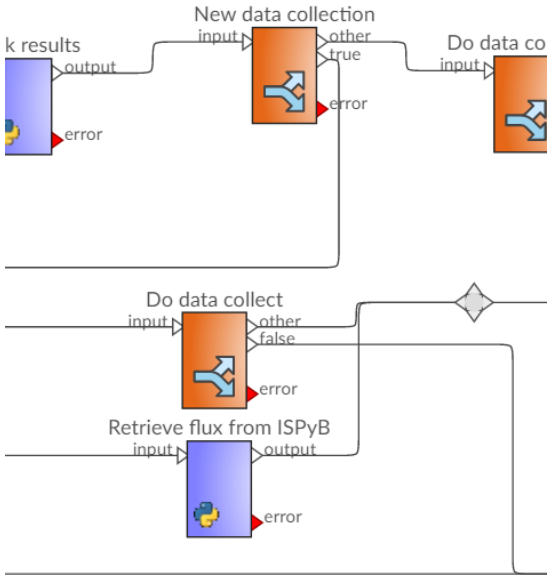


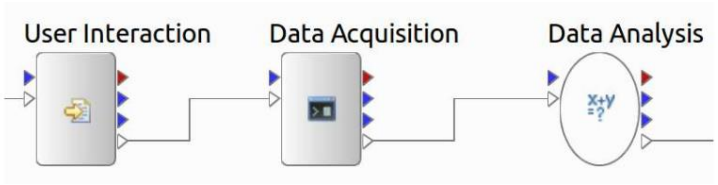
Workflows and ISPyB/EXI

Olof Svensson
Data Analysis Unit
ISDD / ESRF

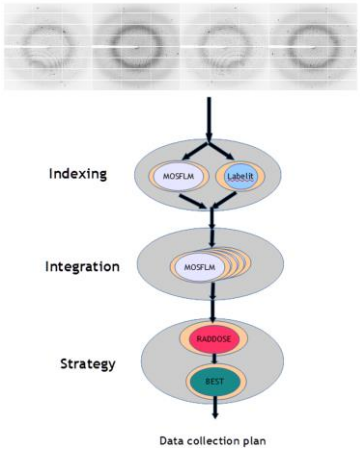
The European Synchrotron | ESRF

EXAMPLES OF COMPLEX DATA COLLECTION

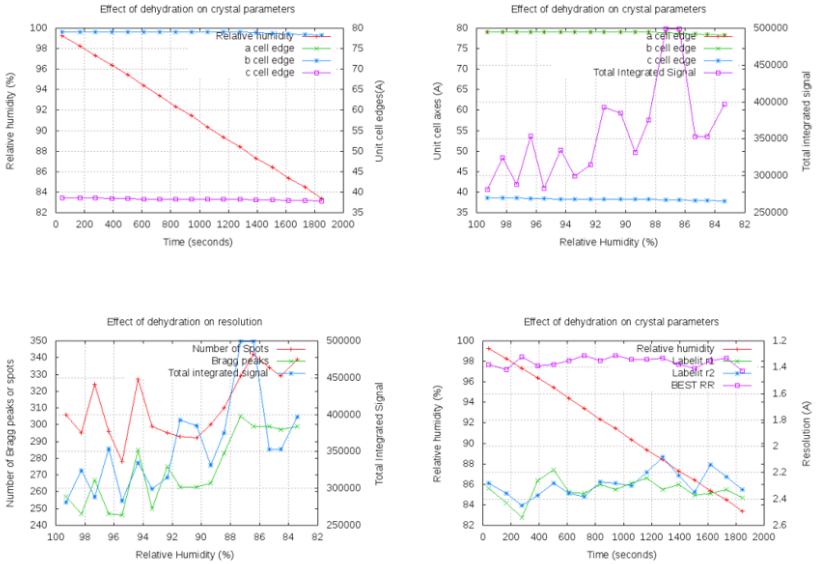
- **Combination of:**
  - Data collection
  - Data processing
  - [User interaction]



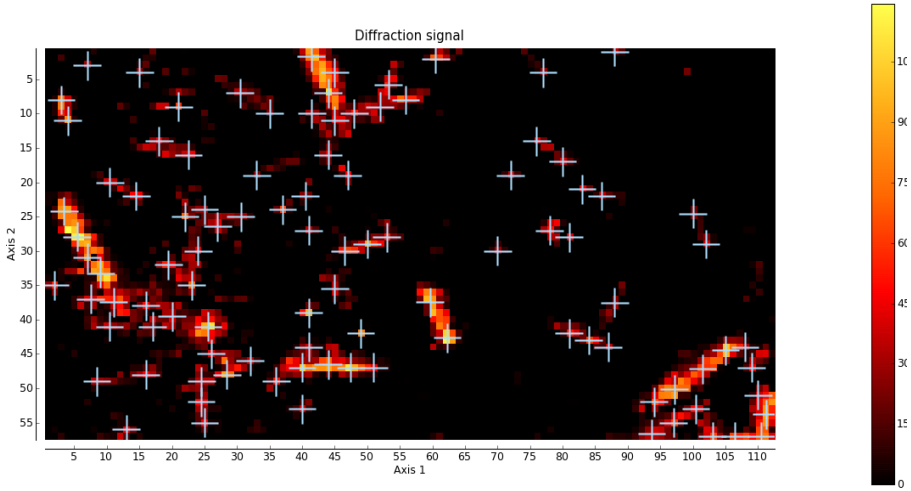
Example: Characterisation



Example: Dehydration

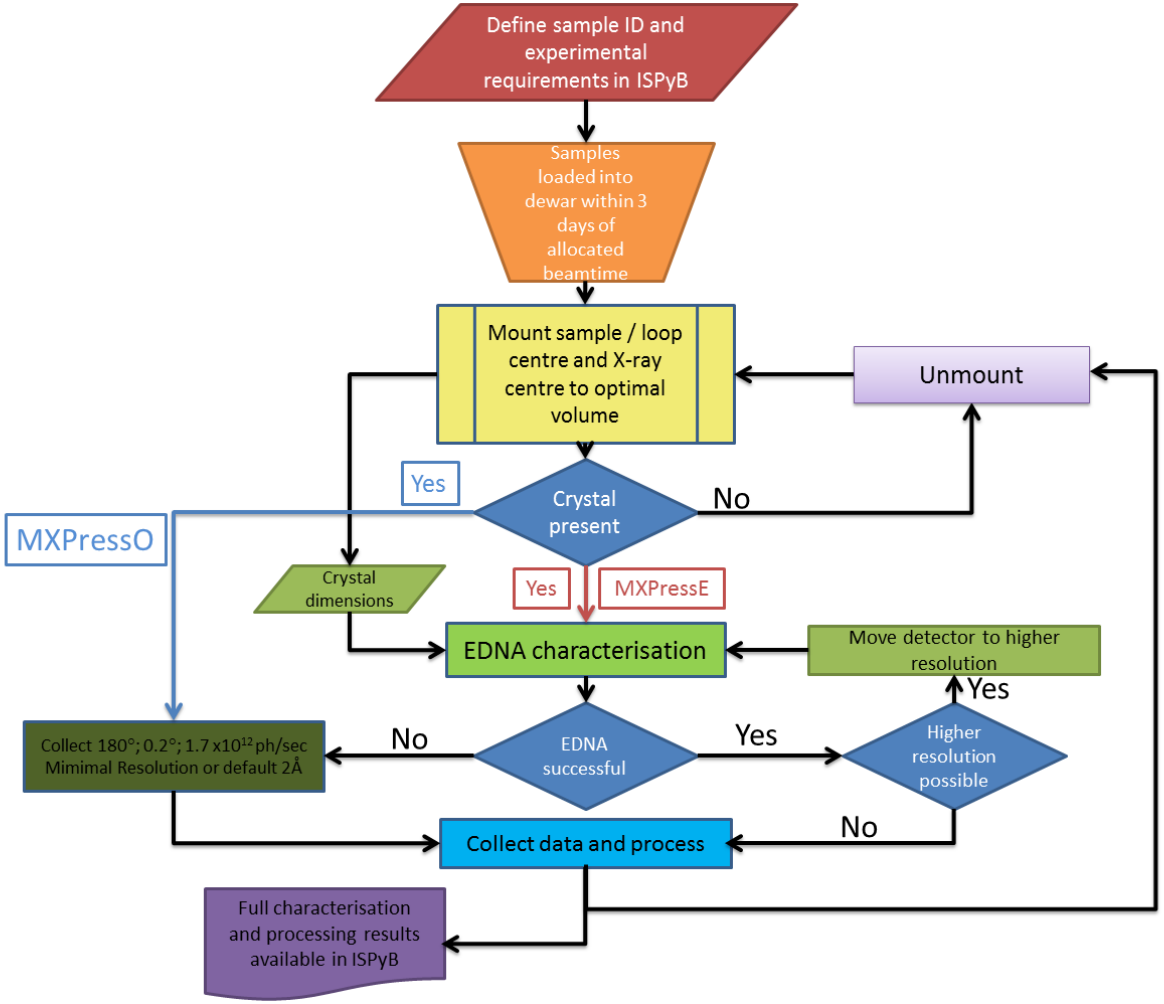


Example: Mesh and collect



THE MOST COMPLEX DATA COLLECTION: MXPRESSA

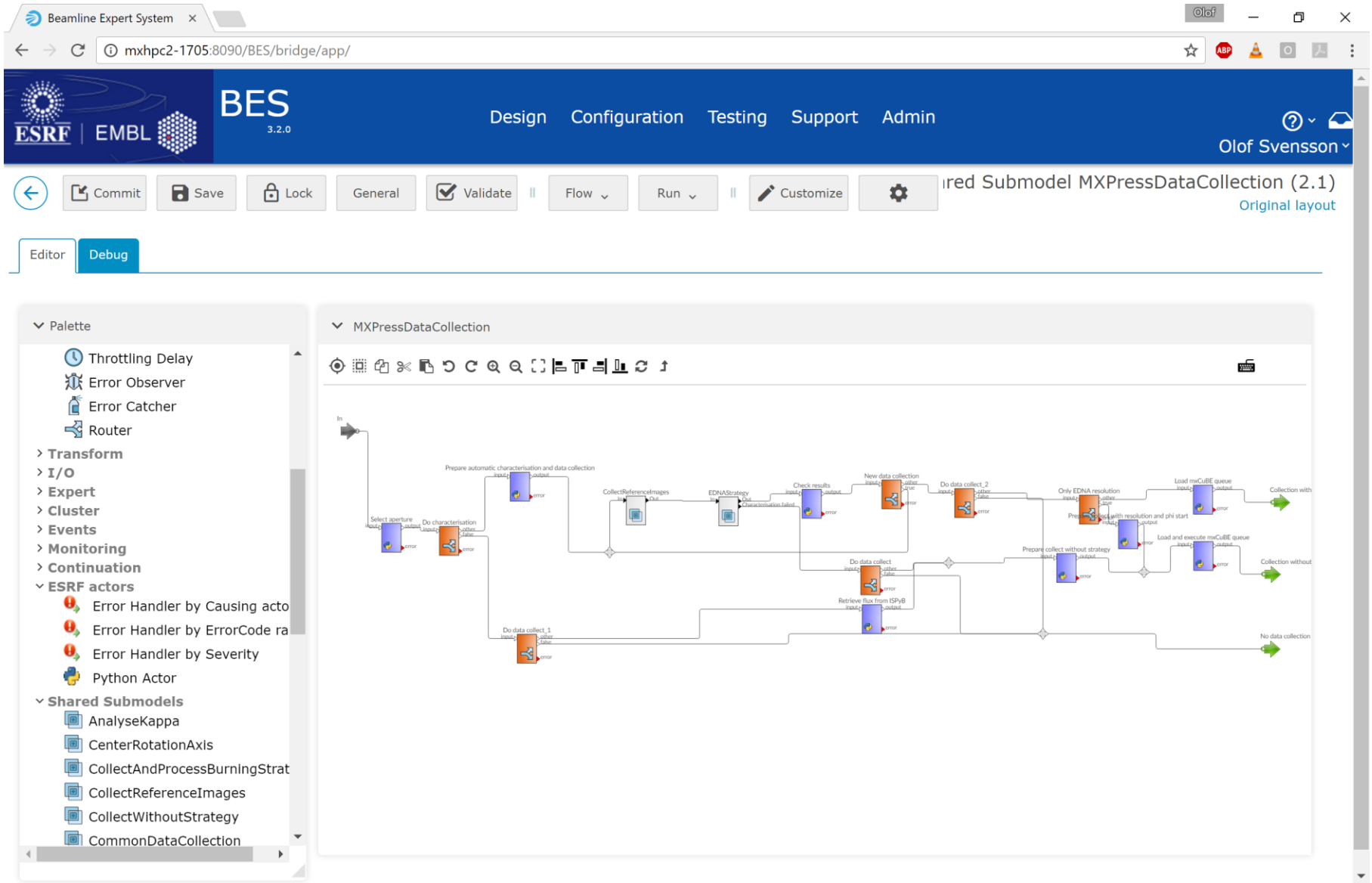
MXPressA (~10000)



THE BEAMLINE EXPERT SYSTEM

- **The BES (Beamline Expert System)**
  - A customized version of Passerelle EDM for the ESRF
  - Similar to the SES (Soleil Expert System)
- **Passerelle EDM:**
  - Based on the open-source passerelle workflow engine, which is based on the Ptolemy II workflow engine (written in Java)
  - Combines the passerelle workflow engie with a workflow design tool
  - Both the engine and the design tool are connected to a data base
  - Accessed via a web GUI (Jboss Application Server 7)
- **Passerelle EDM is a commercial product:**
  - Initial one-off purchase followed by one-year maintenance contracts
  - ESRF license costs shared between the ESRF and EMBL Grenoble
  - Contact person: Erwin de Ley (erwin.de.ley@isencia.be)

BES DESIGN TOOL



BES – WORKFLOW EXECUTION LOG

Beamline Expert System

mxhpc2-1705:8090/BES/bridge/app/

ESRF | EMBL

BES 3.2.0

Design Configuration Testing Support Admin

Olof Svensson

Refresh Export Flow

Request details

Request (id= 4734179, type = MXPressA)

Reference: 9070
Id: 4734179
Type: MXPressA
Creation Date: 2016-12-09 16:02:31
Status: Finished
External Reference: fx57
Correlation Id:
Initiator: id30a1
Executor:

7 rows 10

Name	Value	Data Type
collection_software	mxCuBE_-_2.0	STRING
directory	/data/visitor/fx57/id30a1/20161209/RAW_DATA/BIPDE1/BIPDE1-CD014403_H04-3_DI01318464	STRING
modelpath	MXPressA	STRING
prefix	BIPDE1-CD014403_H04-3_DI01318464	STRING
run_number	1	STRING
sample_lims_id	575403	STRING
sample_node_id	119	STRING

Task | Result Item | Event

Enter search string here

Id: Status: Created Error Finished Cancelled Started Timeout
Type: Discriminator: Initiator:

View 101 rows 400

Id	Creation Date	Duration(ms)	Status	Type	Initiator
4734180	2016-12-09 16:02:31	8	Finished	Start	actor:/MXPressA.Start
4734291	2016-12-09 16:02:31	632	Finished	eu.esrf.services.python.v2.6	actor:/MXPressA.Beamline%20setup
4734292	2016-12-09 16:02:32	576	Finished	eu.esrf.services.python.v2.6	actor:/MXPressA.Prepare%20MXPressA
4734293	2016-12-09 16:02:32	1850	Finished	eu.esrf.services.python.v2.6	actor:/MXPressA.CommonPrepareExperiment.Init%20workflow
4734294	2016-12-09 16:02:34	768	Finished	eu.esrf.services.python.v2.6	actor:/MXPressA.CommonPrepareExperiment.Read%20motor%20positions
4734295	2016-12-09 16:02:35	299	Finished	eu.esrf.services.python.v2.6	actor:/MXPressA.CommonPrepareExperiment.Default%20parameters
4734296	2016-12-09 16:02:35	2295	Finished	eu.esrf.services.python.v2.6	actor:/MXPressA.MXPressMultiCrystal.Interpretation%20of%20diffraction%20p
4734297	2016-12-09 16:02:37	315	Finished	eu.esrf.services.python.v2.6	actor:/MXPressA.MXPressMultiCrystal.Contact%20user%20if%20first%20acro

WORKFLOW EXECUTION STATISTICS

- **Shipped with Passerelle EDM:**

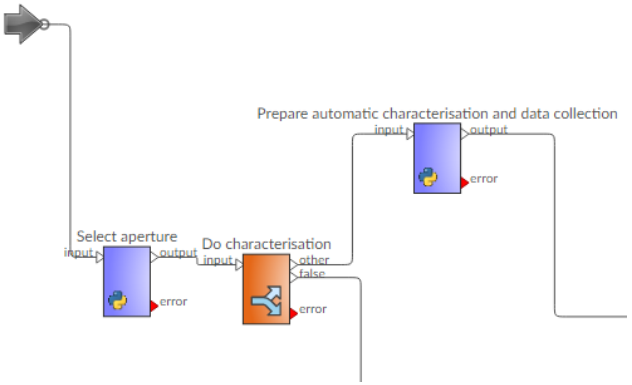
Enter search string here

Type: Day(Creation Day): 01/12/2016 To 07/12/2016

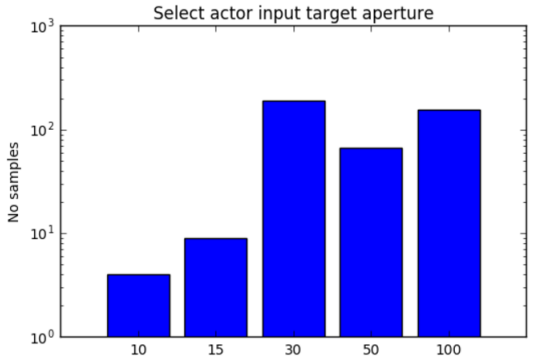
Initiator: bm14 id23eh1 id23eh2 id29 id30a1 id30a2 id30a3 id30b

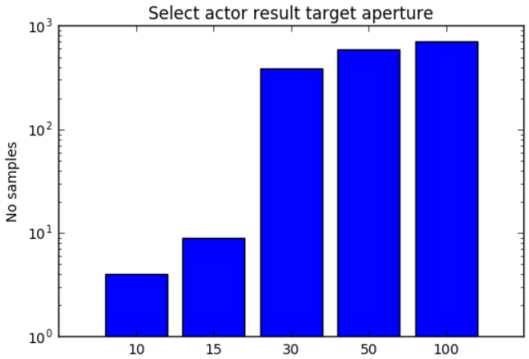
71 rows

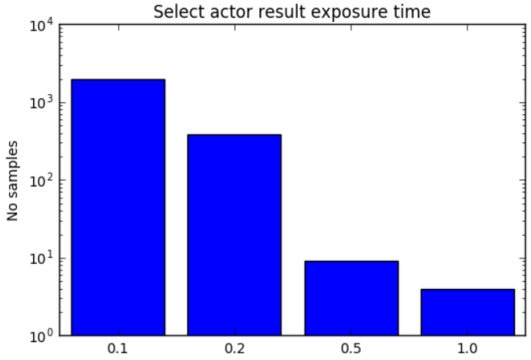
Creation Date/Day	dozor_OAR	autoPROC_OAR	XrayCentring	XIA2_DIALS_OAR	XDSAPP_OAR	VisualReorientation	TroubleShooting	MeshScan	MeshAndCollect	MXPressA_aperture_30um	MXPressA	KappaReorientation	HelicalCharacterisation	EDNA_proc_OAR	CreateThumbnails	Characterisation	Sum
01/12/2016	303	113	1	139			2	91			29			139	2565	154	3,536
02/12/2016	408	107	17	107			1	93			40		2	107	3371	300	4,553
03/12/2016	253	127	18	140			2				94			140	2787	117	3,678
04/12/2016	406	81	1	81						29	11			80	1974	325	2,988
05/12/2016	580	108	2	108		10	1	28	9		5			108	1799	463	3,221
06/12/2016	162	49		49	6			5						42	518	111	942
07/12/2016	276	66	38	71			1	11			8	6	1	68	1821	203	2,570
Sum	2,388	651	77	695	6	10	7	228	9	29	187	6	3	684	14,835	1,673	21,488



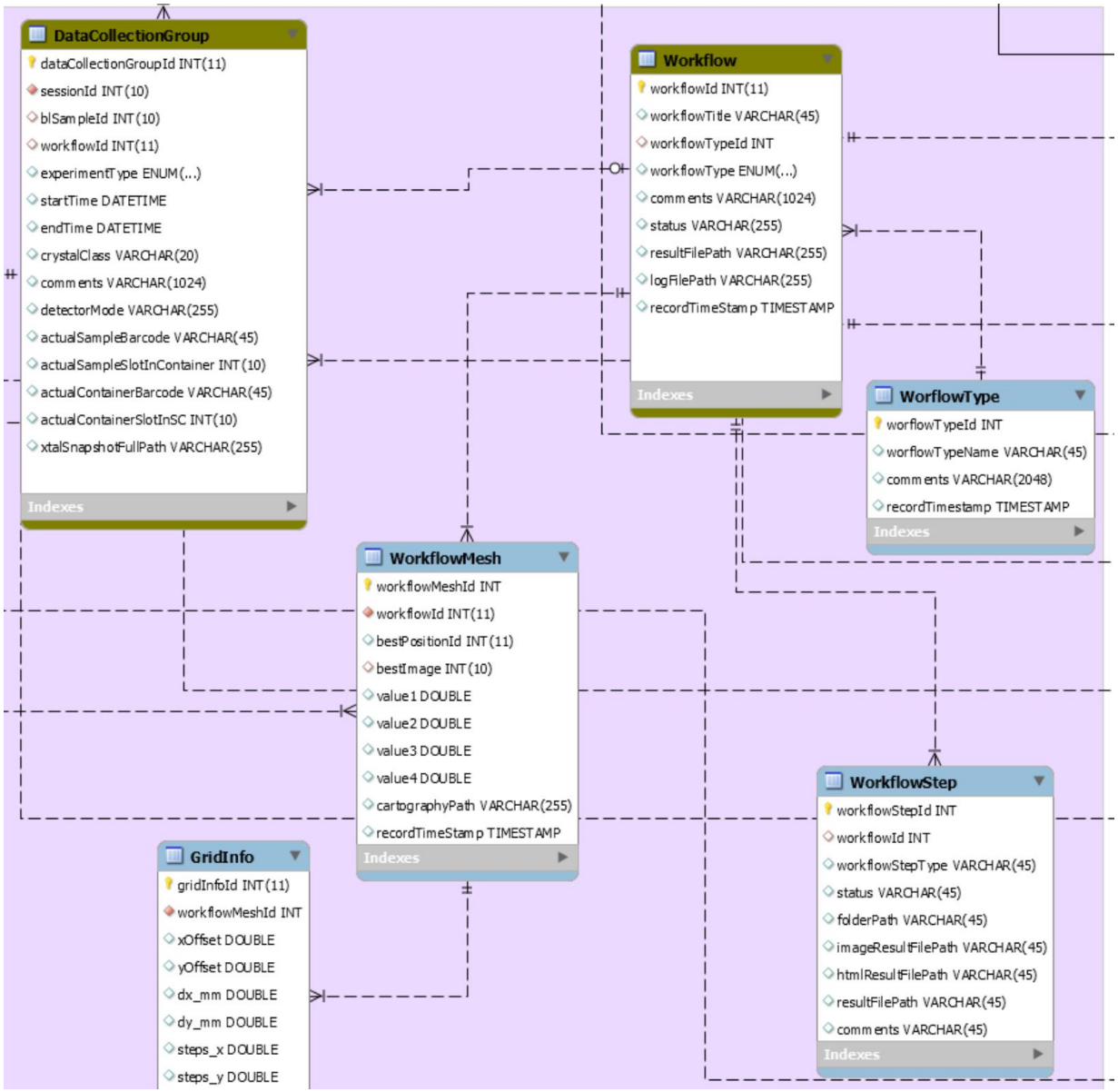
- **Custom statistics with Python script accessing BES database:**





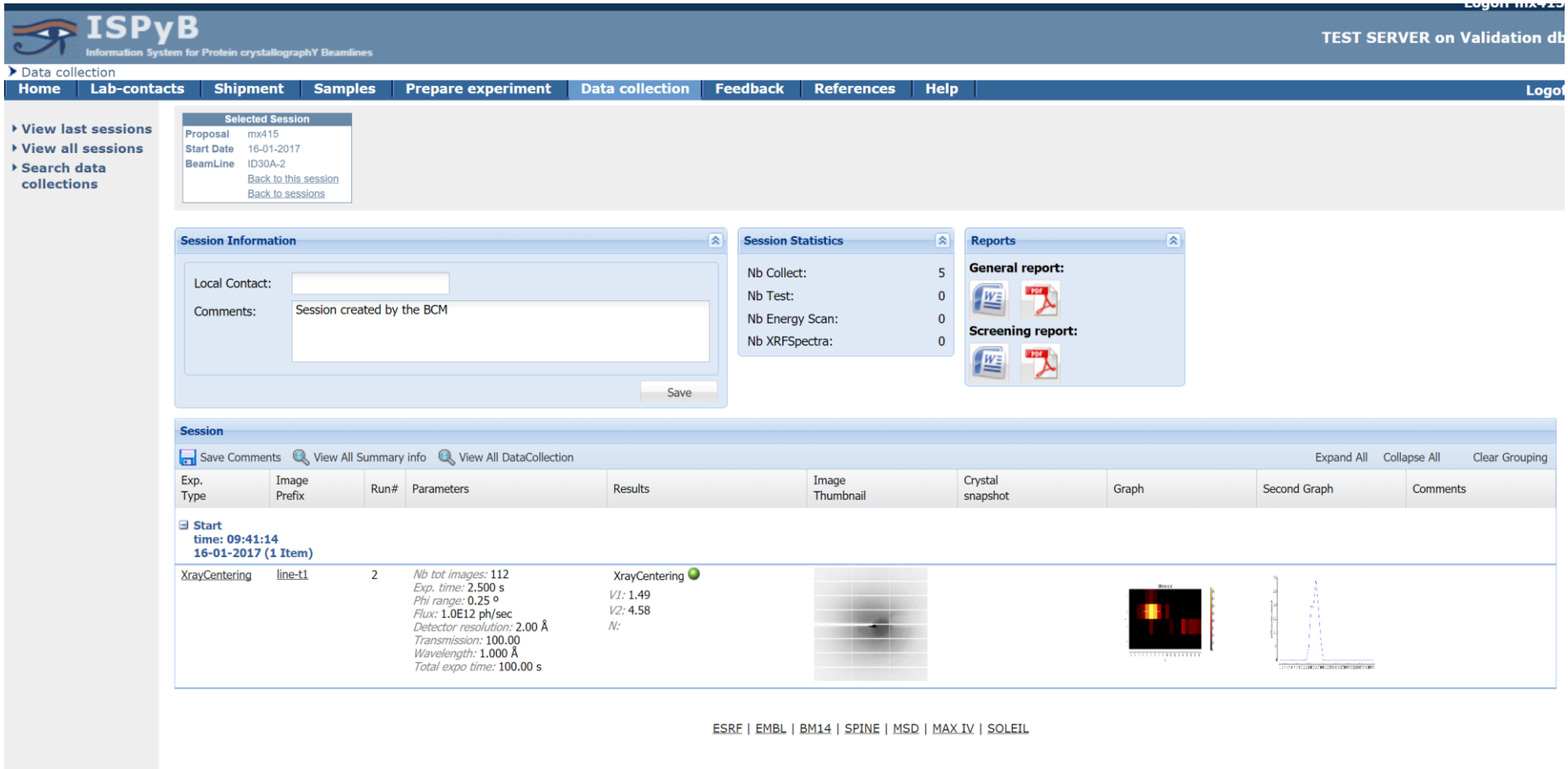


ISPYB DATA BASE AND WORKFLOWS



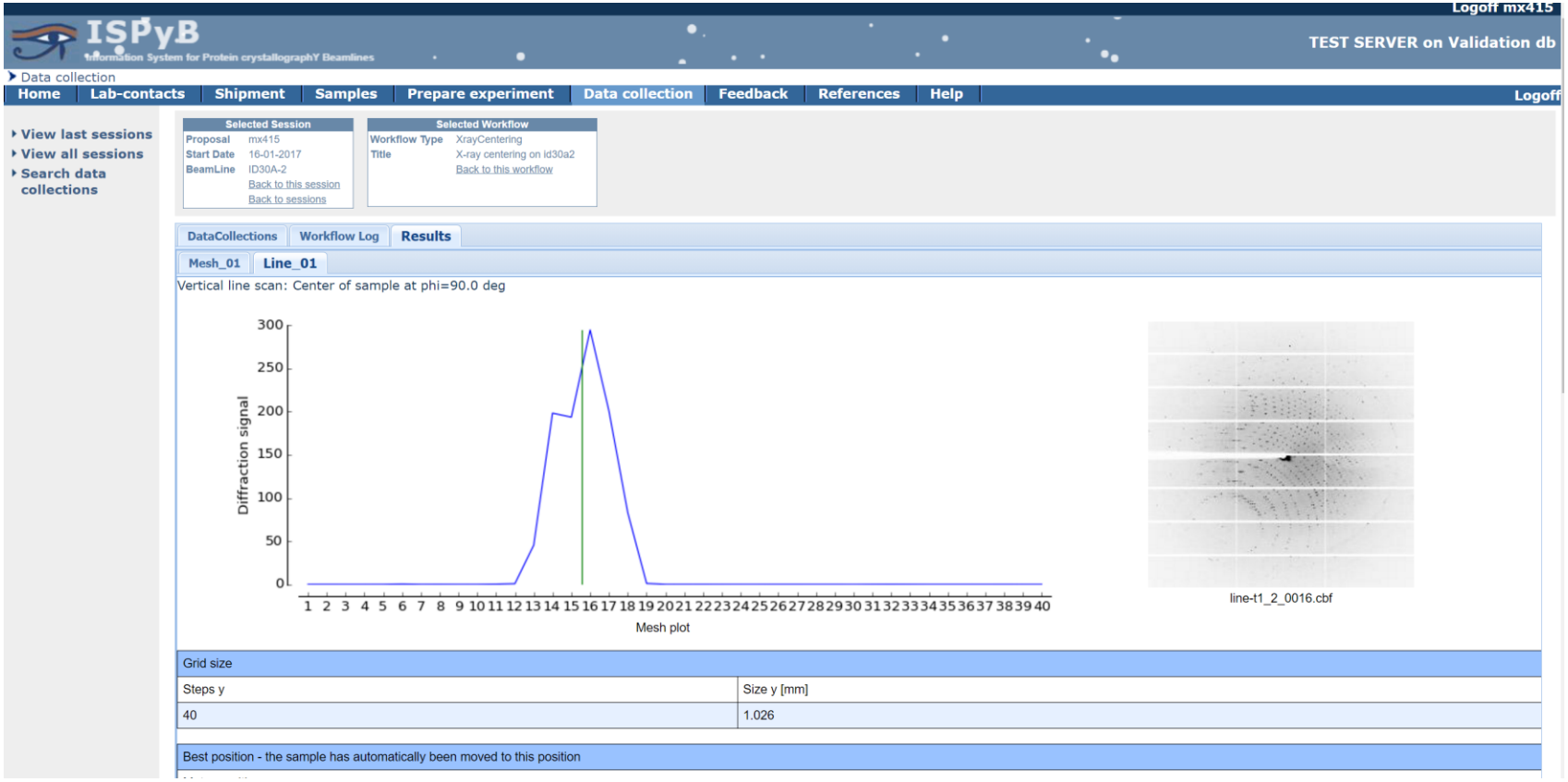
ISPYB GUI

- **The results of the workflows are accessed from the summary page**



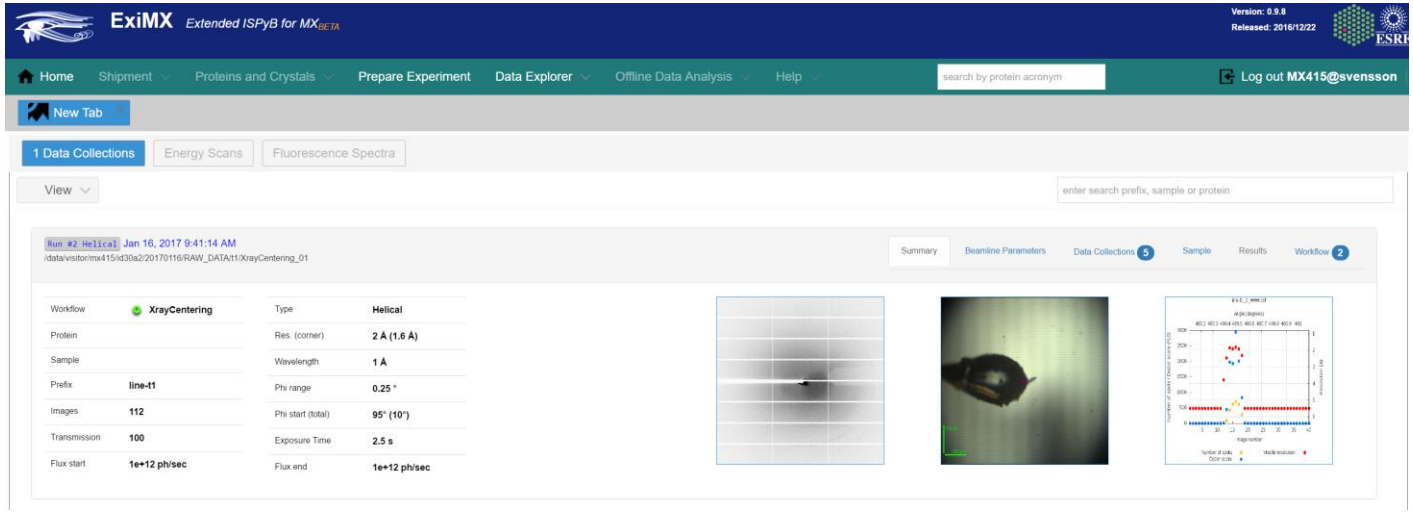
ISPYB GUI

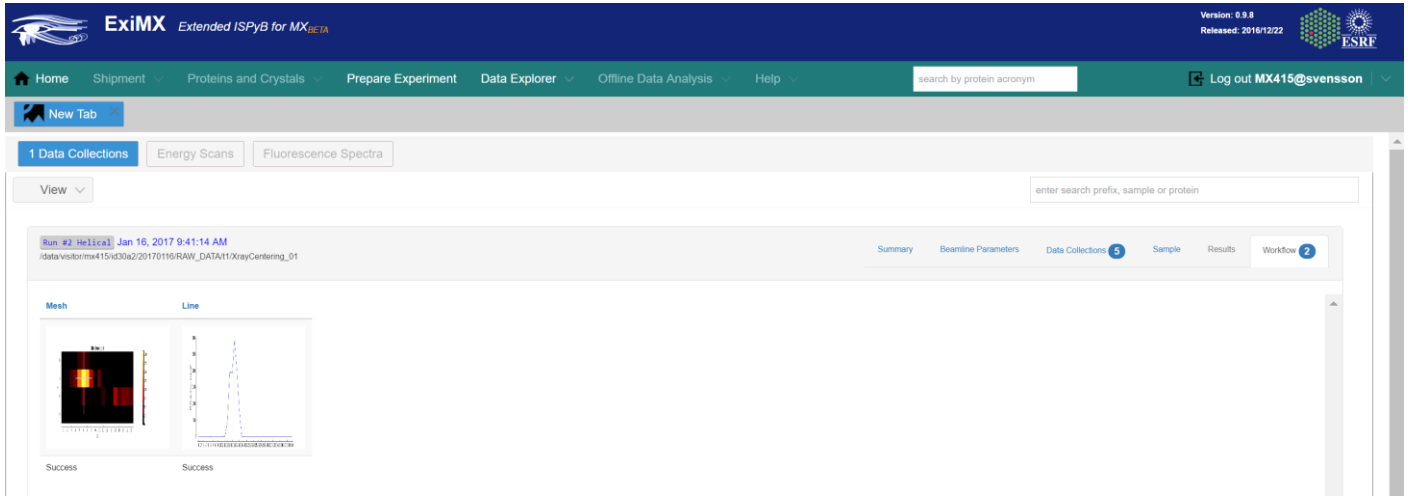
- **The results of web pages are presented as HTML pages**
  - The directory location of the web pages is stored in Workflow.resultFilePath. Every directory inside this directory that contains a “index.html” page is rendered as a tab, the name of the tab is the name of the directory



EXI GUI

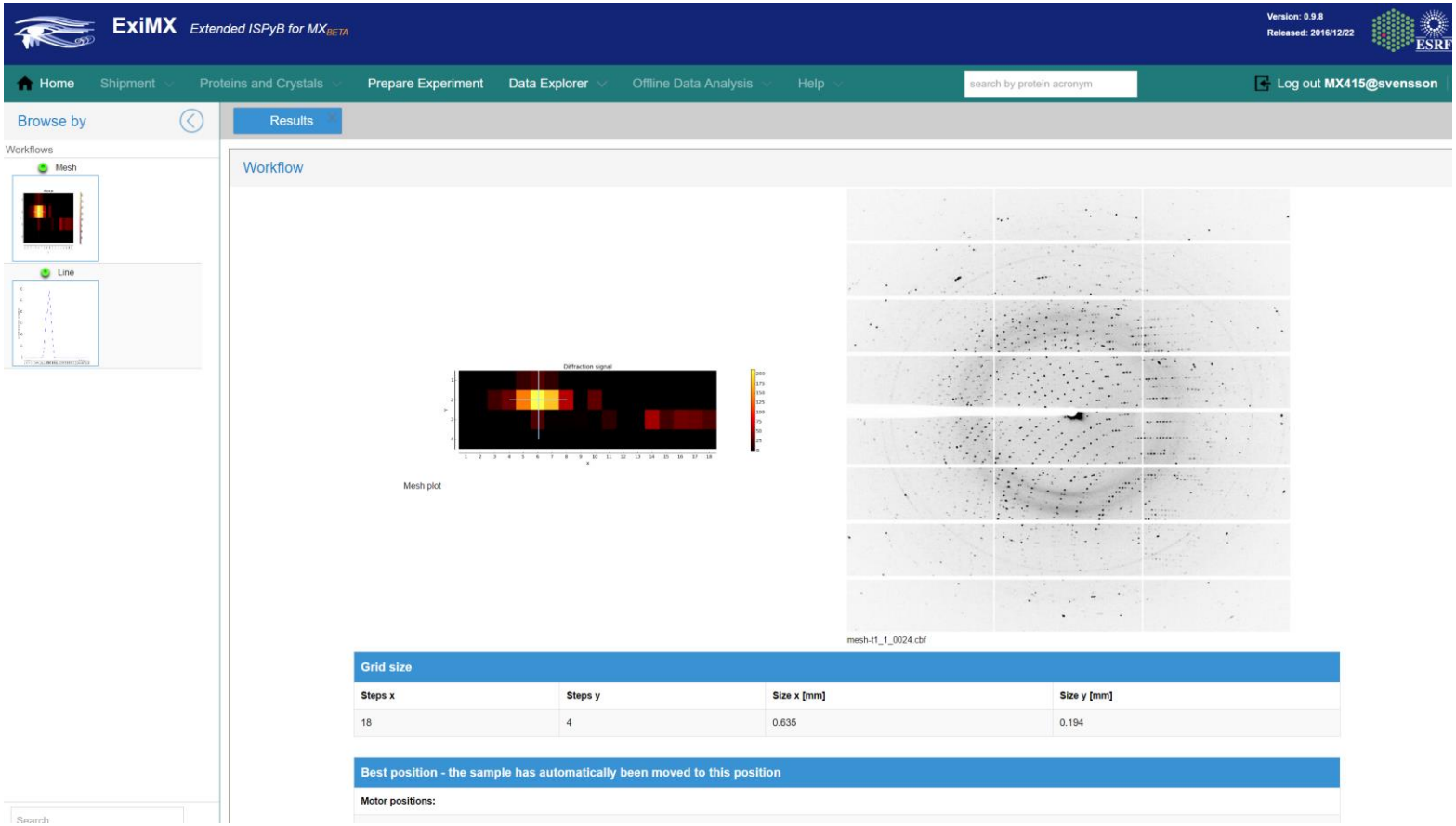
- **Results of workflows are accessed from the summary page**





DIFFERENCE BETWEEN ISPYB GUI AND EXI

- **EXI:**
  - The results are stored as json files.
  - The json file contains all data (text, tables and encoded images)
  - The location of the json file is stored in WorkflowStep.resultFilePath



ACKNOWLEDGEMENTS

- David von Stetten, Max Nanao, Sasha Popov, Daniele de Sanctis, Stéphanie Monaco, Didier Nurizzo, Matias Guijarro, Solange Delagenière, Alejandro de Maria, Marcus Oskarsson (ESRF)
- Matthew Bowler (EMBL Grenoble)
- Erwin de Ley and Koen Heunick (Isencia, Belgium)
- The ESRF Data Analysis Unit, Beamline Control Unit and Structural Biology group

Further reading:

- **Fully automatic characterization and data collection from crystals of biological macromolecules**, Svensson, O., Malbet-Monaco, S., Popov, A., Nurizzo, D., & Bowler, M. W. (2015). Acta Crystallographica Section D: Biological Crystallography, 71(Pt 8), 1757–1767. http://doi.org/10.1107/S1399004715011918
- **MeshAndCollect: an automated multi-crystal data-collection workflow for synchrotron macromolecular crystallography beamlines**, Ulrich Zander · Gleb Bourenkov · Alexander N Popov · Daniele De Sanctis · Olof Svensson · Andrew A Mccarthy · Ekaterina Round · Valentin Gordeliy · Christoph Mueller-Dieckmann · Gordon A Leonard, ACTA CRYSTALLOGRAPHICA SECTION D BIOLOGICAL CRYSTALLOGRAPHY 71(11):2328-2343 · NOVEMBER 2015
- **Facilitating best practices in collecting anomalous scattering data for de novo structure solution at the ESRF Structural Biology Beamlines**, Daniele de Sanctis · Marcus Oscarsson · Alexander Popov · Olof Svensson · Gordon Leonard, Acta Cryst. (2016). D72, 413-420
- **The use of workflows in the design and implementation of complex experiments in macromolecular crystallography**, S. Brockhauser, O. Svensson, M. W. Bowler, M. Nanao, E. Gordon, R. M. F. Leal, A. Popov, M. Gerring, A. A. McCarthy and A. Gotz,, Acta Cryst. (2012). D68, 975-984