



Generating and publishing datasets using AiiDA and Materials Cloud

Kristjan Eimre
EPFL, Switzerland

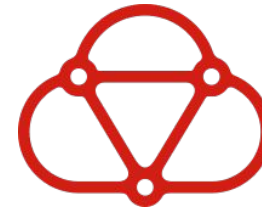
CECAM workshop: Machine Learning Interatomic Potentials and Accessible Databases
Grenoble
11.09.2024

Outline



**Automated Interactive Infrastructure and
Database for Computational Science**

- Recent usability improvements



**MATERIALS
CLOUD**

**Web platform for seamlessly sharing data
and resources**

- Materials Cloud 3D crystals database updates
- Archive-OPTIMADE integration



<http://www.aiida.net>

Computational science infrastructure that provides

- scalable workflow engine
- built-in support for HPC
- flexible plugin system
- automatic full data provenance

Language: implemented and API in python

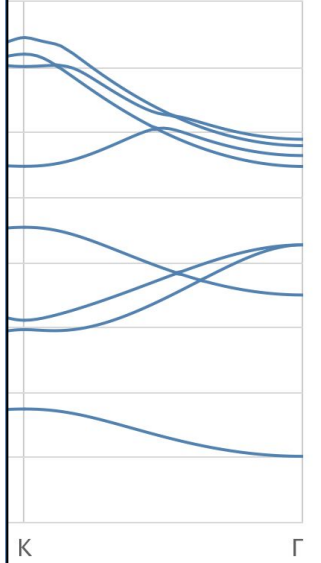
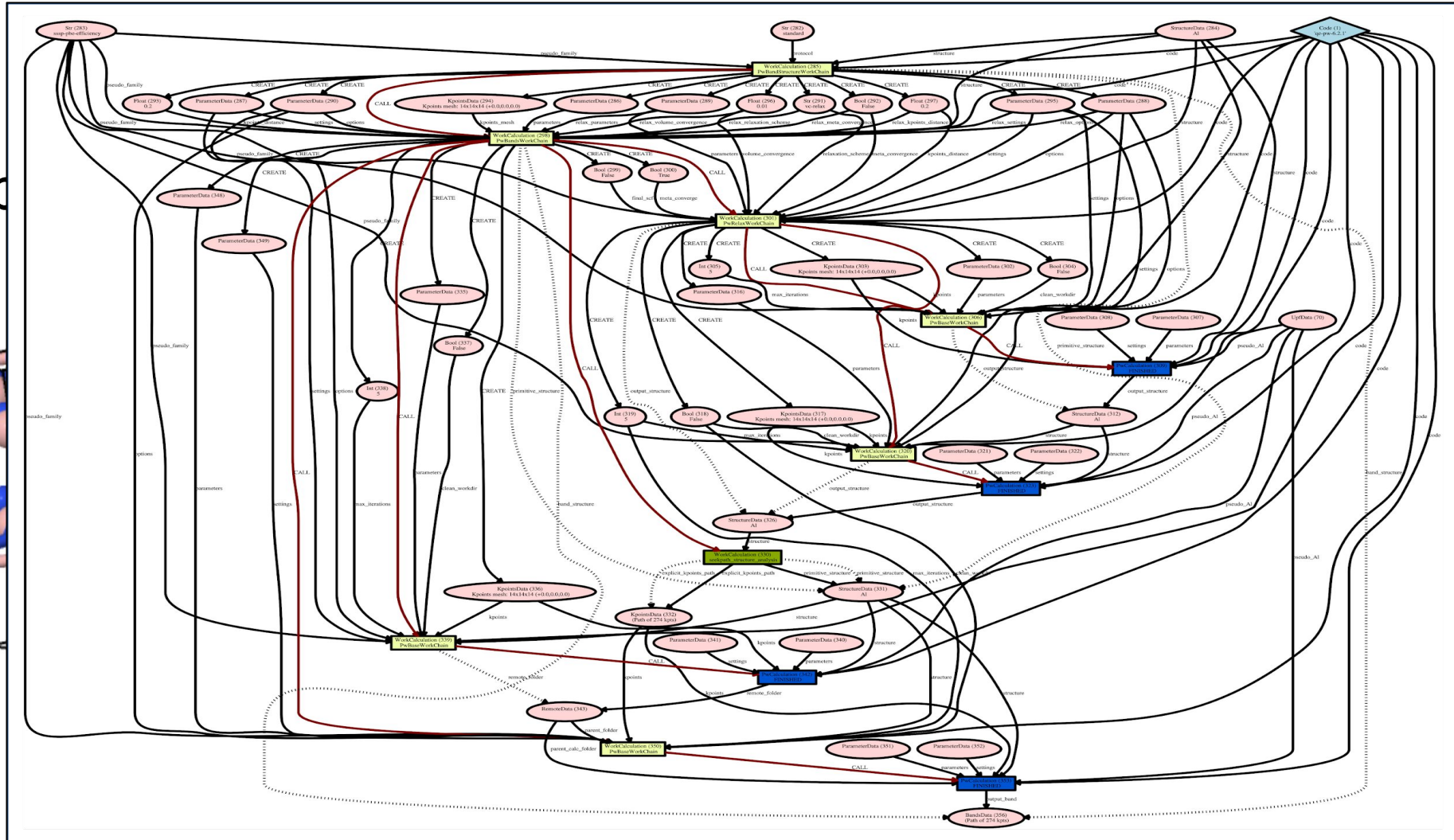
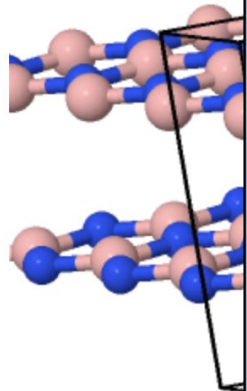
Open source: <https://github.com/aiidateam/aiida-core>

G. Pizzi et al., Comp. Mat. Sci. 111, 218-230 (2016)

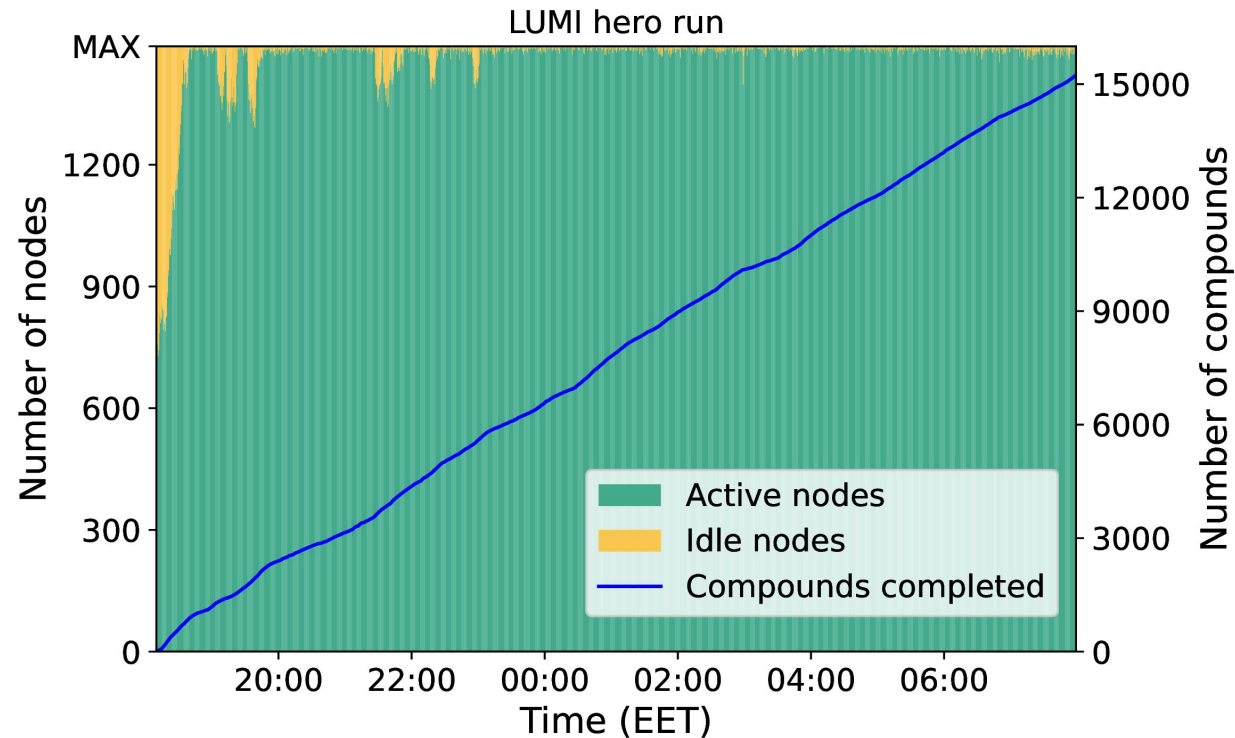
S.P. Huber et al., Scientific Data 7, 300 (2020)

Workflow engine

Example



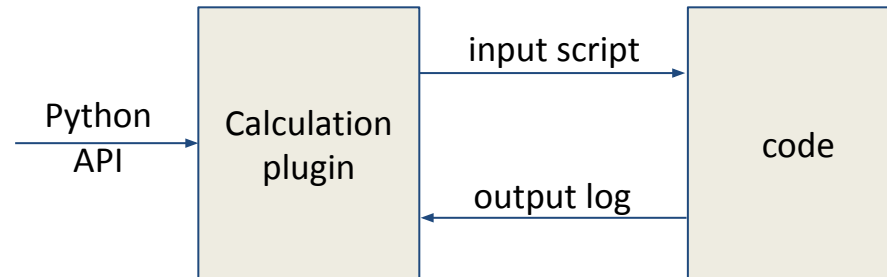
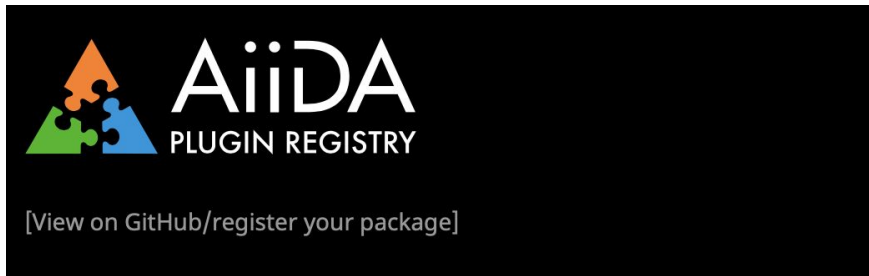
Scalable workflow engine



- Geometry relaxations of 15,324 inorganic compounds in 13 hours
- Full partition of LUMI-C: 1,500 nodes with 128 cores each
- 7887 issues dealt with on the fly

Plugins

<https://aiidateam.github.io/aiida-registry/>



Registered plugin packages: 91

Calculations	131 plugins in 53 packages
Parsers	109 plugins in 54 packages
Data	103 plugins in 30 packages
Workflows	135 plugins in 39 packages
Console scripts	26 plugins in 15 packages
Other	99 plugins in 26 packages

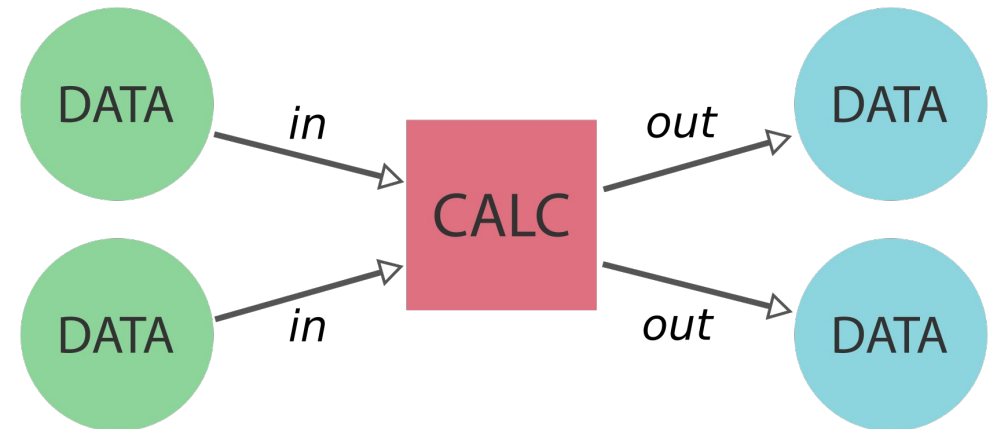
- Plugins collected in the AiiDA plugin registry
- 130+ codes currently supported
- Many are community-contributed

Data provenance

What you get as a bonus: full data provenance!

AiiDA stores:

- data transformations/calculations
- inputs
- outputs
- inter-connections

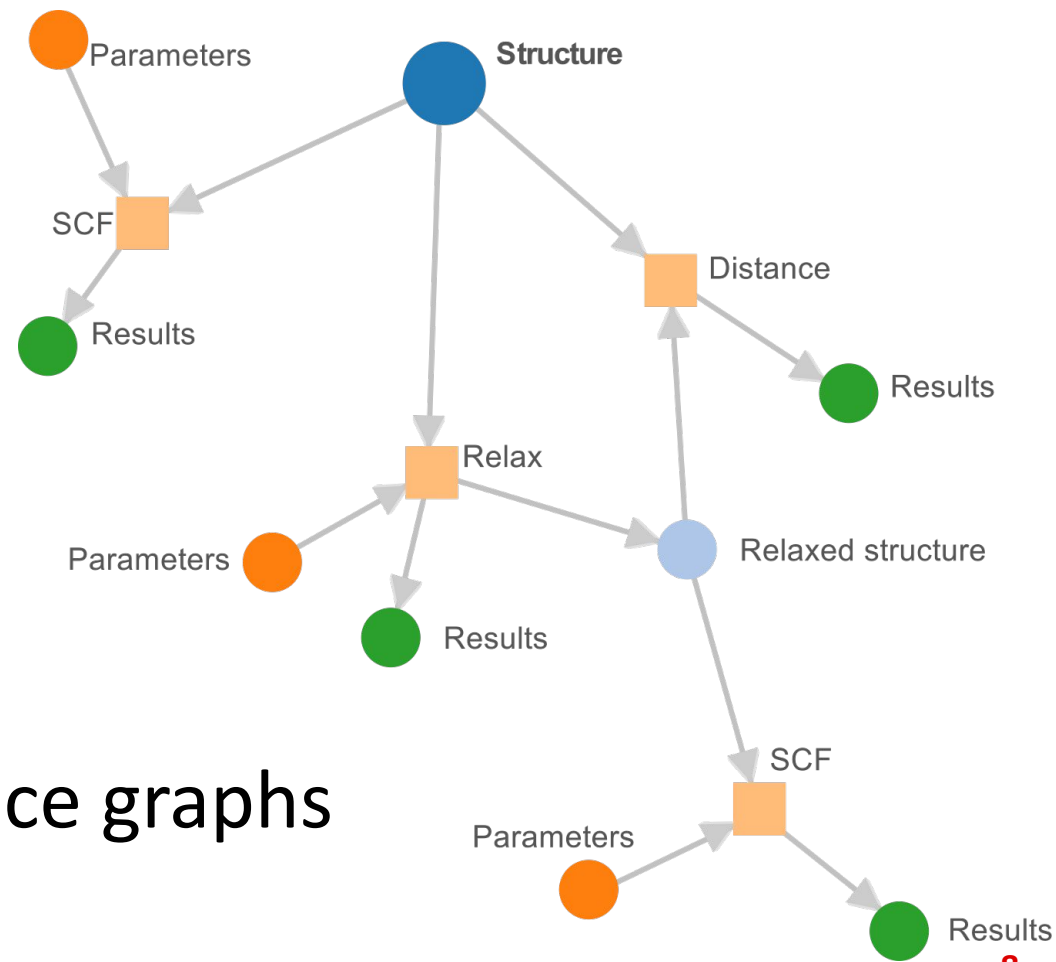


Data provenance

What you get as a bonus: full data provenance!

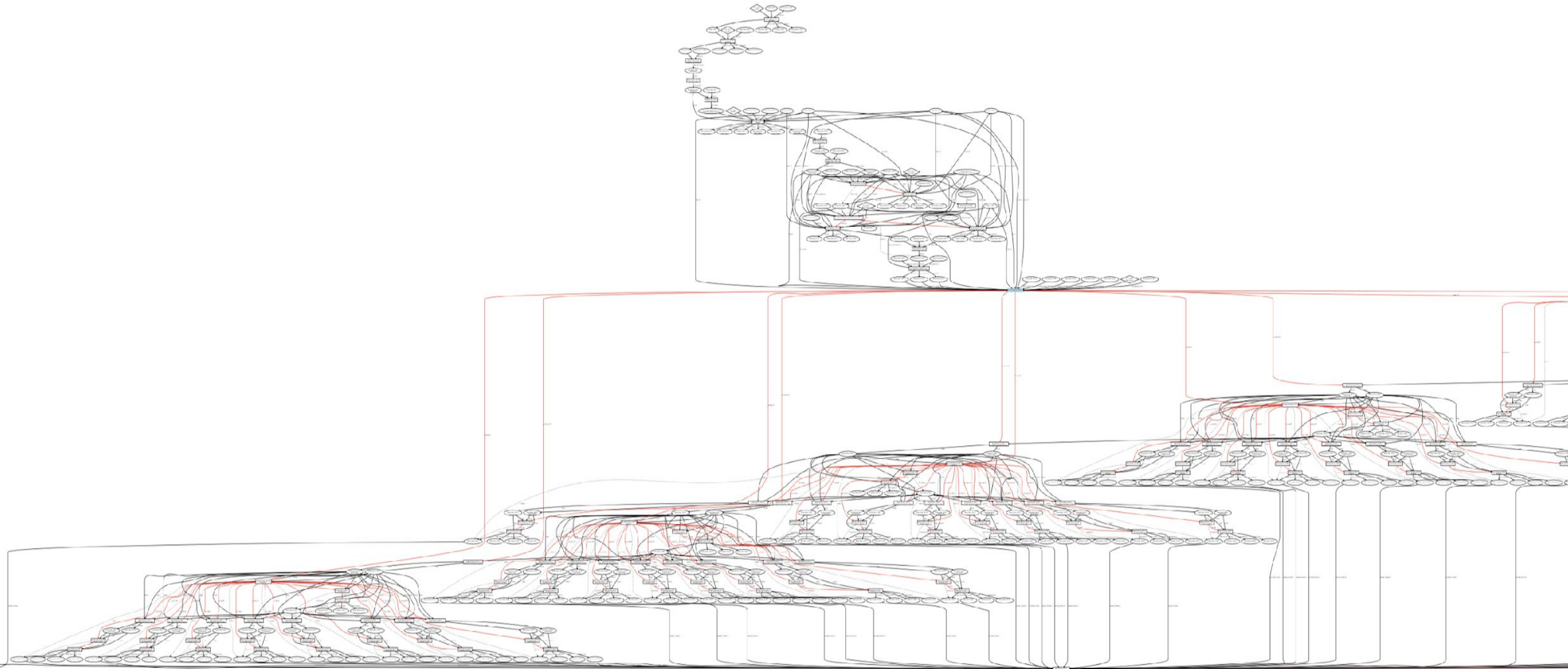
AiiDA stores:

- data transformations/calculations
- inputs
- outputs
- inter-connections



Provenance graphs

Data provenance



Molecular dynamics study of Lithium in a solid electrolyte

AiiDA improvements

Recent effort to improve the usability and reduce the learning curve

AiiDA improvements: installation

Traditional (Complete) installation

- Requires
 - PostgreSQL
 - RabbitMQ

```
$ pip install -q aiida-core
$ verdi profile setup core.psql_dos
Report: enter ? for help.
Report: enter ! to ignore the default and set no
Profile name: new_profile
Set as default? [Y/n]: n
Email Address (for sharing data) [kristjan.eimre@
First name [Kristjan]:
Last name [Eimre]:
Institution [EPFL]:
Use RabbitMQ? [Y/n]:
PostgreSQL engine [postgresql_psycopg2]:
PostgreSQL hostname [localhost]:
PostgreSQL port [5432]:
PostgreSQL username:
```

verdi presto (v2.6.1, July 1st)

```
$ pip install -q aiida-core
$ verdi presto
Report: Option `--use-postgres` not enabled: config
Report: RabbitMQ server not found (Failed to connect
Report: See https://aiida-core.readthedocs.io/en/st
Report: Initialising the storage backend.
Report: Storage initialisation completed.
Success: Created new profile `presto-3`.
Success: Configured the localhost as a computer.
$ verdi status
✓ version:      AiiDA v2.6.2
✓ config:       /home/kristjan/.aiida
✓ profile:      presto-3
✓ storage:      SqliteDosStorage[/home/kristjan/.ai
● broker:      No broker defined for this profile:
● daemon:      No broker defined for this profile:
$
```

Not suitable for high-throughput

AiiDA improvements: aiida-shell

aiida-shell: run any executable through AiiDA without a plugin

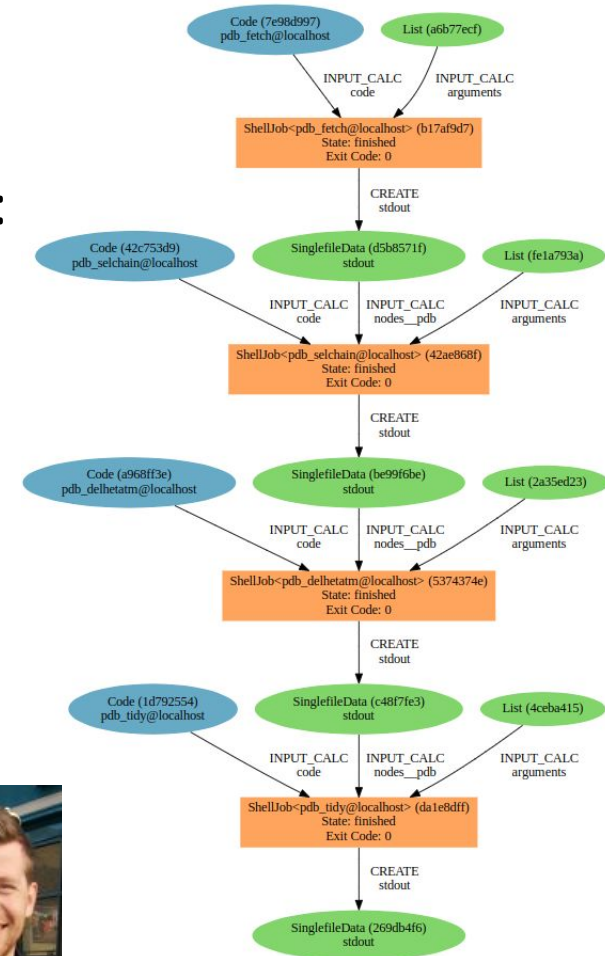
Example shell commands (manipulating Protein Data Bank files):

```
$ pdb_fetch 1brs | pdb_selchain -A,D | pdb_delhetatm | pdb_tidy > 1brs_AD_noHET.pdb
```

Run via aiida-shell:

```
from aiida_shell import launch_shell_job

results, node = launch_shell_job('pdb_fetch', '1brs')
results, node = launch_shell_job('pdb_selchain', '-A,D {pdb}', {'pdb': results['stdout']})
results, node = launch_shell_job('pdb_delhetatm', '{pdb}', {'pdb': results['stdout']})
results, node = launch_shell_job('pdb_tidy', '{pdb}', {'pdb': results['stdout']})
```



<https://github.com/sphuber/aiida-shell>

Dr. Sebastiaan
P. Huber

AiiDA improvements: WorkGraph

Workflow: $(x + y) * z$ extending to $(x + y) * z + a$

Calculations

```
@calcfunction
def add(x, y):
    return Int(x + y)

@calcfunction
def multiply(x, y):
    return Int(x * y)
```

WorkChain

```
class AddMultiplyWorkChain(WorkChain):

    @classmethod
    def define(cls, spec):
        super().define(spec)
        spec.input('x')
        spec.input('y')
        spec.input('z')
        spec.outline(
            cls.add,
            cls.multiply,
            cls.results,
        )
        spec.output('result')

    def add(self):
        self.ctx.sum = add(self.inputs.x,
self.inputs.y)

    def multiply(self):
        self.ctx.product =
multiply(self.ctx.sum, self.inputs.z)

    def results(self):
        self.out('result', self.ctx.product)
```

WorkGraph

```
wg = WorkGraph("add_multiply")
wg.tasks.new(add, name="add")
wg.tasks.new(multiply, name="multiply",
x=wg.tasks["add"].outputs["result"])
```

<https://github.com/aiidateam/aiida-workgraph>



Dr. Xing
Wang

AiiDA improvements: WorkGraph

Workflow: $(x + y) * z$ extending to $(x + y) * z + a$

WorkChain

WorkGraph

```
class AddMultiplyAddWorkChain(WorkChain):  
  
    @classmethod  
    def define(cls, spec):  
        super().define(spec)  
        spec.expose_inputs(AddMultiplyWorkChain,  
namespace='add_multiply')  
        spec.input('a')  
        spec.outline(  
            cls.add_multiply,  
            cls.add2,  
            cls.results,  
        )  
        spec.output('result')  
  
    def add_multiply(self):  
        future =  
self.submit(AddMultiplyWorkChain,  
**self.exposed_inputs(AddMultiplyWorkChain,  
'add_multiply'))  
        return ToContext(add_multiply=future)  
  
    def add2(self):  
        self.ctx.sum =  
add(self.ctx.add_multiply.outputs.result,  
self.inputs.a)  
  
    def results(self):  
        self.out('result', self.ctx.sum)
```

```
class AddMultiplyWorkChain(WorkChain):  
  
    @classmethod  
    def define(cls, spec):  
        super().define(spec)  
        spec.input('x')  
        spec.input('y')  
        spec.input('z')  
        spec.outline(  
            cls.add,  
            cls.multiply,  
            cls.results,  
        )  
        spec.output('result')  
  
    def add(self):  
        self.ctx.sum = add(self.inputs.x,  
self.inputs.y)  
  
    def multiply(self):  
        self.ctx.product =  
multiply(self.ctx.sum, self.inputs.z)  
  
    def results(self):  
        self.out('result', self.ctx.product)
```

```
wg = WorkGraph("add_multiply")  
wg.tasks.new(add, name="add")  
wg.tasks.new(multiply, name="multiply",  
x=wg.tasks["add"].outputs["result"])
```

```
wg.tasks.new(add, name="add2",  
x=wg.tasks["multiply"].outputs["result"])
```

<https://github.com/aiidateam/aaida-workgraph>



Dr. Xing
Wang

AiiDA improvements

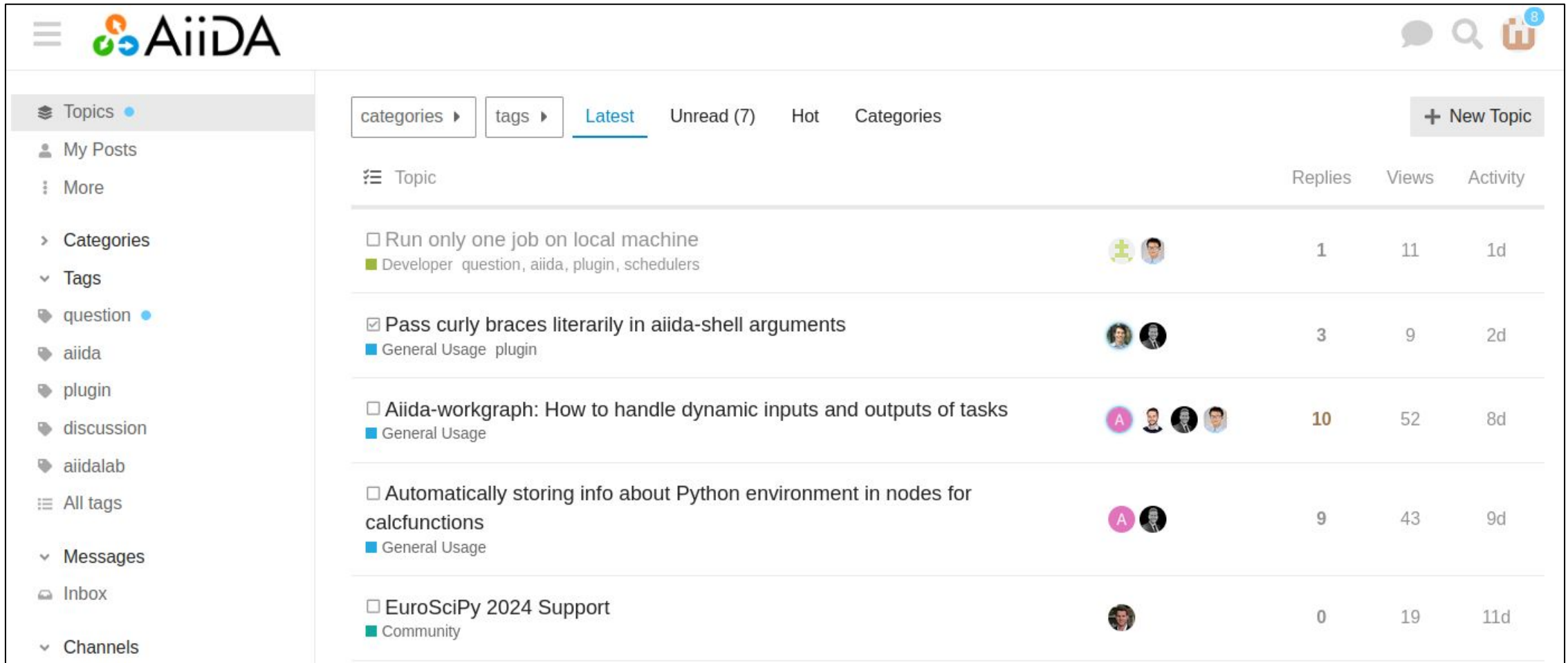
- verdi storage backup <destination>
 - incremental backup of the AiiDA profile
 - supports remote paths directly
- verdi process dump <uuid/pk>
 - dumps all input and output scripts of the selected aiiida process into a human-readable folder structure

```
> verdi process dump 7727b19f-5c77-4d2a-bbb0-71a8cf87b4b7
Success: Raw files for WorkChainNode <485293> dumped into
> cd dump-PwBaseWorkChain-485293/
> tree
.
├── 01-create_kpoints_from_distance
│   └── inputs
│       └── source_file
├── 02-iteration_01-PwCalculation
│   ├── inputs
│   │   ├── _aiidasubmit.sh
│   │   └── aiida.in
│   ├── node_inputs
│   │   └── pseudos
│   │       ├── Ge
│   │       │   └── ge_pbesol_v1.4.uspp.F.UPF
│   │       └── P
│   │           └── P.pbesol-n-rrkjus_psl.1.0.0.UPF
│   └── outputs
│       ├── _scheduler-stderr.txt
│       ├── _scheduler-stdout.txt
│       ├── aiida.out
│       └── data-file-schema.xml
└── README.md

9 directories, 10 files
```

AiiDA discourse forum

<https://aiida.discourse.group/>

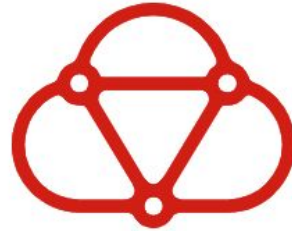


The screenshot displays the AiiDA Discourse forum interface. On the left is a navigation sidebar with options like Topics, My Posts, More, Categories, Tags, question, aiida, plugin, discussion, aiidalab, All tags, Messages, Inbox, and Channels. The main content area shows a list of topics under the 'Latest' filter. Each topic entry includes a title, tags, author avatars, and statistics for replies, views, and activity time.

Topic	Replies	Views	Activity
<input type="checkbox"/> Run only one job on local machine Developer question, aiida, plugin, schedulers	1	11	1d
<input checked="" type="checkbox"/> Pass curly braces literarily in aiida-shell arguments General Usage plugin	3	9	2d
<input type="checkbox"/> Aiiida-workgraph: How to handle dynamic inputs and outputs of tasks General Usage	10	52	8d
<input type="checkbox"/> Automatically storing info about Python environment in nodes for calcfuctions General Usage	9	43	9d
<input type="checkbox"/> EuroSciPy 2024 Support Community	0	19	11d

Materials Cloud

<https://www.materialscloud.org/>



MATERIALSCLOUD



LEARN



WORK



DISCOVER



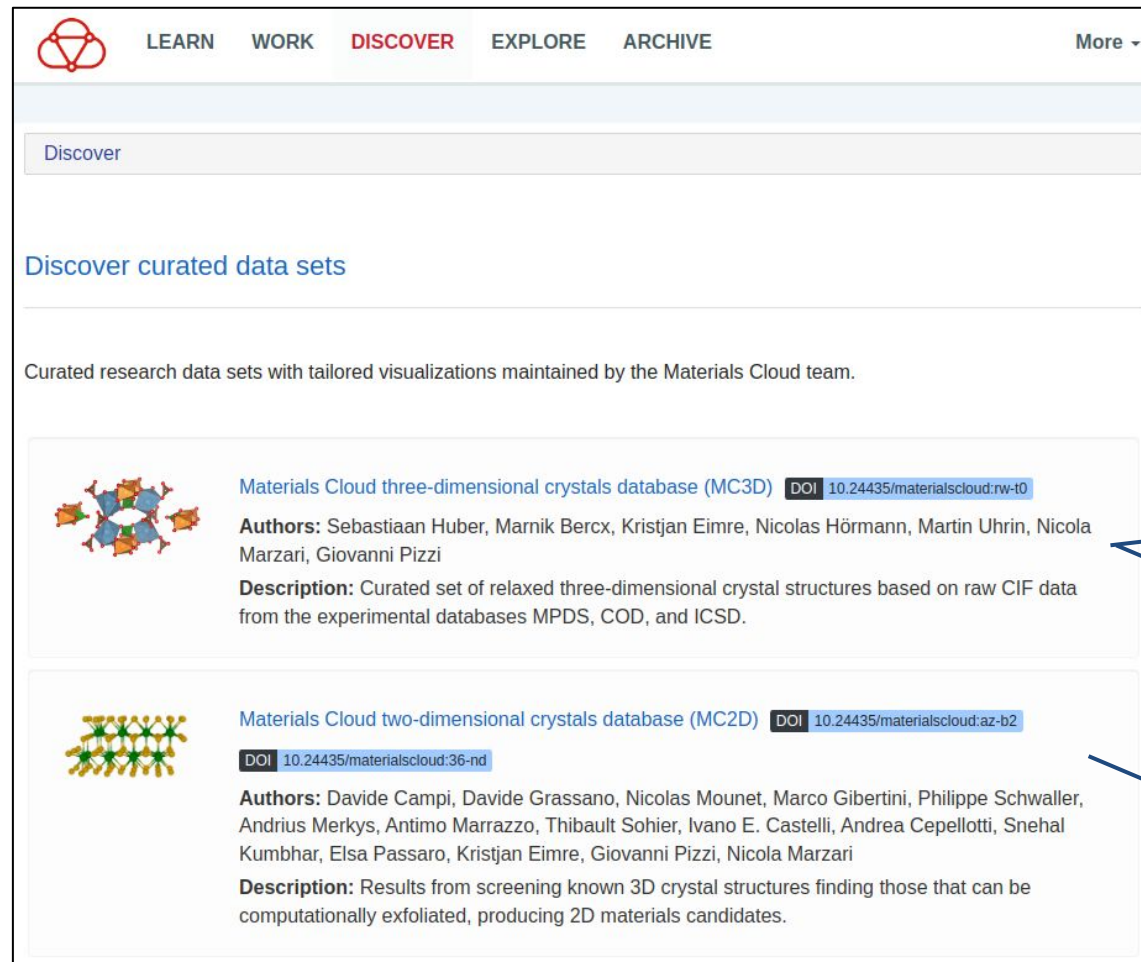
EXPLORE



ARCHIVE

Discover and explore

Discover: curated web interface

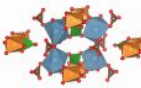
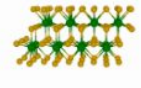


LEARN WORK **DISCOVER** EXPLORE ARCHIVE More ▾

Discover

Discover curated data sets

Curated research data sets with tailored visualizations maintained by the Materials Cloud team.

-  **Materials Cloud three-dimensional crystals database (MC3D)** DOI: 10.24435/materialscloud.rw-t0
Authors: Sebastiaan Huber, Marnik Bercx, Kristjan Eimre, Nicolas Hörmann, Martin Uhrin, Nicola Marzari, Giovanni Pizzi
Description: Curated set of relaxed three-dimensional crystal structures based on raw CIF data from the experimental databases MPDS, COD, and ICSD.
-  **Materials Cloud two-dimensional crystals database (MC2D)** DOI: 10.24435/materialscloud.az-b2
DOI: 10.24435/materialscloud.36-nd
Authors: Davide Campi, Davide Grassano, Nicolas Mounet, Marco Gibertini, Philippe Schwaller, Andrius Merkys, Antimo Marrazzo, Thibault Sohier, Ivano E. Castelli, Andrea Cepellotti, Snehal Kumbhar, Elsa Passaro, Kristjan Eimre, Giovanni Pizzi, Nicola Marzari
Description: Results from screening known 3D crystal structures finding those that can be computationally exfoliated, producing 2D materials candidates.

Explore: visualize AiiDA provenance

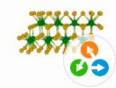


LEARN WORK DISCOVER **EXPLORE** ARCHIVE More ▾

Explore


Explore the full provenance

An interactive browser for exploring AiiDA provenance graphs uploaded to the Materials Cloud Archive.


-  **Browse your own AiiDA database**
Description: Use the REST API built into AiiDA to connect to your own AiiDA database. Your data stays inside your browser and is not transmitted to Materials Cloud.
-  **Materials Cloud three-dimensional crystals database (MC3D) (PBE-v1)**
DOI: 10.24435/materialscloud.rw-t0
Authors: Sebastiaan Huber, Marnik Bercx, Nicolas Hörmann, Martin Uhrin, Nicola Marzari, Giovanni Pizzi
Description: Curated set of relaxed three-dimensional crystal structures based on raw CIF data from the experimental databases MPDS, COD, and ICSD. Calculations are performed with the PBE-v1 methodology.
-  **Materials Cloud three-dimensional crystals database (MC3D) (PBEsol-v1)**
DOI: 10.24435/materialscloud.rw-t0
Authors: Sebastiaan Huber, Marnik Bercx, Nicolas Hörmann, Martin Uhrin, Nicola Marzari, Giovanni Pizzi
Description: Curated set of relaxed three-dimensional crystal structures based on raw CIF data from the experimental databases MPDS, COD, and ICSD. Calculations are performed with the PBEsol-v1 methodology.
-  **Materials Cloud two-dimensional crystals database (MC2D)** DOI: 10.24435/materialscloud.az-b2
DOI: 10.24435/materialscloud.36-nd
Authors: Davide Campi, Davide Grassano, Nicolas Mounet, Marco Gibertini, Philippe Schwaller, Andrius Merkys, Antimo Marrazzo, Thibault Sohier, Ivano E. Castelli, Andrea Cepellotti, Giovanni Pizzi, Nicola Marzari
Description: Results from screening known 3D crystal structures finding those that can be computationally exfoliated, producing 2D materials candidates.

Materials Cloud 3D crystals database

MPDS




ICSD
FIZ Karlsruhe



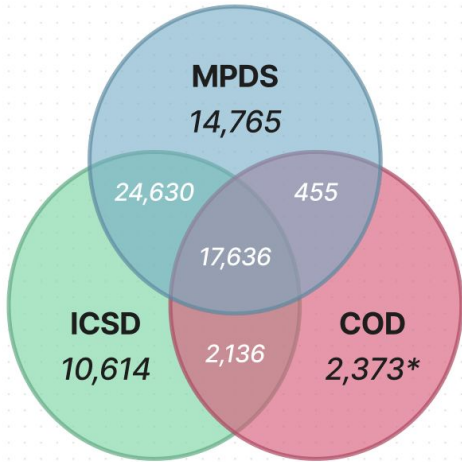
901,210
Experimental structures



 AiiDA

- Filter non-stoichiometric, organic, ...
- Check uniqueness

TOTAL: 72,609*



MC3D-source (e.g. mc3d-123)



MC3D/PBE-v1

Relaxation with `PBE-v1` methodology.
34,487 relaxed structures



MC3D/PBEsol-v1

Relaxation with `PBEsol-v1` methodology.
33,674 relaxed structures



Properties/screening:

- thermodyn stability
- spectroscopies
- electriles
- superconductors

Materials Cloud 3D crystals database

<https://mc3d.materialscloud.org/>

Use [About](#) [REST API](#)

Select a methodology: PBEsol-v1 ?

- PBEsol-v1
- PBE-v1

Elements filtering mode:

- Include/exclude ?
- Only selected

Showing 33674 entries out of 33674 ? [Reset column filters](#) [Show columns](#)

ID	Formula	Num. of atoms/cell	Space group number	Is source theoretical?	Is source high pressure?	Is source high temperature?	Total magnetization (μB/cell)
mc3d-10/pbesol-v1	Ge ₂₆ Ir ₆ Y ₆	40	223	no	no	no	-
mc3d-10000/pbesol-v1	FeSbV	3	216	no	no	no	-
mc3d-10004/pbesol-v1	Cs ₃ F ₆ Fe	10	225	no	no	yes	5.00
mc3d-10023/pbesol-v1	Ga ₁₂ Mn ₄ Zr ₆	22	59	no	no	no	7.09
mc3d-10030/pbesol-v1	B ₇ Ca ₉ F ₂ Li ₅ O ₂₁	44	1	no	no	no	-
mc3d-10031/pbesol-v1	C ₄ N ₈ Se ₄	16	62	no	no	no	-
mc3d-10044/pbesol-v1	Nb ₁₂ Ni ₆ Si ₄	24	227	no	no	no	-
mc3d-10048/pbesol-v1	Ca ₆ Ga ₄ P ₈	18	15	no	no	no	-
mc3d-10049/pbesol-v1	Cs ₄ F ₂₀ O ₂ Sb ₄	30	13	no	no	no	-
mc3d-1005/pbesol-v1	H ₈ O ₁₈ P ₄ Rh ₂	32	5	no	no	no	3.96

Page Size: 20 1 to 20 of 33,674 < < Page 1 of 1,684 > >

[Download filtered entries](#) ?

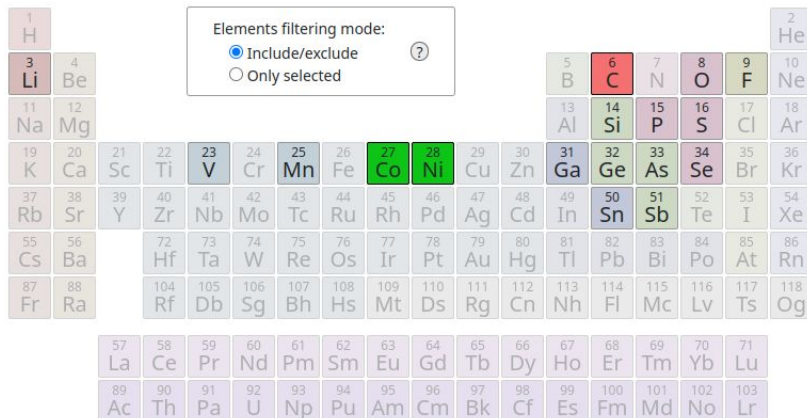
Materials Cloud 3D crystals database

<https://mc3d.materialscloud.org/>

Use [About](#) [REST API](#)

Select a methodology: PBEsol-v1 ?

Elements filtering mode:
 Include/exclude ?
 Only selected



Showing 9 entries out of 33674 ?

Reset column filters Show columns

ID	Formula	Num. of atoms/cell	Space group number	Is source theoretical?	Is source high pressure?	Is source high temperature?	Total magnetization (μB/cell)
mc3d-10103/pbesol-v1	CoNi	2	221	yes	no	no	
mc3d-22117/pbesol-v1	Co ₃ Li ₉ Mn ₃ Ni ₃ O ₁₈	36	151	no	no	no	
mc3d-31847/pbesol-v1	Co ₂ GaNi	4	139	yes	no	no	
mc3d-3377/pbesol-v1	Co ₆ Ni ₃ V ₃	12	166	no	no	no	
mc3d-40385/pbesol-v1	Co ₃ Ni	4	221	yes	no	no	
mc3d-43549/pbesol-v1	CoF ₆ Ni	8	148	no	no	no	
mc3d-43590/pbesol-v1	Co ₂ GaNi	4	225	yes	no	no	
mc3d-69079/pbesol-v1	CoGeMnNi	4	216	no	no	no	4.63
mc3d-72144/pbesol-v1	Co ₂ GaNi	4	123	no	no	no	3.02

Total magnetization of the ferromagnetic solution, if it was found.
Apply filter:
Greater than
0.1
 AND OR
Equals
Filter...
Reset Apply

Page Size: 20 1 to 9 of 9 < > Page 1 of 1 > >

Download filtered entries ?

Materials Cloud 3D crystals database

<https://mc3d.materialscloud.org/>

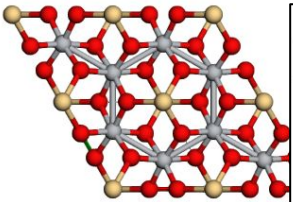
Materials Cloud three-dimensional crystals database (MC3D)

DOI: 10.24435/materialscloud:rw-t0

CdO₆Ti₂ (mc3d-10114/pbesol-v1)

General overview

Structure



Info
Formula: CdO₆Ti₂

70e7946c-ae04-42b5-a3c8-87628be2594c


Click to interact

Supercell: 2 2 2

Bonds Packed cell Atom labels

Structural details

General



Explore provenance 

Cell

	x [Å]	y [Å]
V ₁	5.1469	0.0000
V ₂	-2.5735	4.4574
V ₃	0.0000	0.0000

Provenance links

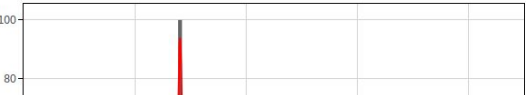
Relevant nodes in the provenance browser:

- Final optimization calculation 
- Final structure 

X-ray diffraction pattern

Select the X-ray source
CuKα

Select peak broadening profile
Gaussian



UID: 70e7946c-ae04-42b5-a3c8-87628be2594c
Type: aiida.calculations.quantumespresso.pw
Created on 24 September 2021
Modified 10 months ago
Creator: Mamik Berxc (EPFL)

CalcJobNode

JOB ID: 308095
SCHEDULER STATE: done
REMOTE WORKING DIRECTORY: /scratch/e1000/mbercx/aiida/70e7946c-ae04-42b5-a3c8-87628be2594c

INPUT FILES

- aiida.in
- _aiidasubmit.sh

OUTPUT FILES

- aiida.out
- _scheduler-stderr.txt
- _scheduler-stdout.txt
- data-file-schema.xml

Node metadata

ACCOUNT: mr0

APPEND_TEXT

CUSTOM_SCHEDULER_COMMANDS

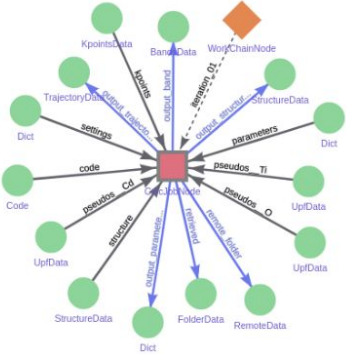
DETAILED_JOB_INFO

```
["retval":0,"stderr":"","stdout":"AlllocCPUS|Account|AssocID|AveCPU|AvePages|AveRSS|AveVMS|Cluster|Comment|CPUTime|CPUTimeRAW|DerivedExitCode|Elapsed|Eligible|End|ExitC  
09-24T01:36:37|2021-09-24T02:41:56|0|31304|mr0|308095|aiida-693144|||||||256|1|nid001265|4294656625|normal|1|128|01:03:50|5-16:10:40|490240|2021-09-  
24T02:40:27|COMPLETED|2021-09-  
24T01:36:37|00:00:00|29:44:40|12:00:00|03:00:53|26364|mbercx|02:31:09|n256|mr0|1793|00:00:00|0|18356K|446972K|eiger|06:19:44|22784||00:01:29|2021-09-24T02:40:27|2021-09-  
24T02:41:56|0|308095.batch|batch|18356K|nid001265|0|446972K|nid001265|0|00:00:00|nid001265|0|256|1|nid001265|1|||256|||2021-09-24T02:40:27|COMPLETED|2021-09-  
24T02:40:27|00:00:00|00:00:133|00:00:00:656||00:00:523|n256|mr0|1793|00:00:00|0|2748K|173004K|eiger|06:19:44|22784||00:01:29|2021-09-24T02:40:27|2021-09-  
24T02:41:56|0|308095.extern|extern|2748K|nid001265|0|173004K|nid001265|0|00:00:00|nid001265|0|256|1|nid001265|1|||256|||2021-09-24T02:40:27|COMPLETED|2021-09-  
24T02:40:27|00:00:00|00:00:00|00:00:00|00:00:00|n256|mr0|1793|00:01:24|0|132675008|2433303904|einer|06:02:40|21760|00:01:25|2021-09-24T02:40:31|2021-09-
```

Last Job Information

AiIDA Provenance Browser

Click on node to browse, drag to animate

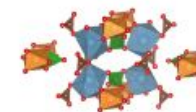


Materials Cloud 3D crystals database

Accessing data directly

Materials Cloud three-dimensional crystals database (MC3D)

DOI [10.24435/materialscloud:rw-t0](https://doi.org/10.24435/materialscloud:rw-t0)



Materials Cloud three-dimensional crystals database is a curated set of computationally relaxed three-dimensional crystal structures and calculated properties. The crystal structures originate from experimental databases.

Use

About

REST API

This section contains an overview of our REST APIs to access the MC3D data.

1. Materials Cloud and AiiDA REST APIs

The MC3D frontend is running on the following APIs:

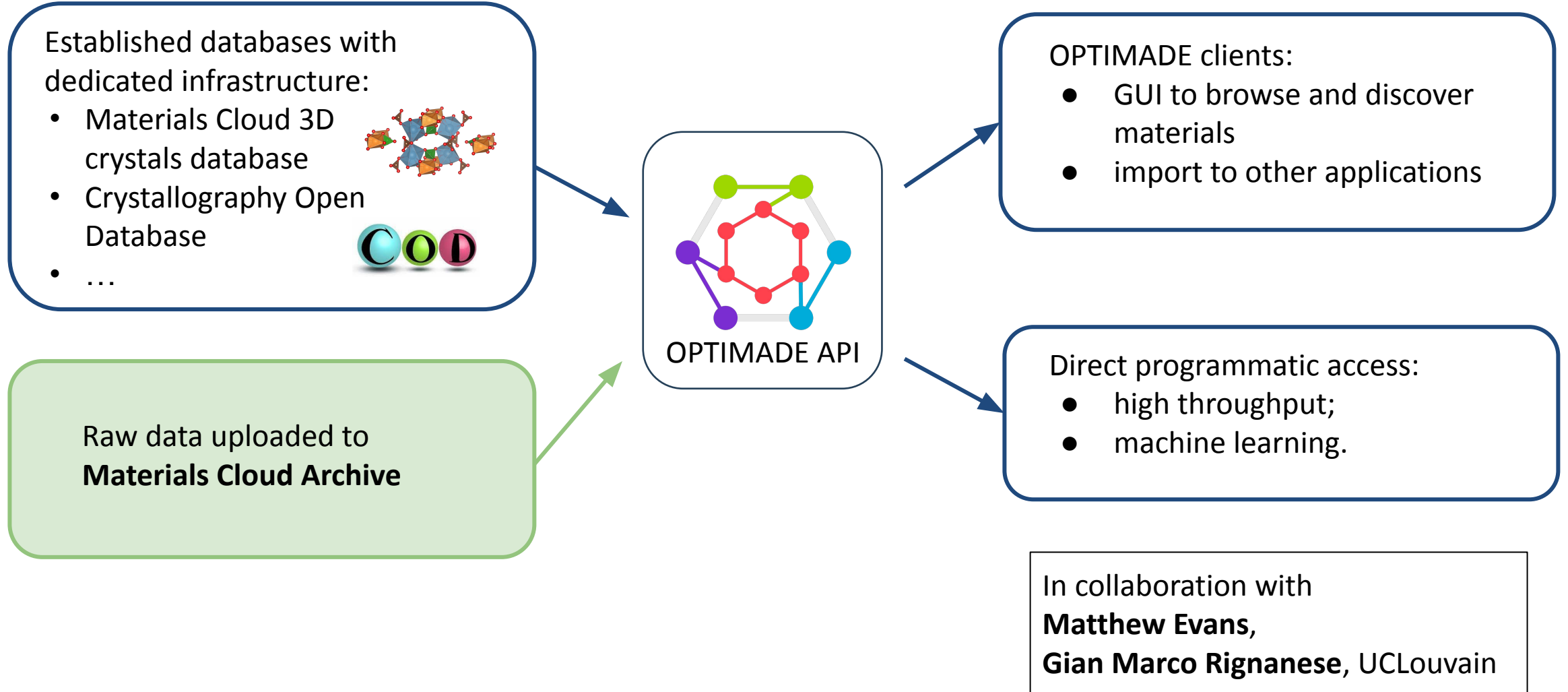
- Index of materials: <https://dev-aiida.materialscloud.org/mc-rest-api/mc3d/pbe-v1/entries>
- Single entry data: <https://dev-aiida.materialscloud.org/mc-rest-api/mc3d/pbe-v1/entries/mc3d-10>
- AiiDA REST API for properties and provenance: <https://dev-aiida.materialscloud.org/mc3d/api/v4>

2. OPTIMADE REST API

The MC3D database can also be accessed via an API following the [OPTIMADE specification](#). This currently only includes the crystal structures and no properties or provenance information is provided. Relevant endpoints are

- <https://aiida.materialscloud.org/mc3d/optimade/v1/info>
- <https://aiida.materialscloud.org/mc3d/optimade/v1/structures>

Archive-OPTIMADE service



Archive-OPTIMADE service

<https://archive.materialscloud.org>

Machine learning-accelerated discovery of A_2BC_2 ternary electrides with diverse anionic electron densities

Zhiqi Wang¹, Yutong Gong¹, Matthew L. Evans^{2*}, Yujing Yan¹, Shiyao Wang¹, Nanxi Miao¹, Ruiheng Zheng¹, Gian-Marco Rignanese^{1,2*}, Junjie Wang^{1*}

¹ State Key Laboratory of Solidification Processing, School of Materials Science and Engineering, Northwestern Polytechnical University, Xi'an, Shaanxi 710072, People's Republic of China.

² IMCN-MODL, Université catholique de Louvain, Chemin des Étoiles, 8, B-1348 Louvain-la-Neuve, Belgium.

DOI: [10.24435/materialscloud:c8-gy](https://doi.org/10.24435/materialscloud:c8-gy) [version v1]

Publication date: Nov 28, 2023

Files

File name	Size	Description
mp_comparison.json.gz	49.0 KiB	A list of pymatgen 'ComputedEntry' containing the results of atome2 'MPGGARelax' calculations to enable direct comparison with the Materials Project's convex hull (as of 20/11/2023).
data.csv	18.5 KiB	Computed ELF max and stability info for each structure
raw.tar.gz	328.2 MiB	Additional raw data, including bandstructures and phonon dispersion curves for every structure.
structures.tar.gz	40.4 KiB	CIFs of every structure considered, with the exact structures used for initial ELFCar calculations, as well as by the MP compatibility relaxation (in the corresponding sub-folders).
scripts.zip	2.9 KiB	A directory containing Python scripts used for the re-relaxation and stability calculations of the structures, alongside a requirements file with the dependencies required for repeating them, as well as the script required to create this archive from the raw data.
README.txt	1.0 KiB	README
optimade.yaml	1.5 KiB	A config file for the MCloud/OPTIMADE integration that allows ingestion of the data into an OPTIMADE API

```
1 entries:
2   - entry_type: structures
3     entry_paths:
4       - file: structures.tar.gz
5         matches:
6           - structures/mp_gga/*.cif
7     property_paths:
8       - file: data.csv
9     property_definitions:
10      - name: elf_max
11        title: Maximum value of the ELF
12        description: The maximum value of the electron
13        unit: dimensionless
14        type: float
```

<https://optimadeclient.materialscloud.io>

Query a provider's database

Materials Cloud Archive
Select a provider
Alexandria
International materials repository (CMR)
Materials Cloud
Materials Cloud Archive
The Materials Project
Material-Property-Descriptor Database
Material Properties Open Database (MPOD)
open database of state
Open Materials Database
2DMapedia
AFLOW
Crystallography Open Database
Materials Platform for Data Science
openmlake
novel materials discovery (NOMAD)
The Open Quantum Materials Database (OQMD)
Joint Automated Repository for Various Integrated Simulations (JARVIS)
Theoretical Crystallography Open Database

Materials Cloud Archive
Databases contributed by the community to the Materials Cloud Archive

MC Archive c8-gy
Z. Wang et al., Machine learning-accelerated discovery of A-BC₂ ternary electrides with diverse anionic electron densities, Materials Cloud Archive 2023.181 (2023) doi: 10.24435/materialscloud:c8-gy

Hide Periodic Table

Results
Ascending nskles
Showing 1-25 of 145 results
Dy4S4Y2 (Id:structures.tar.gz/structures/mp_c...)

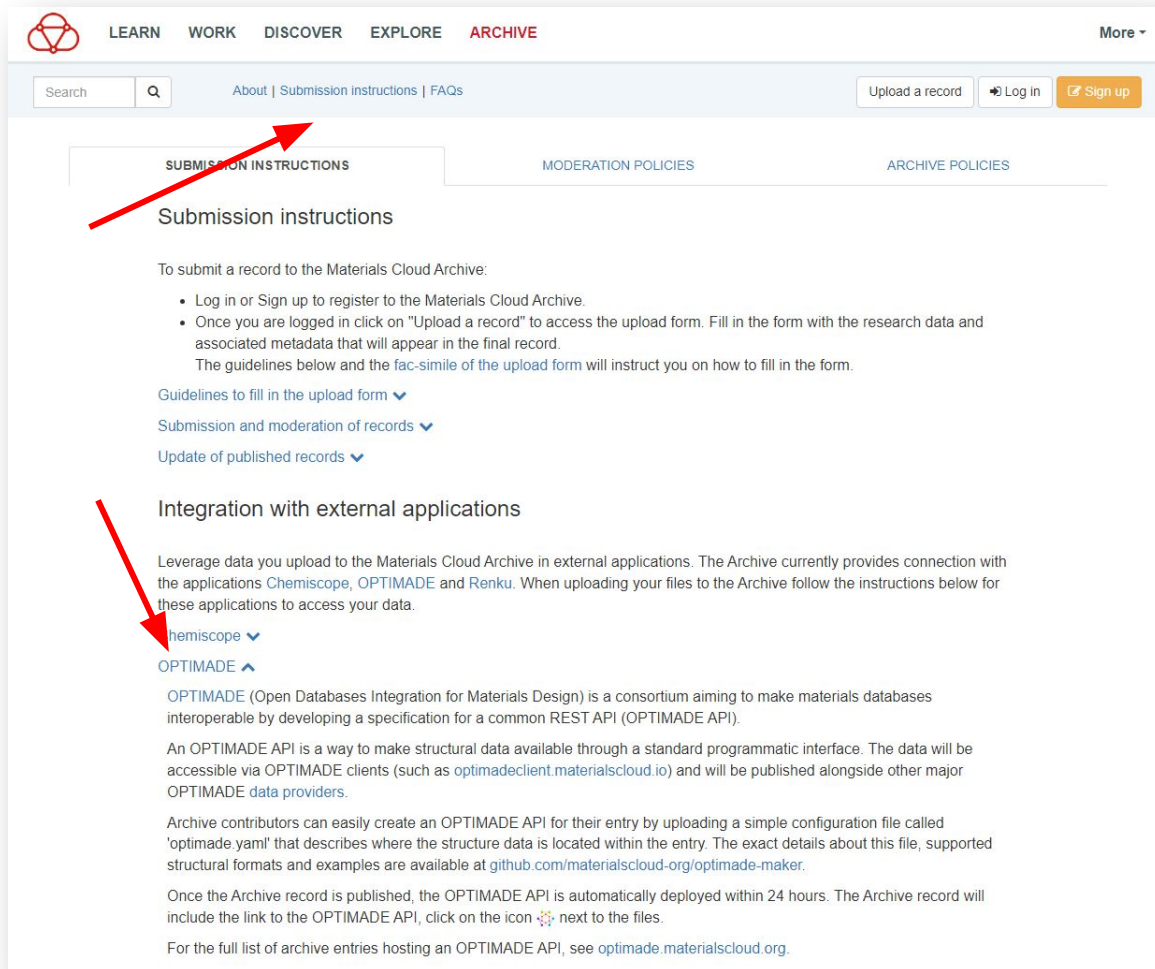
Crystallographic Information File v1.0 (via ASE) (.cif) Download

Use In QC Input Generator

Archive-OPTIMADE service

<https://archive.materialscloud.org/>

<https://optimade.materialscloud.org/>



The screenshot shows the Materials Cloud Archive website. The navigation bar includes 'LEARN', 'WORK', 'DISCOVER', 'EXPLORE', and 'ARCHIVE'. A search bar is present, along with links for 'About', 'Submission instructions', and 'FAQs'. There are buttons for 'Upload a record', 'Log in', and 'Sign up'. The 'SUBMISSION INSTRUCTIONS' section is highlighted with a red arrow. Below it, there are sections for 'Submission instructions' and 'Integration with external applications', both also highlighted with red arrows. The 'Integration with external applications' section lists 'chemiscope' and 'OPTIMADE' as options.

Materials Cloud Archive OPTIMADE server



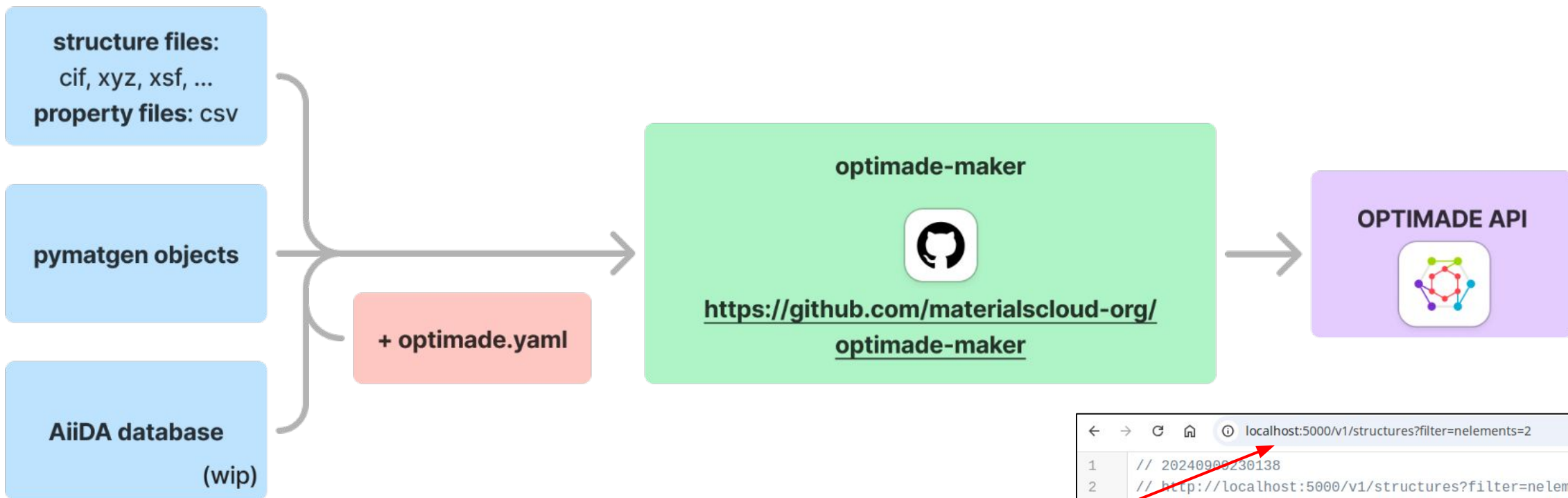
Index meta-database:

- </archive/index/v1/info>
- </archive/index/v1/links>

Available databases:

Date	Archive entry	OPTIMADE endpoint
2024.08.05	V. Trinet et al., <i>Optical materials discovery and design with federated databases and machine learning</i> , Materials Cloud Archive 2024.114 (2024), doi: 10.24435/materialscloud:5p-vq	/archive/5p-vq
2024.04.26	L. Kahle et al., <i>High-throughput computational screening for solid-state Li-ion conductors</i> , Materials Cloud Archive 2024.65 (2024), doi: 10.24435/materialscloud:vg-ya [version 2]	/archive/vg-ya
2023.11.28	Z. Wang et al., <i>Machine learning-accelerated discovery of A_2BC_2 ternary electrides with diverse anionic electron densities</i> , Materials Cloud Archive 2023.181 (2023), doi: 10.24435/materialscloud:c8-gy	/archive/c8-gy

optimade-maker



```
> ls
data.csv  optimade.yaml  structures.zip
> optimake serve
2024-09-09 22:50:24 - optimake - INFO - optimade.jsonl doesn't exist. Converting archive.
Parsing structures files: 100%|██████████████████████████████████████████████████| 2/2 [00:00<00:00, 77.36]
Constructing OPTIMADE structures entries: 2it [00:00, 731.29it/s]
Parsing properties for structures entries: 100%|██████████████████████████████████████████████████| 1/1 [00:00<00:00, 87.14]
INFO: Started server process [31378]
INFO: Waiting for application startup.
INFO: Application startup complete.
INFO: Uvicorn running on http://0.0.0.0:5000 (Press CTRL+C to quit)
```

```
localhost:5000/v1/structures?filter=nelements=2
1 // 20240909230138
2 // http://localhost:5000/v1/structures?filter=nelements=2
3
4 {
5   "data": [
6     {
7       "id": "set2/102",
8       "type": "structures",
9       "attributes": {
10        "immutable_id": "structures.zip/cifs/set2/102.cif",
11        "last_modified": "2024-09-09T22:50:24.651087",
12        "elements": [
13          "C",
14          "Sr"
15        ],
16        "nelements": 2,
```

Acknowledgements

People involved in AiiDA, AiiDALab, and Materials Cloud



Giovanni Pizzi (PSI)

Marnik Bercx (PSI)

Miki Bonacci (PSI)

Edan Bainglass (PSI)

Xing Wang (PSI)

Jusong Yu (PSI)

Ali Khosravi (PSI)

Alexander Goscinski (PSI)

Michail Minotakis (PSI)

Yuhao Jiang (PSI)

Timo Reents (PSI)

Nataliya Paulish (PSI)



Nicola Marzari (EPFL, PSI)



Sebastiaan P. Huber (Microsoft)



Chris Sewell (EPFL)



Francisco F. Ramirez (EPFL)



Kristjan Eimre (EPFL)



Daniel Hollas (U. Bristol)



Leopold Talirz (Schott)



Aliaksandr Yakutovich (Empa)



Carlo Pignedoli (Empa)



Valeria Granata (EPFL)



Joost VandeVondele (ETHZ,CSCS)



Thomas Schulthess (ETHZ,CSCS)



swissuniversities



Materials Science and Technology



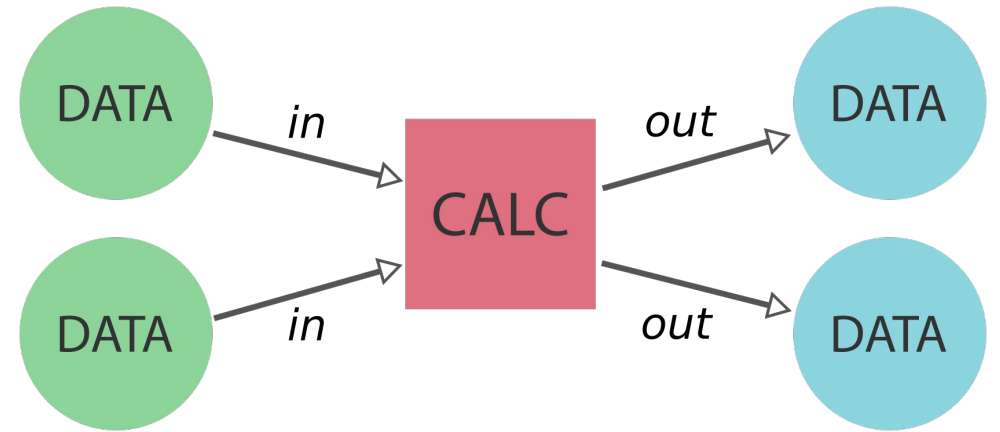
Summary

- Recent AiiDA improvements:
 - verdi presto - easy installation
 - aiida-shell - run executables without AiiDA plugins
 - WorkGraphs - easy combining of AiiDA processes
- Materials Cloud 3D crystals database
 - 'PBEsol-v1' subdatabase
 - Improved web interface
- Materials Cloud Archive - OPTIMADE integration
 - automatic OPTIMADE APIs for contributed raw data
 - optimade-maker

Data provenance

Simple recipe - store:

- data transformations/calculations
- inputs
- outputs
- inter-connections



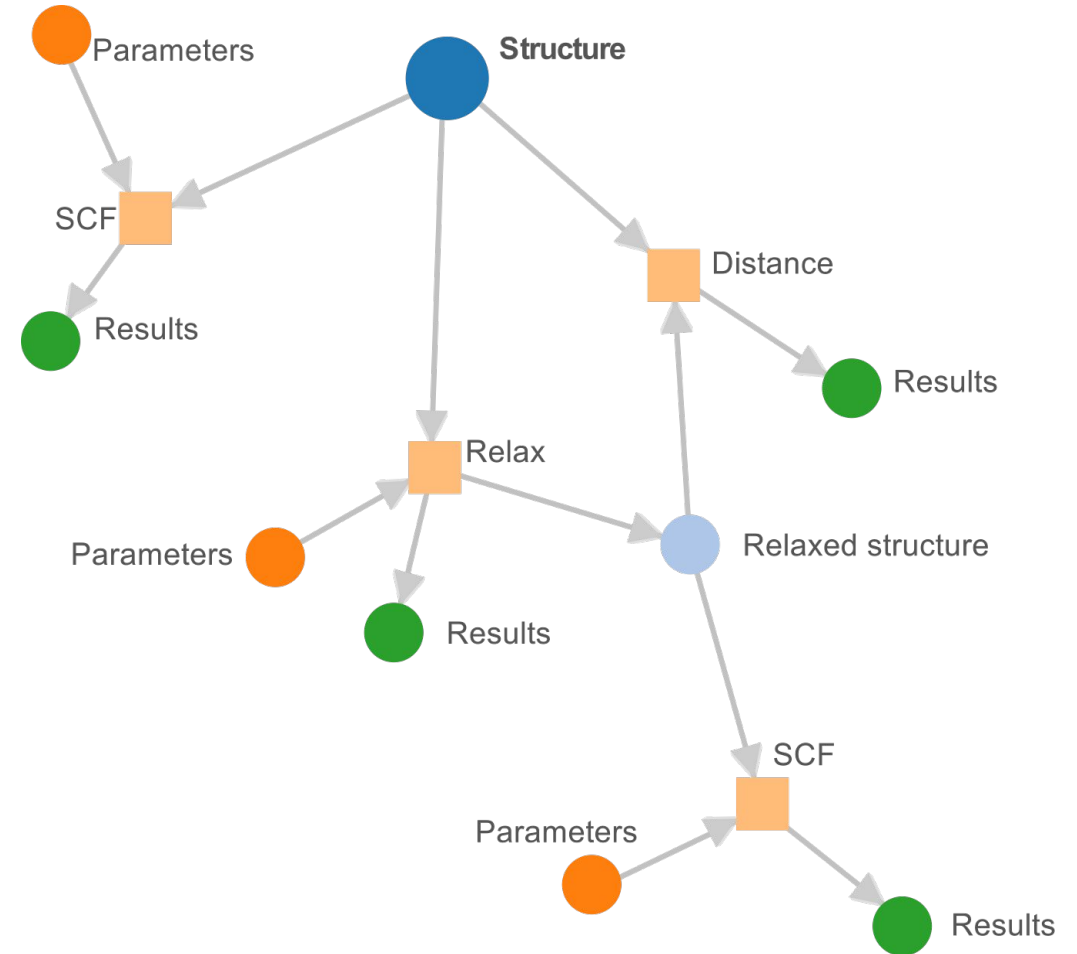
Data provenance

Simple recipe - store:

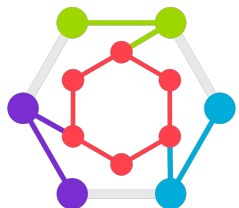
- data transformations/calculations
- inputs
- outputs
- inter-connections

Provenance graphs

- when data gets reused, a directed graph is created
- that quickly grows in complexity

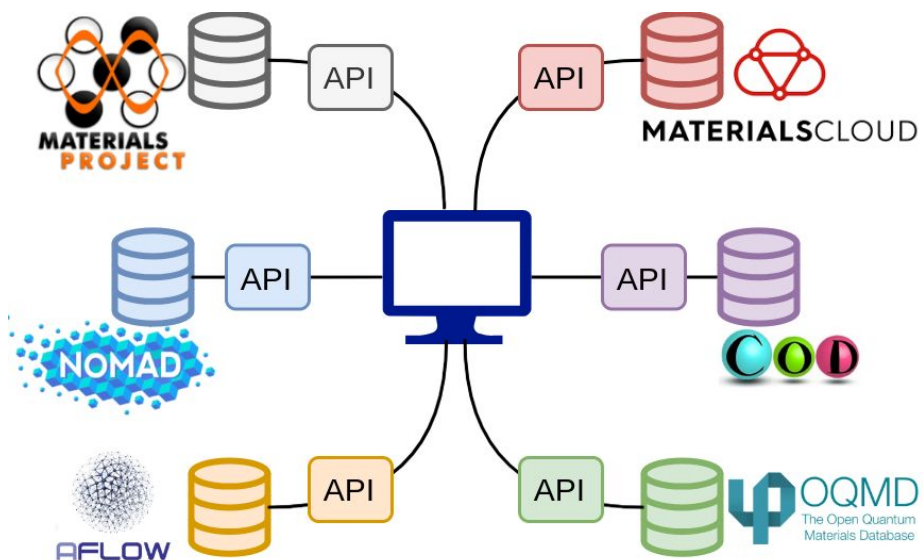


OPTIMADE

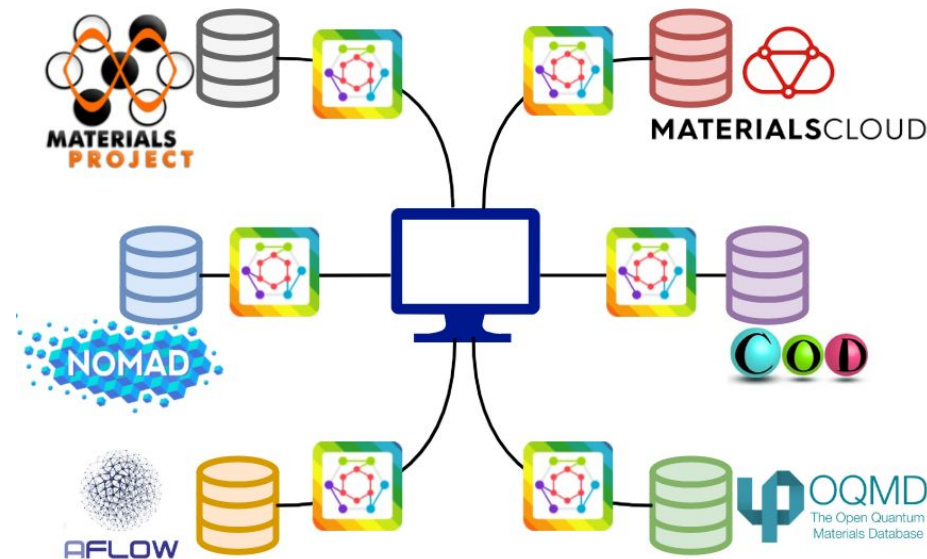


OPTIMADE
Open Databases Integration
for Materials Design

Consortium to make materials databases
interoperable via a **common REST API**, 15+ partners

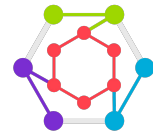


Andersen et al, Sci. Data 8, 217 (2021)



Evans et al, arXiv 10.48550/arXiv.2402.0057 (2024)

OPTIMADE

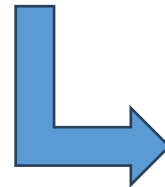


OPTIMADE API

base URL specific to the DB

universal OPTIMADE query

[https://aiida.materialscloud.org/mc3d/optimade/v1/structures?filter=elements HAS ALL "C", "Si"](https://aiida.materialscloud.org/mc3d/optimade/v1/structures?filter=elements HAS ALL 'C', 'Si')



machine-read
able json

Supports queries for/based on

- chemical composition;
- custom properties;
- references;
- & more, with new functionality constantly being added

```
{
  "data": [
    {
      "id": "417",
      "type": "structures",
      "links": null,
      "meta": null,
      "attributes": {
        "immutable_id": "6dd18556-a5e6-442d-a0f2-da0ae5beb716",
        "last_modified": "2022-07-01T21:18:58Z",
        "elements": [
          "C",
          "Si"
        ],
        "nelements": 2,
        "elements_ratios": [ ... ], // 2 items
        "chemical_formula_descriptive": "C9Si9",
        "chemical_formula_reduced": "CSi",
        "chemical_formula_hill": "C9Si9",
        "chemical_formula_anonymous": "AB",
      }
    }
  ]
}
```

OPTIMADE



OPTIMADE Clients

<https://optimadeclient.materialscloud.io/>

<https://www.materialscloud.org/work/tools/qeinputgenerator>

The image displays two overlapping web interfaces. The background interface is the OPTIMADE client, showing search results for the chemical formula Co8Ge12Li12O48 (id=mp-1013807). It includes a search bar, a periodic table with selected elements (Co, Ge, Li, O), and a 3D ball-and-stick model of the crystal structure. The foreground interface is the Quantum ESPRESSO input generator, which provides instructions on how to use the crystal structure data to generate input files for Quantum ESPRESSO calculations. It includes a 'Generate the P' button and a 3D visualization of the crystal structure with a 'Supercell' control.

Quantum ESPRESSO input generator

- About the Quantum ESPRESSO input generator and structure
- Instructions
- Acknowledgements

The crystal structure has been successfully uploaded. Adapt the parameters below and then press the "Generate" button.

Select here the pseudopotential library:
Select here the magnetism/smearing:^[?]
Select here the k-points distance ($1/\text{\AA}$)^[NOTE]
(and smearing (eV) in case of fractional occupations):

By continuing, you agree with the terms of the license.

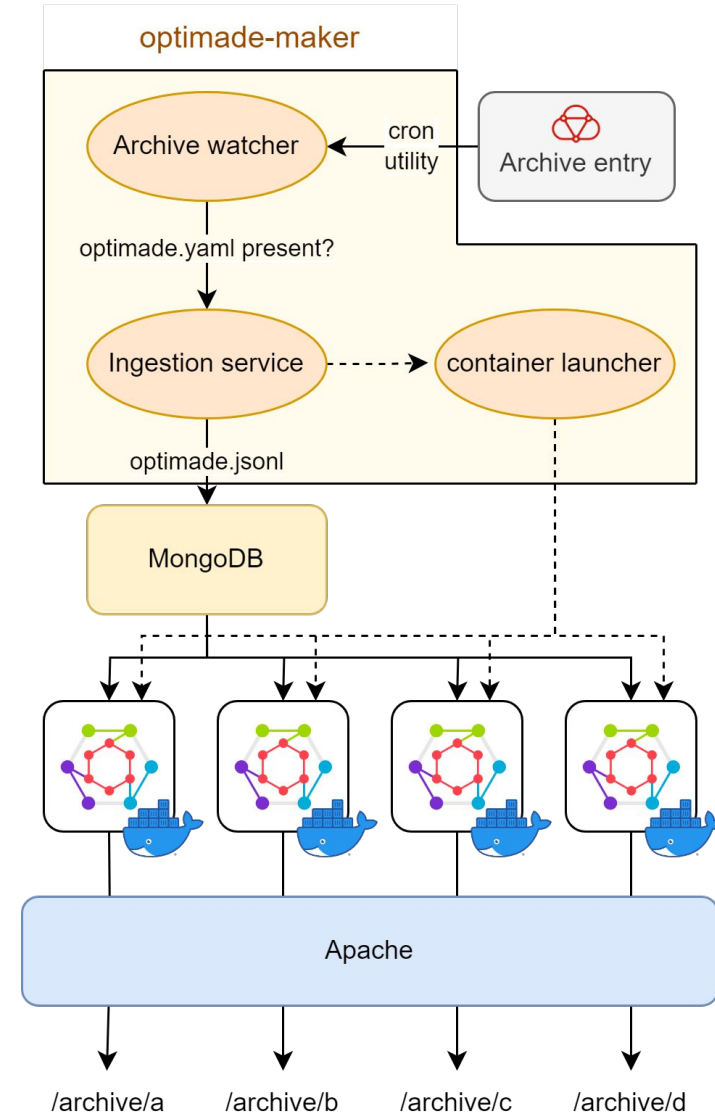
Generate the P

Supercell: 2 2 2 UPDATE RESET 2x2x2 CELL

Archive-OPTIMADE service

Implementation

- Archive watcher detects optimade.yaml
- Pipeline to populate a MongoDB
- Docker containers for each API
 - OPTIMADE Python tools
- Apache as reverse proxy.
- Tools open source: optimade-maker
 - <https://github.com/materialscloud-org/optimade-maker>





Materials Cloud OPTIMADE servers

<https://www.materialscloud.org/optimade>

LEARN WORK DISCOVER EXPLORE ARCHIVE More ▾

Home > OPTIMADE APIs

Materials Cloud OPTIMADE APIs

MAIN

Provider name: mcloud

Databases corresponding to the Materials Cloud Discover and Explore sections.

Index meta-database:
</optimade/main/v1/info>
</optimade/main/v1/links>

ARCHIVE

Provider name: mcloudarchive

Databases served automatically from the entries of the Materials Cloud Archive.

Index meta-database:
</optimade/archive/v1/info>
</optimade/archive/v1/links>

[Go to a table of the entries!](#)