How is Computation reshaping Materials Chemistry?

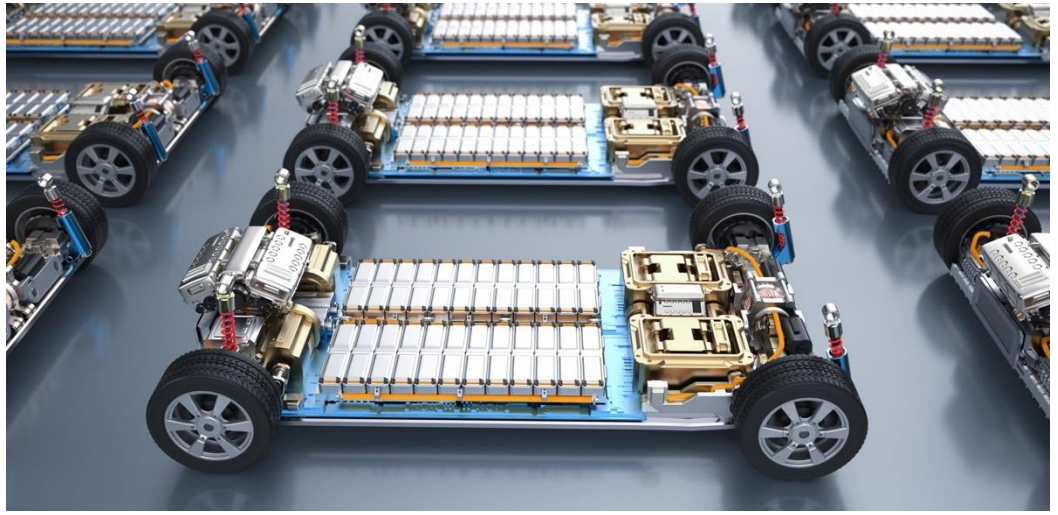
From trial and error to autonomous laboratories

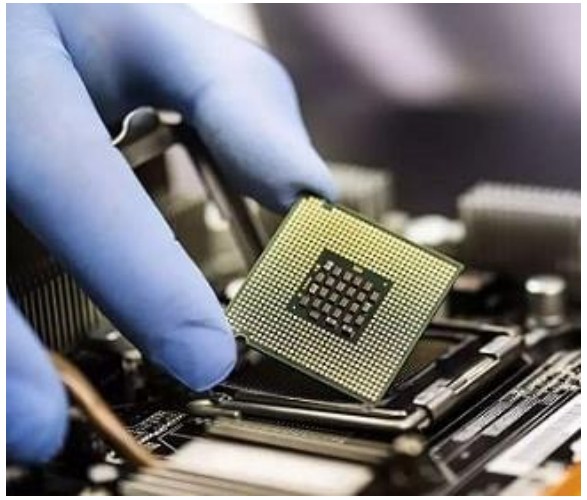
Graduate Student Seminar Talk

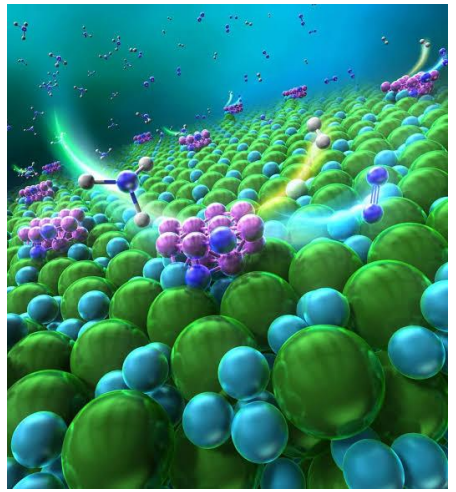
Presented by – Nikhil Singh (2023CYZ8251)
Supervisor – Prof. Dibyajyoti Ghosh

Almost every modern technology is waiting on a material

	Waiting for?
Electric Cars	Better Battery electrolyte
Solar Power	Cheaper, Stable Absorbers
Faster Chips	New semiconductors and dielectrics
Green Ammonia	More efficient Catalyst materials







The bottleneck is not the idea; it is discovering the material that makes it possible.

I. Old age methods of Materials Discovery?

The age of Luck! Or Bad-Luck?

Accidents and brute-force testing

Vulcanised rubber

Mauveine

Teflon

Post-it adhesive

When luck ran out, chemists simply tried everything

~6,000 Plant and Fibre materials carbonised and tested by Edison's team in the hunt for a lamp filament

Hit and Trial !!

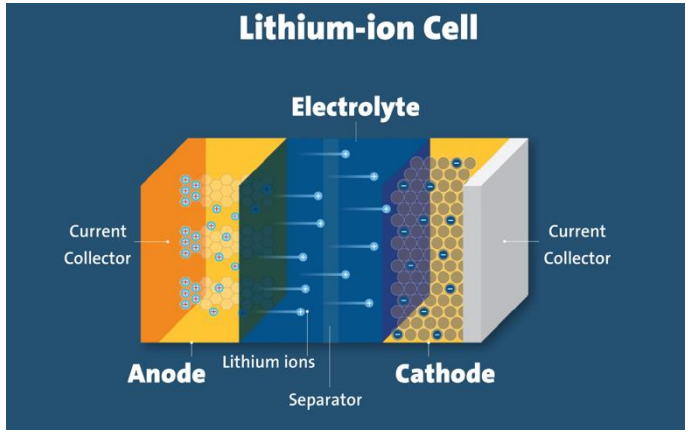
- Guess a composition from chemical intuition
- Grind, heat, wait, and hope a single phase forms
- Measure the property you care about
- Change one variable and repeat, for years

I. Old age methods of Materials Discovery?

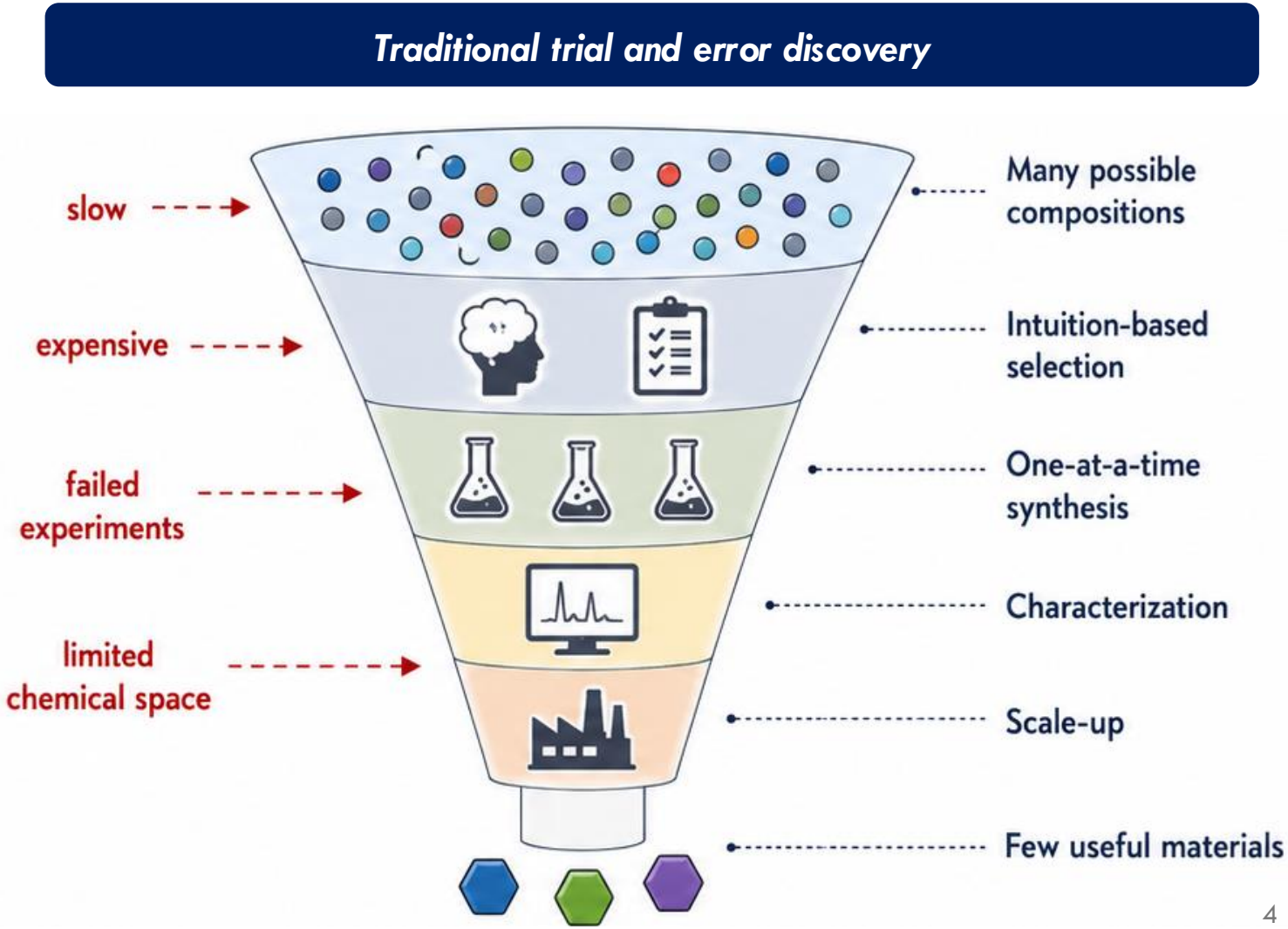
We paid in time and money. But mostly we paid in the materials we never made.

Typical time from discovering a new material to a shipping product: **10 – 20 years**

A realistic rate of genuinely new compounds made and characterized : **2-3 materials per PhD student**



Lithium-ion chemistry was demonstrated in the 1970s and reached consumers in 1991.



II. Age of Theory and Computation

In principle, chemistry was solved in 1926

$$H\Psi = E\Psi$$

Give me the atoms and their positions, and this equation contains every property of the material.

…and in practice, it was not.

Kinetic Energy Terms **Potential Energy Terms**

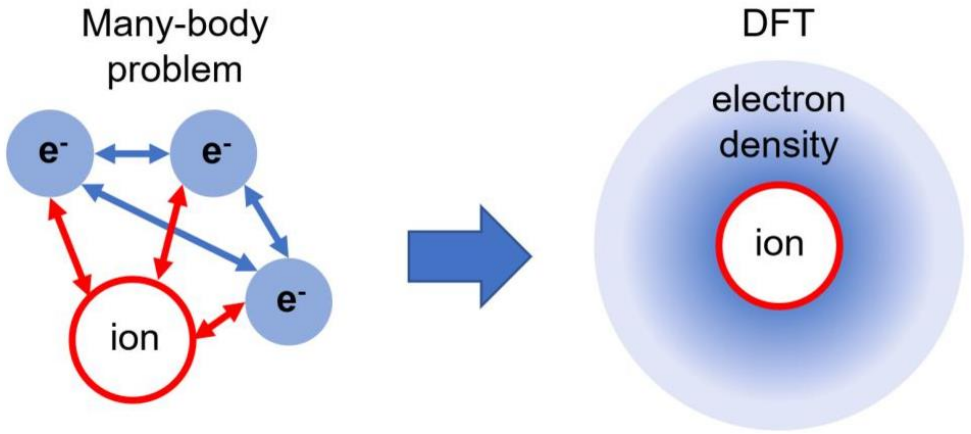
$$\hat{H} = -\boxed{\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2} - \boxed{0} - \boxed{\sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}}} + \boxed{\sum_{i=1}^{N}\sum_{j>1}^{N}\frac{1}{r_{ij}}} + \boxed{\text{Const.}}$$

KE Elec. **KE Nucl.** **Nucl.-Elec. Attract.** **Elec.-Elec. Repuls.** **Nucl.-Nucl. Repuls.**

'n' number of different theories and approximations:
(Developed over many decades)

1. Hartree Fock
2. Density Functional Theory *(Perfect Balance)*
3. Møller-Plesset Theory
4. Coupled Cluster Theory
5. Configuration Interaction

Density Functional Theory



"Stop tracking every electron. Track only the electron density"

Nobel Prize 1998 (Chemistry)

Photo from the Nobel Foundation archive.
Walter Kohn

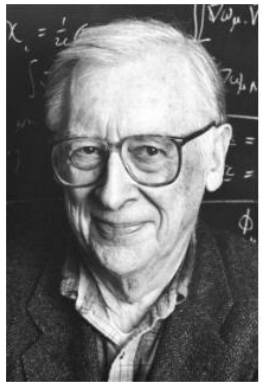


Photo from the Nobel Foundation archive.
John A. Pople

II. Age of Theory and Computation

What this gave a chemist: (Without touching lab equipment)

- **Structure** → Arrangement of atoms in a lattice
- **Stability** → Whether the compound survives or decomposes into something else
- **Band gap** → Metal/Semiconductor/Insulator?
- **Response** → Under light, strain, pressure or a field (E or M)

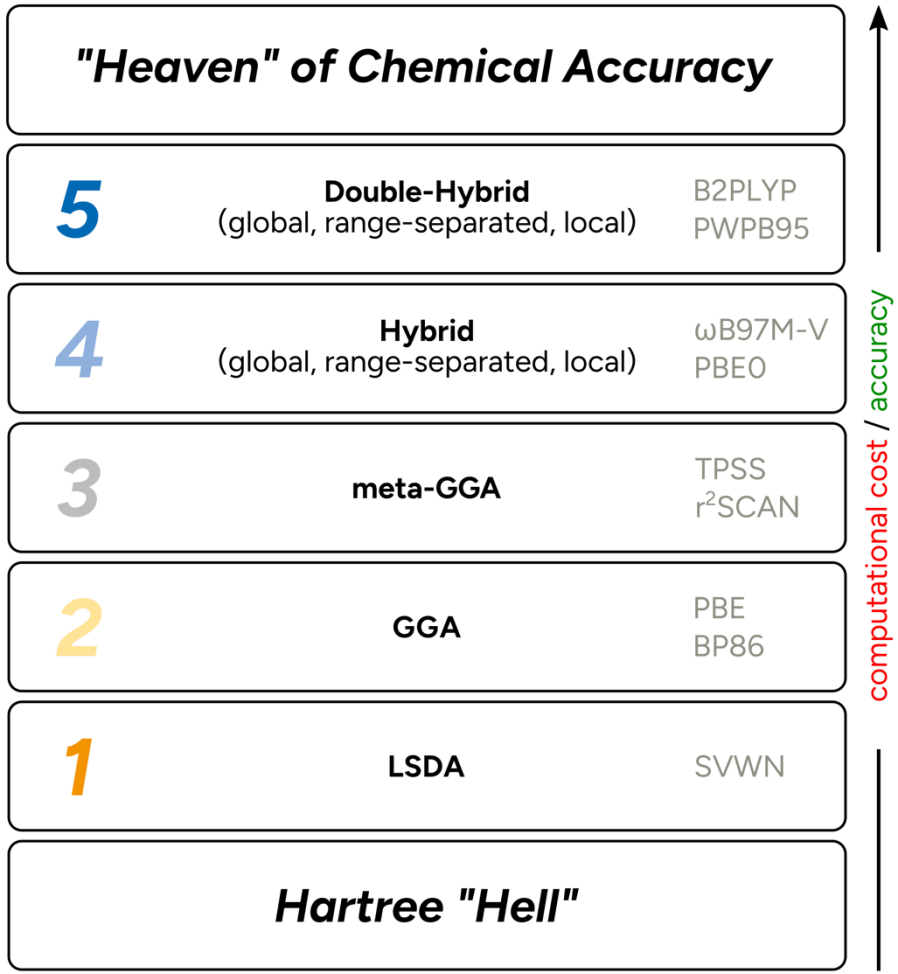
VASP

QUANTUMESPRESSO

CP2K

Gaussian, Inc.

Accuracy is purchased with computer time



https://orca-manual.mpi-muelheim.mpg.de/contents/modelchemistries/DensityFunctionalTheory.html

II. Age of Theory and Computation

PHYSICAL REVIEW VOLUME 140, NUMBER 4A 15 NOVEMBER 1965

Self-Consistent Equations Including Exchange and Correlation Effects*

W. KOHN AND L. J. SHAM
University of California, San Diego, La Jolla, California
(Received 21 June 1965)

From a theory of Hohenberg and Kohn, approximation methods for treating an inhomogeneous system of interacting electrons are developed. These methods are exact for systems of slowly varying or high density. For the ground state, they lead to self-consistent equations analogous to the Hartree and Hartree-Fock equations, respectively. In these equations the exchange and correlation portions of the chemical potential of a uniform electron gas appear as additional effective potentials. (The exchange portion of our effective potential differs from that due to Slater by a factor of $\frac{2}{3}$.) Electronic systems at finite temperatures and in magnetic fields are also treated by similar methods. An appendix deals with a further correction for systems with short-wavelength density oscillations.

PHYSICAL REVIEW B VOLUME 55, NUMBER 23 15 JUNE 1997-I

First-principles spin-polarized calculations on the reduced and reconstructed TiO_2 (110) surface

P. J. D. Lindan and N. M. Harrison
Daresbury Laboratory, Warrington WA4 4AD, United Kingdom

M. J. Gillan
Physics Department, Keele University, Staffordshire ST5 5BG, United Kingdom

J. A. White
Cavendish Laboratory, Madingley Road, Cambridge CB3 0HE, United Kingdom
(Received 21 January 1997)

PHYSICAL REVIEW B VOLUME 62, NUMBER 10 1 SEPTEMBER 2000-II

First-principles calculations of defects in oxygen-deficient silica exposed to hydrogen

Peter E. Blöchl
IBM Research, Zurich Research Laboratory, CH-8803 Rüschlikon, Switzerland
(Received 14 January 2000)

Hydrogen-related defects and oxygen vacancies in silica are analyzed using first-principles density-functional calculations. Energetics, structures, charge-state levels, and hyperfine parameters are determined. These calculations identify the hydrogen bridge related to the E_4' center as the defect responsible for the stress-induced leakage current, a forerunner of dielectric breakdown of gate oxides in transistors.

2572 *J. Phys. Chem. A* **1998,** *102,* 2572−2578

Accurate Method for Obtaining Band Gaps in Conducting Polymers Using a DFT/Hybrid Approach

U. Salzner,*,† P. G. Pickup, and R. A. Poirier
Department of Chemistry, Memorial University of Newfoundland, St. John's, NF A1B 3X7, Canada

J. B. Lagowski
Department of Physics and Physical Oceanography, Memorial University of Newfoundland, St. John's, NF A1B 3X7, Canada

Received: May 16, 1997; In Final Form: August 28, 1997

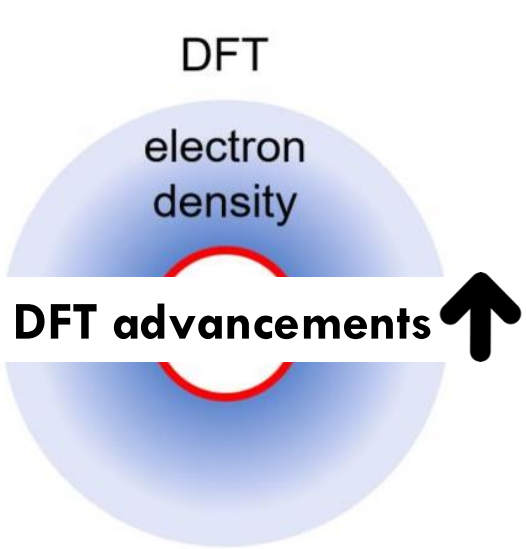
II. Age of Theory and Computation

All this was made possible due to computational advancements over the past two decades!

Isolated Computational studies

3/4 materials per study

Over 20 years



10^4 - 10^5

Materials/study

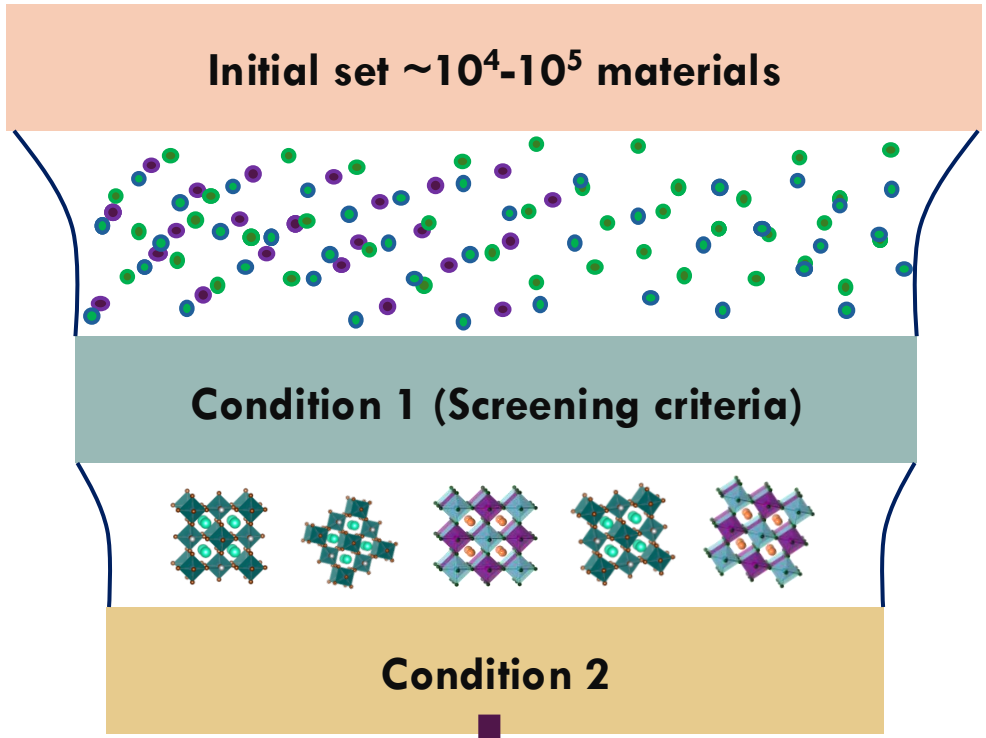
"Lineshine" at NSC, Shenzhen

"El Capitan" at LLNL, USA

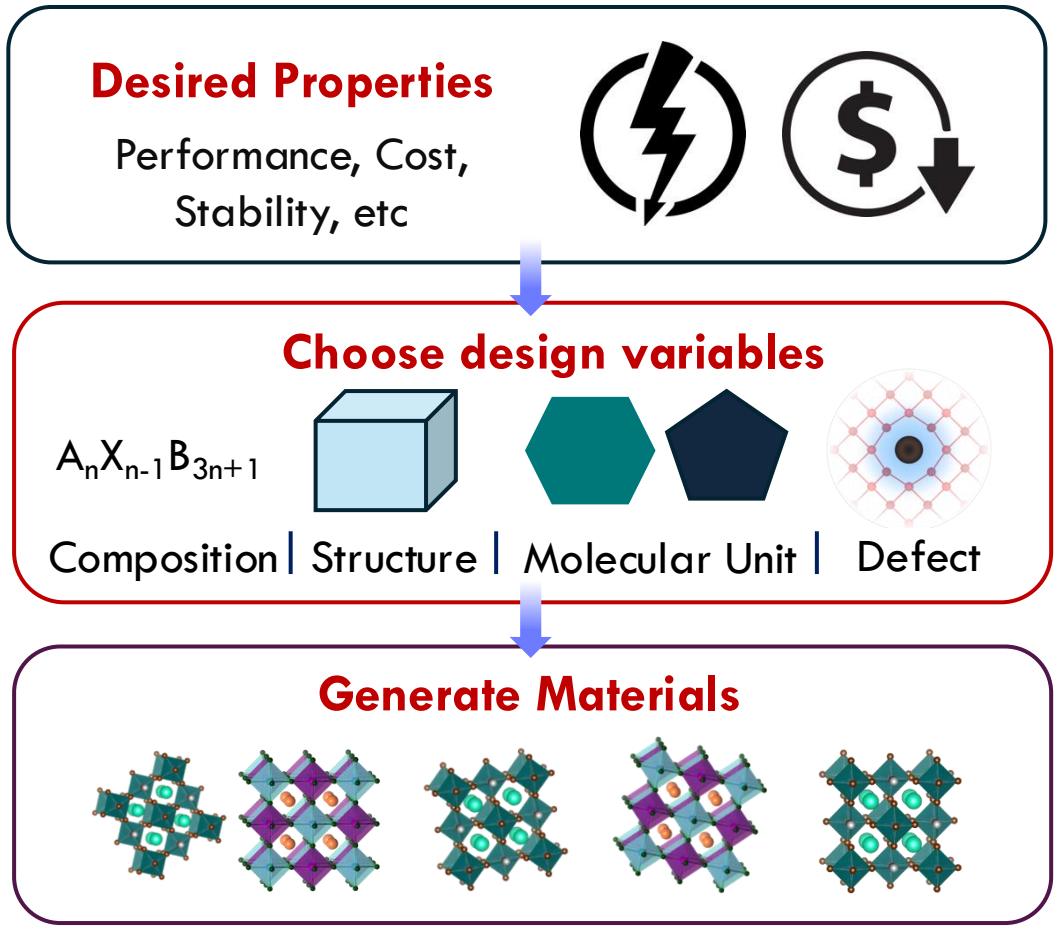
III. Modern approaches for materials discovery

Two different approaches

Top-Down

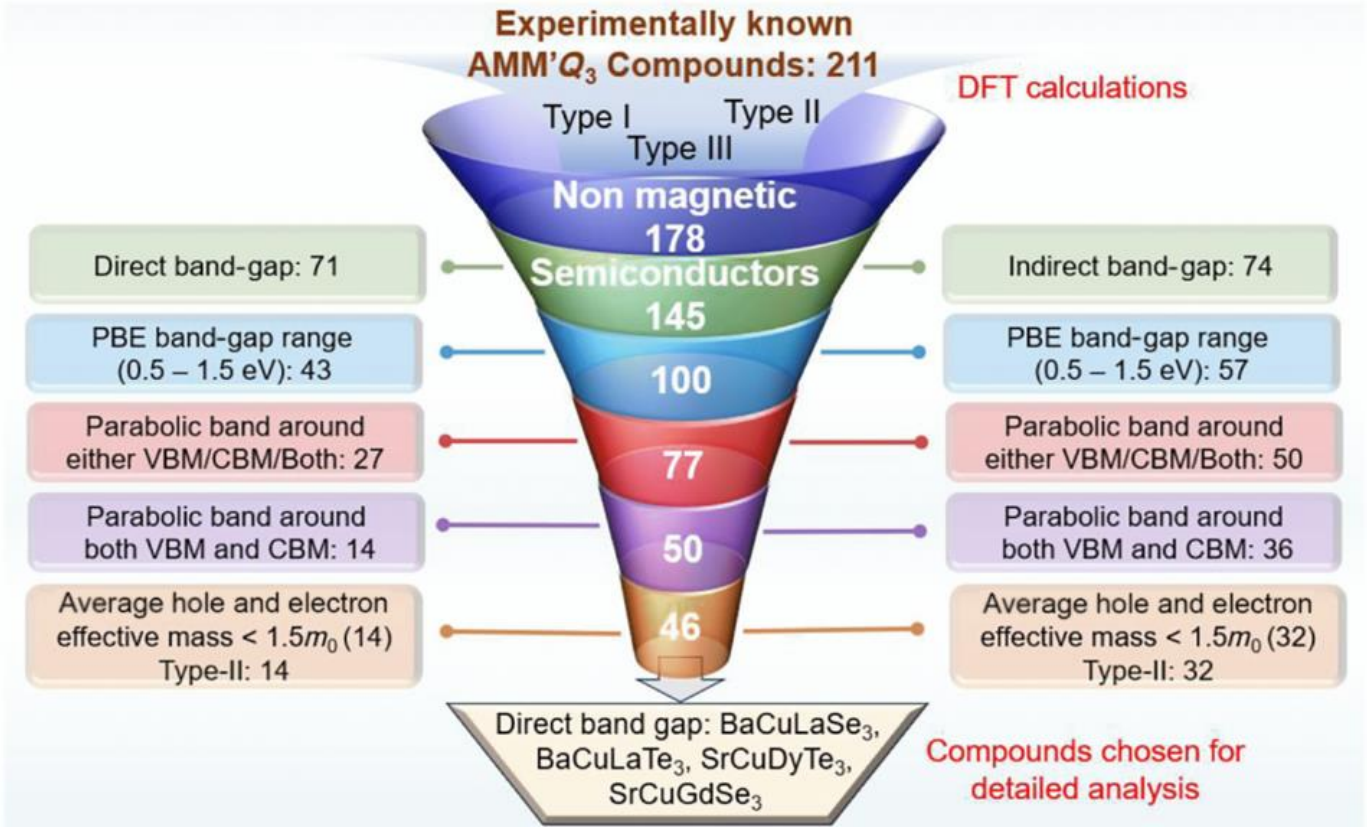


Bottom-Up

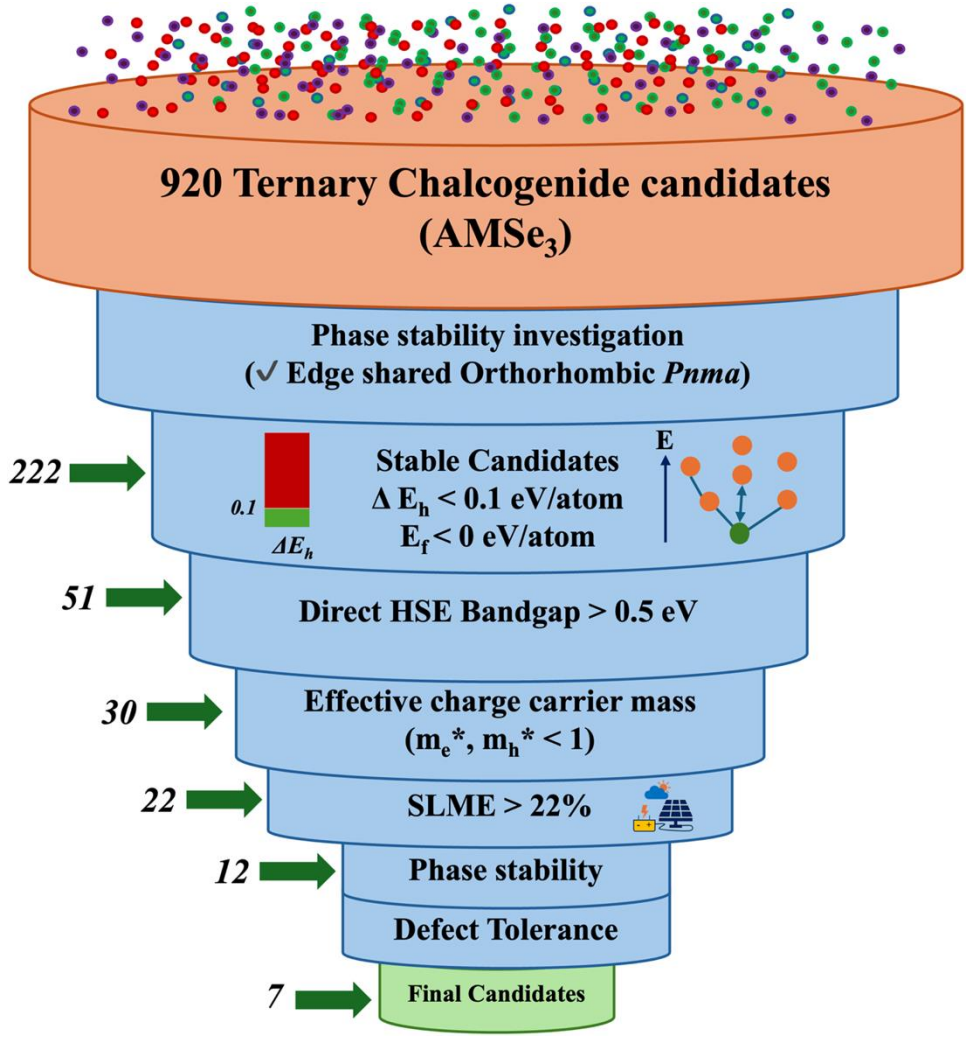


III. Modern approaches for materials discovery

High Throughput studies (Top-Down)



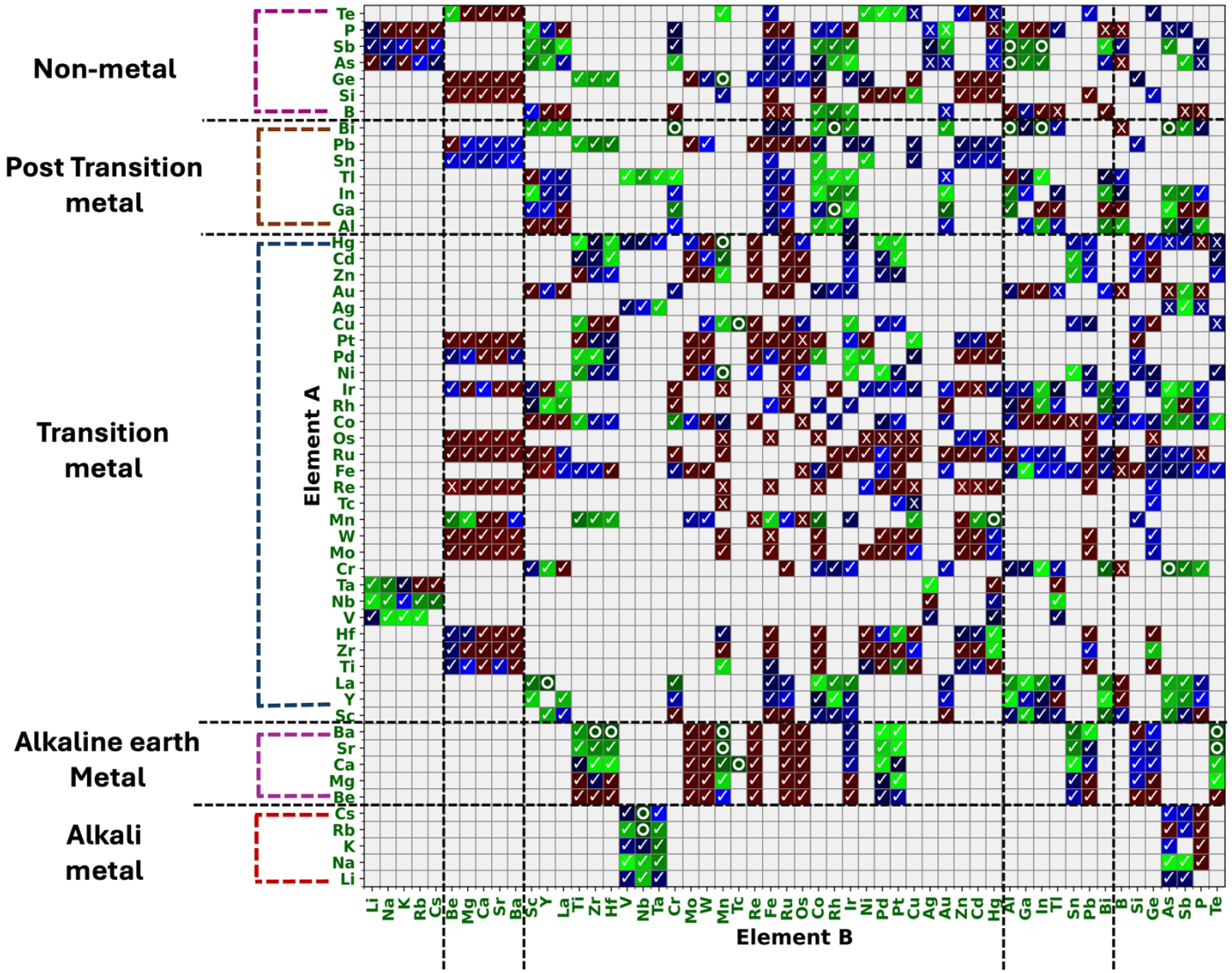
Singh et al., **Phys. Rev. Applied,** 2026, 25,034073

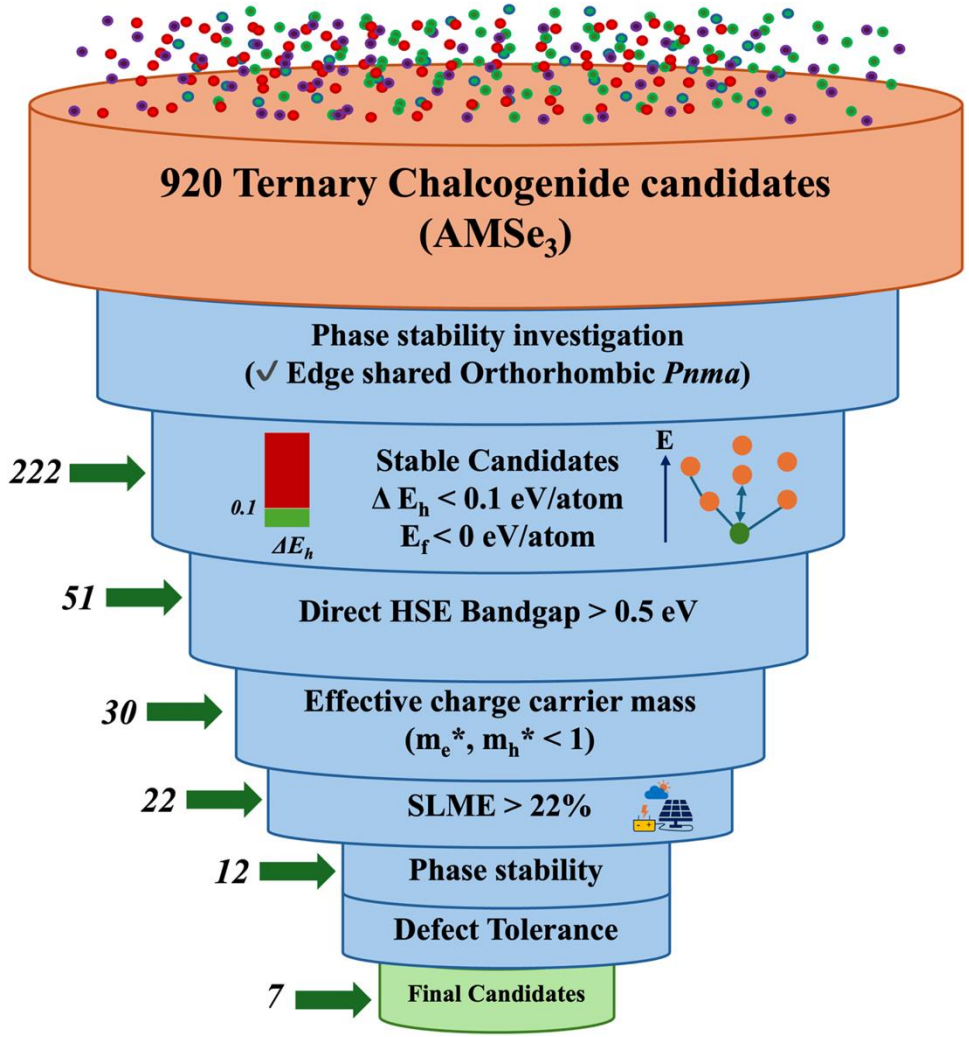


*Singh et al, **J. Mater. Chem. A** 2025; 13 (13): 9192–9210*

III. Modern approaches for materials discovery

High Throughput studies (Top-Down)

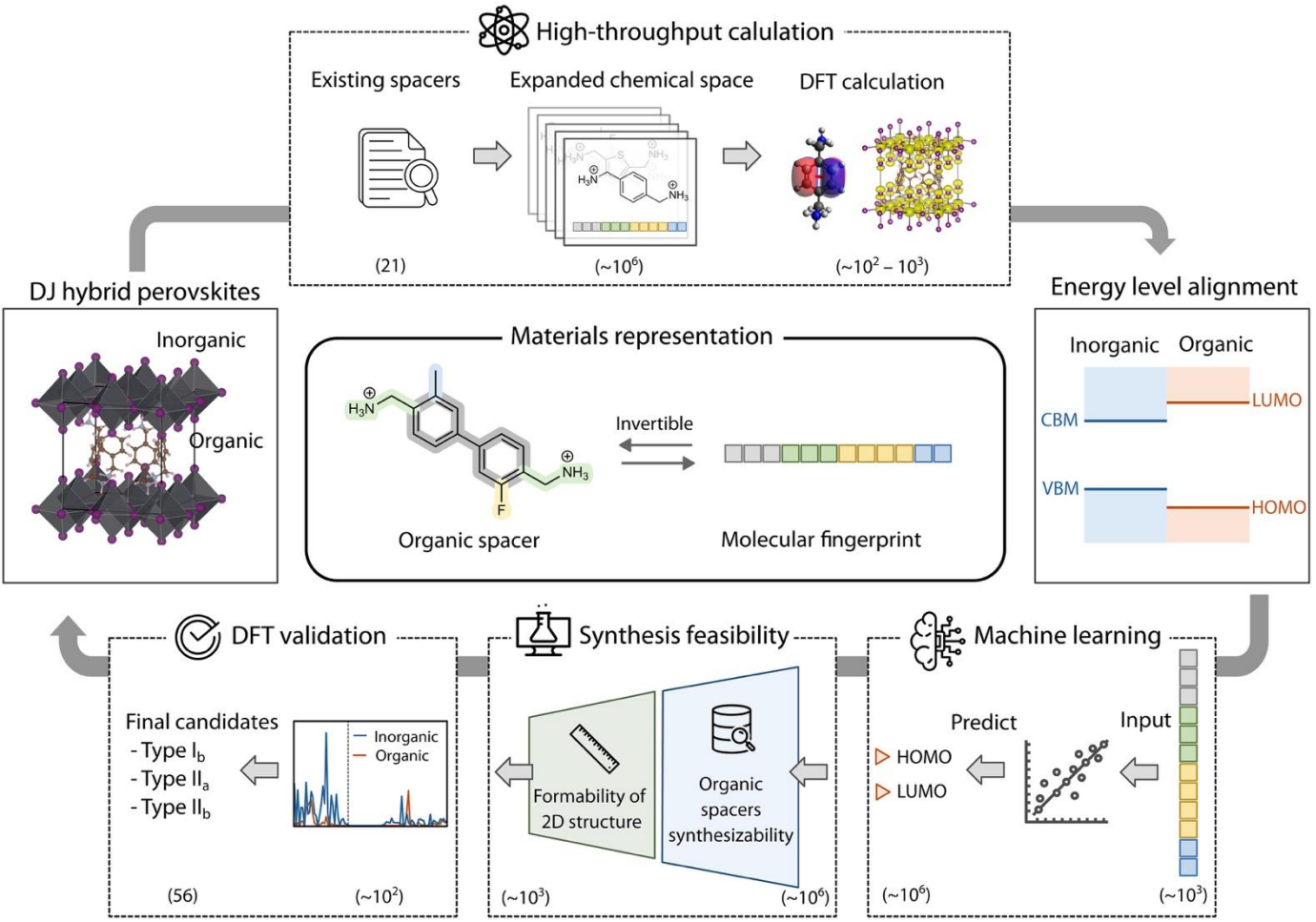


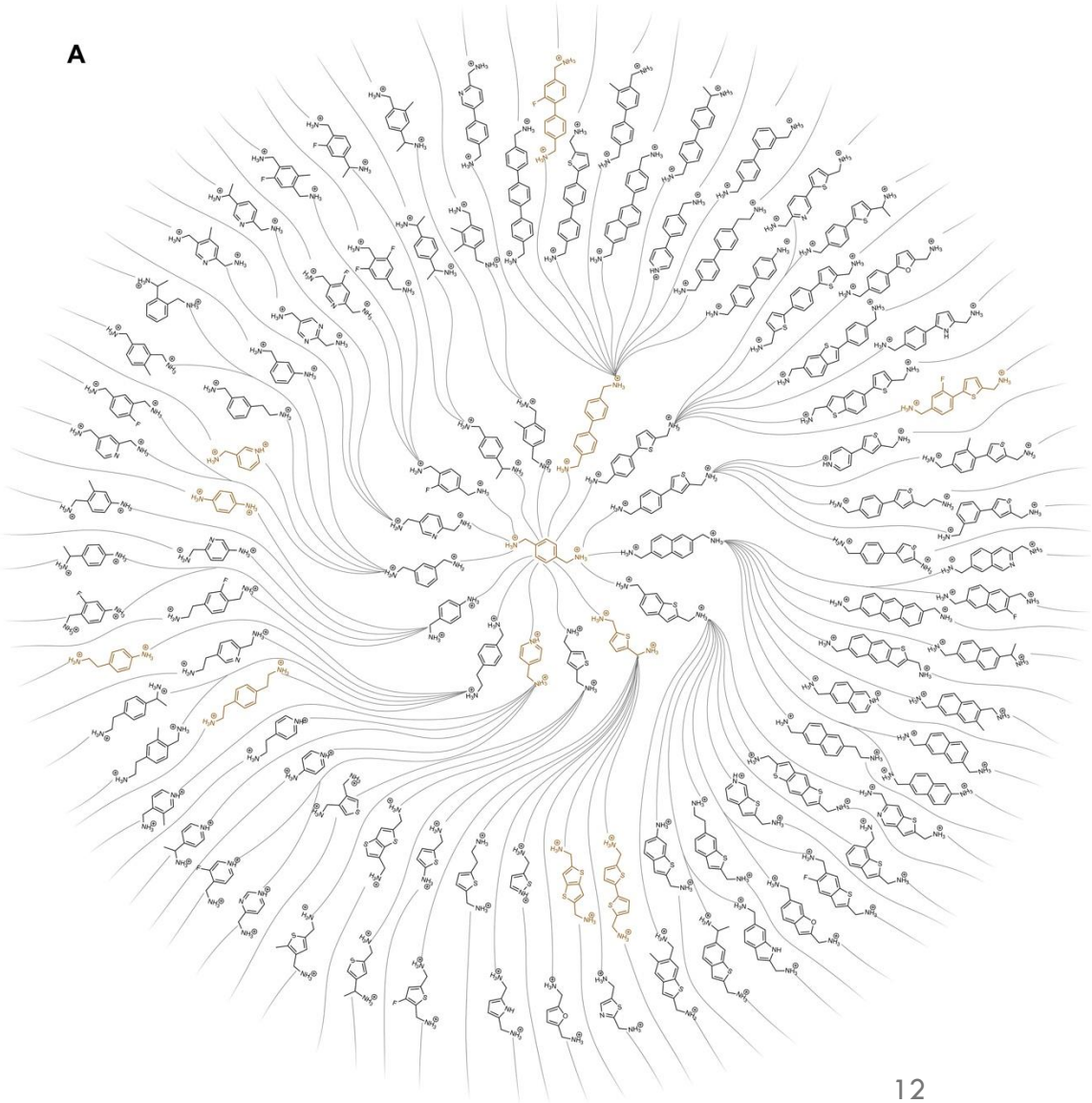


Singh et al, J. Mater. Chem. A 2025; 13 (13): 9192–9210

III. Modern approaches for materials discovery

Reported Studies – Inverse Design (Bottom Up)





Yongxin Lyu *et al.*, Fingerprinting organic molecules for the inverse design of two-dimensional hybrid perovskites with target energetics. ***Sci. Adv.*** **12**,eaeb4144(2026)

IV. The present and future of Materials (ML & AI)

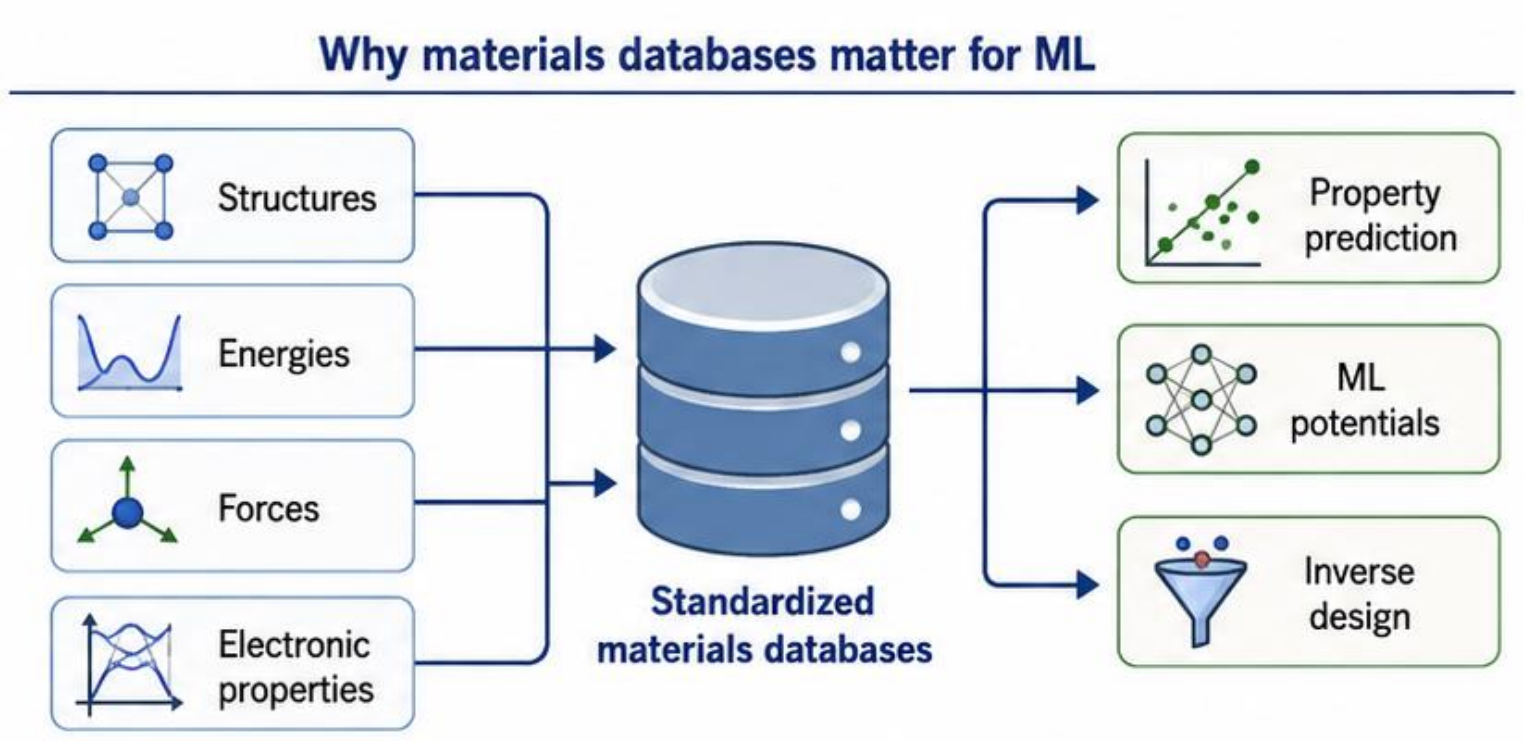
We ran DFT millions of times and left the results lying around

The Materials Project
materialsproject.org

NOMAD

OQMD
The Open Quantum Materials Database

AFLOW
Automatic - FLOW for Materials Discovery



DFT calculations are no longer single-use results. Reusable and input for ML

Challenges:
1. **Different computational settings**
2. **Incomplete details**

IV. The present and future of Materials (ML & AI)

Why Machine Learning was needed:

Computations got faster and required less researcher time

But ……

CPU time was still an issue! And, We wanted to go even faster

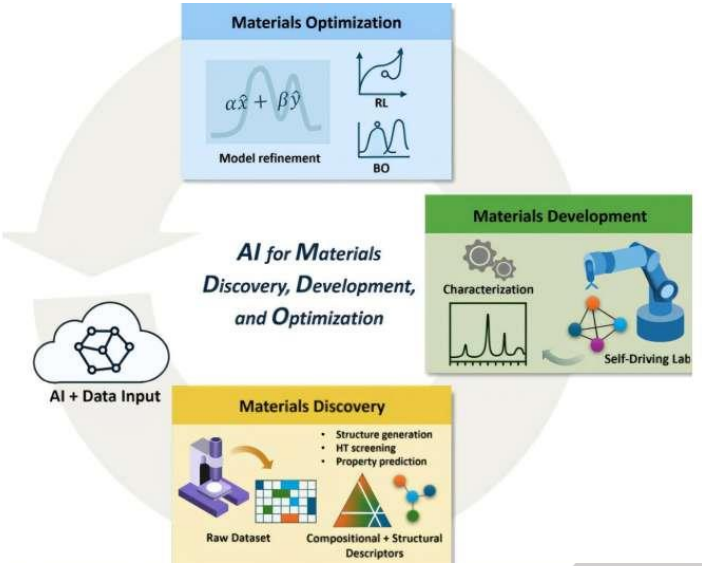
Pyzer-Knapp *et al.* Accelerating materials discovery using artificial intelligence, high performance computing and robotics. *npj Comput Mater* **8,** 84 (2022)

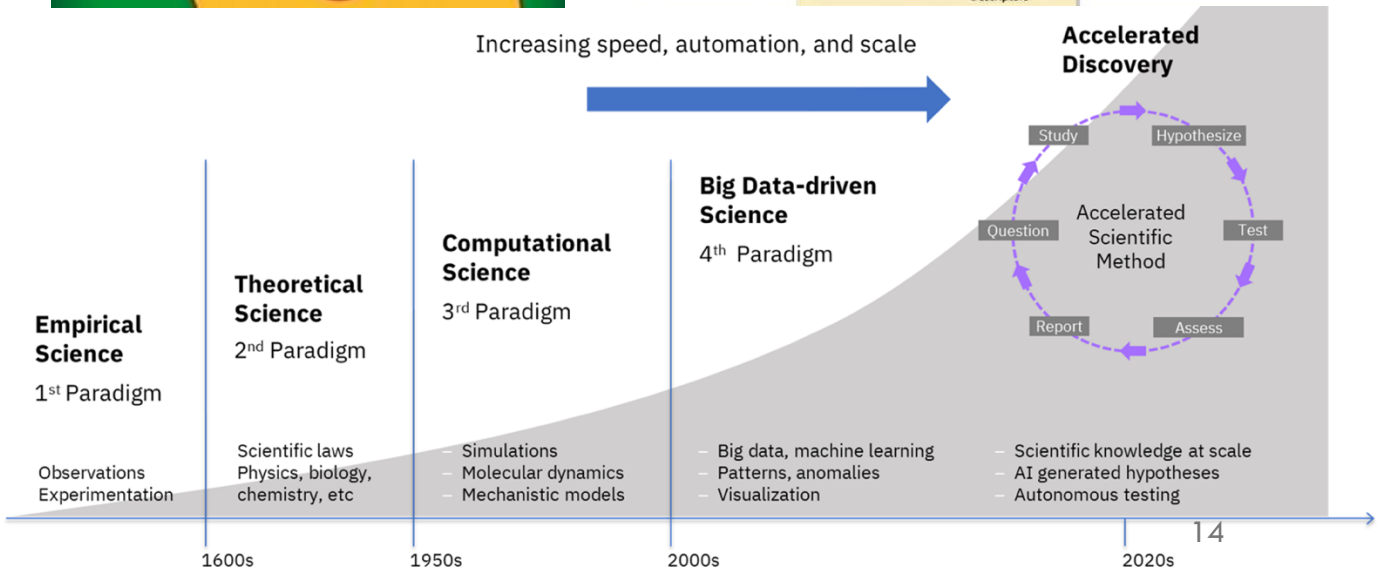
The age of Artificial Intelligence

Me: *uses machine learning*

Machine: *learns*

Me:



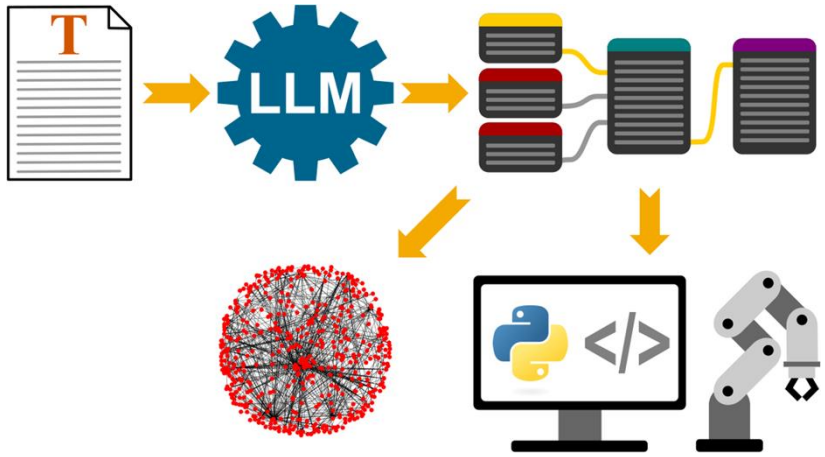


IV. The present and future of Materials (ML & AI)

UNDERSTAND

Materials Knowledge Extraction

- Mine Literature
- Extract composition, synthesis and properties
- Employ synthesis conditions in ML based pipeline

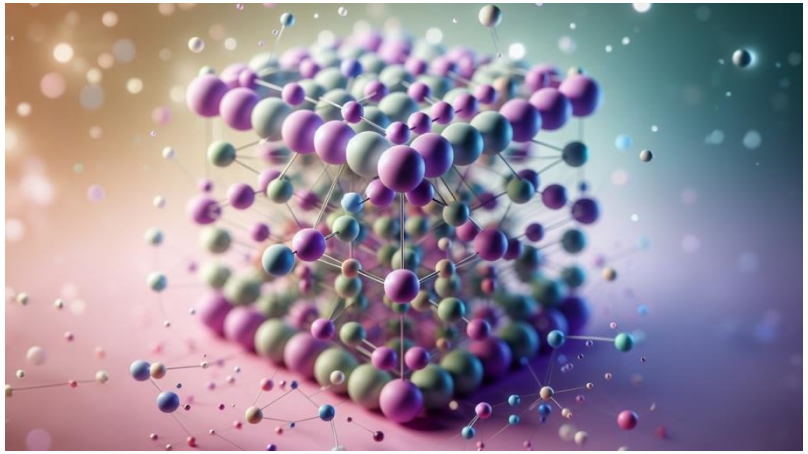


Literature → Structured materials knowledge

GENERATE

Generative and Inverse Design

- Diffusion models for novel candidates
- Property Conditioned candidate design
- Inverse design from the target performance

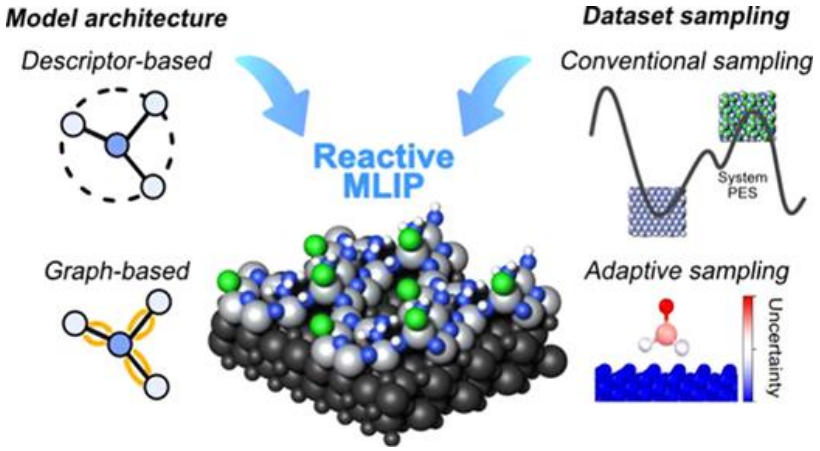


Desired properties → Candidate materials

PREDICT & SIMULATE

Physics-Aware Materials Models

- GNN and Foundational models
- ML interatomic potentials
- Fast prediction of properties, dynamics, etc



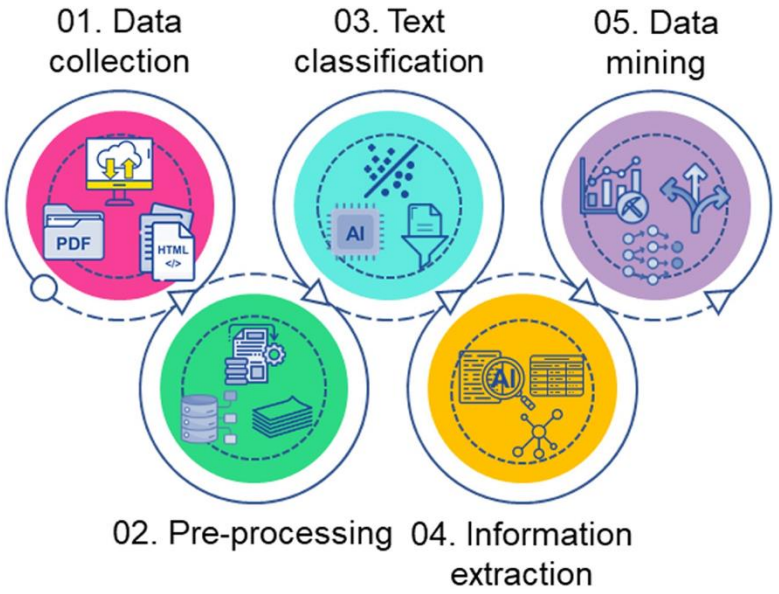
Atomic structure → Properties and dynamics

Three Frontiers of AI for Materials Science

IV. The present and future of Materials (ML & AI)

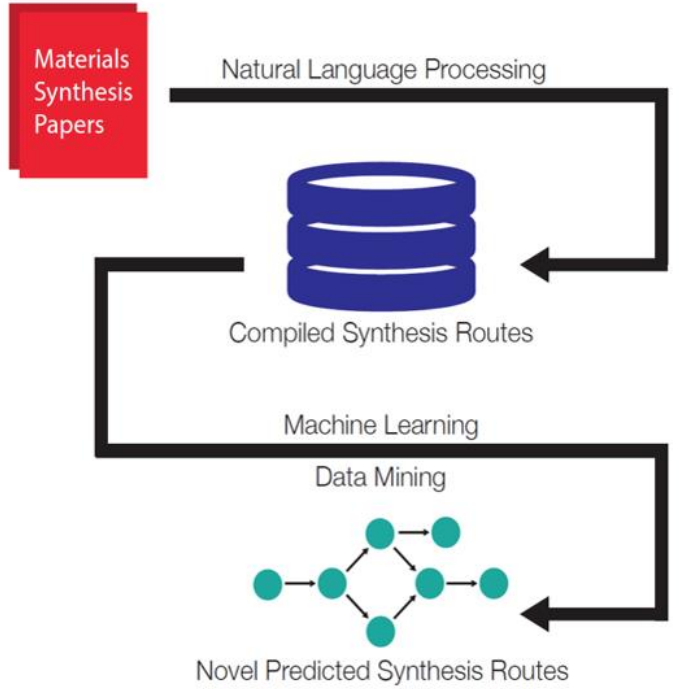
Two centuries of chemistry sits in prose, not in tables.

Papers to read "someday"



Language models now extract synthesis recipes : Precursors, temperatures, atmospheres, durations from millions of papers, and turn them into something a computer can reason over.

AI can read quickly!



IV. The present and future of Materials (ML & AI)

Natural Language Processing

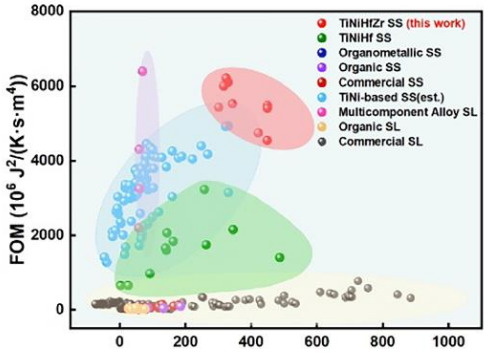
Complexity

Application

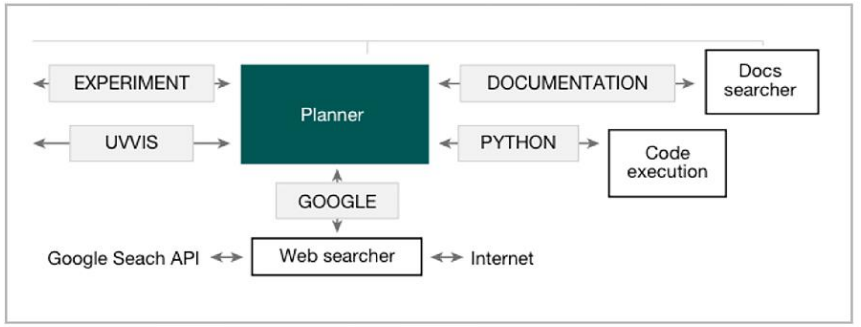
Materials tasks

Technique

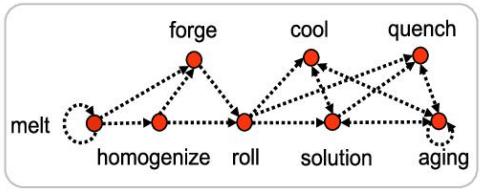
New materials discovery



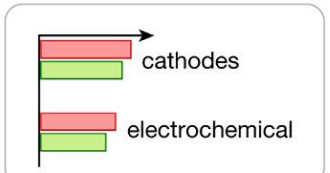
Autonomous research



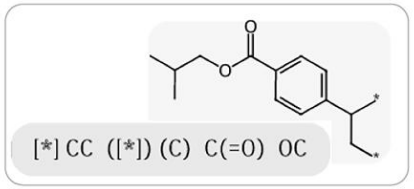
Information extraction



Static similarity



Sequential representation



Prompt engineering

Prompt:
Answer the question as truthfully as possible using the provided context.

Natural Language Understanding

Natural Language Processing

Natural Language Generation

Handcrafted Rules Based (1950s)
Machine Learning (1980s)
Neural Language Model (2001)
Multitask Learning (2008)
Word Embedding (2013)
Deep Learning for NLP (2013)
Sequence to Sequence (2014)
Attention Mechanism (2015)
Large Language Models (2018-present)

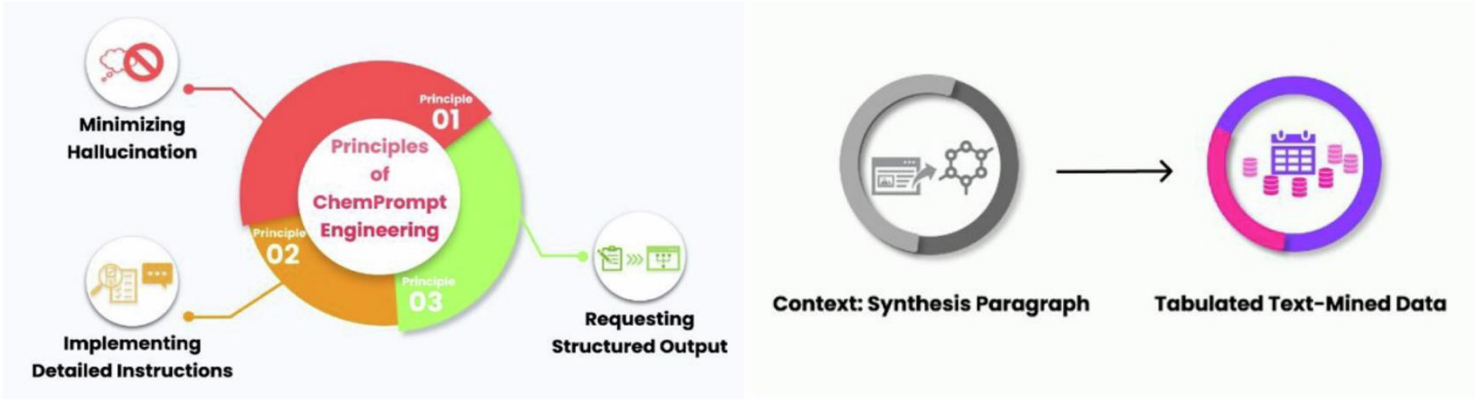
Techniques evolution

Jiang, X., Wang, W., Tian, S. *et al*. **Applications of natural language processing and large language models in materials discovery.** ***npj Comput Mater*** **11**, 79 (2025)

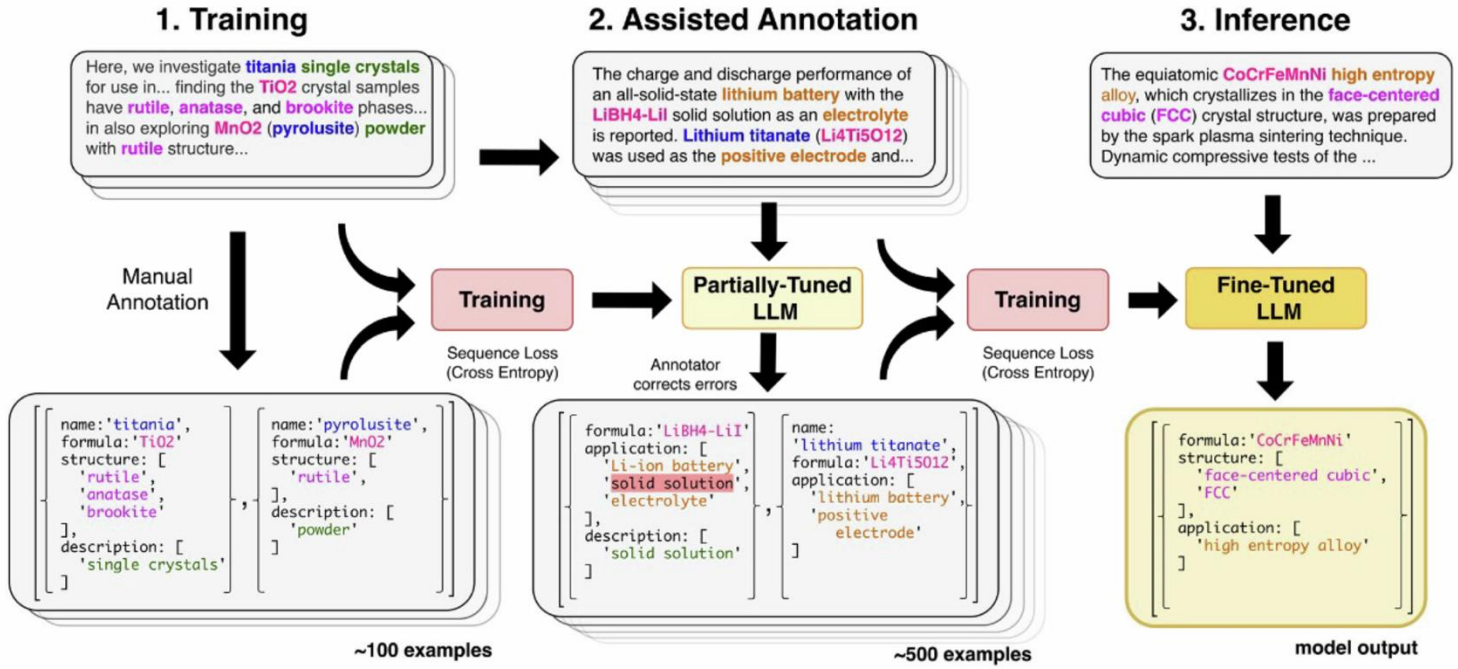
IV. The present and future of Materials (ML & AI)

Natural Language Processing

a.



b.

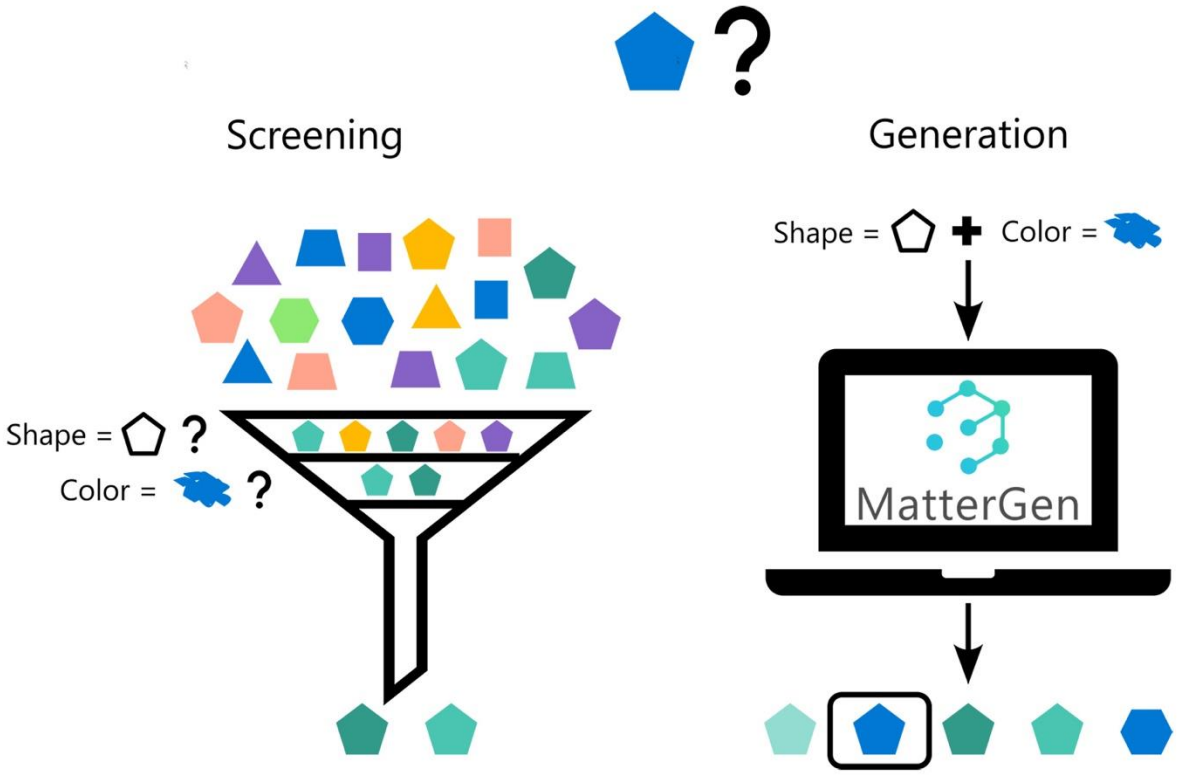


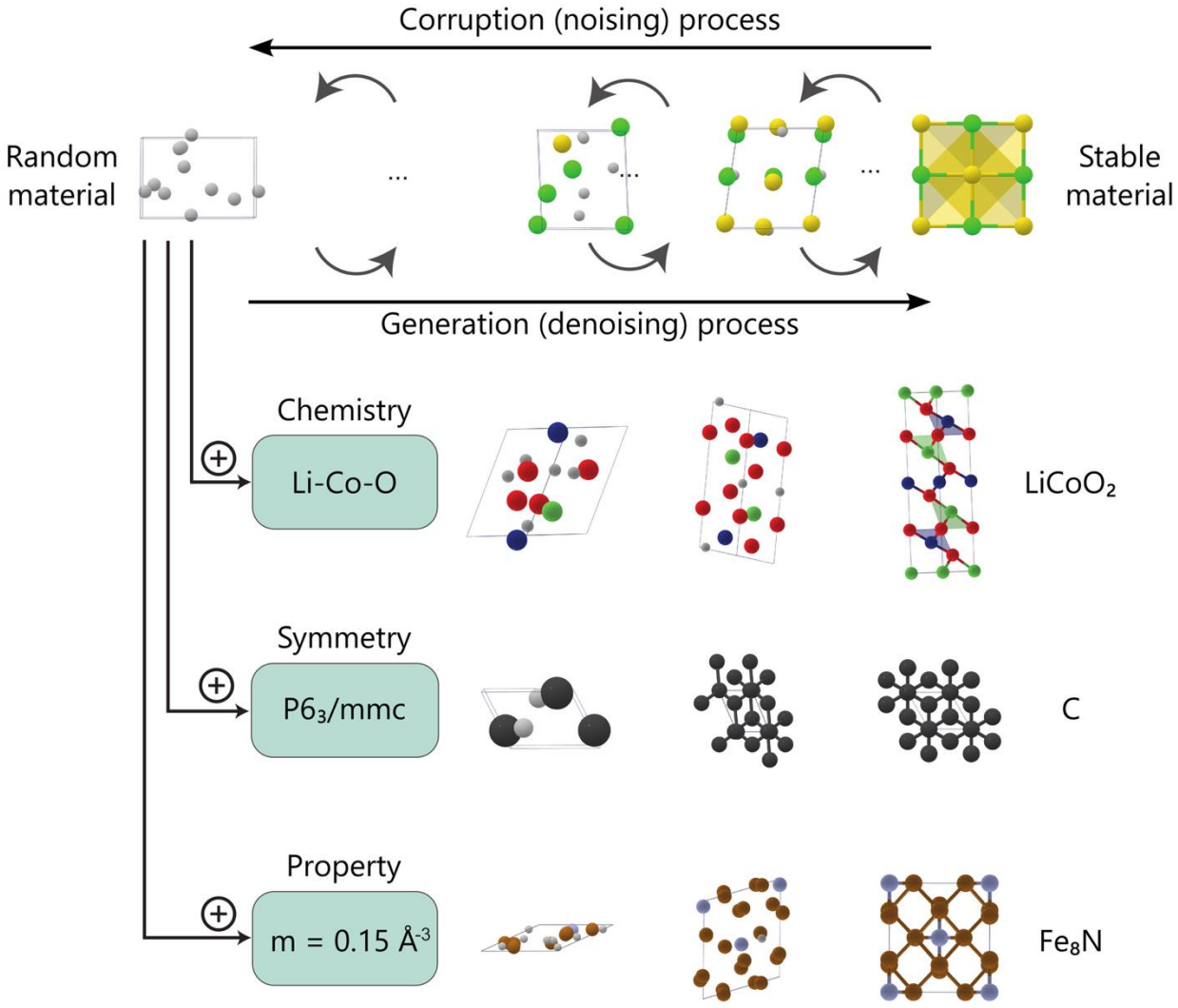
Jiang, X., Wang, W., Tian, S. *et al*. **Applications of natural language processing and large language models in materials discovery.** ***npj Comput Mater*** **11,** 79 (2025)

IV. The present and future of Materials (ML & AI)

Dreaming Up Crystal Structures!

MatterGen™



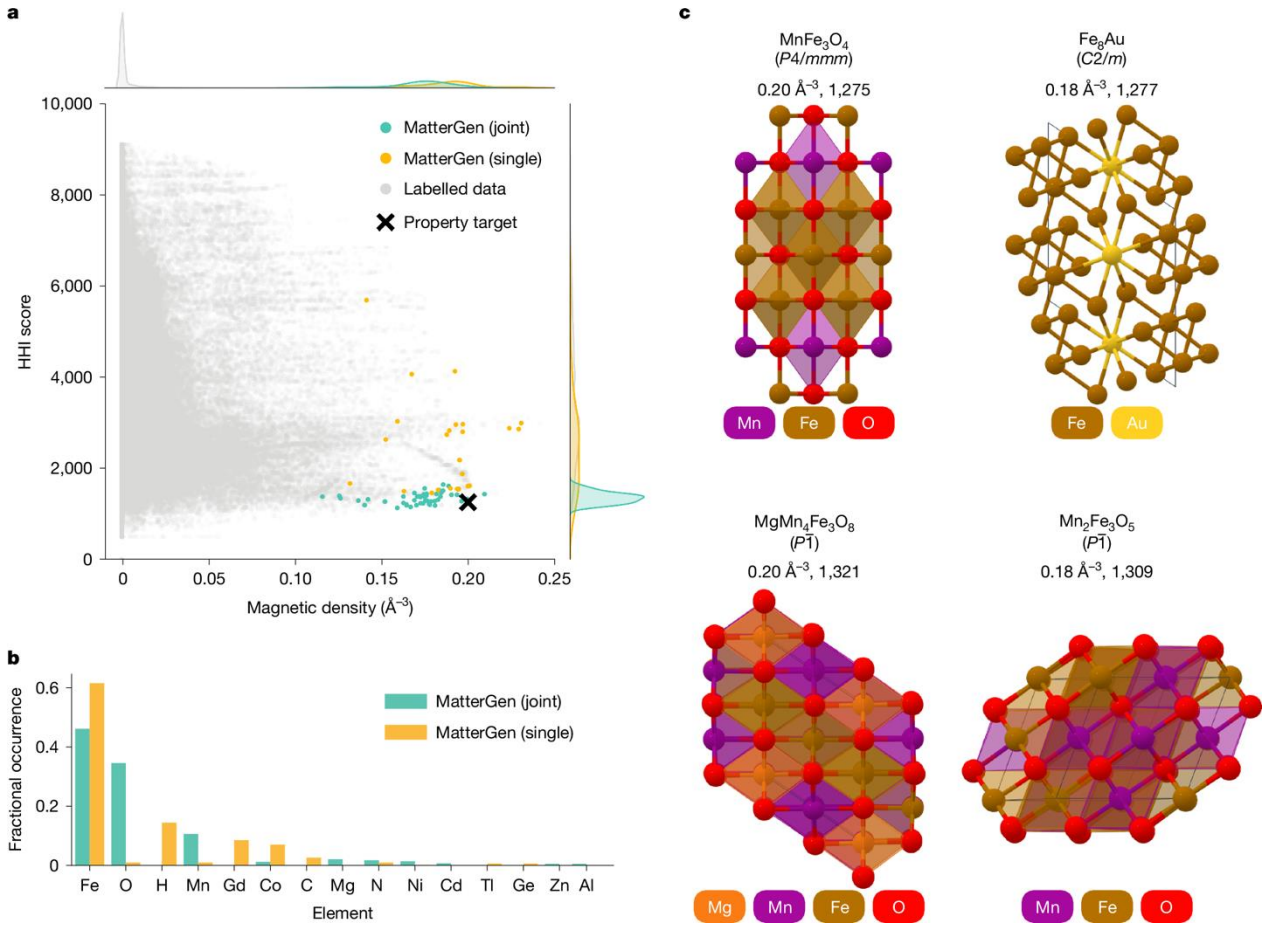


Zeni, C. *et al*. A generative model for inorganic materials design. *Nature* **639**, 624–632 (2025).

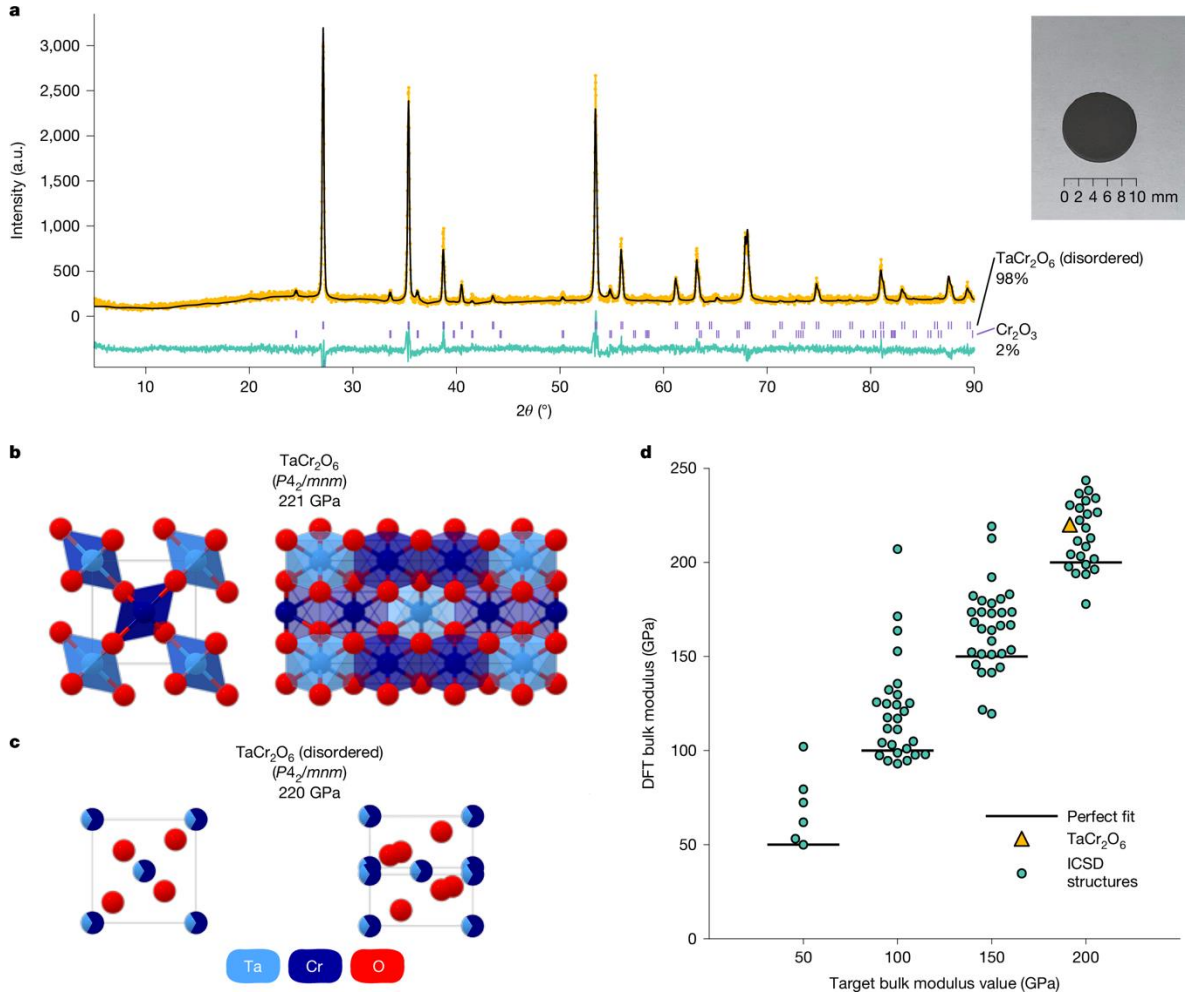
IV. The present and future of Materials (ML & AI)

Dreaming Up Crystal Structures!

Designing low-supply-chain-risk magnets



Experimental validation of generated structures



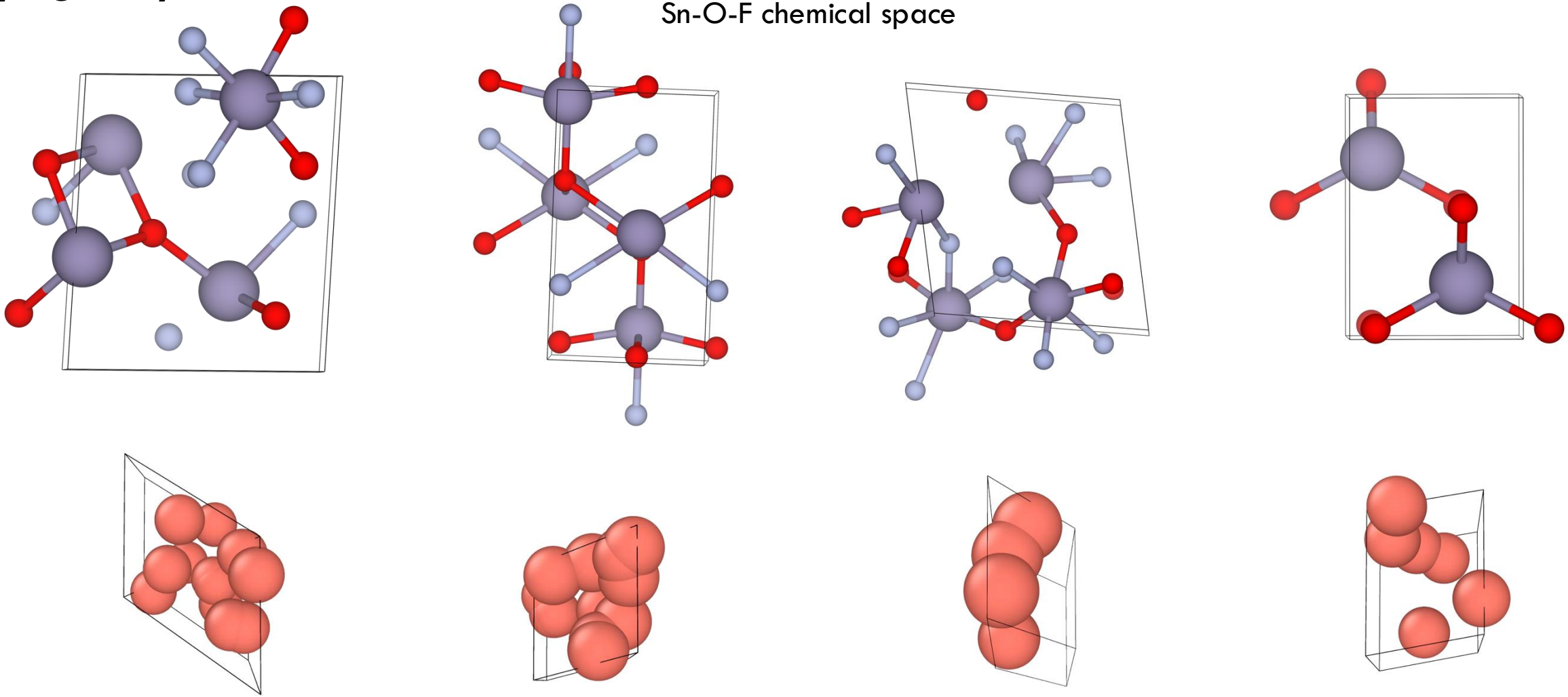
Zeni, C. *et al*. A generative model for inorganic materials design. *Nature* **639**, 624–632 (2025).

IV. The present and future of Materials (ML & AI)

Trying it myself

Dreaming Up Crystal Structures!

Sn-O-F chemical space



V. Amalgamation of Experiments and Theory (ML)

The Age of Robots

We can now predict far faster than we can make

~400,000 structures predicted stable by one model, in one release, in one year (Google Deepmind)

However, genuinely new compounds a skilled experimentalist makes and characterises in the same year **~ 10-15**

For most of history, the shortage was ideas. The shortage is now hands.

V. Amalgamation of Experiments and Theory (ML)

Early landmark case study

nature

Explore content ⌄ About the journal ⌄ Publish with us ⌄

nature › articles › article

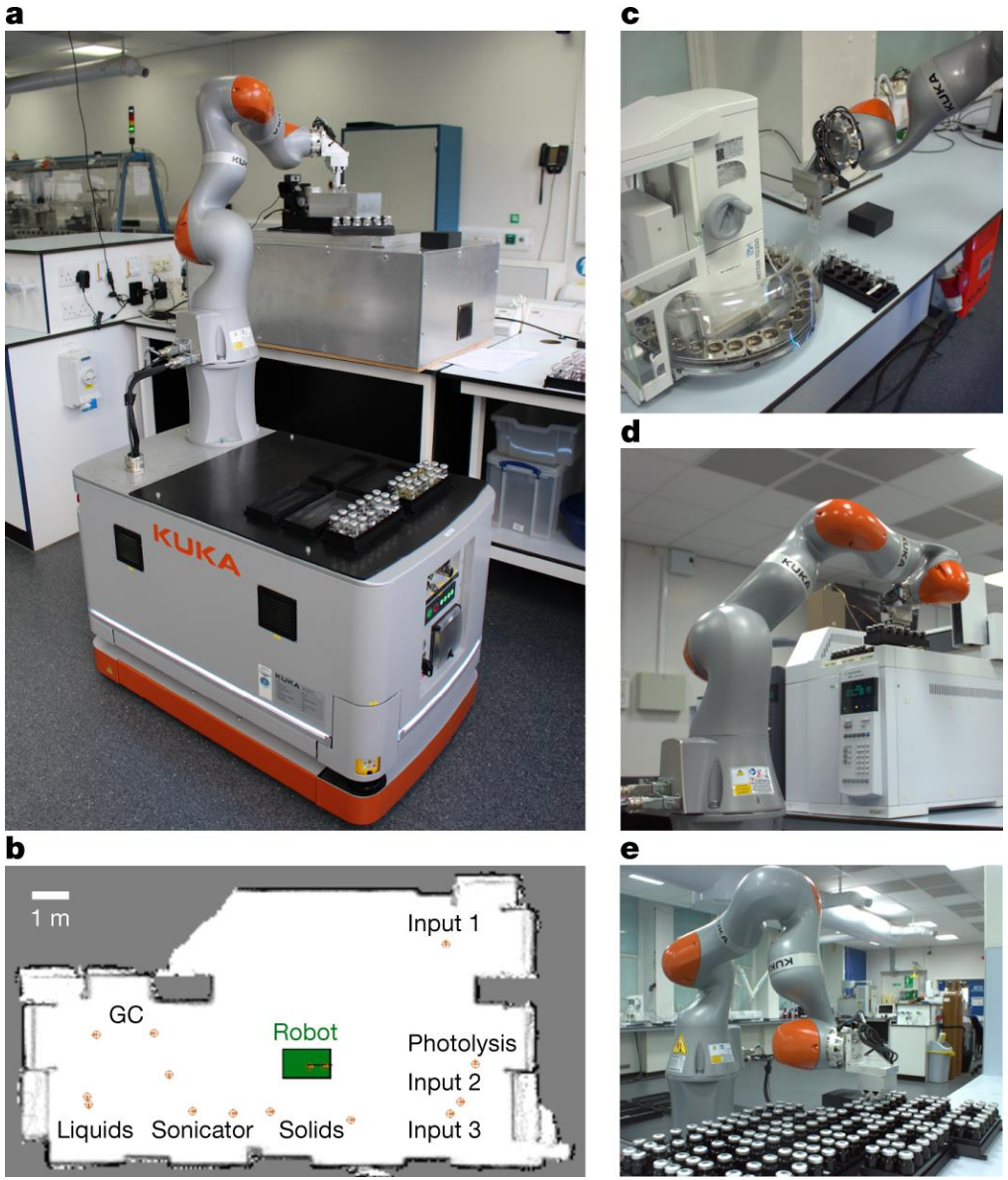
Article | Published: 08 July 2020

A mobile robotic chemist

Benjamin Burger, Phillip M. Maffettone, Vladimir V. Gusev, Catherine M. Aitchison, Yang Bai, Xiaoyan Wang, Xiaobo Li, Ben M. Alston, Buyi Li, Rob Clowes, Nicola Rankin, Brandon Harris, Reiner Sebastian Sprick & Andrew I. Cooper ✉

Nature **583**, 237–241 (2020) | Cite this article

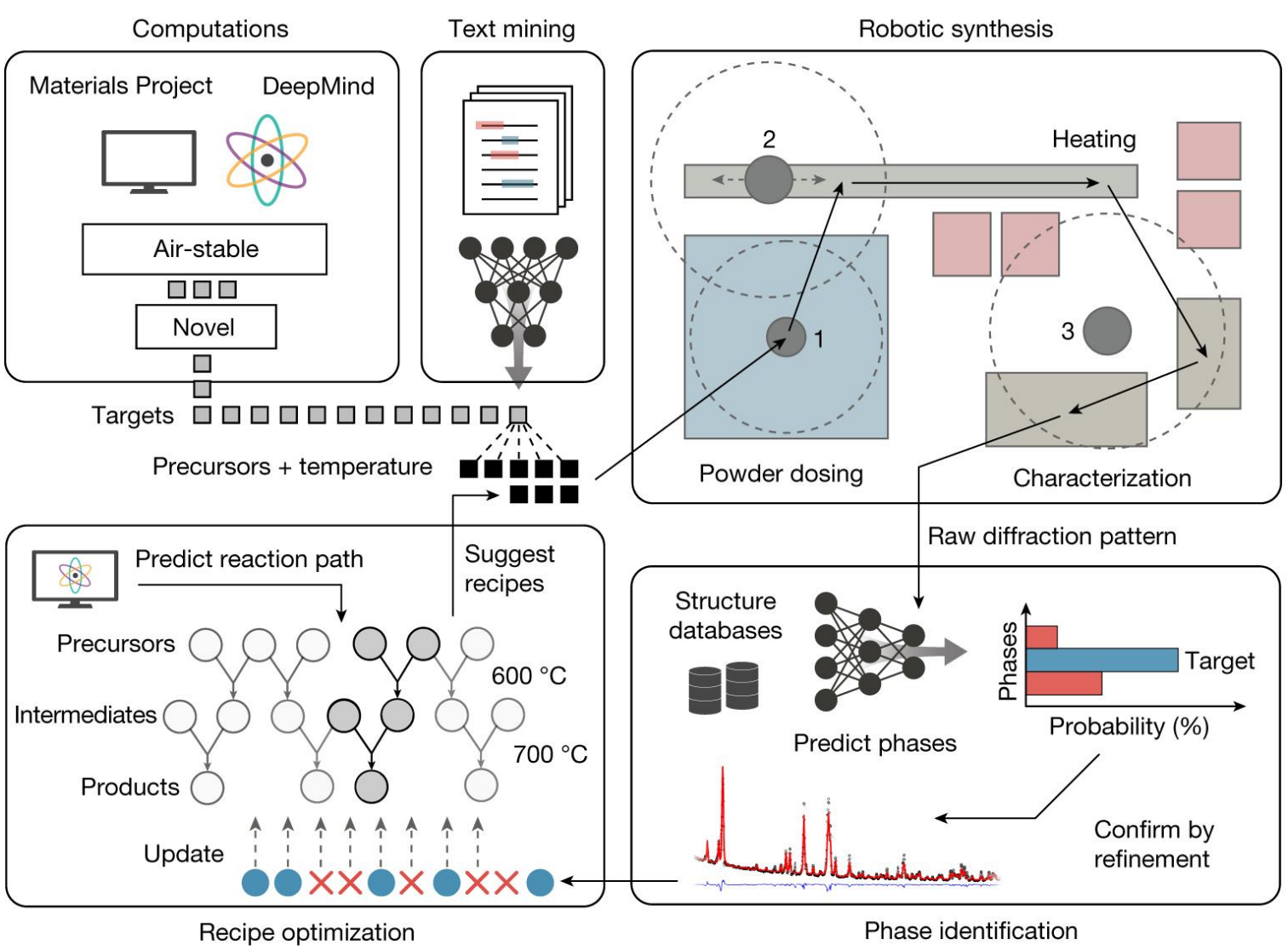
A free-roaming robot performed **688 experiments in eight days** across a ten-variable space and identified photocatalyst formulations producing approximately **six times greater hydrogen-evolution activity** than the initial mixtures.



V. Amalgamation of Experiments and Theory (ML)

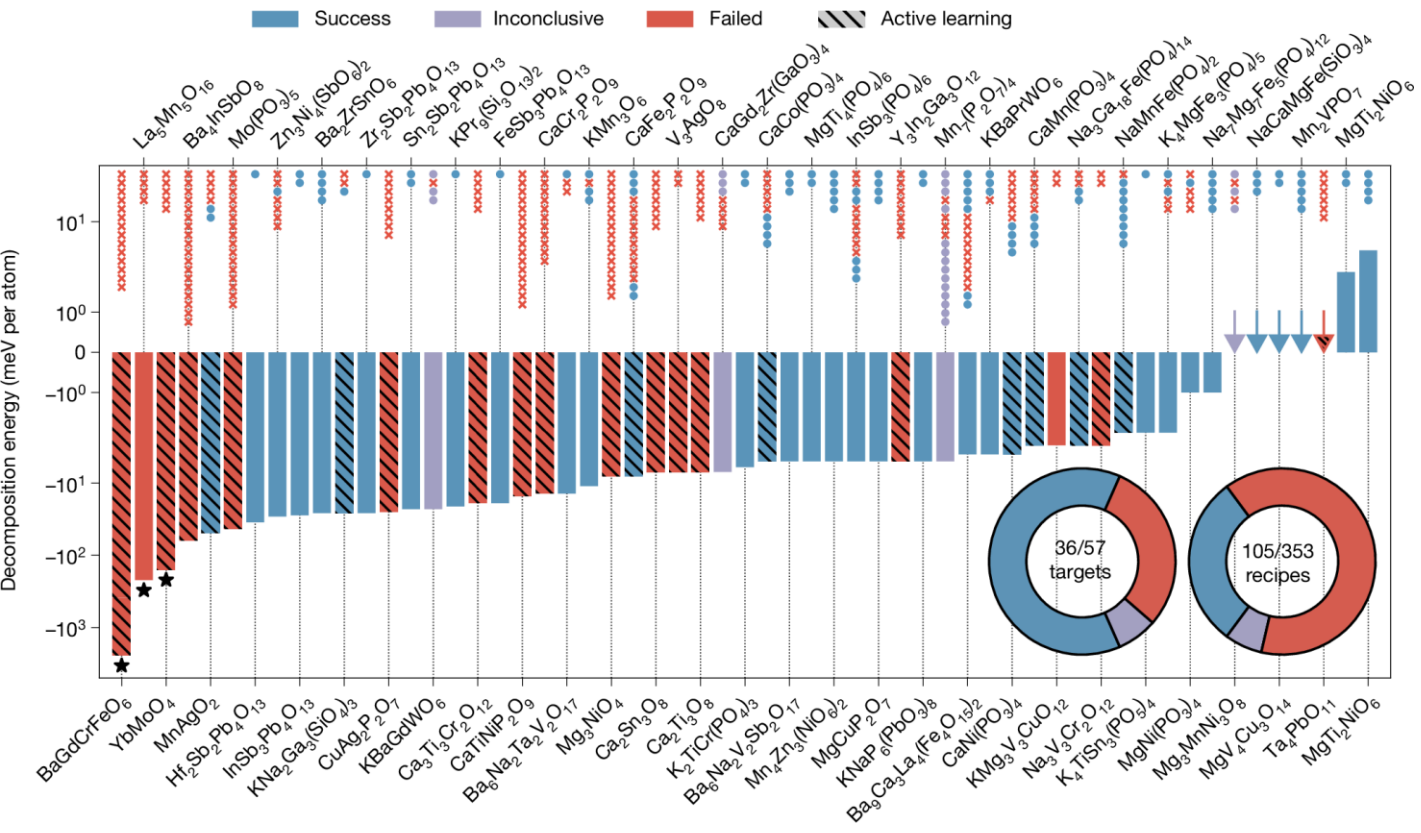
Early landmark case study

V. Amalgamation of Experiments and Theory (ML)

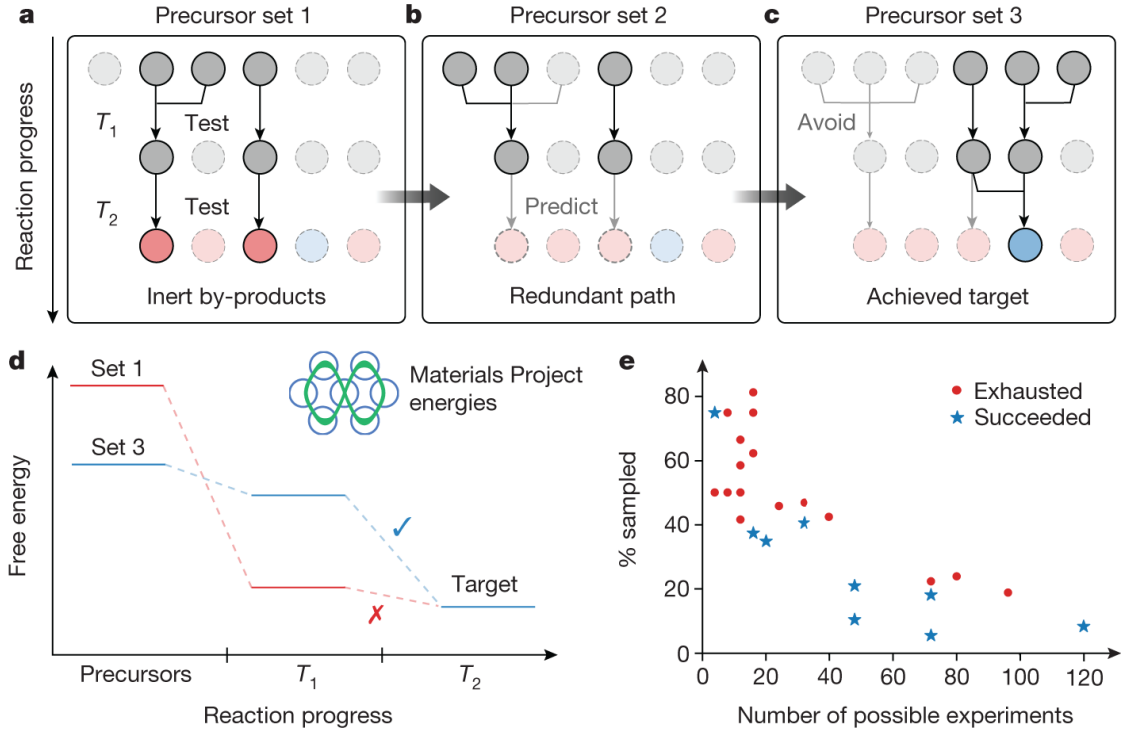


Szymanski, N.J. *et al*. An autonomous laboratory for the accelerated synthesis of inorganic materials. *Nature* **624,** 86–91 (2023)

V. Amalgamation of Experiments and Theory (ML)



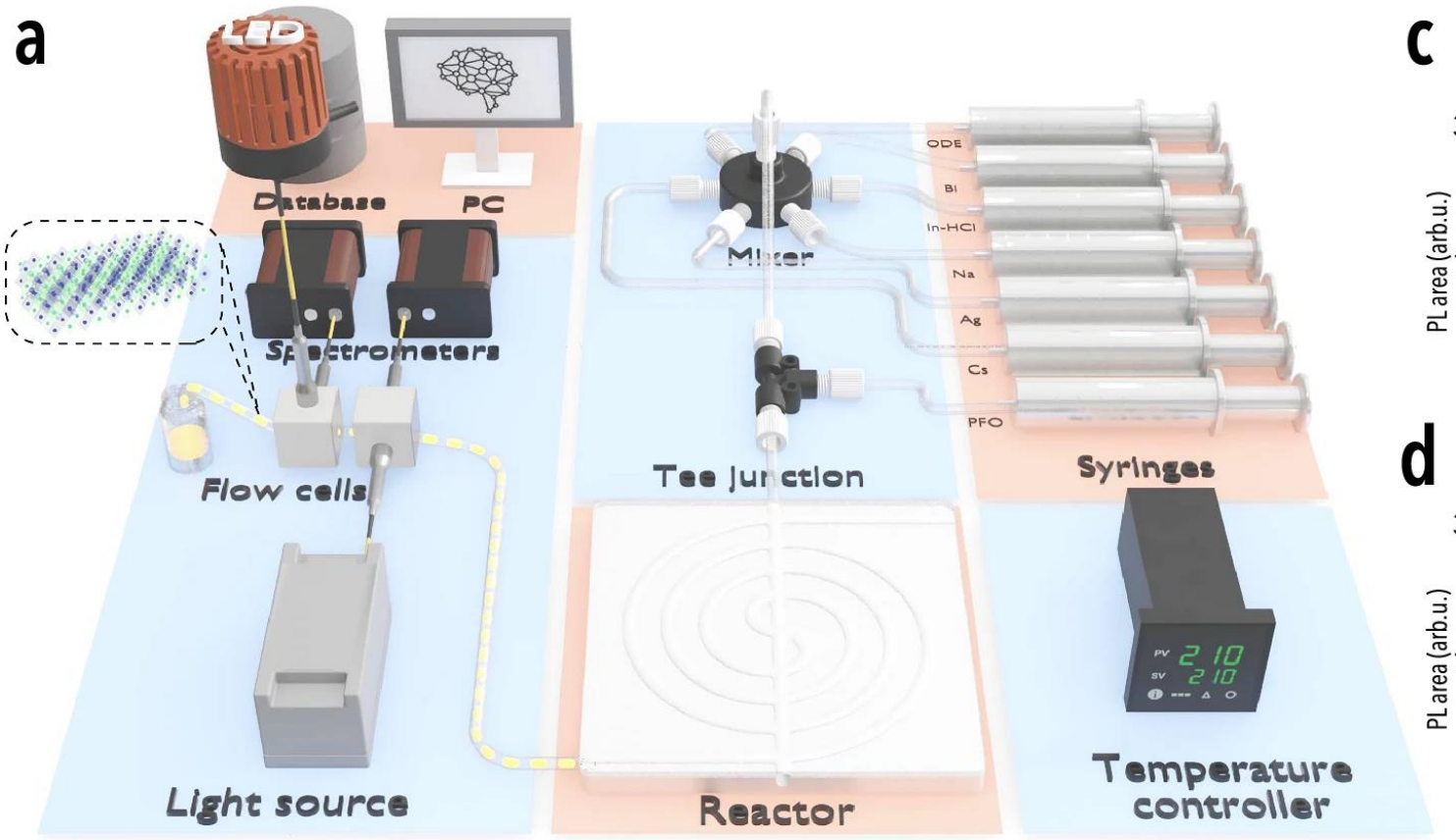
Targeted syntheses of 57 inorganic materials

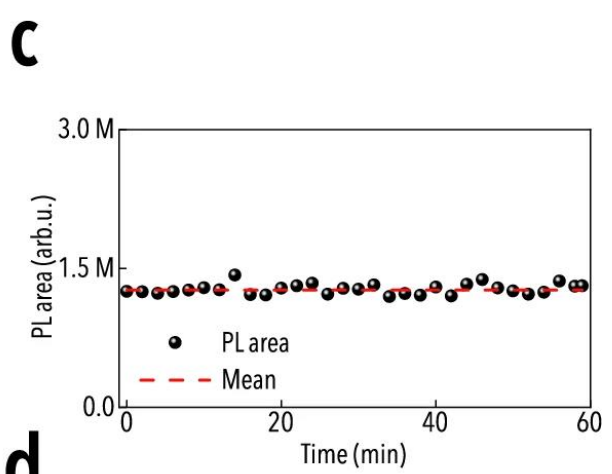


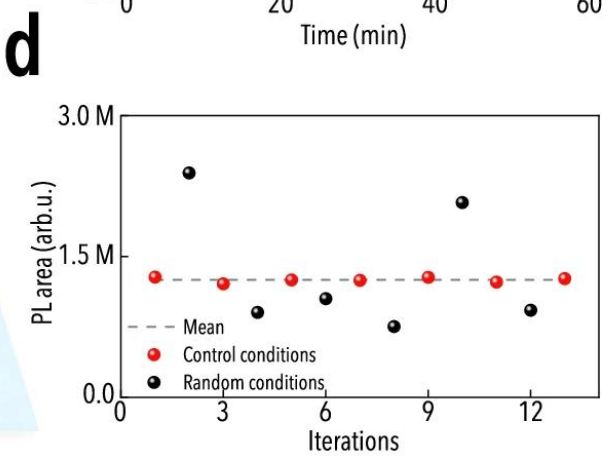
Active learning with pairwise reaction analysis

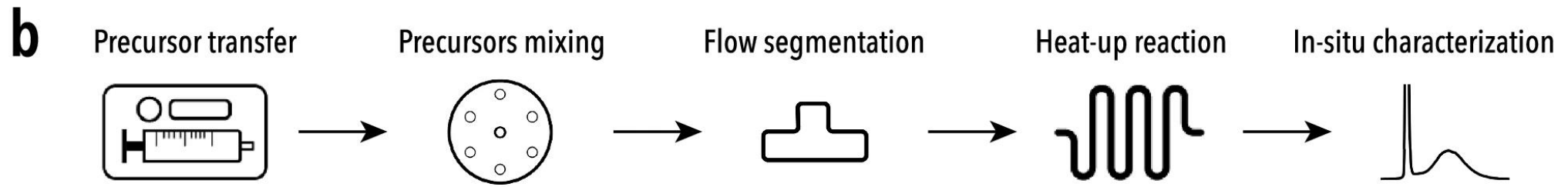
Szymanski, N.J. *et al*. An autonomous laboratory for the accelerated synthesis of inorganic materials. *Nature* **624**, 86–91 (2023)

V. Amalgamation of Experiments and Theory (ML)



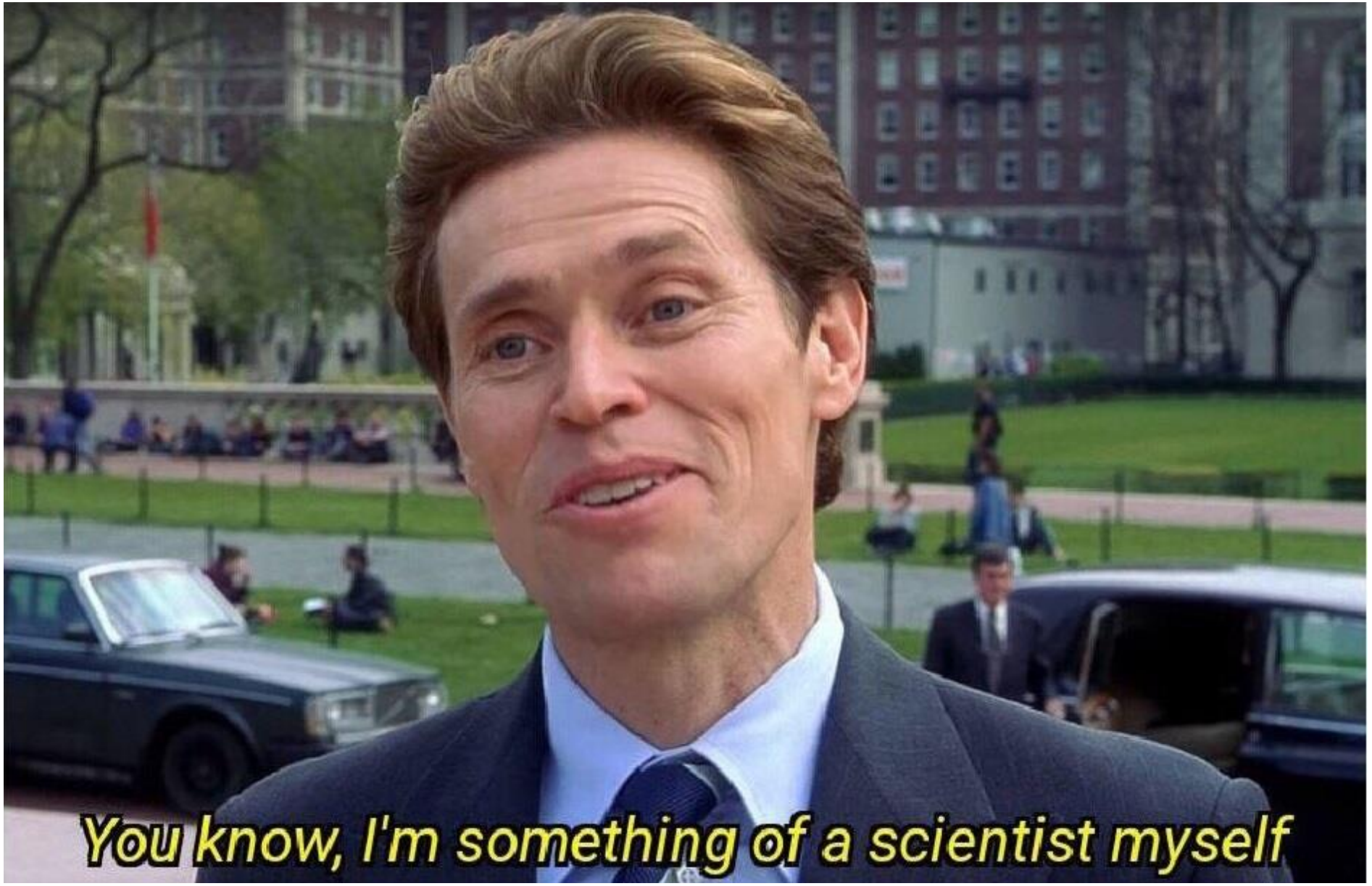






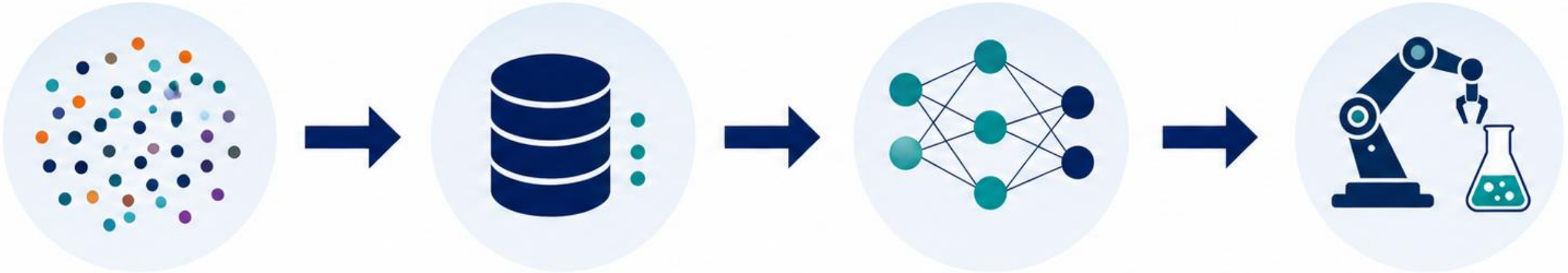
Li *et al*. Autonomous microfluidic experimentation for exploring reaction inference and synthesizing double perovskite nanoplatelets. *Nat Commun 17, 4514 (2026)*

Robots and Machine Learning after doing all this :)



So, has computation changed Materials Chemistry?

Yes, but not simply by making calculations faster.



Larger Scale of exploration

Calculations → reusable data

Numerous Machine Learning models

Prediction connected to automated synthesis

"Let's try to synthesize a material?"

1. What to make?
2. Why will it work?
3. Is it relevant to real-world use?

If you torture the data enough, nature will always confess

- Roland Coase