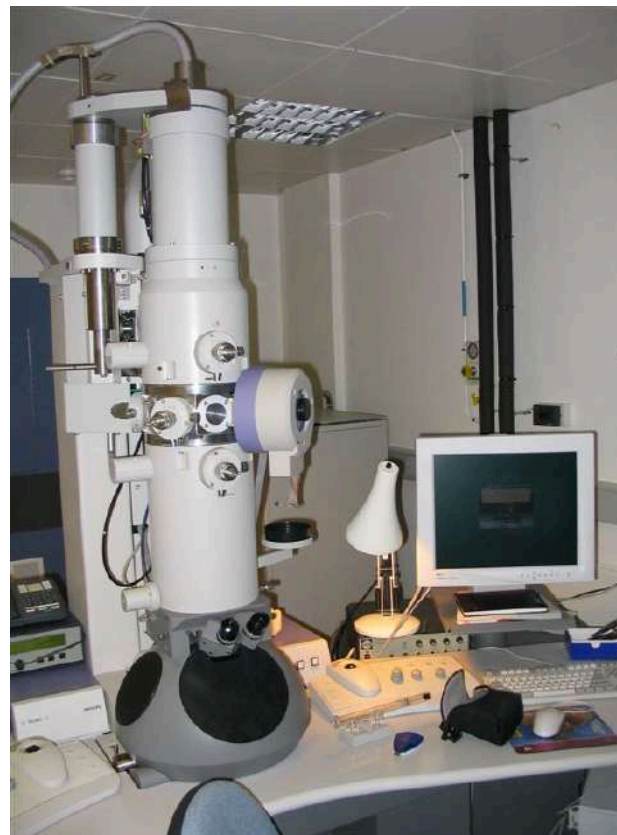
Reproducibility in Experiments and Synthesis

- **Experiment**

- *Instrumentation*
- *Sample* Preparation
- Sample Environment
- Monitors and Detectors
- Data Processing
- *Notes*

- **Synthesis**

- *Sample* History
- *Processes*
- *Characterisation*
- *Notes*



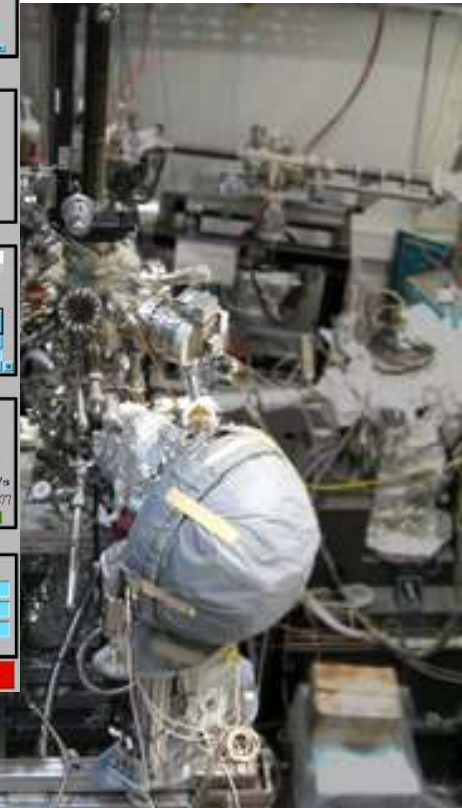
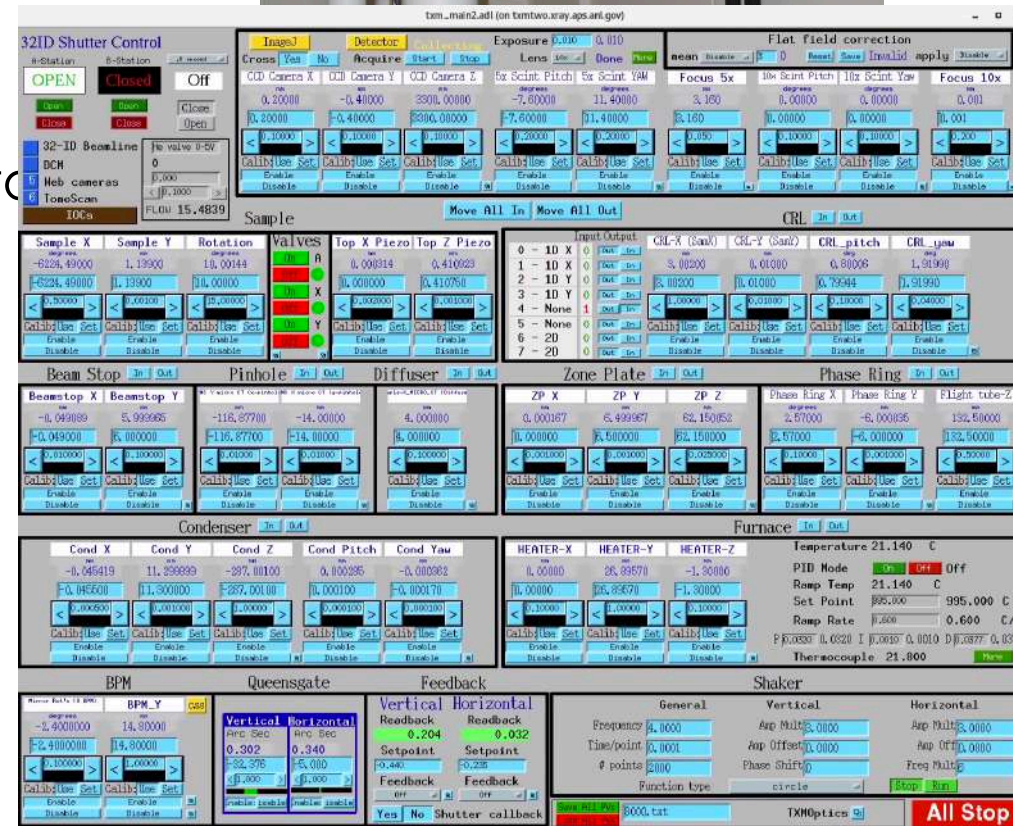
Reproducibility in Experiments and Synthesis

• Experiment

- *Instrumentation*
- *Sample* Preparation
- Sample Environment
- Monitors and Detectors
- Data Processing
- *Notes*

• Synthesis

- *Sample* History
- *Processes*
- *Characterisation*
- *Notes*



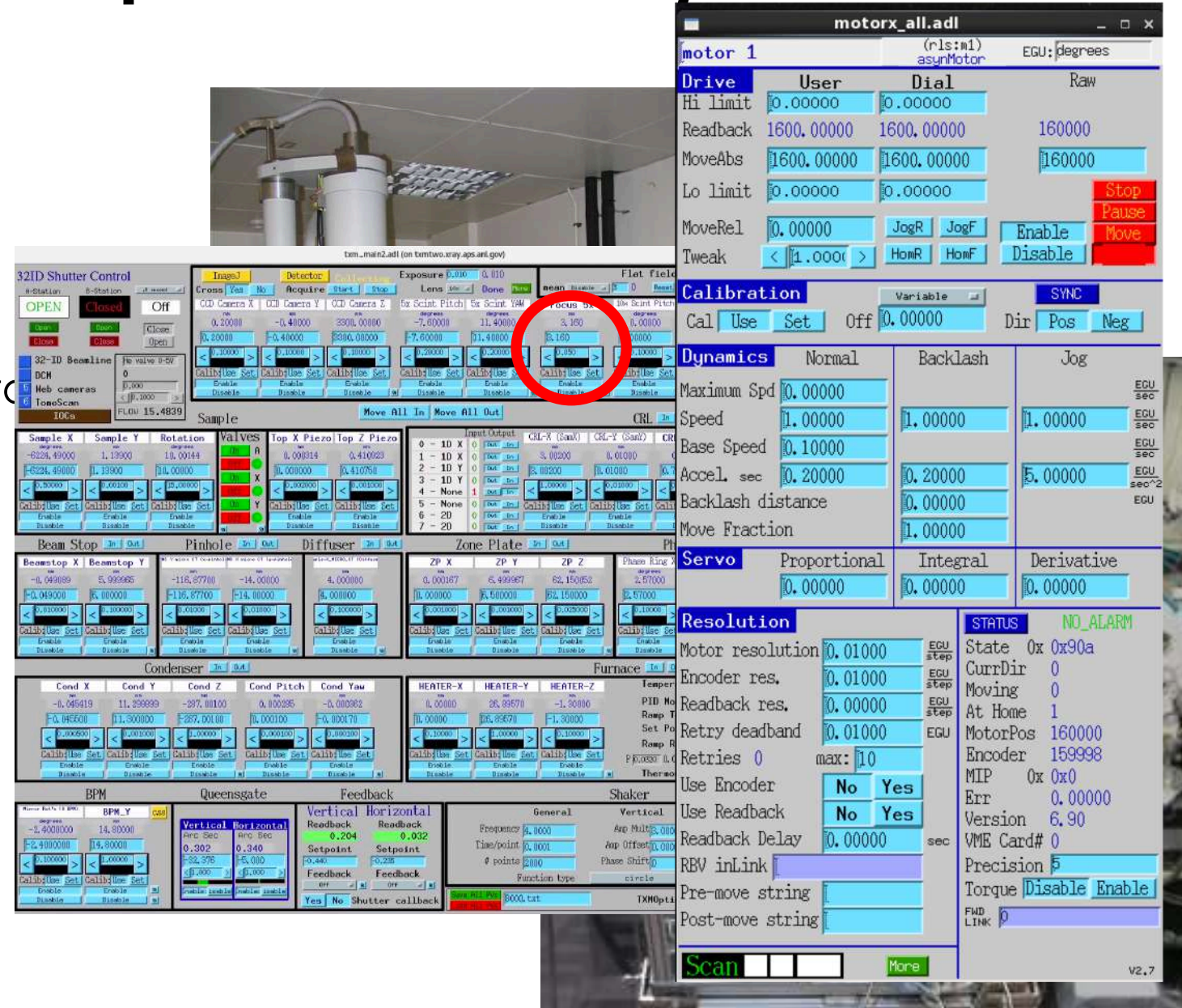
Reproducibility in Experiments and Synthesis

• Experiment

- *Instrumentation*
- *Sample* Preparation
- Sample Environment
- Monitors and Detectors
- Data Processing
- *Notes*

• Synthesis

- *Sample* History
- *Processes*
- *Characterisation*
- *Notes*

The screenshot displays a complex control interface for a motor system. Key sections include:

- motor 1 (rls:m1) asynMotor EGU:degrees**

Drive	User	Dial	Raw
Hi limit	0.00000	0.00000	
Readback	1600.00000	1600.00000	160000
MoveAbs	1600.00000	1600.00000	160000
Lo limit	0.00000	0.00000	
MoveRel	0.00000	JogR JogF	Enable
Tweak	< 1.0000 >	HomR HomF	Disable
- Calibration**

Variable	Value	Dir
Cal Use	Set	Off
Cal	0.00000	Pos Neg
- Dynamics**

Parameter	Normal	Backlash	Jog
Maximum Spd	0.00000		
Speed	1.00000	1.00000	1.00000
Base Speed	0.10000		
Accel. sec	0.20000	0.20000	5.00000
Backlash distance		0.00000	
Move Fraction		1.00000	
- Servo**

Mode	Proportional	Integral	Derivative
Mode	0.00000	0.00000	0.00000
- Resolution**

Parameter	Value	Unit
Motor resolution	0.01000	EGU/step
Encoder res.	0.01000	EGU/step
Readback res.	0.00000	EGU/step
Retry deadband	0.01000	EGU
- STATUS NO_ALARM**
 - State 0x 0x90a
 - CurrDir 0
 - Moving 0
 - At Home 1
 - MotorPos 160000
 - Encoder 159998
 - MIP 0x 0x0
 - Err 0.00000
 - Version 6.90
 - VME Card# 0
 - Precision 5
 - Torque Disable Enable
 - FWD LINK 0

Reproducibility in Experiments and Synthesis

• Experiment

- *Instrumentation*
- *Sample* Preparation
- Sample Environment
- Monitors and Detectors
- Data Processing
- *Notes*

• Synthesis

- *Sample* History
- *Processes*
- *Characterisation*
- *Notes*



Reproducibility in Experiments and Synthesis

• Experiment

- *Instrumentation*
- *Sample* Preparation
- Sample Environment
- Monitors and Detectors
- Data Processing
- *Notes*

• Synthesis

- *Sample* History
- *Processes*
- *Characterisation*
- *Notes*



motorx_all.adl

Drive	User	Dial	Raw
Hi limit	0.00000	0.00000	
Readback	1600.00000	1600.00000	160000
MoveAbs	1600.00000	1600.00000	160000
Lo limit	0.00000	0.00000	
MoveRel	0.00000	JogR JogF	Enable
Tweak	< 1.0000 >	HomR HomF	Disable

32ID Shutter Control

steps or ticks

↓ EGU's

↓ Raw

Logic Diagram:

```

    graph TD
        RVAL --> X1((X))
        MRES --> X1
        X1 --> DVAL
        X1 --> X2((X))
        RRES --> X2
        RDBL Link --> X2
        X2 --> RRBV
        RRBV --> X3((X))
        REP --> X3
        RMP --> X3
        X3 --> DRBV
        URIP == NO --> X3
        X3 --> X4((X / +))
        DIR: +/- 1 --> X4
        X4 --> RBV
        RBV --> OFF
    
```

Reproducibility in Experiments and Synthesis

Reproducibility



motorx_all.adl			
motor 1		(r1s:m1)	EGU: degrees
		asynMotor	
Drive	User	Dial	Raw
Hi limit	0.00000	0.00000	
Readback	1600.00000	1600.00000	160000
MoveAbs	1600.00000	1600.00000	160000
Lo limit	0.00000	0.00000	

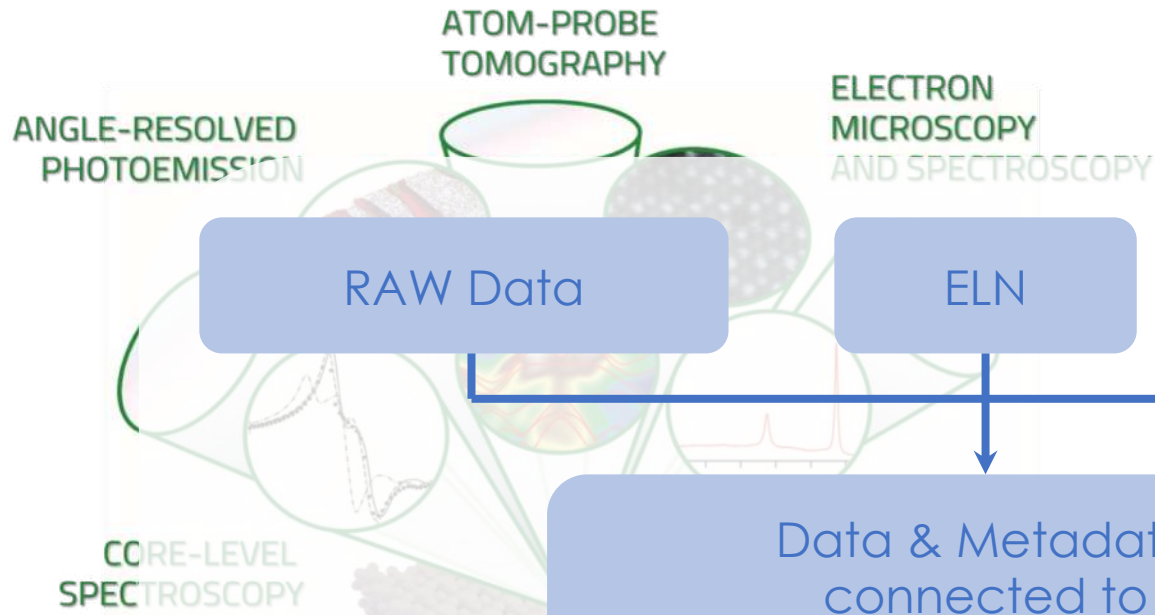
- **Synthesis**

- *Sample* History
- *Processes*
- *Characterisation*
- *Notes*

Cond X	Cond Y
-0.045419	11.333333

BPM_Y	Vertical	Horizontal
14.00000	0.302	0.340

Data Models for Experiments



Ellipsometry

Electron Microscopy

Atom Probe Microscopy

Photoemission Spectroscopy

Temperature dependent I-V measurements



NXellipsometry

Status:

application definition, extends `NXObject`

Description:

...tion for ellipsometry measurements
...tropic ellipsometry. In this applic
...her with a UTC offset.

Data Analysis

Variables used throughout the document, e.g. dimensions

`N_wavelength`: Size of the energy / wavelength vector us

`N_variables`: How many variables are saved in a measur
Mueller matrix.

`N_angles`: Number of incident angles used

`N_p1`: Number of sample parameters scanned

`N_time`: Number of time points measured

Groups cited:

`NX_TIME`, `NXaperture`, `NXcollection`, `NXdata`, `NXdetector`,
`NXsample`, `NXsubentry`, `NXtransformations`, `NXuser`

Structure:

ENTRY: (required) `NXentry`

"This is the application definition describing ellipsome
experiments may be as simple as identifying how a re
single wavelength changes its polarization state, to a
ellipsometry experiment. The application definition de
experimental instrument - calibration information if av
the state of the sample - sample description"

definition: (required) `NX_CHAR`

An application definition for ellipsometry.

Obligatory value: `NXellipsometry`

@version: (required) `NX_CHAR`

Version number to identify which definition of
used for this entry/data.

<https://fairmat-experimental.github.io/nexus-fairmat-proposal/>



Characterisation Methods in FAIRmat

Area B:

- Optical Spectroscopy, EM, AREPS, XPS, Atomprobe Microscopy

Area A:

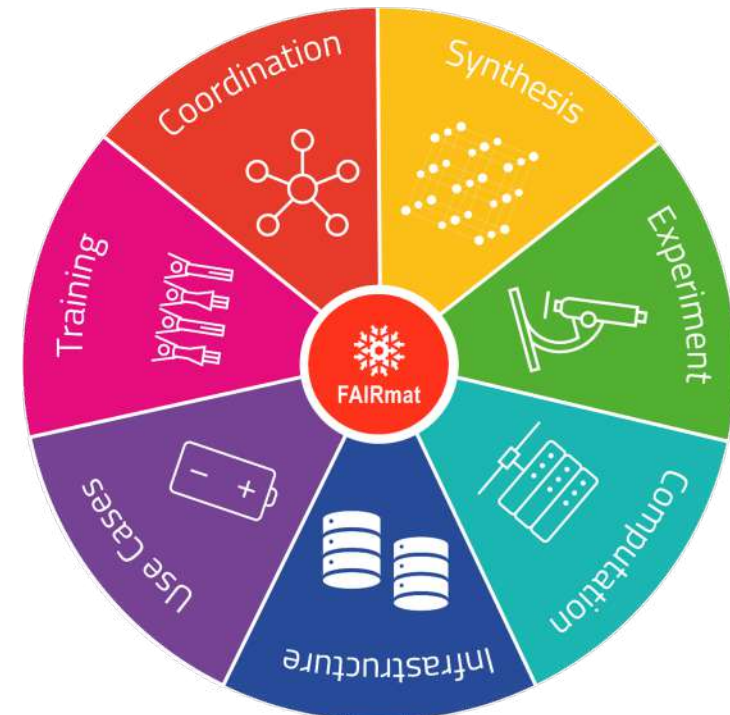
- Magnetic Resonance, Magnetometry, Scattering and Imaging, Electrical and Mechanical methods
- Synthesis processes and Sample description

Area C connection:

- Theory models and Simulation results

Area D Support

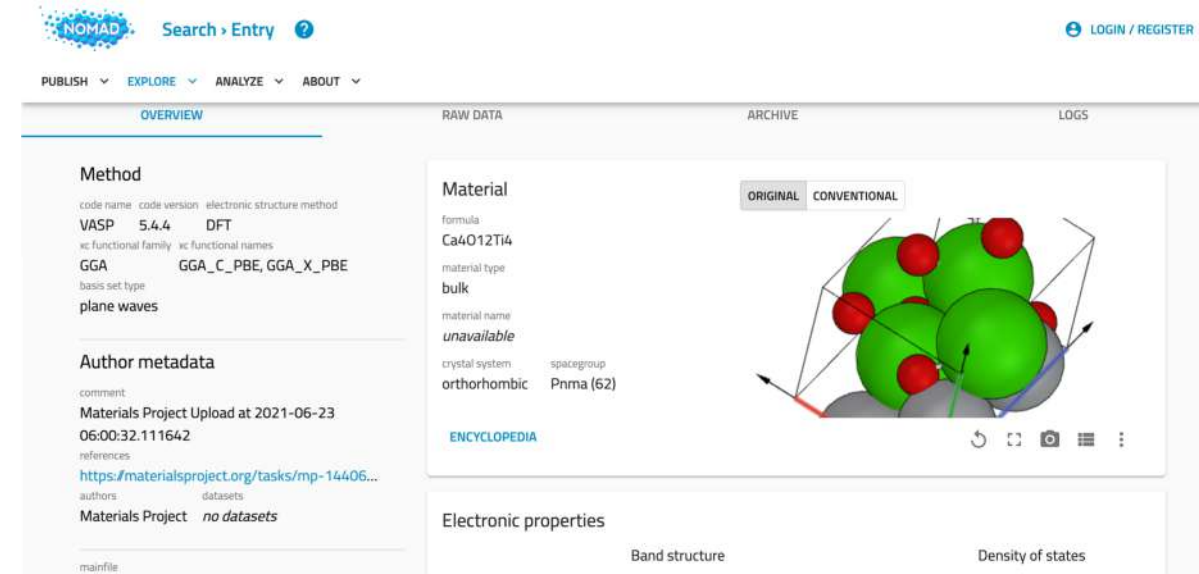
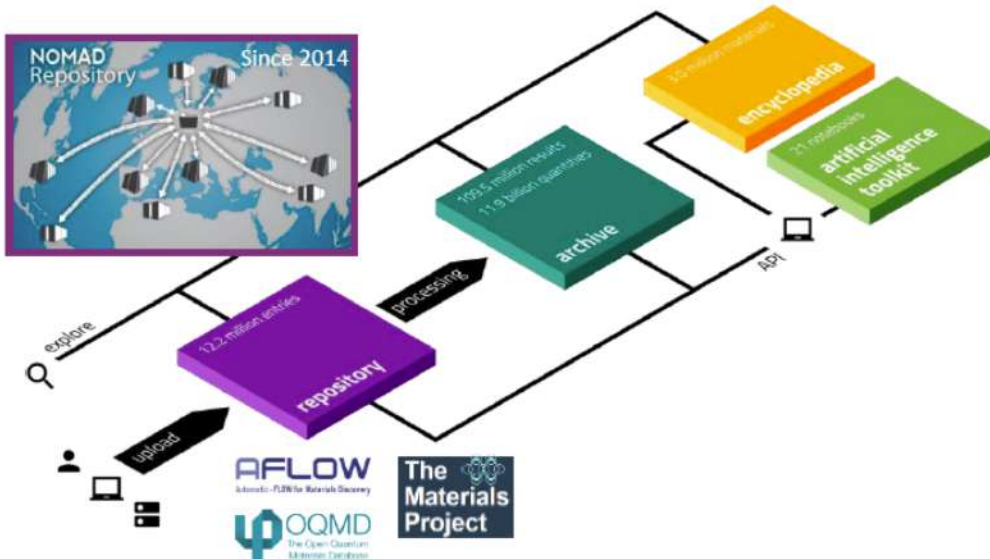
- Integration at Control System level
- Infrastructure for Data Management



FAIR Theory and Computations



Biggest database in Materials Science (<https://nomad-lab.eu>)
more than 100 million high-quality calculations



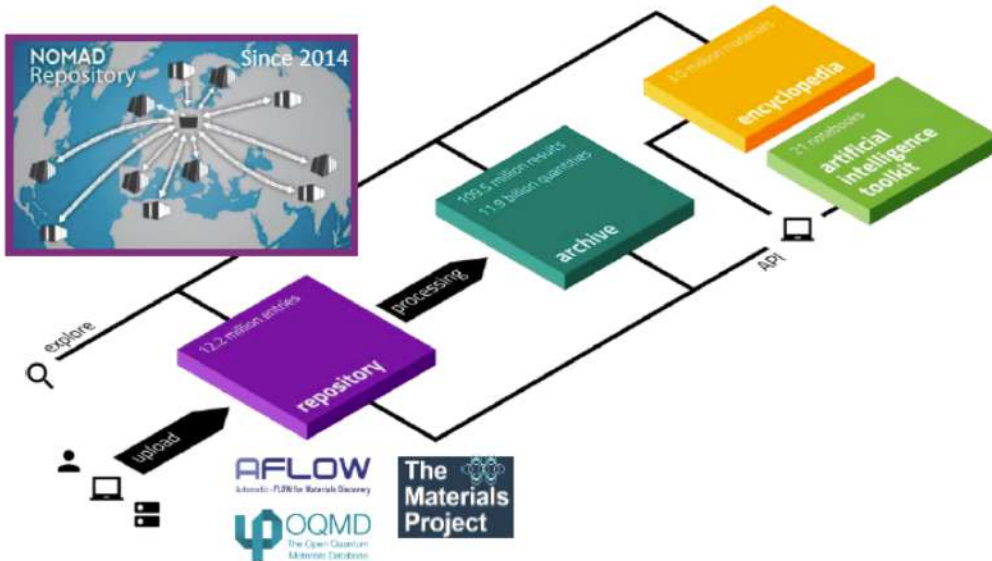
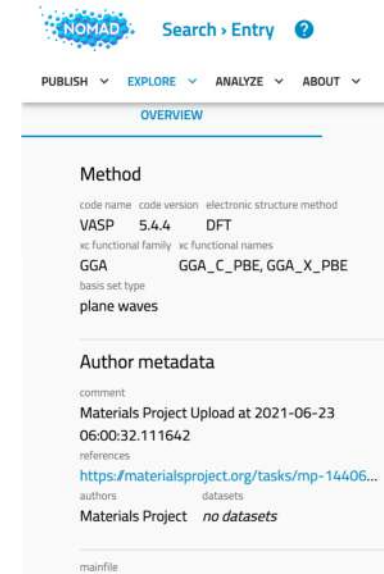
The screenshot shows the NOMAD web interface for a specific entry. The top navigation bar includes 'NOMAD Search > Entry' and 'LOGIN / REGISTER'. Below the navigation are tabs for 'PUBLISH', 'EXPLORE', 'ANALYZE', and 'ABOUT'. The main content area is divided into several sections:

- OVERVIEW**: Contains 'Method' details (code name: VASP, version: 5.4.4, electronic structure method: DFT), 'Author metadata' (Materials Project Upload at 2021-06-23, 06:00:32.111642), and 'references' (link to <https://materialsproject.org/tasks/mp-14406...>).
- RAW DATA**: Contains 'Material' details (formula: Ca4O12Ti4, material type: bulk, material name: unavailable, crystal system: orthorhombic, spacegroup: Pnma (62)).
- ARCHIVE**: Contains a 3D visualization of the material structure, with 'ORIGINAL' and 'CONVENTIONAL' views.
- LOGS**: Contains 'Electronic properties' (Band structure, Density of states).

FAIR Theory and Computations



Biggest database in Materials Science (<https://nomad-lab.eu>)
more than 100 million high-quality calculations

NOMAD Search > Entry

PUBLISH EXPLORE ANALYZE ABOUT

OVERVIEW

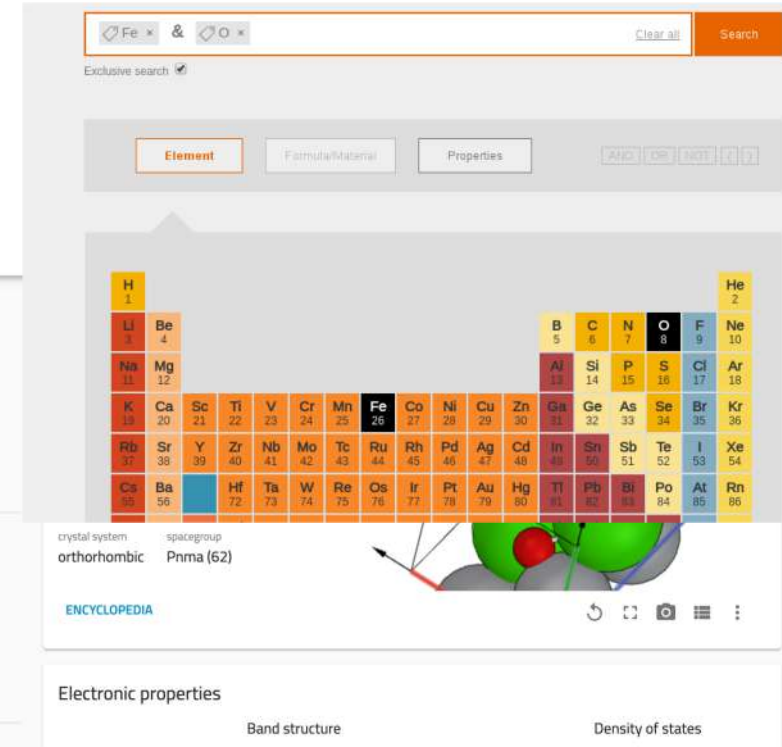
Method

code name code version electronic structure method
 VASP 5.4.4 DFT
 vc: functional family vc: functional names
 GGA GGA_C_PBE, GGA_X_PBE
 basis set type
 plane waves

Author metadata

comment
 Materials Project Upload at 2021-06-23
 06:00:32.111642
 references
<https://materialsproject.org/tasks/mp-14406...>
 authors datasets
 Materials Project no datasets

mainfile



Search results for Fe & O. The page shows a periodic table with Fe and O highlighted. Below the table, the crystal system is identified as orthorhombic with space group Pnma (62). The page also displays electronic properties, including band structure and density of states.

crystal system: orthorhombic
 spacegroup: Pnma (62)

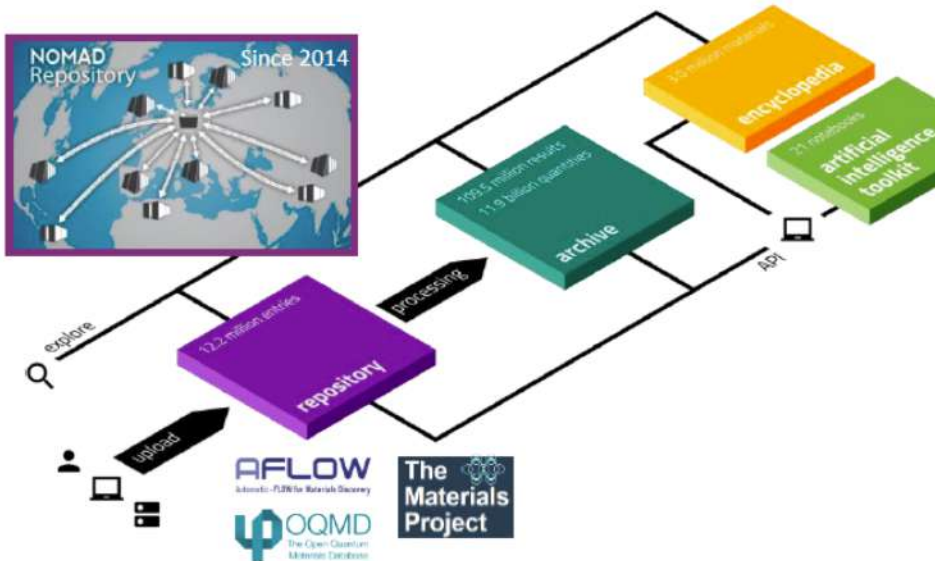
ENCYCLOPEDIA

Electronic properties
 Band structure Density of states

FAIR Theory and Computations



Biggest database in Materials Science (<https://nomad-lab.eu>)
more than 100 million high-quality calculations



The screenshot shows the NOMAD web interface. At the top, there is a search bar with 'Fe' and 'O' entered. Below it, there are filters for 'Element', 'Formula/Material', and 'Properties'. The main content area displays results for 'AgFeO₃ - space group 221'. The interface is divided into several panels:

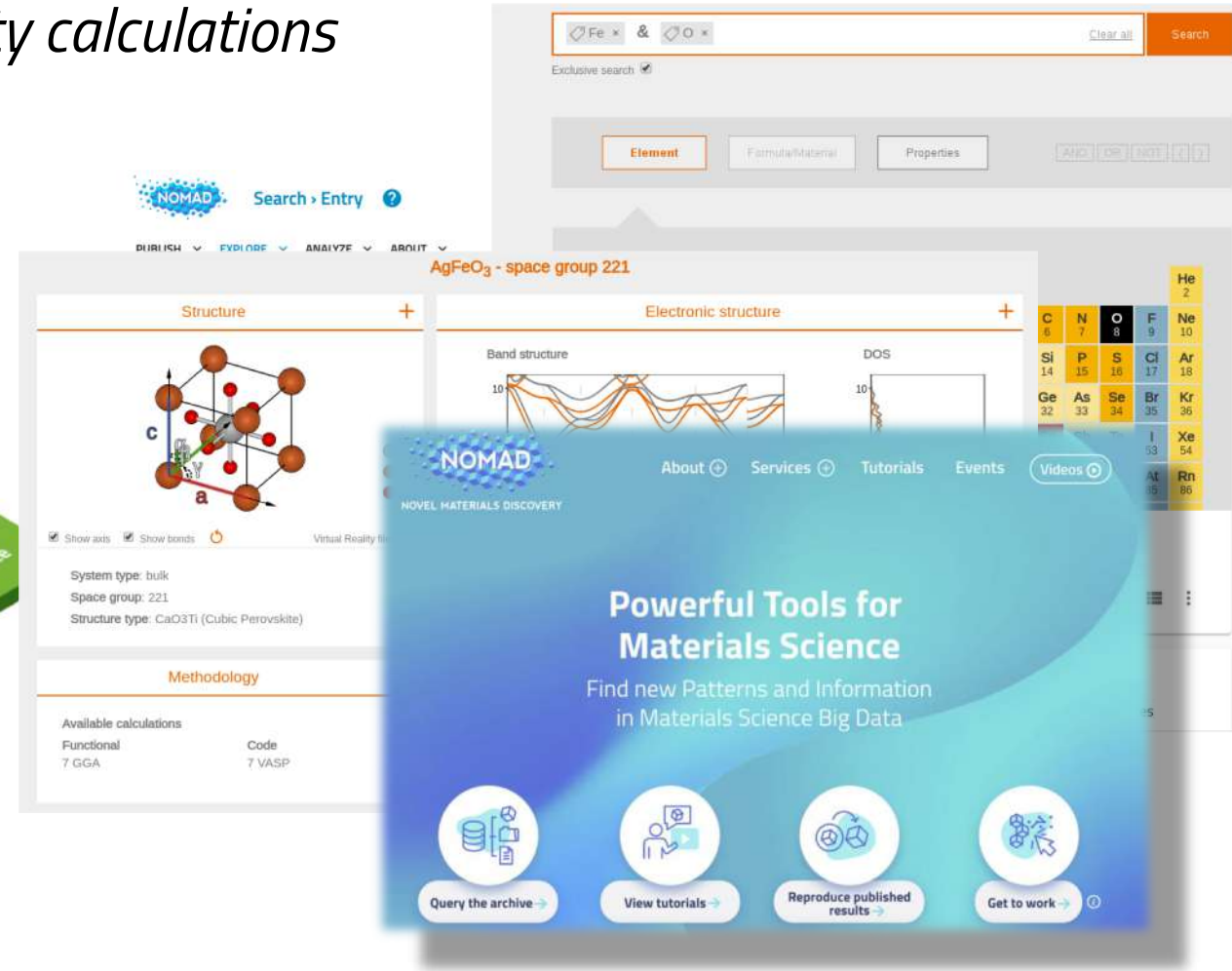
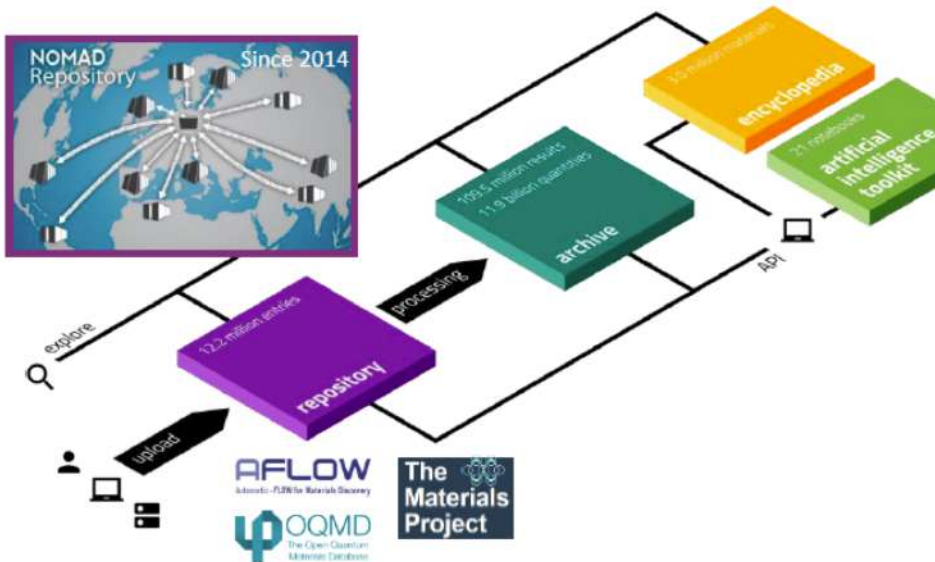
- Structure:** Shows a 3D ball-and-stick model of the crystal structure. The legend indicates Ag (grey), Fe (orange), and O (red). The structure type is identified as CaO3Ti (Cubic Perovskite).
- Electronic structure:** Contains two plots: 'Band structure' and 'DOS (states/eV/cell)'. The band structure plot shows energy levels from -5 to 10 eV along high-symmetry paths (Γ-X-M-Γ-R-X|M-R). The DOS plot shows the density of states from 0 to 12 eV.
- Methodology:** Lists available calculations: 7 GGA and 7 VASP.

On the right side of the interface, a periodic table is visible, highlighting the elements Ag, Fe, and O.

FAIR Theory and Computations



Biggest database in Materials Science (<https://nomad-lab.eu>)
more than 100 million high-quality calculations

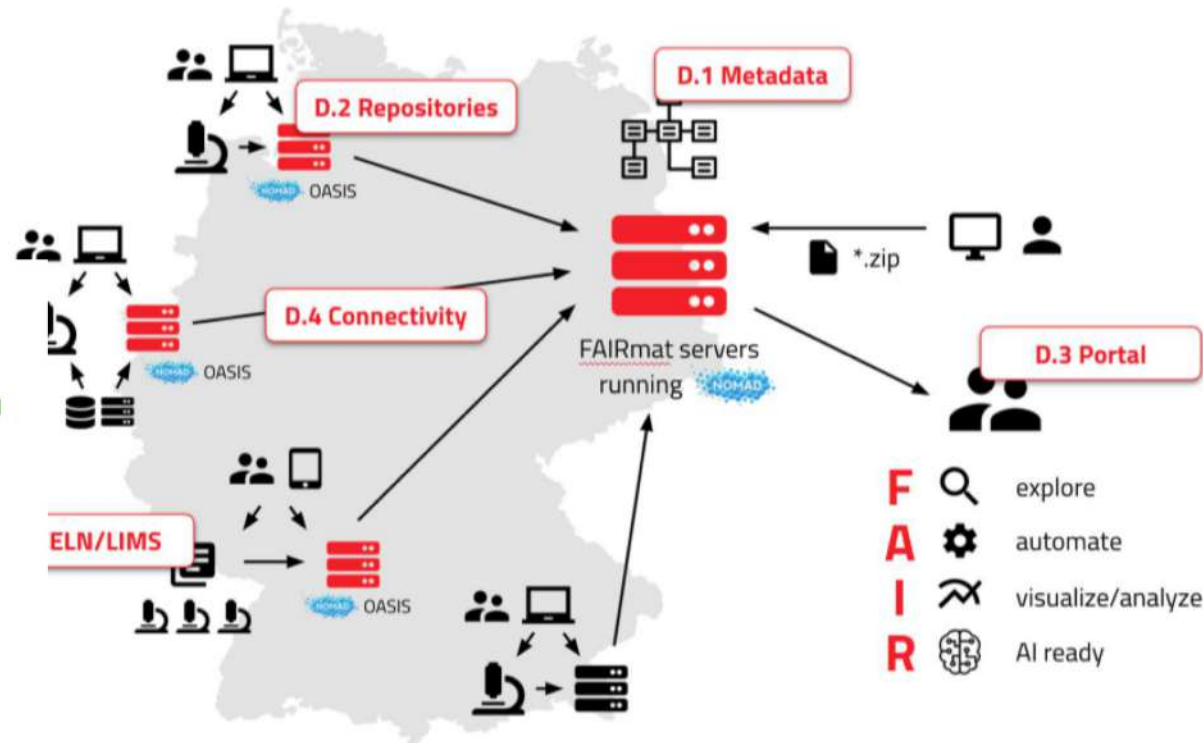
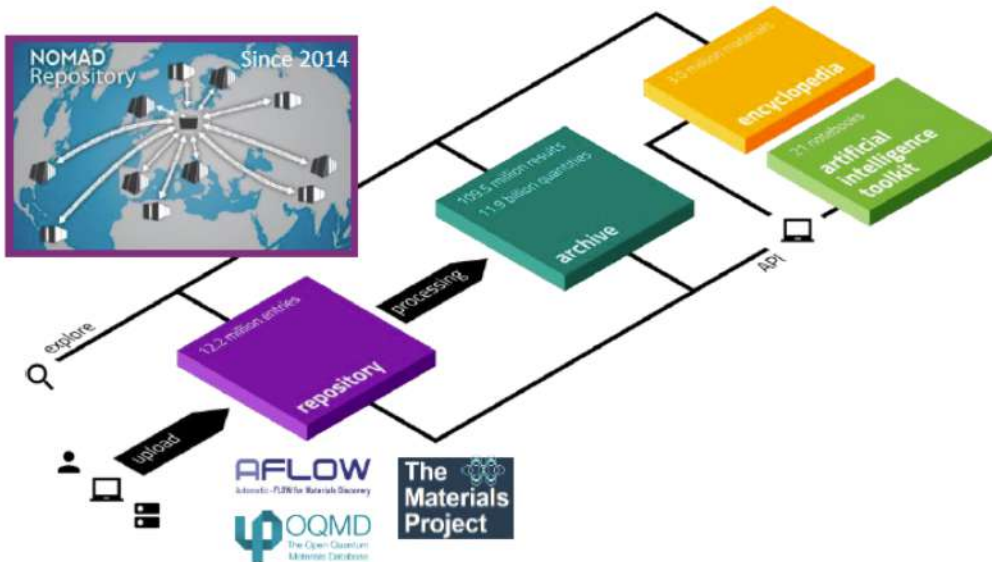


The screenshot shows the NOMAD web interface for the entry 'AgFeO₃ - space group 221'. The interface includes a search bar at the top with 'Fe' and 'O' entered. Below the search bar are buttons for 'Element', 'Formula/Material', and 'Properties'. The main content area is divided into two panels: 'Structure' and 'Electronic structure'. The 'Structure' panel shows a 3D ball-and-stick model of the crystal structure. The 'Electronic structure' panel shows a band structure plot and a Density of States (DOS) plot. Below these plots, the system type is listed as 'bulk', the space group as '221', and the structure type as 'CaO3Ti (Cubic Perovskite)'. The 'Methodology' section lists 'Available calculations' as 'Functional' (7 GGA) and 'Code' (7 VASP). A navigation menu at the top right includes 'About', 'Services', 'Tutorials', 'Events', and 'Videos'. A blue overlay at the bottom right contains the text 'Powerful Tools for Materials Science' and 'Find new Patterns and Information in Materials Science Big Data', with four icons representing 'Query the archive', 'View tutorials', 'Reproduce published results', and 'Get to work'.

FAIR Theory and Computations



Biggest database in Materials Science (<https://nomad-lab.eu>)
more than 100 million high-quality calculations

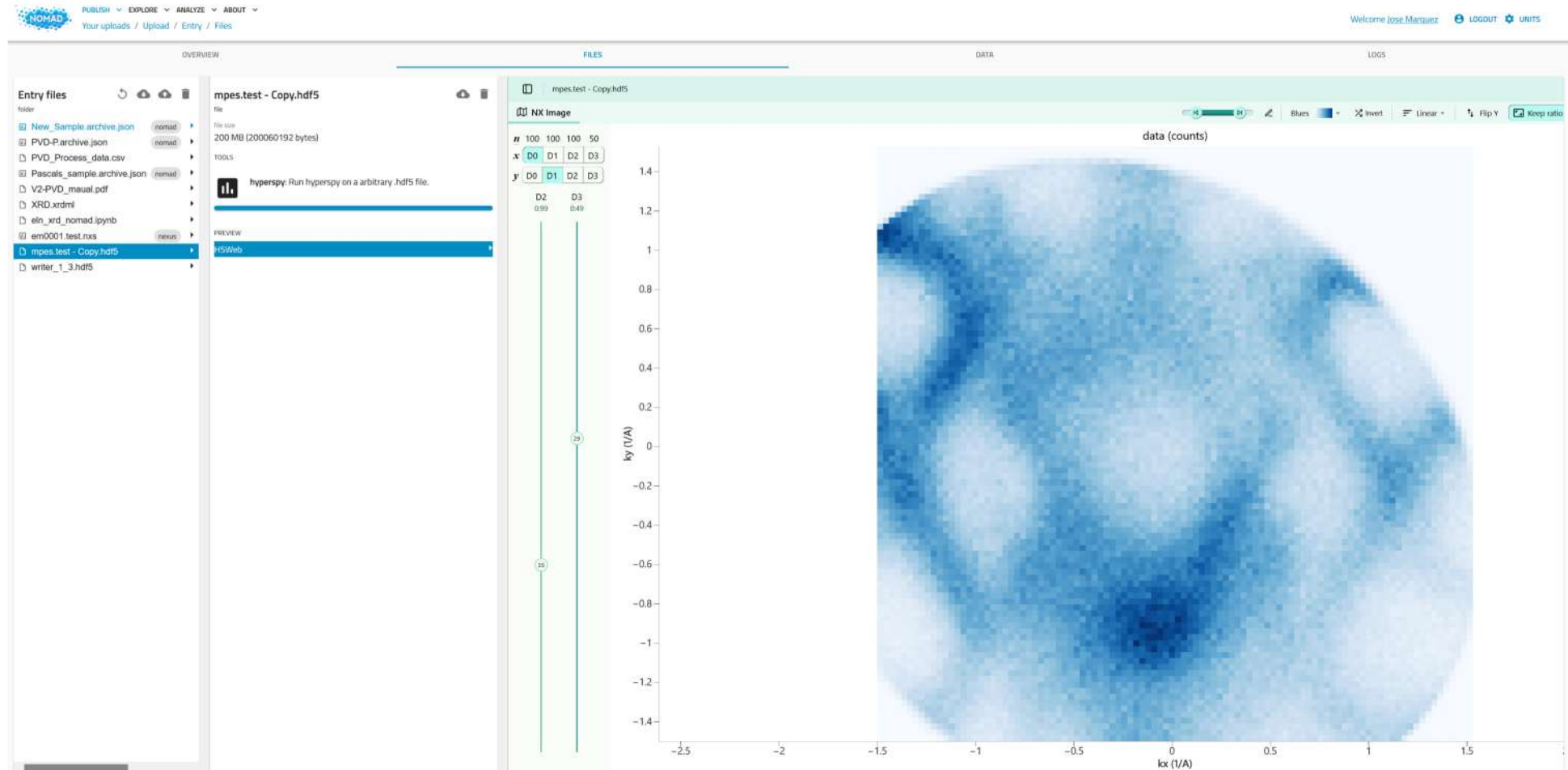




FAIR Theory and Computations



Biggest database in Materials Science (<https://nomad-lab.eu>)
more than 100 million high-quality calculations

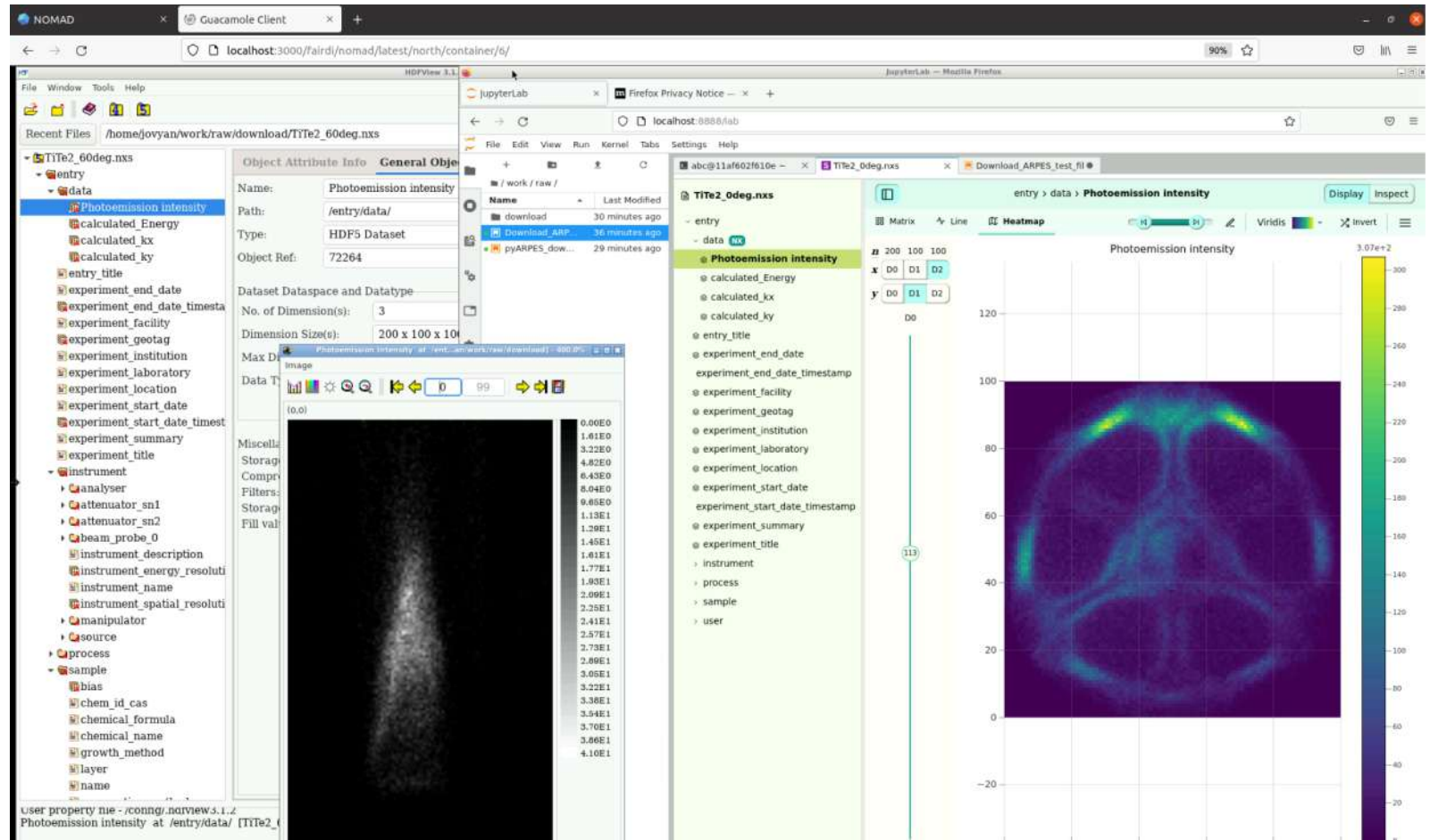




FAIR Theory and Computations



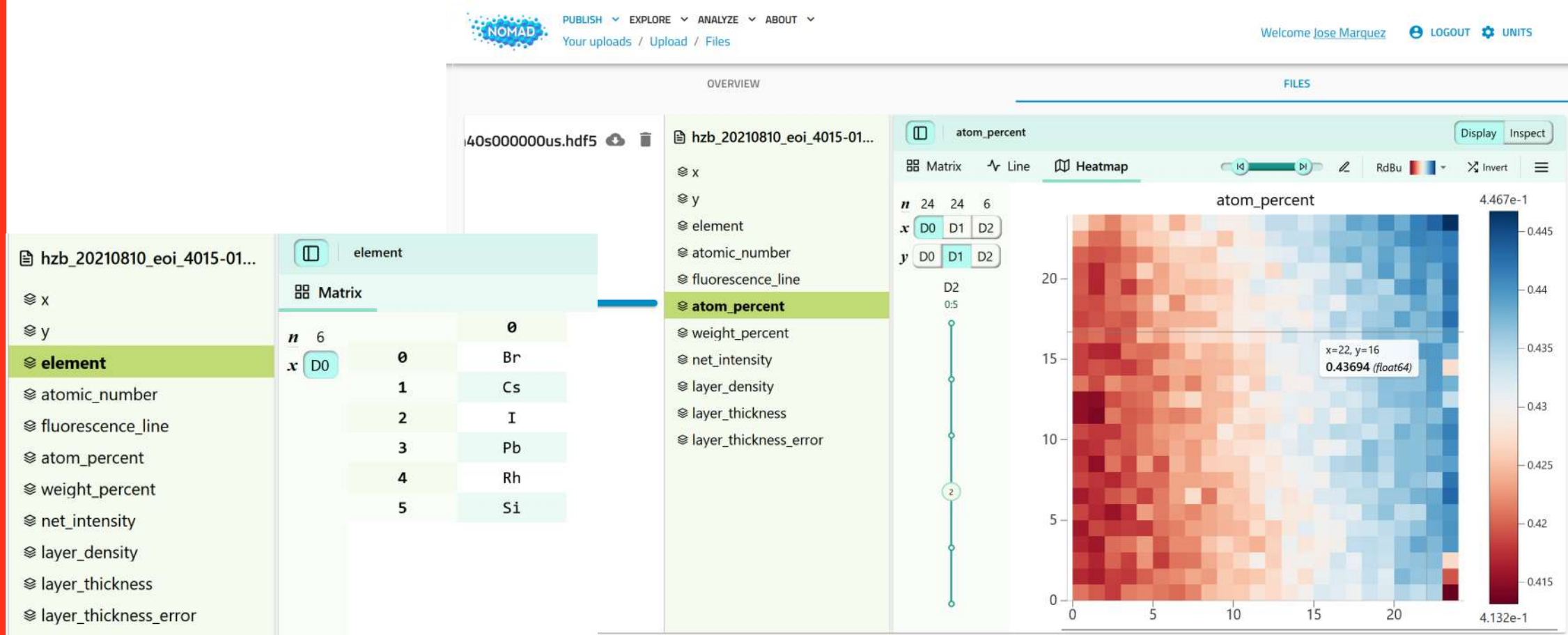
Biggest database in Materials Science (<https://nomad-lab.eu>)
more than 100 million high-quality calculations



FAIR Theory and Computations



Biggest database in Materials Science (<https://nomad-lab.eu>)
more than 100 million high-quality calculations



The screenshot displays the NOMAD web interface. The top navigation bar includes links for PUBLISH, EXPLORE, ANALYZE, and ABOUT. The user is logged in as Jose Marquez. The main content area is divided into several panels:

- Overview Panel:** Shows a file named `40s000000us.hdf5` and a selected file `hzb_20210810_eoi_4015-01...`. A list of variables is shown, with `atom_percent` selected.
- Matrix Panel:** Displays a matrix for the selected file. The matrix has 6 rows and 6 columns. The first row is labeled `0` and the first column is labeled `D0`. The elements are:

	0	1	2	3	4	5
0	0	Br				
1		Cs				
2			I			
3				Pb		
4					Rh	
5						Si
- Heatmap Panel:** Shows a heatmap visualization of the `atom_percent` data. The x and y axes range from 0 to 24. A tooltip indicates a value of `0.43694 (float64)` at `x=22, y=16`. The color scale ranges from `4.132e-1` (dark red) to `4.467e-1` (dark blue).



FAIR Theory and Computations



Biggest database in Materials Science (<https://nomad-lab.eu>)
more than 100 million high-quality calculations

The screenshot displays the NOMAD web interface with the following components:

- Navigation:** PUBLISH, EXPLORE, ANALYZE, ABOUT, Your uploads / Upload / Entry / Data. User: Welcome Jose Marquez, LOGOUT, UNITS.
- Search:** search
- Entry Panel:** Includes sections for results, metadata, and data.
- Sample Panel:** Details for 'Sample' including owner (Pepe Marquez), ID (1988), creator (16/03/2022 18:44), and institution (HU Berlin).
- Processes Panel:** Lists sub-sections like pvd_evaporation, pld_deposition, ebeam_evaporation, hotplate_annealing, tube furnace_annealing, rtp_annealing, spin_coating, and chemical_bath_deposition.
- PVDEvaporation Panel:** Shows parameters such as Operator (Pepe Marquez), Date time (02/03/2022 18:45), Instrument (PVD-P), and a checked option 'Creates layer'. It also includes a comments editor.
- substrate_temperature Panel:** Features a plot of substrate temperature (K) vs process_time (s). The plot shows a temperature rise from ~306 K to a peak of ~316 K at 4000s, followed by a gradual decrease. Below the plot is a data table with values ranging from 305.5632 to 305.9654 K.



FAIR Theory and Computations



Biggest database in Materials Science (<https://nomad-lab.eu>)
more than 100 million high-quality calculations

Screenshot of the NOMAD Metainfo Browser interface. The browser address bar shows the URL: `localhost:3000/faird/nomad/latest/gui/analyze/metainfo/EntryArchive/nexus/nx_application_ellipsometry/nx_group_ENTRY/nx_field_definition/nx_attribu...`

The interface displays a search results page for the term "nexus". The search results are organized into columns, each representing a different section of the data model:

- nexus**: Sub-section definition. Includes sub-section definitions like `nx_application_apm`, `nx_application_archive`, `nx_application_arpes`, `nx_application_canSAS`, `nx_application_directtof`, `nx_application_tofraw`, `nx_application_ellipsometry`, `nx_application__TIME`, `nx_application_em_nion`, `nx_application_fluo`, `nx_application_indirecttof`, `nx_application_lqproc`, `nx_application_laueetof`, `nx_application_monopd`, `nx_application_mpes`, `nx_application_electronanalyser`, `nx_application_energydispersion`, `nx_application_calibration`, `nx_application_mx`, `nx_application_refscan`, `nx_application_retof`, `nx_application_sas`, `nx_application_sastof`, `nx_application_scan`, `nx_application_spe`, `nx_application_sqom`, `nx_application_stxm`, `nx_application_tas`, `nx_application_tofmpd`, `nx_application_tofsingle`, `nx_application_tomo`, `nx_application_tomophase`, and `nx_application_tomoproc`.
- nx_application_ellipsometry**: Sub-section definition. Description: "Draft application definition for ellipsometry measurements, including complex systems up to variable angle spectroscopic ellipsometry. In this application definition, times should be specified always together with a UTC offset." Includes links to "nexus manual", properties like `nx_kind: group` and `nx_optional: false`, and a "SHOW USAGE" button.
- nx_group_ENTRY**: Sub-section definition. Description: "This is the application definition describing ellipsometry experiments. Such experiments may be as simple as identifying how a reflected beam of light with a single wavelength changes its polarization state, to a variable angle spectroscopic ellipsometry experiment. The application definition defines: - elements of the experimental instrument - calibration information if available - parameters used to tune the state of the sample - sample description" Includes links to "nexus manual", properties like `nx_kind: group` and `nx_optional: false`, and a "SHOW USAGE" button.
- nx_field_definition**: Sub-section definition. Description: "An application definition for ellipsometry." Includes links to "nexus manual", properties like `nx_kind: field`, `nx_type: NX_CHAR`, and `nx_optional: false`, and a "SHOW USAGE" button.
- nx_attribute_version**: Sub-section definition. Description: "Version number to identify which definition of this application definition was used for this entry/data." Includes links to "nexus manual", properties like `nx_kind: attribute` and `nx_optional: false`, and a "SHOW USAGE" button.

NeXus-FAIRmat

Proposal of NeXus expansion for FAIRmat data.

[Watch](#)

Navigation

FAIRmat-NeXus Proposal

NeXus Documentation

- NeXus: User Manual
- Examples of writing and reading NeXus data files
- NeXus: Reference Documentation
- NeXus Community
- Installation
- NeXus Utilities
- About these docs

MPES Structure

Ellipsometry Structure

Electron Microscopy Structure

Atom Probe Microscopy Structure

Quick search

[Go](#)

Google search

[Go](#)

global NeXus manual [Go](#)



Application Definition

- NXuser
- NXinstrument
 - NXexcitation
 - NXdetection
- NXsample
 - Wavelength_detection[detection]
 - Measured_data[time,parameter,excitation,detection]
 - Timepoints[time]
 - Angle_of_incidence[angles]
 - Sample_stage
 - ...
- plot



doc: "Draft application definition for general optical experiment
symbols:

doc: Variables used throughout the document, e.g.

N_wavelength_detection: Size of the energy / wavelength

N_wavelength_excitation: "Size of the array of wavelengths used for
excitation in the case of photoluminescence"

N_parameters: Number of sample parameters specified

N_time: Number of time points measured

(NXopt_experiment):

(NXentry):

doc: |

This is an application definition describing general optical experiments

A general optical experiment consists of a light source, a detector, a
beam path, a sample + stage + environment, and various types of measurements,
such as reflection or transmission measurements, Raman spectroscopy,
infrared spectroscopy etc.

This application definition defines:

- * information on excitation and detection
- * description of the beam path
- * sample description
- * N-dimensional data field

definition: