2020 Atomistic Simulations for Industrial Needs Workshop (online)

August 6, 2020

The Platform for the Interatomic Potentials Validation

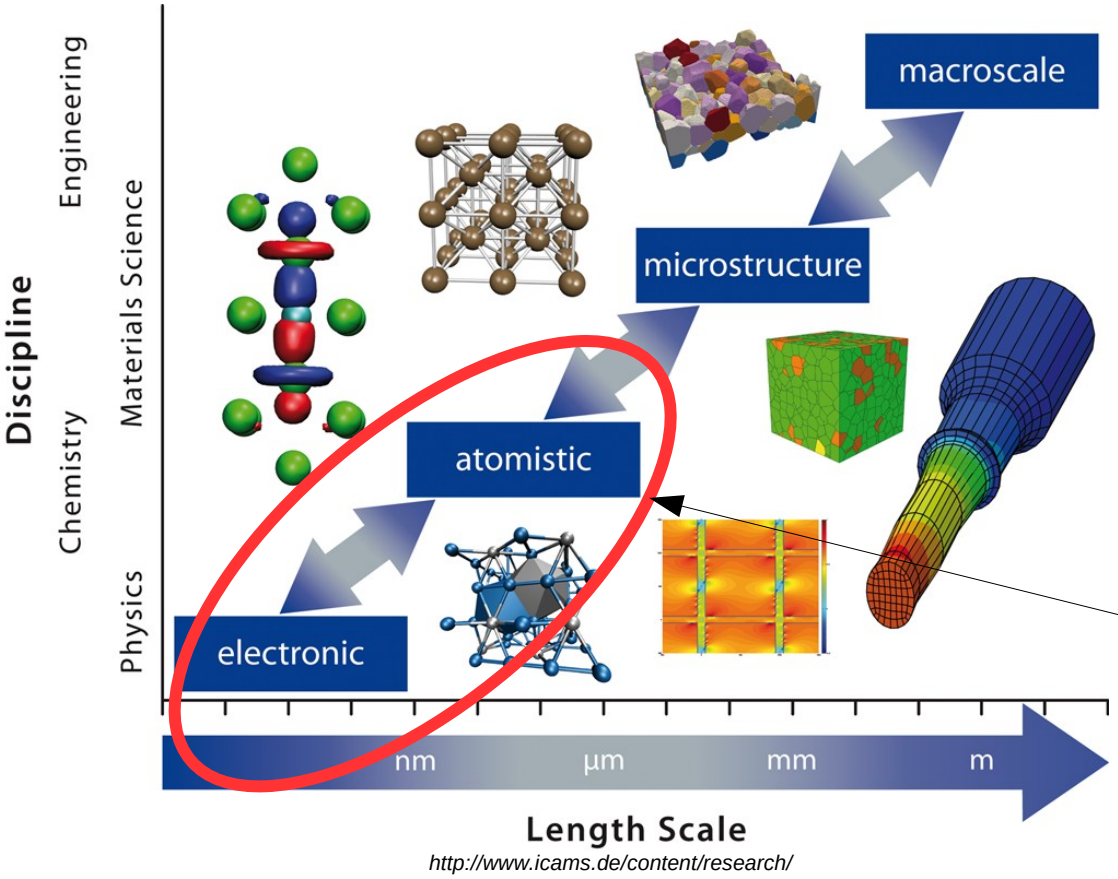
Yury Lysogorskiy
Thomas Hammerschmidt
Matous Mrovec
Ralf Drautz

Interdisciplinary Centre for Advanced Materials Simulation (ICAMS)
Ruhr-Universität Bochum, Germany
yury.lysogorskiy@rub.de

ICAMS
INTERDISCIPLINARY CENTRE FOR ADVANCED MATERIALS SIMULATION

RUHR UNIVERSITÄT BOCHUM
RUB

Atomistic simulations



http://www.icams.de/content/research/

ab initio and molecular dynamic simulations

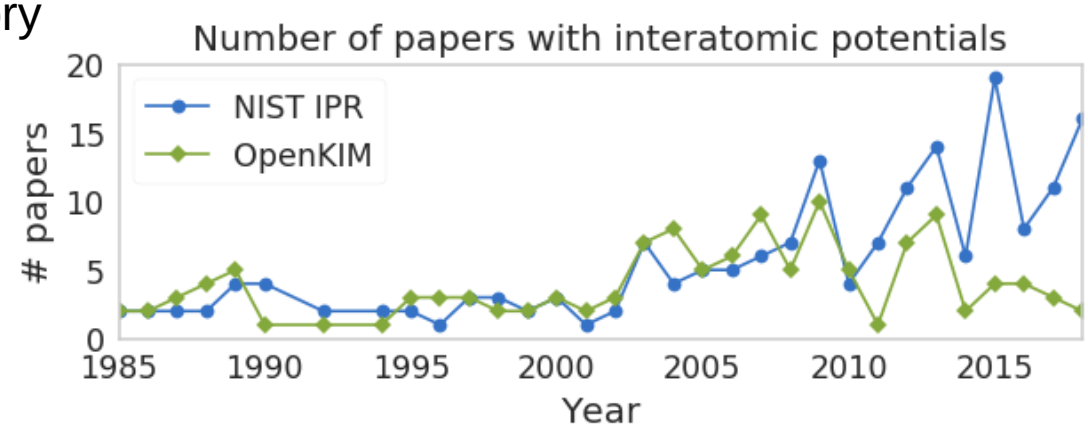
Interatomic potentials (as of 2019)

- NIST interatomic potentials repository

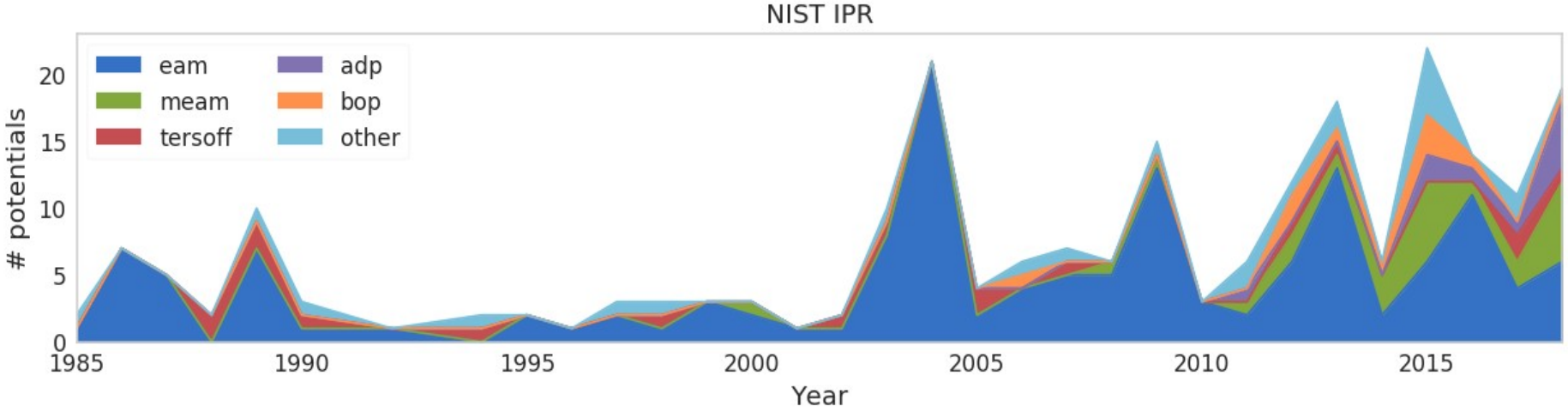
NIST IPR

- OpenKIM

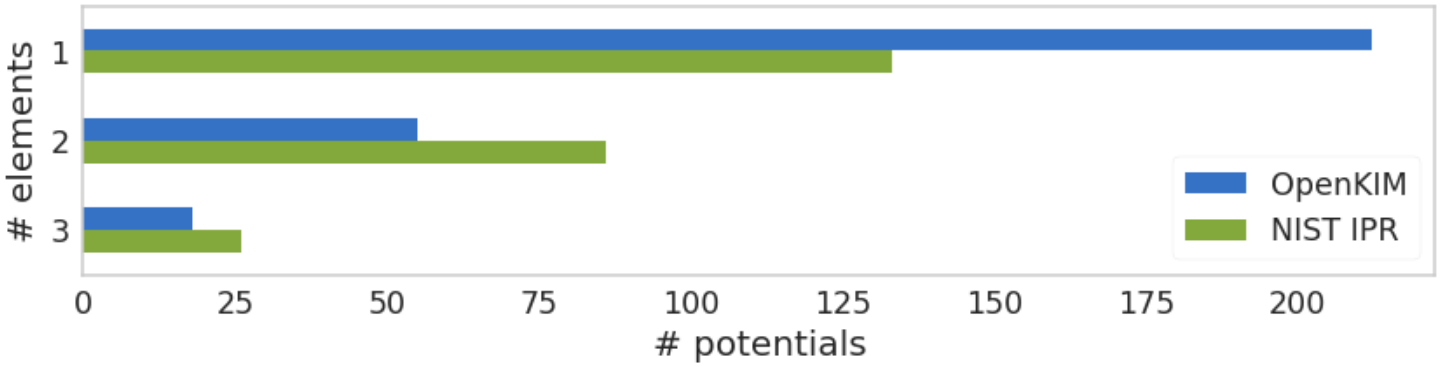
OpenKIM



Interatomic potential types (LAMMPS pair_style)

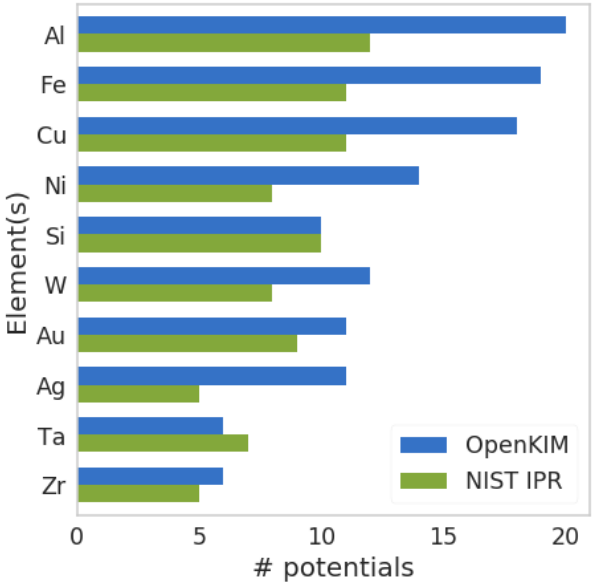


Interatomic potentials statistics (as of 2019)

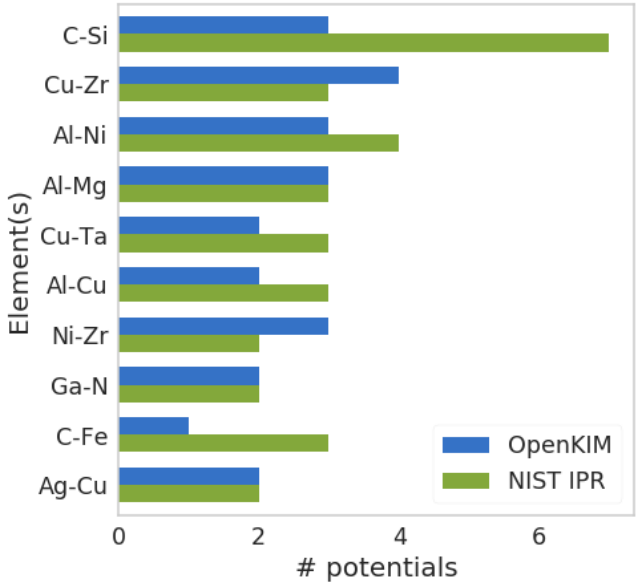


Single-specie interatomic potentials are the most common

Most frequent elements in single-species potentials



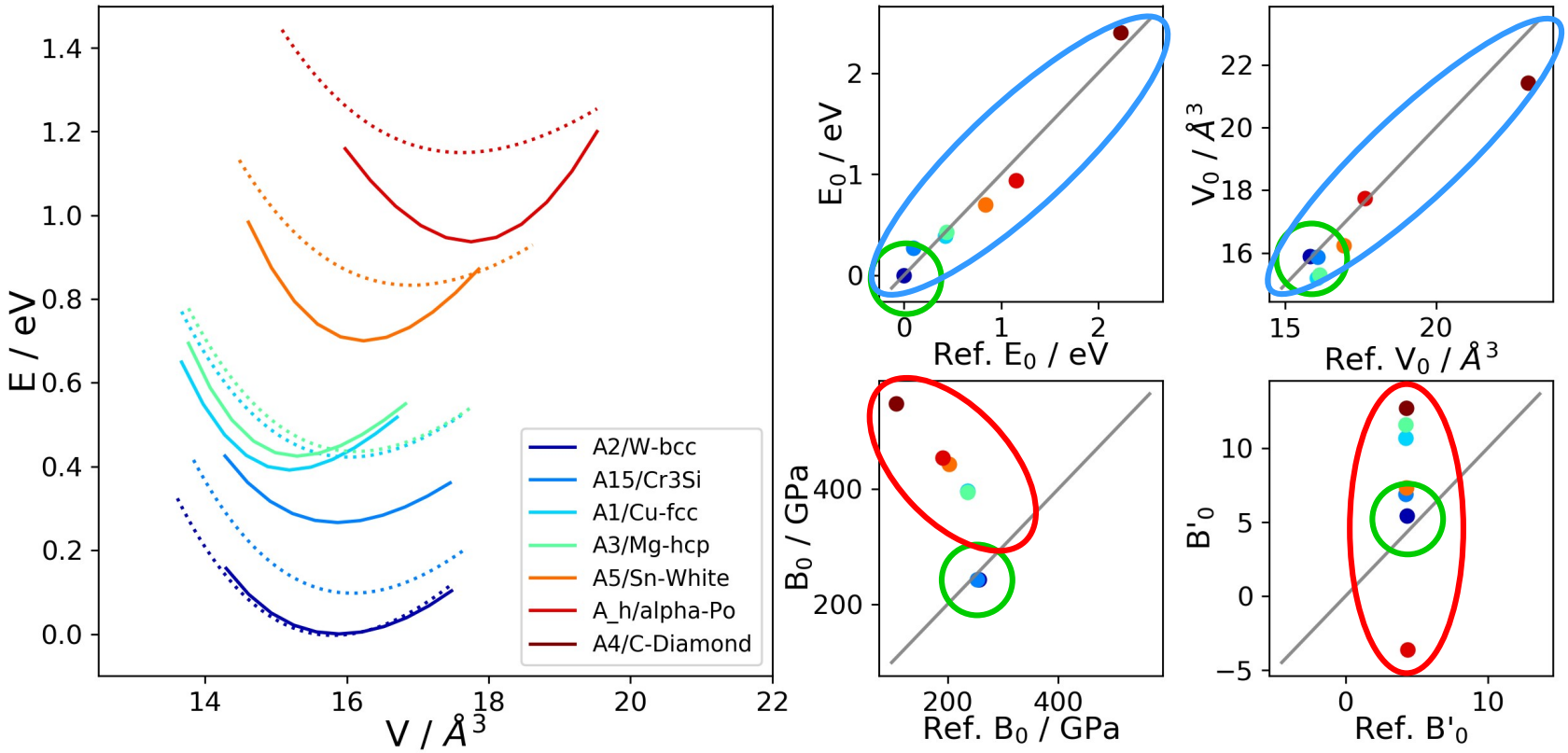
Most frequent binaries in two-species potentials



Transferability

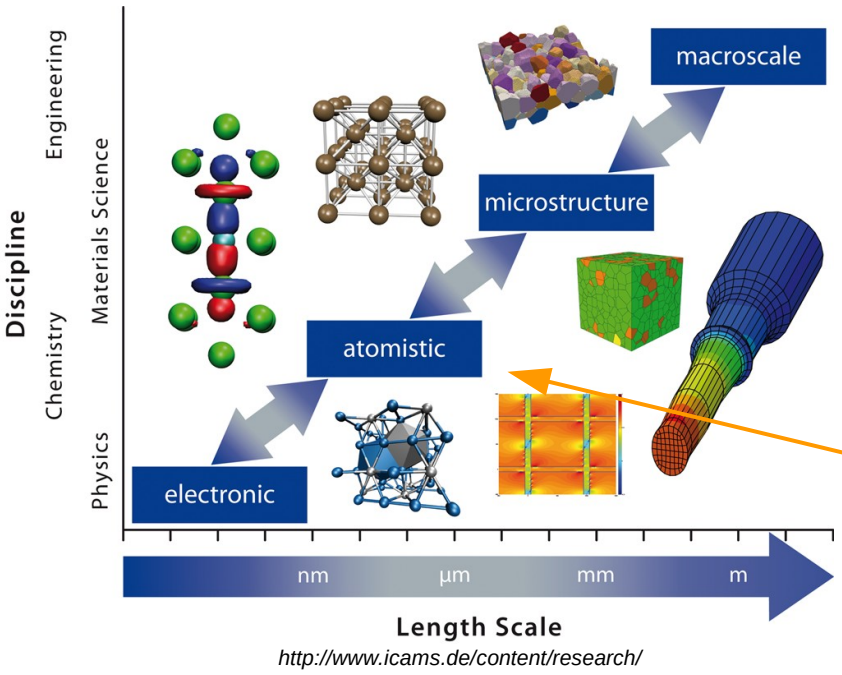
Transferability of interatomic potentials – **reliability** of calculated **properties** for structures which are **not** from potential **training set**

Example:solid line - MEAM Park for Mo [*H.Park et al. Phys. Rev. B 85, 214121]*,
dotted – reference,DFT-PBE



Properties of ground-state (A2/BCC) have small error
Level I properties (energy,volume) – good transferability
Level II properties (B_0, B') – bad transferability

Interatomic potentials



http://www.icams.de/content/research/

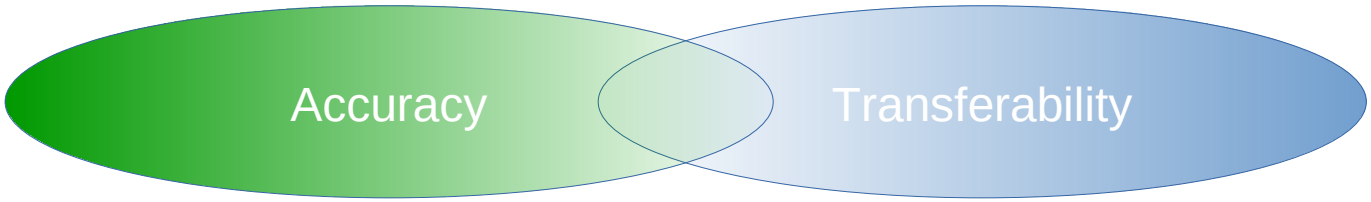
Training set: **DFT**/experimental structures/properties

Potential functional form: MEAM, ADP, BOP, …, ML

Potential fitting (energy-matching, force-matching, property-matching)

Interatomic potential

How reliable are these potentials?



Transferability of interatomic potentials – **reliability** of calculations for structures and/or properties that are **not** from the **training set**

Validate by comparison to reference data

The (basic) needs of atomistic simulation community

Roles

Molecular dynamics simulation user

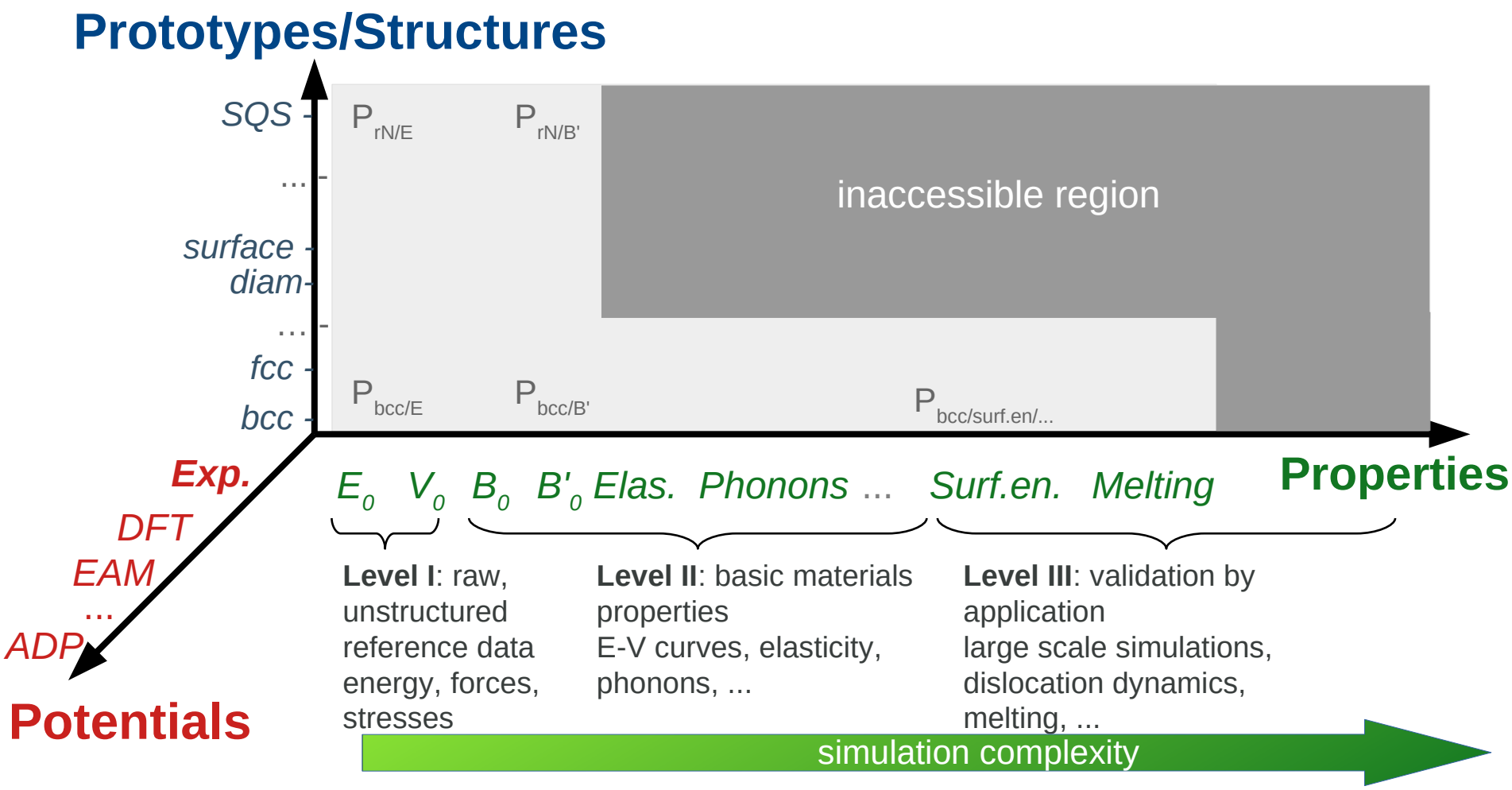
- What is the best/appropriate potentials for my particular purpose ?
- What properties could I simulate (and how)?
- What would be the error of my MD simulations?

Interatomic potentials developer

- How my new potential would perform?
- Are there standard benchmarks which I can use for my new potential?
- Is my potential better than the previous ones?

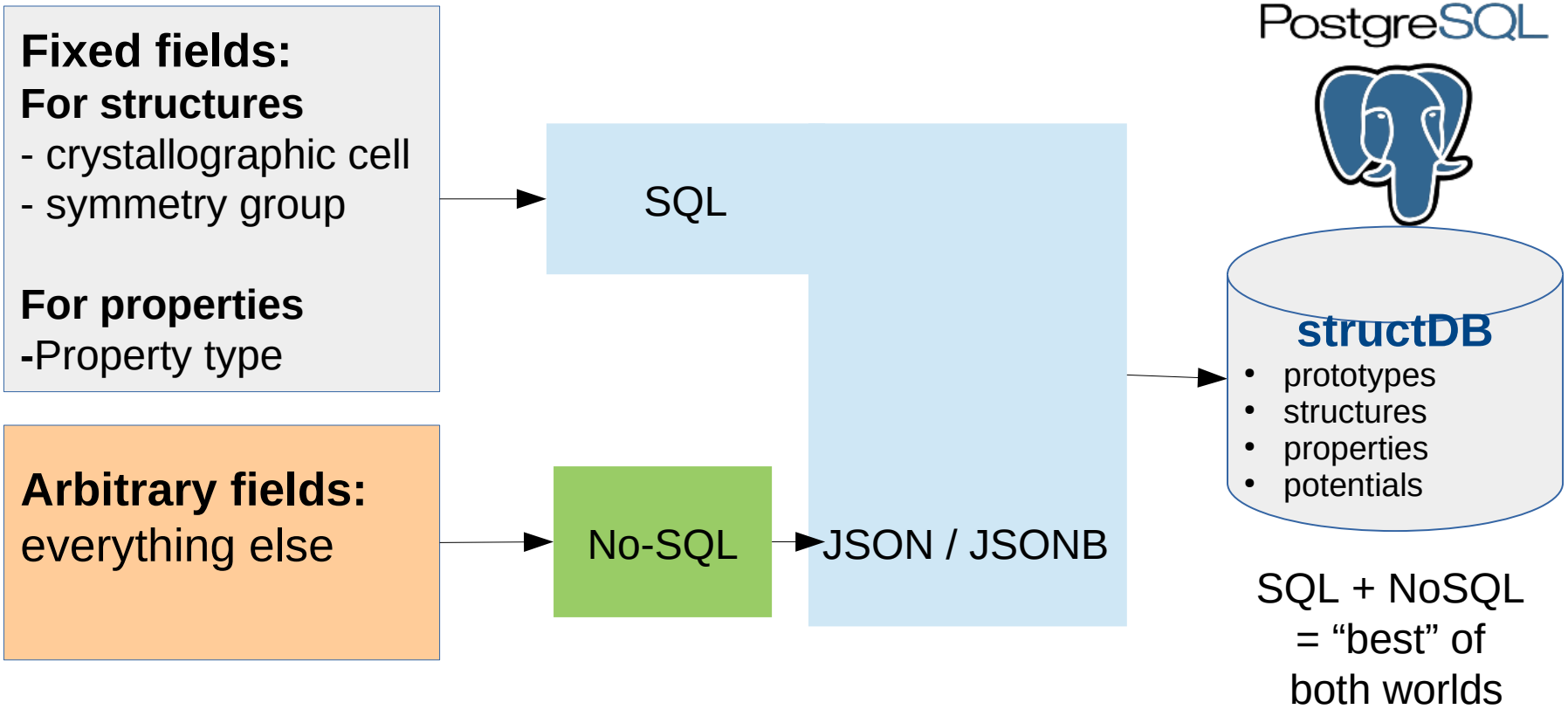
Atomistic simulations PPP-space

Abstract space for the “navigation” in atomistic simulations

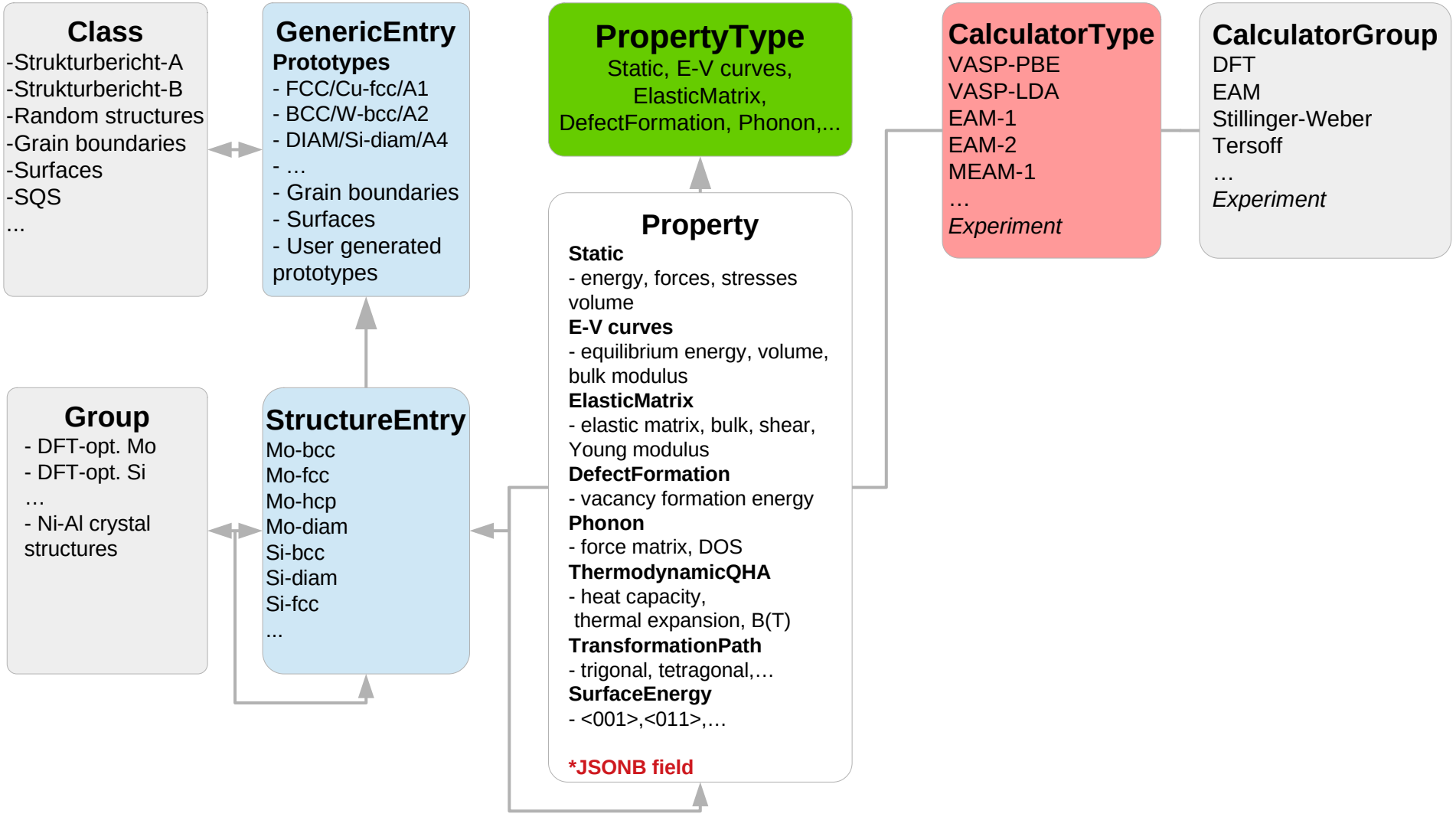


Data types of atomistic simulations

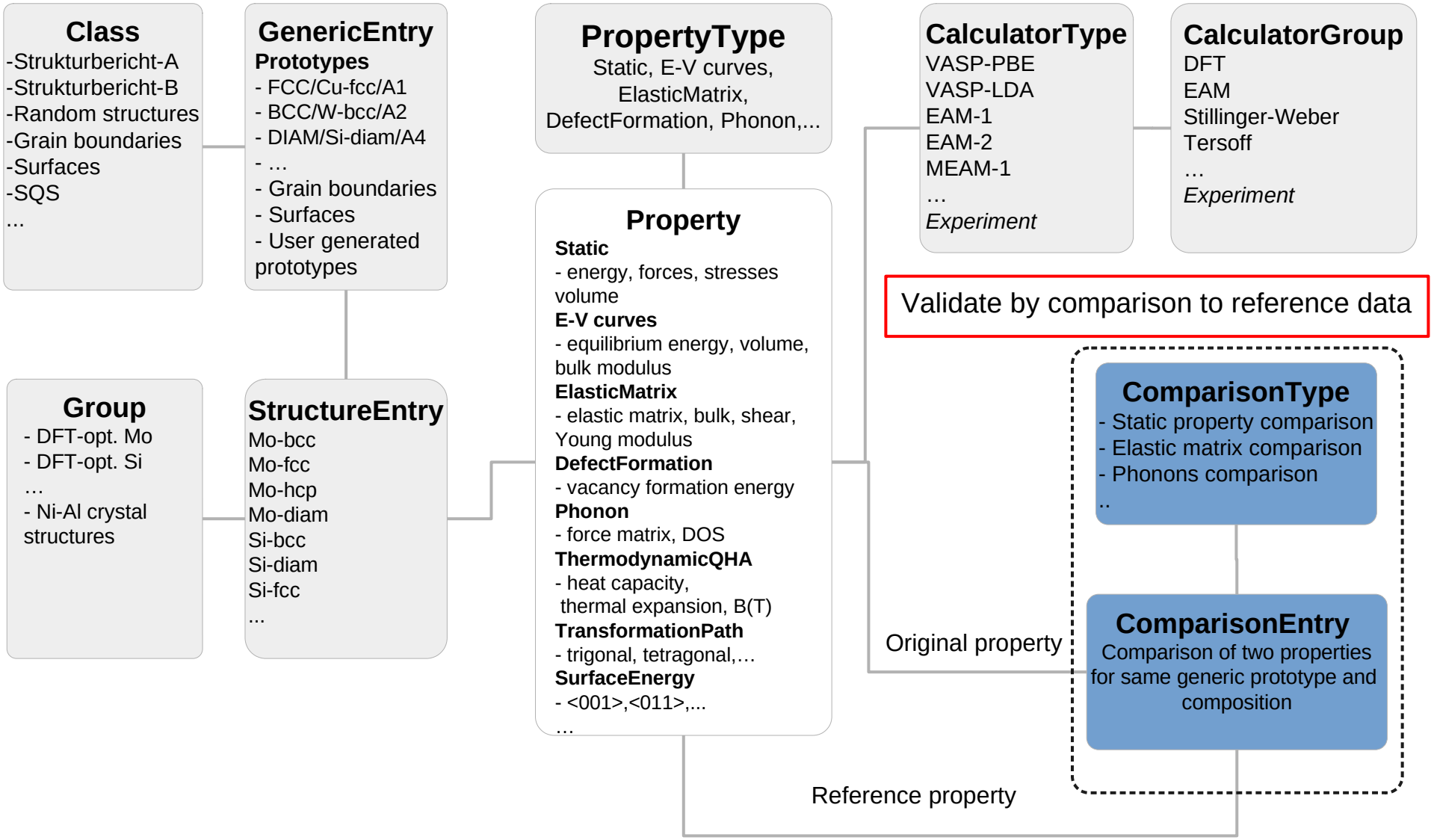
- **Atomic structures:** crystallographic cell, symmetry group, atomic coordinates, composition, etc...
- **Basic data:** energies, forces, volumes, etc...
- **Featured data:** elasticity, defect formation energy, phonon DOS, Cv, Cp, thermal expansion, transformation paths, etc...



Database design

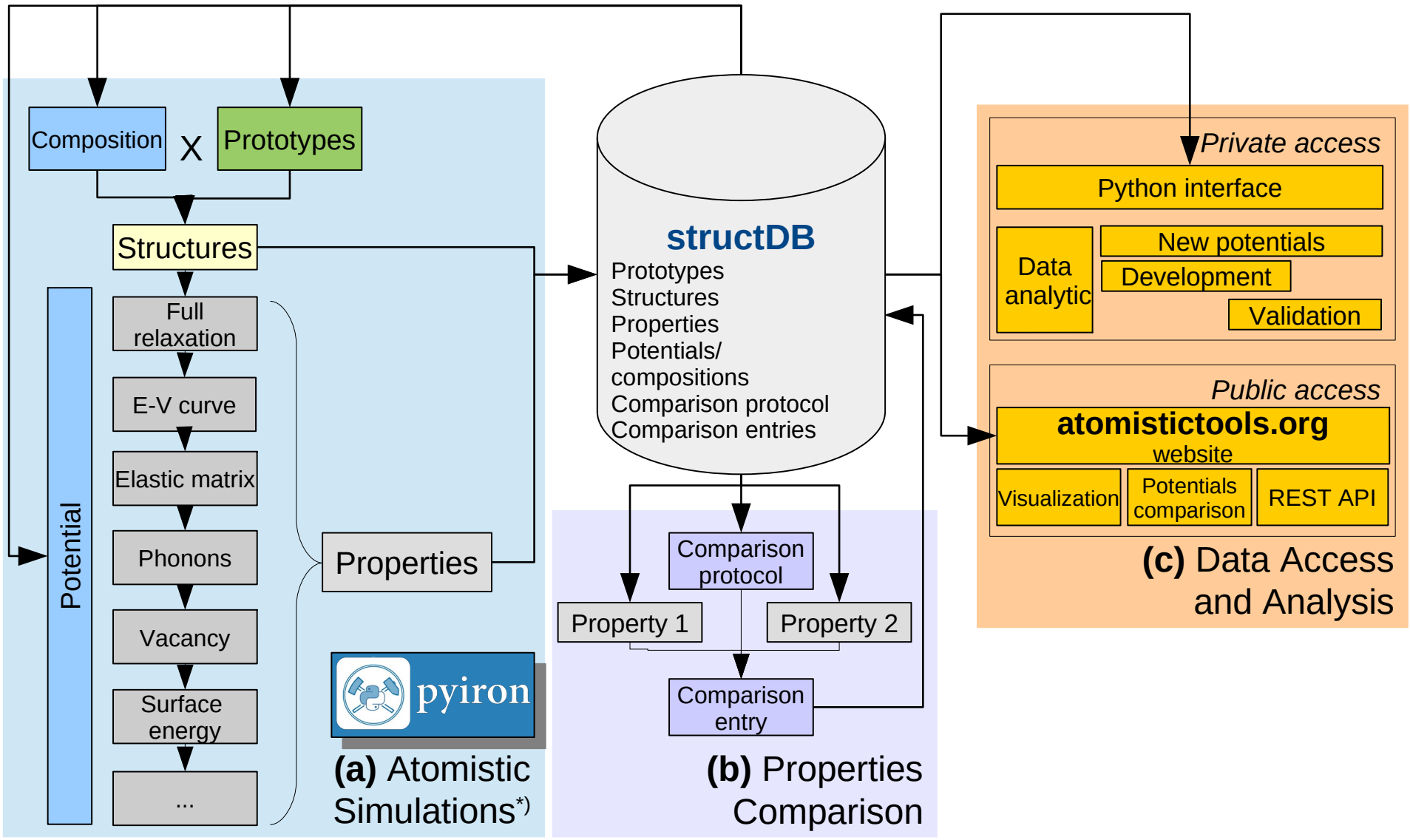


Data management: structDB



Interatomic potentials validation

Data management: structDB

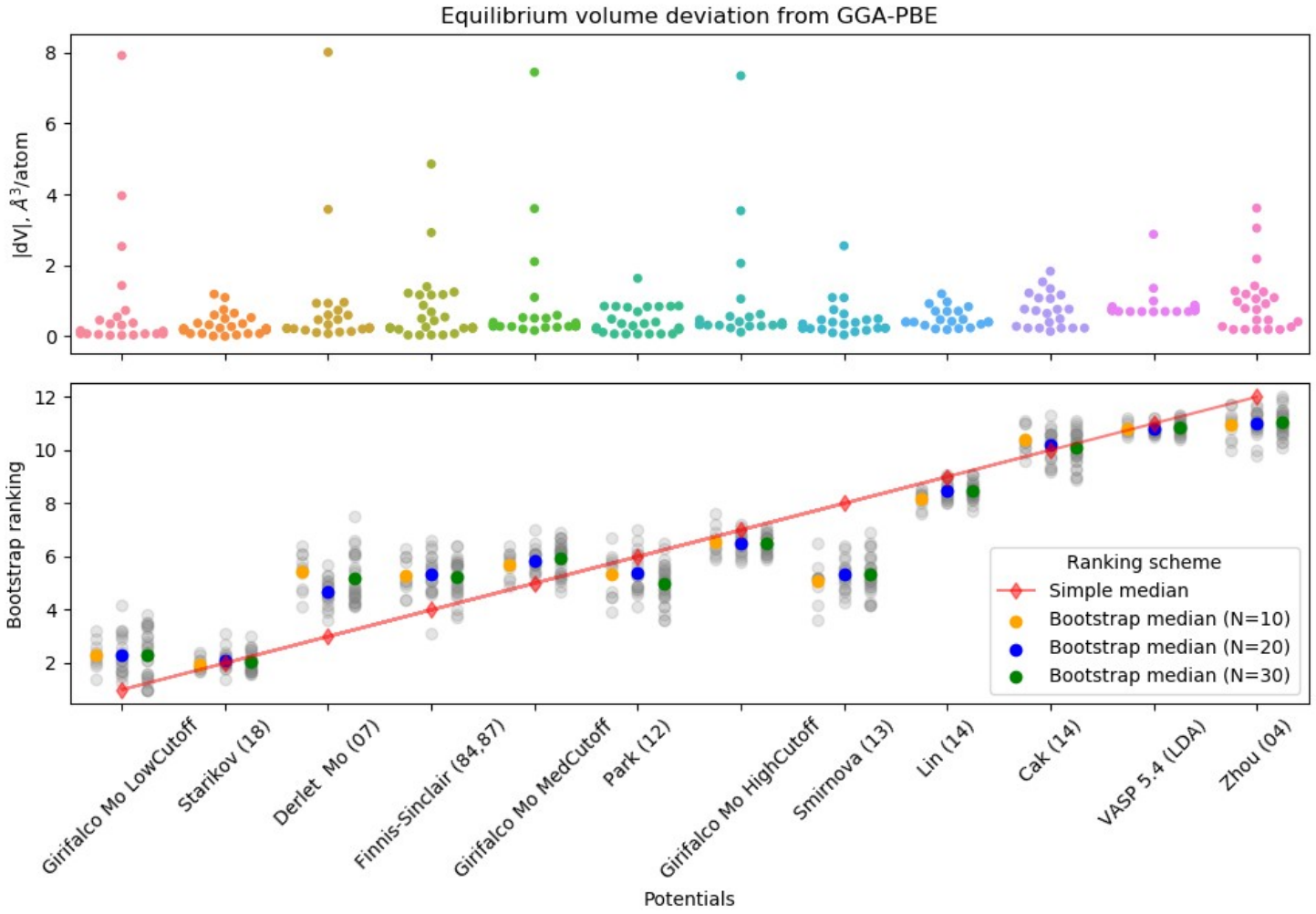


*) calculations are done with *pyiron*

Robust comparison of properties error distributions

Problem: Stability and robustness of potentials comparison wrt. outliers and incomplete sampling

Solution: Bootstrapping of error distributions → averaging → robust continuous ranking



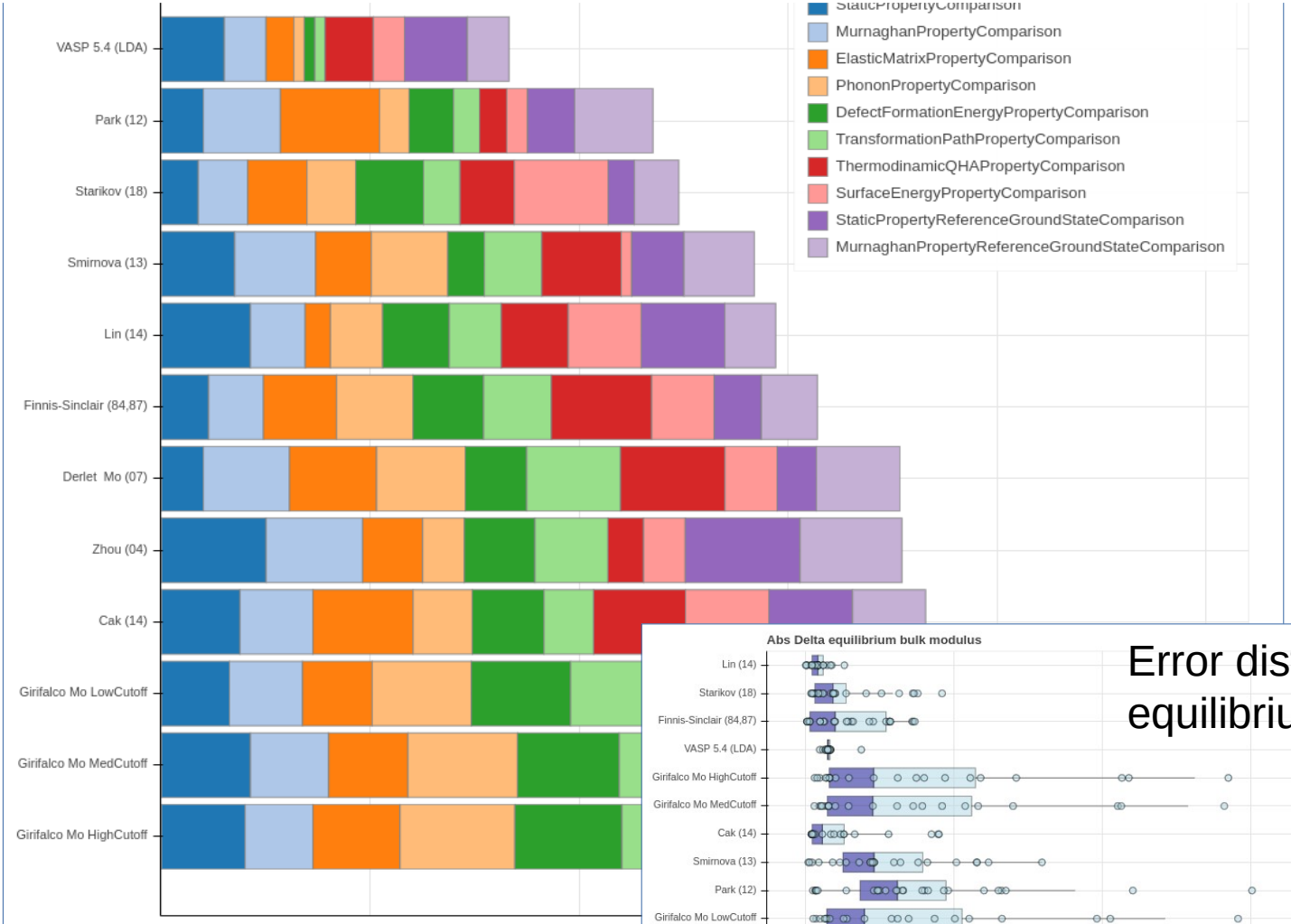
Points - errors for different crystal prototypes
(up to 29 points for the Strukturbericht A group)

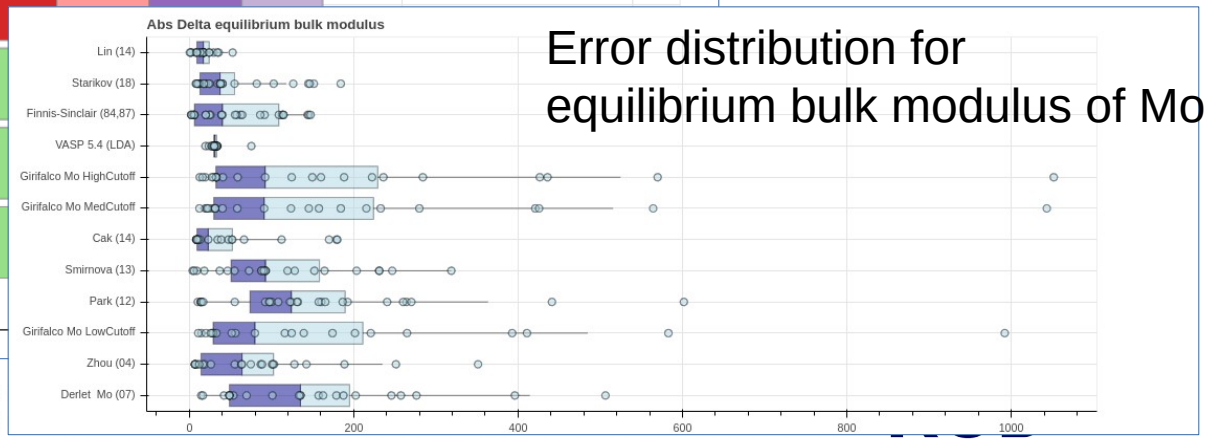
- Bootstrap error distributions 10, 20 and 30 times
- Convergence of the bootstrap ranking
- “Continuous” average rank

Atomistictools.org demo

atomistictools.org

Interatomic potentials performance in comparison wrt. DFT-PBE for Mo





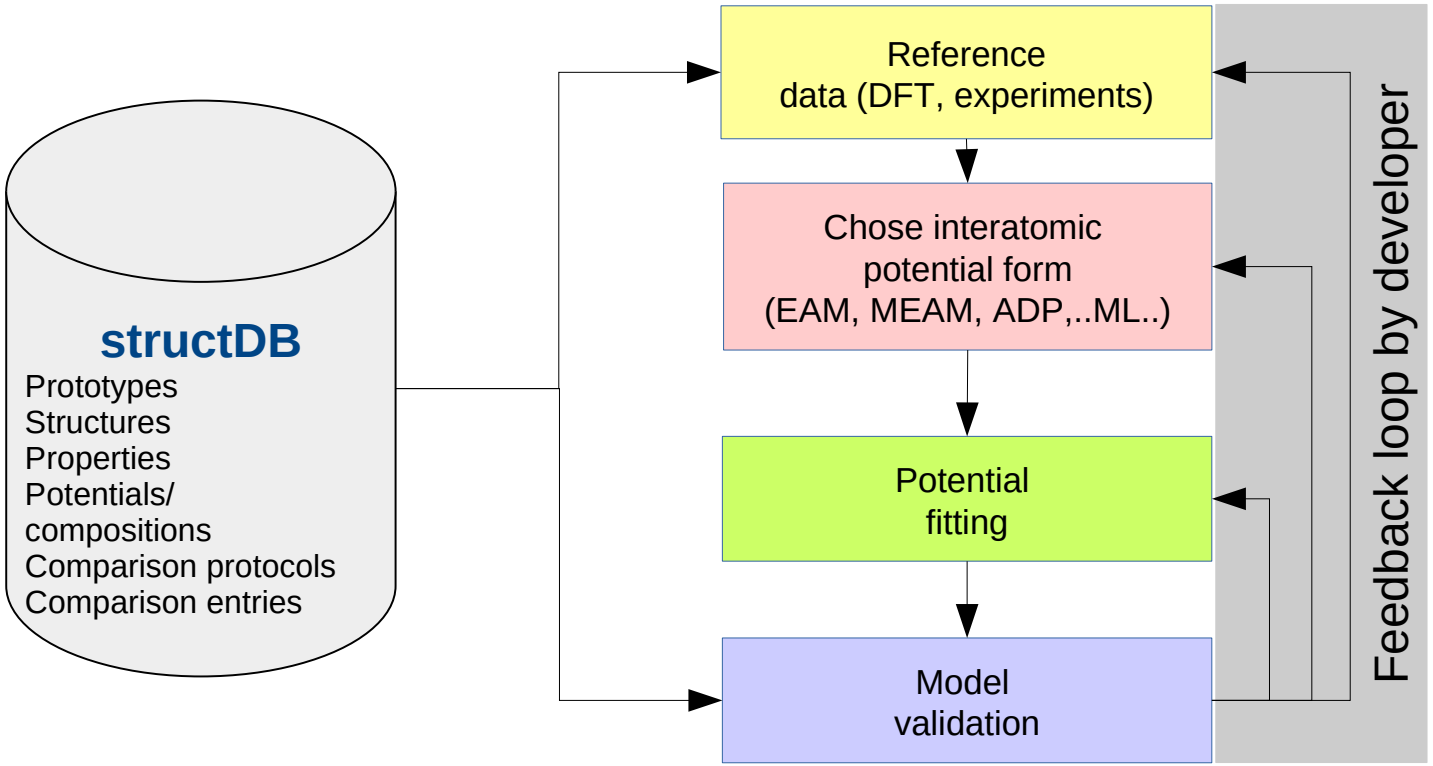
Error distribution for equilibrium bulk modulus of Mo

***YL**, Thomas Hammedschmidt, Ralf Drautz, Validation of the interatomic potentials for the unary compounds, in preparation*

Interatomic potentials development

with Alberto Ferrari and Sergei Starikov

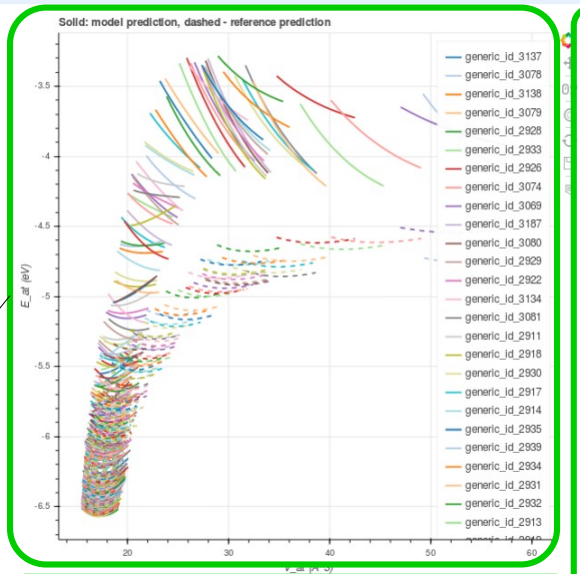
Interatomic potentials development process



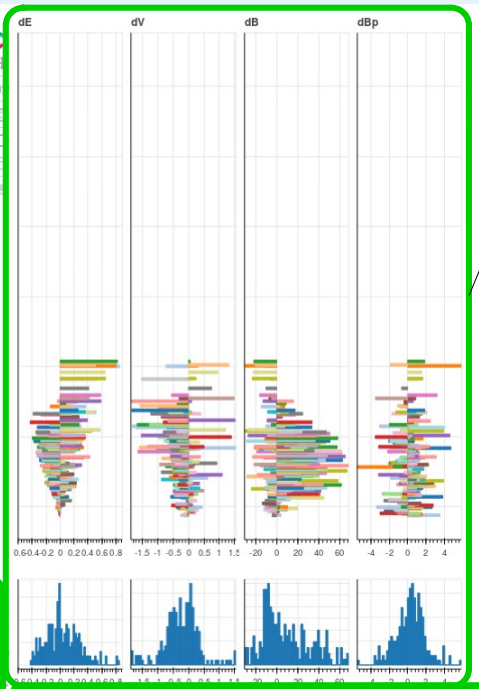
Bond order potential for Ti (by Alberto Ferrari)

Work environment screenshot

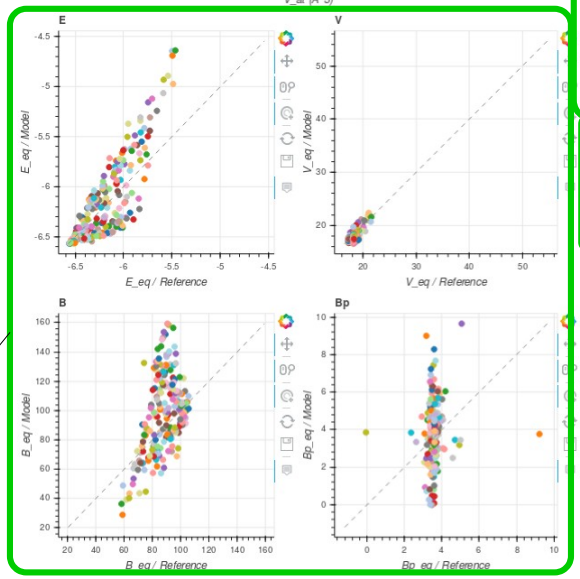
E-V curves for ~200 random structures



Error distributions for diff. properties



#	Stats	Mean	Std.	Min	Max
0	dE	0.191467948456...	0.160001832851...	0.001084865303...	0.8543508699
1	dV	0.455704448278...	0.424010841633...	0.001288255078...	1.8683717418
2	dB	19.06055198150...	16.38544426729...	0.019944958006...	68.707641602
3	dBp	1.253239861371...	1.057479813722...	0.006307481397...	5.8291087496



#	names	E_mod	E_ref	dE	V_mod	V_ref	dV	B_mod	B_ref	dB	Bp_mo	Bp_ref	dBp
0	gen...	-6.5...	-6.5...	-0.0...	17....	17....	-0.1...	101...	105...	-4.2...	4.8...	3.4...	1.4...
1	gen...	-6.5...	-6.5...	-0.0...	16....	17....	-0.2...	109...	105...	4.2...	6.9...	3.4...	3.5...
2	gen...	-6.5...	-6.5...	-0.0...	17....	17....	-0.0...	105...	104...	1.0...	5.0...	3.4...	1.5...
3	gen...	-6.5...	-6.5...	-0.0...	17....	17....	-0.1...	99....	103...	-4.3...	4.0...	3.4...	0.5...
4	gen...	-6.5...	-6.5...	-0.0...	17....	17....	-0.0...	100...	104...	-3.7...	2.8...	3.4...	-0.5...
5	gen...	-6.5...	-6.5...	-0.0...	17....	17....	-0.0...	104...	104...	0.8...	4.7...	3.3...	1.3...
6	gen...	-6.5...	-6.5...	-0.0...	17....	17....	-0.0...	100...	103...	-3.2...	0.4...	3.4...	-2.9...
7	gen...	-6.5...	-6.5...	0.0...	17....	17....	0.2...	98....	100...	-2.4...	3.0...	4.0...	-1.0...
8	gen...	-6.5...	-6.5...	-0.0...	17....	17....	0.1...	101...	98....	2.4...	2.4...	3.1...	-0.7...
9	gen...	-6.5...	-6.5...	-0.0...	17....	17....	-0.0...	95....	103...	-8.3...	2.6...	3.4...	-0.7...
10	gen...	-6.5...	-6.5...	-0.0...	17....	17....	0.0...	105...	103...	1.4...	0.8...	3.3...	-2.5...
11	gen...	-6.5...	-6.5...	-0.0...	17....	17....	0.0...	99....	103...	-3.6...	2.2...	3.4...	-1.2...
12	gen...	-6.5...	-6.5...	-0.0...	17....	17....	0.0...	95....	102...	-6.3...	2.4...	3.4...	-0.9...
13	gen...	-6.5...	-6.5...	0.0...	17....	17....	0.3...	91....	103...	-12....	1.9...	3.4...	-1.4...

Potential quality metrics

Potential-vs-DFT properties pairplots

*A. Ferrari, M. F. Schröder, **YL**, et. al., Phase transitions in titanium with an analytic bond-order potential MSMSE, 27, 085008, (2019)*

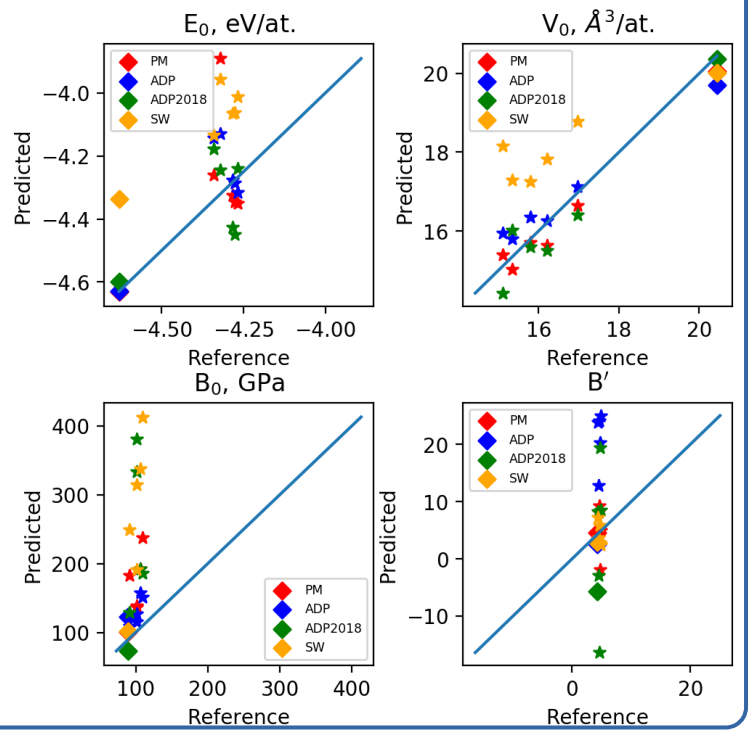
Angular dependent potential potential for Si - validation (by Sergei Starikov)

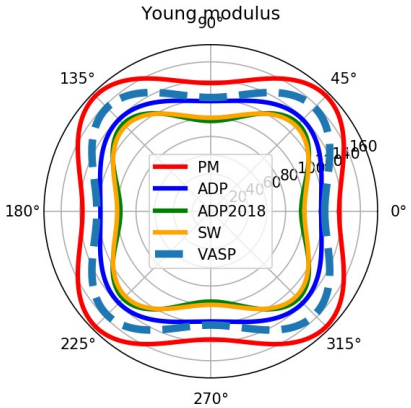
Independent validation of the new potential for Si wrt. reference DFT data (VASP) and comparison with existing potentials

New potential:
ADP2018: ADP Starikov 2018

Existing potentials:
PM: 2017—Purja-Pun-G-P-Mishin-Y--Si
ADP: ADP_Starikov_2017
SW: Three_Body_Stillinger_Weber_Si__MO_405512056662_003

E-V curves for low-energy structures



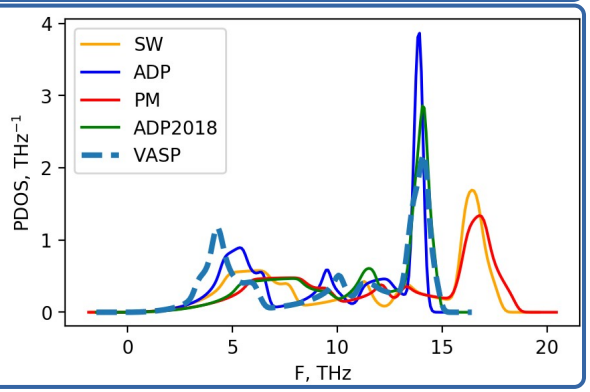


Elastic properties for Si-diam

SHORT_NAME	d_C	d_BH	d_EH	d_abs_Young
PM	41.0267	11.7134	17.1394	17.1507
SW	56.5568	12.5396	26.9793	26.5838
ADP2018	58.3122	16.2173	24.2326	24.7337
ADP	102.143	32.405	16.0617	15.0352

Phonon density-of-states for Si-diam

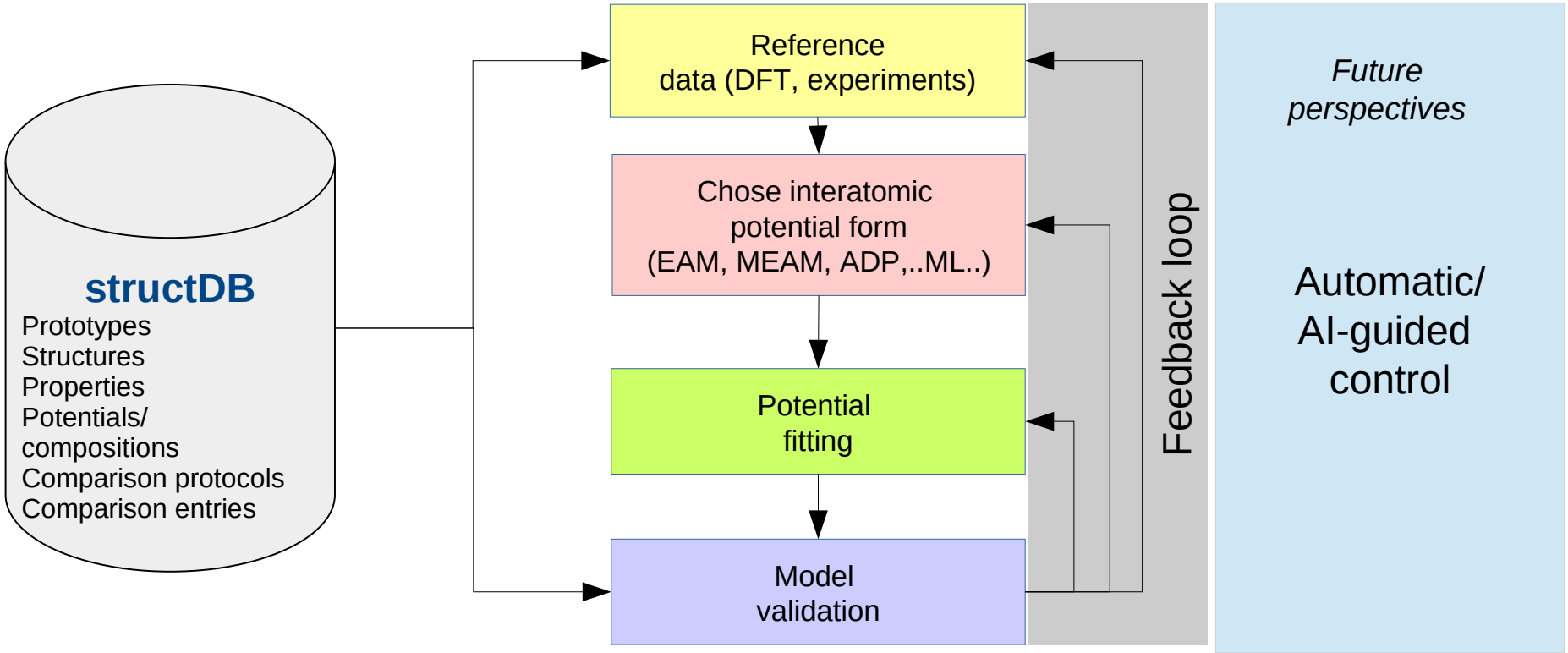
SHORT_NAME	d_DOS_MAE	d_freq_MAE	d_Cv_MAE
ADP	0.0309421	0.377617	0.0341284
ADP2018	0.0283281	1.11505	0.309852
SW	0.0486096	1.68497	0.451345
PM	0.0513959	2.39533	0.647455



*S. Starikov, I. Gordeev, **YL**, L. Kolotova, S. Makarov, Optimized interatomic potential for study of structure and phase transitions in Si-Au and Si-Al systems, Comp.Mat.Sci. 184 (2020): 109891.*

Interatomic potentials development - future perspectives

Automatic/ML/AL/AI-guided fitting procedure based on the potentials quality metric



Use ML/AI to create physical models

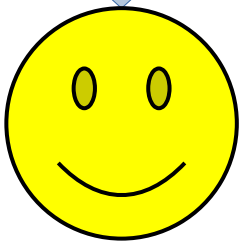
Future perspectives: distributed resources

Future perspectives: use-cases and roles

Atomistic simulations distributed resources
(NIST IPR, OpenKIM, atomistictools.org, ...)

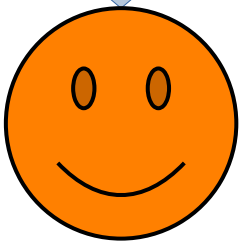
Web interface, REST-API, API

- **Interatomic potentials (IP)**
- **Models validation**
- Protocols validation



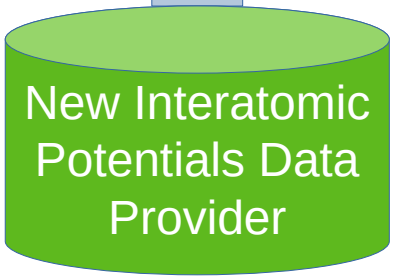
Molecular dynamics simulation user

- **Provide new IP**
- Get test protocols for new IP
- Get reference data for fitting and validation

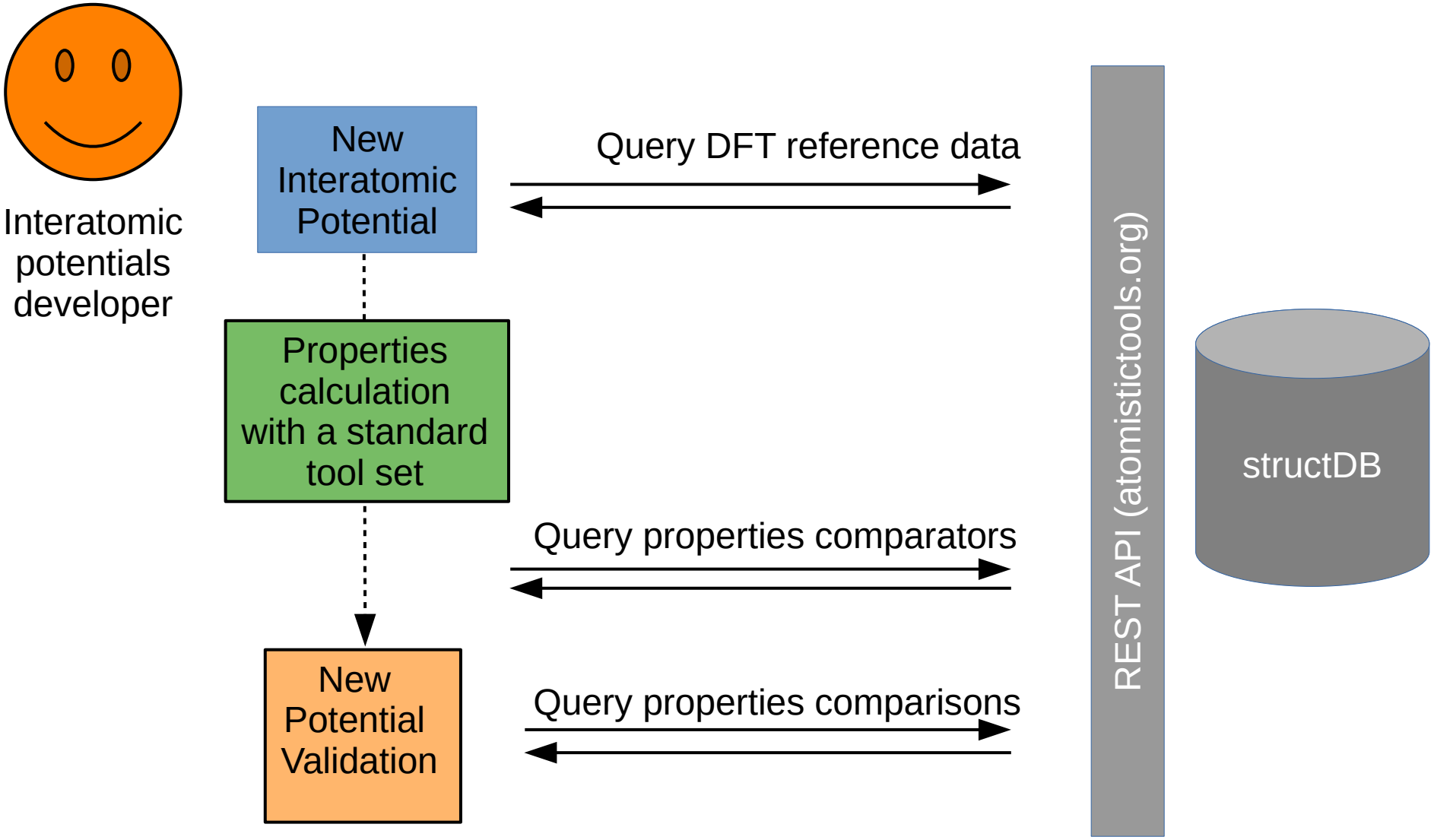


Interatomic potentials developer

Provide new data, validation tests, resources, protocols, etc.



REST API use scenario: new potential validation



REST API interface (proof-of-concept)

Provide functionality for querying the interatomic potentials validation database without direct connection to database (via WEB)

- Crystal prototypes (generic)

```
generics=db_rester.query_generic(prototype_strukturbericht=”A2”)
```

- Structures

```
structures=db_rester.query_structure(prototype_strukturbericht=”A2”,
                                     calculator_name="VASP%PBE%",
                                     structure_composition="Mo%")
```

- Computed properties (i.e. reference DFT data)

```
reference_data=db_rester.query_property(property_type_name="murnaghan",
                                        structure_composition="Si%",
                                        prototype_strukturbericht="%",
                                        calculator_name="VASP%PBE%")
```

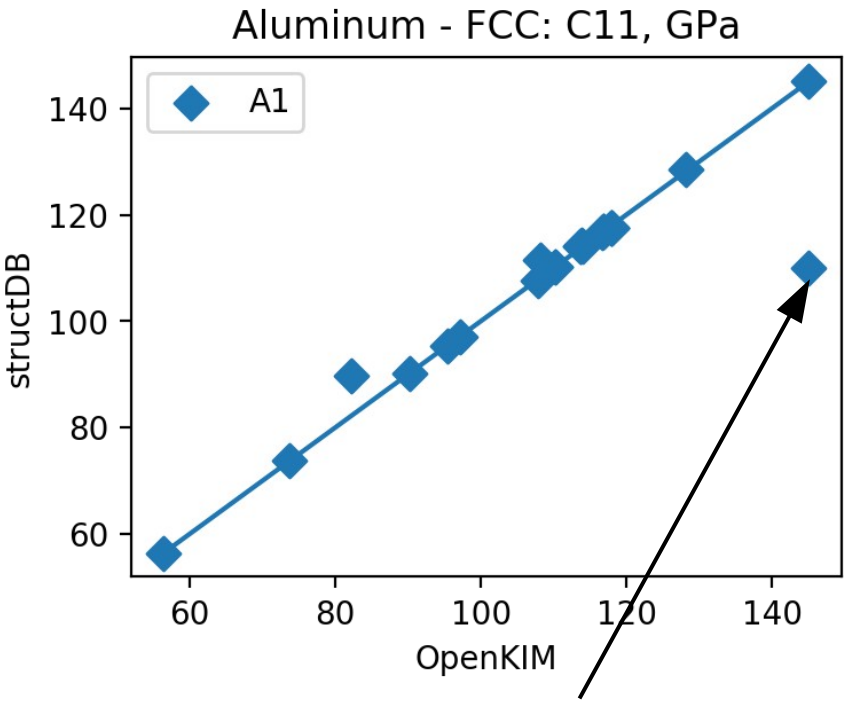
- Properties comparison protocol (comparator)

```
comparator=db_rester.query_comparator(property_type_name="phonons")
```

Data consistency? (open question)

Distributed query to OpenKIM & atomistictools.org*

Query elastic constants of fcc aluminum

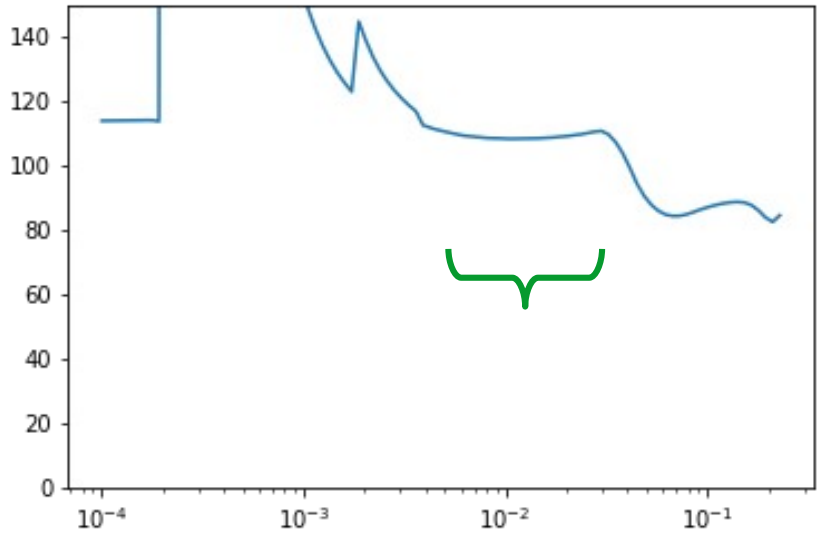


EAM_Dynamo_ZhouJohnsonWadley_2004_Al

- OpenKIM uses small strain range (~1e-8) and numerical derivatives
- Numerical noise of the model implementation
- Instability of properties computational protocol

Each property calculation protocol has it's "hyperparameters", i.e. strain range, number of points, fit order for elastic matrix calculations

EAM_Dynamo_ZhouJohnsonWadley_2004_Al MO_131650261510_005



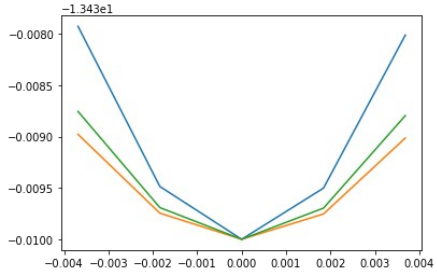
Strain range: 10^{-5} - $2*10^{-1}$

*) *proof-of-concept*

Properties calculation hyper-parameters

Example: elastic matrix property via energy-strain curves

ElasticMatrixCalculator(
num_of_point=5,
eps_range=0.005,
fit_order=2,
optimize_deformed_structure=True,
fmax=0.005)



“hyper-parameters”

- Physically-inspired suggestions (max **num_of_point**, min **eps_range**, min **f_max**) not always correct, due to numerical noise, etc.

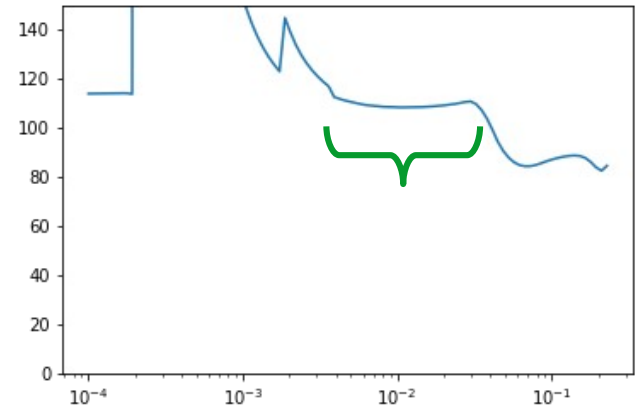
- Sensitivity analysis should be done for each parameter and property type:
  - vary hyperparameters combinatorially
  - compute property
  - find “safety spot” in hyper-parameter space (stable predictions)
- Does this “spot” is transferable to another potentials types and elements?

Properties calculation hyperparameters

Each property calculation protocol has it's "hyperparameters".
Ex.: Elastic matrix calculation: strain range, number of points, fit order, ...

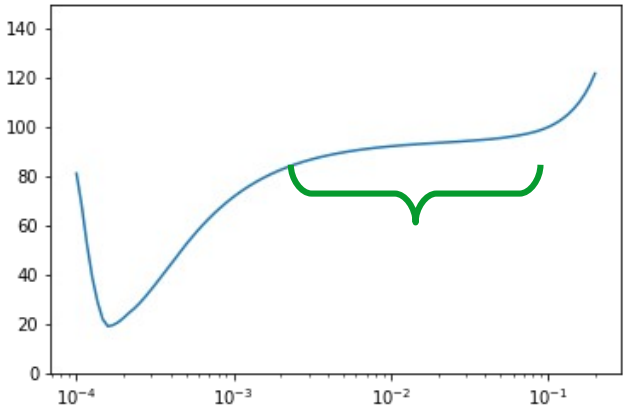
Aluminum-fcc: C11 constant

EAM_Dynamo_
ZhouJohnsonWadley_
2004_Al
MO_131650261510_005



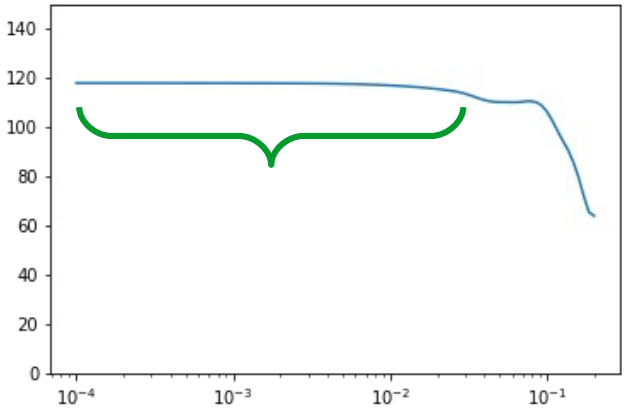
Strain range: 10^{-5} - $2*10^{-1}$

Morse_Shifted
GirifalcoWeizer_
1959HighCutoff_Al
MO_140175748626_003



Strain range: 10^{-5} - $2*10^{-1}$

EAM_
ErcolessiAdams
1994_Al
MO_324507536345_002



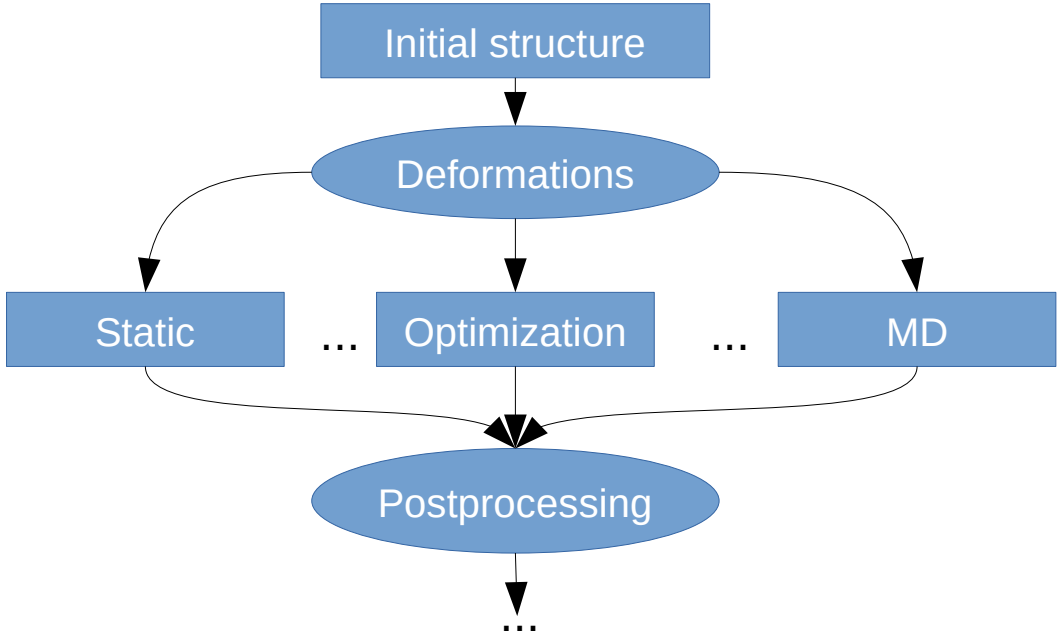
Strain range: 10^{-5} - $2*10^{-1}$

There are an overlapping of "safety-spots" of hyperparameters for properties calculations for different potentials and elements

Generalization of properties calculation protocols

1. Properties calculation protocols are mostly general:

- Initial structure(s)
- Deformations (optional)
- perform
  - Static calculation
  - Optimization
  - Molecular Dynamic
- Post-processing
- … repeat (optional)



2. Different execution "backends/engines":

- ASE, Fireworks/Atomate, AiiDA, AFLOW, pyiron, Atomman, …

Separation between **protocol** and **computation engine(s)** is needed for unification!

- Unified "**language**" to describe the **atomic properties** calculation protocol
- Mapping to different execution "backends"
- Way to achieve interoperability between data producers and data storage
- **Standardize** atomistic simulations

Concluding remarks

Concluding remarks

- Interatomic potentials validation by high-throughput **properties comparison (transferability)**
- Common protocols & tests for interatomic potentials validation and DFT
- Testing of publicly available potentials:
  - **OpenKIM** & **NIST** potentials repositories
- Potential comparison as a service for community:
  - **atomistictools.org** (testing stage, password: "**icams"**)
- Comparison wrt. reference data: usually DFT (VASP 5.4 PBE) or experiment (Pearson Crystallographic Database, Landolt-Boernstein, …)
- REST API with compatible interfaces for unified queries

Further development

- Production stage of **atomistictools.org** (for unaries)
- Increase the amount of properties & comparison metrics
- Sampling strategies for binaries and ternaries compounds
- Data analysis of the potentials transferability, identification of the key metrics
- Testing for ML potentials

Thank you for attention