

Announcing the ACA Best Practices for Data Analysis & Archiving Scientific Interest Group

John P. Rose¹, Suzanna Ward² and Nicholas Sauter³, ¹Department of Biochemistry & Molecular Biology, University of Georgia, Athens, GA 30602 USA, ²Cambridge Crystallographic Data Centre, Cambridge, CB21EZ UK and ³Molecular Biophysics & Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, MS 33-345, Berkeley, CA 94720 USA.

Data, along with its generation, transmission, immediate storage, analysis, documentation, reproducibility, and long-term curation, has been identified as a key topic of interest by methods developers, scientific end users and funding sources. The current generation detectors used in X-ray diffraction and CryoEM can produce data at amazing rates, leading to data sets of considerable size. These large data sets are taxing both Facility and user resources for data transport, storage and analysis. Another area facing similar challenges is in the use of Hybrid Methods (combining data from a number of different experimental techniques) to answer complex scientific questions. Again, there is strong community interest in, not only preserving the raw data that the analysis was built on but also preserving all aspects as to how the data were combined and analyzed. The challenges of “Big Data” are not restricted to macromolecules. They are common to other techniques such as material science and complex small molecule systems.

The American Crystallographic Association recognizing the challenges outlined above has recently established a new Scientific Interest Group - Best Practices for Data Analysis & Archiving. The SIG hopes to host a one-day Session next year at the ACA meeting in Covington, Kentucky. An introduction to the SIG and tentative plans for next year’s Session will be presented.