09/15/25 6TH GLOBAL RESEARCH PLATFORM WORKSHOP

SCIENTIFIC COMPUTATION AND AI AT ADVANCED PHOTON SOURCE

WEIJIAN ZHENG
Postdoc Appointee
Data Science and Learning Division
Argonne National Laboratory

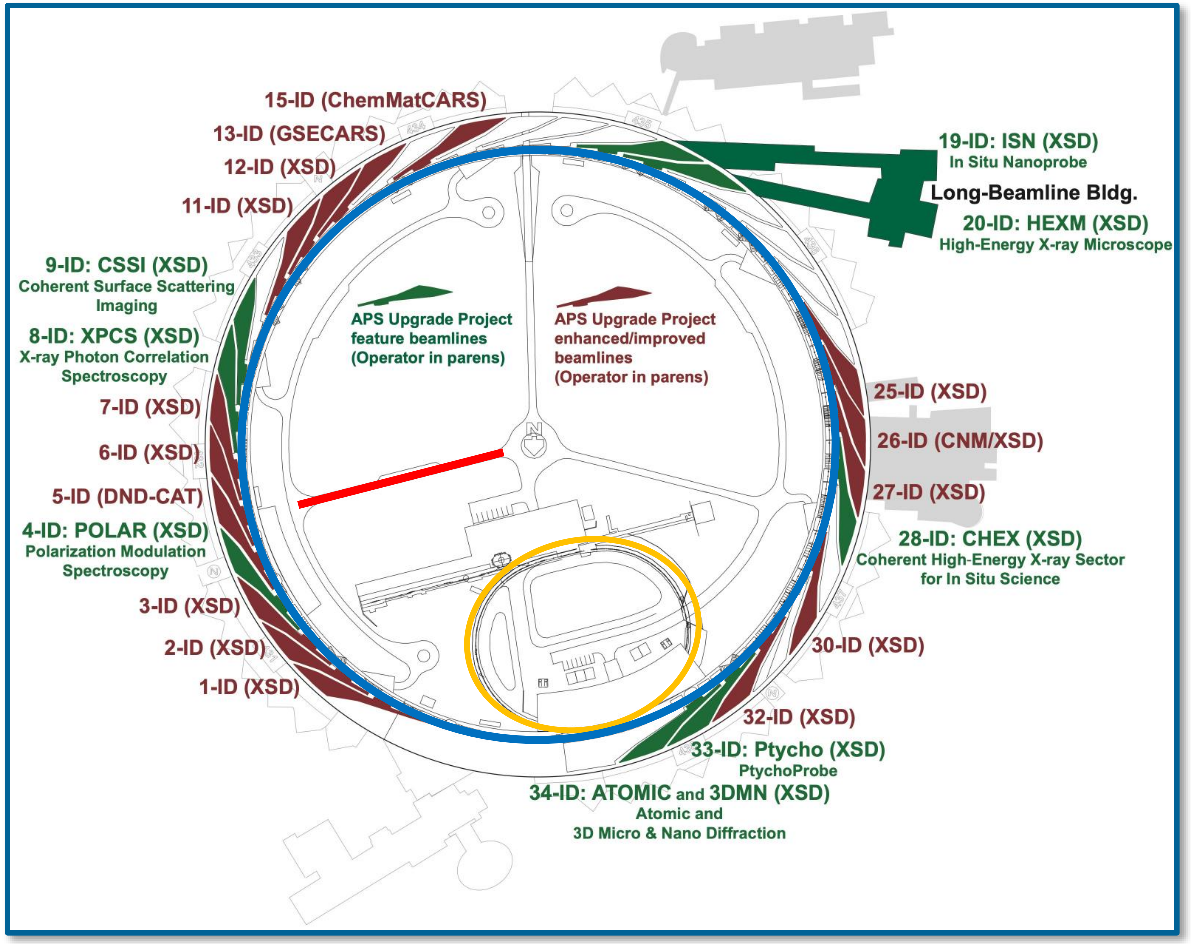
RAJKUMAR KETTIMUTHU
Computer Scientist and Group Leader
Data Science and Learning Division
Argonne National Laboratory

ANTONINO MICELI
Physicist, Group Leader
X-ray Science Division
Argonne National Laboratory

OUTLINE

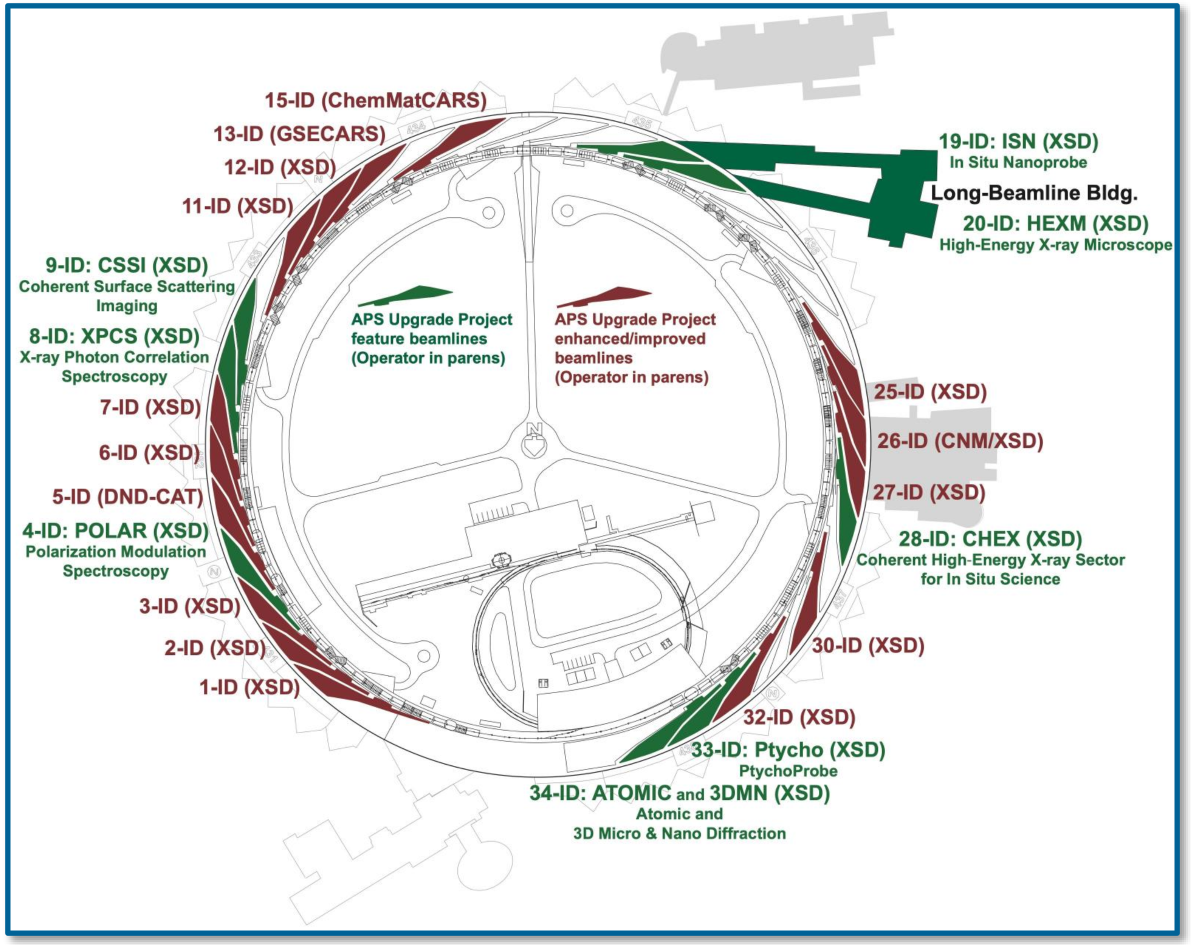
- Background
  - X-ray synchrotron
  - APS upgrade project
- Scientific computing, HPC, and AI at APS
  - Computing goals
  - HPC and network/computing structure
  - ML/AI at APS
  - Challenges
- Conclusions

WHAT IS X-RAY SYNCHROTRON:



- Electrons are generated and injected into a **linear accelerator (linac)** or a smaller **booster** ring
- They are then accelerated to very high energies and injected into the main synchrotron **storage ring**.
  - In the **storage ring**, electrons circulate in a nearly perfect vacuum at close to the speed of light.
  - When electrons are deflected by magnets, they emit synchrotron radiation, very bright, highly collimated X-rays.
- The emitted X-rays are guided into **beamlines**, experimental stations equipped with optics and detectors.

APS UPGRADE PROJECT



- New and updated insertion devices for optimum spectral performance.
- 9 new feature beamlines + Long Beamline Building (green)
- 15 enhanced and improved beamlines (red)
- Coupling with high performance computing & AI/ML for real-time data analysis and visualization

44 Petaflops

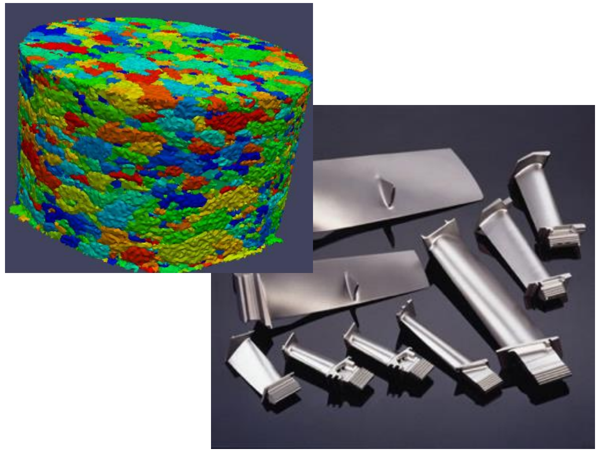
~2 Exaflop (2000 Petaflops)

HEXM: High Energy X-ray Microscopy (20-ID)

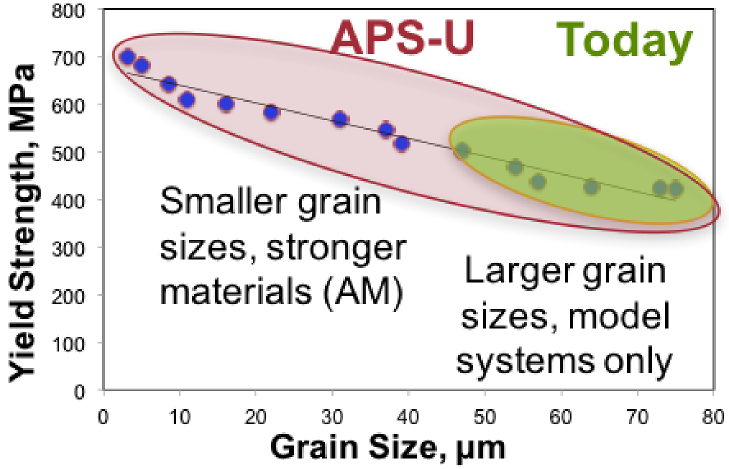
Long beamline for 3D materials characterization of engineering materials

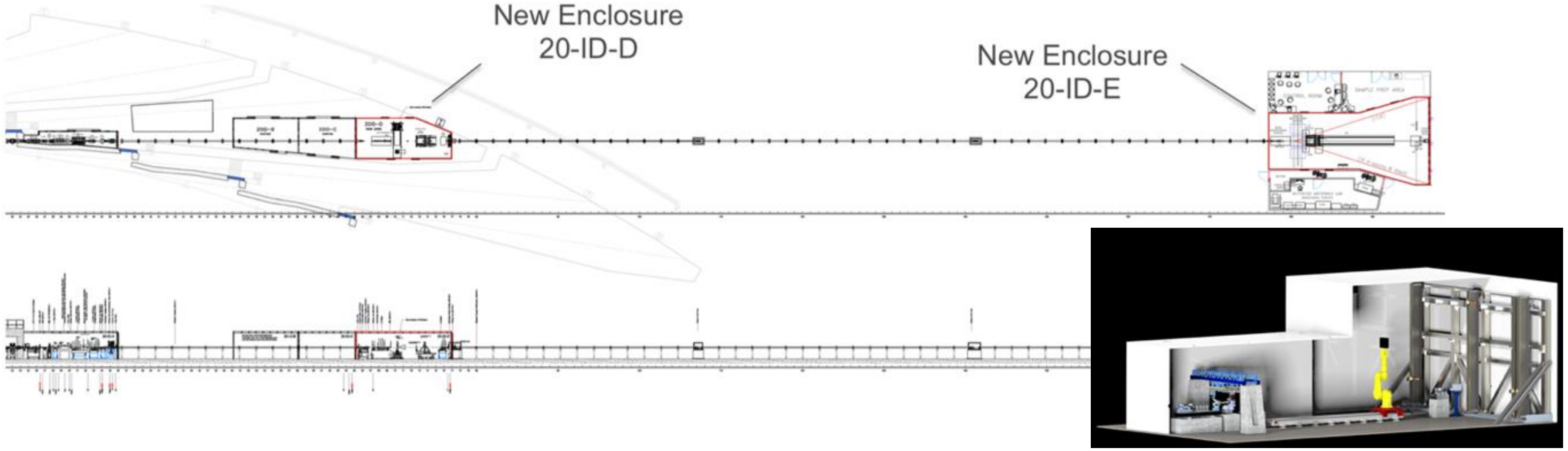
HEXM will give new information:

- Intra-granular structure (eg, ‘see’ crack initiation)
- Smaller grained materials (eg, energy storage, additive manufacturing)
- Dynamic processes (eg, fatigue)



Process-enhanced properties for airfoils





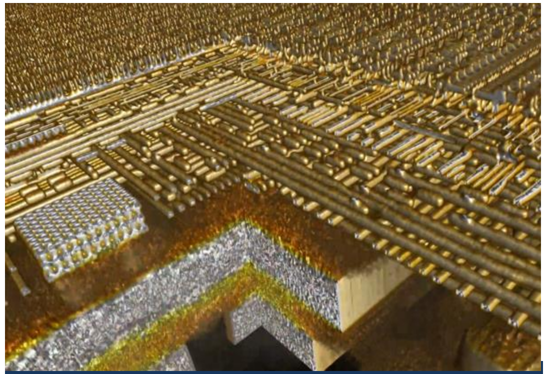
Activated Materials Laboratory (AML)

APS UPGRADE PROJECT

- APS-U beamlines are coming on-line providing transformative capabilities for understanding the spatial and temporal properties of materials using x-rays.
- The greatly increased data rates and volumes will require new analysis, storage, computation strategies to obtain real-time scientific insight.

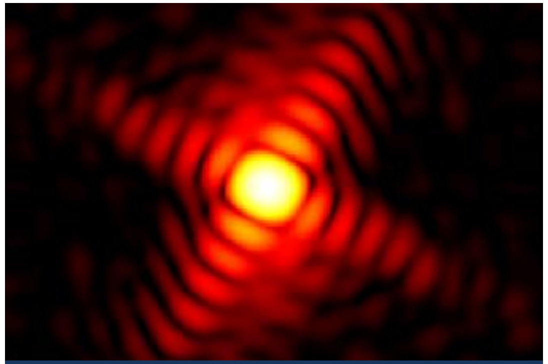
HIGH ENERGY

Penetrate bulk materials and operating systems



BRIGHTNESS

Macroscopic fields of view with nanometer resolution



COHERENCE

Highest possible spatial resolution

DATA SCIENCES

Real-time data analysis and decision making

The Role of Scientific Computing at the APS

Advanced computational approaches and technologies are essential to fully unlocking the scientific potential of the new facility

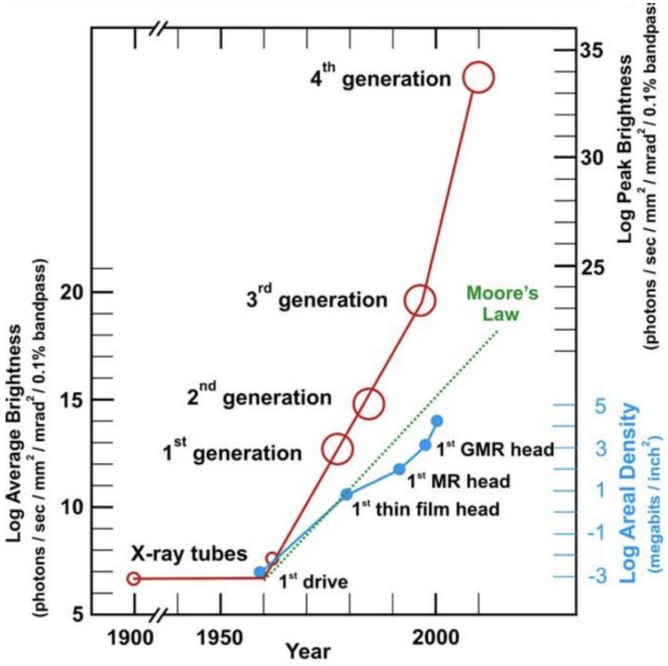
- The upgraded source opens the door for new measurement techniques and increases in throughput, which, coupled to technological advances in detectors, new multi-modal data, and advances in data analysis algorithms, including artificial intelligence and machine learning (AI/ML), will open a new era of synchrotron light source enabled research
- In particular, the high-brightness, and increase in coherent x-ray flux at the new APS is leading to significant increases in data rates and experiment complexity that can only be addressed with advanced computing

APS Strategic Computing Goals

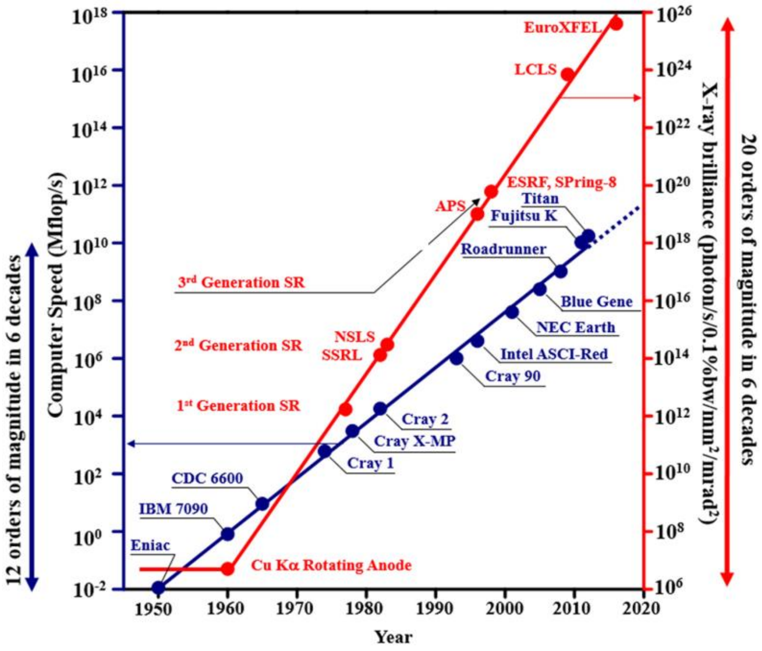
1. Enable High-Speed Discovery	*Harness the unprecedented power of computing and data science expertise at Argonne to reduce the time to science*
2. Unlock Exceptionally Challenging Experiments	*Apply Argonne's strengths in AI/ML and autonomous experiments, large-scale simulations, and x-ray experiments to unlock otherwise unreachable complex scientific solutions*
3. Leverage Data for New Science	*Take advantage of the opportunities created by the explosion of data at the APS coupled with AI to advance new science that is only achievable using combined knowledge*
4. Empower Users to Realize the Full Potential of the APS	*Provide APS users with the cutting-edge tools needed to produce world leading science*
5. Expand Collaborations	*Leverage partnerships and initiatives to amplify the impact of Argonne and APS resources*

WHY HPC+AI?

Light source Compute needs outpace moore's law



http://archive.synchrotron.org.au/images/AOF2017/Boland---AOF---Future-light-sources-2017-05-29.pdf



https://www.physics.ucla.edu/research/imaging/research_CDI.html

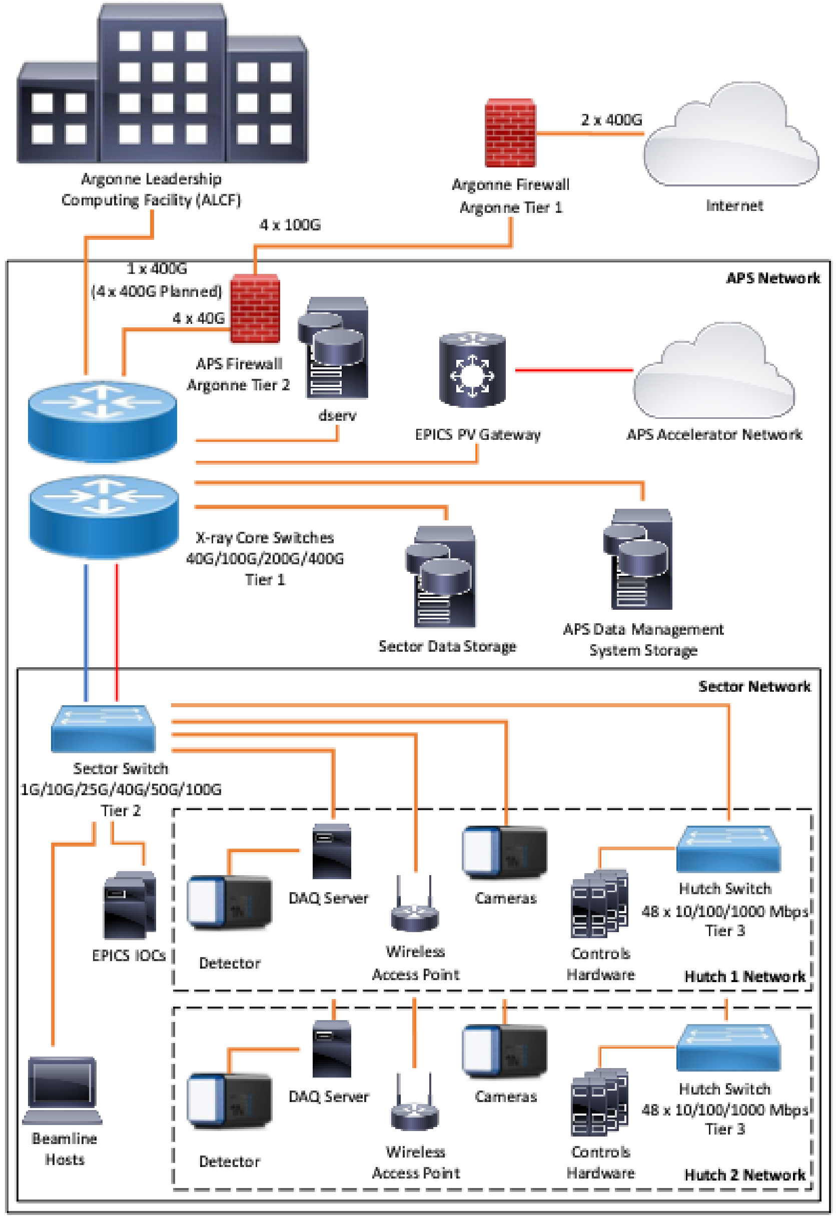
Oleg Shpyrko Ph.D thesis: https://arxiv.org/abs/cond-mat/0407333

We need to rethink *how* we do data analysis → NN models 100-1000X faster than conventional methods

Network Infrastructure

State-of-the-art network performance, and security and resiliency

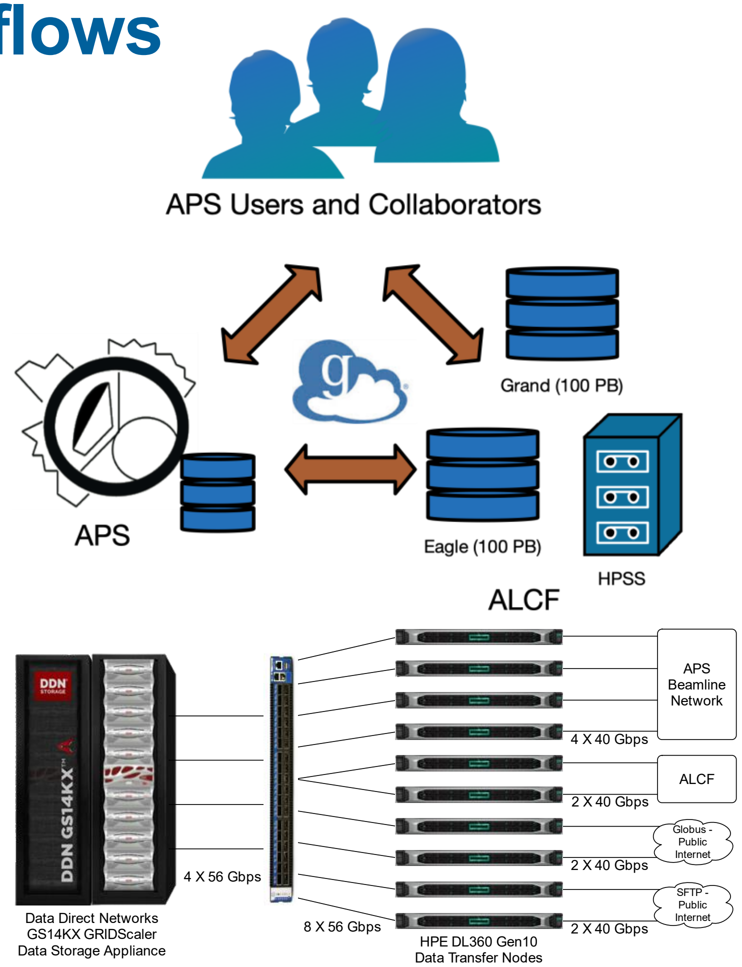
- 3-tier network infrastructure: facility, sector, enclosure
- Supervisory control and data acquisition (SCADA) architecture to better support controls, data, and regular network traffic
- New switches for APS sectors capable of multiple 100 Gbps links
- New core network switches capable of multiple 400 Gbps links
- Upgraded the APS ↔ ALCF network connection to 4 x 100 Gbps
- Installed new fiber plant for all APS beamlines; 768 pairs of new single-mode fiber from the APS computer room to beamlines



Data Management, Workflows, and Science Portals

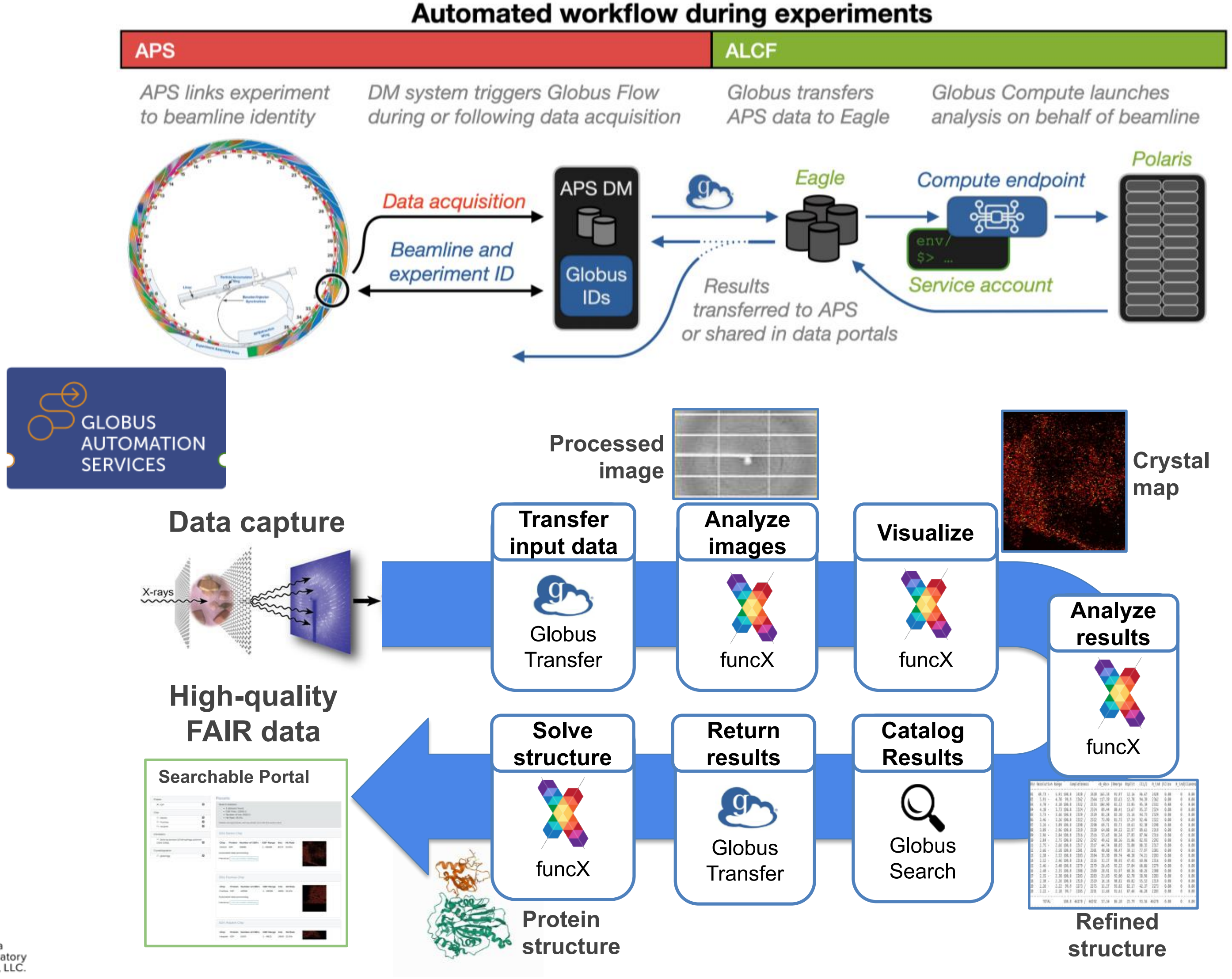
The APS Data Management System — facility-wide software and hardware for managing data and computing workflows

- Tools to automate transfer of data between acquisition devices, computing resources, and data storage systems
- Data from detectors is written directly to central storage (~8 PB) or to local storage and transferred after acquisition; long-term storage on tape
- Ownership and access permissions are granted to the users signed-up to perform a particular experiment
- Workflow tools automate data processing via plug-ins and integration with Globus Compute
- Users can download data at their home institutions using Globus Transfer (globus.org)



Data Management, Workflows, and Science Portals

Leveraging Globus as a computational fabric to enable computing

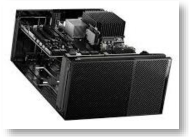


Computing Infrastructure

Multi-tiered approach to computing from beamlines to supercomputers

Instrument / Beamline **APS** **ALCF**

Compute Infrastructure



Edge Devices

Workstations
GPGPU-equipped

Local Cluster
~20 TFLOP/s CPU cores

Polaris Supercomputer
44 petaflop/s (4 petaflop/s for on-demand)

Aurora Supercomputer
>1 exaflop/s

Storage infrastructure

Voyager
(~8 PB)

New Data Buffer
(~50 PB)

Grand
(100 PB)

Eagle
(100 PB)

HPSS
Tape Storage

Network infrastructure

New beamline and instrument networks; new Tier 2 and 3 switches (in progress)

New APS fiber plant and Tier 1 switches (completed)

Terabit per second APS <-> ALCF network (planned)

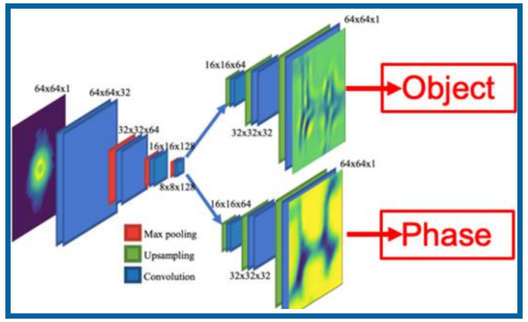
New high-speed ALCF network (planned)

Artificial Intelligence / Machine Learning (AI/ML)

Data Reduction

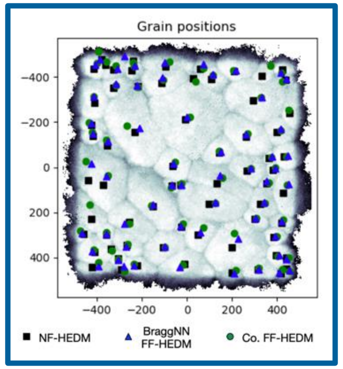
PtychoNN: Machine learning ptychography reconstruction

- 100s of times faster and requires up to 5 times less data than conventional iterative approaches



BraggNN: Machine learning method for determining Bragg peak locations from far-field high-energy diffraction microscopy data

- >200 times faster than conventional pseudo-Voigt profiling approach

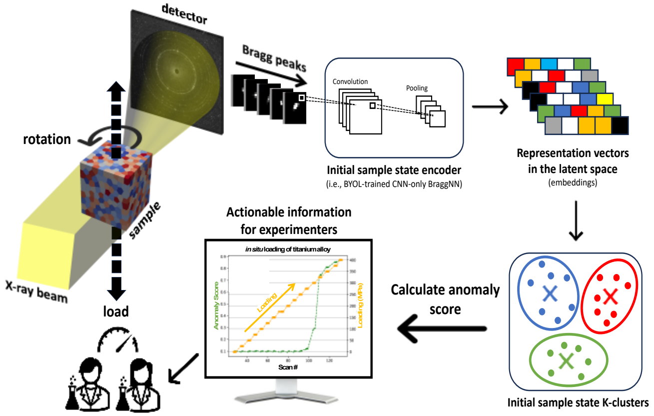


A. Babu, T. Zhou et al., Nature Comm., 14, 7059 (2023) .

Liu, Z., Sharma, H., Park, J. S., Kenesei, P., Miceli, A., Almer, J., Kettimuthu, R., Foster, I., “BraggNN: fast X-ray Bragg peak analysis using deep learning,” IUCrJ 9, 104-113 (2022).

Experiment Steering

REI workflow: A new method to determine the onset of plastic deformation in bulk metallic polycrystalline engineering materials as they experience external stimuli. As compared to previous microstructure reconstructions, the new approach is fully automated, computationally faster by 50×, and works for up to 7× sparser datasets.

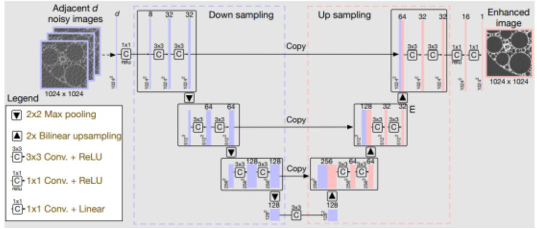


Zheng, W., Park, J.S., Kenesei, P., Ali, A., Liu, Z., Foster, I., Schwarz, N., Kettimuthu, R., Miceli, A. and Sharma, H., 2024. Rapid detection of rare events from in situ X-ray diffraction data using machine learning. IUCrJ, 57(4).

Knowledge Extraction

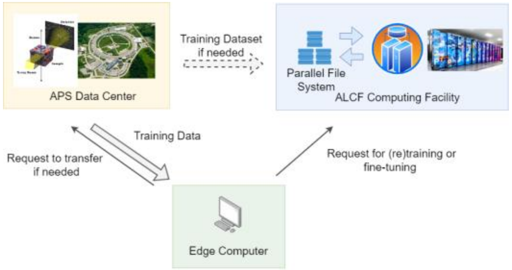
TomoGAN: Generative adversarial network improves the quality of tomographic reconstructions

- Uses up to 1/16th less dose or projections



Liu, Z., Bicer, T., Kettimuthu, R., Gursoy, D., De Carlo, F. and Foster, I., 2020. TomoGAN: low-dose synchrotron x-ray tomography with generative adversarial networks: discussion. Journal of the Optical Society of America A, 37(3), pp.422-434.

M2ML: Model and Data Management for Machine Learning: Integrating Instruments, Edge and HPC for Accelerated Machine Learning

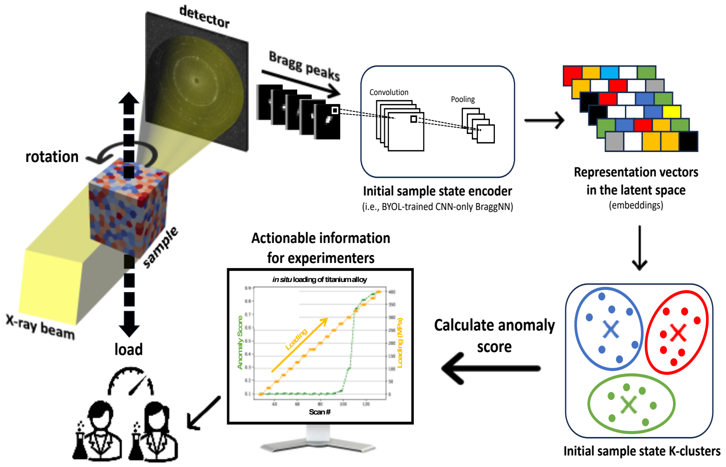


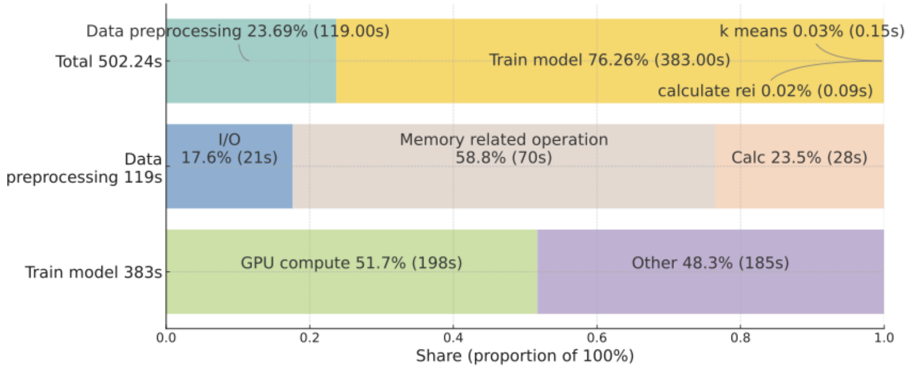
Zheng, W., Sharma, H., Chard, R., Kenesei, P., Park, J.S., Schwarz, N., Miceli, A., Foster, I.T. and Kettimuthu, R., 2024, December. Model and Data Management for Machine Learning (M2ML): Integrating Instruments, Edge and HPC for Accelerated Machine Learning. In 2024 IEEE International Conference on Big Data (BigData) (pp. 4275-4282). IEEE.

CHALLENGES FOR CURRENT TOOLS

Use the REI anomaly detection workflow as an example

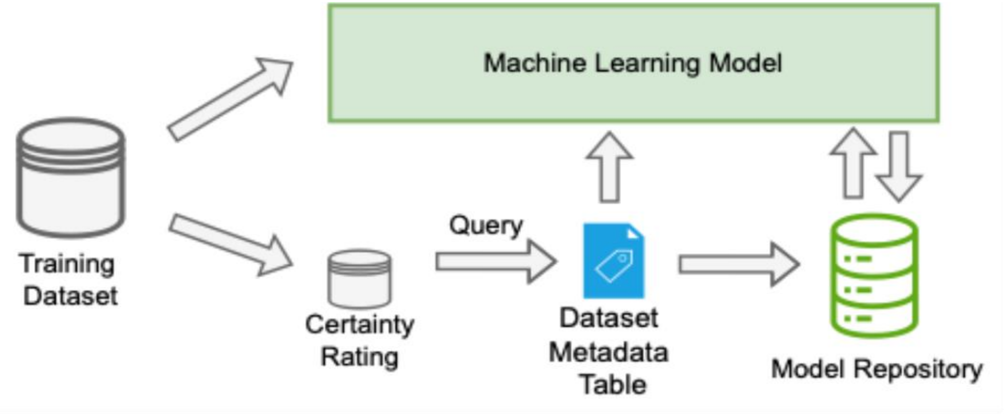
- Current data processing time: 8 minutes for each scan of 1440 frames
  - Each scan of data is around 12 GB
  - With major bottleneck of data preprocessing
- Two major upgrade will happen soon:
  - Increased detector size: from 2k-by-2k -> 4k
  - Faster detector speed: 7 fps -> 700 fps

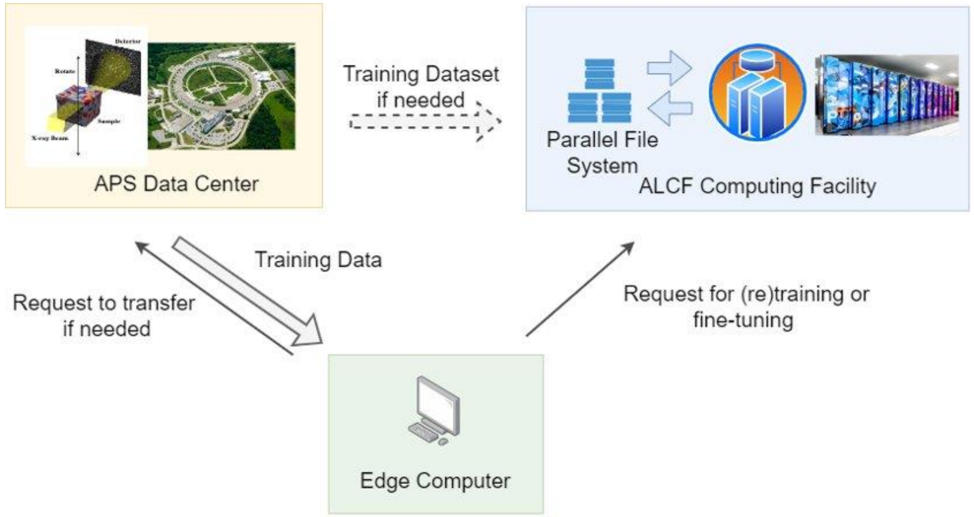




MODEL AND DATA MANAGEMENT FOR MACHINE LEARNING (M2ML)

- A data and model management framework for machine learning,
  - Abstracts it into three key tasks: the 1) initialization task, the 2) (re)training task, and the 3) analysis task
  - To enhance the usability of the framework
- Optimize the performance
  - By introducing a metadata table to replace the traditional database systems





SUMMARY

- Big data problem very real at APS, existing and next generation light sources
- Hundreds of Pb per year generated at APS after the upgrade
- Complexity, multi-modality, operando science is becoming the norm
- Exploit full computing continuum, including learning from real time data
- Delivering rapid analysis at scale is critical and will provide competitive advantage
- Many other opportunities, for example:
  - Accelerator control and fault detection with edge devices
  - Coupling simulations, advances in surrogate models with experimental science

Argonne
NATIONAL LABORATORY