ISPyB Collaboration Meeting 4th June 2018

Next meeting 2nd July 2018 (14:00 UK, 15:00 European Time)

Contents:

Participants

Organisation	Name
Diamond Light Source (DLS)	
	Neil Smith (NS)
European Synchrotron Radiation Facility (ESRF)	
	Solange Delageniere (SD)
	Stephanie Malbet-Monaco (SM)
	Alex De Maria (AM)
	Olof Svenssson (OS)
Global Phasing	
	Rasmus Fogh (RF)
	Gerard Bricogne (GB)
	Claus Flensburg (CF)
	Clemens Vonrhein (CV)
EMBL Hamburg	
	Ivars Karpics (IK)

Updates from Members

DLS: Work is ongoing to develop ISPyB web services for the SynchLink mobile application. A Swagger/OpenAPI specification is being developed – generated from the ISPyB REST annotations. DLS will be attending ESRF to discuss shipping API to support INSTRUCT project with University of Oulu.

ESRF: No longer in shutdown. Added links and objects to BLSubSample and BLSampleImage tables to link subsamples to images, in preparation for entering plate data.

GPhL: Questions about added fields and documentation, as given in circulated email.

Hamburg: Have made the first attempt to connect Hamburg data collection with a Grenoble remote CRIMS database.

Conceptual modelling, Documentation and Data model

Database tables

NS mentioned that KL has made some progress on documenting the MX tables. A "work in progress" page is available here: https://github.com/DiamondLightSource/ispyb-database/wiki

Also a script for generating documentation from database tables and columns (using inline comments) is available here:
https://github.com/DiamondLightSource/ispyb-database/blob/master/doc/list_of_tables_and_columns.rst

Database definitions

CV asked whether there was an authoritative reference to the definitions and formulae that described what information should be entered for each database field. The answer was that the database fields had been created as a push from the various processing pipelines, and that there was no central set of definitions. It was agreed that such a set of definitions would be very useful, and how they might be calculated differently in different programs. It was suggested that Global Phasing was probably in the best position to produce such a set of definitions.

Database modelling issues, ALL

STARANISO

Prior to the meeting RF circulated a proposal describing the value of storing anisotropic diffraction parameters. Notably because anisotropic diffraction limits are now available through GPhL's STARANISO program.

GB cited published discussions on anisotropic diffraction; highlighting the challenge that reusing existing database slots for the new parameters will appear misleading and cause trouble with reviewers. The STARANISO server has now been running for 2.5 years, and users are increasingly asking for the corresponding output data to be available, so that they can be incorporated in the internal databases of pharma companies, and to be added to the standard statistics stored in the PDB.

SM thinks it is a good idea to store these data, which may well become generally used in the future, but notes that GPhL software is the only pipeline producing these data now.

OS described the necessary work into two parts. In order to store the data it is necessary to create the relevant fields in the database, gather the information, and populate the ISPyB database. This requires relatively little effort, especially since GPhL is already providing the

information in their ISPyB input xml file. In order to make the data visible to normal synchrotron users, it is also necessary also to modify the user interface displays, which is more demanding and harder to prioritise. Especially, as SM notes, ESRF is in the middle of a major move to a new display interface (EXI).

The data are currently available to users in the summary and log file output from GPhL software. GPhL's view is that this is not additional data, but a fundamentally different and better way of evaluating data quality, and as such should be presented to the user up front, on the ISPyB web page.

NS adds that DLS currently working on the autoprocessing tool display, and that he has forwarded the point on anisotropic diffraction to DLS domain experts for comments. GB comments that DLS has asked for a license to integrate STARNISO in their own pipelines, so presumably is interested in displaying these data to users.

In a more detailed discussion (at the start of the meeting) KL (via SN) reports that the form of the proposed column names, specifically for the ellipsoid axis vectors and their lengths is technically appropriate and consistent with normal ISPyB practice (insofar as this is defined).

KL also raised the point whether the isAnomalous Boolean and the completeness/anomalousCompleteness were not mutually redundant. To which CV answered that the isAnomalous switch was obsolete for modern practices, and that both completeness and anomalous completeness were normally both calculated within each single processing job.

In conclusion, capture of the anisotropic diffraction data was viewed favourably. The database changes should be raised as an issue on Github (ACTION: RF (subsequently completed)). The inclusion in the ISPyB data presentation was generally seen positively – but priorities and resource constraints at large scale facilities are an ongoing issue. Global Phasing offered to contribute to both efforts as much as possible.

SOAP vs REST for EM

NS queried whether the ESRF EM interface used only REST or SOAP protocols. At ESRF the SOAP service is used to push data into ISPyB while the REST service is used to support the user interface.

Any Other Business

IK raised interest in the subject of serial crystallography. From a follow up e-mail, it would be a good topic for the next face to face meeting at Elettra, Trieste.

Date of Next Meeting

The next meeting is planned for 2nd July 2018.