Issues For Other Software Developers
Alun Ashton
Computing and Data Acquisition
Group www.diamond.ac.uk

Scope
.... issues that are interesting to those who write software that might follow data collection and reduction....
.... the essence of it is the data, not the format...
.... at the end of data reduction there should be an mmCIF with all useful meta data in addition to the "real" data of F's and sigmas.... during the experiment and in the use of "dna," [or equivalent] we will hold useful metadata within a database system...
.... what data will you want, and in what form would you like to have them?

Format?
• XML needs defining to be useful.
• mm/imgCIF is at least a recognised standard.

• You don't want EVERY file to carry all the info - ideally there would be a data base for each structure, but realistically, people like to keep some of the essential info on hand.

issues that are interesting to those who write software that might follow data collection and reduction......
• Can we assume (correct) data collection starts at the bioinformatics stage and not at the collecting diffraction images stage?
• Can we assume that a persistent metadata catalogue will always be available, associated and 'with' the data?

The last word.
• Anything as long as its correct.
• After much though and deliberation we have come to the conclusion that the most important thing that could be included would be the phases. They would make life a lot easier.
NeXus

- Based on a well established and supported format - HDF
- Allows storing of data, sample, environment, information together
- There is an experienced user community
- A common data format for synchrotron, neutron and muon data (and from other sources)
- The format must worry about the implementation and support details

diamond: www.diamond.ac.uk
DNA: www.dna.ac.uk
eHTPX: www.e-htpx.ac.uk
GDA: www.gda.ac.uk
ICAT: Information CATalogue
NeXus: www.nexus.anl.gov/index.html
SRB: Storage Resource Broker
www.sdsc.edu/srb/index.php/Main_Page

ImgCIF at Diamond

- CCD detectors that produce ImgCIF specified at time of tender.
- ADSC Q315.
- Chris Neilson and Harry Powell with help from others producing a working version.
- Diamond wanted images in an open standard
- Other than ImgCIF it is intended that NeXus will be the file format for all DLS beamlines.

Phase I (2007)

- 3 MX (0.9 – 2.5 Å optimised for 0.98 Å) with double crystal monochromator: Kirkpatrick-Baez horizontal and vertical focusing mirrors; Focal spot size ~94 μm (f/3) x 1.7 mm (y) (FWHM); estimated flux at 12.6 keV: 3.5 x 10^{12} ph/s; fully automated sample handler; cryo cooling; CCD detector
- One station will have containment three facility for pathogenic samples

Phase II

- Microfocus beam line
- Fixed wavelength side station (0.96 Å) (MR & ligand binding studies)
- Long wavelength side station for Sulphur anomalous (1.5 – 2.5 Å)
Software for all DLS beamlines