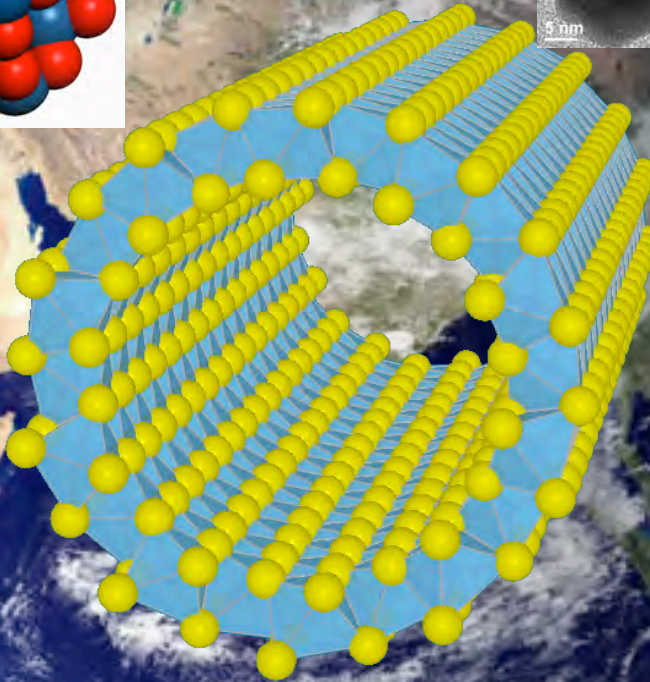
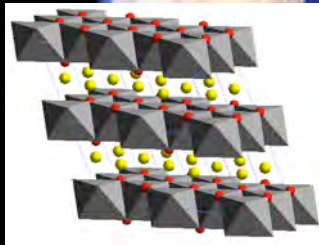
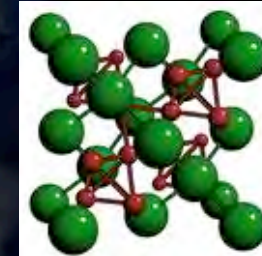
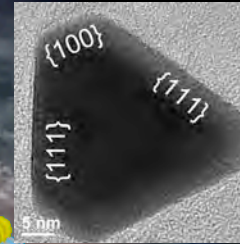
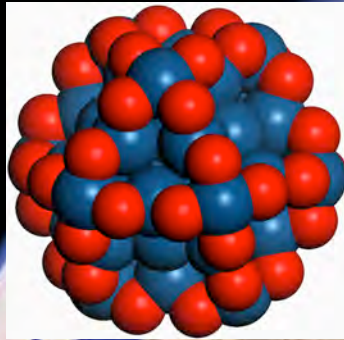


# Designing Materials for our Energy Future



MIT-NSF workshop Sept 8, 2008

G. Ceder, Department of Materials Science and Engineering

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

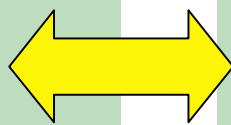
# The group

*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

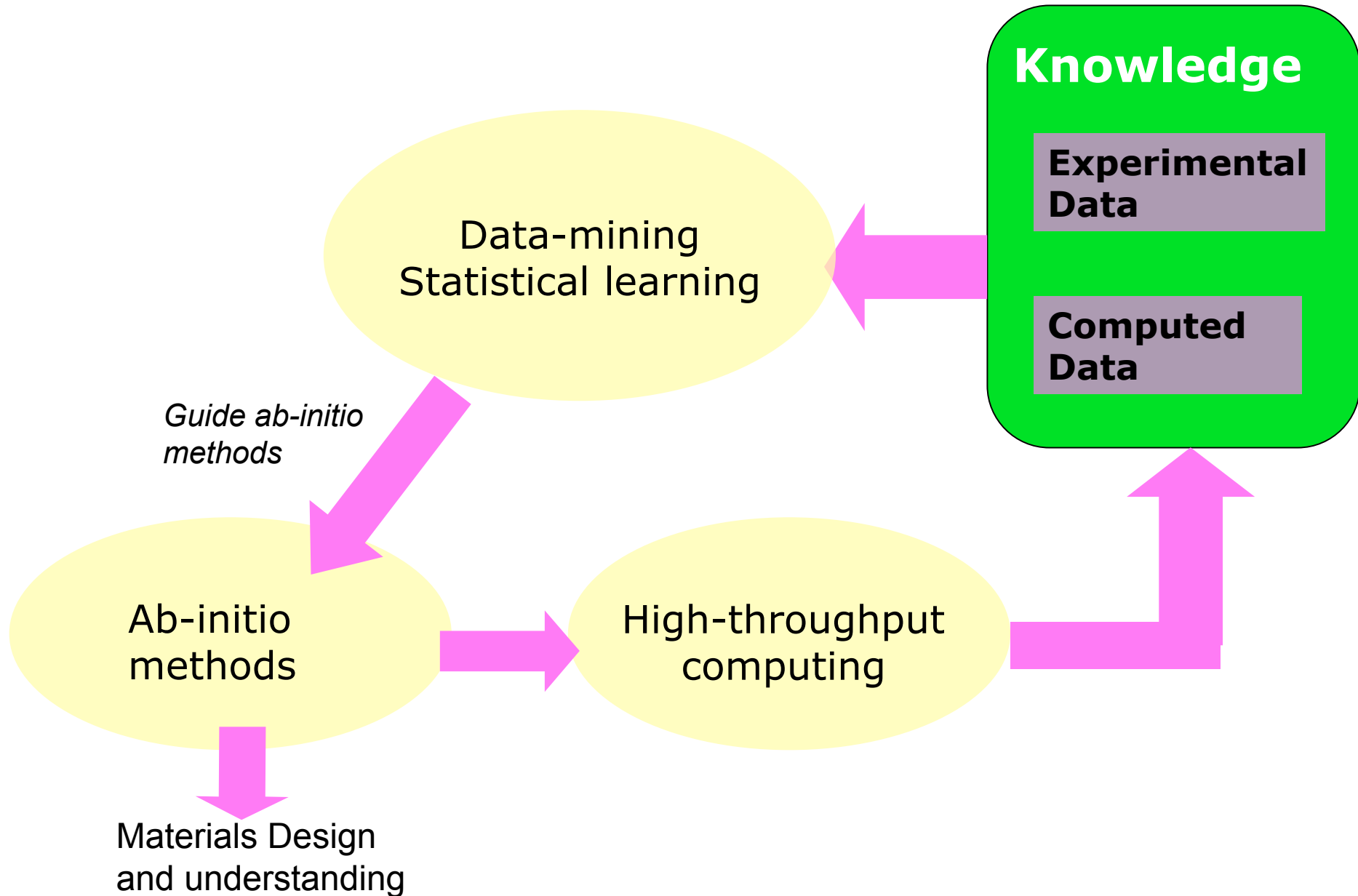


## Applied Problems in Energy

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

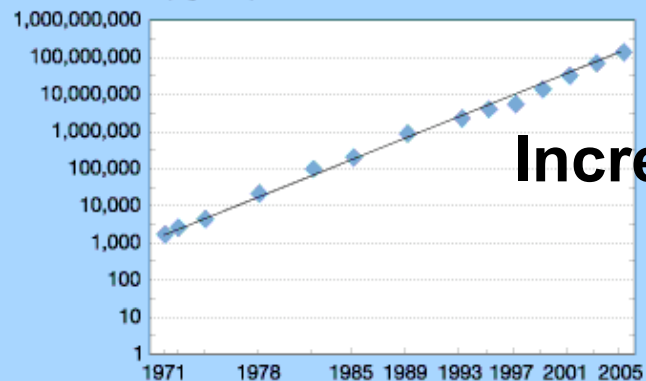
- Equations that describe matter can be solved accurately and stable

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

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

Figure 8-1.  
Moore's Law: 1971–2005

Transistor Count (log scale)



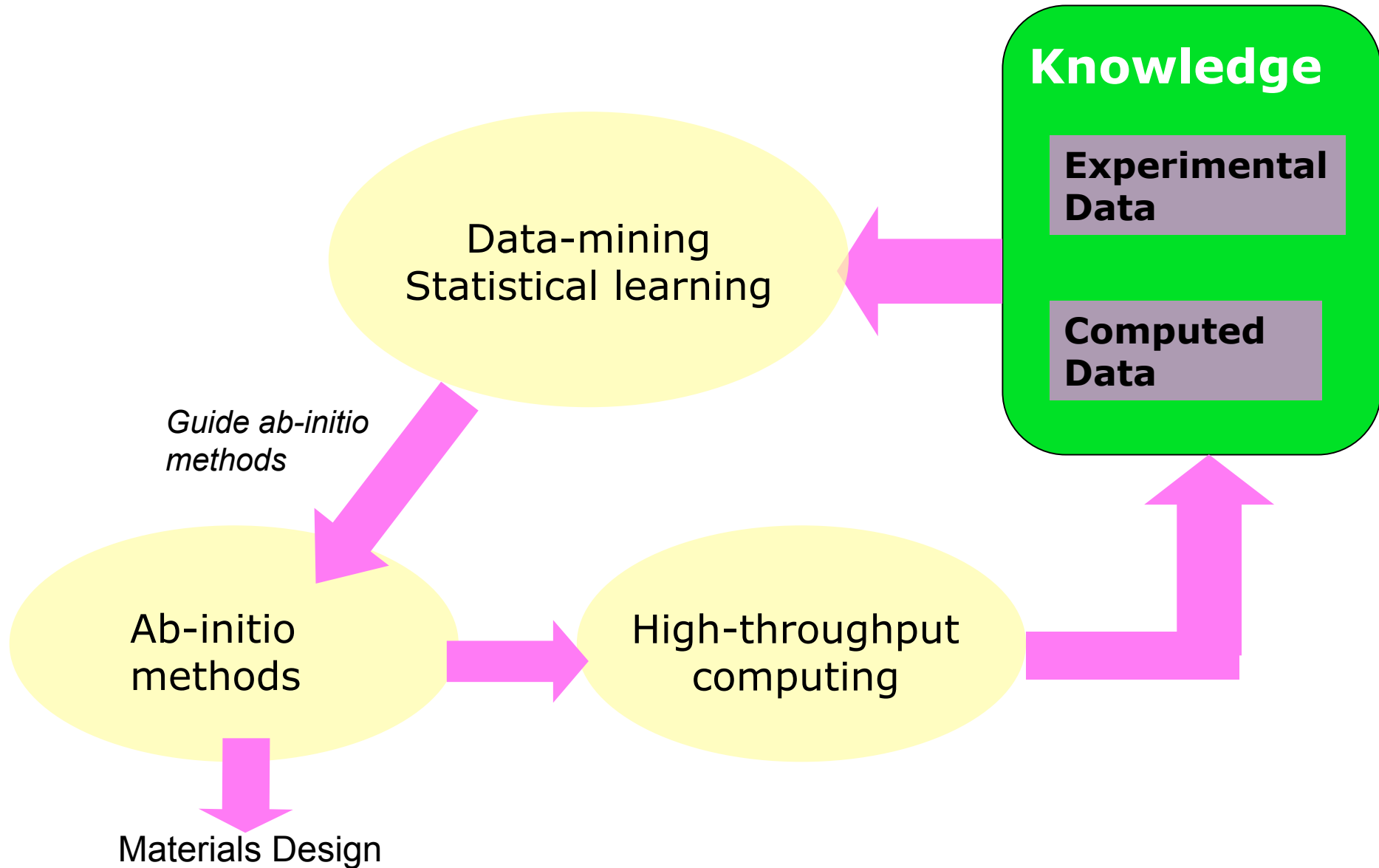
NOTES: The line on the graph represents the trend that defines Moore's Law. The data points reflect actual (1971–2001) and projected (2003–2005) data.

See appendix table 8-1. *Science & Engineering Indicators – 2002*

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

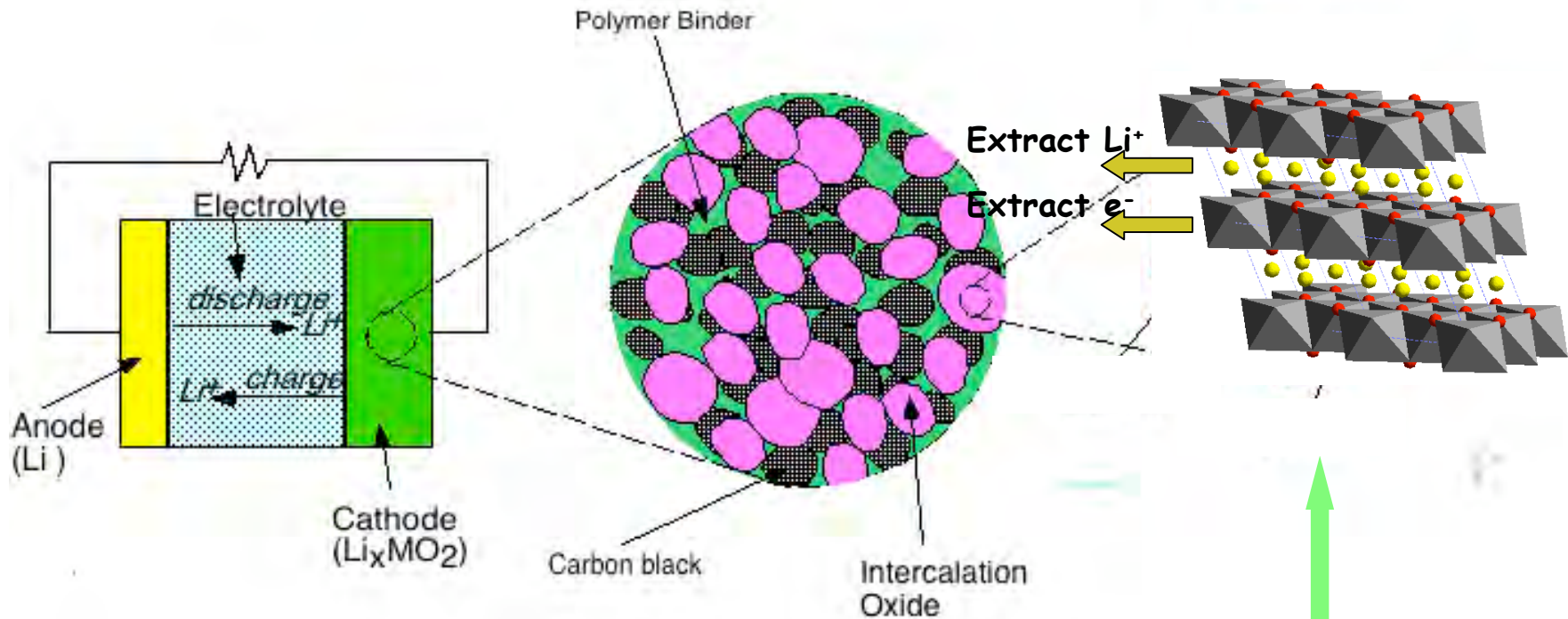


# Data-mining and High-Throughput *ab-initio* computing



**An example of “ab initio” computational materials design**

# Example: Design of high power Li battery materials



$\text{Li} + \text{Metal-Oxide} \rightarrow \text{Li-Metal-Oxide}$

**Cathode needs to host and exchange large amounts of  $\text{Li}^+$  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

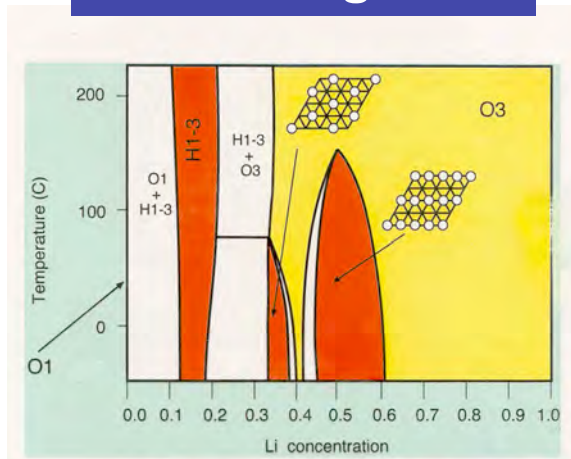
Voltage

→ energy density

	Exp	Calcul.
$\text{LiCoO}_2$	4.2	4.25
$\text{LiFePO}_4$	3.5	3.45
$\text{LiNiO}_2$	3.75	3.71

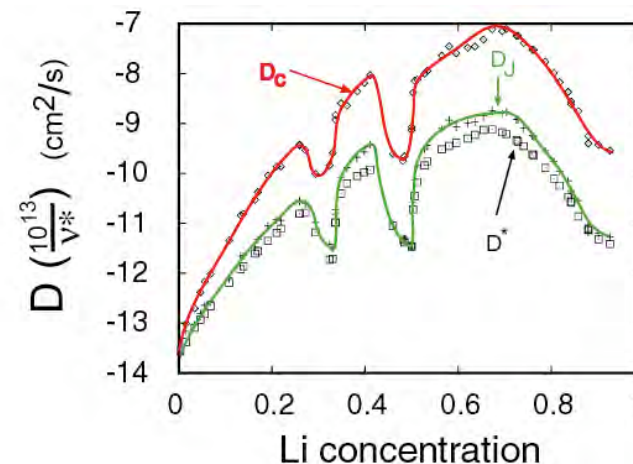
lifetime ←

Phase diagram



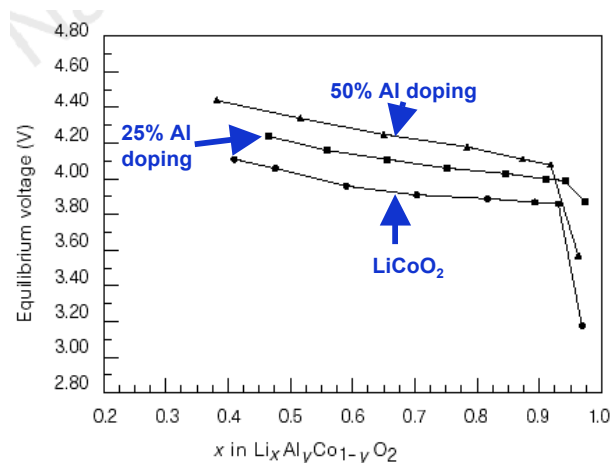
Li diffusivity

→ power



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

## Al-doping effect in $\text{LiCoO}_2$

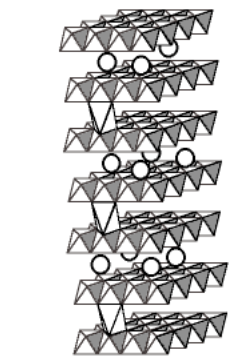


## Predict Interesting materials

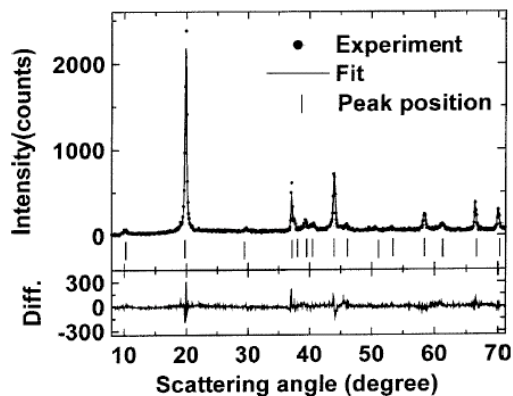
**$\text{LiNiPO}_4$ : 5.1 V**

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

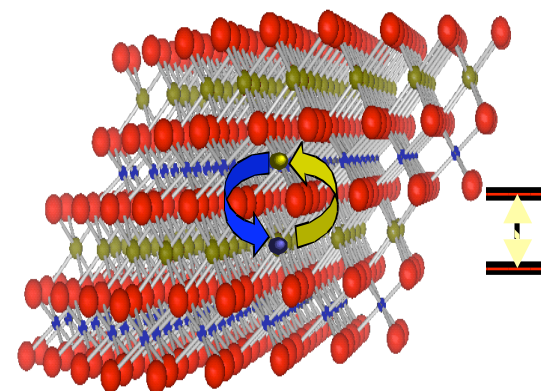
## H1-3 phase of $\text{LiCoO}_2$ at high charge



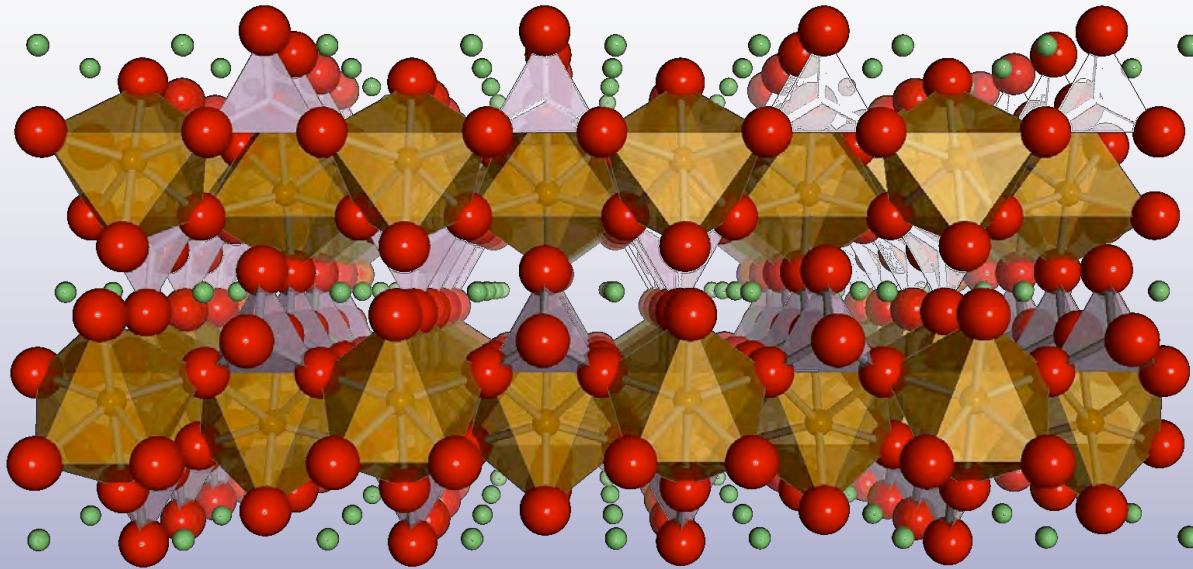
H1-3 (stage II compound)



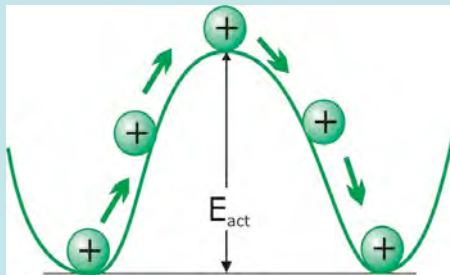
## High rate $\text{Li}(\text{Ni}_{0.5}\text{Mn}_{0.5})\text{O}_2$



# *LiFePO<sub>4</sub> designed for extreme rate behavior*



## Li migration barrier calculation

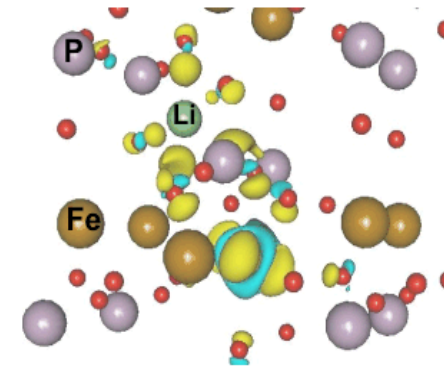
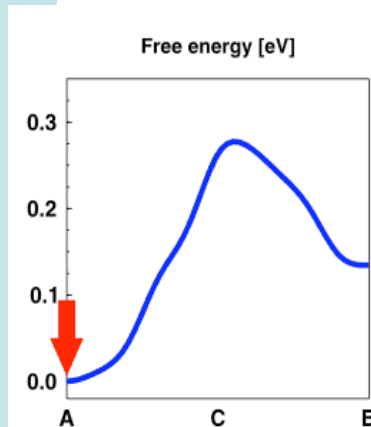


Along the *a*-axis:  $E_a > 1$  eV

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

Along the *b*-axis:  $E_a$  is low

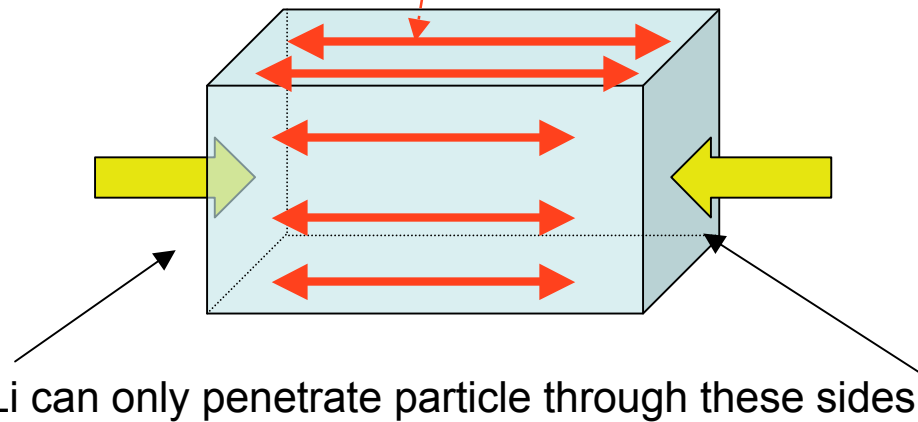
## Electron migration barrier calculation



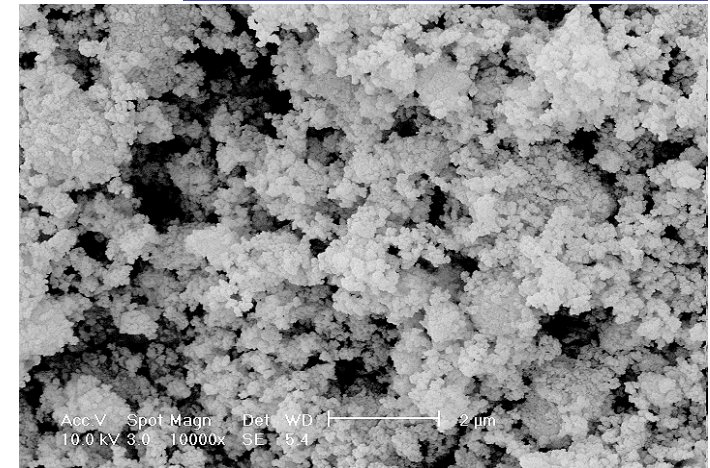
# LiFePO<sub>4</sub> with optimized surface structure



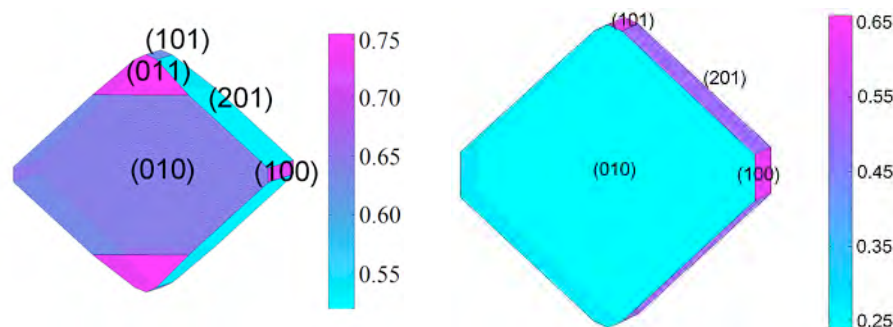
On other surfaces Li has to move over surface to the sides



Make nanoparticles

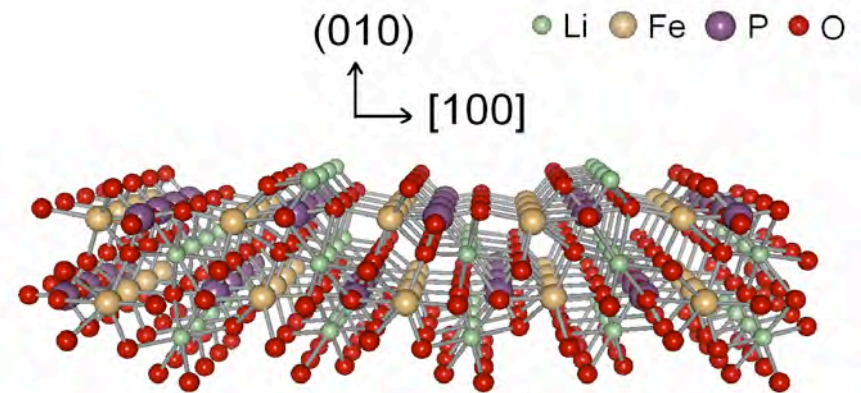


Optimize particle shape (surface energies !)

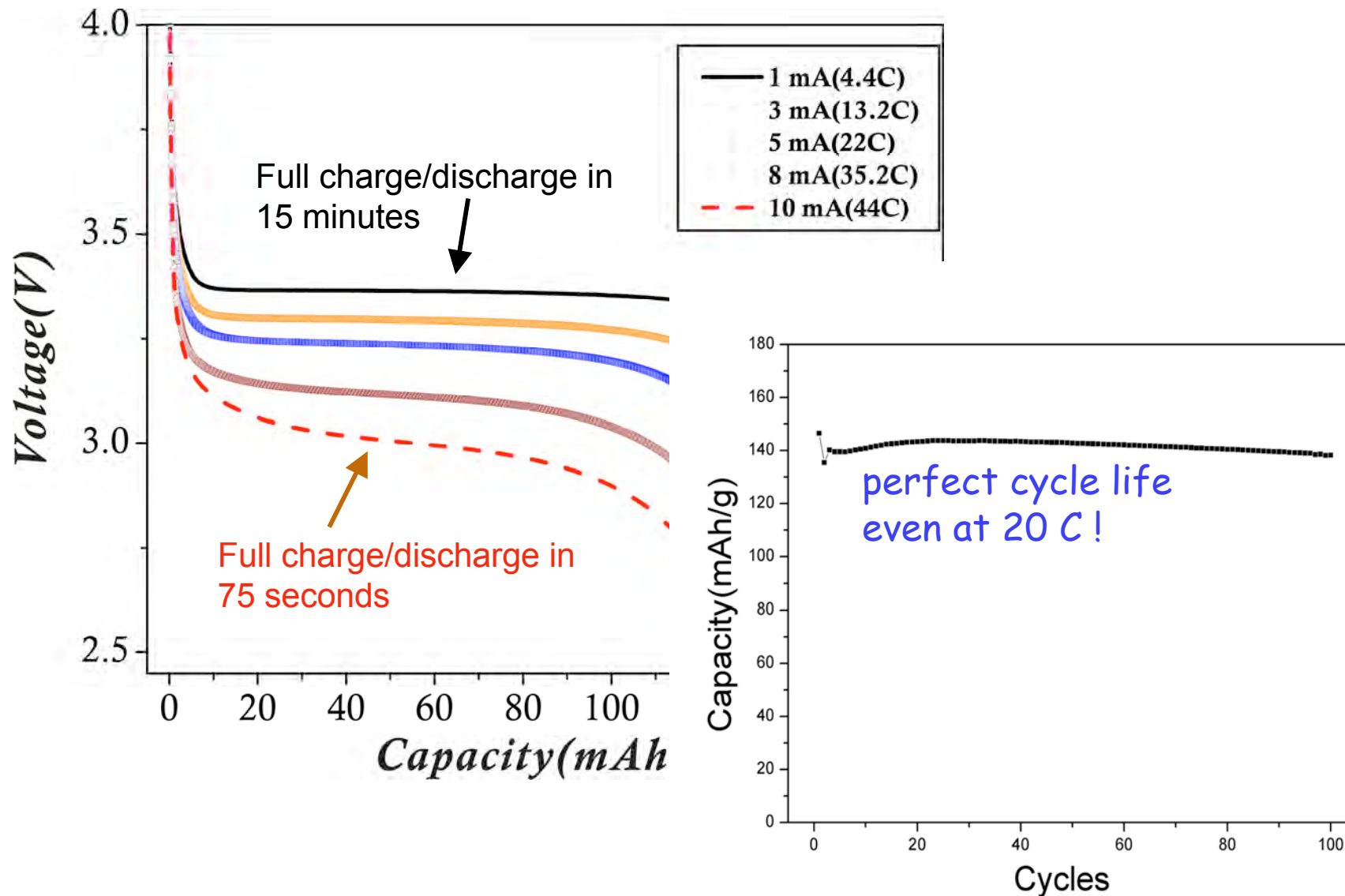


Platelet shape

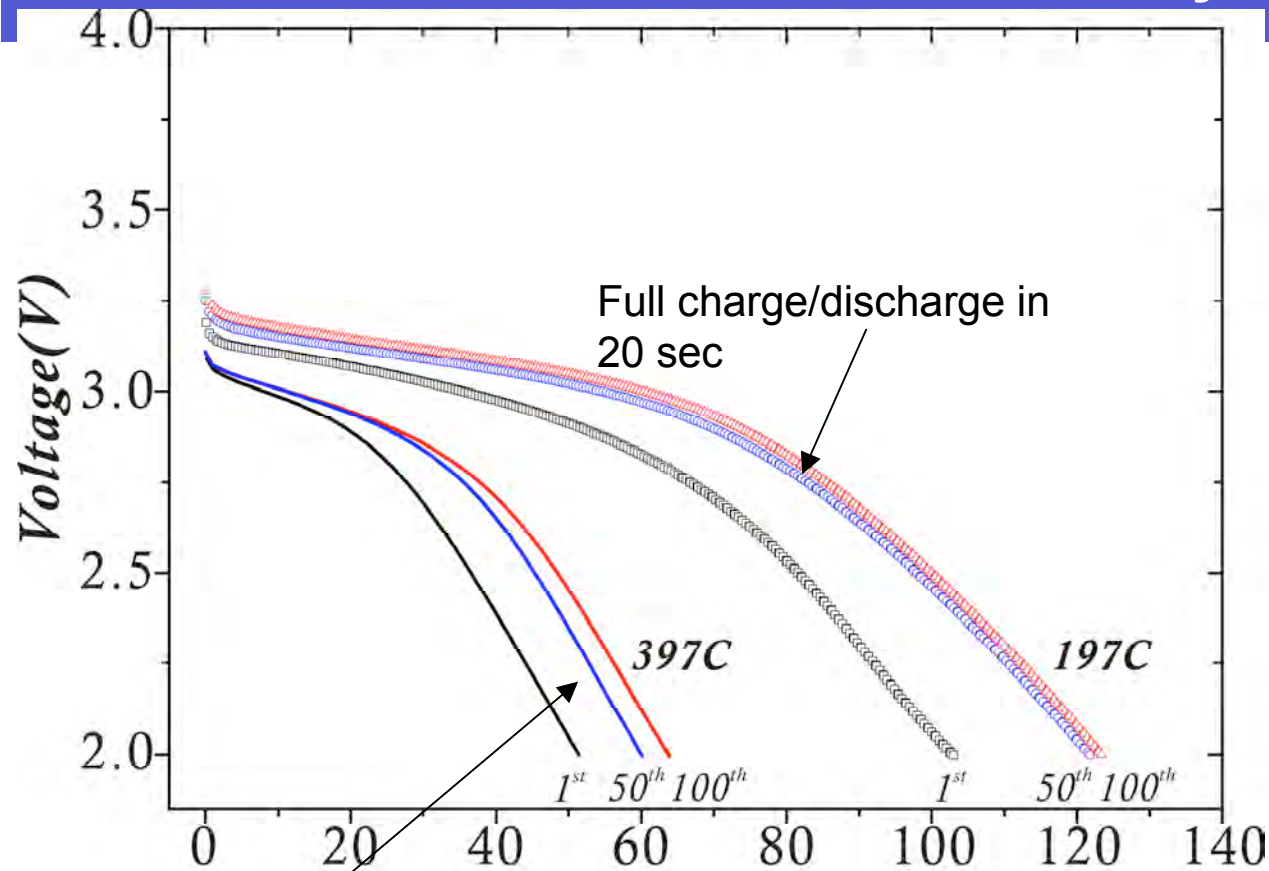
Modify surface chemistry so that Li ions can move rapidly over surface



# 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**

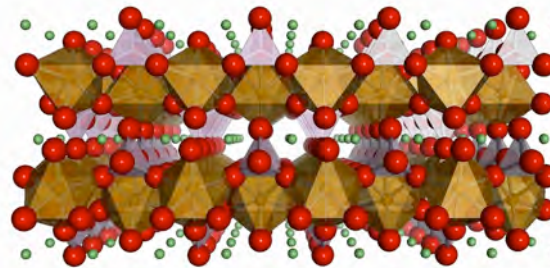
Power density:

**175 kW/liter**

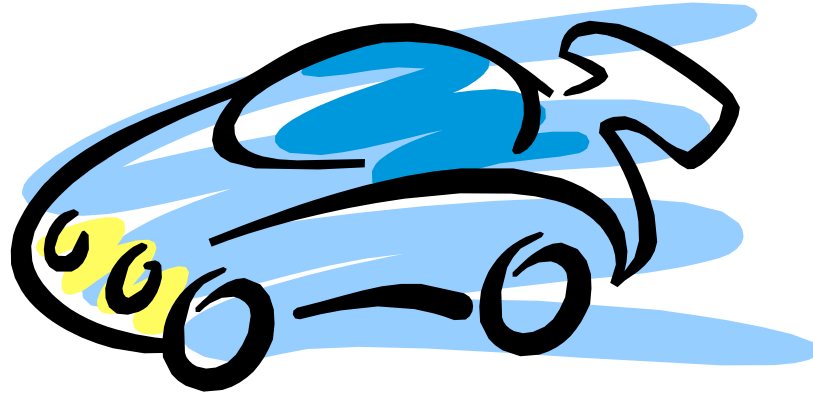
**90kW/kg**



Capac  
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 computation of crystal structures

We can now ask “meta” questions

Automatic  
Input file tra  
Error Chec  
run automa  
Automated  
implement

Run on E  
Developed  
numbers  
parallel e

The screenshot displays a 'Database search engine v1.0' window with search criteria for 'Li&O' in the 'ICSD' database. The search results table shows 1684 entries, with the first few rows as follows:

article title	db key	journal year	spacegroup hmna...	sum formula	chemical formula
The scattering pow...	22402-ICSD	19261925	F m -3 m	Li2 O1	Lithium Oxide
Ueber die Kristallst...	24143-ICSD	1953	P -6	Li2 O2	Lithium Peroxide
Die Kristallstruktur...	25530-ICSD	1957	P 63/m m c	Li2 O2	Lithium Peroxide
Preparation et etud...	282-ICSD	1976	I -4 m 2	Cu1 Li1 O1	Lithium Copper(I) ...
Neutron diffraction ...	288-ICSD	1976	I m -3	Li0.36 O3 W1	Lithium Tungstate *
The crystal structur...	493-ICSD	1976	C 1 2/c 1	Li1 O8 Ta3	Lithium Tritantalate
Structure cristalline...	638-ICSD	1976	P 1 n 1	Li1 O3 P1	Lithium Catena-ph...
Zur Kristallstruktur ...	1037-ICSD	1977	P m m n Z	Al1 Li5 O4	Pentalithium Alumi...
On the structure of	1043-ICSD	19772002	P n n m	Li2 O	

Overlaid on the search results are three windows:

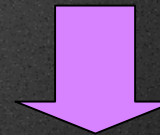
- Properties settings:** A window with three tabs: 'Search Properties', 'Display Properties', and 'Detailed Properties'. It contains various checkboxes for search and display options.
- Connection setting:** A window for configuring the database connection, showing fields for 'Location' (cedercomputer.mit.edu), 'Username' (highthroughput), and 'Password' (masked).
- Details on entry 2288:** A window providing detailed information for the selected entry, including journal volume (29), year (1926), mineral name (null), spacegroup (F m -3 m), and chemical formula (Lithium Oxide).



## 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

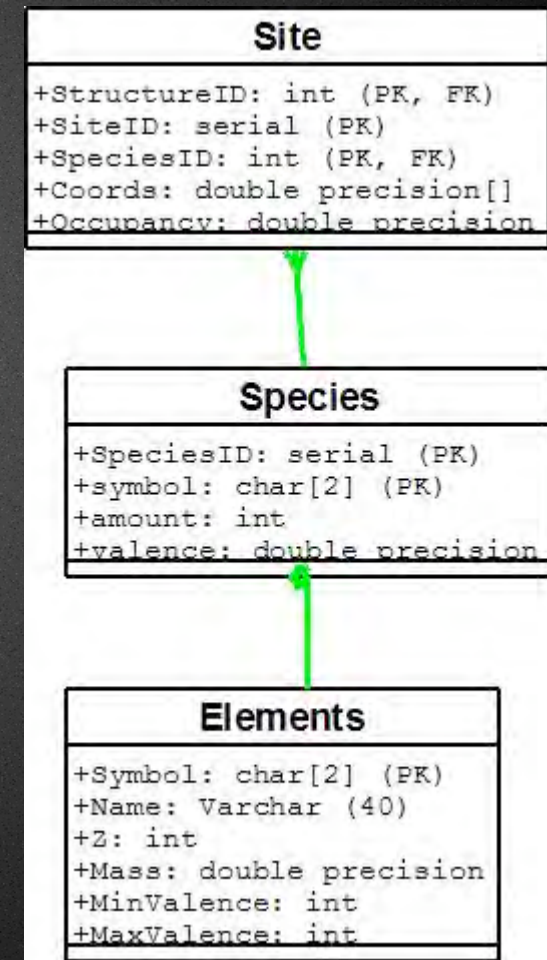
## putting the 'relation' in relational

- 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

\*\*\*

Structure 91305 - Li2 Mv1 02



**BASIC DB PROPERTIES**

name	Li2 Mv1 02
structure id	91305
entry id	9470
entry physical location	Documents\Battery\Li2 Mv1 02\Li2 Mv1_02\proj\analysis\21
db key	207050041141Mv1-02-016404-VASPI

**EXTENDED VASP INFORMATION**

initial charge	2.00000000
initial potential	1000

**SOURCE INFORMATION (VASP)**

user nickname	Abnab
comp creation time	2007-05-10 09:43:14
generation scheme (description)	PP: clean all the content of TM 0 structure
generation scheme (version)	1.0

**STRUCTURE INFORMATION**

prototype id	1001
prototype key	1
initial structure id	1001
initial potential	1000
structure group	1
description	Li2 Mv1 02 (Li2 Mv1 02) (207050041141Mv1-02-016404-VASPI) (structure Li2 Mv1 02) (L)
volume	-0.00000000

**cell**

0.21461412	0.00000000	0.00000000
-1.9994077E	2.7146649E	0.00000000
-0.00000000	-0.00000000	5.10720289

**atoms**

0.51333660	0.00000000	0.00000000	0.00000000
0.00000000	0.00000000	0.00000000	0.00000000
-0.00000000	-0.00000000	-0.00000000	0.00000000
0.00000000	0.00000000	0.00000000	0.00000000
0.00000000	0.00000000	0.00000000	0.00000000

**TRANSFORMATIONS**

**INVERSE TRANSFORMATIONS**

**PARAMETERS (VASP)**

To show VASP Parameters of all the basis the bottom of the screen

**CALCULATION FAMILY**

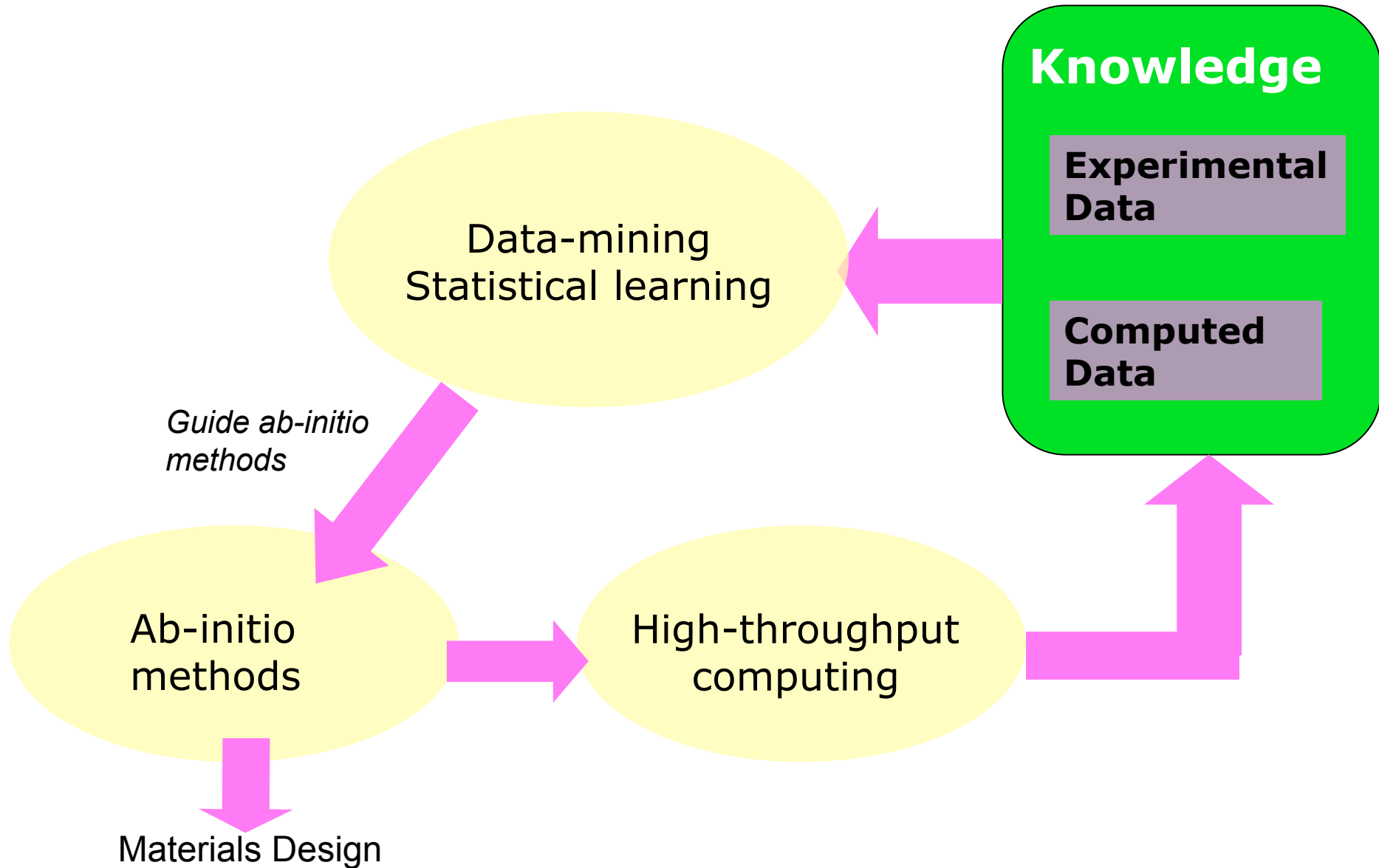
Calculation Name	Li2 Mv1 02
------------------	------------

**BATTERY PROPERTIES**

Average Voltage	3.4759465
Capacity (mAh/g)	7107.960577114
Capacity (mAh/g)	1402.218175318
Energy Density (mAh/g)	1702.318175318
Energy Density (mAh/g)	487.852467814
V01 change (mAh, percent)	1.8304777111E
V01 change (mAh, percent)	25.476221442175
Volume change (%)	26.77033668841
Volume change (mAh)	3.05635401541342
Volume change (mAh)	10.117466250000
Structure ID -> Name	925 0

**REMARKS**

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



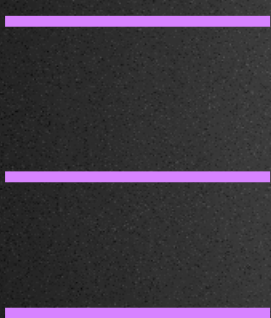
## Data mining — Li insertion Cathodes

Picture removed



# Data mining - Li insertion Cathodes

5V  
4V  
3V



Picture removed

2+

3+

4+

5+

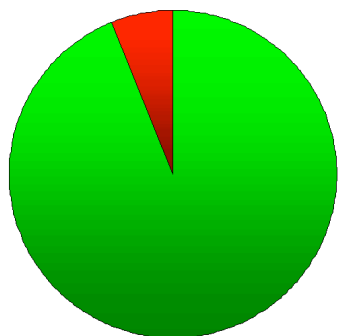
6+



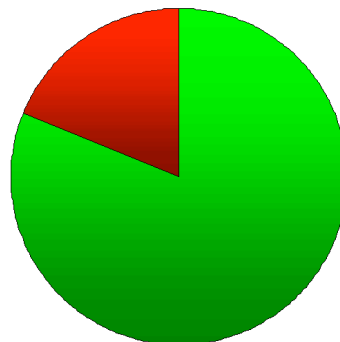
## Substantial added value of data

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

elements 94%  
(255)

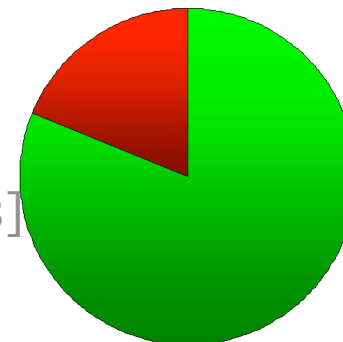


binaries 81%  
(4949)



ternaries 80%  
(10247)

[still in  
progress]



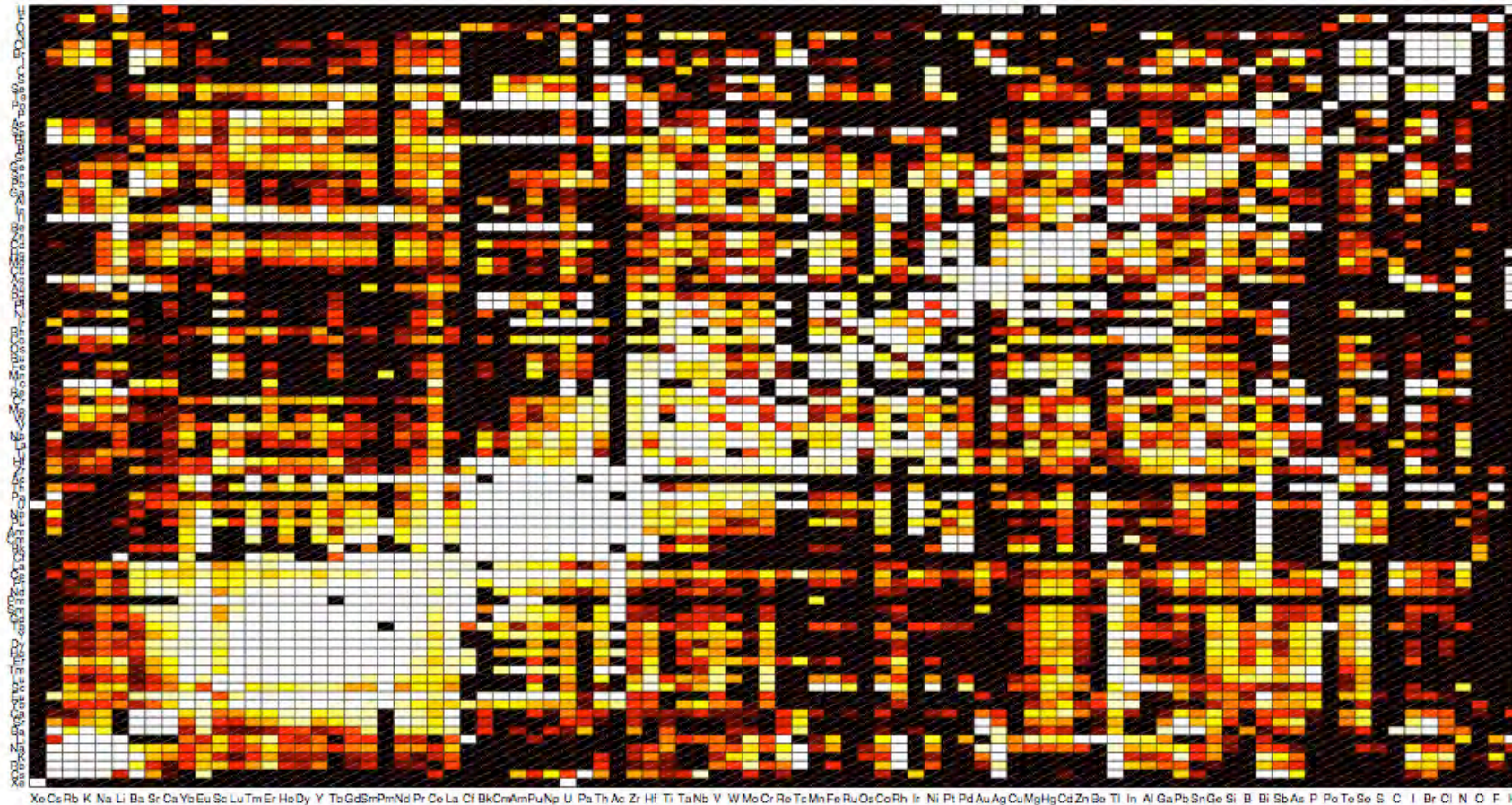
## Substantial added value of data

- **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**

# Ab-initio approach Energy + Optimization

Smart searching ?

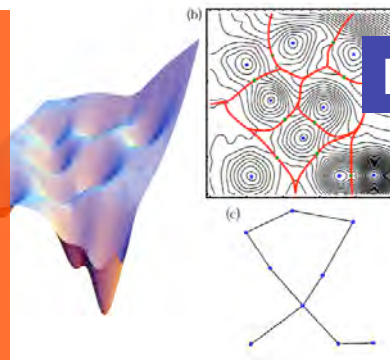
*Ground states only*

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*

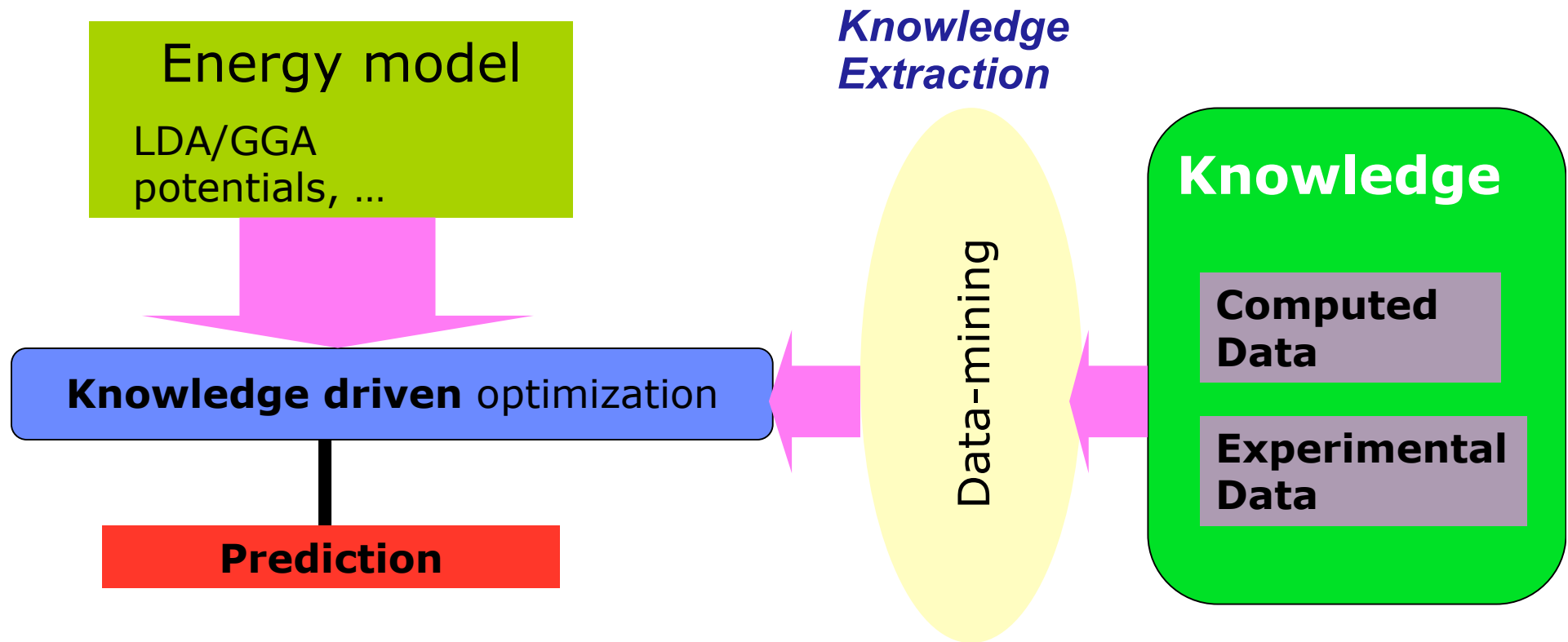
• *Can we learn about underlying physics in nature without making it explicit ? Do so in mathematically rigorous way*

**Mathematical "intuition"**

- Statistical learning
- Data mining
- Artificial intelligence

# Concept

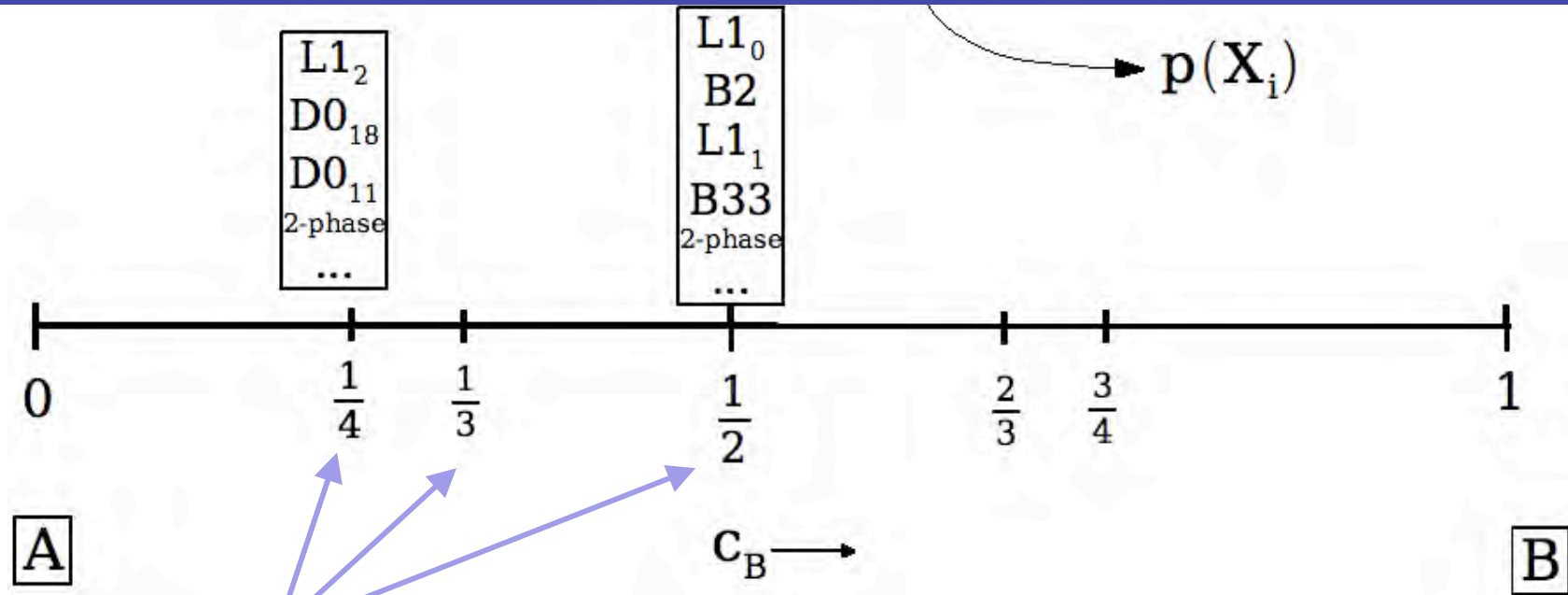
# Knowledge driven search



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

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

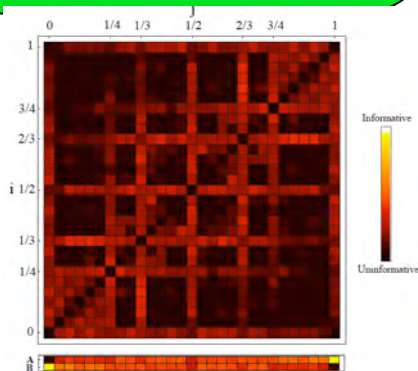
ground state set of the alloy

$P(\mathbf{X})$

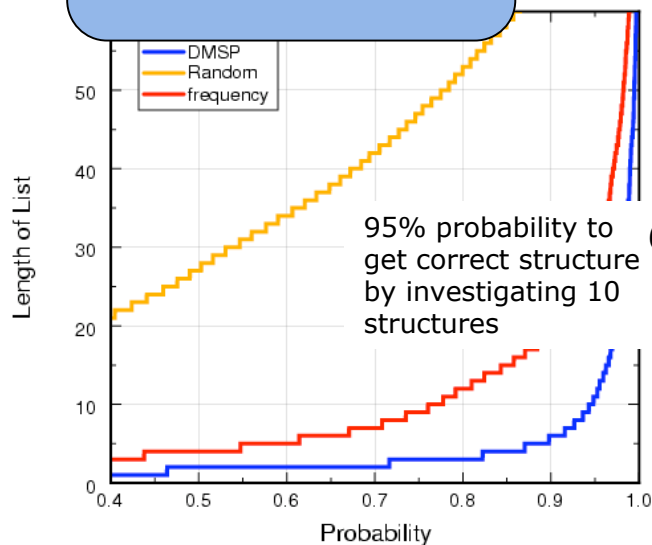
= knowledge of all Nature

# A first success: Data Mining to Predict Crystal Structure

DATABASE:  
ALL KNOWN  
CRYSTAL  
STRUCTURE  
ASSIGNMENTS



$P(X)$   
approximation  
from exp data

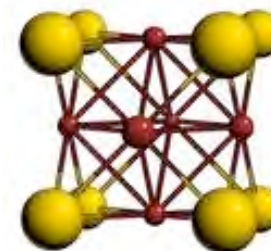


List of  
candidate  
structures

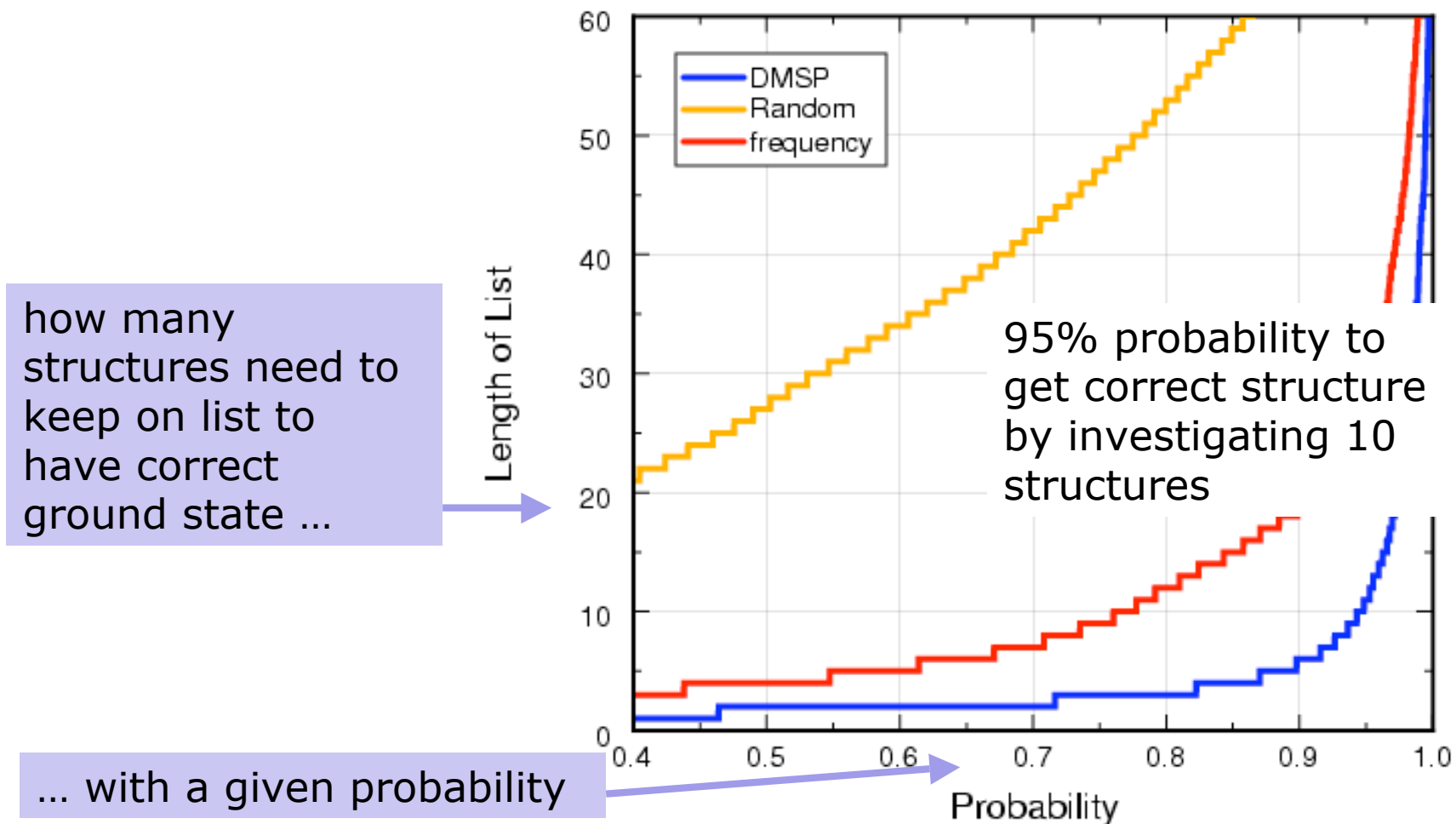
First  
Principles  
Computation



Correct Structure



Cross-validation prediction on > 5000 structures



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**

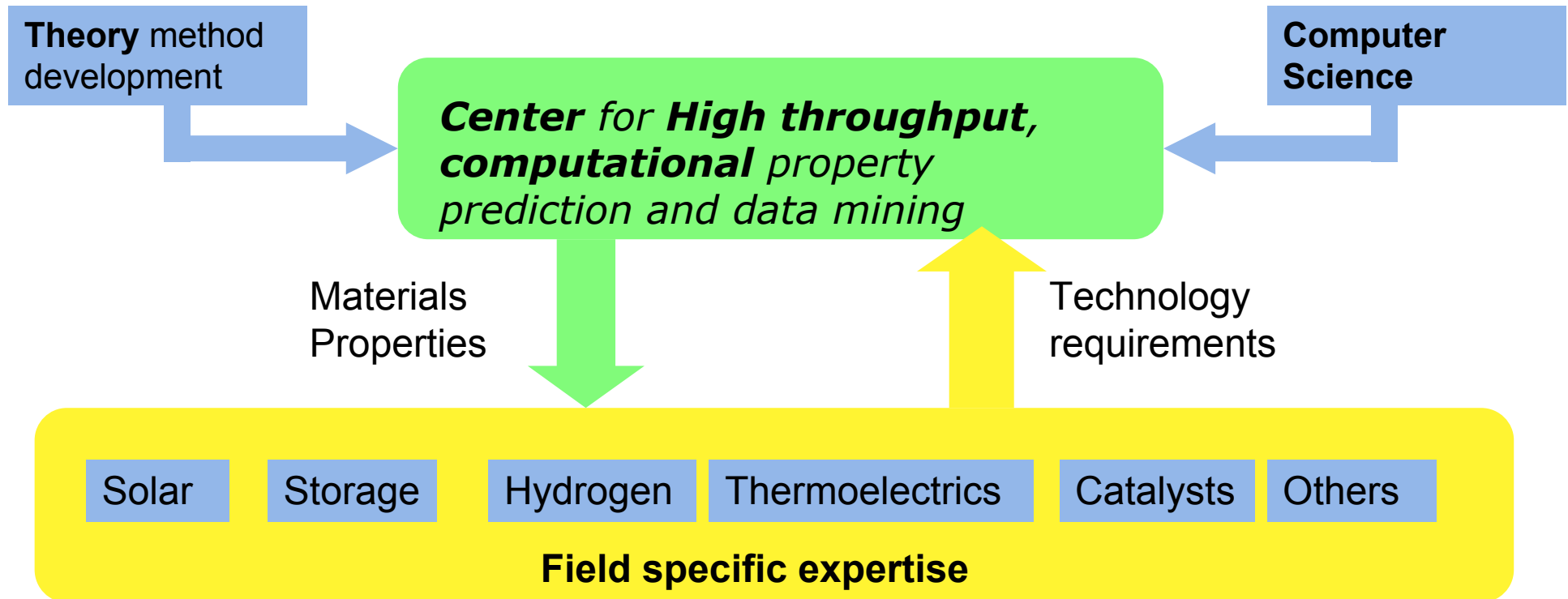
# Synergy between applications when data is preserved

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**