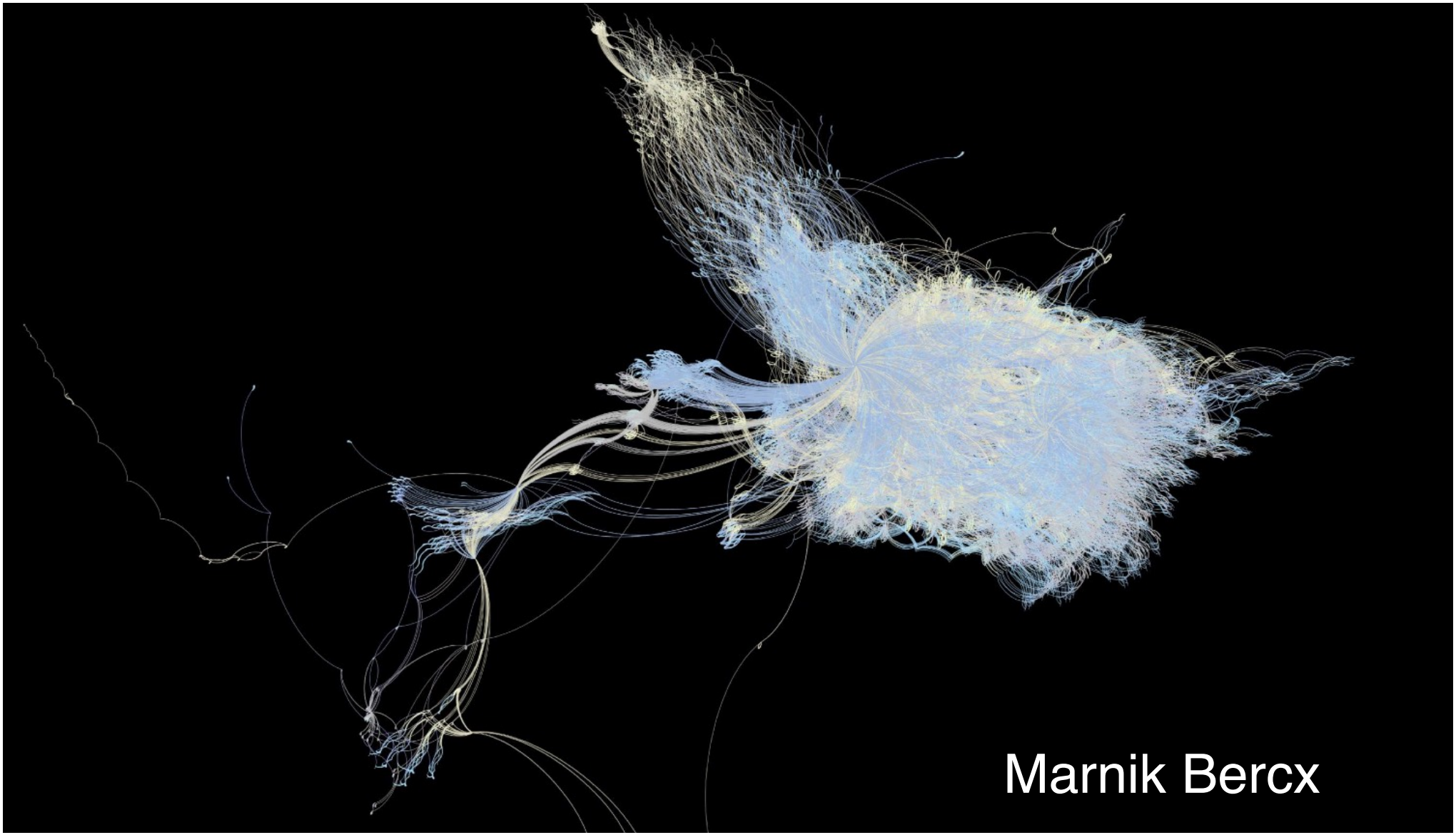
MaX School on Advanced Materials and Molecular Modelling with Quantum ESPRESSO

AiiDA

MATERIALSCLOUD



FNSNF

MARVEL
NATIONAL CENTRE OF COMPETENCE IN RESEARCH

MAX DRIVING THE EXASCALE TRANSITION

swissuniversities

EPFL

Challenges in high-throughput HPC

- **Workflow automation**
  - Need tools to define complex workflows with advanced error handling
  - An automated, robust and scalable engine to run the workflows
- **Data management**
  - Data should be stored reliably and efficiently
  - Stored data should be interoperable and queryable
- **Reproducibility**
  - All produced data should be reproducible by storing the full provenance

AiiDA

- **COMPUTATIONAL SCIENCE INFRASTRUCTURE**
- **FOR HIGH THROUGHPUT WORKFLOWS**
- **WITH FULL DATA PROVENANCE**

Language: implemented and API in python

License: MIT open source **http://www.aiida.net/**

Source: https://github.com/aiidateam/aiida-core

MIT LICENSED

open source initiative

Scalable workflow engine

Automated full data provenance

Built-in support for HPC

Flexible plugin system

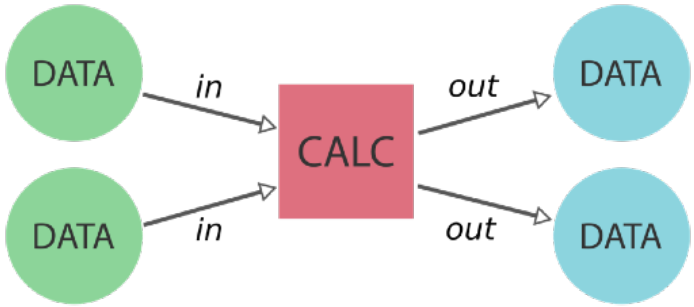
G. Pizzi et al., Comp. Mat. Sci. 111, 218-230 (2016)

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

Data provenance

Simple recipe

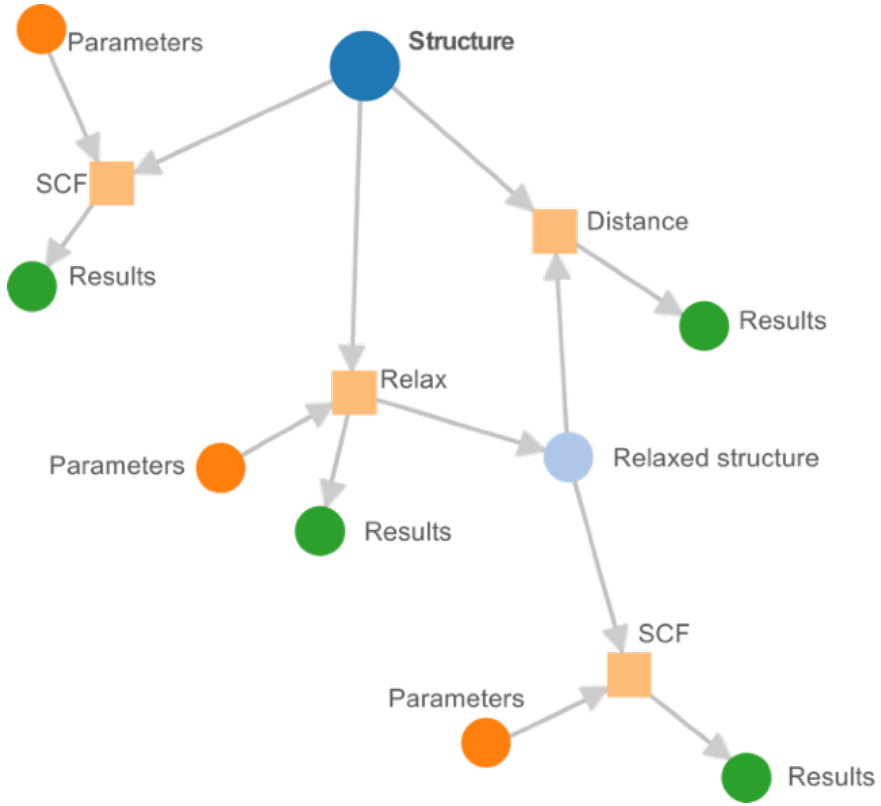
- Store data transformations or '**calculations**'
- Store its **inputs** and their metadata
- Store its **outputs** and their metadata
- Most **crucially** store the **inter-connections**



Data provenance

Simple recipe

- Store data transformations or '**calculations**'
- Store its **inputs** and their metadata
- Store its **outputs** and their metadata
- Most **crucially** store the **inter-connections**



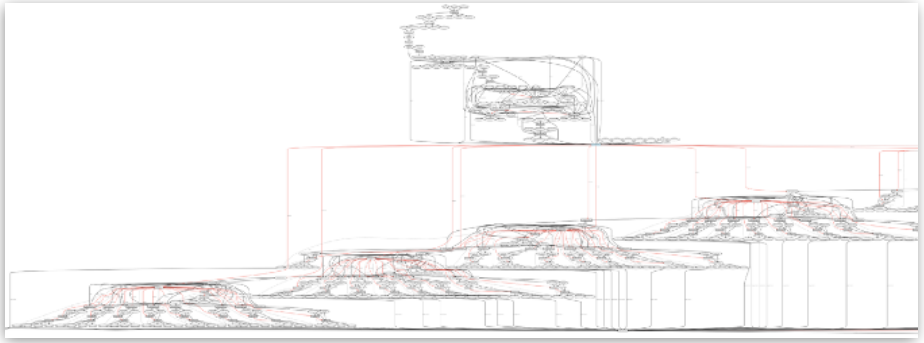
Provenance graphs

- When data gets reused, a directed graph is created
- That quickly grow in complexity even for "simple" workflows

Data provenance

Simple recipe

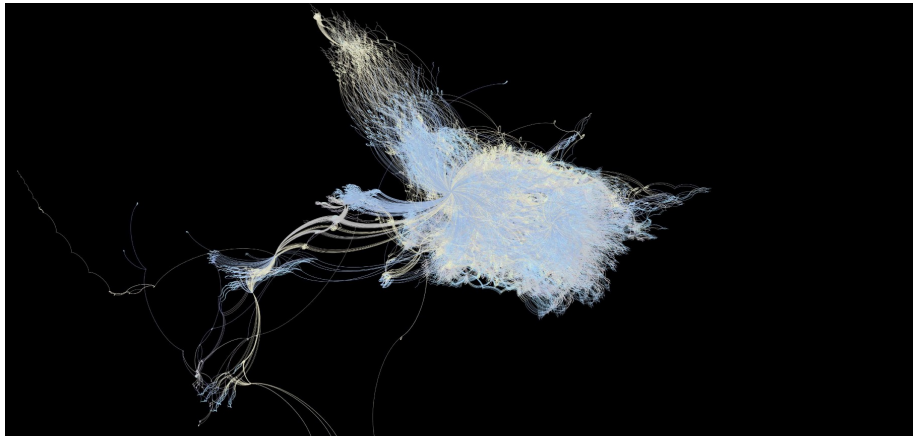
- Store data transformations or '**calculations**'
- Store its **inputs** and their metadata
- Store its **outputs** and their metadata
- Most **crucially** store the **inter-connections**



Molecular dynamics study of Lithium in a solid electrolyte.

Graph requirements

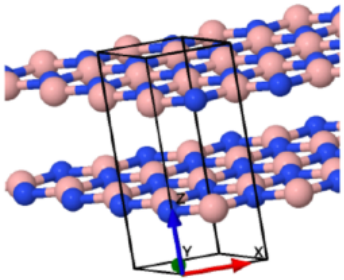
- Needs to be automated
- Needs to be stored *as data is created*

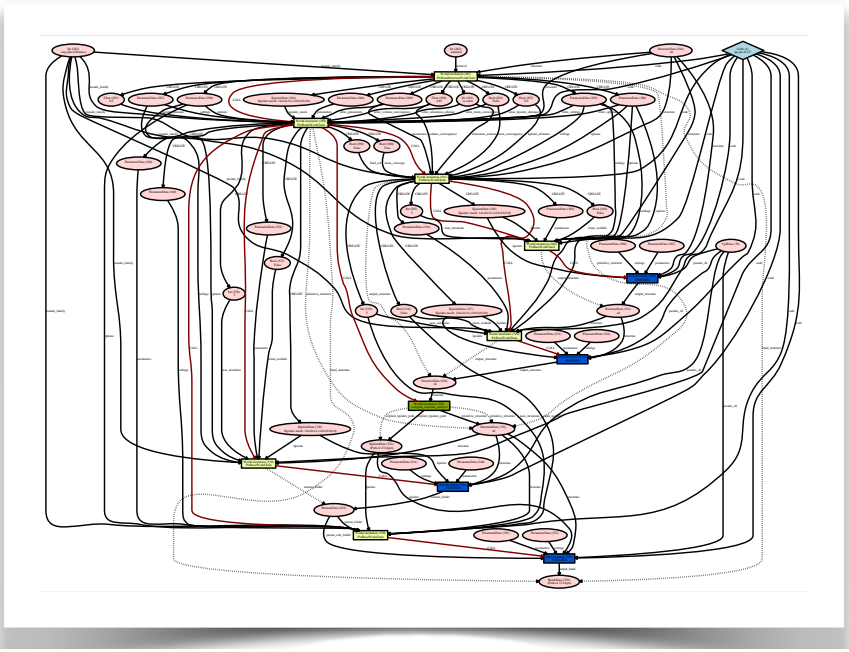


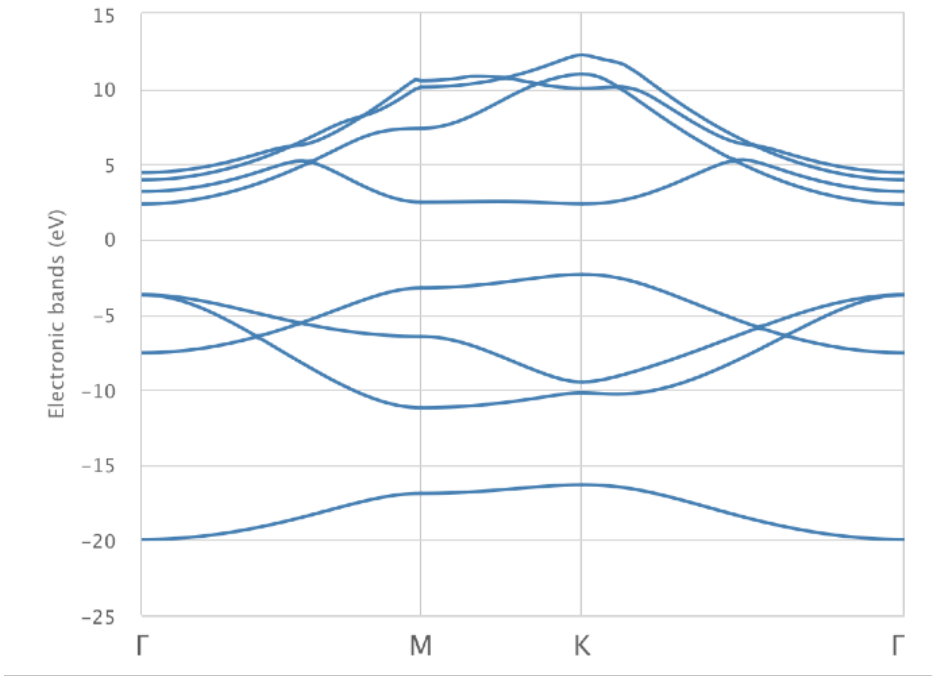
Graphical representation of actual AiiDA database

Turn-key workflows in AiiDA

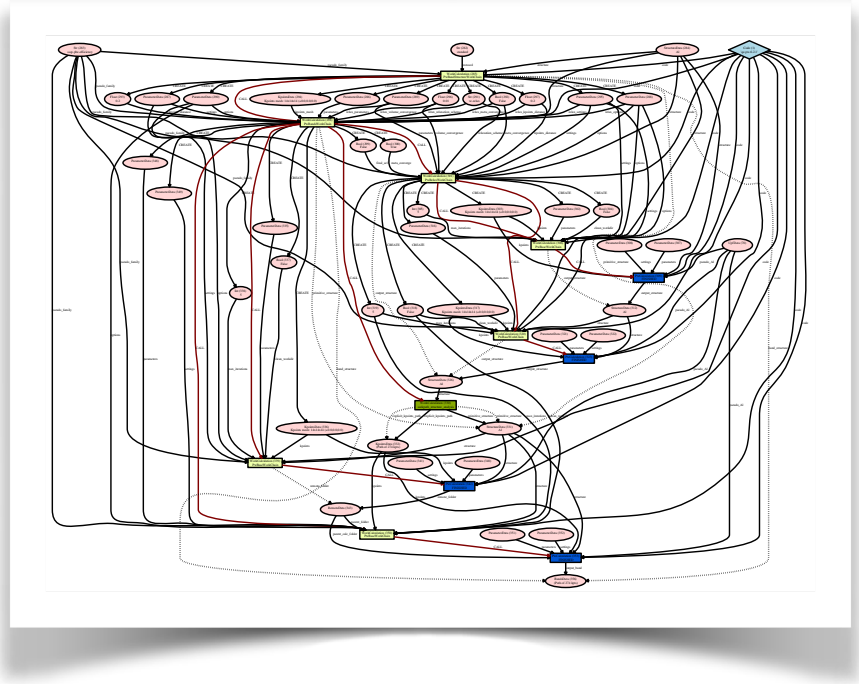
- Given a material, we often need to compute advanced quantities
- These are often non-trivial and result from a complex workflow

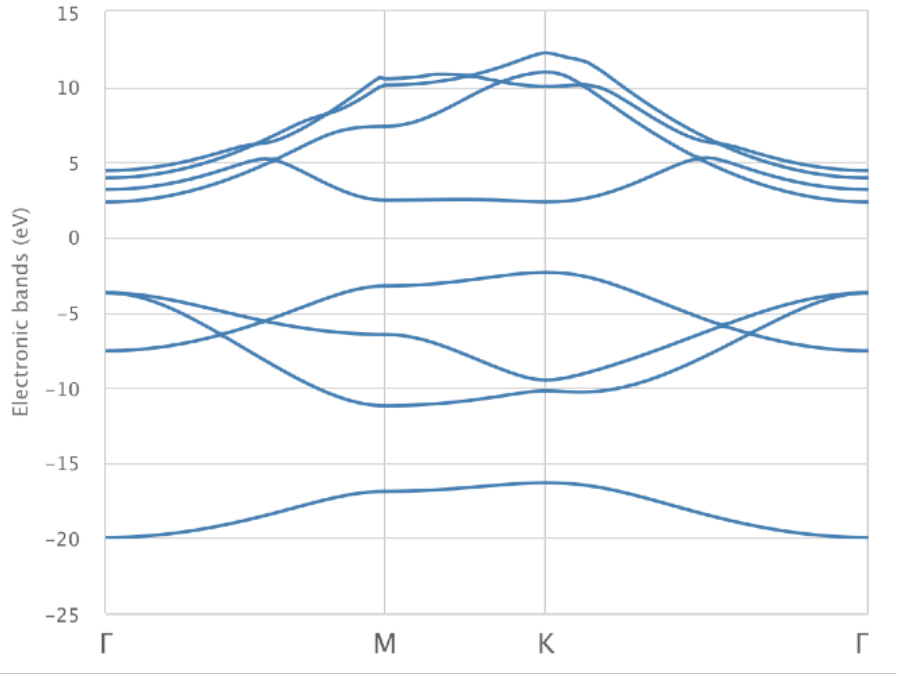






Turn-key workflows in AiiDA





- The AiiDA provenance graph allows to know how the structure was computed and to **reproduce that single specific calculation**:
  log of "what happened in the past"
- We need also an **easy way to re-run the same calculation again** with different parameters or for a different material: **turn-key workflows**

Turn-key workflows in AiiDA

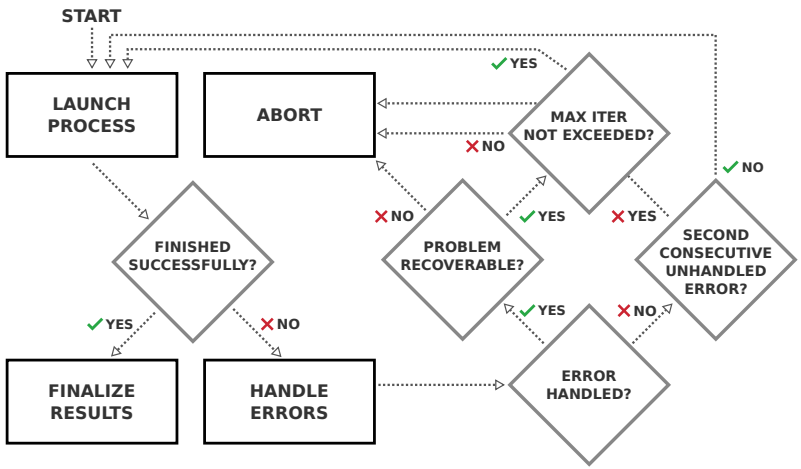
```
class BandsWorkChain(WorkChain):
    @classmethod
    def define(cls, spec):

        spec.input(
           'structure',
           valid_type=StructureData, ...)
        spec.expose_inputs(
                    PwRelaxWorkChain,
           namespace='relax', ...)
        ...

        spec.outline(
            cls.setup,
            if_(cls.should_run_relax)(
                cls.run_relax,
                cls.inspect_relax,
            ),
            if_(cls.should_run_seekpath)(
                cls.run_seekpath,
            ),
            cls.run_scf,
            cls.inspect_scf,
           ...
        )
```

Workflows:

- Encode scientists' knowledge on how to calculate properties
- "Self-documenting" inputs and outputs
- Modular
- Input validation
- Robustness: error handling

Turn-key workflows in AiiDA

Error recovery - Try to "fix" failing calculations by changing the inputs.



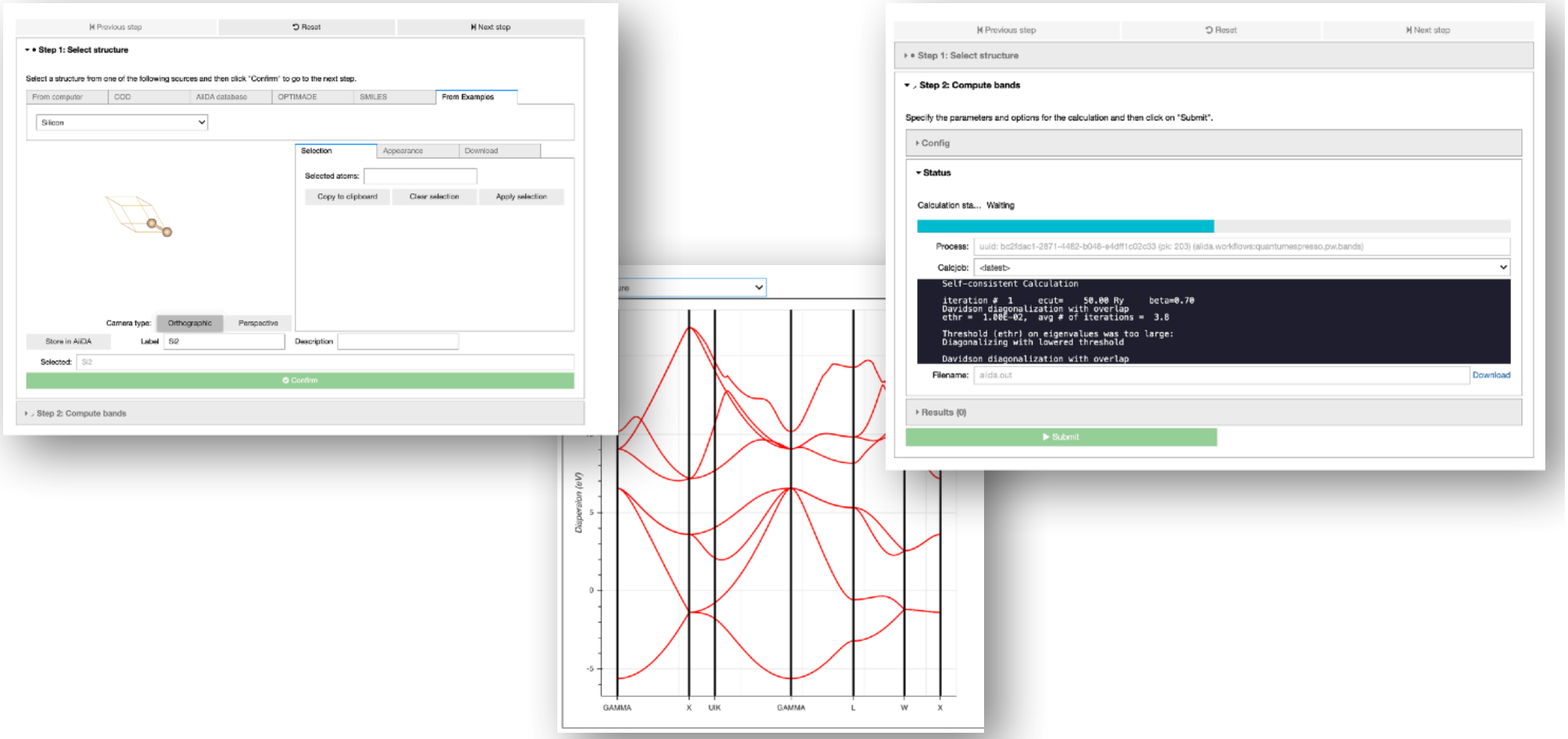
or...

Protocols - Default computation parameters with reasonable precision.

```
In [1]: PwBandsWorkChain = WorkflowFactory('quantumespresso.pw.bands')

In [2]: builder = PwBandsWorkChain.get_builder_from_protocol(
   ...:     code=load_code('qe-v6.7-pw'),
   ...:     structure=load_node(87),
   ...:     protocol='fast'
   ...: )

In [3]: from aiida.engine import submit

In [4]: submit(builder)
Out[4]: <WorkChainNode: uuid: ... (pk: 105) (aiida.workflows:quantumespresso.pw.bands)>
```

Easy access to these simulations on the cloud

- **Easy access to these advanced HPC capabilities to everybody**: AiiDAlab



Sharing in AiiDA: codes, plugins and workflows





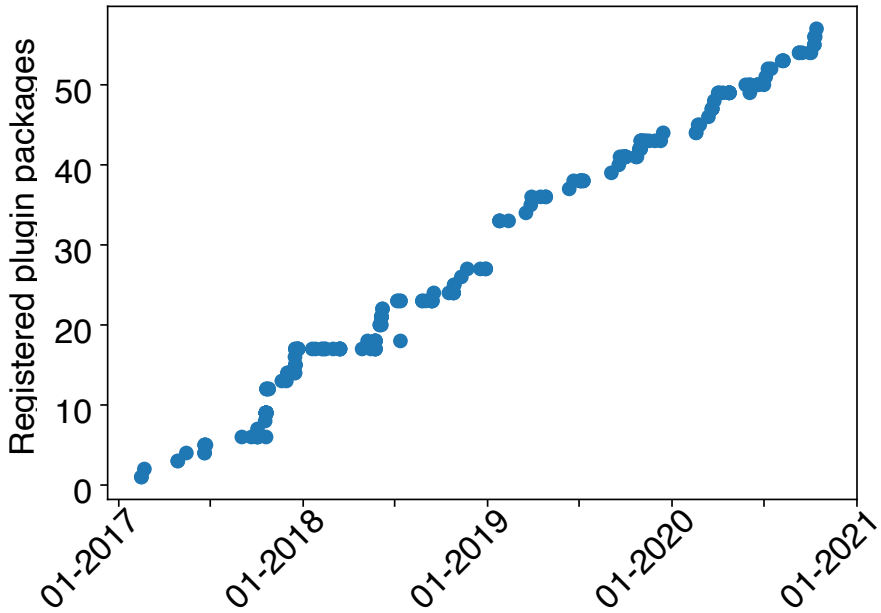

Calculation | **Data** | **Parsers** | **Transport and scheduler** | **Workflows** | **Importers & exporters**

Registered plugin packages: 58

Calculations	98 plugins in 39 packages
Parsers	84 plugins in 40 packages
Data	79 plugins in 24 packages
Workflows	95 plugins in 27 packages
Console scripts	19 plugins in 14 packages
Other	95 plugins in 26 packages

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

- Plugins collected in the AiiDA plugin registry
- **Almost 100 codes currently supported,** >90 workflows
- Many are **community-contributed**



Code interoperability: common workflow interfaces

- **Long term goal**: robust, cross-validated and accessible simulations with unified interface.

Common workflows for computing materials properties using different quantum engines

Sebastiaan P. Huber,[1,*] Emanuele Bosoni,[2] Marnik Bercx,[1] Jens Bröder,[3,4] Augustin Degomme,[5] Vladimir Dikan,[2] Kristjan Eimre,[6] Espen Flage-Larsen,[7] Alberto Garcia,[2] Luigi Genovese,[5] Dominik Gresch,[8] Conrad Johnston,[9] Guido Petretto,[10] Samuel Poncé,[1] Gian-Marco Rignanese,[10] Christopher J. Sewell,[1] Vasily Tseplyaev,[3,4] Martin Uhrin,[1] Aliaksandr V. Yakutovich,[11,1] Austin Zadoks,[1] Pezhman Zarabadi-Poor,[12,13] Bonan Zhu,[14,13] Nicola Marzari,[1] and Giovanni Pizzi[1,†]

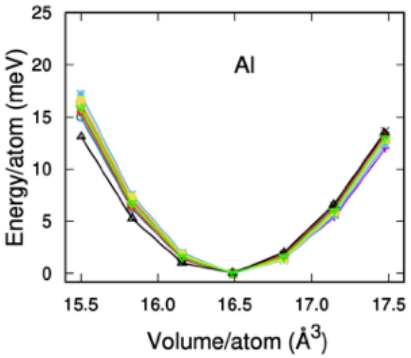
Submitted to npj Computational Materials

https://github.com/aiidateam/aiida-common-workflows/

```
$ aiida-common-workflows launch eos siesta --structure=Al --protocol=precise
```

Implementations for **11 quantum codes**

EOS only for the 9 with PBC; relax and dissociation for all 11

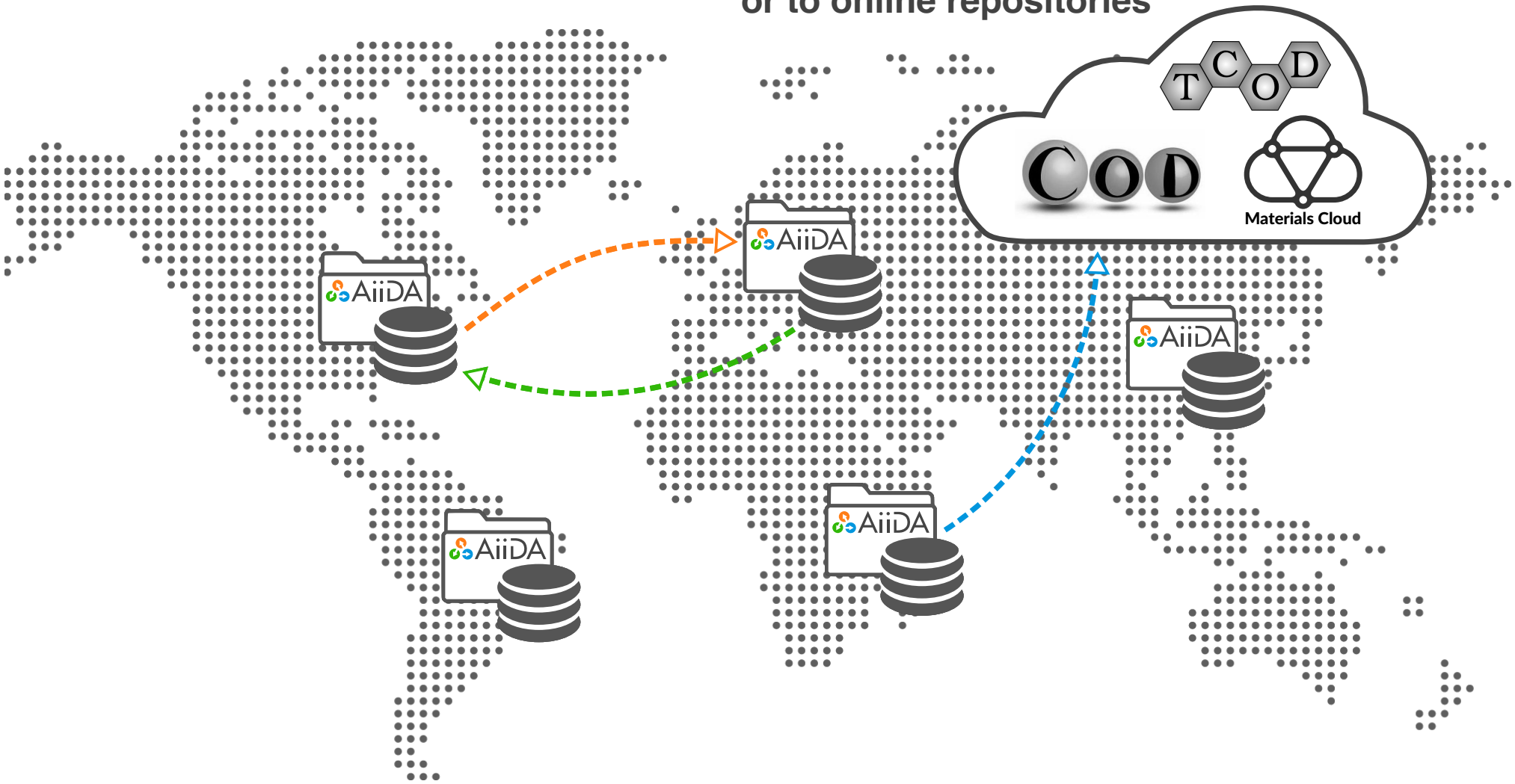


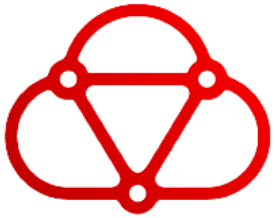
Sharing in AiiDA: data and graphs

- *Private* AiiDA instances
- UUIDs to uniquely identify nodes

- Data can be shared to other AiiDA repositories

or to online repositories





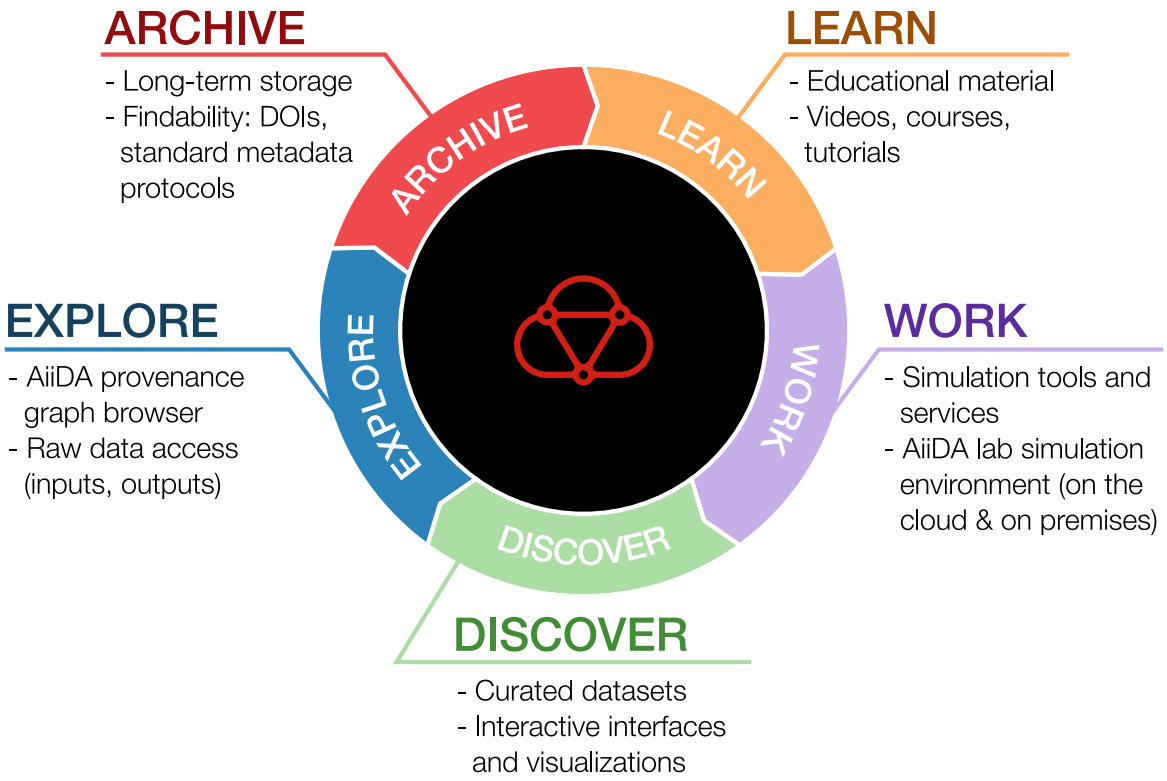
MATERIALSCLOUD

https://www.materialscloud.org

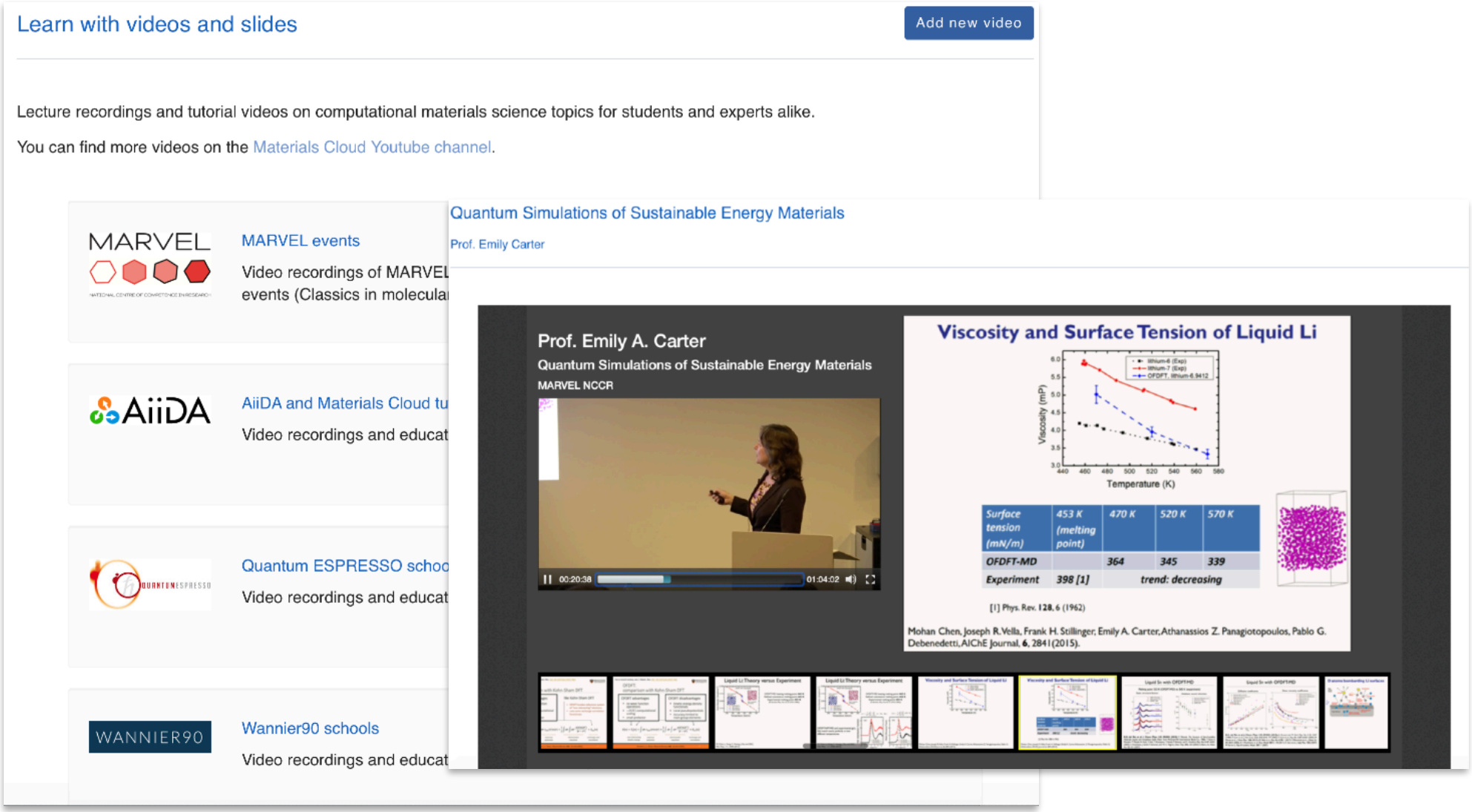
L. Talirz et al., Scientific Data 7, 299 (2020)

Materials Cloud

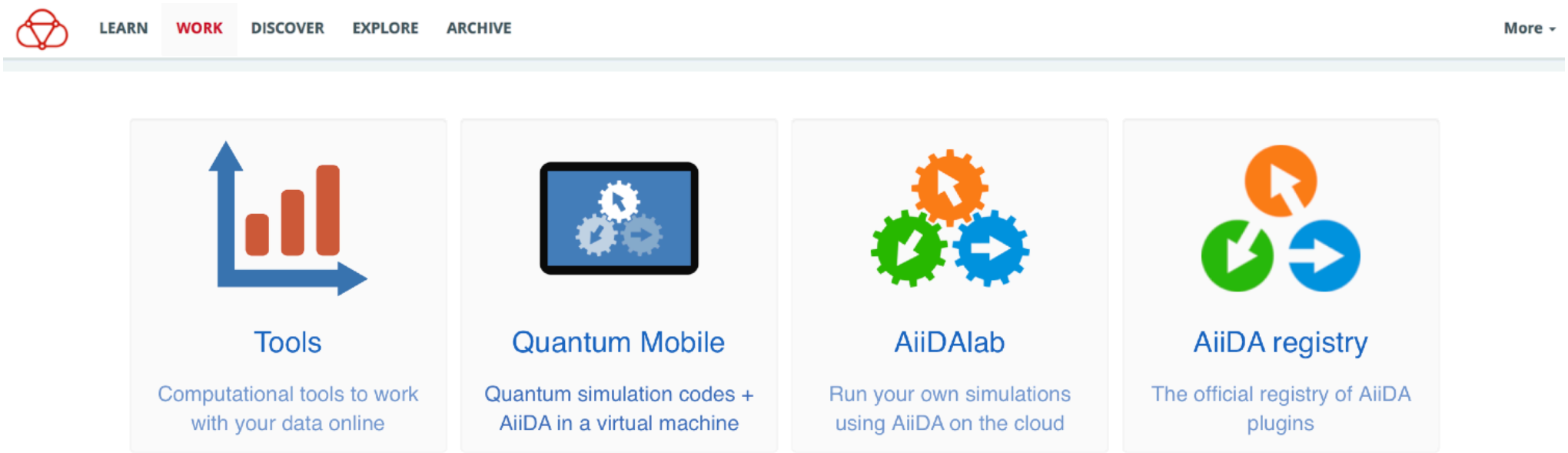
- **AiiDA** is the 'engine', like **Git -** used in production *since 2015*
- **Materials Cloud** is the dissemination platform (like **GitHub**) **and more** (cloud computing and data generation platform) - online since *Dec 2017*



Materials Cloud Learn: Educational platform



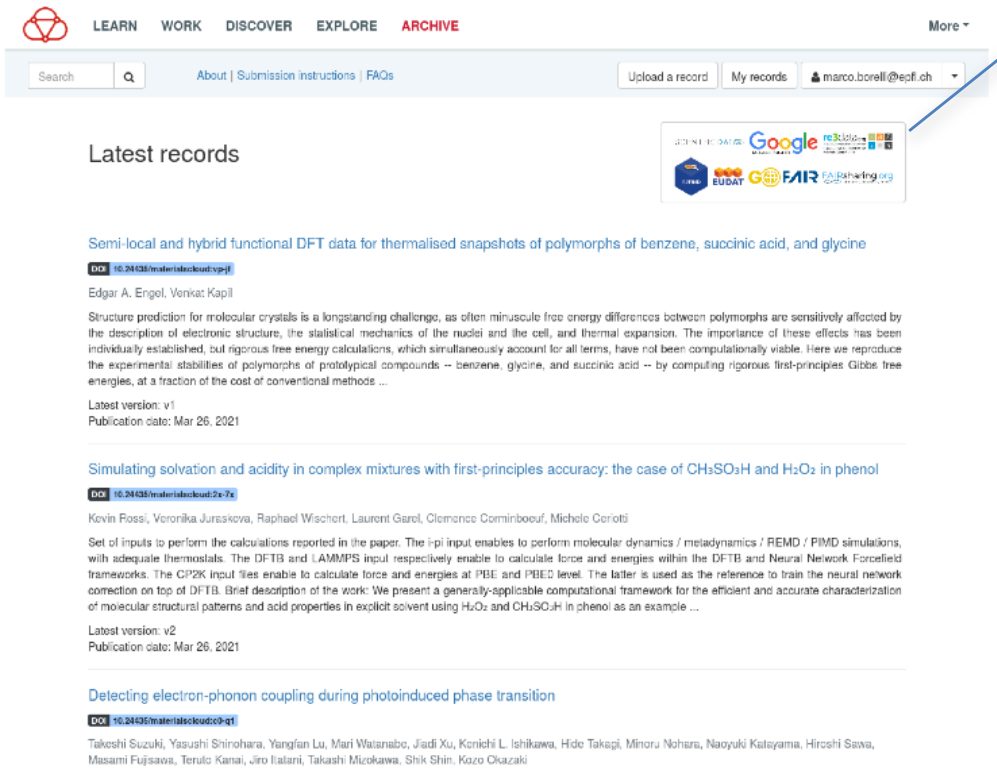
Data generation: Materials Cloud **Work**



Quantum Mobile

- **Downloadable VM** with preinstalled **AiiDA and codes** like QE, Yambo, Fleur, Siesta, CP2K, …
- Includes **same AiiDAlab apps environment** as on Materials Cloud
- Ideal for **education** (courses, tutorials, ...)

Open data sharing: **Archive, Discover, Explore**



Recommended data repository
by Nature's journal Scientific Data

Indexed by Google Dataset Search
and by EUDAT/EOSC's B2FIND

Registered on FAIRsharing.org
and re3data.org

New! Recommended by the new "Open Research Europe" journal

European Commission

Research and Innovation

Open Research Europe

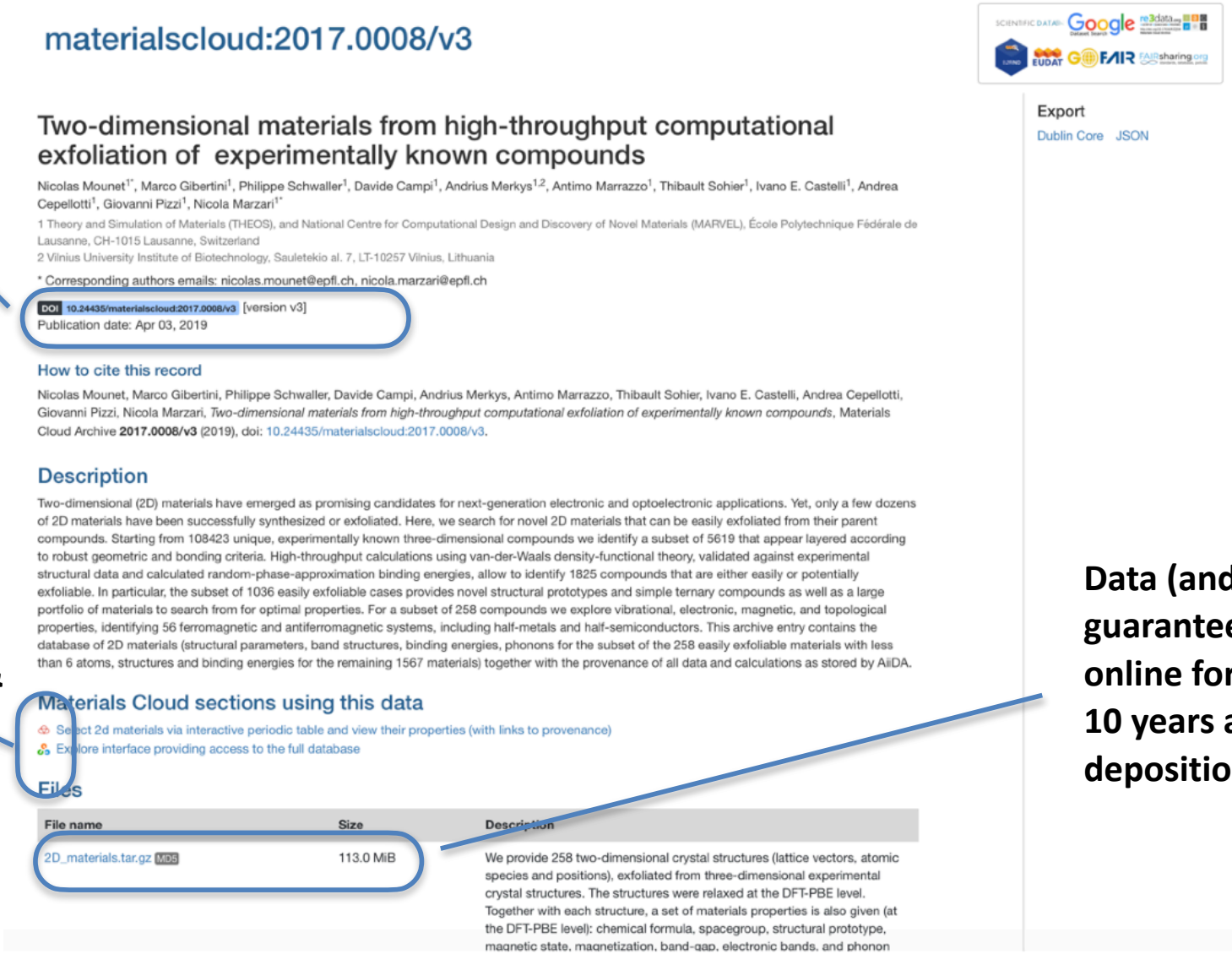
https://open-research-europe.ec.europa.eu/for-authors/data-guidelines

Open data sharing: Archive, Discover, Explore

DOIs assigned

Direct links to Discover & Explore

Data (and metadata) guaranteed to be online for at least 10 years after deposition



materialscloud:2017.0008/v3

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

Nicolas Mounet[1*], Marco Gibertini[1], Philippe Schwaller[1], Davide Campi[1], Andrius Merkys[1,2], Antimo Marrazzo[1], Thibault Sohier[1], Ivano E. Castelli[1], Andrea Cepellotti[1], Giovanni Pizzi[1], 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

DOI 10.24435/materialscloud:2017.0008/v3 [version v3]
Publication date: Apr 03, 2019

Export
Dublin Core JSON

How to cite this record

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*, Materials Cloud Archive **2017.0008/v3** (2019), doi: 10.24435/materialscloud:2017.0008/v3.

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 binding energies, allow to identify 1825 compounds that are either easily or potentially exfoliable. In particular, the subset of 1036 easily exfoliable cases provides novel structural prototypes and simple ternary compounds as well as a large portfolio of materials to search from for optimal properties. For a subset of 258 compounds we explore vibrational, electronic, magnetic, and topological properties, identifying 56 ferromagnetic and antiferromagnetic systems, including half-metals and half-semiconductors. This archive entry contains the database of 2D materials (structural parameters, band structures, binding energies, phonons for the subset of the 258 easily exfoliable materials with less than 6 atoms, structures and binding energies for the remaining 1567 materials) together with the provenance of all data and calculations as stored by AiiDA.

Materials Cloud sections using this data

Select 2d materials via interactive periodic table and view their properties (with links to provenance)
Explore interface providing access to the full database

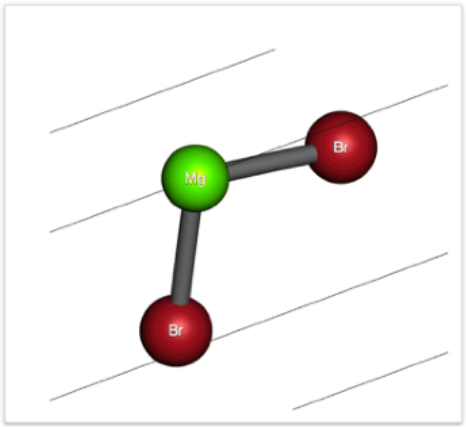
Files

File name	Size	Description
2D_materials.tar.gz MD5	113.0 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

Open data sharing: **Archive, Discover, Explore**

DISCOVER

Compound: $MgBr_2$



Info and properties

See definitions...

Formula: $MgBr_2$
Spacegroup: P-3m1
Pointgroup: -3m
Prototype: CdI2
Band gap [eV]: 4.8

Magnetic properties:
- Magnetic State: non-magnetic
- Tot. Magnetization [μB/cell]: -
- Abs. Magnetization [μB/cell]: -

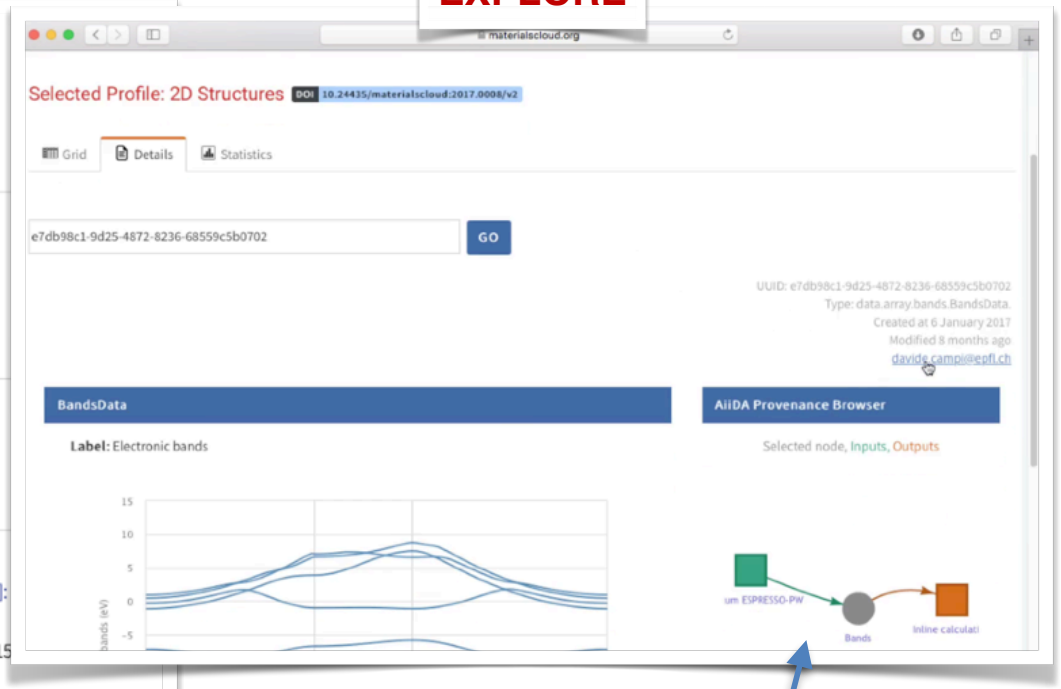
Binding Energies:
- DF2-C09 Binding energy [meV/Å^2]:
  (From parent COD 9009107)
- rVV10 Binding energy [meV/Å^2]: 15
  (From parent COD 9009107)

Delta in interlayer distance (vdW vs revPBE):
- Δ_{DF2} [%]: 17.1 (From parent COD 9009107)
- Δ_{rVV10} [%]: 18.3 (From parent COD 9009107)

Band structure

UUID links to jump to the provenance graph in the EXPLORE section

EXPLORE

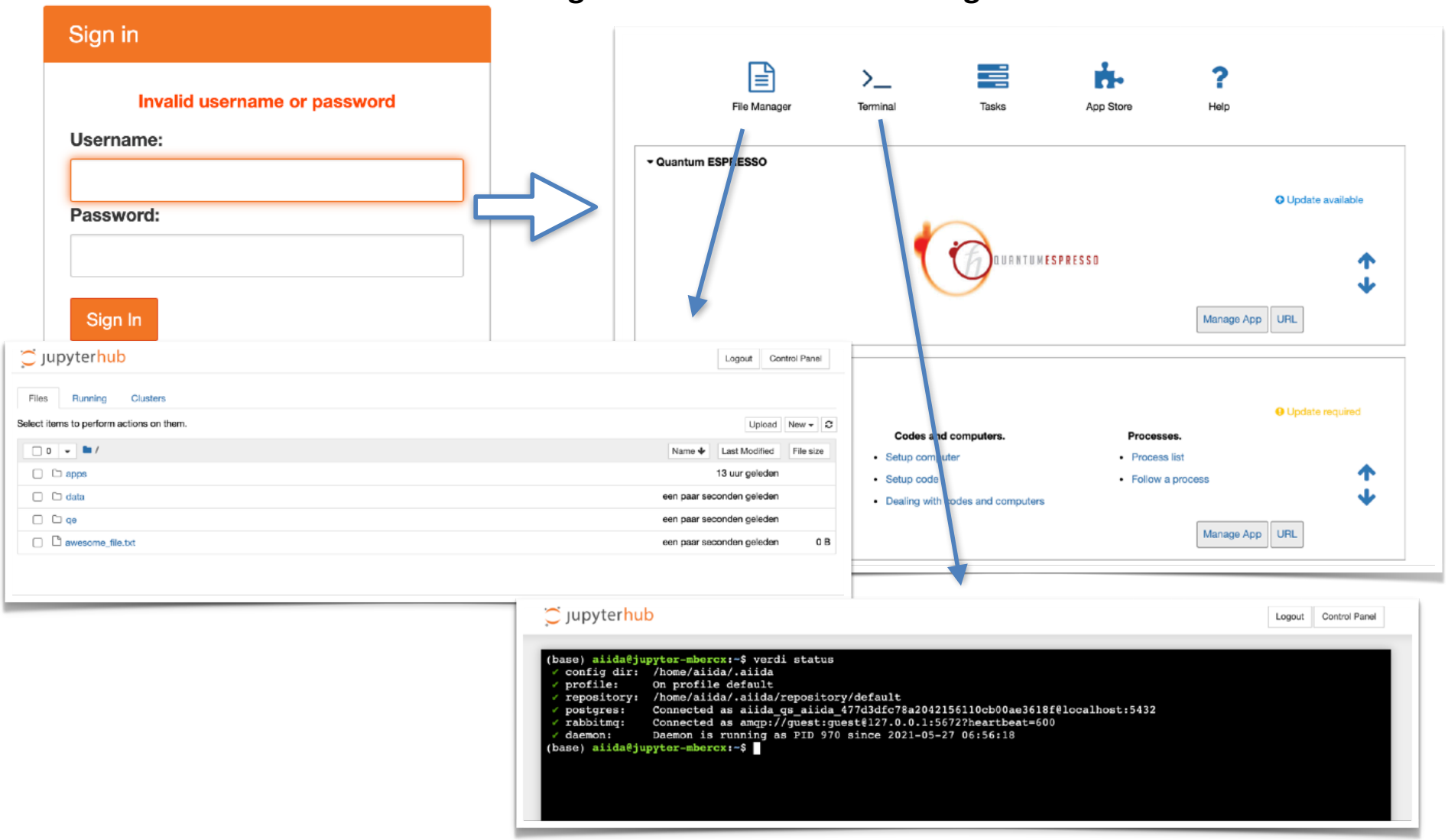


Browse the full AiiDA provenance graph (inputs, outputs, …) at any level

Today's tutorial: Step 1 - Log in

https://qe-school.aiida-tutorials.net/

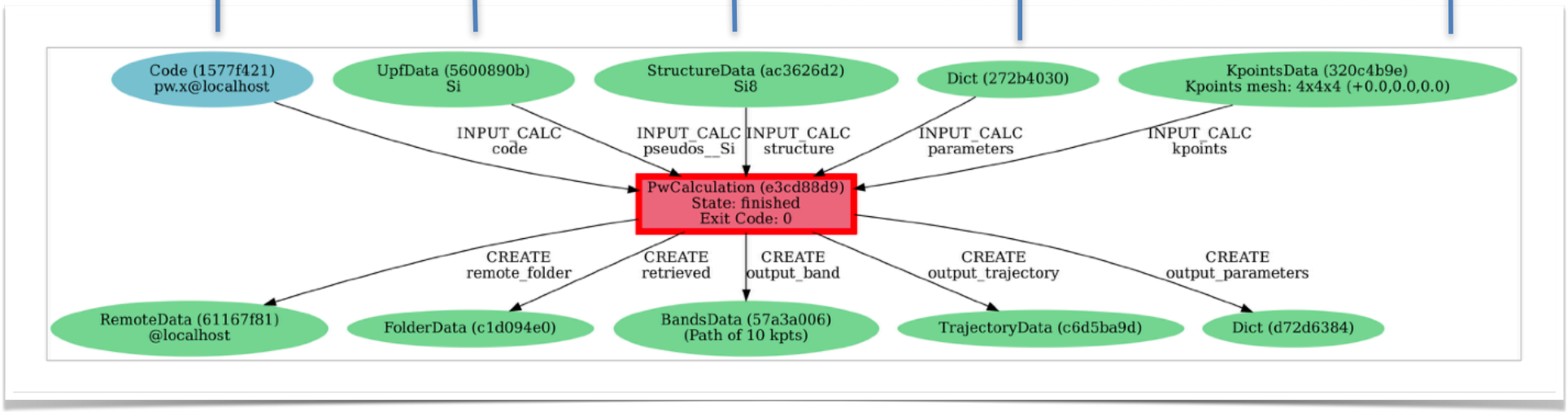
Ignore this erroneous message



Today's tutorial: Step 2 - Calculations

We'll start by running a simple `pw.x` calculation through AiiDA and learn:

- *Importing a structure*
- *Installing pseudos*
- *Setting up a code*
- *Specifying the k-point mesh*
- *Setting up input parameters*



- *Generate a provenance graph for the `pw.x` calculation*
- *Analyse the outputs*

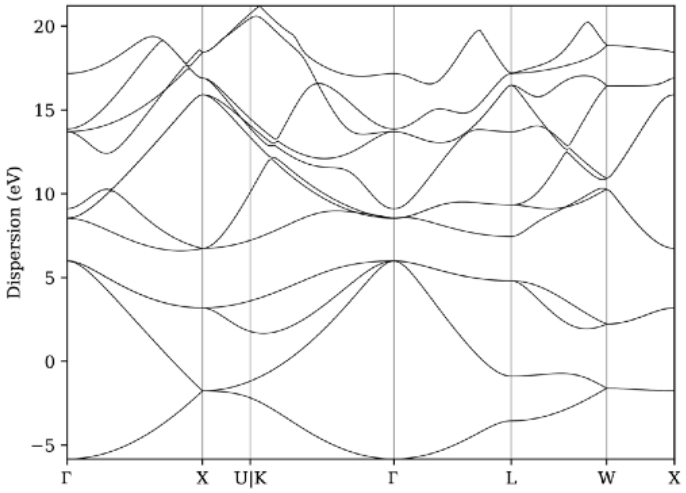
Today's tutorial: Step 3 - Workflows

Next we'll run the `PwBandsWorkChain` to calculate the band structure.

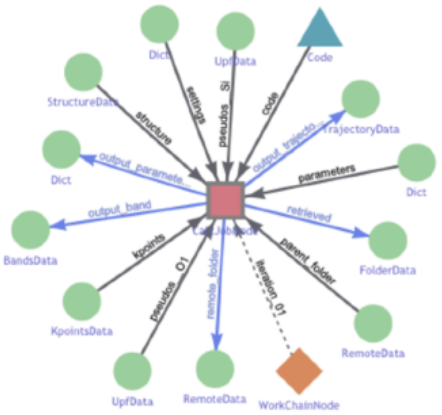
Use the protocol:

```
In [1]: PwBandsWorkChain = WorkflowFactory('quantumespresso.pw.bands')

In [2]: builder = PwBandsWorkChain.get_builder_from_protocol(
   ...:     code=load_code('qe-v6.7-pw'),
   ...:     structure=load_node(87),
   ...:     protocol='fast'
   ...: )

In [3]: from aiida.engine import submit

In [4]: submit(builder)
Out[4]: <WorkChainNode: uuid: ... (pk: 105) (aiida.workflows:quantumespresso.pw.bands)>
```

Get the band structure:



Explore the provenance:

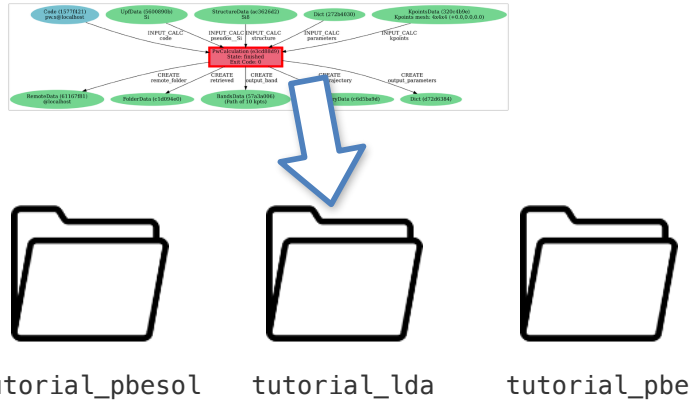


MATERIALSCLOUD

Today's tutorial: Step 4 - Managing data

Finally we'll learn how to manage and query our data!

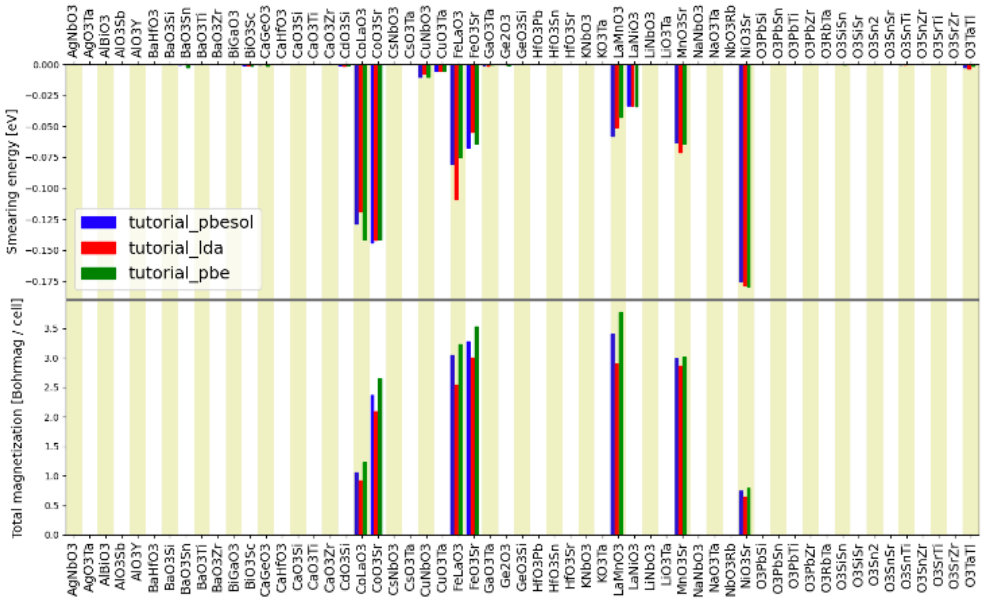
- *Importing data*
- *Organising your data with groups*




```
$ verdi group list --count
Info: to show groups of all types, use the `-a/--all` option.
  PK  Label            Type string    User                 Node count
----  ---------------  -------------  -----------------  ------------
   5  tutorial_pbesol  core           aiida@localhost              57
   6  tutorial_lda     core           aiida@localhost              57
   7  tutorial_pbe     core           aiida@localhost              57
```

- *Using the* `Querybuilder`

```
In [1]: qb = QueryBuilder()

In [2]: qb.append(
   ...:     Group,
   ...:     filters={'label': 'tutorial_pbesol'},
   ...:     tag='group'
   ...: )

In [3]: ...
```



AiiDA and Materials Cloud teams

The Materials Cloud And AiiDA teams


Carl Simon Adorf (EPFL)


Casper W. Andersen (EPFL)


Marco Borelli (EPFL)


Valeria Granata (EPFL)


Sebastiaan P. Huber (EPFL)


Leonid Kahle (EPFL)


Snehal P. Kumbhar (EPFL)


Elsa Passaro (EPFL)


Marnik Bercx (EPFL)


Francisco F. Ramirez (EPFL)


Leopold Talirz (EPFL)


Aliaksandr Yakutovich (EPFL)


Chris Sewell (EPFL)


Giovanni Pizzi (EPFL)


Berend Smit (EPFL)


Joost VandeVondele (ETHZ,CSCS)


Thomas Schulthess (ETHZ,CSCS)


Nicola Marzari (EPFL)

Contributors for the 40+ plugins: Quantum ESPRESSO, Wannier90, CP2K, FLEUR, YAMBO, SIESTA, VASP, CASTEP, CRYSTAL, …

Contributors to aiida-core and former AiiDA team members —
Oscar Arbelaez, Michael Atambo, Valentin Bersier, Marco Borelli, Jocelyn Boullier, Jens Bröder, Ivano E. Castelli, Andrea Cepellotti, Keija Cui, Vladimir Dikan, Marco Dorigo, Y.-W. Fang, Fernando Gargiulo, Marco Gibertini, Davide Grassano, Dominik Gresch, Conrad Johnston, Rico Häuselmann, Daniel Hollas, Eric Hontz, Jianxing Huang, Christoph Koch, Espen Flage-Larsen, Ian Lee, Daniel Marchand, Antimo Marrazzo, Andrius Merkys, Simon Pintarelli, Nicolas Mounet, Tiziano Müller, Gianluca Prandini, Philip Rüßmann, Riccardo Sabatini, Ole Schütt, Phillippe Schwaller, Andreas Stamminger, Atsushi Togo, Daniele Tomerini, Nicola Varini, Martin Uhrin, Jason Yu, Austin Zadoks, Bonan Zhu, Mario Zic, Spyros Zoupanos

Acknowledgements and funding

H2020 Centre of Excellence "MaX"
Scaling towards exascale machines and high-throughput efficiency

SNSF NCCR "MARVEL"
Discovery of new materials via simulations and dissemination of curated data

Swissuniversities P-5 "Materials Cloud"
Scaling the web platform, extending to more disciplines

Moreover:

BIG MAP