

Otkrivanje novih materijala na računalu pomoću kvantne fizike i strojnog učenja

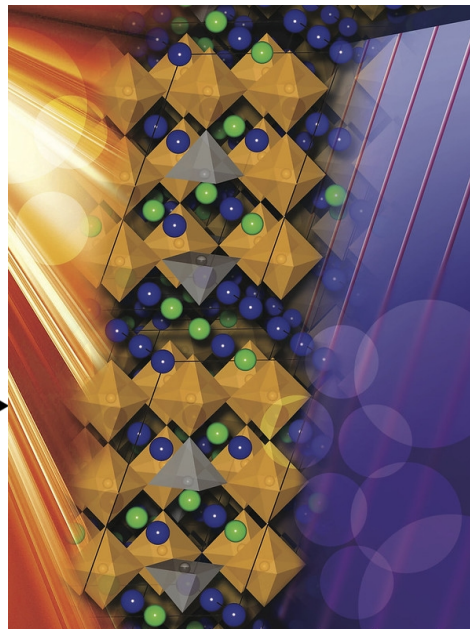
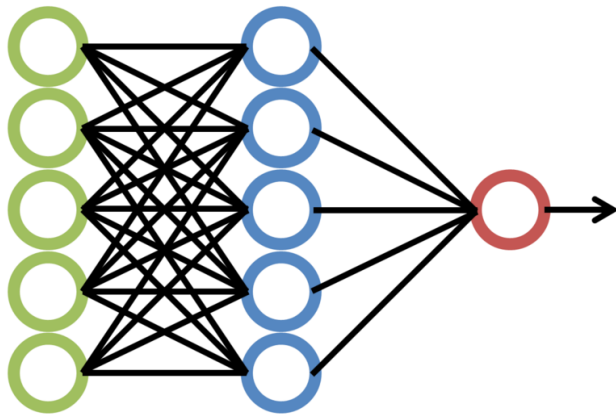
Ivor Lončarić

Institut Ruđer Bošković

Zagreb, Croatia



$$\hat{H}\Psi = E\Psi$$



HRZZ

Hrvatska zaklada
za znanost

UIP-2020-02-5675



srce

Sveučilište u Zagrebu
Sveučilišni računski centar

Uvod

Kameno doba



Brončano doba



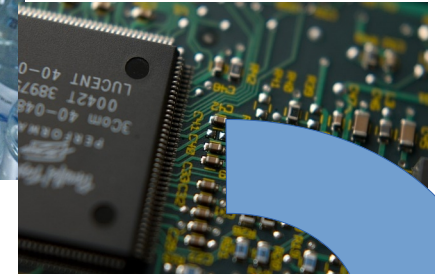
Doba čelika



Doba plastike



Doba silicija



Održivi razvoj kroz nove funkcionalne materijale

“Boston Consulting Group has estimated that nature co-design will affect more than \$30 trillion of economic activity over the next 30 years, the equivalent of 40% of current global GDP”

Uvod

Zakoni fizike za predikciju bilo kojeg svojstva materijala su poznati
~100 godina



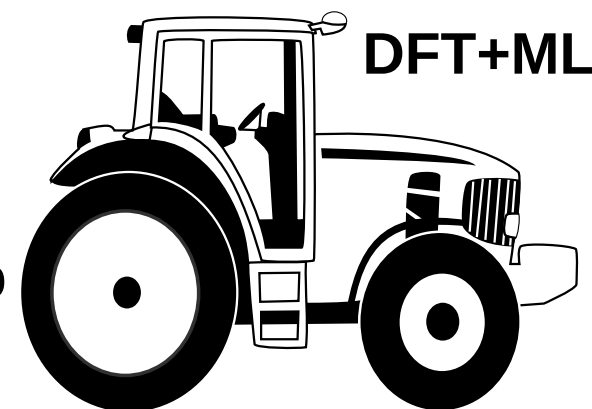
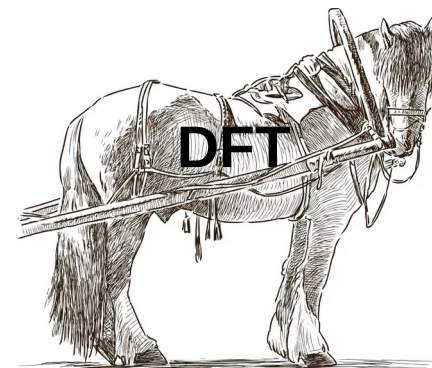
Ali njihovo rješavanje za konkretan materijal je gotovo nemoguće, a
i dobre aproksimacije su računalno jako skupe

Potreba za superračunalima - na
akademske superračunalima >50%
za fiziku/kemiju materijala (bez
superračunala svjetska irelevantnost)



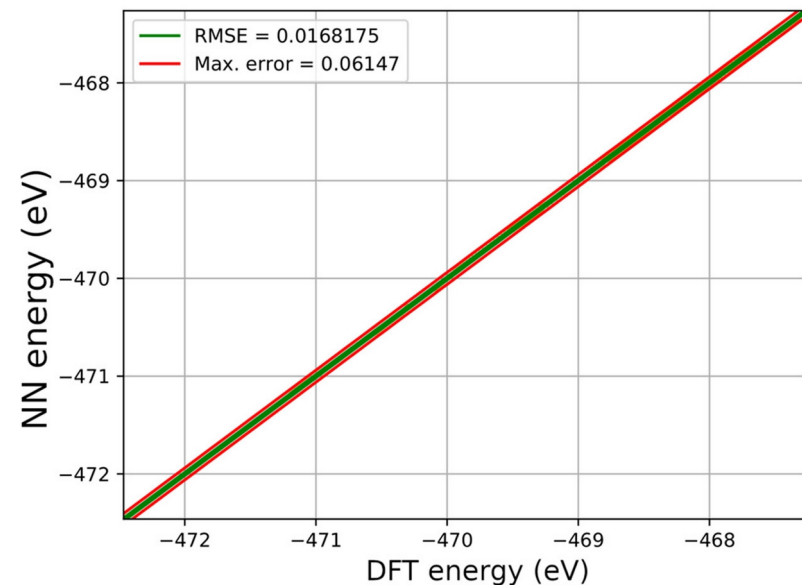
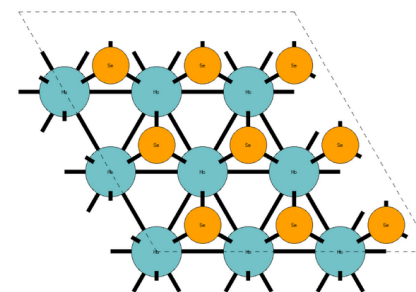
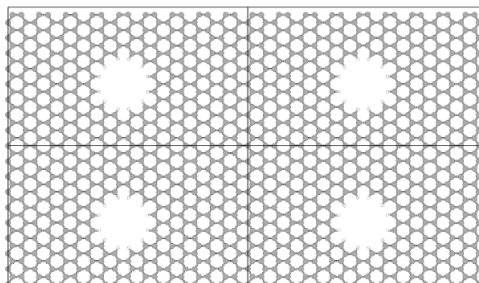
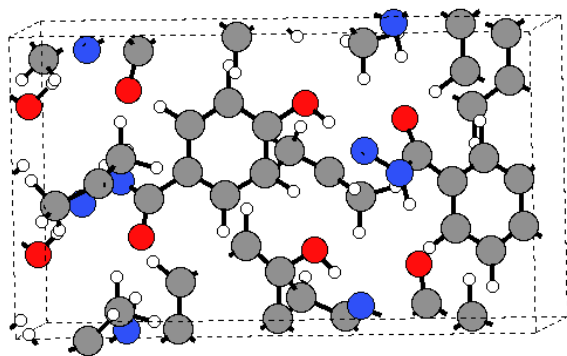
Što moja grupa radi

- Teorija funkcionala gustoće (DFT) uz svoje aproksimacije ima najbolji omjer točnosti i računalne zahtjevnosti pa je i najraširenija metoda danas za proučavanje materijala
- Zbog skaliranja n^3 ograničena upotreba na male sustave (~ 100 atoma) i kratke dinamike (nekoliko ps)
- Moguće je naučiti rezultate DFT-a strojnim učenjem na DFT bazi podataka te dalje preskočiti skupe DFT račune - glavna ideja UIP-a
- Mi kao grupa sada znamo napraviti ML modele koji imaju jednaku točnost kao DFT
- Moguće je napraviti univerzalne modele koji potpuno zamjenjuju DFT i mogu se iskoristiti za potragu za novim funkcionalnim materijalima



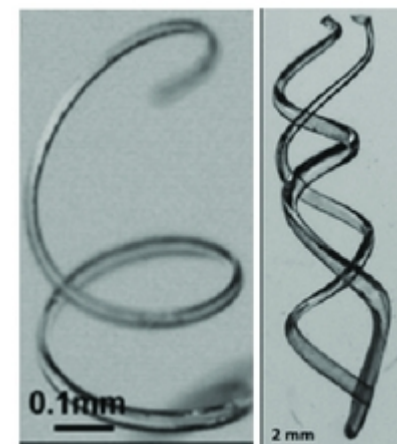
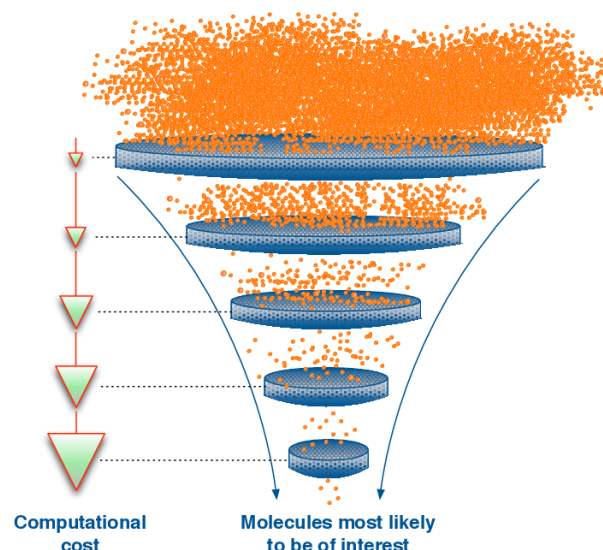
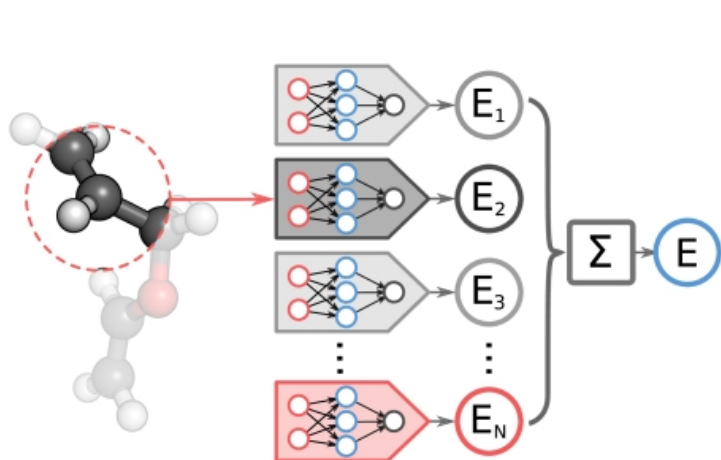
Što moja grupa sad radi a prije se nije moglo

- Tipični workflow: Stvoriti bazu podataka DFT izračuna (SRCE CPU resursi)
- Trenirati ML model na podacima (SRCE GPU resursi)
- Iterativno popraviti model (CPU+GPU)
- Modelirati materijale (CPU+GPU): npr. fazni dijagrami (p, T), mehanička i termalna svojstva, dugotrajna dinamika (kataliza), optička svojstva...

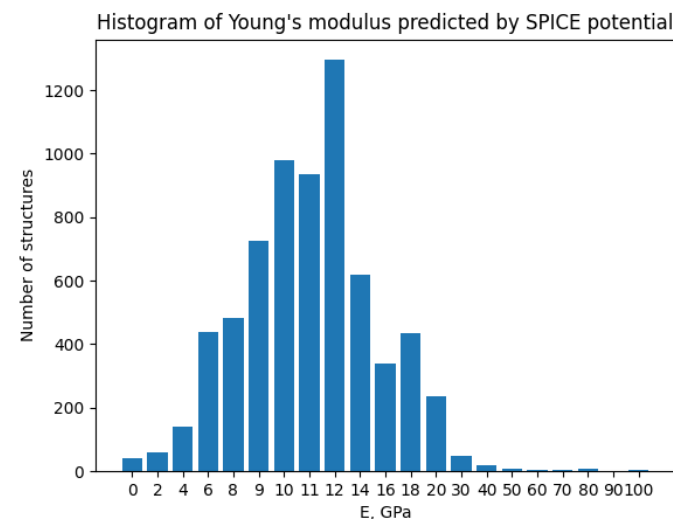


Dizajn materijala

- Moj UIP cca prije 5 godina: cilj napraviti univerzalni ML model za molekularne kristale koristeći uglavnom podatke za molekule te naći one sa zanimljivim mehaničkim svojstvima



- Imamo state-of-the art model koji omogućava planirano pretraživanje
- Potrebno je poboljšanje u vidu podataka za same molekularne kristale (računalno skupo) - novi HrZZ doktorand koji će raditi na tome uz pomoć SRCE resursa



Što rade drugi

- Za inorganske materijale je lakše stvoriti bazu podataka, a potencijalne mogućnosti su šire - trenutno utrka za najvećom bazom i najboljim modelom

Model	F1 ↑	DAF ↑	Prec ↑	Acc ↑	TPR ↑	TNR ↑	MAE ↓	RMSE ↓	R ² ↑	KSRME ↓	Training Set	Model Params	Targets	Date Added
eqV2 M	0.917	6.05	0.924	0.975	0.91	0.986	0.02	0.072	0.848		3M (102.4M) (OMat24...	86.6M	EFS _D	2024-10-18
ORB	0.88	6.04	0.924	0.965	0.841	0.987	0.028	0.077	0.824	1.73	3M (32.1M) (MPtrj+Alex)	25.2M	EFS _D	2024-10-11
MatterSim	0.859	5.65	0.863	0.957	0.856	0.975	0.026	0.08	0.812		17M (MatterSim)	182.0M	EFS _D	2024-06-16
GNoME	0.829	5.52	0.844	0.955	0.814	0.972	0.035	0.085	0.785		6M (89.0M) (GNoME)	16.2M	EF _C	2024-02-03
eqV2 S DeNS	0.815	5.04	0.771	0.941	0.864	0.953	0.036	0.085	0.788	1.67	146K (1.6M) (MPtrj)	31.2M	EFS _D	2024-10-18
ORB MPtrj	0.765	4.7	0.719	0.922	0.817	0.941	0.045	0.091	0.756	1.73	146K (1.6M) (MPtrj)	25.2M	EFS _D	2024-10-14
SevenNet	0.724	4.25	0.65	0.904	0.818	0.919	0.048	0.092	0.75	0.767	146K (1.6M) (MPtrj)	842.4K	EFS _C	2024-07-13
MACE	0.669	3.78	0.577	0.878	0.796	0.893	0.057	0.101	0.697	0.647	146K (1.6M) (MPtrj)	4.7M	EFS _C	2023-07-14
CHGNet	0.613	3.36	0.514	0.851	0.758	0.868	0.063	0.103	0.689	1.72	146K (1.6M) (MPtrj)	412.5K	EFS _C M	2023-03-03
M3GNet	0.569	2.88	0.441	0.813	0.803	0.813	0.075	0.118	0.585	1.41	63K (188.3K) (MPF)	227.5K	EFS _C	2022-09-20

- Trenutno najbolje modele/baze podataka imaju



- Promjena paradigme modeliranja materijala te dizajna novih materijala uz potrebu za dovoljnim resursima -> SRCE

Zaključak

- Predstoji promjena paradigme dizajna i modeliranja materijala
- Moderno (atomističko) modeliranje materijala već ne postoji bez superračunalnih resursa
- Moderni dizajn materijala će se u budućnosti sve više oslanjati na (ML) modele koji iako bitno brži od kvantnih modela ipak zahtijevaju ozbiljne resurse uz povećanje broja korisnika



UIP-2020-02-5675

