



# Machine learning for battery discovery

Gareth Conduit

# Alchemite™ for materials design



Train from **sparse** datasets

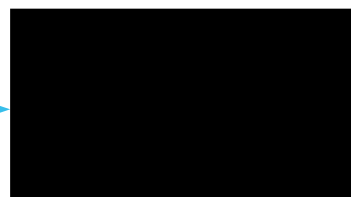
**Merge** simulations, physical laws, and experimental data

**Reduce** the need for expensive experimental development

**Accelerate** materials and drugs discovery

**Generic** with **proven** applications in materials discovery and drug design

# Black box machine learning for materials design



# Training machine learning



Cycles

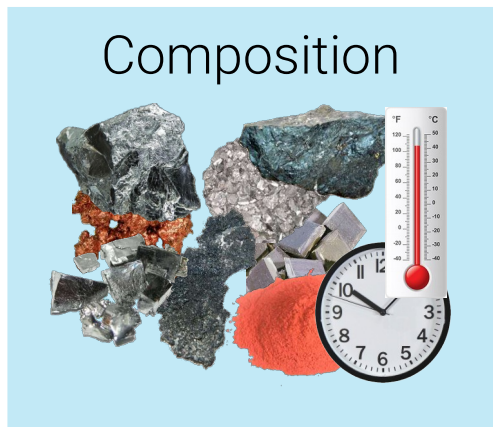
Charge

Potential

Cost

A vertical stack of images including a digital multimeter, an analog ammeter, a digital voltmeter, and a pile of US dollar bills, representing electrical and cost parameters.

# Machine learning for materials design



Cycles 

Charge 

Potential 

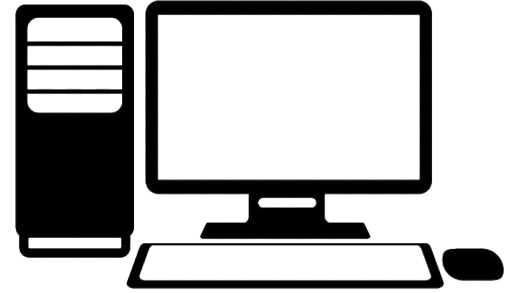
Cost 

# Two sources of information



## Experiment

Accurate  
Quantities of interest  
Lack of data  
Expensive



## Computational

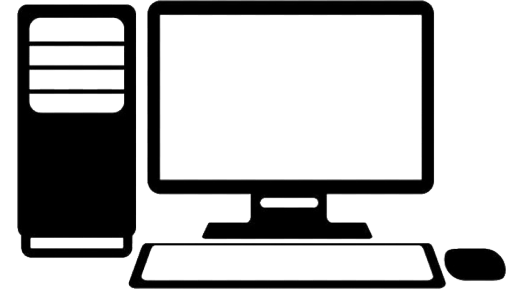
Less accurate  
Atom level insights  
Perform on demand  
Cheap to perform

# Merge information with machine learning



## Experiment

Accurate  
Quantities of interest  
Lack of data  
Expensive



## Computational

Less accurate  
Atom level insights  
Perform on demand  
Cheap to perform

# Nickel-Cobalt-Manganese (NCM) battery materials





# Design variables and target properties



Concentration  
of Ni, Mn, Co  
Location of  
atoms



Charge cycles  
Voltage  
Total charge  
Volume change  
Li migration  
Ground state  
Charge rate

# Design variables and target properties with DFT

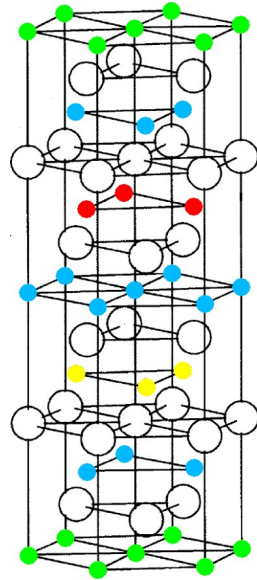


Concentration  
of Ni, Mn, Co  
  
Location of  
atoms

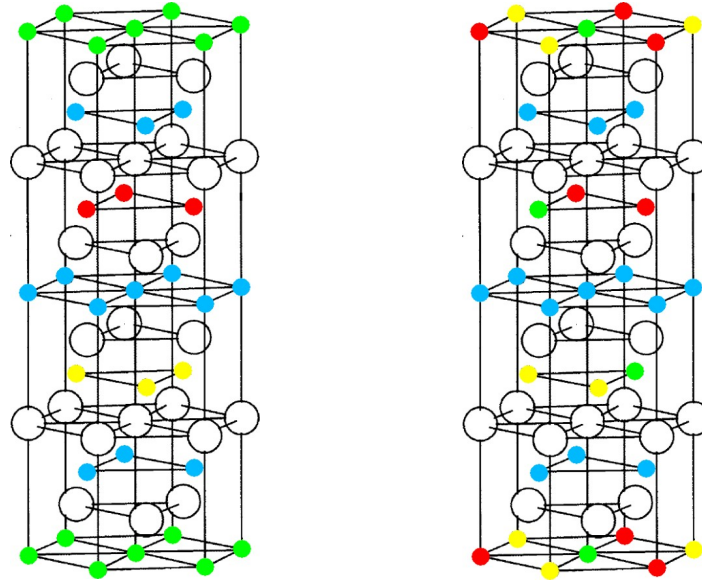


Volume change  
Li migration  
Voltage  
Ground state

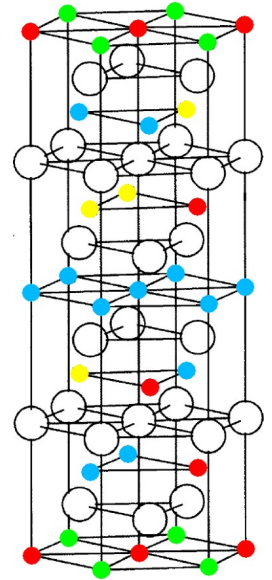
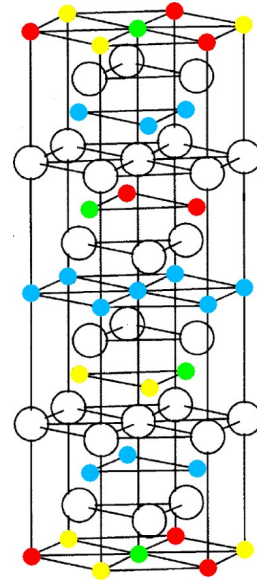
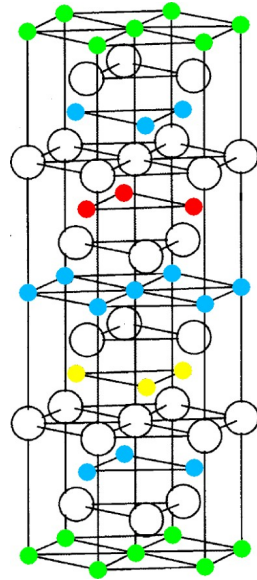
# Approach: exhaustive exploration of unit cells



# Approach: exhaustive exploration of unit cells



# Approach: exhaustive exploration of unit cells

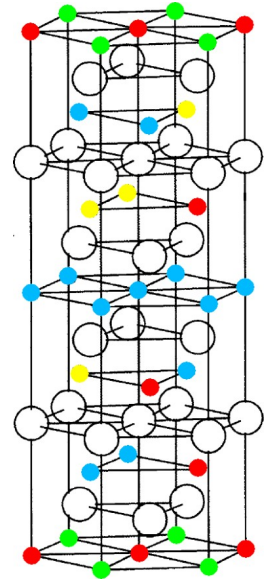
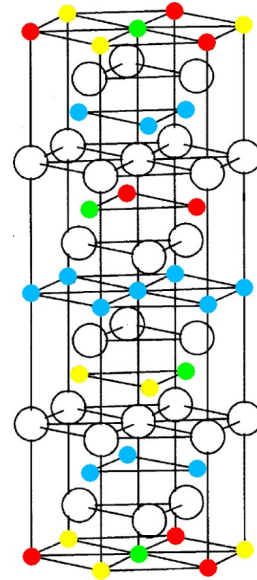
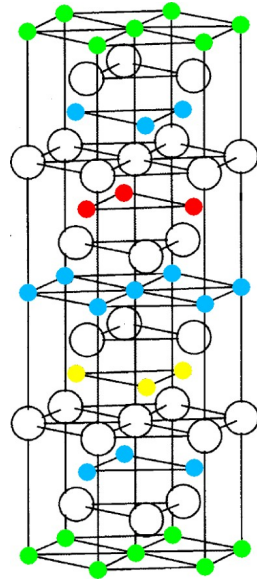


# Approach: exhaustive exploration of unit cells

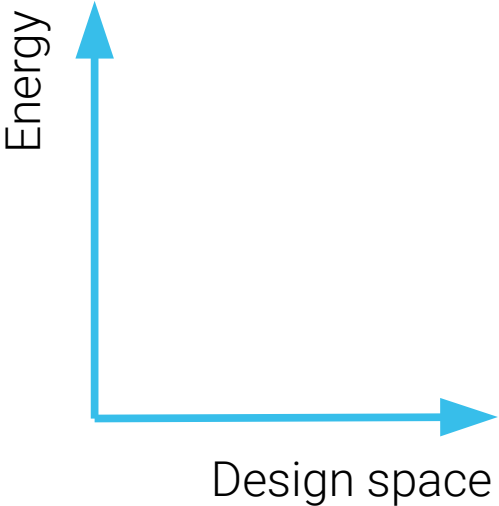
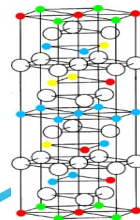
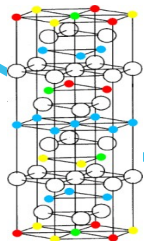
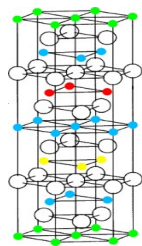


153153000  
permutations  
=42000 years

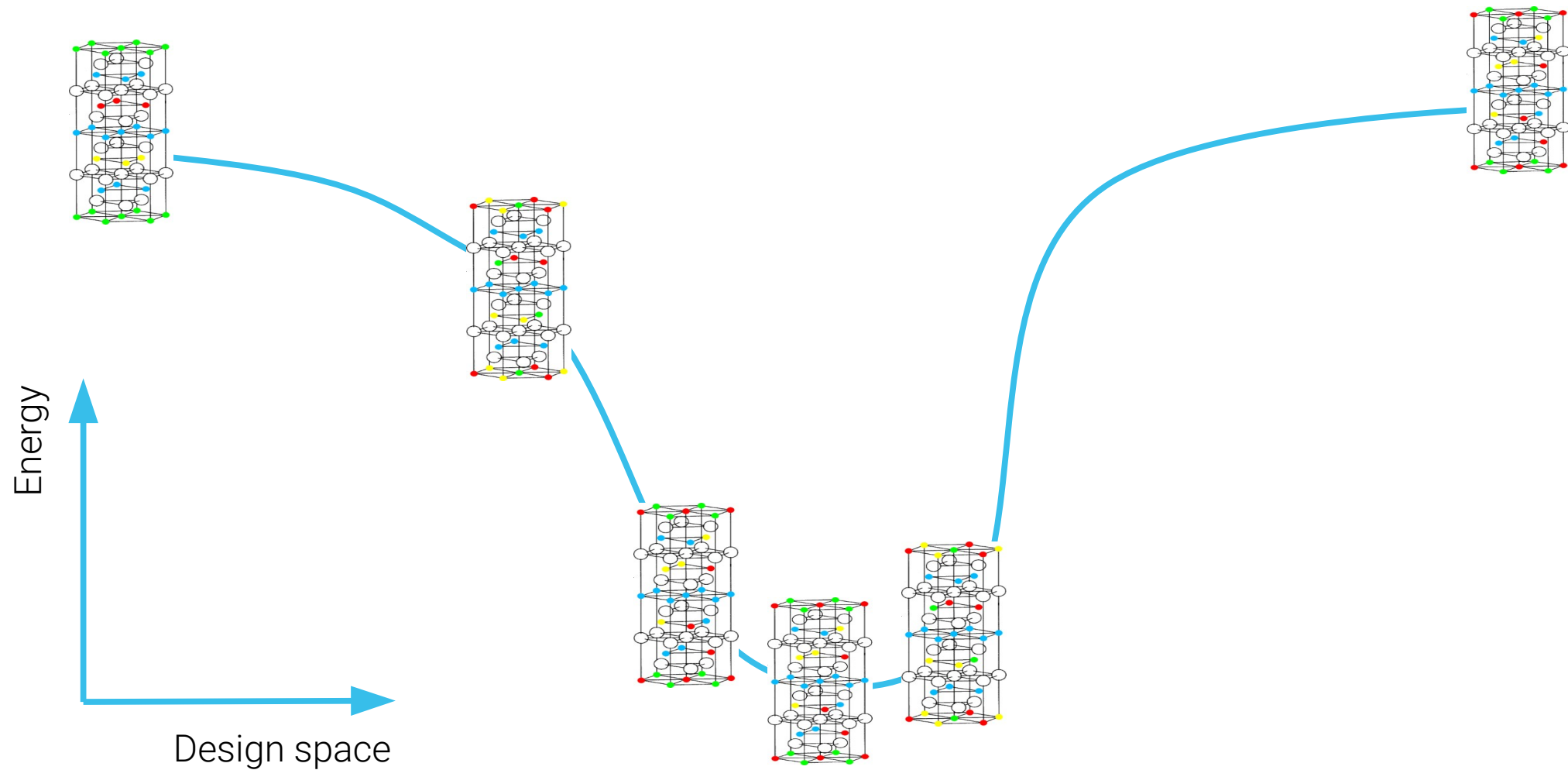
Only examine  
order that fits  
into the unit cell



# Train on initial results

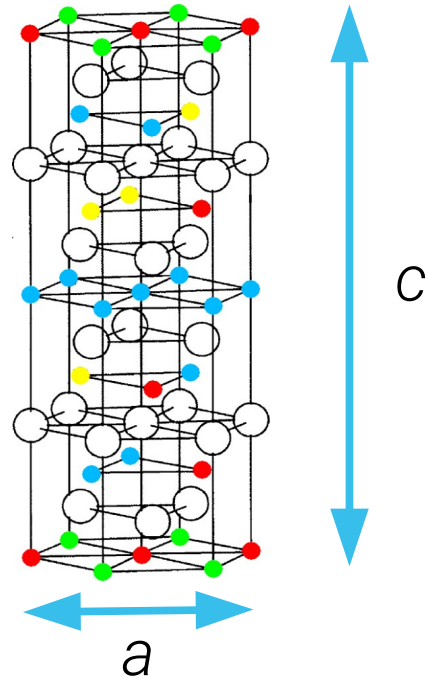


# Guided calculation for recursive learning

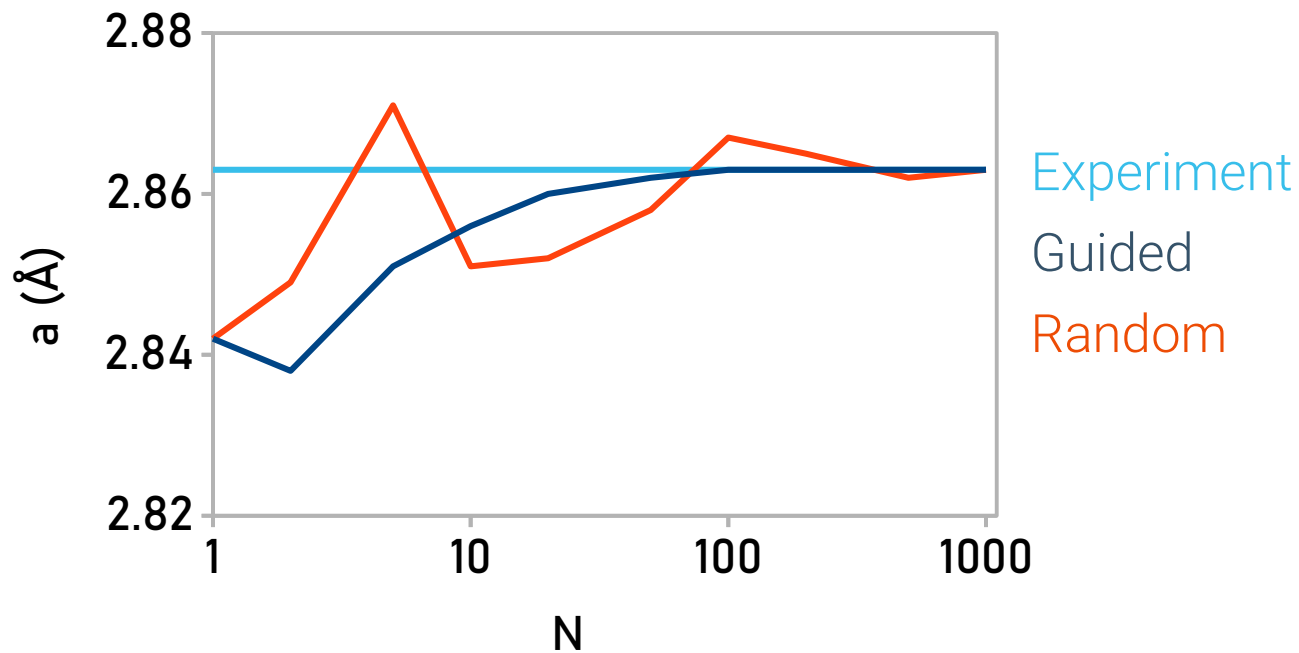




# Lattice constants

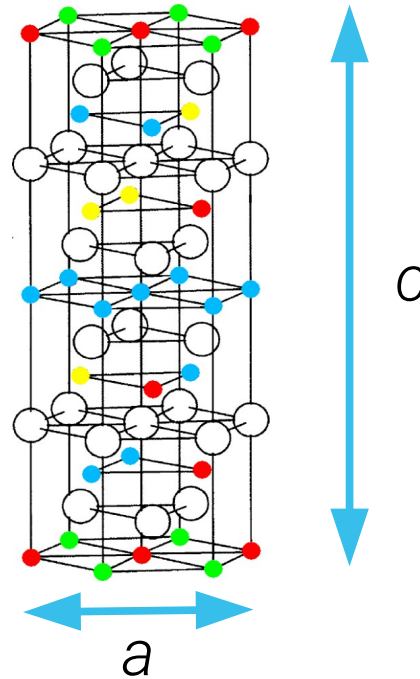


# How many calculations are required



Machine learning guidance requires **5-times fewer** calculations

# Predicting the lattice constant from DFT



Structure

$\text{LiNi}_{0.4}\text{Co}_{0.2}\text{Mn}_{0.4}\text{O}_2$  prediction

$\text{LiNi}_{0.4}\text{Co}_{0.2}\text{Mn}_{0.4}\text{O}_2$  experiment

$a$  (Å)

2.863

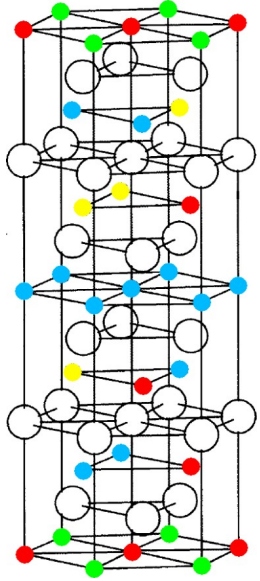
2.866

$c$  (Å)

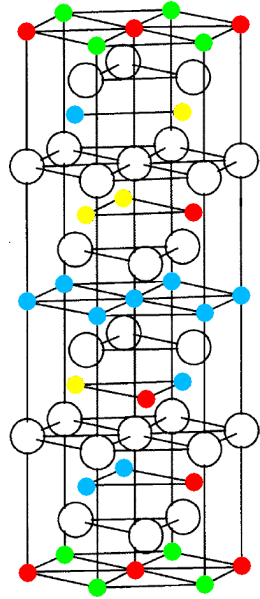
14.257

14.254

# Tracking Li migration

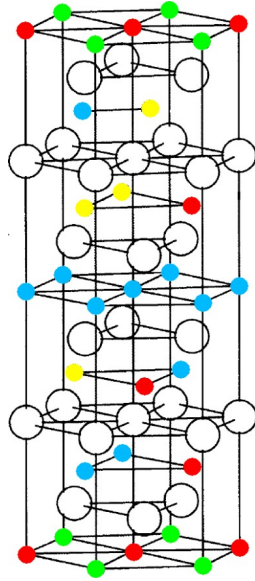


Original structure



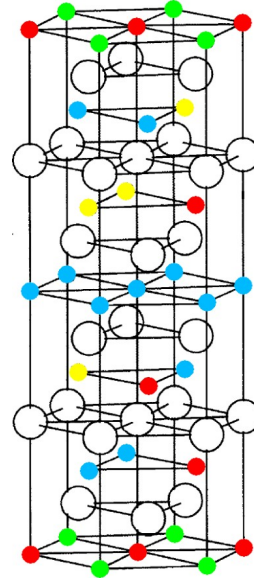
Remove Li

+ Li 



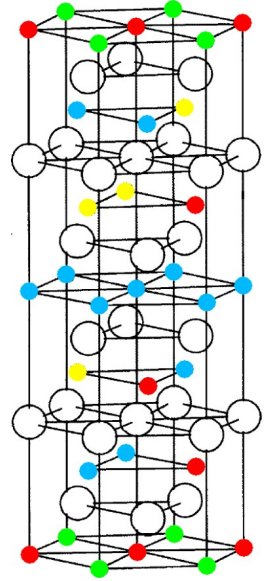
Relax atoms

+ Li 



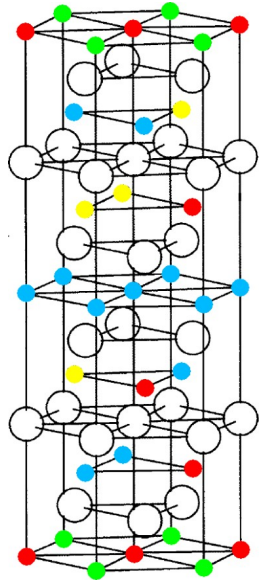
Reinsert Li

+ Li 



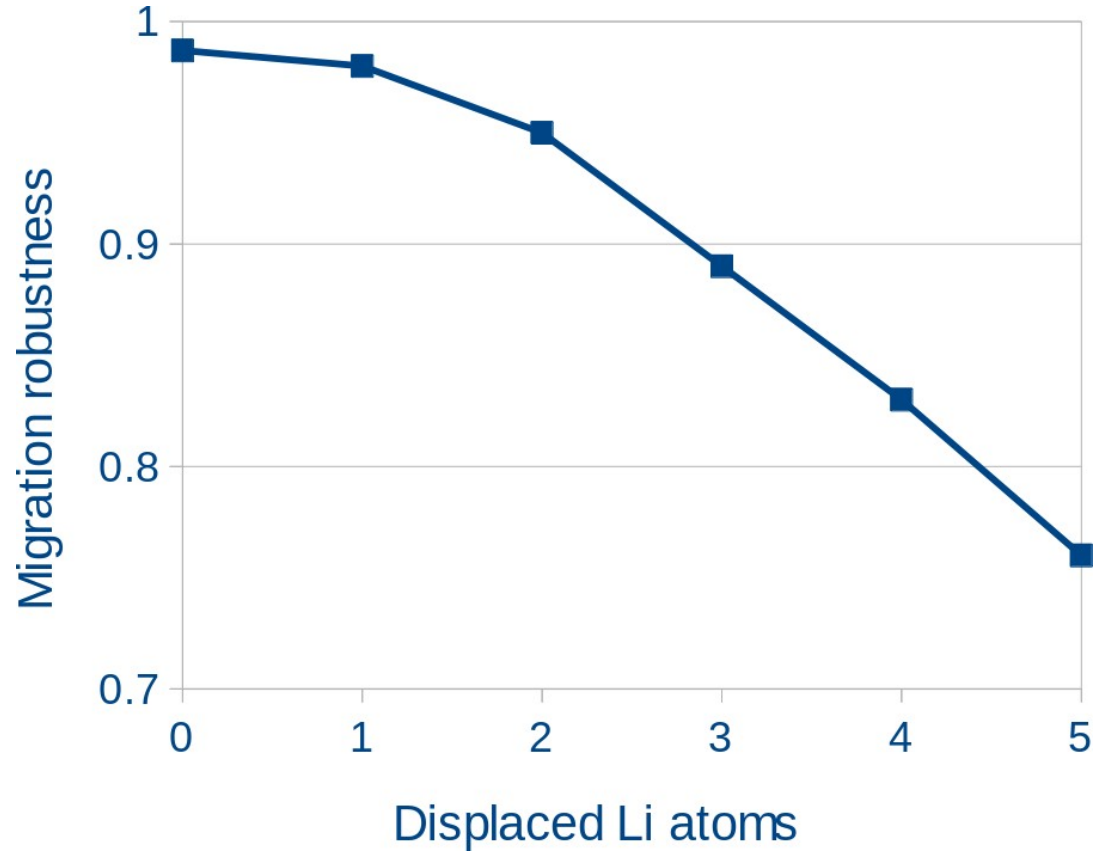
Relax atoms

# Li migration optimal structures

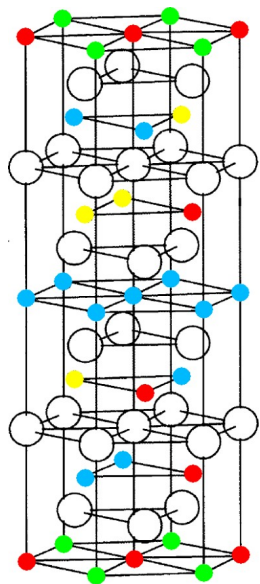


Ground state

82% robust

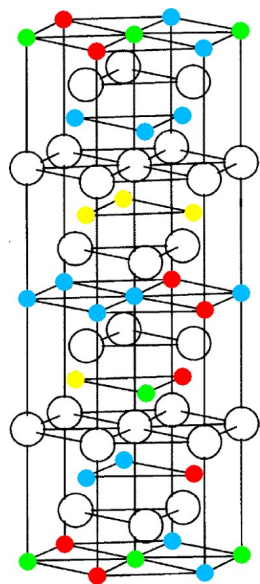


# Li migration optimal structures displacing 4xLi



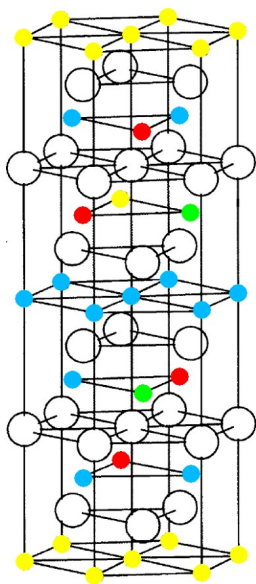
Ground state

82% robust



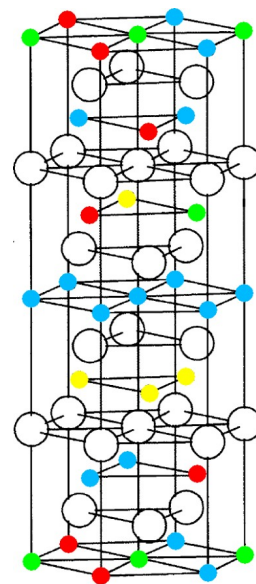
Configuration 1

100% robust



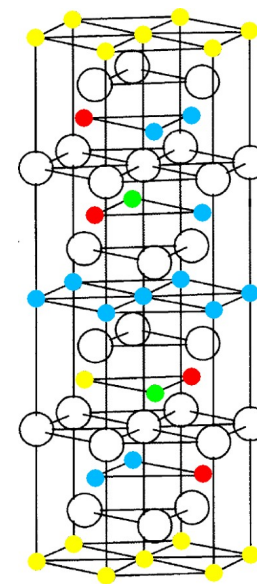
Configuration 2

100% robust



Configuration 3

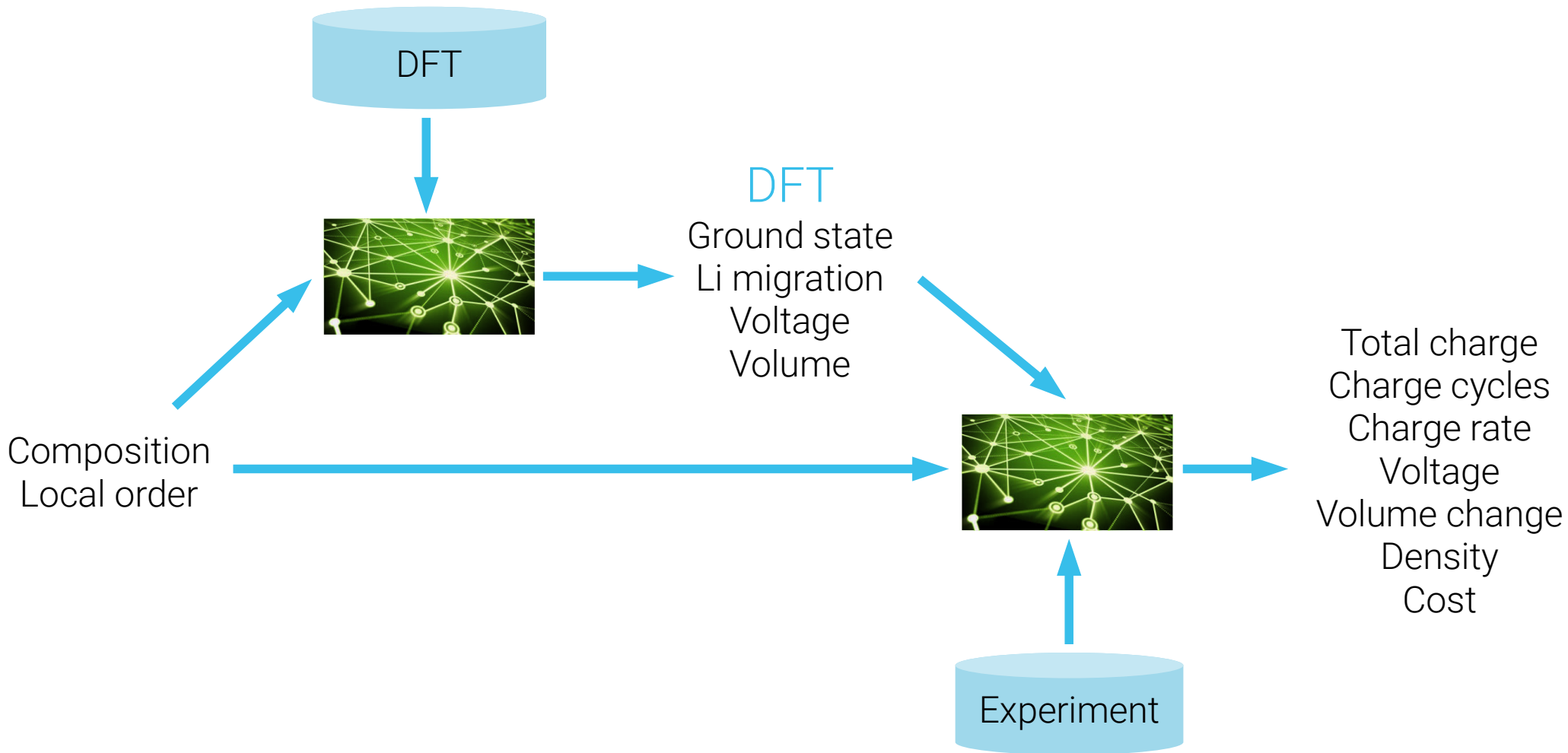
100% robust



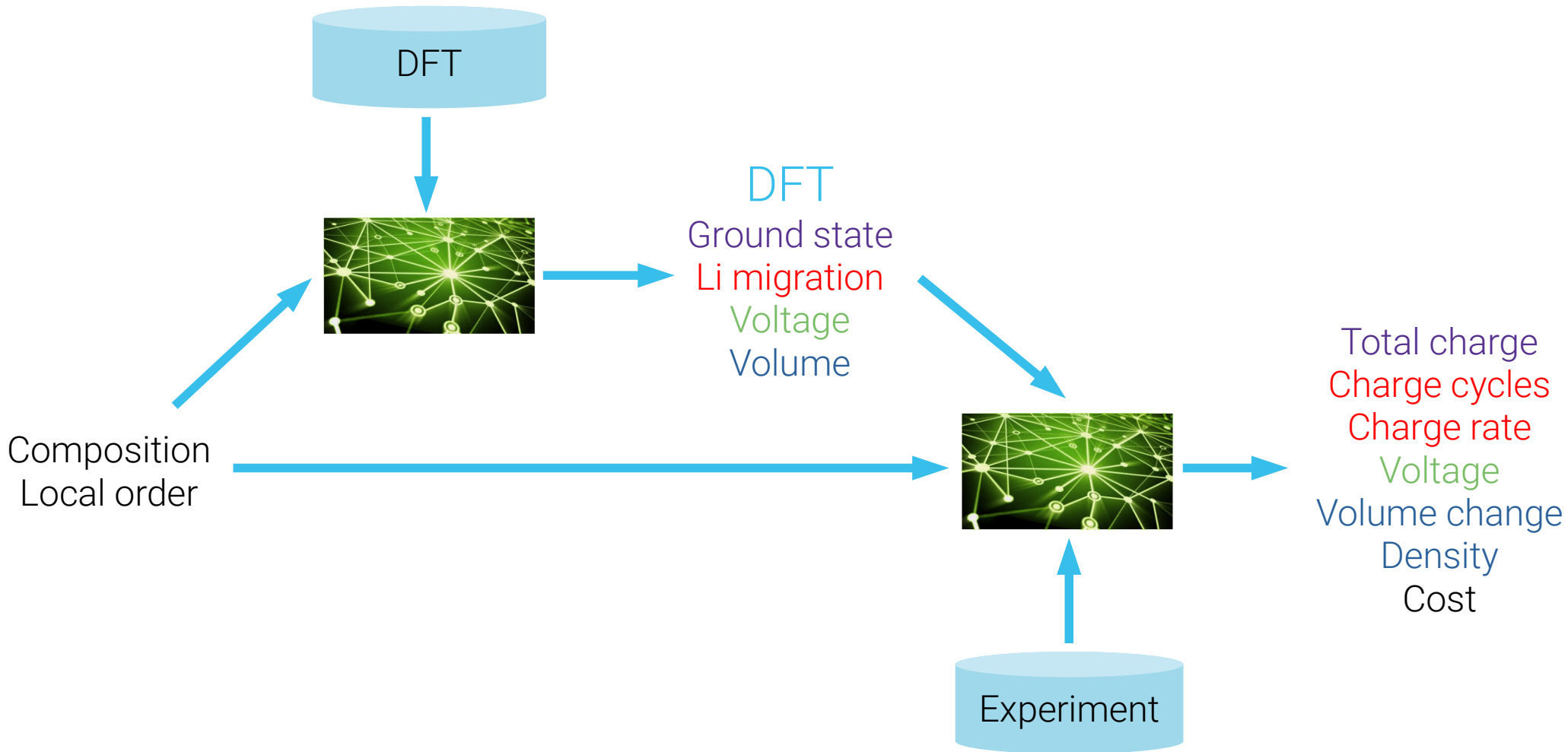
Configuration 4

100% robust

# Merge computational and experimental data



# Merge computational and experimental data





# Battery management system



Juxtapose physics-based modeling with machine learning

In-service data from a particular battery and others deployed to make bespoke predictions of remaining useful life

Model that spans time-scales to permit simultaneous state-of-health and state-of-charge predictions

Data from testing in first few cycles to predict long-term battery performance



## Predicting the state of charge and health of batteries using data-driven machine learning

Man-Fai Ng<sup>1</sup>, Jin Zhao<sup>2</sup>, Qingyu Yan<sup>2</sup>, Gareth J. Conduit<sup>3</sup> and Zhi Wei Seh<sup>4</sup>

Machine learning is a specific application of artificial intelligence that allows computers to learn and improve from data and experience via sets of algorithms, without the need for reprogramming. In the field of energy storage, machine learning has recently emerged as a promising modelling approach to determine the state of charge, state of health and remaining useful life of batteries. First, we review the two most studied types of battery models in the literature for battery state prediction: the equivalent circuit and physics-based models. Based on the current limitations of these models, we showcase the promise of various machine learning techniques for fast and accurate battery state prediction. Finally, we highlight the major challenges involved, especially in accurate modelling over length and time, performing in situ calculations and high-throughput data generation. Overall, this work provides insights into real-time, explainable machine learning for battery production, management and optimization in the future.

With rising concerns about global warming, electrification of transport has recently emerged as an important vision in many countries. The successful development of electric vehicles (EVs) depends highly on the cycling performance, cost and safety of the batteries. Rechargeable lithium-ion (Li-ion) batteries are currently the best choice for EVs due to their reasonable

where  $C_{cur}$  is the capacity of the battery in its current state,  $C_{full}$  is the capacity of the battery in its fully charged state,  $C_{nom}$  is the nominal capacity of the brand-new battery<sup>2</sup>.

In essence, SOC denotes the capacity of the battery in its current state compared to the capacity in its fully charged state (equivalent of a fuel gauge), while SOH describes the capacity of the battery

Predicting the State of Charge and Health of Batteries using Data-Driven Machine Learning  
Nature Machine Intelligence 2, 161 (2020)

# Battery component specification



Improved understanding of battery properties

Empowers specification of the optimal components

Bespoke battery design for each customer



## Predicting the state of charge and health of batteries using data-driven machine learning

Man-Fai Ng<sup>1</sup>, Jin Zhao<sup>2</sup>, Qingyu Yan<sup>2</sup>✉, Gareth J. Conduit<sup>3</sup>✉ and Zhi Wei Seh<sup>4</sup>✉

Machine learning is a specific application of artificial intelligence that allows computers to learn and improve from data and experience via sets of algorithms, without the need for reprogramming. In the field of energy storage, machine learning has recently emerged as a promising modelling approach to determine the state of charge, state of health and remaining useful life of batteries. First, we review the two most studied types of battery models in the literature for battery state prediction: the equivalent circuit and physics-based models. Based on the current limitations of these models, we showcase the promise of various machine learning techniques for fast and accurate battery state prediction. Finally, we highlight the major challenges involved, especially in accurate modelling over length and time, performing in situ calculations and high-throughput data generation. Overall, this work provides insights into real-time, explainable machine learning for battery production, management and optimization in the future.

With rising concerns about global warming, electrification of transport has recently emerged as an important vision in many countries. The successful development of electric vehicles (EVs) depends highly on the cycling performance, cost and safety of the batteries. Rechargeable lithium-ion (Li-ion) batteries are currently the best choice for EVs due to their reasonable

where  $C_{\text{cur}}$  is the capacity of the battery in its current state,  $C_{\text{full}}$  is the capacity of the battery in its fully charged state,  $C_{\text{nom}}$  is the nominal capacity of the brand-new battery<sup>2</sup>.

In essence, SOC denotes the capacity of the battery in its current state compared to the capacity in its fully charged state (equivalent of a fuel gauge), while SOH describes the capacity of the battery

# Other materials designed



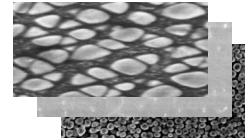
Steel welding consumables



Titanium additive manufacturing



High temperature alloys



Lubricants



Journal of Chemical Physics 153, 014102 (2020)  
Fluid Phase Equilibria 501, 112259 (2019)  
Materials & Design 168, 107644 (2019)  
Computational Materials Science 147, 176 (2018)

Physical Review Applied 12, 034024 (2019)  
Matter 1, 219 (2019)  
Scripta Materialia 146, 82 (2018)  
Materials & Design 131, 358 (2017)



## API for integration

intellegens

Search...

- PUT Train a model
- POST Load model into memory
- PUT Unload model from memory
- PUT Impute missing data
- PUT Validate given data
- PUT Predict given and missing data
- PUT Find the outlying values in a dataset
- PUT Suggest which missing values to provide from the training dataset to improve future predictions.
- GET Get all optimize jobs for given model ID
- POST Optimize for specified

Suggest which missing values to provide from the training dataset to improve future predictions.

Get a specified number of suggestions for additional measurements which are currently omitted from the data used to train the model. These measurements, if provided, would best improve subsequent predictions for a given list of 'targetColumns'.

AUTHORIZATIONS: `oauth (alchemiteapi.models.suggest)`

PATH PARAMETERS

→ `id` required string <uuid>  
 Example: 00112233-4455-6677-8899-aabbccddeeff  
 Unique identifier for the model.

REQUEST BODY SCHEMA: `application/json`

→ `numberOfSuggestions` integer  
 Default: 10  
 Request the top numberOfSuggestions values that will most improve predictions for the requested targetColumns.

→ `targetColumns` Array of strings  
 A list of column headers which all appear in the training data. Suggested measurements will be targeted to best improve predictions for these columns. If not given then targetColumns will be treated as being all columns.

## Within the browser

intellegens Alchemite™ Analytics + Create model

M0 Cheddar/Akawi ha...  
432 rows 9 cols

Analytics All properties

ACTUAL VS PREDICTED PROPERTY VS PROPERTY IMPORTANCE

HARDNESS R<sup>2</sup>: 0.8206

3D Scatter Plot: Predicted vs Actual vs Spinel (wt. %)

Powered by Alchemite API v0.17.4

# Prospects for the future



**Merge** computational simulations and experimental data

**Design** battery materials

**Guided** simulations and experiments leads to **5x** speedup

**Embedded** battery management software