

Tuesday Evening Poster Sessions, October 31, 2017

Scanning Probe Microscopy Focus Topic

Room Central Hall - Session SP-TuP

Scanning Probe Microscopy Poster Session

SP-TuP-1 Pycroscopy – A Community-Driven Software Package for Analyzing Microscopy Data, *S Somnath, Chris Smith, S Jesse, R Vasudevan, N Laanait*, Oak Ridge National Laboratory

Microscopy and material science are undergoing profound changes, driven by experimental datasets that are rapidly growing in dimensionality and size, increased accessibility to high-performance computing (HPC) resources, and more sophisticated computer algorithms than ever before. These changes are most pronounced in the functional imaging of materials. However, the softwares supplied with instruments such as microscopes are typically very expensive, do not provide access to advanced or user-defined data analysis routines, and store data in proprietary formats. Furthermore, these proprietary software and data formats not only impede data analysis but also hinder continued research and instrument development, especially in the age of “big data”. Therefore, moving to the forefront of data-intensive materials research requires general and unified data curation and analysis platforms that are community driven and HPC-ready.

We have developed a platform called Pycroscopy that uses community-driven approaches for analyzing and storing data. Pycroscopy is freely available via popular software repositories, and therefore lifts any financial burden for handling data. Pycroscopy uses an intuitive data structure stores data in and hierarchical data format (HDF) files that can be interrogated using any programming language, scales well from kilobyte to terabyte sized datasets, and can readily be used in HPC environments unlike proprietary data formats. More crucially, Pycroscopy uses a universal data format that is curation-ready and therefore both meets the guidelines for data sharing issued to federally funded agencies and satisfies the implementation of digital data management as outlined by the United States Department of Energy. This instrument-independent data format has also greatly simplified the correlation of data acquired from multiple instruments, necessary for comprehensive studies of materials. Unlike many other open-source packages that focus on analytical or processing routines specific to an instrument, the general definition of the Pycroscopy data format can be readily adopted for different microscopy techniques. Furthermore, the generality of Pycroscopy provides material scientists access to a vast and growing library of community-driven data processing and analysis routines that far exceed those provided by instrument manufacturers and are desperately needed in the age of big data. In summary, Pycroscopy can greatly accelerate materials research and discovery through the realms of big, deep, and smart data.

Author Index

Bold page numbers indicate presenter

— J —

Jesse, S: SP-TuP-1, 1

— L —

Laanait, N: SP-TuP-1, 1

— S —

Smith, C: SP-TuP-1, **1**

Somnath, S: SP-TuP-1, 1

— V —

Vasudevan, R: SP-TuP-1, 1