

OCTOBER 25, 2023

CATALYZING SCIENTIFIC ADVANCEMENTS CUTTING-EDGE WORKFLOWS AND COLLABORATIVE COMPUTING AT THE ADVANCED PHOTON SOURCE

HANNAH PARRAGA

Software Engineering Associate
Argonne National Laboratory

**JOHN HAMMONDS, STEVEN HENKE,
SINIŠA VESELI, NICHOLAS SCHWARZ,**



Argonne National Laboratory is a
U.S. Department of Energy laboratory
managed by UChicago Argonne, LLC.



X-Ray Light Sources

Indispensable tools in the exploration of matter

- Wavelengths of the emitted photons span a range of dimensions from atoms to cells
- Scientific user facilities
- 5 funded by the US Department of Energy - Basic Energy Sciences
- Serves thousands of users per year from academia, government, and industry
- Diverse communities: materials research, biology, geosciences, life sciences, security, and many more



Advanced Photon Source (APS) at
Argonne National Laboratory

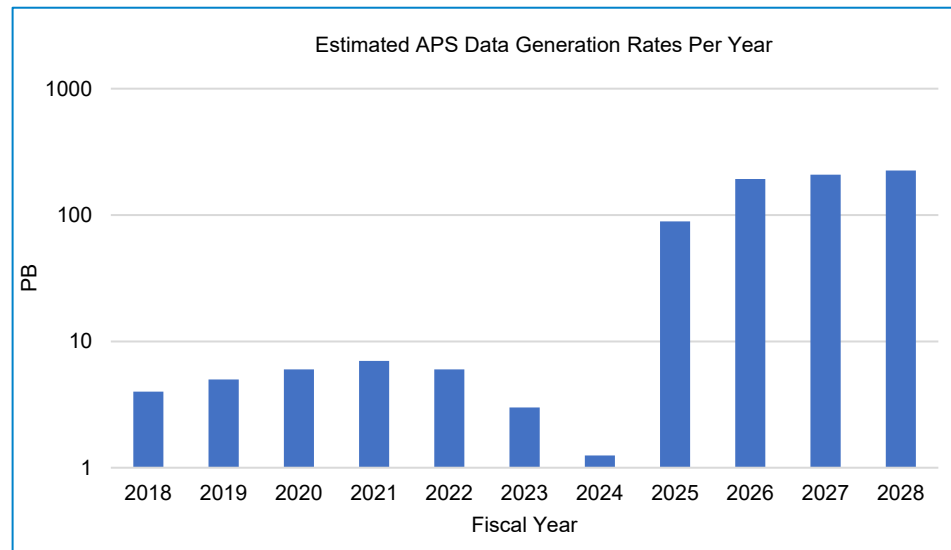
THE UPGRADED ADVANCED PHOTON SOURCE

Upgraded light source up to 500x brighter than current APS for higher resolution

Can image materials in less than a billionth of a second

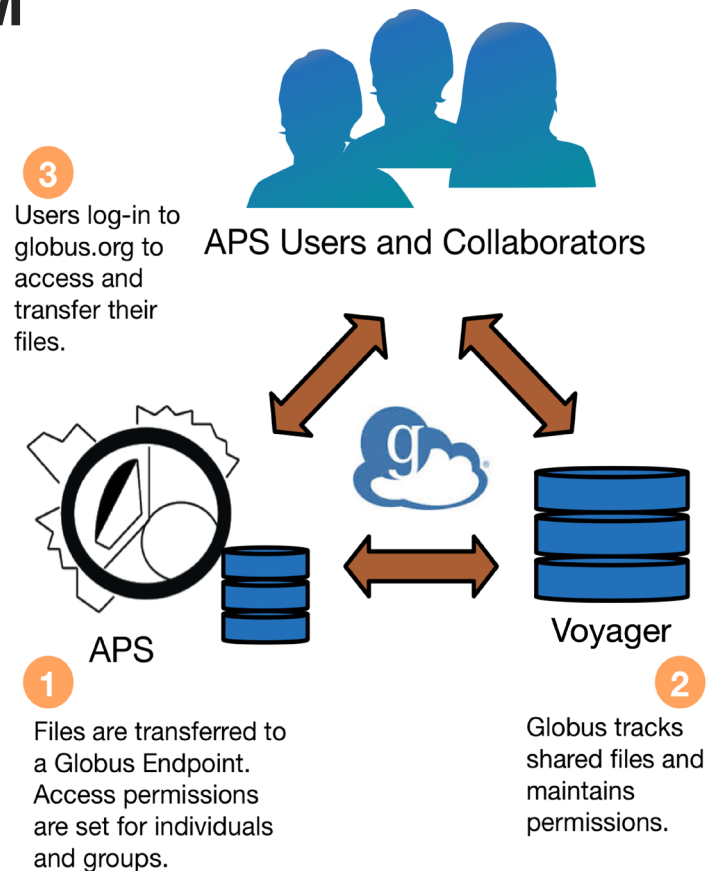
Over the next decade the APS will

- Generate 100s of petabytes (PBs) of raw data per year
- Require 10s of petaflop/s of on-demand computing power for first pass data processing, reconstruction, and reduction



APS DATA MANAGEMENT SYSTEM

- DM software is designed to help beamlines with various data management tasks
- *Storage area management*: movement of acquired data from local storage to a more permanent location, data archival, etc.
- *Enabling users and applications to easily find and access data*: metadata and replica catalog, remote data access tools
- *Facilitating data processing and analysis with automated (in real-time) or user-initiated processing workflows*



LEVERAGING GLOBUS SERVICES

- **Globus Connect Server 5** installed at central storage nodes allows secure access to data
- **Globus Transfer** provides data movement within flows and by users for post processing tasks
- **Globus Compute** offloads heavy computation to remote resources such as HPC
- **Globus Flows** executes complex data analysis pipelines
- **Globus Groups** secures access to experiment results in data portals
- **Globus Auth** executes flows as a beamline service account



BLUESKY INTEGRATION



- Bluesky is an experiment specification and orchestration engine for control and collection of scientific data and metadata
- Bluesky interacts with hardware through a high-level abstraction
- Cross-facility, international collaboration
- Integration with the APS Data Management System means an end-to-end pipeline from detector to data portal



PARTNERING WITH THE ARGONNE LEADERSHIP COMPUTING FACILITY

- Service accounts for each beamline allow submitting jobs without each user maintaining their own compute allocations
- Trusted ALCF and APS Globus credentials to manage shared data access
- On-demand computing with preemption quickly executes APS jobs
- Dedicated log-in nodes for Polaris run Globus Compute Endpoints

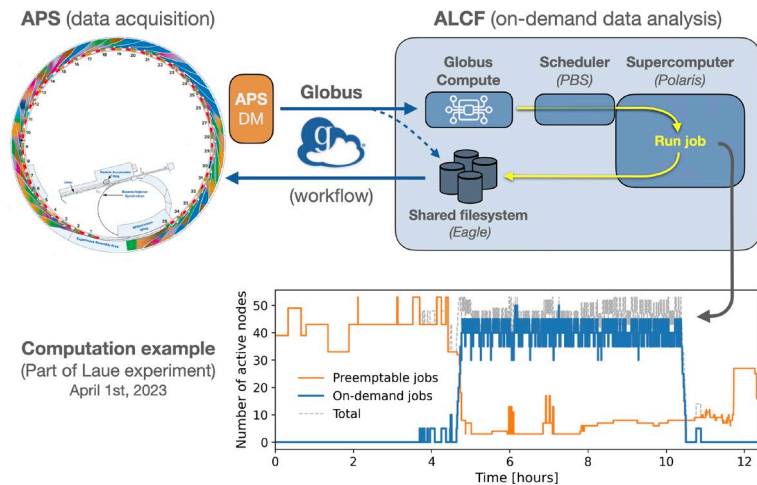
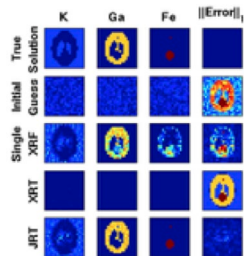
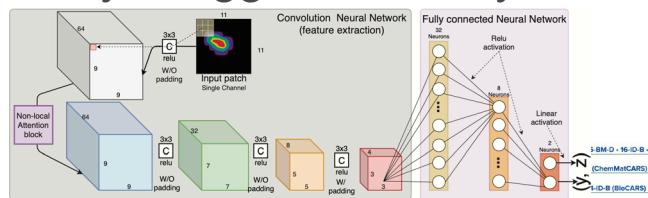
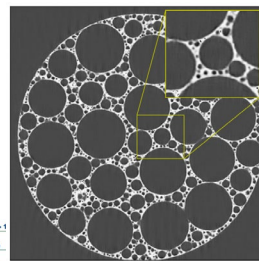
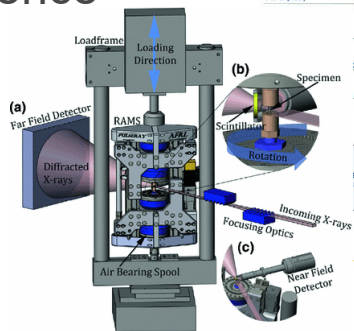


Diagram courtesy of Benoit Cote

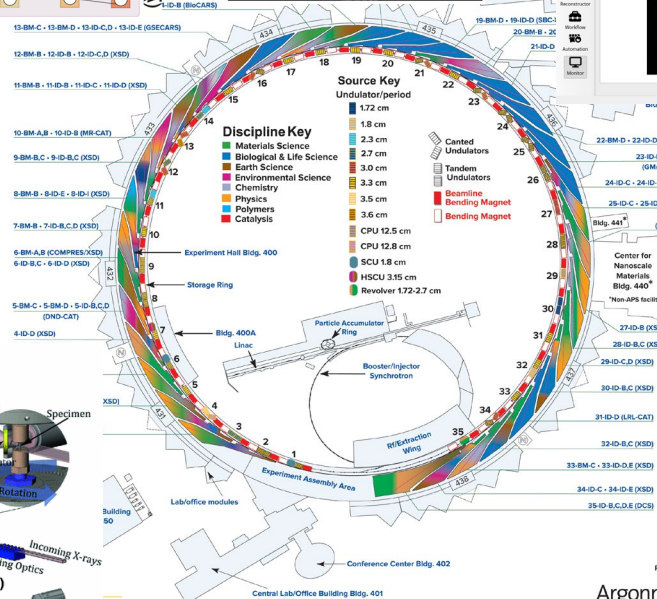
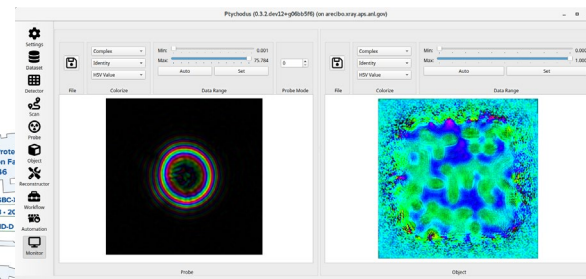
Fast X-ray Bragg Peak Analysis



X-ray Fluorescence Spectroscopy

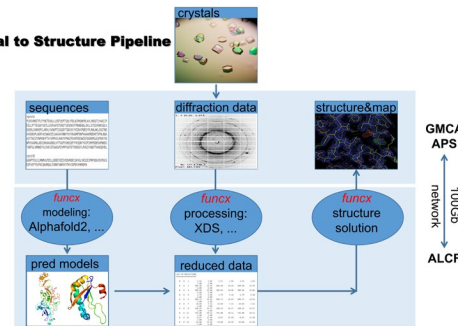


Tomography



Ptychography

Crystal to Structure Pipeline



Crystallography (GM/CA)

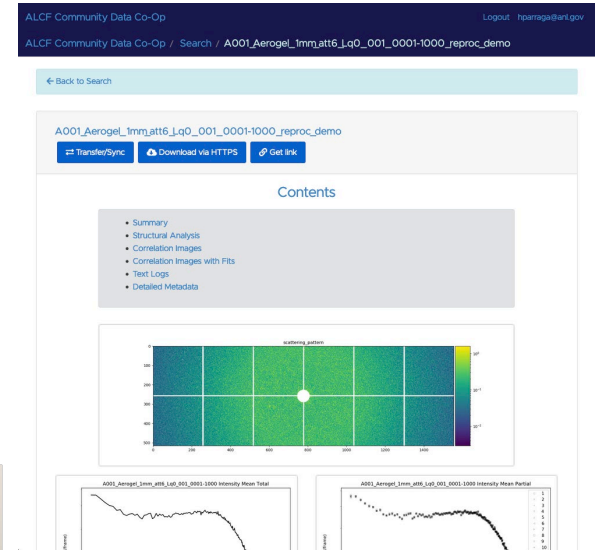
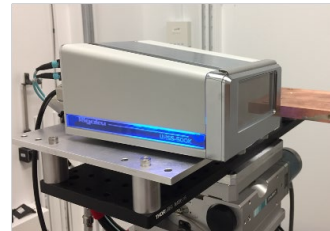
High-energy Diffraction Microscopy

Laue Micro-diffraction

X-RAY PHOTON CORRELATION SPECTROSCOPY AT BEAMLINE 8-ID-I

- Research outcomes include study of cartilage and joints, solar cells, flexible electronics, plastics, batteries, glass, effect of stress on the body, viruses and more...
- Hundreds of experiment users each year
- Over 400 TB data generated so far
- Hundreds of experiments processed

XPCS Data Portal
<https://acdc.alcf.anl.gov/>



RIGAKU 0.5 MPix
Ultra High Speed Sensor
Q. Zhang *et al.*, J. Synchrotron Rad. (2021)

Adopting Digital Object Identifiers














- A Digital Object Identifier (DOI) is a string of numbers, letters and symbols used to permanently identify an article or document and link to it on the web
- The Office of Science and Technology Policy's 2022 Public Access Memo states that federally funded research requires a unique digital persistent identifier for all scientific research and development awards
- DOIs created for 2023 beamtime awards

Award DOI Service

OSTI.GOV
U.S. Department of Energy
Office of Scientific and Technical Information

[Submit/Edit](#) [Dev Tools & Resources](#) [About](#) [FAQs](#) [Contact](#) [Create Account](#) [Sign In](#)

Structure determination of proteins from Sars-CoV-2

AWARD Award DOI: https://doi.org/10.46936/APS-179063/60009021	Award Description Covid-19 Research Structural Analysis of proteins from Sars-Cov-2: Determine the crystal structures of listed proteins, in the complexes with small potential drug molecules, RNA, or other proteins from Sars-Cov-2. NSP3_PLPro_HSCUDV53 NSP9 NSP15 NSP16 NSP10 NSP8 NSP7 SGR_a SGR_b ADRP 3CLpro PLP Fabs mAbs Investigator: Gade, Priyanka  ^[1] ; Chang, Changsoo  ^[2] ; Finckrock, Yanhui  ^[2] ; Malteeva, Natalia  ^[2] ; Michaleka, Karolina  ^[2] ; Joachimiak, Andrzej  ^[2] ; Kim, Youngchang  ^[2] ; Lazarski, Krzysztof  ^[2] ; Osipiuk, Jerzy  ^[2] ; Tan, Kevin  ^[2] <div><div>1. The University of Texas at El Paso ^K</div><div>2. Argonne National Laboratory ^K</div></div> Awarding Organization: Advanced Photon Source (APS)  ^K Award Type: facilities Site Award Number: 179063 Award Start Date: 2022-10-03
---	---

ACKNOWLEDGEMENTS

ALCF:

- William Allcock
- Benoit Cote
- Jennifer Francis
- Carissa Holohan
- Ryan Milner
- Michael Papka
- Paul Rich
- George Rojas
- Haritha Siddabathuni Som

DSL:

- Tekin Bicer
- Ian Foster
- Rajkumar Kettimuthu

Globus:

- Rachana Ananthakrishnan
- Benjamin Blaiszik
- Ryan Chard
- James Pruyne
- Nickolaus Saint
- Rafael Vescovi

ACKNOWLEDGEMENTS

APS:

- Jonathan Almer
- Olga Antipova
- Mathew Cherukara
- Andrew Chihpin
- Miaoqi Chu
- Francesco De Carlo
- Junjing Deng
- Barbara Frosik
- Arthur Glowacki
- Doga Gursoy
- John Hammonds
- Steven Henke
- Pete Jemian
- Yi Jiang
- Peter Kenesei
- Jonathan Lang
- Alex Lavens
- David Leibfritz
- Lu Xi Li
- Antonino Miceli
- Suresh Narayanan
- Jun-Sang Park
- Michael Prince
- Arvind Ramanathan
- Nathan Rogers
- Alec Sandy
- Nicholas Schwarz
- Roger Sersted
- Hemant Sharma
- Dina Sheyfer
- Kenneth Sidorowicz
- Joseph Sullivan
- Jonathan Tischler
- Susan White De Pace
- Max Wyman
- Sinisa Veseli
- Stefan Vogt
- Qingping Xu
- Fang Zhang
- Xuan Zhang
- Tao Zhou

REFERENCES

- R. Pokharel, "Overview of High-Energy X-Ray Diffraction Microscopy (HEDM) for Mesoscale Material Characterization in Three-Dimensions," *Materials Discovery and Design By Means of Data Science and Optimal Learning*, Springer, 2018, 167 - 201. DOI: 10.1007/978-3-319-99465-9_7
- D. Sheyfer, Q. Zhang et al., Phys. Rev. Lett. 125, 125504 (2020)
- Z. Liu, T. Bicer, R. Kettimuthu, D. Gursoy, F. De Carlo, and I. Foster. "TomoGAN: Low-Dose X-Ray Tomography with Generative Adversarial Networks." preprint arXiv:1902.07582 (2019)
- Z. Liu, H. Sharma, J. Park, P. Kenesei, J. Almer, R. Kettimuthu, I. Foster. arXiv <https://arxiv.org/abs/2008.08198>
- S. Veseli, N. Schwarz and C. Schmitz, J. Synchrotron Rad. 25 1574 (2018)
- Q. Zhang et al., J. Synchrotron Rad. (2021)
- Daniel Allan, Thomas Caswell, Stuart Campbell & Maksim Rakitin (2019) Bluesky's Ahead: A Multi-Facility Collaboration for an *a la Carte* Software Project for Data Acquisition and Management, Synchrotron Radiation News, 32:3, 19-22, DOI: [10.1080/08940886.2019.1608121](https://doi.org/10.1080/08940886.2019.1608121)
- <https://www.whitehouse.gov/wp-content/uploads/2022/08/08-2022-OSTP-Public-Access-Memo.pdf>
- <https://www.doi.org/the-identifier/what-is-a-doi/>

REFERENCES

- <https://www.aps.anl.gov/Dynamics-and-Structure>
- <https://acdc.alcf.anl.gov/>
- <https://git.aps.anl.gov/DM/dm-docs>
- <https://github.com/globus-gladier/gladier-xpcs>
- https://github.com/AdvancedPhotonSource/boost_corr
- <https://github.com/AdvancedPhotonSource/ptychodus>
- <https://github.com/AdvancedPhotonSource/XRF-Maps>

Argonne National Laboratory's work was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under contract DE-AC02-06CH11357. Use of the Advanced Photon Source was supported by the U. S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357. This research used resources of the Argonne Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC02-06CH11357. The submitted manuscript has been created by UChicago Argonne, LLC, Operator of Argonne National Laboratory ("Argonne"). Argonne, a U.S. Department of Energy Office of Science laboratory, is operated under Contract No. DE-AC02-06CH11357. The U.S. Government retains for itself, and others acting on its behalf, a paid-up nonexclusive, irrevocable worldwide license in said article to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, by or on behalf of the Government. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan. <http://energy.gov/downloads/doe-public-access-plan>.



U.S. DEPARTMENT OF
ENERGY