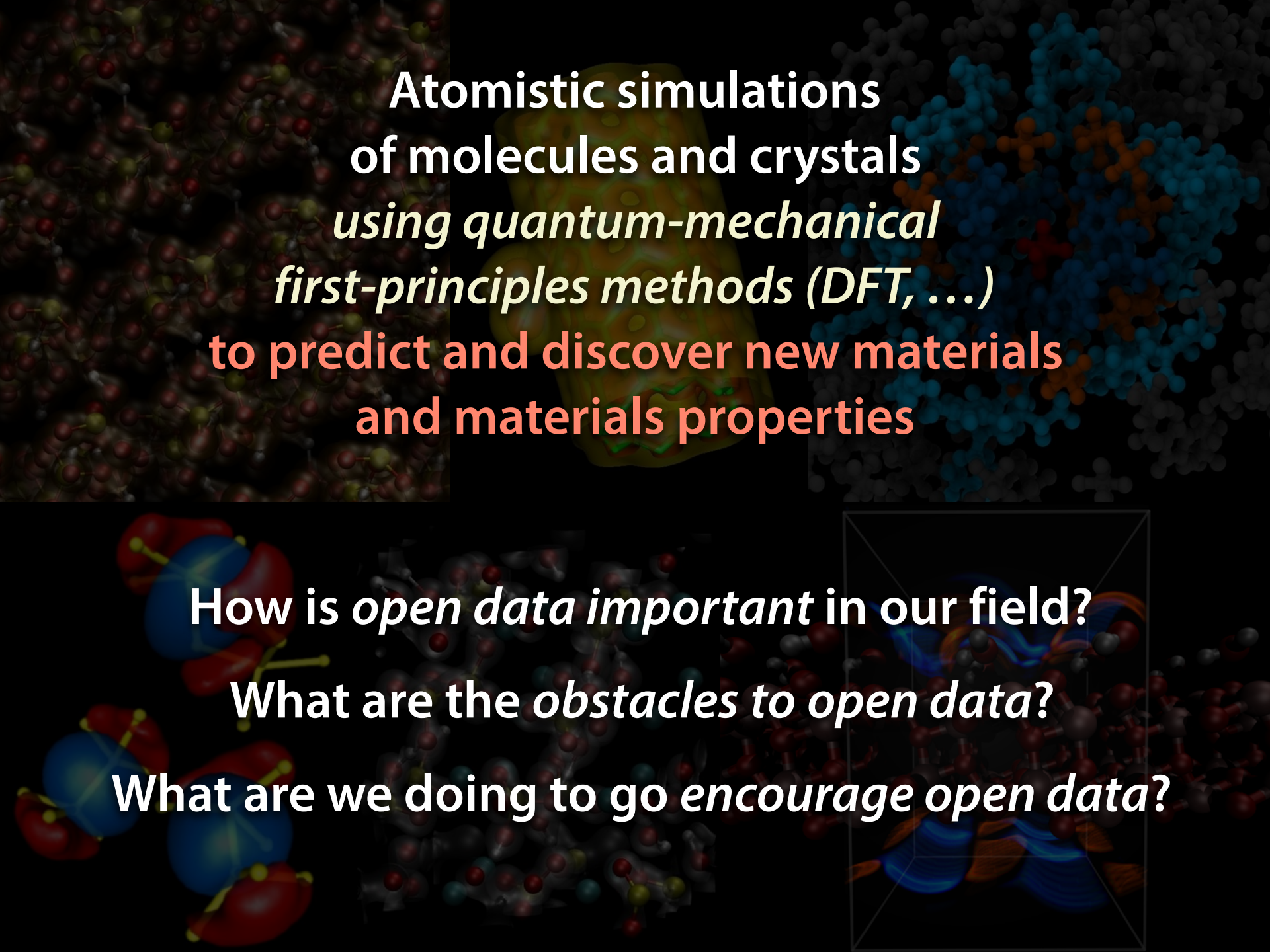
Discovering novel materials and sharing materials data: Benefits and challenges

Giovanni Pizzi
Theory and Simulation of Materials, EPFL, Switzerland

THEOS
THEORY AND SIMULATION OF MATERIALS

MARVEL

EPFL
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE



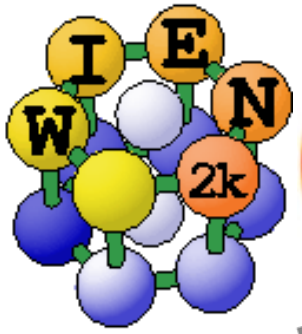
Atomistic simulations
of molecules and crystals
using quantum-mechanical
first-principles methods (DFT, …)
to predict and discover new materials
and materials properties

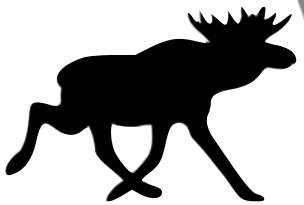
How is *open data important* in our field?

What are the *obstacles to open data*?

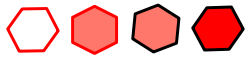
What are we doing to go *encourage open data*?

Simulation codes

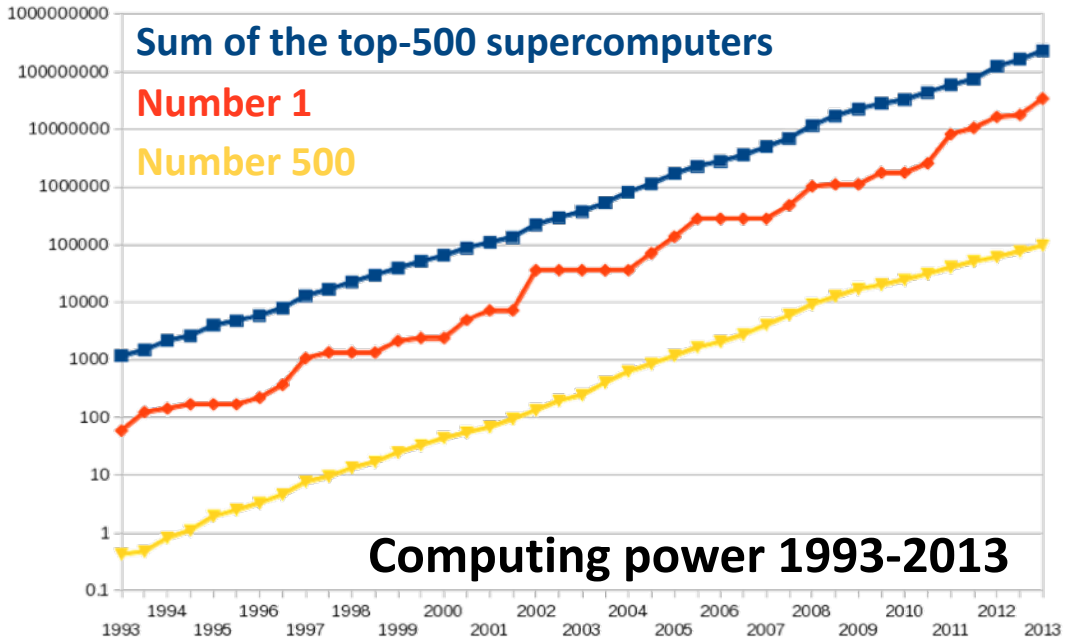




Accuracy and *predictive power*
of quantum engines



Exponential increase in computational power in the past years (“Moore’s law”)



150,000x increase
in the past 20 years

1 month (1993)

↓

10 seconds (2015)

Result: materials design and discovery via
high-throughput computations

...but did we think at how to manage simulations and share the data?

...but did we think at how to manage simulations and share the data?

What happens to the data?

```
0000000 89 50 4e 47 0d 0a 1a 0a 00 00 00 0d 49 48 44 52
0000010 00 00 08 9b 00 00 0c 1f 08 06 00 00 00 8d 23 34
0000020 73 00 00 01 19 69 43 43 50 49 43 43 20 50 72 6f
0000030 66 69 6c 65 00 00 28 91 63 60 60 52 48 2c 28 c8
0000040 61 12 60 60 c8 cd 2b 29 0a 72 77 52 88 88 8c 52
0000050 60 7f c0 c0 c6 c0 cb c0 cf a0 c7 60 96 98 5c 5c
0000060 e0 18 10 e0 c3 00 04 30 1a 15 7c bb c6 c0 08 a2
0000070 2f eb 82 cc c2 94 c7 0b b8 52 52 8b 93 81 f4 1f
0000080 20 ce 4e 2e 28 2a 61 60 60 cc 00 b2 95 cb 4b 0a
0000090 40 ec 1e 20 5b 24 29 1b cc 5e 00 62 17 01 1d 08
00000a0 64 6f 01 b1 d3 21 ec 13 60 35 10 f6 1d b0 9a 90
00000b0 20 67 20 fb 03 90 cd 97 04 66 33 81 ec e2 4b 87
00000c0 b0 05 40 6c a8 bd 20 20 e8 98 92 9f 94 aa 00 f2
00000d0 bd 86 a1 a5 a5 85 26 89 7e 20 08 4a 52 2b 4a 40
00000e0 b4 73 7e 41 65 51 66 7a 46 89 82 23 30 a4 52 15
00000f0 3c f3 92 f5 74 14 8c 0c 0c cd 18 18 40 e1 0e 51
0000100 fd 39 10 1c 9e 8c 62 67 10 62 08 80 10 9b 23 c1
0000110 c0 e0 bf 94 81 81 e5 0f 42 cc a4 97 81 61 81 0e
0000120 03 03 ff 54 84 98 9a 21 03 83 80 3e 03 c3 be 39
0000130 c9 a5 45 65 50 63 18 99 8c 19 18 08 f1 01 69 f3
0000140 4a 7d e7 76 61 f4 00 00 00 09 70 48 59 73 00 00
0000150 2e 23 00 00 2e 23 01 78 a5 3f 76 00 00 40 00 49
0000160 44 41 54 78 01 ec dd 0d 7c 55 57 9d ef ff af be
0000170 a0 6d 3a c6 5e ec 1c c7 46 1b 8a b4 06 70 c2 4b
0000180 d3 48 bc e4 5f 09 8a 86 44 6e a6 85 b4 43 34 03
0000190 53 1a 3a fc 1b e8 58 43 7d 48 a5 46 b9 cd f4 81
00001a0 d8 3a 90 5e 5e 92 d2 a1 c6 c2 94 80 de dc c8 83
00001b0 d2 21 d8 49 a7 a1 69 ec 10 85 60 1b 69 a3 46 6d
00001c0 d4 dc 9a fe 1b 2b e7 f5 9a ff 3a cf 7b ef f3 90
00001d0 73 92 73 f2 70 f2 39 af 17 3d fb 61 ed b5 d7 7a
00001e0 af bd f7 39 e9 fa 9d b5 de f6 5f e6 25 5e 08 20
00001f0 80 00 02 08 20 80 00 02 08 20 80 00 02 08 20 80
0000200 00 02 08 20 80 00 02 08 20 80 00 02 08 20 80 00
0000210 02 08 20 80 00 02 08 c4 21 f0 f6 38 d2 90 04 01
0000220 04 10 40 00 01 04 10 40 00 01 04 10 40 00 01 04
0000230 10 40 00 01 04 10 40 00 01 04 10 40 00 01 04 10
0000240 40 00 01 04 10 40 00 01 af 00 c1 26 5c 08 08 20
0000250 80 00 02 08 20 80 00 02 08 20 80 00 02 08 20 80
0000260 00 02 08 20 80 00 02 08 20 80 00 02 08 20 80 00
0000270 02 08 20 80 00 02 71 0b 10 6c 12 37 15 09 11 40
0000280 00 01 04 10 40 00 01 04 10 40 00 01 04 10 40 00
0000290 01 04 10 40 00 01 04 10 40 00 01 04 10 40 00 01
00002a0 04 10 40 00 01 04 08 36 e1 1a 40 00 01 04 10 40
00002b0 00 01 04 10 40 00 01 04 10 40 00 01 04 10 40 00
00002c0 01 04 10 40 00 01 04 10 40 00 01 04 10 40 00 01
00002d0 04 10 88 5b 80 60 93 b8 a9 48 88 00 02 08 20 80
00002e0 00 02 08 20 80 00 02 08 20 80 00 02 08 20 80 00
00002f0 02 08 20 80 00 02 08 20 80 00 02 08 20 80 00 02
```

PHYSICAL REVIEW B **69**, 014304 (2004)

Intrinsic lifetimes and anharmonic frequency shifts of long-wavelength optical phonons in polar crystals

G. Deinzer, M. Schmitt, A. P. Mayer, and D. Strauch
Institut für Theoretische Physik, Universität Regensburg, D-93040 Regensburg, Germany
(Received 13 June 2003; revised manuscript received 14 October 2003; published 30 January 2004)

Quantitative calculations of phonon lifetimes due to anharmonic three-phonon processes require knowledge of cubic anharmonic coupling coefficients. In order to determine the temperature dependence of phonon frequencies, anharmonic force constants of up to fourth order are needed. In polar crystals, the macroscopic electric field gives rise to nonanalytic terms in these coefficients. It is shown how these non-analytic terms can be determined from other physical quantities including higher-order dipole moments, Raman coefficients, and nonlinear susceptibilities. The contribution of these terms to the intrinsic damping of the long-wavelength optical phonon modes in GaAs has been determined by an *ab initio* calculation.

DOI: 10.1103/PhysRevB.69.014304 PACS number(s): 63.20.Kr, 63.20.−e, 63.10.+a, 78.30.−j

I. INTRODUCTION

In pure insulating crystals, the intrinsic lifetimes of phonons are due to anharmonic multi-phonon processes. If energy and momentum conservation allow for three-phonon processes to occur, these are usually the dominant ones. The finite lifetimes of long-wavelength optical phonons give rise to nonzero widths of the Raman lines and are relevant in other physical situations too. In polar semiconductors, the lifetimes of longitudinal optical (LO) phonons play an important role in the dynamics of highly excited free carriers.[1] Precise knowledge of the decay efficiencies of long-wavelength optical phonons into acoustic phonons is also required for a quantitative interpretation of experiments carried out on nonequilibrium phonon dynamics in GaP.[2]

For covalent semiconductors, the width of the Raman line has been calculated *ab initio* in very good agreement with experiment.[3–6] The cubic anharmonic coupling coefficients needed for these calculations have been determined in two different ways: In their pioneering work, Debernardi *et al.*[3] have applied the "$2n+1$ theorem" by Gonze and Vigneron[7] to determine the cubic coefficients directly in the framework of density-functional perturbation theory (DFPT). Lang *et al.*[5] have combined DFPT with the frozen phonon approach to calculate cubic coupling coefficients as numerical derivatives of dynamical matrices with respect to a frozen-in displacement pattern corresponding to an optical zone-center mode. This method was also used to determine lifetimes of vibrational adlayer modes.[8] In this way, fourth-order anharmonic force constants have also been determined that are needed for a calculation of the temperature dependence of zone-center phonon frequencies.[5,8] Debernardi[6] has combined these two methods in his calculation of fourth-order coupling coefficients by taking numerically a first derivative of third-order force constants determined by DFPT for non-equilibrium configurations.

In the case of polar crystals, both approaches have to be modified to account for nonanalytic terms that occur in the anharmonic coupling coefficients due to the long-range Coulomb interactions giving rise to macroscopic electric fields. They are analogous to the term in the dynamical matrices of polar crystals that leads to the Lyddane-Sachs-Teller splitting.[9] In the following section, we derive the form of these nonanalytic terms and relate them to other physical quantities. In the case of cubic anharmonic coupling coefficients, these are the dielectric tensor, the second-order nonlinear susceptibility tensor, the first-order and second-order dipole moments, and the first-order Raman tensor. Once these quantities are known, the cubic coupling coefficients relevant for the lifetimes of long-wavelength LO and TO phonons may be calculated using either of the two approaches. In Sec. III, we present *ab initio* data for the difference of LO and TO damping functions in GaAs. The influence of the macroscopic field on the temperature dependence of the zone-center optical frequencies is briefly discussed in Sec. IV, which is followed by concluding remarks.

II. CUBIC ANHARMONIC COUPLING COEFFICIENTS

The dynamical matrix of a polar crystal contains a contribution from the long-range Coulomb interaction of the ions, which is not analytic as a function of the wave vector $\mathbf{q}$ at $\mathbf{q}=\mathbf{0}$ (Ref. 9). As a consequence of this contribution, the limit $\mathbf{q}\to\mathbf{0}$ is not unique and depends on the direction from which the center of the Brillouin zone is approached. Analogous terms exist in the Fourier-transformed anharmonic force constants. They are the subject of this work.

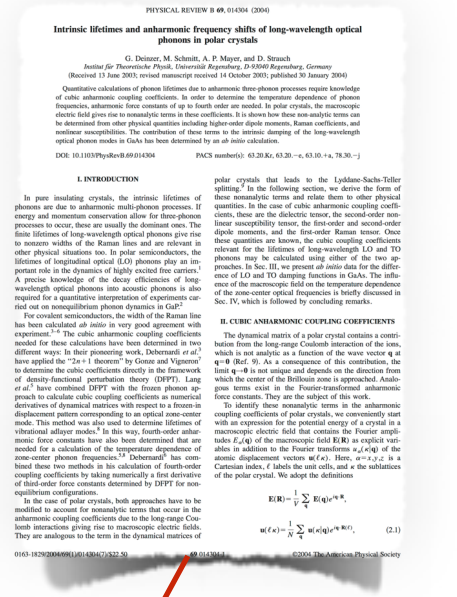
To identify these nonanalytic terms in the anharmonic coupling coefficients of polar crystals, we conveniently start with an expression for the potential energy of a crystal in a macroscopic electric field that contains the Fourier amplitudes $E_\alpha(\mathbf{q})$ of the macroscopic field $\mathbf{E}(\mathbf{R})$ as explicit variables in addition to the Fourier transforms $u_\alpha(\kappa|\mathbf{q})$ of the atomic displacement vectors $\mathbf{u}(\ell\kappa)$. Here, $\alpha=x,y,z$ is a Cartesian index, ℓ labels the unit cells, and κ the sublattices of the polar crystal. We adopt the definitions

$$\mathbf{E}(\mathbf{R})=\frac{1}{V}\sum_{\mathbf{q}}\mathbf{E}(\mathbf{q})e^{i\mathbf{q}\cdot\mathbf{R}},$$

$$\mathbf{u}(\ell\kappa)=\frac{1}{N}\sum_{\mathbf{q}}\mathbf{u}(\kappa|\mathbf{q})e^{i\mathbf{q}\cdot\mathbf{R}(\ell)}, \qquad (2.1)$$

5-star open data

(5stardata.info)



PHYSICAL REVIEW B 69, 014304 (2004)

Intrinsic lifetimes and anharmonic frequency shifts of long-wavelength optical phonons in polar crystals

G. Deinzer, M. Schmitt, A. P. Mayer, and D. Strauch

Institut für Theoretische Physik, Universität Regensburg, D-93040 Regensburg, Germany

(Received 13 June 2003; revised manuscript received 14 October 2003; published 30 January 2004)

Quantitative calculations of phonon lifetimes due to anharmonic three-phonon processes require knowledge of cubic anharmonic coupling coefficients. In order to determine the temperature dependence of phonon frequencies, anharmonic force constants of up to fourth order are needed. In polar crystals, the macroscopic electric field gives rise to nonanalytic terms in these coefficients. It is shown how these non-analytic terms can be determined from other physical quantities including higher-order dipole moments, Raman coefficients, and nonlinear susceptibilities. The contribution of these terms to the intrinsic damping of the long-wavelength optical phonon modes in GaAs has been determined by an *ab initio* calculation.

DOE: 10.1103/PhysRevB.69.014304 PACS number(s): 63.20.Kr, 63.20.−e, 63.10.+a, 78.30.−j

I. INTRODUCTION

In pure insulating crystals, the intrinsic lifetimes of phonons are due to anharmonic multi-phonon processes. If energy and momentum conservation allow for three-phonon processes to occur, these are usually the dominant ones. The finite lifetimes of long-wavelength optical phonons give rise to nonzero widths of the Raman lines and are relevant in other physical situations too. In polar semiconductors, the lifetimes of longitudinal optical (LO) phonons play an important role in the dynamics of highly excited free carriers.[1] A precise knowledge of the decay efficiencies of long-wavelength optical phonons into acoustic phonons is also required for a quantitative interpretation of experiments carried out on nonequilibrium phonon dynamics in GaP.[2]

For covalent semiconductors, the width of the Raman line has been calculated *ab initio* in very good agreement with experiment.[3–6] The cubic anharmonic coupling coefficients needed for these calculations have been determined in two different ways: In their pioneering work, Debernardi *et al.*[3] have applied the "2*n*+1 theorem" by Gonze and Vigneron[7] to determine the cubic coefficients directly in the framework of density-functional perturbation theory (DFPT). Lang *et al.*[5] have combined DFPT with the frozen phonon approach to calculate cubic coupling coefficients as numerical derivatives of dynamical matrices with respect to a frozen-in displacement pattern corresponding to an optical zone-center mode. This method was also used to determine lifetimes of vibrational adlayer modes.[8] In this way, fourth-order anharmonic force constants have also been determined that are needed for a calculation of the temperature dependence of zone-center phonon frequencies.[5,6] Debernardi[6] has combined these two methods in his calculation of fourth-order coupling coefficients by taking numerically a first derivative of third-order force constants determined by DFPT for nonequilibrium configurations.

In the case of polar crystals, both approaches have to be modified to account for nonanalytic terms that occur in the anharmonic coupling coefficients due to the long-range Coulomb interactions giving rise to macroscopic electric fields. They are analogous to the term in the dynamical matrices of polar crystals that leads to the Lyddane-Sachs-Teller splitting.[2] In the following section, we derive the form of these nonanalytic terms and relate them to other physical quantities. In the case of cubic anharmonic coupling coefficients, these are the dielectric tensor, the second-order nonlinear susceptibility tensor, the first-order and second-order dipole moments, and the first-order Raman tensor. Once these quantities are known, the cubic coupling coefficients relevant for the lifetimes of long-wavelength LO and TO phonons may be calculated using either of the two approaches. In Sec. III, we present *ab initio* data for the difference of LO and TO damping functions in GaAs. The influence of the macroscopic field on the temperature dependence of the zone-center optical frequencies is briefly discussed in Sec. IV, which is followed by concluding remarks.

II. CUBIC ANHARMONIC COUPLING COEFFICIENTS

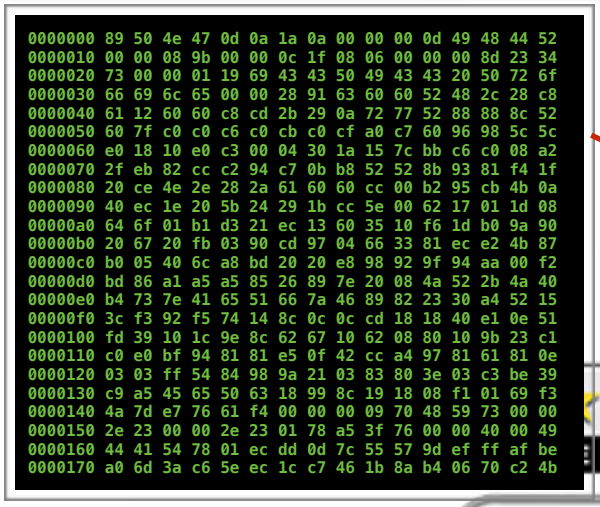
The dynamical matrix of a polar crystal contains a contribution from the long-range Coulomb interaction of the ions, which is not analytic as a function of the wave vector $\mathbf{q}$ at $\mathbf{q}=\mathbf{0}$ (Ref. 9). As a consequence of this contribution, the limit $\mathbf{q}\to\mathbf{0}$ is not unique and depends on the direction from which the center of the Brillouin zone is approached. Analogous terms exist in the Fourier-transformed anharmonic force constants. They are the subject of this work.

To identify these nonanalytic terms in the anharmonic coupling coefficients of polar crystals, we conveniently start with an expression for the potential energy of a crystal in a macroscopic electric field that contains the Fourier amplitudes $E_\alpha(\mathbf{q})$ of the macroscopic field $\mathbf{E}(\mathbf{R})$ as explicit variables in addition to the Fourier transforms $u_\alpha(\kappa|\mathbf{q})$ of the atomic displacement vectors $\mathbf{u}(\ell\kappa)$. Here, $\alpha=x,y,z$ is a Cartesian index, ℓ labels the unit cells, and κ the sublattices of the polar crystal. We adopt the definitions

$$\mathbf{E}(\mathbf{R})=\frac{1}{V}\sum_{\mathbf{q}}\mathbf{E}(\mathbf{q})e^{i\mathbf{q}\cdot\mathbf{R}},$$

$$\mathbf{u}(\ell\kappa)=\frac{1}{N}\sum_{\mathbf{q}}\mathbf{u}(\kappa|\mathbf{q})e^{i\mathbf{q}\cdot\mathbf{R}(\ell)}, \qquad (2.1)$$

0163-1829/2004/69(1)/014304(7)/$22.50 69 014304-1 ©2004 The American Physical Society



```
0000000 89 50 4e 47 0d 0a 1a 0a 00 00 00 0d 49 48 44 52
0000010 00 00 08 9b 00 00 0c 1f 08 06 00 00 00 8d 23 34
0000020 73 00 00 01 19 69 43 43 50 49 43 43 20 50 72 6f
0000030 66 69 6c 65 00 00 28 91 63 60 60 52 48 2c 28 c8
0000040 61 12 60 60 c8 cd 2b 29 0a 72 77 52 88 88 8c 52
0000050 60 7f c0 c0 c6 c0 cb c0 cf a0 c7 60 96 98 5c 5c
0000060 e0 18 10 e0 c3 00 04 30 1a 15 7c bb c6 c0 08 a2
0000070 2f eb 82 cc c2 94 c7 0b b8 52 52 8b 93 81 f4 1f
0000080 20 ce 4e 2e 28 2a 61 60 60 cc 00 b2 95 cb 4b 0a
0000090 40 ec 1e 20 5b 24 29 1b cc 5e 00 62 17 01 1d 08
00000a0 64 6f 01 b1 d3 21 ec 13 60 35 10 f6 1d b0 9a 90
00000b0 20 67 20 fb 03 90 cd 97 04 66 33 81 ec e2 4b 87
00000c0 b0 05 40 6c a8 bd 20 20 e8 98 92 9f 94 aa 00 f2
00000d0 bd 86 a1 a5 a5 85 26 89 7e 20 08 4a 52 2b 4a 40
00000e0 b4 73 7e 41 65 51 66 7a 46 89 82 23 30 a4 52 15
00000f0 3c f3 92 f5 74 14 8c 0c 0c cd 18 18 40 e1 0e 51
0000100 fd 39 10 1c 9e 8c 62 67 10 62 08 80 10 9b 23 c1
0000110 c0 e0 bf 94 81 81 e5 0f 42 cc a4 97 81 61 81 0e
0000120 03 03 ff 54 84 98 9a 21 03 83 80 3e 03 c3 be 39
0000130 c9 a5 45 65 50 63 18 99 8c 19 18 08 f1 01 69 f3
0000140 4a 7d e7 76 61 f4 00 00 00 09 70 48 59 73 00 00
0000150 2e 23 00 00 2e 23 01 78 a5 3f 76 00 00 40 00 49
0000160 44 41 54 78 01 ec dd 0d 7c 55 57 9d ef ff af be
0000170 a0 6d 3a c6 5e ec 1c c7 46 1b 8a b4 06 70 c2 4b
```

How to make
this process **easy**?

Why is it *not* so easy?

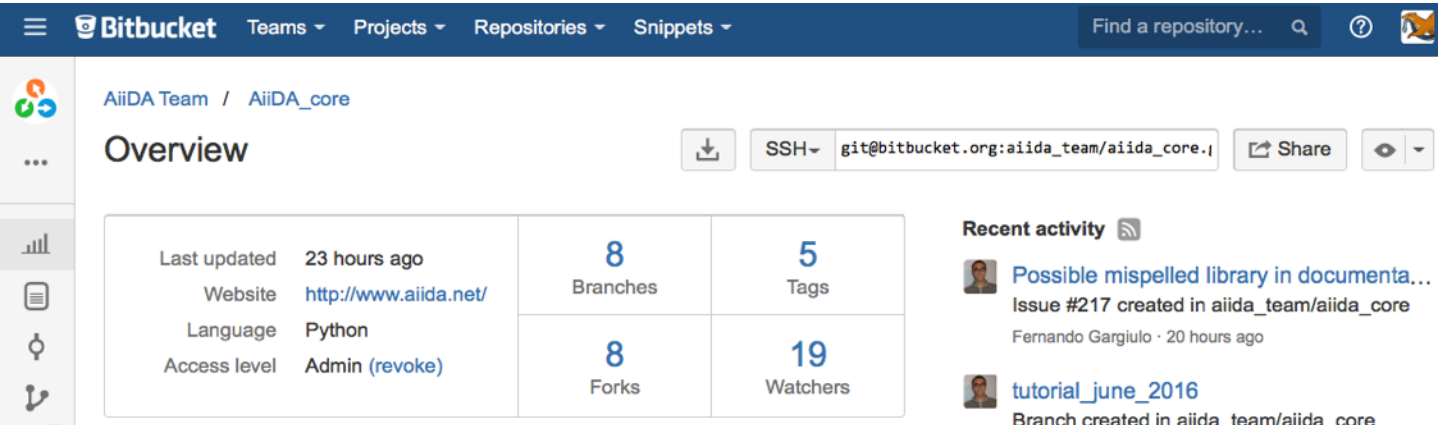
- Researchers' main interests are publications (and citations)
  Spending time to "make order" is not perceived as giving direct benefit
- I don't want to share my data before publication (or patent)
  I want to publish as soon as possible; After I'll spend time to "prepare" the data only if it won't take too much time
- There is no easy tool to help me organise my computational research data
- Even if I organise my data, how do I share it so that it is useful for others?

Our (open-source) tool: AiiDA

We run *computer* laboratories:
we should manage *simulations*
and *data* with a dedicated platform
to make the data share-ready
already during its generation

AiiDA
Automated Interactive Infrastructure and Database for Computational Science

http://www.aiida.net



MIT License

open source initiative

Approved License

What are our driving requirements for computational science? It should be…

reproducible

Often not possible from the data reported in papers

searchable

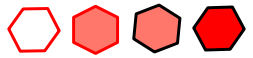
Find and reuse existing calculations

reliable

Results in repositories, automated verification procedures

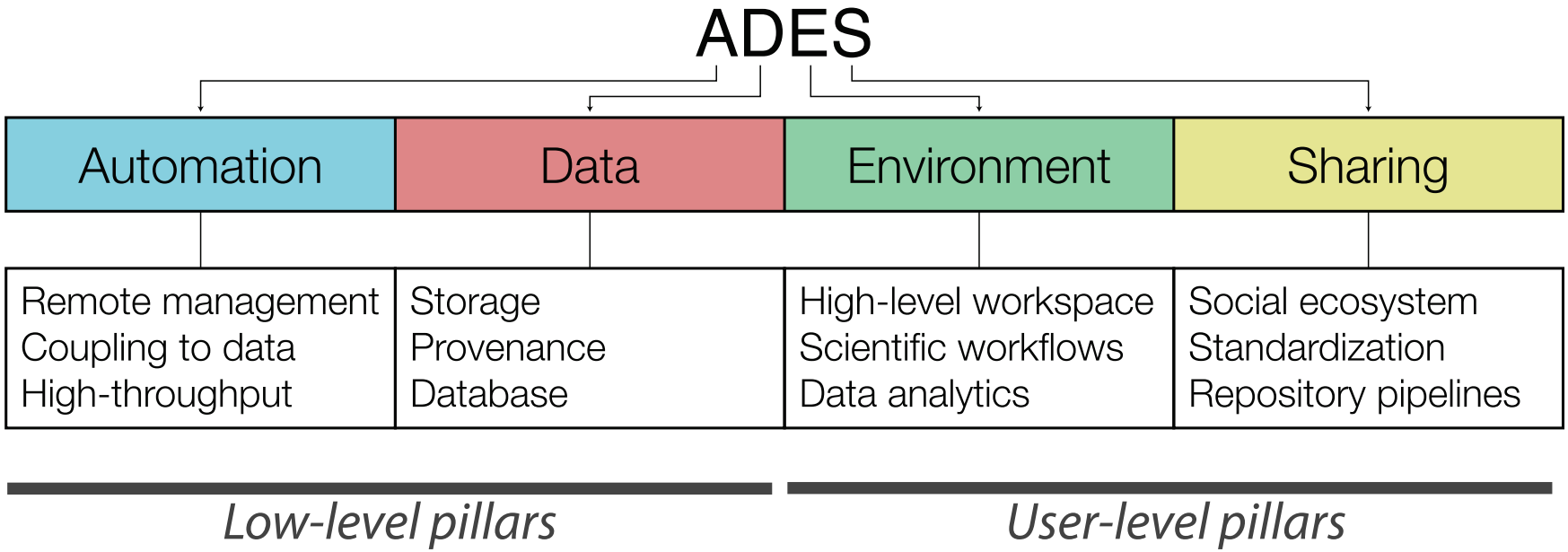
shareable

towards "open data", but private before publication



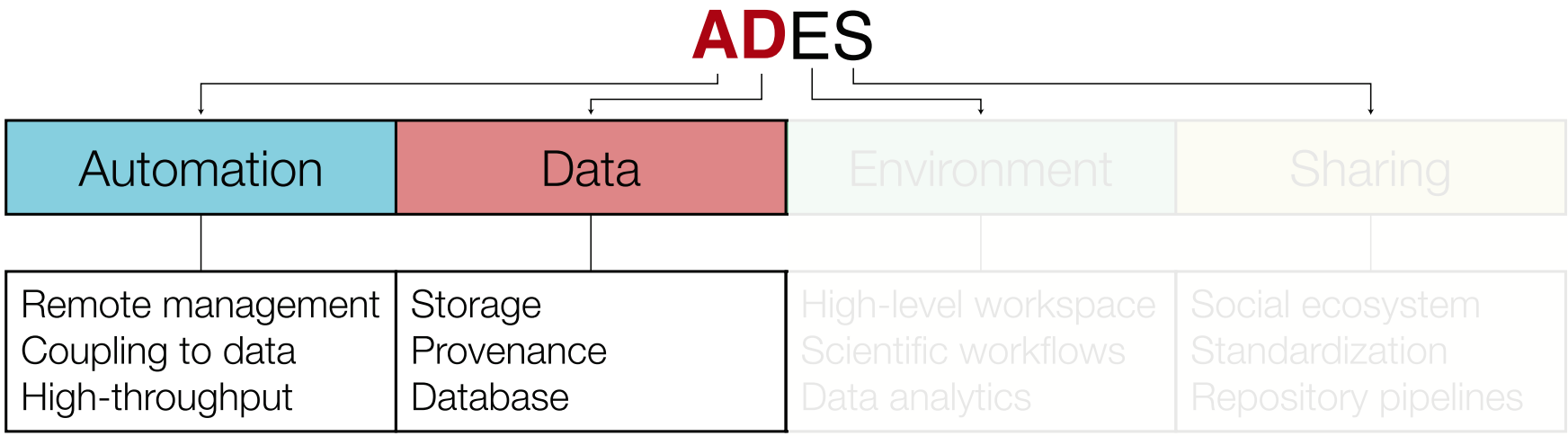
The ADES model

We have encoded these requirements in the four pillars for a computational science infrastructure



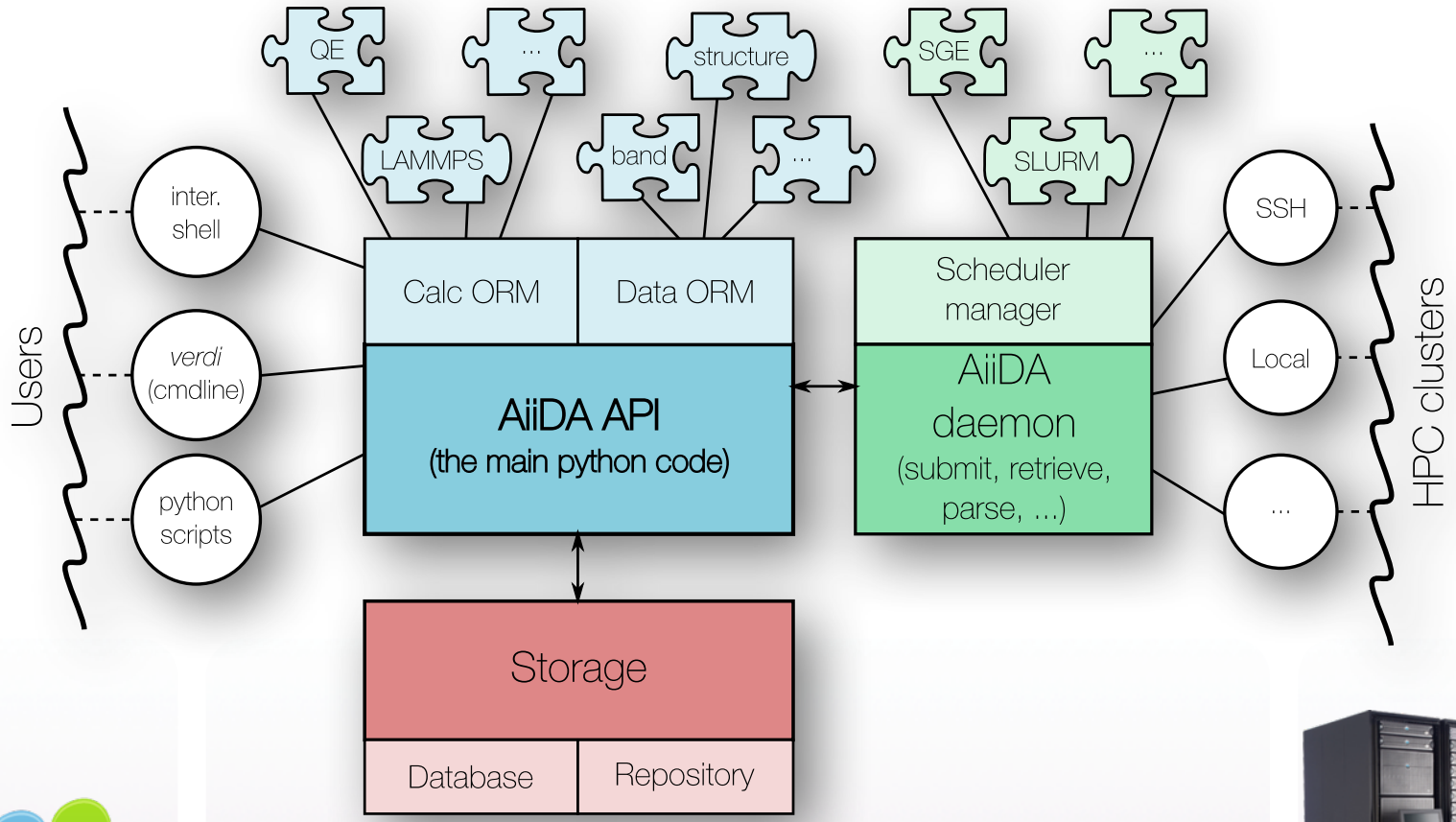
G. Pizzi et al., Comp. Mat. Sci 111, 218-230 (2016)

Automation and Data in AiiDA

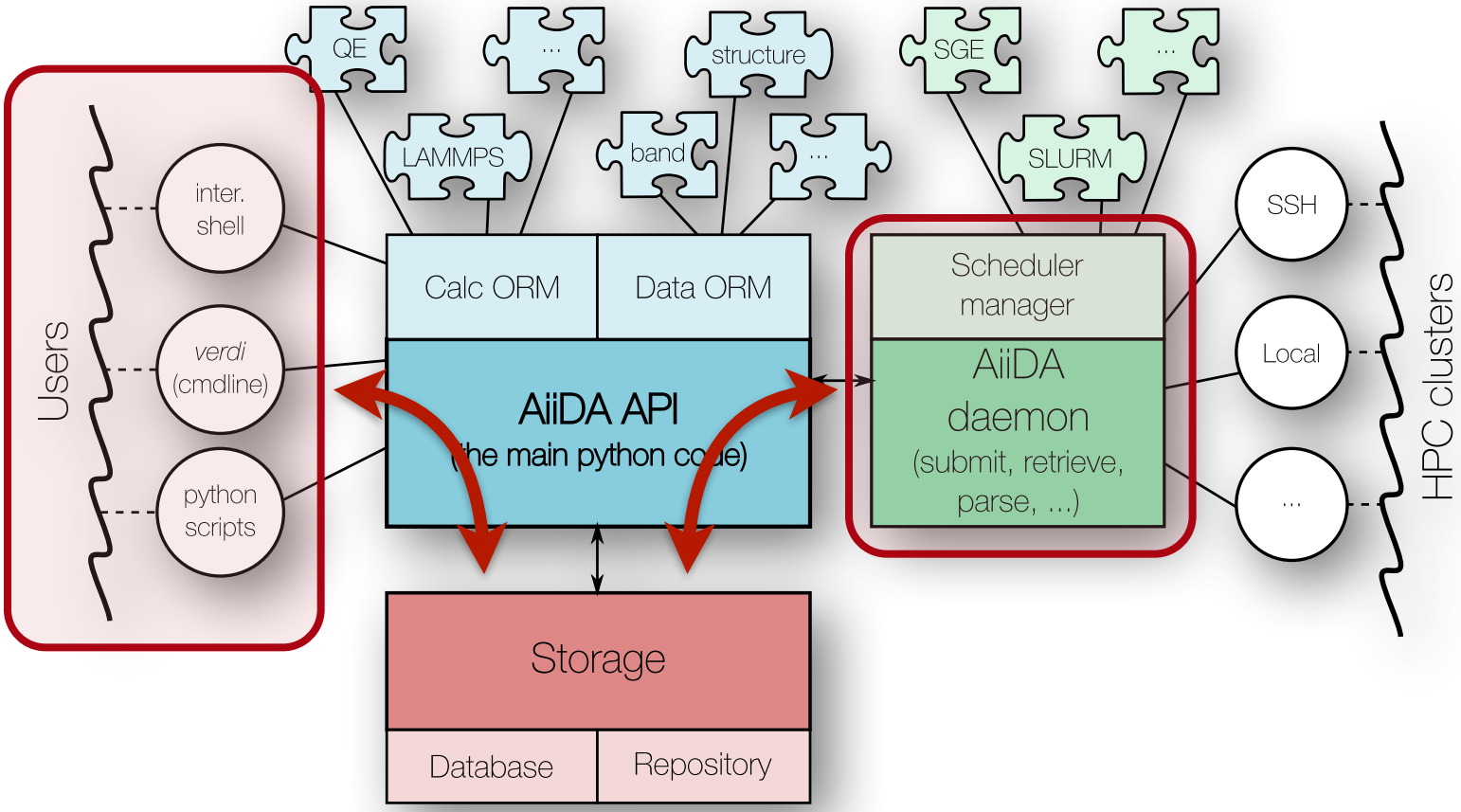


G. Pizzi et al., Comp. Mat. Sci 111, 218-230 (2016)

What is AiiDA?



Automation: coupling to data



- ***Coupling* automation to data:**
  - *uniformity* of the input data, usage of codes and computers
  - full *reproducibility* of calculations (data is stored first)

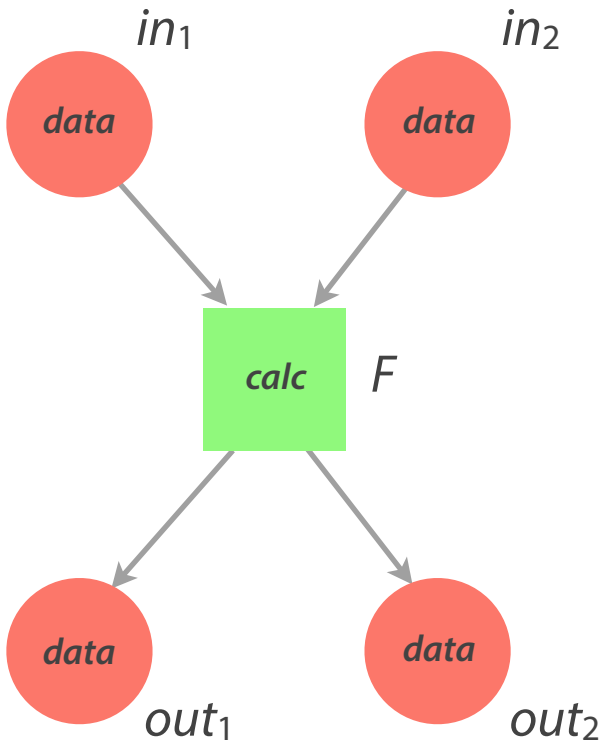
Storage and provenance

- How do we store simulations **preserving the connected structure between them?**

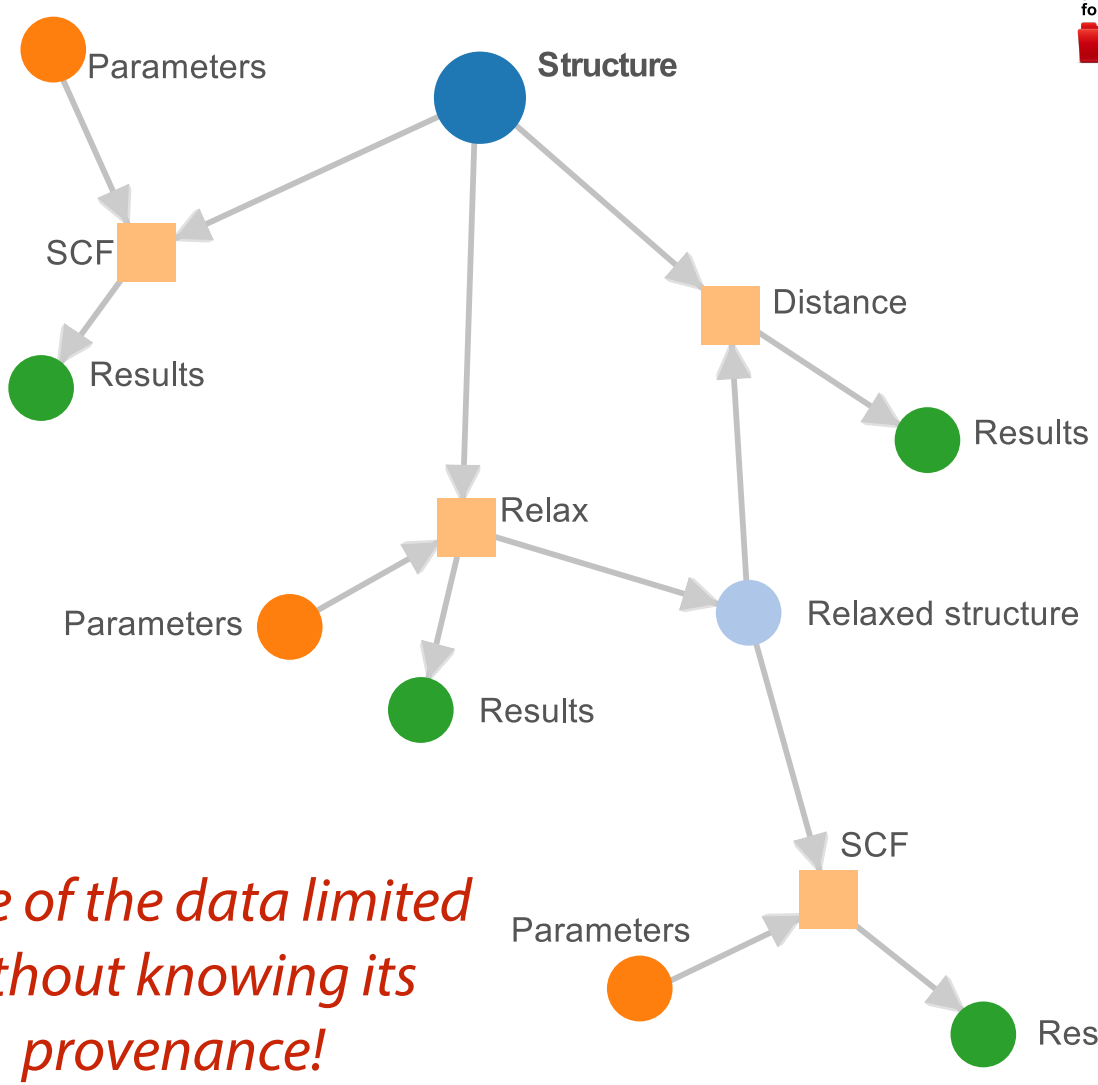
- Inspiration from the *open provenance model*

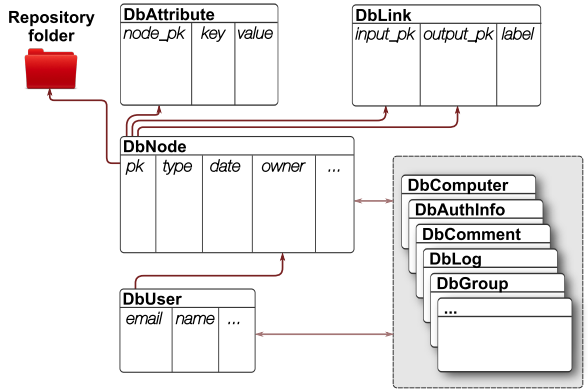
- **Any calculation: a function,** converting inputs to outputs:

$$out_1, out_2 = F(in_1, in_2)$$



Data provenance: Directed Acyclic Graphs



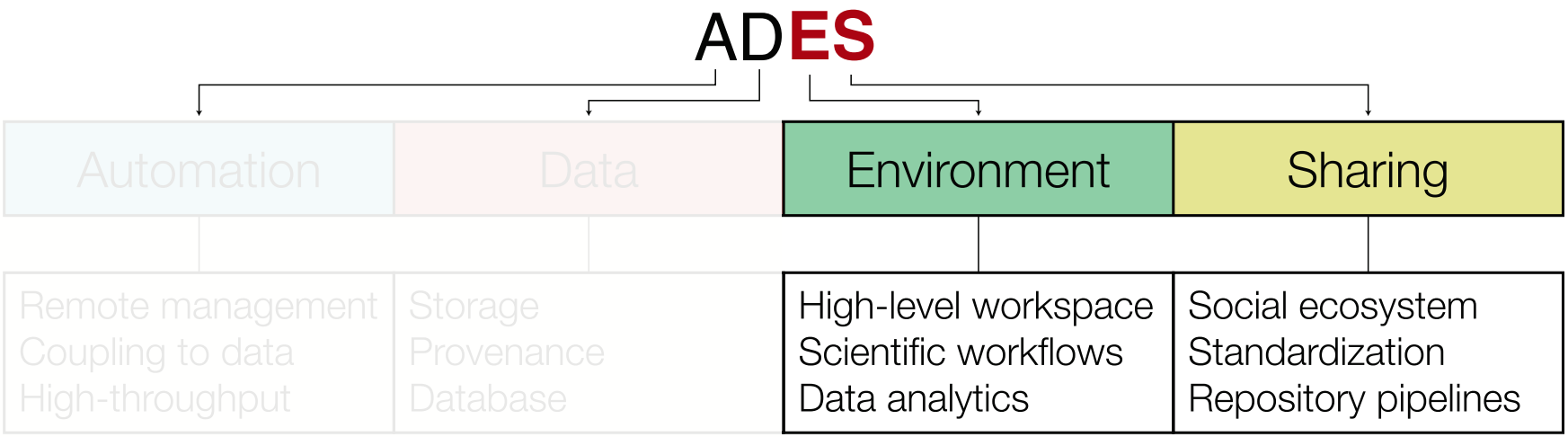


G. Pizzi et al., Comp. Mat. Sci. 111, 218-230 (2016)

Need of advanced databases to store and efficiently query

Value of the data limited without knowing its provenance!

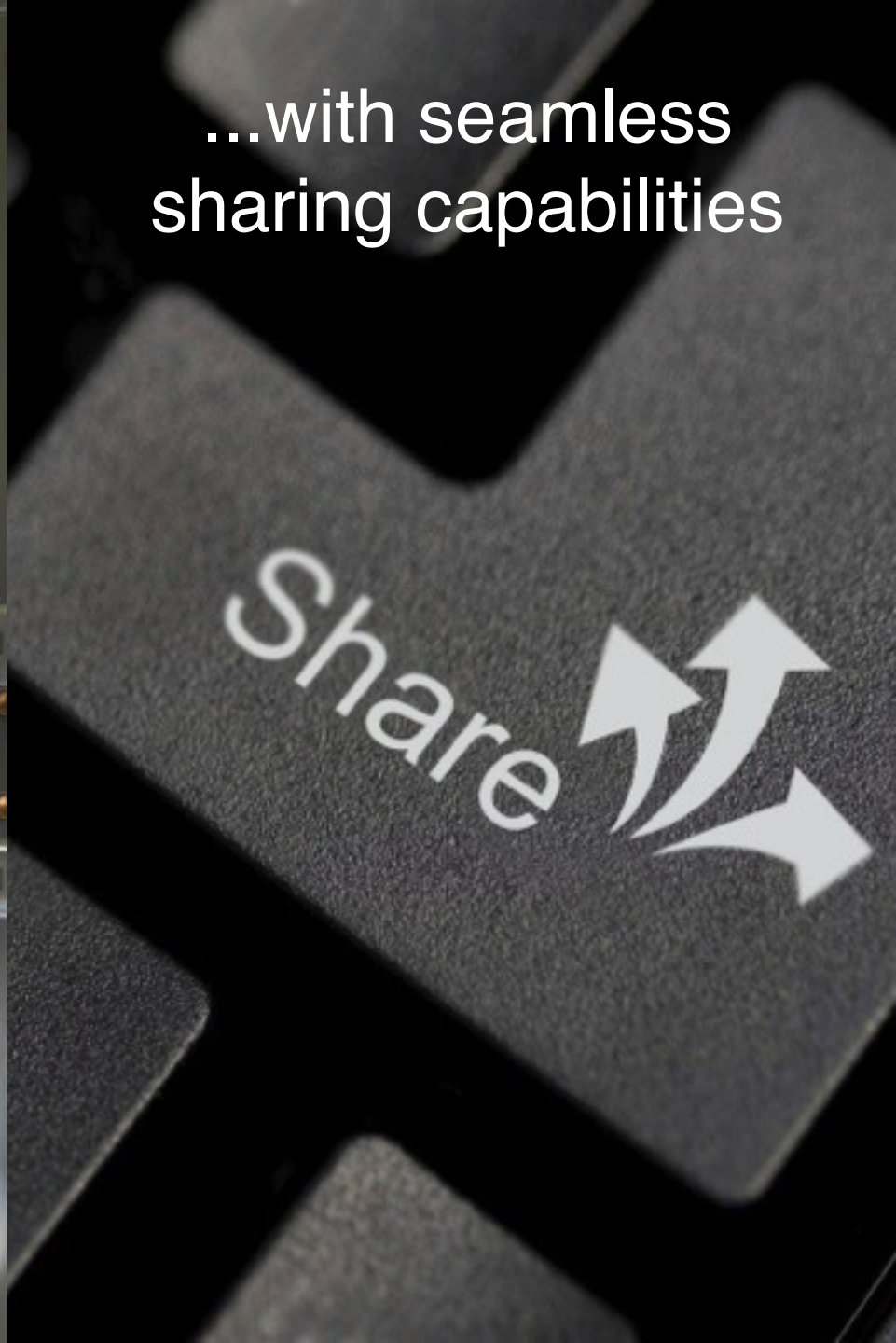
Environment and Sharing in AiiDA



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

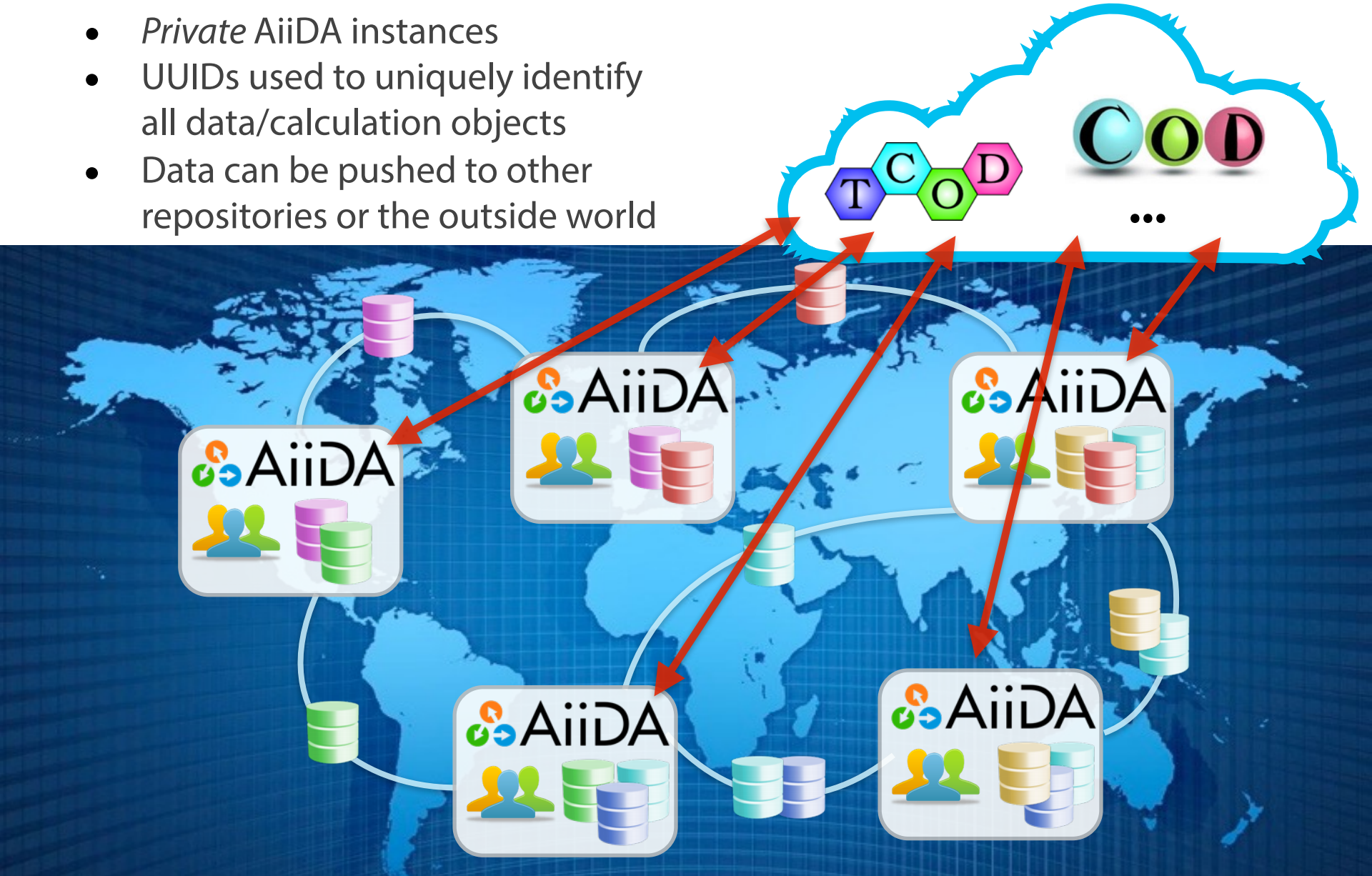
Keep data private...

...with seamless sharing capabilities



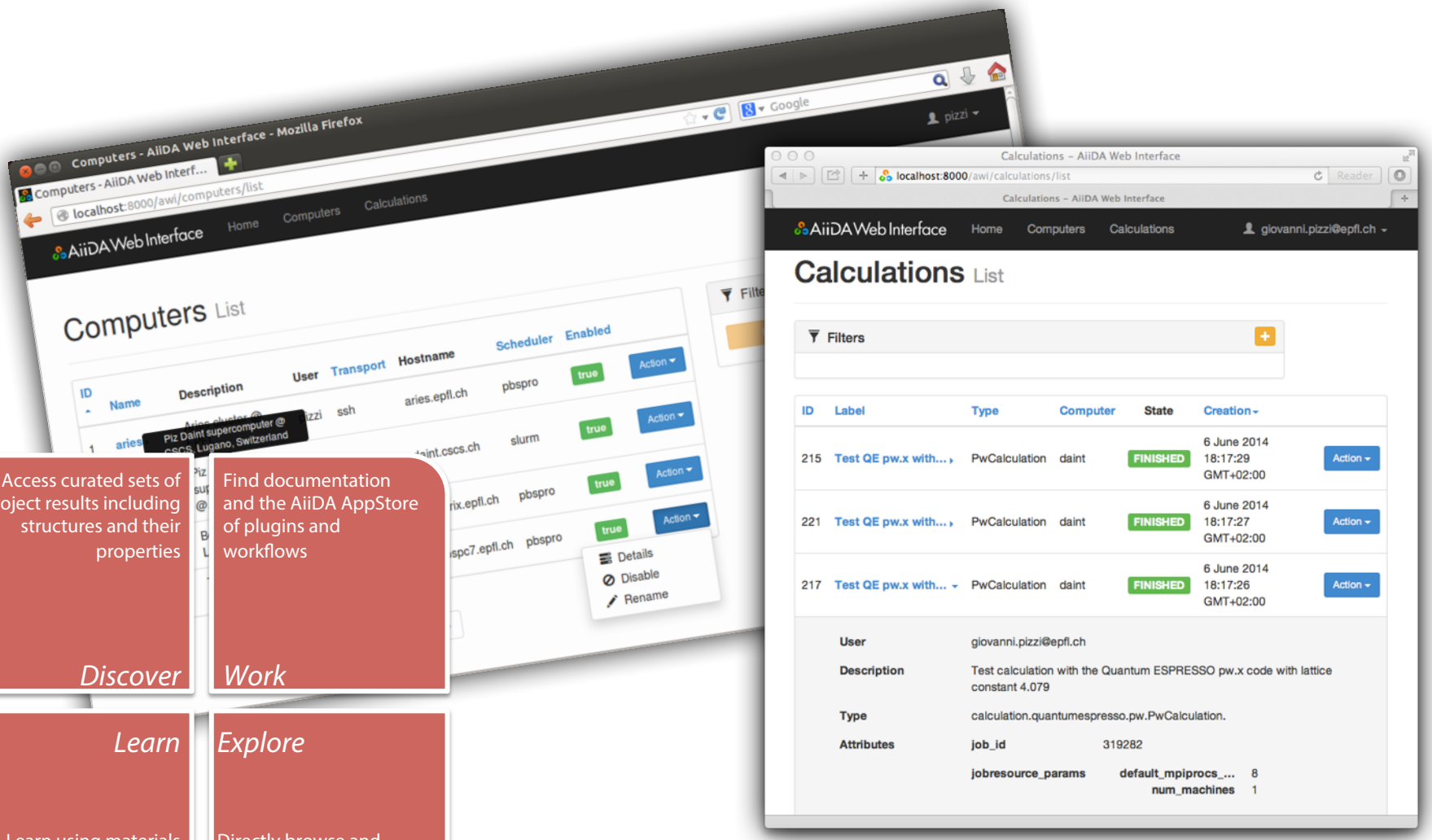
Sharing in AiiDA

- *Private* AiiDA instances
- UUIDs used to uniquely identify all data/calculation objects
- Data can be pushed to other repositories or the outside world



Materials Cloud

www.materialscloud.org



Access curated sets of project results including structures and their properties

Discover

Find documentation and the AiiDA AppStore of plugins and workflows

Work

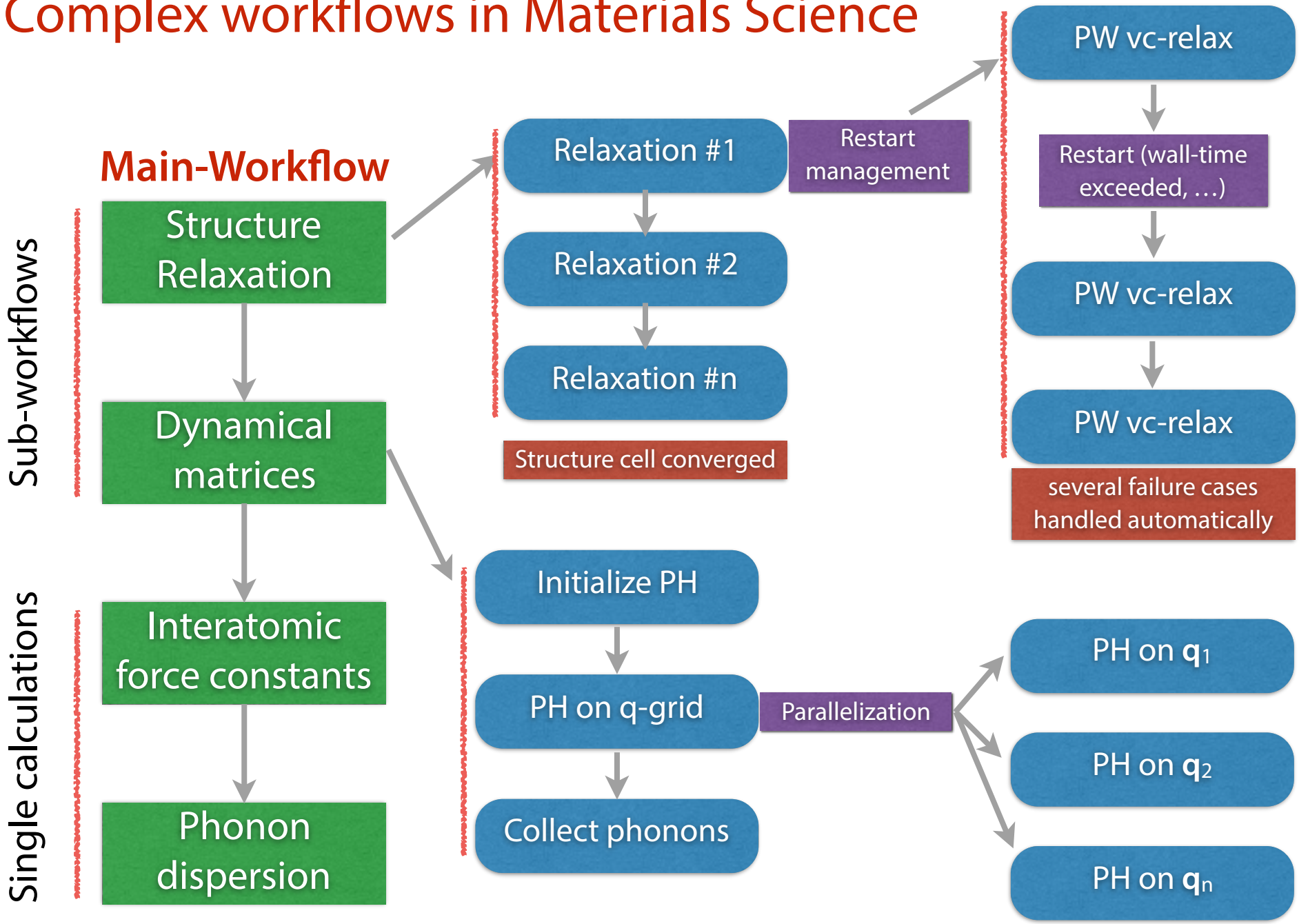
Learn

Learn using materials including lectures and video tutorials related to materials science

Explore

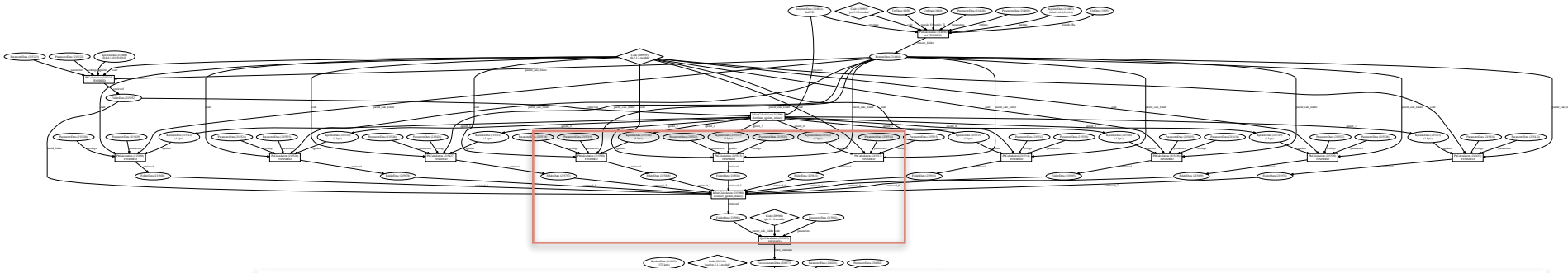
Directly browse and visualise the AiiDA database

Complex workflows in Materials Science



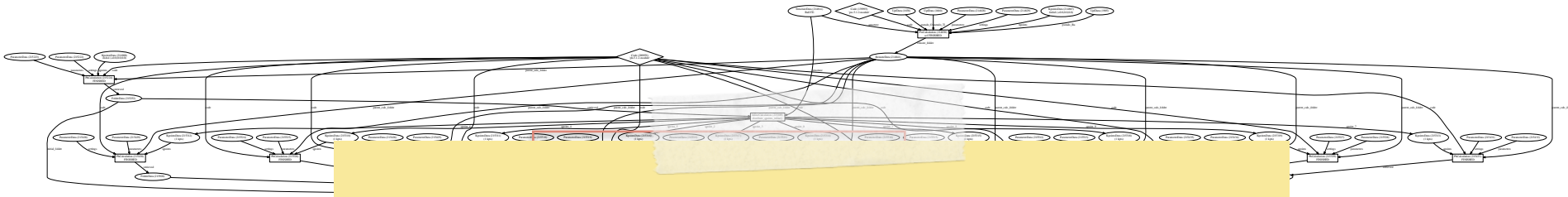
Encoding the scientists' knowledge in workflows

```
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'}
          }
wf = asyncd(PhBandsWorkflow, **params)
```



Encoding the scientists' knowledge in workflows

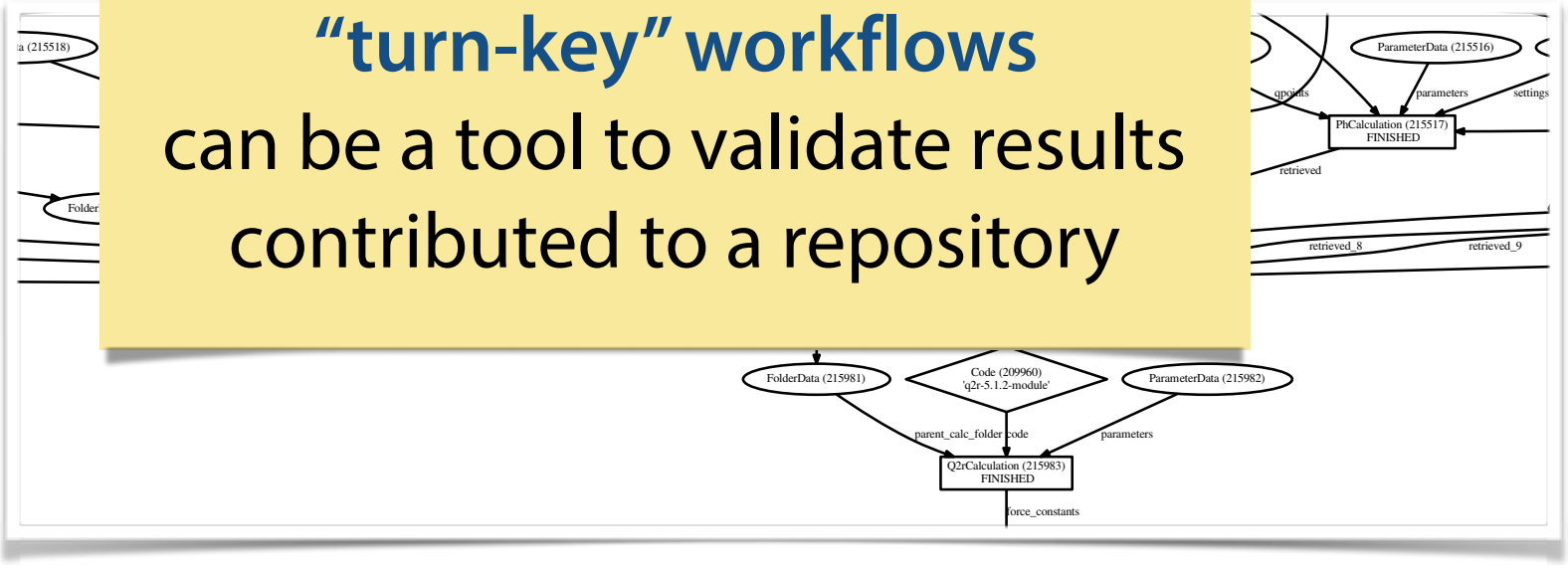
```
params = {'input': {'kpoints_density': 0.2,
                    'convergence': 'tight'},
          'structure': structure,
          'pseudo_family': pseudo_family,
          'machinename': 'mycluster',
          'pw_input':{'volume_conv_threshold': 5e-2},
          'pw_parameters': {
          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



Simple and reliable

"turn-key" workflows

can be a tool to validate results contributed to a repository



App-store model: open tools for open data

"App-store"-like model for plugins & workflows, e.g.

- **Computers**: automatically setup a new cluster or supercomputer
- **DB importers**: load structures and data from COD, ICSD, …
- **Calculations**: plugin for your simulation software
- **Turn-key solutions**: workflows to compute a desired property
- …

The MaterialsCloud AiiDA app-store

Simulation codes

Quantum ESPRESSO

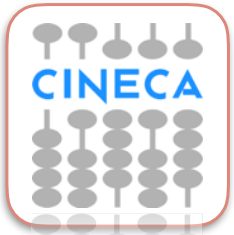
Yambo

CP2K

Fleur

Computing clusters

CSCS (CH)

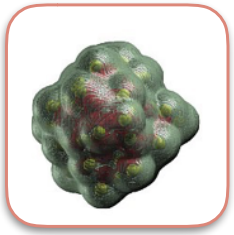


CINECA (IT)

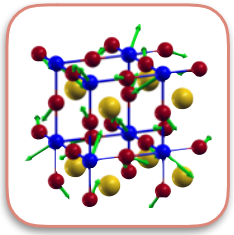
BSC (ES)

Jülich (DE)

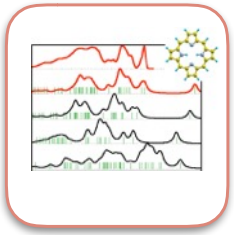
Turn-key workflows



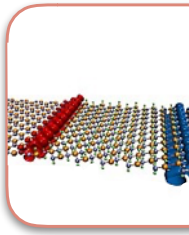
Electronic density



Phonons



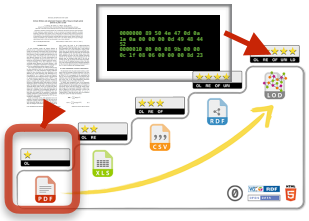
Optical properties



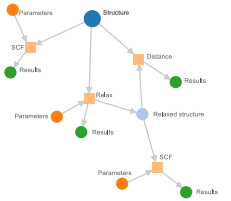
Transport

Some final thoughts

- **Publishers and funding agencies:** How much should they support/enforce publication of 'raw' data? (Things are moving: e.g. Scientific Data@NPG, …)

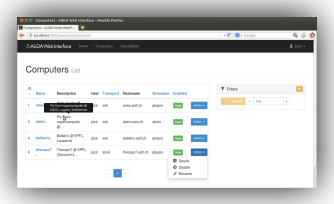


- **Quality of metadata:** not only raw data, but structure: provenance, reproducibility, and open/documented format



- **Ontologies and interoperability**: Domain-specific ontologies (e.g. in RDF) for crystal structures, energy, forces, bulk moduli, …; not easy to "convince" researchers… But how much should we really invest in this?

- **Open repository**: curated or not? Limited submission or accepts anything? How to verify the quality of data?



- **Access to data**: Fully open with no barriers, or via registration +terms of use? How to know the actual data usage (e.g. to justify project with funding agencies)? What about potential ethical implications?

Acknowledgements: people

The AiiDA team


Giovanni Pizzi (EPFL)

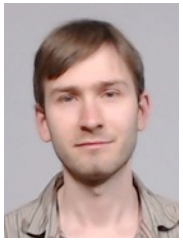

Andrea Cepellotti (EPFL)


Nicolas Mounet (EPFL)


Riccardo Sabatini (EPFL)


Boris Kozinsky (BOSCH)


Nicola Marzari (EPFL)


Martin Uhrin (EPFL)


Leonid Kahle (EPFL)


Spyros Zoupanos (EPFL)


Snehal Waychal (EPFL)


Fernando Gargiulo (EPFL)


Andrius Merkys (Vilnius)

Plugin contributors — **Quantum ESPRESSO NEB**: Marco Gibertini (EPFL); **Quantum ESPRESSO DOS, PDOS; Wannier90**: Daniel Marchand (EPFL); **CP2K**: Aliaksandr Yakutovich (EMPA), Uli Schauer (ETHZ), Tiziano Müller, Andreas Gloess, Patrick Seewald (UZH); **FLEUR**: Jens Broeder, Gregor Michalicek, Daniel Wortmann (Jülich); **Exciting**: Anton Kozhevnikov (CSCS); **YAMBO**: Andrea Ferretti, Giovanni Borghi, Daniele Varsano (CNR-NANO), Gianluca Prandini (EPFL); **SIESTA**: Victor Garcia Suarez (Uni Oviedo); **i-PI**: Venkat Kali (EPFL); **VASP**: Mario Zic (Trinity College Dublin).
Contributors — Prof. Christoph Koch (EPFL); Valentin Bersier, Philippe Schwaller (THEOS EPFL); Marco Dorigo (ICAMS - Bochum); Eric Hontz (MIT & Bosch RTC)
Early beta testers — Giovanni Borghi, Ivano Castelli, Marco Gibertini (THEOS EPFL); Prateek Mehta (Bosch RTC)

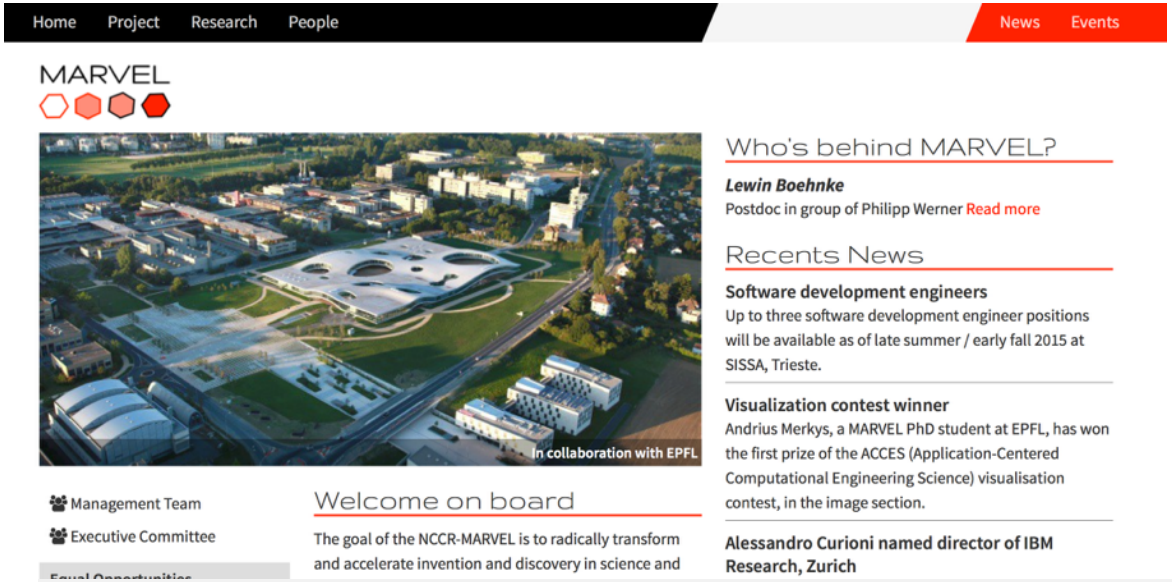
Acknowledgements: funding

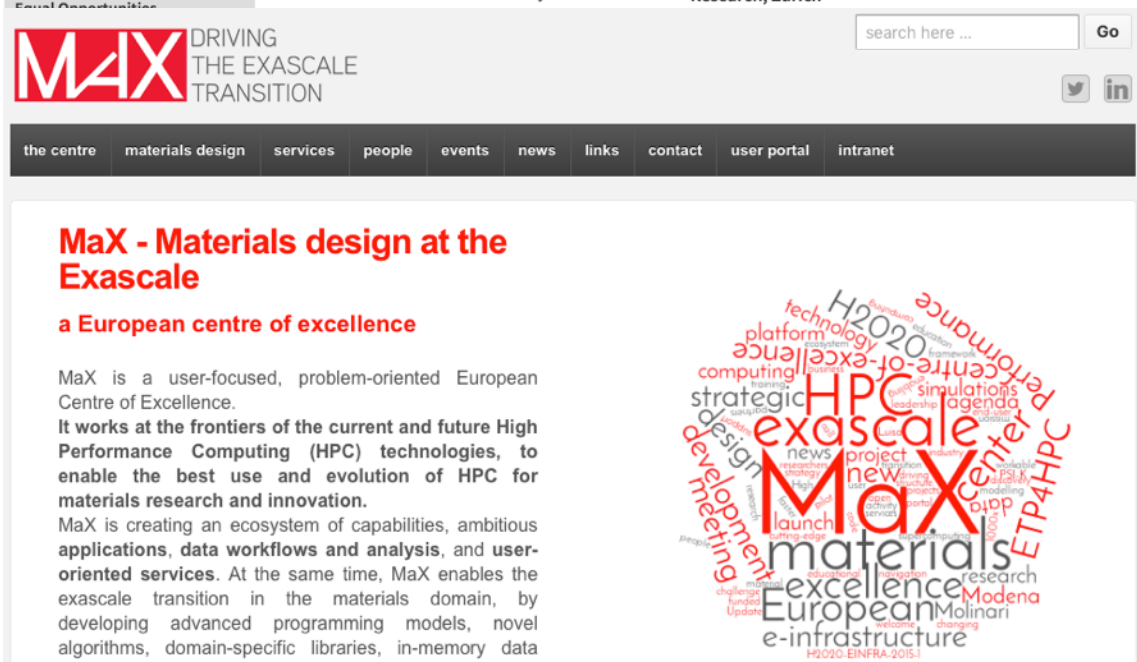
- **MARVEL NCCR**
  http://marvel-nccr.ch
  Swiss National Centre of Competence on Computational Design and Discovery of Novel Materials
  Started May 2014, funded 2014-2026.

  33 professors from 11 Swiss Institutions.

- **MaX**
  http://max-center.eu
  "MAterials design at the eXascale": EU H2020 e-infrastructure project
  From Sep 2015 for 30 months.

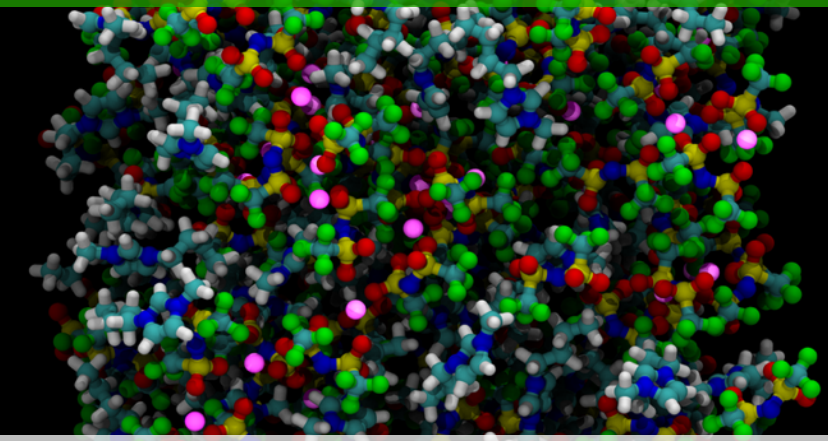
  Groups from Modena, Trieste, Lausanne, Barcelona, Julich
  + 5 supercomputing centers: CINECA, CSCS, BCS, Julich, KTH

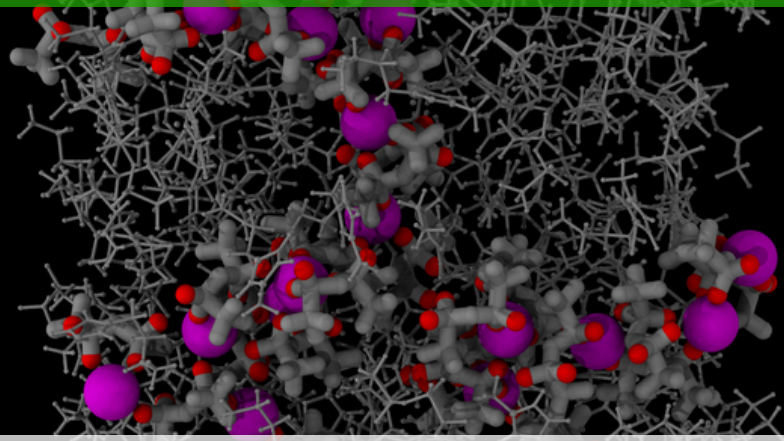




Shift the scientist's focus from the **management** of simulations...

…to the **discovery** of new materials…





…making it **seamless to share** research data

Contacts and info:

Website: http://www.aiida.net

Docs: http://aiida-core.readthedocs.io

Git repo: https://bitbucket.org/aiida_team/aiida_core/


https://www.facebook.com/aiidateam


@aiidateam