Scientific Software Engineering & Data Management
X-ray Science Division (XSD), Advanced Photon Source
Effort Plan for FY24 and FY25

The APS Scientific Software Engineering & Data Management group provides leadership and scientific software engineering expertise in the areas of data analysis, data management, high-performance computing, visualization, and workflow and orchestration applications in support of world-class photon sciences at the APS. This mission is realized through the creation of a core software application portfolio in prioritized areas, including coherence, imaging, and high-energy techniques, as well with software tools for data access and management, and data streaming for real-time feedback. Effort is aligned with facility priorities and strategies, which at this time include scientific software and data management tools critical to the techniques enabled by the APS-U, and is used in a transparent, flexible, and well-documented manner.

SDM believes it can uniquely enable great science (and x-ray synchrotron techniques) by creating great scientific and data management software. This is a key component in the creation of beamlines of the future, and is critical to the success of the APS-U Era facility. Well-formed collaborative teams of software engineers, computer scientists, algorithm developers, beamline staff, users, peer groups, and other facilities and institutions develop the SDM group's software portfolio. The group works closely with the APS-CXS group, the CELS-DSL and CELS-MCS divisions, other facilities, and CAMERA to implement new algorithms and mathematical methods, with the XSD-BC group and engineers at other facilities to integrate data analysis with beamline data acquisition systems, and with the APS-IT group, the CELS-DSL division, the Globus team, and other facilities to develop data management solutions. This document will be reviewed and updated at least once a year.

Five-year Goals

The SDM group's goals are directed at creating and deploying software tools enabling the full benefit of the portfolio of anticipated future beamlines, including the APS-U beamlines.

1. Creation and deployment of a set of high-performance computing (HPC) enabled software tools that address cross-cutting critical technique domain areas needed by future beamlines. This includes software in the areas of coherence, imaging, high-energy, and multi-modal techniques.
2. Continue development and deployment of data management tools at XSD beamlines.
3. Integration of general-purpose data streaming, feedback, and verification tools with beamline control software and HPC data analysis software.
4. Integration of APS beamline data workflows with ALCF computing and storage systems.
5. Achieve goals for DOE Lab 23-3030 awards, and the EXCEL MIE proposal/project.

Goals and Action Plan for FY24 & FY25

- Continue to support HPC-enabled tools aligned with APS strategy for techniques such as XPCS, XRF, and ptychography by completing highest-priority goals in each area below.
- Meet FY24 and FY25 goals for the AI/ML for SUFs awards, and LDRDs.
- In collaboration with the Globus team, deliver science data portals and workflow automation tools at APS beamlines.
- Continue development and support of the APS Data Management System at APS beamlines.
- Continue development and support of remote access tools for APS beamlines.
- Create specifications and plans for a new beamline data storage buffer aligned with APS Operations and the EXCEL MIE proposal/project.

Project	Summary	FY24 SDM FTE	FY25 SDM FTE
APS Data Management System	Continue application of analysis workflow, and data management and distribution tools at APS beamlines.	2.59	2.75
APS Remote Access System	Continue support for and development of a web application for managing remote access to beamline computers for remote experiments.	0.25	0.25
Beamline Data Pipeline (BDP) Project	Explore, identify, demonstrate, and deploy a portable data pipeline solution that addresses APS-U Era beamline data needs.	See description below	See description below
Bragg Coherent Diffraction Imaging (CDI) Software	Provide ongoing support for the Bragg CDI reconstruction tools.	1.00	1.00
Coherent Surface Scattering Imaging (CSSI) Software	Implementation of high-performance CSSI and GISAXS software applications.	0.50	0.50
Continuation of A Smart Data and Computational Fabric Pilot for the APS and APS-U	Develop and support data science portals for select APS beamlines and techniques.	See description below	See description below
Correlation Toolkit	Develop a real-time HPC-enabled set of tools for time-based correlation data analysis.	0.75	0.75
General-Purpose Reciprocal-Space Mapping (RSM) Tools	Continue development and support for high-performance RSM tools.	0.25	0.25
In-operando AC scattering software	Continue development of software for in-operando AC scattering experiments.	1 Research Aide Technical; 0.2 FTE	1 Research Aide Technical; 0.2 FTE
Laue Diffraction	Develop high-performance computing tool kit for the new Laue depth reconstruction algorithm.	0.75	0.75
LLMs for APS Scientific Processes	Investigate and prototype the use of LLMs in APS scientific processes, workflows, and experiment automation.	0.15	0.15
MIDAS Tomography Frontend	Refactor TomoPy user-facing GUI to work with MIDAS and improve post-processing capabilities.	0.15	0.15
Ptychography Software	Continue development of HPC ptychography reconstruction software.	0.50	0.50
Ptychography+XRF Software	Develop software that uses ptychography results to enhance XRF data products.	0.25	0.25
X-ray Fluorescence Mapping (XFM) Software	High-performance computing (HPC) enabled fitting library and tools for fast elemental mapping of x-ray fluorescence microscopy software.	0.90	0.90

A Collaborative Machine Learning Platform for Scientific Discovery	Support AI/ML for SUFs award, A Collaborative Machine Learning Platform for Scientific Discovery.	1.00 (postdoc)[*]	1.00 (postdoc)[*]
A Pipeline for Autonomous Comprehensive Reduction and Analysis of Sequential 2-Dimensional Diffraction and Scattering Data	Support a pipeline for autonomous comprehensive reduction and analysis of sequential 2-Dimensional diffraction and scattering data LDRD.	TBD	TBD
Auto Parameter Tuning for XRF	Application and/or development of a supervised learning neural network model and train it on collected XRF datasets. Auto Parameter Tuning for XRF LDRD.	1.00 (postdoc); 0.10[*]	0.10 (postdoc); 0.10[*]
LYNX	Develop workflows and analyze experiment data for the LYNX ptychography-laminography instrument.	1.5[*]	1.5[*]
Simulation-Guided High-Energy X-ray Diffraction Experiment Framework for Understanding Irradiation-Induced Slip Anisotropy in Structural Metals	Support simulation-guided high-energy x-ray diffraction experiment framework for understanding irradiation-induced slip anisotropy in structural metals LDRD.	TBD[*]	TBD[*]
Real-time Feedback & Data Acquisition System for APS-U Accelerator	Software framework and tools for the collection of data used for controls, statistics, and diagnostics of technical systems for the MBA accelerator.	0.16[**]	0.00[**]
Multi-modal Diffraction Tomography	Develop robust near real-time software for diffraction tomography once algorithm development is complete.	0.00[***]	0.00[***]

[*] Funded external to Operations budget
[**] APS-U project funded – Pending FY24 and FY25 ERA
[***] Additional funding/effort levels will enable the XSD-SDM group to take on additional projects and add further capabilities at the APS.

SWOT Analysis for Scientific Software

Strengths	Weaknesses
• World-leading software efforts in several scientific areas. • World-class beamline staff and user groups contribute new algorithms and software that expand the scientific productivity of the APS. • Highly productive internal group of professional scientific software engineers. • Close collaborations with APS users and staff, and the XSD-CXS group to provide algorithms and with the XSD-BC group to provide integration with beamline workflows.	• Current funding situation does not allow for the APS to meet its entire mission-critical data analysis software needs. • Most current generation data analysis tools are not suited to stream data in HPC environments needed to keep up with anticipated data rates. • Many scientist-developed packages lack professional software engineering needed to make them more productive • Lower facility productivity due to lack of data analysis tools.
Opportunities	**Threats**
• Collaborations with ANL expertise will help bring state-of-the-art HPC applications to the APS. • Collaborations with DOE facilities and resources could amplify development efforts and provide needed software in a cost-effective manner for the entire DOE complex. • The APS Upgrade-enabled techniques may be fully realized, answering new scientific questions; the APS maintains its position as the most productive light source.	• Without further investment and collaboration in this area, the APS will not fully realize the scientific potential of the APS Upgrade. • User groups may seek to perform cutting-edge experiments at other light sources where better software support is available. • Other domestic and international light sources have considerably larger and more active software and algorithm development programs that can leapfrog APS leadership.

SWOT Analysis for Data Management & Distribution

Strengths	Weaknesses
• World-leading expertise at ANL in data sciences, data management and transfer (e.g. Globus Services team). • APS is one of the DOE's largest data collecting user facility, producing a wealth of scientifically valuable data. • Collaborative efforts continue to form between the APS and expertise elsewhere at ANL.	• Preponderance of existing unique solutions at beamlines involving manual, inefficient management steps; no common user experience. • Current manual methods cannot keep pace with increasing data rates. • Lowered productivity due to time taken away from staff and users to address tasks that may be automated.
Opportunities	**Threats**
• Leverage expertise from CLS, UoC, and the Globus Services team. • Reduce cost by leveraging outside software resources and expertise. • Consistent data management user experience. • Increase scientific productivity through automation of data management tasks.	• The full potential of the APS Upgrade cannot be realized without managed data workflows. • Lowered scientific productivity due to an inability to keep up with increases in data. • International light sources that have invested heavily in data management software may overtake the APS in terms of scientific productivity.

Appendix - Project Details

APS Data Management System

Summary: Continue application of analysis workflow, and data management and distribution tools at APS beamlines.

Team: John Hammonds (SDM), Steven Henke (SDM), Hannah Parraga (SDM), Sinisa Veseli (SDM), APS-IT, et al.
FY24 SDM Effort: 2.59 FTE
FY25 SDM Effort: 2.75 FTE
Out Years: TBD

As data rates and volumes increase due to a combination of advances in detector technologies, increased use of multi-modal acquisition techniques, and the planning benefits of the APS-U project, current manual data workflow and management mechanism will not be sufficient. The APS has a need for tools and infrastructure that automates analysis pipelines, maintains and tracks data ownership, catalogs metadata, provides data distribution endpoints and Software as a Service (SaaS) web interfaces for data analysis, etc.

The APS team will place great emphasis on leveraging best-in-class tools, rather than on developing new systems. For example, they will continue to work closely with the Globus Services team in order to not duplicate effort and best leverage DOE and ANL resources. Open-source tools will be used in order to best meet the needs of the APS in an efficient and cost-effective manner.

Goals include:
1. Continue support for existing beamlines.
2. Deploy system at additional beamlines, where possible.
3. Interface with Argonne Leadership Computing Facility (ALCF) tape-based storage.
4. Develop mechanisms to enable writing data directly to a storage system from a detector.
5. Add additional features to support eBERLight.
6. Support DOI effort.
7. Streamline deployment process with deployment tools and instructional materials.
8. Add features in support of beamline data confidentiality.
9. Develop web portal and a graphical user interface for workflow and data processing job management. Add monitoring features for tracking overall system health and usage.
10. Add features in support of direct streaming of detector data to HPC resources.
11. Interface with Globus automation tools where appropriate.

APS Remote Access System

Summary: Continue support for and development of a web application for managing remote access to beamline computers for remote experiments.

Team: John Hammonds (SDM), AES-IT, et al.
FY24 SDM Effort: 0.25 FTE
FY25 SDM Effort: 0.25 FTE
Out Years: TBD

Goals include:
1. Continue to provide support and development for the existing system.
2. Implement scheme to automatically add/remove remote users to specific beamline computers.

3. Enforce automatically removing user access after a set period.
4. Enforce removal of staff that have left Argonne.
5. Explore multi-factor authentication mechanisms.
6. Add end-of-run reporting.

Beamline Data Pipeline (BDP) Project

Summary: Explore, identify, demonstrate, and deploy a portable data pipeline solution that addresses APS-U Era beamline data needs.

Team: Tekin Bicer (DSL/CXS), Miaoqi Chu (SDM), Arthur Glowacki(SDM), Tejas Guruswamy (DET), John Hammonds (SDM), Steven Henke (SDM), Pete Jemian (BC), Keenan Lang (BC), Tom Naughton (IT), Hannah Parraga (SDM), Roger Sersted (IT), Hemant Sharma (CXS), Sinisa Veseli (SDM), David Wallis (IT), Mary Westbrook (IT), et al.
FY24 SDM Effort: SDM effort accounted for in Workflow & Data Management Tools and other projects
FY25 SDM Effort: SDM effort accounted for in Workflow & Data Management Tools and other projects
Out Years: TBD

This project is creating a template for APS beamlines and support groups to follow when deploying new instruments or data pipelines. The template is being validated first in a laboratory setting using a testbed, and then by the successful application to the APS-U Feature Beamlines and APS-U Enhancement Projects. This project is also identifying gaps in implementations and capabilities and suggesting action items needed to fill the identified gaps.

Goals include:
1. Continue cross-divisional collaboration to prepare data infrastructure for expected APS-U data rates and volumes.
2. Maintain productive collaboration with APS-U Feature Beamlines to get feedback on the proposed solution and ensure that beamline-specific concerns are addressed.
3. Continue adapting the template data pipeline for each of the APS-U Feature Beamlines.
4. Deploy production-ready data pipelines at each of the APS-U Feature Beamlines.
5. Increase the use of BDP tools and best practices throughout the facility.

Bragg Coherence Diffraction Imaging (CDI) Software

Summary: Provide ongoing support for the Bragg CDI reconstruction tools. Run these tools on ALCF supercomputers. Support the process of creating new algorithms in the toolset.

Team: Barbara Frosik (SDM), Ross Harder (MIC), et al.
FY24 SDM Effort: 1.00 FTE
FY25 SDM Effort: 1.00 FTE
Out Years: TBD

Attaining a robust image of a sample in a computation time nearer the data acquisition time will allow nearer real-time feedback to experiment parameters. The experimenter may then begin to perform guided, carefully executed experiments. Currently, most Bragg CDI users will benefit from near real-time phase retrieval for their data. It will also open the instrument up to far less sophisticated CDI users. This technique will be critical to one or more APS-U Era beamlines.

The cohere tool is a high-performance package that can complete phase retrieval of a typical image in less than 1 minute. With different cases this may vary, as data sizes may be large, or users may use different features based on data the reconstruction may benefit from, such as the genetic algorithm, or multiple

reconstructions. These tools integrate an AI start option for very fast phase retrieval. The ability to reconstruct multipeak data was recently added. The preprocessing phase integrates the AutoAlien algorithm that may be used to replace manual artifact removal. These tools provide a user-friendly GUI that is used for configuration and to control the reconstruction process. The tools run on both CPUs and GPUs. The tool can be installed with pypi package on all platforms: Linux, Windows, and macOS. Users have a choice of back-end libraries to use: numpy, cupy, and pytorch. The cohere tool supports a cluster configuration to provide more memory for the reconstruction operation. As data rates are growing due to APS-U and new detectors, there is a need to utilize the ALCF supercomputer to have access to greater hardware resources. The new capabilities are being added for diverse scenarios. The process of gathering data, routing to ALCF resources, and reconstruction will be developed.

Goals include:
1. Provide access to code and scheme to easily add new functionality and document the process.
2. Extend choice of libraries that can be utilized to run tools.
3. Provide and maintain documentation.
4. Collaborate with ALCF resources to enable running the tools on ALCF supercomputers.
5. Add new features to keep up with current research.

Coherent Surface Scattering Imaging (CSSI) Software

Summary: Implementation of high-performance CSSI and GISAXS software applications.

Team: Peco Myint (SDM), Miaoqi Chu (SDM), Ashish Tripathi (DYS/CXS), Zhang Jiang (DYS), Jin Wang (TRR), Joseph Strazalka (DYS), et al.
FY24 SDM Effort: 0.50 FTE
FY25 SDM Effort: 0.50 FTE
Out Years: TBD

CSSI is a coherent X-ray technique that can non-destructively probe 3D surface structures with very high resolution. A featured beamline for CSSI will be built as a part of the APS-U project. Due to the unique geometry of CSSI experiments (reflection at grazing incident angles), the scattering process involves strong dynamical scattering effects that are absent in the traditional transmission-type coherent diffractive imaging/ptychography. A collaborative team with physics, applied mathematics, and computer science backgrounds is working to identify the best algorithms to reconstruct CSSI data. Software support is required to implement the algorithms with high efficiency and maintainability.

Goals include:
1. Develop a data-processing GUI for users to view the diffractions, do Fresnel multislicing simulations to quickly compare with data, and do image reconstructions by preprocessing data for an already existing software such as ptychodus.
2. Develop a beamline control GUI for sample alignment and automating multi-hour or multi-day long ptychography or tomography scans in CSSI reflection geometry.
3. Explore the physical nature that distinguishes between the CSSI's surface scattering and transmission type scattering.
4. Develop physics-based forward simulation packages to generate high-quality data to test and benchmark reconstruction algorithms.
5. Implement reconstruction algorithms that account for the surface scattering geometry that can yield high-fidelity CSSI reconstructions. Optimize the algorithms on HPC to enable reconstruction on large input datasets. Explore the integration of CSSI reconstruction algorithms with Tike, a general-purpose ptychography reconstruction package.
6. Develop a modern data reduction and visualization package for the GIXS instrument, which shares the same beamline and hardware with CSSI.

Continuation of A Smart Data and Computational Fabric Pilot for the APS and APS-U

Summary: Develop and support data science portals for select APS beamlines and techniques.

Team: Ian Foster (CELS-DSL), Rachana Ananthakrishnan (Globus), Ryan Chard (Globus), Ben Blaiszik (CELS-DSL), Nick Saint (CELS-DSL), Rafael Vescovi (CELS-DSL), Nicholas Schwarz (SDM), Hannah Parraga (SDM), John Hammonds (SDM), Steven Henke (SDM), Sinisa Veseli (SDM), Junjing Deng (MIC), Olga Antipova (MIC), Si Chen (MIC), Lu Xi Li (MIC), Barry Lai (MIC), Peter Kenesei (MPE), Arthur Glowacki (SDM), Miaoqi Chu (SDM), Peco Myint (SDM), Suresh Narayanan (DYS), Jun-Sang Park (MPE), Hemant Sharma (CXS), et al.
FY23 SDM Effort: SDM effort accounted for in Workflow & Data Management Tools and other projects
FY24 SDM Effort: SDM effort accounted for in Workflow & Data Management Tools and other projects

Goals include:
1. Meet goals in the Continuation of A Smart Data and Computational Fabric Pilot for the APS and APS-U award.
2. Plan and develop new data portals that takes advantage of modern web technology such as plotly, to enable users to inspect the data anywhere with high interactivity.

Correlation Toolkit

Summary: Develop a real-time HPC-enabled set of tools for time-based correlation data analysis.

Team: Miaoqi Chu (SDM), Eric Dufresne (DYS), Jay Horwath (CXS), Suresh Narayanan (DYS), Qingteng Zhang (DYS), et al.
FY24 SDM Effort: 0.75 FTE
FY25 SDM Effort: 0.75 FTE
Out Years: TBD

Time-based correlations are an important analysis tool used to study the dynamic nature of complex materials. The recent development and application of higher-frequency detectors allows the investigation of faster dynamic processes enabling novel science in a wide range of areas resulting in the creation of greater amounts of image data that must be processed within the time it takes to collect the next data set to guide data collection. The increased brightness afforded by the APS-U project will compound this data processing challenge by producing data with higher count rates.

Goals include:
1. Continue supporting the current tools, fix bugs, improve performance, and add new features. The tools include:
    a. pysimplemask: GUI software to draw masks and make q-partitions.
    b. webgui: web-based tool that allows users the view the XPCS results in a web browser.
    c. boost-corr: high-performance GPU correlation toolkit.
    d. xpcs-viewer: full-fledged GUI tool to view XPCS results in an interactive way.
2. Perform R&D for new streaming-based XPCS correlation and visualization algorithms.
3. Work with collaborators to deploy the tools to the Globus Science Portal and the Data Management platform.
4. Collaborate with the XPCS user community and APS beamline staff to research and implement novel correlation functions such as higher-order correlations and angular correlations.
5. Extend the current tools, which are mostly designed for small-angle XPCS, to wide-angle XPCS. Provide support to other beamlines that will perform XPCS measurements in the APS-U Era, such as 9-ID and 12-ID.
6. Explore novel algorithms in the AI/ML domain to analyze XPCS datasets that will increase significantly in size in the APS-U Era.

General-Purpose Reciprocal-Space Mapping (RSM) Tools

Summary: Continue development and deployment of high-performance RSM tools.

Team: John Hammonds (SDM), Hannah Parraga (SDM), Michael Prince (SDM), Jonathan Tischler (SSM), Zhan Zhang (SSM), et al.
FY24 SDM Effort: 0.25 FTE
FY25 SDM Effort: 0.25 FTE
Out Years: TBD

This project aims to continue development of a general-purpose tool for reciprocal-space mapping at the APS. The tool allows users to examine a volume of data and select portions on which to apply transformations that convert detector pixel locations from diffractometer geometry to reciprocal-space units, and then map pixel data on to a 3D reciprocal-space grid. It can map data acquired using 4- and 6-circle diffractometers, and with scans taken over angles or energy, and can operate via a graphical user interface, or in batch processing mode. Data too big to fit entirely into memory at one time is processed in smaller chunks and reassembled to form the final output volume, allowing users to process arbitrarily large input datasets.

This tool has the potential to serve an even larger number of APS beamlines, and will be critical to a number of APS-U beamlines and high-energy diffraction experiments. It is currently in regular use for scattering and diffraction experiments at the 33-BM and 33-ID beamlines, for micro-diffraction analysis at 34-ID, and for time-resolved diffraction work at 7-ID. Development is underway for WA-XPCS analysis at 8-ID, and for data exploration with inelastic x-ray measurements at 30-ID. Fast tools for reciprocal-space mapping using distributed computing resources are needed to make nearer real-time decisions regarding the next set of data that is collected. This work will leverage effort related to workflow and management tools, and data streaming and analysis tools for real-time analysis.

Goals include:
1. Continue supporting the current tool and adding new features as requested.
2. Further parallelize tools to greatly improve performance.
3. Develop automated file-based and streaming-based workflows for processing RSM data locally and at the ALCF on Polaris.

In-operando AC Scattering Software

Summary: Continue development of software for in-operando AC scattering experiments.
Proposed Team: Research Aide Technical (SDM), Peco Myint (SDM), John Hammonds (SDM), Philip Ryan (MM), et al.
FY24 SDM Effort: 1 Research Aide Technical, 0.2 FTE
FY25 SDM Effort: 1 Research Aide Technical, 0.2 FTE
Out Years: TBD

Goals include:
1. Continue development of the xPlotUtil in-operando AC scattering software.
2. Continue space-mapping GUI development.
3. Integrate real-time processing at the beamline and Jupyter notebooks.

Laue Diffraction

Summary: Develop high-performance computing tool kit for the new Laue depth reconstruction algorithm.

Team: Hannah Parraga (SDM), Michael Prince (SDM), Doga Gursoy (CXS), Dina Sheyfer (SSM), Jon Tischler (SSM), et al.
FY24 SDM Effort: 0.75 FTE
FY25 SDM Effort: 0.75 FTE
Out Years: TBD

Processing Laue micro- and nano-diffraction microscopy data generally consists of three main steps in the following order: depth reconstruction, peak searching and indexing, and q-space histogram generation. The depth reconstruction process generates new images corresponding to the scattering observed from a single depth. Peak searching and indexing finds all the peaks and indexes them to get the crystal orientation of a Laue pattern. With energy scans, a 1D or 3D histogram of intensity in q-space may also be generated.

In order to improve data collection time in the APS-U era, the APS is developing coded-aperture scans that may replace the current wire-scans for obtaining depth reconstructions. Along with higher-performance implementations of the contemporary wire scan mode data, high-performance implementations of coded-aperture reconstruction methods are being developed.

Goals include:
1. Construct frontend to allow scientists to launch and manage reprocessing workflows on HPC resources.
2. Develop indexing workflow and integrate into user-facing reprocessing UI.
3. Refactor real-time workflows for robustness and that reprocess data.
4. Continue optimization of reconstruction algorithm to reduce resource usage and decrease turnaround time.

LLMs for APS Scientific Processes

Summary: Investigate and prototype the use of LLMs in APS scientific processes, workflows, and experiment automation.

Team: Mathew Cherukara (CXS), Michael Prince (SDM), Henry Chan (CNM), Varuni Katti (NST), Tao Zhou (NST), Aikaterini Vriza (ALCF), Matthew Dearing (BIS), et al.
FY24 SDM Effort: 0.15 FTE
FY25 SDM Effort: 0.15 FTE
Out Years: TBD

Over the last year the field of large language models (LLMs) has developed rapidly. Many of the capabilities of these models including knowledge extraction, summarization, and agentic tool use may be well suited for scientific workflows and automation at the APS. This project intends to explore the use of LLMs at the APS and to develop prototype deployments of such applications. A prototype has already been developed with access to documentation from the APS, CNM and ALCF user facilities. Additionally, a simple agent-based system has been developed to control diffractometers through SPEC. Model deployment is handled via a collaboration with BIS and fine-tuning may be possible with resources at ALCF.

Goals include:
1. Productionize and deploy current application for internal use at the APS.
2. Explore additional capabilities and LLM prompting techniques as they are developed.
3. Experiment with agentic and tool-use workflows for control and automation of experiments.
4. Integrate with Bluesky.

MIDAS Tomography Frontend

Summary: Refactor TomoPy user-facing GUI to work with MIDAS and improve post-processing capabilities.

Team: Peter Kenesei (MPE), Michael Prince (SDM), Hemant Sharma (CXS), Austin Yunker (DSL), et al.
FY24 SDM Effort: 0.15 FTE
FY25 SDM Effort: 0.15 FTE
Out Years: TBD

1-ID conducts tomography to examine samples and inform the analysis of far-field HEDM experiments. To improve ease-of-use and encourage more users to analyze their data during an experiment, a prototype tomography UI was constructed with TomoPy and the ASTRA toolbox for CUDA acceleration. MIDAS's reconstruction backend is better tuned for the needs of 1-ID. This project intends to expand and productionize this UI to incorporate MIDAS reconstructions and develop features to better accommodate the needs of HEDM experiments. The first of which would be a post-processing pipeline to denoise and segment reconstructions. Once software is more productized, it may be considered for use in other beamlines that perform tomography scans.

Goals include:
1. Integrate MIDAS into user-facing tomography UI.
2. Add post-processing segmentation pipeline to identify voids and cracks in samples.
3. Productionize software including documentation, extensibility, and setup processes.
4. Continue to incorporate feature requests and feedback from the 1-ID beamline.

Ptychography Software

Summary: Continue development of HPC ptychography reconstruction software.

Team: Steven Henke (SDM), Mathew Cherukara (CXS), Daniel Ching (CXS), Ash Tripathi (CXS), Peco Myint (SDM), Junjing Deng (MIC), Stefan Vogt (XSD-ADMIN), et al.
FY24 SDM Effort: 0.50 FTE
FY25 SDM Effort: 0.50 FTE
Out Years: TBD

Ptychography is one of the exemplar APS-U enabled techniques and will be one of the largest data producing techniques post APS-U. Proper support and development of existing tools, including enhancements to support additional experiment geometries, support for reconstructing thicker samples, and the development of workflows that improve reconstruction quality, robustness, and efficiency are needed. Provide support and new feature development for the Tike ptychography reconstruction library, PtychoNN neural network, and Ptychodus analysis application.

Goals include:
1. Continue supporting the Tike ptychography reconstruction library.
2. Improve Tike reconstruction quality via innovative reconstruction techniques as well as systematic comparisons with other ptychography tools.
3. Incorporate additional physics into PtychoNN to improve inference quality and generalizability
4. Continue developing the Ptychodus ptychography analysis application.
5. Add support for multislice ptychography to enable reconstructing thick samples
6. Add support for reflection geometry ptychography (e.g., Bragg Ptychography and Coherent Surface Scattering Imaging)

7. Implement reconstruction parameter optimization workflow
8. Grow a user base of APS ptychography tools by improving software quality and maintaining productive collaborations with beamline scientists.

Ptychography+XRF Software

Summary: Develop software that uses ptychography results to enhance XRF data products.

Team: Steven Henke (SDM), Arthur Glowacki (SDM), Junjing Deng (MIC), Wendy Di (MCS), Olga Antipova (MIC), Si Chen (MIC), Luxi Li (MIC), Tim Mooney (BC), Sarah Wieghold (MIC), et al.
FY24 SDM Effort: 0.25 FTE
FY25 SDM Effort: 0.25 FTE
Out Years: TBD

Ptychography and X-Ray Fluorescence (XRF) are complementary experimental techniques that can be simultaneously employed for structural and elemental sample characterization. This software, called PtychoXRF, uses the estimated probe and corrected scan positions from the ptychographic reconstruction to computationally enhance the resolution of XRF data products. Input and output data formats are compatible with existing APS software, including Ptychodus, XRF-Maps, uProbeX, TomoCuPy, and TomoPy so PtychoXRF can be seamlessly inserted into existing processing toolchains.

Goals include:
1. Implement state-of-the-art XRF resolution enhancement algorithms to ensure that APS scientists have access to the best available tools.
2. Develop a user interface that facilitates user-guided data processing.

X-ray Fluorescence (XRF) Microscopy Software

Summary: High-performance computing (HPC) enabled fitting library and tools for fast elemental mapping of x-ray fluorescence microscopy software.

Team: Arthur Glowacki (SDM), Olga Antipova (MIC), Si Chen (MIC), Wendy Di (MCS/CXS), Barry Lai (MIC), Lu Xi Li (MIC), Yi Jang (MIC), Stefan Vogt (XSD/ADM), at al.
FY22 SDM Effort: 0.90 FTE
FY23 SDM Effort: 0.90 FTE
Out Years: TBD

XRF imaging typically involves the creation and analysis of 3D data sets, where at each scan position the full spectrum is recorded. This allows one to later process the data in a variety of different approaches, e.g., by spectral region-of-interest (ROI) summation with or without background subtraction, principal component analysis, or fitting. Additionally, it is possible to sum up the per pixel spectra over selected spatial ROIs to improve the photon statistics in such a spectrum.

The XRF microscopy technique is a staple technique that will be used by many APS-U beamlines in combination with other x-ray acquisition modalities, such as fluorescence tomography and fluorescence ptychography. The increase in intensity and smaller spot size due to benefits of the APS-U will drastically increase data size and data rates for this technique. To facilitate real-time data analysis and fast feedback for experiment steering, HPC-enabled implementations of common elemental mapping algorithms and data I/O schemes that facilitate streaming data, and appropriate user interfaces are required. This work will leverage effort related to workflow and management tools, and data streaming and analysis tools for real-time analysis, and can serve in conjunction with existing tomography software to provide analysis code for fluorescence tomographic reconstructions.

Goals include:

1. Continue developing a graphical user interface for XRF-Maps. Adding new features requested by beamline scientist and end users.
2. Integrate a workflow for XRF-Tomography data analysis.
3. Implement/Integrate new alignment algorithms for XRF-Tomography datasets. .
4. Continue implementing software for APS-U era needs and extend data analysis workflows to be processed on ALCF and other supercomputing resources.

A Collaborative Machine Learning Platform for Scientific Discovery 2.0

Summary: Support BES AI/ML for SUFs award: A Collaborative Machine Learning Platform for Scientific Discovery 2.0.

Team: Postdoc (SDM), Nicholas Schwarz (SDM), et al.
FY24 SDM Effort: 1.00 postdoc
FY25 SDM Effort: 1.00 postdoc

Due to the tremendous increase in data generated by the scientific user facilities, building a software ecosystem that utilizes the power of machine learning technology is a logical step moving forward. The goals of the platform include, but not limited to, data labeling to automatically extract features, build inferences with experiments through similarities/differences, developed search capabilities to help users navigate through their experiments, etc. The platform will be equipped with ML algorithms, trained ML models, and standard tagged/labeled data by the world domain experts to aid users with low barriers of entry across all the DOE SUF facilities.

Goals include:

1. Develop classification capabilities for distinguishing different types of data. For example, transmission and reflection data can be distinguishable in scattering experiments.
2. Develop feature extraction from multi-dimensional data.
3. Place previous work into operational use.
4. Construct a foundational platform carrying the capabilities mentioned in 1 and 2, enabling low barrier of entry to access by users. In the platform, users can use the pre-trained models, make contribution to the platform, and provide feedback regarding the model or data integrities. Do this by integrating the platform for use at the APS and integrate ML approaches, for techniques such as ptychography, into the system.

A Pipeline for Autonomous Comprehensive Reduction and Analysis of Sequential 2-Dimensional Diffraction and Scattering Data

Summary: Support a pipeline for autonomous comprehensive reduction and analysis of sequential 2-Dimensional diffraction and scattering data LDRD.
Team: Wenqian Xu (SRS), James Weng (SRS), Brian Toby (CXS), Nicholas Schwarz (SDM), et al.
FY24 SDM Effort: TBD
FY25 SDM Effort: TBD
Out Years: TBD

Auto Parameter Tuning for XRF

Summary: Application and/or development of a supervised learning neural network model and train it on collected XRF datasets. Auto Parameter Tuning for XRF LDRD.

Team: Postdoc (SDM), Arthur Glowacki (SDM), Yi Jian (MIC), Wendy Di (CXS), Hannah Parraga (SDM), et al.

FY24 SDM Effort: 1.00 postdoc, 0.10 FTE
FY25 SDM Effort: 1.00 postdoc, 0.10 FTE
Out Years: TBD

Create a supervised learning neural network for XRF parameter tuning. Train the network on existing datasets.

Goals include:
1. Generate training datasets.
2. Generate and evaluate different NN models.

LYNX

Summary: Develop workflows and analyze experiment data for the LYNX ptychography-laminography instrument.

Team: Steven Henke (SDM), Hanna Ruth (SDM), et al.
FY24 SDM Effort: 1.50 FTE
FY25 SDM Effort: 1.50 FTE
Out Years: TBD

Simulation-Guided High-Energy X-ray Diffraction Experiment Framework for Understanding Irradiation-Induced Slip Anisotropy in Structural Metals

Summary: Support simulation-guided high-energy x-ray diffraction experiment framework for understanding irradiation-induced slip anisotropy in structural metals LDRD.

Team: Jun-Sang Park (MPE), Tejas Guruswamy (DET), Stephan Hruszkewycz (MSD), Mark Messner (AMD), Siniša Veseli (SDM), Xuan Zhang (NSE)
FY24 SDM Effort: TBD
FY25 SDM Effort: TBD
Out Years: TBD

This project will enable seamless in situ characterization of bulk polycrystalline materials covering macroscopic to nanoscopic scales through an integrated multi-modal imaging framework. This Coherence Enhanced High Energy Diffraction Microscopy (CE-HEDM) framework will enhance the resolution of conventional HEDM technique by exploiting the post APS-U x-ray coherence and brilliance and integrating High-Energy BCDI as a zoom-in companion to HEDM. Crystal plasticity finite element simulation capabilities will be an integral part of the framework to identify critical grains in near real-time and map them by CE-HEDM. The new imaging platform will be applied to study irradiation-enhanced plasticity anisotropy in stainless steels.

Real-time Feedback & Data Acquisition System for APS-U Accelerator

Summary: Software framework and tools for the collection of data used for controls, statistics, and diagnostics of technical systems for the MBA accelerator.

Proposed Team: Sinisa Veseli (PSC/SDM), Guobao Shen (CTL), et al.
FY24 SDM Effort: 0.16 FTE (Oct - 0.50 FTE, Nov - 0.50 FTE, Dec - 0.50 FTE, Jan - 0.30 FTE, Feb - 0.05 FTE, Mar - 0.05 FTE, 0.00 - FTE for remainder of FY24)

The real-time feedback and data acquisition (RTFB/DAQ) system is a software framework and associated tools that enable fast data collection for controls, statistics, and diagnostics associated with the state-of-the-art embedded controllers utilized by the APS-U project MBA-based accelerator design. The DAQ software interfaces with several technical subsystems to provide time-correlated and synchronously sampled data that can be used for commissioning, troubleshooting, performance monitoring, and early fault detection. The key features of the system include capability to acquire data from multiple subsystems at various sample rates, support for continuous data acquisition, and the ability to route data to any number of applications. Future work will focus on developing high-level applications and utilities for processing of DAQ data, both in real-time (streaming) and batch (file-based) mode.

Multi-modal Diffraction Tomography

Summary: Develop robust near real-time software for diffraction tomography.

Proposed Team: TBD SDM Member, et al.
FY24 SDM Effort: 0.00 FTE**
FY25 SDM Effort: 0.00 FTE**
Out Years: TBD

Algorithmic work is currently underway in the XSD-CXS group. Performance and engineering work will commence once algorithmic proof-of-concept is complete.