Infrastructure to enable user contributions to Materials Project

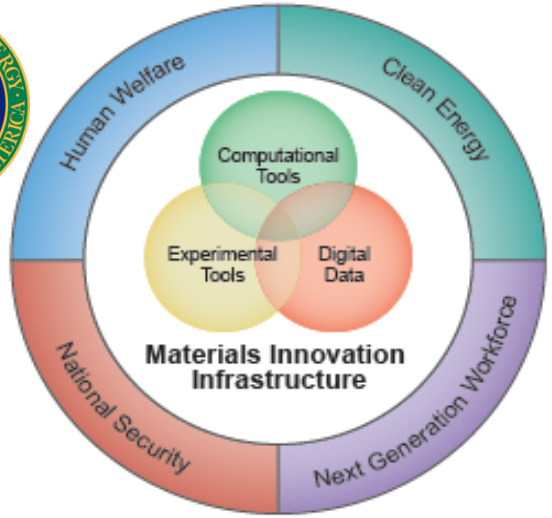
MPComplete & MPContribs

MGI & Materials Project: Achievements to Date

Materials Genome Initiative

2011: "*make the process of* ***discovery & development of advanced materials faster, less expensive, more predictable***"

"solutions in most pressing areas require advanced materials"



High-Quality Materials DATA

- > **60,000 relaxed compounds:** validated energy, phase diagrams
- > **70,000 Pourbaix diagrams:** world's largest set
- > **28,000 band structures** + higher accuracy 2,700 band gaps
- > **1,300 elastic tensors:** world's largest data set

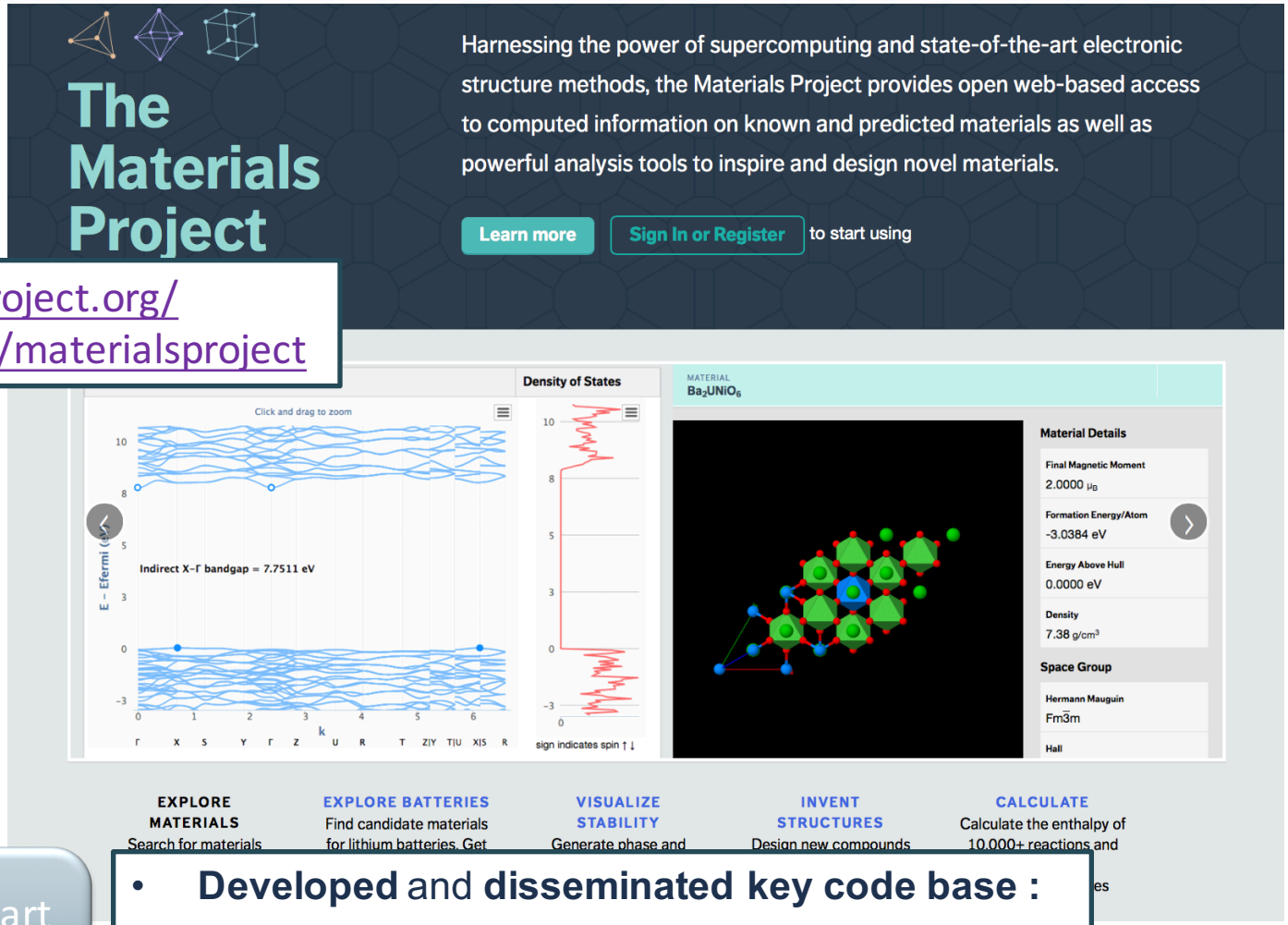
Rapid DISSEMINATION

- **Ten Apps** enabling material searching and design
- **First Materials data API** ; community download > 8 million data
- **MPContribs framework:** platform for data sharing
- **Over 14,000 registered users !**

DESIGN

Design of **novel functional materials** (photocatalysts, thermoelectrics, cathodes/electrolytes)

MP Web Site – A Science Gateway

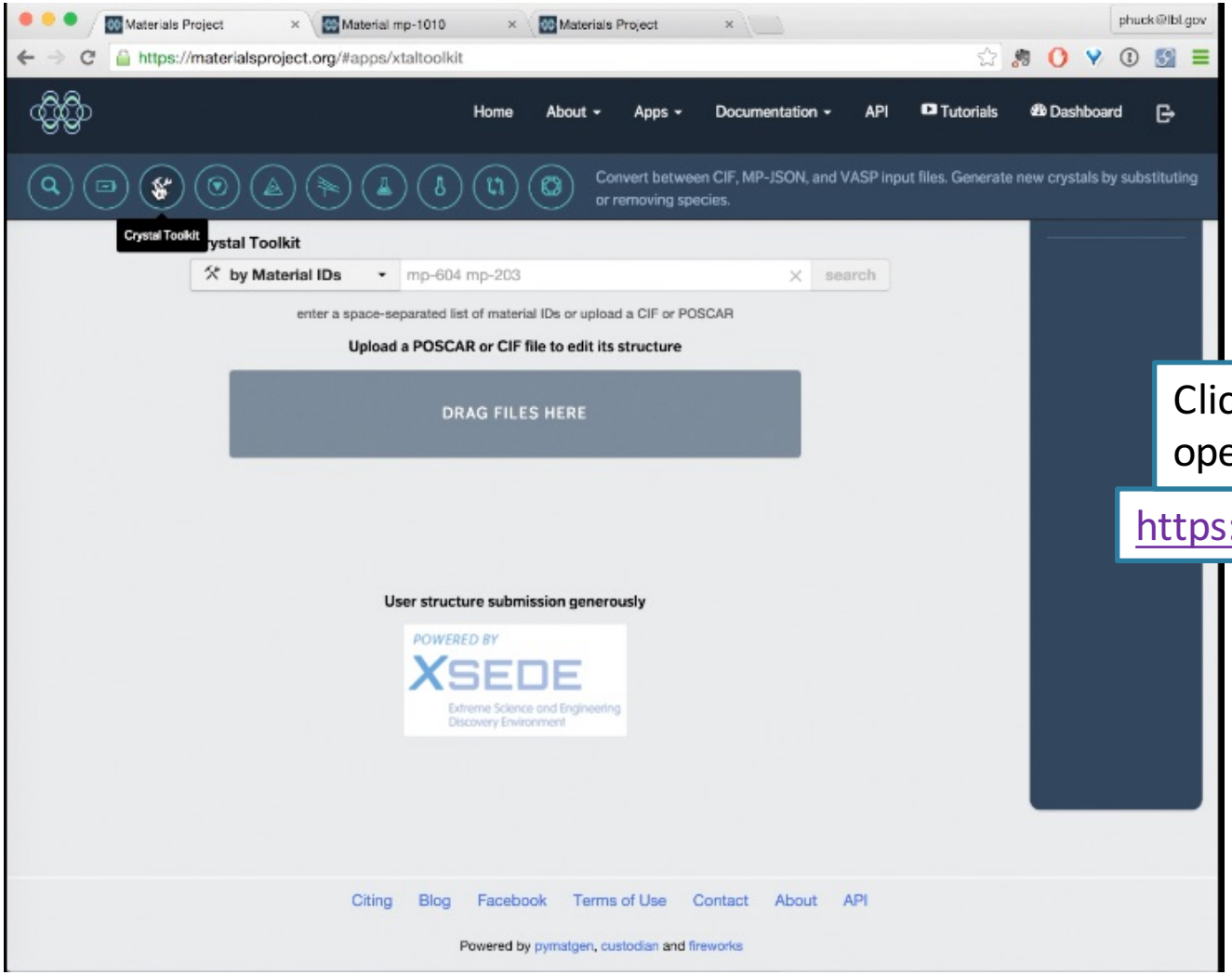


https://materialsproject.org/
https://github.com/materialsproject

State-of-the-art OPEN SOURCE CODES

- **Developed** and **disseminated key code base :**
- **FireWorks** workflow
- **pymatgen**; comprehensive analysis code
- **Custodian** failure recovery

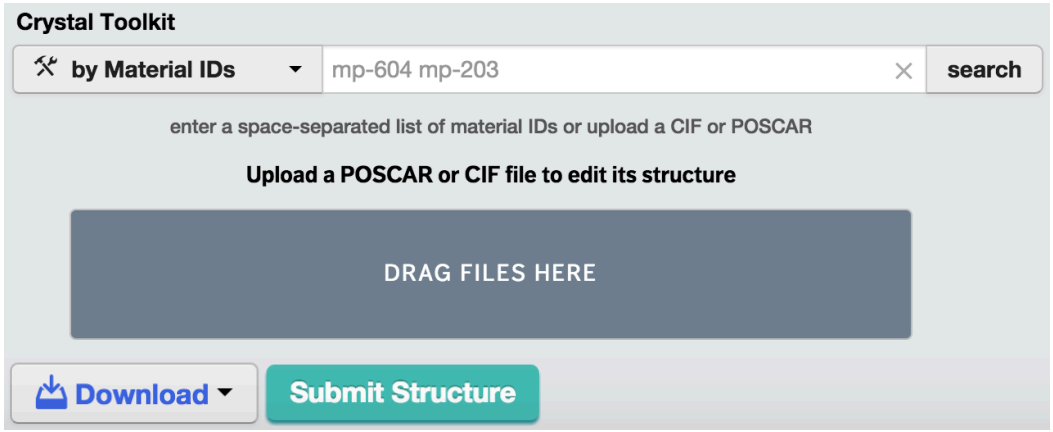
Crystal ToolKit App & MPComplete Demo

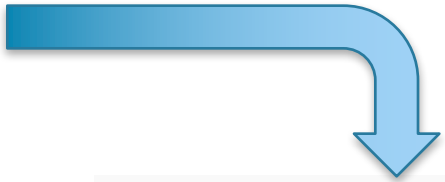


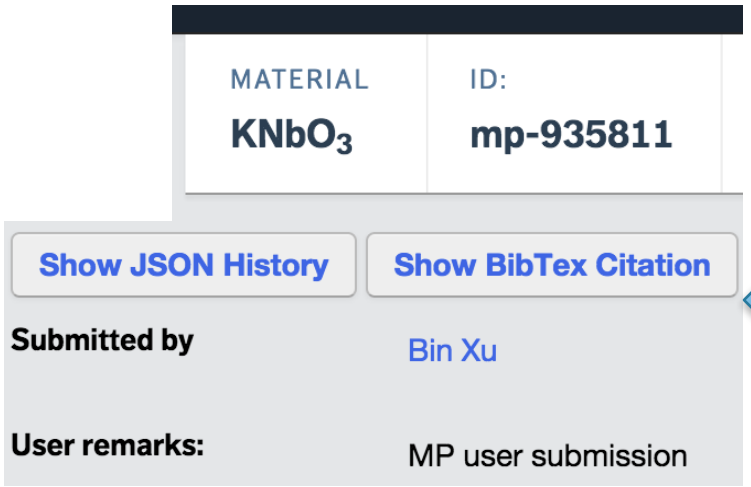
Click image or URL to open demo on YouTube

https://youtu.be/k6mEA-sDSig

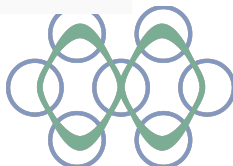
Production Workflow / FireWorks Dash





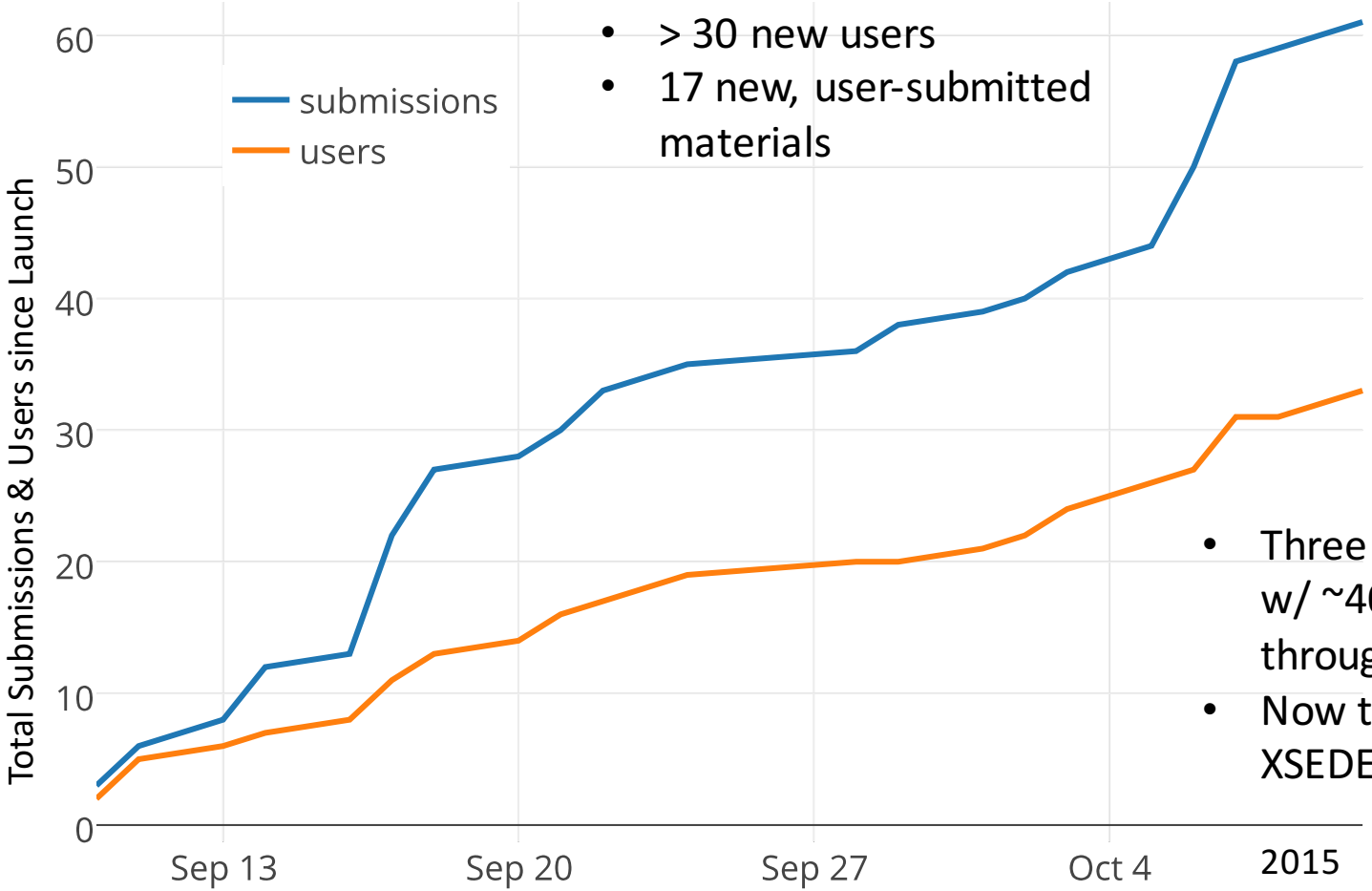


http://fireworks.dash.materialsproject.org/wf/1445229



MP-Complete Trial Run @ XSEDE

- One-month trial run on XSEDE resources
- > 60 submissions
- > 30 new users
- 17 new, user-submitted materials



- Three bulk submissions w/ ~400 structures run through NERSC
- Now transitioning back to XSEDE after extension

Roles of MPComplete & MPContribs

Suggest new structures for calculation by MP

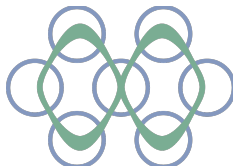
MPCOMPLETE

- Assigns [new] MP identifier
- Runs calculations
- Adds material to core database

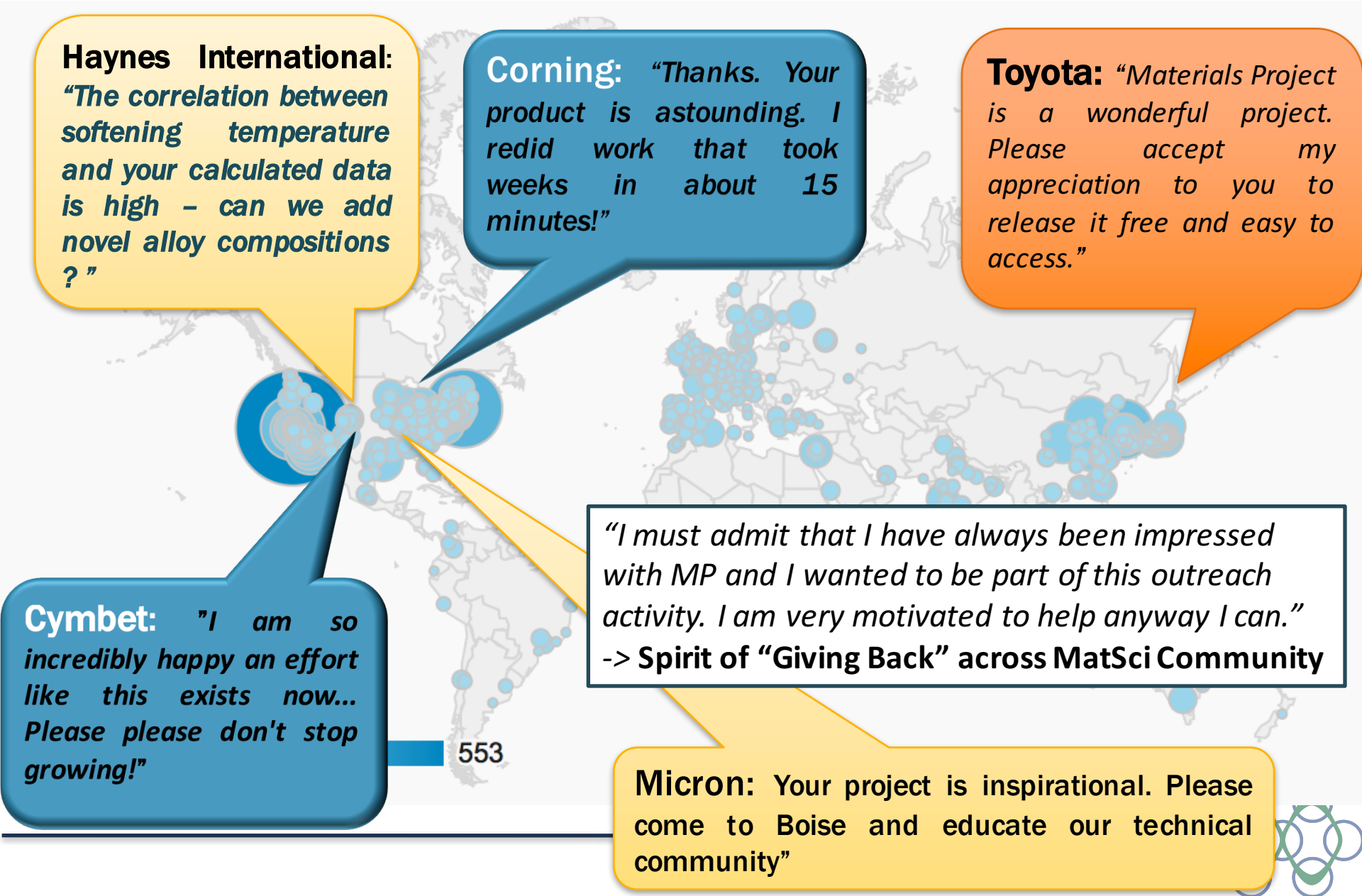
Contribute data and disseminate it through MP

MPCONTRIBS

- Annotate materials already existing in core database
- Disseminate complementary theoretical and experimental data to MP users



MP is a World-Wide Resource



MPContribs: Collaborative Platform for sharing Data

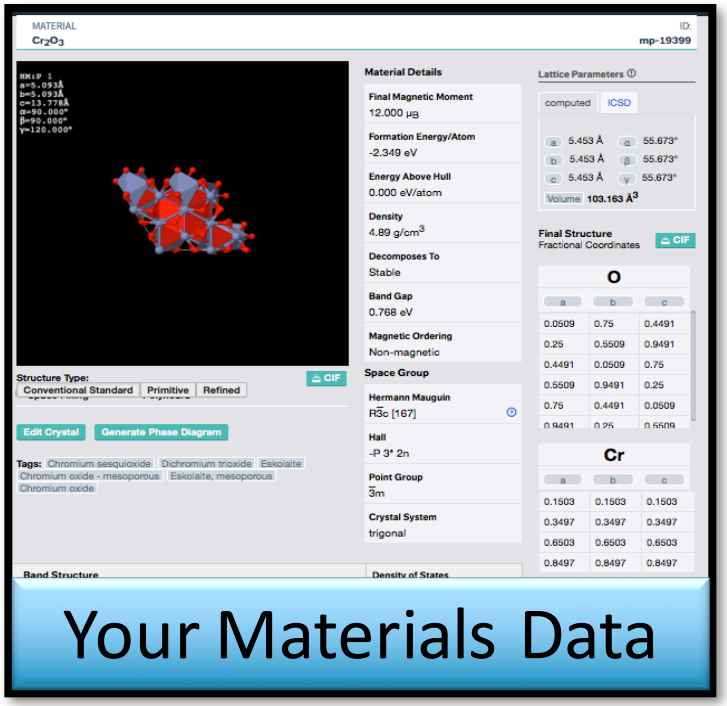
"help sharing datasets with the world"

A. T. N'Diaye (ALS, LBNL):

- measured XAS/XMCD spectra
- properties of rare earth substitutes
- processing of instrumental data
- integration w/ MP phase diagrams

D. Morgan, H. Wu (SI2, UW):

- computed diffusion coefficients
- automated VASP data extraction and integration with MP



Your Materials Data

also aim to develop a customized web app driven by MPCONTRIBS

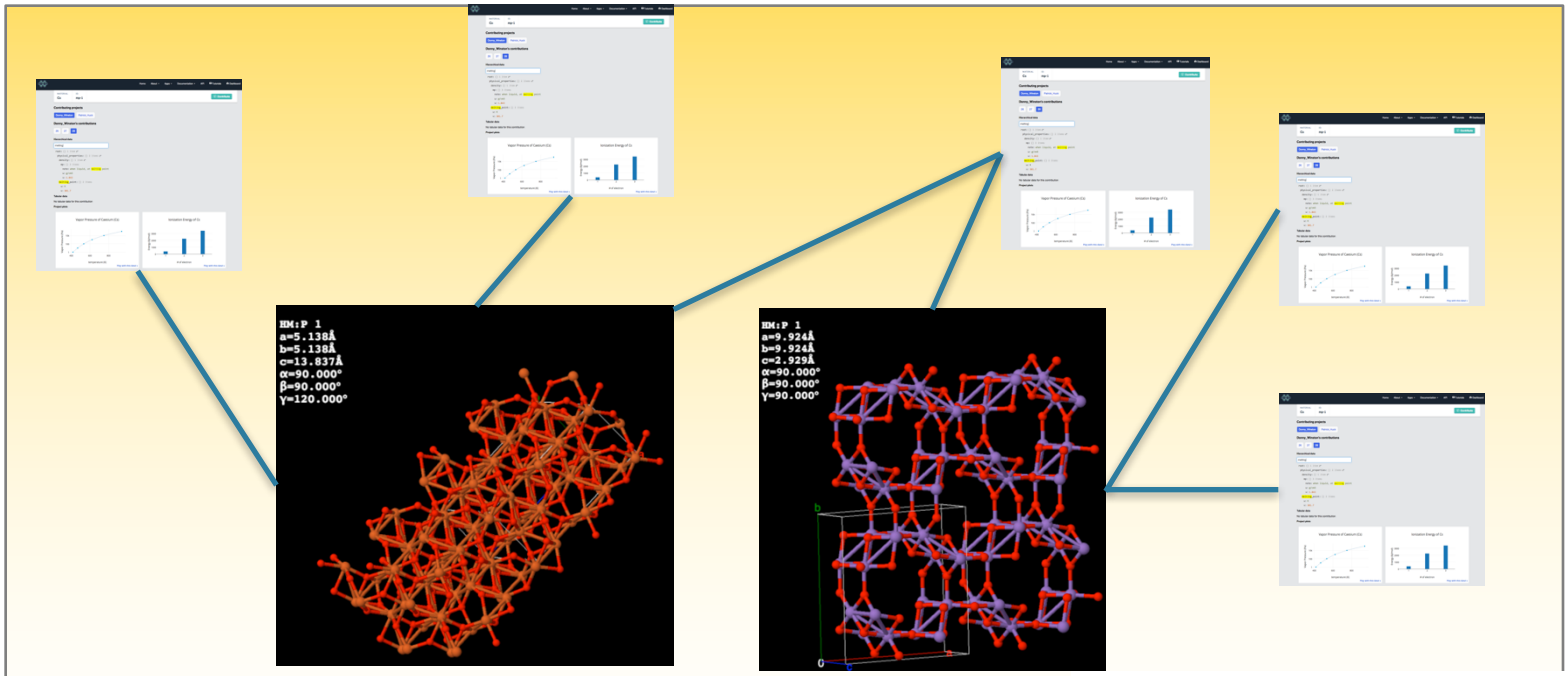
- **Establish MatSci Hub for theoretical and experimental Data**
- **Effective dissemination for exposure to large MP user community**

-> new platform to efficiently integrate data from disparate sources

Huck P. *et al.*; eScience15; GCE15; Poster

full contributors list

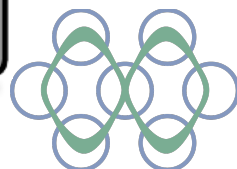
Collective Annotation of Core Database



hierarchical material categorization to cover disparate scopes of experimental and theoretical contributions

raw user files
experimental instruments
theoretical calculations

Categories [Identifiers]
Structure [mp-24972]
< Composition [Fe2O3]
< Chem. System [Fe*O*]



Correlating Contributions w/ Core Data

Hierarchical Material Categorization:

raw user files
experimental instruments
theoretical calculations

Categories [Identifiers]
Structure [mp-24972]
< Composition [Fe2O3]
< Chem. System [Fe*O*]

Symmetry + Lattice

materials/mp-24972/contributions
(mostly theoretical contributions)

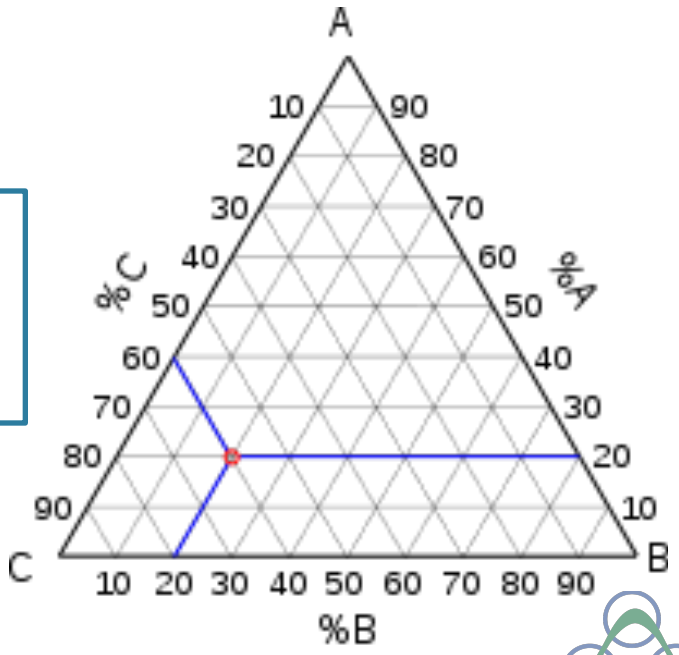
links to

Specific compositions

compositions/Fe2O3/contributions
(experimental contributions)

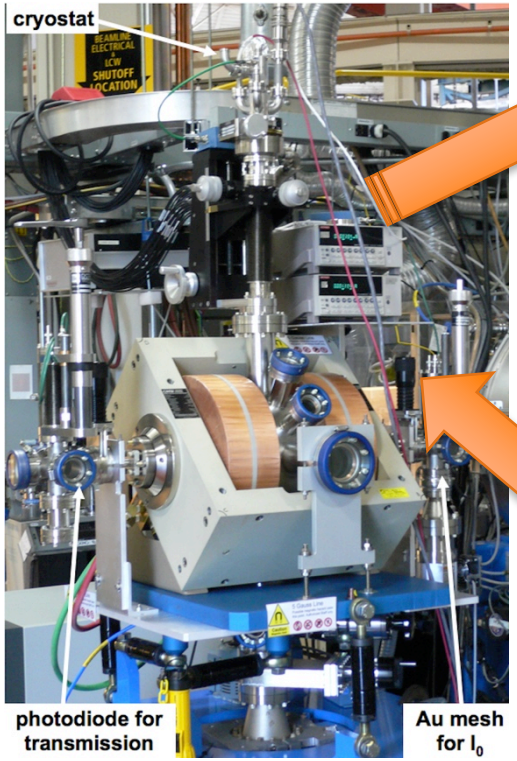
multiple compositions

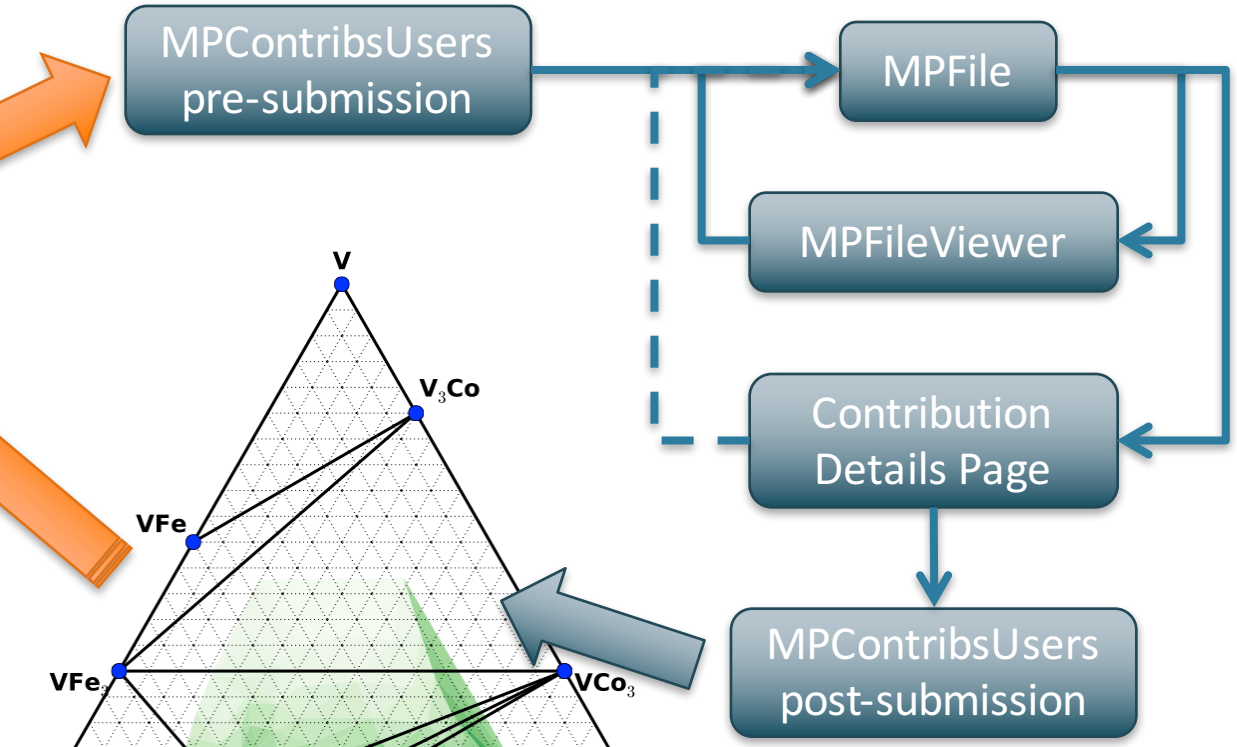
ternary plot representation, e.g. phase diagrams



SHOWCASE: **Unified Experimental and Theoretical XAS Application** [see paper at 10th GCE workshop]

MPContribs User Workflow





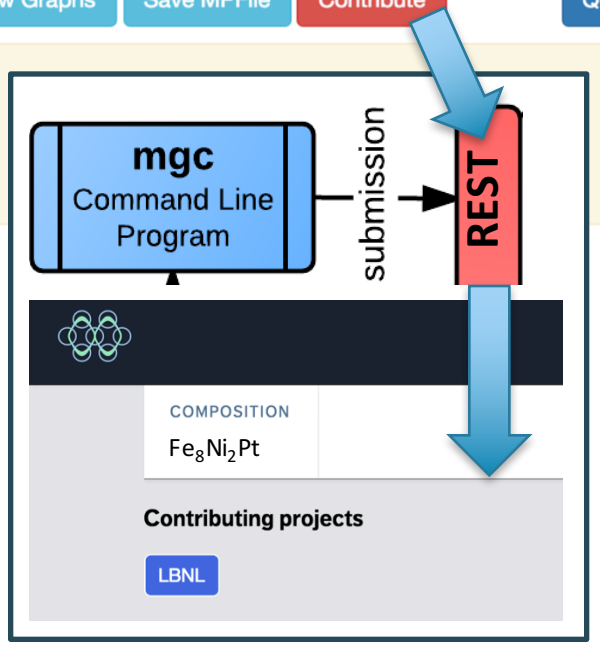
- pre-processing of user output data and conversion into MPFile
- visual and iterative checking of MPFile by user ("get data in shape")
- MPFile submission via command line or web portal (through REST)
- contributed data appears instantly on the MP portal

Offline Web UI – “Get Data in Shape”

Home | ArchieML | Choose MPFile ... | Load MPFile | View MPFile | View Graphs | Save MPFile | Contribute | Quit

locally process contribution #0 ... OK.
check consistency of contribution #0 ... OK.
build contribution #a898b40 into Fe8Ni2Pt ... OK.
1 contributions successfully processed.

Go to contribution ...



Fe8Ni2Pt » Patrick Huck » #a898b40

Hierarchical Data

- (root): {} 3 Items
  - Experiment: {} 4 Items
    - Preparation: {} 1 Item
    - Sample: {} 5 Items
      - Material_Name: Platinum doped Permalloy
      - Form: ~20nm film on Si wafer
      - Thickness: ca. 20nm with 2-3 nm Au-capping (nominally)
      - Grower: Ales Hrabec
      - Authors: Ales Hrabec, Alpha T. N'Diaye, Elke Arenholz, Christopher Marrows
    - Measurement: {} 5 Items
    - Beamline: {} 7 Items
  - Ni_XMCD: {} 2 Items
    - get_xmcd: {} 1 Item
    - xas_normalization_to_min_and_max: {} 3 Items
      - energy_range: 800 1000
      - normalization_factor: 0.952002315041
      - offset: 0.358620768783
  - Fe_XMCD: {} 3 Items

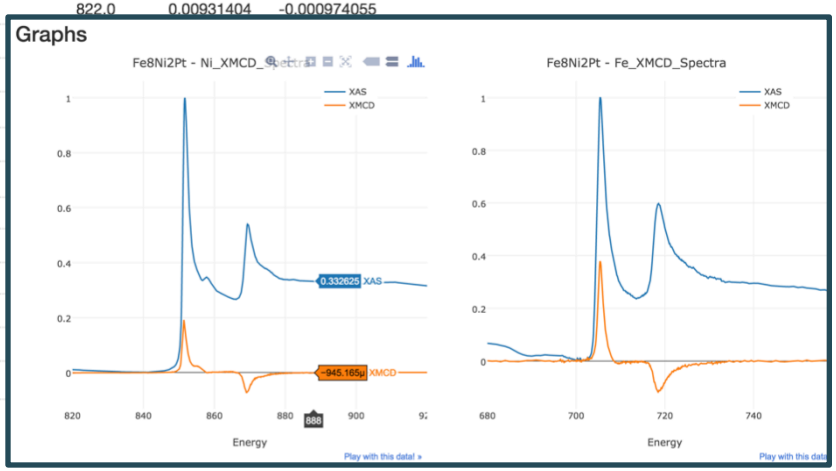
Tables

data_Ni_XMCD_Spectra Collapse/Expand

Search:

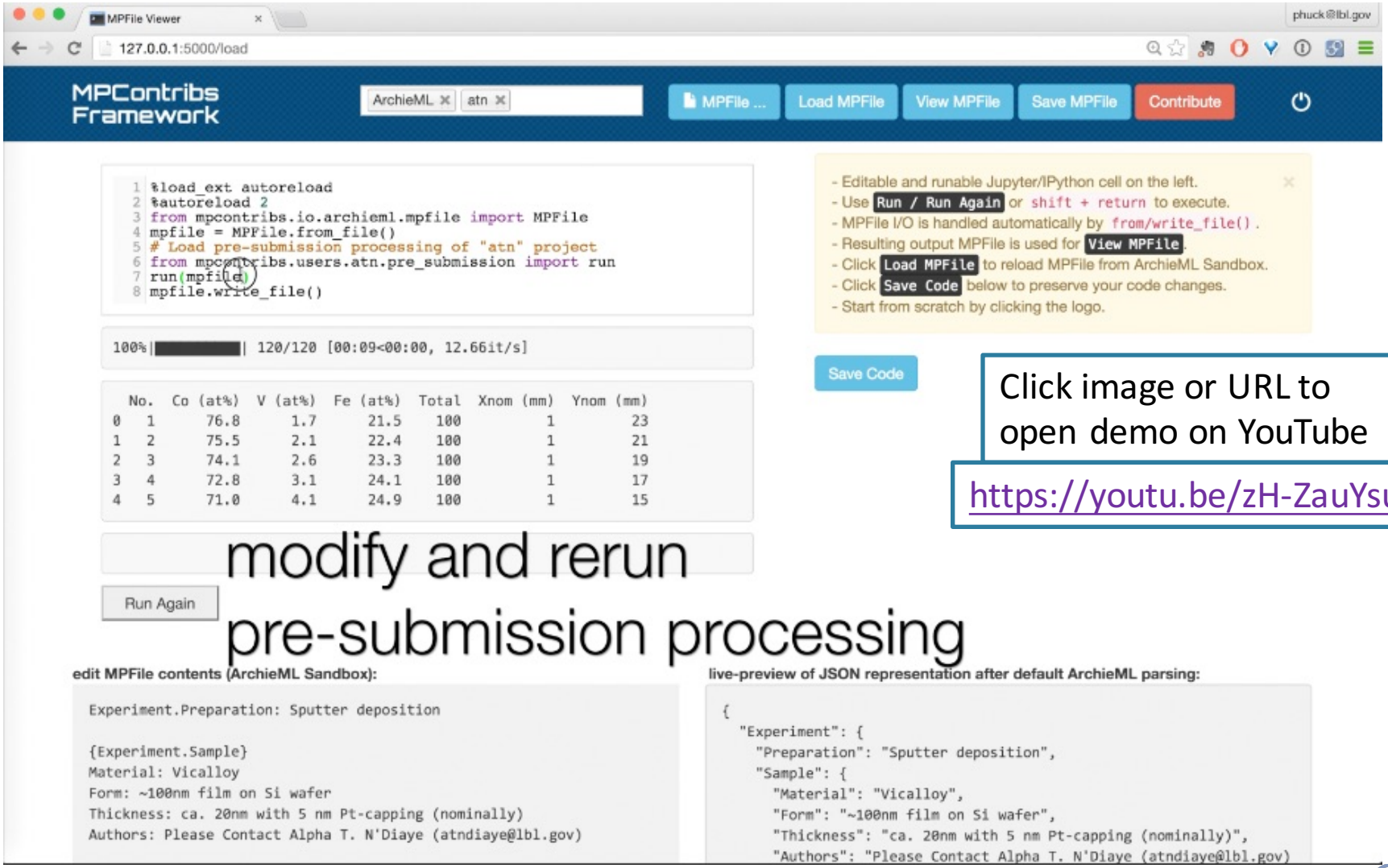
Energy	XAS	XMCD
821.0	0.0104183	-0.000451802
822.0	0.00931404	-0.000974055

Graphs



Demo: https://youtu.be/xlwttmXSpHg (Full Playlist)

Offline Web UI – Release-Ready Demo



Click image or URL to open demo on YouTube

https://youtu.be/zH-ZauYsu64

ALS MPFile: Labbook & Processing Control

```
Experiment.Preparation.Description: Sputter deposition

{Experiment.Sample}
Material_Name: Platinum doped Permalloy
Form: ~20nm film on Si wafer
Thickness: ca. 20nm with 2-3 nm Au-capping (nominally)
Grower: Ales Hrabec
Authors: A. Hrabec, A. T. N'Diaye, E. Arenholz, C. Marrows

{Experiment.Measurement}
Detection: total electron yield
Temperature: RT
Orientation: 30° grazing incidence
Date: 2015-06-24
Measured_by: Alpha T. N'Diaye

{Experiment.Beamline}
Beamline: ALS-6.3.1
Method: Soft x-ray XAS and XMCD
Polarization: circular, positive (ca. 60%)
Magnet_Field: 0.8T switching point by point, parallel to beam
Count_Time: 1s
Delay_Time: 0.5s
Monochromator.Exit_Slit: 20µm
Monochromator.Grating: 600l/mm
```

Shared Meta-Data (from Labbook)

MPFile = Config + Input + Output

```
{Ni20Fe80Pt10}
Ni_XMCD.get_xmcd.energy_range: 800 1000
Fe_XMCD.get_xmcd.energy_range: 600 800

{Ni20Fe80Pt10.Ni_XMCD.xas_normalization_to_min_and_max}
energy_range: 800 1000
normalization_factor: 0.952002315041
offset: 0.358620768783

{Ni20Fe80Pt10.Fe_XMCD.scaling_preedge_to_1}
preedge_range: 690 700
xas_minus_factor: 0.348231766387
xas_plus_factor: 0.349333591384

{Ni20Fe80Pt10.Fe_XMCD.xas_normalization_to_min_and_max}
energy_range: 600 800
normalization_factor: 1.00964185927
offset: 0.984095999176
```

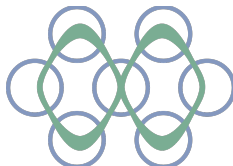
Raw Data Processing Instructions

Results of Processing Steps

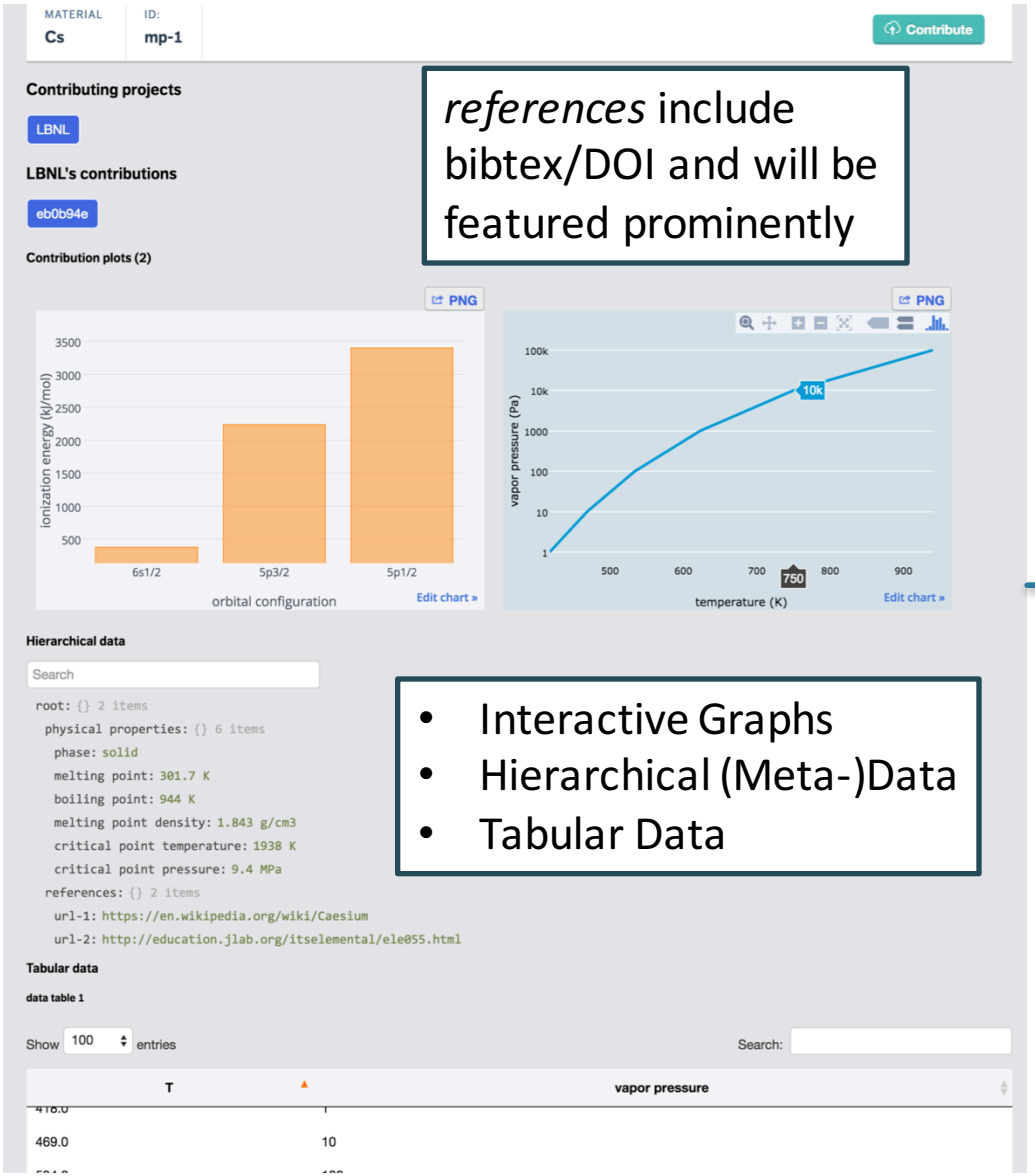
```
[+Ni20Fe80Pt10.Ni_XMCD_Spectra]
Energy,XAS,XMCD
820,0.0104944,-0.00140602
821,0.0104183,-0.000451802
822,0.00931404,-0.000974055
823,0.00821621,-0.00083305

[+Ni20Fe80Pt10.Fe_XMCD_Spectra]
Energy,XAS,XMCD
```

Final XAS Spectra (CSV)

- Default interactive graphs from each data table
- Check results & reiterate via **Offline Web UI**
- *Future:* guide user in what to include as Meta-Data



MPFile & Generic Contribution Details Page



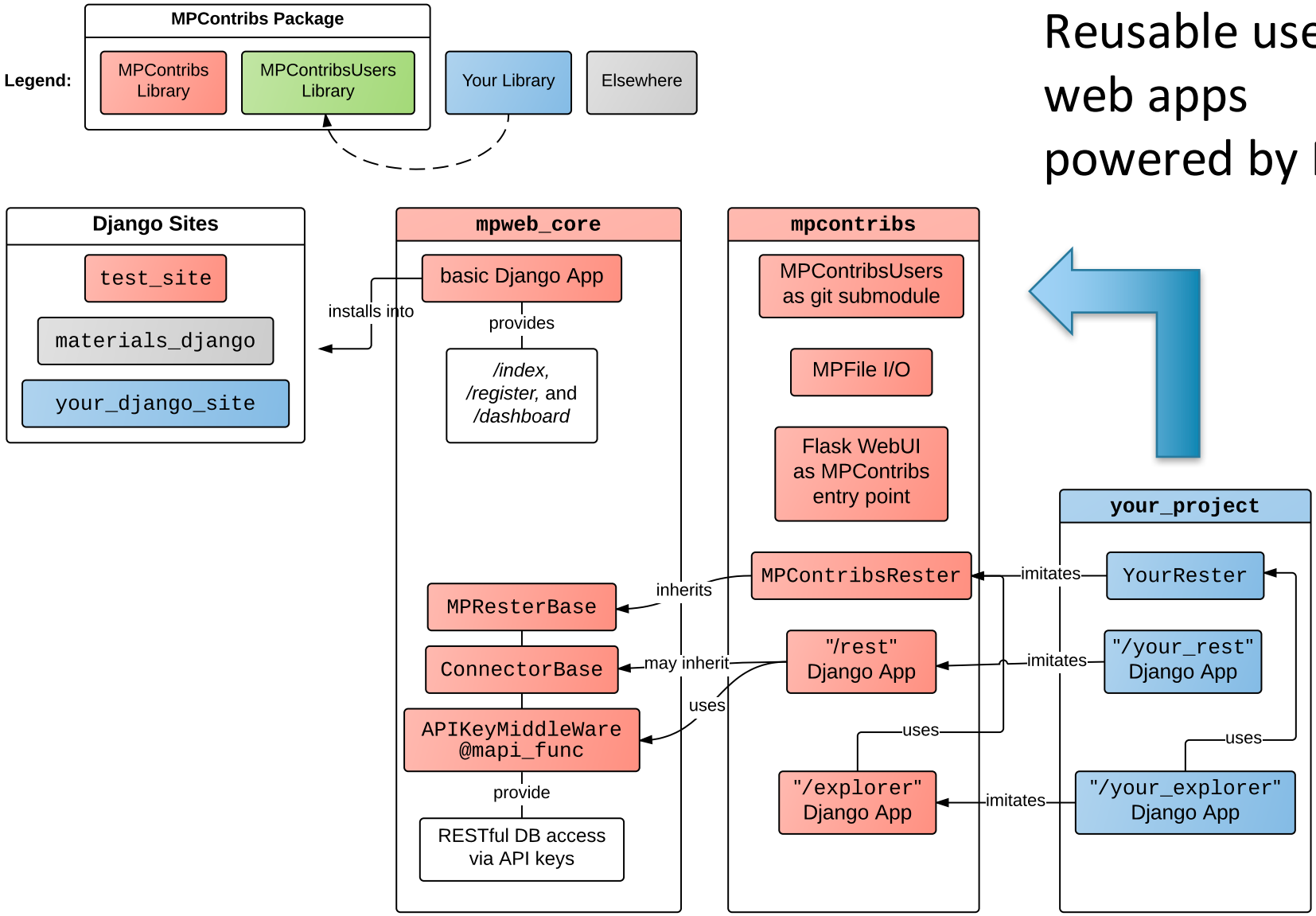
references include bibtex/DOI and will be featured prominently

- Interactive Graphs
- Hierarchical (Meta-)Data
- Tabular Data

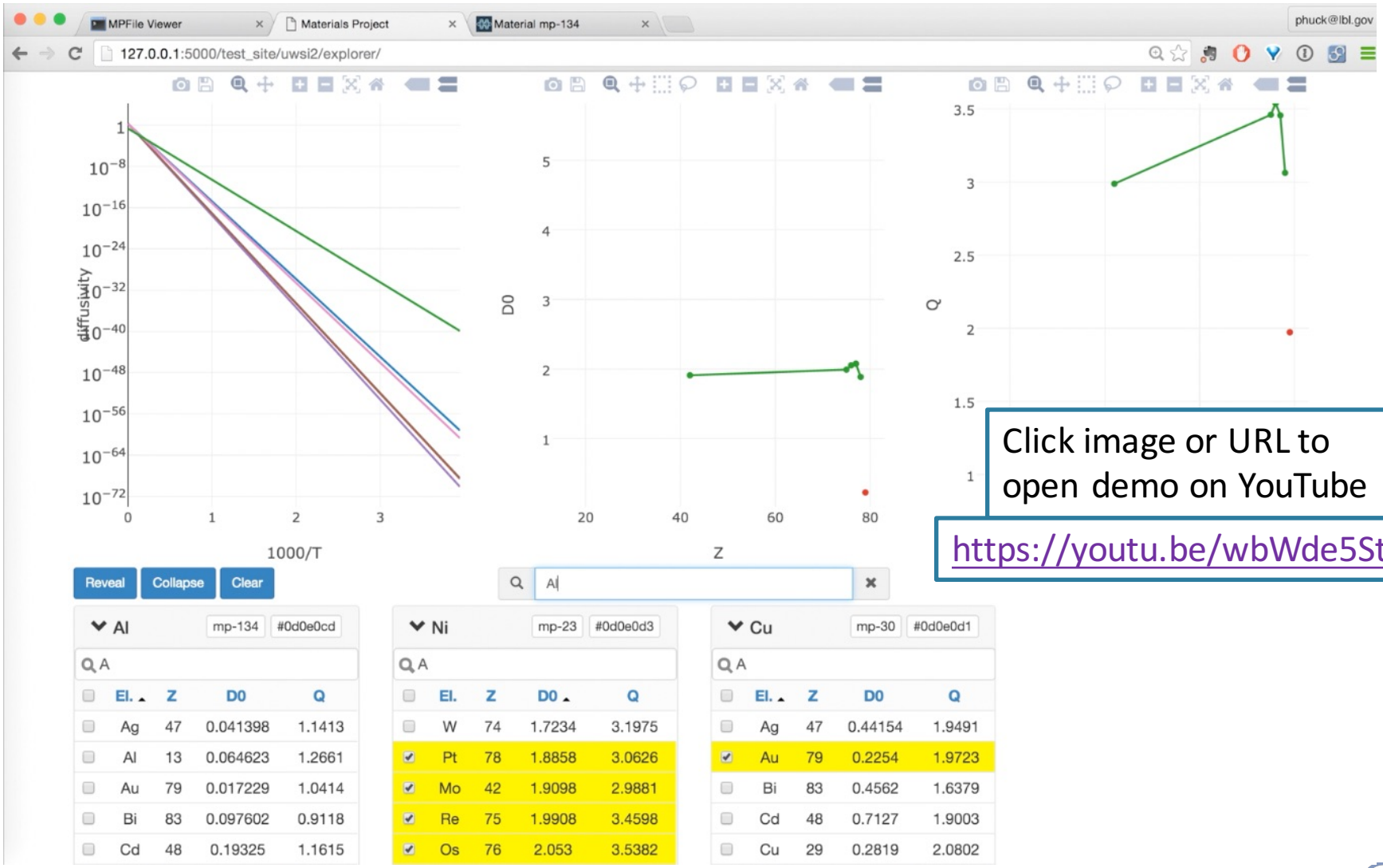
```
{mp-1.physical_properties} for caesium
phase: solid
melting_point: 301.7 K
boiling_point: 944 K
melting_point_density: 1.843 g/cm3
critical_point_temperature: 1938 K
critical_point_pressure: 9.4 MPa

[mp-1.references] list of url, bibtex, or doi
* https://en.wikipedia.org/wiki/Caesium
* http://education.jlab.org/itselemental/ele055.html

{mp-1.plots.default_data_table2} overwrite graph props
x: configuration
y: ionization energy
kind: bar
table: table2

[+mp-1.table1] can be named freely
T, vapor pressure
418,1
469,10
534,100
623,1000
750,10000
940,100000

[+mp-1.table2]
electron number, ionization energy, configuration
1,375.7,6s1/2
2,2234.3,5p3/2
3,3400,5p1/2

[+mp-2] bare data section for palladium
temperature (K), vapor pressure (Pa)
```

ArchieML as MPFile format

MPContribs Software Components



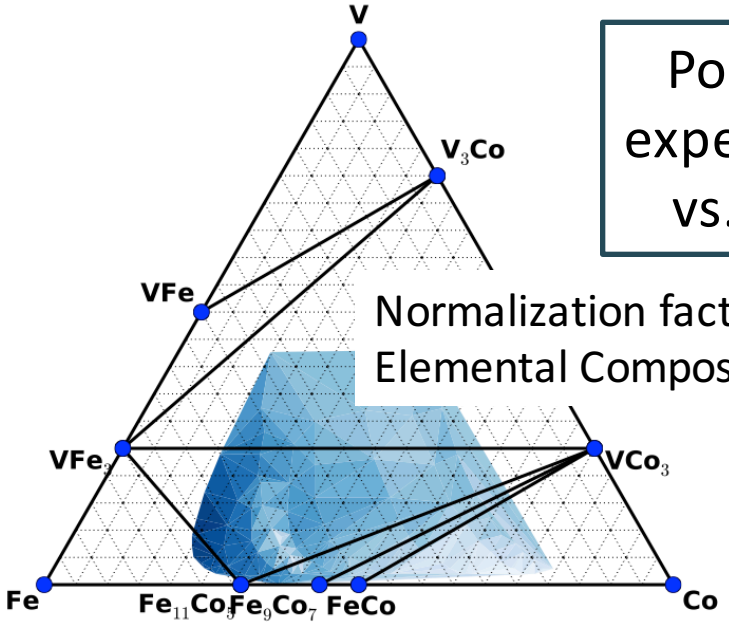
UW/SI2 Diffusion App - Demo



Click image or URL to open demo on YouTube

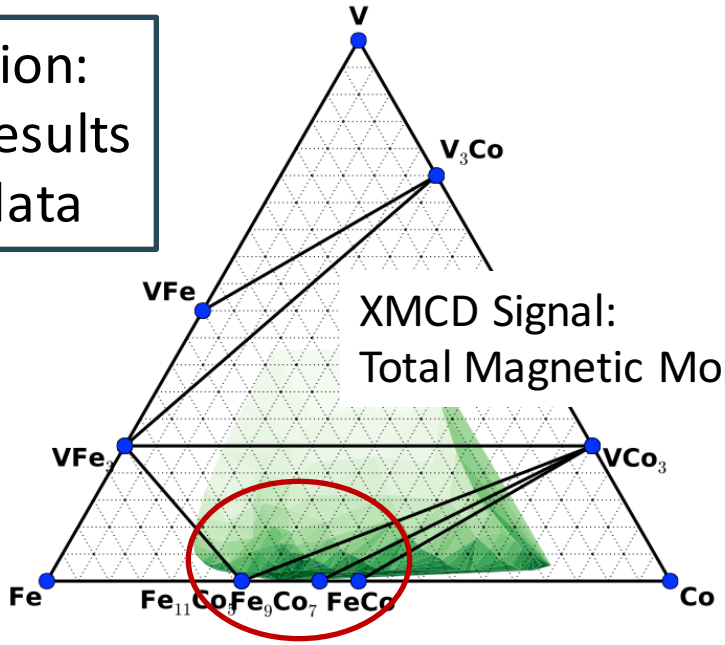
https://youtu.be/wbWde5StHnU

XMCD/Magnetic Moment – Phase Diagram



Post-submission:
experimental results
vs. core MP data

Normalization factors:
Elemental Composition

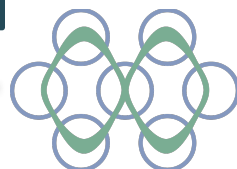


XMCD Signal:
Total Magnetic Moment

Does data coincide w/ any phase transitions or stable compounds?

MPCONTRIBS helps contributor to

1. guide the planning of (follow-up) experiments
2. understand results of unknown materials
3. provide reference of well understood materials (exp. data hub)
4. significantly reduce manual repetitive analysis work



Conclusions

User-contributed data will play an important role in expanding MP towards serving a larger community (wait-list!).

MPComplete:

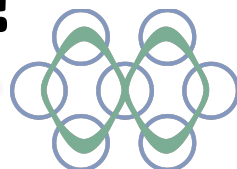
- Underpins MP's Crystal ToolKit App
- Enables user suggestions of new structures for calculation by MP
- MP production at NERSC, but user submissions at XSEDE

MPContribs:

- Software platform for users to share data with >14k users
- Extend core data with user-curated and -contributed data
- Data preparation, interactive graphing & integrated analyses
- Combine contributions and instantly serve on generic frontend
- Basis for other projects to build custom web applications

Docs: http://pythonhosted.org/mpcontribs/
GitHub: https://github.com/materialsproject/MPContribs

Thanks!



Demo: https://youtu.be/xlwttmXSpHg (Full Playlist)