

ISPyB Developers' meeting

July 9 2019

DRAFT

Participants:

ALBA: Daniel Sanchez

Diamond: Karl Levik, Neil Smith,

ESRF: Olof Svensson, Solange Delageniere, Alejandro de Maria, Maxim

Global Phasing: Clemens Vonnrhein, Gerard Bricogne, Rasmus Fogh

Max IV: Alberto Nardella

SOLEI: Tatiana Isabet

1. Minutes and matters arising

The minutes of the April developers' meeting were agreed. No matters arising

2. Minutes and presentations from the Lund meeting

The minutes of the Lund meeting were not, after all, visible at <https://ispyb.github.io/ISPyB/>. Participants will send presentations and minutes to RF, who will make a limited meeting report from them by the end of the week.

3. Status reports

Max IV: Working on FastDP autoprocessing plots and statistics pages. Next task will be to put latest ISPyB into production.

ALBA: Nothing to report

ESRF: Working on reprocessing facility for EXI2 (react framework). Reprocessing jobs will be input from ISPyB and sent to the new system. Work is now on a prototype, it is planned to reach production in September, with support for various programs (EDNA, XDSAPP, autoPROC, ...). For now, geared to single-sweep data collections but will be expanded later to multi-sweep and multi-crystal collections.

SOLEIL: Nothing new to report, severe manpower shortage.

Diamond: Business as usual. Work on SynchWeb for sample registration, EM integration, SynchWeb workflows, and software infrastructure such as modernising servers and moving to Python 3.

Global Phasing: CV has been swallowed up by refinement, so the move of MRFANA to open source has been delayed but **not** abandoned. Documentation for using existing binaries will be disseminated soonest. MXCuBE integration has proceeded successfully, recent tests at ALBA producing successful data on real-life crystal systems.

4. Multi-scan data collection

GPhL raised the question of how to store multi-sweep data collections and the associated processing results in ISPyB. For the specific GPhL use case a set of sweeps are collected on a single crystal specifically to be processed together, and all links, shared information etc. are handled by the GPhL workflow and processing software. The question is how to store and view this information. The issue overlaps with work done at ESRF for reprocessing (within ISPyB), and also considerations for handling multi-crystal data in a separate non-ISPyB system for reprocessing in Diamond pipelines. GPhL asked whether the DataCollectionGroup could provide a basis for a solution to this problem, and ESRF and Diamond plan to rely on this in the first instance, but it is defined mainly as a grouping for data **collection** and is used among other things for mesh scans; it can not seamlessly be used also for defining a grouping at the processing stage. One structural impediment is also that the DataCollectionGroup is linked to a single BLSample, which precludes its use for multi-crystal collections (but not for multi-sweep collections on one crystal). Diamond, accordingly, intends to handle multi-crystal data collection in SynchWeb outside ISPyB. ESRF also raised the possibility of attaching multi-sweep processing to a single one of the sweeps involved. CV raised the question of having/getting a system that can handle multi-wavelength and multi-crystal as well as multi-sweep collections, and KL confirms that multi-wavelength handling is integral to the current system.

It is agreed that the discussion will be taken up again at the next web meeting, and that ESRF will present a draft of their ongoing work at this meeting (**ACTION:** ESRF). It was requested that a version be presented for comments and/or additional use cases while changes are still possible.

5. Refactoring plans and face-to-face meeting

There have been no developments on this point since the end-of-April web meeting. This meeting agrees that refactoring is necessary and important (though in competition for resources with adding new features) and that the proposed face-to-face meeting should go ahead in order to speed up long-accumulated refactoring needs. Tatiana confirms that SOLEIL will handle the organisation and send out mails with practical details (**ACTION** SOLEIL). It is agreed that the meeting should take place on 12-13 September, starting at lunch and ending in the evening.

KL (**ACTION**): volunteers to prepare draft change proposals for the next web meeting covering the areas of DataCollection harmonisation, Energy Scan, XRF, and also protein/components, and sample composition.

ESRF (**ACTION**) volunteers to prepare draft change proposals for merging the tables for autoproccessing programs (with attachments) and phasing program runs (with attachments), and to clean up the little used tables for screening.

Further subjects mentioned as worth refactoring were multi-sweep/multi-crystal data collection and processing (see previous point), and fully harmonising the handling of basic tables such as sessions.

Since there is a limit to how many areas can be realistically be refactored at once, the question on what to take up at the face-to-face meeting and the timing of resulting refactorings will be taken up again at the next web meeting.

6. Any Other Business

KL raised the question of which license ISPyB was under, because Diamond wanted to distribute derivative work under a compatible license. It was clarified that ISPyB was released under LGPL, and emphasised (Solange) that the location of license files, license statements at the top of files, and lists of involved developers in the documentation needed to be kept up to date.

Next Meeting

The next meeting should be at the start of September to start preparations for face-to-face meeting; RF has set up a Doodle poll (<https://doodle.com/poll/rf7x6pqf9wxiwexd>).