

■ National Synchrotron Light Source II

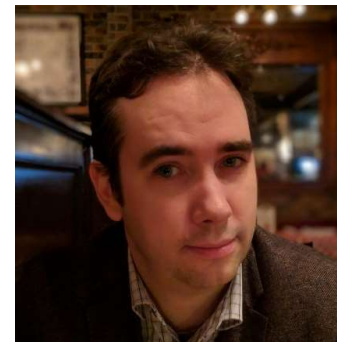
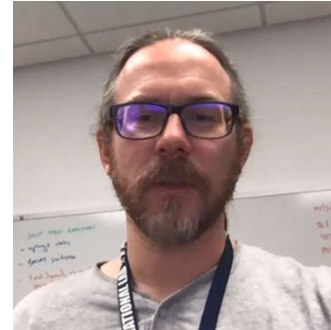
Learning to Play the Game of Science

Dr. Daniel Olds
PDF Beamline Scientist
National Synchrotron Light Source II

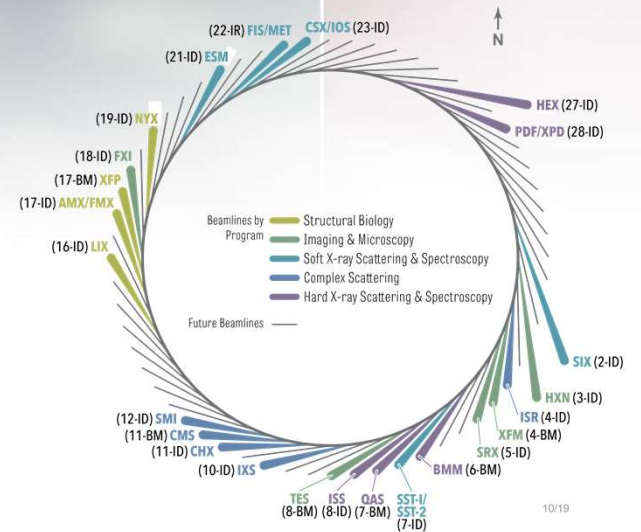
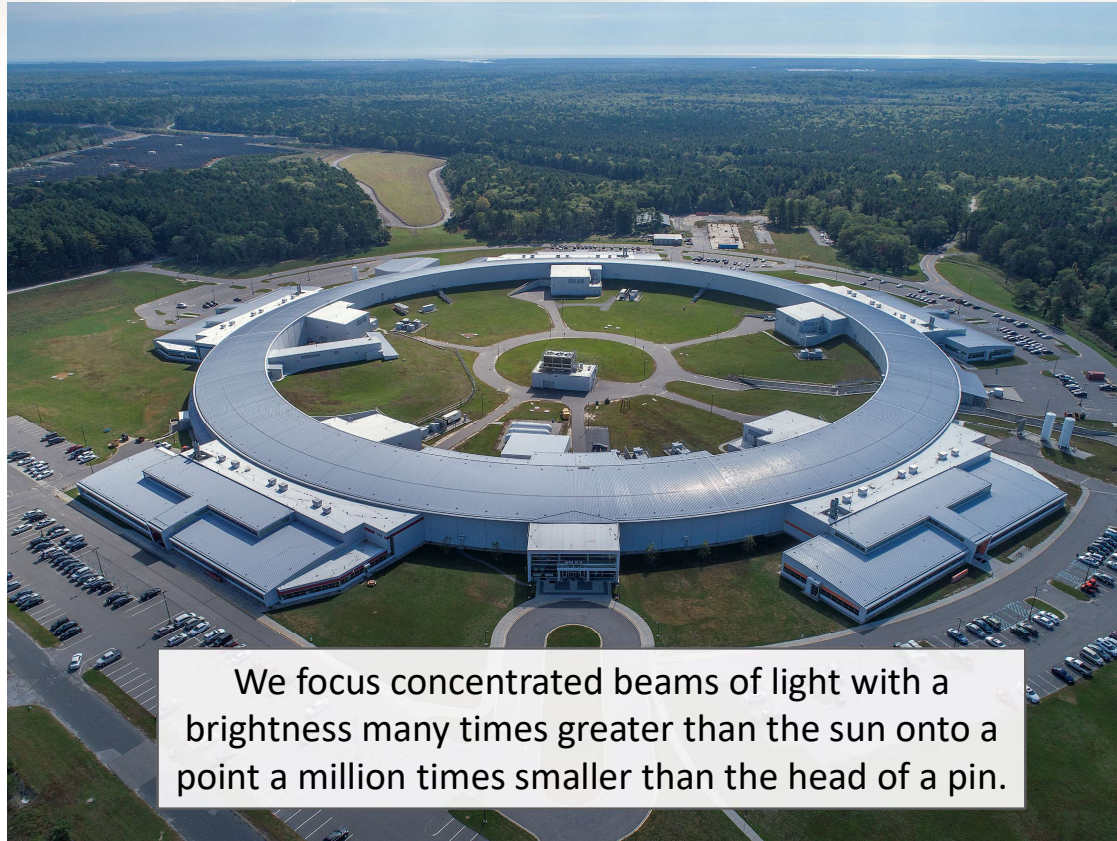


Collaborators

- Dr. Stuart Campbell
- Ms. Clara Cook
- Mr. Josh Lynch
- Dr. Phillip Maffettone
- Dr. Tom Caswell

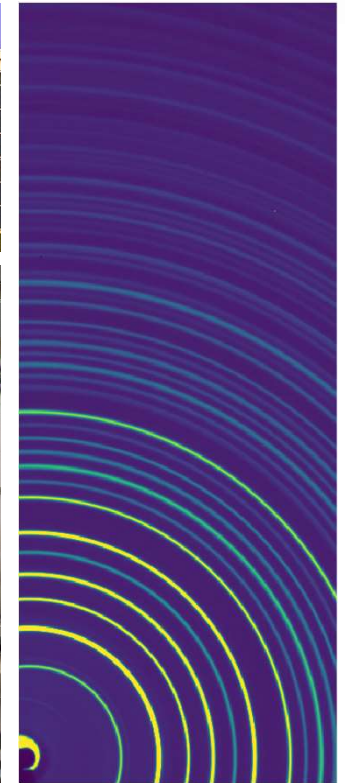
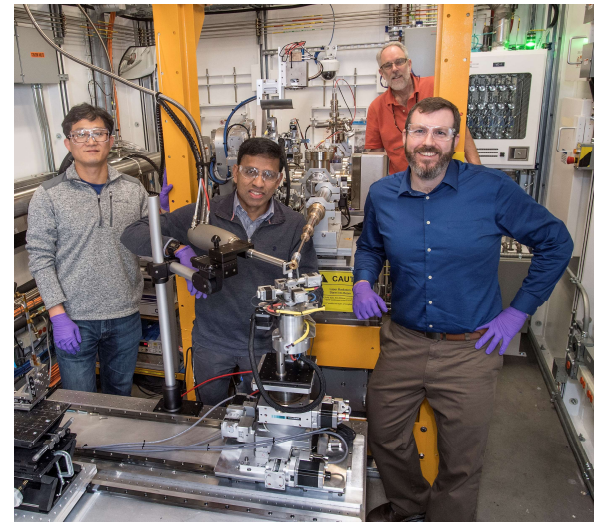


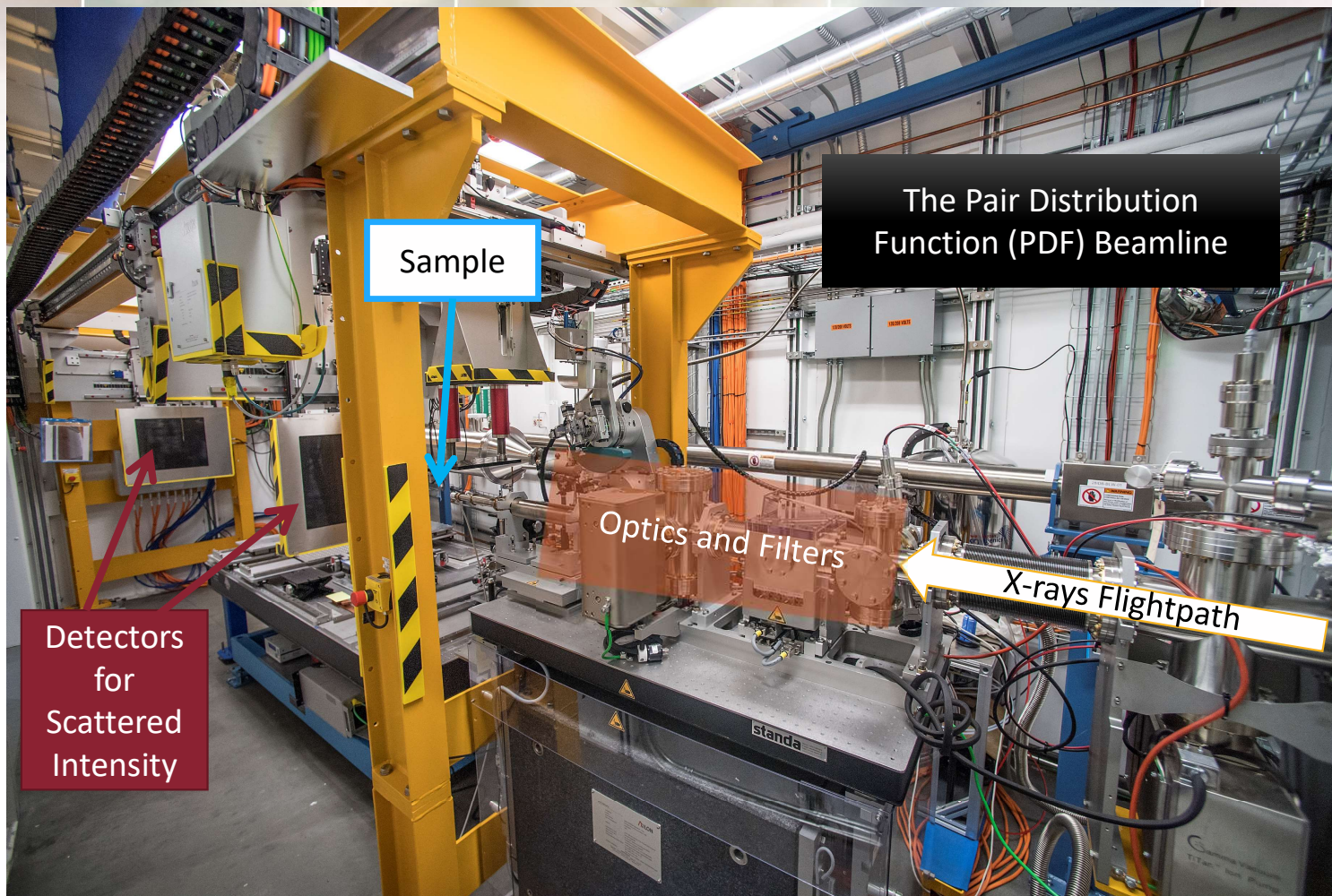
NSLS-II : The brightest around



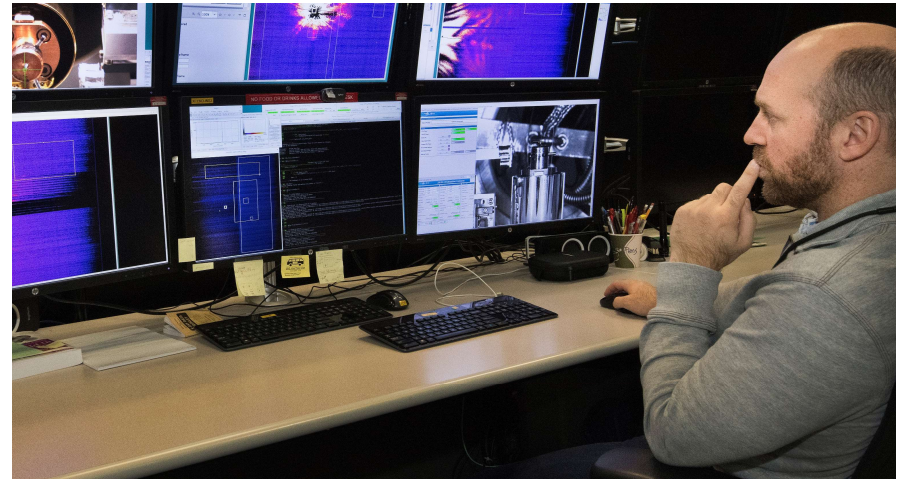
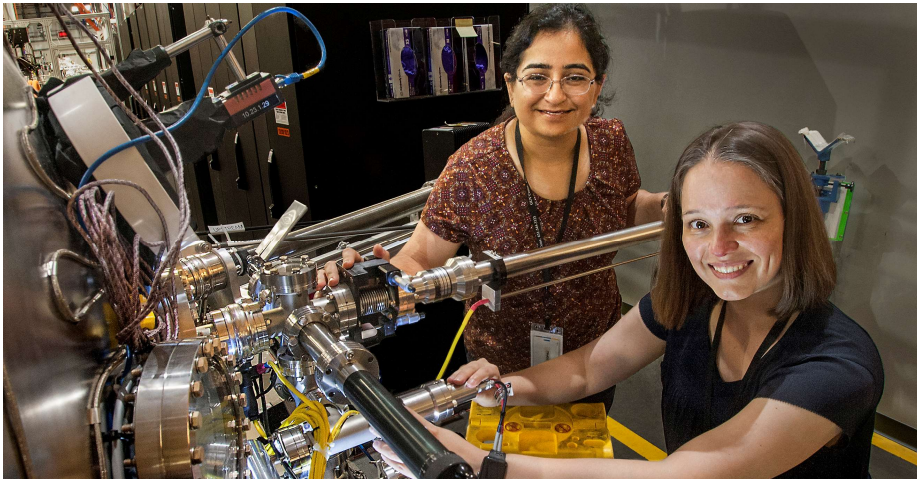
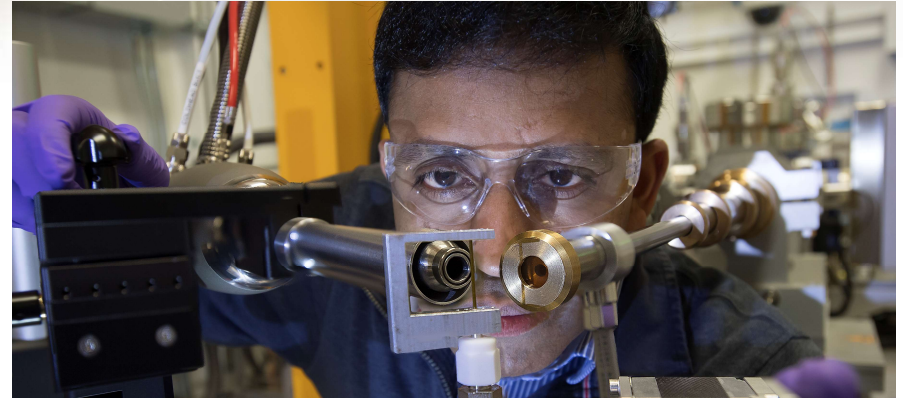
What is a Beamline?

- NSLS-II generates the widest spectrum of light in the DOE complex – from infrared to high-energy X-rays.
- Beamlines use this light to reveal the structure and function of things - each specialized for a specific technique.
- Who can use a beamline? – **Anyone.** Beamtime is awarded based on general user program, proposals evaluated by external review panels.
- Each beamline is staffed by scientists, engineers, and support staff facilitate and assist users in realizing their goals.

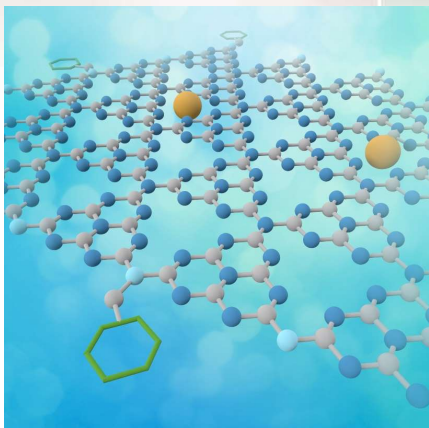




Busy, busy, busy beamlines.

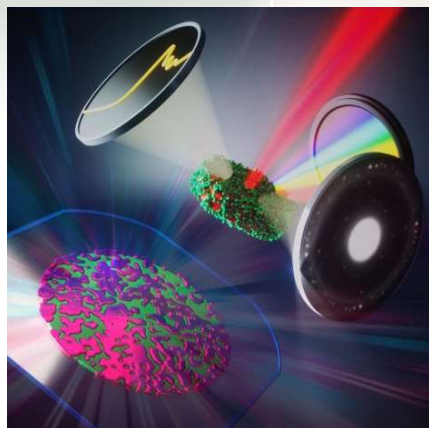


Science!



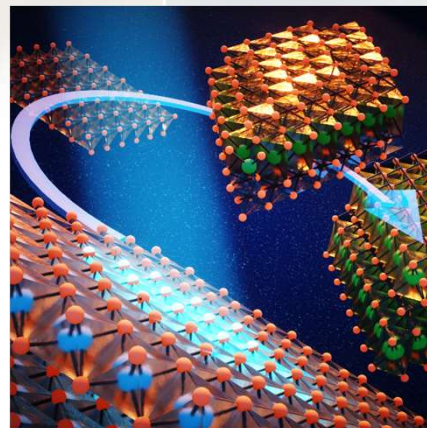
Higher efficacy co-catalysts on a nanosheet for environmentally friendly water purification.

C. Chu, et. al., PNAS 117 (12) 6376 (2020).



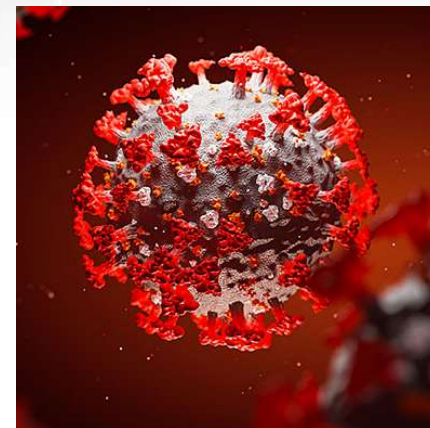
Solid-state interfacial dealloying for applications in catalysis, sensing, and energy storage.

C. Zhao, et. al. Mater. Horiz., 6, 1991 (2019)



Understanding the reaction mechanisms in the electrodes of Na-TiS₂ batteries.

C.-H. Lin, et. al. J. Mater. Chem. A, 8, 12339 (2020).

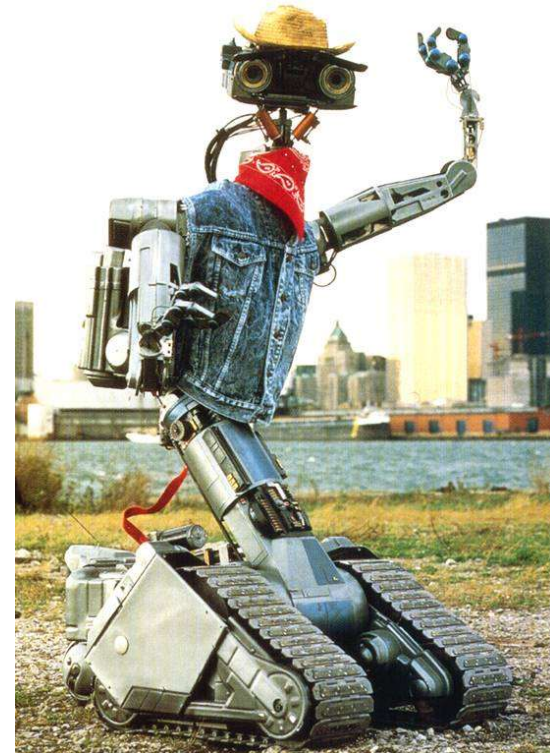


Targeting a critical molecular switch in COVID-19 to inhibit replication pathways.

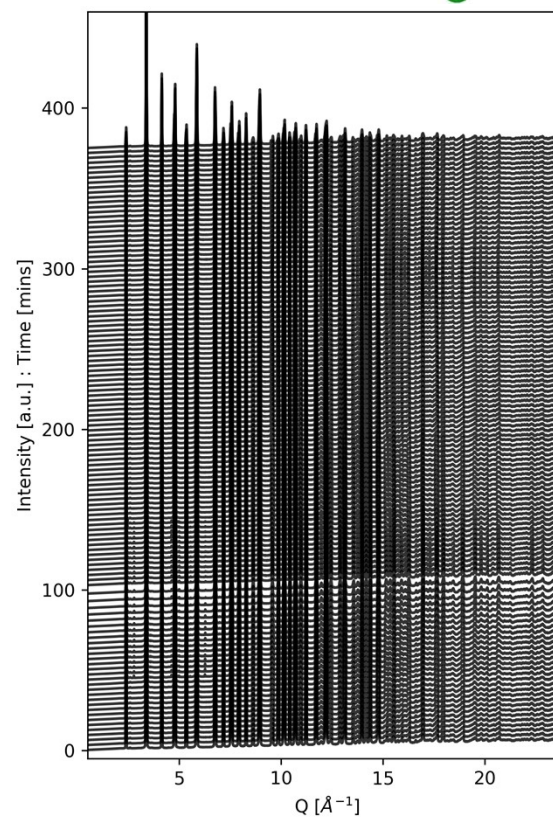
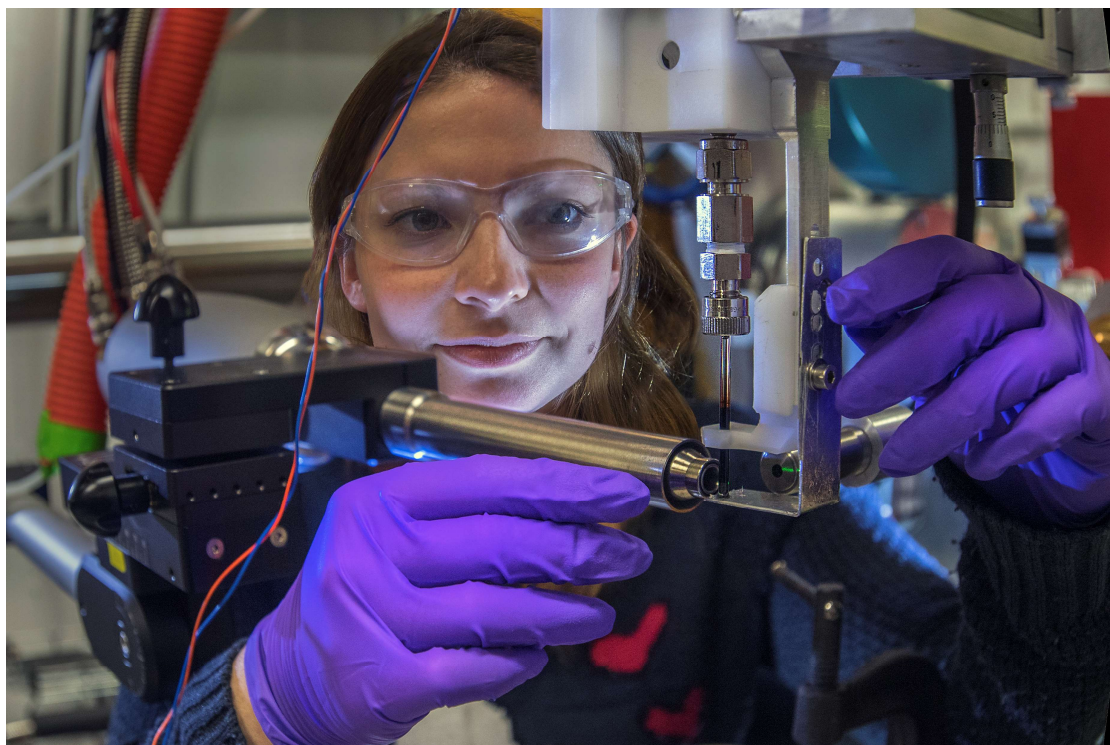
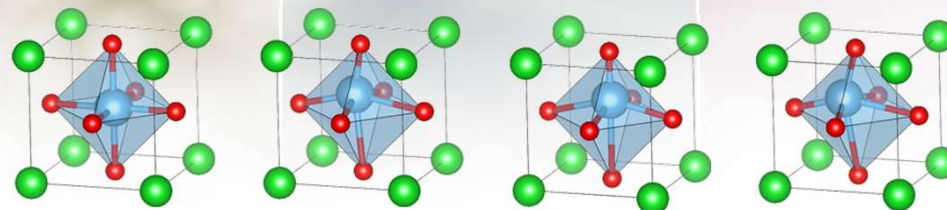
J. A. Kelly, et. al., J. Biol. Chem. jbc.AC120.013449

Users of the NSLS-II utilize our coherent, intense synchrotron light to perform materials discovery. Our portfolio of scientific impact includes: **Energy conversion and storage**, **Complex and nanomaterials**, **Quantum information science**, **Life and environmental science**, and **Physical sciences and chemistry**.

What does the coming AI-revolution mean for our users?



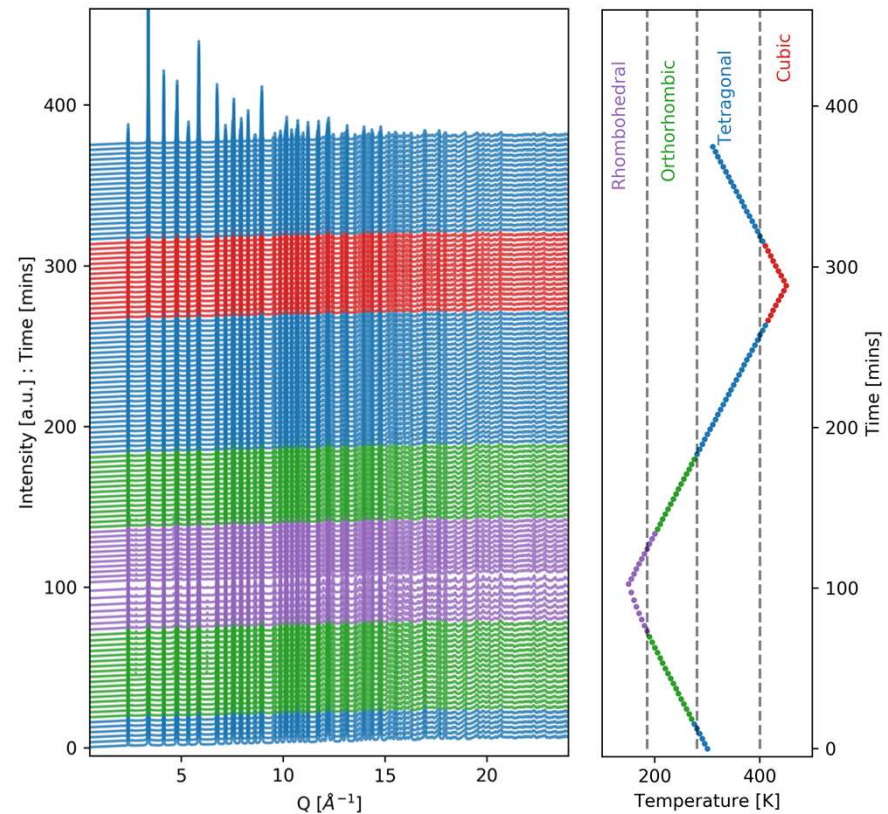
In situ Sample Heating



Unsupervised Learning for Data Comprehension

- Simple clustering algorithms can help users understand data as it's measured
- New users may not have expertise to quickly understand data themselves
- Why wait to get your data home to begin your analysis and understanding?
- Tools for at-the-beamline analysis don't need to be fancy, just helpful
- More efficient use of beamtime leads to accelerated materials discovery

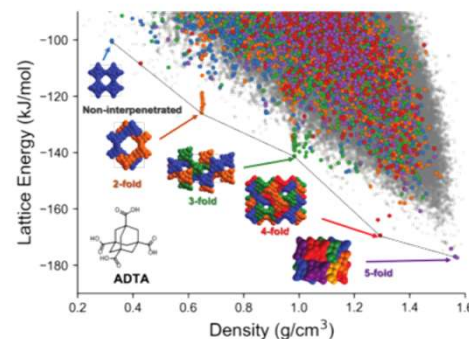
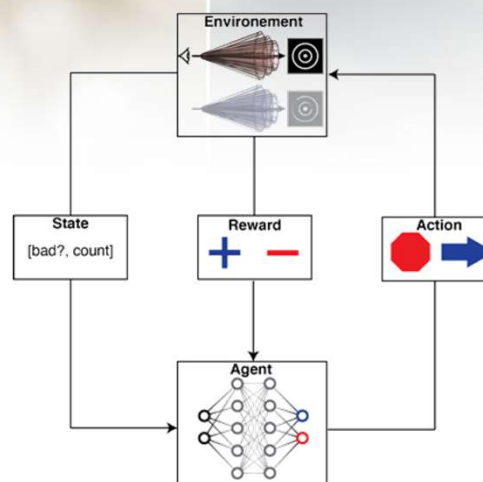
Campbell, S., *et. al.*, "Outlook for Artificial Intelligence and Machine Learning at the NSLS-II", *MLST*. 2 (1), 013001 (2021)



What AI and automation are bringing to the NSLS-II

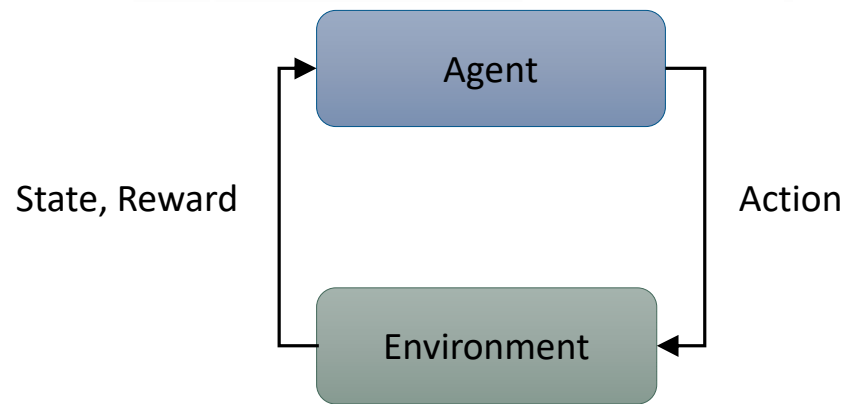
- Automation of data collection process straightforward in Bluesky
- Adaptive *in situ* studies focusing on features-of-interest during scans
- Methods to provide critical insights to researchers' mid-experiment
- Accelerate analysis to match speeds of data collection at modern beamlines
- Efficient use of beamtime via 'gamification' of beamtime decision logic with Reinforcement Learning

Maffettone, P.M., *et al.* "Crystallography companion agent for high-throughput materials discovery." *Nat. Comput. Sci.* **1**, 290-297 (2021)



Maffettone, P.M., *et al.* "Gaming the Beamlines - Employing Reinforcement Learning to Maximize Scientific Outcomes at Large-Scale User Facilities" *MLST 2*, 025025 (2021).

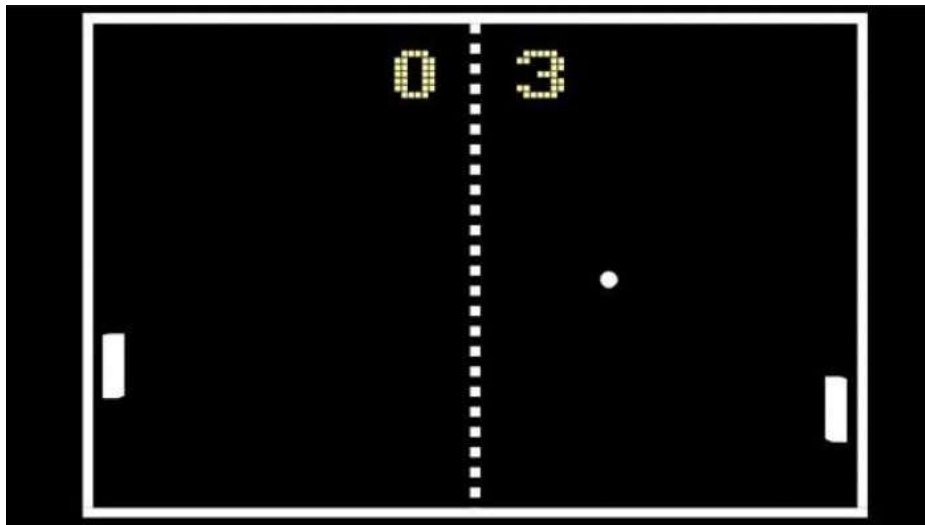
Reinforcement Learning



Goal is for Agent to develop an optimal policy that maximizes rewards as it interacts with environment. Details of the methods are a much larger talk.



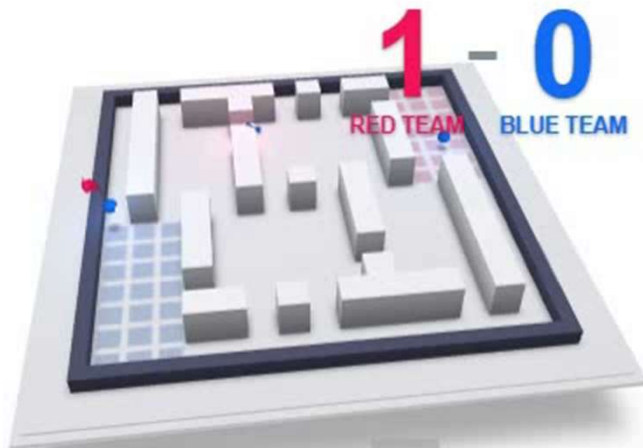
Pong



- Agent has three actions (move up / move down / don't move)
- +1 points for winning a round
- -1 points for losing a round
- Game ends at fixed number of rounds (21?)

- Good policy – Don't let opponent score (track ball)
- Better policy – Score on opponent while defending own goal!

Capture the Flag



- Reinforcement learning applied to game of capture the flag played on a computer.
- Advanced implementation with many agents training in parallel, competing and evolving to maximize rewards (win games).
- Agents demonstrated learning several novel and successful tactics over time.

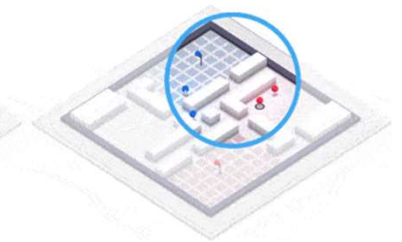
Home Base Defence



Opponent Base Camping



Teammate Following



<https://deepmind.com/blog/article/capture-the-flag-science>

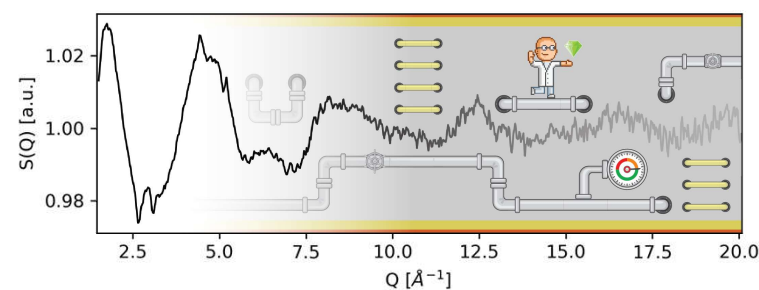
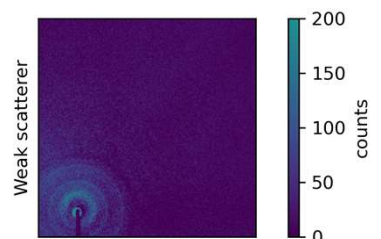
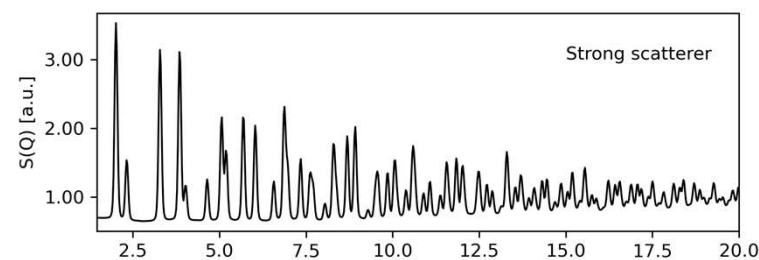
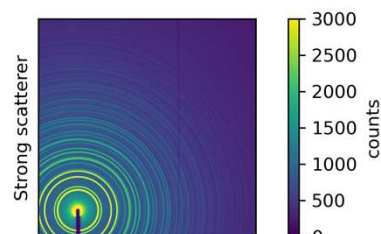
Jaderberg, M., *et. al.*, *Science* 364(6443) 859-865

Gamification of Beamline Science

- Many tasks we perform in daily 'science' activities can be described in the language of games in context of maximizing output.
 - +1 for papers published
 - -1 for breaking sample
- RL offers language to convert those rules and objectives into a domain that AI understands.
- Can we use what we've learned from AI developments (and years of study on *Super Nintendo*) to gamify our research goals and maximize scientific productivity on the beamlines?

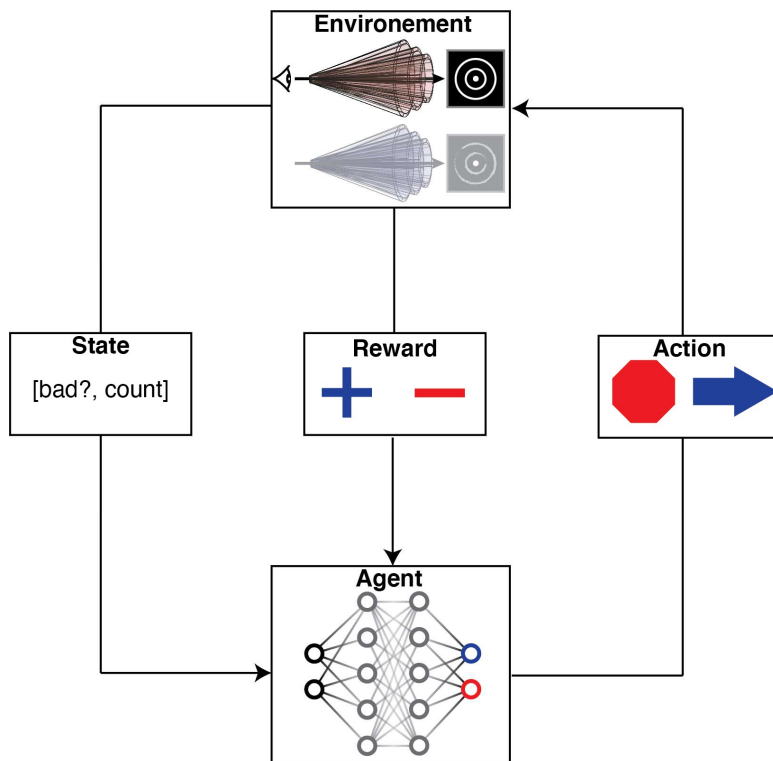


How to optimize data collection?



Step 1. You need to define the game!

BadSeed

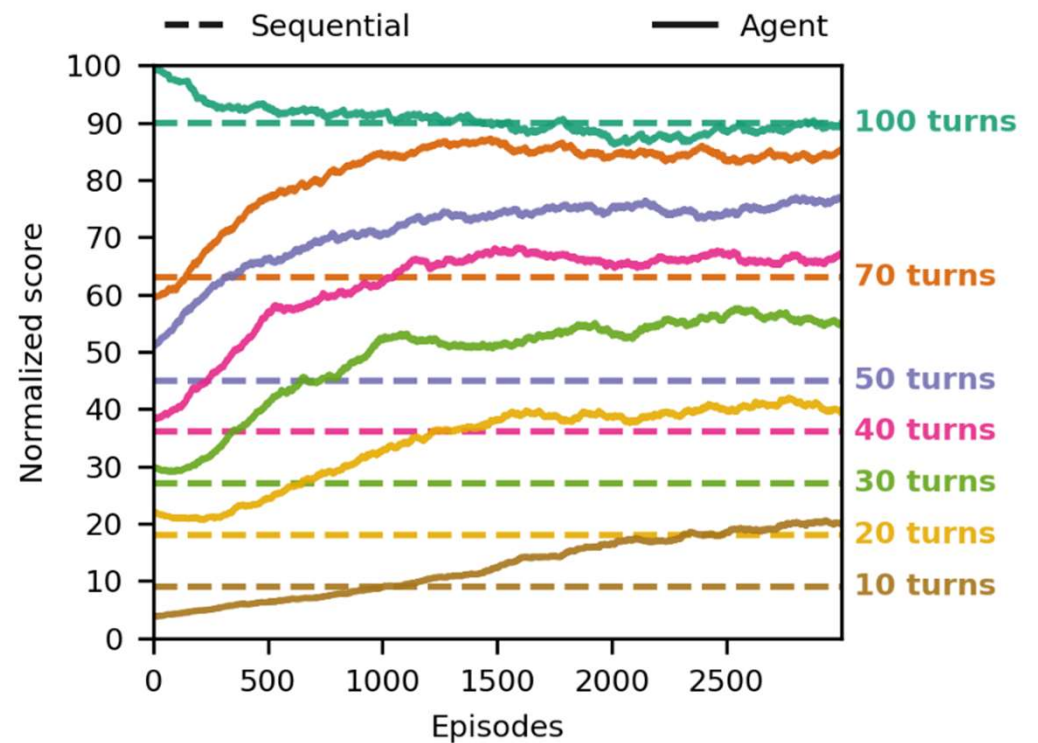


N samples, with m 'bad' scatterers
Bad scatters require '10' times more measurements
Beamline can advance to next sample, or stay and measure
External heuristic is assumed to identify 'bad' data

+1 point every measurement on bad-sample up to 10
+0 point otherwise
Game can end at anytime (end of beamtime, beamdump)
'Won' if maximum points achieved ($m \times 10$) points

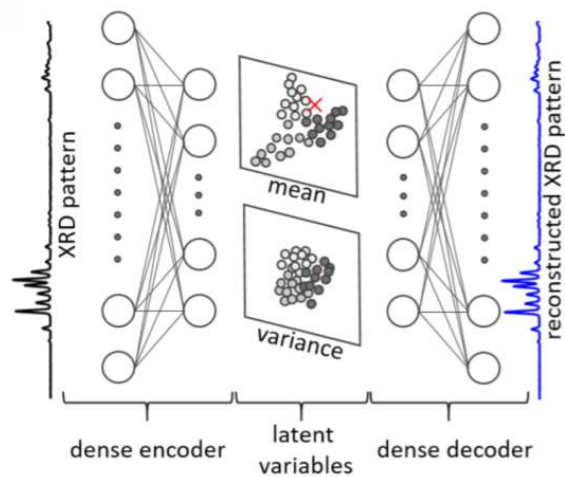
Avg Score of Agent vs. Sequential Plan

- Compare to 'default' behavior of measuring sequentially.
- Agent quickly learns to outperform default plan
- On long time-scales, policy doesn't matter (lots of statistics to go around)
- For 'abrupt' beamtimes, agent demonstrates >100% improvement in score than default behavior.



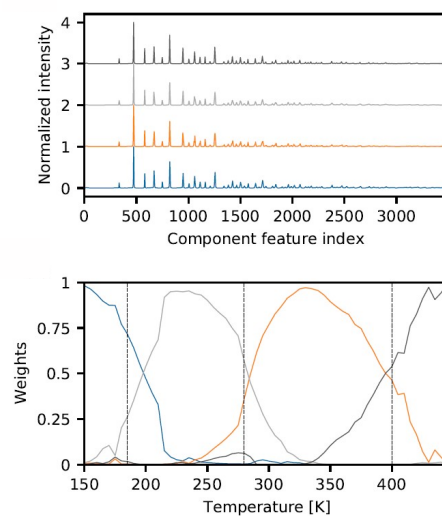
The right tool for the right job...

Variational Autoencoders
for novelty detection



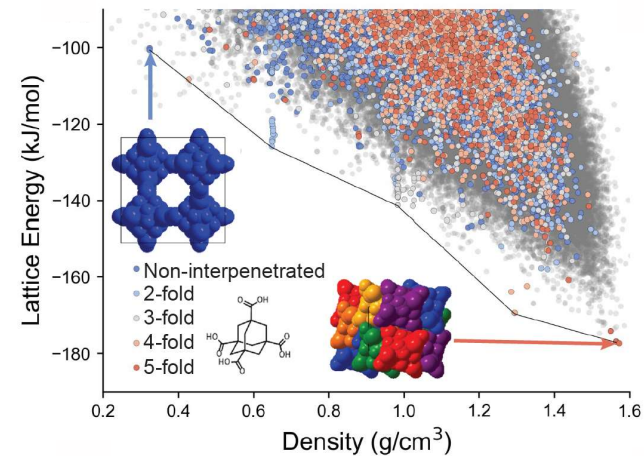
arXiv: 2104.04392

Unsupervised interpretation of
phase transitions



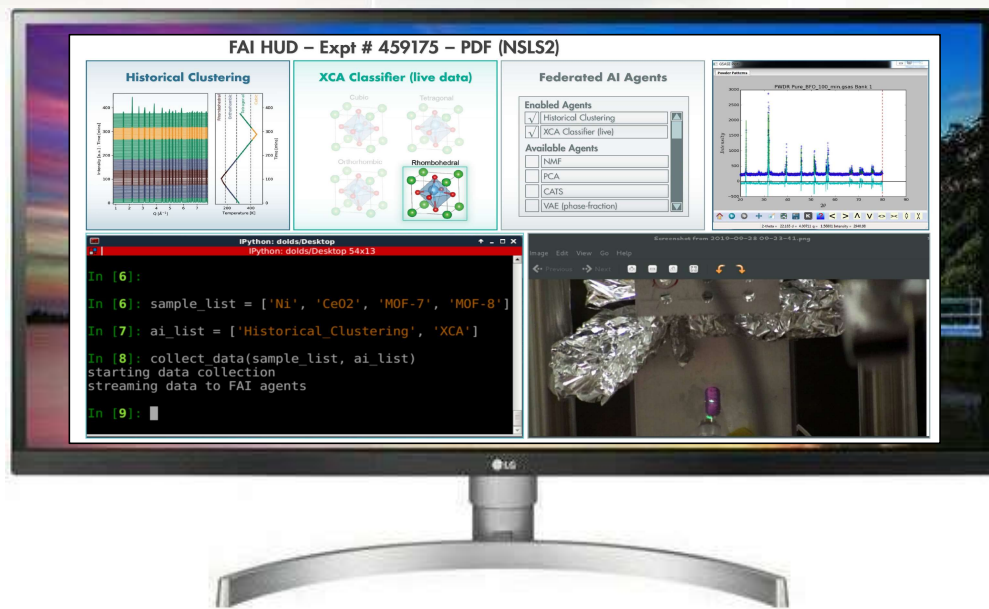
arXiv: 2104.00864

Ensemble supervised learning for
probabilistic characterization



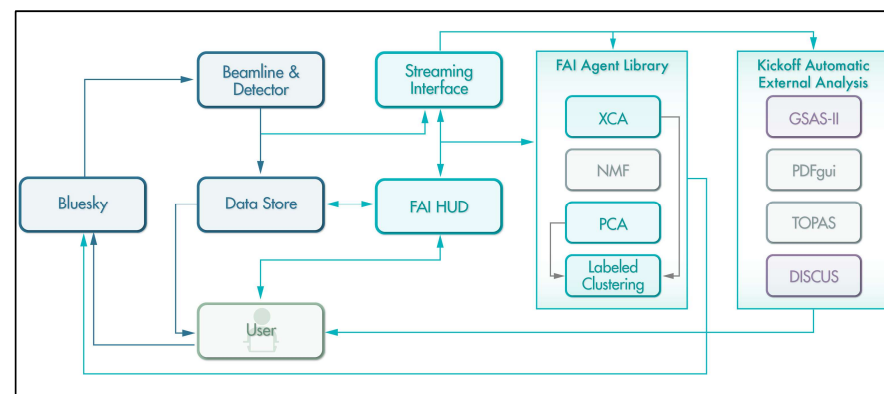
Nat. Comp. Sci.
In Press (2021)

Putting ALL the Tools together with Federated Artificial Intelligence (FAI)



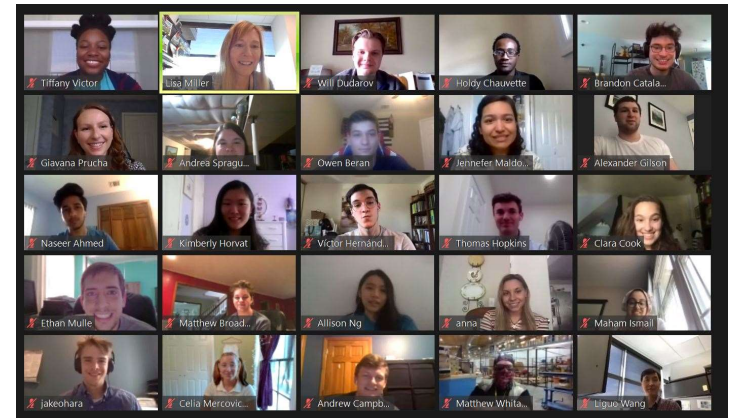
FAI employs a swarm of different methods (both generalized data agnostic and technique-specific) and presents the collective results to researchers in real-time.

Just as a pilot does not rely on a single gauge to fly an airplane, the core design of FAI is that a researcher should not be limited to a single analysis method when trying to understand their data, especially when time or data collection is precious!



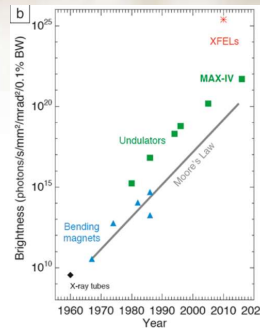
Wanted: Next Generation of Tool Makers

- Previously AI development largely driven by defined commercial and industrial needs.
- The scientific community has leveraged many developments initially driven by commercial needs, but few currently widely adopted.
- NSLS-II is ahead of the pack for lightsources and beamlines worldwide in employing AI methods to accelerate and enable new science.
- Helping to foster next generation of workforce that will be fluent in these methods for applications in physical sciences and beyond.



Conclusions

- At synchrotron lightsources, our measurement speeds have long surpassed our analysis capabilities, and the problem is **only going to get worse!**
- We are nearly to peak-hype on AI, so it's ok to be skeptical... maybe better to think about the marriage of **applied statistics** and **big data**.
- Lots AI and robotics can't do – but we're not aiming to replace older methods that still work.
- We can start thinking about new ways of applying these technologies towards future, previously intractable scientific endeavors.



Thanks for your attention!