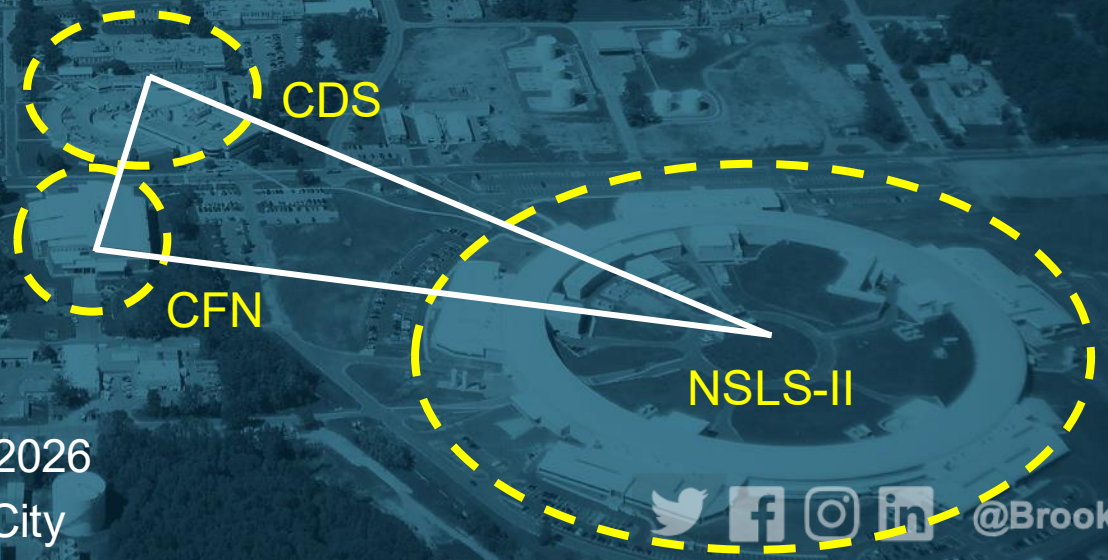
Physics-Informed AI Pipeline for Real-Time Interpretation of X-ray Absorption Spectra

Deyu Lu

Center for Functional Nanomaterials

Brookhaven National Laboratory



New York Scientific Data Summit 2026

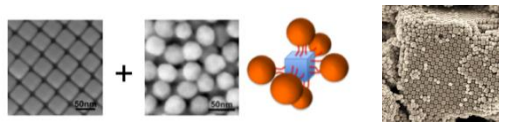
SUNY Global Center, New York City

Sep. 28, 2026

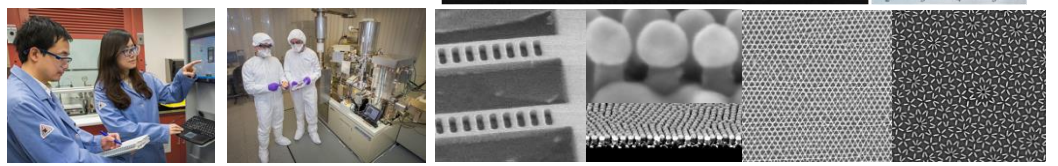
@BrookhavenLab

CFN: a DOE nano science user facility

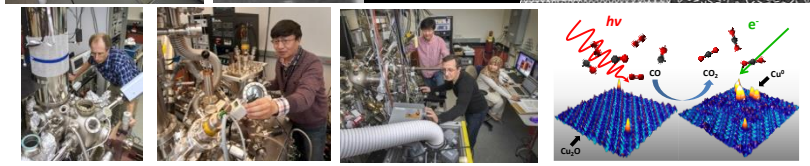
Materials Synthesis



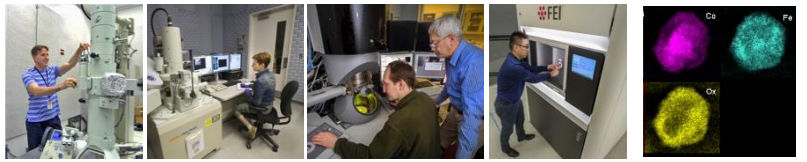
Nanofabrication



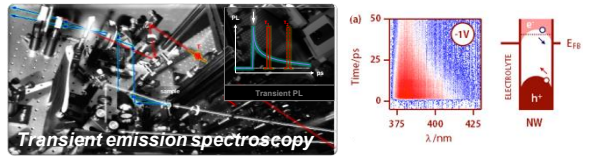
Proximal Probes



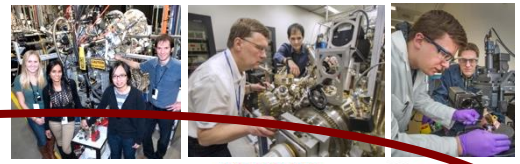
Electron Microscopy



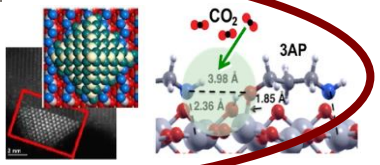
Optical Spectroscopy and Microscopy



Advanced X-ray and UV Probes @ NSLS II

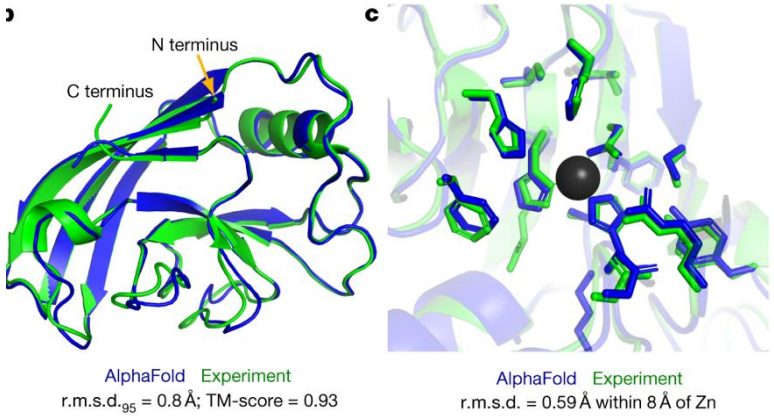


Theory and Computation

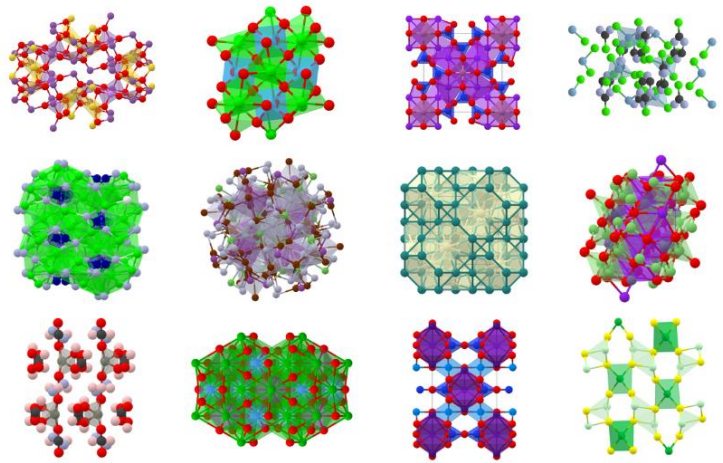




AI in materials synthesis & discovery

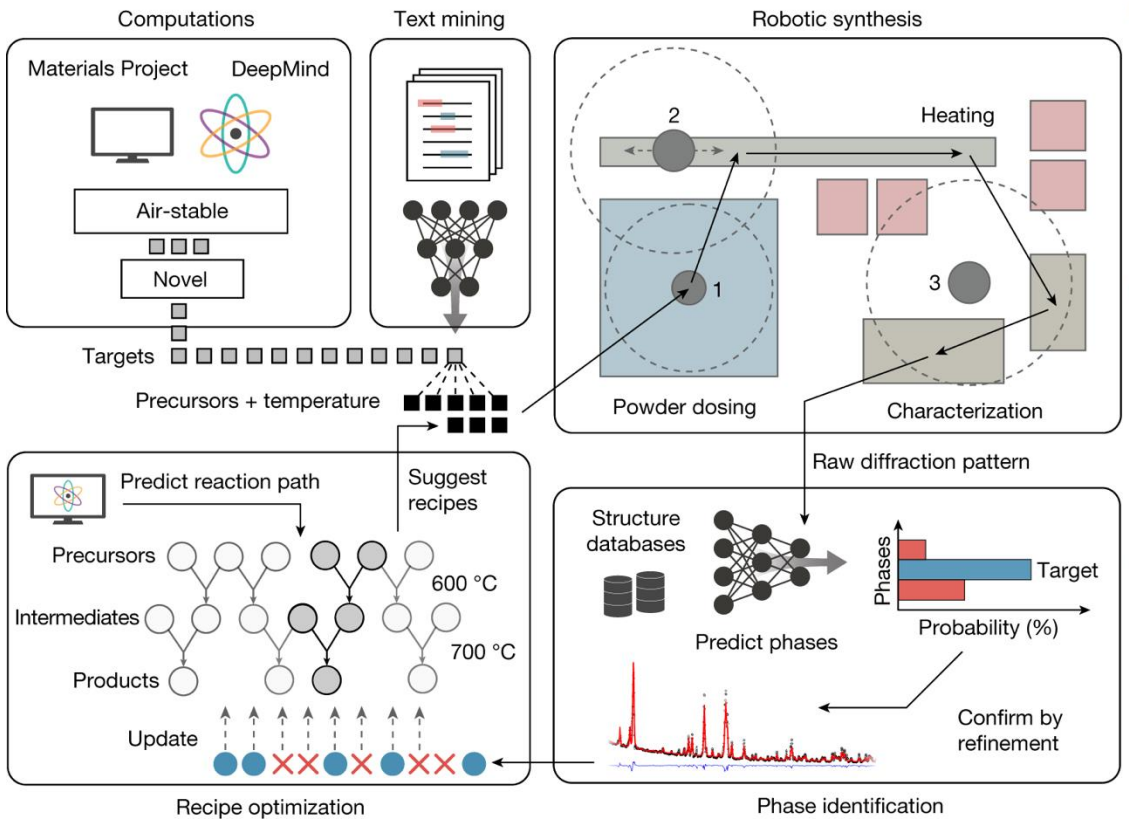


Google DeepMind, *Nature* **596**, 583 (2021)



https://newscenter.lbl.gov/2023/11/29/google-deepmind-new-compounds-materials-project/

A-Lab





Szymanski, N.J., Rendy, B., Fei, Y. *et al*. *Nature* **624**, 86 (2023)

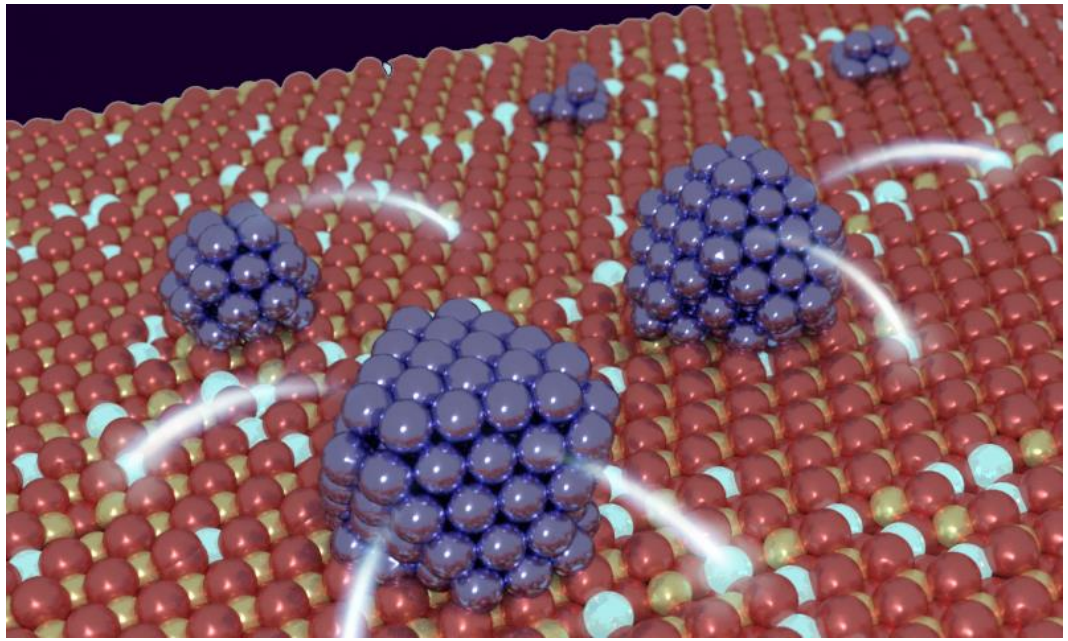
Interaction of functional materials and environment: catch atoms in motion

Catalysis:

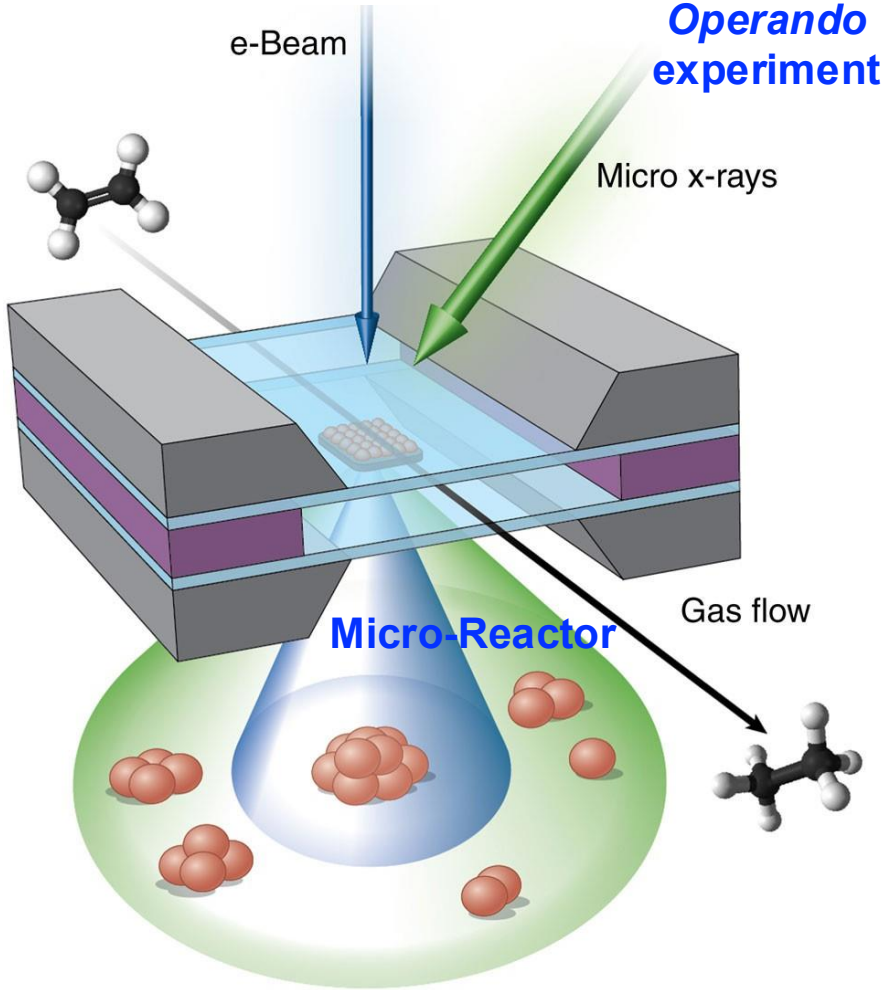
- Control the dynamic evolution of catalysts

Battery:

- Optimize the electrochemical processes



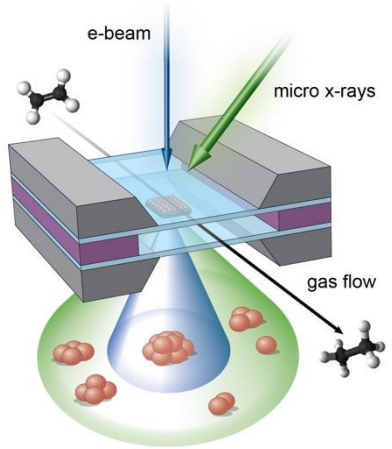
Phys.org, Dec 14, 2015. Credit: Sergey Kozlov & Oriol Lamiel



Li, Zakharov, Zhao, Tappero, Jung, Elsen, Baumann, Nuzzo, Stach & Frenkel, Nat. Comm. 6, 7583 (2015)

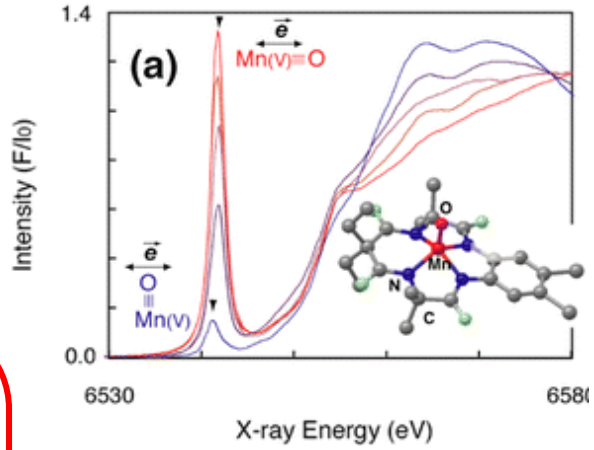
X-ray absorption spectroscopy

catalysis



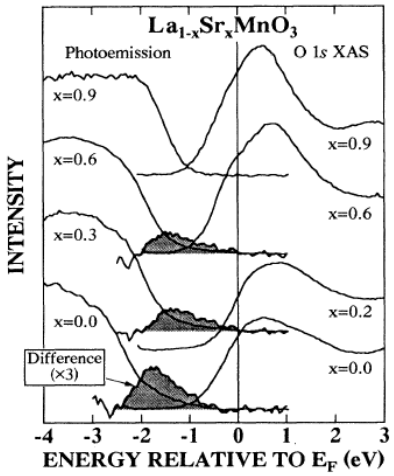
Nat. Comm., 6, 7583 (2015)

photosynthesis system

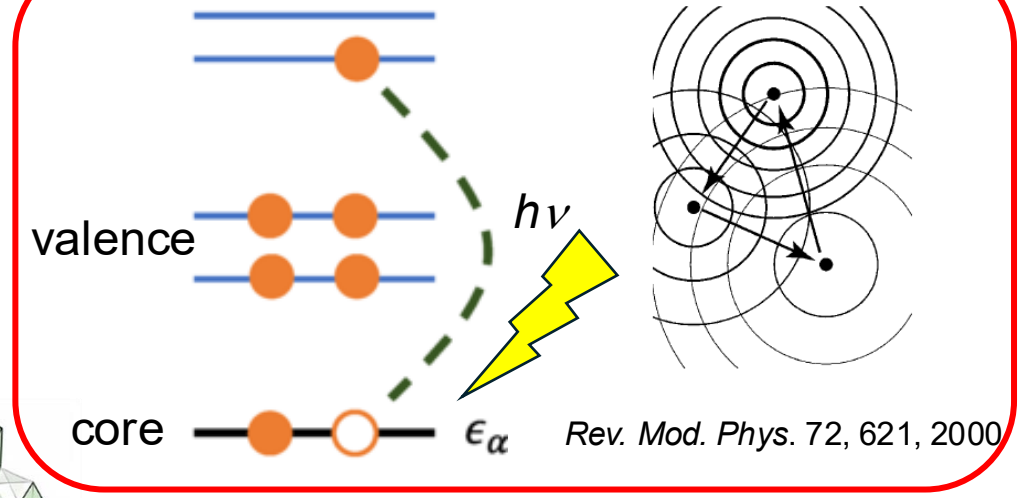


Photosynthesis Research **102**, 241 (2009)

condensed matter physics

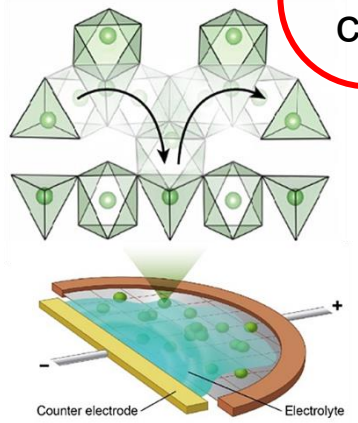


Phys. Rev. B 51, 13942 (1995)



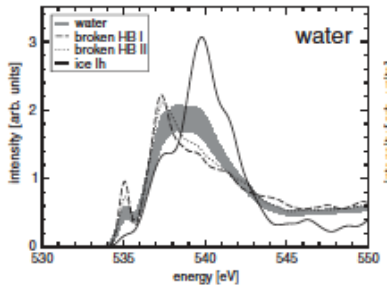
Rev. Mod. Phys. 72, 621, 2000

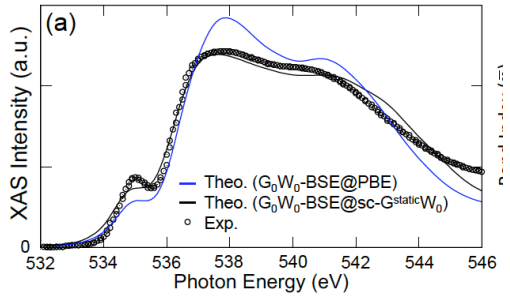
lithium ion battery

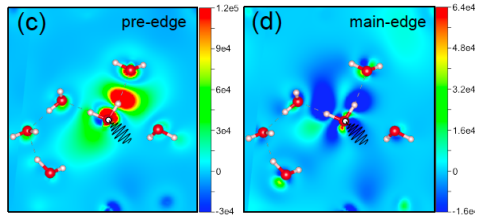


Science 367, 1030 (2020)
https://www.bnl.gov/newsroom/news.php?a=217247

liquid

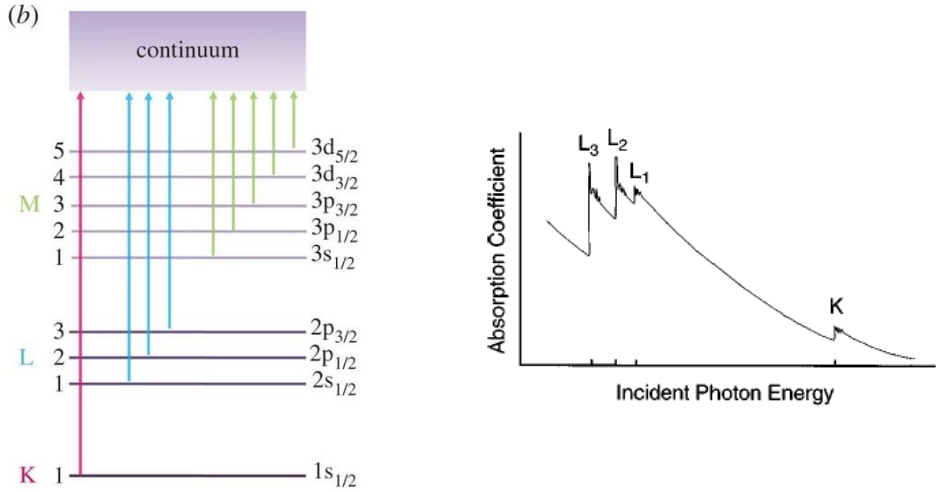




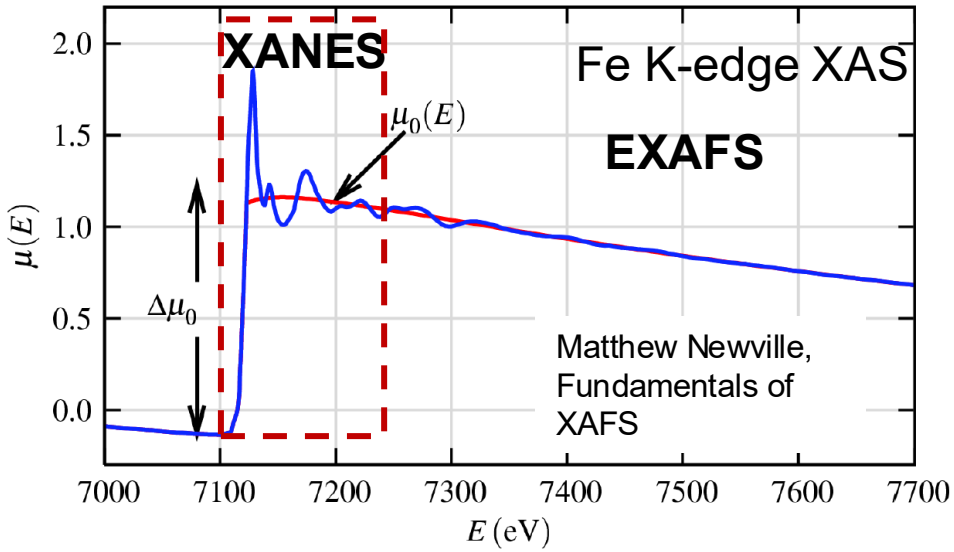


Phys. Rev. Lett., 96, 215502 (2006) *PNAS*, 119, e2201258119 (2022)

X-ray absorption spectroscopy

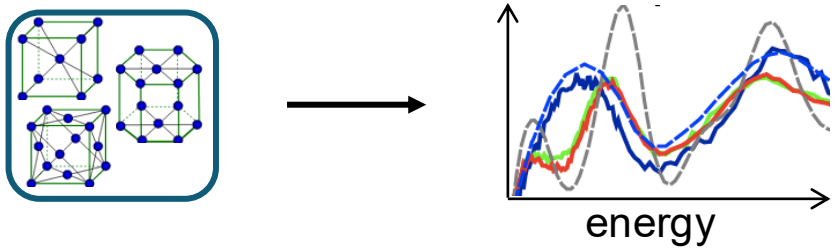


Phil. Trans. R. Soc. A 371.1995 (2013)



X-ray absorption near edge structure (XANES)

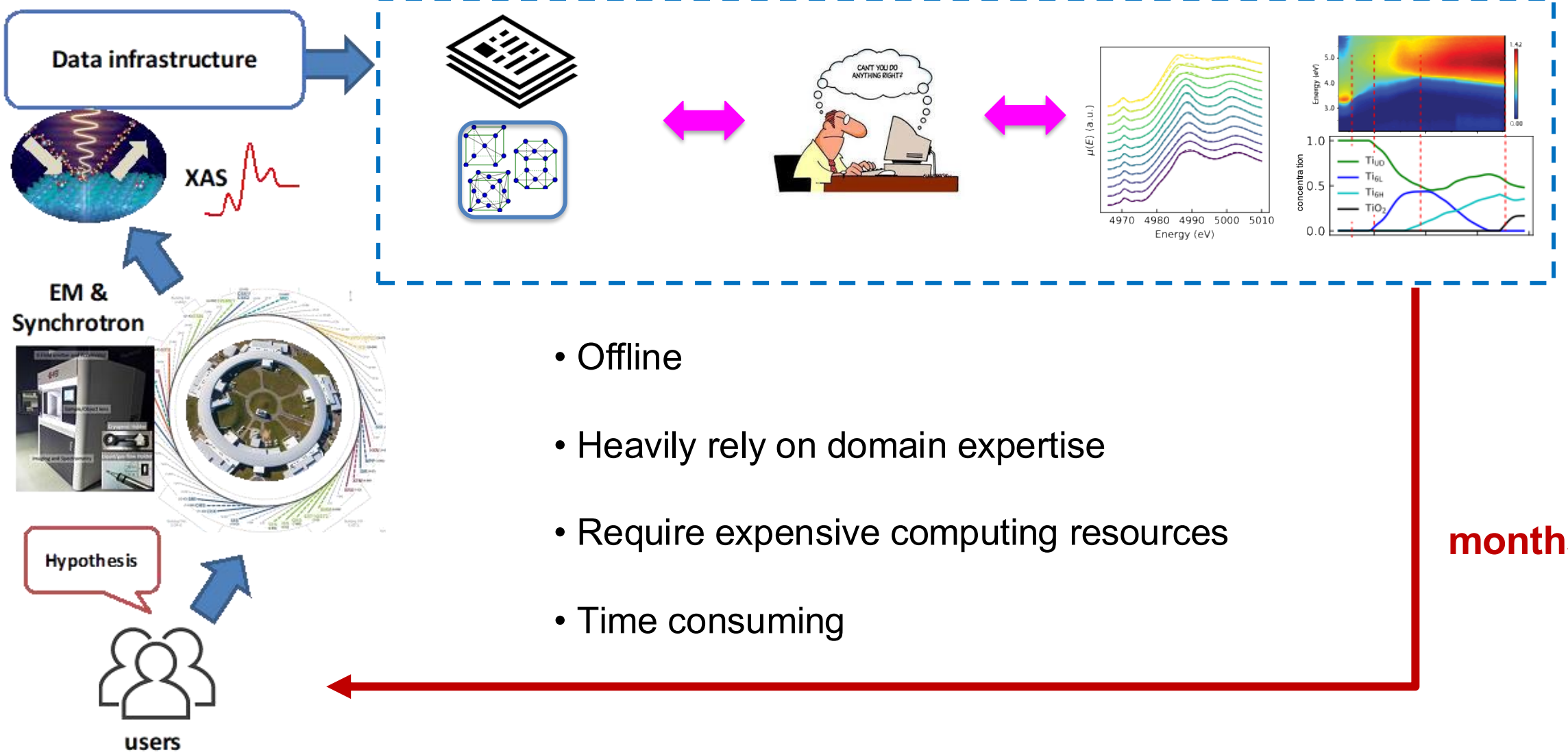
- Element specific
- Sensitive to electronic structure
- Sensitive to the local chemical environment (e.g., symmetry, coordination number and charge state)
- Great time/energy resolution



Fermi's golden rule

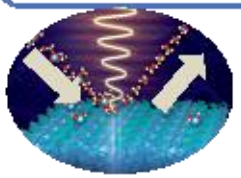
$$\sigma(\omega) = 4\pi^2 \frac{\omega}{c} \sum_F \left|\langle 0|\hat{O}|F\rangle\right|^2 \delta(\omega + E_0 - E_F)$$

Traditional XAS analysis is inefficient



Addressing the bottleneck: speed and timing

Data infrastructure



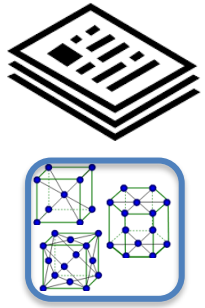
XAS

EM & Synchrotron

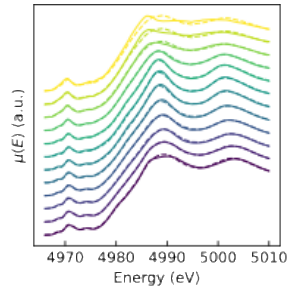


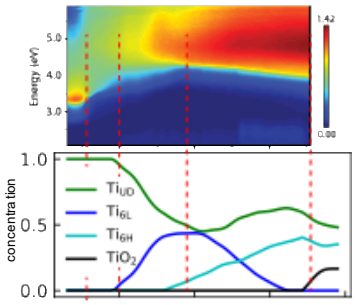
Hypothesis

users









- 1 XAS measurement: ~30 s
- Traditional workflow: ~ 3 months offline
- **Analysis latency: 260,000x longer than acquisition**
- Bridging the gap requires both AI-driven speedup and integrated workflows that deliver insight in time to guide decisions.

months

AI-enabled real-time analysis & autonomous experiment

Data infrastructure

AI AGENT

XAS

EM & Synchrotron



Hypothesis

users

Opt.: {*c*, P, T, t}

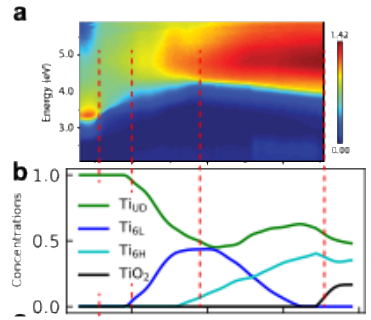
- **Phase**
- **Local structural motif**
- **Charge state**

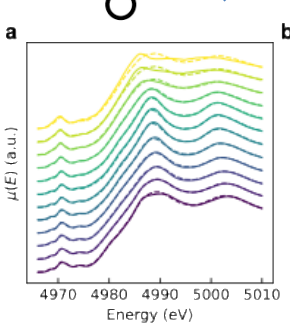
…

> Q: Show me candidate atomic structures.

Ai

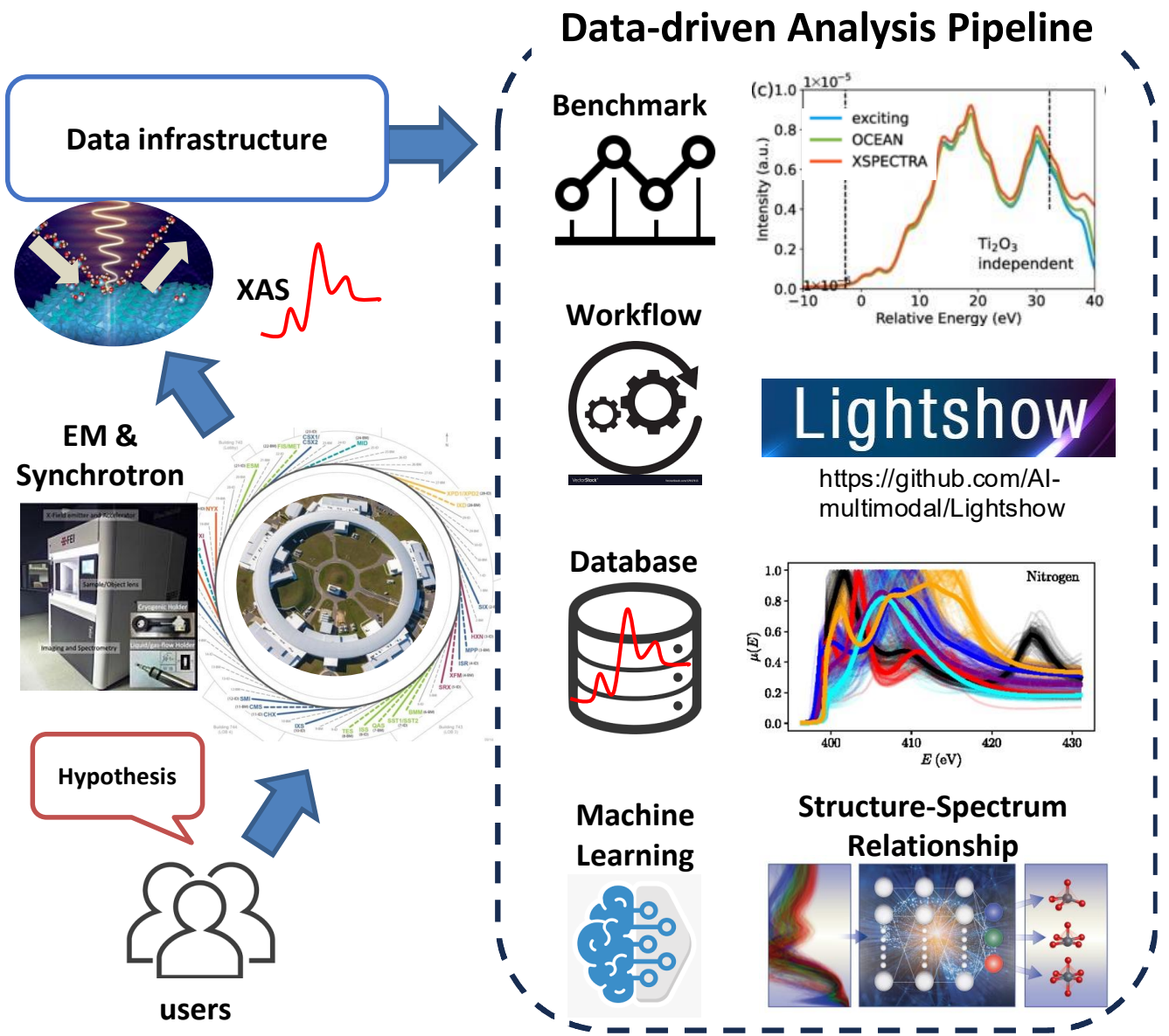
generative





Real-time feed back

Data-driven XAS analysis pipeline



- Quantify the **effects of key approximations and implementations** in simulation
- Validate theory against experiment and identify **areas of for improvement**

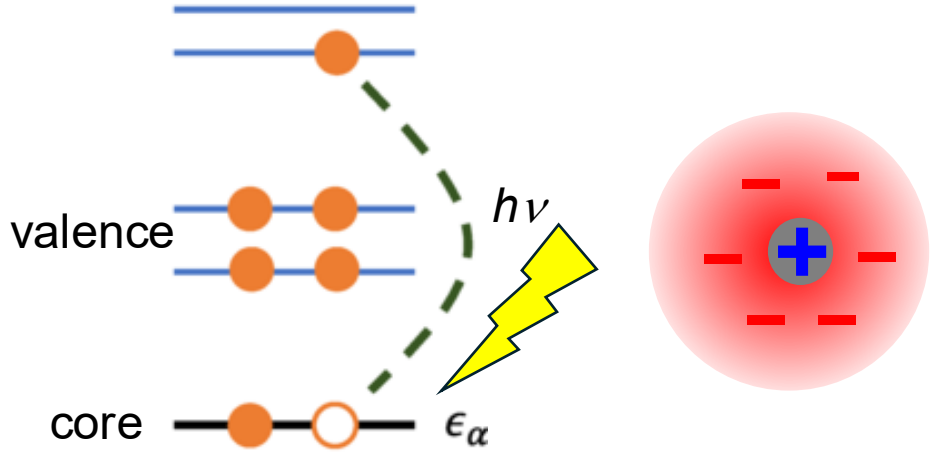
- Standardize *input* file generation for high throughput simulations
- Lower the barrier for non-expert
- Ensure the **reproducibility** and **interoperability**

- Spectral datasets for extracting structure-spectrum relationship
- **Training set** for ML surrogates
- Starting point for transfer learning models

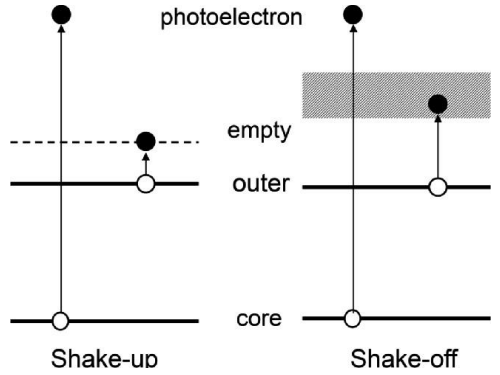
- **Forward model**: spectra prediction from structure
- **Inverse model**: physical descriptor extraction
- **Interpretable model**

Important physics in K-edge XANES

Core-hole final state effects

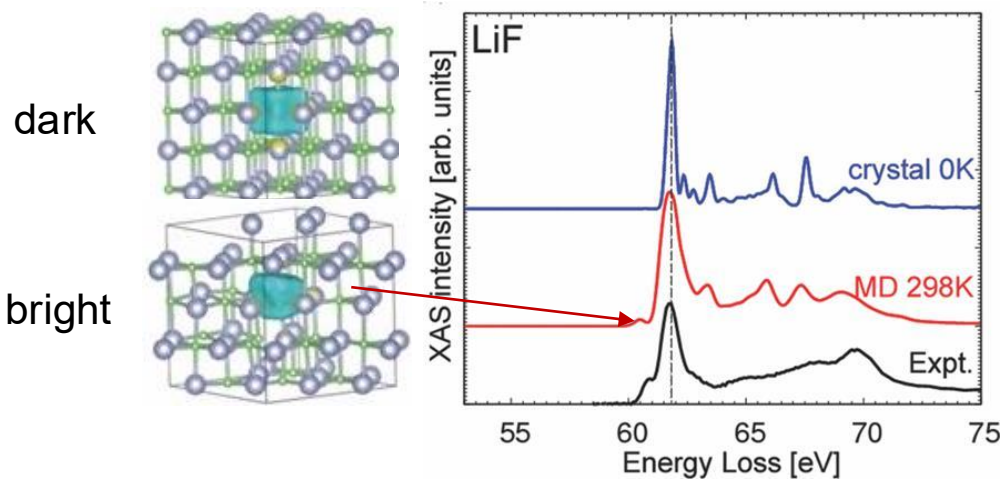


satellites

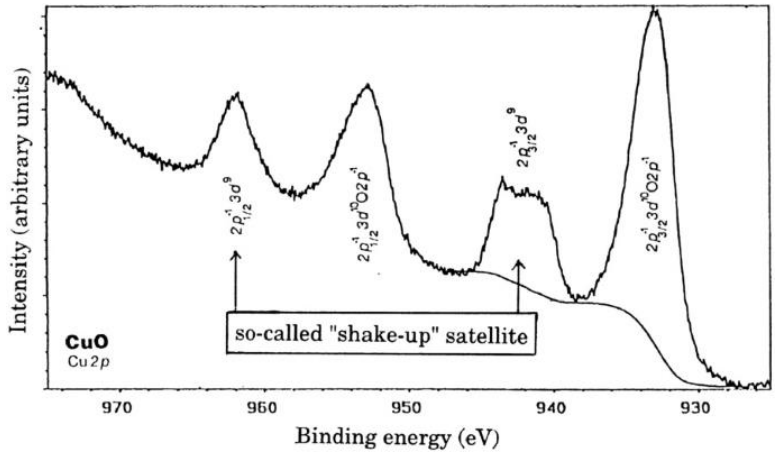


Kawai, J. (2022). Chemical Effects of Multiply Ionized Satellites. In *X-Ray Spectroscopy for Chemical State Analysis*

Finite temperature effects: e-p coupling



J. Chem. Phys. 140, 034107 (2014)



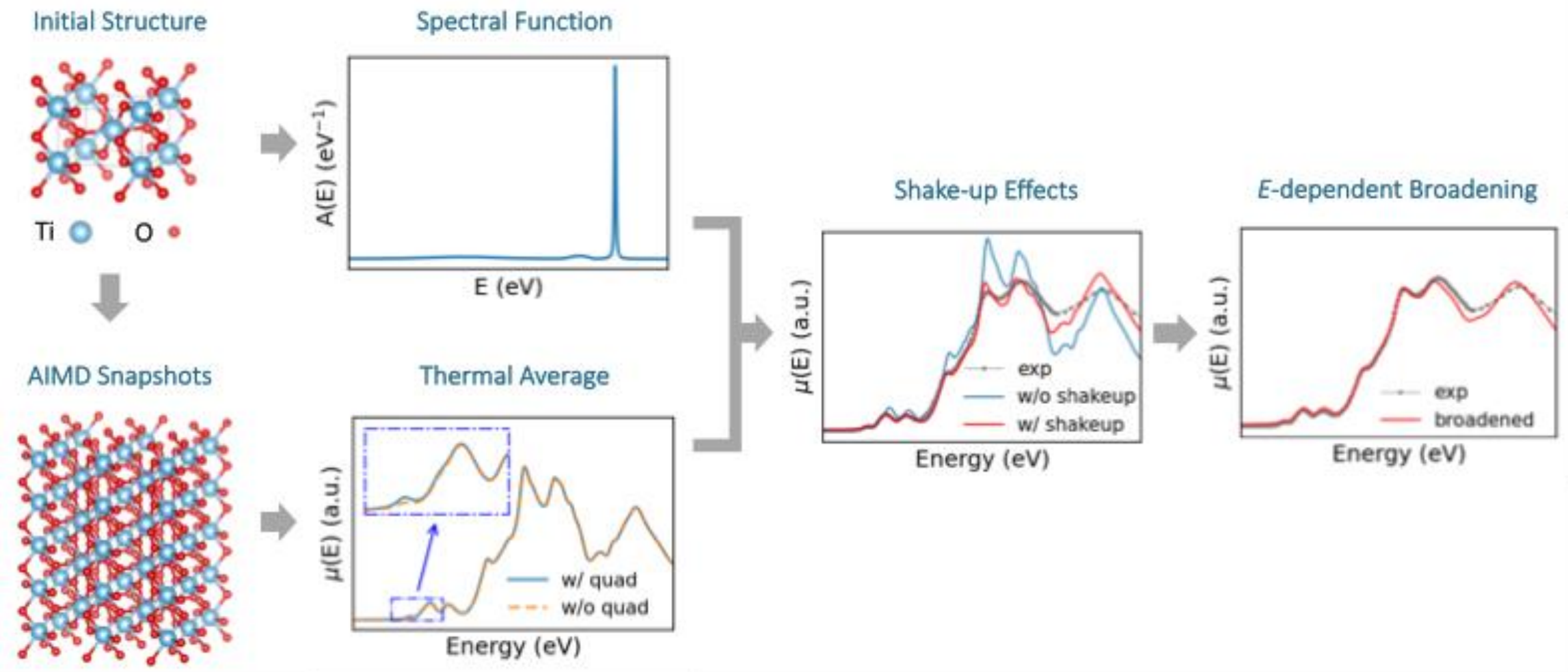
Adv. X-Ray Chem. Anal. Jpn. **21**, 149 (1990)

Workflow of XANES benchmark



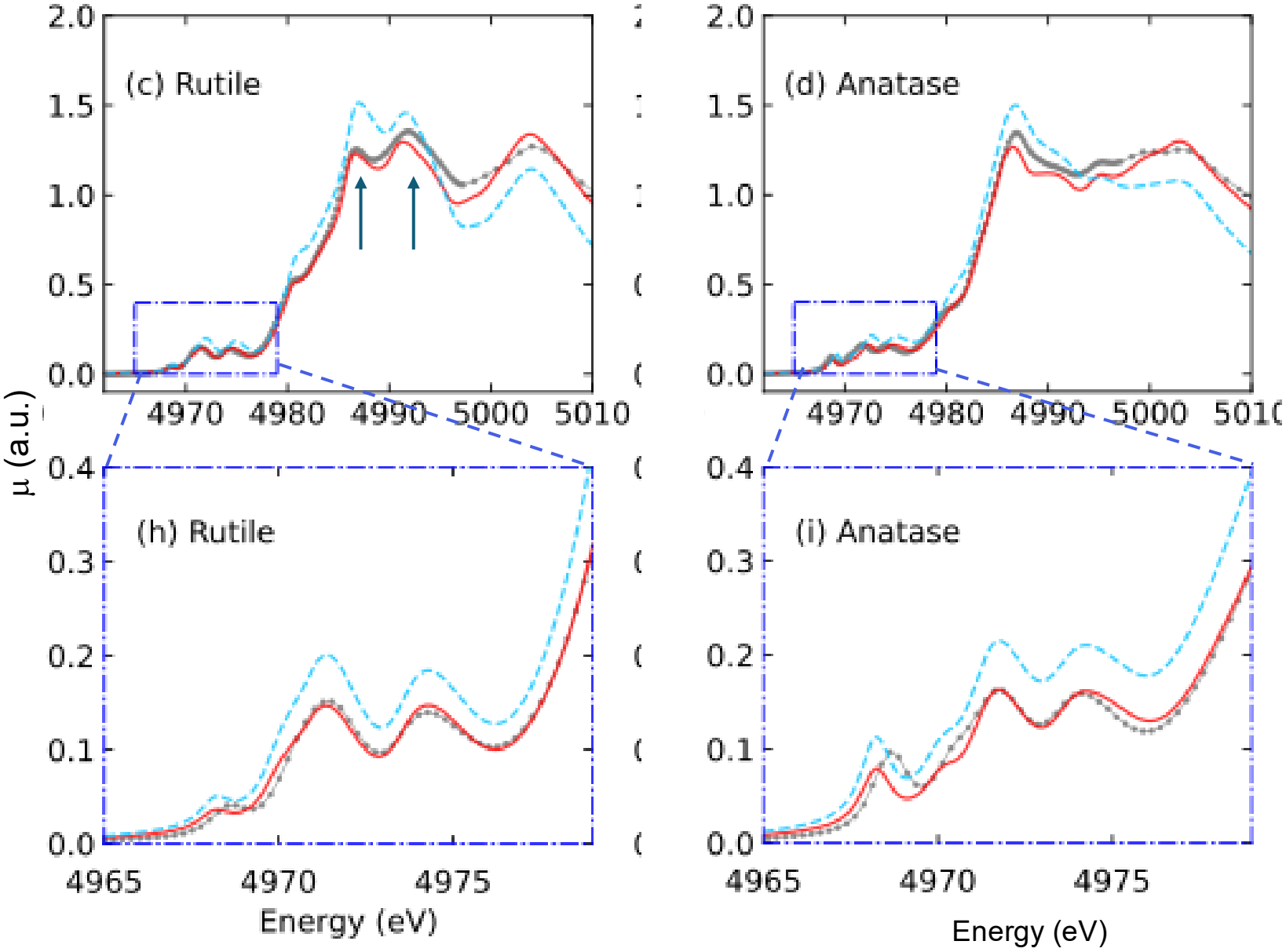
Chuntian Cao

Ti K-edge benchmark: in coll. w/ **John Vinson, Bruce Ravel** (NIST), **Josh Kas** (Univ. Washington)



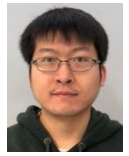
How accurate is theory?

- Shake-up satellite shifts the oscillator strength to higher energy.
- It has the opposite effect of the BSE, which pulls the oscillator strength to lower energy.
- Shake-up satellite enhances the high-energy peak in the main edge doublet.



Meng et al. and DL, *Phys. Rev. Mater.* 8, 013801 (2024); C. Cao, J. J. Kas, K. Dyro, B. Ravel, J. Vinson, and **DL**, *submitted*.

Lightshow: a Python package for streamlining XAS simulations

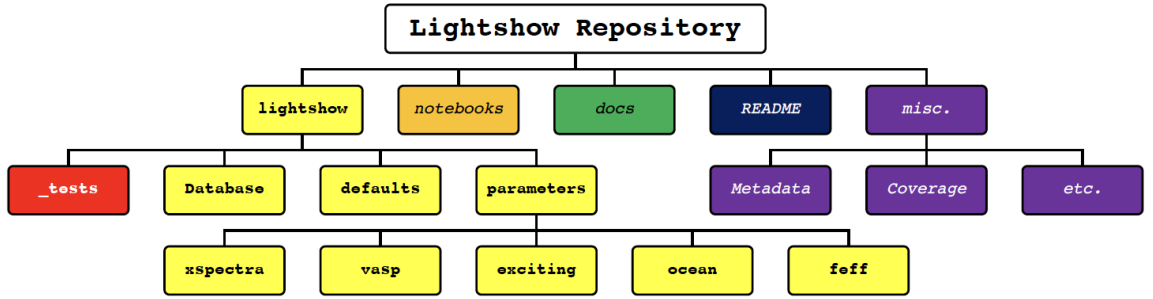
Matthew Carbone Fanchen Meng

Lightshow

Code: github.com/AI-multimodal/Lightshow
Docs: ai-multimodal.github.io/Lightshow
```
> pip install lightshow
```

- Lightshow is a Python library for easily generating computational spectroscopy input files and post analysis.





M. R. Carbone, F. Meng *et al.*: *The Journal of Open Source Software, 8, 5182 (2023).*

Spectra databases

FEFF Databases

- C, O and N K-edge QM9 molecular XANES (~1.2 M)
- 3d transition metal K-edge XANES: (Ti – Cu: ~57 K)

https://zenodo.org/records/10551358
https://zenodo.org/records/13851174)
https/doi.org/10.17172/NOMAD/2023.03.24-1
https://doi.org/10.24435/materialscloud:6z-qm
https://doi.org/10.24435/materialscloud:kn-mb
https://doi.org/10.24435/materialscloud:85-8x

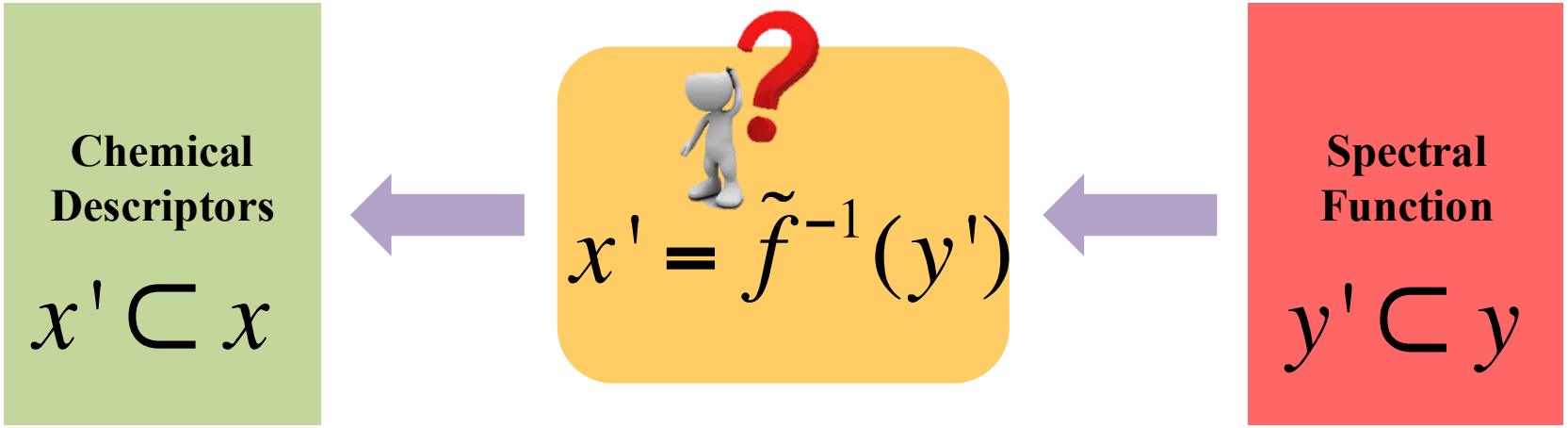
VASP Databases

- Crystalline and amorphous LPS S K-edge XANES (~2.7 K)
- Delithiated LPS P (~2 K) and S K-edge XANES (~8K)
- Ti (~4 K) and Cu K-edge XANES (~3 K)

Phys. Rev. Res., 5, 013180 (2023)
Scientific Data, 10, 349 (2023)
Cell Rep. Phys. Sci.., 5, 101909 (2024)
Phys. Rev. Mater. 9 (4), 043803(2025)

Solve local chemical environment from XAS

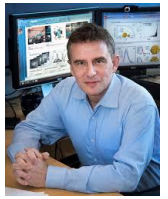
the inverse problem, often ill-conditioned



Solving 3D metal nanoparticle structure

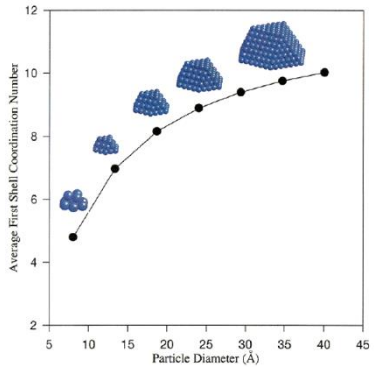


Janis Timoshenko (SBU)

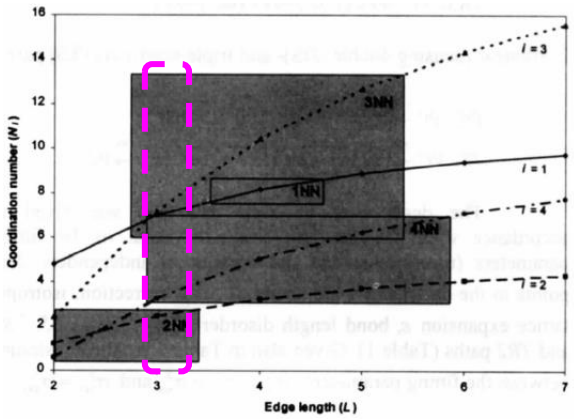


Anatoly Frenkel (SBU/BNL)

Average coordination numbers

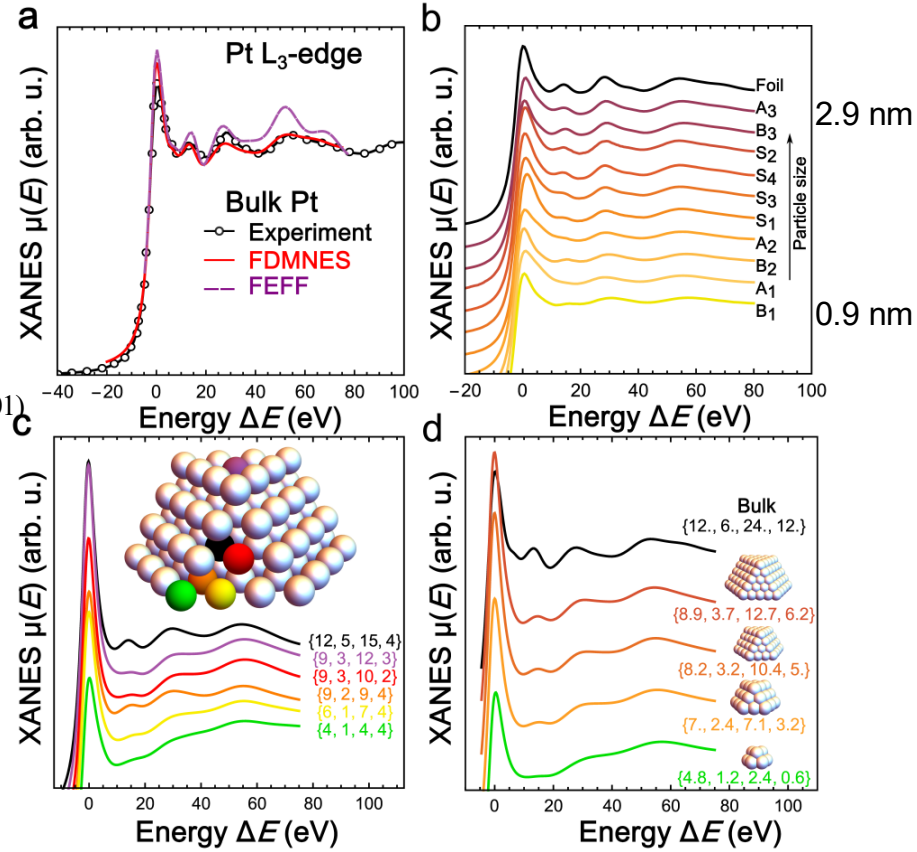


Frenkel, Hills, and Nuzzo, *J. Phys. Chem. B,* 105, 12689 (2001)

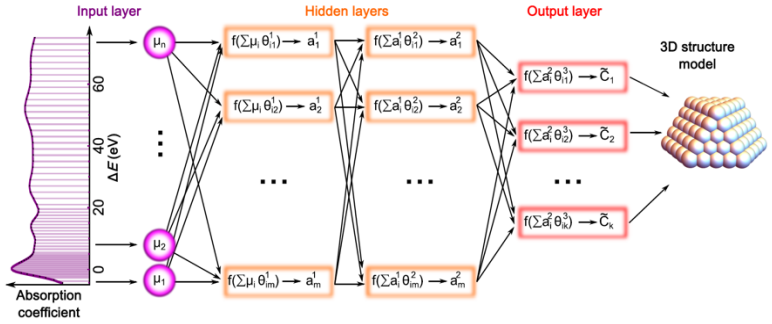


Frenkel, *J. Synch. Radiat.*, 6, 293 (1999)

XAS Spectral contrast



Supervised ML



Sample	τ_1 [a]	τ_2 [a]	τ_3 [a]	τ_4 [a]	d_{TEM} [a,b] (nm)	Model NPs CNs[c]	Model NPs size[c] (nm)
Foil	11.6(2)	5.8(2)	23(1)	11.1(8)	-	{12, 6, 24, 12}	∞
A_3	9.1(3)	4.3(3)	11(2)	8(1)	3(1)	{9.4, 4.0, 14.4, 7.1}	2.8
S_4	8.9(3)	4.2(4)	10(2)	7(1)	1.2(2)	{8.5, 3.2, 11.5, 5.0}	1.2
S_2	8.1(3)	3.7(4)	8(2)	4.5(8)	1.2(3)	{7.8, 3.3, 9.6, 4.1}	1.1
S_3	7.7(4)	3.8(4)	4(2)	3.9(9)	0.9(2)	{7.7, 3.1, 9.2, 3.8}	1.1
S_1	7.4(4)	2.0(3)	3(1)	6(1)	1.1(2)	{7.4, 2.6, 8.0, 3.3}	1.2
A_2	6.6(4)	2.3(4)	3(1)	5(1)	1.1(3)	{6.6, 2.1, 6.0, 2.9}	1.4
A_1	6.3(3)	1.5(3)	2(1)	5(1)	0.9(2)	{6.2, 1.9, 5.1, 2.4}	1.1

Timoshenko, **DL**, Lin and Frenkel, *J. Phys. Chem. Lett.,* 8, 5091 (2017)

Local coordination number of 3d TM



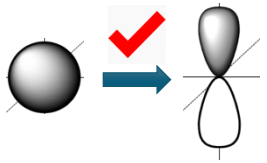


Matthew Carbone Mehmet Topsakal

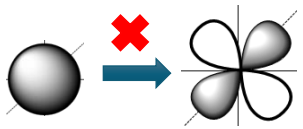
3d transition metal K-edge

Section rule: Δl=±1

with inversion symmetry

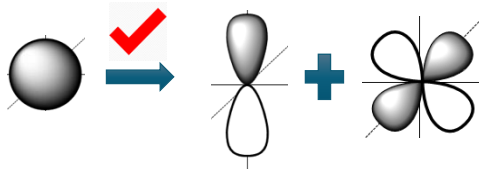


dipole allowed

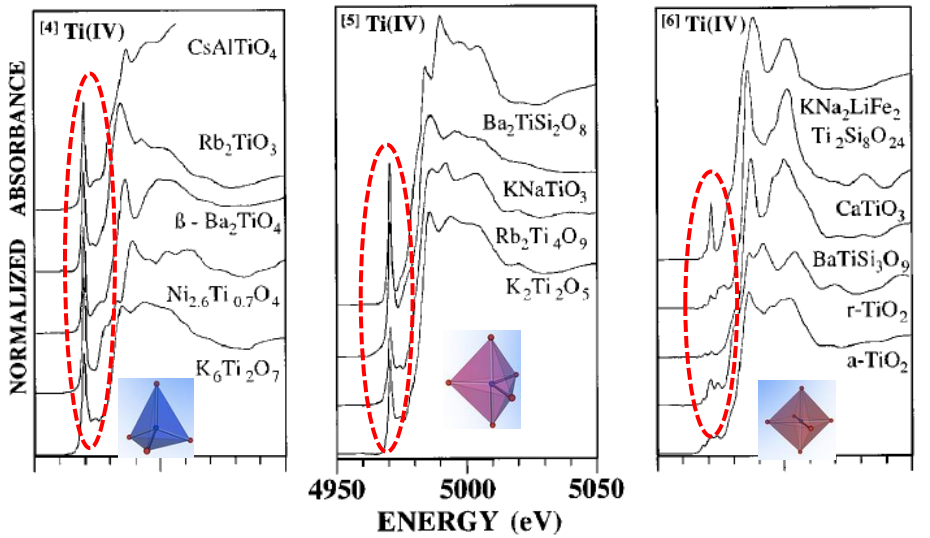


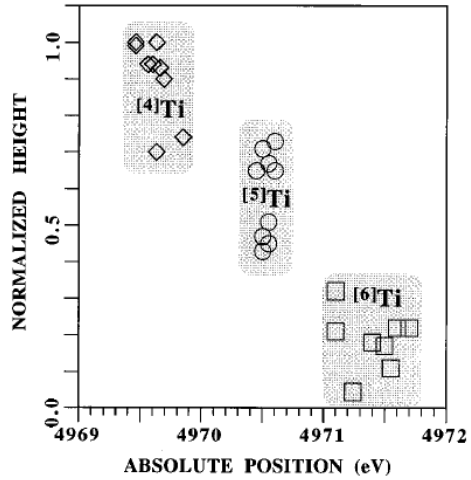
dipole forbidden

inversion symmetry breaking

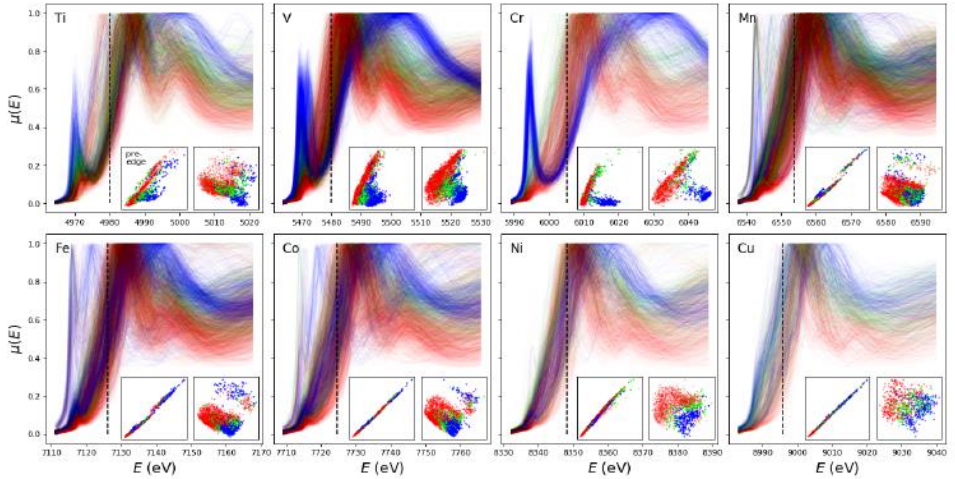


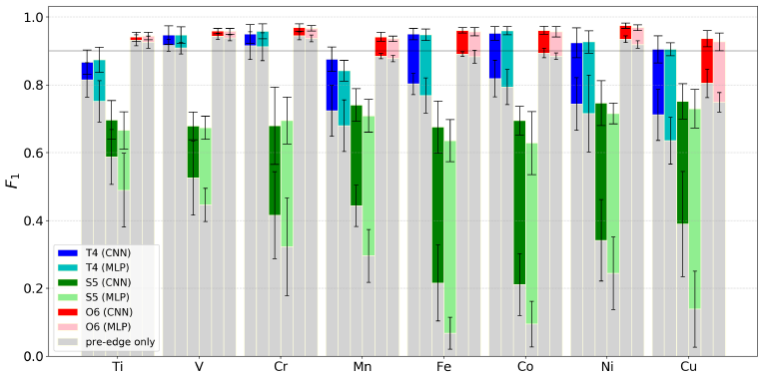
p and *d* orbitals can mix; dipole allowed





Farges, Brown and Rehr, *Phys. Rev. B,* 56 1809 (1997)

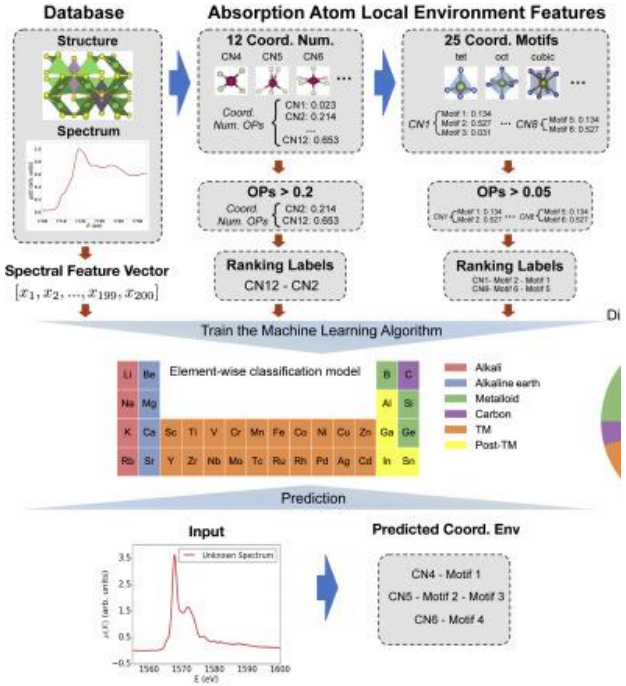




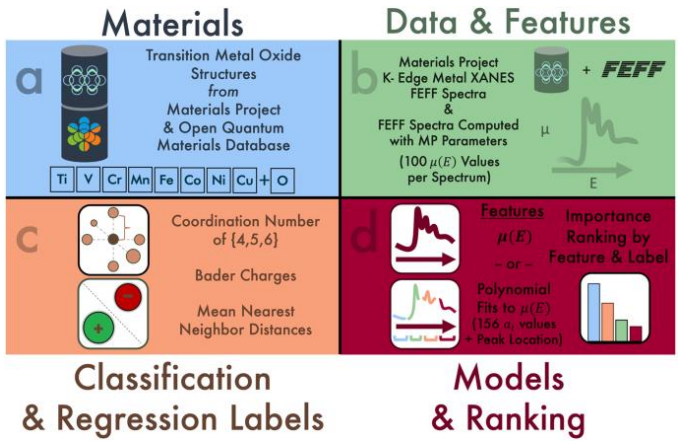
Carbone, Yoo, Topsakal and **DL**, *Phys. Rev. Mater*, 3, 033604 (2019)

Classification & regression

coord. motif

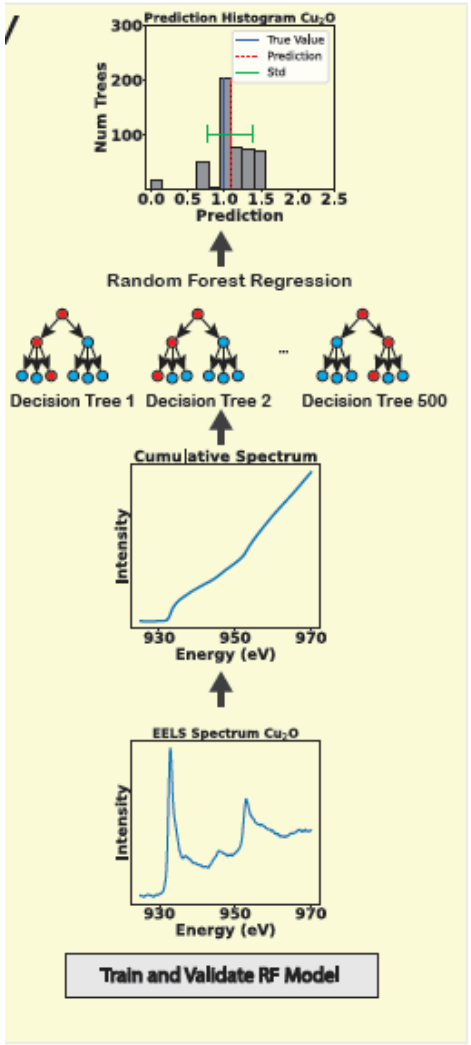


Zheng, Chen, Chen, and Ong, *Patterns*, *1*, 100013 (2020)



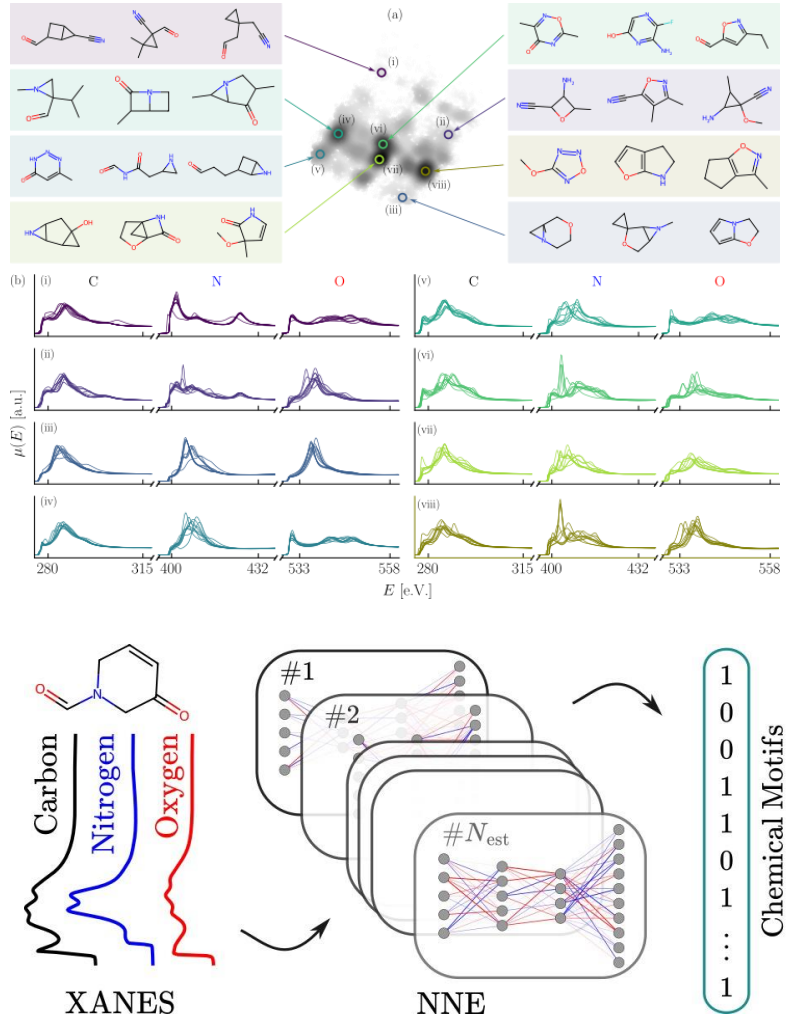
Torrisi *et al., npj Comput. Mater.*, **6**, 109 (2020)

charge state



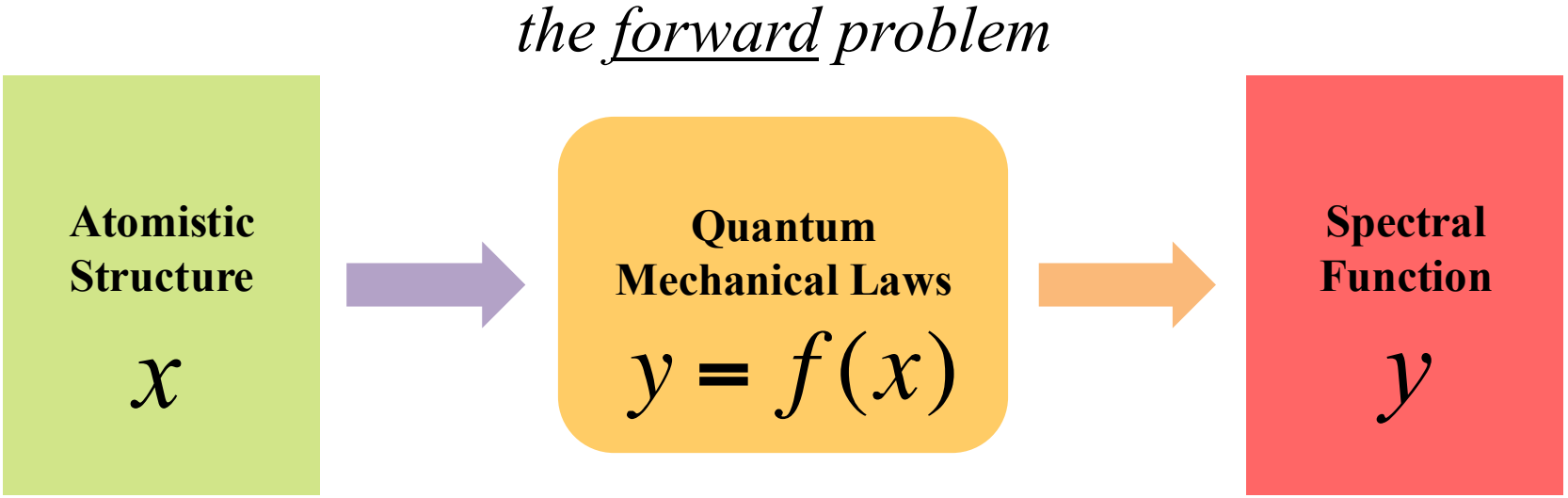
Gleason, **DL**, Ciston, *npj Comput. Mater.*, 10, 221 (2024)

chemical group

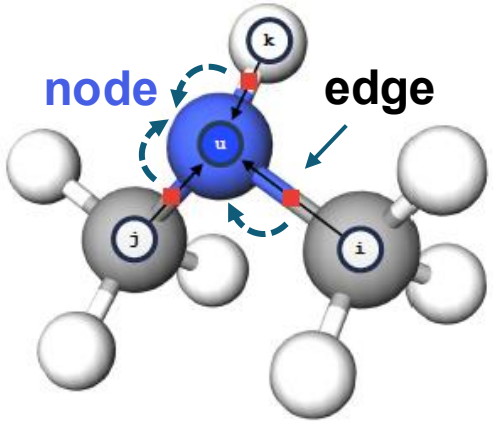


Carbone, Maffettone, Qu, Yoo and **DL**, *J. Phys. Chem. A*, 128, 1948 (2024)

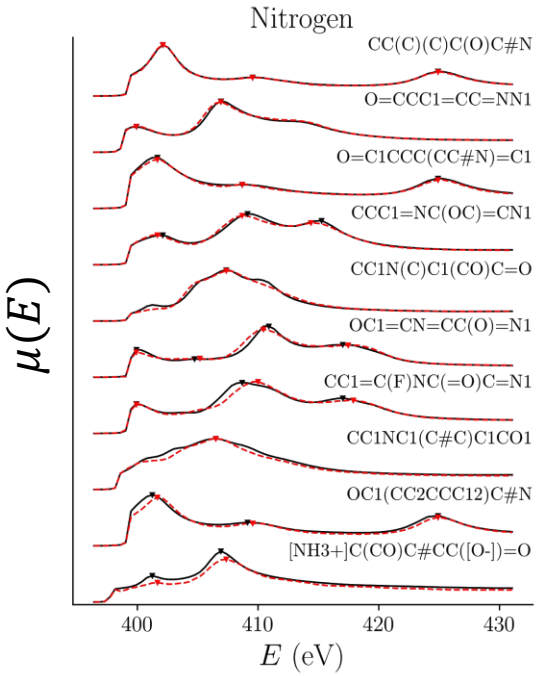
Solve local chemical environment from XAS



XAS spectral prediction

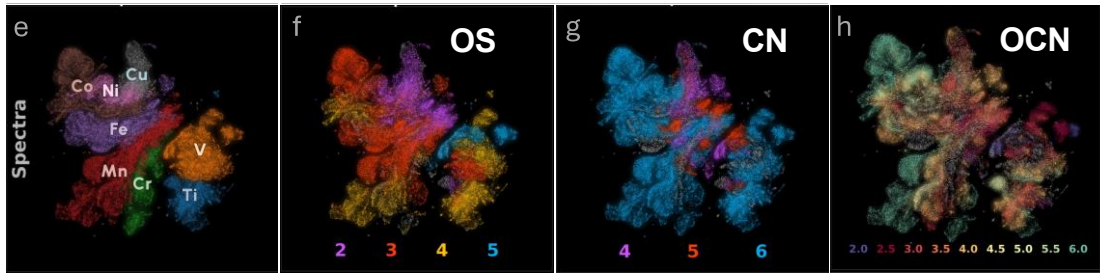


molecular graph
message passing



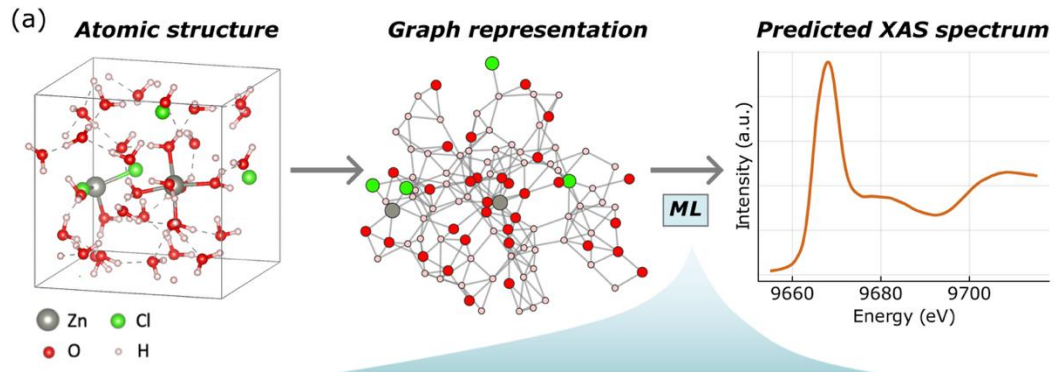
Carbone, Topsakal, **D.L.**, Yoo, *Phys. Rev. Lett.* 124, 156401 (2020)

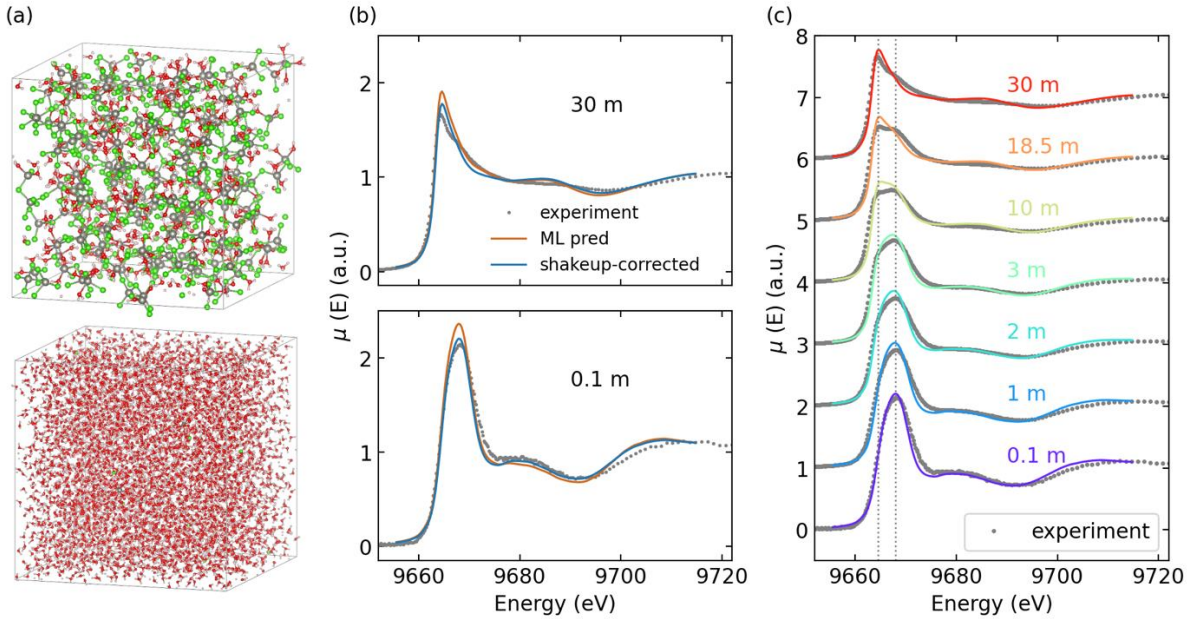
OmniXAS: 3d TM (Ti, V, Cr, Mn, Fe, Co, Ni, Cu)



Kharel, Meng, Qu, Carbone, **DL**, *Phys. Rev. Mater.*, 9, 043803 (2025)

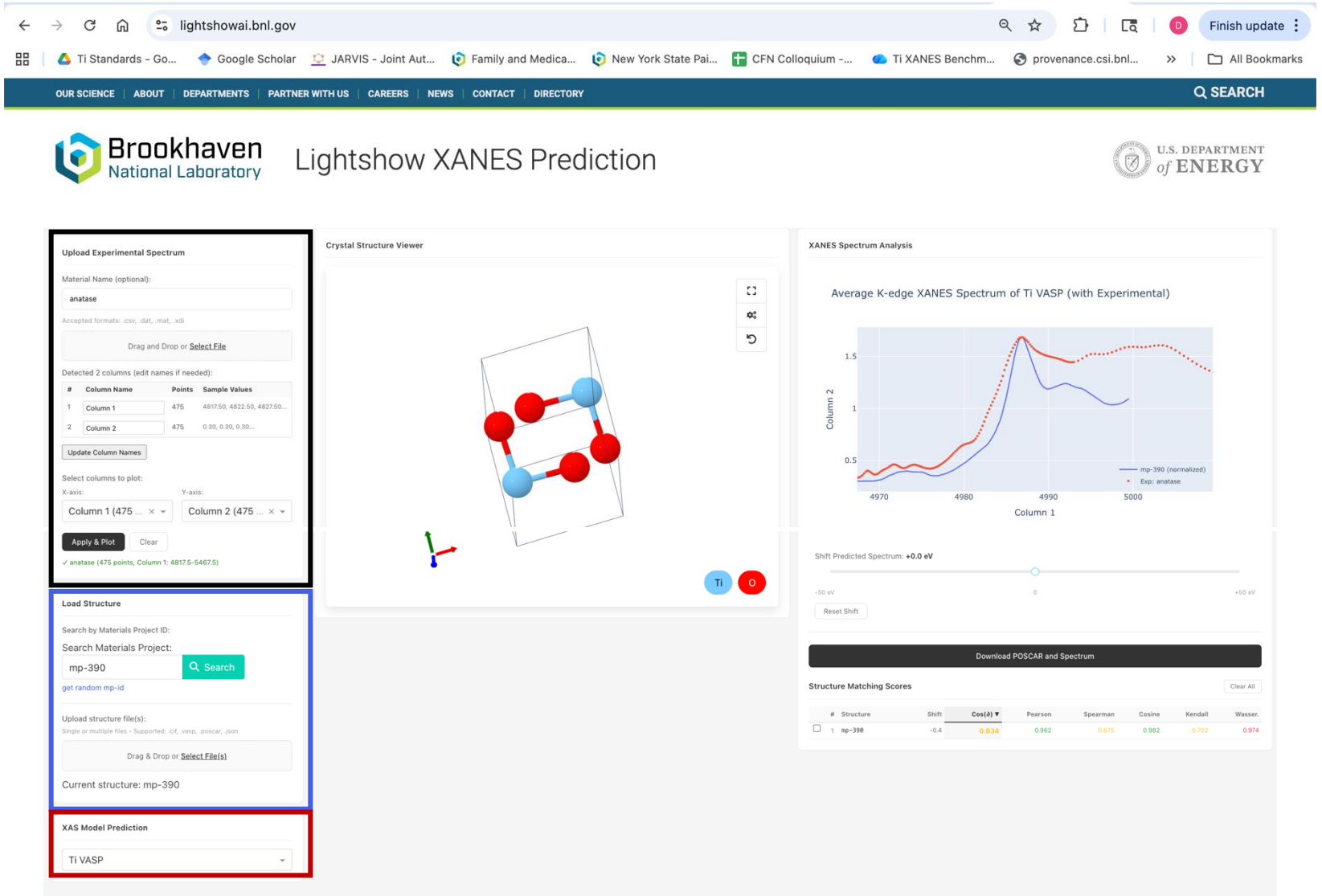
Zn K-edge in $ZnCl_2$ solution





Cao, Li, Campos, Pace, Kas, Wu, Ma, Yang, Xu, Yoo,. Takeuchi, Takeuchi, Yan, Marschilok, **DL**, *JPCB* 130, 5084 (2026)

lightshowai.bnl.gov



- ✓ Cloud-based end-to-end XAS analysis
- ✓ Spectral prediction with GNN models
- ✓ Structure model visualization and atom-site specificity
- ✓ ML model vs exp spectral matching
- ✓ Real-time data streaming at light source end station
- ✓ Many more …

LightshowAI agentic chatbot



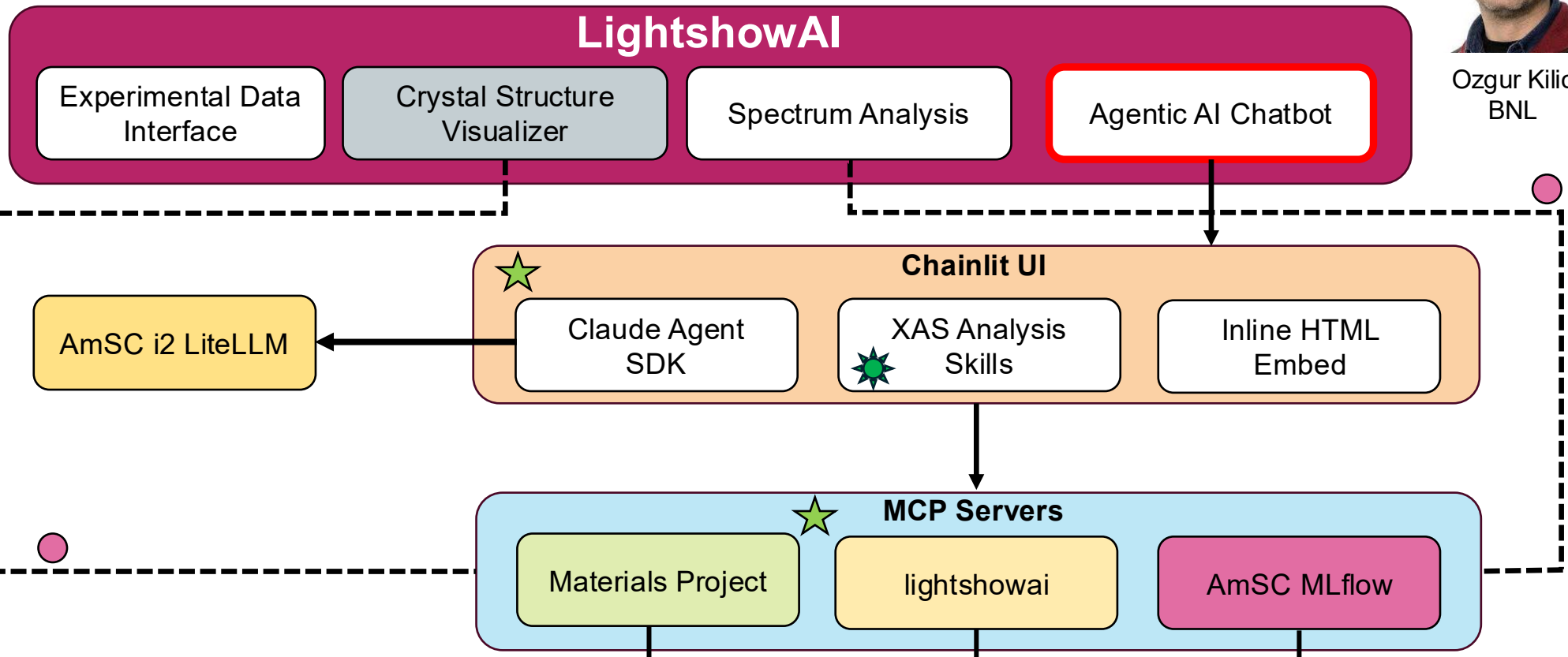
Huihuo Zheng
ANL



Ozgur Kilic
BNL



Sairam Sri
Vatsavai
BNL



Huihuo Zheng, AmSC-ANL

Ozgur Kilic, Sairam Sri Vatsavai, BNL

BNL+AmSC

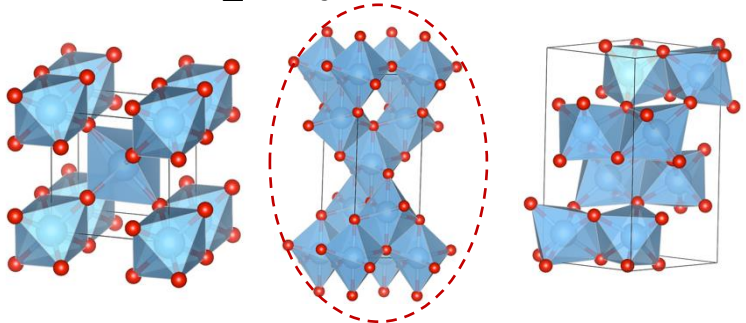
AmSC Model Services

AmSC AI Services (LiteLLM Gateway, Claude Subscription)

Materials Project (LBNL)

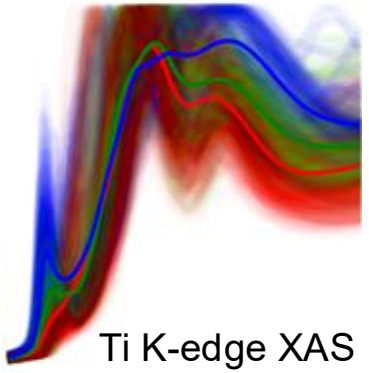
Material structure identification with AI-assisted real-time analysis in experiment

TiO_2 crystalline phases



Rutile Anatase Brookite

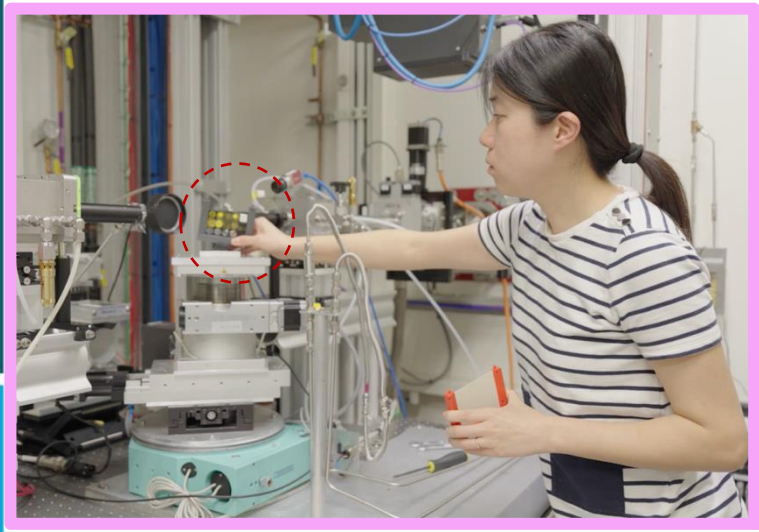
- Properties are strongly phase-dependent
- Rapid structure identification is important

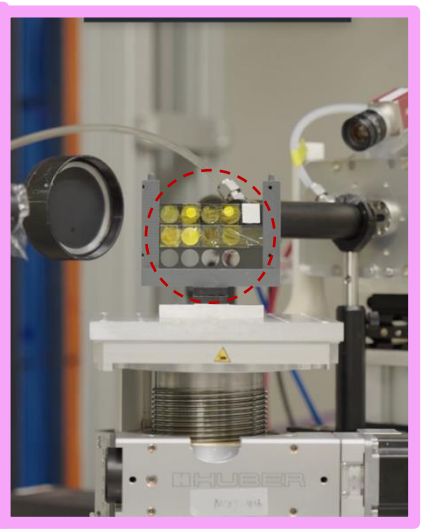


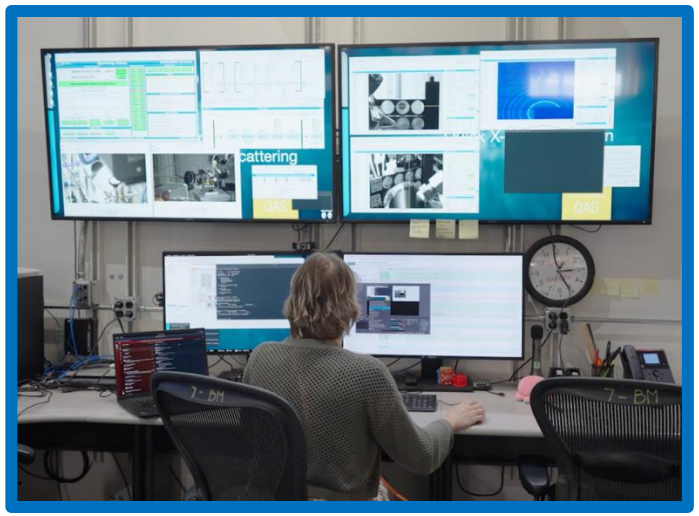
Ti K-edge XAS

- Rich spectral features
- Fingerprints of the structure

beamline measurement → **data collection** → **AI-assisted analysis**



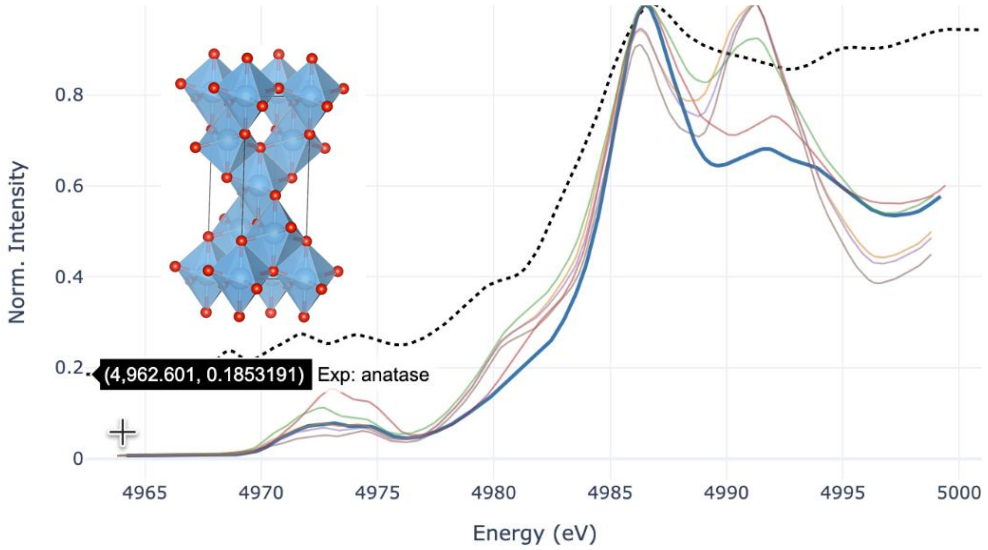




Determine the TiO_2 crystalline phase from its Ti K-edge XAS in real time using *LightshowAI*

LightshowAI agent: domain-expertise-informed scientific reasoning

Please find the best structure what match the experimental spectrum here, using VASP model



XANES analysis skills

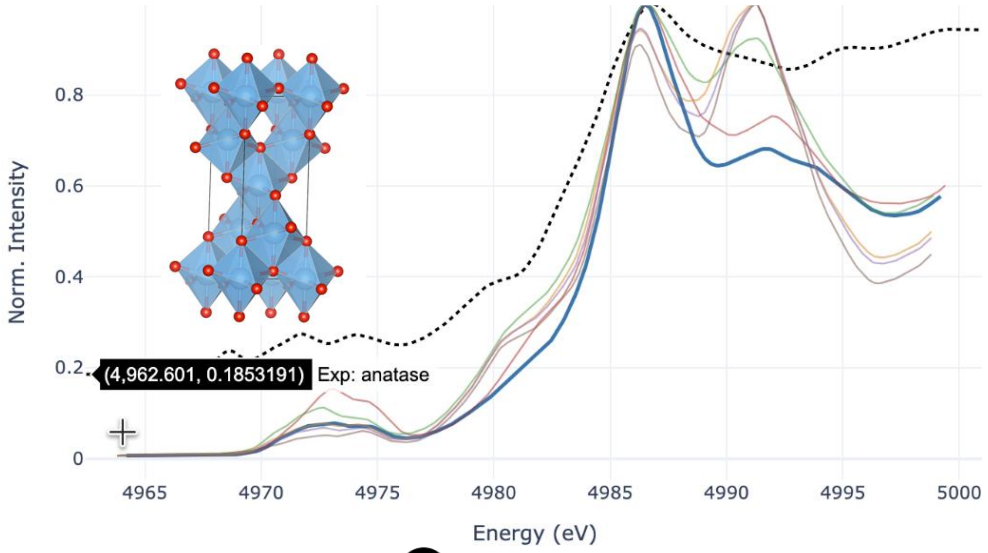
Run an end-to-end Ti K-edge XANES benchmarking workflow.

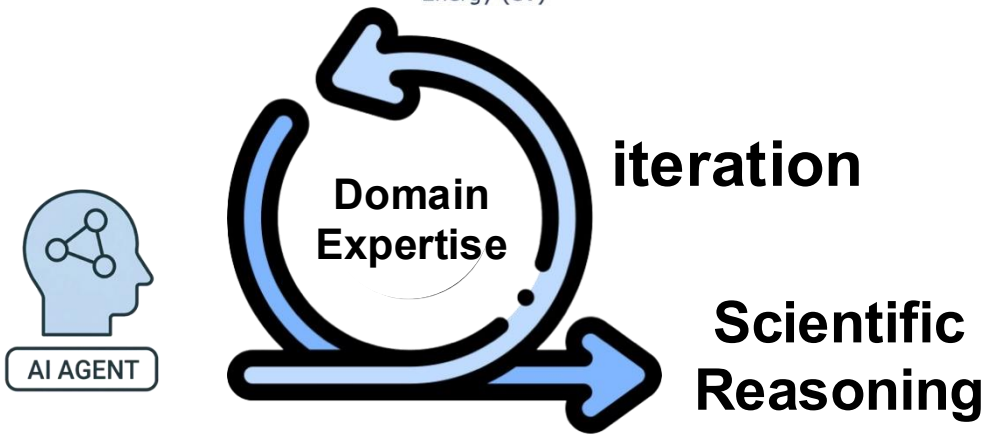
- ✓ Retrieve candidate structures from the Materials Project
- ✓ Predict spectra with *LightshowAI* (VASP or FEFF model)
- ✓ Compare to experimental .dat standards using shift-optimized Pearson r, Spearman ρ, Cos(∂), Cosine, and Kendall τ
- ✓ Rank structures per experimental standard and generate an interactive comparison figure.

https://github.com/AI-multimodal/LightshowAI/tree/agentic

LightshowAI agent: domain-expertise-informed scientific reasoning

Please find the best structure what match the experimental spectrum here, using VASP model





"Following expert feedback from Deyu (scientist), the agent updated the analysis from RMSD-only to include Pearson r and Spearman ρ with per-structure energy shift optimization…"

Results vs. `anatase.dat` (sorted by Pearson r)

#	MP-ID	Phase	Shift	Cos(∂)	Pearson	Spearman	Co
1	**mp-390**	**Anatase**	**−0.3**	—	**0.9714**	—	—
2	mp-554278	Monoclinic	—	—	0.9681	~best	—
3	mp-1439	TiO_2-II	—	—	0.9597	—	—
4–6	mp-1840/775938/2657	Brookite/ Pbcn-2/ Rutile	—	—	lower	—	—

⚠️ **Near-tie warning**: Anatase and Monoclinic (C2/m) are within Δ Spearman ≈ 0.0013 — well inside the noise band. Cos(∂) is the recommended tiebreaker here since it is most sensitive to the pre-edge doublet shape unique to anatase.

Summary & outlook

- ML models can already successfully extract structural / chemical descriptors in materials in proof-of-concept.
- Practical applications on experimental data in the diverse material space require high-quality training data and accurate theory.
- Agentic AI and next-generation ML models can bring XAS analysis near the instrument, which leads to a tremendous speed-up and the new capability of dynamic experiment control.
- Future user facilities may operate under a new paradigm by providing AI tools as a service.



Acknowledgment

BNL

Zhu Liang
Fanchen Meng
Wei Chen
Xuance Jiang
Ruoshui Li
Sencer Selcuk
Mehmet Topsakal
Xiaohui Qu
Mingzhao Liu
Mark Hybertsen
Sara Mason
Dario Stacchiola

Brian Lee
Judith Yang
Matthew R. Carbone
Chuntian Cao
Shubha Kharel
Shinjae Yoo
Yuewei Lin

BNL/SBU
Anatoly Frenkel
Amy C Marschilok
Shan Yan
Esther Takeuchi
Ken Takeuchi

Eli Stavitski
Seongmin Bak
Yonghua Du
Lu Ma
Sizhan Liu
Ke Chen
Sairam SriVatsavai
Christopher Kelly
Ozgur Kilic
Daniel Allan
Stuart Wilkins
Kari Barry
Yevgen Matviychuk

Columbia University
Jagriti Singh
Shekhawat
Cem Komurcuoglu
Haoyue Guo
Kerry Sun
Alexander Urban
Dan Steingart

DOE CSGF
Nina Cao (MIT)
Pavan Ravindra (Columbia Univ.)

Student Assistant
Anton Yanaky (SBU)

ANL
Feng Wang
Huihuo Zheng

Univ. Washington
Joshua Kas

NIST
John Vinson
Bruce Ravel
Conan Weiland

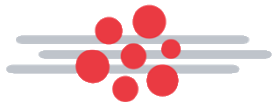
Humboldt-Universität zu Berlin
Claudia Draxl
Benedikt Maurer
Fabian Peschel

Univ. Chicago
Christian Vorwerk

Utrecht University
Nongnuch Artrith

LBL
Jim Ciston
David Prendergast

LLNL
Tuah Anh Pham
Tim Hsu
Hyuna Kwon



Brookhaven National Laboratory

U.S. DEPARTMENT OF ENERGY

Genesis Mission | AMERICAN SCIENCE CLOUD

AIMM
Artificial Intelligence for MultiModal Analysis