Duke UNIVERSITY

AFLOW

Automatic-FLOW for Materials Discovery

a distributed materials genome properties repository from high-throughput

From meta-data to meta-properties with only a bit of meta-physics

Stefano Curtarolo

Materials Science - Duke University – USA

Center for Materials Genomics – USA

C. Toher, C. Oses, R. Friedrich, D. Hicks, F. Rose

DEPARTMENT OF THE NAVY ONR Science & Technology

NSF

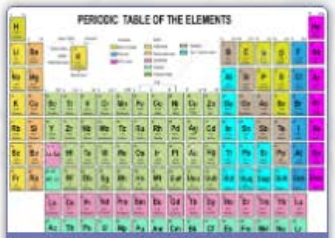
U.S. DEPARTMENT OF ENERGY

UC San Diego

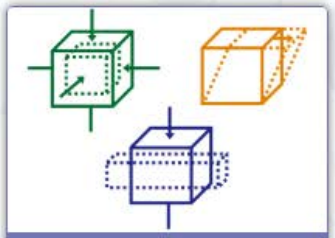
AFLOW.org: data/ metadata and applications

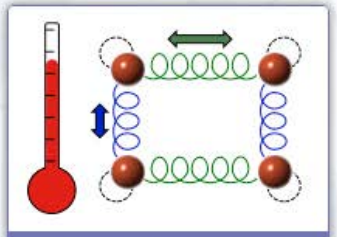
aflow.org

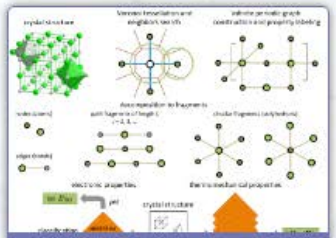
Apps and Docs



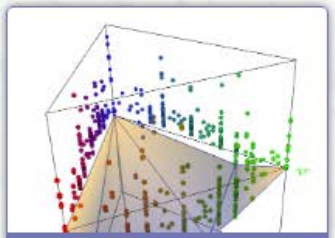
MendeLIB search


Elastic properties

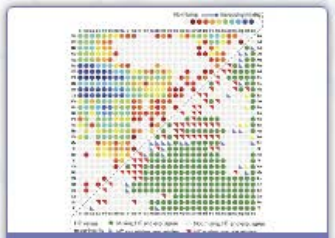

Thermal properties

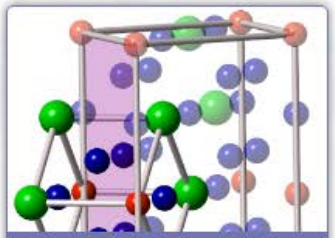

AFLOW-ML

Crystal prototypes


AFLOW-CHULL


AFLOW-online


Binary alloy library


Superalloys search

AFLOW REST-API WIKI

REST-API Docs

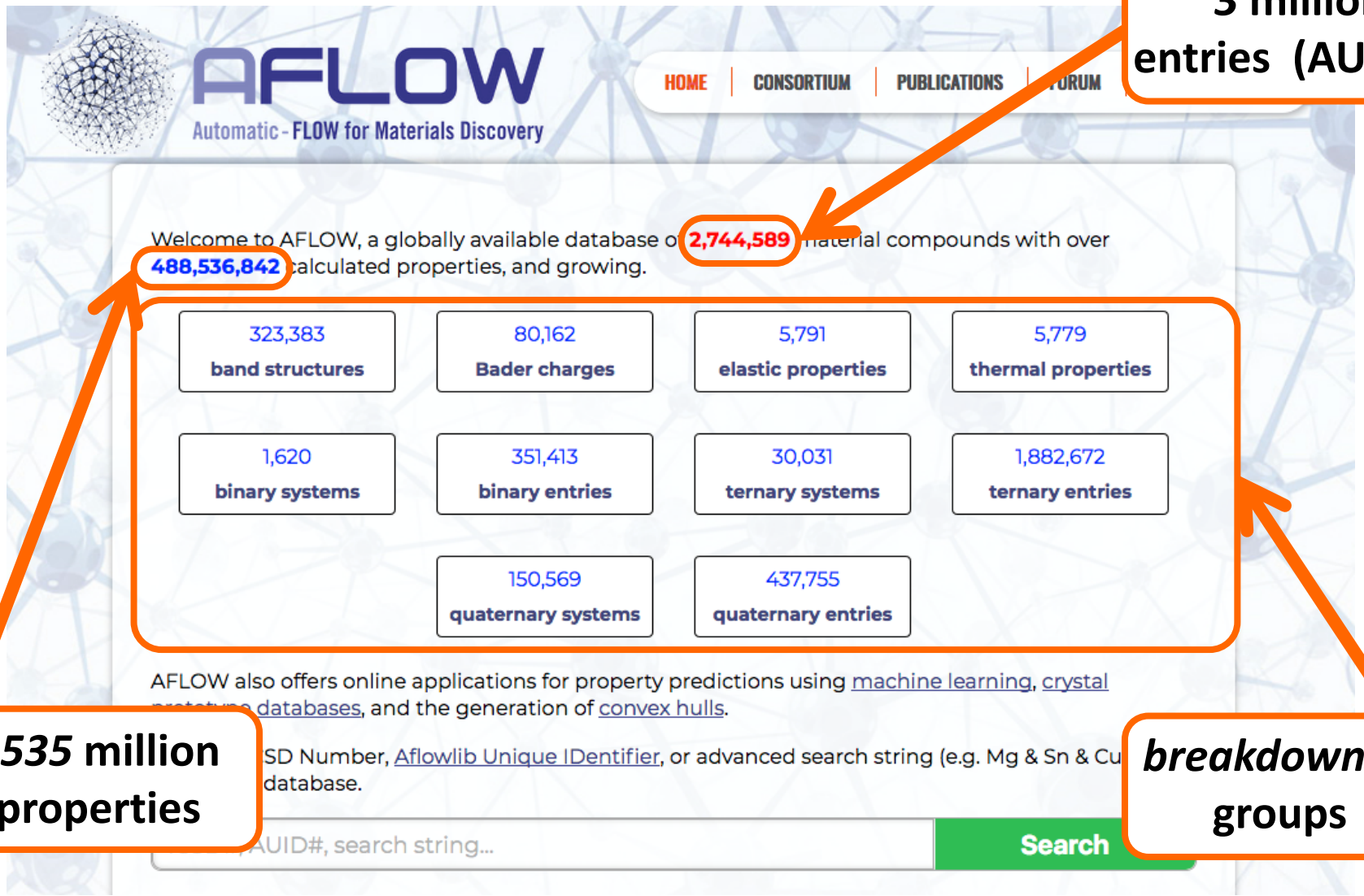
AFLOWπ

PAOFLOW

Curtarolo et al., Comput. Mater. Sci. ***58****, 218 (2012); Calderon et al., Comp. Mat. Sci.* ***108A****, 233 (2015)*

AFLOW.org: web portal

aflow.org



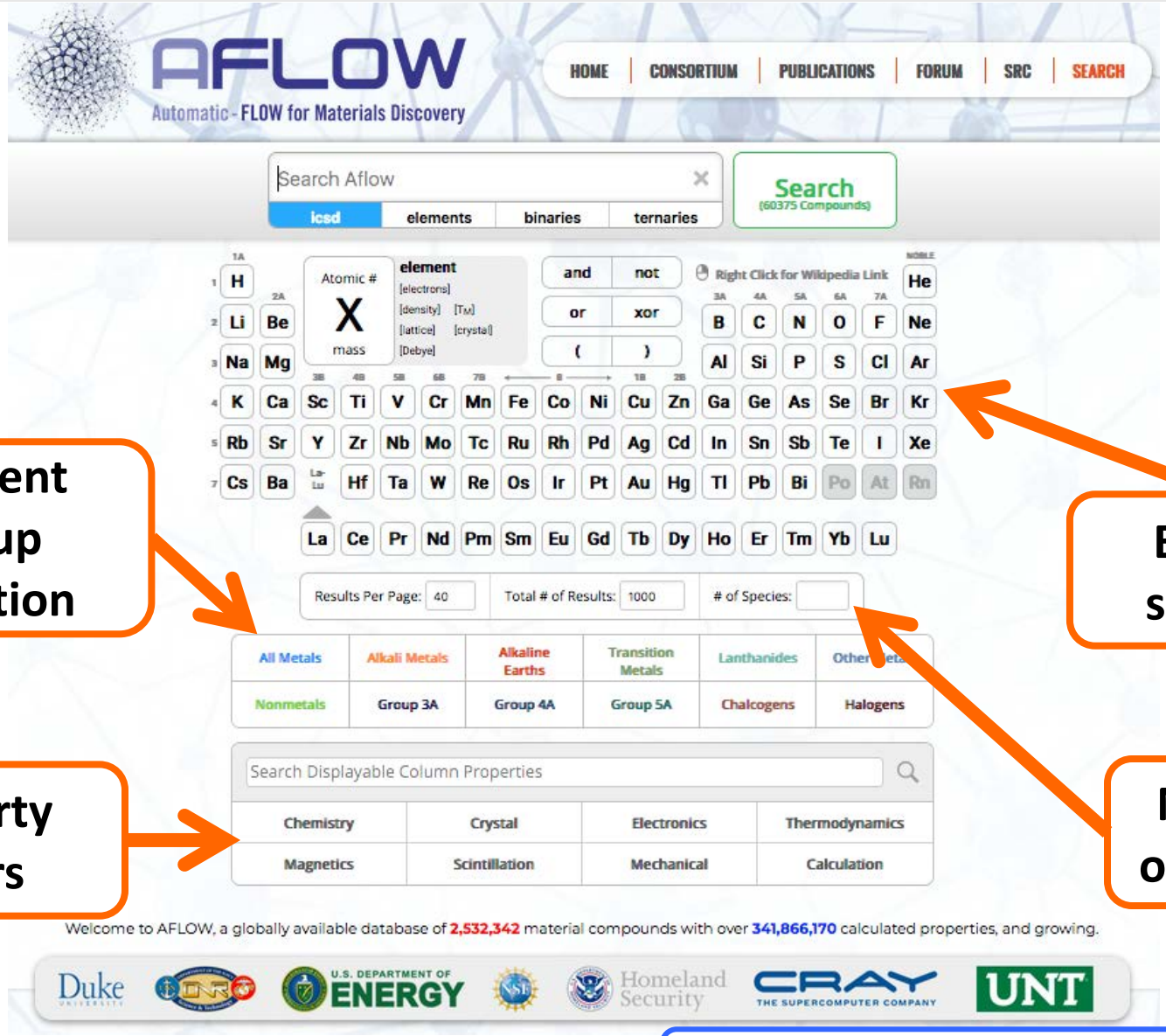
~3 million entries (AUIDs)

~535 million properties

***breakdown* of groups**

Curtarolo et al., Comput. Mater. Sci. ***58****, 218 (2012); Calderon et al., Comp. Mat. Sci.* ***108A****, 233 (2015)*

AFLOW.org: search page



Element group selection

Element selection

Property Filters

Number of species

aflow.org/advanced.php

R. H. Taylor *et al.*, Comput. Mater. Sci. **93**, 178-192 (2014)

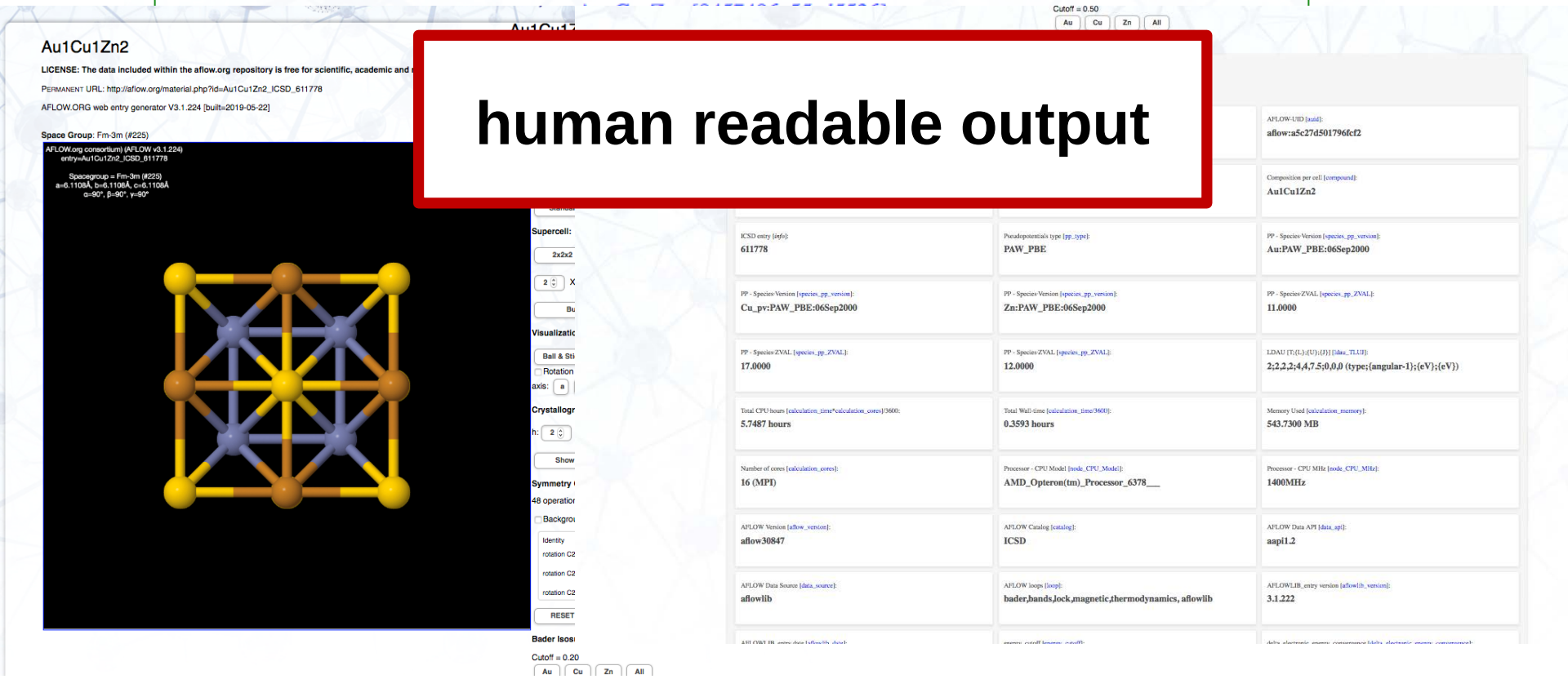
Web portal for HUMANS

searches

[First]

ENTRY
$As_2Cu_1O_8Zn_2$ [75f66067131a579e]
$As_4Cu_4H_4O_{20}Zn_4$ [521293f6fc9719f4]
$Au_1Cu_1Zn_2$ [a5c27d501796fcf2]

human readable output



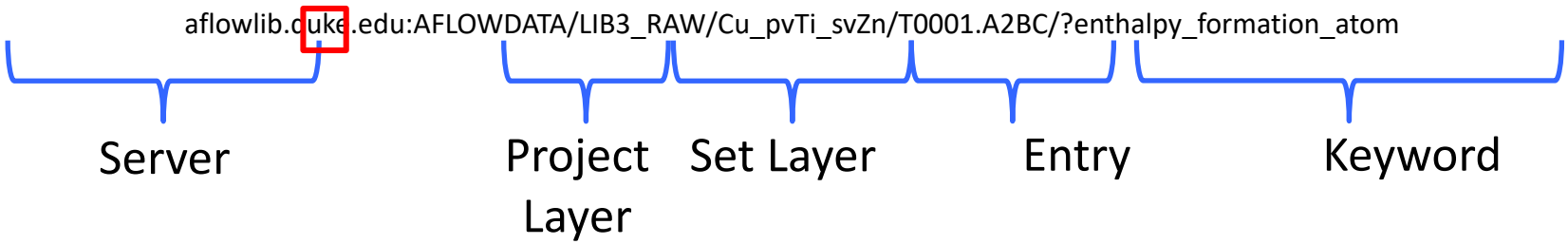
Curtarolo et al., Comput. Mater. Sci. ***58****, 218 (2012); Calderon et al., Comp. Mat. Sci.* ***108A****, 233 (2015)*

objects for machines: REST-API

- The AFLOW REST-API AFLOW URL (AURL) has the form:

<server>:AFLOWDATA/<project>/<set>/<entry>/?<keyword>

- *The AURL for the formation enthalpy per atom of the ternary Heusler structure T0001.A2BC with composition Cu_2TiZn is:*



- This becomes the AURL:

aflowlib.duke.edu/AFLOWDATA/LIB3_RAW/Cu_pvTi_svZn/T0001.A2BC/?enthalpy_formation_atom

- The full list of entries in the set Cu-Ti-Zn can be retrieved using:

aflowlib.duke.edu/AFLOWDATA/LIB3_RAW/Cu_pvTi_svZn/?aflowlib_entries

R. H. Taylor *et al.,* Comput. Mater. Sci. **93**, 178-192 (2014)

primitive metadata: REST-API

- Example python script to access API & download space groups (see tutorial script *restapi*.py *from* af*low_api.txz*):

```
import json
from urllib.request import urlopen

SERVER = 'http://aflowlib.duke.edu/AFLOWDATA'
PROJECT = '/LIB3_RAW'
SET = '/Cu_pvTi_svZn'
ENTRIES = '/?aflowlib_entries'

response = urlopen(SERVER+PROJECT+SET+ENTRIES).read().decode('utf-8')

entry_list = response.rstrip().split(',')


for entry in entry_list:
    entry_sg_url = SERVER + PROJECT + SET + '/' + entry + '/?spacegroup_relax' # spacegroup relax URL
    entry_sg = urlopen(entry_sg_url).read().decode('utf-8')
    print('{0}:   {1}'.format(entry, entry_sg))
```

Project: Ternary alloys

Set: CuTiZn alloy system

Keyword: List of entries in this alloy system

Download list of entries with urlopen

Loop through list of entries

Download space group for each entry

R. H. Taylor *et al.*, Comput. Mater. Sci. **93**, 178-192 (2014)

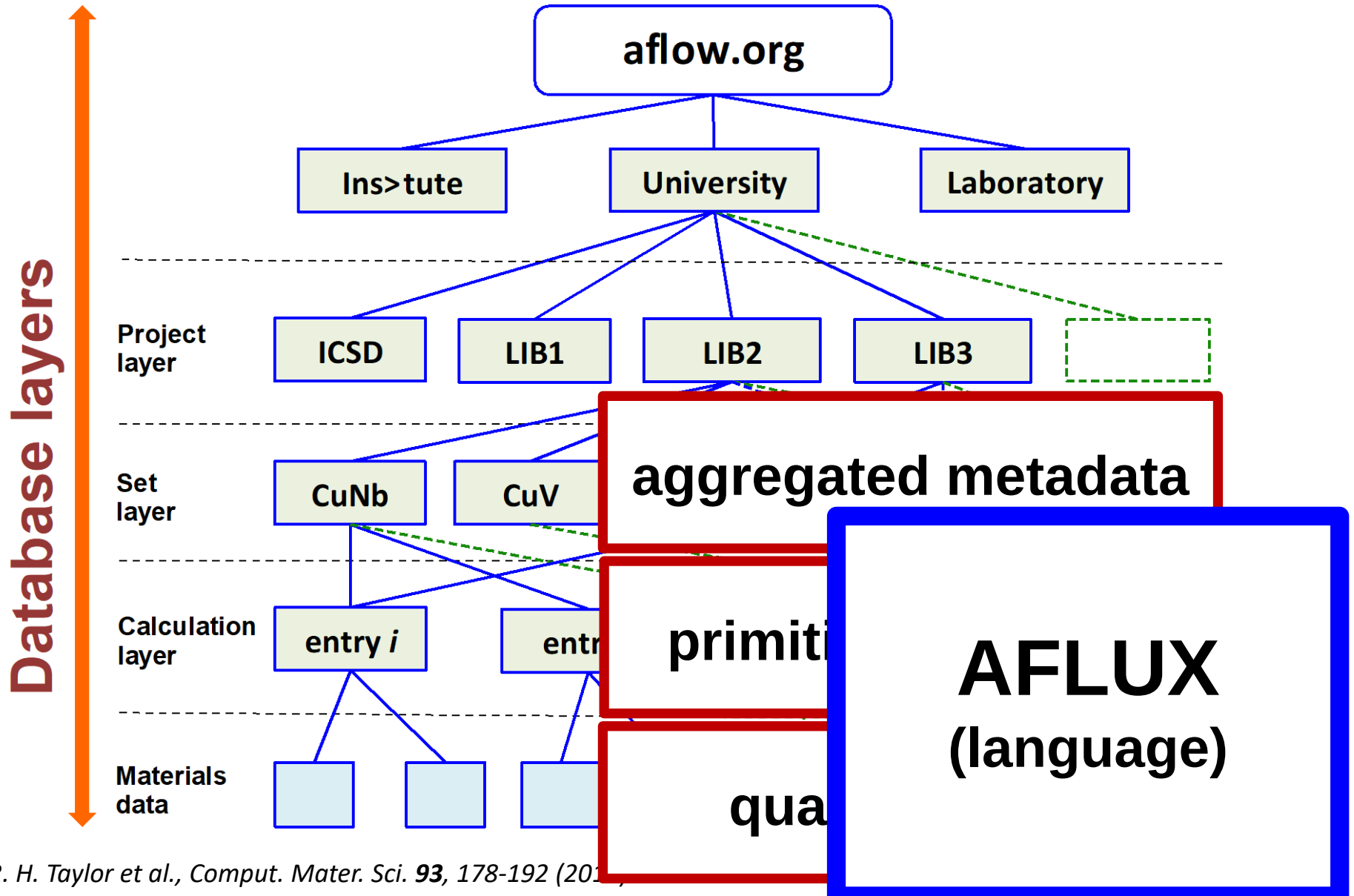
primitive metadata for SYMMETRY: AFLOW-SYM

- *C*alculates symmetry descriptions for
  - *l*attice : lattice points
  - *superlattice* *: atomic sites are decorated with single species*
  - *reciprocal lattice* *: reciprocal lattice points*
  - *c*rystal : lattice points + atomic sites
  - *crystal spin* *: lattice points + atomic sites + magnetic moment*

- *Provides the following operator representations:*
  - *o*perator type
  - *Hermann-Mauguin*
  - Schoenflies
  - *transformation matrix (Cartesian)*
  - *t*ransformation matrix (fractional)
  - *generator matrix*
  - *s*o(3) coefficients($\mathbf{L}_x$,$\mathbf{L}_y$,$\mathbf{L}_z$)
  - *angle*
  - *a*xis
  - *quaternion (vector)*
  - *quaternion (2 x 2 matrix)*
  - *quaternion (4 x 4 matrix)*
  - *s*u(2) coefficients (Pauli)
  - *inversion boolean*
  - *i*nternal translation (Cartesian)
  - *internal translation (fractional)*
  - *atom index map*
  - *atom type map*
  - *l*attice translation (Cartesian)
  - *lattice translation (fractional)*

D. Hicks *et al.,* Acta Cryst. A**74**, 184-203 (2018)

AFLOW.org: database organization



R. H. Taylor et al., Comput. Mater. Sci. ***93****, 178-192 (201*

AFLUX: language to search for data and primitive metadata

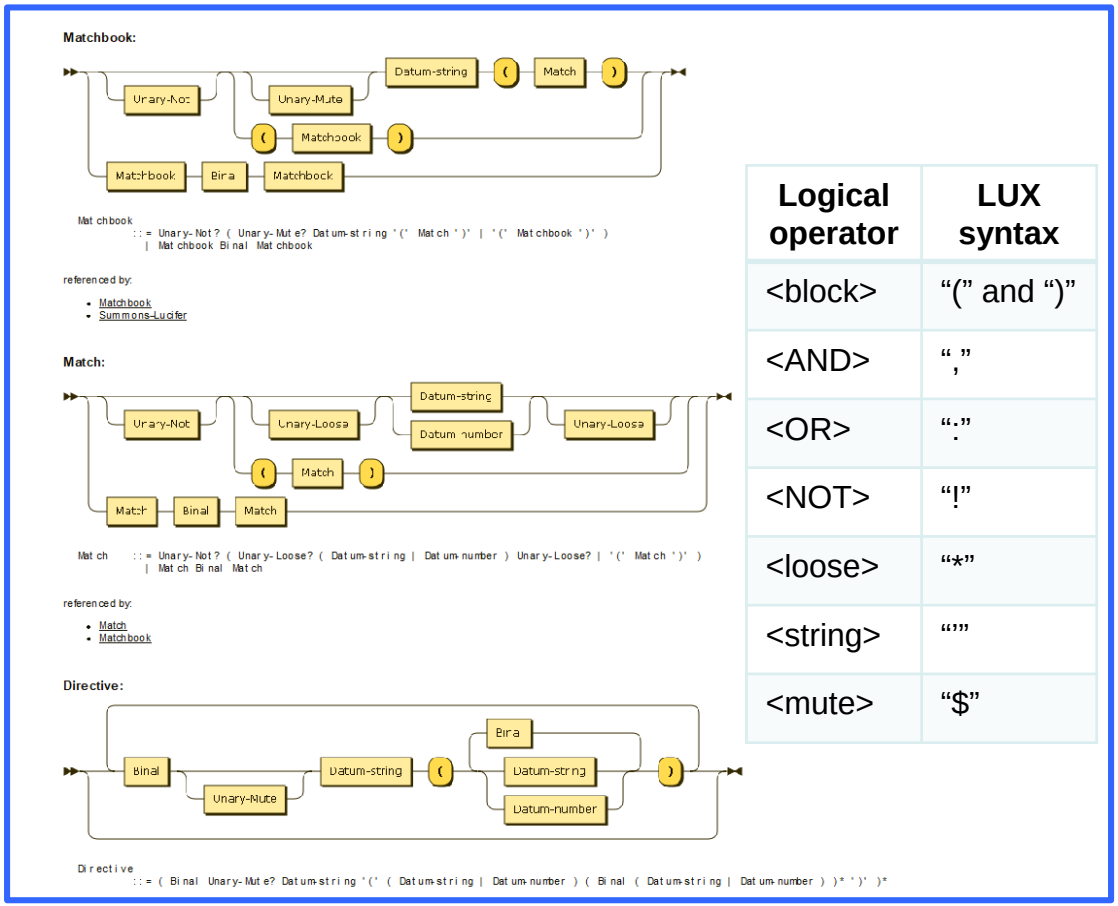
+

```
aurl=aflowlib.duke.edu:AFLOWDATA/LIB3_RAW/AgCoMn_pv/T0002.A2BC!
auid=aflow:AgCoMn_pv/T0002.A2BC:PAW_PBE!
aapi=1.0!
keywords=aurl,auid,aflow_api_version,code,compound,prototype,nspecies,...!
aflowlib_entry_date=20140130_20:34:00_GMT-5!
aflowlib_entry_version=30794!
aflow_version=aflow30293!
calculation_cores=1!
calculation_memory=639!
calculation_time=18347.2!
corresponding=Stefano_Sanvito_sanvitos@tcd.ie!
loop=thermodynamics,bands,magnetic!
node_CPU_Cores=12!
node_CPU_MHz=2661!
node_CPU_Model=Intel(R)_Xeon(R)_CPU_X5650_@_2.67GHz!
node_RAM_GB=24!
code=vasp.4.6.35!
composition=2,1,1!
compound=Ag2CoMn1!
density=8.94193!
eentropy=0!
eentropy_atom=0!
Egap=0!
energy=-20.4051!
energy_atom=-5.10128!
enthalpy=-20.4051!
enthalpy_atom=-5.10128!
enthalpy_formation=1.51248!
```

mandatory **keywords**

optional control **keywords**

optional materials **keywords**



Logical operator	LUX syntax
<block>	“(” and “)”
<AND>	“,”
<OR>	“:”
<NOT>	“!”
<loose>	“*”
<string>	“”
<mute>	“$”

- Aim: Programatically expose the same functionality as the web search interface at http://aflow.org/advanced.php

F. Rose *et al.*, Comput. Mater. Sci. **137**, 362 (2017).

AFLUX: AFLOW Search-API

- AFLUX enables search functionality through query part of URI
- Supports use of several logical operators
- Operator scope can be inter-property and/or intra-property

Logical operator	AFLUX syntax	Operator scope
<block>	"(" and ")"	Intra- and inter-property
<AND>	","	Intra- and inter-property
<OR>	":"	Intra- and inter-property
<NOT>	"!"	Intra-property
<loose>	"*"	Intra-property
<string>	"'"	Inter-property
<mute>	"$"	Intra-property

F. Rose *et al.*, Comput. Mater. Sci. **137**, 362 (2017).

AFLUX: AFLOW Search-API

http://aflowlib.duke.edu/search/API/?<matchbook>,<directives>

Search API Server | Query

- Matchbook:
  - Materials keywords with arguments
  - Example: <server>/?species((Na:K),Cl),nspecies(2),Egap(1*,*5),energy_cell
  - Keyword list available from schema directive: <server>/?schema

- Directives:
  - Formatting instructions with arguments
  - “format” takes arguments “json” and “html”
  - “catalog” takes arguments “icsd”, “lib1”, “lib2”, ...
  - “paging” controls number of entries and page displayed, sorted in ascending order of first materials keyword – order can be reversed by using a negative page number in the argument

F. Rose *et al.*, Comput. Mater. Sci. **137**, 362 (2017).

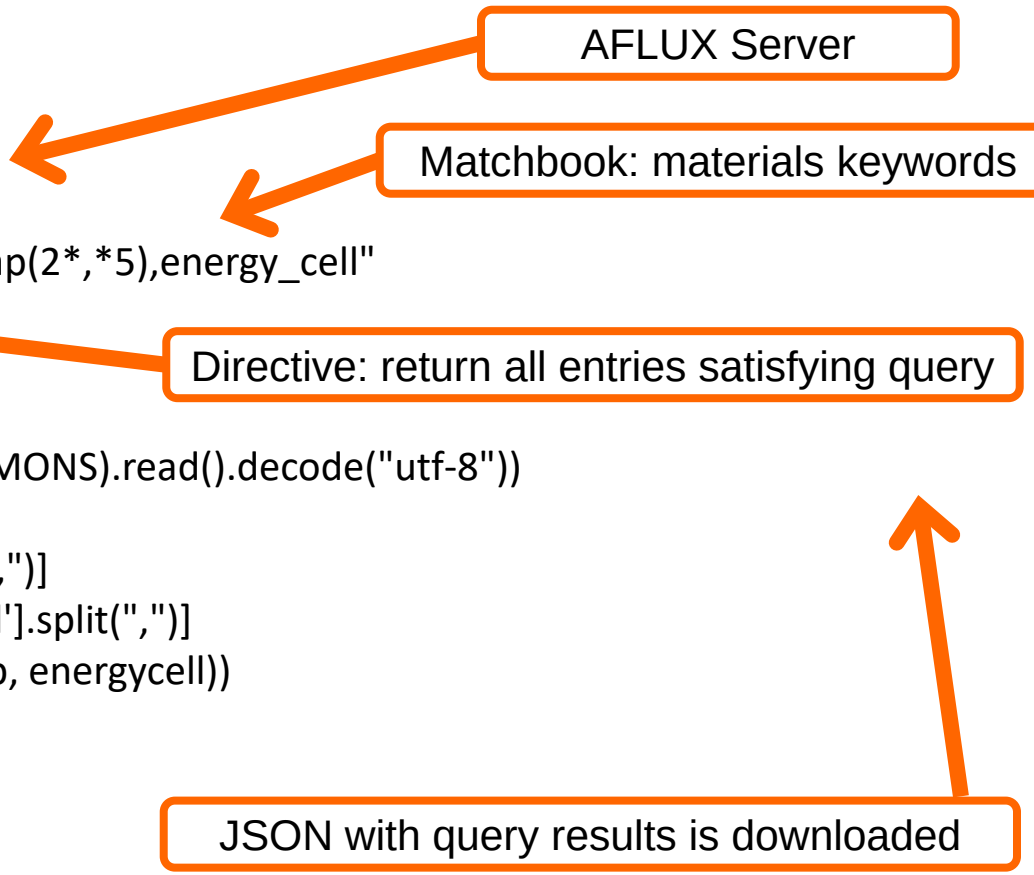
APE: AFLOW Python Environment: search-API

- Example python script to access AFLUX search-API (see tutorial script afluxapi.py from af*low_api.txz*):

```
#!/usr/bin/env python
import json, sys, os
from urllib.request import urlopen

SERVER="http://aflowlib.duke.edu"
API="/search/API/?"
MATCHBOOK="species((Na:K),Cl),nspecies(2),Egap(2*,*5),energy_cell"
DIRECTIVES="$paging(0)"
SUMMONS=MATCHBOOK+","+DIRECTIVES

response=json.loads(urlopen(SERVER+API+SUMMONS).read().decode("utf-8"))
for datum in response:
   bandgap=[float(x) for x in datum['Egap'].split(",")]
   energycell=[float(x) for x in datum['energy_cell'].split(",")]
   print ("{}, {}, {}".format( datum['auid'], bandgap, energycell))
```



F. Rose *et al.*, Comput. Mater. Sci. **137**, 362 (2017).

What is aggregate data: ontological problem

AGGREGATED METADATA

set of calculations giving a particular property

e.g. elasticity, phonons, entropy, synthesizability, etc..

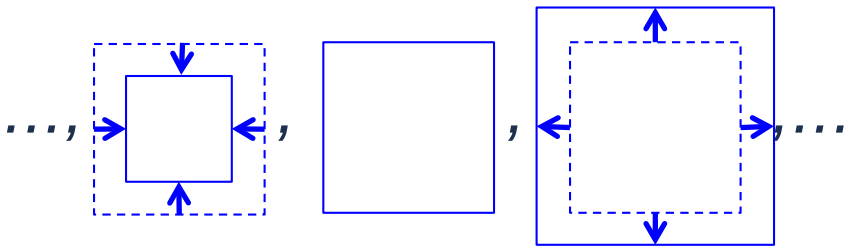
aggregated metadata requires <u>workflow of workflows</u>

- **one workflow (e.g. band-structure) => one AUID**
- **workflow of workflows (phase stability) => AUID of AUIDs**

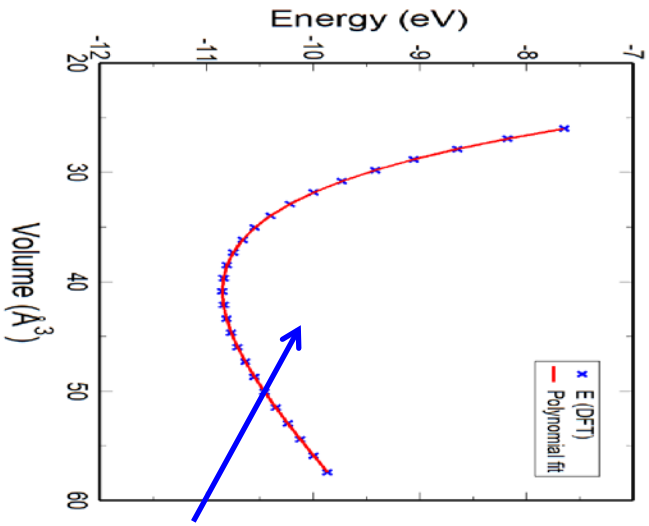
STATIC vs DYNAMIC AGGREGATED METADATA

STATIC METADATA: Debye-Grüneisen AFLOW-AGL

Different volume cells



E(V) data from DFT (VASP, QE, etc.)



Fit with polynomial or eqn of state: bulk modulus, B(V)

Debye temperature: $\theta_D(V)$

Grüneisen parameter: γ

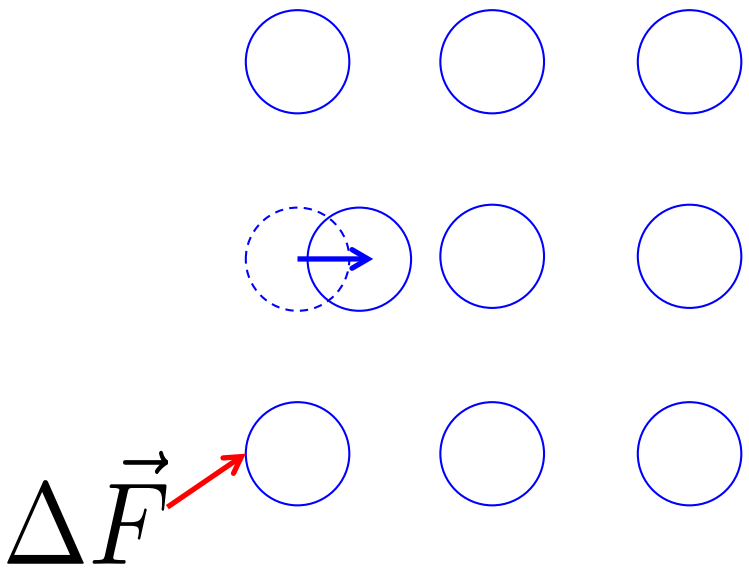
Heat capacity: C_V

Lattice thermal conductivity, κ_L

C. Toher *et al.*, Phys. Rev. B **90**, 174107 (2014); M. A. Blanco *et al.*, Comp. Phys. Commun. **158**, 57 (2004)

STATIC METADATA: Phonon calculations AFLOW-APL

AFLOW Phonon Library

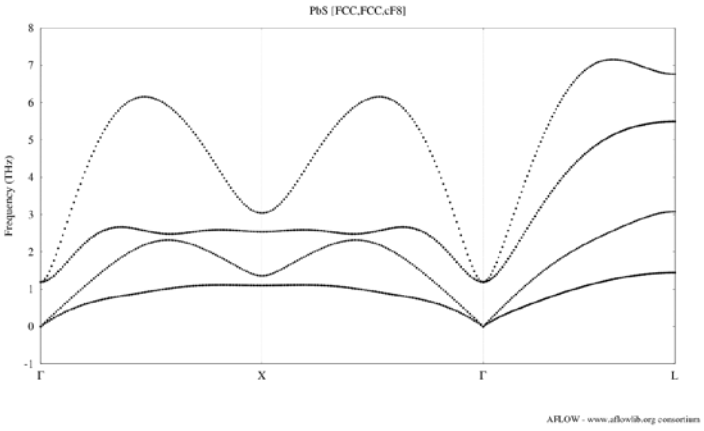


Perturb atom, obtain change in force on others

Interatomic force constants

$$\Phi_{ij}^{\alpha\beta} = \frac{\partial^2 V}{\partial r_i^\alpha \partial r_j^\beta}$$

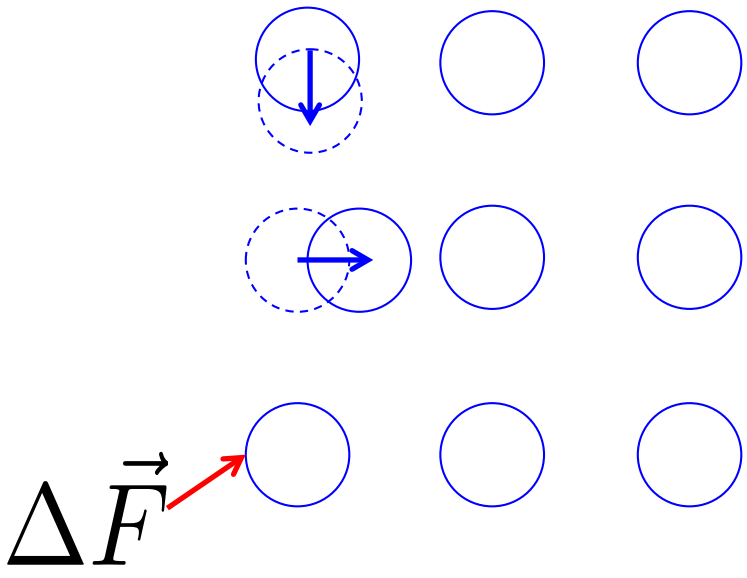
Phonon dispersion



Curtarolo *et al.*, Comput. Mater. Sci. **58**, 218 (2012); Curtarolo *et al.*, Comput. Mater. Sci. **58**, 227 (2012)

STATIC METADATA: Thermal Conductivity AFLOW-AAPL

AFLOW Anharmonic Phonon Library



Perturb atoms, obtain change in force on others

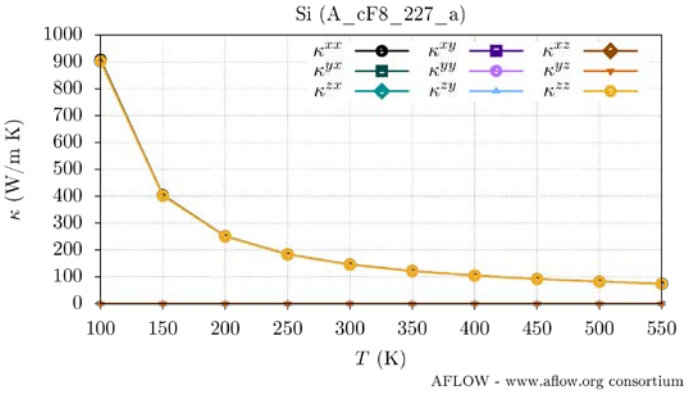
Interatomic force constants

$$\Phi_{ijk}^{\alpha\beta\gamma} = \frac{\partial^3 V}{\partial r_i^\alpha \partial r_j^\beta \partial r_k^\gamma}$$

Boltzmann Transport Equation

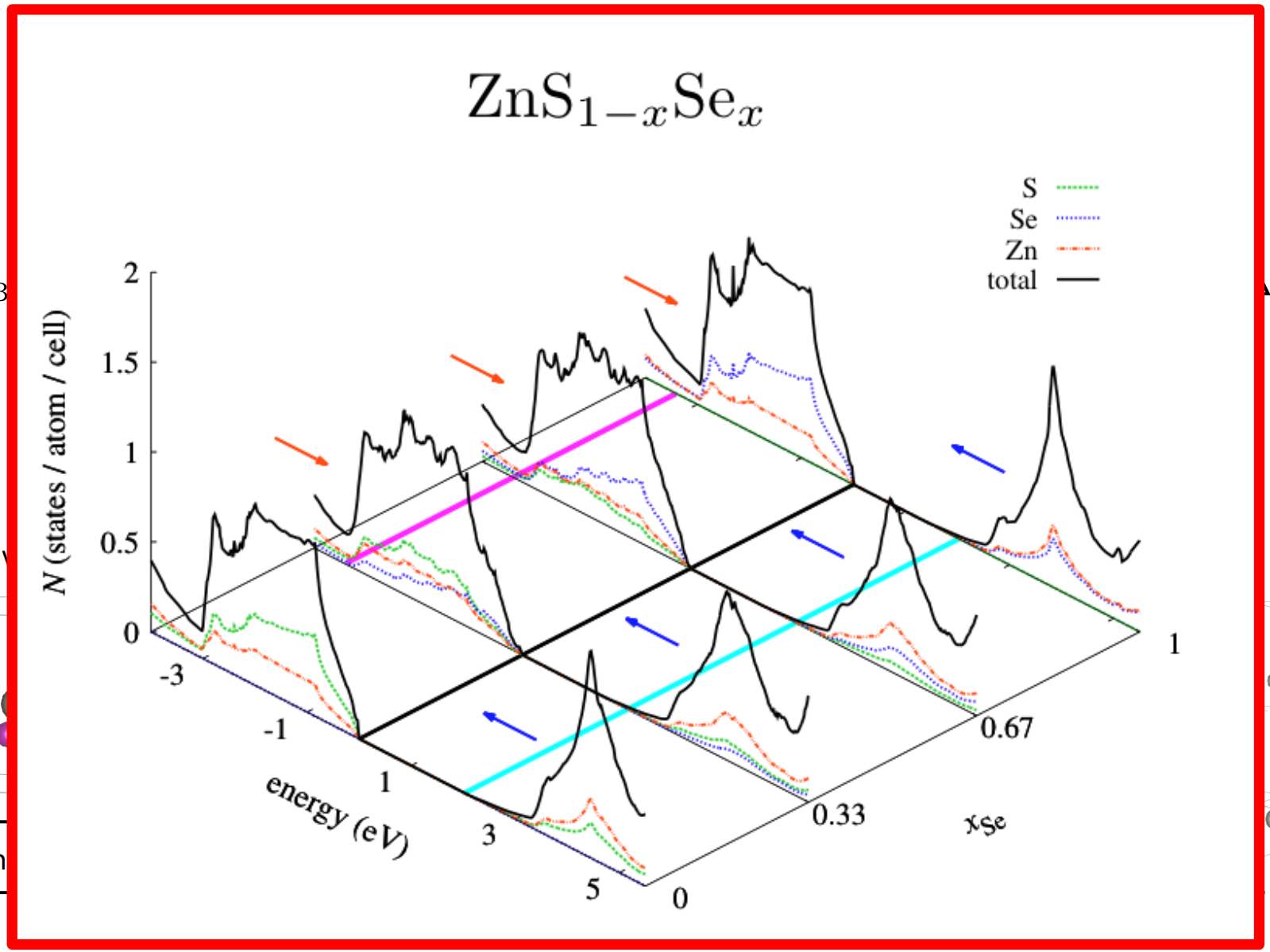
$$\kappa^{\alpha\beta} = \sum_\lambda C_\lambda v_\lambda^\alpha F_\lambda^\beta$$

Thermal conductivity



Plata et al., npj Comput. Mater. **3**, 45 (2017)

STATIC METADATA: disorder as sum of order - AFLOW-POCC



k_B ... $H_{f,i})$

conv

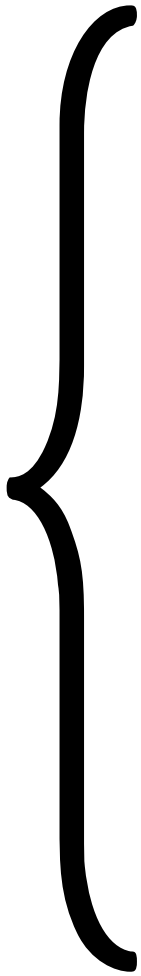
Zn

$H_{f,4}$

K. Yang, C. Oses, and S. Curtarolo, Chem. Mater. **28**(18), 6484 (2016)

STATIC METADATA and IDENTIFIERS

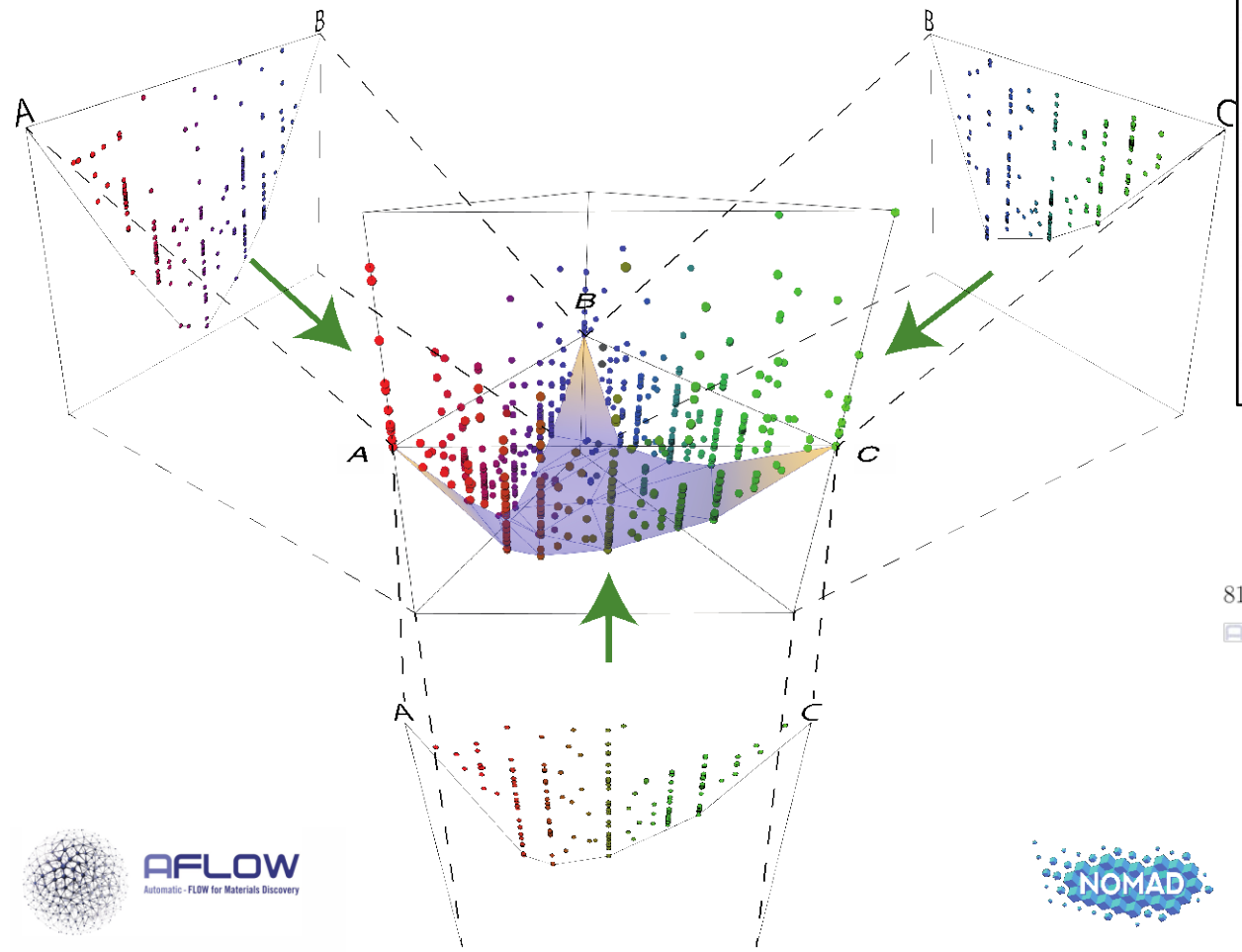
aflow:e0ac186639d0a150

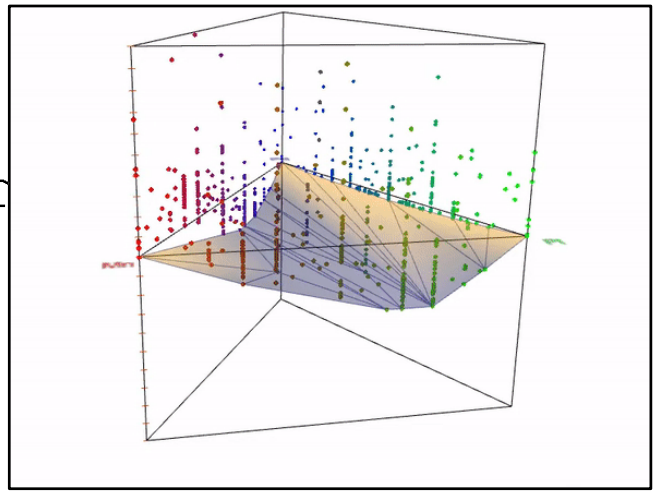


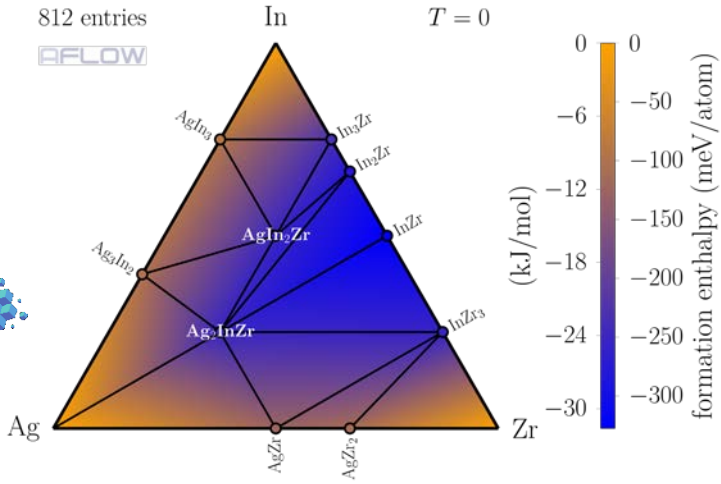
```
aflow_type=aggregate
method=aflow_apl
version=aflow_4
aggregate_parameters=…..
aggregate_content=[
aflow:e49dd5cbb21edec5
aflow:ac91a692fbd41751
aflow:828099c6035556a5
aflow:2776de8554edceff
aflow:db62c43f06264fb3
aflow:cec73e2f39cbd4da
aflow:464f0aa2d7ebf15a
aflow:0ce04a4d6541e0e2
aflow:e6c66d8f0e02e9e4
aflow:03a40b741db4d6c4
aflow:fa78e94bd638dbef
..]
```

PYTHON-IZABLE: http://aflowlib.org/material.php?id= aflow:e0ac186639d0a150

DYNAMIC METADATA: Convex Hull Phase AFLOW-CHULL







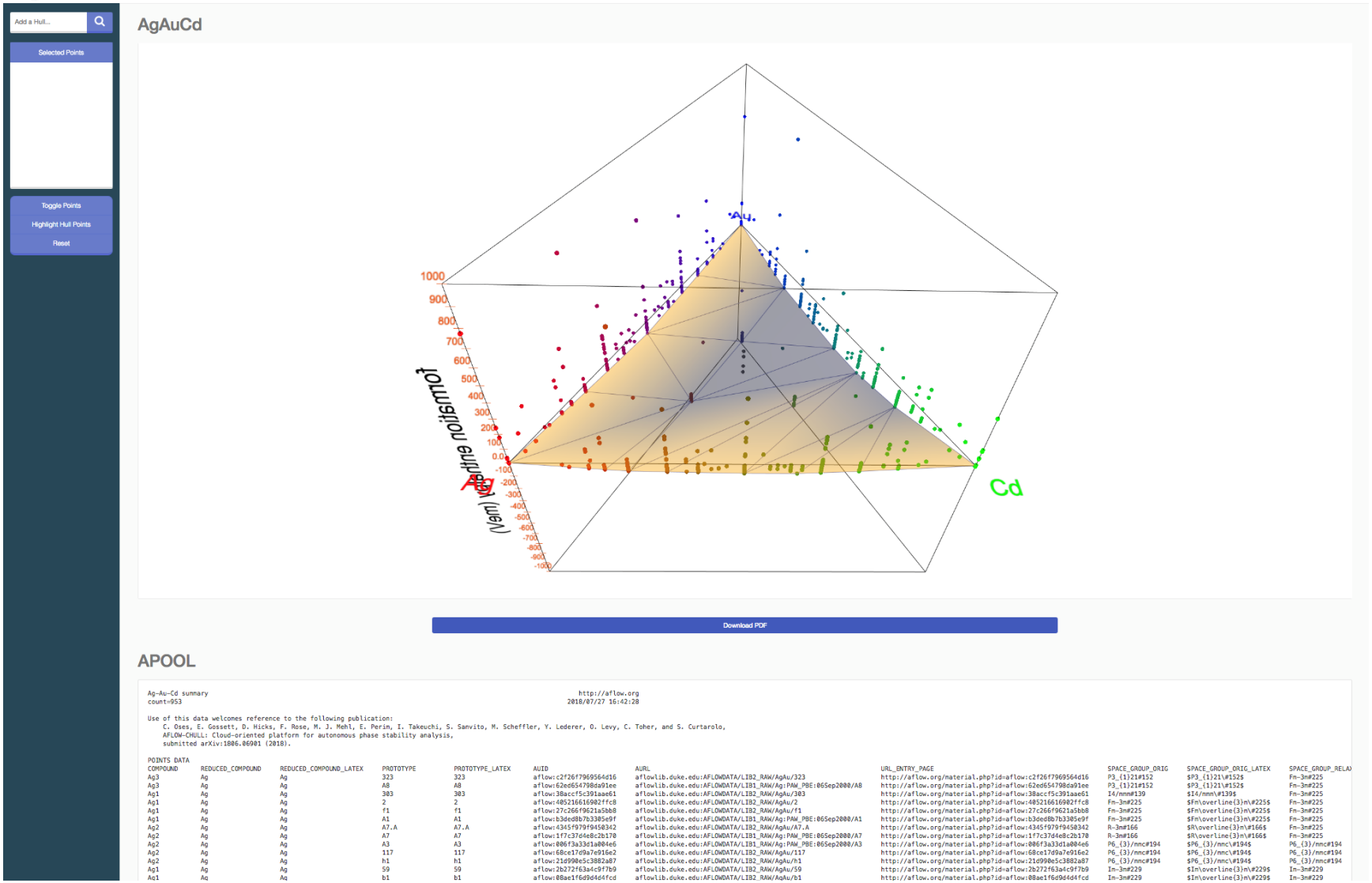
AFLOW
Automatic - FLOW for Materials Discovery

NOMAD

C. Oses *et al.,* J. Chem. Inf. Model. **58**(12), 2477-2490 (2018)

DYNAMIC METADATA: Convex Hull Phase AFLOW-CHULL

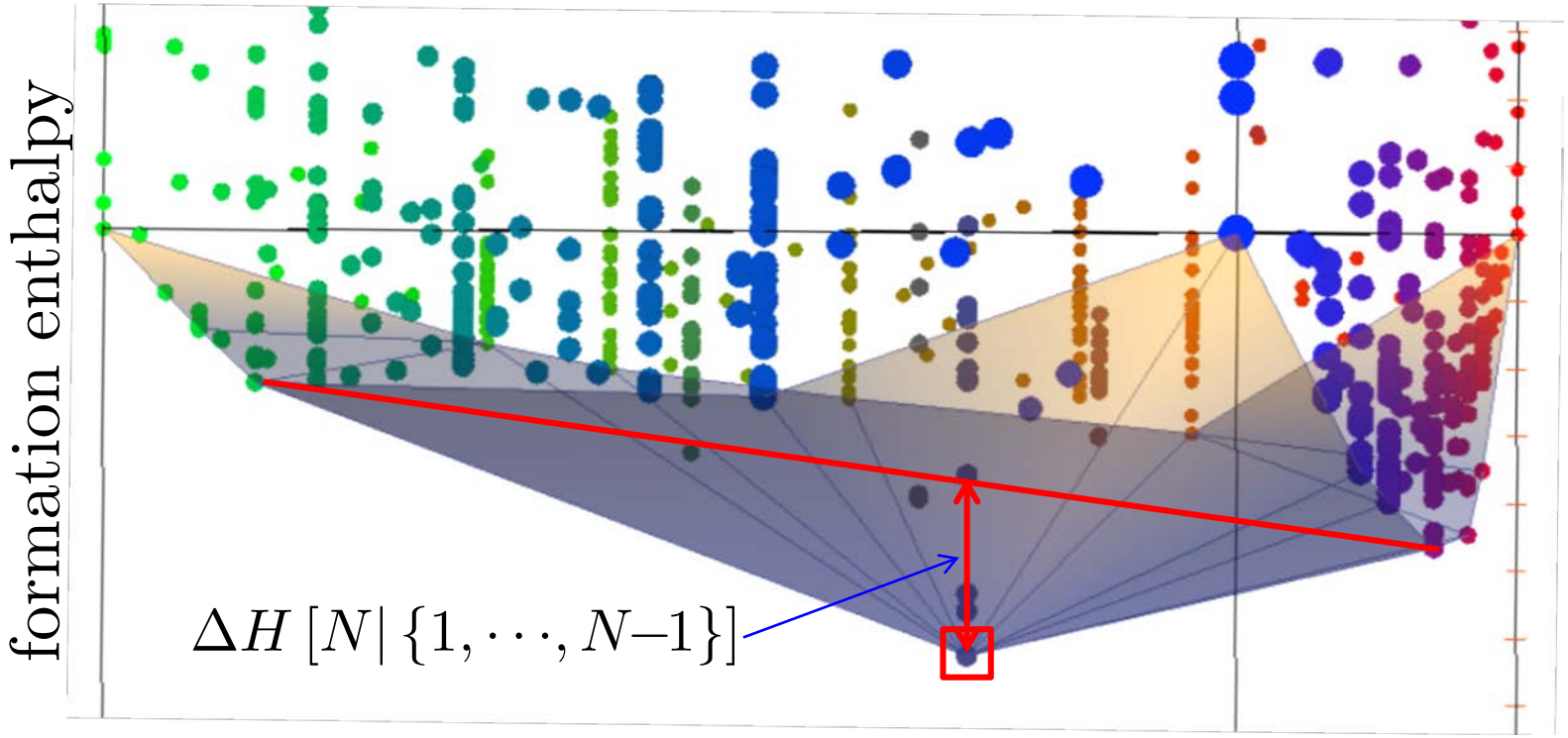
Go to aflow.org/aflow-chull and type/select AgAuCd



C. Oses *et al.*, J. Chem. Inf. Model. **58**(12), 2477-2490 (2018)

DYNAMIC METADATA: Δ Convex Hull - AFLOW-CHULL

"Enthalpy gain": energetic distance to hull formed from N-1 species components: $\Delta H\left[N|\left\{1,\cdots,N{-}1\right\}\right]$

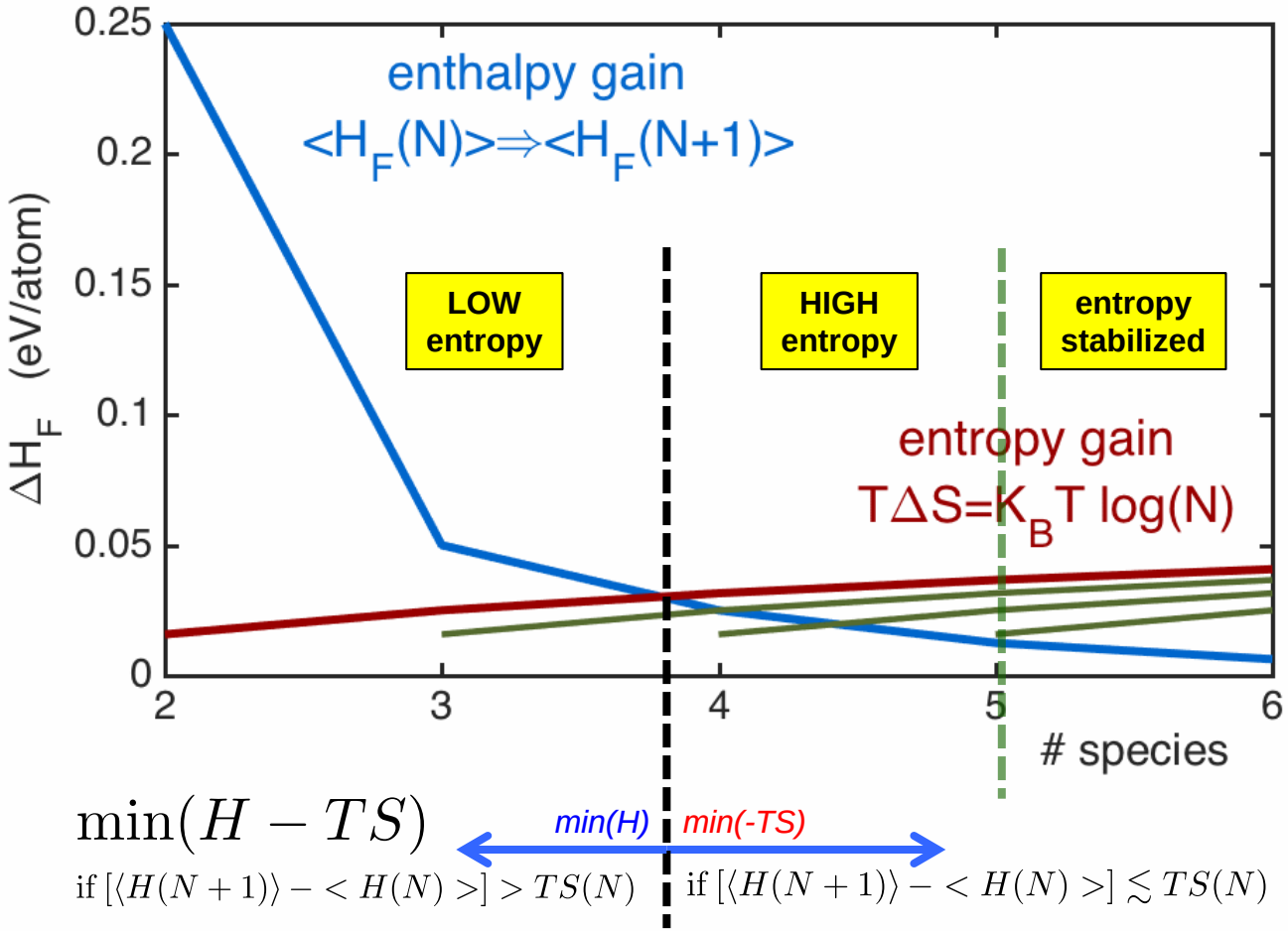


$$\Delta H\left[N|\left\{1,\cdots,N{-}1\right\}\right] \equiv H_{\text{hull}}\left[\left\{1,\cdots,N{-}1\right\}\right] - H\left[N\right]$$

if $H\left[N\right] < H_{\text{hull}}\left[\left\{1,\cdots,N{-}1\right\}\right]$; otherwise is zero.

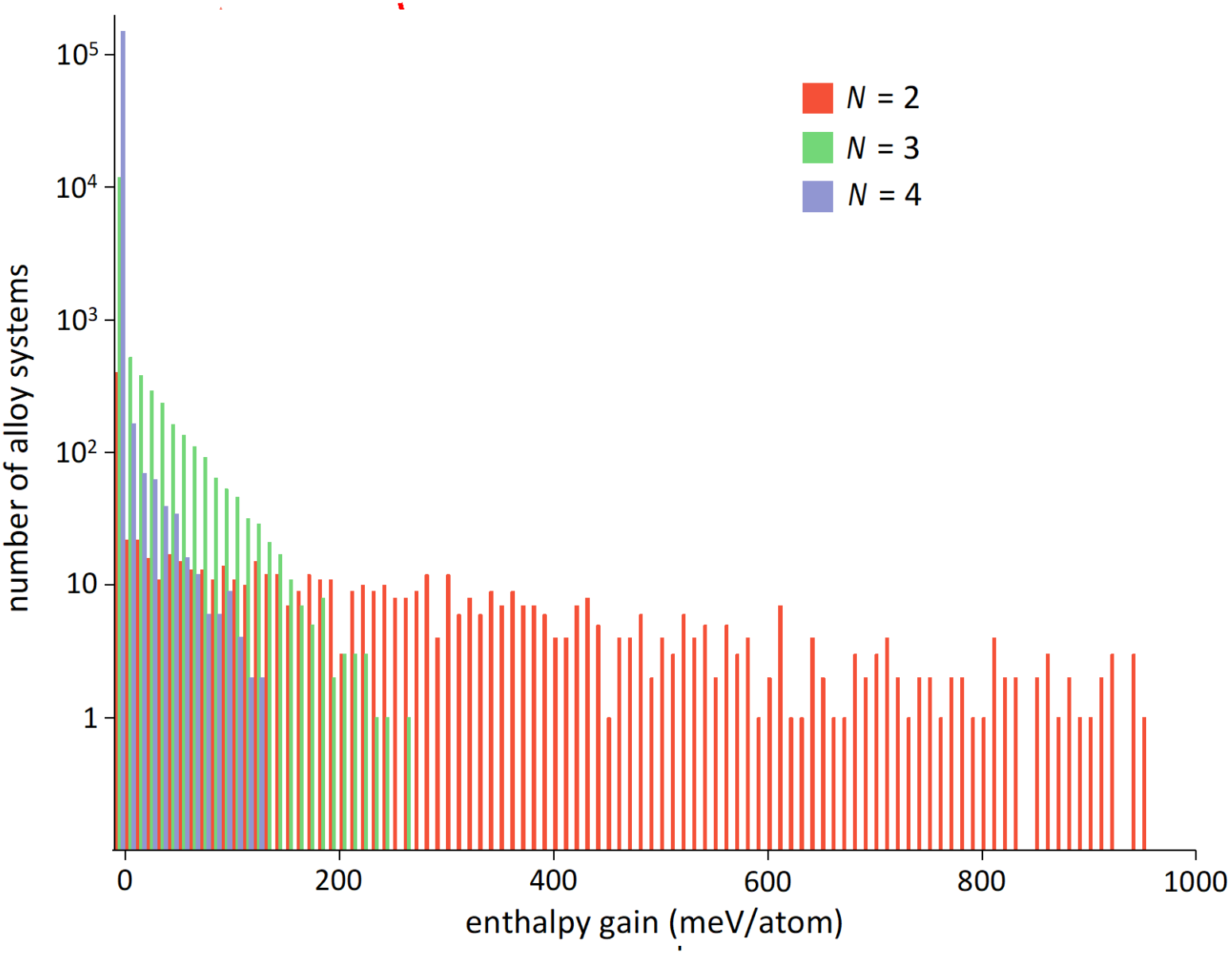
C. Toher, C. Oses, D. Hicks, S. Curtarolo, NPJ Comput. Mater. **in press** (2019)

Theory from METADATA: N+1 theorem



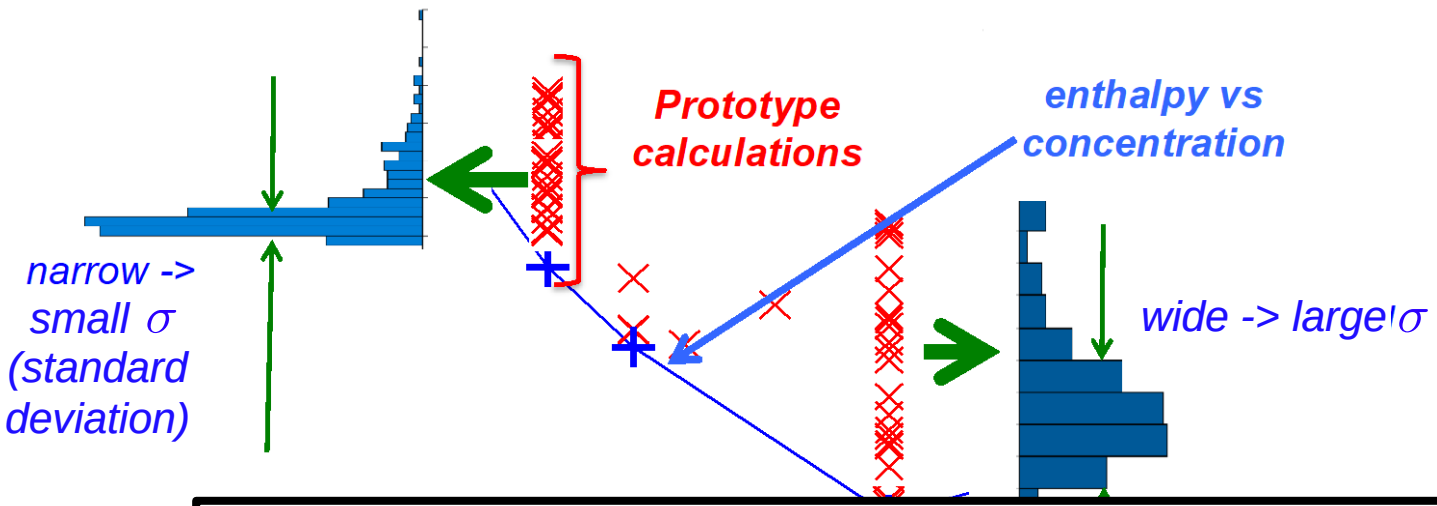
C. Toher, C. Oses, D. Hicks, S. Curtarolo, NPJ Comput. Mater. **in press** (2019), M. Widom, *J. Mater. Res.* **33**, 2881 (2018)

Materials from METADATA: The ground-state-less truth



C. Toher, C. Oses, D. Hicks, S. Curtarolo, NPJ Comput. Mater. **in press** (2019)

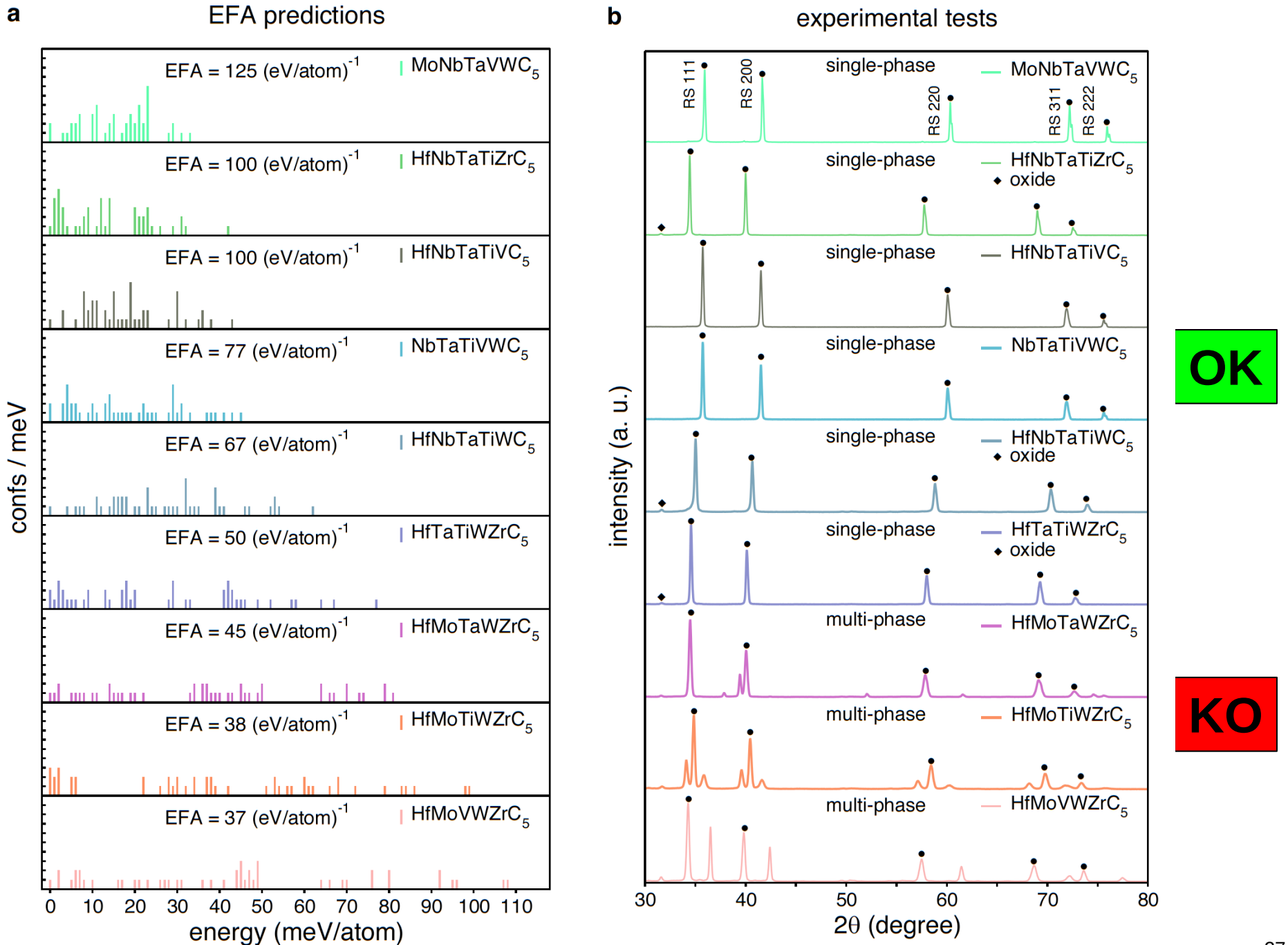
DYNAMIC METADATA: crude ML measuring entropy



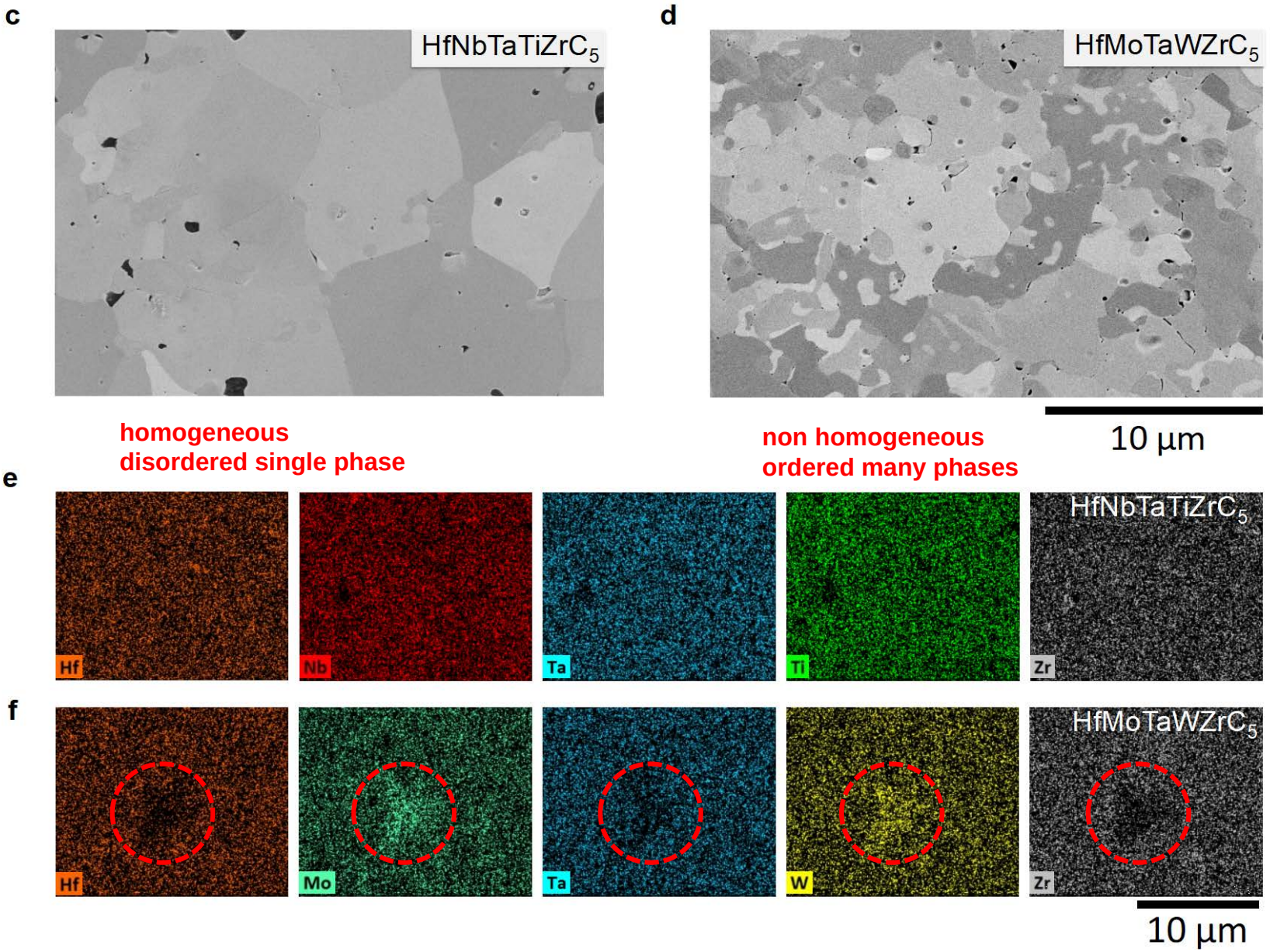
AFLOW-POCC: combinations
AFLOW-SYM: degenerations

$$\mathrm{EFA}(N) \quad \sigma\{H_i(N)\} = \sqrt{\frac{\sum_{i=1}^{n} g_i \left(H_i - H_{\mathrm{mix}}\right)^2}{\left(\sum_{i=1}^{n} g_i\right) - 1}}, \quad -1$$

P. Sarker, T. Harrington *et al.,* Nature Comms. **9,** 4980 (2018)
E. Perim *et al.,* Nature Comms. **7**, 12315 (2016).
D. Hicks *et al.,* Acta Cryst. A**74**, 184-203 (2018)


a
EFA predictions
EFA = 125 (eV/atom)^-1
MoNbTaVWC5
EFA = 100 (eV/atom)^-1
HfNbTaTiZrC5
EFA = 100 (eV/atom)^-1
HfNbTaTiVC5
EFA = 77 (eV/atom)^-1
NbTaTiVWC5
EFA = 67 (eV/atom)^-1
HfNbTaTiWC5
EFA = 50 (eV/atom)^-1
HfTaTiWZrC5
EFA = 45 (eV/atom)^-1
HfMoTaWZrC5
EFA = 38 (eV/atom)^-1
HfMoTiWZrC5
EFA = 37 (eV/atom)^-1
HfMoVWZrC5
confs / meV
0 10 20 30 40 50 60 70 80 90 100 110
energy (meV/atom)
b
experimental tests
RS 111
RS 200
RS 220
RS 311
RS 222
single-phase
MoNbTaVWC5
single-phase
HfNbTaTiZrC5
oxide
single-phase
HfNbTaTiVC5
single-phase
NbTaTiVWC5
single-phase
HfNbTaTiWC5
oxide
single-phase
HfTaTiWZrC5
oxide
multi-phase
HfMoTaWZrC5
multi-phase
HfMoTiWZrC5
multi-phase
HfMoVWZrC5
intensity (a. u.)
30 40 50 60 70 80
2θ (degree)
OK
KO

5-metal carbides: Hf Mo Nb Ta Ti V W Zr + C



P. Sarker, T. Harrington *et al.,* Nature Comms. **9,** 4980 (2018)

CONCLUSIONS

AFLOW
Automatic-FLOW for Materials Discovery
a distributed materials genome properties repository from high-throughput

AFLOW tries to be as FAIR as possible.

- Findable: AUIDs, rich, searcheable
- Accessible: retrievable, AFLUX-able
- Interoperable: accessible, shared and simple vocabulary
- Reusable: through python, APE, and web-apps