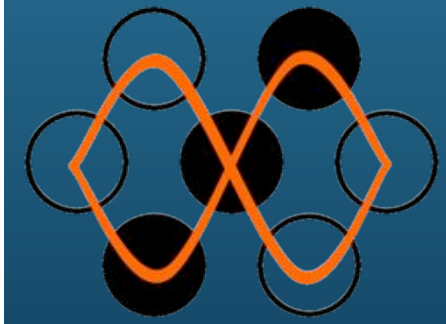
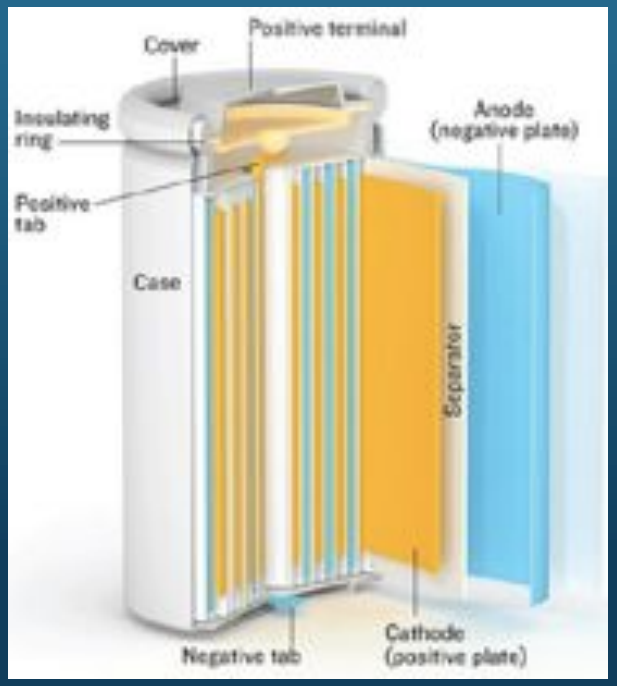


The Materials Genome Initiative and the Materials Project

Gerbrand Ceder
Massachusetts Institute of Technology

MGI Second Anniversary June 25 2013

Materials are Important for Society



Materials Play a Strategic Role Today

Sept 7, 2010

Japan arrest Chinese boat captain

Sept 22, 2010

China blocks shipments of **Rare Earth Metals** to Japan

Sept 24, 2010

Japan releases captain

DOE launches rare earth metals research hub

Fri, 01/11/2013 - 10:37am

Japan and Vietnam Join Forces to Exploit Rare Earth Elements

Short supplies of the metals used in high-tech applications have prompted the nations to establish a joint-research center

Rare Earths Fall as Toyota Develops Alternatives: Philips developing rare earth substitute for LEDs, says CEO

Huge rare earth deposits found in Pacific –Japan experts

News | Videos | Video Details | Disambiguation | Images | Related Links | Twitter | BBC Results | Shopping

The Times of India 2011-07-04 : TOKYO: Vast deposits of rare earth minerals, crucial in making high-tech electronics products, have been found on the floor of the Pacific Ocean and can be readily extracted, Japanese scientists said on Monday. "The deposits have a heavy concentration of rare earths. Just one square kilometre (0.4 square mile) of deposits will be able to provide one-fifth of the current global annual consumption," said Yasuhiro Kato, an associate professor of earth science at... more »

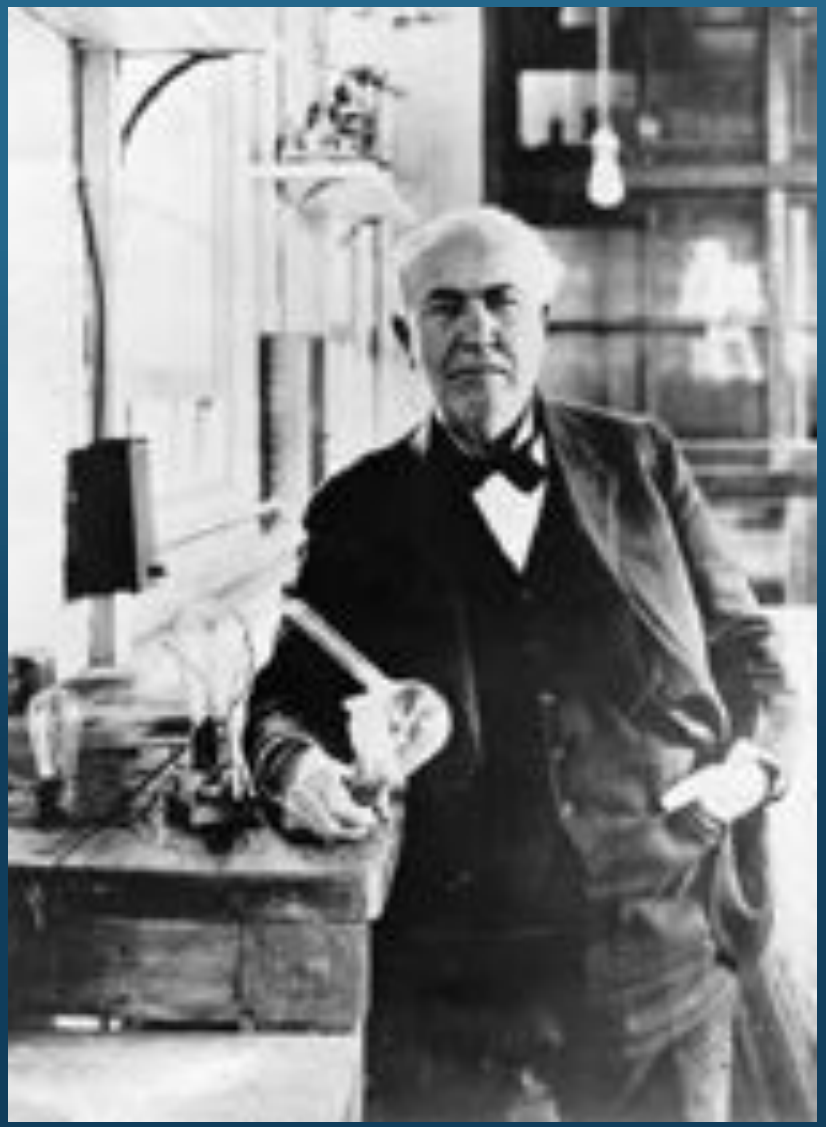
How does one design or “find” new materials ?

“Edison Style”

When looking for a light bulb filament, Edison tried about 3,000 materials

…

And he didn’t find the best one …



2010: Materials Design in Hollywood

Finding a new material in the information age

Finding a new material in the information age

… but we know nothing

There are about **50,000 to 200,000** known inorganic compounds

Elastic constants: about 200 compounds

Super conductors ≈ 1000

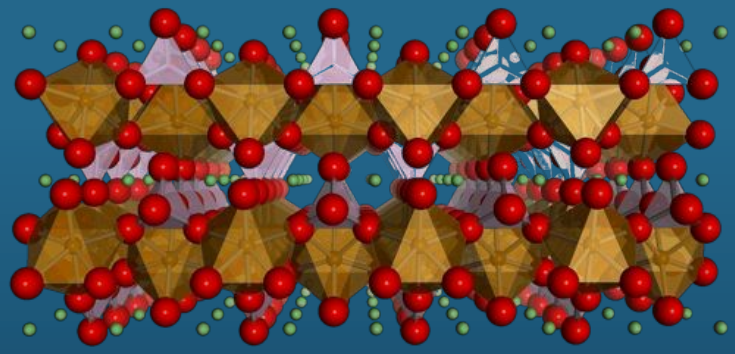
Dielectric constant ≈ 300-400

For almost every property we are below 1% in coverage ….

We know nothing about most materials …

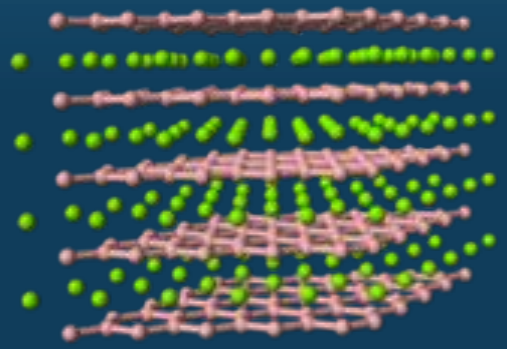
Example: LiFePO4

- First synthesized in 1977
- Only identified as useful battery cathode in 1997



Example: MgB_2

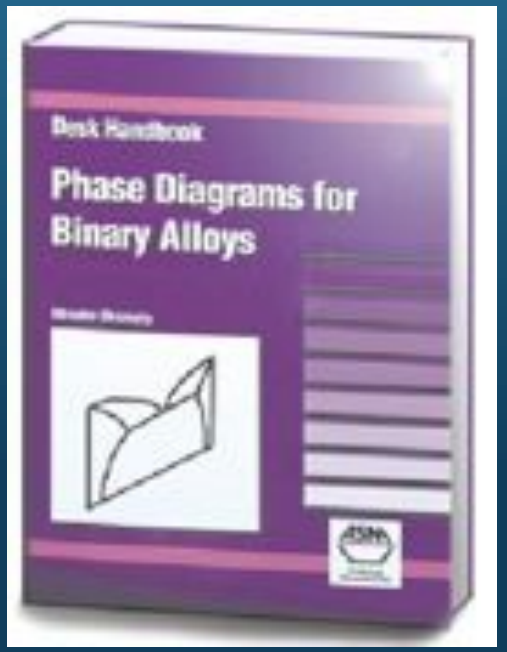
- First synthesized in 1953
- Only identified as high temperature superconductor in 2001



We do not have the basic “roadmaps” for materials processing

≈ 35,000 phase diagrams known, most of them very incomplete

- 150,000 possible ternaries
- ≈ 4 Million possible quaternaries



We are working with this …

When we should have this

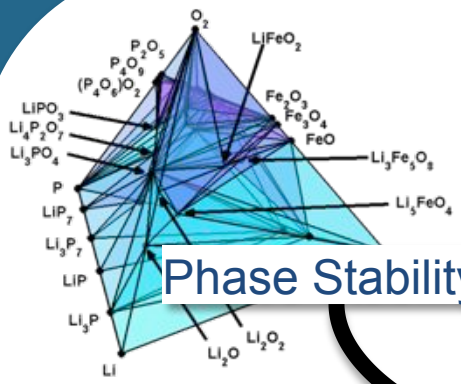
Imagine …

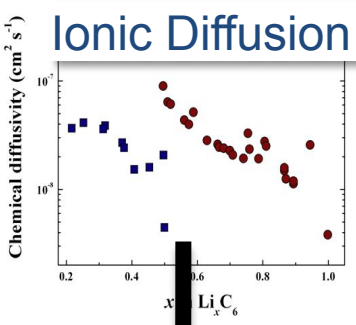
That one knows for every known compound

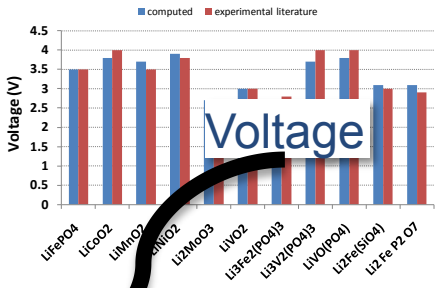
- The bandgap, the effective masses, the Seebeck coefficient
- The elastic constants, the dielectric constants, the absorption spectra
- The point defect energies, solubilities, electronic signature of defects, mobility of defects and impurities, The stacking fault energy, the Peierls dislocation barrier
- Surface energies, in common environments (aqueous, air, …)

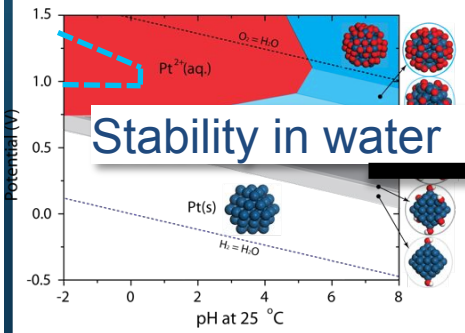
That you had the capability to query these properties for novel compounds upon request …

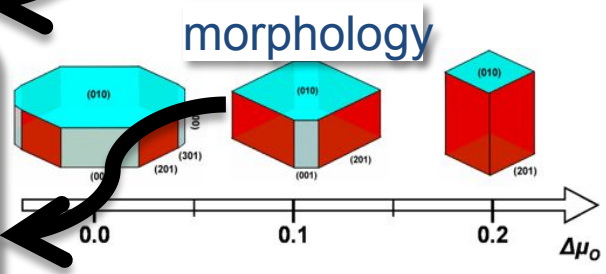
Many Properties are Computable



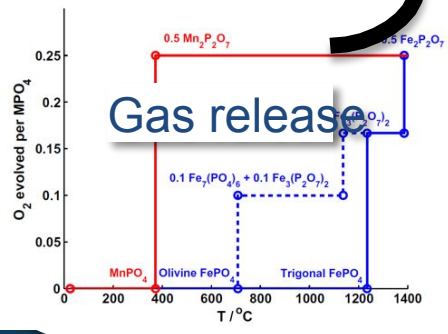


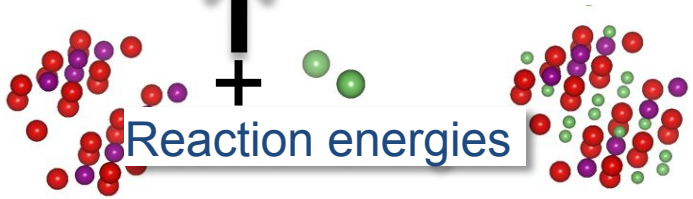




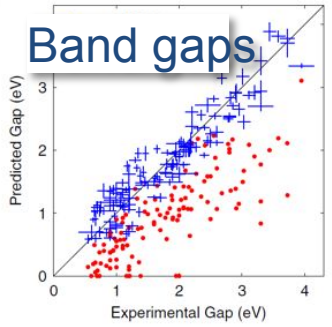


Computational Capability Leveraged for Many Applications!



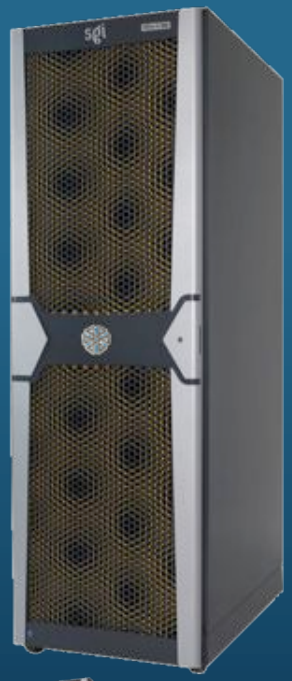


$\Delta H = [\, E(X) + E(Y) \,] - E(XY)$



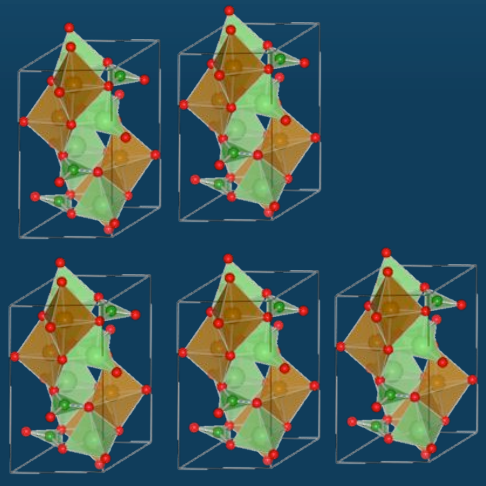
Computations are scalable

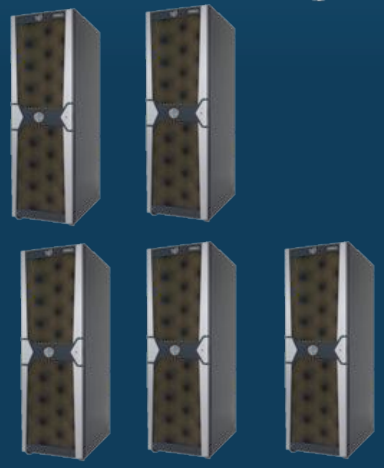
+



+

$$H = \sum_{i=1}^{N_e} \nabla_i^2 + \sum_{i=1}^{N_e} V_{nuclear}(r_i) + \frac{1}{2} \sum_{i}^{N_e} \sum_{j \neq i}^{N_e} \frac{1}{|r_j - r_i|}$$



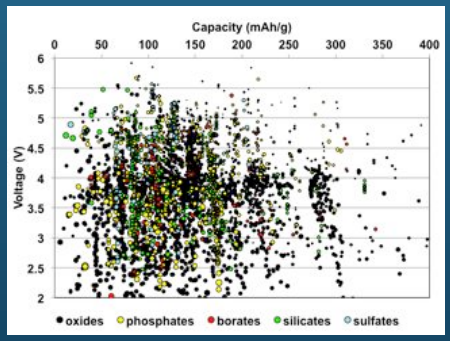


Thought provoking: NERSC flagship Hopper could crunch through ~20k structures in 1 day…

The Materials Project: Vision

www.materialsproject.org

Compute the basic properties ("genes") of materials across the world of compounds and disseminate that information to the materials community to enable rapid materials searching and design.



Lawrence Berkeley National Laboratory

Kristin Persson
Anubhav Jain
Daniel Gunter
Maciej Haranszyk
Wei Chen

Massachusetts Institute of Technology

Gerbrand Ceder
Shyue Ping Ong

National Energy Research Scientific Computing Center

David Skinner
Shreyas Cholia
Jack Deslippe

University of California, Berkeley

Mark Asta

National University of Singapore

Stefan Adams

The University of Wisconsin Madison

Dane Morgan

Duke University

Stefano Curtarolo

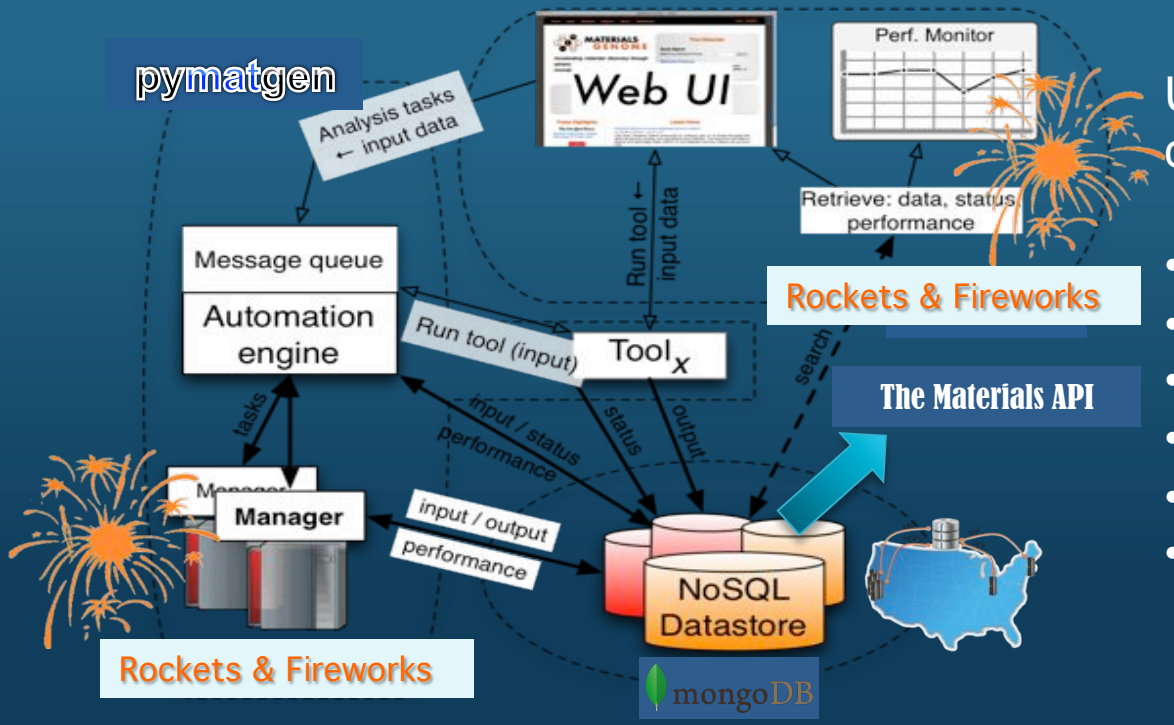
University of Kentucky

Alan Dozier
Raphael Fink

UCL Université Catholique de Louvain, Belgium

Geoffroy Hautier

Large Software Infrastructure Structure

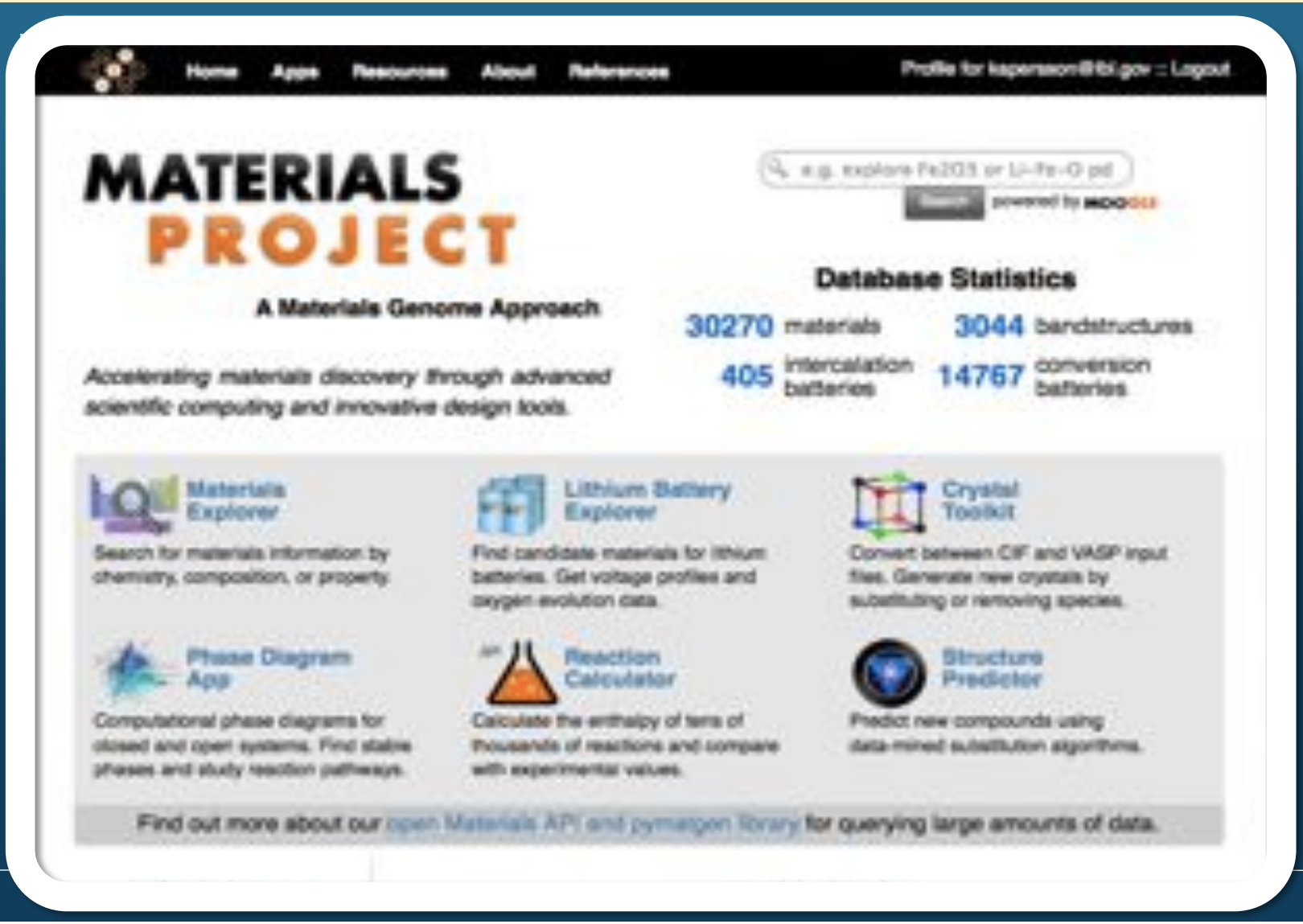


Use high-throughput ab initio computing to determine:

- Energies
- Crystal structures
- Band structures
- Defect energies
- Absorption spectra
- …

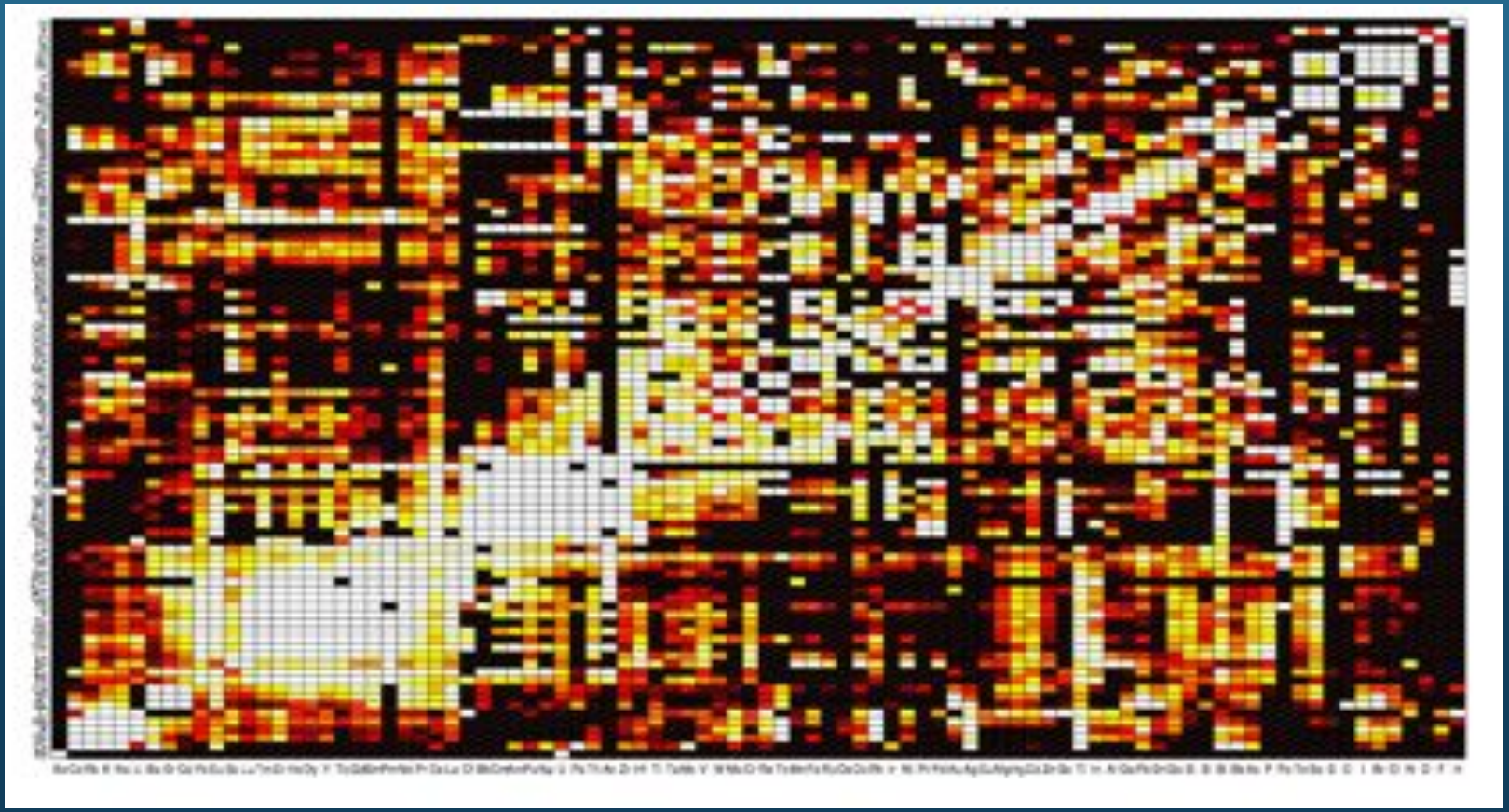
Today's Status:

- ❑ Over 30,000 compounds and growing monthly
- ❑ Multiple tools based on computed data



Machine learned substitution rules

which elements can substitute for each other ?

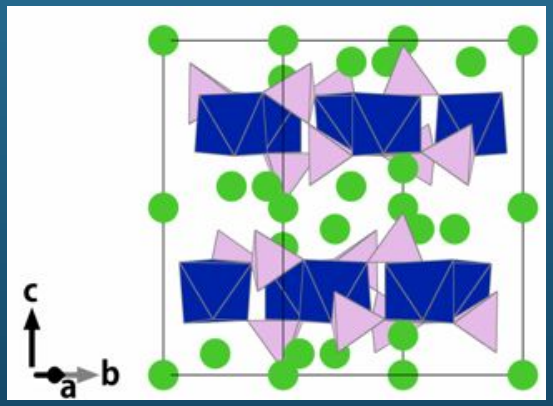


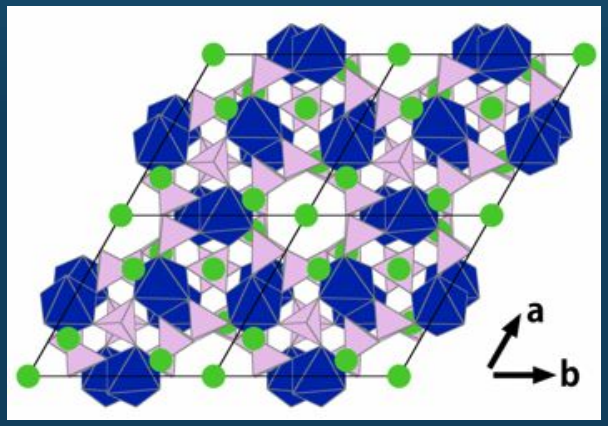
Data mined from all known inorganic compounds

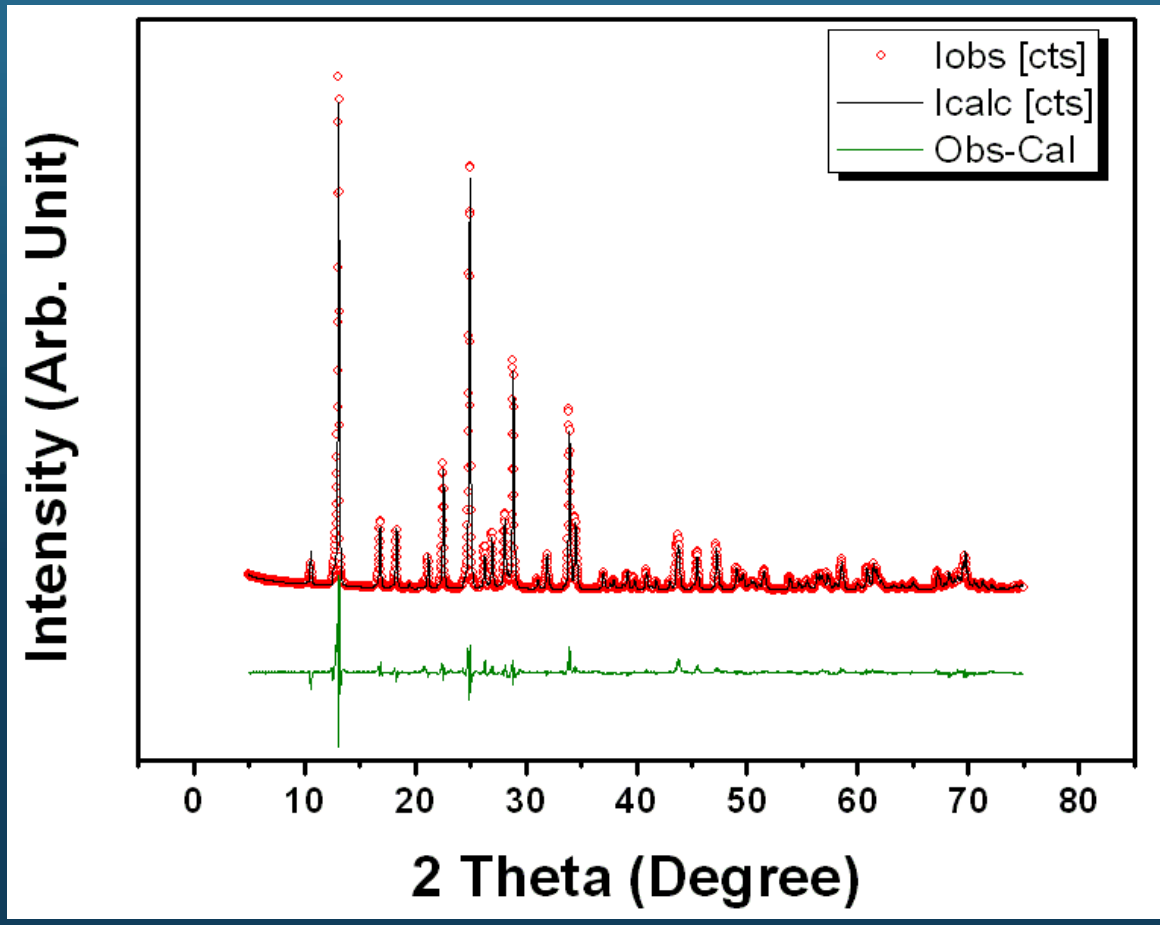
G. Hautier, et al, *Data Mined Ionic Substitutions for the Discovery of New Compounds*, Inorganic Chemistry, 50 (2), 656-663 (2011).

New Phosphate Discovered as Battery Material: $Li_9V_3(P_2O_7)_3(PO_4)_2$

This compounds was not known to exist

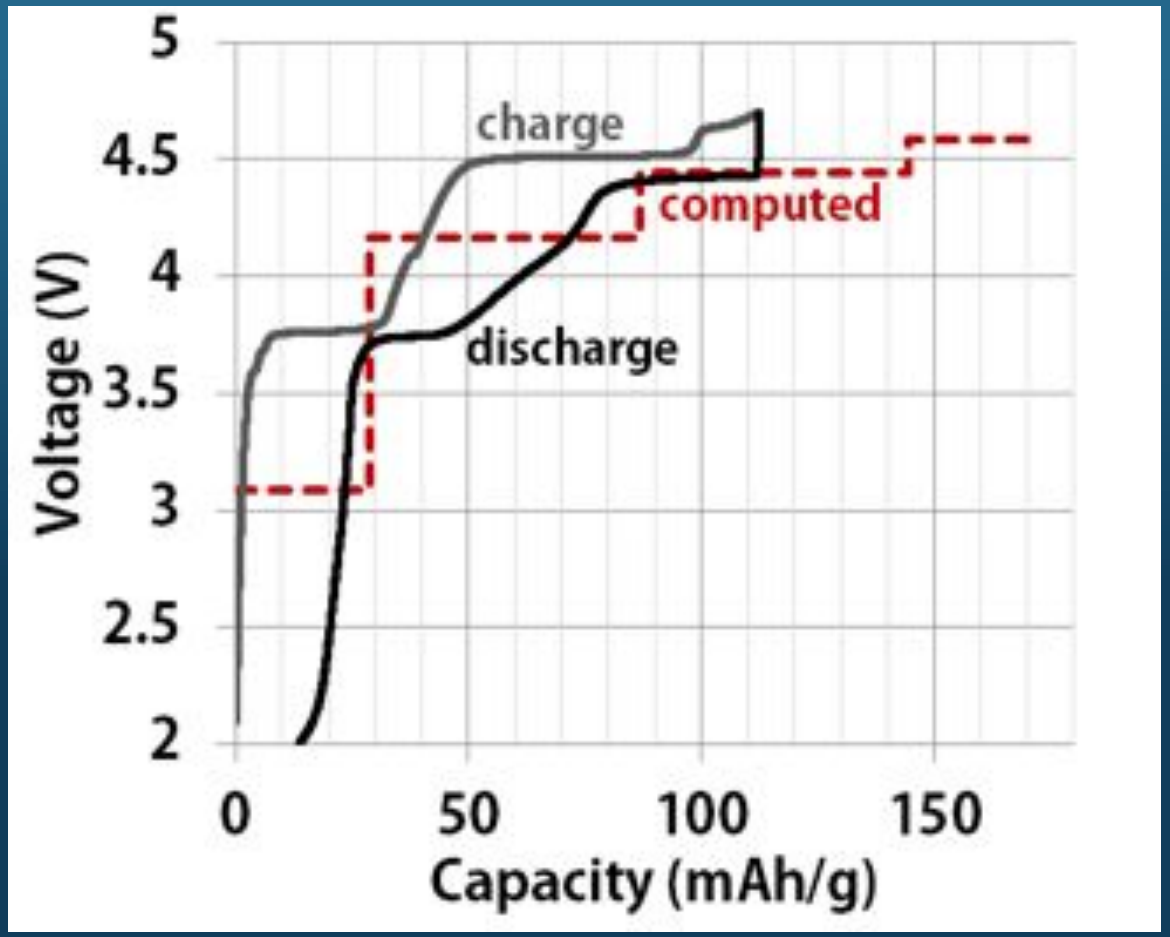






A. Jain et al., J. Electrochemical Society, 159 (5), pp. A622-A633 (2012).

New Phosphate $Li_9V_3(P_2O_7)_3(PO_4)_2$ performs well



Open Source under the hood

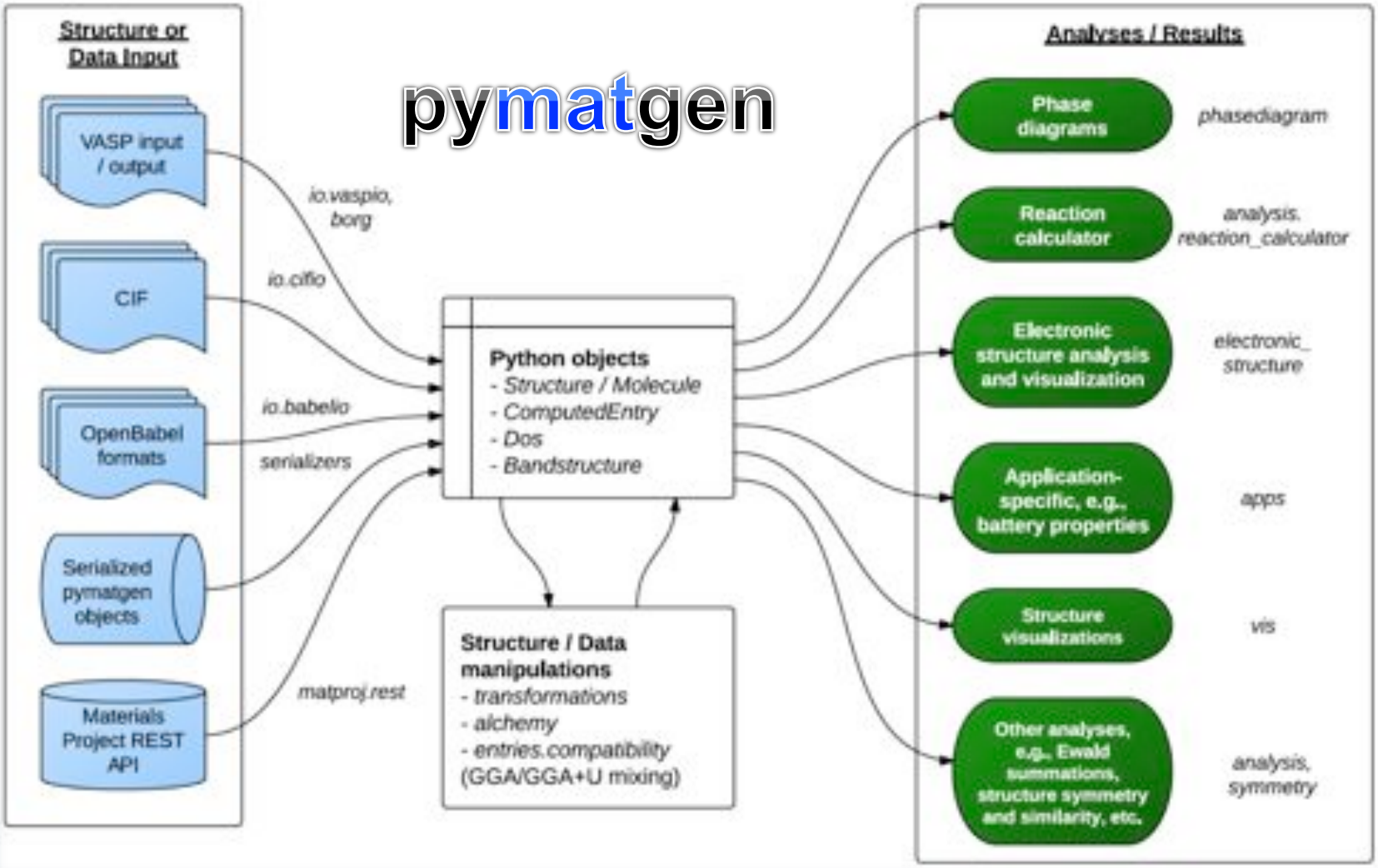
django

pymatgen

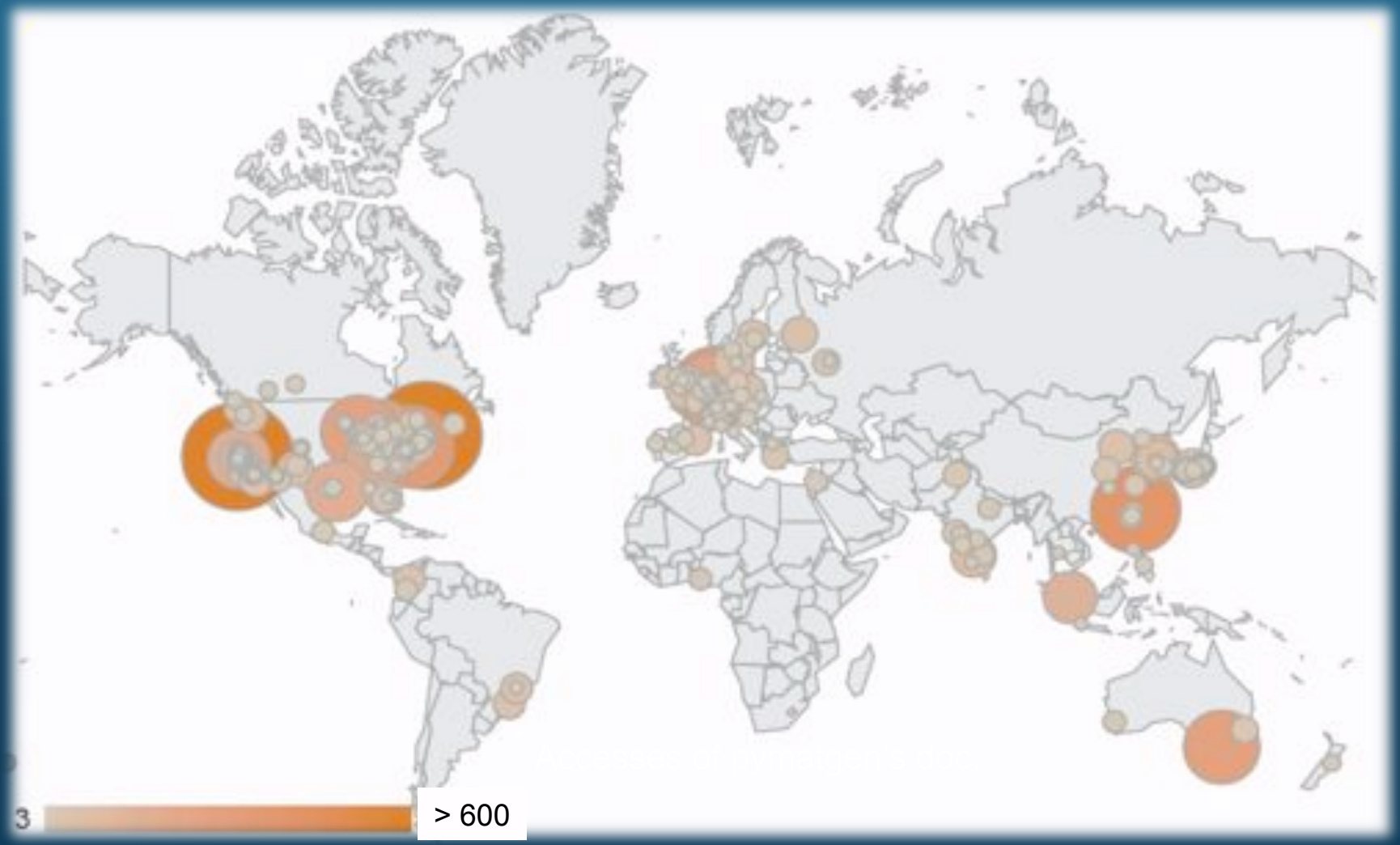
- Database interface
- Run management (fireworks)

Defines core extensible Python objects for materials data representation.

pymatgen

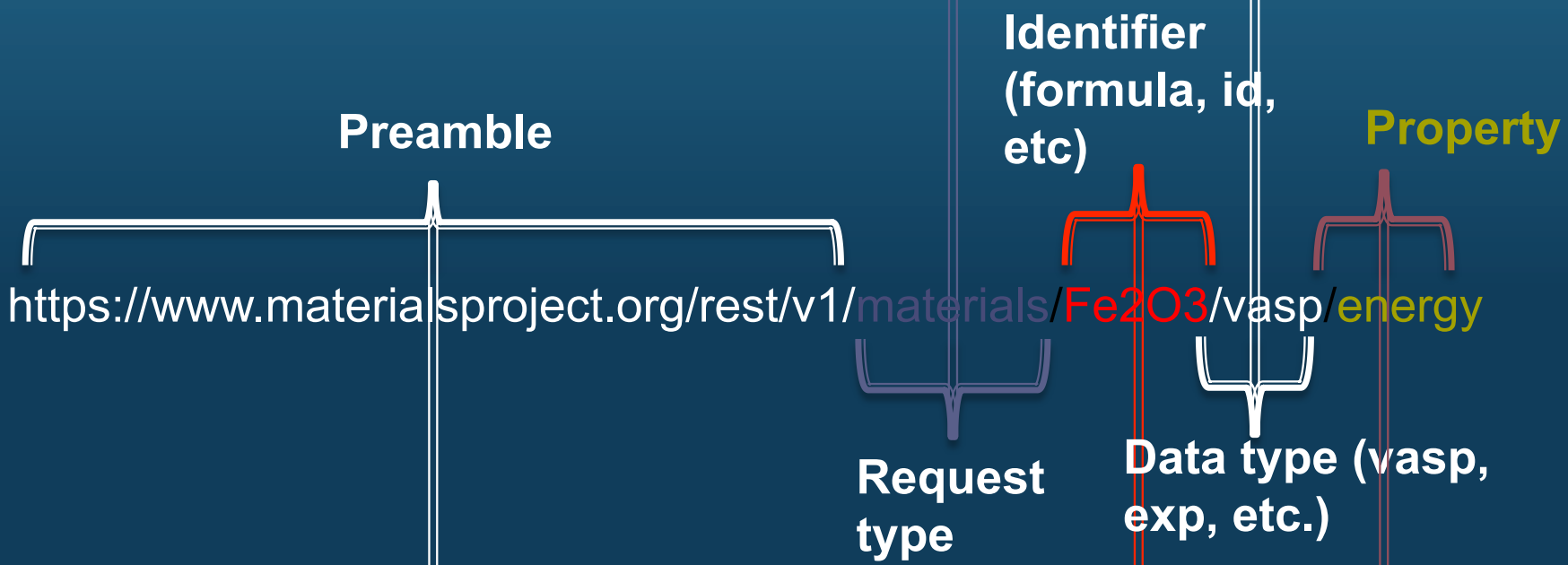

Structure or Data Input
VASP input / output
CIF
OpenBabel formats
Serialized pymatgen objects
Materials Project REST API
io.vaspio, borg
io.cifio
io.babelio
serializers
matproj.rest
Python objects
- Structure / Molecule
- ComputedEntry
- Dos
- Bandstructure
Structure / Data manipulations
- transformations
- alchemy
- entries.compatibility
(GGA/GGA+U mixing)
Analyses / Results
Phase diagrams
phasediagram
Reaction calculator
analysis. reaction_calculator
Electronic structure analysis and visualization
electronic_ structure
Application-specific, e.g., battery properties
apps
Structure visualizations
vis
Other analyses, e.g., Ewald summations, structure symmetry and similarity, etc.
analysis, symmetry

The Codes are Freely Available



The Materials API allows access at the "deep" data levels

An open platform for accessing LOTS of Materials Project data



API keys available at
https://www.materialsproject.org/profile

An open platform for accessing data based on REpresentational State Transfer (REST) principles

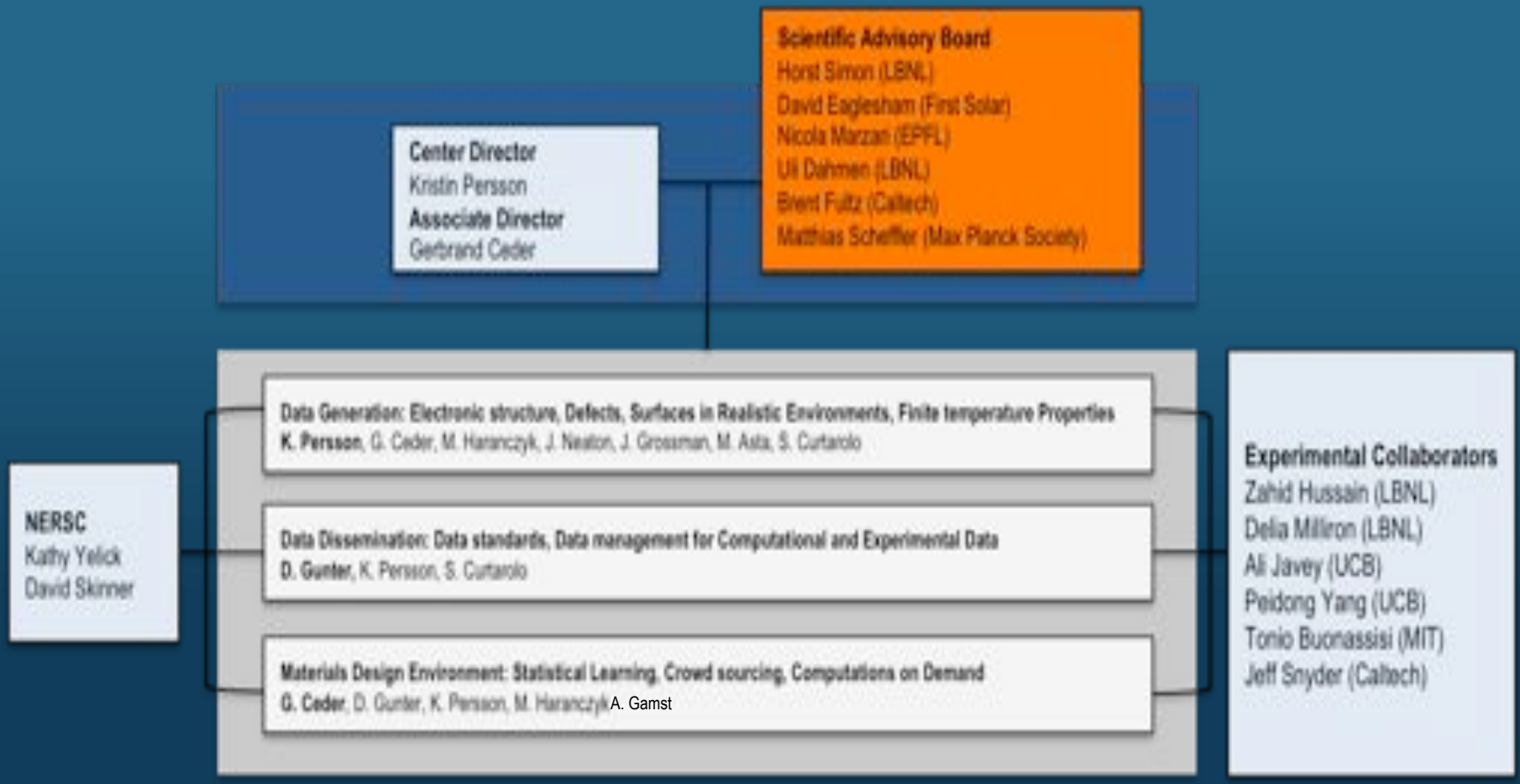
Improved accessibility of data

More developers of analyses and apps

Increased data value

Ongoing Work

The Materials Project Center; DOE funded

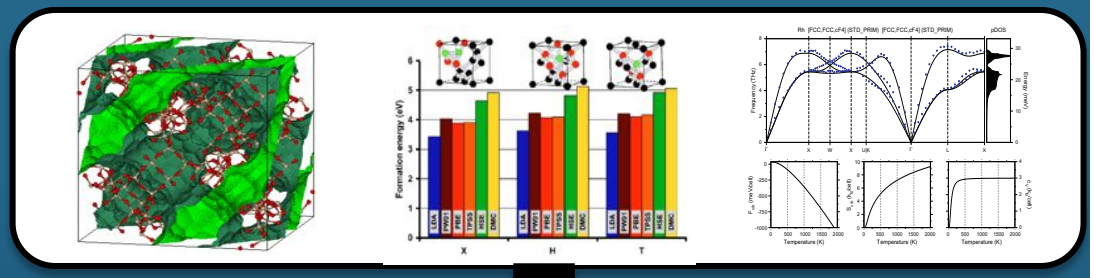


A. Gamst

From Data to Materials Design

Thrust 1: Generate Materials Data

- Higher order electronic structure
- Defects
- Surfaces
- Finite temperature properties

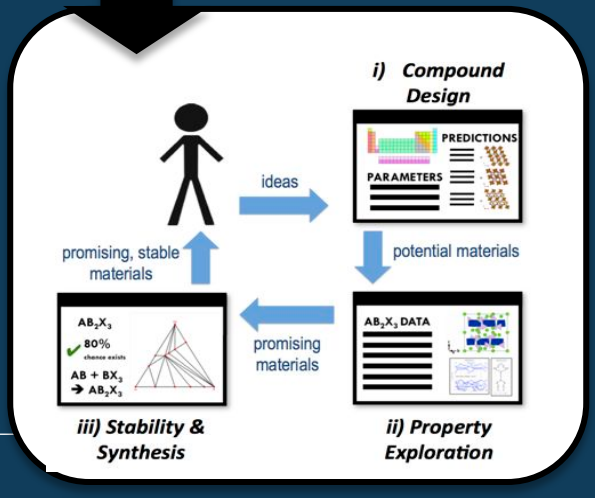


Thrust 2: Disseminate Data

- Populate the database
- User-friendly web and REST interface
- Experimental data

Thrust 3: A Materials Design Environment

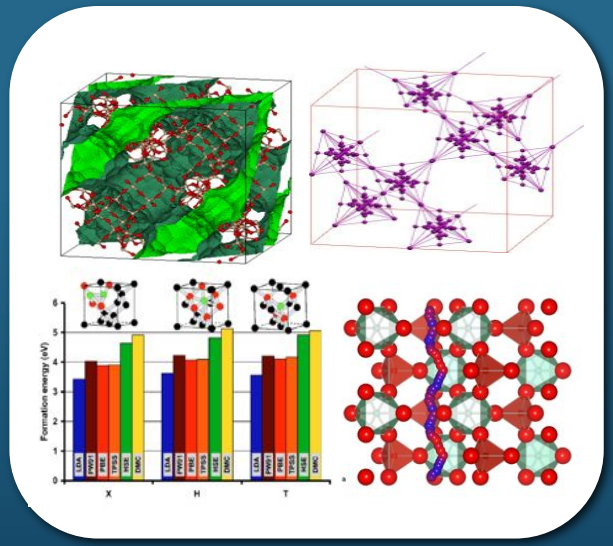
- Data mining tools
- "Computations on Demand"



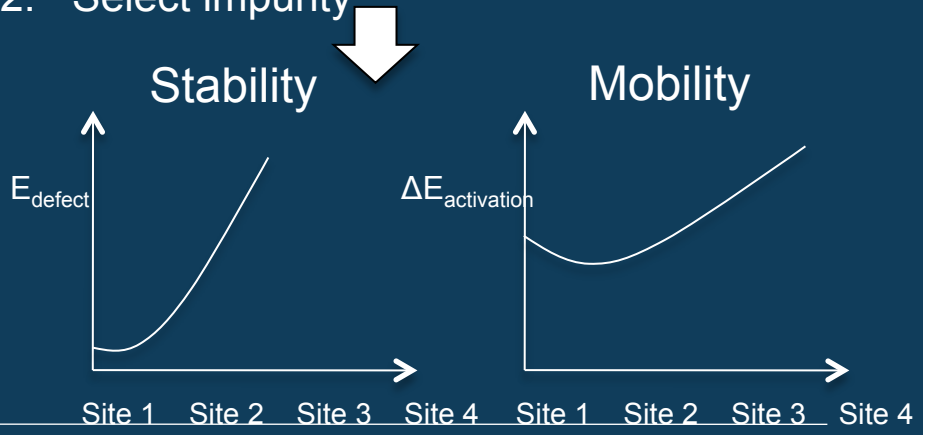
Data Generation : Defects

FINDING/RANKING POSSIBLE DEFECTS:

- Topological algorithms
- A rough ranking using empirical potentials, accounting for simple electrostatics and elastic interaction
- Benchmarking



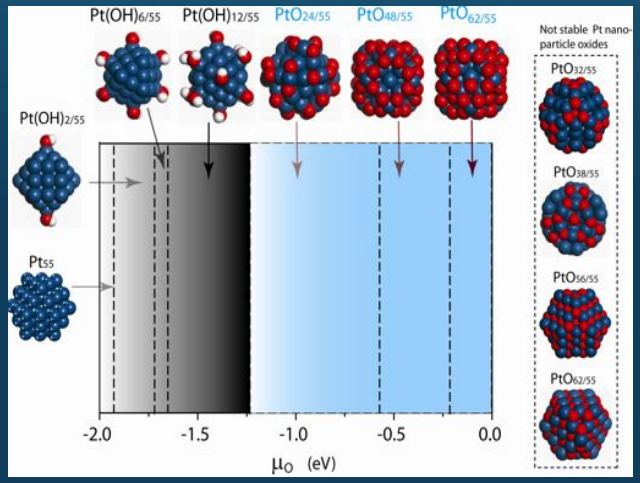
1. Select compound (crystal structure + chemistry)
2. Select impurity

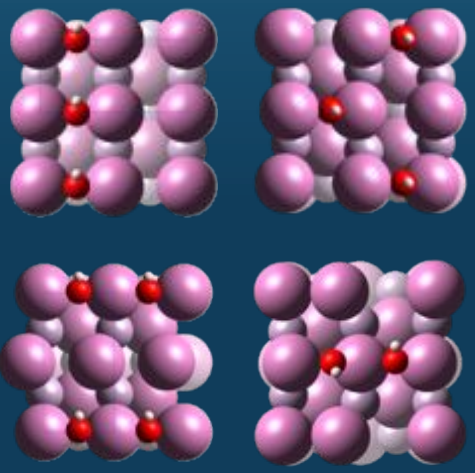


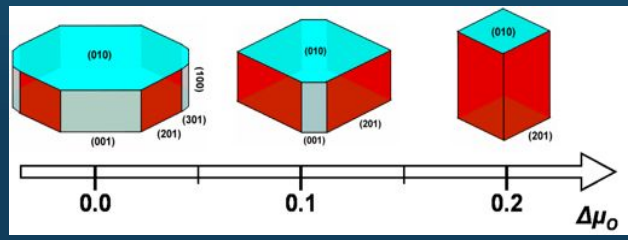
Data Generation : Surfaces

- Tools to generate surfaces; dipole removal;
- Site and coverage rules for O, CO_2, H, OH, H_2O absorbents
- Aqueous stability of bulk and surfaces

GOAL: Computed surface energies, structures and Wulff shapes for most inorganic compounds; Pourbaix Diagrams

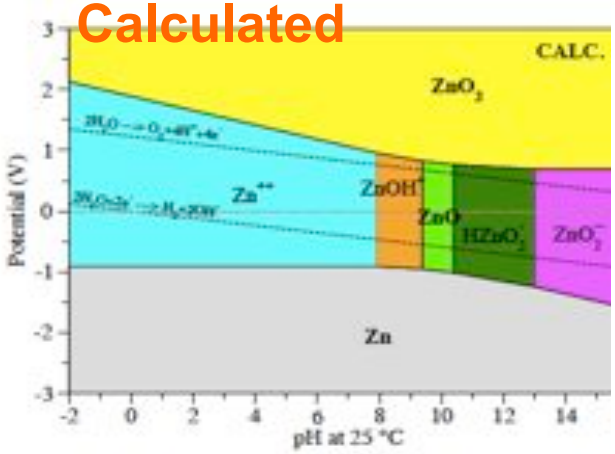


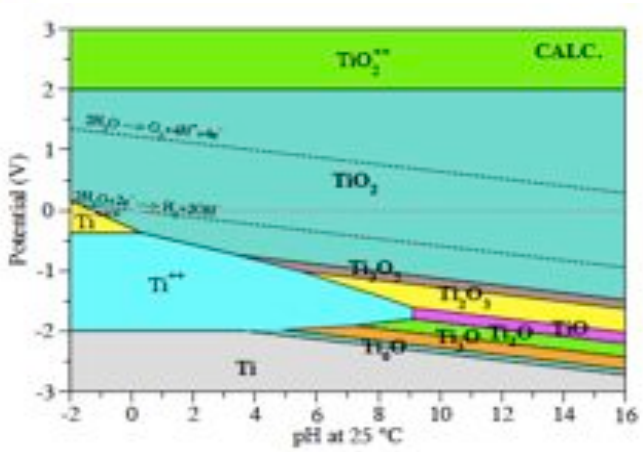


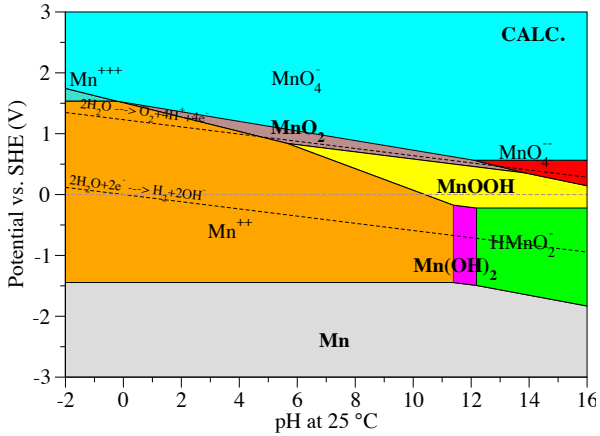


Bulk Pourbaix Maps

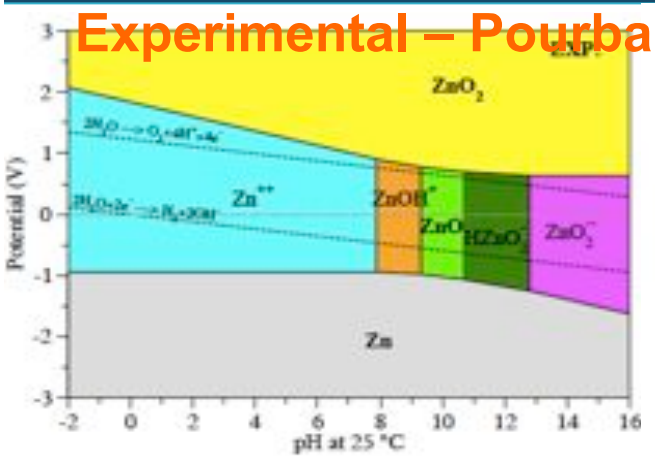
Calculated

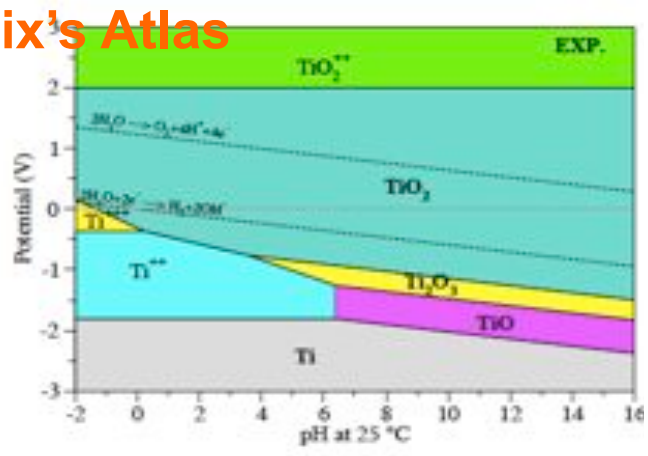


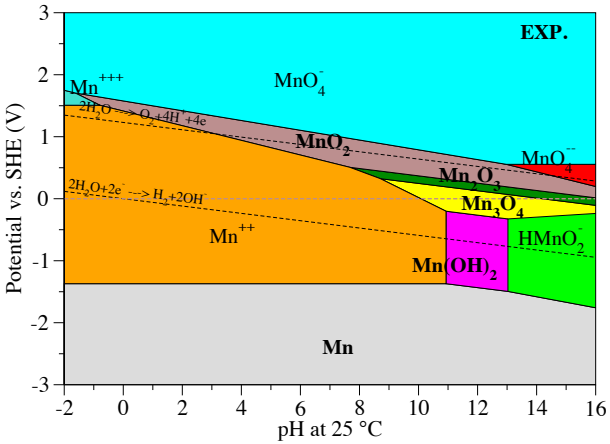




Experimental – Pourbaix's Atlas

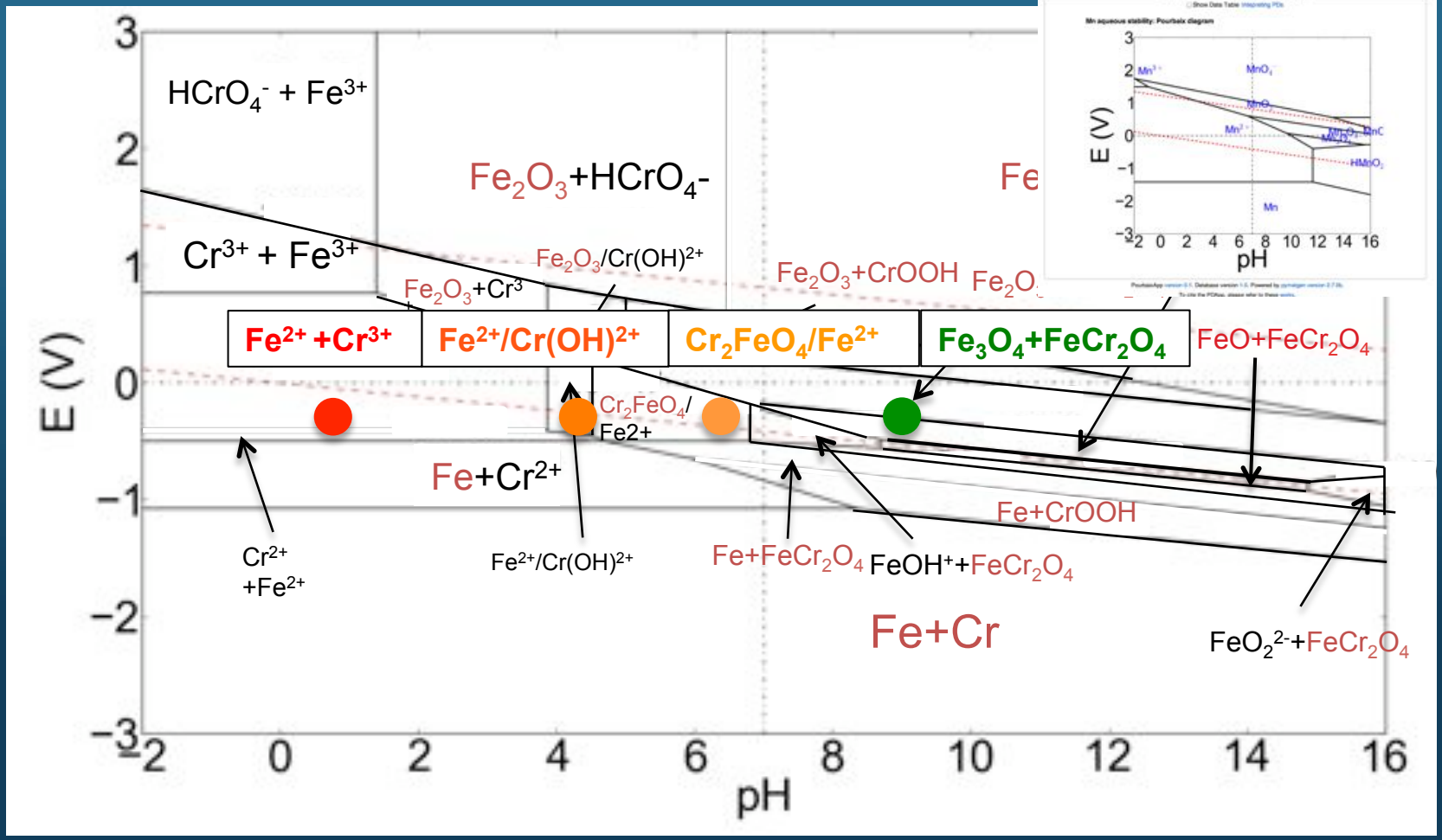






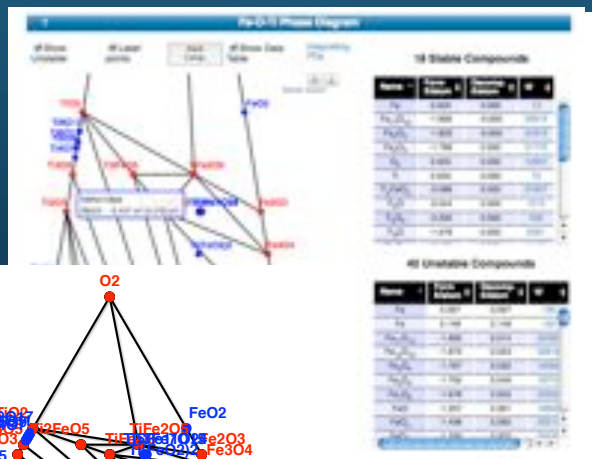
Persson et al, PHYSICAL REVIEW B**, 005400 (2012),** *Prediction of solid-aqueous equilibria: Scheme to combine first-principles calculations of solids with experimental aqueous states*

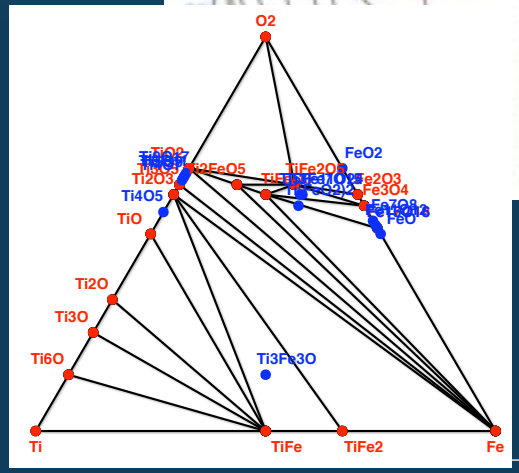
Multi-component Analysis

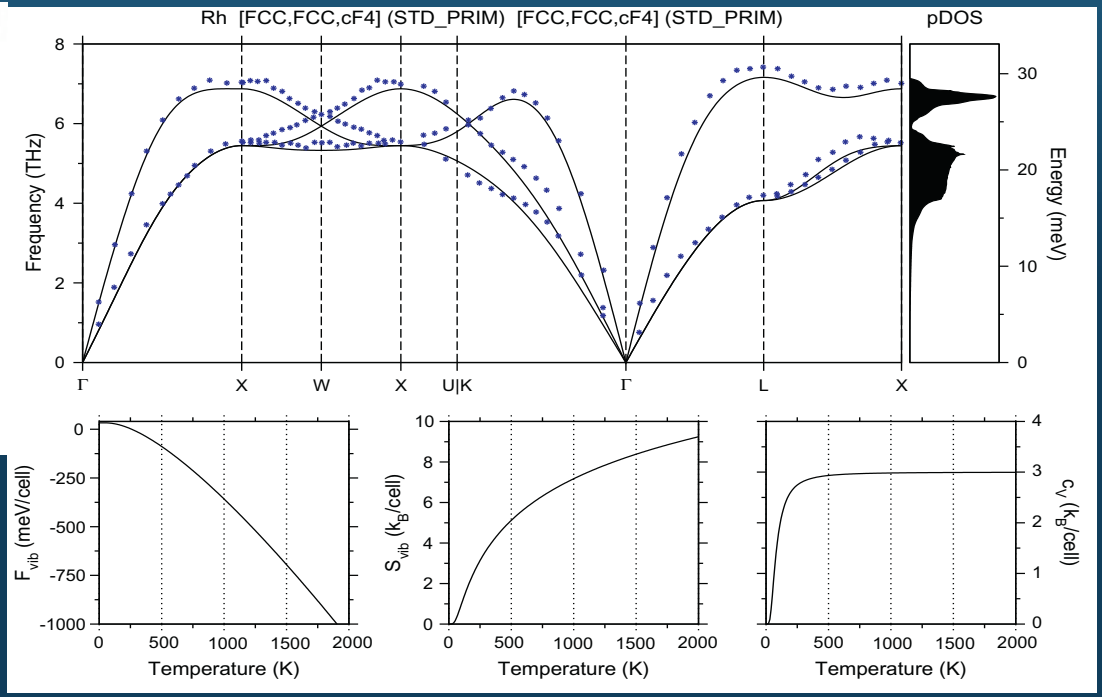


Data Generation: Finite Temperature Properties

- vibrational spectra, electron-phonon coupling, heat capacities, free energies, Debye temperature, Gruneisen parameters, vibrational heat conductivities and thermal expansion…
- Finite temperature solid-state phase diagrams
- Comparison with available experiments



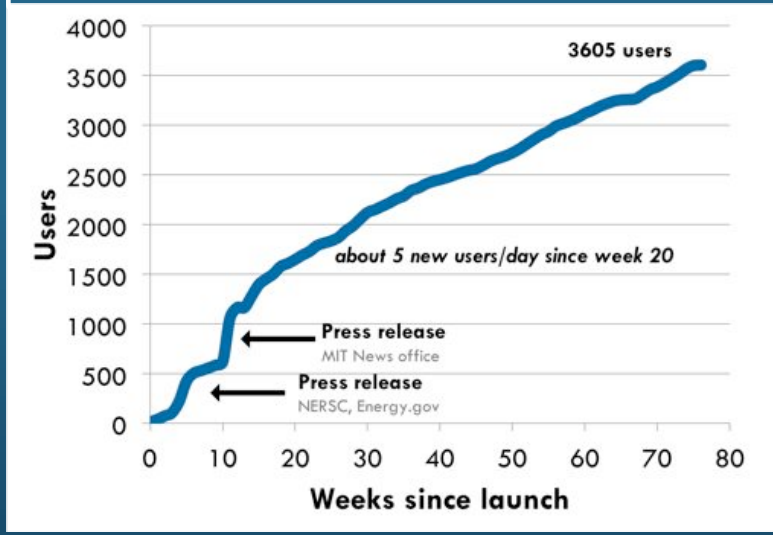




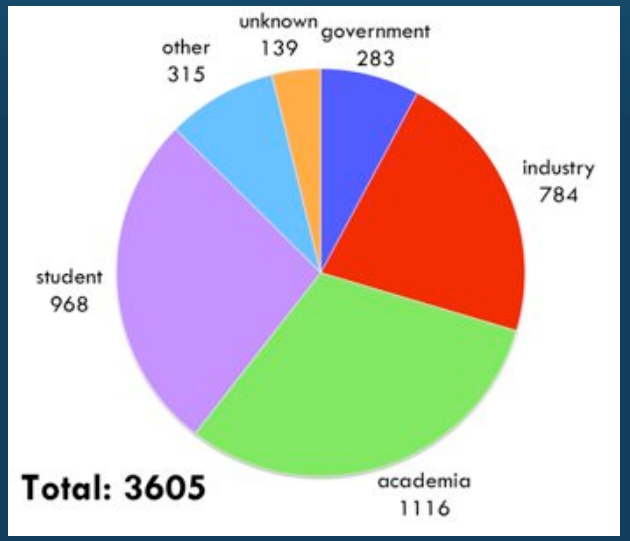
Build it and They will Come ?

The Materials Project: Status

- Early infrastructure from MIT fully recoded at LBNL + NERSC
- Launched online Oct 2011
- $11M/5 year DOE Center grant
- > 4,000 registered users
- > 200,000 materials records downloaded in 8 months through RESTful interface



- **Companies:** Toyota, Sony, Bosch, 3M, Honda, Samsung, LG Chem, Dow Chemicals, GE Global Research, Intermoleular, Applied Materials, Energizer, Advanced Materials, General Motors, Corning, DuPont, Nippon Steel, L'Oreal USA, Caterpillar, HP, Unilever, Lockheed Martin, Texas Instruments, Ford, Bose, Sigma-Aldrich, Siemens, Raytheon, Umicore, Seagate, Intermolecular, …

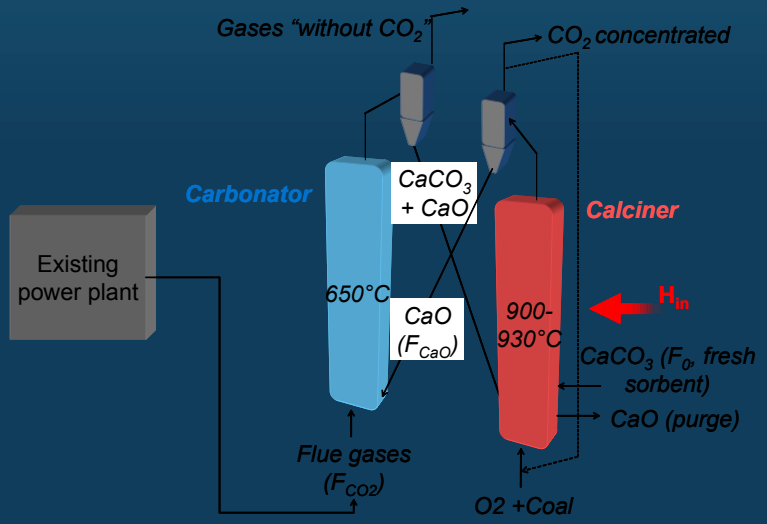


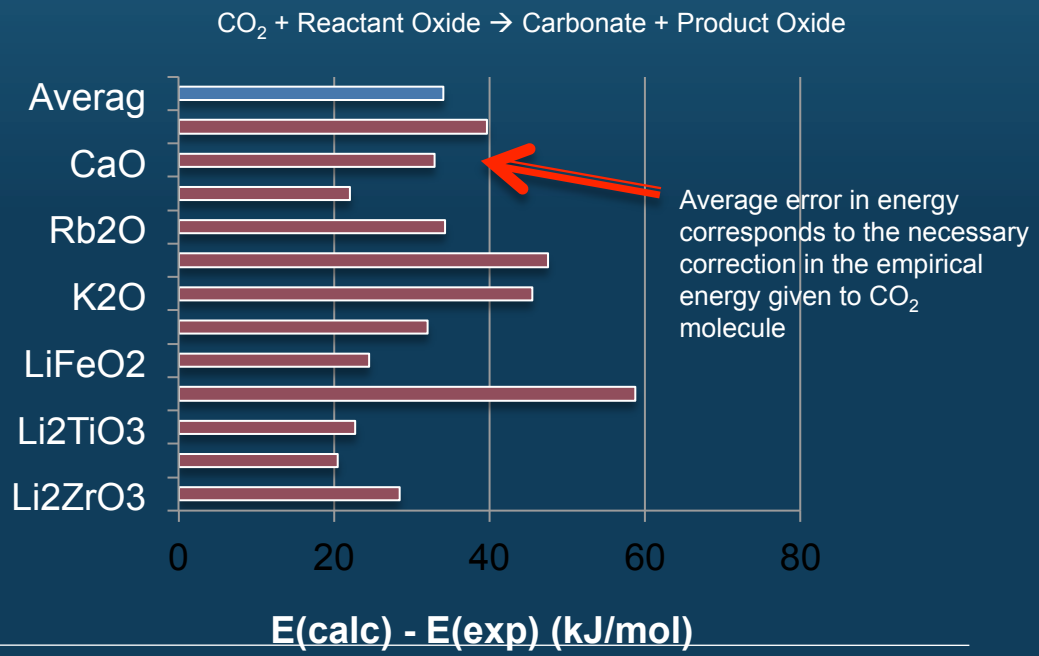
- **Battery startups:** Envia systems, Nanoexa, Pellion, Sion Power, Planar Energy Devices, Phostech, PolyPlus

UK: Screening for CO_2 sorbents

Calculating reaction energies of thousands of oxides with CO_2

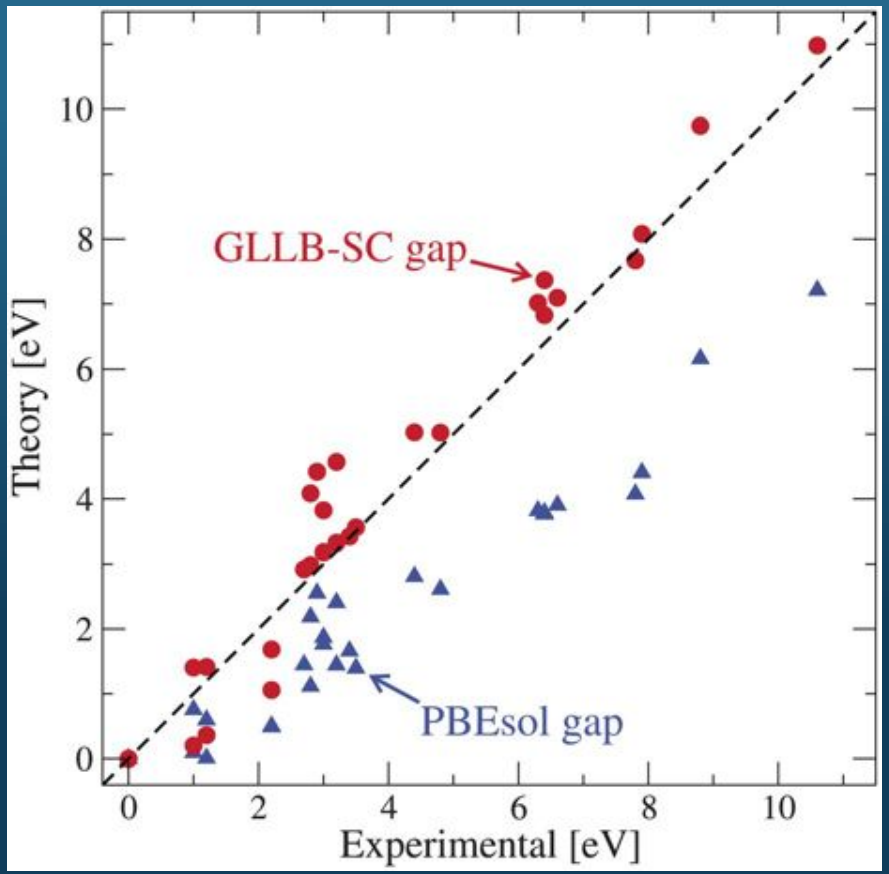
- High selectivity and absorption capacity for CO2 at elevated temperatures (400-900°C)
- Good absorption/desorption kinetics - preferably under a wide range of $p(CO_2)$
- Good cyclability for absorption/desorption processes
- Good hydrothermal and mechanical properties





Denmark: New Methodologies

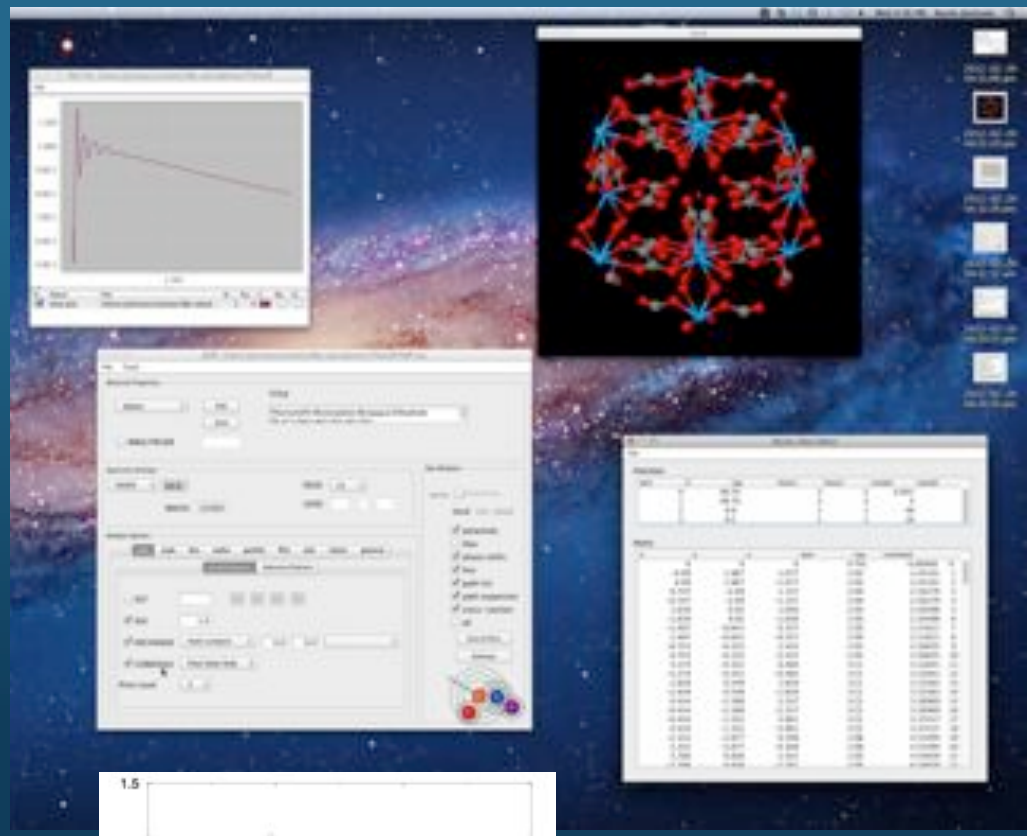
- Testing new functional: GLLB-SC
- GLLB-SC improves performance on band gaps?
- 20,000 optimized structures from the Materials Project allow for comprehensive testing of GLLB-SC

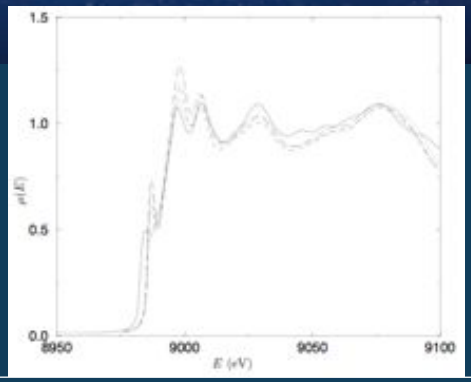


I.E. Castelli, T. Olsen, S. Datta, D.D. Landis, S. Dahl, K.S. Thygesen, and K.W. Jacobsen, Energy & Environmental Science 1–6 (2012).

USA: Absorption/Emission Spectra

- computed K-edge cross sections
- L-edge in future…
- XANES, EXAFS, EELS, …
- cluster of atoms – not periodic solid



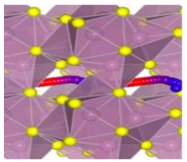


Cu XANES spectrum from exp and FEFF; *Inelastic Losses in X-ray Absorption Theory*, Thesis, Luke Campbell (2002).

JCESR: A Genomic Approach to Electrolytes

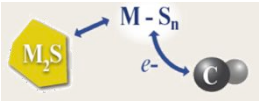
Substitution, mutation and selection to arrive at optimal electrolytes and redox molecules

Multivalent Intercalation



Host Mobility

Chemical Transformation

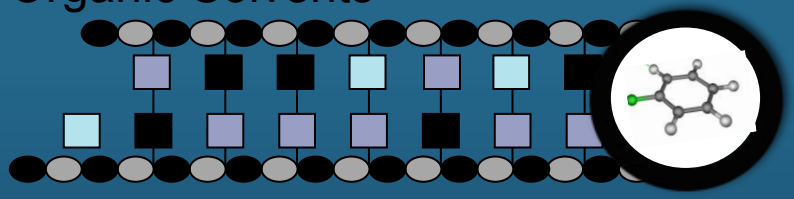


Phase Transformation

Non-aqueous Redox Flow

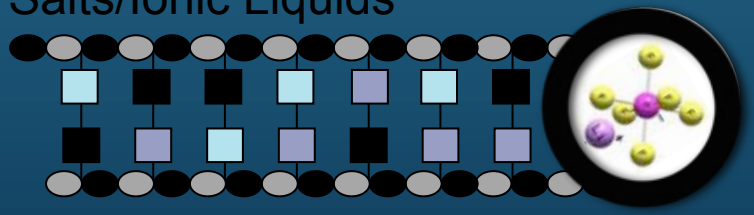
Novel Redox Species

Organic Solvents



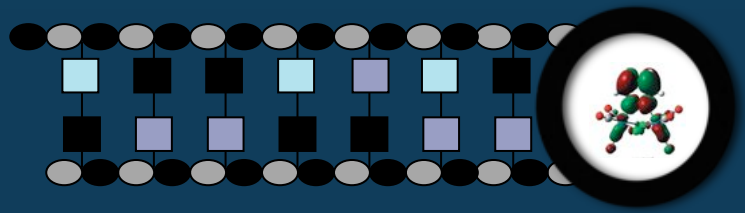
- electrochemical window
- decomposition
- viscosity
- thermal stability

Salts/Ionic Liquids



- electrochemical window
- dissociation strength
- viscosity

Redox Molecules

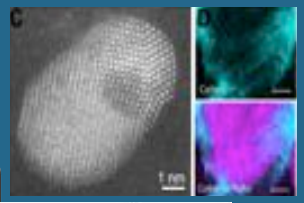


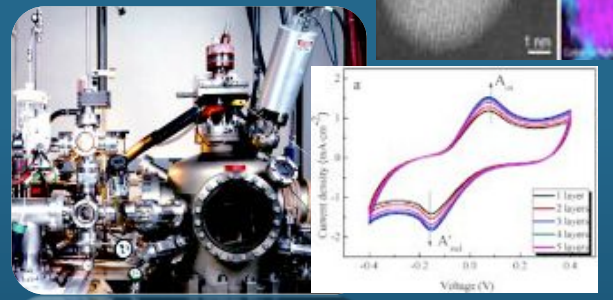
- redox potential
- solubility

BERKELEY LAB

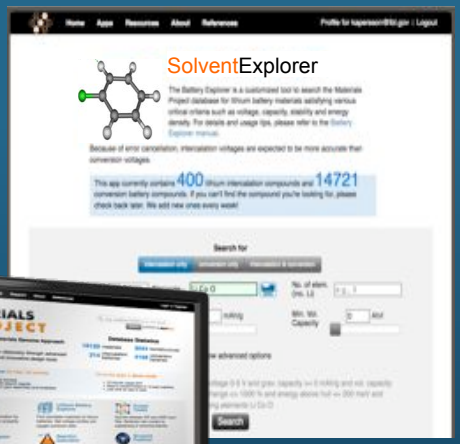
High-throughput Screening and Interactive Design of Electrolytes

Test/Validate

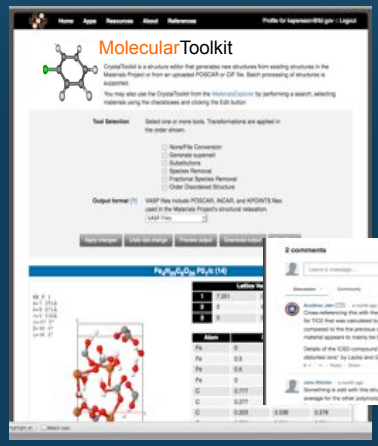




Search through thousands of solvents, salts, redox molecules



Design



Automatic routines generating new molecules + individual online ‘calculations-on-demand’

Search

High-throughput data generation

Calculate

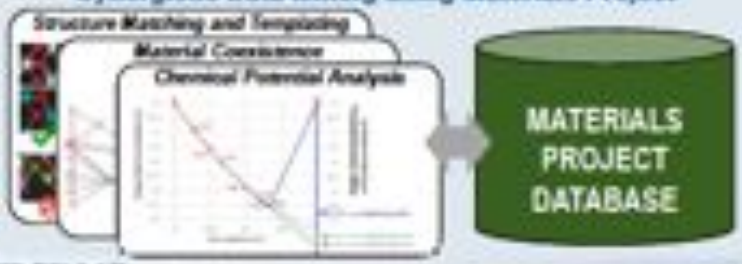
Interactive JCESR community through web commenting

Successful uses of Materials Project in Industry

Materials Project Leveraged by Intermolecular's Combinatorial Acceleration Platform

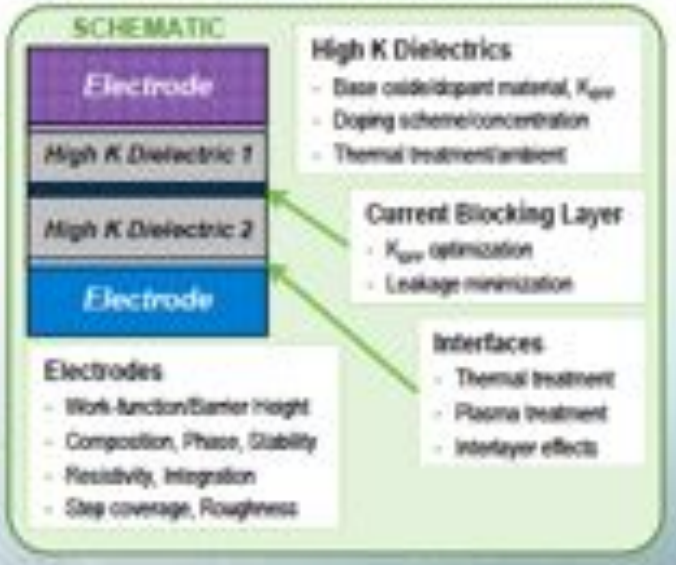
Accelerated development of new materials and manufacturing processes

Intermolecular Informatics Infrastructure
Data Management & Analysis



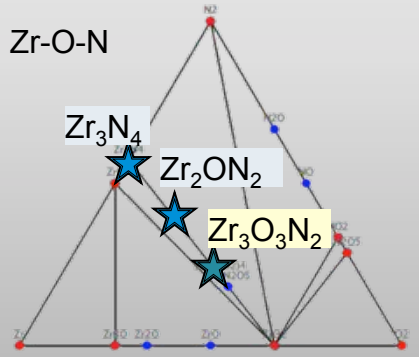
INDUSTRIAL CHALLENGE EXAMPLE

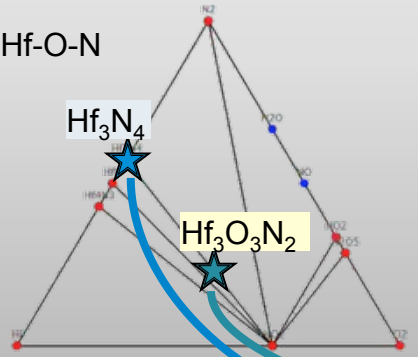
Dynamic Random Access Memory (DRAM) IC cell needs integration of many complex layers



Intermolecular Experiment Screens for New Materials Predicted by High-Throughput Computation

Superior new materials for water-splitting photocatalysis predicted theoretically by the members of Materials Project





Photocatalytic materials:

- Known
- Predicted

Non-photocatalytic materials:

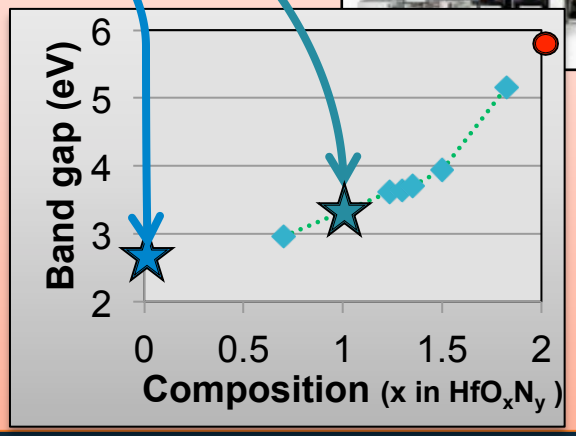
- Stable
- Known metastable

- ✓ Band gap suitable for UV photocatallysis
- ✓ Level alignment perfect for water splitting

Y.Wu et al., Energy Environ. Sci. 6, 157 (2013)

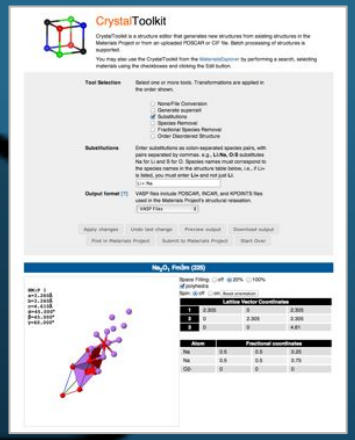
Materials developed by Intermolecular show properties matching these predictions

- ✓ Focus on materials with multiple applications
- ✓ Synthesis methods developed
- ✓ Range of O:N compositions scanned
- ✓ Optical characterization data match the predicted band gap
- ✓ Future workflow could focus on optimizing the water splitting capabilities



Next Steps: Materials Design as an Open Source Cloud Service

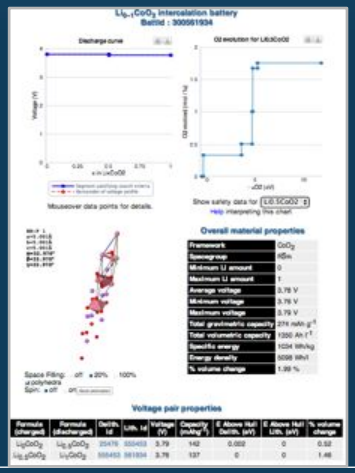
Materials Design Center



User creates novel compounds and submits them for computation.

Materials API

MP & User Apps for Analytics



User creates new web apps or uses existing MP apps to analyze data.

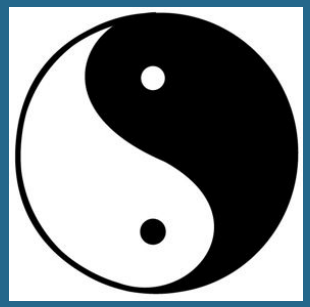
Fully-automated HT computing

MGI – Broadening its impact

Recommendations: Incentives

- Create **incentives** for people to participate in data sharing. Dissemination is a form of impact and needs to be counted as such
- Data and data application tools can have more impact than papers
- Example: MP is thinking of providing contributors with a counter of how much their data/applets has been used

Recommendations: Education



- Complexity of Materials Science requires computation and modeling. Yin and Yang
- Traditional MS&E is not used to having lots of quantitative information
- Continue to educate our young scientists and workforce.
- Engage the educational institutions more in MGI. They will only do this if MGI is reasonably permanent

Recommendations: Test Cases for MGI driven Materials Design

- Create **teams** of experimentalists/theory + MGI efforts to demonstrate how data can be used in real materials design. Do this in different fields
- Feedback from design exercises to MGI effort is valuable. To design materials, field-specific expertise is needed
- E.g. team on thermoelectrics, carbon capture, light weight alloys, etc.

Recommendations: Computational Resources

- Most DOE computing centers are non-receptive to high-throughput computing. Computing grandstanding is put before the science impact
- DOE (and other agencies) need to support these new forms of high-throughput computing. Almost all forms of "leadership" computing facilities are based on 70'ies style problems (single large problems versus tens of thousands of small problems)

EXAMPLE: INCITE program gives away 5 Billion CPU hours in 100Million hour chunks. MP so far has used about 20 Million CPU hours in its lifetime

Recommendations: Data as Infrastructure

- Maybe data becomes part of infrastructure.
- Is it time for a user facility to support data production and management for materials science and engineering ? (i.e. not just computing)

Recommendations: Theory/Science

- MGI will require continued efforts in basic science to keep the pipeline of computational impact going
- Better predictive methods at all length scales. (e.g. better quantum mechanics, better microstructure models, …
- A theory of materials synthesis ? We are foreseeing a future where we can design what is optimal, but can we make it ? A theory of synthesis and processing

Recommendations: Experimental Data

Even Google can not find most experimental data

Widespread Production of Extracellular Superoxide by Heterotrophic Bacteria

Julia M. Diaz,[1]* Colleen M. Hansel,[1,2]†‡ Bettina M. Voelker,[3] Chantal M. Mendes,[1] Peter F. Andeer,[2] Tong Zhang[2]

Science, June 7, 2013

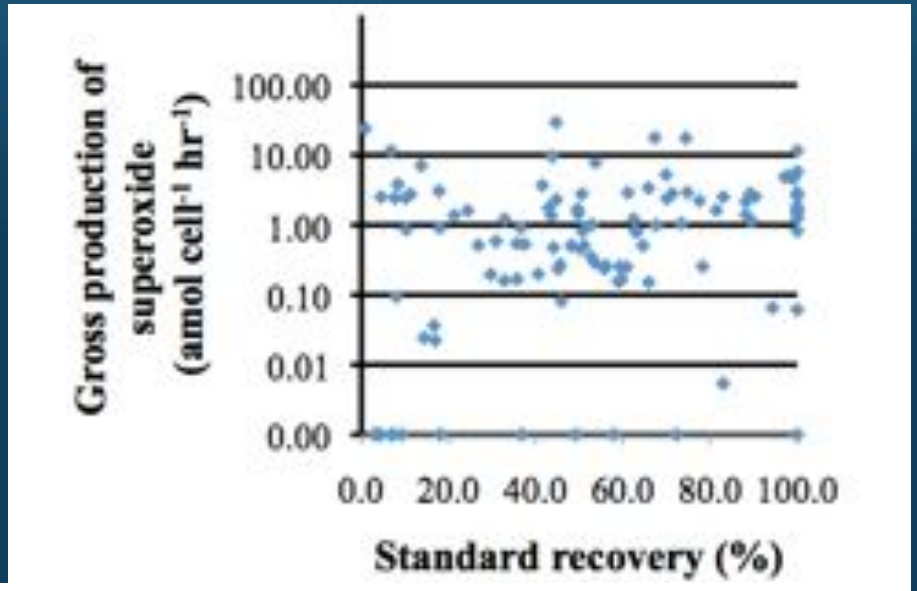


Figure S3: Standard recovery vs. gross production of extracellular superoxide.

Recommendations: Experimental Data

- Experimental data is poorly searchable. Start requiring Data Object Identifiers in research papers.

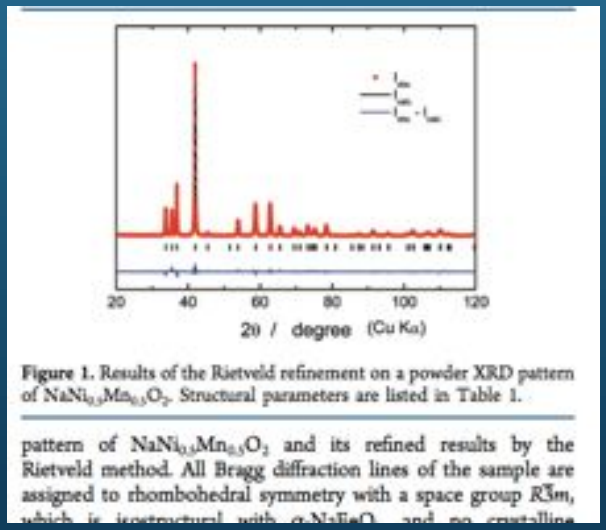
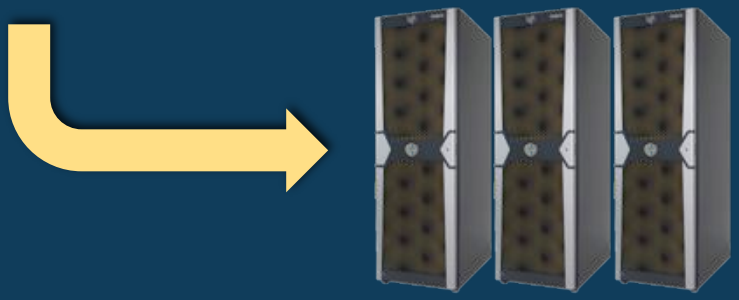


Figure 1. Results of the Rietveld refinement on a powder XRD pattern of $NaNi_{0.5}Mn_{0.5}O_2$. Structural parameters are listed in Table 1.

pattern of $NaNi_{0.5}Mn_{0.5}O_2$ and its refined results by the Rietveld method. All Bragg diffraction lines of the sample are assigned to rhombohedral symmetry with a space group $R\bar{3}m$,

Data Object Identifier XYZ6457

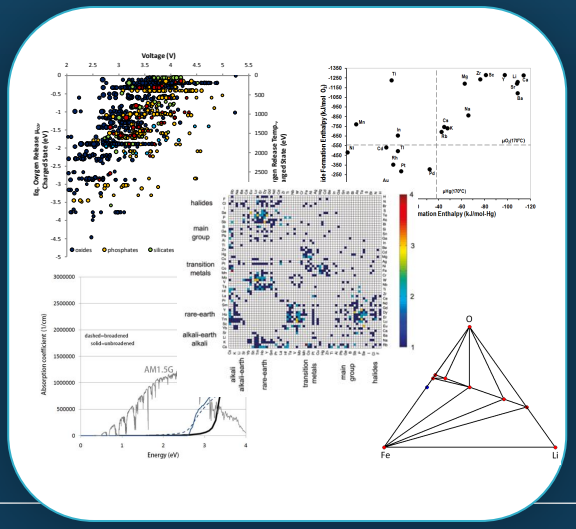


Recommendations: Experimental Data

- Experimental data is poorly searchable. Start requiring Data Object Identifiers in research papers.
- Don’t waste time on trying to “catalogue” and “categorize” data. As long as it is identifiable and findable, the internet will do its work
- At a minimum federally funded research should have data available from published research papers.
- Long term benefits to experimental data identification.

Charting the Materials Genome is Possible

- Within ten years most basic properties of inorganic compounds can be determined computationally
- Materials researchers will have access to all this data. What you do with it will determine your competitive advantage
- Researchers will be able to sit behind a terminal and request the computed properties of modified compounds



… towards a materials genome

Maybe Tony Stark was right …

Feedback and Comments

"Co
fi

Thank you for your attention

Thanks to sponsor:



Anubhav Jain, Shyue Ping Ong, Kristin Persson

and thanks to the team:

Anubhav Jain Alan Dozier Gerbrand Ceder Stefano Curtarolo Daniel Gunter Jeff Grossman Dane Morgan Rafae Fink

Shyue Ping Ong David Skinner Stefan Adams Mark Asta Anthony Gamst Wei Chen Qimin Yan Bharat Mehdsani Geoffroy Hautier

Shreyas Cholia Maciej Haranszyk Will Richards Jeff Neaton Maarten De Jong Sai Jayaram n Richards

reset

"I am so incredibly happy an effort like this exists now... I have been lamenting for years that despite the importance of materials we have remained relatively unaided by the information age. Please please don't stop growing!" Cymbet