Bridging computation and experiment for energy materials discovery

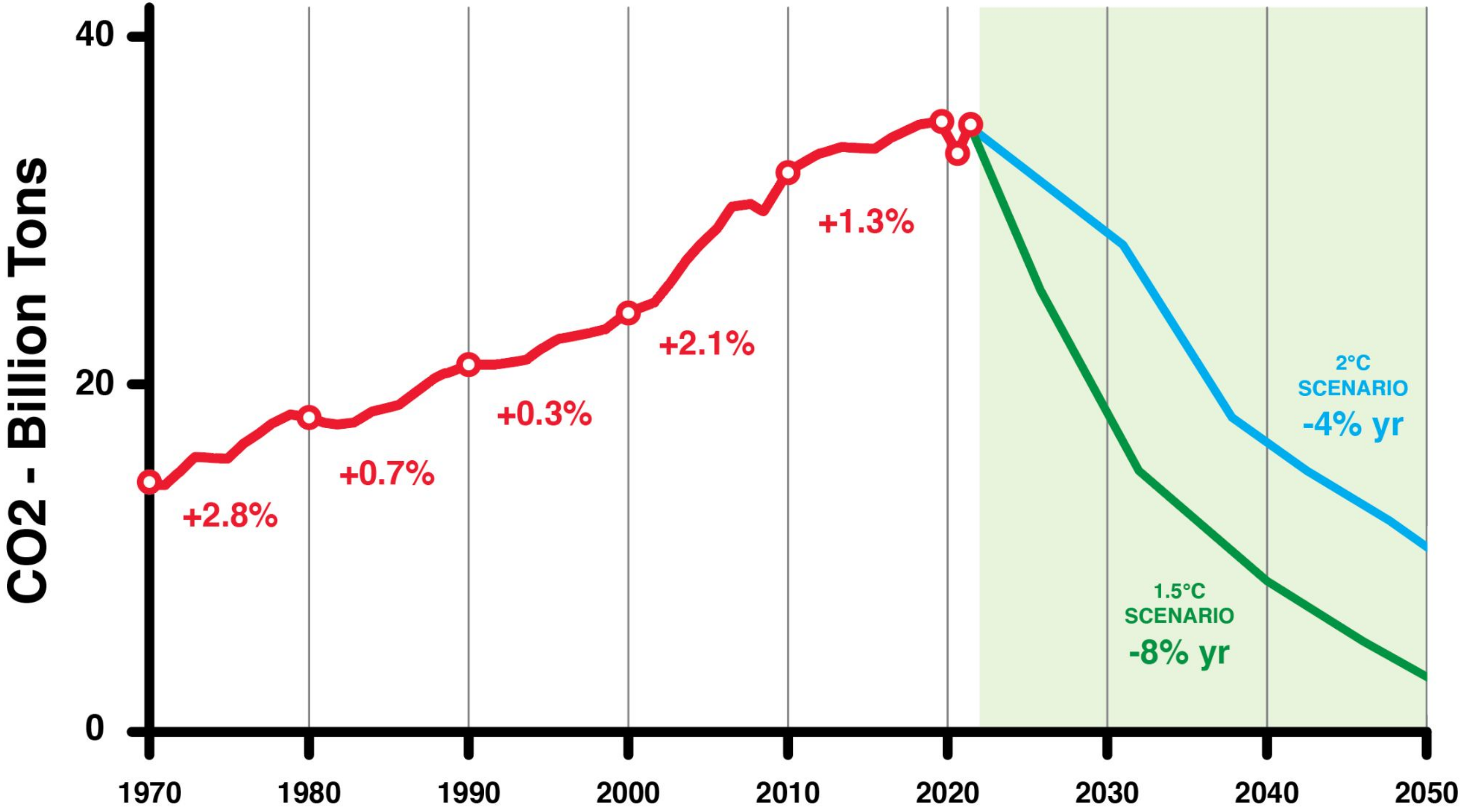
TOYOTA
RESEARCH INSTITUTE

Linda Hung
Manager, Energy & Materials Division, TRI

- TRI overview
- Battery lifetime
- Novel materials

Complex Scientific Workflows at Extreme Computational Scales
IPAM, May 2, 2023

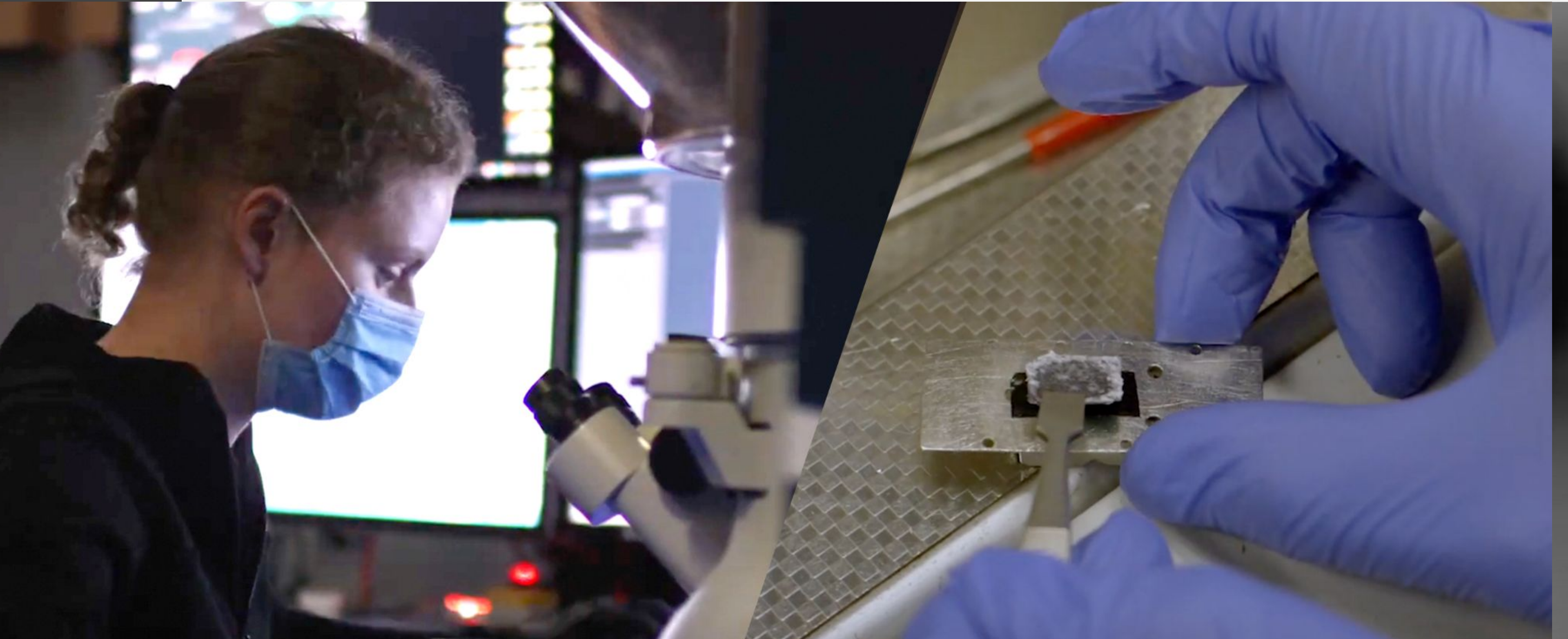
Challenge - Global carbon emissions



Challenge - Resources and sustainability

The world needs:

Better materials



The world needs:

More batteries

The world needs:

Diverse solutions

Energy and Materials Division

Materials Discovery

Better materials are needed for breakthrough performance

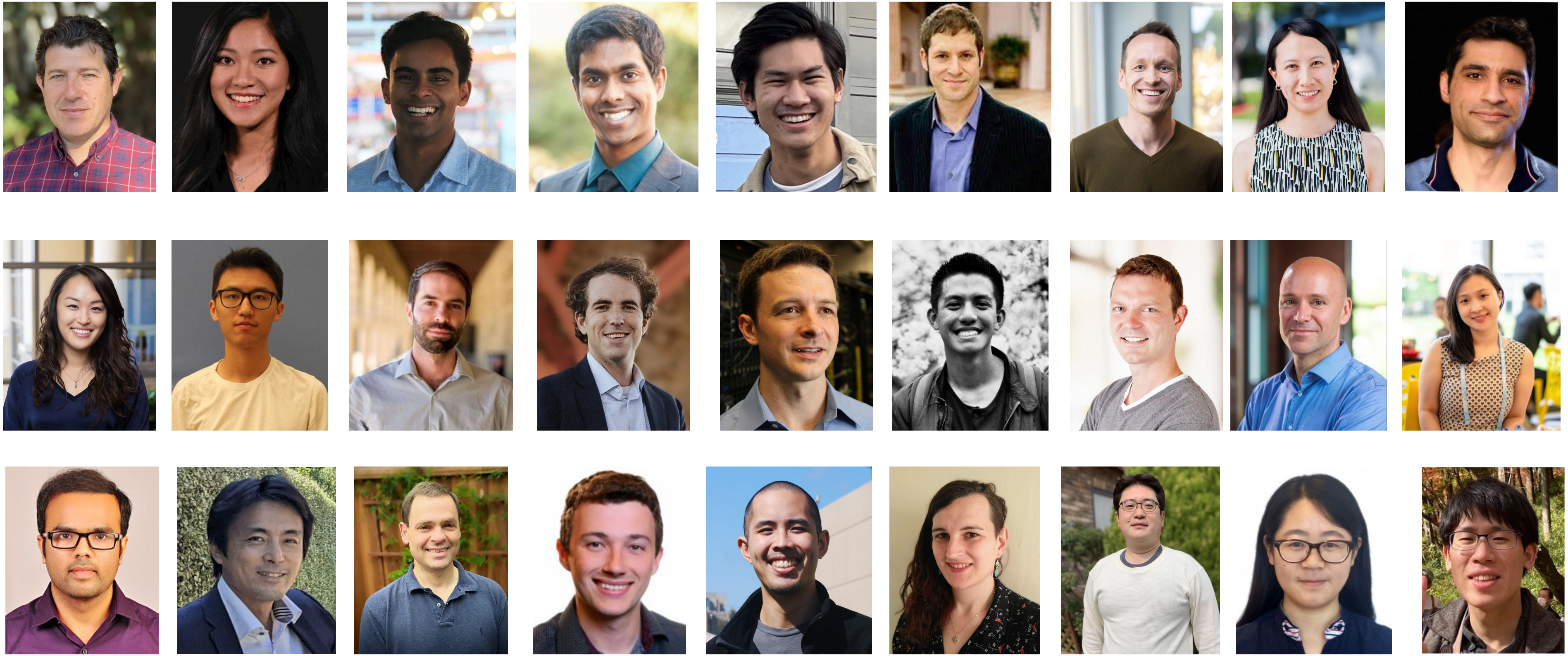
Battery Manufacturing

To make **more batteries**, we must perfect battery manufacturing.

Carbon Neutrality

Which **diverse solutions** are viable depends on economics, society, and policy.

E&M team members



TOYOTA RESEARCH INSTITUTE

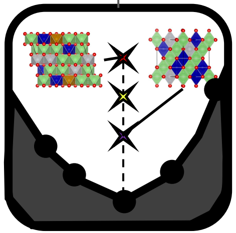

212 Northern Avenue

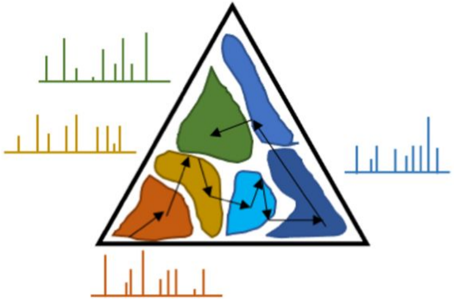
MATERIALS DISCOVERY

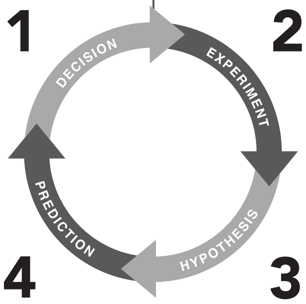
A typical project

Unique	Tailored	AI-powered	
1 *Data Generating Process*	**2** *Data*	**3** *Analysis*	**4** *Feedback*

At universities, national labs, or internal to TRI







Applications: batteries, fuel cells, fundamental materials

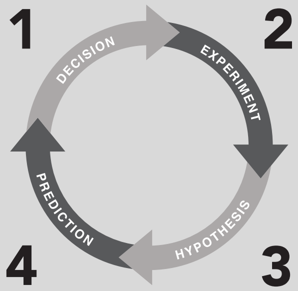
TOYOTA RESEARCH INSTITUTE

Our tools for compute

Challenges with accelerating materials discovery

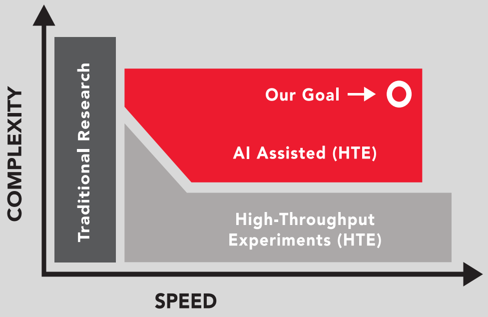
Efficiency is too low

Our approach
Automation & closed loop discovery.



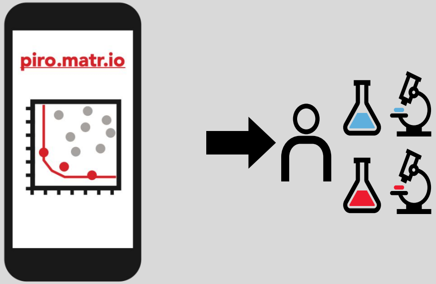
Experimental feedback is very difficult

Our approach
AI powered data analysis.



Lack of theory to enable design

Our approach
Actionable input to experiments

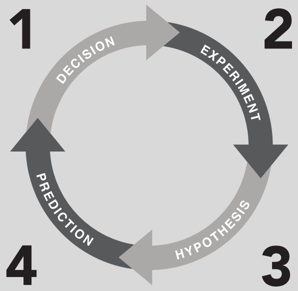


TOYOTA RESEARCH INSTITUTE

Challenges with accelerating materials discovery

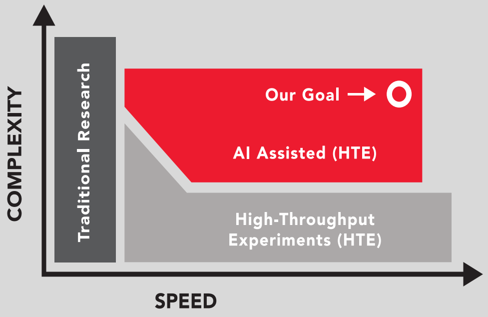
Efficiency is too low

Our approach
Automation & closed loop discovery.



Experimental feedback is very difficult

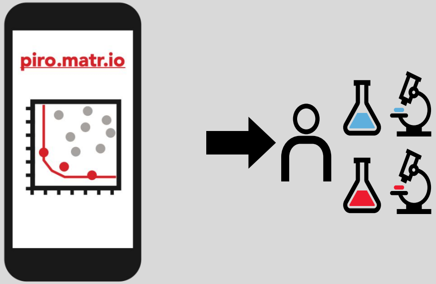
Our approach
AI powered data analysis.



Battery lifetime

Lack of theory to enable design

Our approach
Actionable input to experiments



Novel materials


Shijing Sun


Amalie Trewartha

Battery state of health

Data-driven and physics-driven approaches

TOYOTA RESEARCH INSTITUTE

How long do we expect a battery to last?

Chemistry

History

Environment

… infinite ways to charge and discharge (use), leading to **diverse** lifespans

Electrical cycling testing provides insight at all stages of the battery life cycle

Cell R&D → Cell Production → Module & Pack Design → System Integration → Consumer Use → 2nd Life & Recycling

- **Cell R&D:** Rapid capacity tests provide **feedback and guidance** to innovation in new battery chemistry
- **Module & Pack Design:** Tight production standards and **quality control** aim to prevent bad batteries from being used
- **Consumer Use:** BMS aims to **prevent failure** from environment and operator behavior through monitoring
- **2nd Life & Recycling:** Cycling data **diagnoses state of health** of batteries to judge suitability for 2nd life

… traditional multi-factor systems optimization requires testing **until end of life**,
This process takes **years**.

TOYOTA RESEARCH INSTITUTE

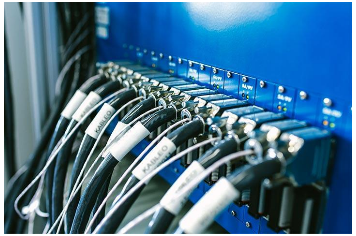
Challenges of battery cycling data

Testing systems are designed for *low-throughput*, *manual* testing

- Each testing system has its own format
- Metadata is vitally important but not stored
- They may not distinguish cycle types - e.g. Diagnostics vs accelerated aging
- Time-series is not necessarily the most helpful format for every problem

- Traditional battery cycling takes **months per battery**
- Feedback and optimization takes **years**

	test_time	date_time	step_time	step_index	cycle_index	current	voltage	charge_capacity	c
1	0.000100	1.494646e+09	-1780.003052	0	0	0.0	3.301788	0.0	
2	9.998300	1.494646e+09	-1770.004883	0	0	0.0	3.301751	0.0	
3	20.002001	1.494646e+09	-1760.001221	0	0	0.0	3.301784	0.0	
4	30.000601	1.494646e+09	-1750.002563	0	0	0.0	3.301805	0.0	
5	30.000700	1.494646e+09	-1750.002441	0	0	0.0	3.301805	0.0	
6	40.001900	1.494646e+09	-1740.001343	0	0	0.0	3.301790	0.0	



TOYOTA RESEARCH INSTITUTE

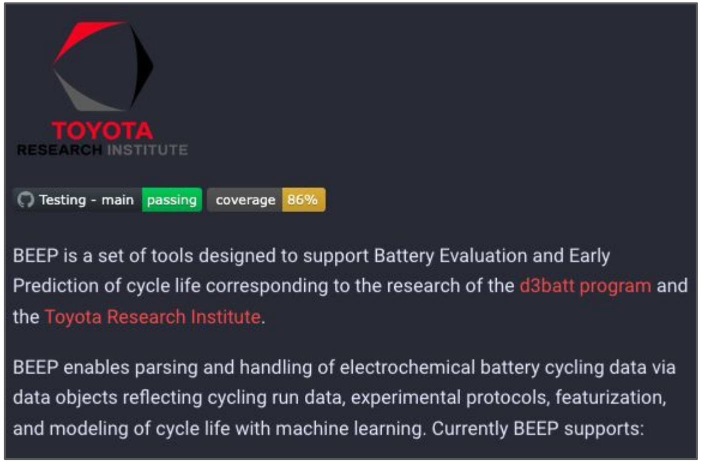
BEEP: Battery Evaluation and Early Prediction

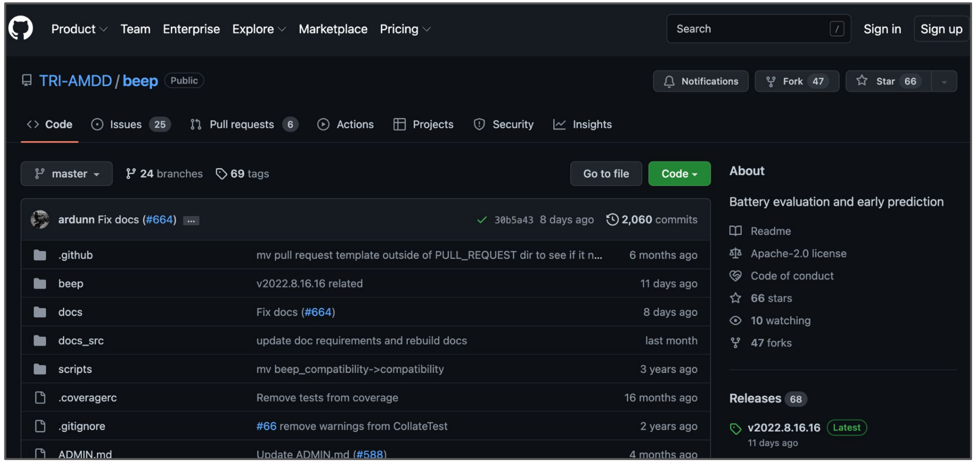
```
pip install beep
```

https://github.com/TRI-AMDD/beep

Support multiple cycler manufacturers including:

- Maccor
- Arbin
- BioLogic
- Neware
- Indigo
- Novonix (**NEW**)

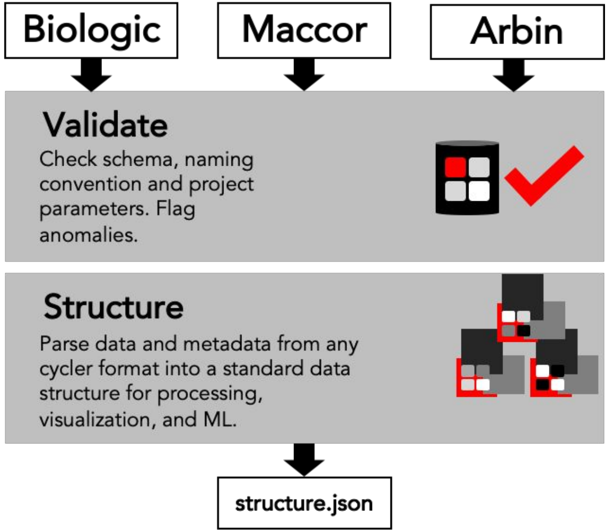


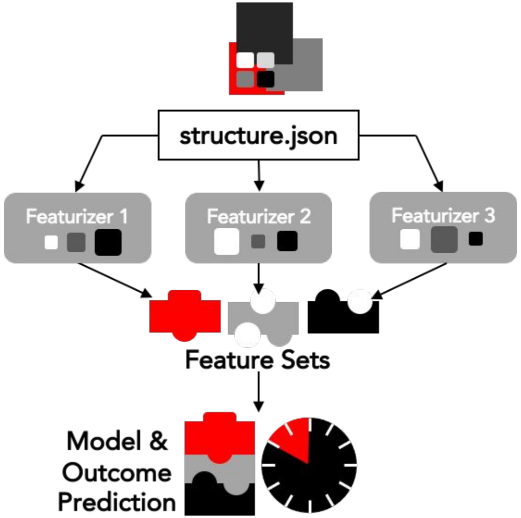


"BEEP: A Python Library for Battery Evaluation and Early Prediction," P.K. Herring *et al.*, SoftwareX **11**, 100506 (2020).

BEEP makes battery cycling data ML ready

Highlight #1 - **structuring**: process data from multiple hardware vendors into one common data structure.



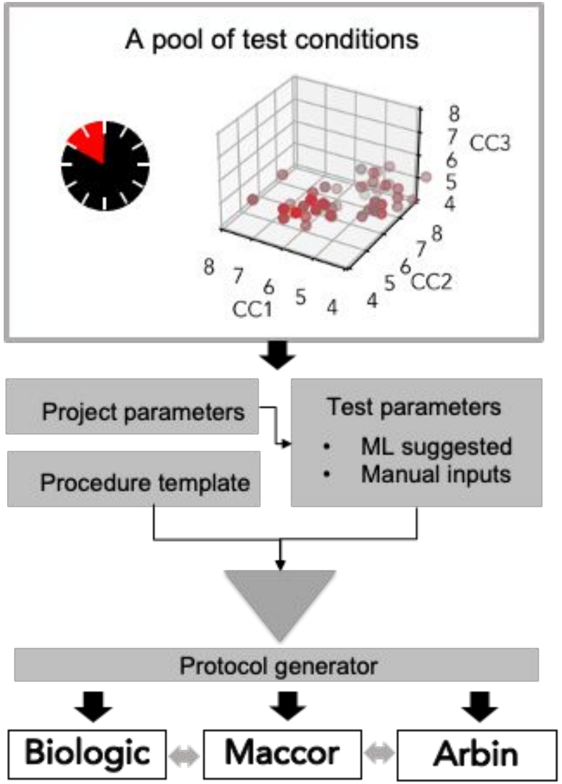


Highlight # 2 – **feature extraction**: transform structured data into ML-ready feature objects, alongside corresponding metadata.

"BEEP: A Python Library for Battery Evaluation and Early Prediction," P.K. Herring *et al.*, SoftwareX **11**, 100506 (2020).

TOYOTA RESEARCH INSTITUTE

BEEP provides programmatic scheduling of testing



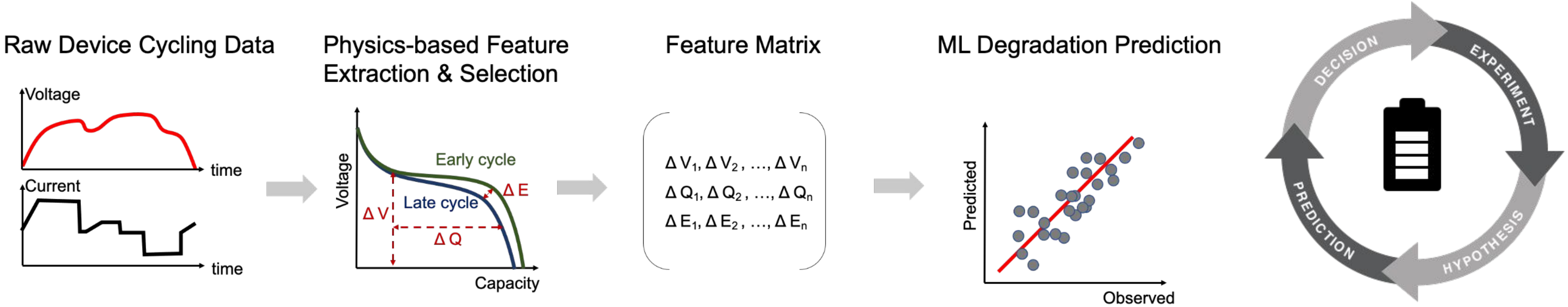
Highlight #3 - **scheduling**:

- Communicate with hardware to launch high-throughput experiments
- Reduce manual errors and expedite scheduling
- Protocols convertible between cyclers

"BEEP: A Python Library for Battery Evaluation and Early Prediction," P.K. Herring *et al.*, SoftwareX **11**, 100506 (2020).

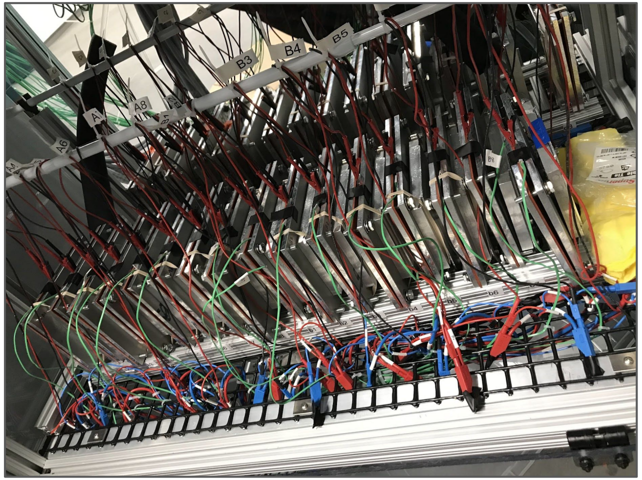
TOYOTA RESEARCH INSTITUTE

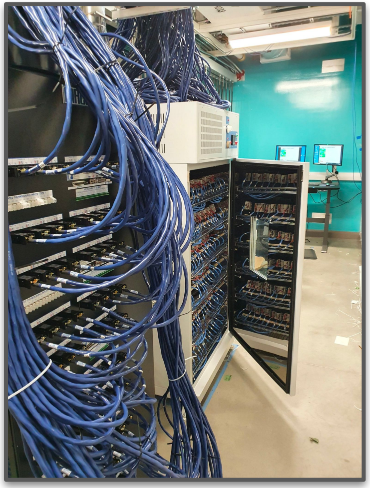
BEEP automation enables data-driven degradation prediction



- Large, diverse cycling dataset collected and processed automatically
- Machine learning based analysis: extracting subtle signals from battery cycling data to offer ***actionable, interpretable*** insights into cell internal state

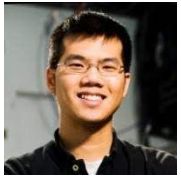
High-throughput testing at SLAC





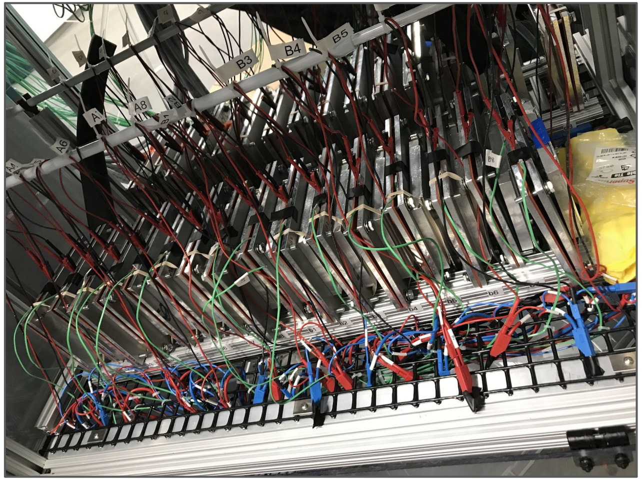
D3BATT

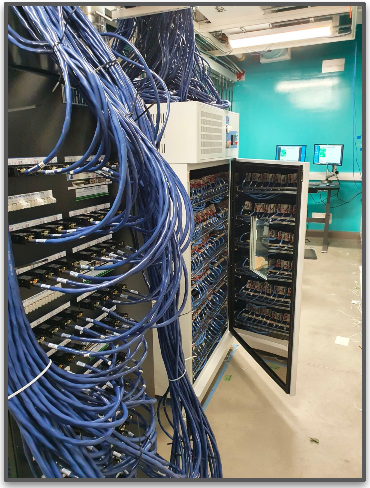
Data-Driven Design of Li-ion Batteries



Professor Will Chueh

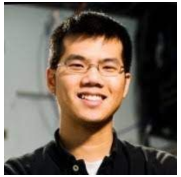
High-throughput testing at SLAC





D3BATT

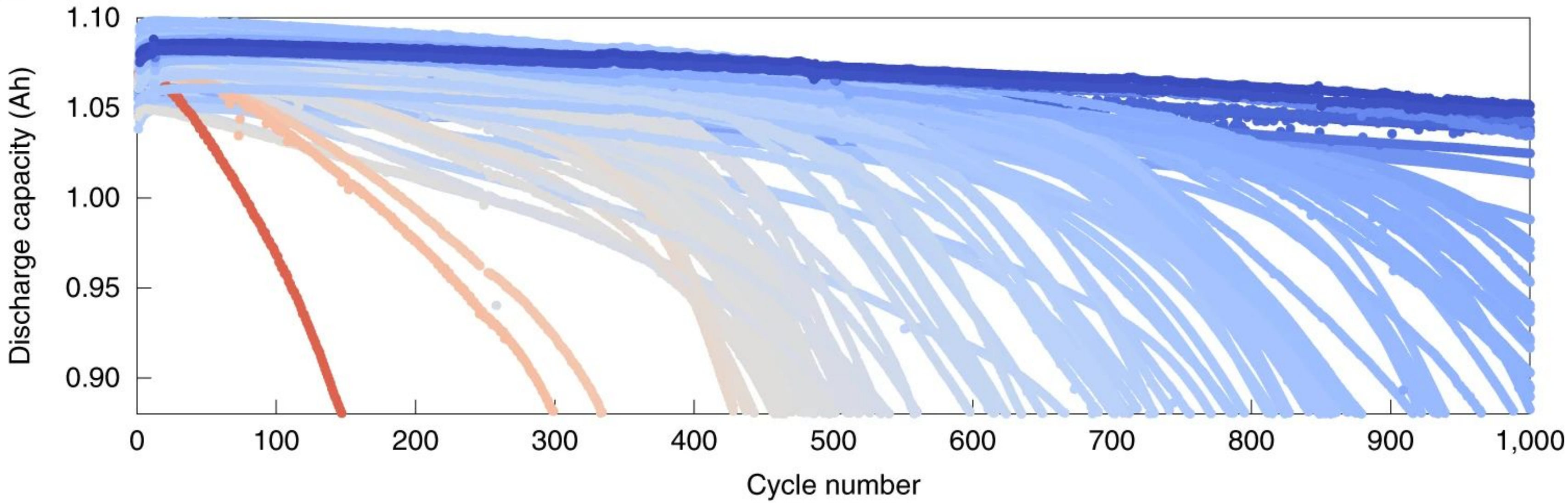
Data-Driven Design of Li-ion Batteries



Professor Will Chueh

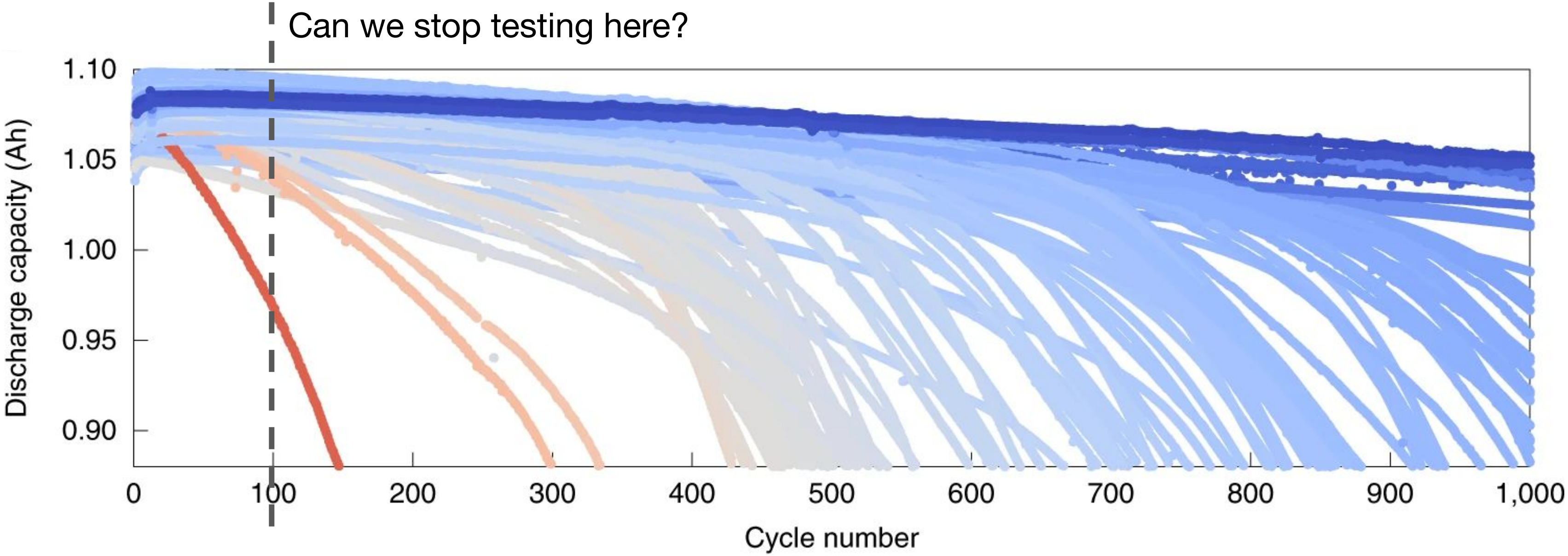
Airflow S3 Sagemaker

Challenge: full aging tests are time-consuming



TOYOTA RESEARCH INSTITUTE

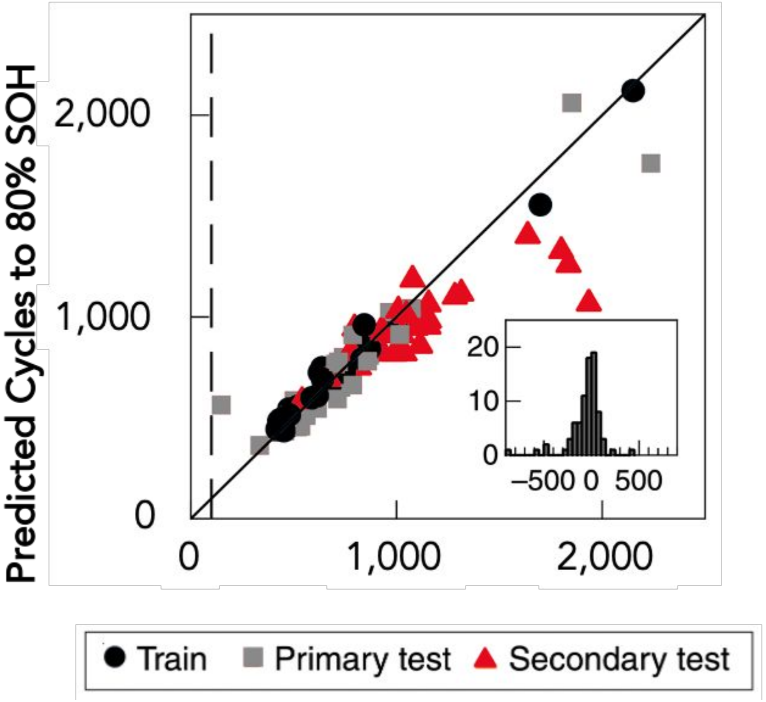
Challenge: full aging tests are time-consuming



"Data-driven prediction of battery cycle life before capacity degradation," K.A. Severson *et al.*, *Nat Energy* **4**, 383 (2019).

Previously at TRI: early-prediction reduces test time by 10x

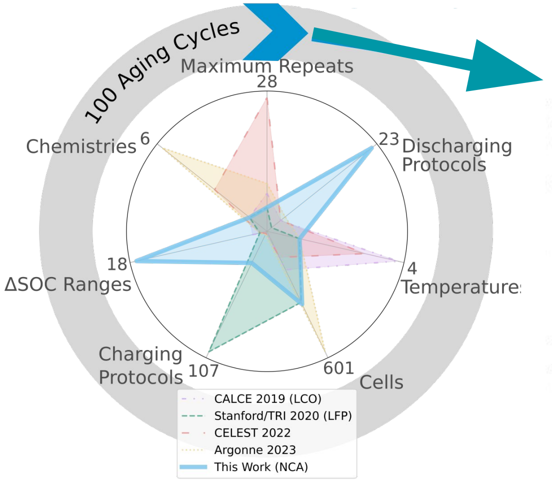
- Features from early cycle predict cycle life
- ML predictions made before observed decrease in capacity
- > 90% accuracy using first 100 aging cycles on LFP cells

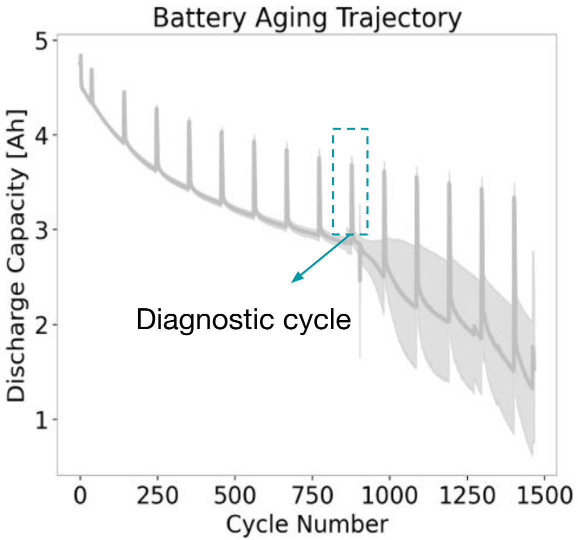


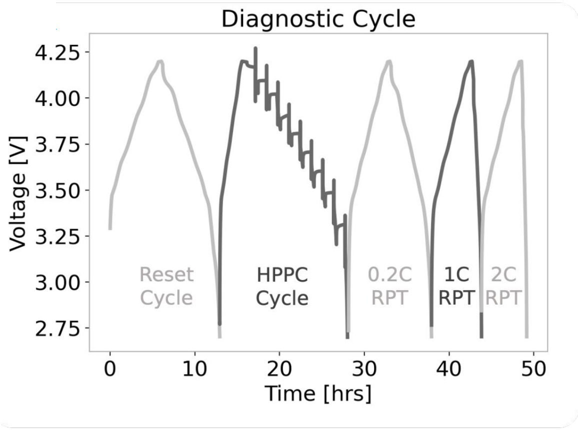
“Data-driven prediction of battery cycle life before capacity degradation,” K.A. Severson *et al.*, *Nat Energy* **4**, 383 (2019).

TOYOTA RESEARCH INSTITUTE

New dataset: diverse usage with standardized diagnostics





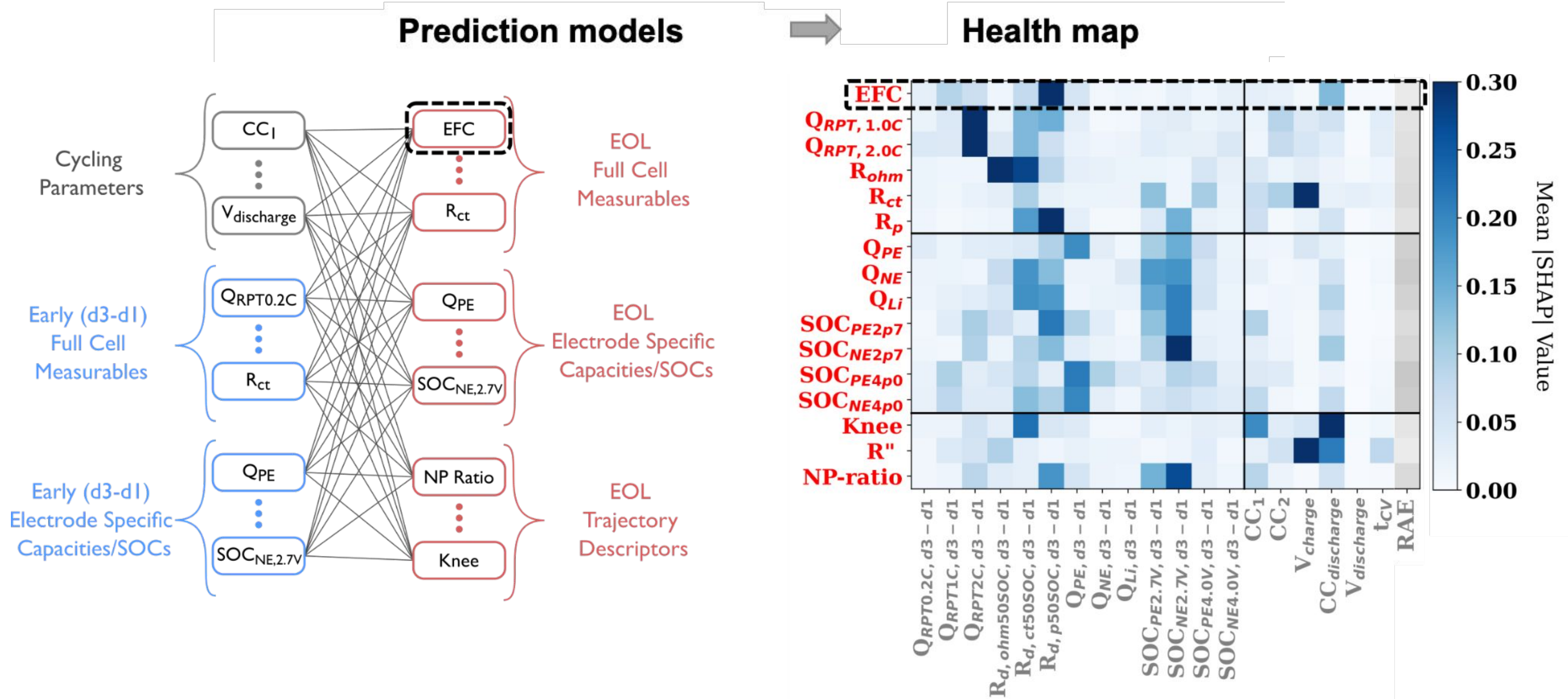


- **363 cells** tested until the end of 1st lifetime on NCA/Gr + SiOx cells[1]
- **218 cycling conditions** with diverse discharge profiles unexplored previously
- **Standardized** diagnostics enable comprehensive health evaluation and comparison

"Interpretable Data-Driven Modeling Reveals Complexity of Battery Aging", Bruis van Vlijmen *et al.*, DOI: 10.26434/chemrxiv-2023-zdl2n (2023).

TOYOTA RESEARCH INSTITUTE

Early-prediction using physically meaningful features



Health map generated from early prediction models provide actionable insights to battery design and optimization

"Interpretable Data-Driven Modeling Reveals Complexity of Battery Aging", Bruis van Vlijmen *et al.*, DOI: 10.26434/chemrxiv-2023-zdl2n (2023).

Battery state of health - continuing challenges

Efficiency is too low

- Early prediction for cycling
- Efficient / fast simulation

Experimental feedback is very difficult

- Datasets: other points of battery lifecycle
- Datasets: diverse battery chemistries
- Characterization

Lack of theory to enable design

- Models that accounting for relevant processes in physics-based models

Joey
Montoya

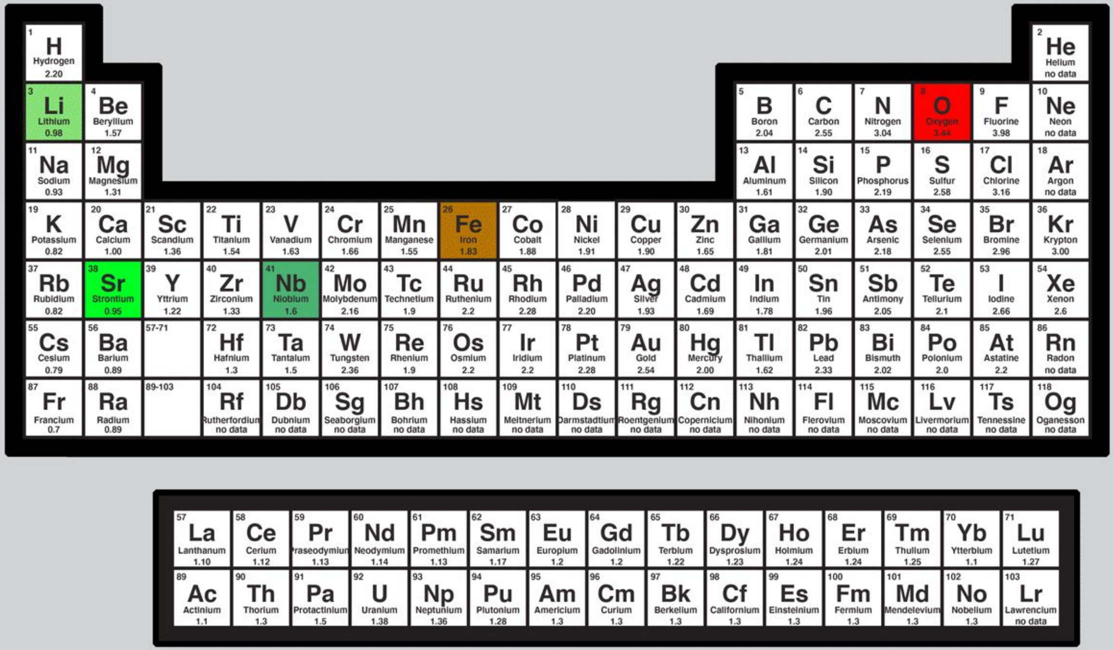
Materials discovery

From simulation to synthesis

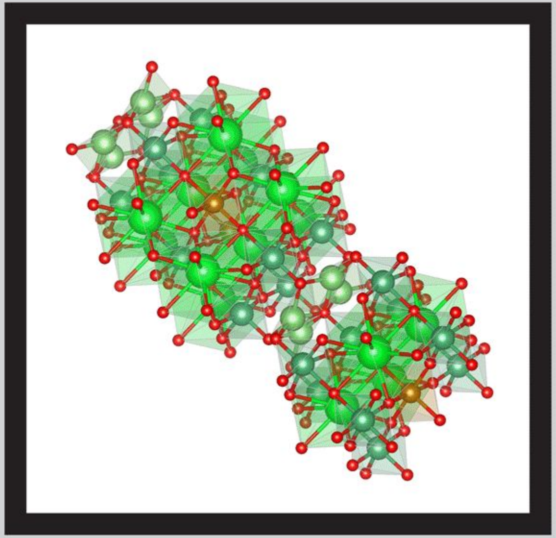
TOYOTA
RESEARCH INSTITUTE

The infinite search space

COMPOSITION

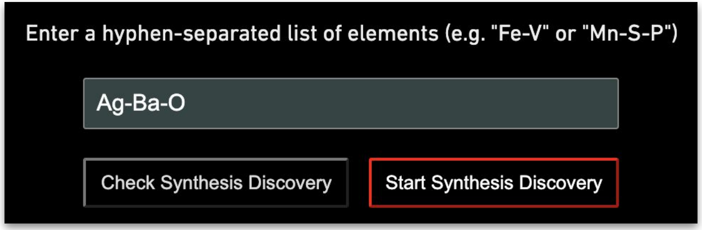


STRUCTURE

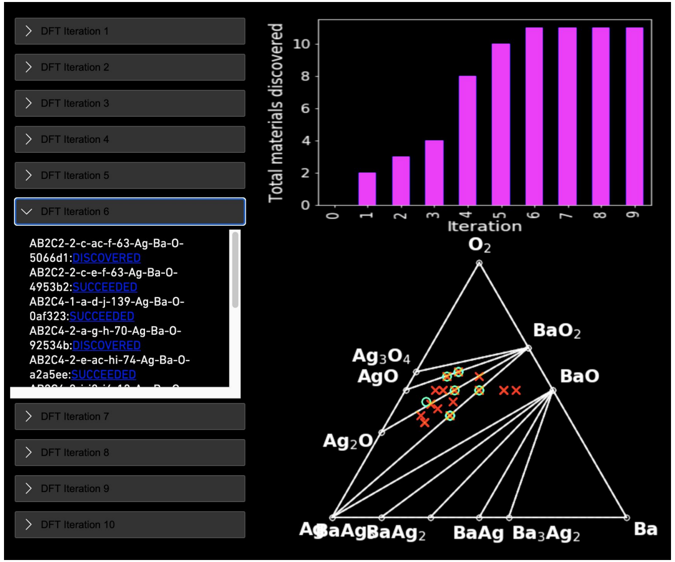


CAMD: Computational Autonomy for Materials Discovery

User **chooses chemical system**, e.g. Ag-Ba-O



CAMD **displays real-time results on frontend**

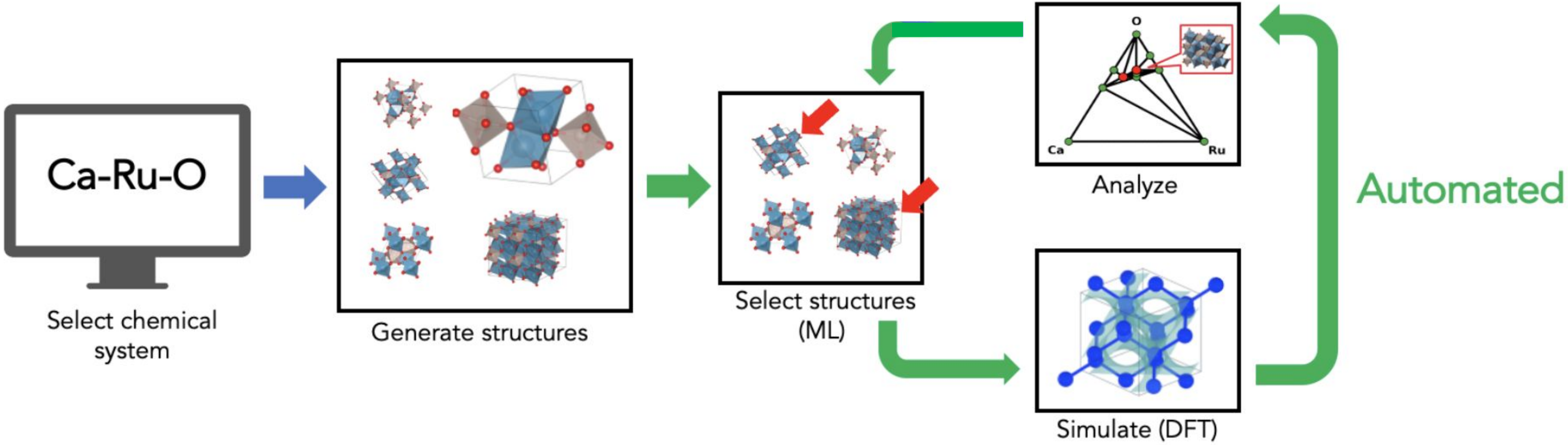


CAMD **stores structures and data**

https://github.com/TRI-AMDD/CAMD

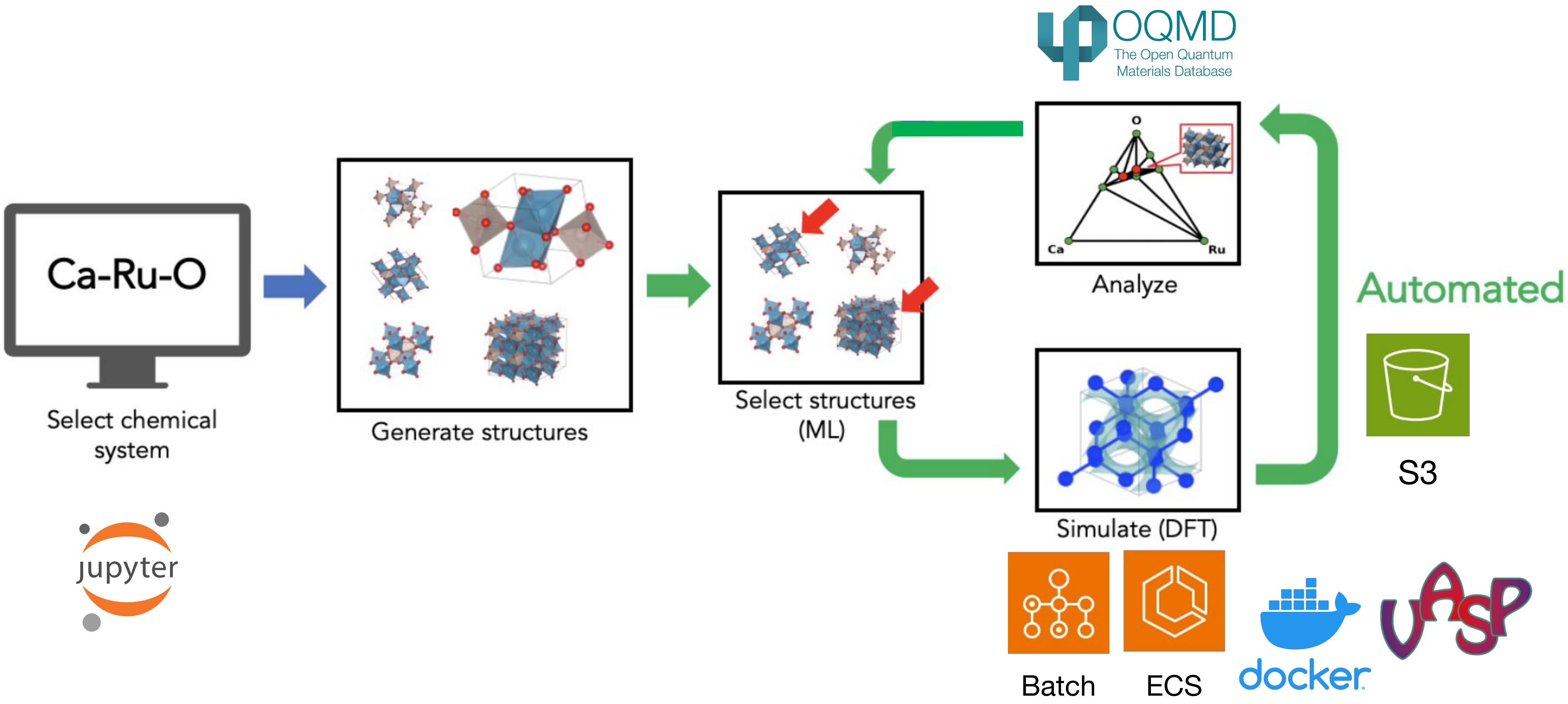
TOYOTA RESEARCH INSTITUTE

Active learning to discover materials



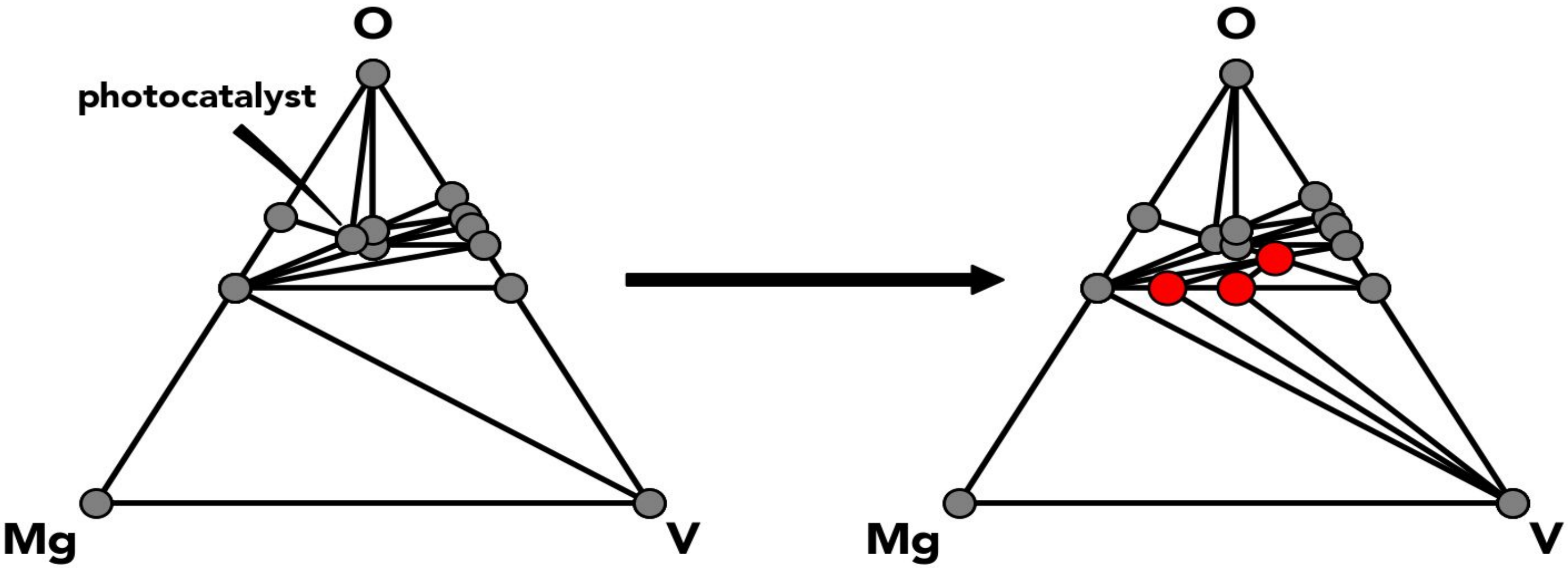
"Autonomous Intelligent Agents for Accelerated Materials Discovery" J.H. Montoya *et al.*, *Chem Sci* **11**, 8517-8532 (2020).

Active learning to discover materials



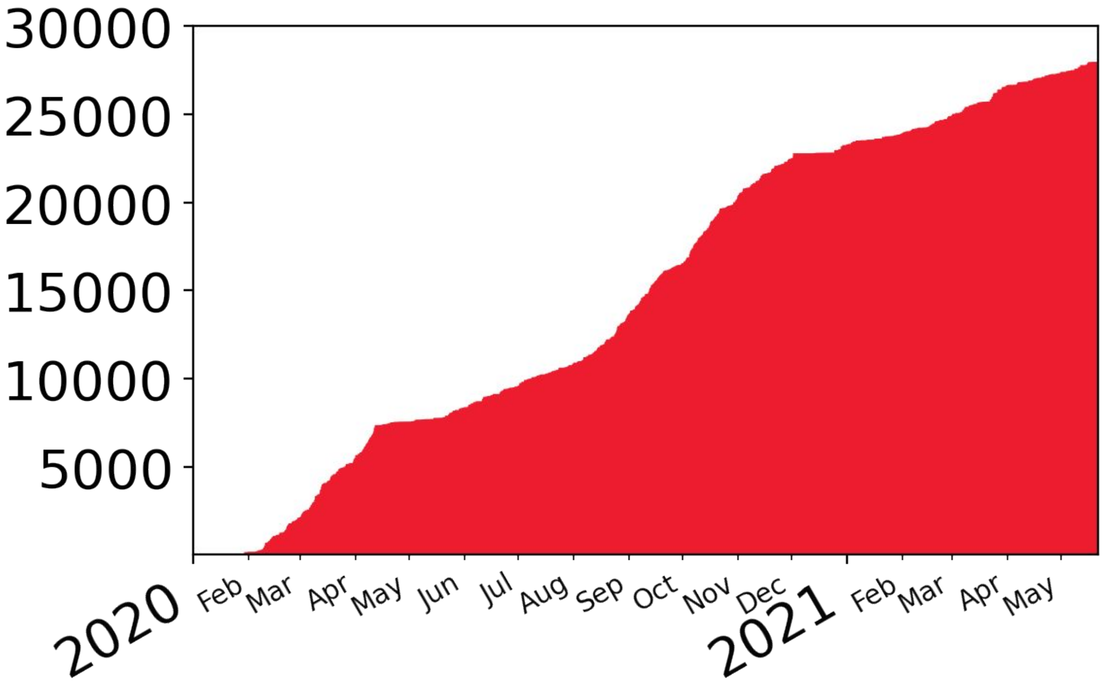
"Autonomous Intelligent Agents for Accelerated Materials Discovery" J.H. Montoya *et al.*, *Chem Sci* **11**, 8517-8532 (2020).

CAMD finds new (meta)stable structures



“Novel inorganic crystal structures predicted using autonomous simulation agents”, W. Ye, X. Lei, *et al.*, *Sci Data* **9**, 302 (2022).

CAMD finds a lot of metastable structures



https://data.matr.io/7/

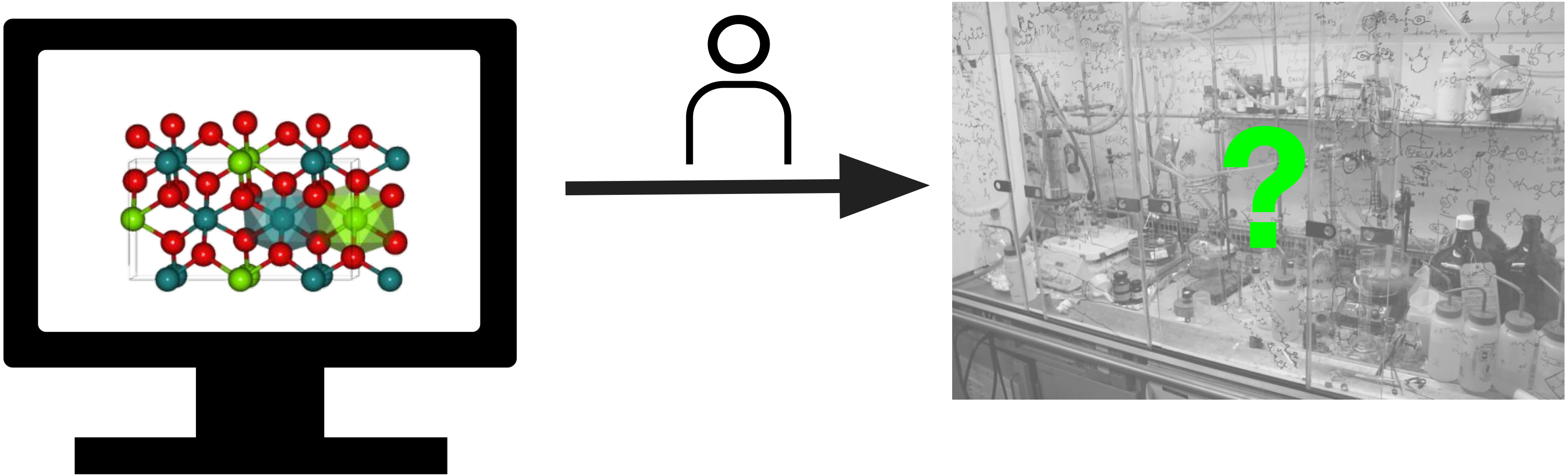
CAMD discovers 2-3 new structures within 200 meV of the convex hull per hour

CAMD spends $3.00 per metastable structure

CAMD's infrastructure is scalable

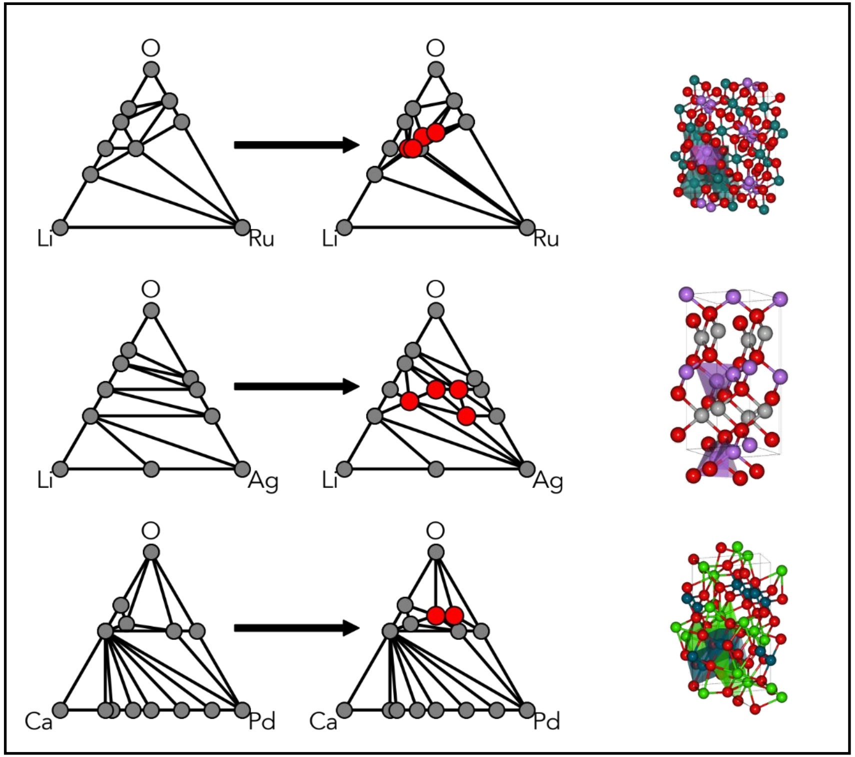
"Novel inorganic crystal structures predicted using autonomous simulation agents", W. Ye, X. Lei, *et al.*, *Sci Data* **9**, 302 (2022).

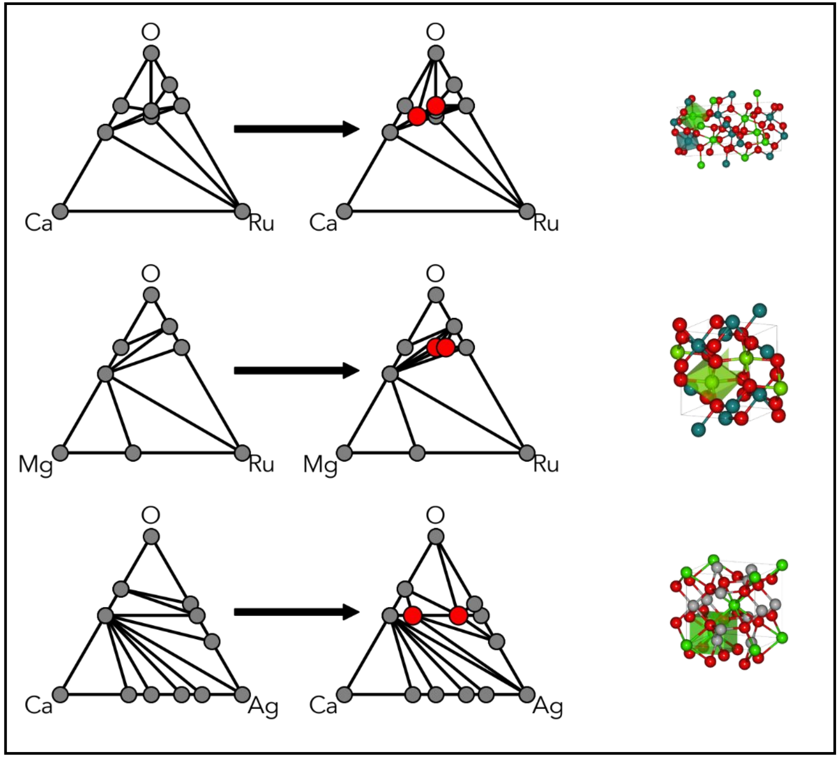
We tried to make CAMD materials



“Computer-assisted discovery and rational synthesis of ternary oxides”, J.H. Montoya *et al.*, DOI: [10.26434/chemrxiv-2023-n4pz9](https://doi.org/10.26434/chemrxiv-2023-n4pz9) (2023).

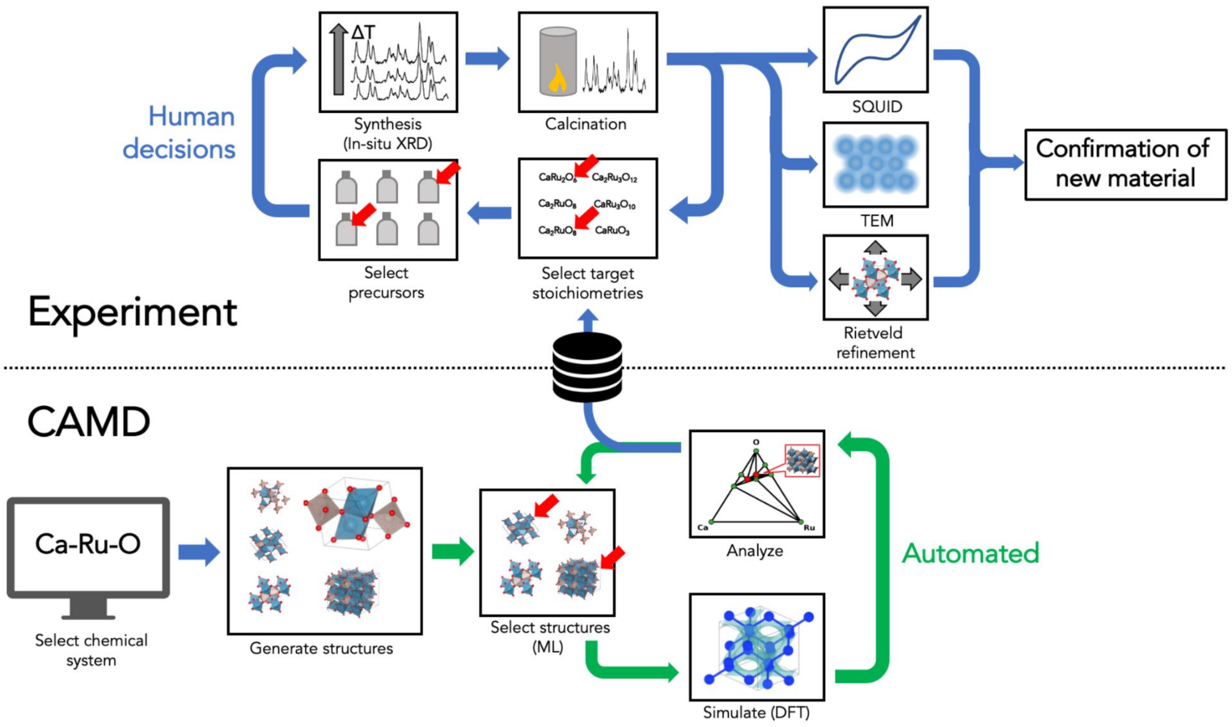
Six CAMD inspired systems were selected





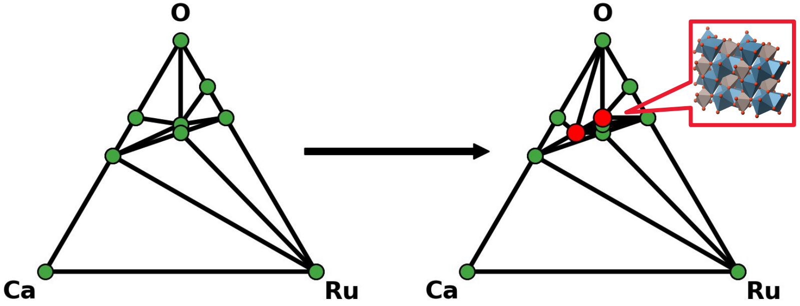
“Computer-assisted discovery and rational synthesis of ternary oxides”, J.H. Montoya *et al.*, DOI: 10.26434/chemrxiv-2023-n4pz9 (2023).

Materials discovery workflow



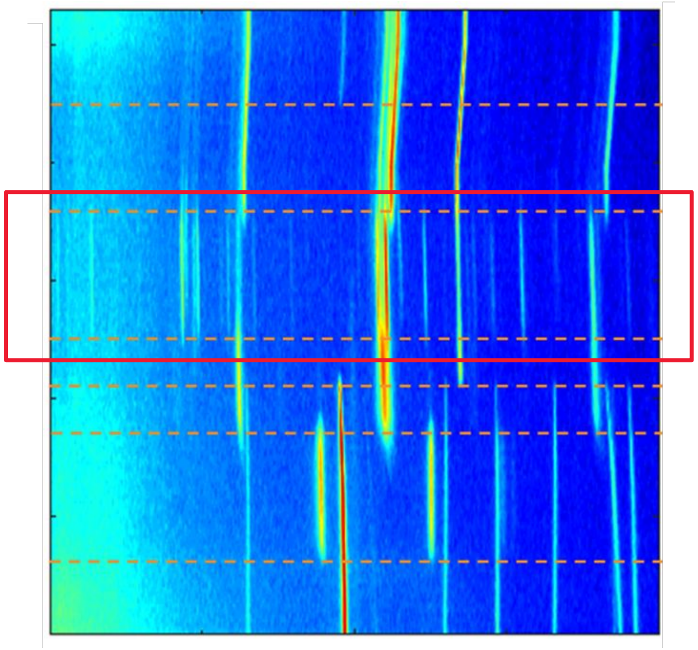
"Computer-assisted discovery and rational synthesis of ternary oxides", J.H. Montoya *et al.*, DOI: 10.26434/chemrxiv-2023-n4pz9 (2023).

A new material predicted by CAMD was synthesized



CAMD predicted a novel material containing Ru, a common element in catalyst materials.

In-situ, variable temp XRD experiments confirm a new phase exists, appears 700-1100°C



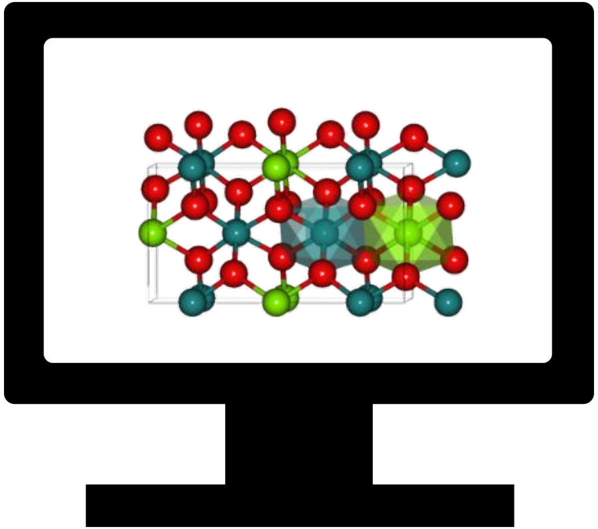
in-situ XRD

“Computer-assisted discovery and rational synthesis of ternary oxides”, J.H. Montoya *et al.*, DOI: 10.26434/chemrxiv-2023-n4pz9 (2023).

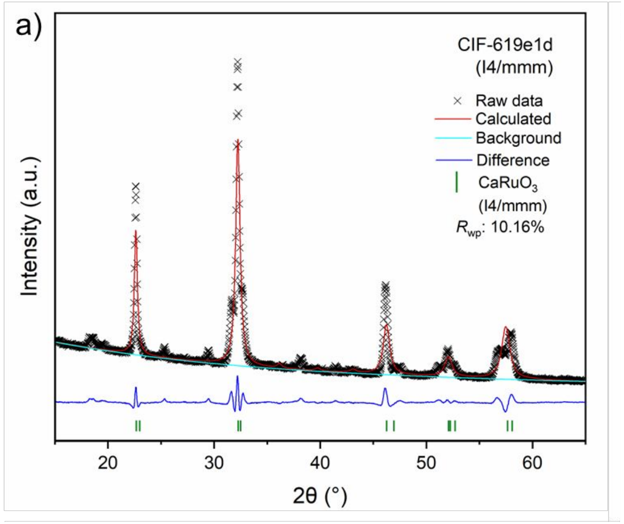
Other examples start to sound the same

10^4 designed on the computer

10^0 confirmed in the lab



~2 years pass



Computation-mediated discovery - challenges

Efficiency is too low

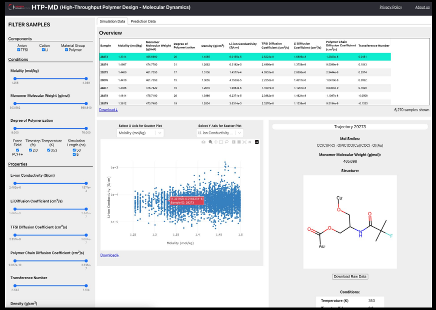
- Automating expt and expt analyses

Experimental feedback is very difficult

- Datasets: dark data
- Data-driven analyses

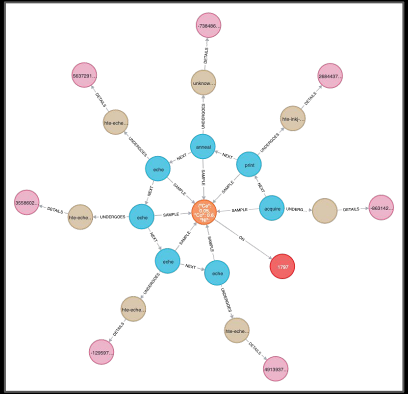
Lack of theory to enable design

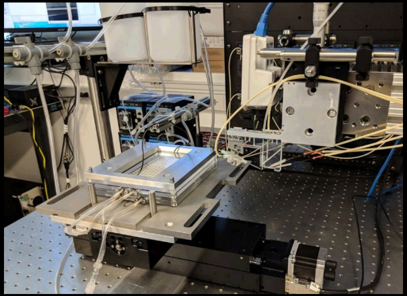
- Datasets: defect structures
- Datasets: kinetic properties
- Integrate data-theory
- Multiscale models



...and more research

Polymers, catalysts, ML





TOYOTA
RESEARCH INSTITUTE

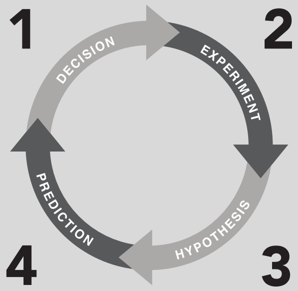
Outlook

Bridging the gap between experiment & simulation

Challenges with accelerating materials discovery

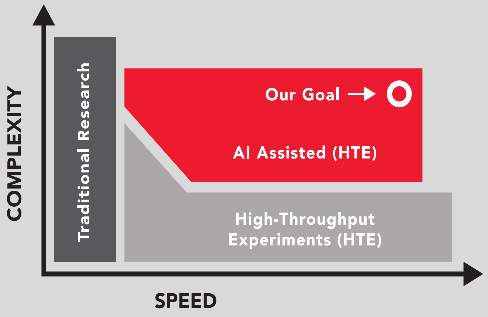
Efficiency is too low

Our approach
Automation & closed loop discovery.



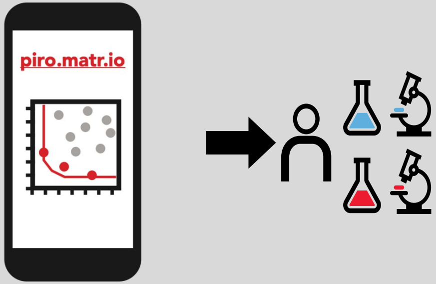
Experimental feedback is very difficult

Our approach
AI powered data analysis.



Lack of theory to enable design

Our approach
Actionable input to experiments



TOYOTA RESEARCH INSTITUTE

Outlook

- Highlights
  - Battery lifetime
  - Novel materials
- Workflow software allows better integration of compute and experiment
- Industry research uses minimal or mid-scale compute -- but also will benefit from exascale compute to address fundamental questions

Thank You