**MS053.O03**

*Treatment of x-ray diffraction data at diamond light source*

James Parkhurst[1], Graeme Winter[1], Richard Gildea[1], Markus Gerstel[1], Karl Levik[1], Irakli Sikharulidze[1], Dave Hall[1], Katherine McAuley[1], Gwyndaf Evans[1], Alun Ashton[1]
[1]*Diamond Light Source, Didcot, United Kingdom*
E-mail: james.parkhurst@diamond.ac.uk

In any experimental discipline, raw data represents the source from which all discoveries are derived. A more strict interpretation in X-ray diffraction experiments may refer to this as primary data since any pixel counts will have been manipulated (e.g. analogue to digital conversion, dark current correction, interpolation of pixels etc.); however the fundamental idea remains: this is the closest it is possible to get to the original experimental measurements.

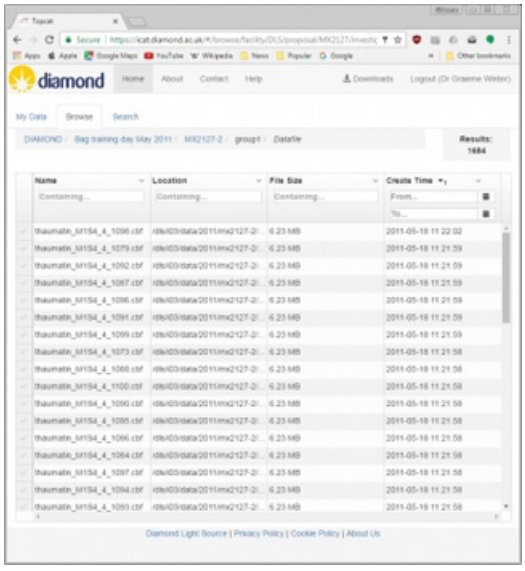
At Diamond Light Source, the principle of secure recording and storage of the primary data was embedded in the data acquisition system from the outset. The general user does not have permission to alter or delete the raw experimental data, and the acquisition system GDA is designed to prevent over-writing of the images. Furthermore, every data file recorded is archived to tape at the ATLAS data store to allow recovery of data from up to a decade ago [Figure 1].

The acquisition of raw data alone is no longer satisfactory for a modern synchrotron facility. Diamond Light Source has pushed the use of automated processing from its inception (Winter & McAuley 2011) and continues to focus on developing automated analysis methods to improve the use of beam time. Analysis of individual diffraction images is performed to show trends observed during the experiment and strategy calculations are used to to guide the experimenter during the data collection. Fast automated data processing is then be used to deliver merging statistics and scaled intensities within a minute of the end of data collection (Winter & McAuley 2011, Winter et al. 2013, Waterman et al. 2013). The results of this processing are then pushed to automated downstream analysis that can provide the user with more experimentally relevant results such as whether a ligand is present or whether the data are suitable experimental phasing.

Occasionally, automated data processing will fail and manual interaction will be needed to successfully process the data. With the shift to remote access data collection and short shifts, many users are now remotely reprocessing their data on Diamond computers using the NX software, to provide a remote desktop. This, combined with cluster based HPC, has to some extent offset the challenges resulting from the deployment of pixel array detectors at all single crystal X-ray diffraction beamlines.

This presentation will review the details of the facilities described above and will consider the likely impact of the next generation of pixel array detectors, the requirements for storing data collected at an ever increasing rate, the use of standard file formats (e.g. HDF5, Nexus), and whether we can continue the same conservative treatment of primary data at Diamond Light Source.

[1] Winter, G. and McAuley, K. E. (2011). Methods. Elsevier Inc., 55(1), 81–93.
[2] Winter, G. et al. (2013). Acta crystallographica. Section D, Biological crystallography, 69(7), 1260–73.
[3] Waterman, D. G. et al. (2013). CCP4 newsletter on protein crystallography, 49, 16–19.




**Keywords:** archiving, automatic processing, remote processing