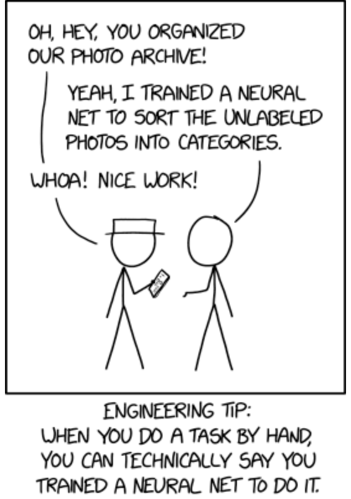



# Technologies, materials, and machine learning

Sai Gautam Gopalakrishnan

Associate Professor, Materials Engineering, Indian Institute of Science

saigautamg@iisc.ac.in; https://sai-mat-group.github.io

Manohar Parrikar Vidnyan Mahotsav, Goa
18 Dec 2025

# Acknowledgments


June 2025

DRDO
इसरो isro
SERB INDIA
IGSTC
INDO-GERMAN SCIENCE AND TECHNOLOGY CENTRE
tcs TATA CONSULTANCY SERVICES
SCHAEFFLER
UKRI
THE ROYAL SOCIETY

Fugaku (Japan)

SERC (IISc)

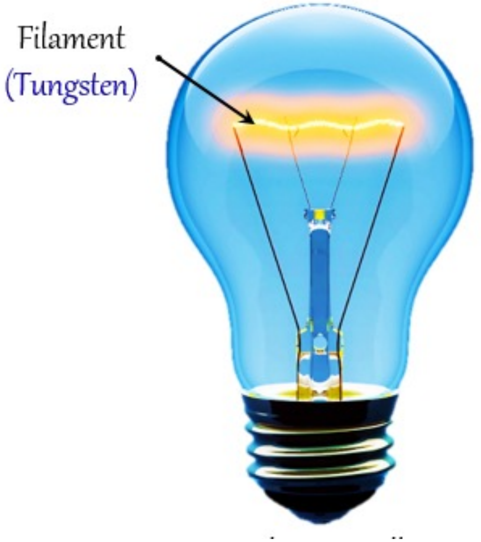
Param Utkarsh (CDAC)


Archer (UK)

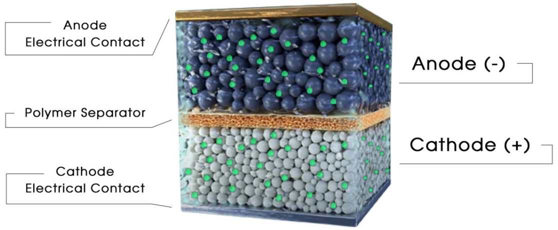

Jureca (Germany)

# What are materials?

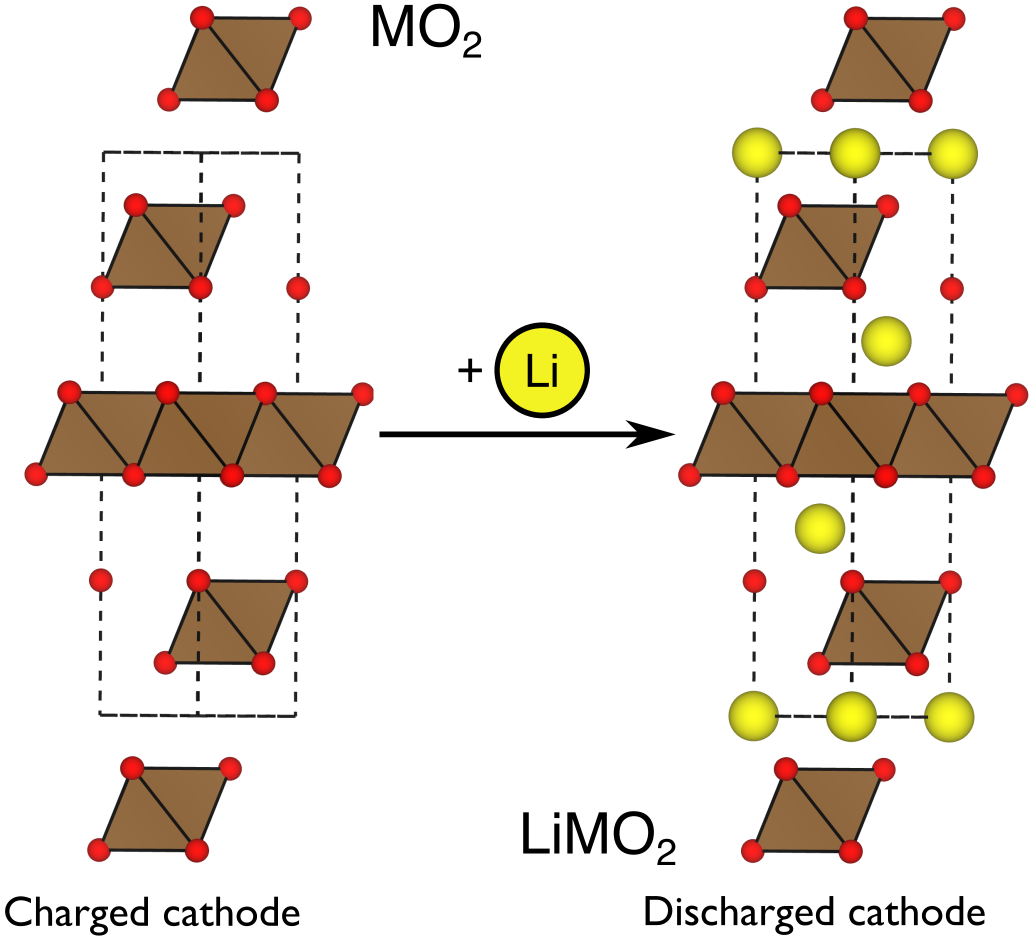
- A substance, typically *solid*, intended for use for a certain (engineering) *application*
- Study of materials: *applied* field intersecting physics, chemistry, and biology with some applied math




**Lithium-Ion Batteries**




# Voltage, capacity, and rate in Li-ion batteries




$$\Delta G_{intercalation} = G_{LiMO_2} - G_{MO_2} - G_{Li}$$

Nernst Equation

$$V = -\frac{\Delta G_{intercalation}}{nF}$$

(Do similar process for anode, take V difference!)

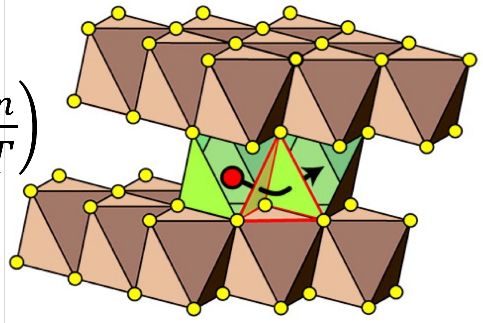
---

1 Li moved = 1 electron stored

Capacity $\propto \frac{\#\ Li\ moved}{\#\ 'Framework'\ atoms}$

---

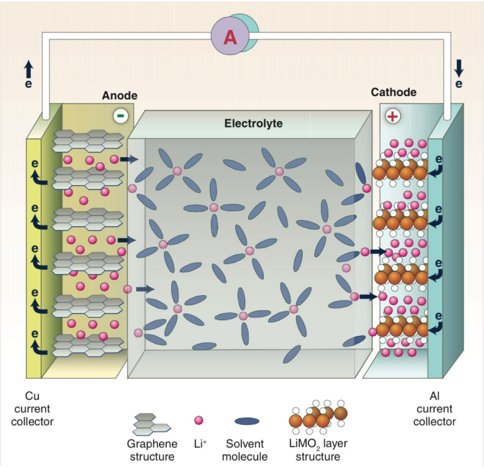
$$Rate \propto \exp\left(-\frac{E_m}{RT}\right)$$

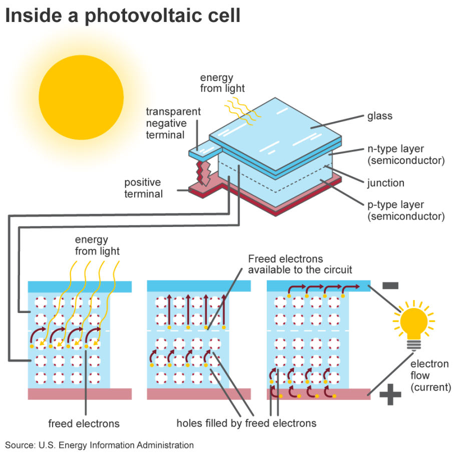


---

Rate: how fast can Li move (or diffuse) within electrode?

# Why design materials?




Key performance bottlenecks in key applications: governed by materials used

Energy and power density of a battery: limited by materials used as electrodes (and at times, electrolytes)

Key material properties: stability, ionic mobility, reaction energies

Usage of better materials (with better properties) → better performance

Efficiency of a photovoltaic: choice of semiconductor used as the light absorber

Key material properties: band gap, stability, resistance to point defects

# Materials are crucial for different technologies

### Energy

- Batteries
- Photovoltaics
- Renewable fuels
- Sensors
- Nuclear fission and fusion

### Healthcare

- Drug delivery
- Hip/knee joints
- Water desalination
- Biomedical devices
- Tissue engineering

Breakthrough in materials is the key enabler or bottleneck of several technologies

### Infrastructure and automotive

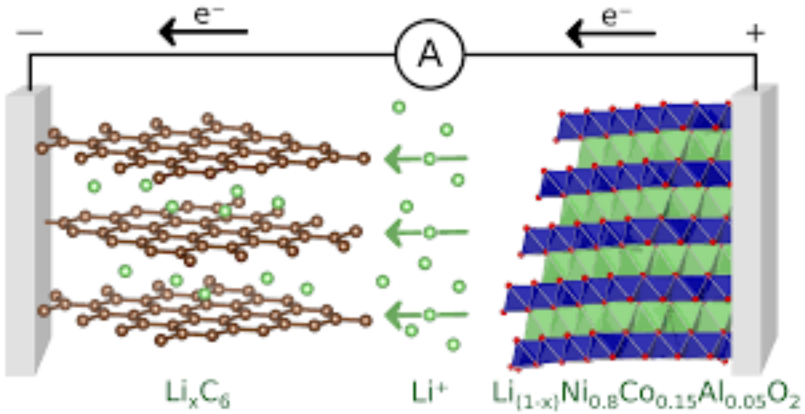
- Alloys for automobiles
- Superalloys for aerospace
- Steel for bridges, flyovers, skyscrapers
- Armor (defense applications)
- Stealth systems (Radars)

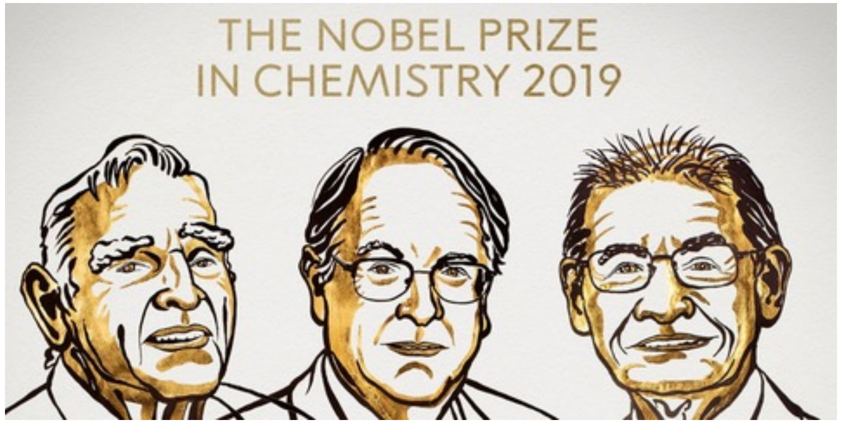
### Everyday applications, internet of things

- Flexible electronics
- Modern sports equipment
- Smaller electronics, nano-chips
- Biodegradable plastics
- Water-repellents

# Typical life of a material scientist/engineer

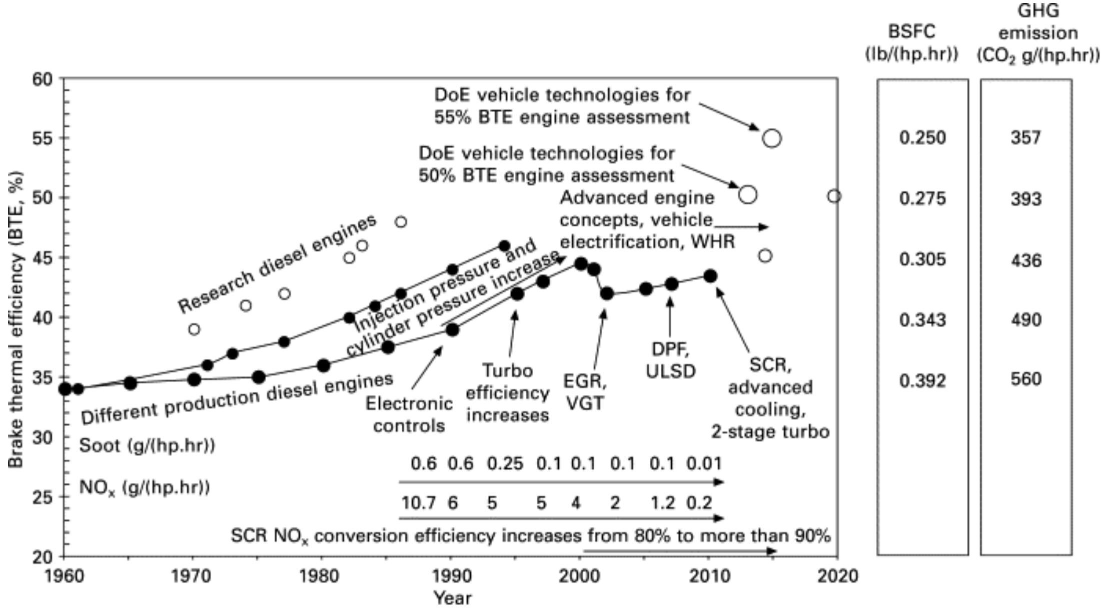
- Identify novel materials
  - And applications for them



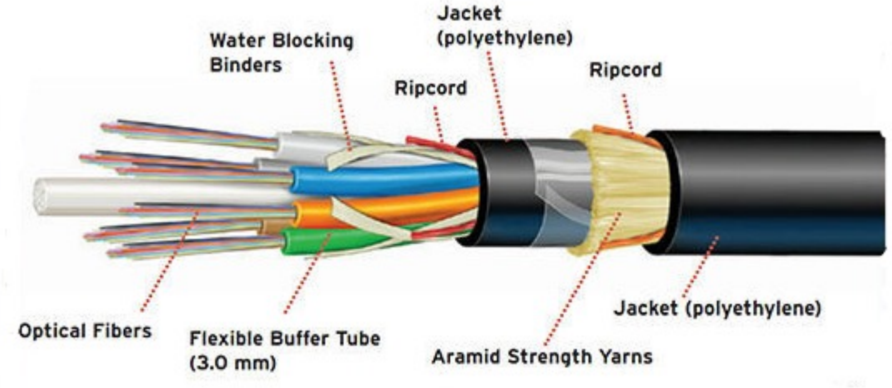
# Typical life of a material scientist/engineer

- Identify novel materials
  - And applications for them
- Improve existing materials
  - Performance, cost, sustainability



# Typical life of a material scientist/engineer

- Identify novel materials
  - And applications for them
- Improve existing materials
  - Performance, cost, sustainability
- Develop better ways to manufacture
  - And "process" or "assemble" components



# Typical life of a material scientist/engineer

- Identify novel materials
  - And applications for them
- Improve existing materials
  - Performance, cost, sustainability
- Develop better ways to manufacture
  - And "process" or "assemble" components
- Prevent or postpone failure
  - Understand mechanisms of failure

# Materials is interdisciplinary

Bridging atoms to rockets, picoseconds to years

Most science in an engineering discipline, most engineering in a science discipline

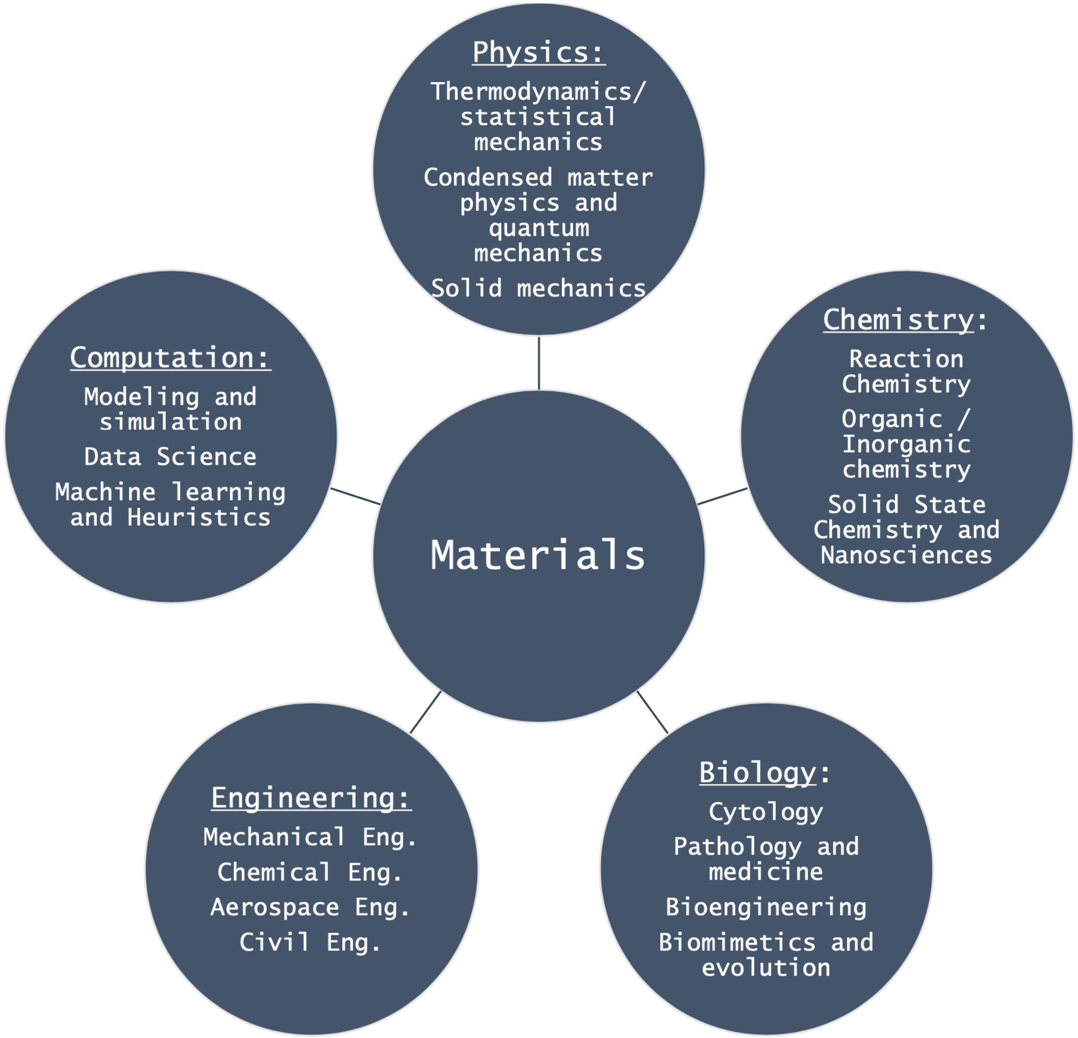
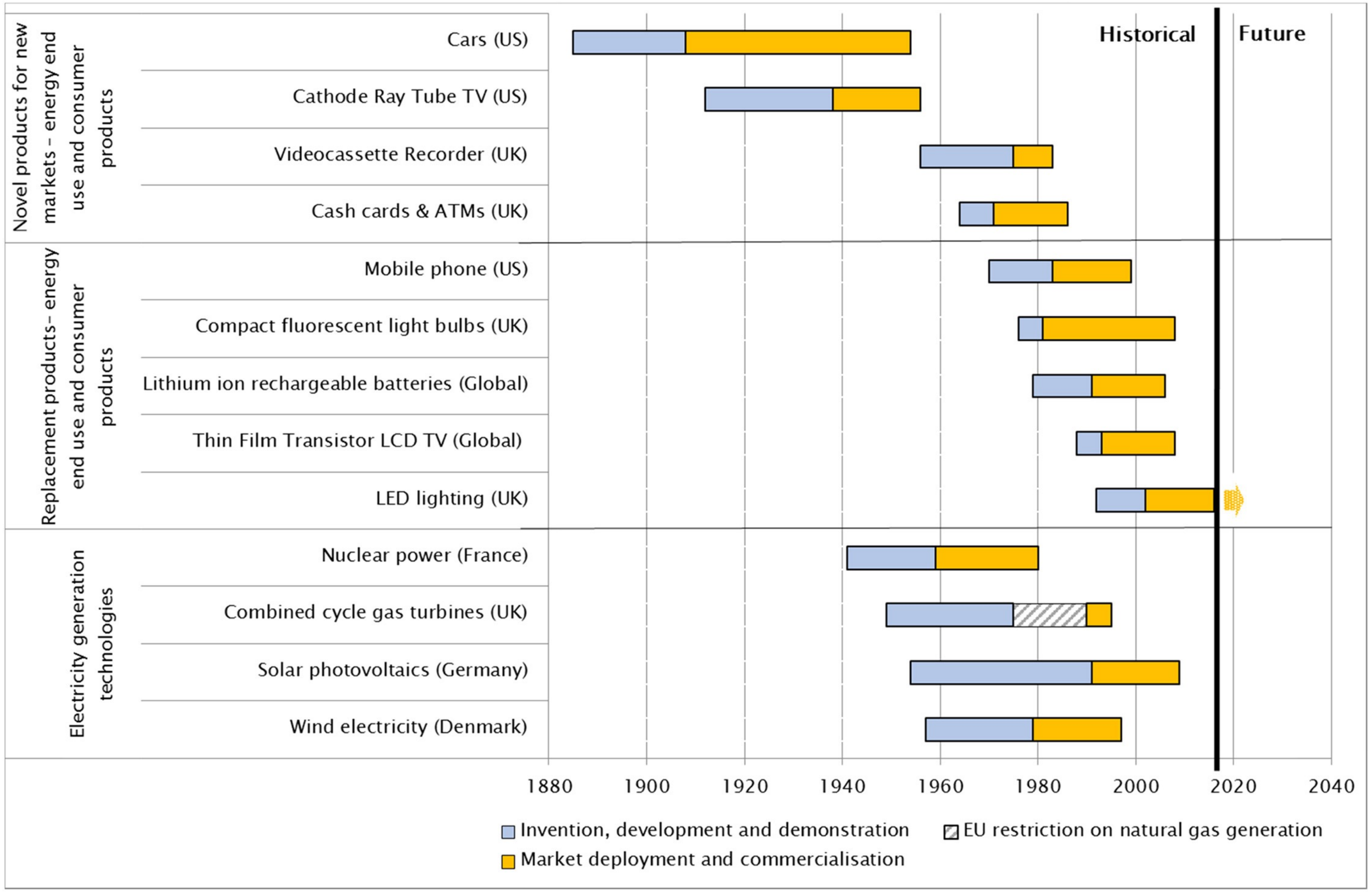



Figure: S Karthikeyan, IISc

# Where does machine learning come in?

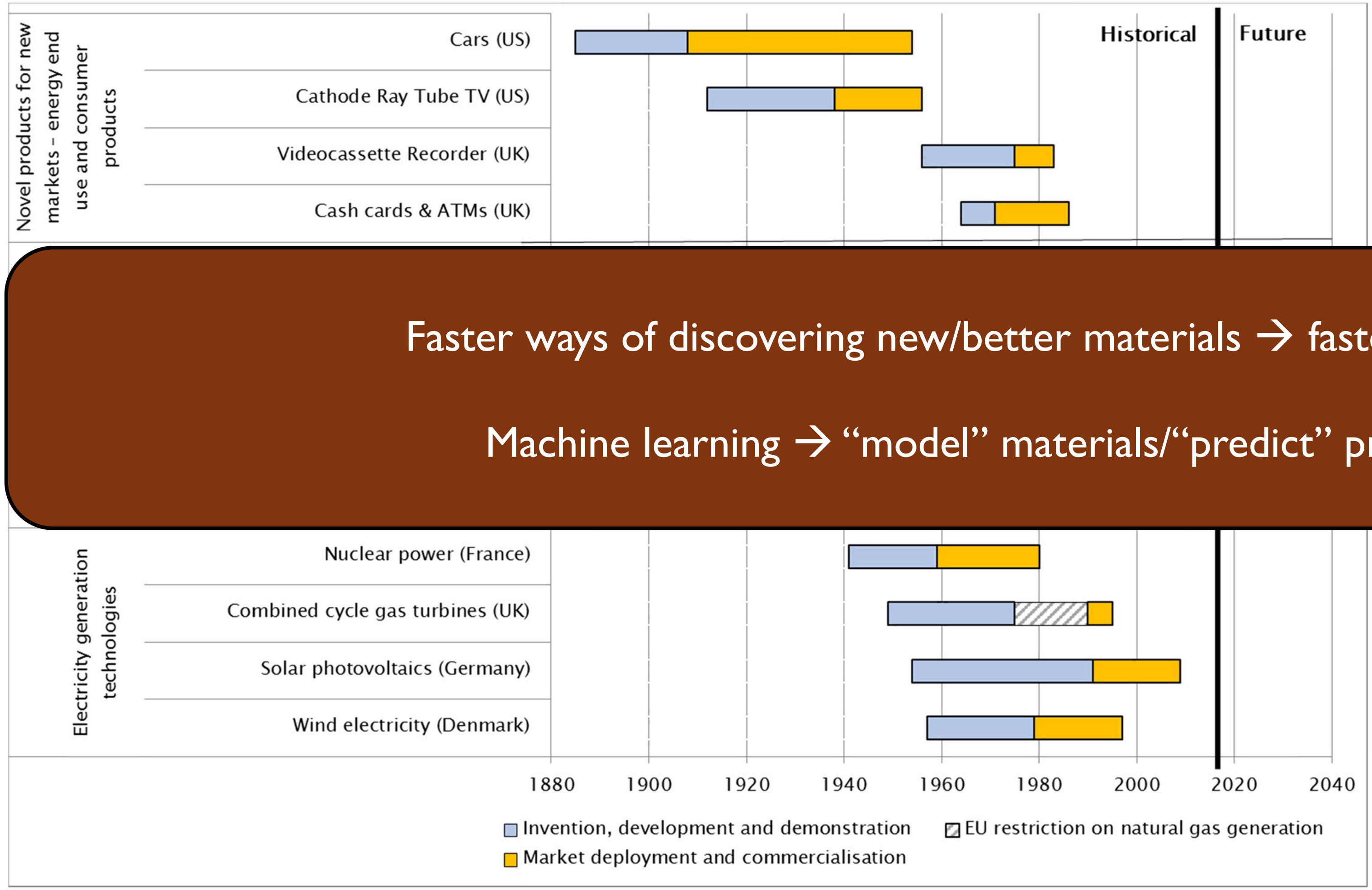
# Why use machine learning (ML) in materials?




Technological innovation and deployment is a 'slow' process: often limited by materials

Innovation is particularly slow in energy sector!

# Why use machine learning (ML) in materials?




Technological innovation and deployment is a 'slow' process: often limited by materials

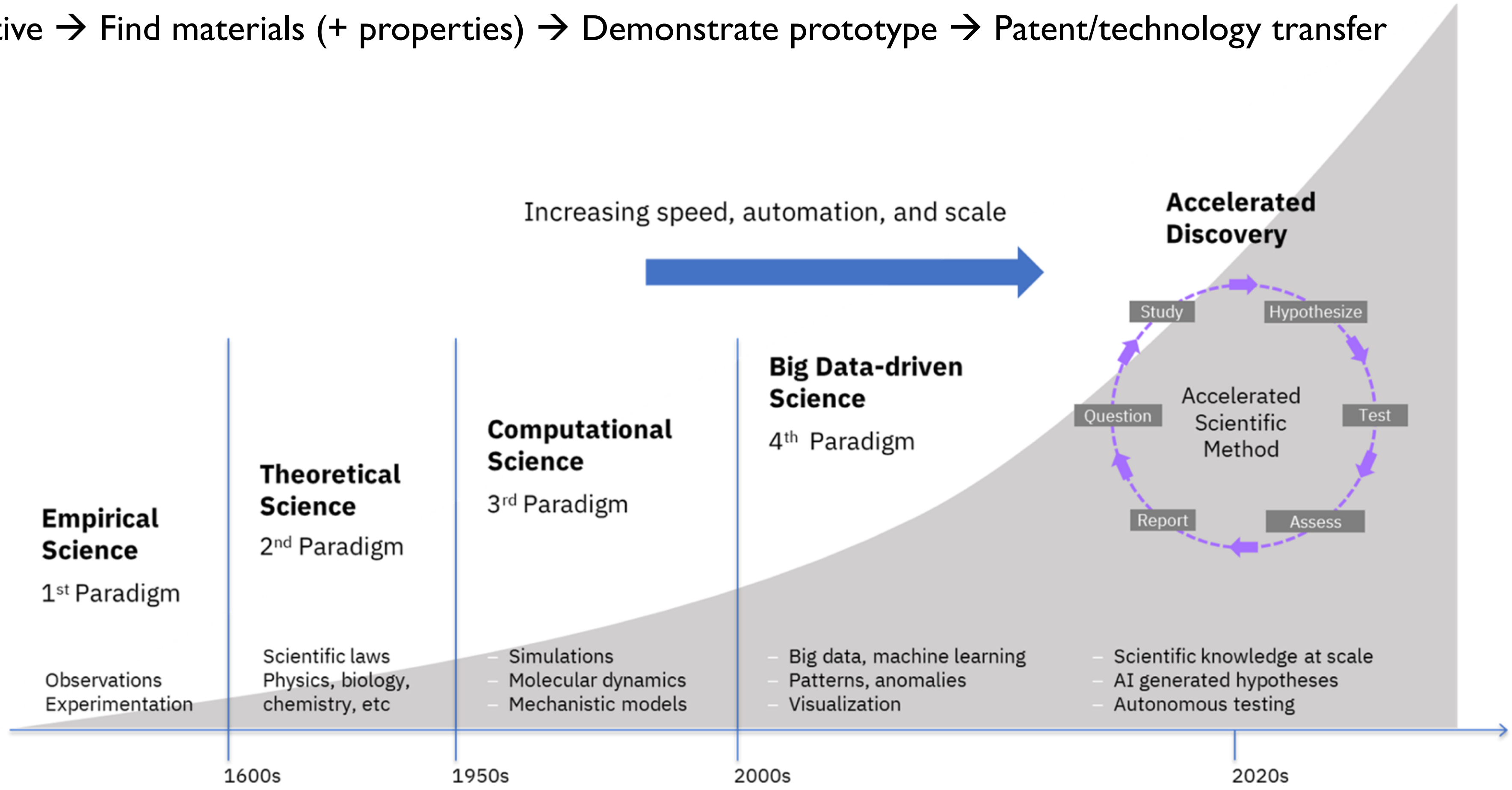
Faster ways of discovering new/better materials → faster innovation cycles

Machine learning → "model" materials/"predict" properties faster

Innovation is particularly slow in energy sector!

# Fourth paradigm of discovery

State objective → Find materials (+ properties) → Demonstrate prototype → Patent/technology transfer

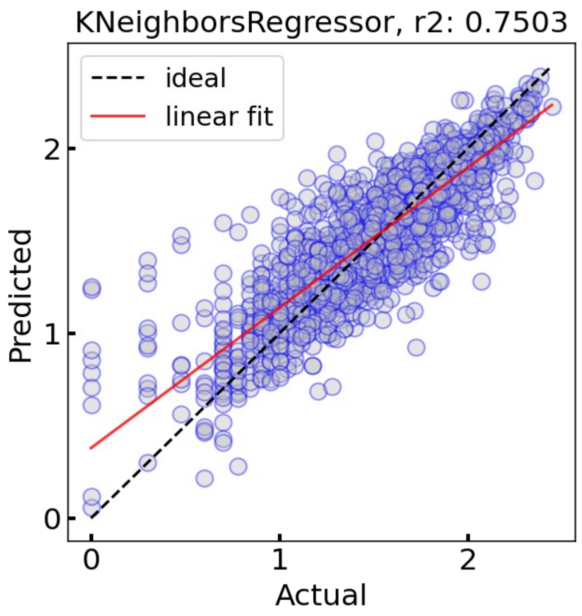




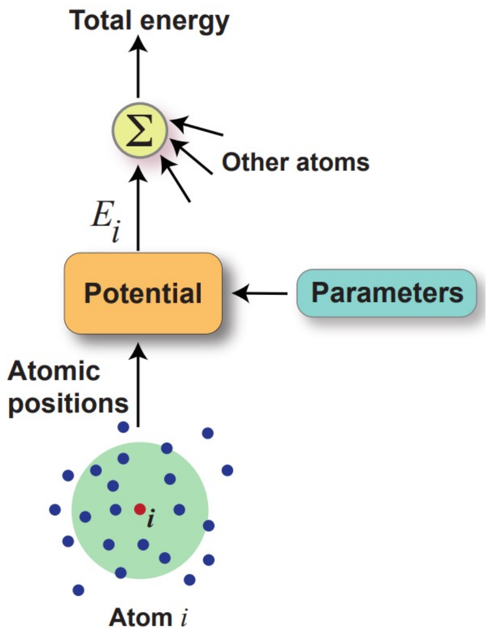

npj Comput. Mater. 8, 84 (2022)

# Types of ML in materials

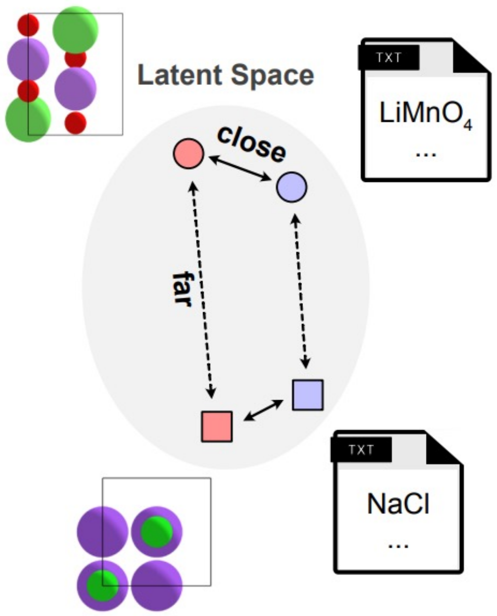
**Regressions**: make property predictions better with 'simple' inputs
(also classifications)




**Interatomic potentials**: describe potential energy surface accurately

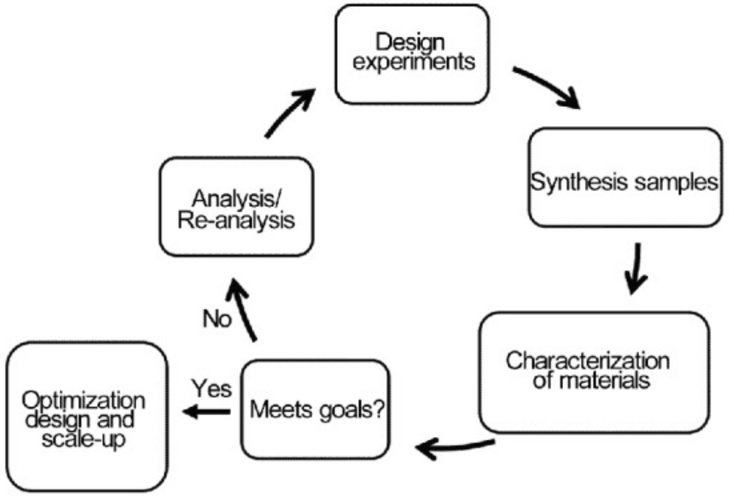



**Advanced models:**
Diffusion (generative) models, language models, transfer learning

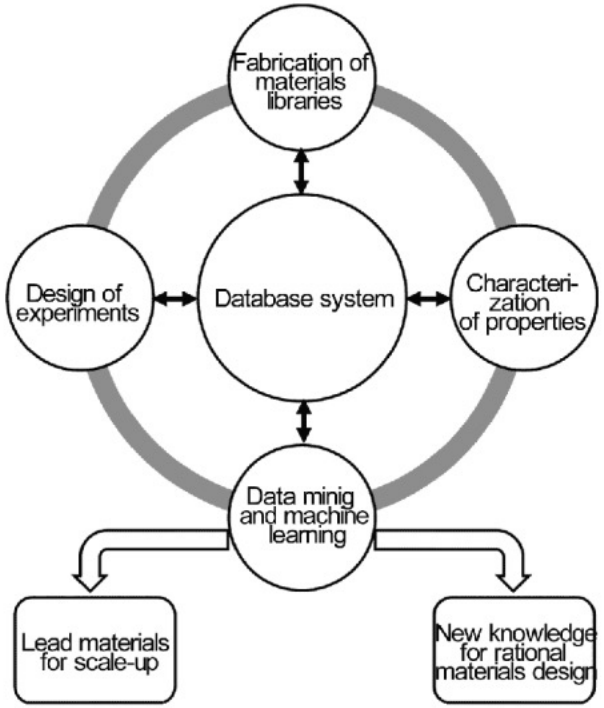


# Where does the data come from?


Data organization: python/API
ML: python




Liu et al., Sci. China Tech. Sci. 62, 4 (2019)



# Where does the data come from?


Home | Search

Home | Benchmark Info | Full Benchmark Data | How To Use | Leaderboards Per Task | Reference

Home

## Leaderboard-Property: General Purpose Algorithms on `matbench_v0.1`

Find more information about this benchmark on the benchmark info page

| Task name | Samples | Algorithm | Verified MAE (unit) or ROCAUC | Notes |
|---|---|---|---|---|
| matbench_steels | 312 | MODNet (v0.1.12) | 87.7627 (MPa) | |
| matbench_jdft2d | 636 | MODNet (v0.1.12) | 33.1918 (meV/atom) | |
| matbench_phonons | 1,265 | MegNet (kgcnn v2.1.0) | 28.7606 (cm^-1) | structure required |
| matbench_expt_gap | 4,604 | MODNet (v0.1.12) | 0.3327 (eV) | |
| matbench_dielectric | 4,764 | MODNet (v0.1.12) | 0.2711 (unitless) | |
| matbench_expt_is_metal | 4,921 | AMMExpress v2020 | 0.9209 | |
| matbench_glass | 5,680 | MODNet (v0.1.12) | 0.9603 | |
| matbench_log_gvrh | 10,987 | coNGN | 0.0670 (log10(GPa)) | structure required |
| matbench_log_kvrh | 10,987 | coNGN | 0.0491 (log10(GPa)) | structure required |
| matbench_perovskites | 18,928 | coGN | 0.0269 (eV/unit cell) | structure required |
| matbench_mp_gap | 106,113 | coGN | 0.1559 (eV) | structure required |
| matbench_mp_is_metal | 106,113 | CGCNN v2019 | 0.9520 | structure required |
| matbench_mp_e_form | 132,752 | coGN | 0.0170 (eV/atom) | structure required |

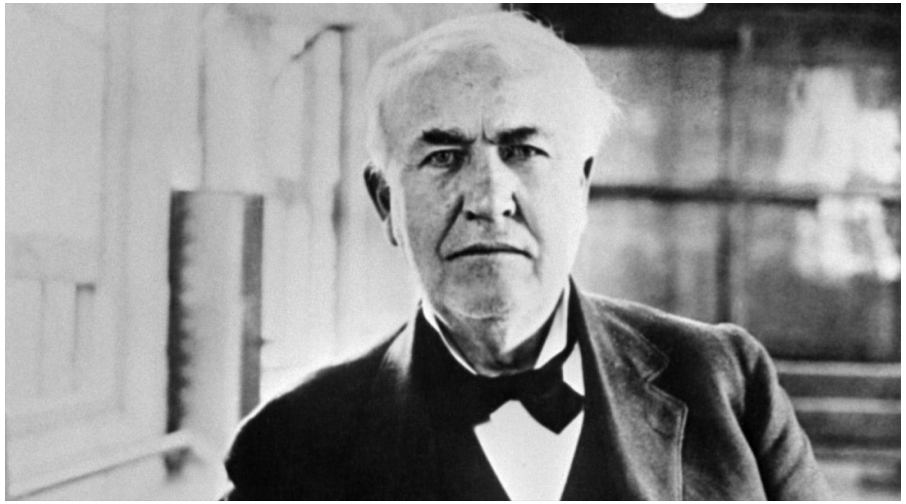
https://matbench.materialsproject.org/

ICSD

NIST National Institute of Standards and Technology

Data organization: python/API
ML: python

# How do we design materials?



**Trial and error** of candidates in a lab

**Simulate and identify** candidates
(on a transparent touch screen preferably)


Density functional theory: (Approximately) predict material properties

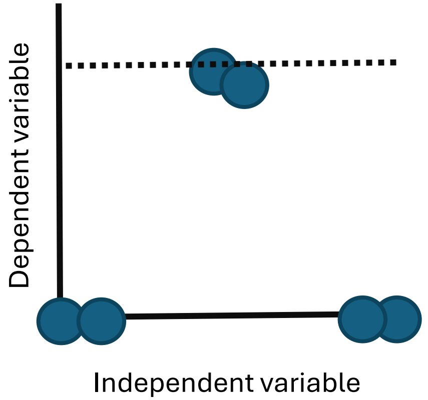
Machine learning: learn from predictions to make better predictions

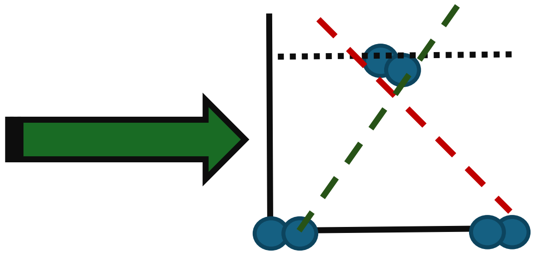
# (Modern) ML in materials and use cases

## Property predictors, interatomic potentials, advanced models

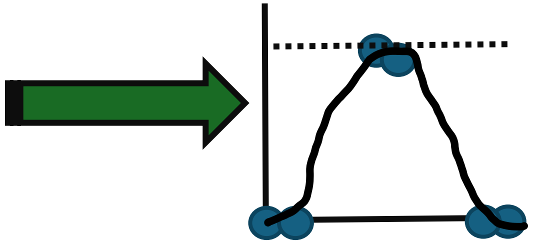
# Neural networks (NNs)

Suppose we want to fit the following data



Fitting a simple regression model

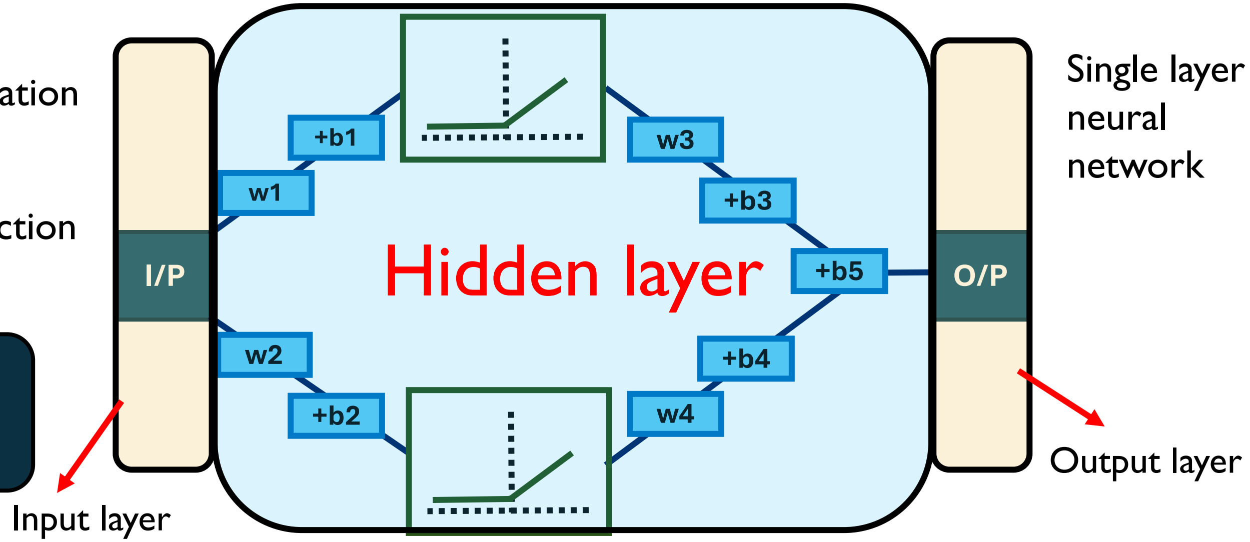


Neural networks can fit a squiggle

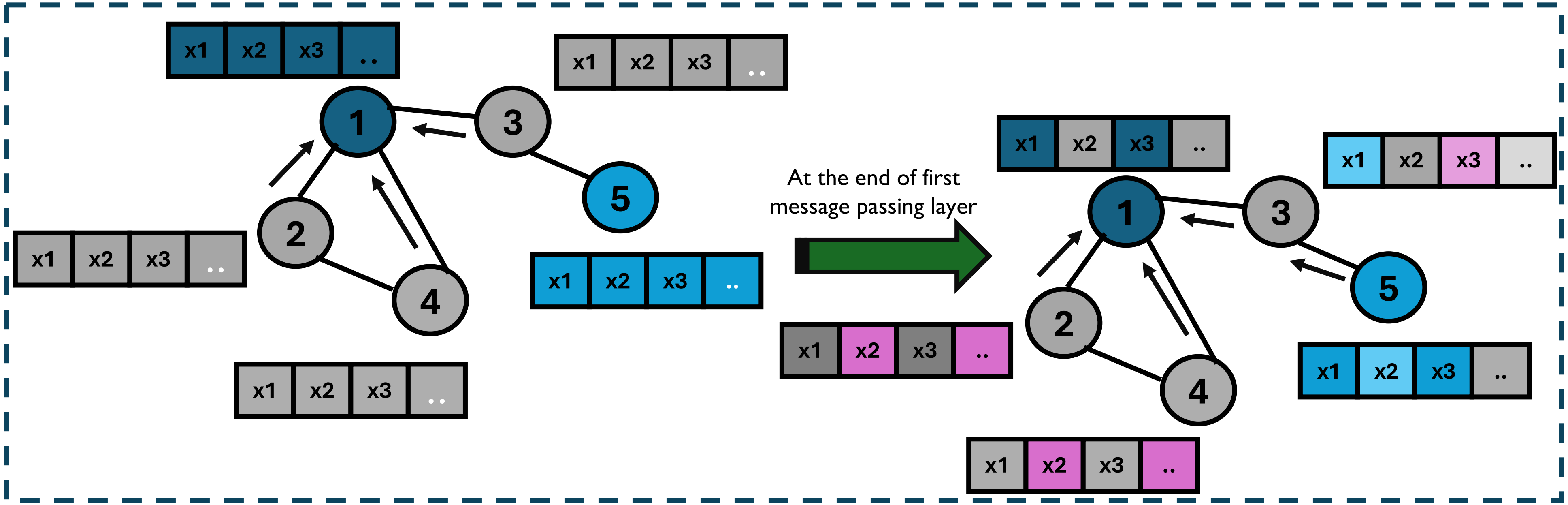
Optimized biases and weights are obtained via back propagation

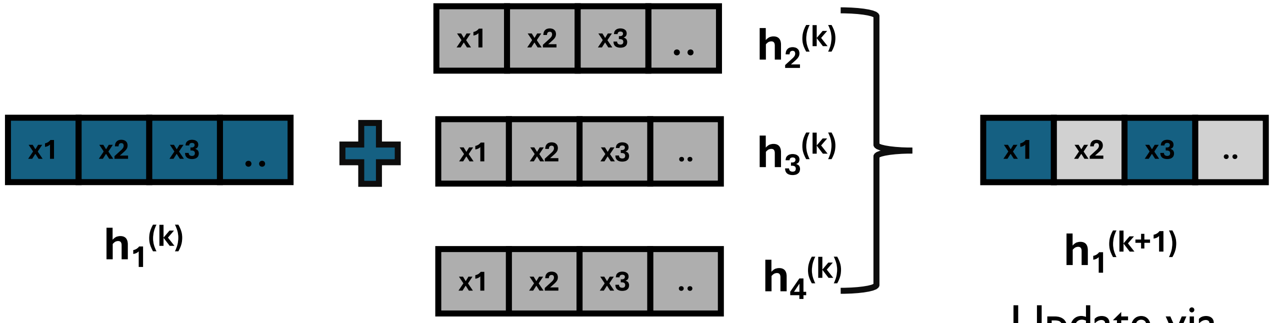
Weights and biases determine the part of the activation function that will contribute to the squiggle

Several types of NNs exist
Graph NNs particularly relevant for materials



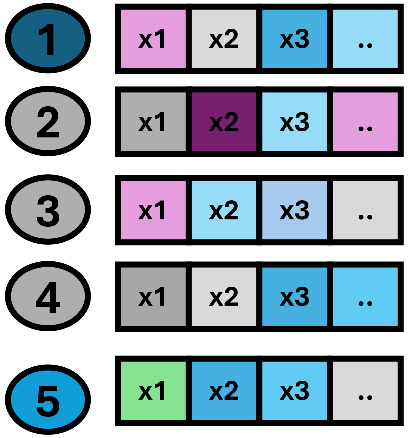
# Graph neural networks




Aggregate

Update via activation



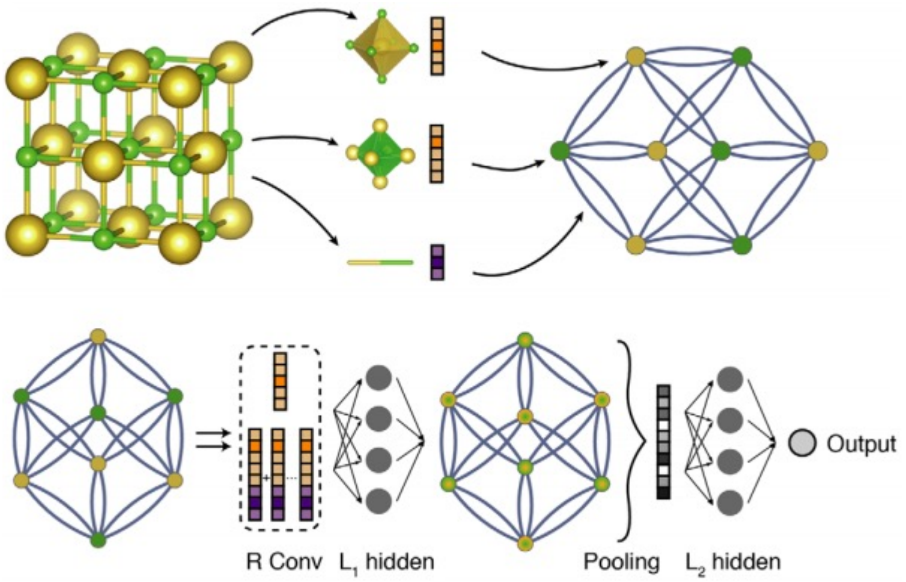

After multiple message passing layers

# Predicting material properties (scratch)

PHYSICAL REVIEW LETTERS **120**, 145301 (2018)

**Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties**

Tian Xie and Jeffrey C. Grossman
*Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*




**Properties** : Formation energy, band gap, Fermi energy, bulk and shear moduli, and Poisson's ratio

**Database**: $10^4$ DFT-calculated datapoints from MP

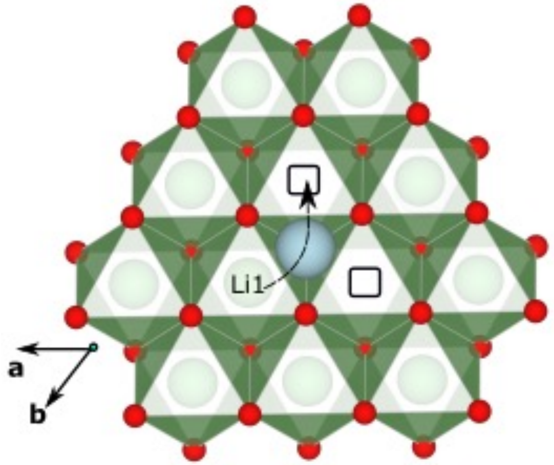
**Model**: Crystal Graph convolutional neural network (CGCNN)

**Performance**:
- Formation energy: 0.039 eV/atom
- Band gap: 0.388 eV
- Fermi energy: 0.363 eV
- Elastic moduli: ~1-2 GPa
- Poisson's ratio: 0.03
- Identified 228 'synthesizable' perovskites out of 18928 in the training database

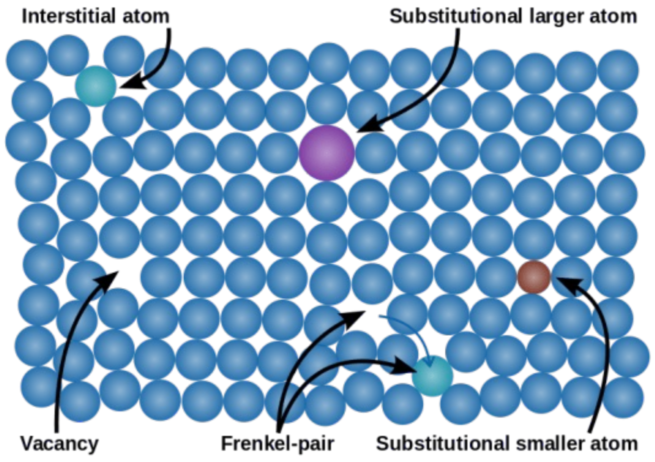
# Predicting material properties (transfer learning)

Several key material properties that govern performance in applications have limited data

- ‘Small’ datasets (< $10^4$ datapoints)
  - Ionic mobilities, defect formation energies, adsorption energies,…
- Limits application of deep learning (DL) frameworks

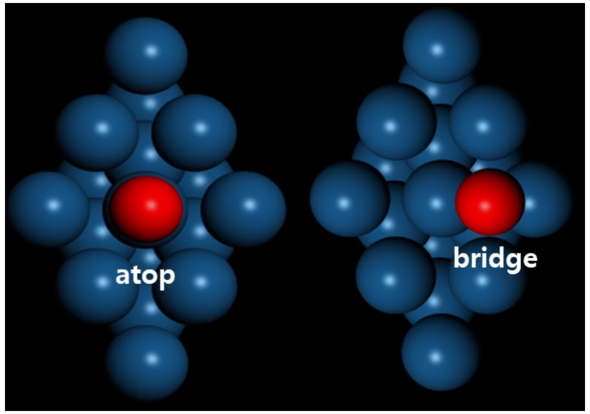



Devi et al., **npj Comput. Mater.** 2022




https://www.differencebetween.com/difference-between-point-defect-and-line-defect/

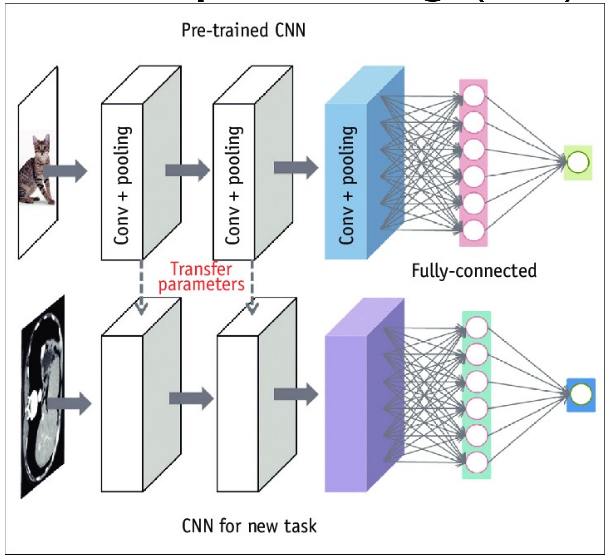
https://sites.psu.edu/



# Predicting material properties (transfer learning)

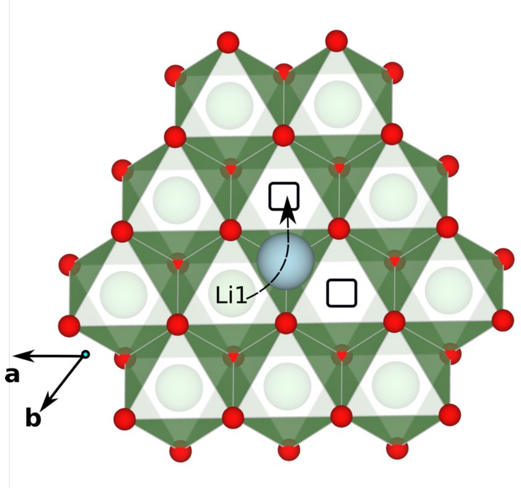
Several key material properties that govern performance in applications have limited data

- ‘Small’ datasets (< $10^4$ datapoints)
  - Ionic mobilities, defect formation energies, adsorption energies,…
- Limits application of deep learning (DL) frameworks




Do et al., **Korean J. Radiol.** 2020

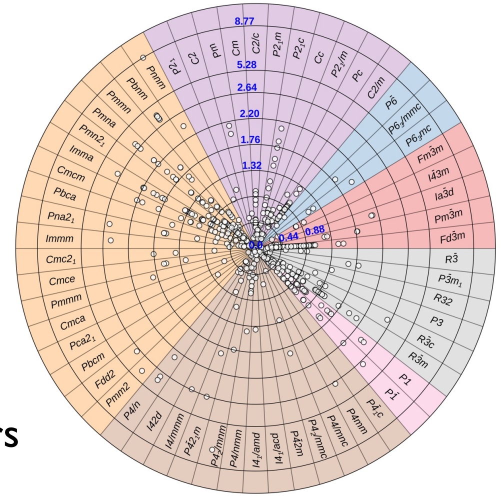
# Transfer learning: battery rate performance

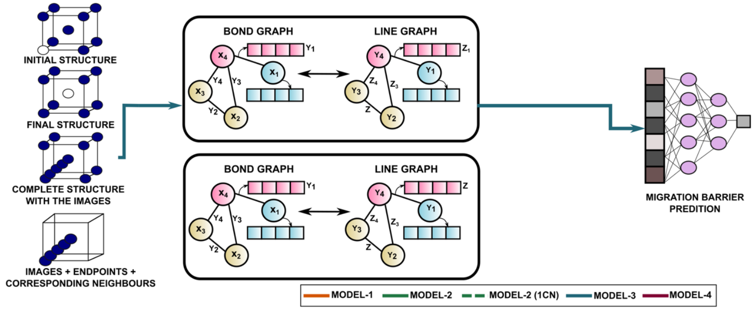



Migration barriers → determine rate performance in batteries

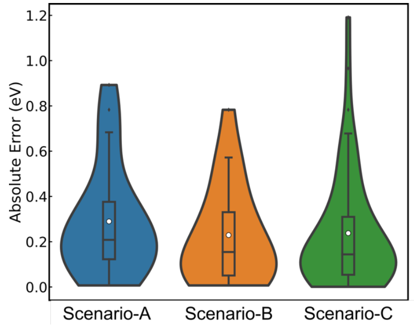
Exponential control on diffusivity

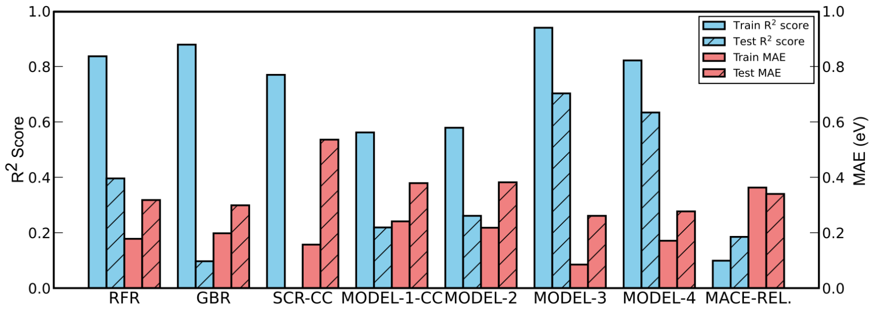
Compile a dataset on calculated barriers






Build a model (graph based architecture)



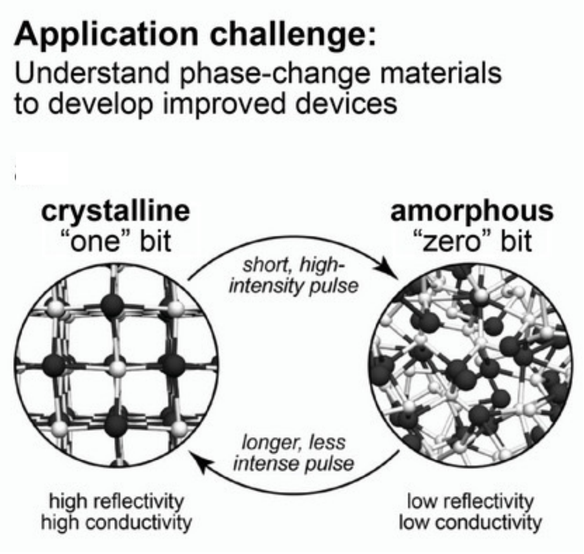

MODEL-3 (transfer learned) is the best!

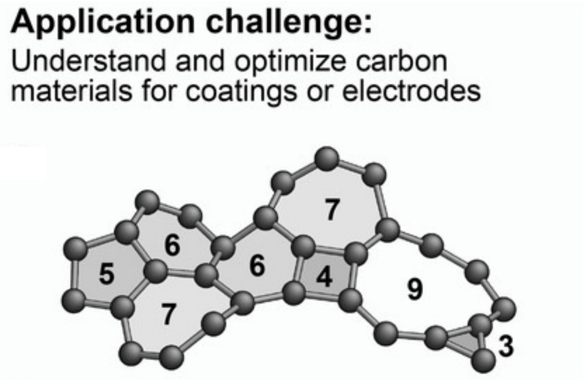
Excellent generalization and classification abilities

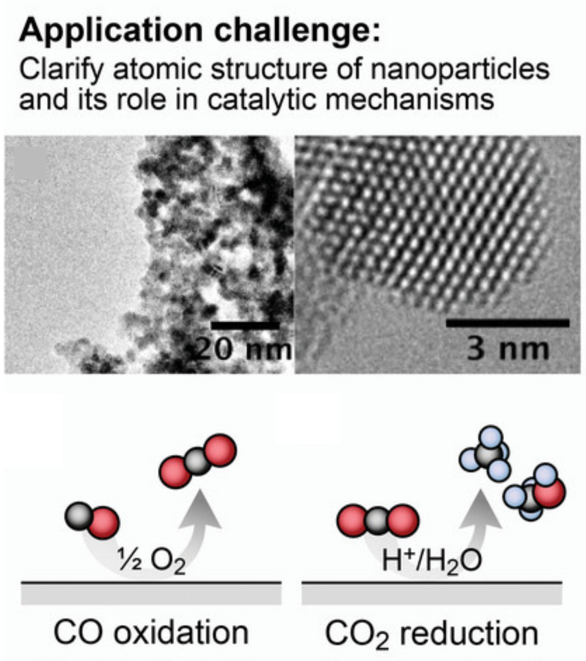


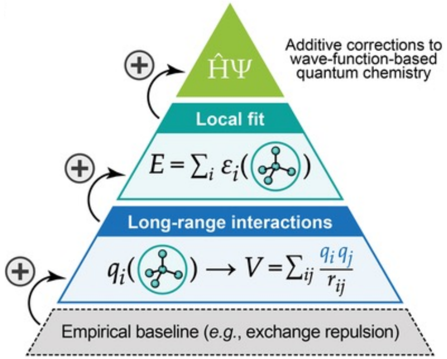
R. Devi, K.T. Butler, and G.Sai Gautam, **arXiv 2025**

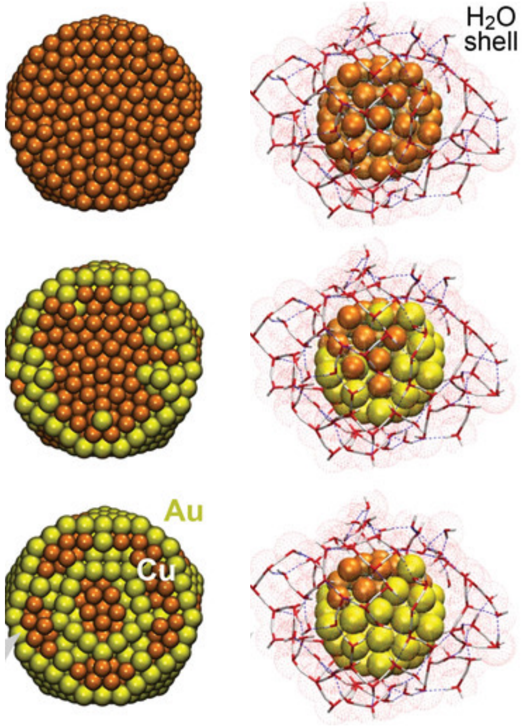
# Graph networks and interatomic potentials

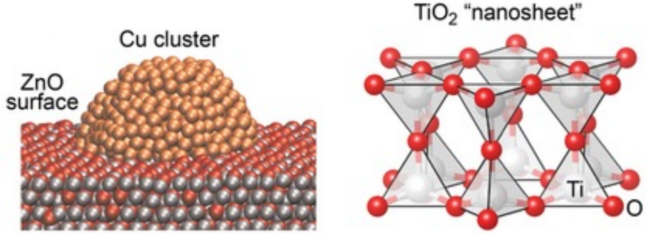


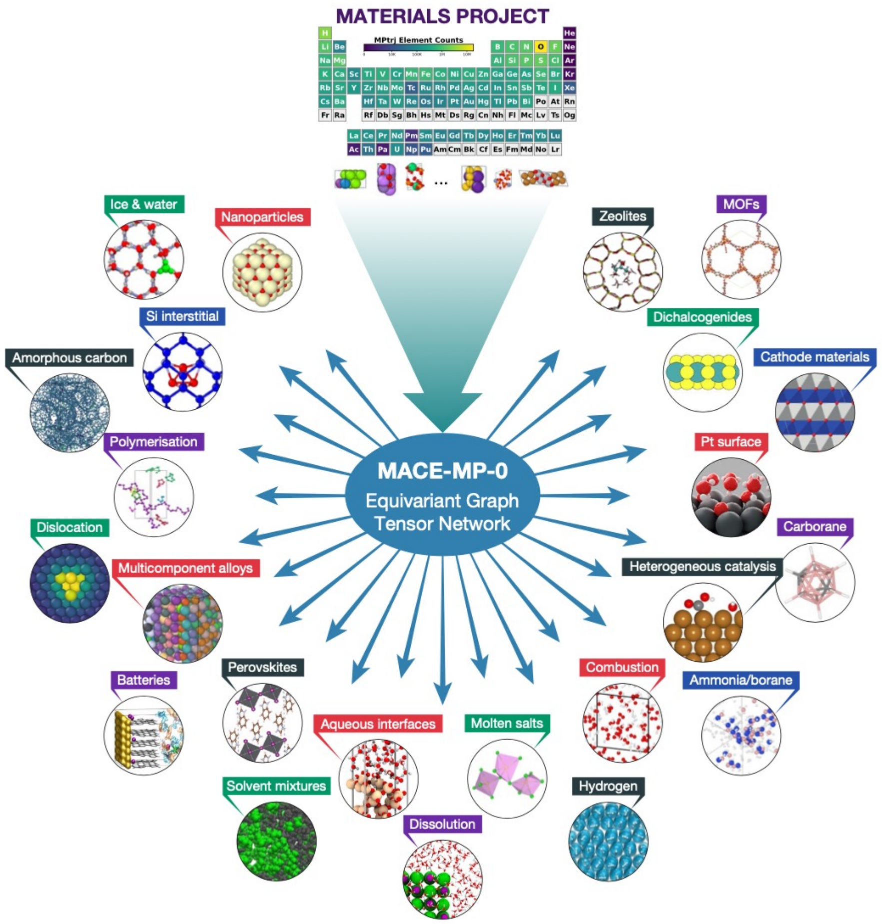



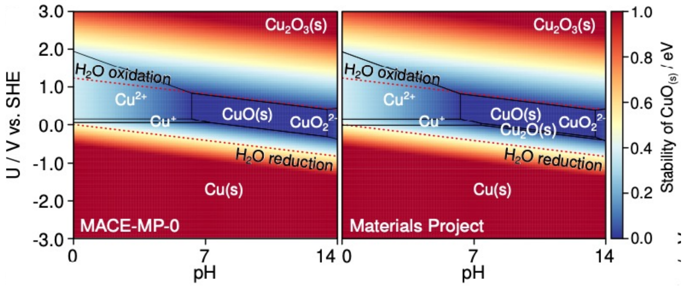

Adv. Mater. 31, 1902765 (2019)

# Foundational interatomic potentials




Trained on Materials Project trajectory dataset (~1.5M structures)

- Stable performance on 30 different property predictions/application areas
- Stable dynamics in solids, liquids, and gases
- GPU; limited system size



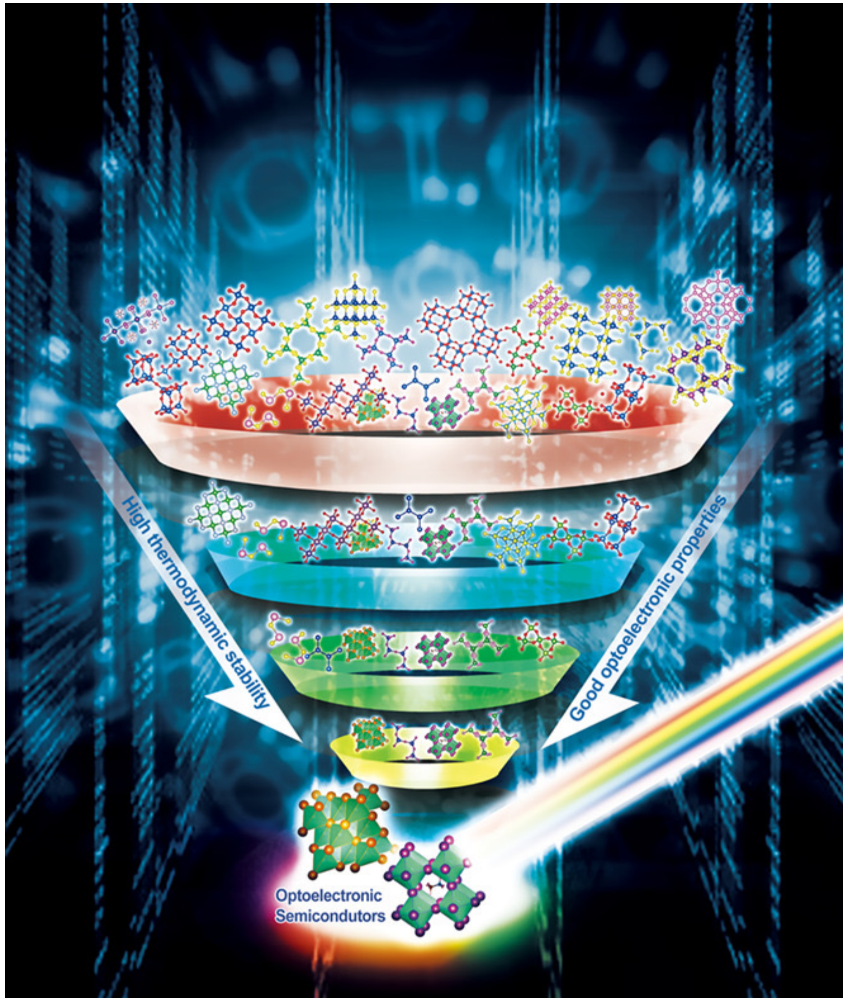

Many more to come…

| Model ⓘ | F1 ↑ | DAF ↑ | Prec ↑ | Acc ↑ | MAE ↓ | $R^2$ ↑ | $\kappa_{SRME}$ ↓ | Training Set |
|---|---|---|---|---|---|---|---|---|
| eqV2 S DeNS | 0.815 | 5.042 | 0.771 | 0.941 | 0.036 | 0.788 | 1.665 | 146k (1.58M) (MPtrj) |
| ORB MPtrj | 0.765 | 4.702 | 0.719 | 0.922 | 0.045 | 0.756 | 1.725 | 146k (1.58M) (MPtrj) |
| SevenNet-l3i5 | 0.76 | 4.629 | 0.708 | 0.92 | 0.044 | 0.776 | 0.55 | 146k (1.58M) (MPtrj) |
| SevenNet-0 | 0.724 | 4.252 | 0.65 | 0.904 | 0.048 | 0.75 | 0.767 | 146k (1.58M) (MPtrj) |
| GRACE-2L (r6) | 0.691 | 4.163 | 0.636 | 0.896 | 0.052 | 0.741 | 0.525 | 146k (1.58M) (MPtrj) |
| MACE-MP-0 | 0.669 | 3.777 | 0.577 | 0.878 | 0.057 | 0.697 | 0.647 | 146k (1.58M) (MPtrj) |
| CHGNet | 0.613 | 3.361 | 0.514 | 0.851 | 0.063 | 0.689 | 1.717 | 146k (1.58M) (MPtrj) |
| M3GNet | 0.569 | 2.882 | 0.441 | 0.813 | 0.075 | 0.585 | 1.412 | 62.8k (188k) (MPF) |

https://matbench-discovery.materialsproject.org/

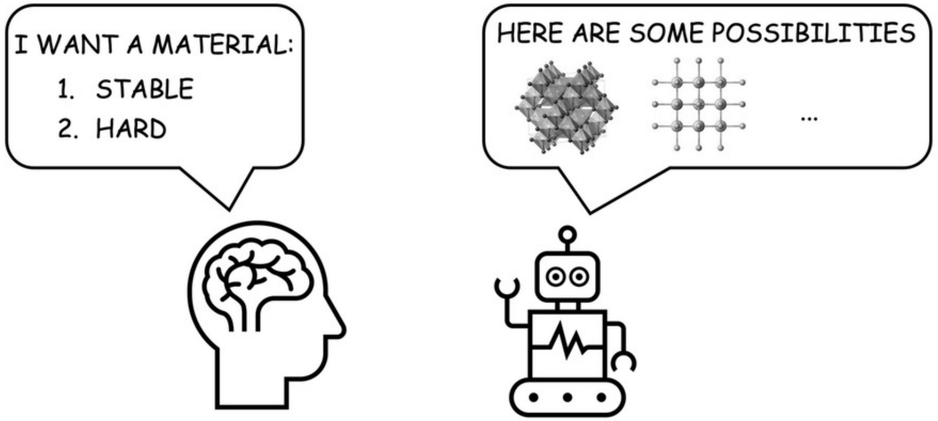

Batatia et al., arXiv, 2401.00096v2 (2024)

# Graph networks and inverse material design




https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/wcms.1489

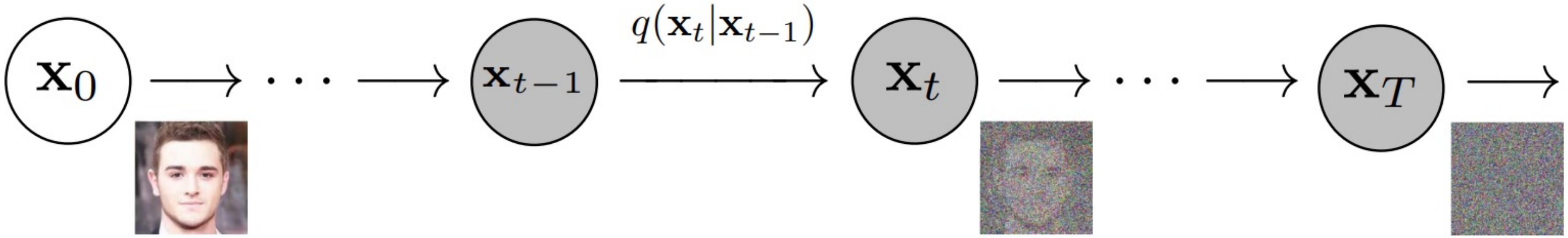
Property → Structure



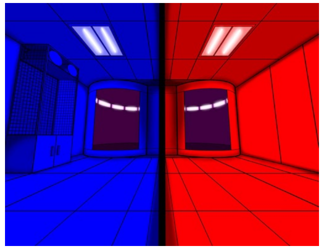

https://news.mit.edu/2022/new-way-perform-general-inverse-design-high-accuracy-0118

# Use diffusion for *structure* generation

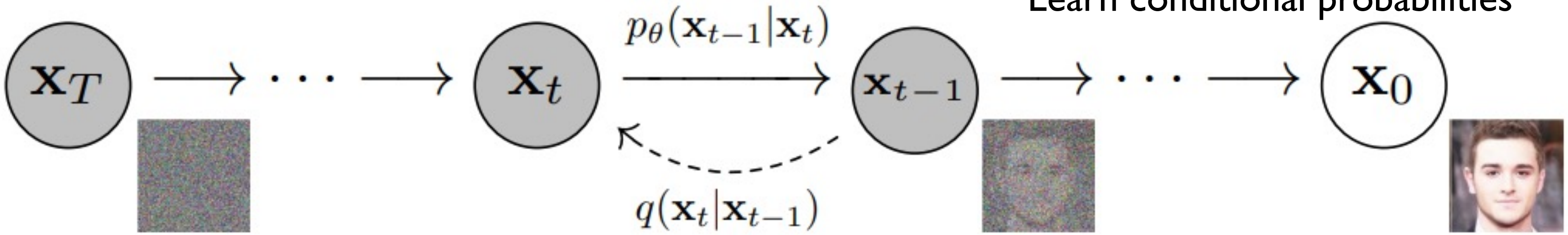
Forward diffusion




Gaussian noise added in a Markov chain



Learn conditional probabilities

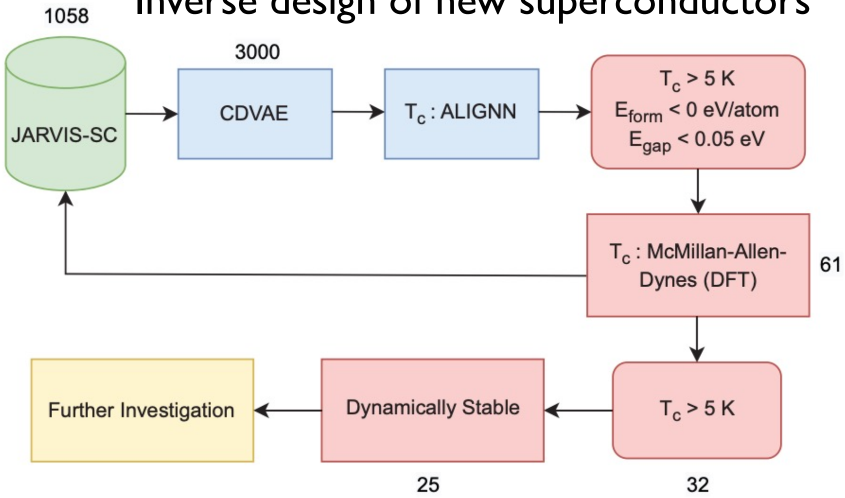



Reverse diffusion

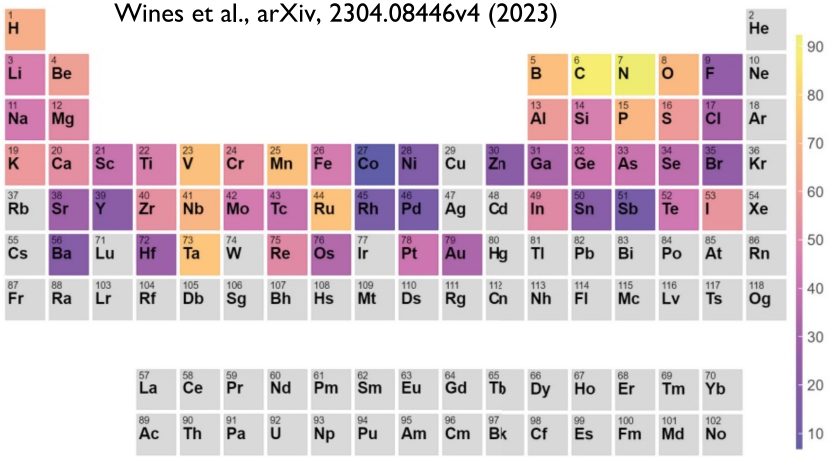
https://www.assemblyai.com/blog/diffusion-models-for-machine-learning-introduction https://medium.com/@luisfelipechary/my-experience-with-diffusion-super-resolution-3386b6574696

# Examples of generative models

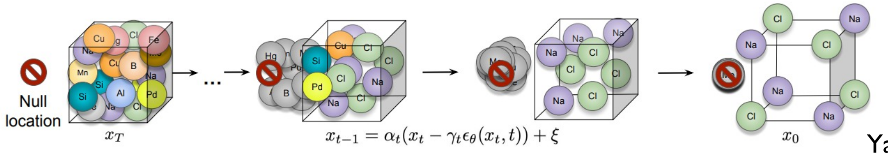
## Inverse design of new superconductors




Wines et al., arXiv, 2304.08446v4 (2023)

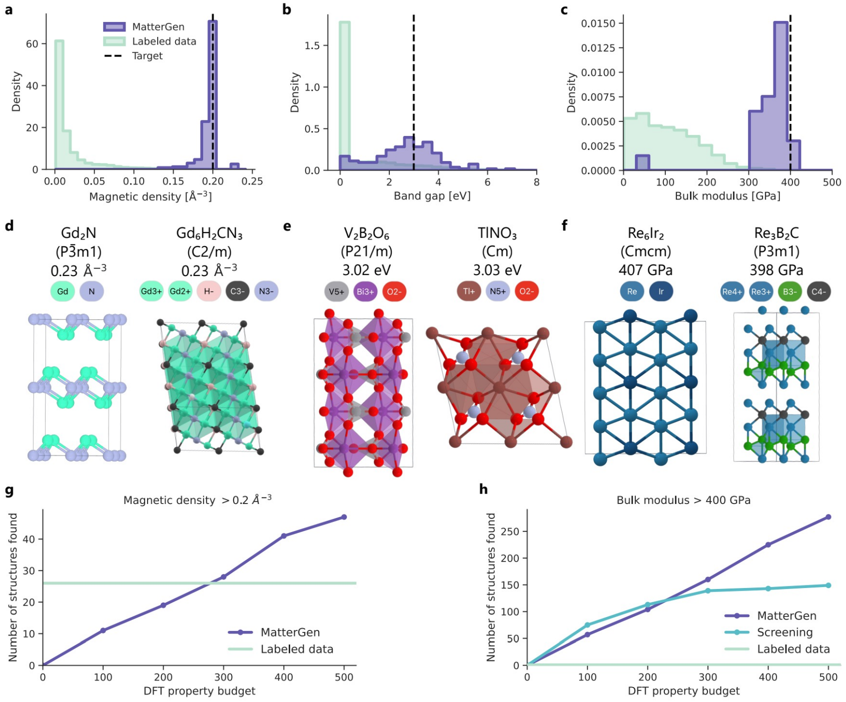



## UniMat + Diffusion (Google Deepmind)




Yang et al., arXiv, 2311.09235v2 (2023)

## MatterGen (Microsoft)




Zeni et al., arXiv, 2312.03687v2 (2024)

# Conclusions

- Designing better materials critical for performance improvement in several applications
  - Computations + ML can significantly accelerate materials design
- Different ways to use ML
  - Property predictions, interatomic potentials, structure generation
- Materials science is a data-limited domain
  - Garbage in = Garbage out; data normalization
  - Real vs. synthetic data
  - What model to choose? Simple models are usually better
  - 'Real' success stories: still few, possibly in development
  - Lots of ongoing work: exciting field!
- General advice
  - Don't do ML just because you can
  - Construct models with care: overfitting, lack of transferability
  - Test and validate, validate and test, and …


saigautamg@iisc.ac.in; https://sai-mat-group.github.io; https://github.com/sai-mat-group