

Towards a Universal Large Atomic Model for Molecular and Materials Simulation

Linfeng Zhang, DP Technology, AI for Science Institute

June 25, 2024

Multi-scale modeling and learning of the microscopic world

VOLUME 55, NUMBER 22

PHYSICAL REVIEW LETTERS

25 NOVEMBER 1985

Unified Approach for Molecular Dynamics and Density-Functional Theory

R. Car

International School for Advanced Studies, Trieste, Italy

and

M. Parrinello

*Dipartimento di Fisica Teorica, Università di Trieste, Trieste, Italy, and
International School for Advanced Studies, Trieste, Italy*

(Received 5 August 1985)

We present a unified scheme that, by combining molecular dynamics and density-functional theory, profoundly extends the range of both concepts. Our approach extends molecular dynamics beyond the usual pair-potential approximation, thereby making possible the simulation of both covalently bonded and metallic systems. In addition it permits the application of density-functional theory to much larger systems than previously feasible. The new technique is demonstrated by the calculation of some static and dynamic properties of crystalline silicon within a self-consistent pseudopotential framework.

PACS numbers: 71.10.+x, 65.50.+m, 71.45.Gm

CPMD, 1985;

-
-
-
-

Multi-scale modeling and learning of the microscopic world

PRL 98, 146401 (2007)

PHYSICAL REVIEW LETTERS

week ending
6 APRIL 2007

Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces

Jörg Behler and Michele Parrinello

Department of Chemistry and Applied Biosciences, ETH Zurich, USI-Campus, Via Giuseppe Buffi 13, CH-6900 Lugano, Switzerland
(Received 27 September 2006; published 2 April 2007)

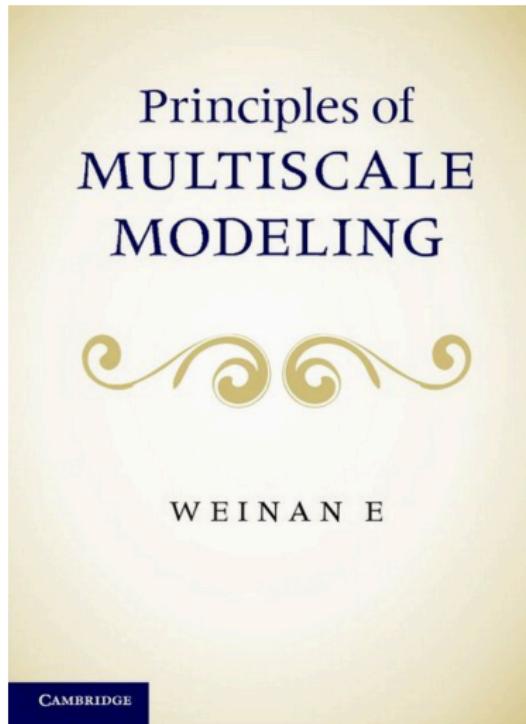
The accurate description of chemical processes often requires the use of computationally demanding methods like density-functional theory (DFT), making long simulations of large systems unfeasible. In this Letter we introduce a new kind of neural-network representation of DFT potential-energy surfaces, which provides the energy and forces as a function of all atomic positions in systems of arbitrary size and is several orders of magnitude faster than DFT. The high accuracy of the method is demonstrated for bulk silicon and compared with empirical potentials and DFT. The method is general and can be applied to all types of periodic and nonperiodic systems.

DOI: 10.1103/PhysRevLett.98.146401

PACS numbers: 71.15.Pd, 61.50.Ab, 82.20.Kh

- CPMD, 1985;
- BPNN, 2007;
-
-

Multi-scale modeling and learning of the microscopic world



- CPMD, 1985;
- BPNN, 2007;
- **W. E's book, 2011;**
-

Multi-scale modeling and learning of the microscopic world

The image shows two side-by-side screenshots of email conversations in an inbox.

Top Screenshot (Roberto Car to Weinan):

- From: Roberto Car <rcar@princeton.edu>
- To: Weinan, me
- Date: Wed, Oct 5, 2016, 9:43 PM
- Subject: meeting
- Message content:

Hi Weinan,

I am late. I will be at your office at 10:20 am. I hope it is still OK with you and Linfeng.

Best,

Roberto

Bottom Screenshot (Linfeng Zhang to Weinan):

- From: Linfeng Zhang <linfengz@princeton.edu>
- To: Weinan
- Date: Fri, Oct 7, 2016, 11:12 PM
- Subject: Neural-Network Representation of High-Dimensional Potential-Energy Surfaces
- Message content:

Dear Prof. E,

It seems there's a series of articles on the topic "Neural-Network Representation of High-Dimensional Potential-Energy Surfaces", by Michele Parrinello, and Jörg Behler, since 2007.

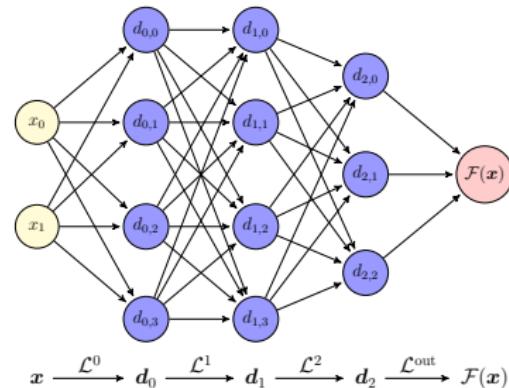
I attached two original PRL articles in 2007 and 2008, and a tutorial review on 2015. Can I have a discussion with you after I digest them? Or in the first place, do they meet your interest?

Thank you!

Best wishes,
Linfeng
- Attachments: 3 attachments (Scanned by Gmail)

- CPMD, 1985;
- BPNN, 2007;
- W. E's book, 2011;
- **systematic opportunity realized, 2016;**

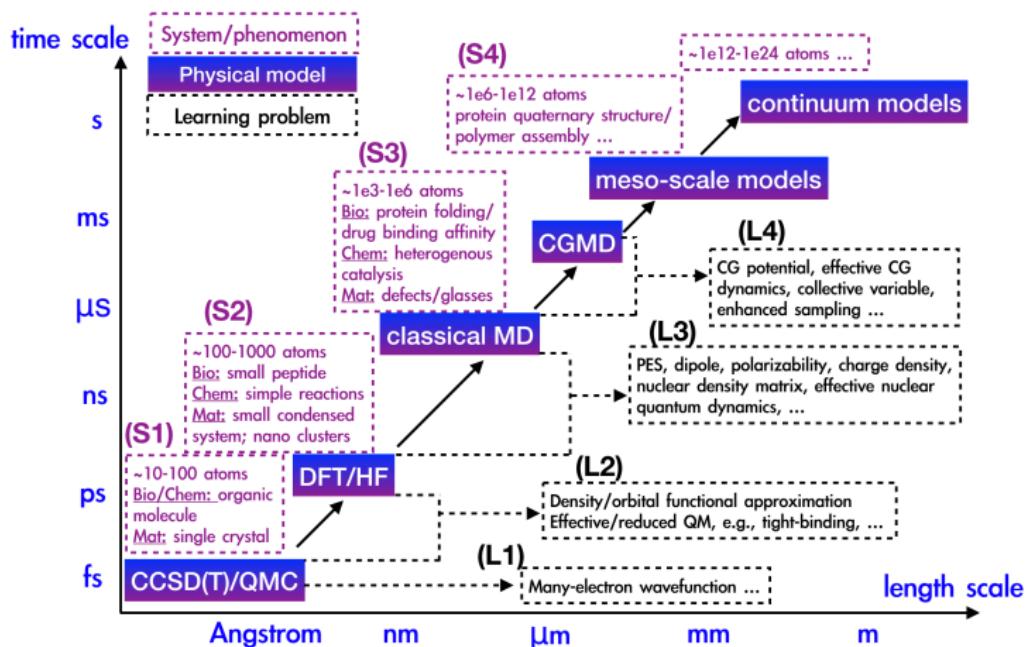
Learning-assisted modeling: Why now?



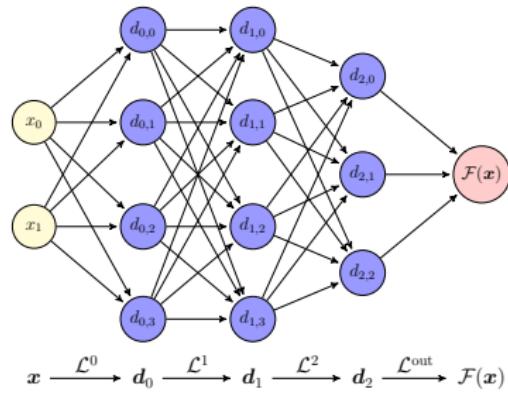
2017: AI gives us tools for representing high-dimensional functions.
2022: ...

Multi-scale Molecular and Materials Modeling

More is different. The existence of hierarchical levels of science, each of which requires its own fundamental principles for advancement (P. W. Anderson);



How to organize joint efforts?



2018: Modern software engineering + open-source community.
2023: ...

DeepModeling: embracing open source & open science

A screenshot of a web browser window. The address bar shows "github.com/deepmodeling/community". Below the address bar, there's a navigation menu with icons for back, forward, and search. The main content area shows the "README.md" file, which contains the DeepModeling Manifesto.

DeepModeling Manifesto

The integration of machine learning and physical modeling is changing the paradigm of scientific research. Those who hope to extend the frontier of science and solve challenging practical problems through computational modeling are coming together in new ways never seen before. This calls for a new infrastructure--new platforms for collaboration, new coding frameworks, new data processing schemes, and new ways of using the computing power. It also calls for a new culture—the culture of working together closely for the benefit of all, of free exchange and sharing of knowledge and tools, of respect and appreciation of each other's work, and of the pursuit of harmony among diversity.

The DeepModeling community is a community of such a group of people.

A screenshot of the DeepModeling GitHub organization page. The header features the "DeepModeling" logo and the tagline "Define the future of scientific computing together". Below the header, there are links to "DeepModeling | Projects | Tutorials | Blog | Docs". The main content area shows the organization's profile information, including its name, creation date, member count, and a summary of its mission. It also displays statistics for repositories, releases, packages, and usage, along with a link to overall issues and pull requests status. At the bottom, there are tabs for "Issues" and "Pull requests".

[DeepModeling](#) | [Projects](#) | [Tutorials](#) | [Blog](#) | [Docs](#)

DeepModeling

Joined GitHub 6 years ago 34 members

About us

Define the future of scientific computing together

53 Repositories (including 8 forks)

Prefers LGPL-3.0 license

497 Releases

17 Packages

2.14 GB used

0 Sponsors

2528 Stargazers

1132 Forkers

197 Watchers

28.4K views in last two weeks

Overall issues and pull requests status

On DeepModeling's repositories

Issues

Pull requests

arXiv > physics > arXiv:2304.09409

Search...
Help | Advanced

Physics > Chemical Physics

[Submitted on 19 Apr 2023]

DeePMD-kit v2: A software package for Deep Potential models

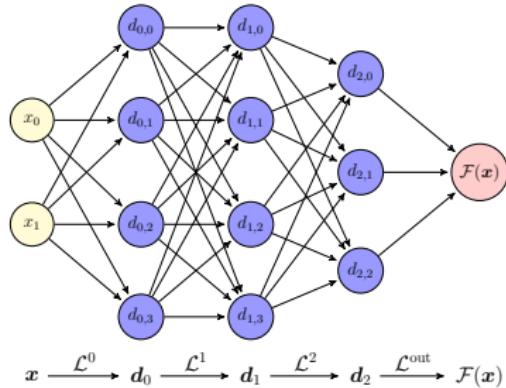
Jinzhe Zeng, Duo Zhang, Denghai Lu, Pinghui Mo, Zeyu Li, Yixiao Chen, Marián Ryník, Liang Huang, Ziyo Li, Shaochen Shi, Yingze Wang, Haotian Ye, Ping Tuo, Jabin Yang, Ye Ding, Yifan Li, Davide Tisi, Qiyu Zeng, Han Bao, Yu Xia, Jiameng Huang, Koki Muraoka, Yibo Wang, Junhan Chang, Fengbo Yuan, Sigrbjörn Leland Bore, Chun Cai, Yinnian Lin, Bo Wang, Jiayan Xu, Jia-Xin Zhu, Chenxing Luo, Yuzhi Zhang, Rhys E. A. Goodall, Wenshuo Liang, Anurag Kumar Singh, Sikai Yao, Jingchao Zhang, Renata Wentzcovitch, Jiequn Han, Jie Liu, Weile Jia, Darrin M. York, Weinan E, Roberto Car, Linfeng Zhang, Han Wang

DP Technology

Current landscape of DeepModeling

The image consists of two main parts. On the left is a screenshot of the GitHub repository for 'DeepModeling'. It shows basic repository statistics: Joined GitHub 6 years ago, 36 members, 50 repositories (including 6 forks), 0 sponsors, 879 releases, 17 packages, 2.81 GB used, and 48.8k views in last two weeks. Below this are sections for 'Overall issues and pull requests status' and 'Most used languages'. The 'Issues' section shows 446 open from communities and 2654 closed. The 'Pull requests' section shows 0 merged, 121 closed, 0 skipped, and 102 skipped. The 'Languages' section lists C++, Python, MATLAB, Jupyter Notebook, Cuda, TeX, Shell, and CMake. On the right is a large diagram titled 'Building the Future of Scientific Computing in the AI4S Era'. The diagram is organized into several sections: 'Industry Products' (Alloy Property Prediction, Semiconductor Property Prediction, Energy Property Prediction, Organic Molecule Property Prediction, Biomolecular Property Prediction), 'Commercial Services' (DPGEN, Model Test, APEX, dpt), 'AI4S Industrial Software' (AIS 101), 'Workflow Framework' (dflow), 'Workflow Square: Data, Models and Workflow' (Model Automation, Model Evaluation, Alloy Property Prediction, Semiconductor Property Prediction, Energy Property Prediction, Organic Molecule Property Prediction, Biomolecular Property Prediction), 'Dynamic Engines' (Quantum Chemistry and DFT Solvers: ABACUS, Molecular Dynamics Solvers: GPUMD, Lammps, Gromacs, Phase Field and FEM Solvers: DeepFlame, Jax-FEM), 'AI4S Modeling Engines' (Quantum Chemistry and DFT Models: DeePKS-kit, DeePTB, LIBRI, Molecular Dynamics Models: DeePMD-kit, RiD-kit, DMFF, NEP, Mesoscopic, Macroscale Models: fealpy, DeepMR), and 'Highly Integrated Computing Platform'. At the bottom, a legend identifies project types: Community Projects (light blue), Ongoing Projects (dashed blue), External Projects (yellow), and Unmatured Solutions (light blue).

Learning-assisted modeling: Why now?



2017: AI gives us tools for representing high-dimensional functions.

2022: AI gives us tools for assimilating truly large-scale data into pre-trained models.

History of GPT: “G”enerative, “P”re-training, “T”ranformer

Attention Is All You Need [2017]

Ashish Vaswani^{*}
Google Brain
avaswani@google.com

Noam Shazeer^{*}
Google Brain
noam@google.com

Niki Parmar^{*}
Google Research
nikip@google.com

Jacob Uszkoreit^{*}
Google Research
uszkoreit@google.com

Llion Jones^{*}
Google Research
llion@google.com

Aidan N. Gomez[†]
University of Toronto
aidan@cs.toronto.edu

Lukasz Kaiser^{*}
Google Brain
lukasz.kaiser@google.com

Illia Polosukhin[‡]
illia.polosukhin@gmail.com

Abstract

The dominant sequence transduction models are convolutional neural networks that include an performing task and also consider the encoder state. We present a new model architecture based solely on attention mechanisms, dispensing entirely. Experiments on two machine translation benchmarks show that our model can learn English to German translation task, improving over the ensemble, by over 2 BLEU. On the WMT 2014 our model establishes a new single-model state-of-the-art of 3.5% improvement over a state-of-the-art models from the literature. We show that it other tasks by applying it successfully to English and limited training data.

Improving Language Understanding by Generative Pre-Training [2018]

Alec Radford^{*} Karthik Narasimhan^{*} Tim Salimans^{*} Ilya Sutskever^{*}
OpenAI OpenAI OpenAI OpenAI
alec@openai.com karthikn@openai.com tim@openai.com ilyas@openai.com

Abstract

Natural language understanding comprises a wide range of diverse tasks such as textual entailment, question answering, semantic similarity assessment, and text generation. In contrast to supervised learning tasks with abundant labeled data for learning these specific tasks is scarce, making it challenging for discriminatively trained models to perform adequately. We demonstrate that large gains on these tasks can be realized by *generative pre-training* of a language model on a diverse corpus of unlabeled text, followed by *task-specific fine-tuning* on each specific task. In our experiments, we make use of task-specific input transformations during fine-tuning to achieve effective transfer while requiring minimal changes to the model architecture. We demonstrate the effectiveness of our approach on a wide range of benchmarks for natural language understanding. Our general task-agnostic model outperforms discriminatively trained models that were specifically trained for each task. Our results are summarized in Table 1, the state of the art in 9 out of the 12 tasks studied. For instance, we achieve absolute improvements of 8.9% on commonsense reasoning (Stories Cloze Test), 5.7% on question answering (RACE), and 1.5% on textual entailment (MultiNLI).

Language Models are Unsupervised Multitask Learners [2019]

Alec Radford^{*} Jeffrey Wu^{*} Rewon Child[†] David Luan[†] Dario Amodei^{**†} Ilya Sutskever^{**†}

Abstract

Natural language processing tasks, such as question answering, machine translation, reading comprehension, and summarization, are typically approached with supervised learning on task-specific datasets. We demonstrate that language models can learn to perform these tasks with any explicit supervision when trained on a new collection of millions of webpages called WebText. When conditioned on a document plus questions, the answers generated by the language model reach 55 F1 on the CoQA dataset, exceeding the 50 F1 of 12 out of 4 baseline systems without using the 127,000+ training examples. The capacity of the language model is essential to the success of zero-shot task transfer and increasing it is fast: it is a 1.6x improvement in 10 minutes of fine-tuning, but still and model refines parag- cusing to its natural

competent generalists. We would like to move towards more general systems which can perform many tasks – eventually without the need to manually create and label a training dataset for each one.

The most common approach to creating ML systems is to collect a dataset of training examples demonstrating correct behavior for a desired task, train a system to imitate these behaviors, and then test its performance on independent and identically distributed (IID) held-out examples. This has served well to make progress on narrow experts. But the often erratic behavior of captioning models (Lake et al., 2015), language models (Radford et al., 2018; Brown et al., 2019), and image classifiers (Aclom et al., 2018) on the diversity and variety of possible inputs highlights some of the shortcomings of this approach.

Language Models are Few-Shot Learners [2020]

Tom B. Brown^{*} Benjamin Mann^{*} Nick Ryder^{*} Melanie Subbiah^{*}
Jared Kaplan^{*} Prafulla Dhariwal^{*} Arvind Neelakantan^{*} Pranav Shyam^{*} Girish Sastry^{*}
Amanda Askell^{*} Sandhini Agarwal^{*} Ariel Herbert-Voss^{*} Gretchen Krueger^{*} Tom Henighan^{*}
Rewon Child^{*} Aditya Ramesh^{*} Daniel M. Ziegler^{*} Jeffrey Wu^{*} Clemens Winter^{*}
Christopher Hesse^{*} Mark Chen^{*} Eric Sigler^{*} Mateusz Litwin^{*} Scott Gray^{*}
Benjamin Chess^{*} Jack Clark^{*} Christopher Berner^{*}
Sam McCandlish^{*} Alec Radford^{*} Ilya Sutskever^{*} Dario Amodei^{*}
OpenAI

“Mystery” of GPT: “Scaling law” and “Emergence”

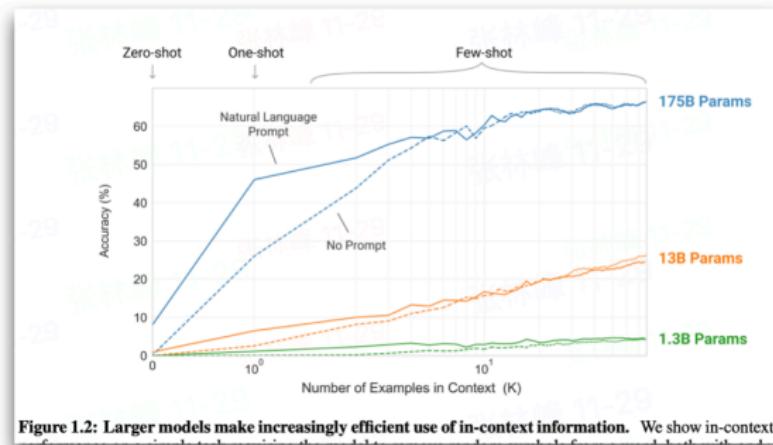
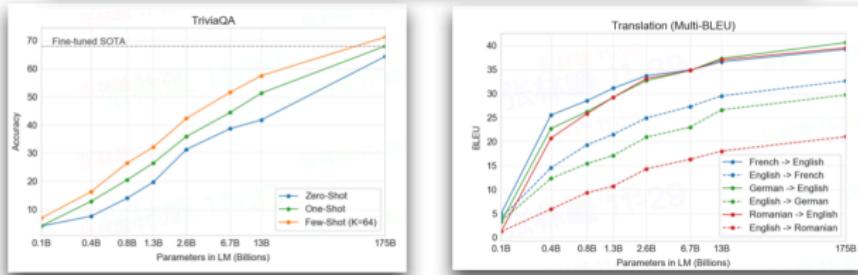


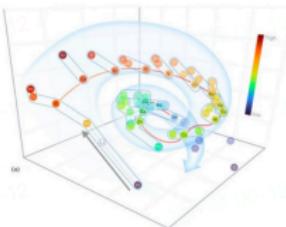
Figure 1.2: Larger models make increasingly efficient use of in-context information. We show in-context



Large science models enable transfer learning and out-of-distribution predictions

DPA-2

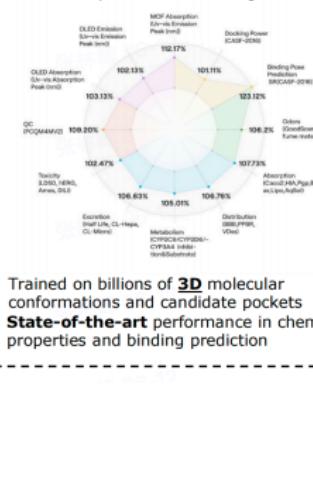
A Universal Large Atomic Model for Molecular and Material Simulation



- Pretrain Machine Learning Potential (MLP) model covering **>70 elements**
 - Can be fine-tuned to a new system for effic./accurate **PES and MD simulations**

Uni-Mol

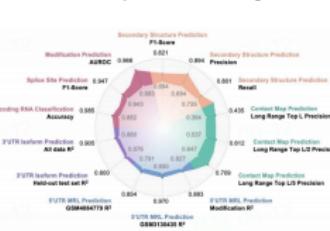
A Large Scale Pretrained Model for 3D Molecular Representation Learning Framework



- Trained on billions of **3D** molecular conformations and candidate pockets
 - **State-of-the-art** performance in chemical properties and binding prediction

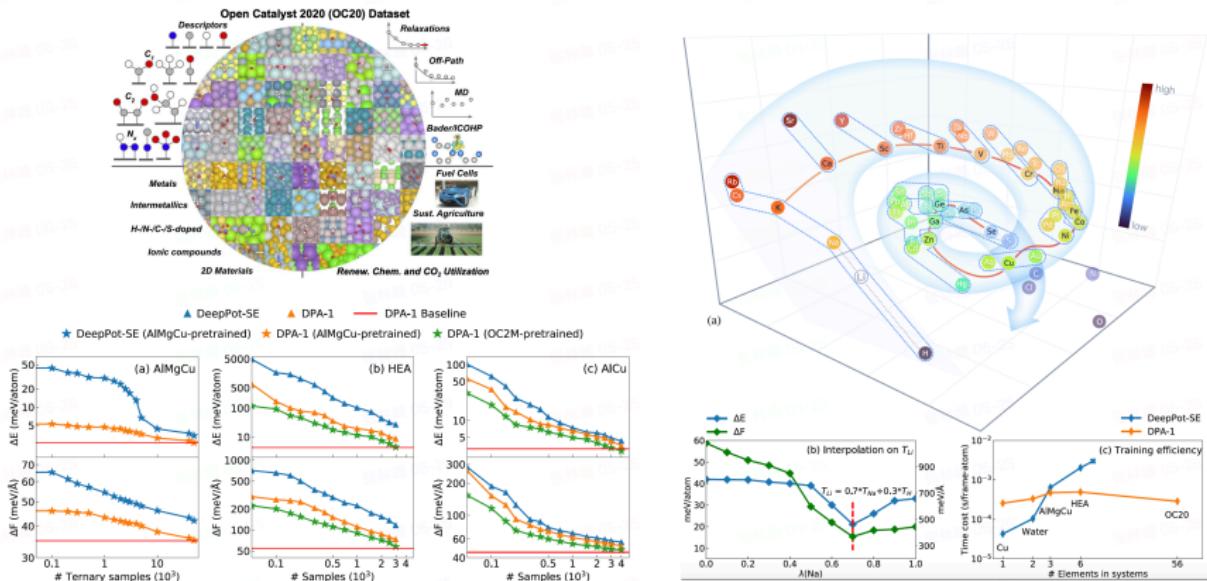
Uni-RNA

A Large Scale Pretrained Model for Nucleic Acid Sequences Embedding



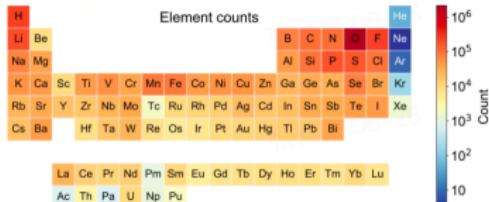
- Trained on **billions** of RNA sequences
 - **State-of-the-art** performance in RNA structure and function prediction tasks

DPA-1: pre-trained model transfers better

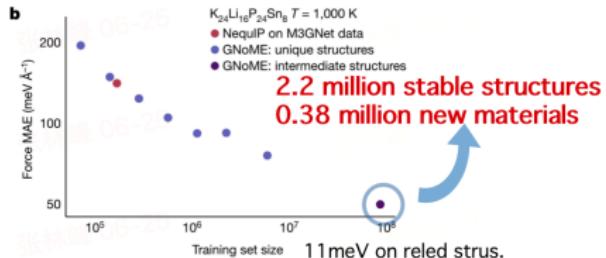


<https://arxiv.org/abs/2208.08236>

More recent efforts towards universal models



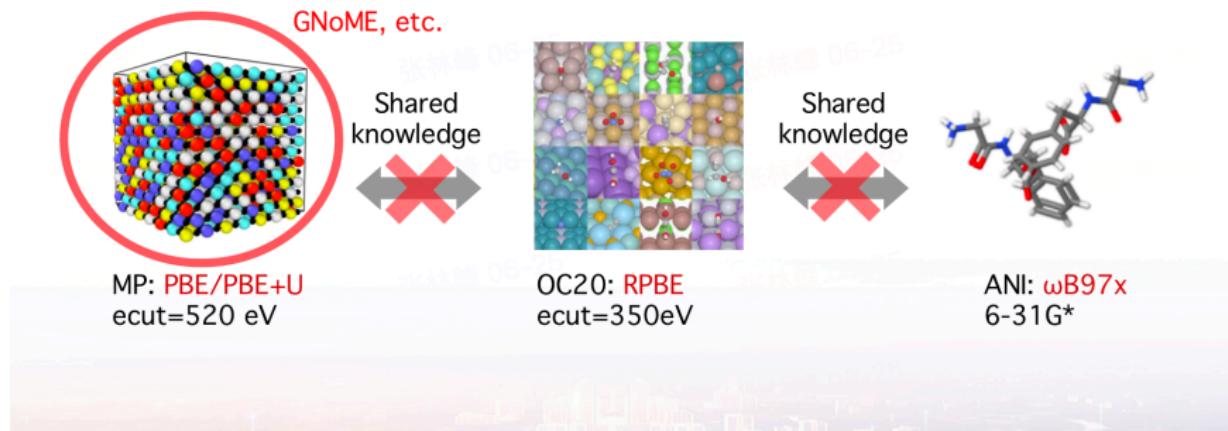
The distribution of the MPF.2021.2.8 dataset.



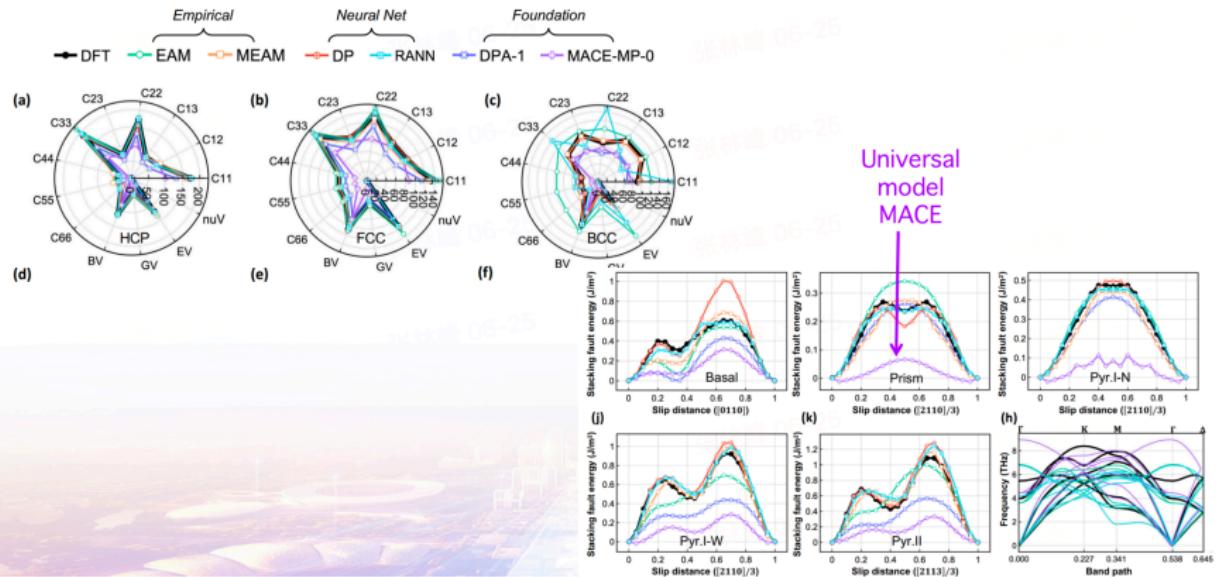
Model	Time	Training data	#data	#element	Training scheme
ALIGNN	2022	JARVIS-DFT	307K	89	Single-task PES
PFP	2022	HME21	9M	45	Single-task PES
M3GNet	2022	MPF.2021.2.8	88K	89	Single-task PES
CHGNet / MACE-MP-0	2023/24	MPtraj	1.6M	89	Single-task PES
GNoME	2023	MP+ 6xAL	~100M	94	Single-task PES
MatterSim	2024	openDBs+AL	~17M	89	Single-task PES

Difficulties of universal models

Not possible to train with multi-disciplinary data.



Example: tested accuracy on Ti

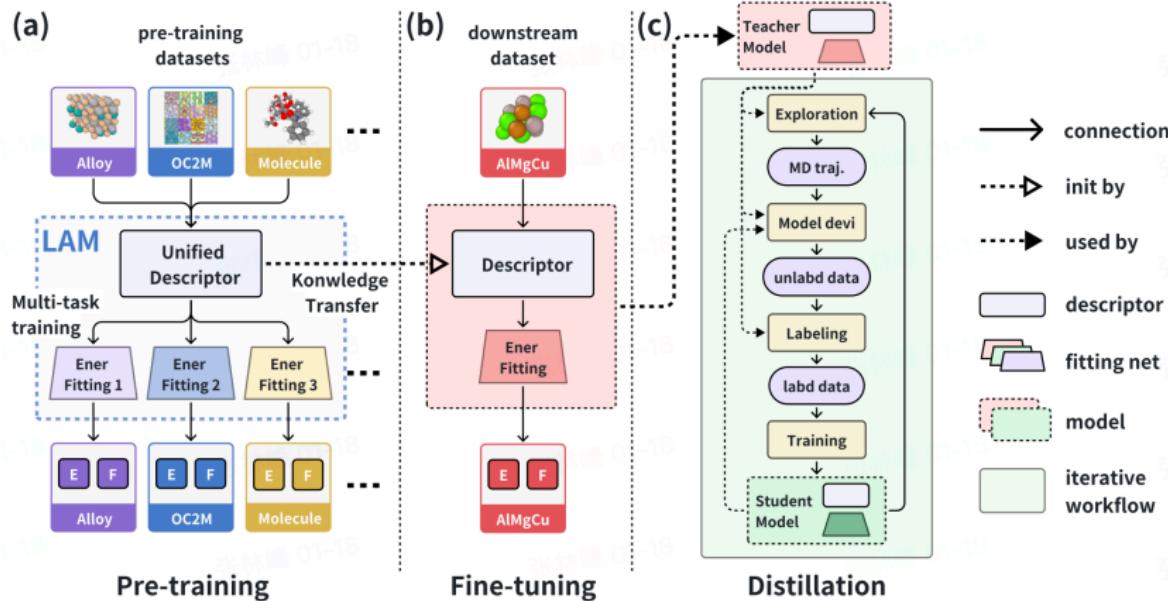


DPA-2: Towards a universal large atomic model for molecular and material simulation

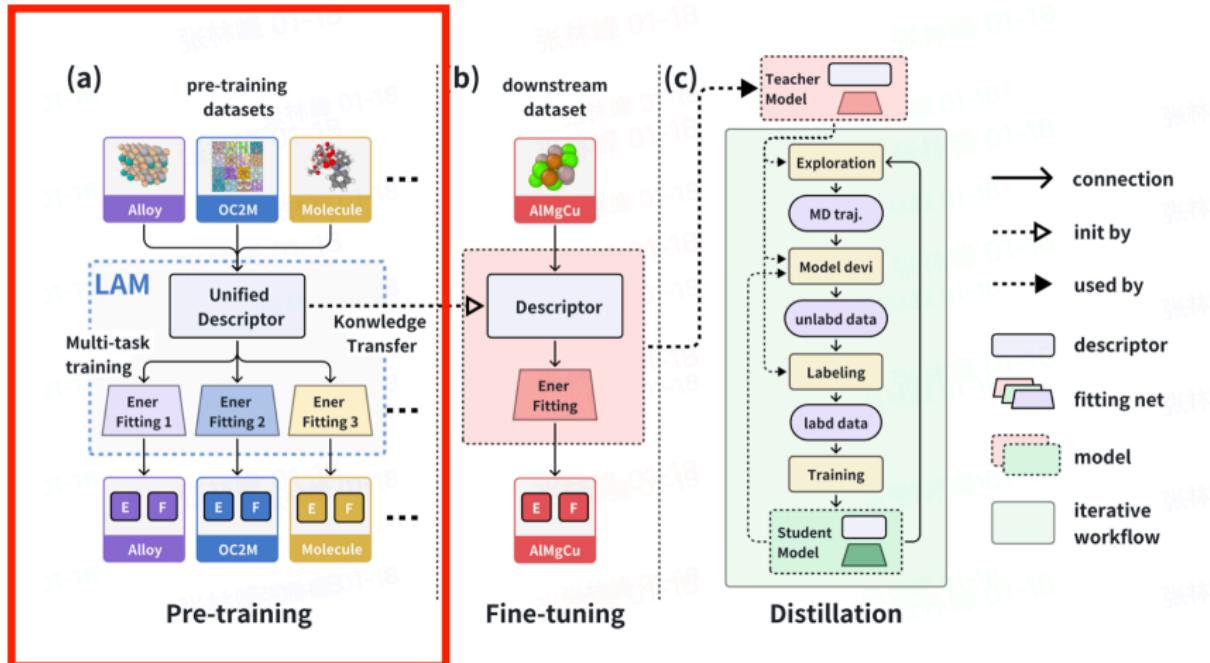
Duo Zhang^{*1,2,3}, Xinzijian Liu^{*1,2}, Xiangyu Zhang^{4,5}, Chengqian Zhang^{2,6}, Chun Cai^{1,2}, Hangrui Bi^{1,2}, Yiming Du^{4,5}, Xuejian Qin^{7,8}, Jiameng Huang^{2,9}, Bowen Li¹⁰, Yifan Shan^{7,8}, Jinzhe Zeng¹¹, Yuzhi Zhang², Siyuan Liu², Yifan Li¹², Junhan Chang^{2,13}, Xinyan Wang², Shuo Zhou^{2,14}, Jianchuan Liu¹⁵, Xiaoshan Luo^{16,17}, Zhenyu Wang^{17,18}, Wanrun Jiang¹, Jing Wu¹⁹, Yudi Yang¹⁹, Jiyuan Yang¹⁹, Manyi Yang²⁰, Fu-Qiang Gong²¹, Linshuang Zhang², Mengchao Shi², Fu-Zhi Dai¹, Darrin M. York¹¹, Shi Liu^{19,22}, Tong Zhu^{10,23,24}, Zhicheng Zhong^{7,8}, Jian Lv¹⁷, Jun Cheng^{21,25,26}, Weile Jia⁴, Mohan Chen^{1,6}, Guolin Ke², Weinan E^{1,27,28}, Linfeng Zhang^{1,2, †}, and Han Wang^{6,29, ‡}

DPA-2: Multi-task pre-training and fine-tuning

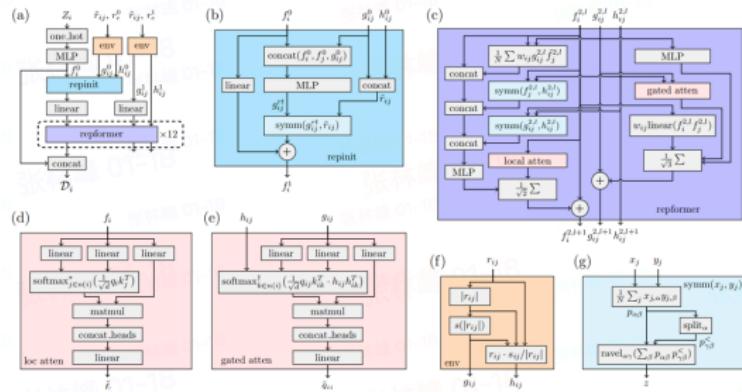
Multi-task matters.



DPA-2: Multi-task pre-training and fine-tuning



DPA-2: Multi-task pre-training and fine-tuning



Pre-training datasets					
Name	element	#train	#test	#total	weight
Alloy	53	71,482	1,240	72,722	2.0
Cathode-P	Li, Na, O, Mn, Fe, Co, Cr, Ni	58,690	6,451	65,141	1.0
Cluster-P	Pd, Ru, Al, Au, Ag, Pt, Si, Cu, Ni	139,200	14,936	154,136	1.0
Drug	H, C, N, O, F, Cl, S, P	1,379,956	24,257	1,404,213	2.0
FerroEle-P	15	6,966	760	7,726	1.0
OC2M	56	2,000,000	999,866	2,999,866	2.0
SSE-PBE-P	Li, P, S, Si, Ge	15,019	755	15,774	1.0
SemiCond-P	14	136,867	14,848	151,715	1.0
H2O-PD	H, O	46,077	2,342	48,419	1.0
Ag/Au-PBE	Ag, Au	16,696	812	17,508	0.2
Al/Mg/Cu	Al, Mg, Cu	24,252	1,145	25,397	0.3
Cu	Cu	14,596	770	15,366	0.1
Sn	Sn	6,449	276	6,725	0.1
Ti	Ti	10,054	474	10,528	0.1
V	V	14,935	738	15,673	0.1
W	W	42,297	2,100	44,397	0.1
Cl2H26	H, C	33,898	1,598	35,496	0.1
HfO2	O, Hf	27,660	917	28,577	0.1
sum	73	4,045,094	1,074,285	5,119,379	13.2

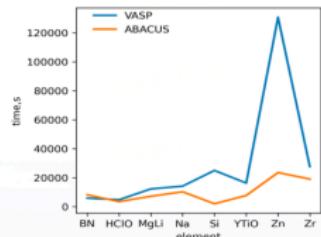
DPA-2: training datasets

327,390 (54%) materials data are labeled by ABACUS.

Name	element	Pre-training datasets		#total	weight
Alloy	53	71,482	1,240	72,722	2.0
Cathode-P	Li,Nu,O,Mn,Fe,Co,Cr,Ni	58,690	6,451	65,141	1.0
Cluster-P	Pd,Ru,Al,Au,Ag,Pt,Si,Cu,Ni	139,200	14,936	154,136	1.0
Drug	H, C, N, O, F, Cl, S, P	1,379,956	24,257	1,404,213	2.0
FerroEle-P	15	6,966	760	7,726	1.0
OC2M	57	2,000,000	999,866	2,999,866	2.0
SSE-PBE-P	Li, P, S, Si, Ge	15,019	755	15,774	1.0
SemiCond-P	14	136,867	14,848	151,715	1.0
H2O-PD	H, O	46,077	2,342	48,419	1.0
Ag Au-PBE	Ag, Au	16,696	812	17,508	0.2
AlUMgCu	Al, Mg, Cu	24,252	1,145	25,397	0.3
Cu	Cu	14,596	770	15,366	0.1
Sn	Sn	6,449	276	6,725	0.1
Ti	Ti	10,054	474	10,528	0.1
V	V	14,935	738	15,673	0.1
W	W	42,297	2,100	44,397	0.1
C12H26	H, C	33,898	1,598	35,496	0.1
HfO2	O, Hf	27,660	917	28,577	0.1
sum	73	4,045,094	1,074,285	5,119,379	13.2

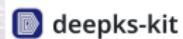


High efficiency



ML models for DFT

<https://arxiv.org/abs/2312.15492>



DPA-2: Multi-task improves zero-shot generalization

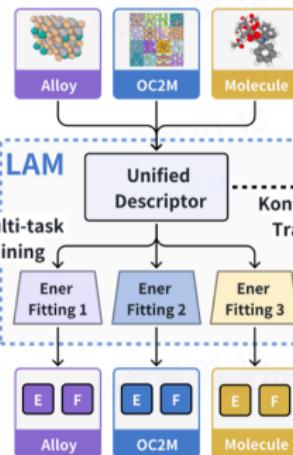
Pre-train	Downstream	Energy RMSE [meV/atom]			Force RMSE [meV/Å]		
		data std.	DPA-2 ST	DPA-2 MT	data std.	DPA-2 ST	DPA-2 MT
Cathode-P	Cathode-D	42.2	39.8	43.8	641.9	339.7	273.9
Cluster-P	Cluster-D	636.0	41.4	40.5	3605.4	238.4	190.5
FerroEle-D	FerroEle-P	43.0	6.3	3.9	868.9	236.6	138.3
SSE-PBE-P	SSE-PBE-D	79.0	40.7	6.2	789.5	635.6	162.4
SSE-PBE-P	SSE-PBESol	84.3	26.1	8.3	810.9	425.0	115.3
SemiCond-P	SemiCond-D	587.6	486.2	175.7	1755.4	1439.4	439.3
Al \cup Mg \cup Cu	AlMgCu-D	383.8	254.3	41.2	1229.5	663.7	111.8
Alloy	AlMgCu-D	383.8	74.9	48.4	1229.5	65.4	112.8
Ag \cup Au-PBE	AgAu-PBED3	906.9	222.9	192.3	878.0	236.9	63.6
H ₂ O-PD	H ₂ O-DPLR	15.6	9.1	9.3	825.2	263.5	263.4
H ₂ O-PD	H ₂ O-SCAN0	12.6	1.1	0.7	2163.2	409.2	162.9
H ₂ O-PD	H ₂ O-PBE0TS	47.0	4.9	4.7	1941.0	58.8	64.4
H ₂ O-PD	H ₂ O-PBE0TS-MD	3.3	0.5	0.6	816.1	37.6	40.8

std.: standard deviation ST: single-task; MT: multi-task

DPA-2: Multi-task pre-training and fine-tuning

(a)

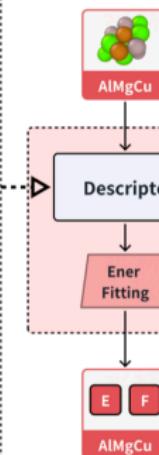
pre-training datasets



Pre-training

(b)

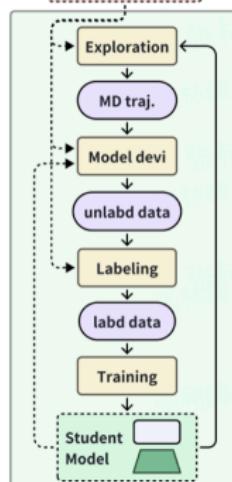
downstream dataset



Fine-tuning

(c)

Teacher Model



→ connection

···→ init by

···→ used by

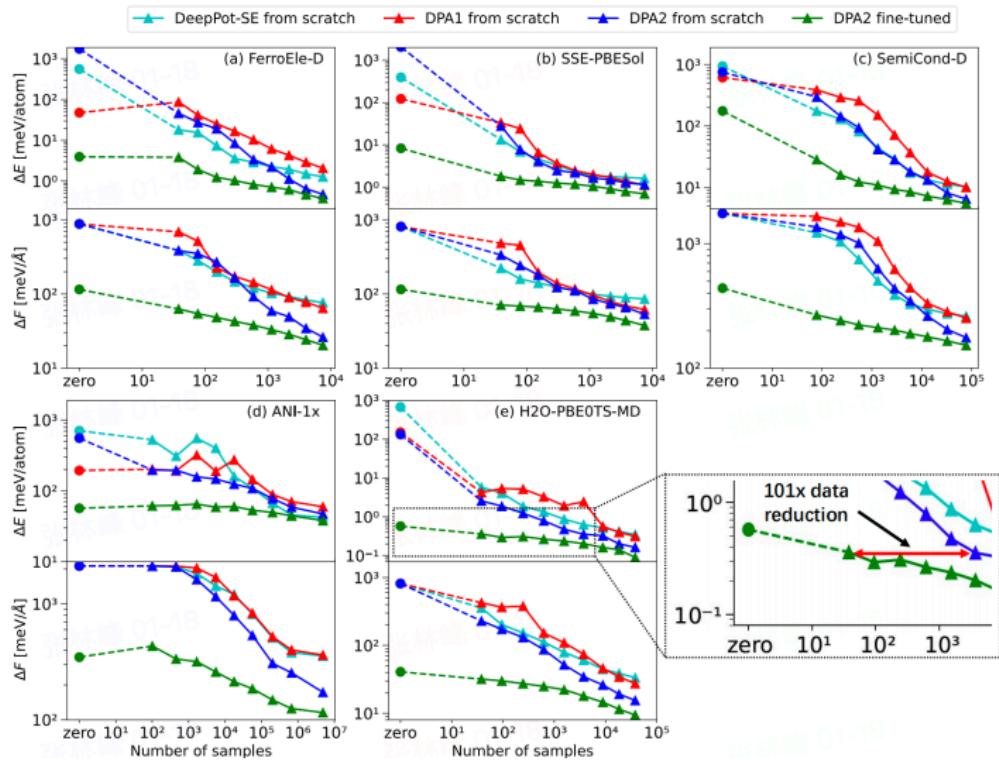
descriptor

fitting net

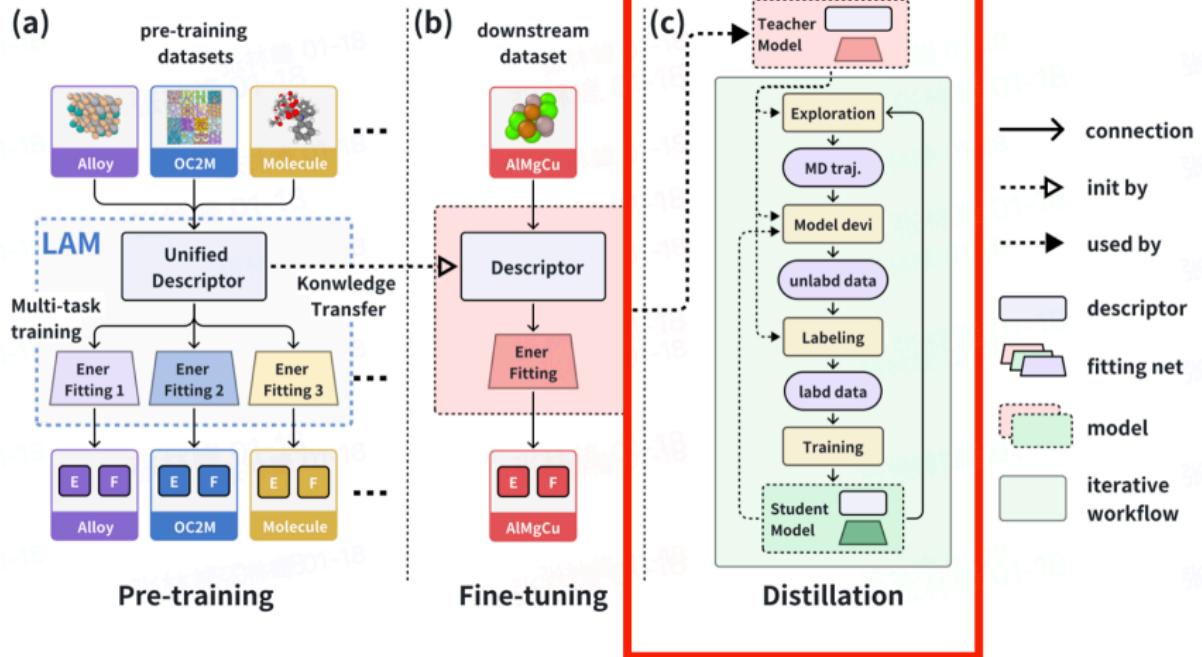
model

iterative workflow

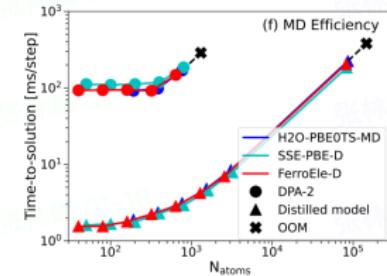
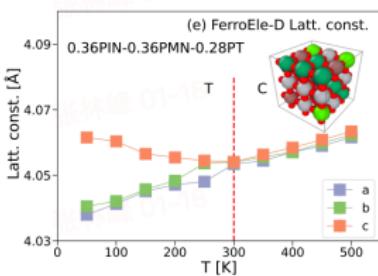
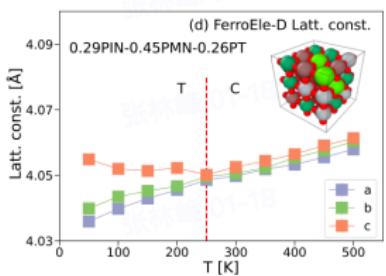
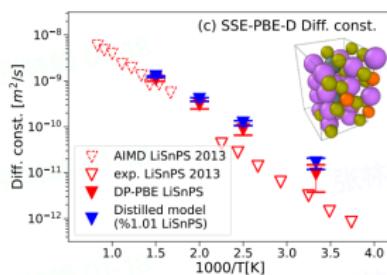
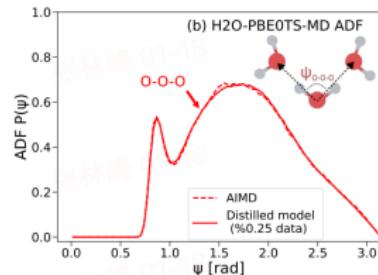
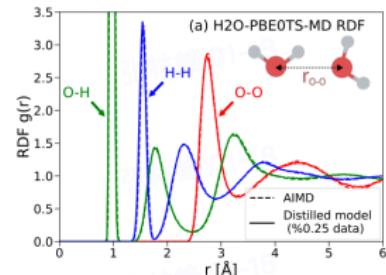
DPA-2: Multi-task pre-training and fine-tuning



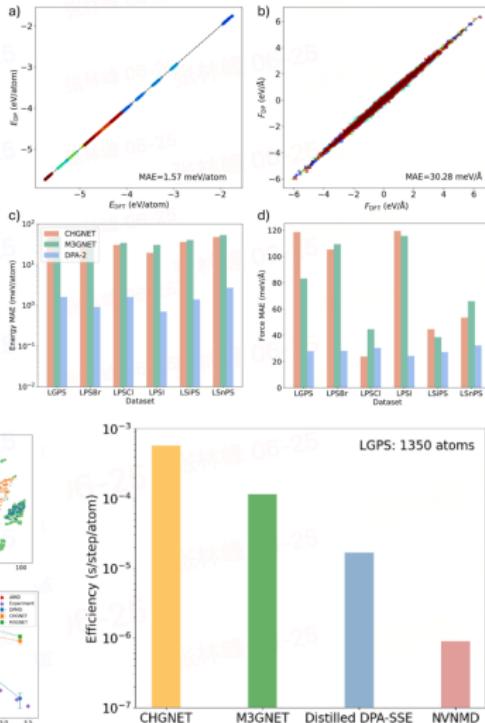
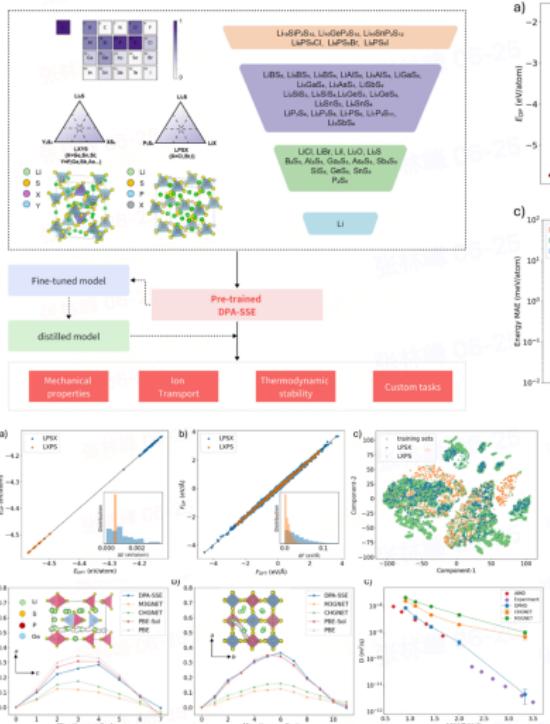
DPA-2: Multi-task pre-training and fine-tuning



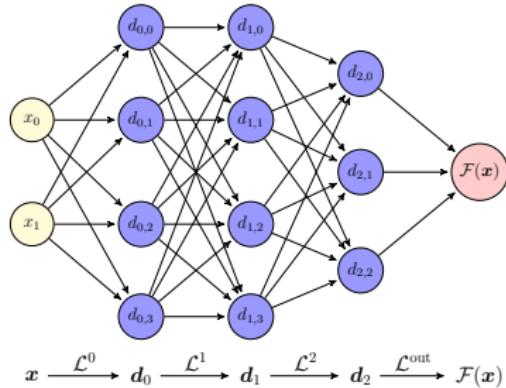
DPA-2: Multi-task pre-training and fine-tuning



Domain-Specific Model for SSE Based on DPA-2



How to organize joint efforts?



2018: Modern software engineering + open-source community.

2023: An open hub sharing data/model/workflows, building LAM together

Hugging Face: the AI community building the future

NEW Create Assistants in HuggingChat

The AI community building the future.

The platform where the machine learning community collaborates on models, datasets, and applications.

Tasks

- Text-to-Image
- Text-to-Video
- Image-to-Text
- Visual Question Answering
- Document Question Answering
- Graph Machine Learning
- Computer Vision
- Depth Estimation
- Image Classification
- Object Detection
- Image Segmentation
- Image-to-Image
- Unconditional Image Generation
- Video Classification
- Zero-Shot Image Classification

Models 469,541

- meta-llama/Llama-2-70B
- stable diffusion/stable-diffusion-v1-base-0.9
- openchat/openchat
- illyasviel/ControlNet-v1-1

Hugging Face: one of the fastest-growing open source projects
<https://github.com/huggingface/transformers/>

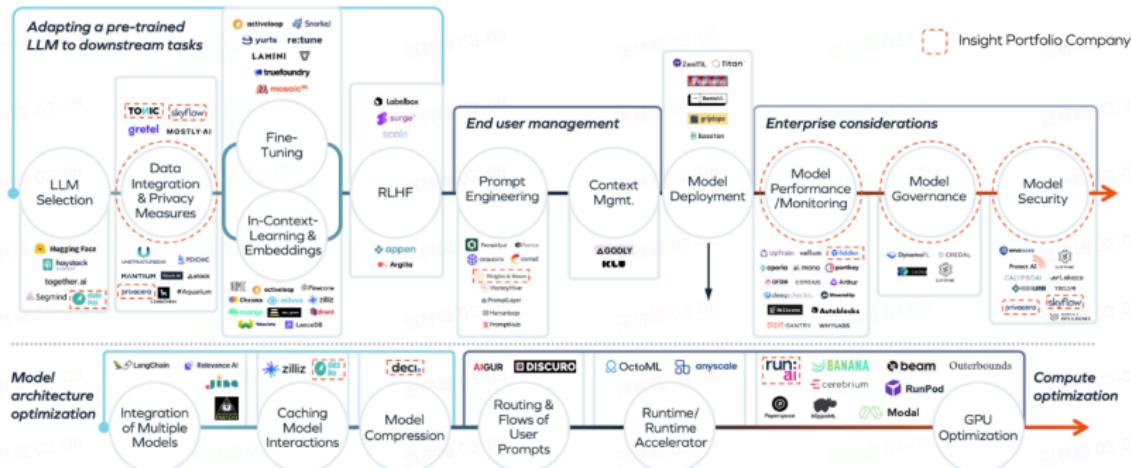
Star History

Star Date	huggingface/transformers	guganliu/transformers	guyuhuylabs/transformers	lambertchen/transformers	openchatbot/transformers	openai/transformers	open-mmlab/transformers	perceptor/transformers	yangxiao/transformers	yangxiao/transformers-gpu
2 years	~10k	~10k	~10k	~10k	~10k	~10k	~10k	~10k	~10k	~10k
6 years	~10k	~10k	~10k	~10k	~10k	~10k	~10k	~10k	~10k	~10k
10 years	~80k	~80k	~80k	~80k	~80k	~80k	~80k	~80k	~80k	~80k

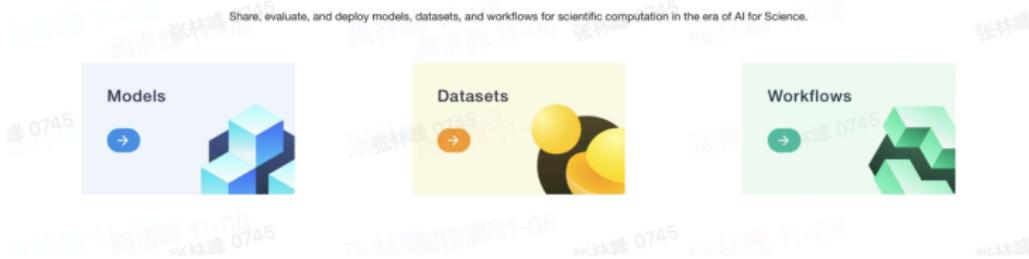
Current landscape of LLMops

LLMops adapts the MLOps tech stack for generative AI use cases

INSIGHT
PARTNERS



AIS Square & OpenLAM



OpenLAM: towards a LAM-centric ecosystem; Conquering the Periodic table! Building foundation models for atoms, not bits!

<https://www.aissquare.com/>

DP Technology ▲■■

More interesting possibilities quickly coming ...

- Integration with electronic structure methods;
- Integration with experimental characterization;
- Integration with enhanced sampling and generative modeling;
- Integration with direct property prediction tasks;
- ...

and more infrastructures quickly updated ...

- NOVA: Nurturing OpenLAM Via ABACUS;
- DeePMD-kit version 3.0;
- dflow: Python framework for constructing AI4S workflows;
- APEX, Alloy Property EXplorer; CALYPSO; DMFF; ...
- 50+ Bohrium APPs ...

Acknowledgements

- **Method development:** Weinan E, Roberto Car, Han Wang, Jiequn Han, Yixiao Chen, ...;
- **DP community:** Wanrun Jiang, Xiaoyang Wang, Yinan Wang, Tongqi Wen, Haidi Wang, Denghui Lu, Jinzhe Zeng, Yuzhi Zhang, Fengbo Yuan, Jianxing Huang, Jiduan Liu, Ping Tuo, etc.;
- **Collaborators:** Jun Cheng, Mohan Chen, Tong Zhu, Marcos Andrade, Hsinyu Ko, Lin Lin, Weile Jia, Lei Wang, etc.
- Colleagues at DP Technology and AI for Science Institute ...