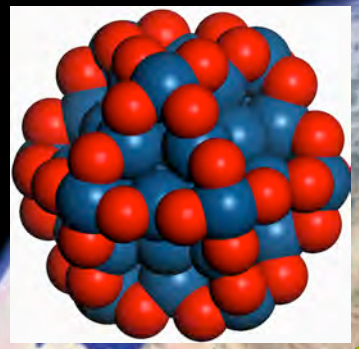
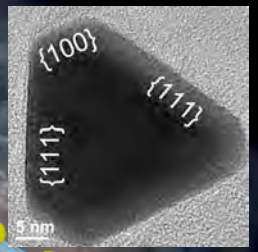
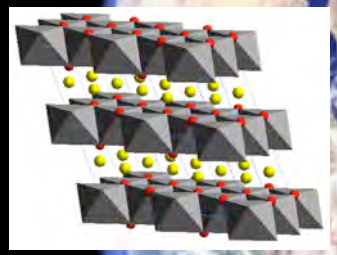
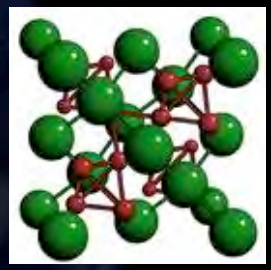
Designing Materials for our Energy Future









MIT-NSF workshop Sept 8, 2008

G. Ceder, Department of Materials Science and Engineering

Chris Fischer, Geoffroy Hautier, Charles Moore, Anubhav Jain, Kristin Persson

The group

From basic science, to applications …

in the fastest possible way

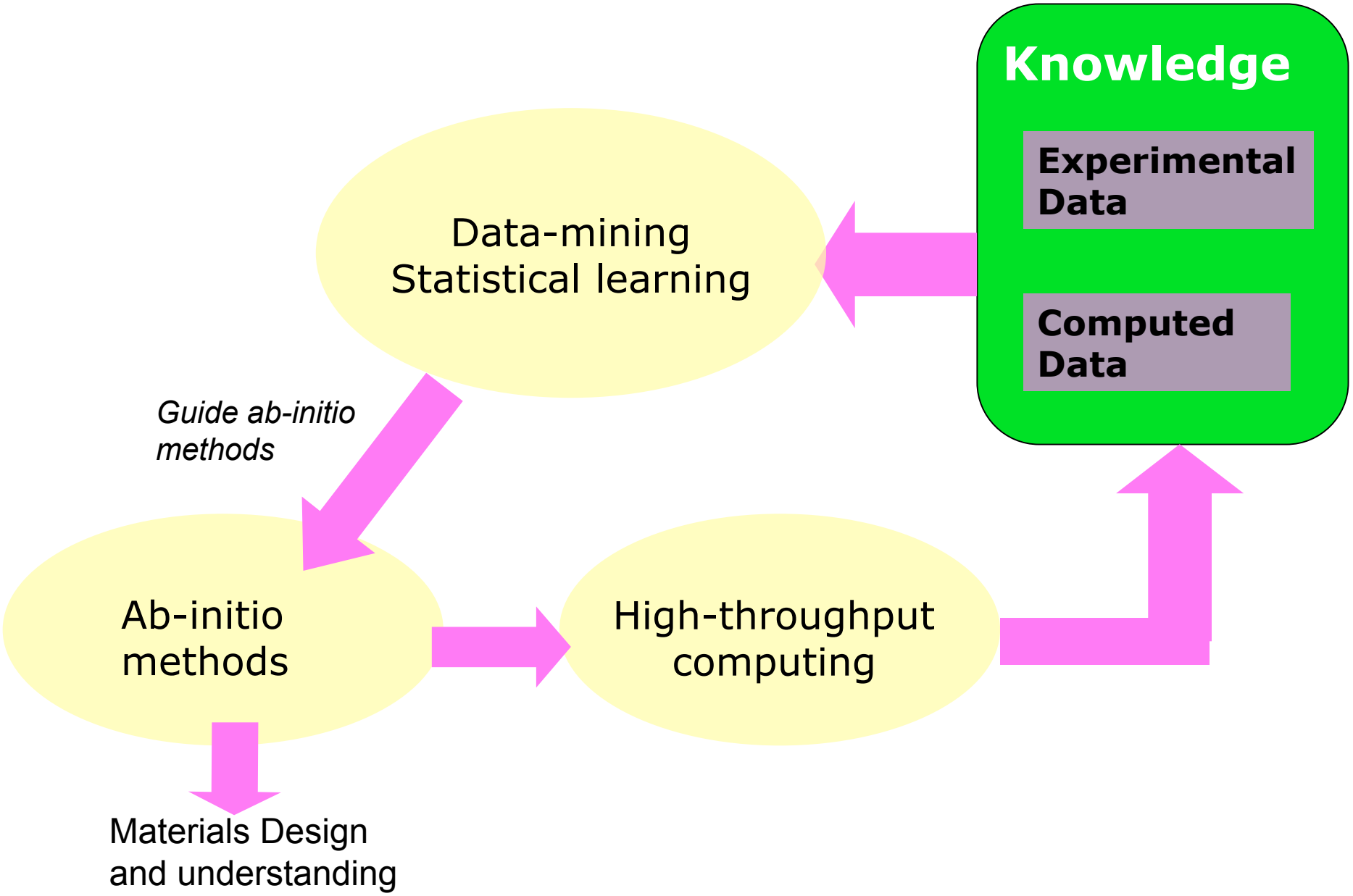
Fundamental

- Phase stability of materials
- Phase transitions and transport
- Electronic Structure of Metal Oxides
- Length scale bridging dynamics
- High-Throughput computation and Data Mining

Applied Problems in Energy

- Electrode materials for Li batteries
- Protons in oxides
- Nanoparticle catalysts
- Thermoelectrics
- Hydrogen Storage
- Nano particle stability and self-assembly

Data-mining and High-Throughput *ab-initio* computing



Materials research has traditionally been slow

Computation is the most scalable research tool

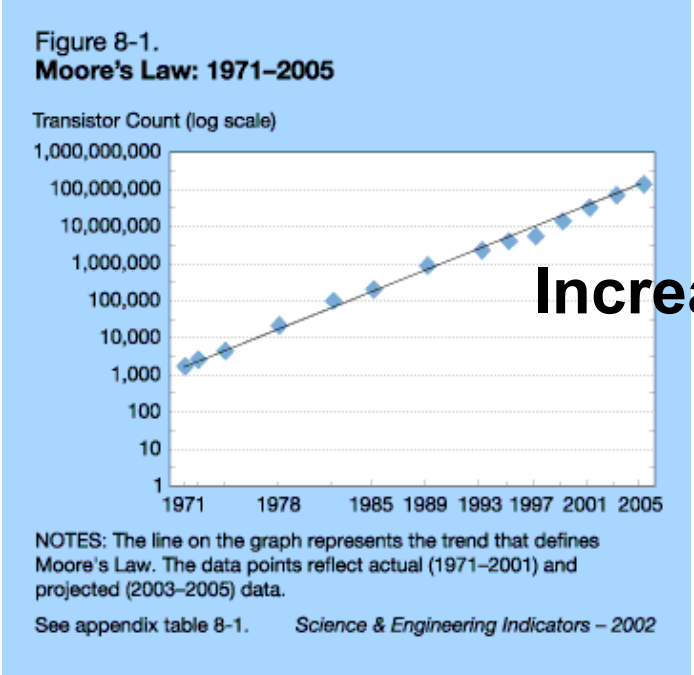
Radical acceleration can only come from high-throughput predictive computational approaches

All components are in place

- **Equations that describe matter can be solved accurately and stable**

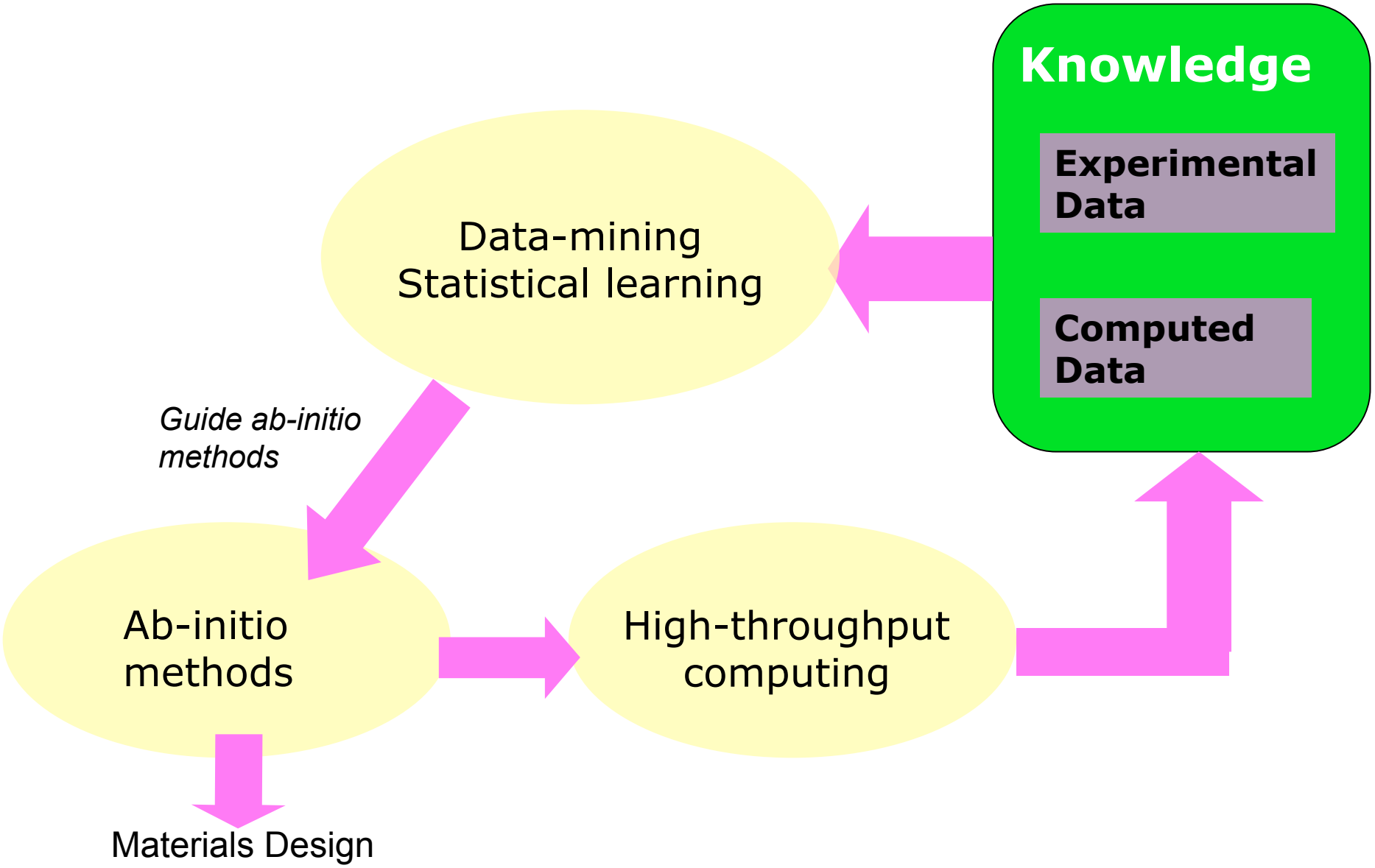
$$H = E_{nuclei}\left(\{\boldsymbol{R}_I\}\right) - \sum_{i=1}^{N_e} \nabla_i^2 + V_{nuclei}(\boldsymbol{r}_i) + \frac{1}{2}\sum_{i \neq j}^{N_e} \frac{1}{|\boldsymbol{r}_i - \boldsymbol{r}_j|}$$

- **Increase in computing/dollar is unparalleled in any other field**



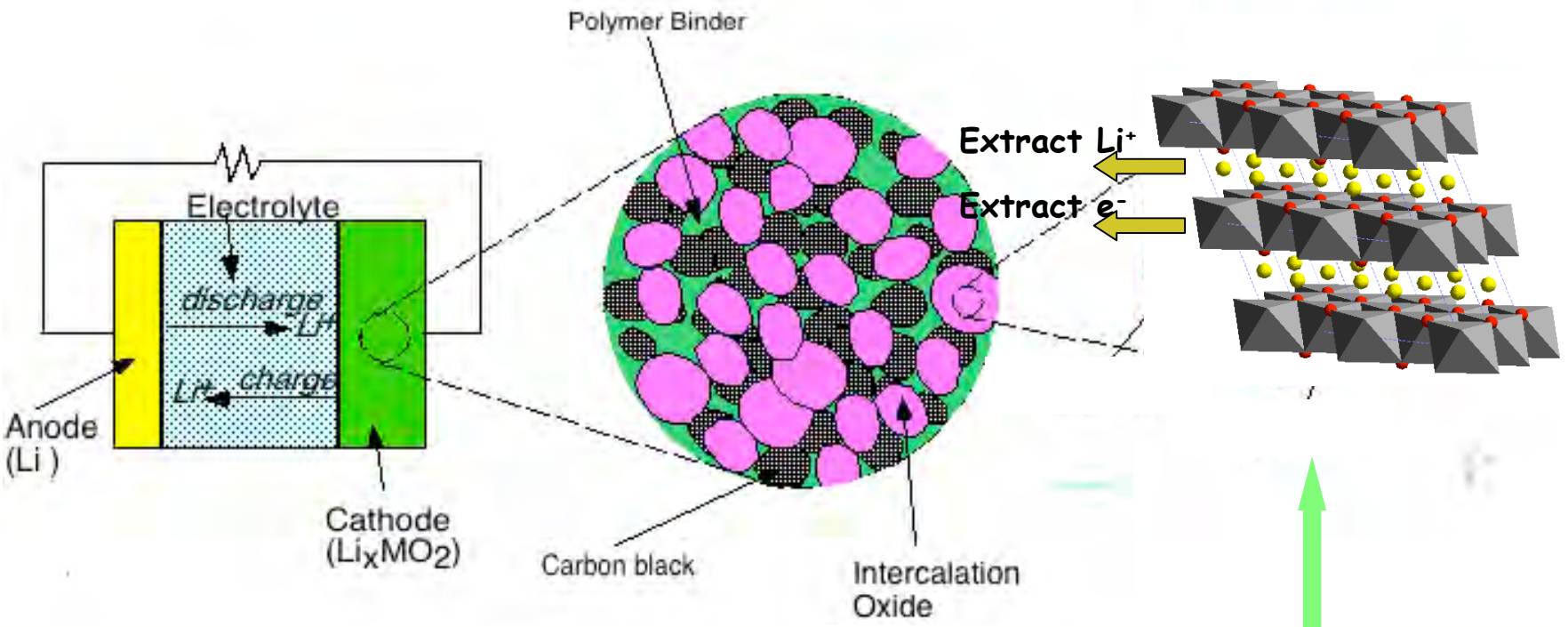
Increase in $\frac{\text{Performance}}{\text{Price}}$ **of 10 million !**

Data-mining and High-Throughput *ab-initio* computing



An example of “ab initio” computational materials design

Example: Design of high power Li battery materials



Li + Metal-Oxide -> Li-Metal-Oxide

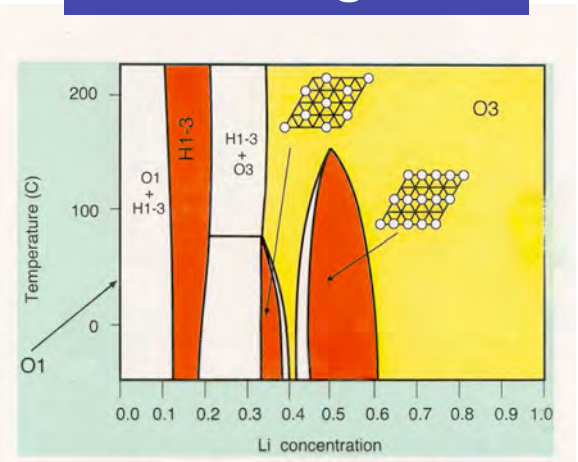
Cathode needs to host and exchange large amounts of Li^+ and electrons at high rate and remain stable

Many properties of the electrode materials can now be predicted with ab initio methods before the material is ever made

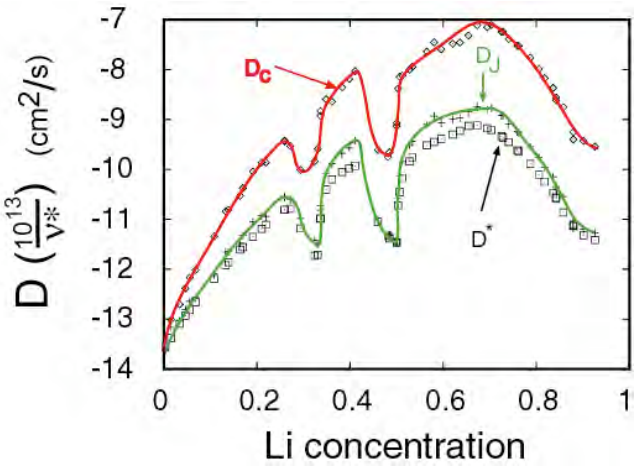
Voltage → energy density

	Exp	Calcul.
$LiCoO_2$	4.2	4.25
$LiFePO_4$	3.5	3.45
$LiNiO_2$	3.75	3.71

lifetime ← **Phase diagram**

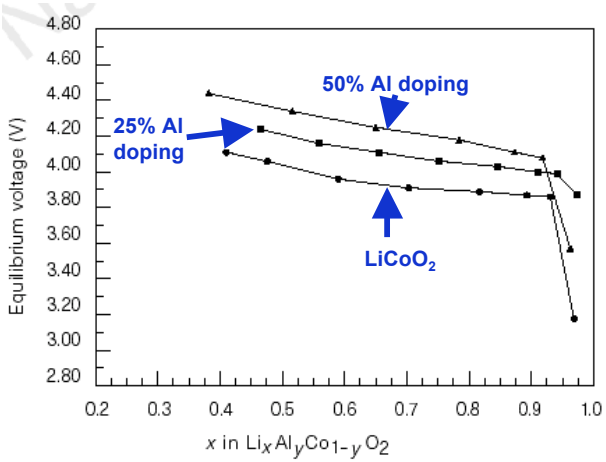


Li diffusivity → power



Track record of ab-initio predictions that were confirmed later with experiment

Al-doping effect in $LiCoO_2$

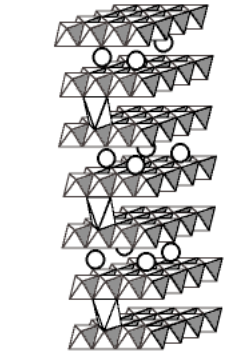


Predict Interesting materials

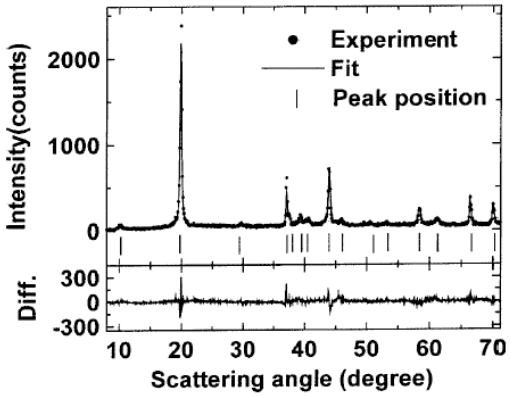
$LiNiPO_4$: 5.1 V

Predicted Spring 2004; Confirmed by experiment [Wolfenstine and Allen, J. Power Sources, 142, 389 2005]

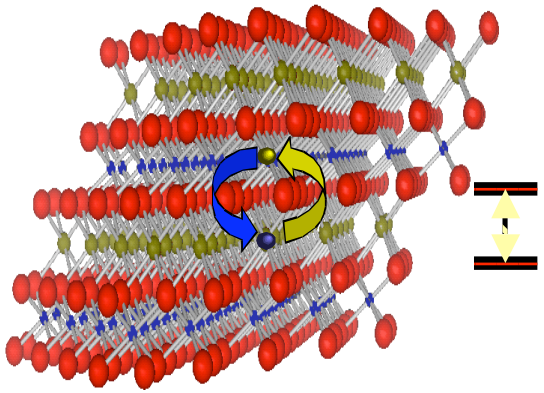
H1-3 phase of $LiCoO_2$ at high charge



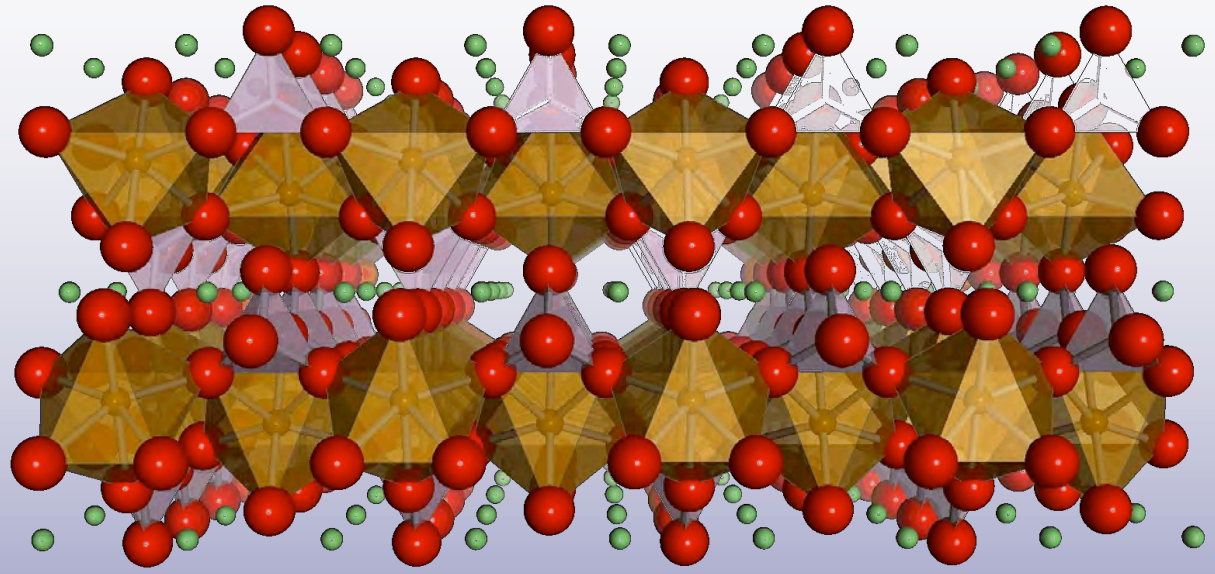
H1-3
(stage II compound)



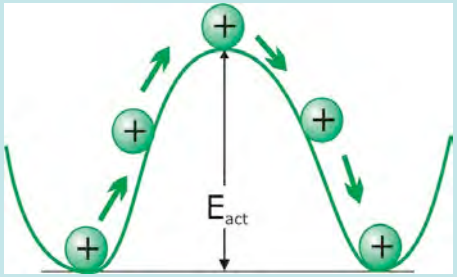
High rate $Li(Ni_{0.5}Mn_{0.5})O_2$



$LiFePO_4$ designed for extreme rate behavior



Li migration barrier calculation

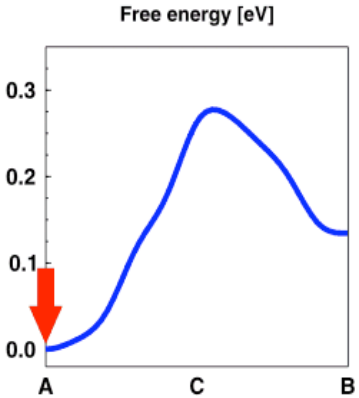


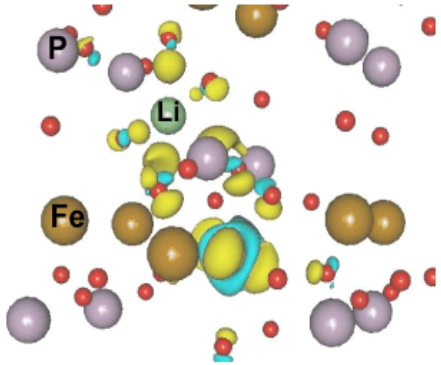
Along the a-axis: $E_a > 1$ *eV*

Along the c-axis: $E_a > 1$ *eV*

Along the b-axis: E_a *is low*

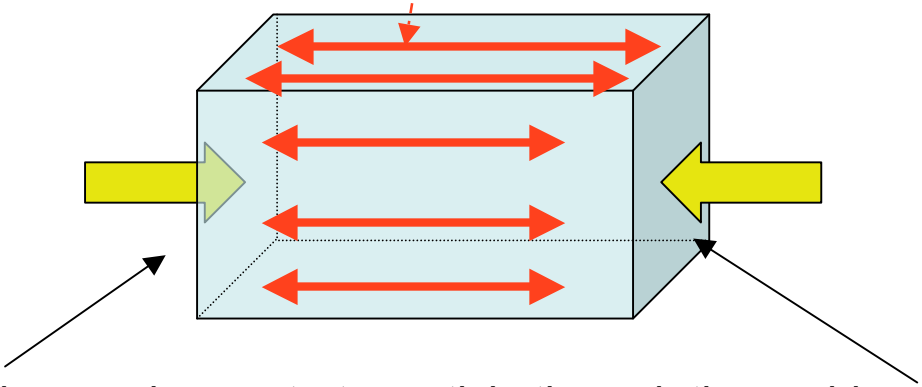
Electron migration barrier calculation





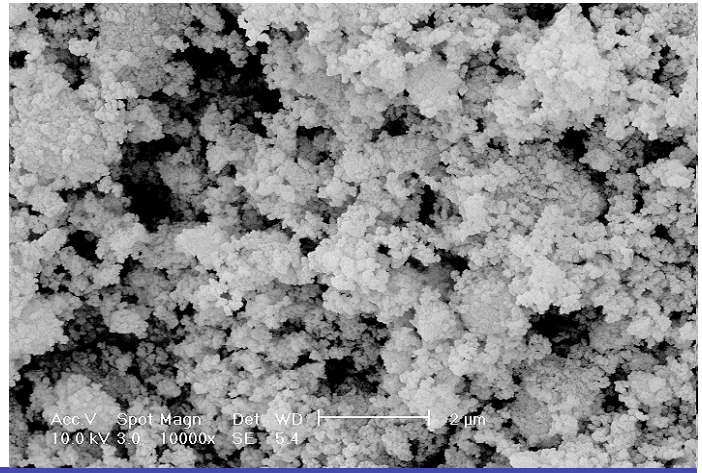
$LiFePO_4$ with optimized surface structure

On other surfaces Li has to move over surface to the sides

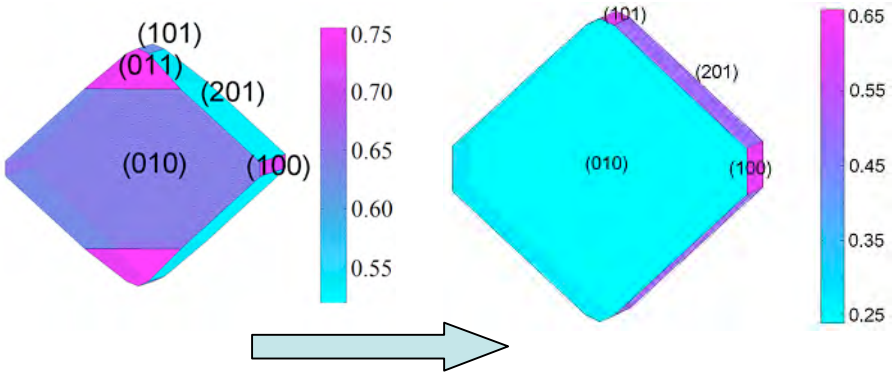


Li can only penetrate particle through these sides

Make nanoparticles

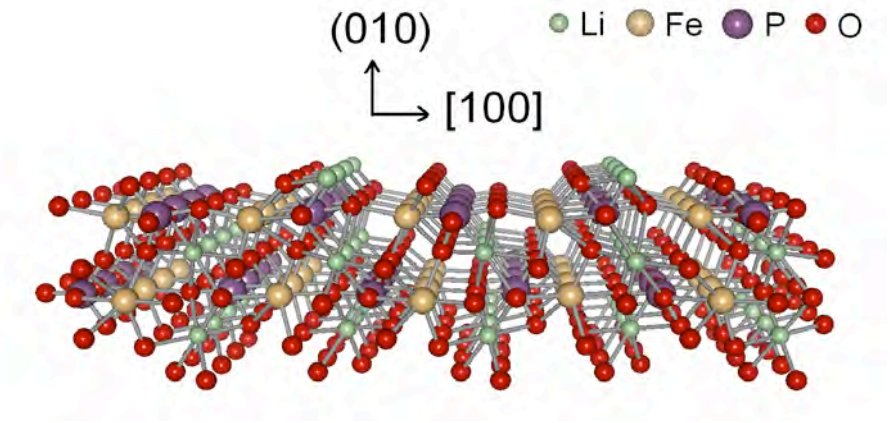


Optimize particle shape (surface energies !)

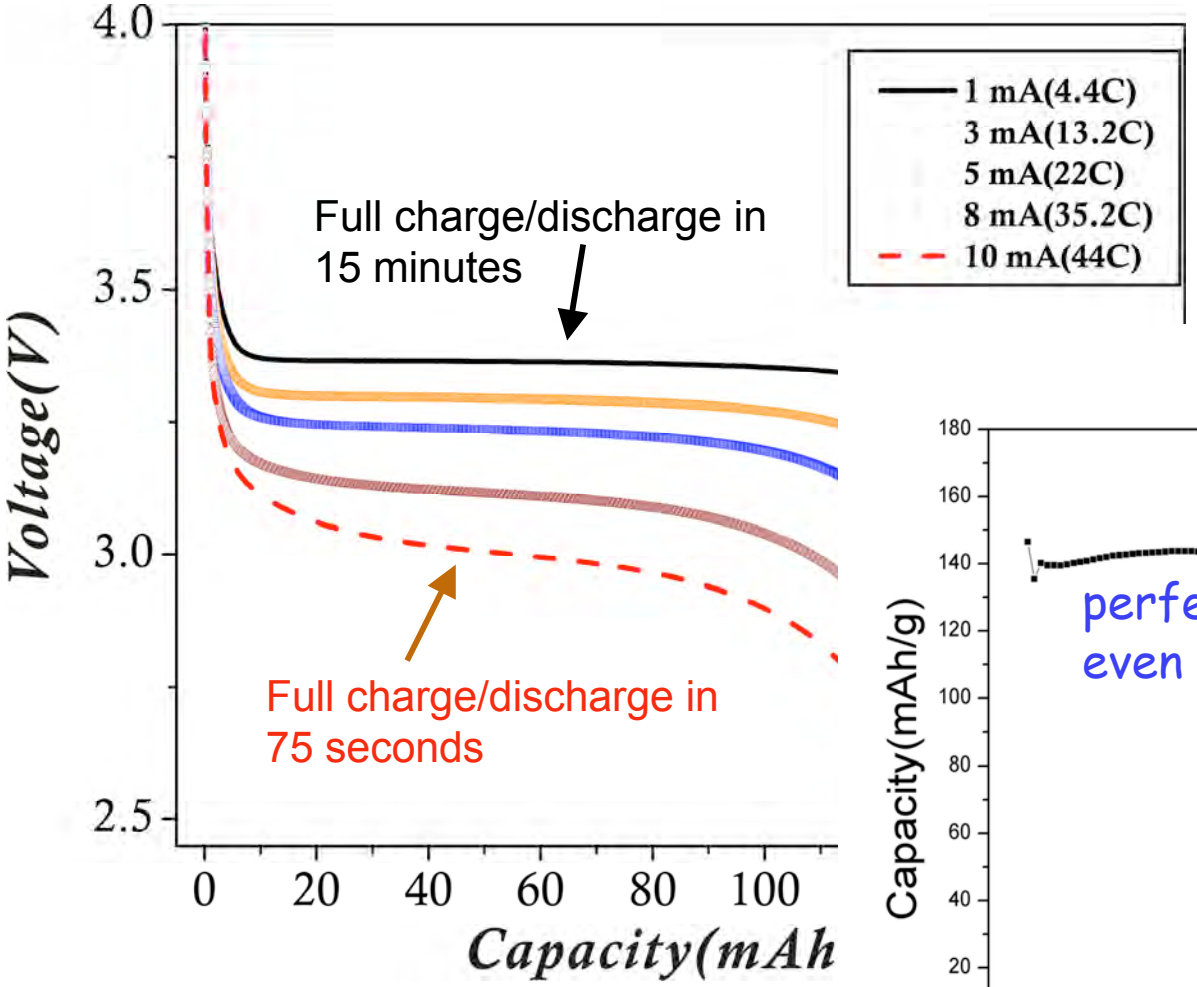


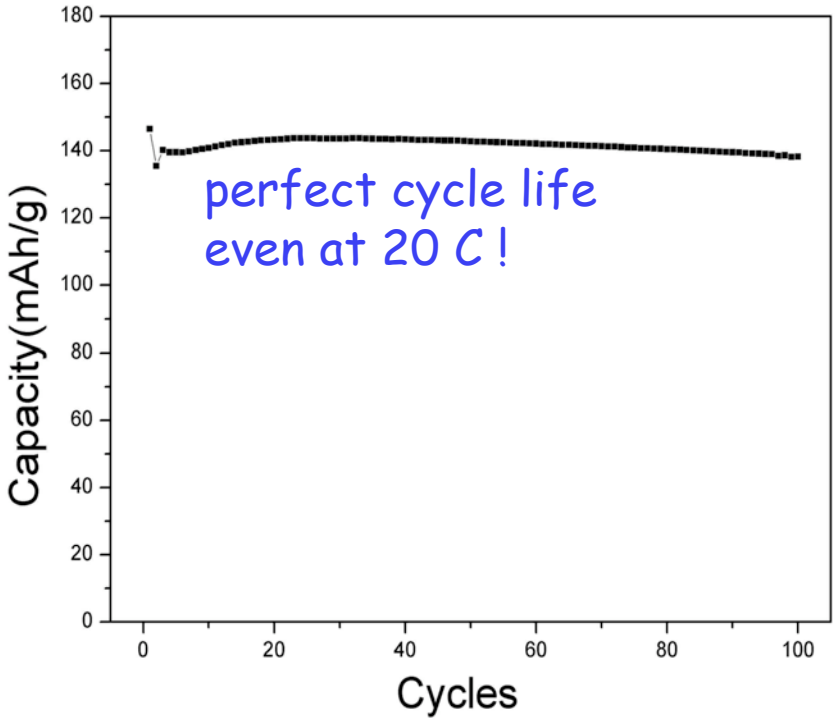
Platelet shape

Modify surface chemistry so that Li ions can move rapidly over surface

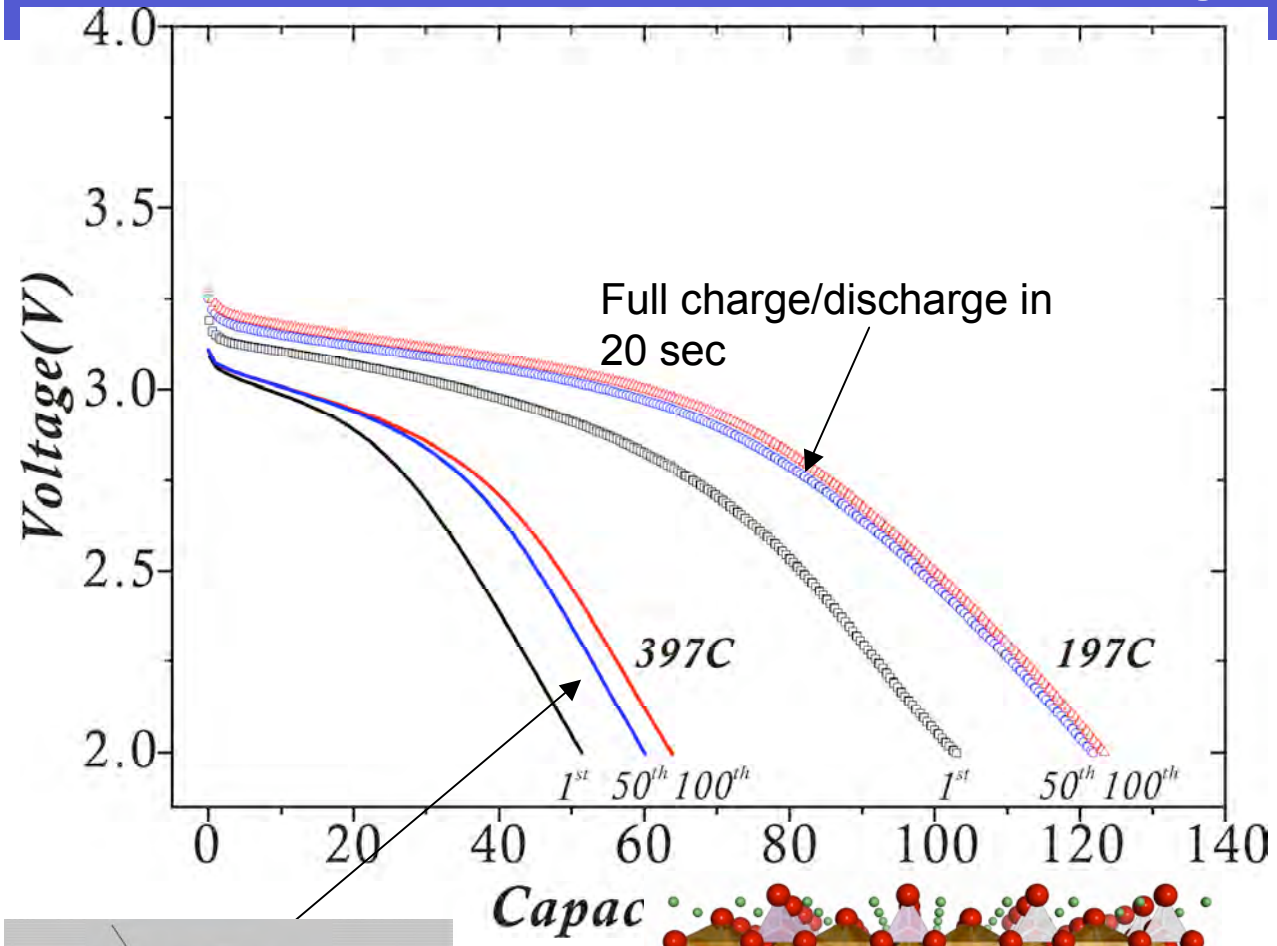


Very high rate material developed





With some electrode modification can obtain highest rate ever observed in a battery material



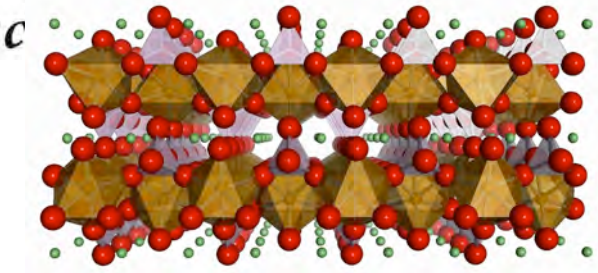
400 C is full battery charge/discharge in **9 seconds**

Power density:

175 kW/liter
90kW/kg

arge in

+



=

If you can compute it -> automate it -> scale it

Components to automation

1. Automated run-time environment: automate everything from setting up the job, queing it, checking it, and import into database
2. The database
3. Mining the data

Created a fully automated environment to do thousands of quantum mechanical calculations

Automated the quantum mechanical energy compu structu

Automatic

Input file tr

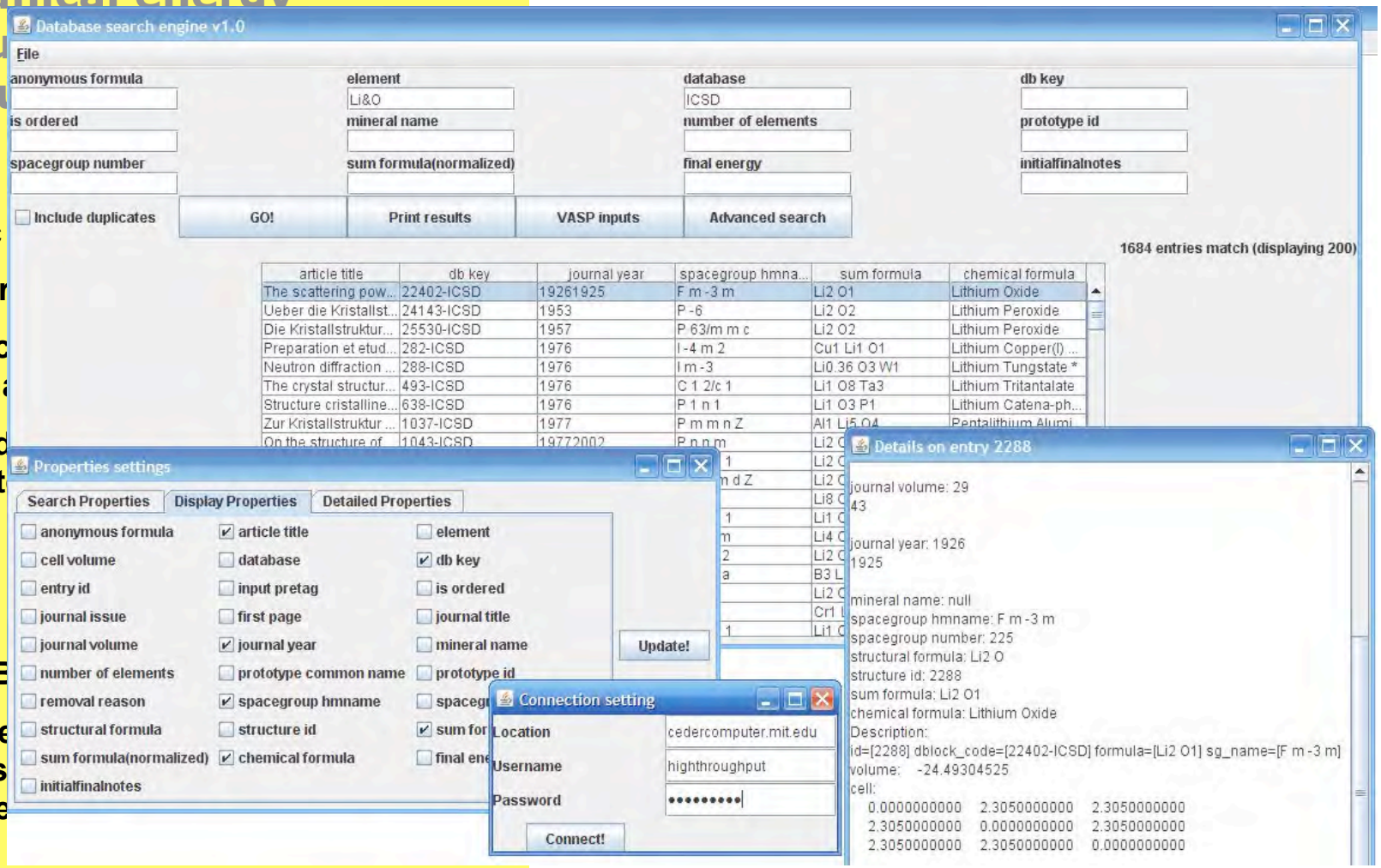
Error Chec run autom

Automated implement

Run on E

Develope numbers parallel e

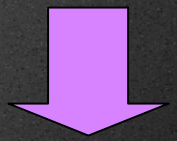
We can now ask “meta” questions



The database

- Database is a *relational database (PostgreSQL)*
- Information stored in tables, data represented as rows

Elements
+Symbol: char[2] (PK)
+Name: Varchar (40)
+Z: int
+Mass: double precision
+MinValence: int
+MaxValence: int



symbol	name	z	mass	minvalence	maxvalence
AA	Generic cation A	0	0	0	0
Ac	Actinium	89	227	3	3
Ag	Silver	47	107.8682	1	3
Al	Aluminum	13	26.9815386	3	3
Am	Americium	95	243	2	6
Ar	Argon	18	39.948	0	0
As	Arsenic	33	74.9216	-3	5
At	Astatine	85	210	7	7
Au	Gold	79	196.966569	1	5
B	Boron	5	10.811	-3	3

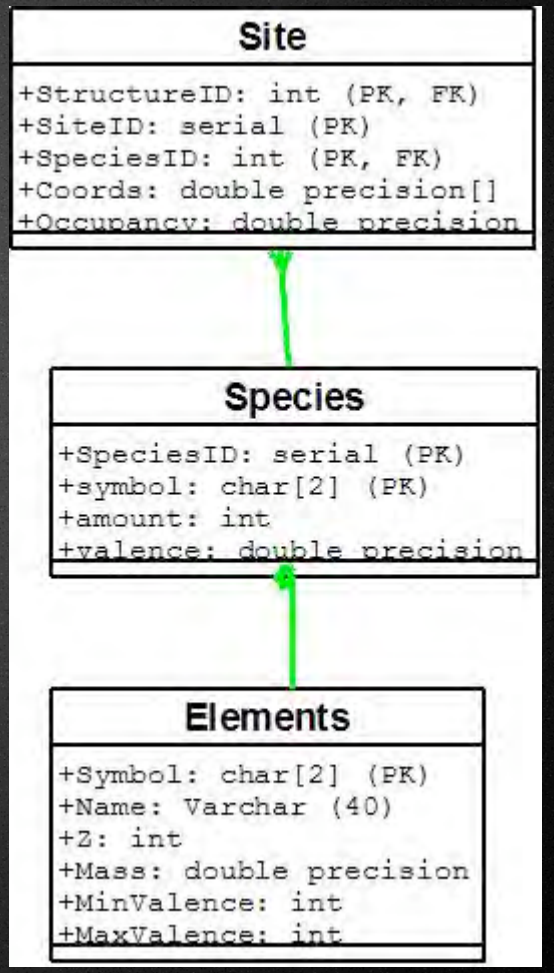
putting the ’relation’ in relational

- Links between tables are relations
- Prevent data redundancy

structureid	siteid	speciesid	coords	occupancy
1151	3007	155	[0:2]={0.25,0.25,0.25}	1
1312	3487	155	[0:2]={0,0,0}	1
1313	3489	155	[0:2]={0,0,0}	1
1314	3491	155	[0:2]={0,0,0}	1
1315	3493	155	[0:2]={0,0,0}	1
1767	5072	155	[0:2]={0,0,0}	1
1973	5880	155	[0:2]={0,0,0}	1
2018	6165	155	[0:2]={0.13386,0.63678,0.33133}	1

speciesid	symbol	amount	valence
155	Na	1	1

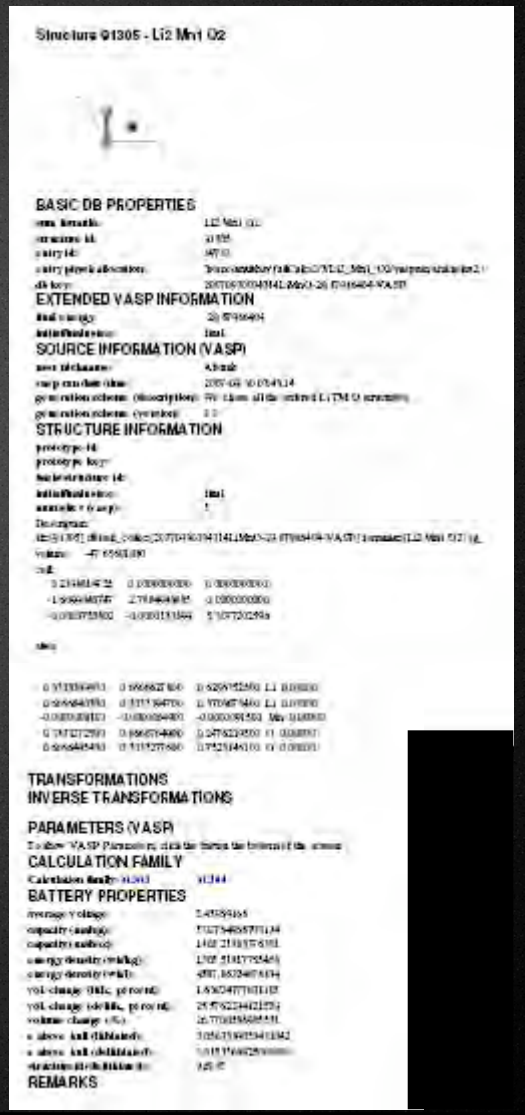
symbol	name	z	mass	minvalence	maxvalence
Na	Sodium	11	22.98976928	1	1



Complete Voltage Analysis of most known compounds

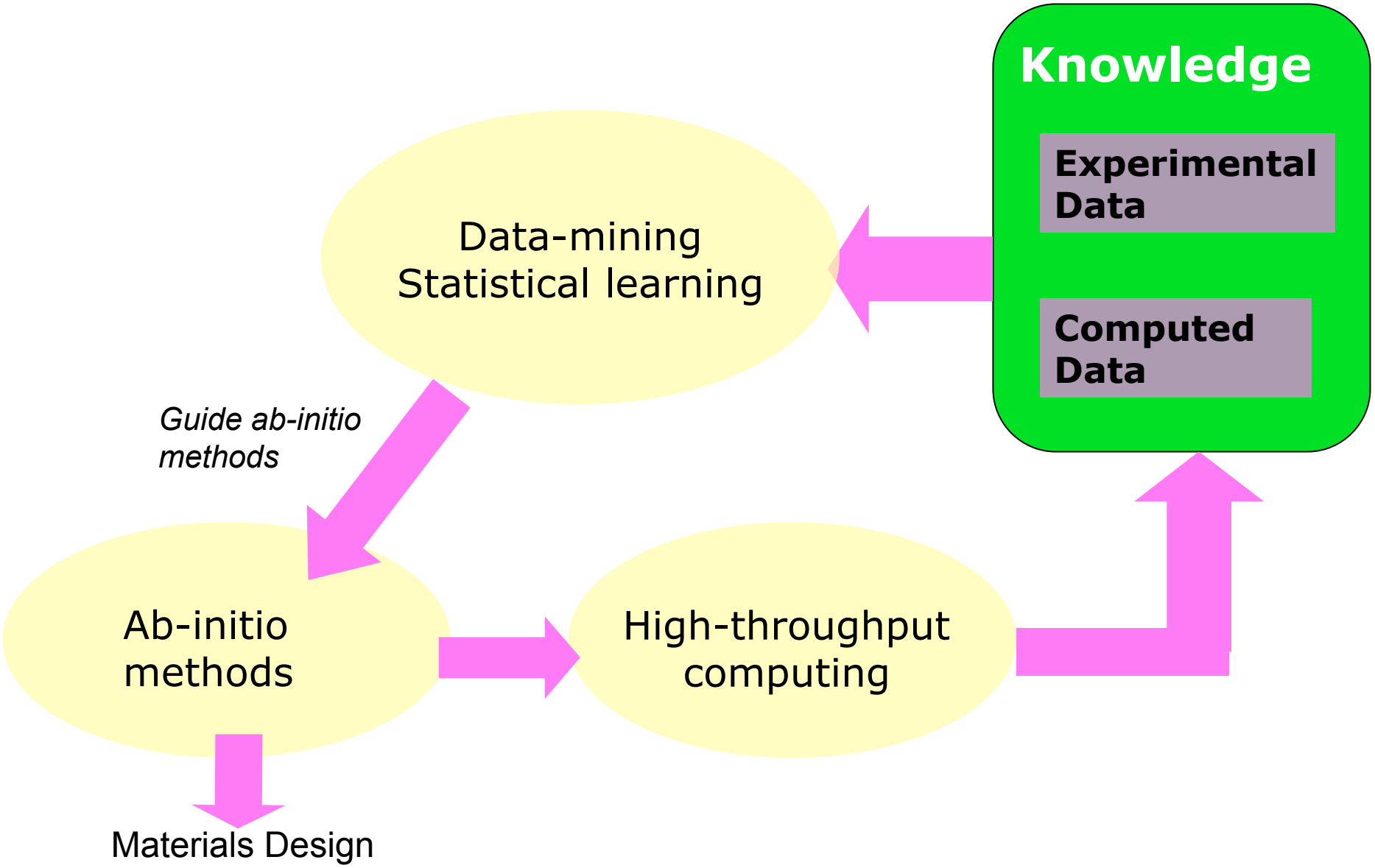
- ~3500 voltage data points

Picture removed



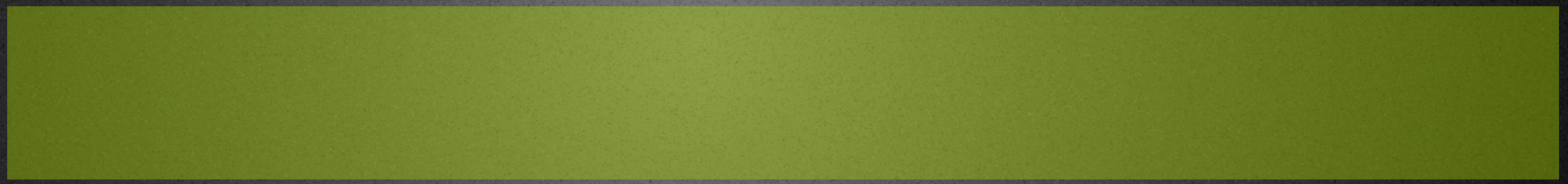
***

Data-mining and High-Throughput *ab-initio* computing

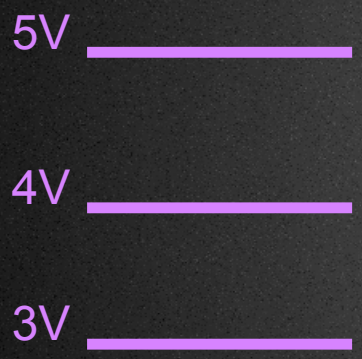


Data mining – Li insertion Cathodes

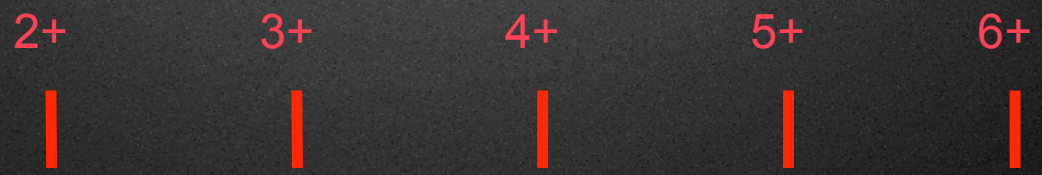
Picture removed



Data mining - Li insertion Cathodes



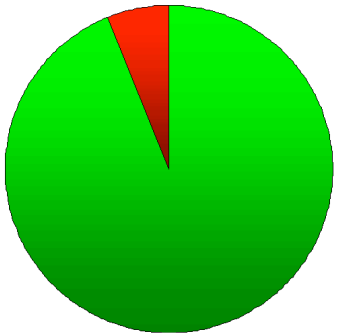
Picture removed



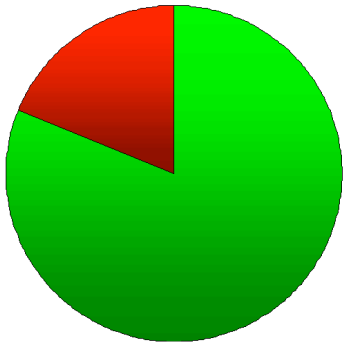
Substantial added value of data

- We have run and converged ~30,000 Ab-initio calculations
- How much of unique ICSD calculated?

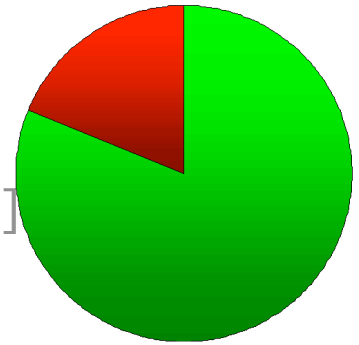
elements 94% (255)



binaries 81% (4949)



ternaries 80% (10247) [still in progress]

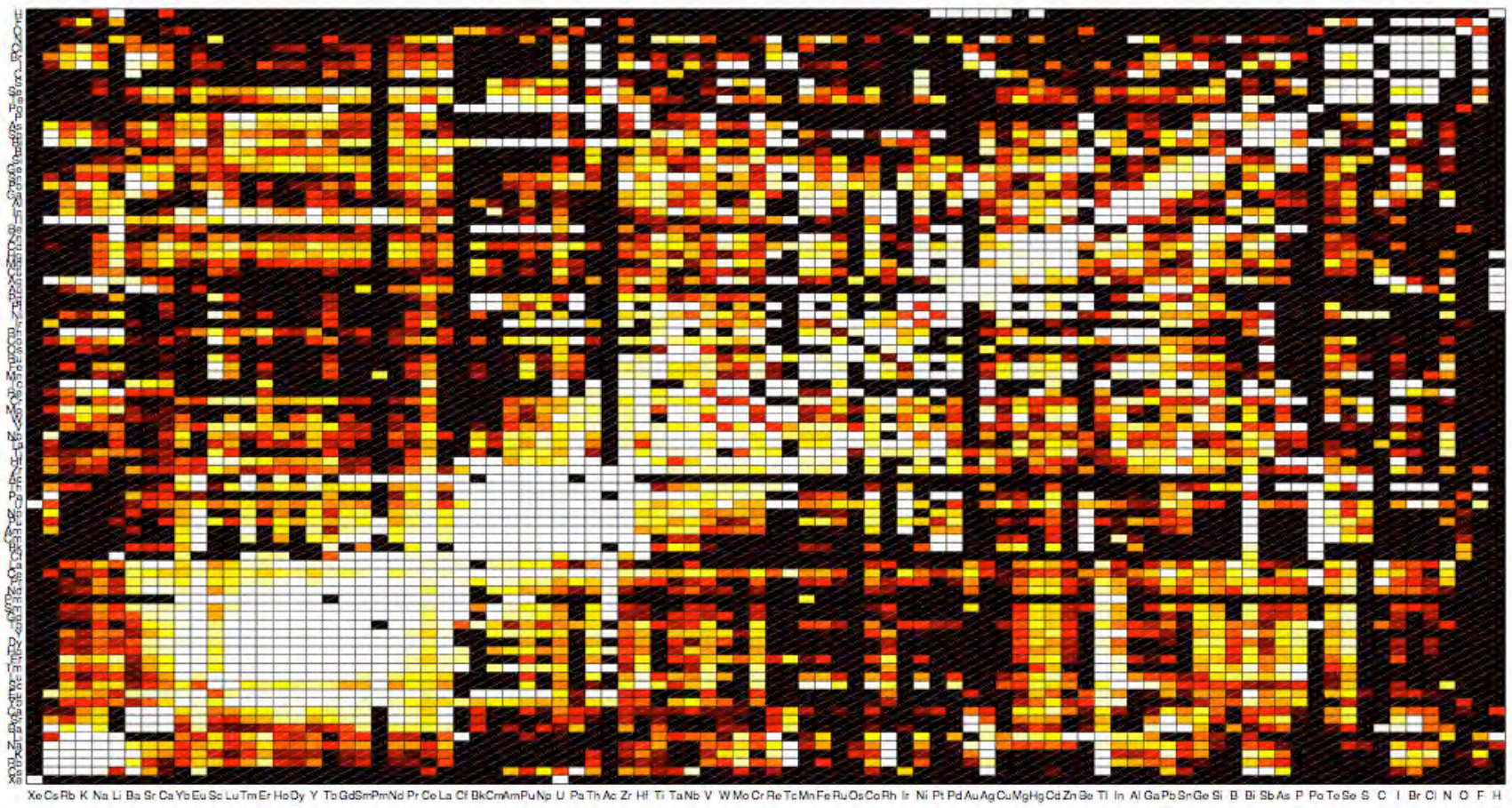


Substantial added value of data

- Can now evaluate phase stability of hypothetical new compounds against separation in binaries, ternaries, etc.
- Identified several hundred potentially new ternary oxide compounds
- Working with LBNL to identify radiation detector materials from these

Learn chemistry and physics ...

which elements can substitute for each other ?



very powerful for materials design

An example of datamining experimental data

funded by NSF-DMR

Example of data mining experimental data: The crystal structure prediction problem

- 1988 Maddox (Nature) described the inability to predict crystal structure as "scandalous" – Not much has changed
- As first principles methods for properties become better, and computers faster, structure problem becomes more limiting to materials design

Without knowing where the atoms are, can't compute much

funded by NSF-DMR and DOE-BES

Ab-initio approach
Energy + Optimization

Smart searching ?

Ground states only

Energy model

LDA/GGA potentials, …

Direct Optimization of E({R})

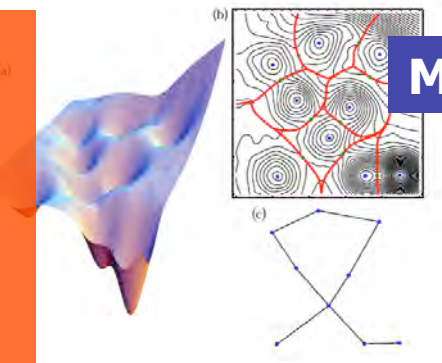
Molecular Dynamics

Simulated Annealing

Genetic Algorithms (e.g. Abraham and Probert, PRB 2006; Oganov & Glass, J. Chem Phys 2006)

Trial and Error

…

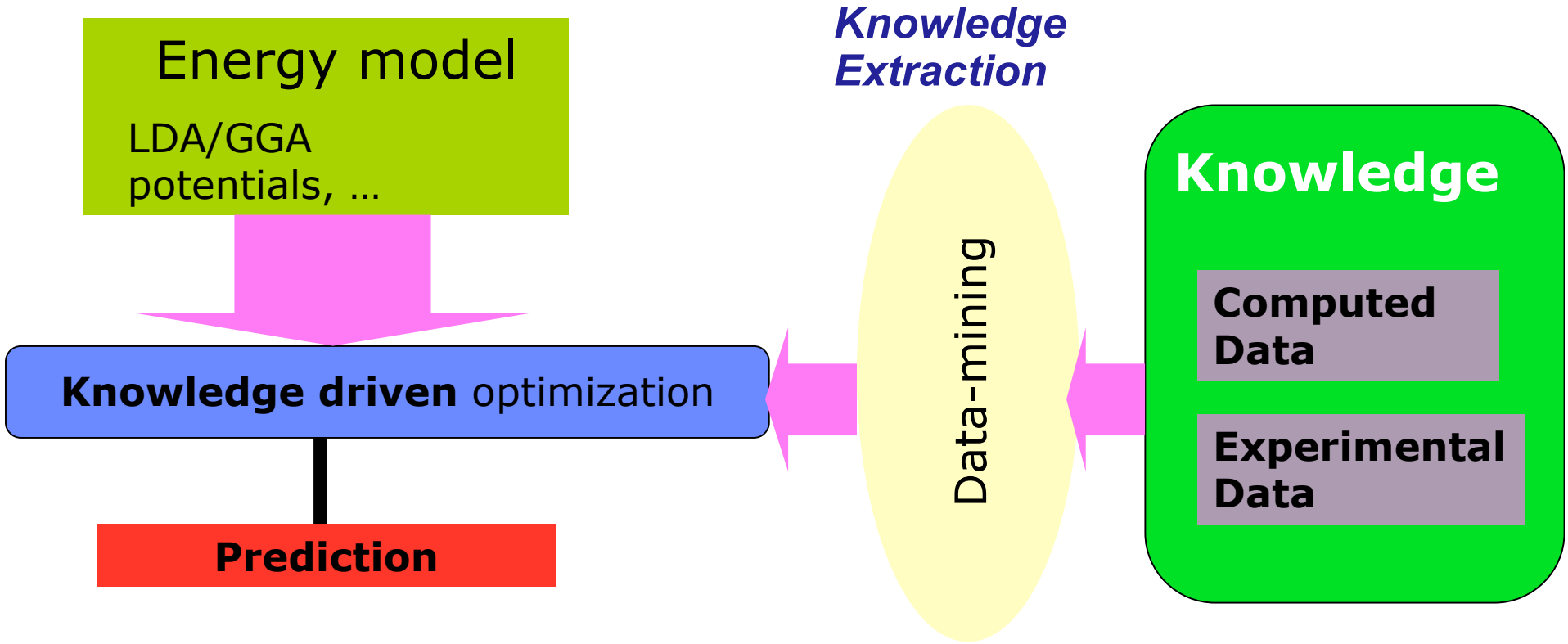


- *Minima in energy space are not random. They result from underlying physical and chemical principles*
- *Can we learn about underlying physics in nature without making it explicit ? Do so in mathematically rigorous way*

Mathematical "intuition"

- Statistical learning
- Data mining
- Artificial intelligence

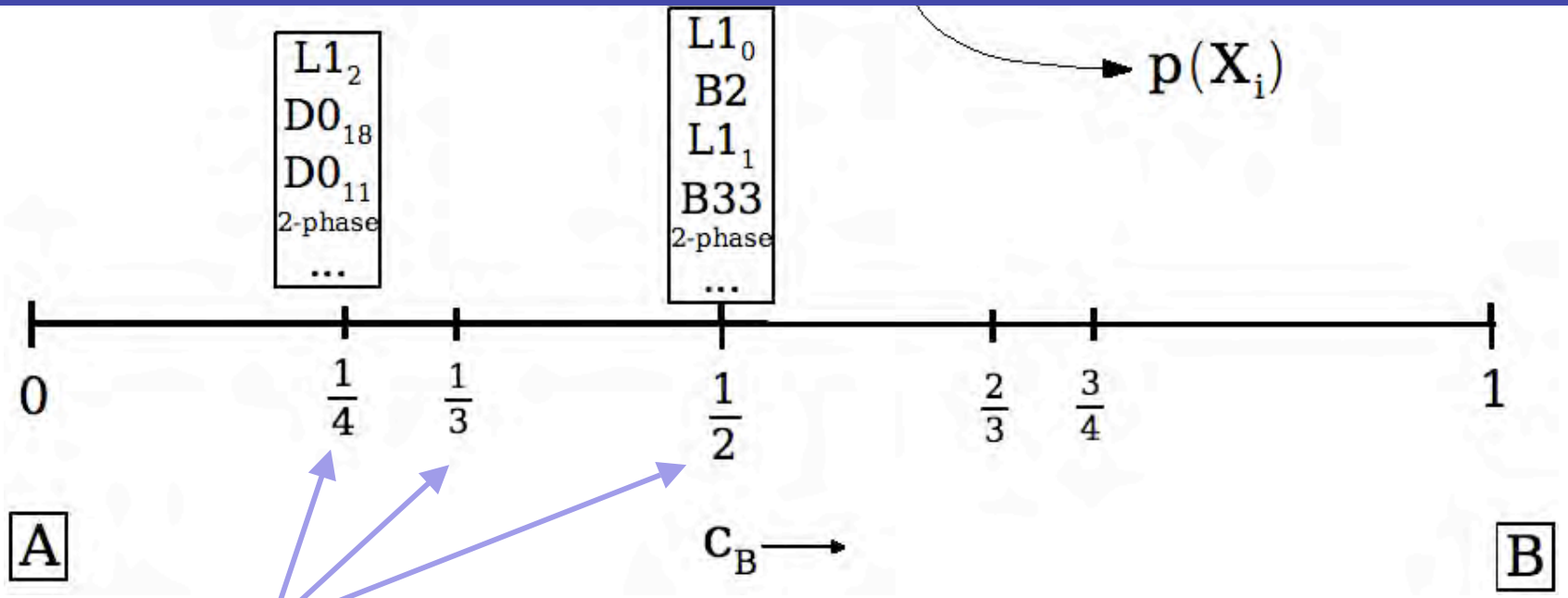
Concept Knowledge driven search



Fischer, C., et al., Predicting Crystal Structure by merging data mining with Quantum Mechanics. Nature Materials, 2006. 5: p. 641.

Curtarolo S, Morgan D, Persson K, Ceder G. Data Mining of Quantum Mechanical Calculations. Phys. Rev. Lett 2003;91:135503 1.

Probability model describes how likely prototypes occur together in system



at each composition *i* variable
x_i = prototype at *i*

x_1 = element A
x_2 = element B

$\mathbf{X} = (x_1, x_2, x_3, \ldots , x_n)$

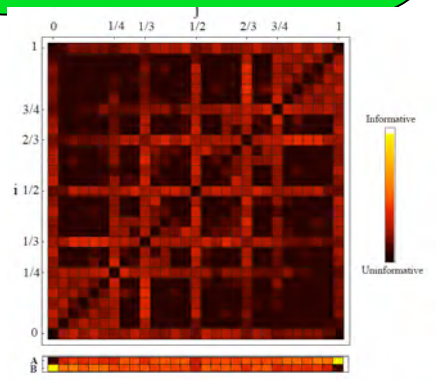
ground state set of the alloy

$P(\mathbf{X})$ → = knowledge of all Nature

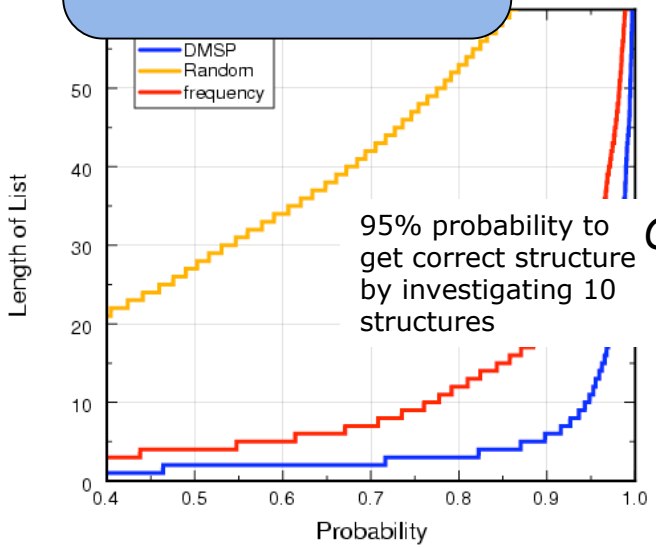
A first success: Data Mining to Predict Crystal Structure

DATABASE:

ALL KNOWN CRYSTAL STRUCTURE ASSIGNMENTS

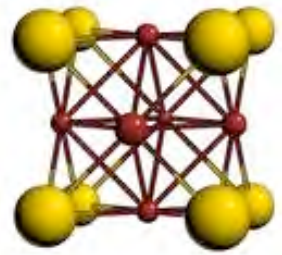


P(X) approximation from exp data



List of candidate structures

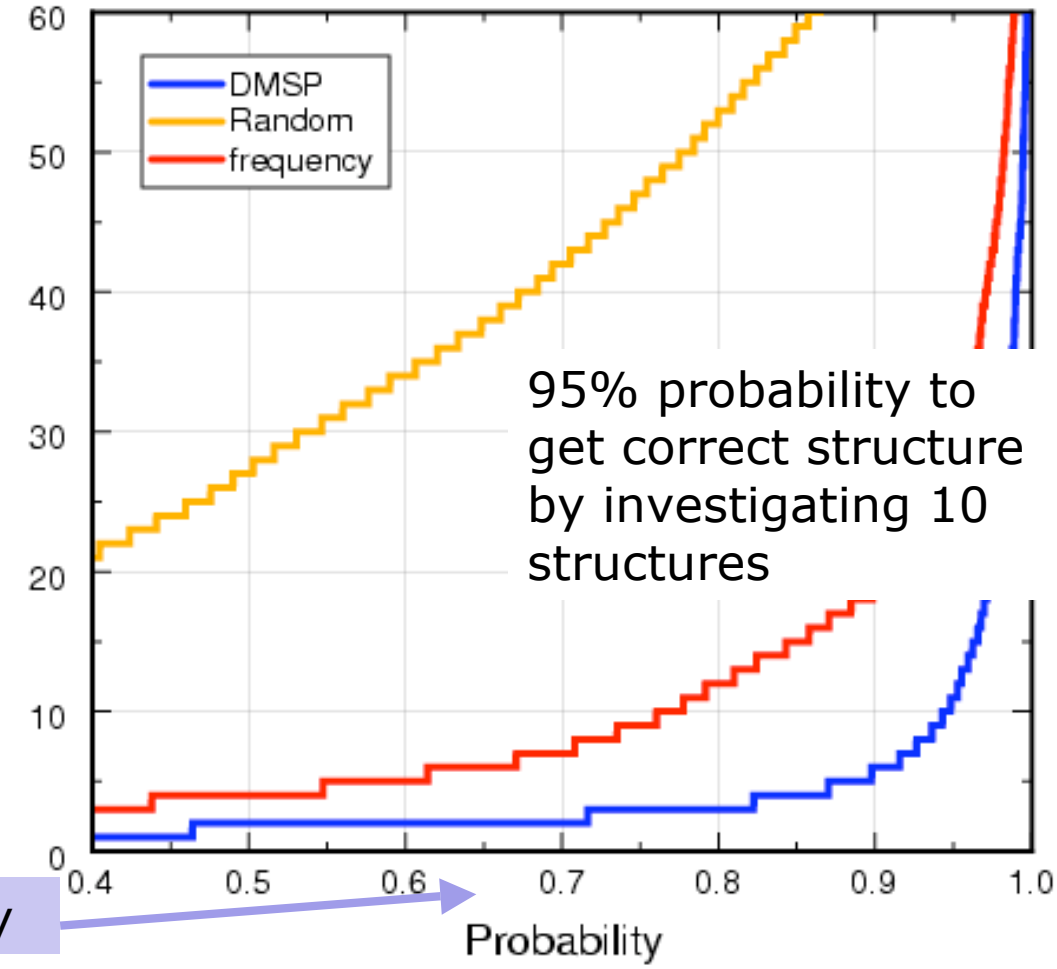
Correct Structure



Probabilistic Model Prediction Capability

Cross-validation prediction on > 5000 structures

how many structures need to keep on list to have correct ground state …



… with a given probability

Fischer, C., Tibetts, K., Morgan, D. & G, C. Predicting Crystal Structure by merging data mining with Quantum Mechanics. **Nature Materials**, **5**, 641(2006).

Note: Solution came from changing the question, not throwing more effort at conventional way of thinking of the problem

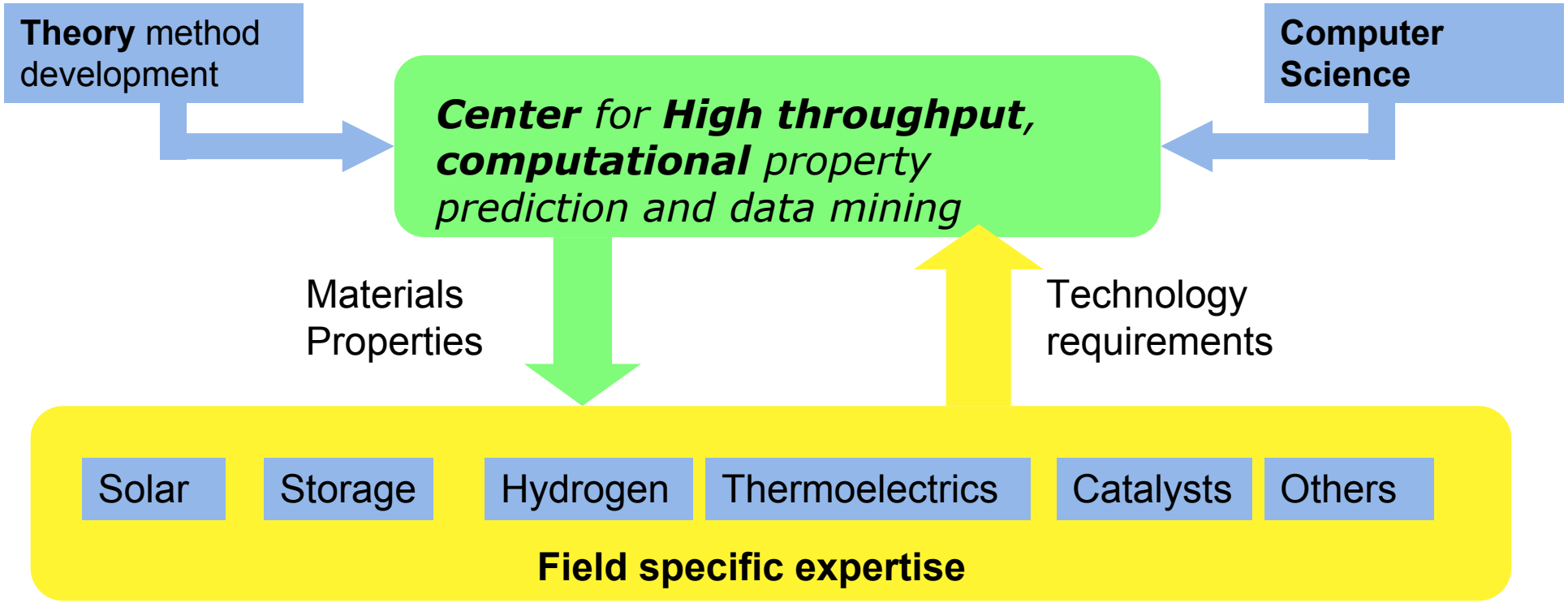
Synergy between applications when data is preserved

recently completed search for potential materials for Hg adsorption from high-temperature coal gasification streams

trying to build capability to include calculated properties relevant to photo-voltaics, photo-synthesis, and thermoelectrics

Vision: Materials Genome Concept

"To rapidly develop data on all materials in nature, relevant to a problem, so that informed and effective choices can be made in research and development programs"



Do for materials development what genome sequencing has done for biology

Issues

Experimental data

- Data not standardized or without proper specification (e.g. many different copper materials when it comes to conductivity)
- No central data reporting
- Most databases proprietary -> often inhibits data mining
- Can we follow the biology model ? NIH demands entry of protein structures in database ?

Computed data

- not all properties can be computed
- accuracy issues
- pollution from poorly converged calculations
- large computing centers sponsored by DOE/NSF/DoD prefer large calculations - not many small ones.

Thank you