Scientific Python for Data Acquisition, Management, and Analysis

Organizers: Daniel Allan (BNL-NSLS-II), Andi Barbour (BNL-NSLS-II), Stuart Campbell (BNL-NSLS-II), Karen Chen-Wiegart (BNL-NSLSL-II), Abigail Giles (BNL-NSLS-II), Max Rakitin (BNL-NSLS-II)

A growing international community of scientific Python users and programmers is building shared open-source software tools for synchrotron science, under the umbrella of the Bluesky Project. This workshop is a hands-on tutorial on using scientific Python to enable more sophisticated data acquisition, access, and analysis during and after an experiment at NSLS-II.

The workshop will be structured around each of the core Bluesky projects and some notable applications built upon them. Sessions will cover the full data lifecycle of acquisition, management, and analysis. This will highlight advancements and opportunities in adaptive (or self-driving) experiments and integrating simulations with experiments. Data management topics will explore the questions of how to search, access, and publish data. Lastly, we will interactively demonstrate data analysis via integration with the larger scientific computing ecosystem, including visualization and parallel processing. These tools can be applied to streamline traditional data analysis techniques and to enable emerging, data-driven methods including artificial intelligence and machine learning (AI/ML).

Speakers will include staff from the NSLS-II Beamline Programs, staff from the NSLS-II Data Science and Systems Integration Program, and their counterparts at peer facilities using related software tools. Sessions will incorporate both step-by-step interactive tutorials and demonstrations of more advanced, specific applications.

Basic familiarity with Python usage will be assumed. Please bring a laptop; this is an interactive session.

Start Time (ET)	Title	Speaker (Affiliation)
9:30 – 10:00 a.m.	Welcome and Introduction	Stuart Campbell, BNL
10:00 – 10:50 a.m.	Interactive Tutorial: Do's and Don'ts with Bluesky	Andi Barbour and Josh Lynch
10:50 – 11:00 a.m.	Break	
11:00 – 12:00 p.m.	Interactive Tutorial: When should I use numpy, pandas, or xarray?	Karen Chen-Wiegart and Dan Allan
12:00 – 1:00 p.m.	Break	
1:00 – 1:30 p.m.	Demo (TBD)	TBD
1:30 – 1:45 p.m.	Demo: Data Workflows with Prefect	Abby Giles and Dylan McReynolds

1:45 – 2:15 p.m.	State of the Bluesky Software Stack	Bluesky Contributors
2:15 – 2:30 p.m.	Vendor Talk: Radiasoft	Boaz Nash and Nathan Cook
2:30 – 2:40 p.m.	Break	
2:40 – 3:30 p.m.	Lightning Talks	various
3:30 – 4:00 p.m.	Discussion, Q&A	
4:00 p.m.	Adjourns	