

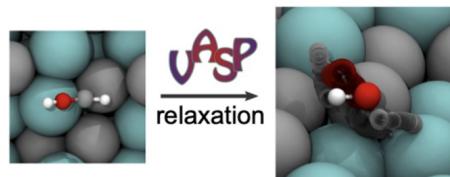


# Hands-on 8 : MLP01 - adsorption configuration optimization

## ▼ Time scale limitation of DFT

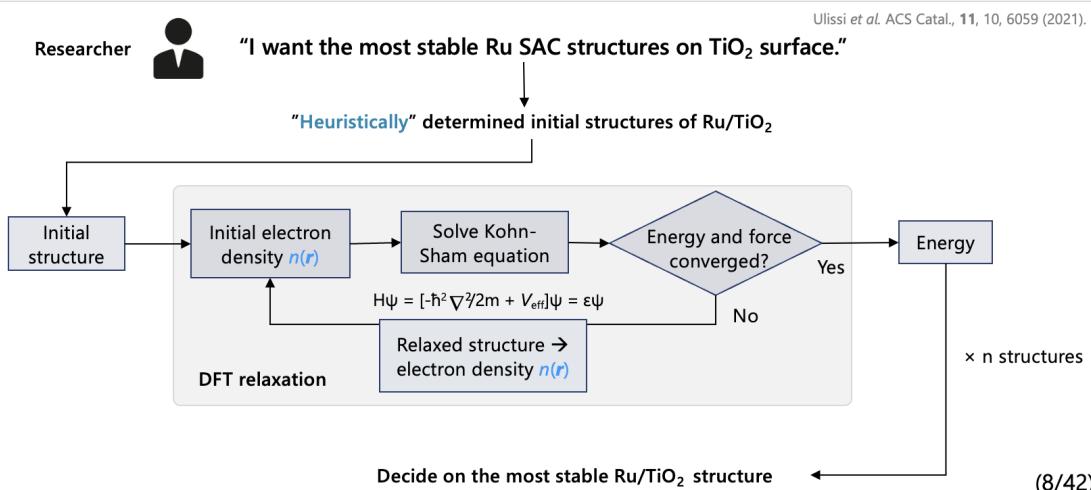
### Challenges of Theoretical Modeling

#### Huge computational cost of DFT calculations

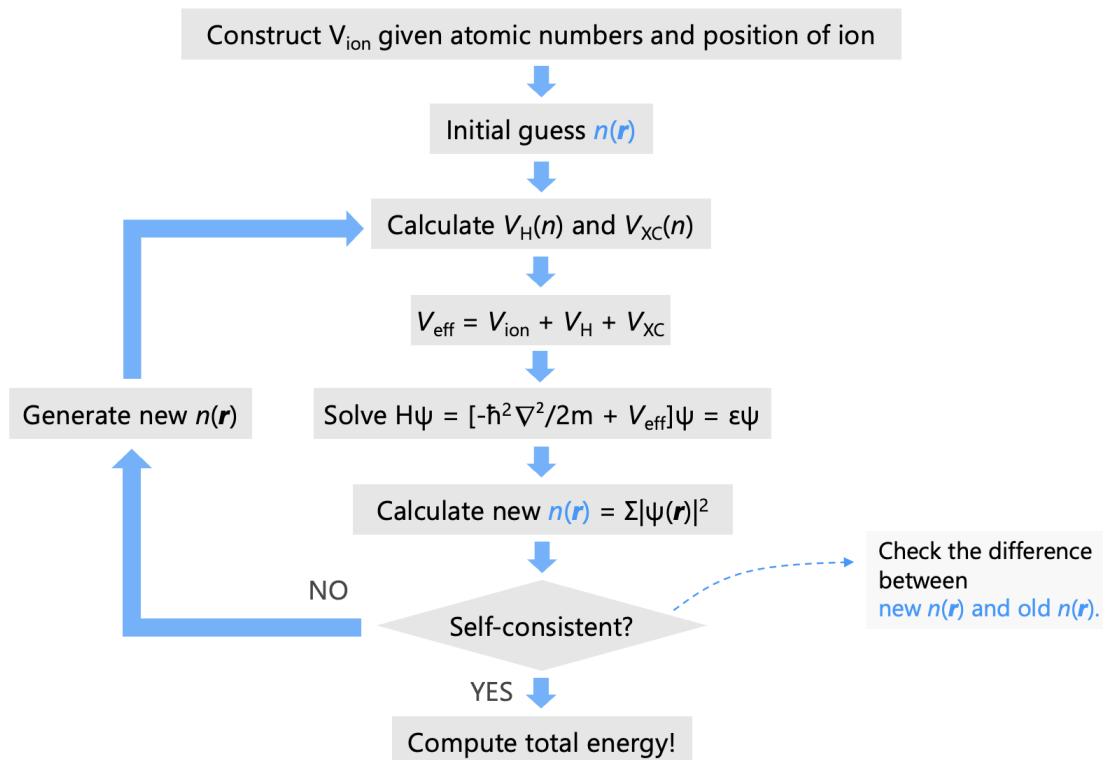


- ✓ The DFT computation scales  $O(n^3)$  with the number of electrons in the system.
- ✓ The DFT computations of 100 atoms approximately take ~1 hour per relaxation on 12 core CPUs.  
→ few hours~ day for get energy of one DFT structure

Ulissi et al. ACS Catal., 11, 10, 6059 (2021).



## DFT calculation optimization process



## ▼ What is Machine learning Potential?

MLP review:



Review

Rise of machine learning potentials in heterogeneous catalysis:  
Developments, applications, and prospects

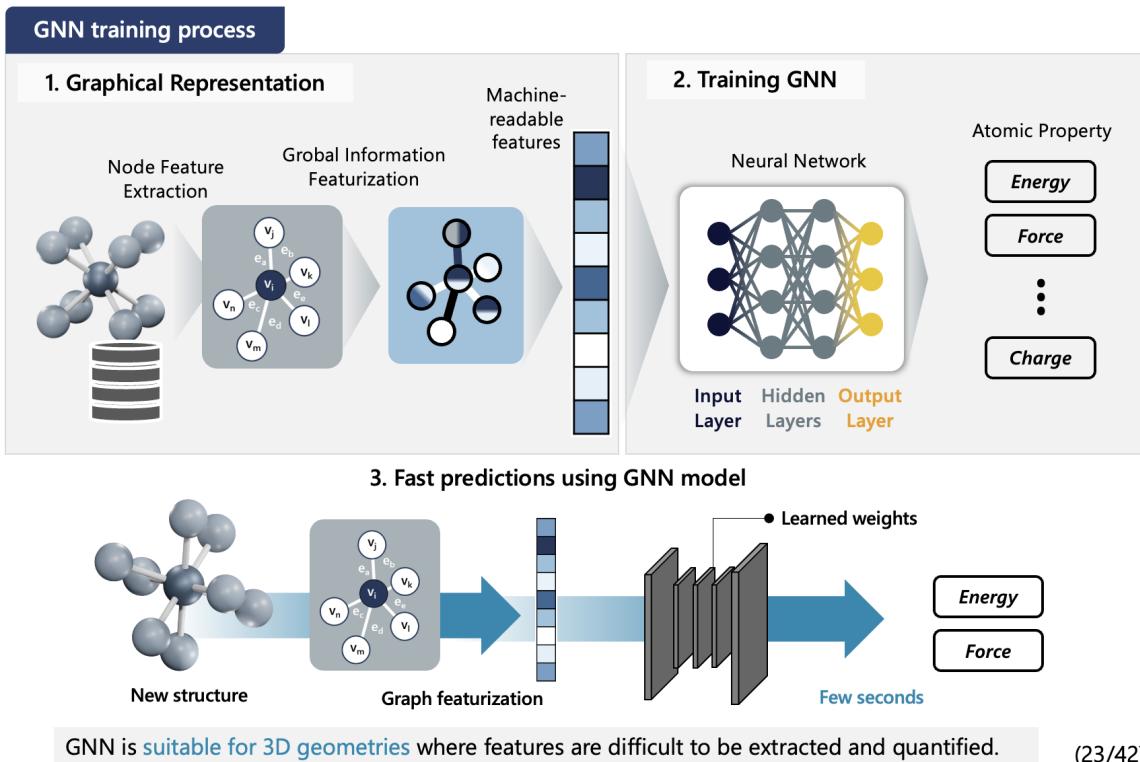
Seokhyun Choung<sup>1</sup>, Wongyu Park<sup>1</sup>, Jinuk Moon<sup>1</sup>, Jeong Woo Han<sup>\*</sup>

Department of Materials Science and Engineering, Research Institute of Advanced Materials, Seoul National University, Seoul, 08826, 1 Gwanak-ro, Gwanak-gu, Republic of Korea



<https://www.sciencedirect.com/science/article/pii/S138589472404244X#da005>

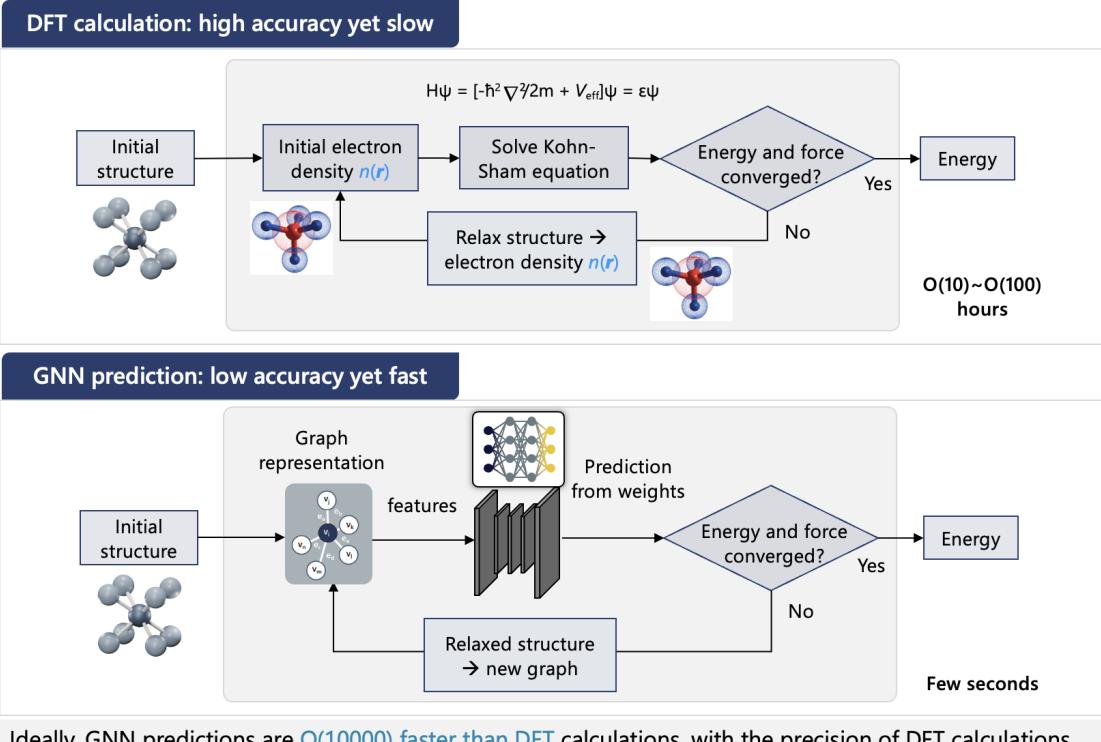
# Graph Neural Network (GNN)



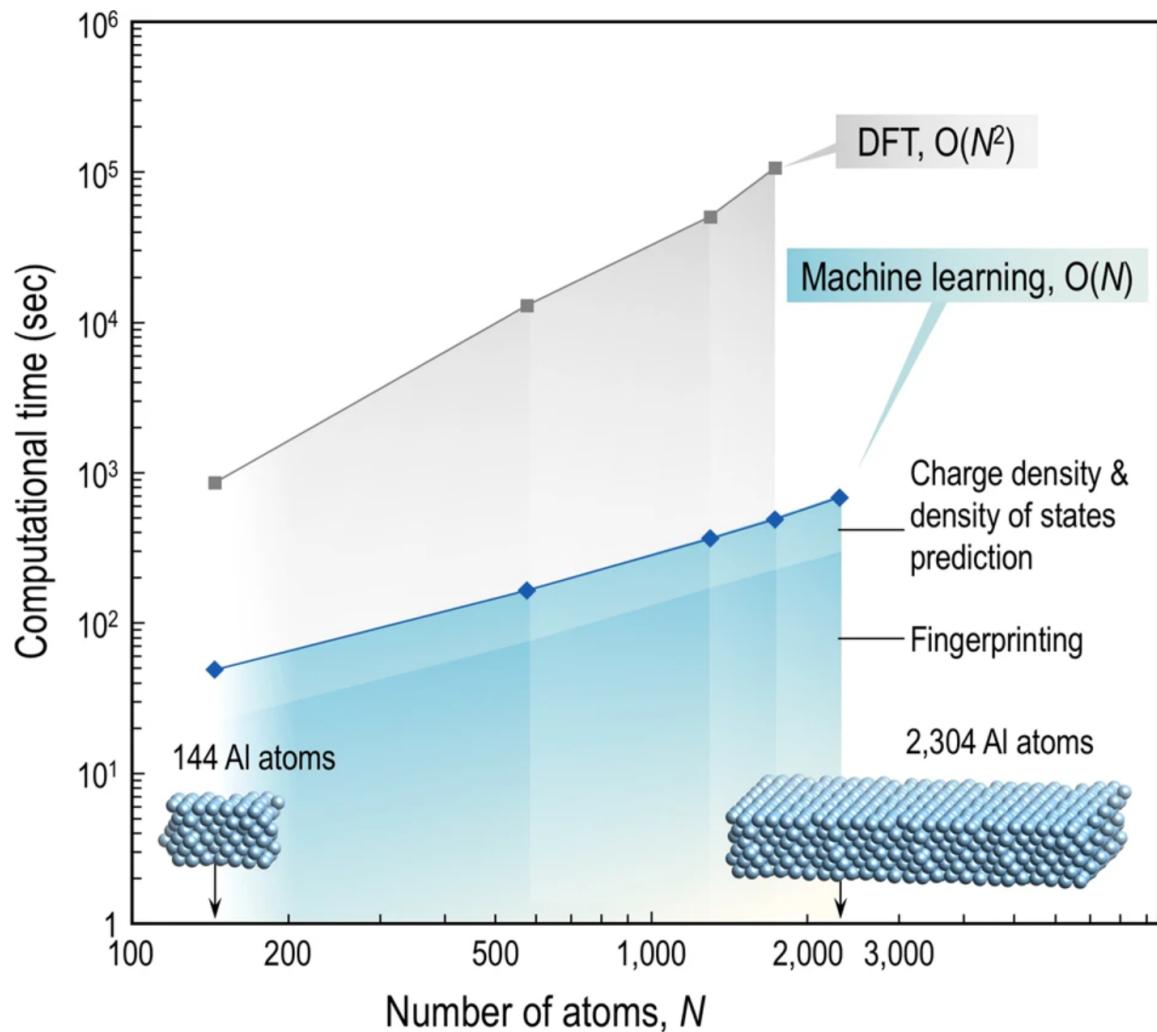
## ▼ Why we use MLP?

Computational cost of MLP compared to DFT

## GNN can Ideally Replace DFT Calculations



Length scale and computational cost of MLP compared to DFT

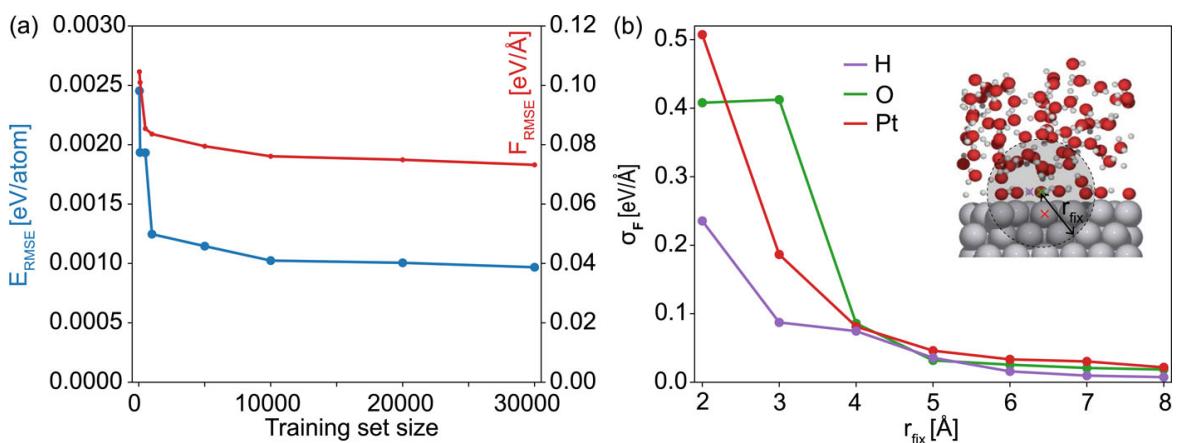
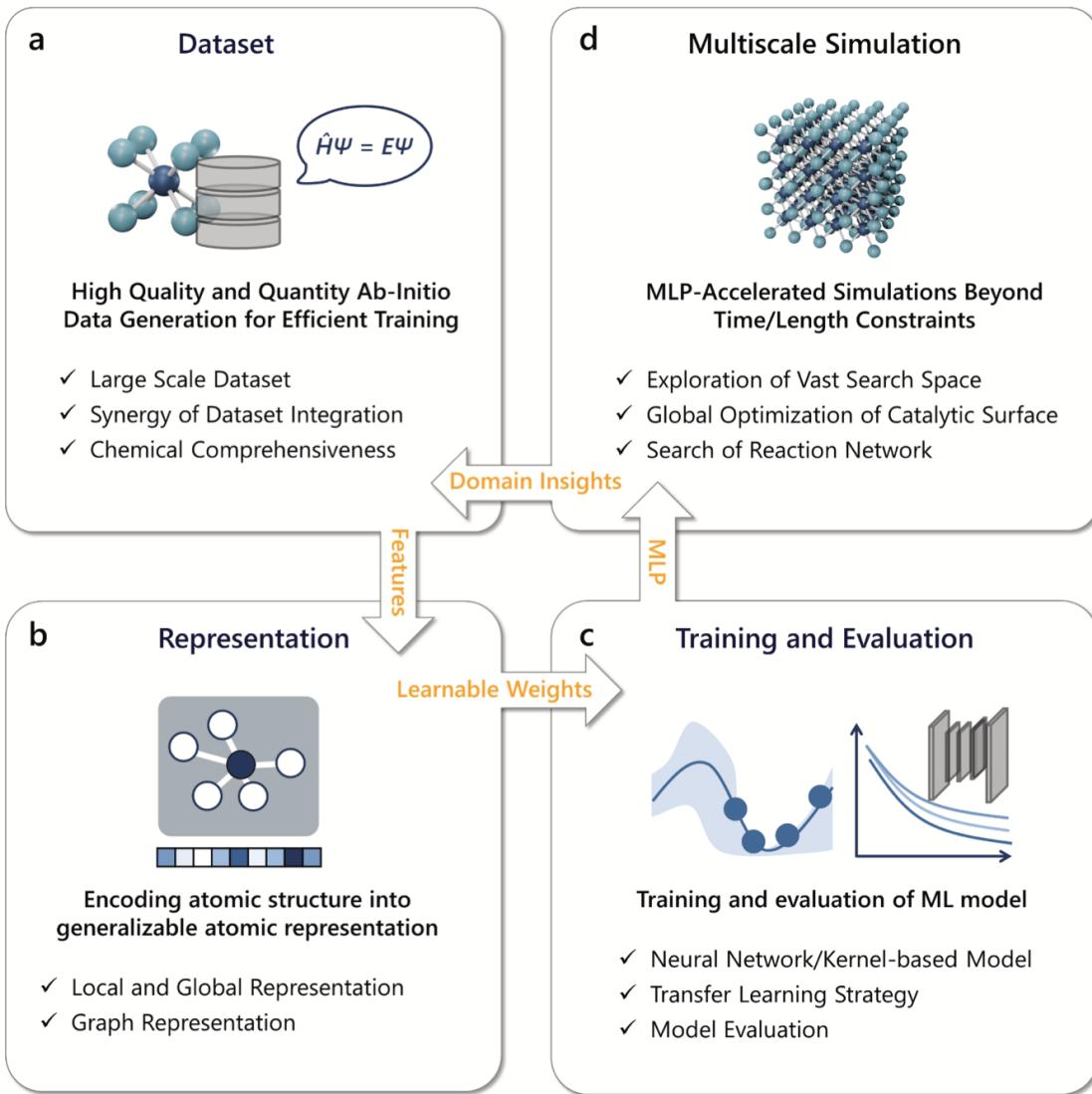


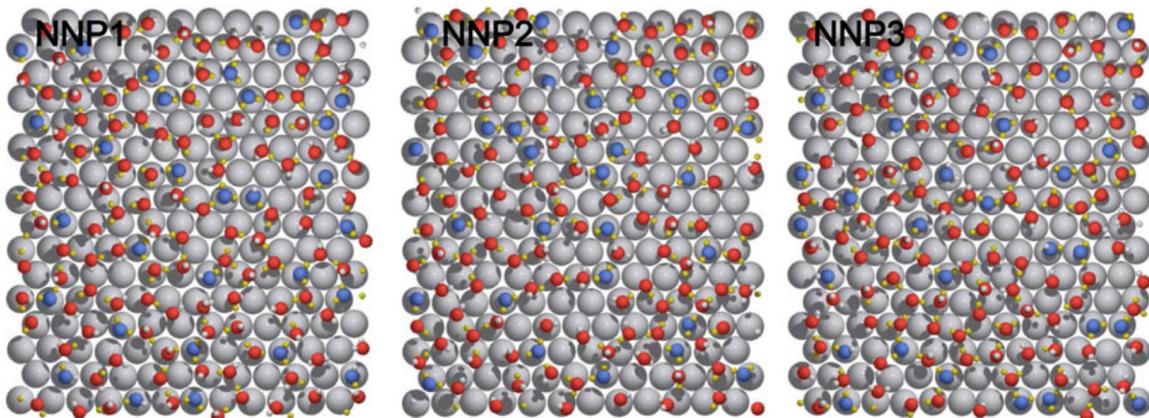
Chandrasekaran *et al.* *npj. Comp. Mater.*, 5, 22, (2019).

**This is because the computational requirements of MLP are lower than those of DFT**

- DFT computational cost =  $O(N^2)$
- MLP computational cost =  $O(N)$

## ▼ Typical workflow of application of MLP

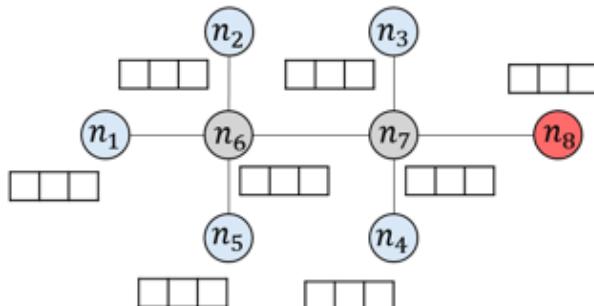
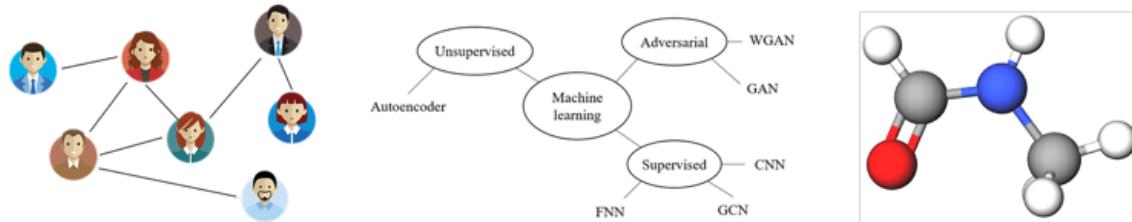




Karsten W. Jacobsen et al. J. Chem. Phys. 155, 224701 (2021)

## ▼ Architectures of MLPs (GNNS)

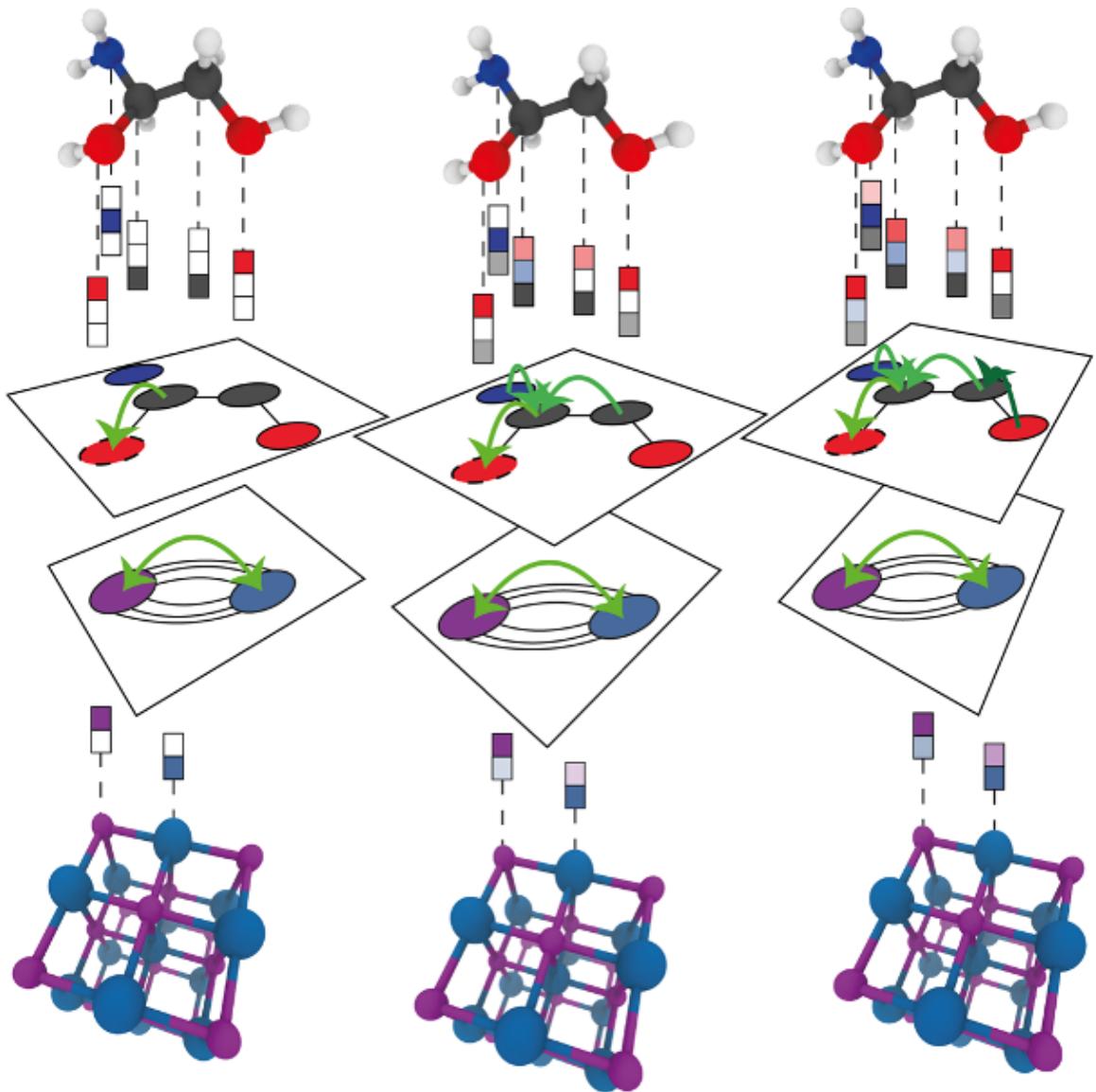
### GNN



$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

<https://untitledtblog.tistory.com/152>

Message passing GNN



Graph-level	Attributes	Description
nodes	atom-type	type of atoms (one-hot)
	chirality	R or S (one-hot or null)
	degree	number of covalent bonds (one-hot)
	radical	number of radical electrons (integer)
	hybridization	sp, sp <sup>2</sup> , sp <sup>3</sup> ... (one-hot)
	aromaticity	part of an aromatic system (binary)
edges	charge	formal charge (integer)
	bond-type	single, double, ... (one-hot)
	conjugation	is conjugated (binary)
	ring	bond is part of a ring (binary)
graph	stereo	None, Any, Z, E (one-hot)
	weight	average atomic weight (float)
	bonds	average bonds per atom (float)

Friederich Comm. Mat. 3, 93 (2022)

## CGCNN

---

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*

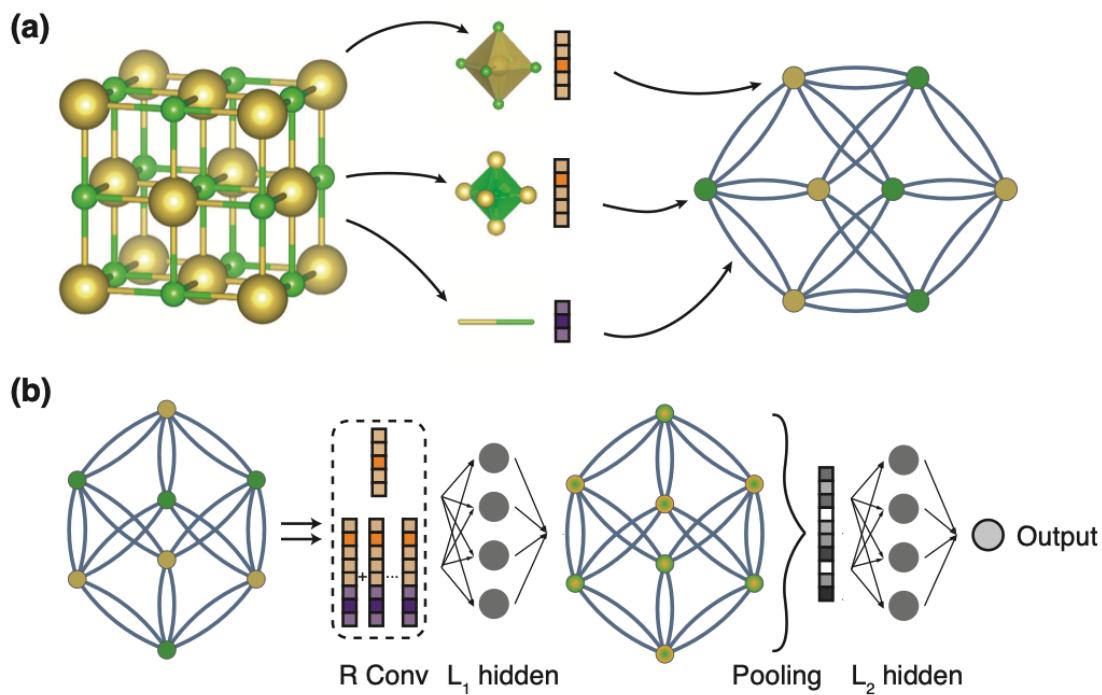
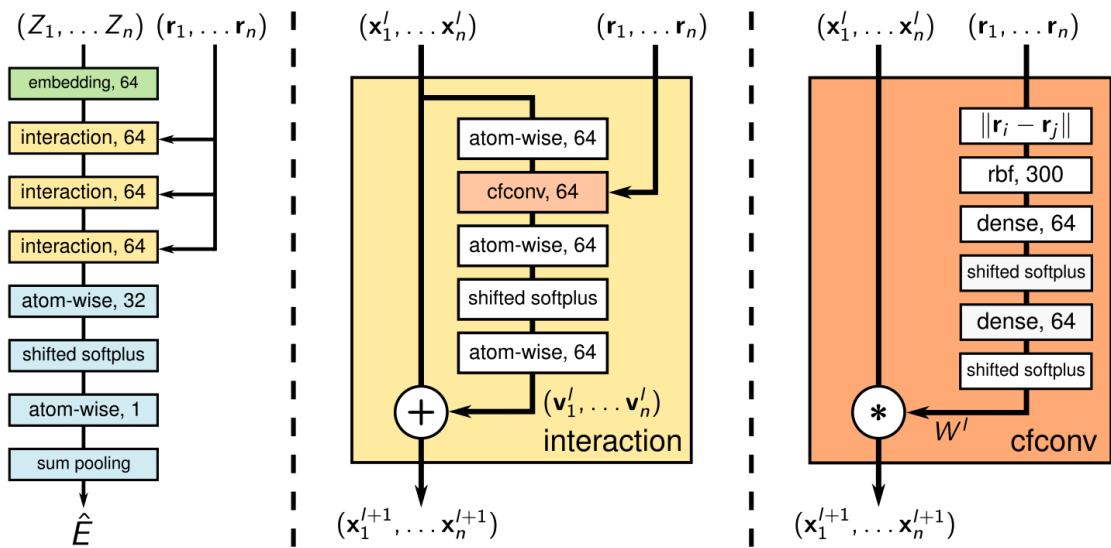
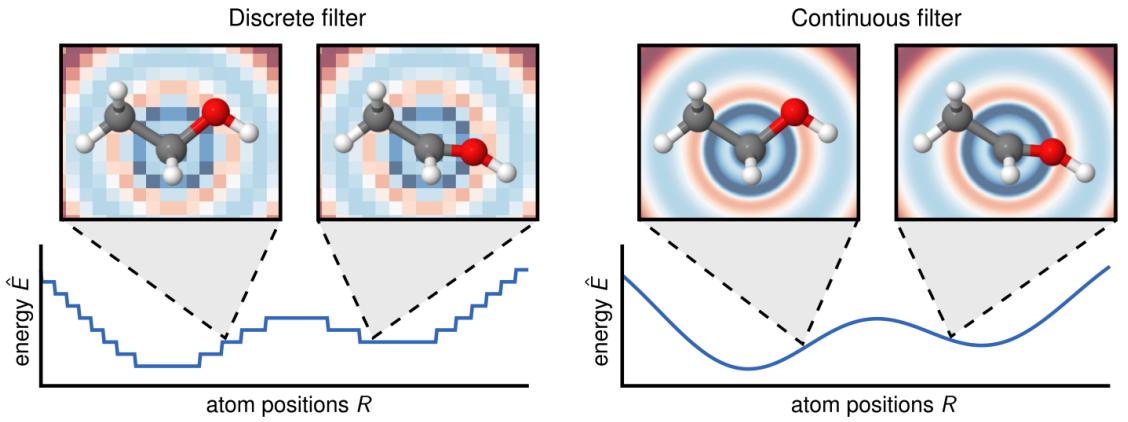


TABLE I. Summary of the prediction performance of seven different properties on test sets.

Property	# of train data	Unit	MAE <sub>model</sub>	MAE <sub>DFT</sub>
Formation energy	28046	eV/atom	0.039	0.081–0.136[18]
Absolute energy	28046	eV/atom	0.072	–
Band gap	16458	eV	0.388	0.6[23]
Fermi energy	28046	eV	0.363	–
Bulk moduli	2041	log(GPa)	0.054	0.050[24]
Shear moduli	2041	log(GPa)	0.087	0.069[24]
Poisson ratio	2041	–	0.030	–

<https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.120.145301>

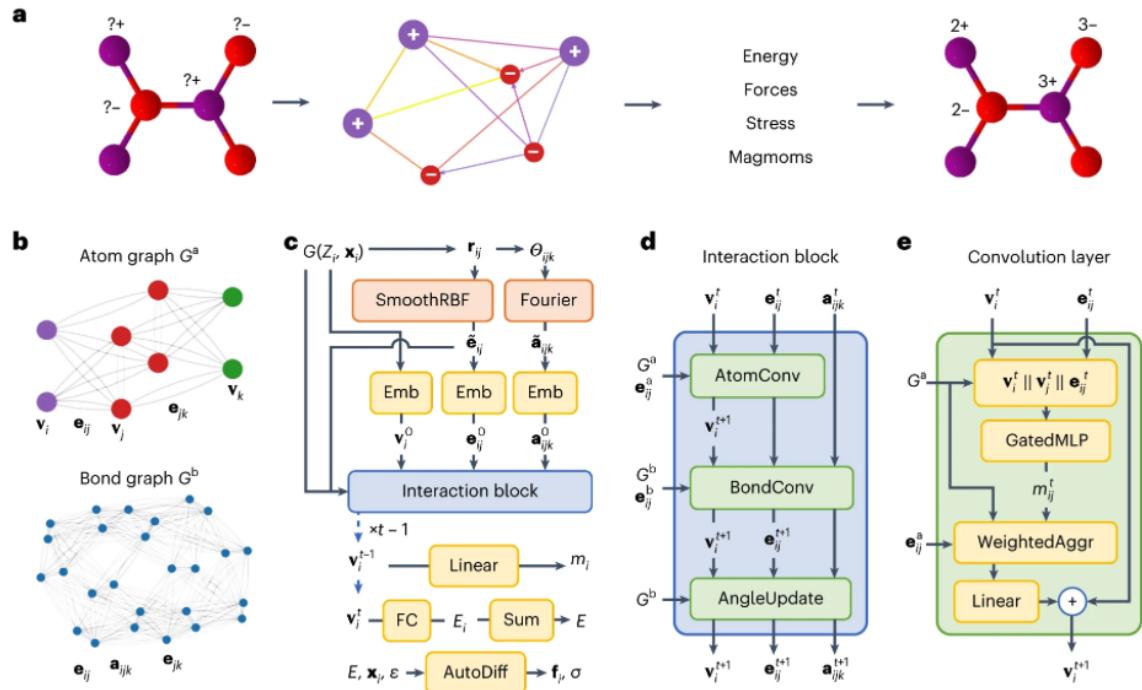
SchNET



		$N = 1,000$			$N = 50,000$		
		GdML [19]	SchNet		DTNN [20]	SchNet	
		forces	energy	both	energy	energy	both
<b>Benzene</b>	<i>energy</i>	<b>0.07</b>	1.19	0.08	<b>0.04</b>	0.08	0.07
	<i>forces</i>	<b>0.23</b>	14.12	0.31	—	1.23	<b>0.17</b>
<b>Toluene</b>	<i>energy</i>	<b>0.12</b>	2.95	<b>0.12</b>	0.18	0.16	<b>0.09</b>
	<i>forces</i>	<b>0.24</b>	22.31	0.57	—	1.79	<b>0.09</b>
<b>Malonaldehyde</b>	<i>energy</i>	0.16	2.03	<b>0.13</b>	0.19	0.13	<b>0.08</b>
	<i>forces</i>	0.80	20.41	<b>0.66</b>	—	1.51	<b>0.08</b>
<b>Salicylic acid</b>	<i>energy</i>	<b>0.12</b>	3.27	0.20	0.41	0.25	<b>0.10</b>
	<i>forces</i>	<b>0.28</b>	23.21	0.85	—	3.72	<b>0.19</b>
<b>Aspirin</b>	<i>energy</i>	<b>0.27</b>	4.20	0.37	—	0.25	<b>0.12</b>
	<i>forces</i>	<b>0.99</b>	23.54	1.35	—	7.36	<b>0.33</b>
<b>Ethanol</b>	<i>energy</i>	0.15	0.93	<b>0.08</b>	—	0.07	<b>0.05</b>
	<i>forces</i>	0.79	6.56	<b>0.39</b>	—	0.76	<b>0.05</b>
<b>Uracil</b>	<i>energy</i>	<b>0.11</b>	2.26	0.14	—	0.13	<b>0.10</b>
	<i>forces</i>	<b>0.24</b>	20.08	0.56	—	3.28	<b>0.11</b>
<b>Naphthalene</b>	<i>energy</i>	<b>0.12</b>	3.58	0.16	—	0.20	<b>0.11</b>
	<i>forces</i>	<b>0.23</b>	25.36	0.58	—	2.58	<b>0.11</b>

[https://arxiv.org/pdf/1706.08566](https://arxiv.org/pdf/1706.08566.pdf)

## CHGNET



*Nature Machine Intelligence* volume 5, pages 1031–1041 (2023)

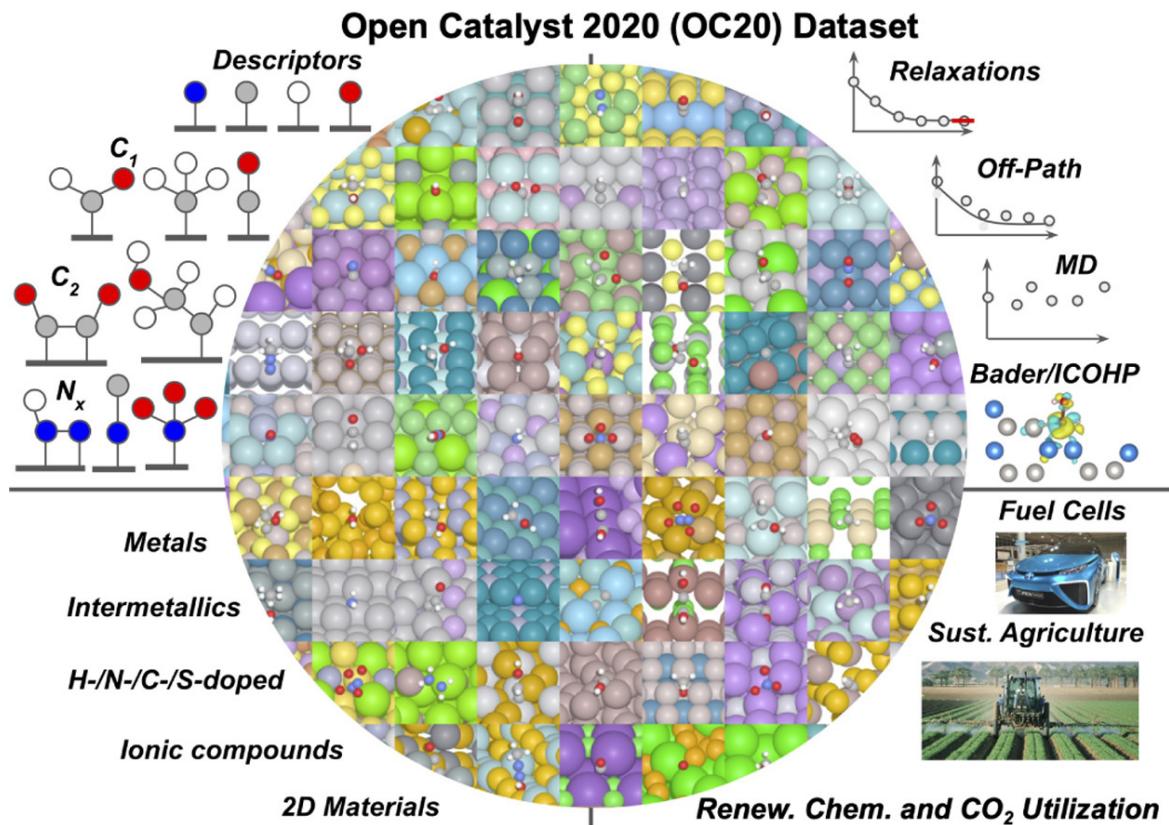
## ▼ Advent of large dataset

### Advent of Large dataset

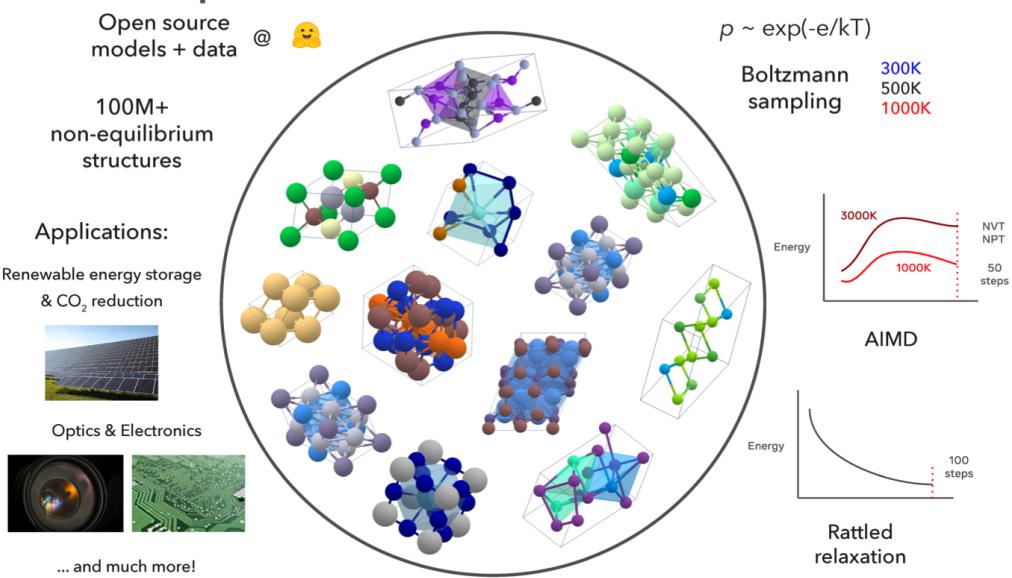
Ab initio datasets of inorganic materials in computational catalysis.

Dataset	Property	Dataset size	Level of theory	Chemical Composition	Accessibility and contribution
Materials Project [34]	Atomic structure, band structure, magnetic properties, and others	$1.5 \times 10^5$	PBE, HSE06, and $r^2$ SCAN	Inorganic Crystals and molecules	Open source, open contribution,
OQMD [35]	Atomic structure, and band structure band gap	$1 \times 10^6$	PBE	Inorganic Crystals	Open source
NOMAD [36]	Atomic structure, and band structure	$3 \times 10^6$	PBE, HSE06, HLE17, and others	Inorganic Crystals	Open source, open contribution
AFLW [37]	Atomic structure, band structure, magnetic properties and phonon spectra	$3.5 \times 10^6$	PBE	Inorganic Crystals	Open source
OC20 [12]	Atomic structure, Bader charge, total energy, and force,	$2.5 \times 10^8$	RPBE	molecular adsorptions on inorganic crystal slabs	Open source, community challenges
OC22 [13]	Atomic structure, total energy, and force	$1 \times 10^7$	PBE	molecular adsorptions on metal oxide slabs	Open source, community challenges
OpenDAC [38]	Atomic structure, total energy and force	$3.8 \times 10^7$	PBE-D3	$\text{CO}_2$ and $\text{H}_2\text{O}$ adsorptions on MOFs	Open source, community challenges
Catalysis-Hub [39]	Atomic structure, total energy, and reaction energy	$1.3 \times 10^5$	PBE, RPBE,PW91 and others	Molecular adsorptions on inorganic crystals	Open source, open contribution,
PFP [40]	Atomic structure, and total energy	$9 \times 10^6$	PBE and PBE + U	Molecules, transition metal slabs, clusters	Partially open
GNoMe [29]	Atomic structure, total energy, and decomposition energy	$3.8 \times 10^5$	PBE and $r^2$ SCAN	Inorganic Crystal	Partially open

### Open Catalyst Project



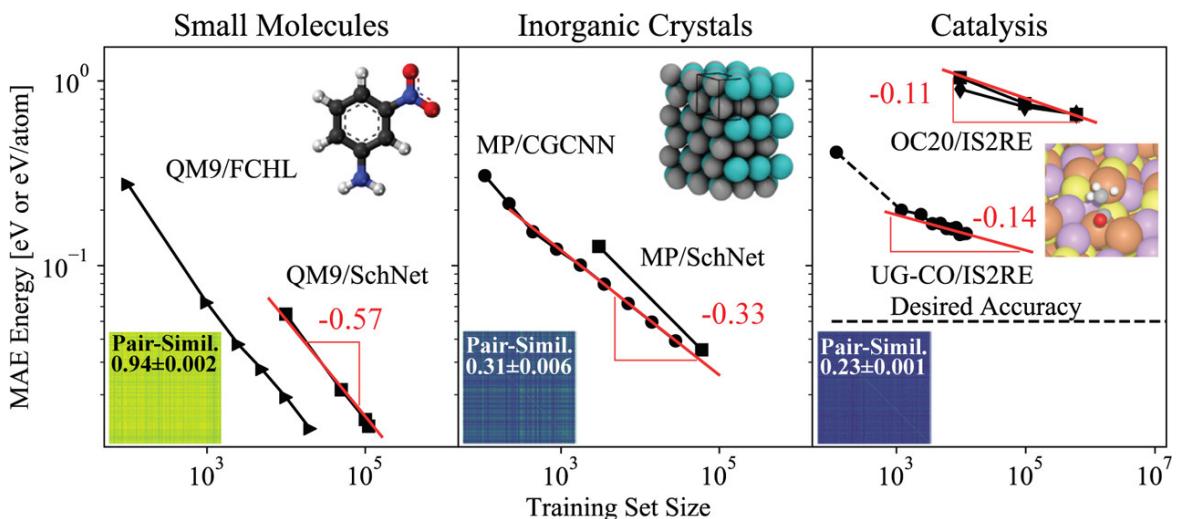
# Open Materials 2024 (OMat24)

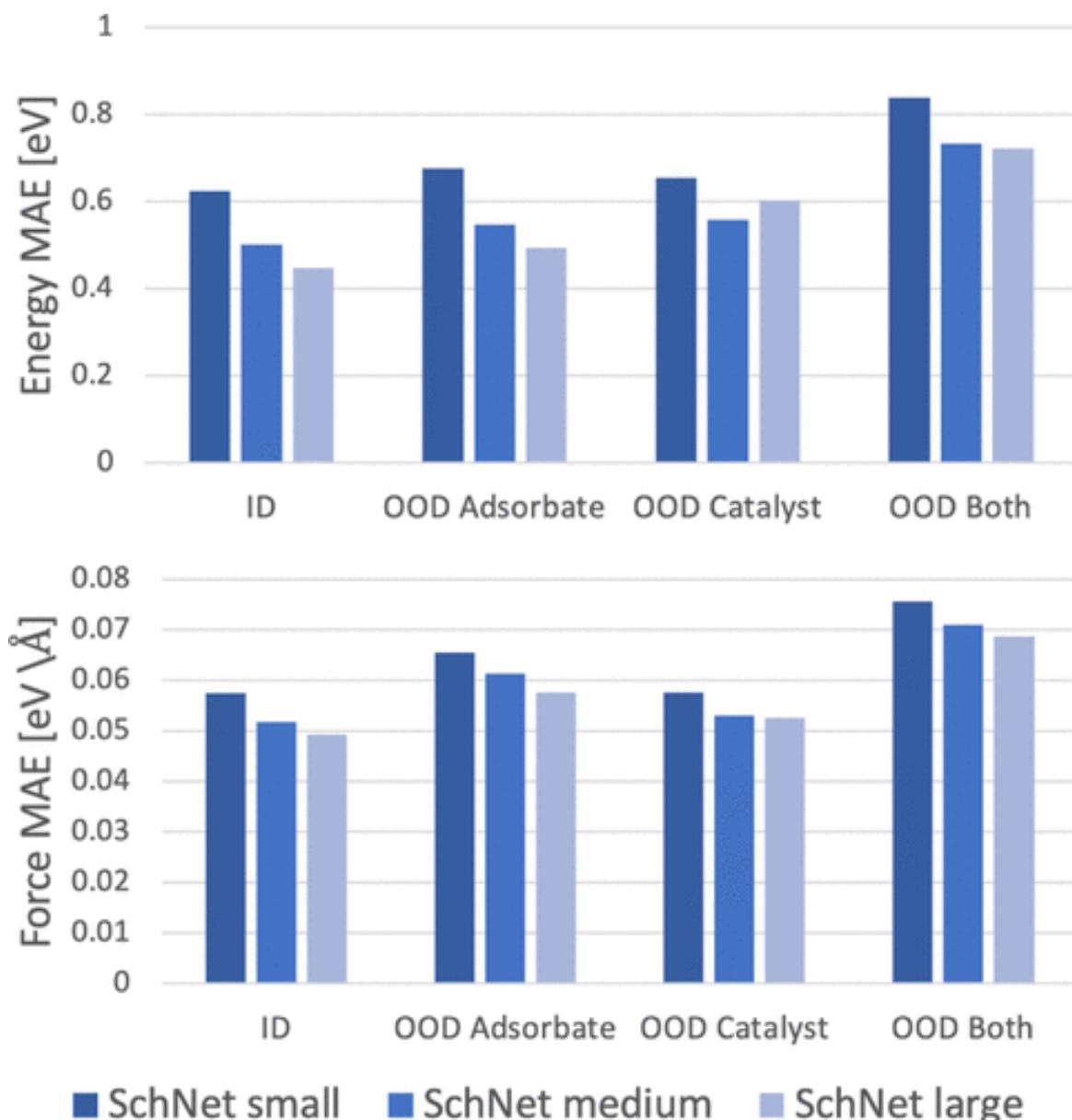


<https://arxiv.org/pdf/2410.12771>

## ▼ Universal MLP

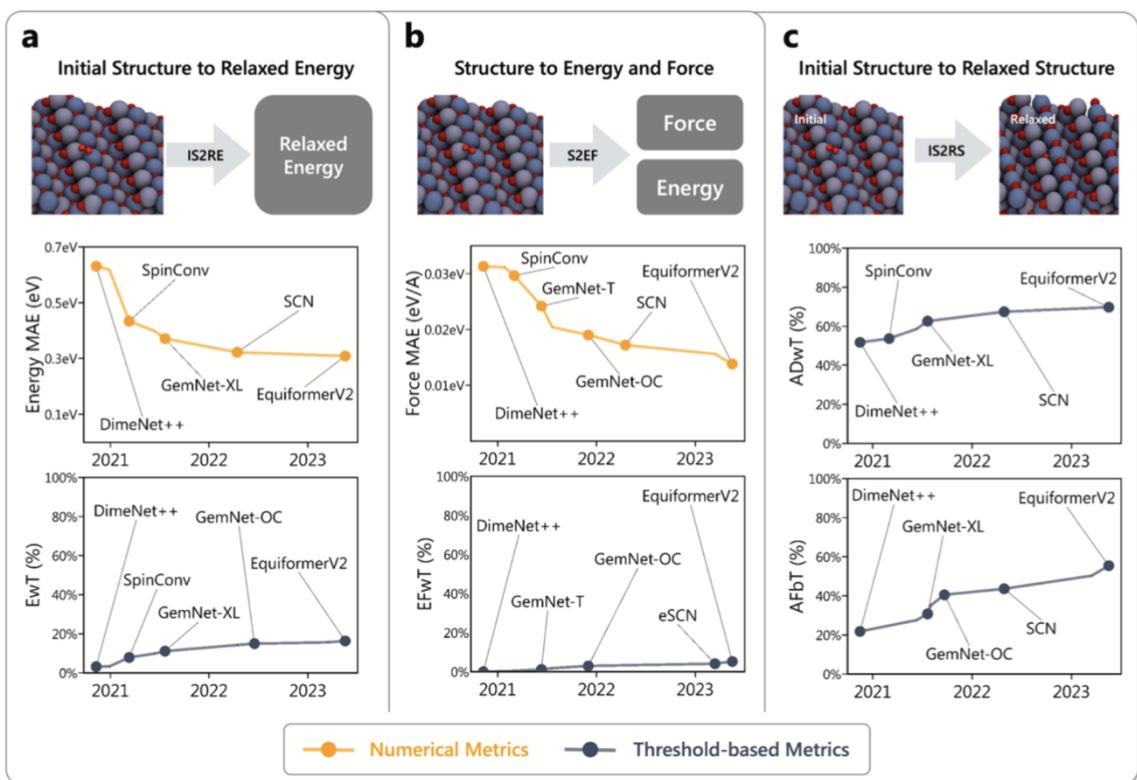
### Scale of database size vs MLP accuracy and transferability





ACS Catal. 2021, 11, 10, 6059-6072

## How we evaluate MLP?



scale matters.

☰ 매일경제

천재들이 했으면 800년 걸릴 일...10년만에 해낸 넌 누구?

국제

## 천재들이 했으면 800년 걸릴 일...10년만에 해낸 넌 누구?

진영태 기자 zin@mk.co.kr

입력 : 2023-11-30 10:38:29 수정 : 2023-11-30 14:23:49

구글 AI 200만개 신소재 구조 발견  
첨단산업 신소재 38만개 생산 가능  
배터리, 태양광 등 활용 범위 넓어

가 □ ☰ ☱

Google DeepMind

Article

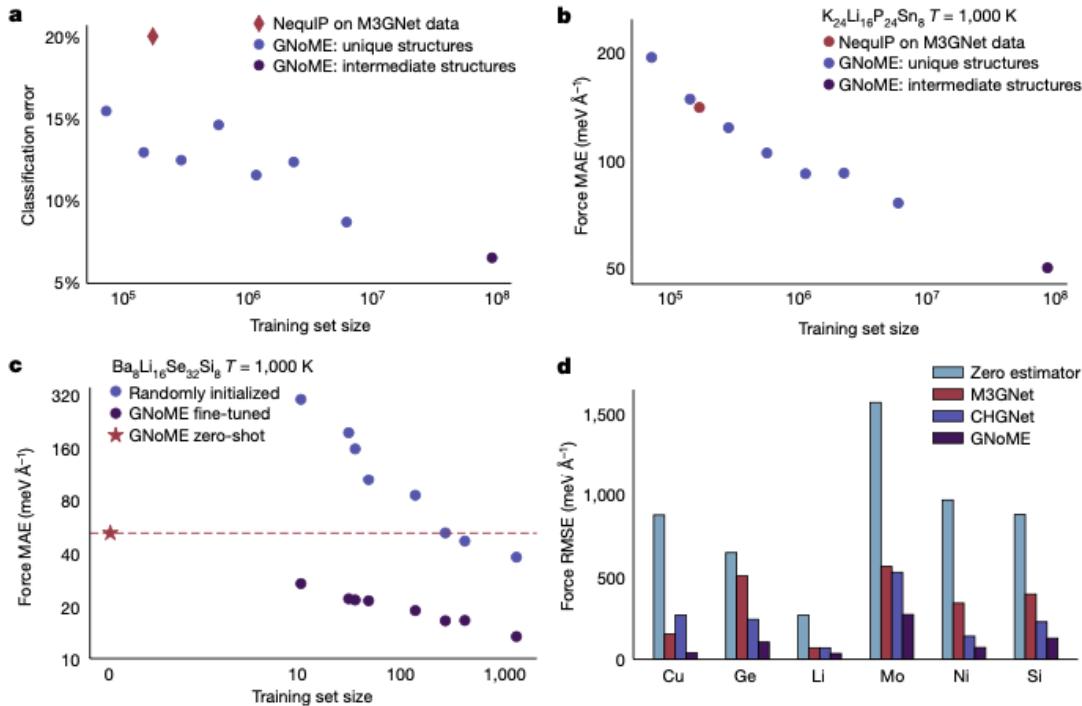
# Scaling deep learning for materials discovery

<https://doi.org/10.1038/s41586-023-06735-9>

Amil Merchant<sup>1,3</sup>✉, Simon Batzner<sup>1,3</sup>, Samuel S. Schoenholz<sup>1,3</sup>, Muratahan Aykol<sup>1</sup>,

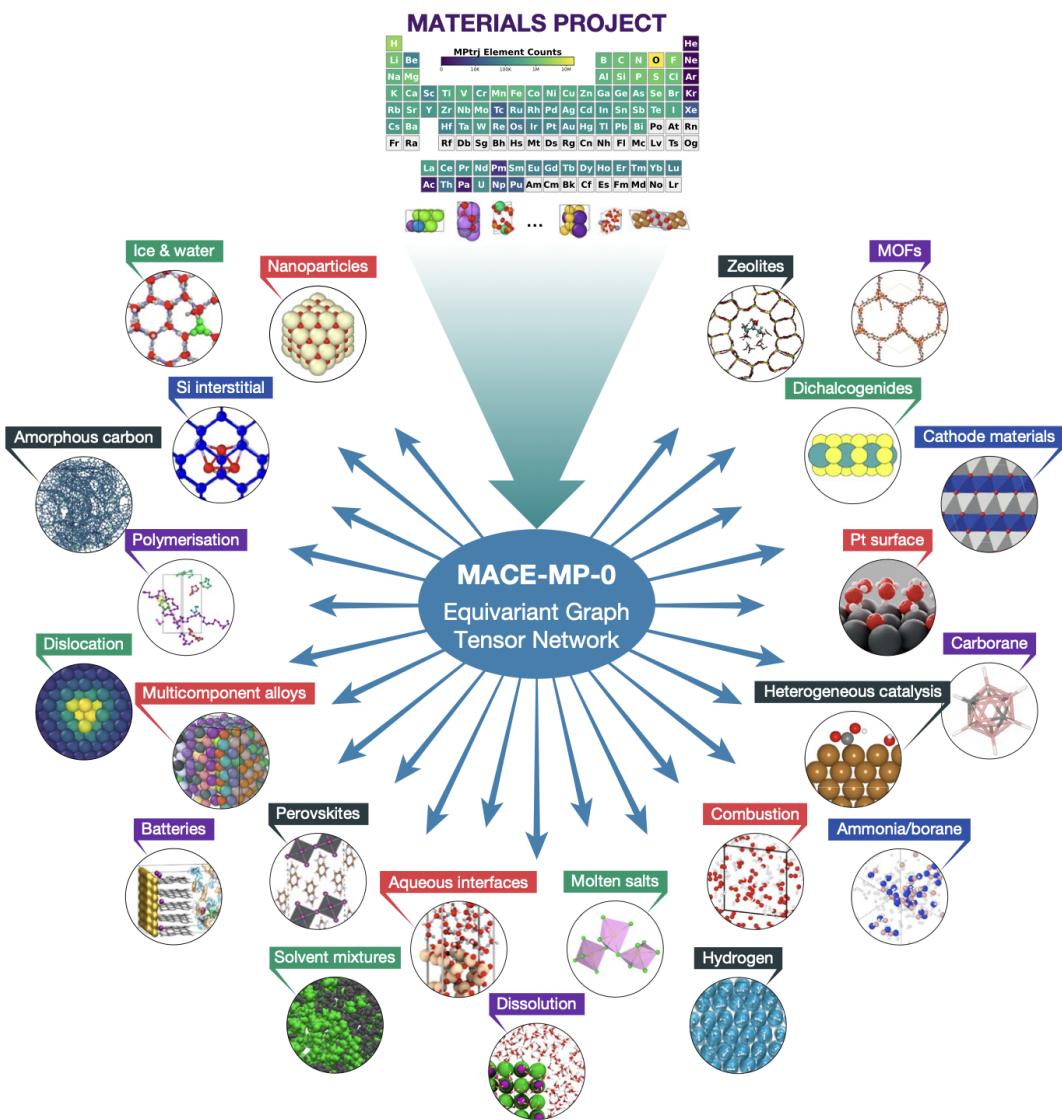
Gwoon Cheon<sup>2</sup> & Ekin Dogus Cubuk<sup>1,3</sup>✉

Received: 8 May 2023



<https://www.nature.com/articles/s41586-023-06735-9>

▼ MACE-MP, one of the 'universal' MLP



<https://arxiv.org/pdf/2401.00096>

## Machine learning potential we will use : MACE

---

## MACE: Higher Order Equivariant Message Passing Neural Networks for Fast and Accurate Force Fields

---

Ilyes Batatia  
Engineering Laboratory,  
University of Cambridge  
Cambridge, CB2 1PZ UK  
Department of Chemistry,  
ENS Paris-Saclay, Université Paris-Saclay  
91190 Gif-sur-Yvette, France  
ilyes.batatia@ens-paris-saclay.fr

Gregor N. C. Simm  
Engineering Laboratory,  
University of Cambridge  
Cambridge, CB2 1PZ UK

Dávid Péter Kovács  
Engineering Laboratory,  
University of Cambridge  
Cambridge, CB2 1PZ UK

Christoph Ortner  
Department of Mathematics  
University of British Columbia  
Vancouver, BC, Canada V6T 1Z2

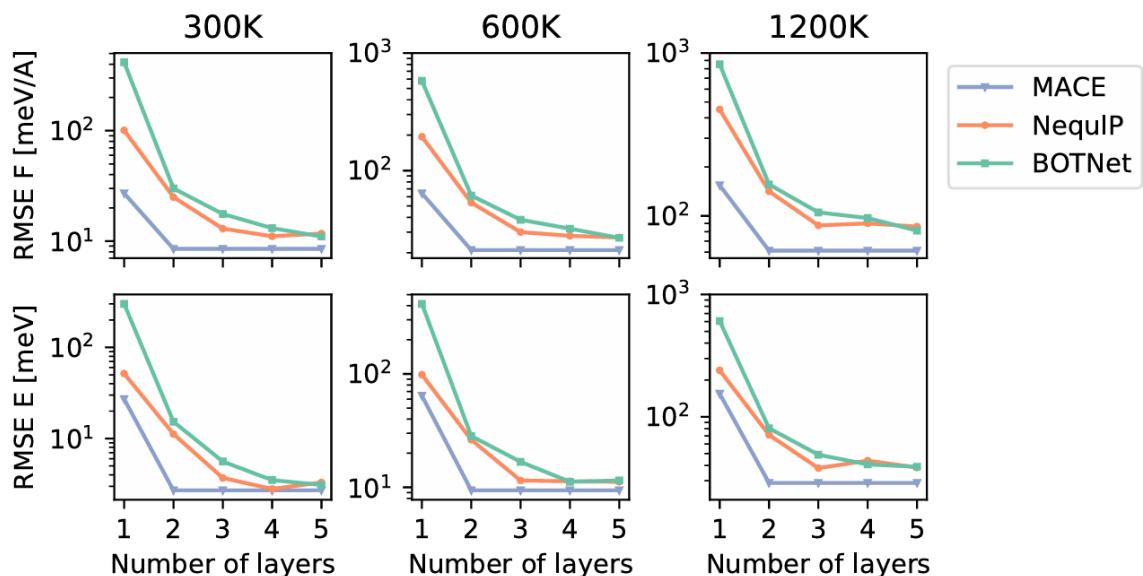
Gábor Csányi  
Engineering Laboratory,  
University of Cambridge  
Cambridge, CB2 1PZ UK

<https://doi.org/10.48550/arXiv.2206.07697>

### Characteristics of MACE

- MPNN Interatomic Potentials
- Equivariant Graph Neural Networks

### Performance of MACE



performace on 3BPA dataset: MACE > NequIP

### Install MACE

## Installation from PyPI

This is the recommended way to install MACE.

```
pip install --upgrade pip  
pip install mace-torch
```



- **MACE-MP : Materials Project Force Fields, trained on bulk dataset**

Example usage in ASE

```
from mace.calculators import mace_mp  
from ase import build  
  
atoms = build.molecule('H2O')  
calc = mace_mp(model="medium", dispersion=False, default_dtype="float32", device='cuda')  
atoms.calc = calc  
print(atoms.get_potential_energy())
```



- **MACE-OFF : Transferable Organic Force Fields, trained on organic molecule dataset**

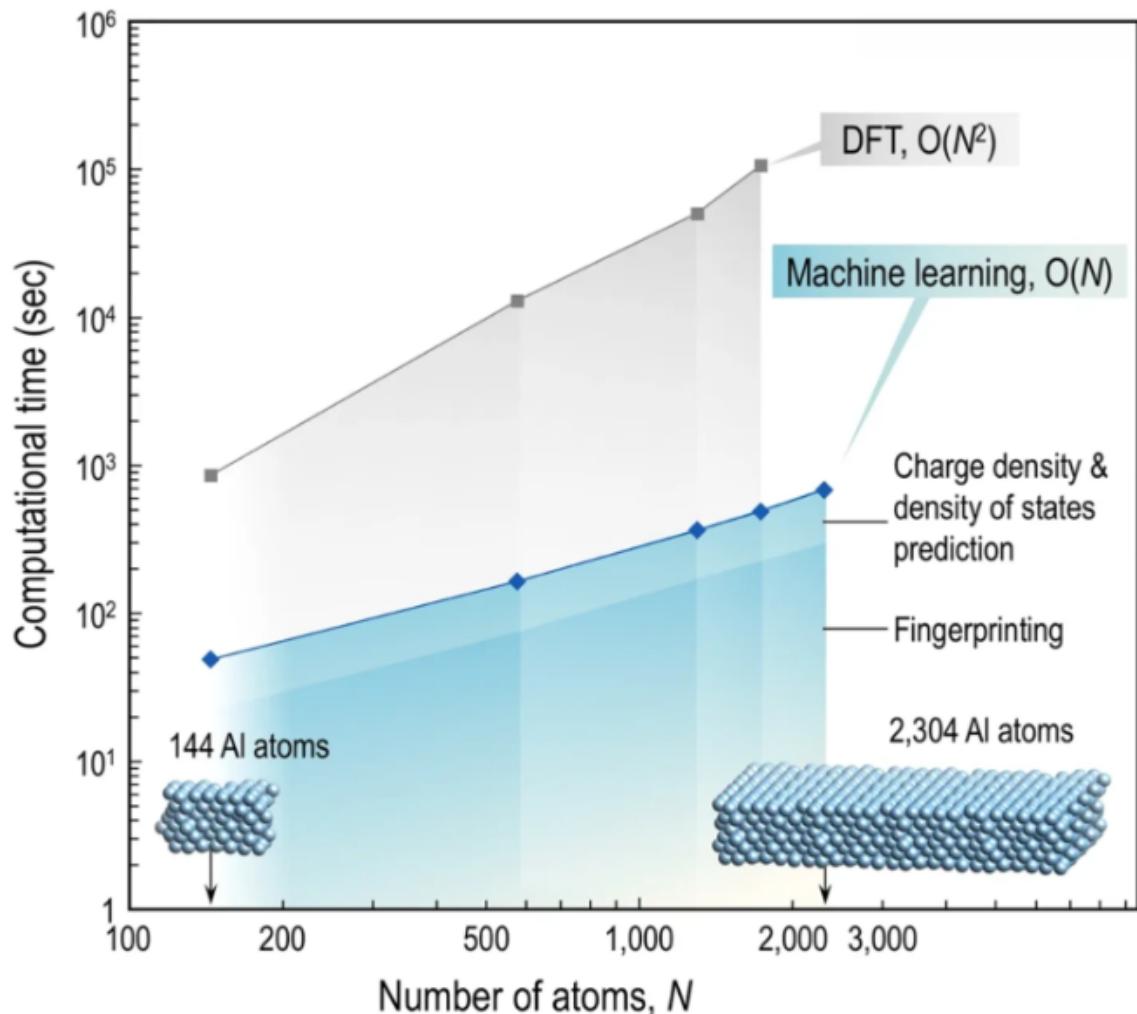
Example usage in ASE

```
from mace.calculators import mace_off  
from ase import build  
  
atoms = build.molecule('H2O')  
calc = mace_off(model="medium", device='cuda')  
atoms.calc = calc  
print(atoms.get_potential_energy())
```



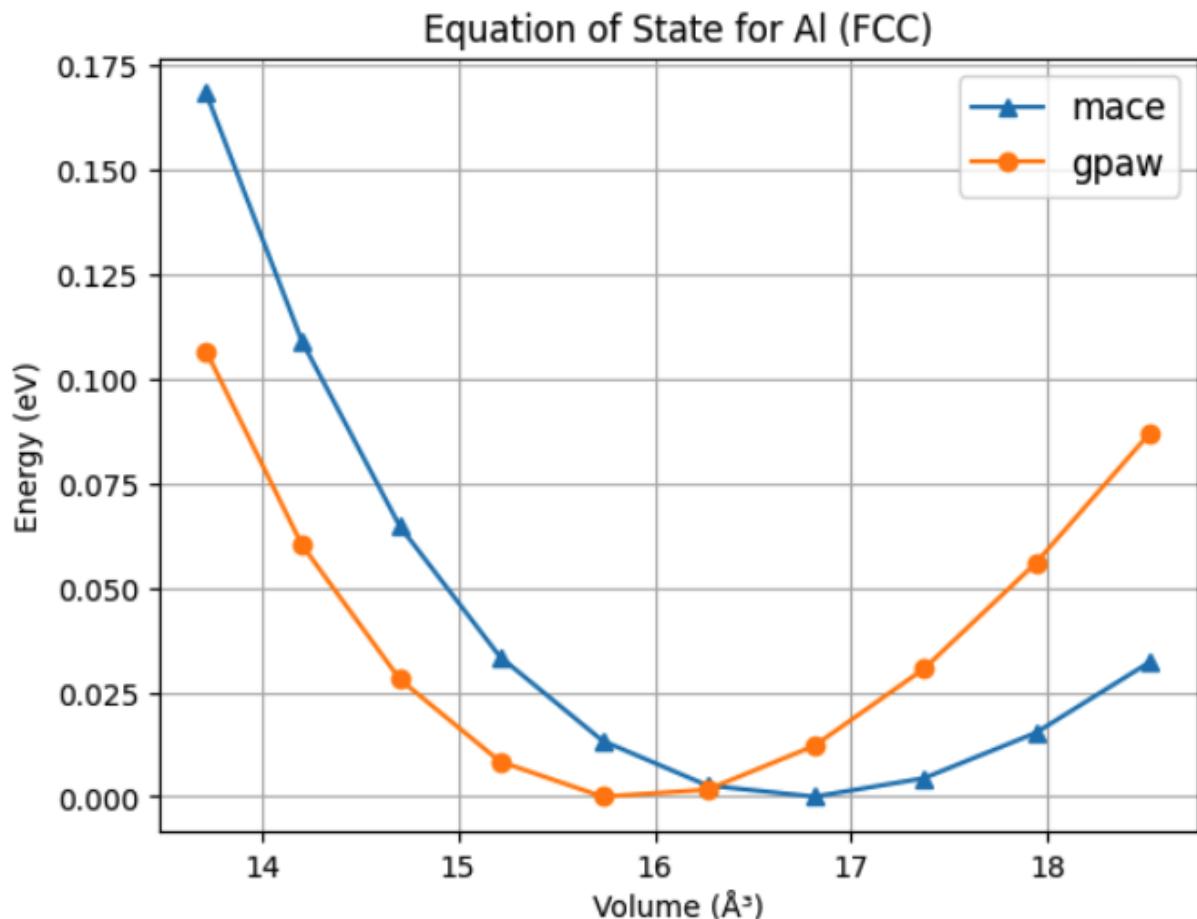
# Hands-on on 12th Nov.

## Task 01: Execution time comparison (revisiting Handson 2)



Let's compare the computation time of DFT and MLP as the number of atoms increases to see if a similar trend in computation time, as shown in the graph, is observed.

## Task 02: Bulk optimization using MACE (revisiting Handson 1)

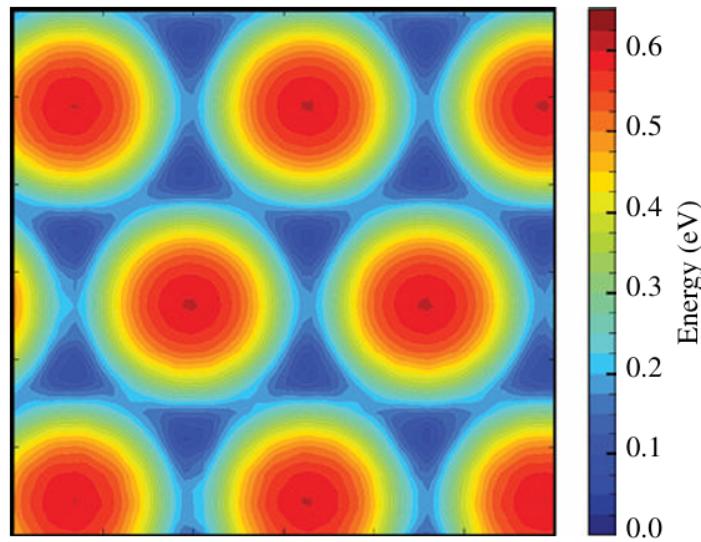


### ▼ Assignment

- Select a metal atom freely, set the number of atoms to [1, 2, 3, 4, 5, 6] for that metal(=number of layers), and compare the time required for energy calculations using DFT and MACE for each system. Plot the time on a log scale
- plot simulation time vs simulating atom scale for MACE pretrained model
- Plot three EOS for three metals and compare them with GPAW result  
ex) Pt EOS, Al EOS, Pd EOS using both GPAW and MACE

### ▼ Upcoming hands-on on 14th Nov.

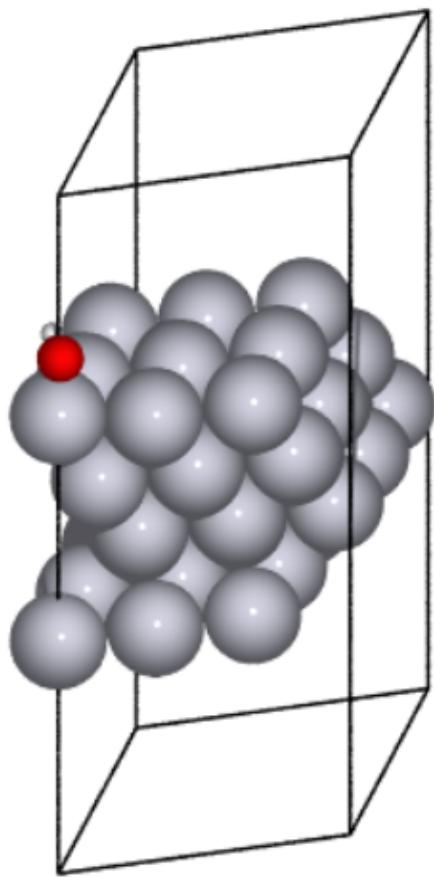
## Task 02 : PES of H diffusion on metal surface



**FIGURE 2.7** PES of H diffusion on Cu(111). The potential energy is plotted over part of the Cu surface area. H adsorbed in the threefold position has been chosen as the reference.

Let's utilize the fast computation speed of MLP to visualize the PES of H diffusion on the Pt(111) surface!

## Task 03 : Finding a optimal adsorption configuration



Let's use a grid search to identify the most stable adsorption site and geometric adsorption structure for OH on the Pt surface.