

## ISPyB Collaboration Meeting 26 March 2018

Next meeting provisionally 3<sup>rd</sup> May 2018 (14:00 UK, 15:00 European Time)

### Contents:

- Participants
- Agenda items
  - Updates from members
  - DLS Web services for mobile application
  - Meta data for plates
  - Conceptual modelling, documentation and data model
  - Database modelling issues
  - Any other business
- Date of Next Meeting

### Participants

Organisation	Name
Diamond Light Source (DLS)	
	Neil Smith (NS)
	Karl Levik (KL)
European Synchrotron Radiation Facility (ESRF)	
	Solange Delageniere (SD)
	Stephanie Malbet-Monaco (SM)
	Alex De Maria (AM)
Global Phasing	
	Rasmus Fogh (RF)
	Gerard Bricogne (GB)
	Peter Keller (PK)

### Updates from Members

- RF raised the point of making a working group to consider a more general reorganisation of the database, as mentioned in the minutes from the Diamond ISPyB meeting. It was agreed that this would be desirable, but that an incremental approach was probably better, and that the time to look at it would be after the choice of modelling tools and the generation of the logical model.
- SD reported that the ISPyB group at the ESRF was mainly working on changes needed for the reimbursement of Dewars at the moment
- NS covered updates from DLS in the next section

### DLS Web services for Mobile Application

Diamond have employed a contractor to help with migration of web services. The aim is to migrate a bespoke program developed to support an iOS application to use the ISPyB REST web services instead. The initial task is focused on documenting the services and endpoints required using swagger definitions. This will help future testing activities.

## **Meta data for plates, DLS**

At Diamond the VMXi beamline performs in-situ data collection using crystallization plates. NS presented some slides to show the meta-data and tables that have been created to support these experiments. Conceptually a plate is a Container (similar to a puck), with each drop being a location represented as a BLSample. Points or regions within the sample are represented as BLSubSample and are used to identify where to shoot the sample.

Data from standard industry library of screens have been captured and imported into the ISPyB database at DLS. The modelling is based around tables for Screens, ScreenComponents, and ScreeningComponentGroups. A 'component' referred to in the tables points to 'protein'.

NS agreed to add more context and explanation to the slides to describe the new information stored in ISPyB.

A discussion about the differences between plates, pucks and multi-pins took place where we tried to clarify the conceptual model behind containers and their relationship to samples. Some confusion may arise from the fact that pins are not treated as containers – rather the pucks that hold the pins are. This is an area where the conceptual modelling activity might help clarify terms used within the database.

## **Conceptual modelling, Documentation and Data model, ESRF**

No detailed information presented on this item (due to workload), however initial impressions are that Draw.io seems to be a better tool for use. The diagrams are not updated yet, but AM will try to prepare them for the next monthly meeting.

It was noted that the logical model is quite different from the physical model (in some areas most of the names seem to differ, for one thing) so it may be necessary to have not only two separate models, but a separate document to take care of the mapping.

After discussion it is agreed that the mapping is probably better done as annotations on the physical layer model. A physical model separate from the logical model is necessary for performance reasons, but in principle the logical model probably ought to come before the physical model.

## **Database modelling issues, ALL**

### *Refactoring*

GB raised a general issue relating to the need for refactoring: ISPyB had grown from being used originally for sample and user tracking into a much more comprehensive system, but there were discontinuities in the data collected that meant that important information present e.g. at crystallisation time (some data?) was not passed on through ISPyB but had to be passed directly to the beamline control system (e.g. MXCuBE).

It was recognised that ISPyB is a critical piece of infrastructure for many facilities, therefore the priority will always be towards users and functional improvements and a lower priority for the kind of refactoring/integration that would make the code reusable and the system maintainable and extensible in the longer term. Refactoring small sets of related tables might be one way to make progress without large scale disruption.

### *Use cases*

AdM asked whether there was a collection of the different kinds of samples and their uses that were relevant across the users of ISPyB. This fitted into a point already raised at the Diamond meeting, on the need to collect and classify use cases.

It was agreed that such a collection would be highly useful, the question being how to collect it and who could provide the resources. GPhL were not unwilling to contribute, but felt that the primary source would have to be people who were actively collecting data; GPhL mentioned Jean-Luc Ferrer and the FIP team, Jose Marquez and the CRIMS team, and Gleb Bourenkov at EMBL Hamburg as potential sources.

It was agreed that there should be a process of gathering requirements from outside scientists, in a manner similar to what had been done for EM, that this should be done under the auspices of the ISPyB Science Committee, and that GB would raise the issue with the Science Committee and stay involved in the process.

#### *Sample clarification*

SM queried whether the BLSamples are mapped to individual crystals within a BLSample, but the answer was that they are merely locations within the Sample, with no grouping into crystals.

#### *Issue #15*

There was a discussion about issue #15, the introduction of 'deleted' columns. The issue started from the fact that deletion of some data would cascade to delete others as well, and that permanently removing information was in general undesirable. Several points were raised: 'Deleted' status could interact with other kinds of object status, e.g. 'isContaminated'; a multi-use status field might limit the number of columns needed, but it was not obvious which kinds of status would be mutually exclusive.

Introduction of a 'deleted' column on some tables might eventually spread to many other tables, raising a scalability issue. And SD had found 'deleted' columns (soft deletion) detrimental in other applications and had gone back to hard deletion. The issue was left open for further discussion.

#### **Any Other Business**

None was raised.

Post meeting there was a discussion between DLS and GP over use of appear.in service. We used appear.in for this VTC and it worked well (apart from some echo on the line when microphones were not muted). PK kindly offered to setup an official ISPyB room under the Global Phasing account so we can host more sites than the free account.

#### **Date of Next Meeting**

The next meeting is tentatively planned for early May due to public holidays.

ESRF proposes: on May 2nd morning, or on May 3rd all day, or on May 4th all day.

To be confirmed by the others.