



aiida.net

Computational Materials Science in the High-Throughput Era

with **AiiDA** and the **Materials Cloud**

Leopold Talirz, Aliaksandr V. Yakutovich, Daniele Ongari

Today's schedule

9:00-10:30

Introductory lecture

10:30-11:00

Coffee break

11:00-12:00

Getting everybody set up
Group A: Room X | Group B: Room Y

12:00-13:00

Lunch break

13:00-17:00

Tutorial & exercises
Group A: Room X | Group B: Room Z

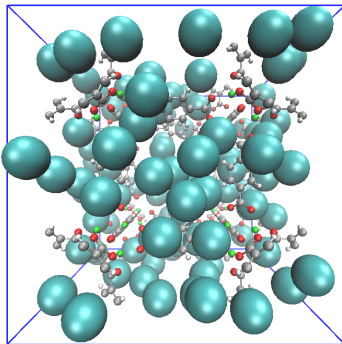
Outline



① Motivation, Architecture



②



③ Topic of today's tutorial

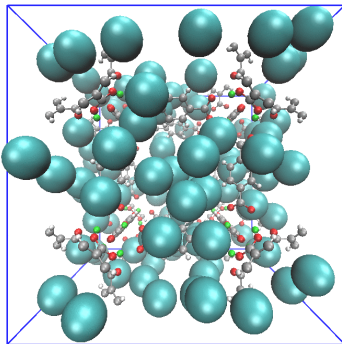
Outline



① Motivation, Architecture



②



③ Topic of today's tutorial

Motivation

Computational Materials Science Challenges

High-Throughput

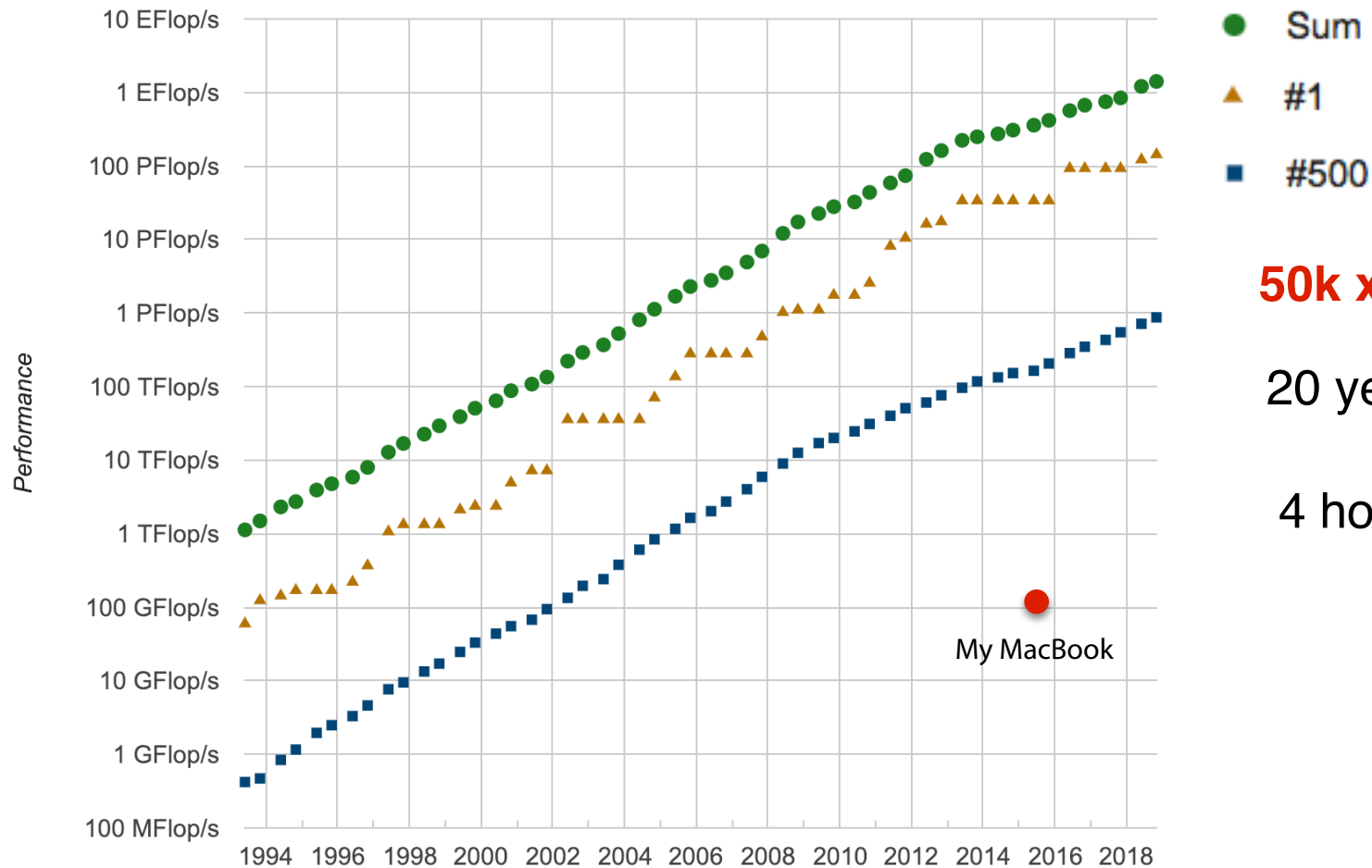
Reproducibility

Open Science

Knowledge Transfer

Challenge 1 – High Throughput

Top 500 Supercomputer Performance



50k x / 20 years

20 years (1998)

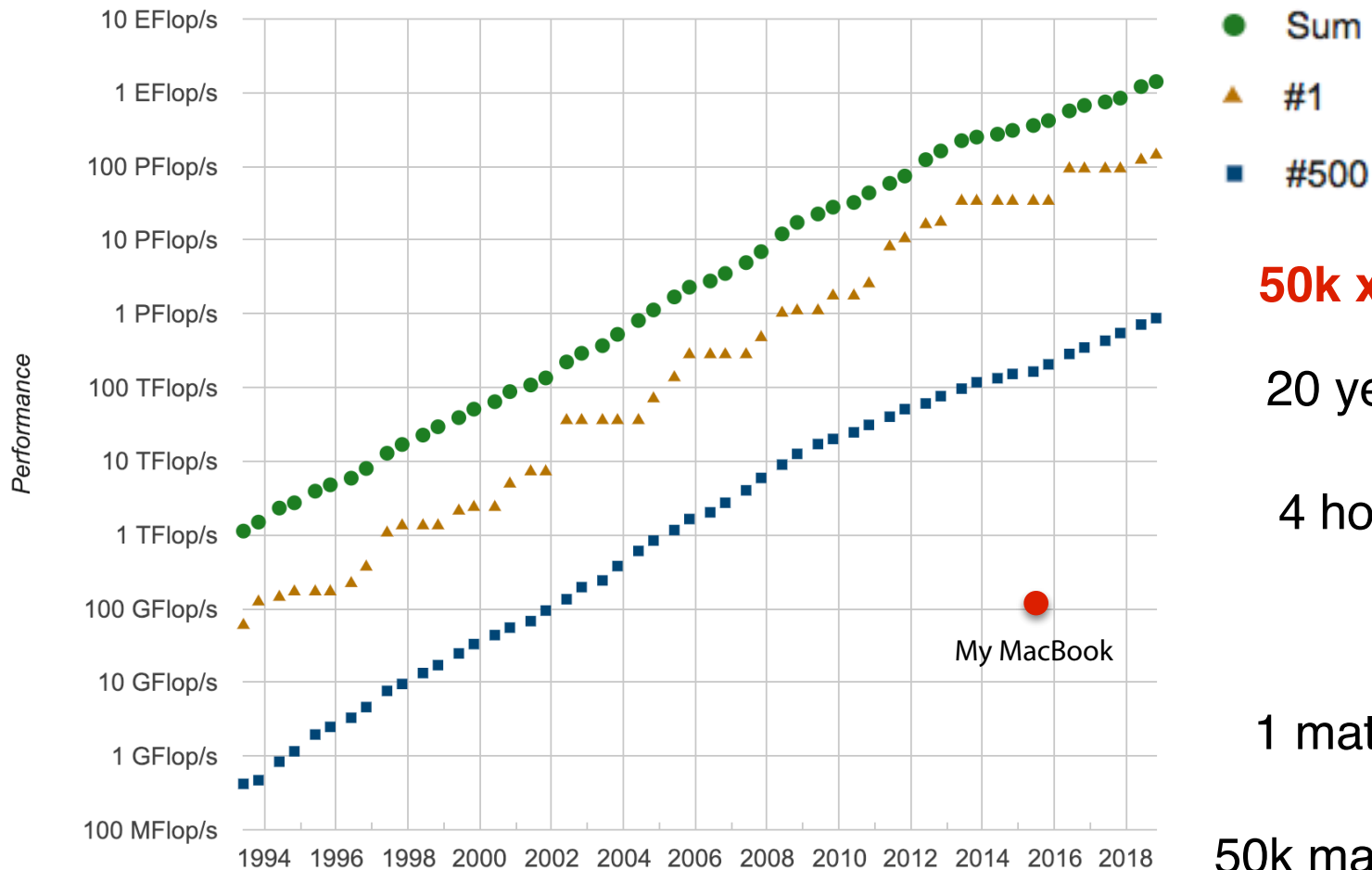


4 hours (2018)

www.top500.org/statistics/perfdevel

Challenge 1 – High Throughput

Top 500 Supercomputer Performance



www.top500.org/statistics/perfdevel

50k x / 20 years

20 years (1998)



4 hours (2018)

OR

1 material (1998)



50k materials (2018)

Motivation

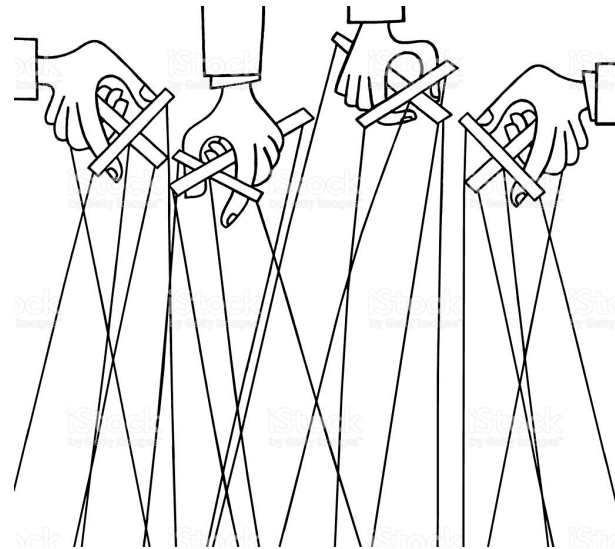
High-Throughput

Reproducibility

Open Science

Knowledge Transfer

- Organize large numbers of calculations
- Deal with corner cases (theory, code, infrastructure)
- Many strings to pull



Source: [istockphoto.com](https://www.istockphoto.com)

Motivation

High-Throughput

Reproducibility

Open Science

Knowledge Transfer

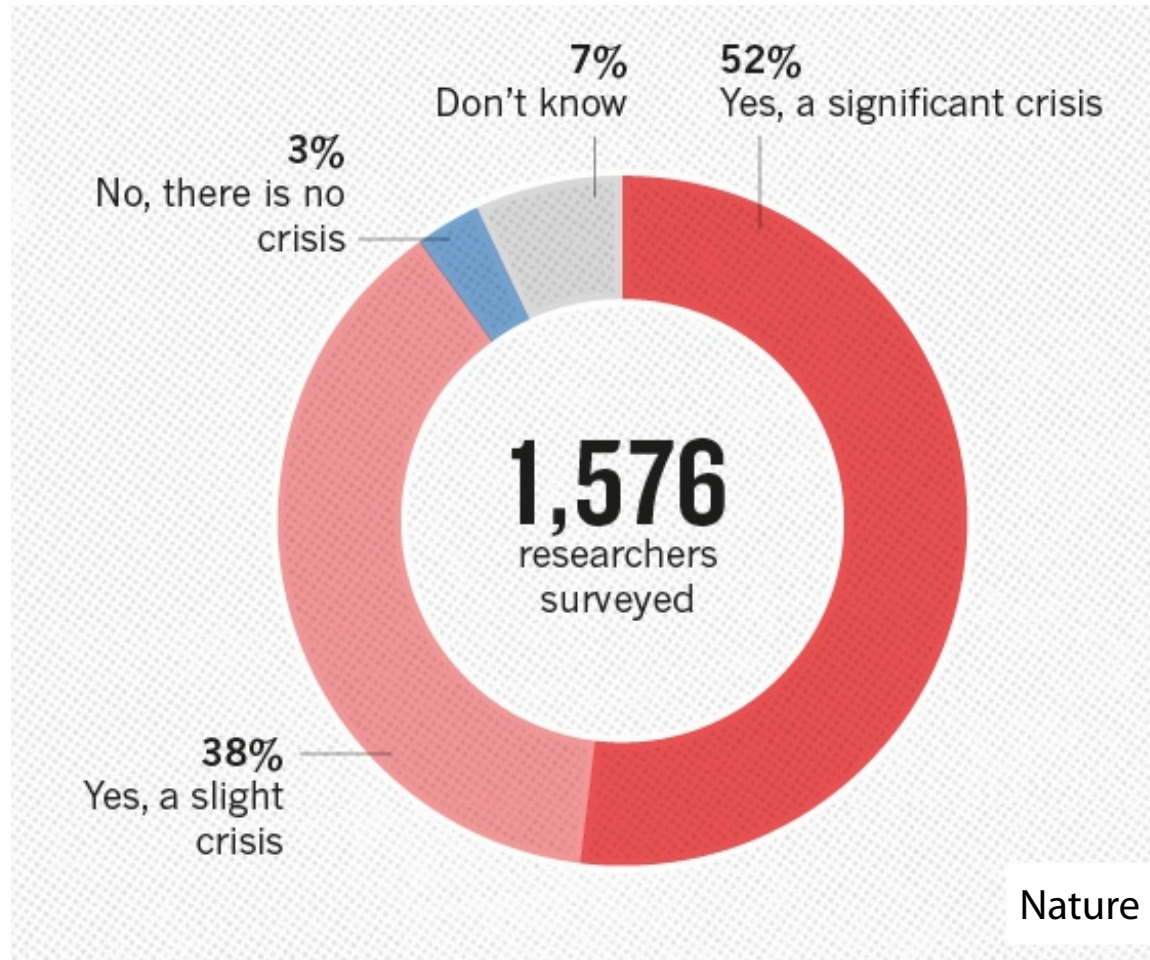
- Keep track of what you calculate
- Keep track of how you did it
- Within a research group:
Can Alice reproduce what Bob computed 1 year ago?



Source: academiccoachingandwriting.org

Challenge 2 – Reproducibility

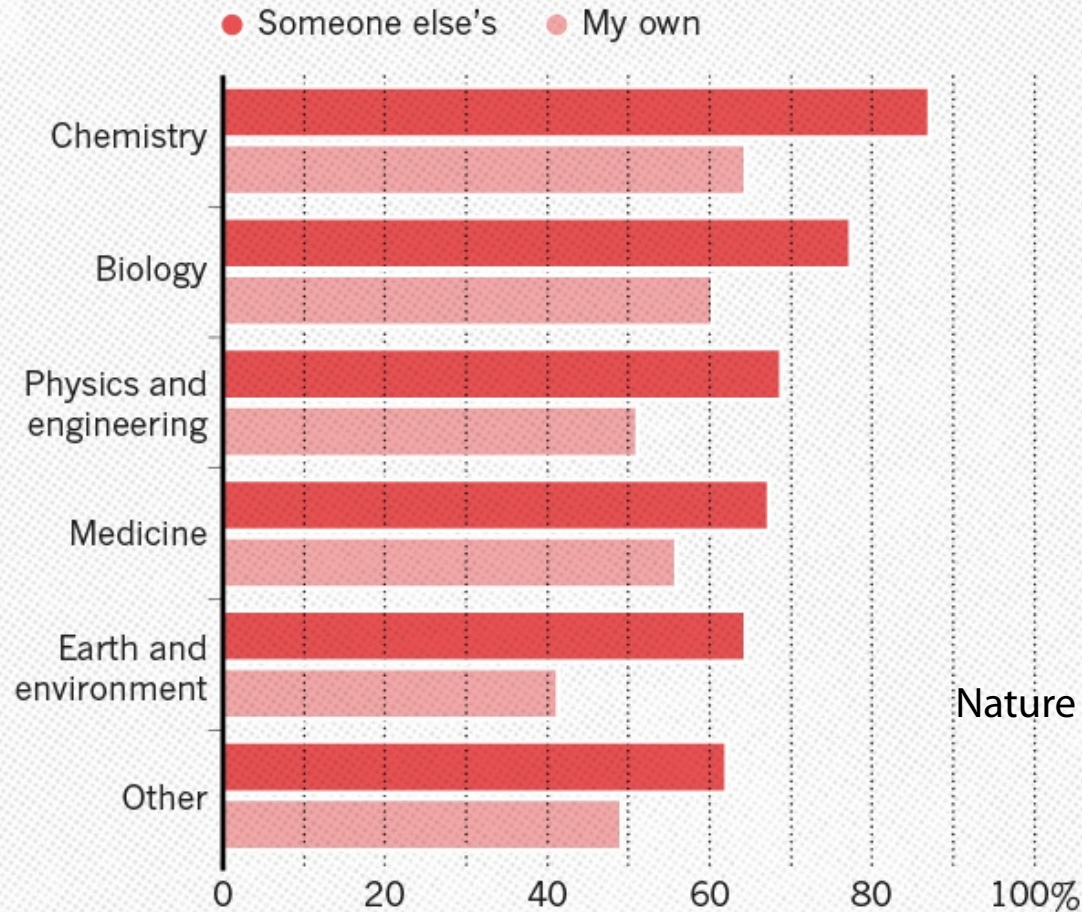
IS THERE A REPRODUCIBILITY CRISIS?



Challenge 2 – Reproducibility

HAVE YOU FAILED TO REPRODUCE AN EXPERIMENT?

Most scientists have experienced failure to reproduce results.



Nature **533**, 452–454 (2016)

Challenge 2 – Reproducibility

HAVE YOU FAILED TO REPRODUCE AN EXPERIMENT?

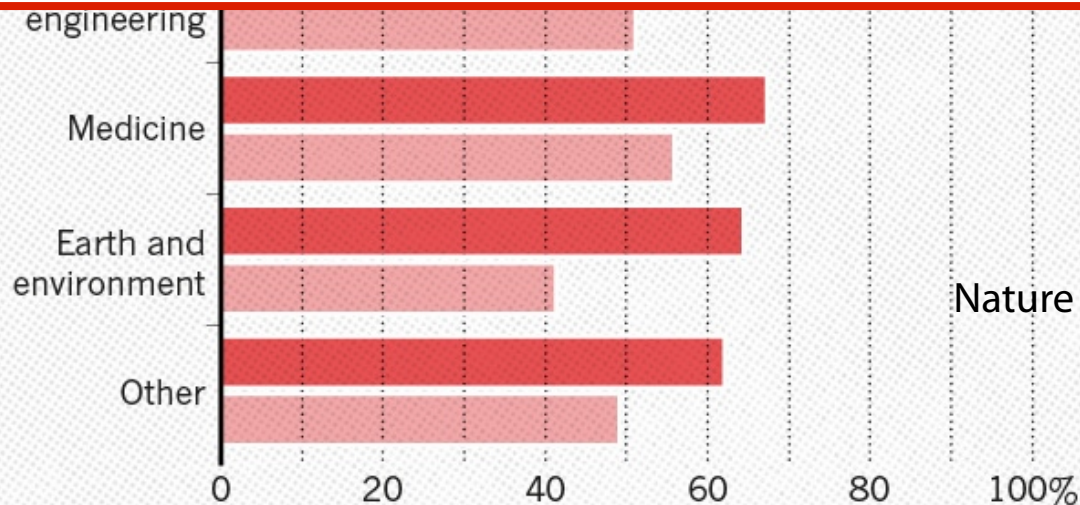
Most scientists have experienced failure to reproduce results.

● Someone else's ● My own



No excuses in computational science

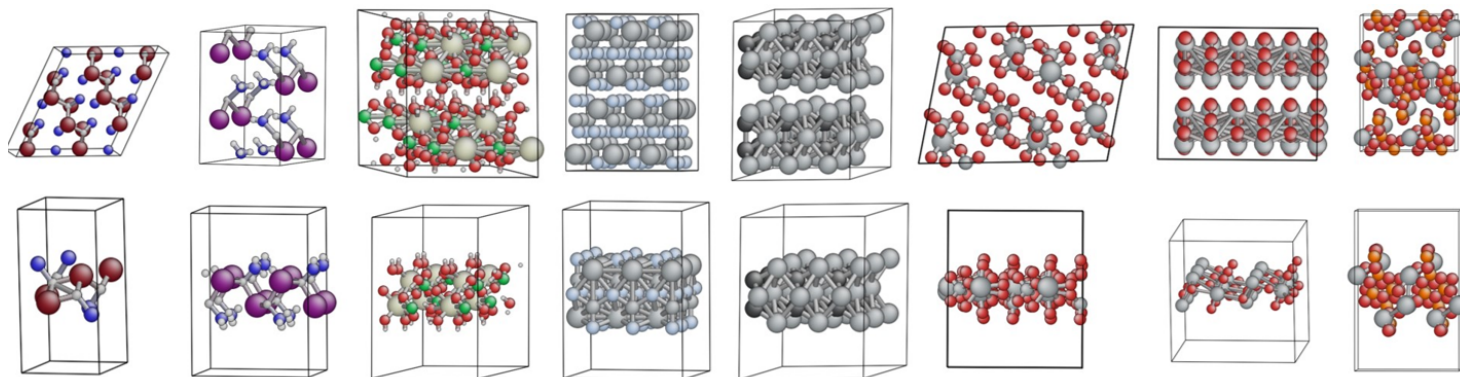
We can and **must** be fully reproducible



Nature **533**, 452–454 (2016)

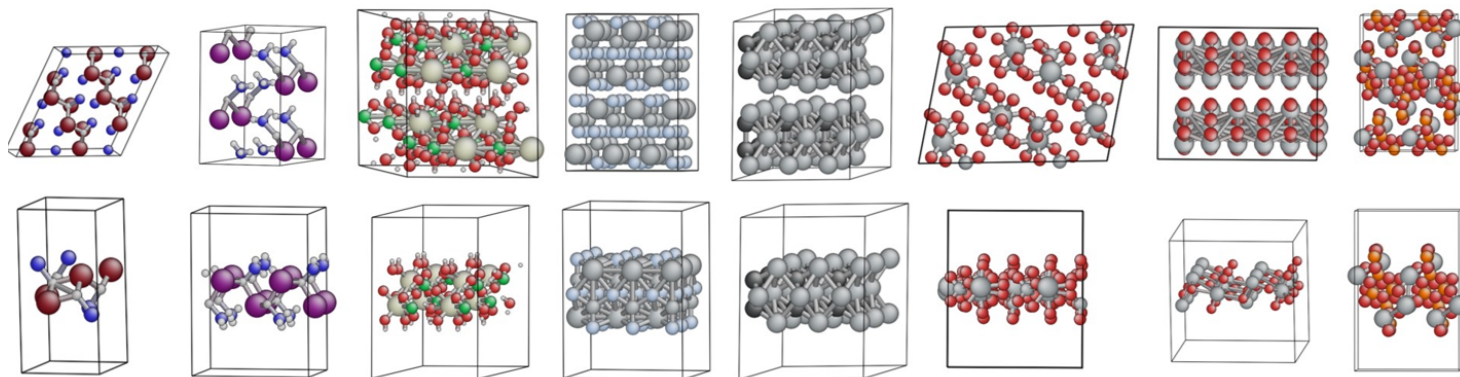
High-throughput Example

DISCOVERING NEW TWO-DIMENSIONAL MATERIALS



High-throughput Example

DISCOVERING NEW TWO-DIMENSIONAL MATERIALS

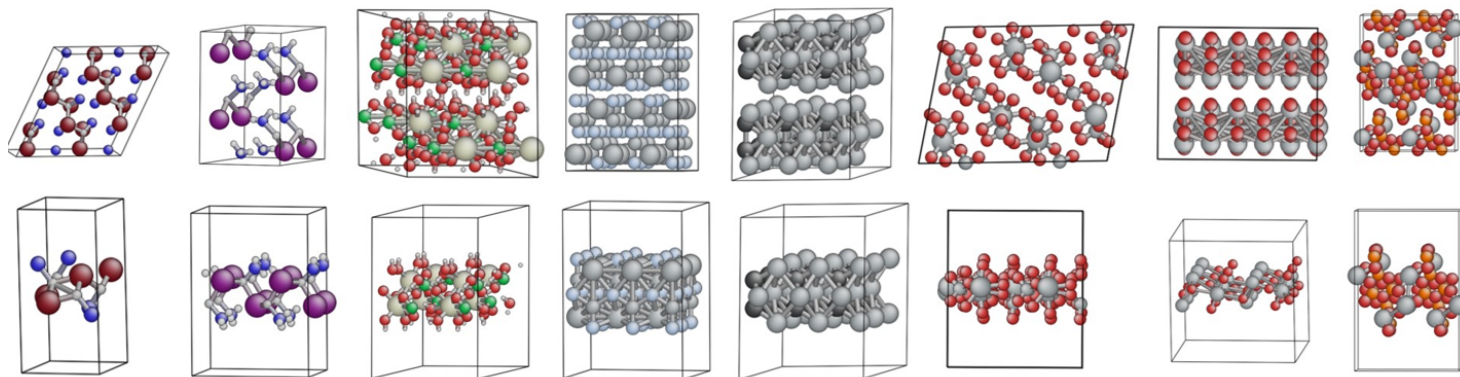


STARTING FROM ICSD/COD DATABASE:

- **108 423** unique 3D structures
- **5619** layered structures
- **>100 000** DFT calculations
- **>30 000** material properties
- **>1·10⁹** attributes

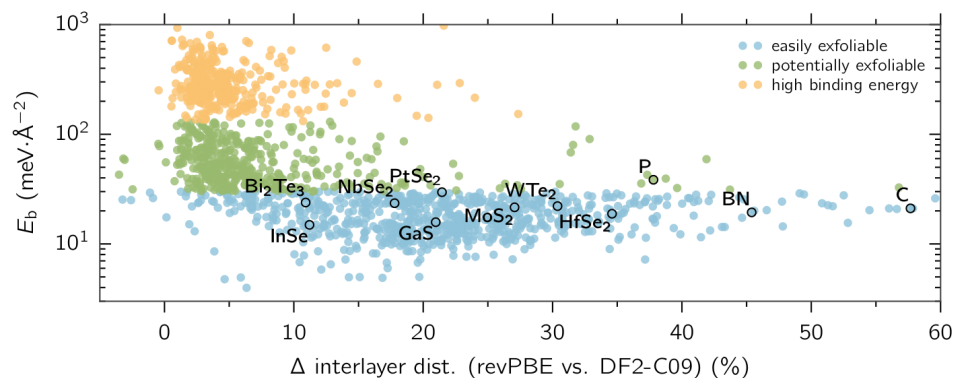
High-throughput Example

DISCOVERING NEW TWO-DIMENSIONAL MATERIALS



STARTING FROM ICSD/COD DATABASE:

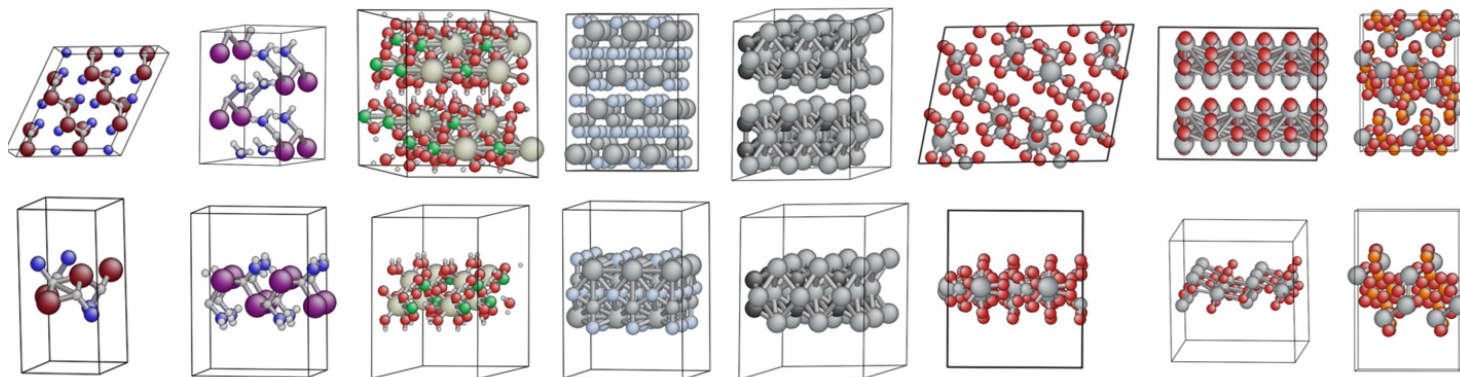
- **108 423** unique 3D structures
- **5619** layered structures
- **>100 000** DFT calculations
- **>30 000** material properties
- **$>1 \cdot 10^9$** attributes



Data needs to be condensed in a few plots

High-throughput Example

DISCOVERING NEW TWO-DIMENSIONAL MATERIALS



STARTING FROM ICSD/COD DATABASE:

- **108 423** unique 3D structures
- **5619** layered structures
- **>100 000** DFT calculations
- **>30 000** material properties
- **>1·10⁹** attributes

III. METHODS

A. Reproducibility and provenance

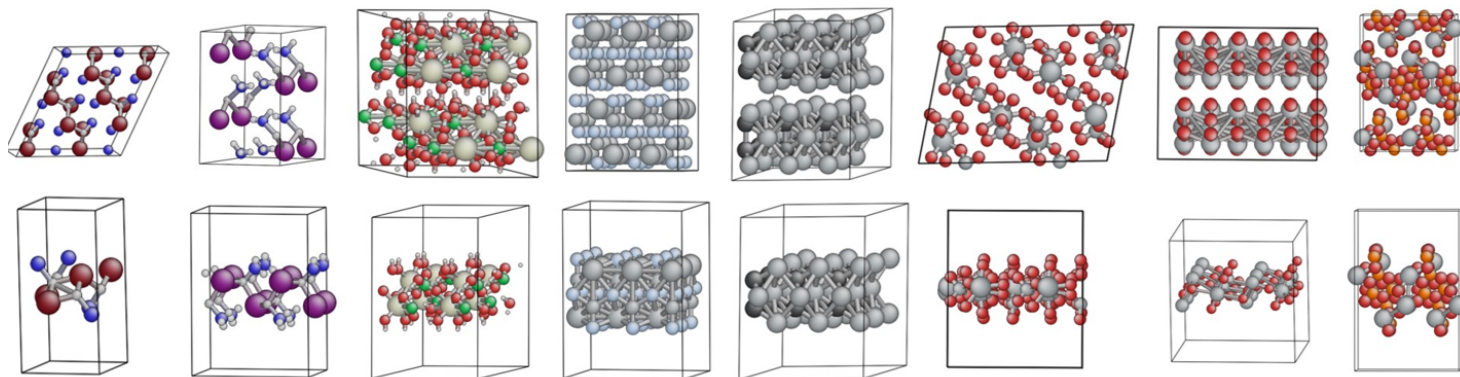
It is an often repeated tenet that results of scientific research must be reproducible. This objective is, however, challenging, especially in high-throughput research, due to the large number of simulations involved and the complex sequence of logical steps needed in the study. To ensure reproducibility we use AiiDA³⁰ as a materials' informatics infrastructure to implement the ADES model of automation, data, environment, and sharing, as dis-

As already mentioned in the main text, in this step we consider only structures coming from experimental measurements, while the source databases are partially populated by purely theoretical structures. This is done using the flags set on the database entries by their curators. We also implemented some heuristics to detect clearly wrong CIF files. For example, we discard structures where the chemical formula provided in the file is inconsistent with the elements in the unit cell. Regardless of all these efforts, it is possible that some incomplete or incorrect structures are still not filtered out from the original databases.

Methods: Impossible to describe every detail

High-throughput Example

DISCOVERING NEW TWO-DIMENSIONAL MATERIALS



STARTING FROM ICSD/COD DATABASE:

- **108 423** unique 3D structures
- **5619** layered structures
- **>100 000** DFT calculations
- **>30 000** material properties
- **>1·10⁹** attributes

III. METHODS

A. Reproducibility and provenance

It is an often repeated tenet that results of scientific research must be reproducible. This objective is, however, challenging, especially in high-throughput research, due to the large number of simulations involved and the complex sequence of logical steps needed in the study. To ensure reproducibility we use AiiDA³⁰ as a materials' informatics infrastructure to implement the ADES model of automation, data, environment, and sharing, as dis-

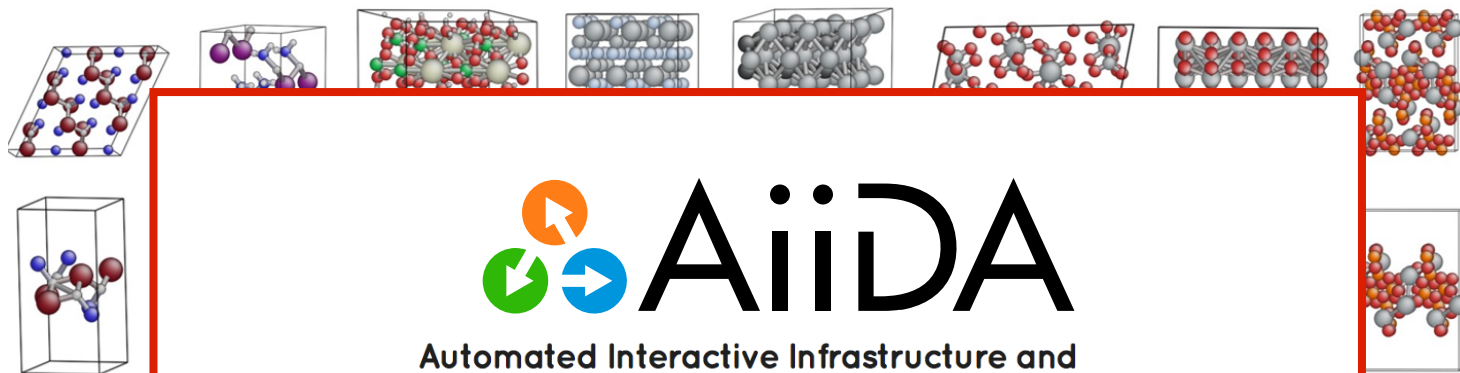
As already mentioned in the main text, in this step we consider only structures coming from experimental measurements, while the source databases are partially populated by purely theoretical structures. This is done using the flags set on the database entries by their curators. We also implemented some heuristics to detect clearly wrong CIF files. For example, we discard structures where the chemical formula provided in the file is inconsistent with the elements in the unit cell. Regardless of all these efforts, it is possible that some incomplete or incorrect structures are still not filtered out from the original databases.

Methods: Impossible to describe every detail

For **authors**, reproducing all data is challenging.
For **peers**, reproducing all data is almost impossible.

High-throughput Example

DISCOVERING NEW TWO-DIMENSIONAL MATERIALS



Automated Interactive Infrastructure and
Database for Computational Science

STARTING FROM

- **108 423** unique structures
- **5619** layered materials
- **>100 000** DFT calculations
- **>30 000** materials
- **>1·10⁹** attributes

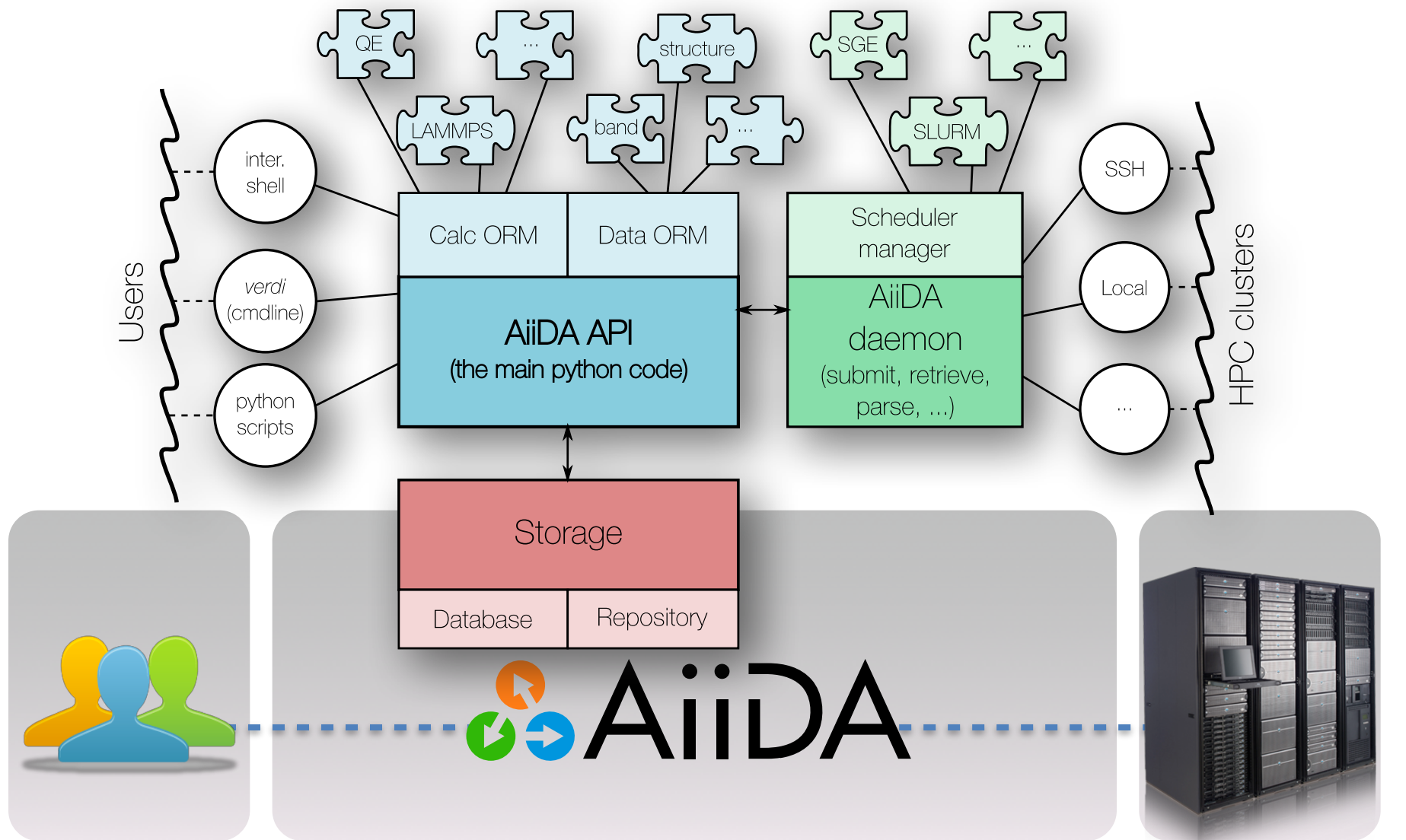
- Computational science **platform**
- for **high-throughput** calculations
- with **automatic data provenance**

ed in the main text, in this step
ctures coming from experimental
he source databases are partially
eoretical structures. This is done
the database entries by their cu-
mented some heuristics to detect
s. For example, we discard struc-
tural formula provided in the file is
lements in the unit cell. Regard-
it is possible that some incomplete
are still not filtered out from the

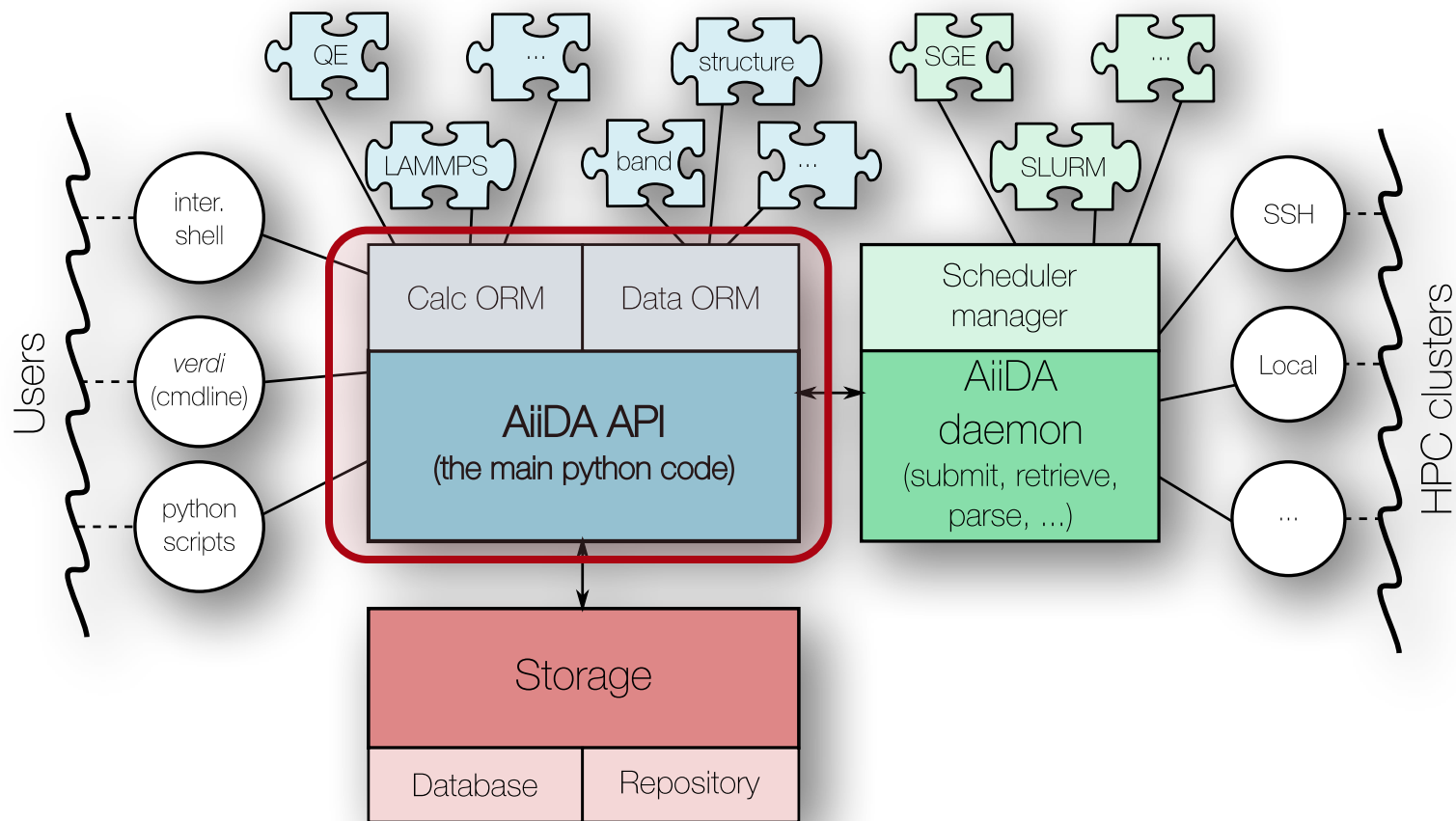
Methods: Impossible to describe every detail

For **authors**, reproducing all data is challenging.
For **peers**, reproducing all data is almost impossible.

AiiDA architecture



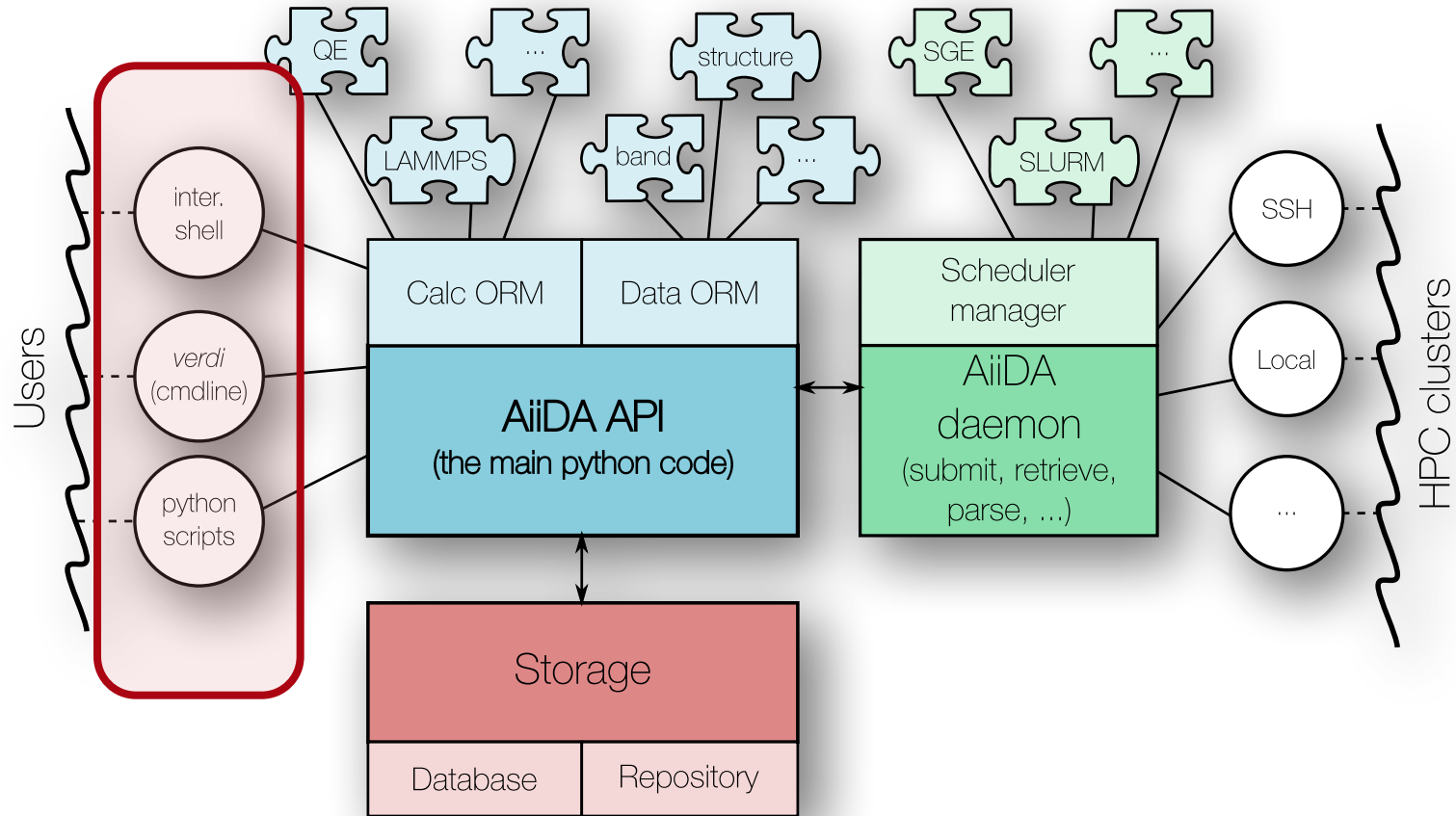
AiiDA architecture



1.

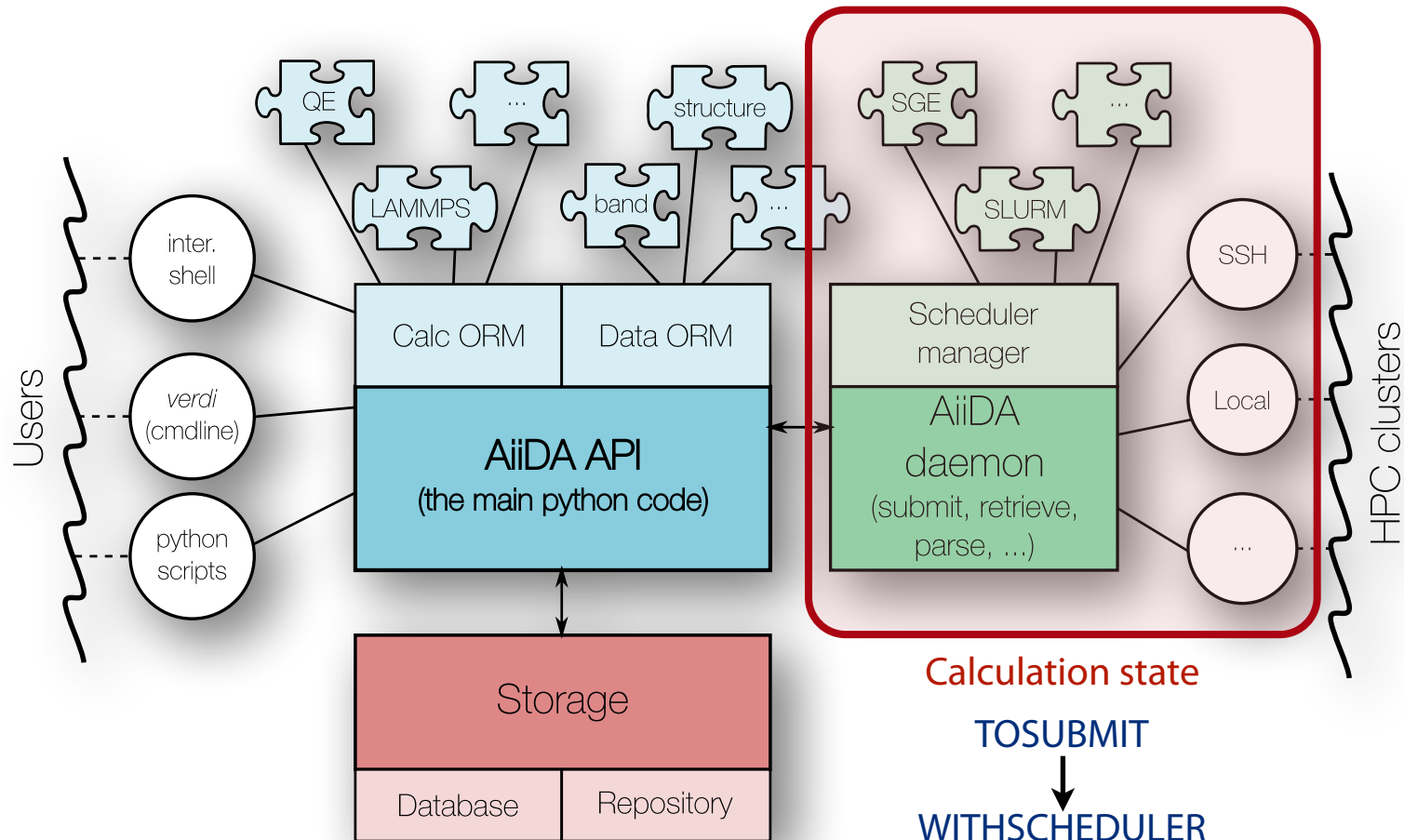
The core: **AiiDA python API**

AiiDA architecture



2. **User interface:**
python scripts, verdi command line tool, verdi shell

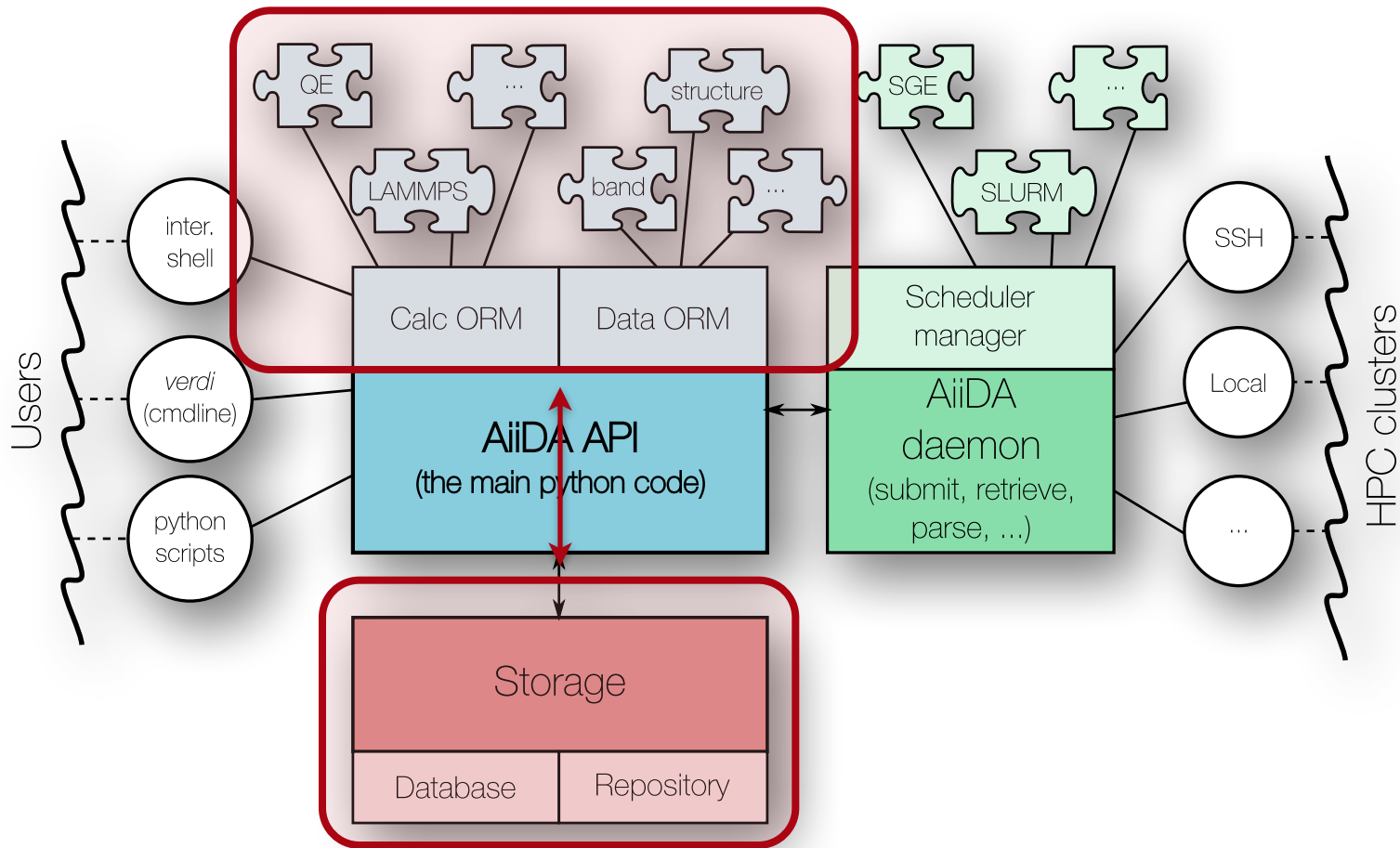
AiiDA architecture



3. AiiDA **daemon**: manage interaction with remote computers without user intervention

Calculation state
TOSUBMIT
↓
WITHSCHEDULER
↓
RETRIEVED
↓
PARSED
↓
FINISHED

AiiDA architecture



4. **AiiDA Object-Relational Mapper (ORM):**
stores data, codes and calculations in local database

AiiDA: Calculation example

```
code = Code.get_from_string('pw-6.3@daint-mr25')
calc = code.new_calc()

calc.set_max_wallclock_seconds(600)
calc.set_resources({"num_machines": 2})

Structure = DataFactory('structure')
structure = Structure(ase = read('TiO2.cif'))

Parameter = DataFactory('parameter')
parameters = Parameter({
    'CONTROL': {
        'calculation': 'scf',
        'restart_mode': 'from_scratch',
    },
    'SYSTEM': {
        'ecutwfc': 40.,
    }
})

Kpoints = DataFactory('array.kpoints')
kpoints = Kpoints(kpoints_mesh = [4,4,4])

calc.use_structure(structure)
calc.use_parameters(parameters)
calc.use_kpoints(kpoints)
calc.use_pseudos_from_family('SSSP_efficiency_v1.0')
calc.store_all()

calc.submit()
```

AiiDA: Calculation example

```
code = Code.get_from_string('pw-6.3@daint-mr25')
calc = code.new_calc()
```

```
calc.set_max_wallclock_seconds(600)
calc.set_resources({"num_machines": 2})
```

```
Structure = DataFactory('structure')
structure = Structure(ase = read('TiO2.cif'))
```

```
Parameter = DataFactory('parameter')
parameters = Parameter({
    'CONTROL': {
        'calculation': 'scf',
        'restart_mode': 'from_scratch',
    },
    'SYSTEM': {
        'ecutwfc': 40.,
    }
})
```

```
Kpoints = DataFactory('array.kpoints')
kpoints = Kpoints(kpoints_mesh = [4,4,4])
```

```
calc.use_structure(structure)
calc.use_parameters(parameters)
calc.use_kpoints(kpoints)
calc.use_pseudos_from_family('SSSP_efficiency_v1.0')
calc.store_all()
```

```
calc.submit()
```

Switch computers in one line
supports different schedulers,
version of codes, ...

AiiDA: Calculation example

```
code = Code.get_from_string('pw-6.3@daint-mr25')
calc = code.new_calc()
```

```
calc.set_max_wallclock_seconds(600)
calc.set_resources({"num_machines": 2})
```

```
Structure = DataFactory('structure')
structure = Structure(ase = read('TiO2.cif'))
```

```
Parameter = DataFactory('parameter')
parameters = Parameter({
    'CONTROL': {
        'calculation': 'scf',
        'restart_mode': 'from_scratch',
    },
    'SYSTEM': {
        'ecutwfc': 40.,
    }
})
```

```
Kpoints = DataFactory('array.kpoints')
kpoints = Kpoints(kpoints_mesh = [4,4,4])
```

```
calc.use_structure(structure)
calc.use_parameters(parameters)
calc.use_kpoints(kpoints)
calc.use_pseudos_from_family('SSSP_efficiency_v1.0')
calc.store_all()
```

```
calc.submit()
```

Switch computers in one line

supports different schedulers,
version of codes, ...

Define (only) necessary inputs

Interface designed by plugin

AiiDA: Calculation example

```
code = Code.get_from_string('pw-6.3@daint-mr25')
calc = code.new_calc()
```

```
calc.set_max_wallclock_seconds(600)
calc.set_resources({"num_machines": 2})
```

```
Structure = DataFactory('structure')
structure = Structure(ase = read('TiO2.cif'))
```

```
Parameter = DataFactory('parameter')
parameters = Parameter({
    'CONTROL': {
        'calculation': 'scf',
        'restart_mode': 'from_scratch',
    },
    'SYSTEM': {
        'ecutwfc': 40.,
    }
})
```

```
Kpoints = DataFactory('array.kpoints')
kpoints = Kpoints(kpoints_mesh = [4,4,4])
```

```
calc.use_structure(structure)
calc.use_parameters(parameters)
calc.use_kpoints(kpoints)
calc.use_pseudos_from_family('SSSP_efficiency_v1.0')
calc.store_all()
```

```
calc.submit()
```

Switch computers in one line
supports different schedulers,
version of codes, ...

Define (only) necessary inputs
Interface designed by plugin

Inputs stored in the DB

AiiDA: Calculation example

```
code = Code.get_from_string('pw-6.3@daint-mr25')
calc = code.new_calc()
```

```
calc.set_max_wallclock_seconds(600)
calc.set_resources({"num_machines": 2})
```

```
Structure = DataFactory('structure')
structure = Structure(ase = read('TiO2.cif'))
```

```
Parameter = DataFactory('parameter')
parameters = Parameter({
    'CONTROL': {
        'calculation': 'scf',
        'restart_mode': 'from_scratch',
    },
    'SYSTEM': {
        'ecutwfc': 40.,
    }
})
```

```
Kpoints = DataFactory('array.kpoints')
kpoints = Kpoints(kpoints_mesh = [4,4,4])
```

```
calc.use_structure(structure)
calc.use_parameters(parameters)
calc.use_kpoints(kpoints)
calc.use_pseudos_from_family('SSSP_efficiency_v1.0')
calc.store_all()
```

```
calc.submit()
```

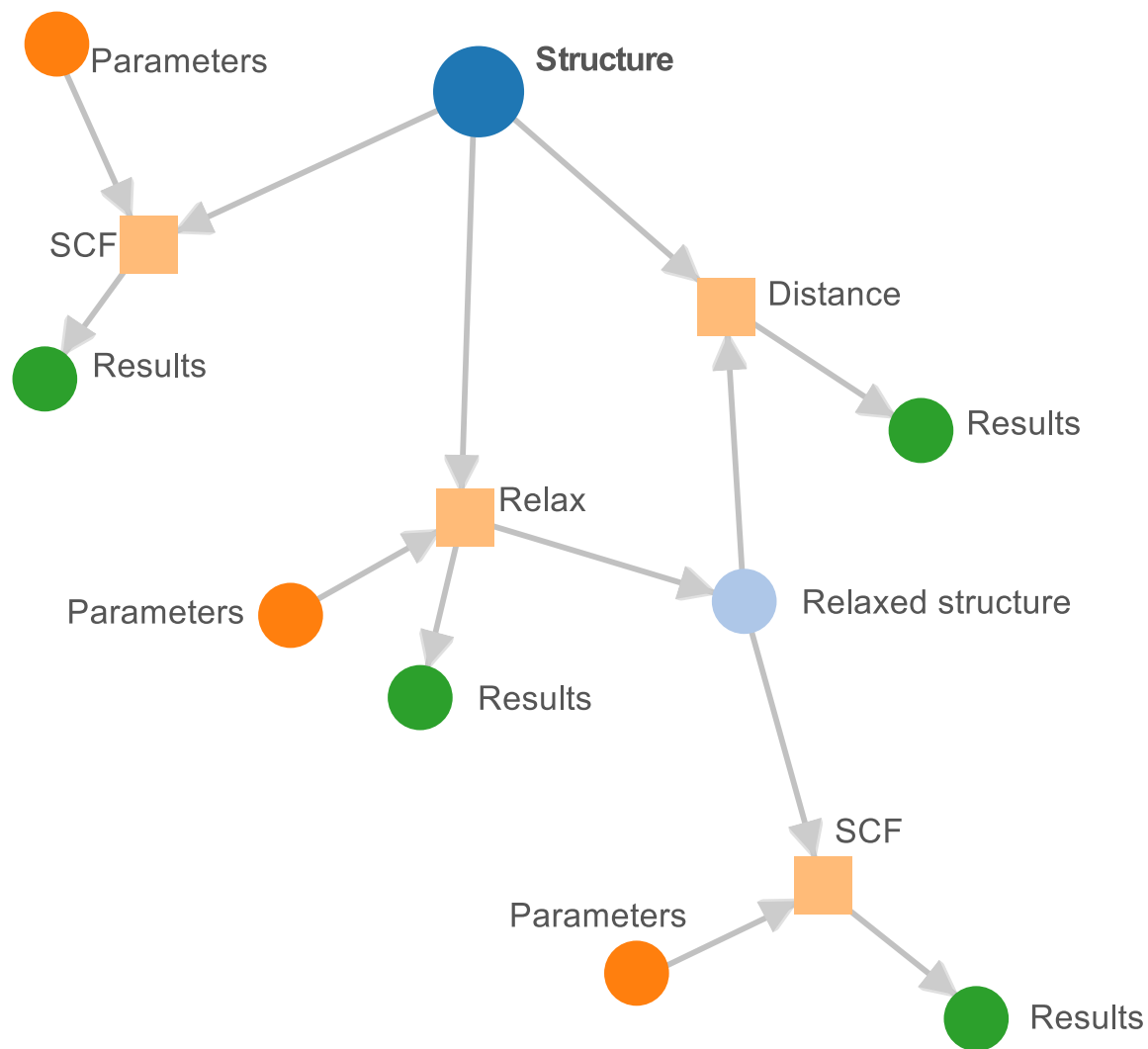
Switch computers in one line
supports different schedulers,
version of codes, ...

Define (only) necessary inputs
Interface designed by plugin

Inputs stored in the DB

Handing over to the daemon

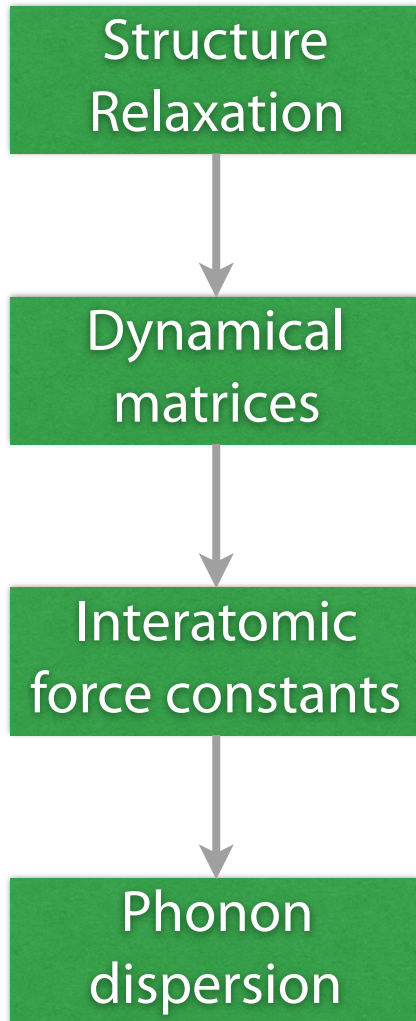
Data provenance: Directed Acyclic Graphs





From calculations to workflows: **phonon dispersion**

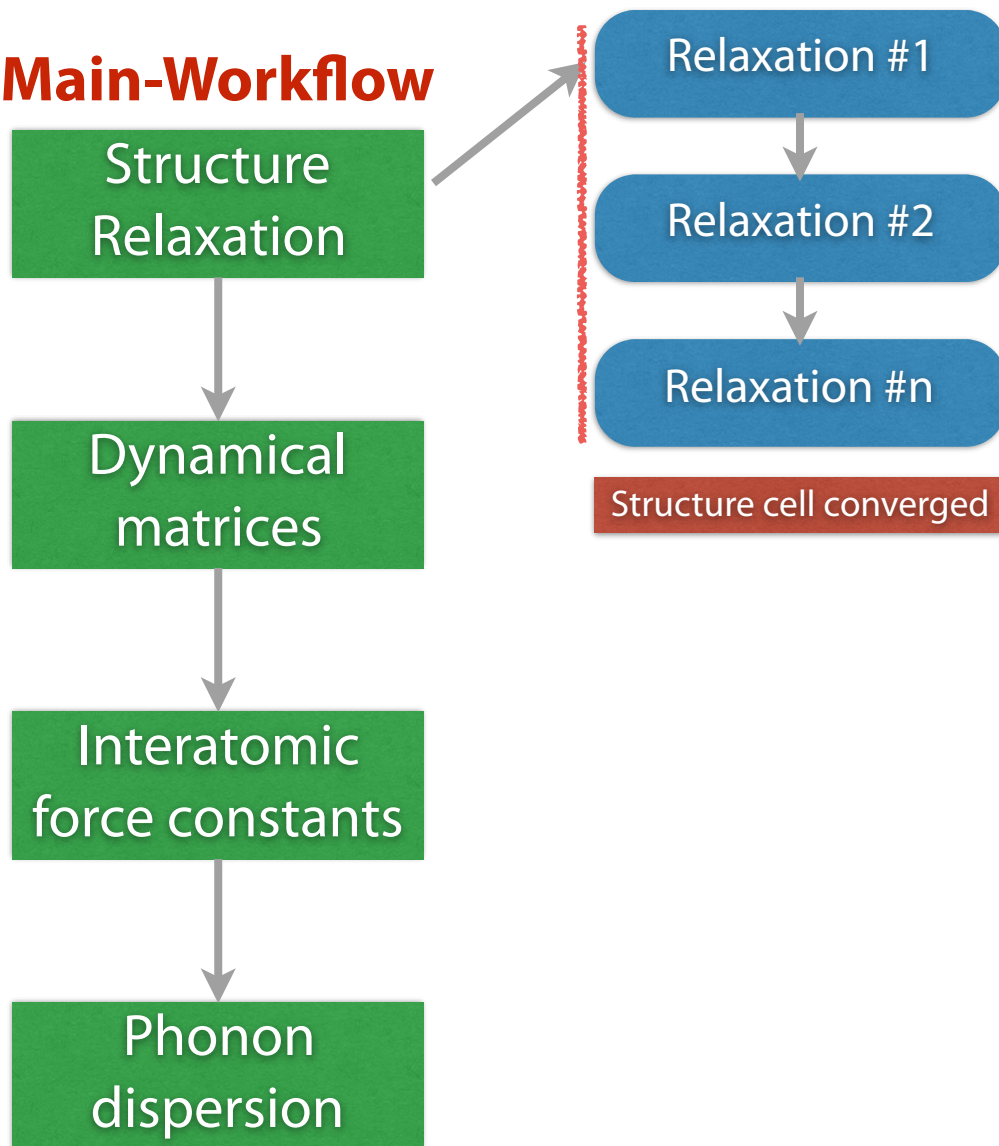
Main-Workflow





From calculations to workflows: **phonon dispersion**

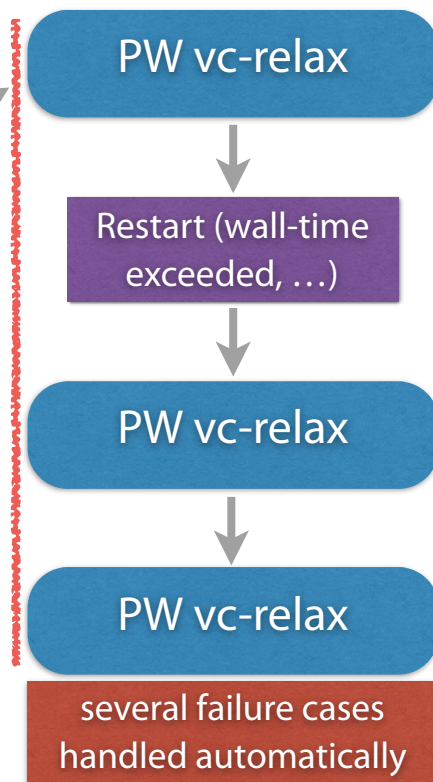
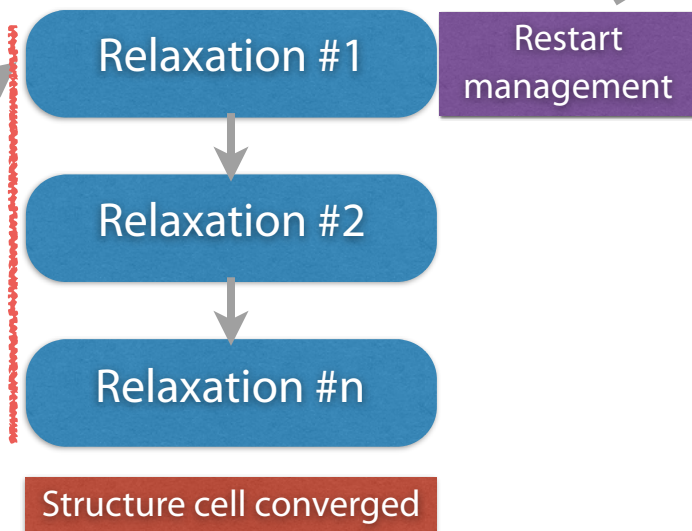
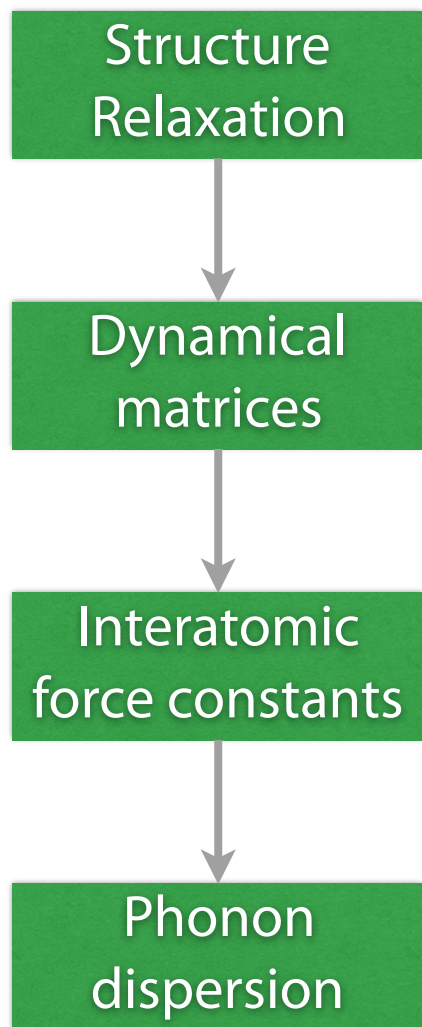
Main-Workflow





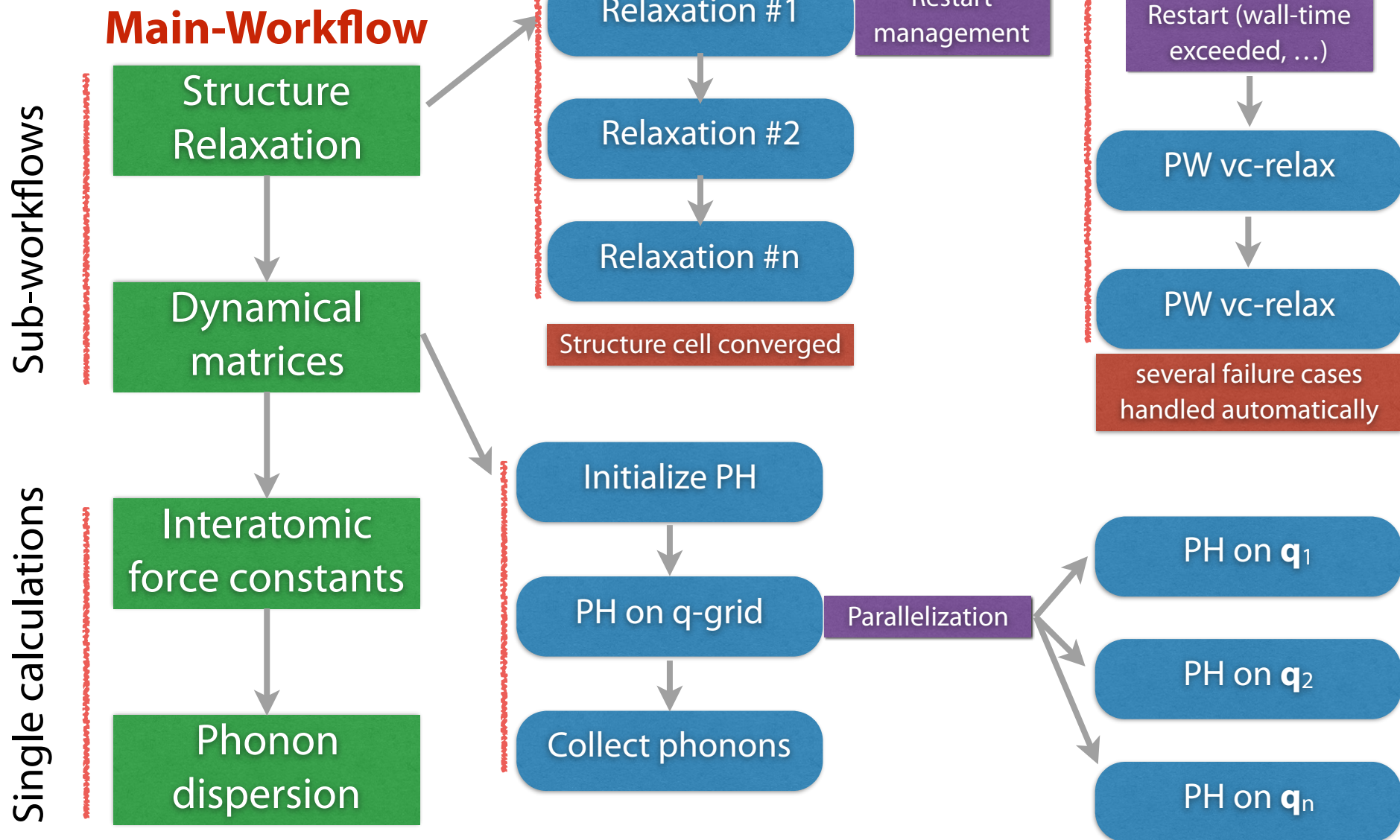
From calculations to workflows: **phonon dispersion**

Main-Workflow





From calculations to workflows: **phonon dispersion**



Workflows = Encapsulation

```
params = {'input': {'kpoints_density': 0.2,
                   'convergence': 'tight'},
          'structure': structure,
          'pseudo_family': pseudo_family,
          'machinename': 'mycluster',
          'pw_input': {'volume_conv_threshold': 5e-2},
          'pw_parameters': {
              'SYSTEM': {'ecutwfc': 30.},
              'ELECTRONS': {'conv_thr': 1.e-10}}
          'ph_input': {
              'distance_kpoints_in_dispersion': 0.005,
              'diagonalization': 'cg'}
          }
future = submit(PhBandsWorkflow, **params)
```

From minimal inputs ...

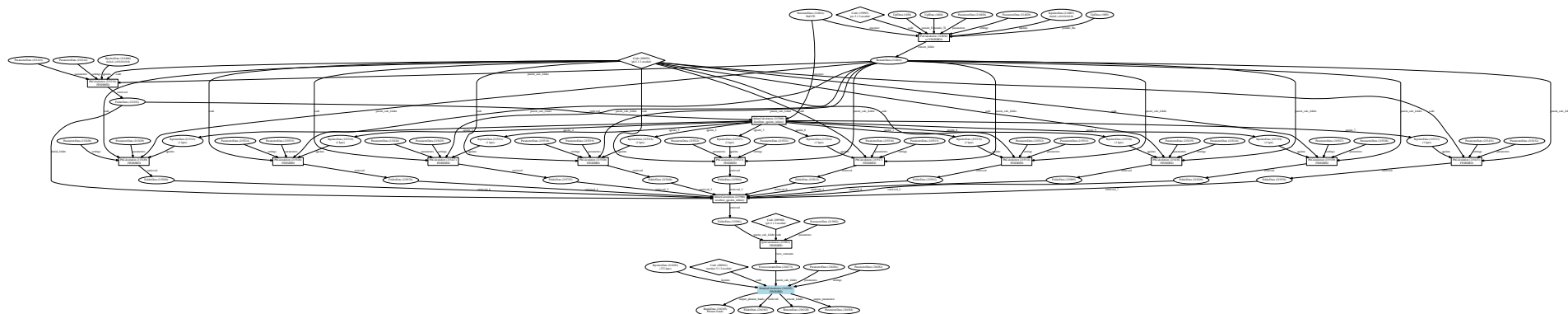
Workflows = Encapsulation

```
params = {'input': {'kpoints_density': 0.2,  
                  'convergence': 'tight'},  
         'structure': structure,  
         'pseudo_family': pseudo_family,  
         'machinename': 'mycluster',  
         'pw_input': {'volume_conv_threshold': 5e-2},  
         'pw_parameters': {'SYSTEM': {'ecutwfc': 30.},  
                           'ELECTRONS': {'conv_thr': 1.e-10}},  
         'ph_input': {'distance_kpoints_in_dispersion': 0.005,  
                      'diagonalization': 'cg'}}  
  
future = submit(PhBandsWorkflow, **params)
```

From minimal inputs ...



... to complex workflows



AiiDA – Development



- **Free and open source (MIT License)**
- github.com/aiidateam/aiida_core
- 36 releases, latest stable version: v0.12.2
- 40 contributors since 2012

Plugin registry: aiidateam.github.io/aiida-registry

AiiDA plugins

ase, castep, codtools, cp2k, crystal17,
fleur, gollum, kkr, lammps, nwchem,
phonopy, quantumespresso, raspa,
siesta, vasp, wannier90, yambo, zeo++,
and more ... (+plugin template)

AiiDA registry of plugins

[\[View on GitHub/register your plugin\]](#)

Global summary of the AiiDA plugin registry

Total number of entries: 25

Calculations	58 plugins in 23 entries
Parsers	52 plugins in 23 entries
Data	24 plugins in 12 entries
Workflows	48 plugins in 9 entries
Other	39 plugins in 11 entries

Available plugins (alphabetically sorted)

[ase: aiida-ase package](#)

The official AiiDA plugin for ASE
Current state: stable

[Plugin code homepage \(hosted on GitHub.com\)](#)

2



MATERIALS
CLOUD



AiiDA lab

Motivation

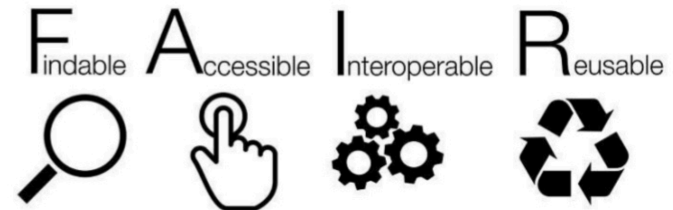
High-Throughput

Reproducibility

Open Science

Knowledge Transfer

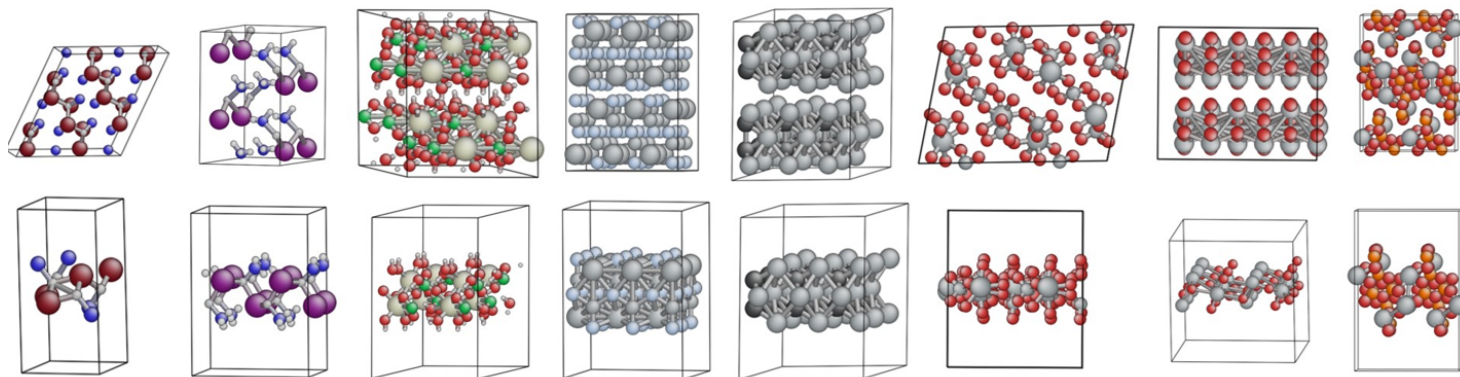
- Open Science
 - ~~Supporting Information~~
 - ~~Just upload everything~~
 - FAIR data
- Making data FAIR is hard, can we make it easier?



Source: Prof. Michel Dumontier

High-throughput Example

DISCOVERING NEW TWO-DIMENSIONAL MATERIALS



STARTING FROM ICSD/COD DATABASE:

- **108 423** unique 3D structures
- **5619** layered structures
- **>100 000** DFT calculations
- **>30 000** material properties
- **>1·10⁹** attributes

III. METHODS

A. Reproducibility and provenance

It is an often repeated tenet that results of scientific research must be reproducible. This objective is, however, challenging, especially in high-throughput research, due to the large number of simulations involved and the complex sequence of logical steps needed in the study. To ensure reproducibility we use AiiDA³⁰ as a materials' informatics infrastructure to implement the ADES model of automation, data, environment, and sharing, as dis-

As already mentioned in the main text, in this step we consider only structures coming from experimental measurements, while the source databases are partially populated by purely theoretical structures. This is done using the flags set on the database entries by their curators. We also implemented some heuristics to detect clearly wrong CIF files. For example, we discard structures where the chemical formula provided in the file is inconsistent with the elements in the unit cell. Regardless of all these efforts, it is possible that some incomplete or incorrect structures are still not filtered out from the original databases.

Methods: Impossible to describe every detail

For **authors**, reproducing all data is challenging.
For **peers**, reproducing all data is almost impossible.

FAIR sharing example

ARTICLES

<https://doi.org/10.1038/s41565-017-0035-5>

nature
nanotechnology

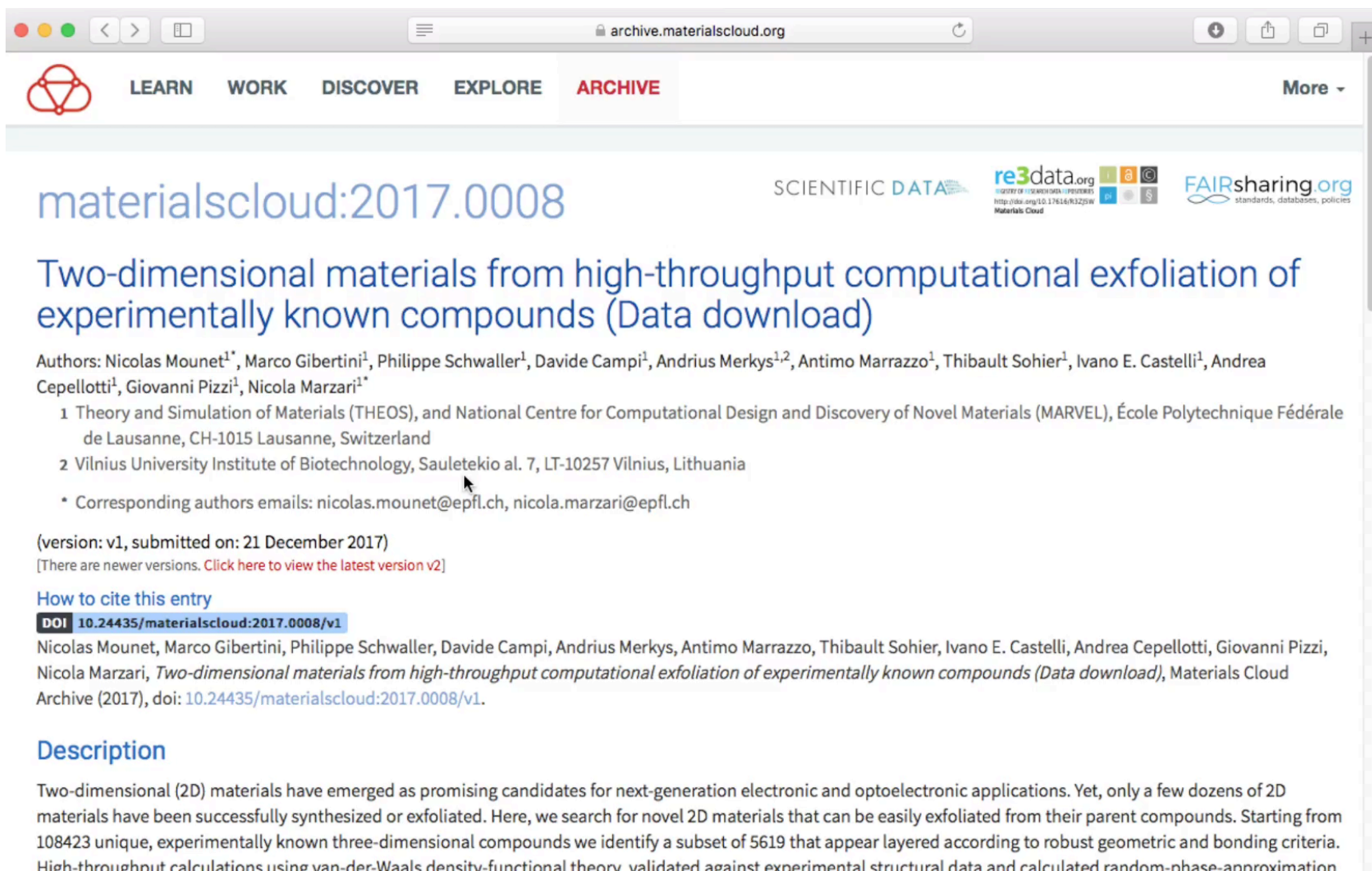
Two-dimensional materials from high-throughput computational exfoliation of experimentally known compounds

order. All results can be fully reproduced thanks to the deployment of the AiiDA infrastructure that tracks the provenance of each data entry. To date, this is the largest available database of 2D compounds and it is available online on the Materials Cloud platform⁵¹.

51. Mounet, N. et al. Two-dimensional materials from high-throughput computational exfoliation of experimentally known compounds (data download). *Materials Cloud Archive* (2017); <https://doi.org/10.24435/materialscloud:2017.0008/v1>

<https://archive.materialscloud.org/2017.0008/v1>

FAIR sharing example



The screenshot shows a web browser window with the address bar displaying `archive.materialscloud.org`. The page header includes navigation tabs: LEARN, WORK, DISCOVER, EXPLORE, and ARCHIVE (which is highlighted). A 'More' dropdown menu is visible on the right. The main content area features the identifier `materialscloud:2017.0008` in large blue text. Below this is the title 'Two-dimensional materials from high-throughput computational exfoliation of experimentally known compounds (Data download)'. The authors listed are Nicolas Mounet^{1*}, Marco Gibertini¹, Philippe Schwaller¹, Davide Campi¹, Andrius Merkys^{1,2}, Antimo Marrazzo¹, Thibault Sohier¹, Ivano E. Castelli¹, Andrea Cepellotti¹, Giovanni Pizzi¹, and Nicola Marzari^{1*}. Two affiliations are provided: 1. Theory and Simulation of Materials (THEOS), and National Centre for Computational Design and Discovery of Novel Materials (MARVEL), École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland; 2. Vilnius University Institute of Biotechnology, Sauletekio al. 7, LT-10257 Vilnius, Lithuania. Corresponding authors' emails are `nicolas.mounet@epfl.ch` and `nicola.marzari@epfl.ch`. The version is noted as v1, submitted on 21 December 2017, with a link to view the latest version v2. A 'How to cite this entry' section provides the DOI `10.24435/materialscloud:2017.0008/v1` and the full citation text. A 'Description' section begins with a paragraph about 2D materials.

materialscloud:2017.0008

Two-dimensional materials from high-throughput computational exfoliation of experimentally known compounds (Data download)

Authors: Nicolas Mounet^{1*}, Marco Gibertini¹, Philippe Schwaller¹, Davide Campi¹, Andrius Merkys^{1,2}, Antimo Marrazzo¹, Thibault Sohier¹, Ivano E. Castelli¹, Andrea Cepellotti¹, Giovanni Pizzi¹, Nicola Marzari^{1*}

1 Theory and Simulation of Materials (THEOS), and National Centre for Computational Design and Discovery of Novel Materials (MARVEL), École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland

2 Vilnius University Institute of Biotechnology, Sauletekio al. 7, LT-10257 Vilnius, Lithuania

* Corresponding authors emails: `nicolas.mounet@epfl.ch`, `nicola.marzari@epfl.ch`

(version: v1, submitted on: 21 December 2017)
[There are newer versions. [Click here to view the latest version v2](#)]

How to cite this entry

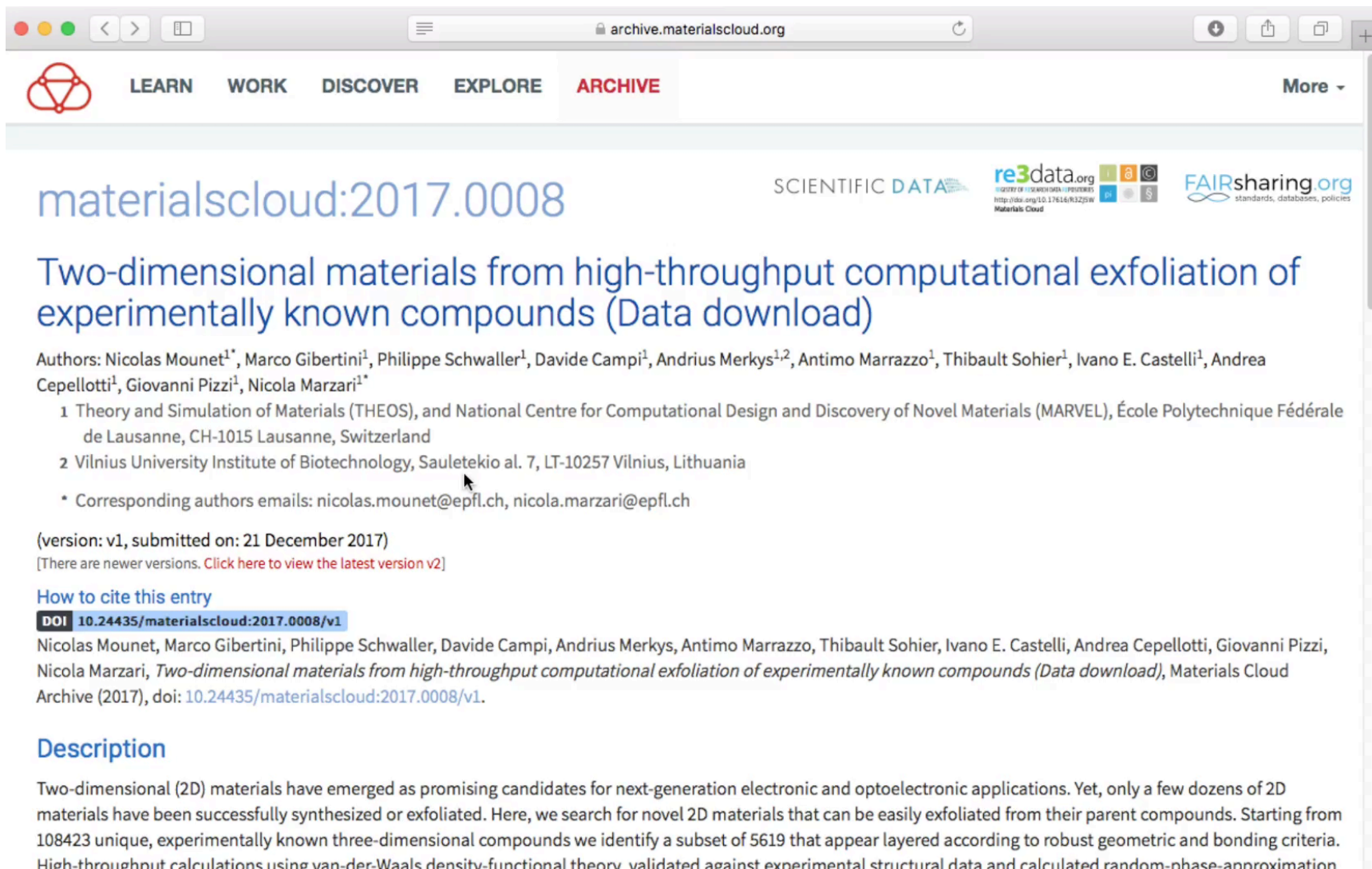
DOI `10.24435/materialscloud:2017.0008/v1`

Nicolas Mounet, Marco Gibertini, Philippe Schwaller, Davide Campi, Andrius Merkys, Antimo Marrazzo, Thibault Sohier, Ivano E. Castelli, Andrea Cepellotti, Giovanni Pizzi, Nicola Marzari, *Two-dimensional materials from high-throughput computational exfoliation of experimentally known compounds (Data download)*, Materials Cloud Archive (2017), doi: `10.24435/materialscloud:2017.0008/v1`.

Description

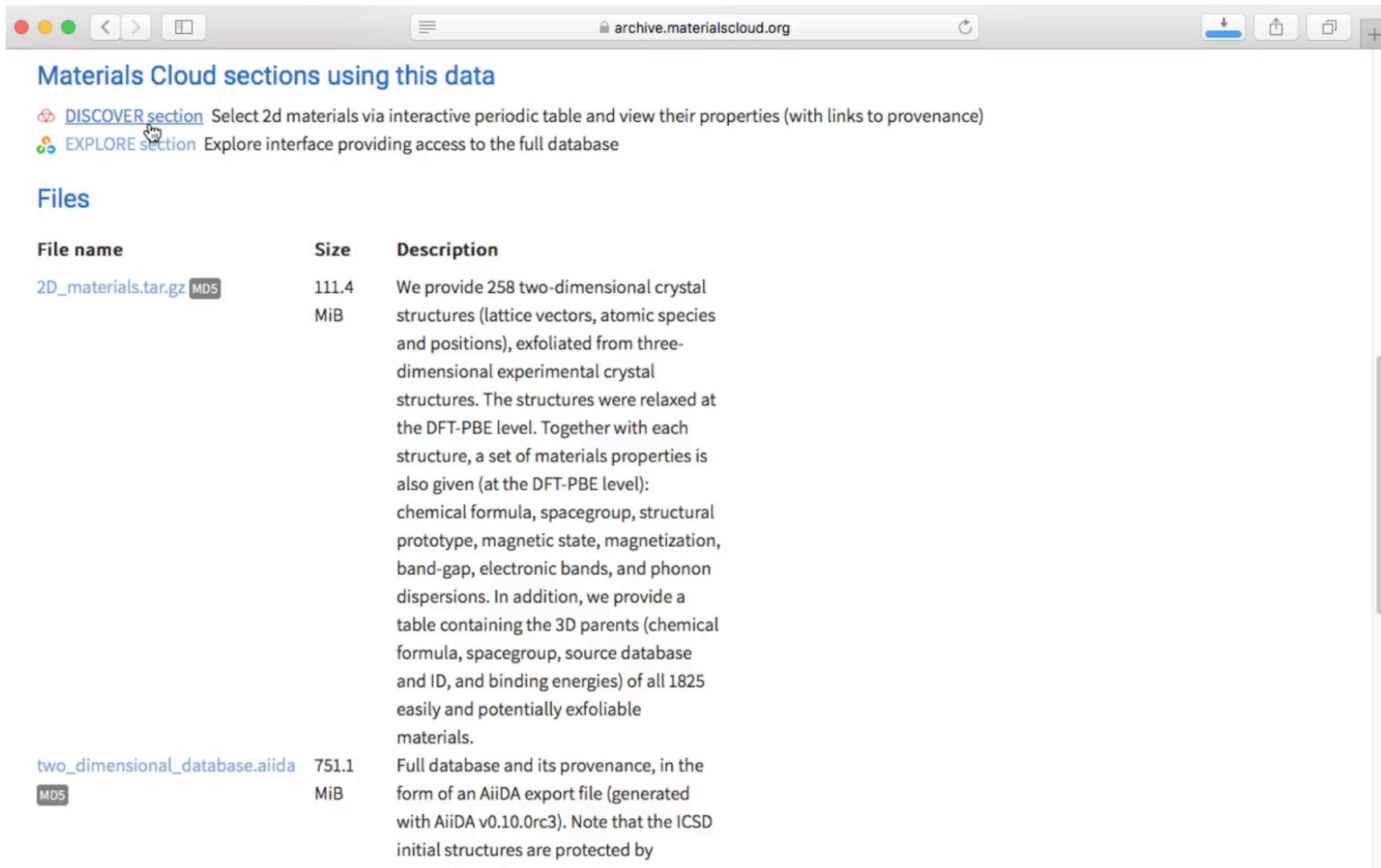
Two-dimensional (2D) materials have emerged as promising candidates for next-generation electronic and optoelectronic applications. Yet, only a few dozens of 2D materials have been successfully synthesized or exfoliated. Here, we search for novel 2D materials that can be easily exfoliated from their parent compounds. Starting from 108423 unique, experimentally known three-dimensional compounds we identify a subset of 5619 that appear layered according to robust geometric and bonding criteria. High-throughput calculations using van-der-Waals density-functional theory, validated against experimental structural data and calculated random-phase-approximation

FAIR sharing example



The screenshot shows a web browser window with the address bar displaying `archive.materialscloud.org`. The page header includes navigation tabs: **LEARN**, **WORK**, **DISCOVER**, **EXPLORE**, and **ARCHIVE** (which is highlighted). A **More** dropdown menu is visible on the right. The main content area features the identifier `materialscloud:2017.0008` in large blue text. To the right of this identifier are logos for **SCIENTIFIC DATA**, **re3data.org**, and **FAIRsharing.org**. Below the identifier is the title **Two-dimensional materials from high-throughput computational exfoliation of experimentally known compounds (Data download)**. The authors listed are Nicolas Mounet^{1*}, Marco Gibertini¹, Philippe Schwaller¹, Davide Campi¹, Andrius Merkys^{1,2}, Antimo Marrazzo¹, Thibault Sohier¹, Ivano E. Castelli¹, Andrea Cepellotti¹, Giovanni Pizzi¹, and Nicola Marzari^{1*}. Two affiliations are provided: 1. Theory and Simulation of Materials (THEOS), and National Centre for Computational Design and Discovery of Novel Materials (MARVEL), École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland; 2. Vilnius University Institute of Biotechnology, Sauletekio al. 7, LT-10257 Vilnius, Lithuania. Corresponding authors' emails are `nicolas.mounet@epfl.ch` and `nicola.marzari@epfl.ch`. The version is noted as **v1**, submitted on **21 December 2017**, with a link to view the latest version **v2**. A section titled **How to cite this entry** provides the DOI `10.24435/materialscloud:2017.0008/v1` and a full citation: Nicolas Mounet, Marco Gibertini, Philippe Schwaller, Davide Campi, Andrius Merkys, Antimo Marrazzo, Thibault Sohier, Ivano E. Castelli, Andrea Cepellotti, Giovanni Pizzi, Nicola Marzari, *Two-dimensional materials from high-throughput computational exfoliation of experimentally known compounds (Data download)*, Materials Cloud Archive (2017), doi: `10.24435/materialscloud:2017.0008/v1`. A **Description** section follows, stating that two-dimensional (2D) materials have emerged as promising candidates for next-generation electronic and optoelectronic applications, and that the authors searched for novel 2D materials that can be easily exfoliated from their parent compounds, starting from 108423 unique, experimentally known three-dimensional compounds and identifying a subset of 5619 that appear layered according to robust geometric and bonding criteria. High-throughput calculations using van-der-Waals density-functional theory, validated against experimental structural data and calculated random-phase-approximation

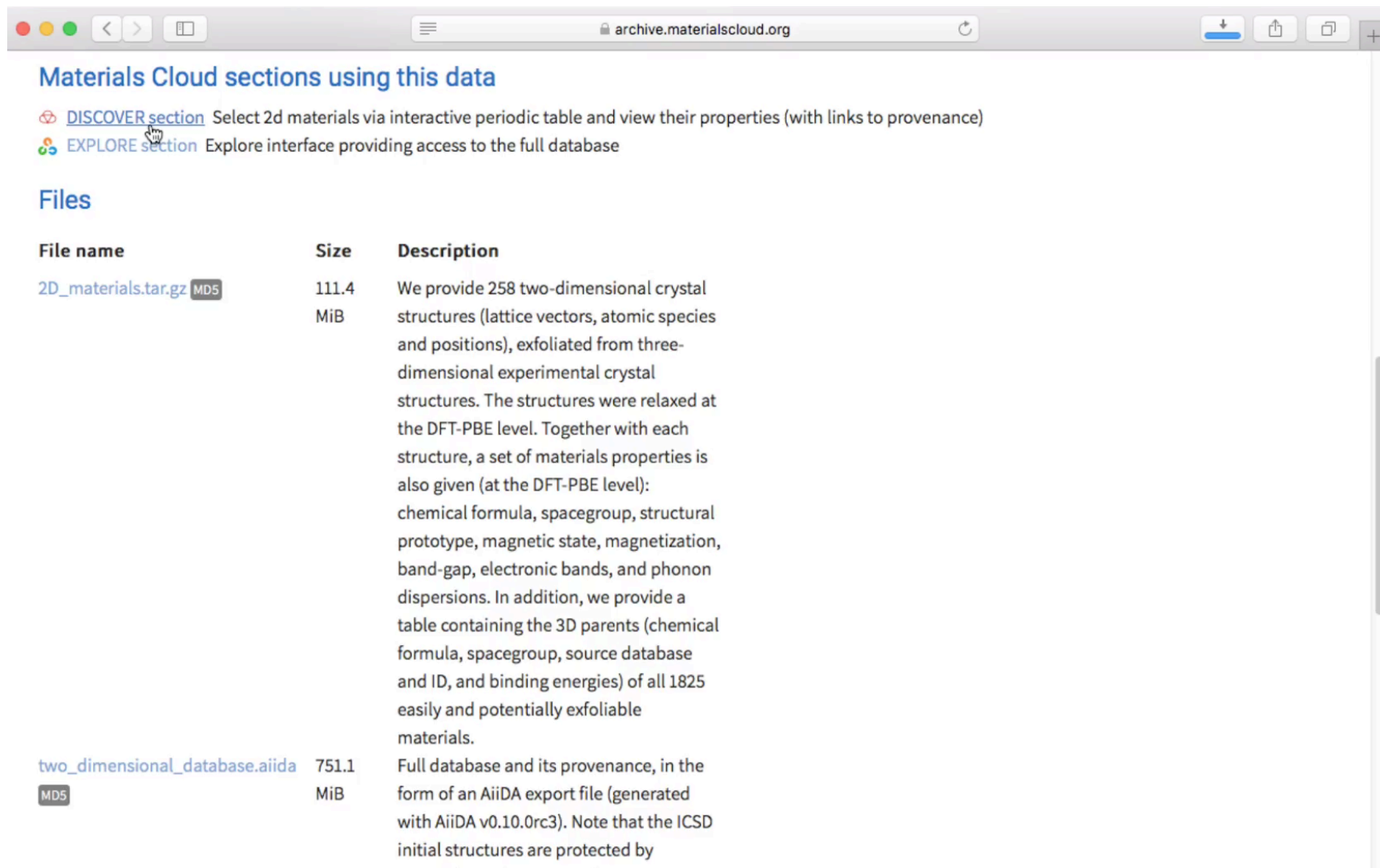
FAIR sharing example



The screenshot shows a web browser window with the address bar displaying `archive.materialscloud.org`. The page title is "Materials Cloud sections using this data". Below the title, there are two links: "DISCOVER section" (with a magnifying glass icon) and "EXPLORE section" (with a person icon). The "EXPLORE section" link is highlighted with a mouse cursor. Below the links, there is a section titled "Files" which contains a table with three columns: "File name", "Size", and "Description".

File name	Size	Description
2D_materials.tar.gz MD5	111.4 MiB	We provide 258 two-dimensional crystal structures (lattice vectors, atomic species and positions), exfoliated from three-dimensional experimental crystal structures. The structures were relaxed at the DFT-PBE level. Together with each structure, a set of materials properties is also given (at the DFT-PBE level): chemical formula, spacegroup, structural prototype, magnetic state, magnetization, band-gap, electronic bands, and phonon dispersions. In addition, we provide a table containing the 3D parents (chemical formula, spacegroup, source database and ID, and binding energies) of all 1825 easily and potentially exfoliable materials.
two_dimensional_database.aiida MD5	751.1 MiB	Full database and its provenance, in the form of an AiiDA export file (generated with AiiDA v0.10.0rc3). Note that the ICSD initial structures are protected by

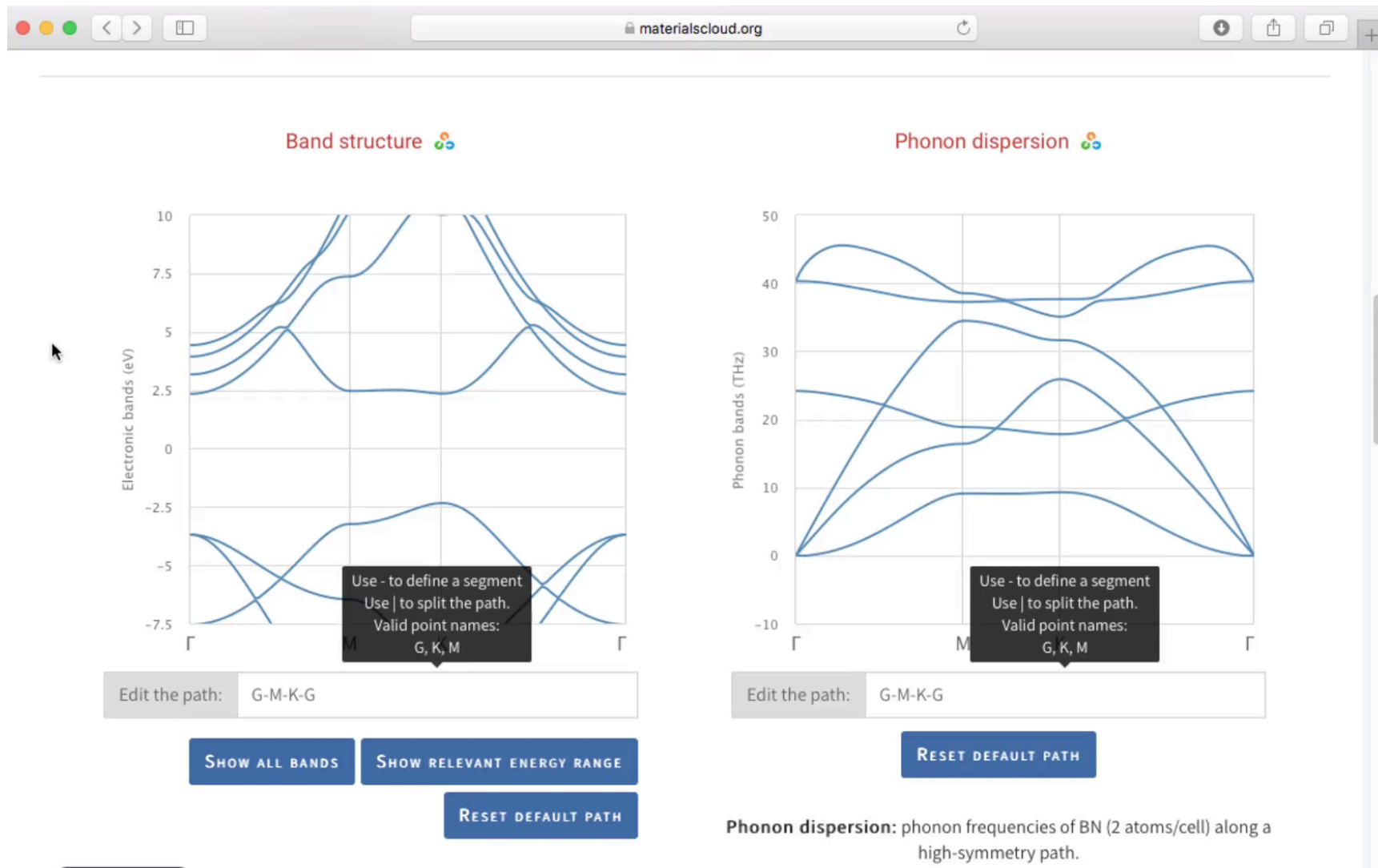
FAIR sharing example



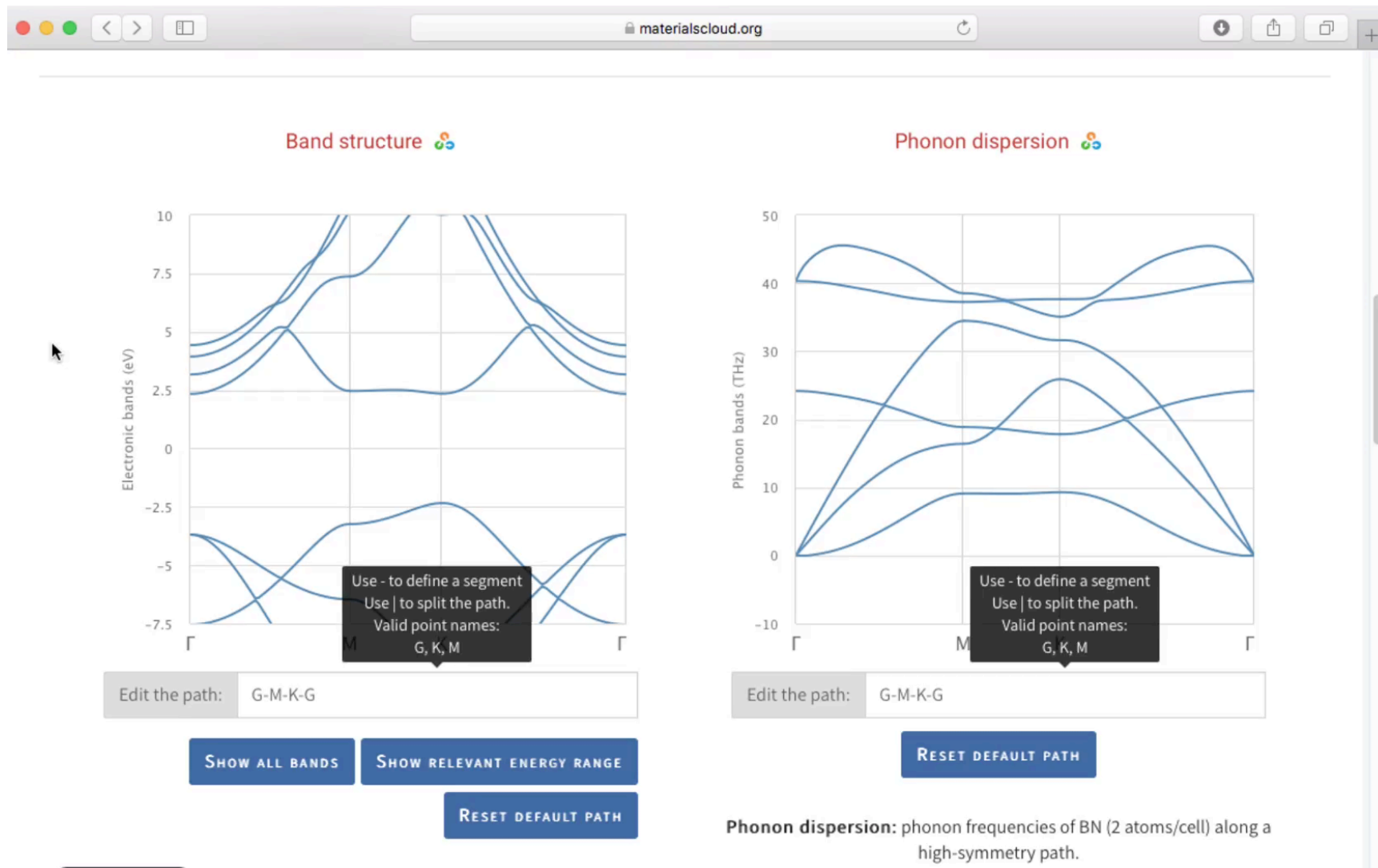
The screenshot shows a web browser window with the address bar displaying "archive.materialscloud.org". The page title is "Materials Cloud sections using this data". Below the title, there are two links: "DISCOVER section" (with a magnifying glass icon) and "EXPLORE section" (with a person icon). The "EXPLORE section" link is highlighted with a mouse cursor. Below the links, there is a section titled "Files" which contains a table with three columns: "File name", "Size", and "Description".

File name	Size	Description
2D_materials.tar.gz MD5	111.4 MiB	We provide 258 two-dimensional crystal structures (lattice vectors, atomic species and positions), exfoliated from three-dimensional experimental crystal structures. The structures were relaxed at the DFT-PBE level. Together with each structure, a set of materials properties is also given (at the DFT-PBE level): chemical formula, spacegroup, structural prototype, magnetic state, magnetization, band-gap, electronic bands, and phonon dispersions. In addition, we provide a table containing the 3D parents (chemical formula, spacegroup, source database and ID, and binding energies) of all 1825 easily and potentially exfoliable materials.
two_dimensional_database.aiida MD5	751.1 MiB	Full database and its provenance, in the form of an AiiDA export file (generated with AiiDA v0.10.0rc3). Note that the ICSD initial structures are protected by

FAIR sharing example



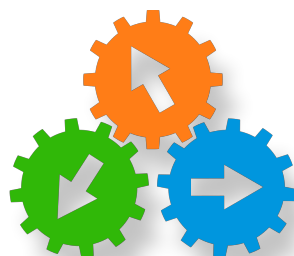
FAIR sharing example



2



MATERIALS
CLOUD



AiiDA lab

Motivation

High-Throughput

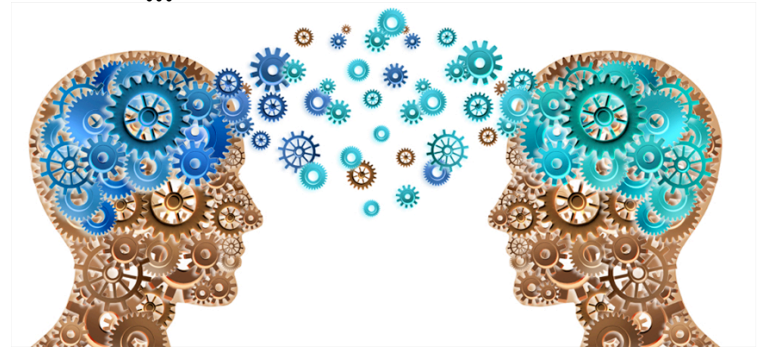
Reproducibility

Open Science

Knowledge Transfer

- How to transfer your insights & expertise, e.g.
 - reporting data sets to an experimental group
 - providing a workflow for your code to a collaborator/ company

• ...



Source: quote.ucsd.edu

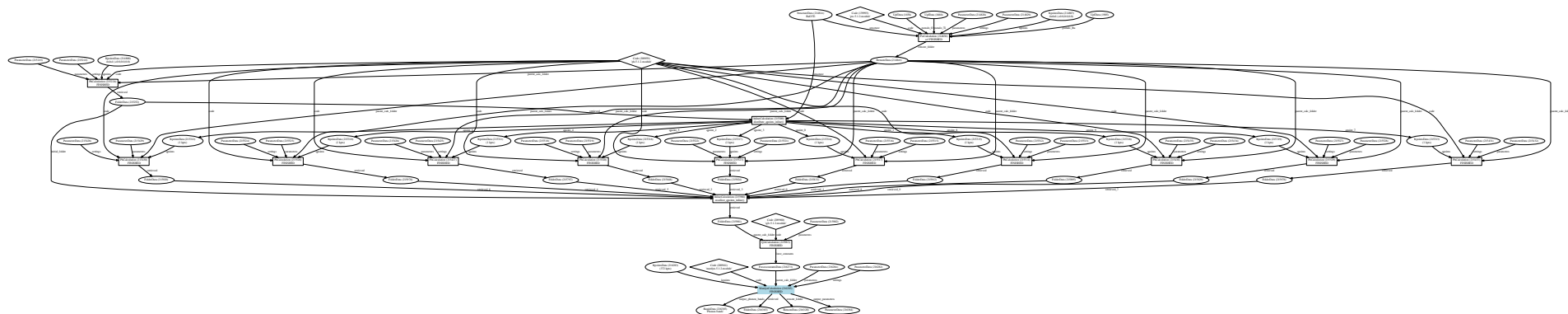
Workflows = Encapsulation

```
params = {'input': {'kpoints_density': 0.2,  
                  'convergence': 'tight'},  
         'structure': structure,  
         'pseudo_family': pseudo_family,  
         'machinename': 'mycluster',  
         'pw_input': {'volume_conv_threshold': 5e-2},  
         'pw_parameters': {'SYSTEM': {'ecutwfc': 30.},  
                           'ELECTRONS': {'conv_thr': 1.e-10}},  
         'ph_input': {'distance_kpoints_in_dispersion': 0.005,  
                      'diagonalization': 'cg'}}  
  
future = submit(PhBandsWorkflow, **params)
```

From minimal inputs ...



... to complex workflows



User base

User	Skills	Goals	Solution
------	--------	-------	----------

User base

User	Skills	Goals	Solution
Computational Scientist	Knows Unix, bash, python	<ul style="list-style-type: none">• run high-throughput calculations• write complex workflows• develop AiiDA plugins	AiiDA on the laptop

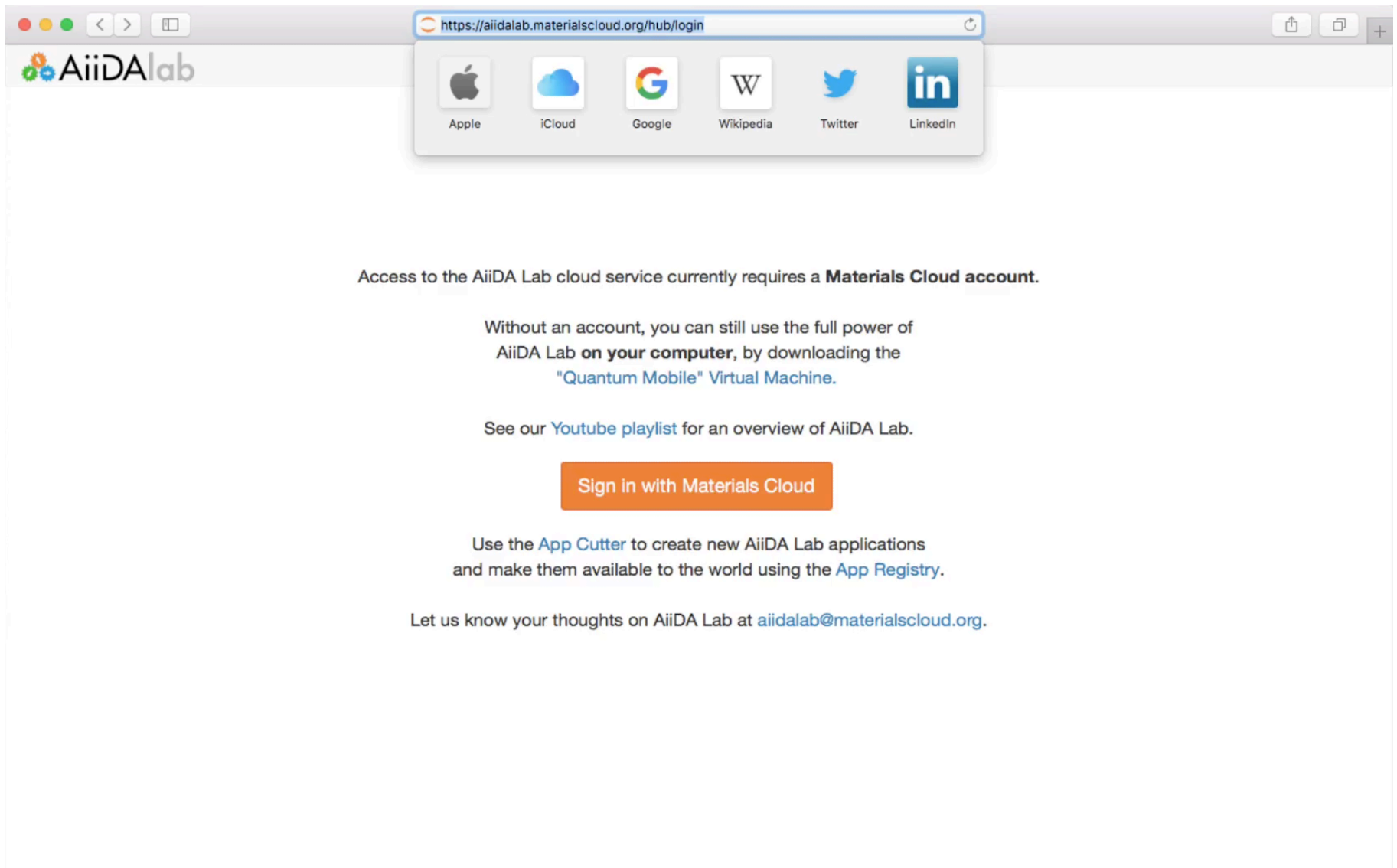
User base

User	Skills	Goals	Solution
Computational Scientist	Knows Unix, bash, python	<ul style="list-style-type: none">• run high-throughput calculations• write complex workflows• develop AiiDA plugins	AiiDA on the laptop
Experimental Scientist	Doesn't know Unix, bash, python	<ul style="list-style-type: none">• run pre-defined workflows• analyze results	AiiDA Lab in the cloud

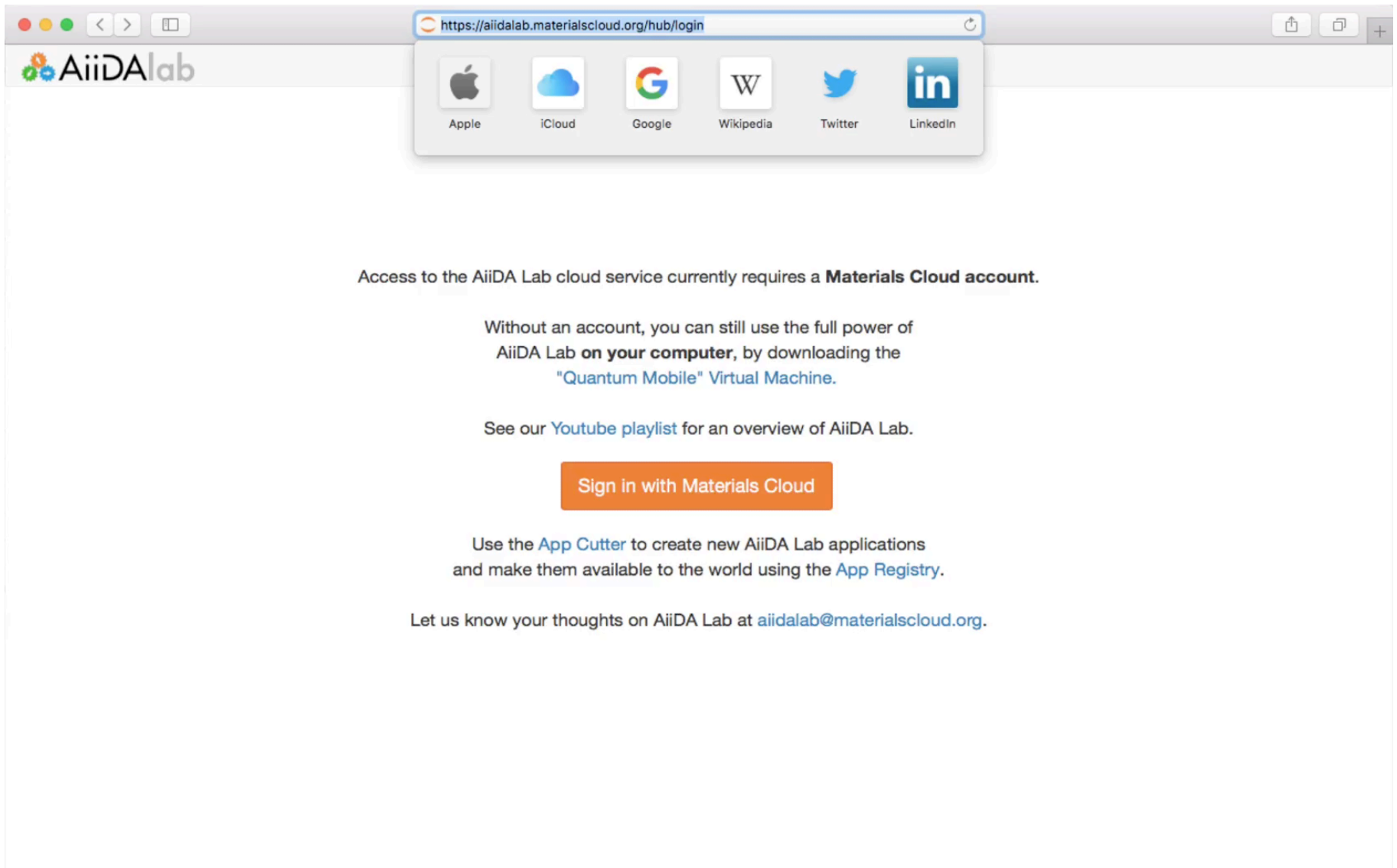
User base

User	Skills	Goals	Solution
Computational Scientist	Knows Unix, bash, python	<ul style="list-style-type: none">• run high-throughput calculations• write complex workflows• develop AiiDA plugins	AiiDA on the laptop
Experimental Scientist	Doesn't know Unix, bash, python	<ul style="list-style-type: none">• run pre-defined workflows• analyze results	AiiDA Lab in the cloud
Student (tutorial/ lecture)	some familiarity with Unix, bash, python	<ul style="list-style-type: none">• learn how to use AiiDA• learn how to use ab-initio codes• take materials home	Quantum Mobile on the laptop

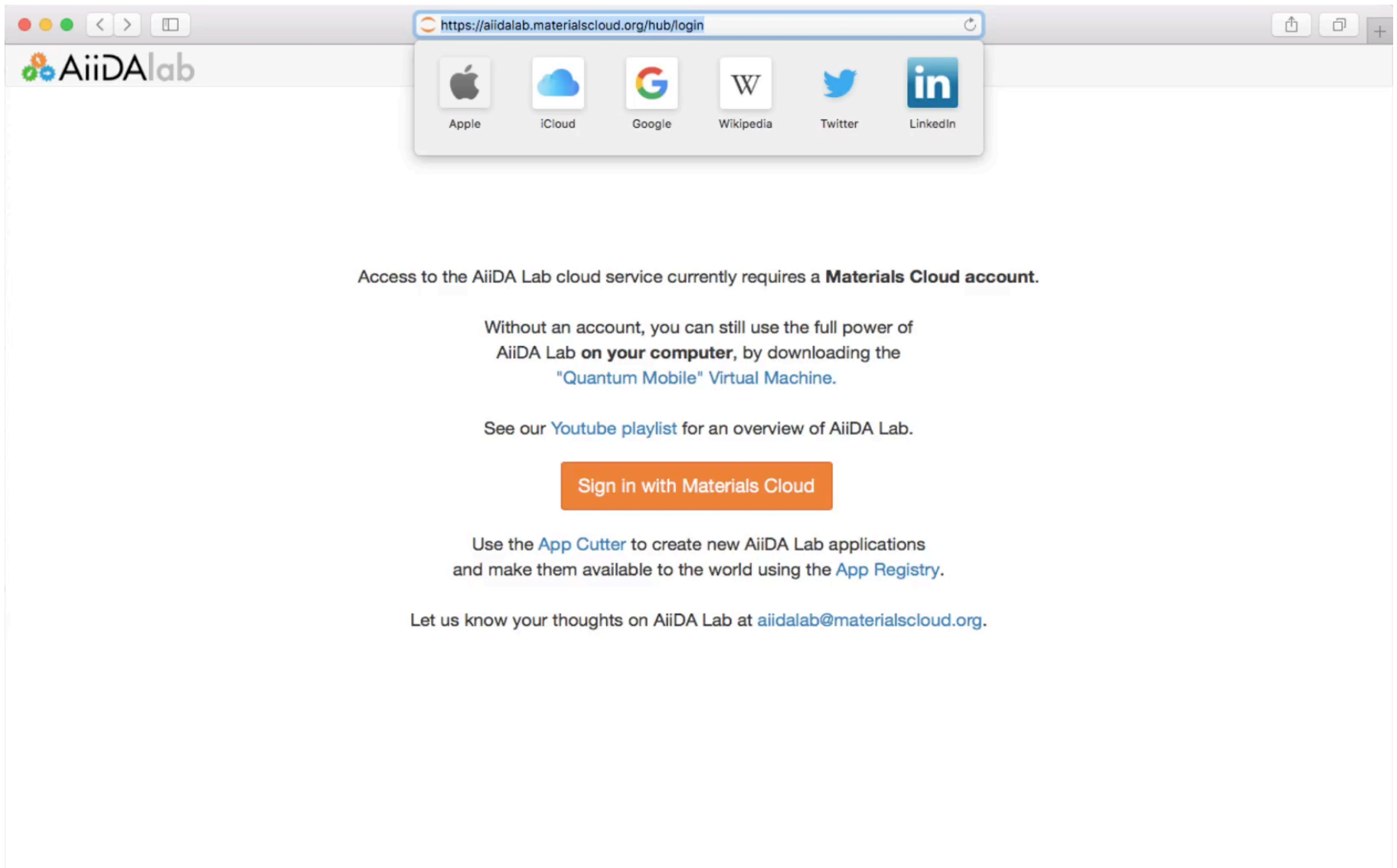
AiiDA Lab demo - login



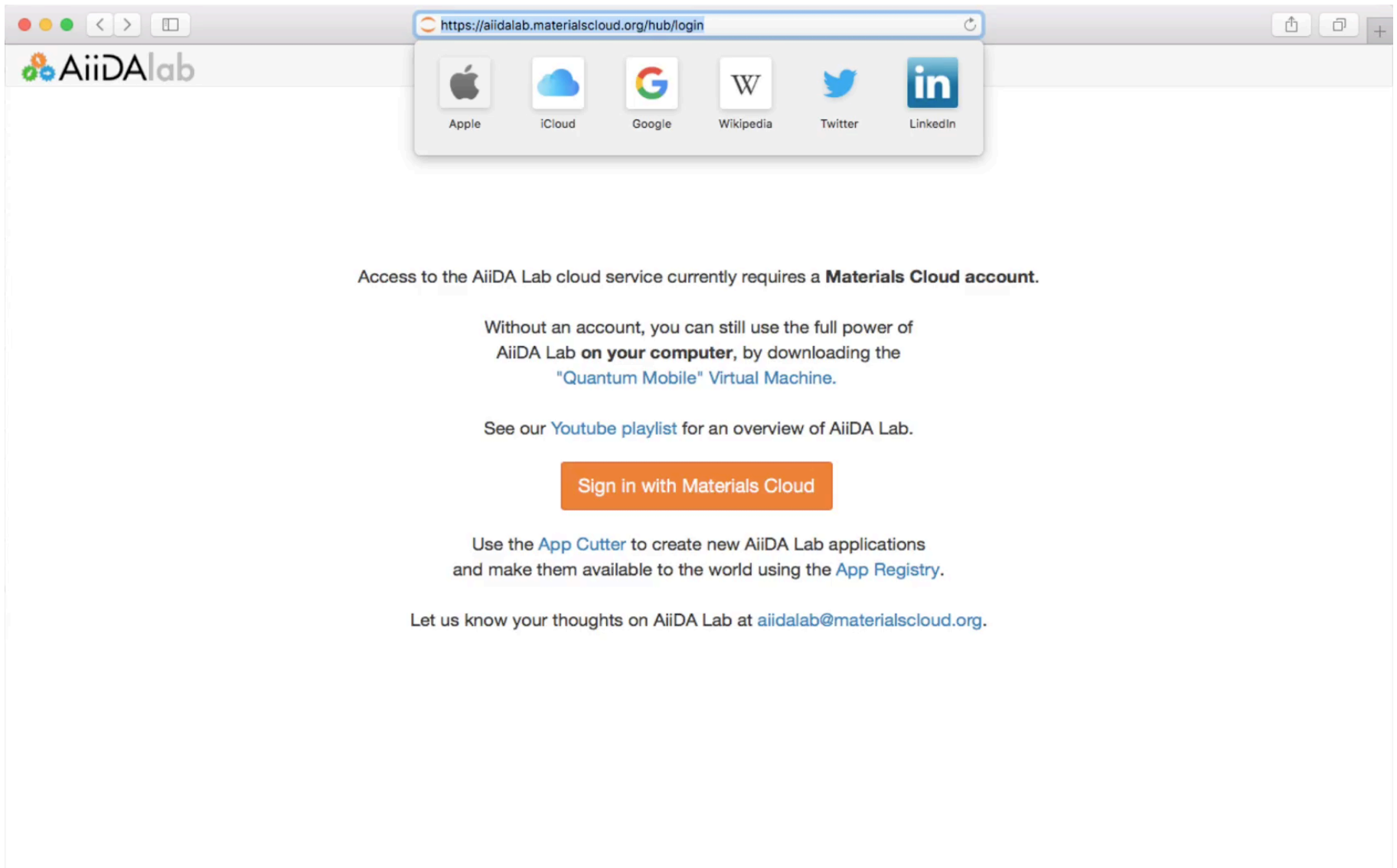
AiiDA Lab demo - login



AiiDA Lab demo - login



AiiDA Lab demo - login



AiiDA Lab demo - appmode

The screenshot shows the AiiDA Lab appmode interface in a web browser. The browser's address bar displays `aiidalab.materialscloud.org`. The page header includes the AiiDA Lab logo and four navigation buttons: "Edit App", "Logout", "Control Panel", and "Materials Cloud".

The main content area is titled "AiiDA Lab" and is organized into four collapsible sections:

- Home**: Contains four large icons for "File Browser", "Terminal", "Tasks", and "Manage Apps".
- AiiDA**: Contains a list of links: "Daemon Status", "Graph Browser", "Workflows", "Delete nodes", and "REST API".
- Quantum Mobile**: Contains two links: "Setup Connection to Quantum-Mobile" and "Setup Codes".
- LSMO apps**: Contains four columns of links:
 - Isotherm**: "Compute one", "Compute multiple"
 - Charges**: "Compute Charges"
 - Pore analysis**: "Compute Pores"
 - Import your data to AiiDA**: "Import database", "Plot imported data"

AiiDA Lab demo - appmode

The screenshot shows the AiiDA Lab appmode interface in a web browser. The browser's address bar displays `aiidalab.materialscloud.org`. The page header includes the AiiDA Lab logo and four navigation buttons: "Edit App", "Logout", "Control Panel", and "Materials Cloud".

The main content area is titled "AiiDA Lab" and is organized into four collapsible sections:

- Home**: Contains four large icons for "File Browser", "Terminal", "Tasks", and "Manage Apps".
- AiiDA**: Contains a list of links: "Daemon Status", "Graph Browser", "Workflows", "Delete nodes", and "REST API".
- Quantum Mobile**: Contains two links: "Setup Connection to Quantum-Mobile" and "Setup Codes".
- LSMO apps**: Contains four columns of links:
 - Isotherm**: "Compute one", "Compute multiple"
 - Charges**: "Compute Charges"
 - Pore analysis**: "Compute Pores"
 - Import your data to AiiDA**: "Import database", "Plot imported data"

AiiDA Lab demo - terminal

The screenshot shows a web browser window with the address bar displaying `aiidalab.materialscloud.org`. The page features a navigation bar with the AiiDA Lab logo and buttons for "Edit App", "Logout", "Control Panel", and "Materials Cloud". The main content area is titled "AiiDA Lab" and is organized into several sections:

- Home**: Contains four large icons for "File Browser", "Terminal", "Tasks", and "Manage Apps".
- AiiDA**: A list of links including "Daemon Status", "Graph Browser", "Workflows", "Delete nodes", and "REST API".
- Quantum Mobile**: A list of links for "Setup Connection to Quantum-Mobile" and "Setup Codes".
- LSMO apps**: A grid of links for "Isotherm" (Compute one, Compute multiple), "Charges" (Compute Charges), "Pore analysis" (Compute Pores), and "Import your data to AiiDA" (Import database, Plot imported data).

Each section has a vertical scrollbar on its right side.

AiiDA Lab demo - terminal

The screenshot shows a web browser window with the address bar displaying `aiidalab.materialscloud.org`. The page features a navigation bar with the AiiDA Lab logo and buttons for "Edit App", "Logout", "Control Panel", and "Materials Cloud". The main content area is titled "AiiDA Lab" and is organized into several sections:

- Home**: Contains four large icons for "File Browser", "Terminal", "Tasks", and "Manage Apps".
- AiiDA**: A list of links including "Daemon Status", "Graph Browser", "Workflows", "Delete nodes", and "REST API".
- Quantum Mobile**: A list of links for "Setup Connection to Quantum-Mobile" and "Setup Codes".
- LSMO apps**: A grid of links for "Isotherm" (Compute one, Compute multiple), "Charges" (Compute Charges), "Pore analysis" (Compute Pores), and "Import your data to AiiDA" (Import database, Plot imported data).

Each section has a vertical scrollbar on its right side.

AiiDA Lab demo - phonon submit

The screenshot shows the AiiDA Lab web interface in a browser window. The address bar displays `aiidalab.materialscloud.org`. The interface includes a top navigation bar with the AiiDA Lab logo and buttons for `Edit App`, `Logout`, `Control Panel`, and `Materials Cloud`.

The main content area is titled **AiiDA Lab** and features a sidebar with expandable sections:

- Home**: Contains icons and links for `File Browser`, `Terminal`, `Tasks`, and `Manage Apps`.
- AiiDA**: Contains a list of links: `Daemon Status`, `Graph Browser`, `Workflows`, `Delete nodes`, and `REST API`.
- Quantum Mobile**: Contains links for `Setup Connection to Quantum-Mobile` and `Setup Codes`.
- LSMO apps**: Contains four columns of links:
 - Isotherm**: `Compute one`, `Compute multiple`
 - Charges**: `Compute Charges`
 - Pore analysis**: `Compute Pores`
 - Import your data to AiiDA**: `Import database`, `Plot imported data`

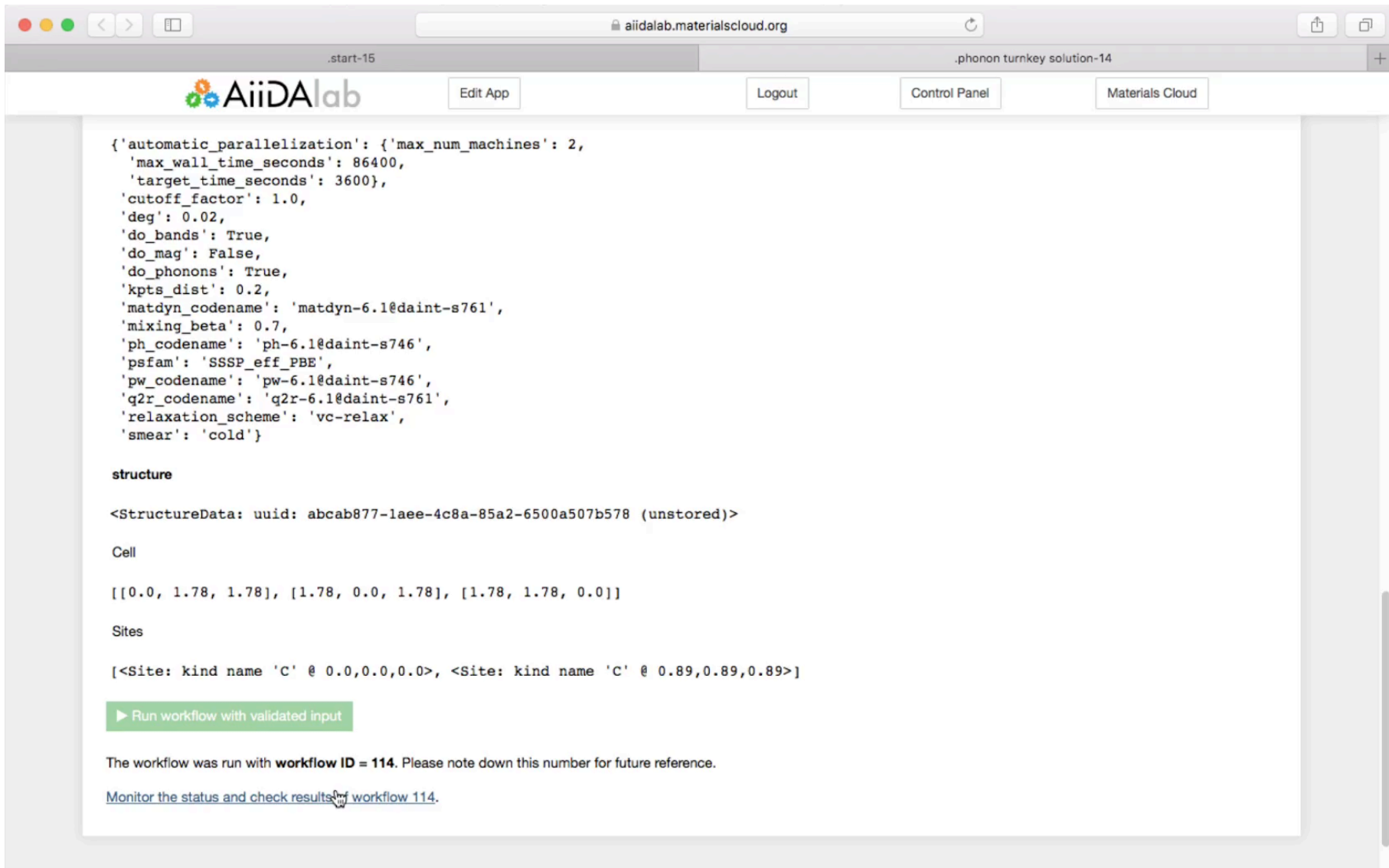
AiiDA Lab demo - phonon submit

The screenshot shows the AiiDA Lab web interface in a browser window. The address bar displays `aiidalab.materialscloud.org`. The interface includes a top navigation bar with the AiiDA Lab logo and buttons for `Edit App`, `Logout`, `Control Panel`, and `Materials Cloud`.

The main content area is titled **AiiDA Lab** and features a dashboard with the following sections:

- Home**: Contains four large icons for `File Browser`, `Terminal`, `Tasks`, and `Manage Apps`.
- AiiDA**: A list of links including `Daemon Status`, `Graph Browser`, `Workflows`, `Delete nodes`, and `REST API`.
- Quantum Mobile**: A list of links including `Setup Connection to Quantum-Mobile` and `Setup Codes`.
- LSMO apps**: A grid of links organized by application type:
 - Isotherm**: `Compute one`, `Compute multiple`
 - Charges**: `Compute Charges`
 - Pore analysis**: `Compute Pores`
 - Import your data to AiiDA**: `Import database`, `Plot imported data`

AiiDA Lab demo - phonon results



The screenshot shows the AiiDA Lab web interface in a browser window. The address bar displays `aiidalab.materialscloud.org`. The page has a header with the AiiDA Lab logo and navigation buttons: "Edit App", "Logout", "Control Panel", and "Materials Cloud". The main content area displays a JSON configuration for a workflow, followed by the structure data and a green button to run the workflow.

```
{ 'automatic_parallelization': { 'max_num_machines': 2,
  'max_wall_time_seconds': 86400,
  'target_time_seconds': 3600 },
  'cutoff_factor': 1.0,
  'deg': 0.02,
  'do_bands': True,
  'do_mag': False,
  'do_phonons': True,
  'kpts_dist': 0.2,
  'matdyn_codename': 'matdyn-6.1@daint-s761',
  'mixing_beta': 0.7,
  'ph_codename': 'ph-6.1@daint-s746',
  'psfam': 'SSSP_eff_PBE',
  'pw_codename': 'pw-6.1@daint-s746',
  'q2r_codename': 'q2r-6.1@daint-s761',
  'relaxation_scheme': 'vc-relax',
  'smear': 'cold' }
```

structure

```
<StructureData: uuid: abcab877-1aee-4c8a-85a2-6500a507b578 (unstored)>
```

Cell

```
[[0.0, 1.78, 1.78], [1.78, 0.0, 1.78], [1.78, 1.78, 0.0]]
```

Sites

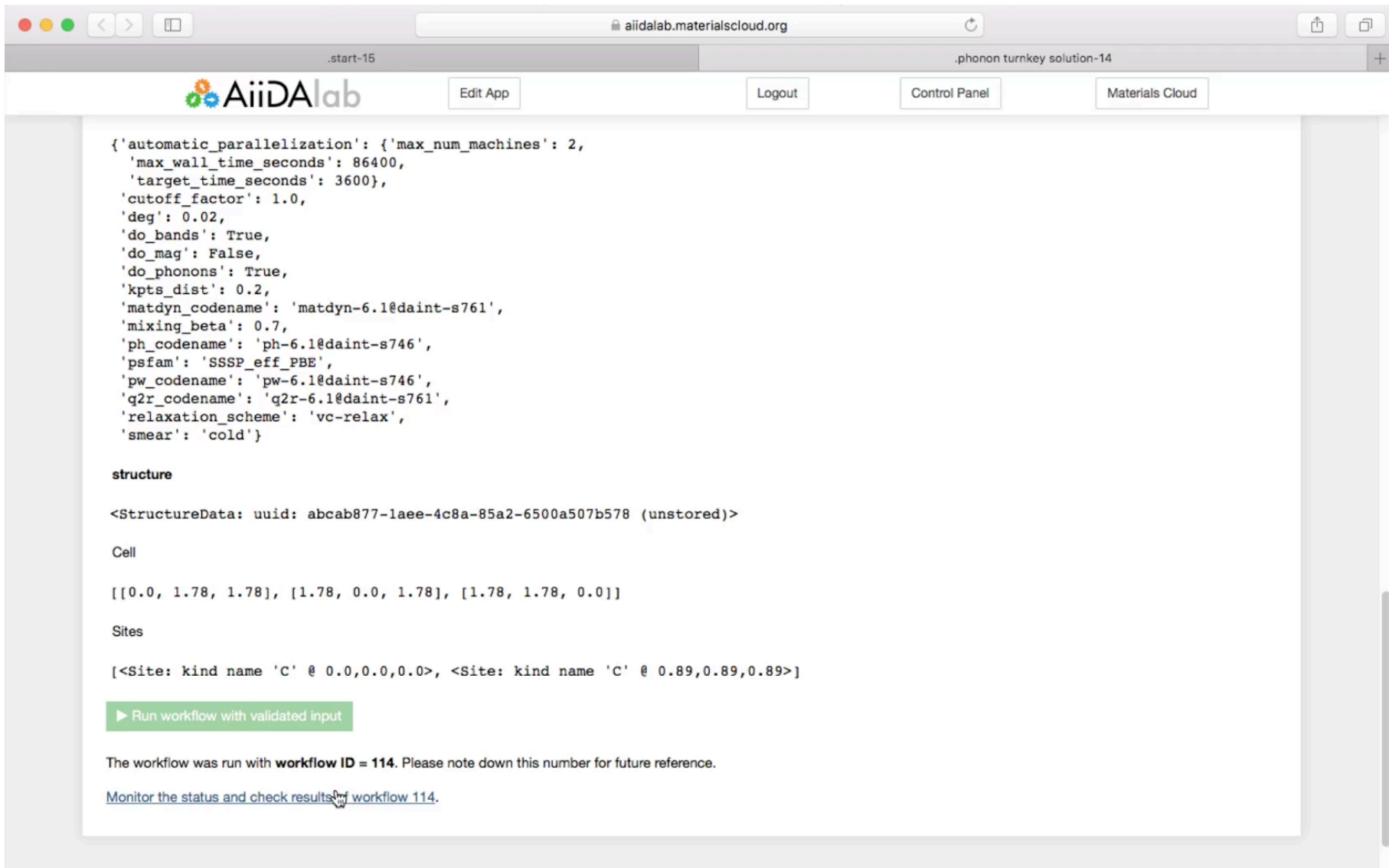
```
[<Site: kind name 'C' @ 0.0,0.0,0.0>, <Site: kind name 'C' @ 0.89,0.89,0.89>]
```

[▶ Run workflow with validated input](#)

The workflow was run with **workflow ID = 114**. Please note down this number for future reference.

[Monitor the status and check results of workflow 114.](#)

AiiDA Lab demo - phonon results



The screenshot shows the AiiDA Lab web interface in a browser window. The address bar displays `aiidalab.materialscloud.org`. The interface includes a navigation bar with the AiiDA Lab logo and buttons for 'Edit App', 'Logout', 'Control Panel', and 'Materials Cloud'. The main content area displays a JSON configuration for a workflow, followed by the structure data and a green button to run the workflow. Below the button, a message states that the workflow was run with ID 114 and provides a link to monitor its status.

```
{ 'automatic_parallelization': { 'max_num_machines': 2,
  'max_wall_time_seconds': 86400,
  'target_time_seconds': 3600 },
  'cutoff_factor': 1.0,
  'deg': 0.02,
  'do_bands': True,
  'do_mag': False,
  'do_phonons': True,
  'kpts_dist': 0.2,
  'matdyn_codename': 'matdyn-6.1@daint-s761',
  'mixing_beta': 0.7,
  'ph_codename': 'ph-6.1@daint-s746',
  'psfam': 'SSSP_eff_PBE',
  'pw_codename': 'pw-6.1@daint-s746',
  'q2r_codename': 'q2r-6.1@daint-s761',
  'relaxation_scheme': 'vc-relax',
  'smear': 'cold' }
```

structure

```
<StructureData: uuid: abcab877-1aee-4c8a-85a2-6500a507b578 (unstored)>
```

Cell

```
[[0.0, 1.78, 1.78], [1.78, 0.0, 1.78], [1.78, 1.78, 0.0]]
```

Sites

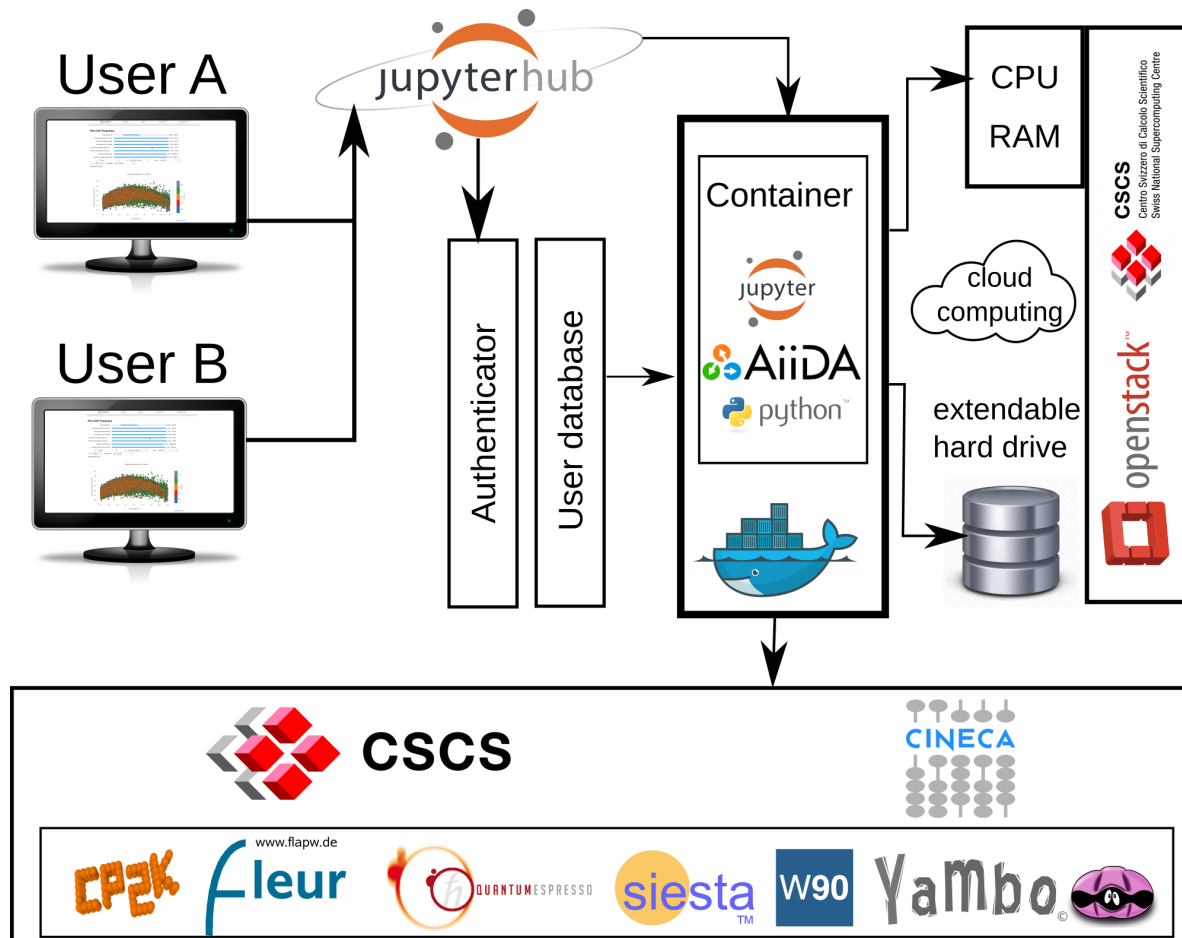
```
[<Site: kind name 'C' @ 0.0,0.0,0.0>, <Site: kind name 'C' @ 0.89,0.89,0.89>]
```

[▶ Run workflow with validated input](#)

The workflow was run with **workflow ID = 114**. Please note down this number for future reference.

[Monitor the status and check results of workflow 114.](#)

AiiDA Lab Architecture



Technologies:

- JupyterHub (multi-user server for Jupyter notebooks)
- Jupyter notebooks (interactive python, + appmode)
- Docker (isolated environment for every user)
- Openstack (software to manage computing clouds)

Conclusions



- Workflows & Daemon help with automation
- AiiDA graph = map of the data jungle
- Your calculations are ready for Open Science



- FAIR sharing of data
- Free: 5 GB/record
(50 GB for AiiDA graphs)



- write simple apps in python
- Run where you like:
 - on your machine (your machine, your setup)
 - on Quantum Mobile (your machine, our setup)
 - on AiiDA Lab (our machine, our setup)



Acknowledgements

Supported by



MARVEL NCCR
www.nccr-marvel.ch



MaX Centre
www.max-centre.eu



SNSF
<http://www.snf.ch>



EPFL
www.epfl.ch



CSCS
www.cscs.ch



PRACE
www.prace-ri.eu

Partners - Research



NFFA
www.nffa.eu



EMMC
www.emmc.info



MarketPlace project
the-marketplace-project.eu



ERC
erc.europa.eu

Developers

The Materials Cloud
And AiiDA teams



Marco
Borelli
(EPFL)



Sebastiaan
P. Huber
(EPFL)



Boris
Kozinsky
(BOSCH)



Snehal P.
Kumbhar
(EPFL)



Leonid
Kahle
(EPFL)



Nicola
Marzari
(EPFL)



Elsa
Passaro
(EPFL)



Giovanni
Pizzi
(EPFL)



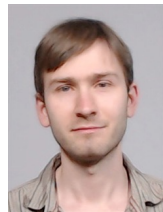
Thomas
Schulthess
(ETHZ,CSCS)



Berend
Smit
(EPFL)



Leopold
Talirz
(EPFL)



Martin
Uhrin
(EPFL)



Joost
VandeVondele
(ETHZ,CSCS)



Casper
Welzel
(EPFL)



Aliaksandr
Yakutovich
(EPFL)



Spyros
Zoupanos
(EPFL)

Contributors for the 23+ plugins: **Quantum ESPRESSO, Wannier90, CP2K, FLEUR, YAMBO, SIESTA, VASP, ...**

Contributors to aiida_core and former AiiDA team members — Valentin Bersier, Jocelyn Boullier, Jens Broeder, Andrea Cepellotti, Fernando Gargiulo, Dominik Gresch, Rico Häuselmann, Eric Hontz, Christoph Koch, Espen Flage-Larsen, Andrius Merkys, Nicolas Mounet, Tiziano Müller, Riccardo Sabatini, Ole Schütt, Phillippe Schwaller

The CSCS support teams

Contacts and info



Website: aiida.net

Docs: aiida-core.readthedocs.io

Git repo: github.com/aiidateam/aiida_core/

Plugin registry: aiidateam.github.io/aiida-registry



facebook.com/aiidateam



[@aiidateam](https://twitter.com/aiidateam)



Materials Cloud: materialscloud.org

- **Archive:** archive.materialscloud.org

- **AiiDA Lab:** aiidalab.materialscloud.org

- **Quantum Mobile:** materialscloud.org/work/quantum-mobile

3 open positions for research software engineers/materials scientists

nccr-marvel.ch

Quantum Mobile Virtual Machine

Terminal window output:

```
(aiida) max@qmobile: ~/codes$ verdi code list
# List of configured codes:
# (use 'verdi code show CODEID' to see the details)
* pk 1 - yambo-4.2.0@localhost
* pk 2 - p2y-4.2.0@localhost
* pk 3 - fleur-0.27-fleur@localhost
* pk 4 - fleur-0.27-mpgen@localhost
* pk 5 - siesta-4.0.1@localhost
* pk 6 - cp2k-5.1@localhost
(aiida) max@qmobile: ~/codes$ verdi calculation list -a
# Last daemon state_updater check: 0h:00m:05s ago (at 15:09:17 on 2017-11-20)
-----
PK      Creation      State      Sched. state      Computer      Type
-----
-----
175      3h ago      FAILED    DONE              localhost     siesta.
siesta
179      3h ago      FINISHED  DONE              localhost     fleur.i
mpgen
195      3h ago      FAILED    DONE              localhost     siesta.
siesta
208      3h ago      FINISHED  DONE              localhost     siesta.
siesta
221      11m ago     FINISHED  DONE              localhost     fleur.i
mpgen
228      11m ago     FINISHED  DONE              localhost     fleur.f
leur
Total results: 6
(aiida) max@qmobile: ~/codes$
```

3D visualization window (MoS2_ML.cif) showing a crystal structure with parameters:

P 1 [P 1]
a=3.188Å
b=3.188Å
c=23.155Å
α=90.0°
β=90.0°
γ=120.0°

Quantum Mobile
MARVEL
MAX

- Ubuntu 16.04 LTS
- **AiiDA, AiiDA Lab**
- **QE, Siesta, fleur, yambo, cp2k, ... + AiiDA plugins**
- Visualization tools (xcrysden, ...)
- used in tutorials & lectures at EPFL, ETH, ...
- **Modular setup: roll your own**

Runs on Linux, MacOS and Windows hosts using VirtualBox

Download: materialscloud.org/work/quantum-mobile