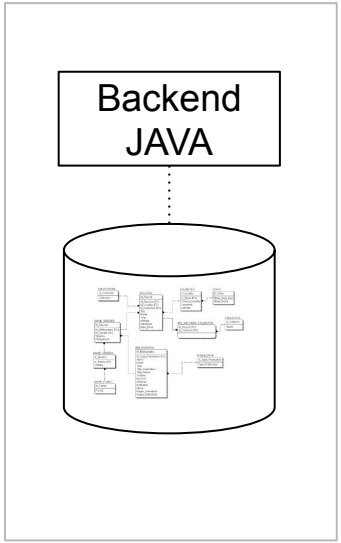
News on the ISPyB collaboration, roadmap and challenges

Alex de Maria Antolinos
Software Engineer
Data Manager@Data Automation Unit
Software Group
ESRF

18/05/2022

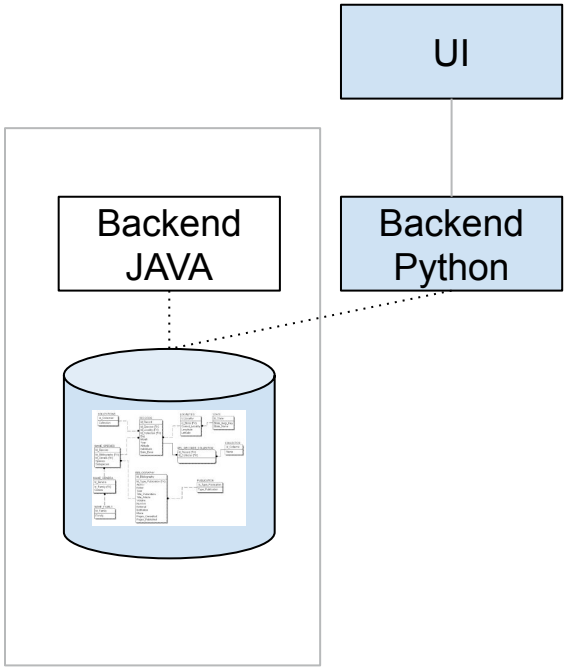
Introduction

- MOU started on 01 January 2017
  - Collaboration around:
    - Data Model
    - Java Backend
  - Participants:
    - ESRF,
    - DLS,
    - SOLEIL,
    - ALBA,
    - MAXIV,
    - HZB,
    - EMBL
    - Global Phasing
- MOU terminated on 31 December 2021
  - Do we want a new MOU?
    - Who? When?



Introduction

- MOU started on 01 January 2017
  - Collaboration around:
    - Data Model
    - Java Backend
  - Participants:
    - ESRF,
    - DLS,
    - SOLEIL,
    - ALBA,
    - MAXIV,
    - HZB,
    - EMBL
    - Global Phasing
- MOU terminated on 31 December 2021
- February 2020 interim collaboration meeting
  - Implement backend in python -> **py-ispyb**
    - EMBL did a prototype https://gitlab.esrf.fr/ispyb/py-ispyb
  - Develop SSX



2022 Collaboration kick-off

- Weekly coordination meetings on Thursday at 9h30
- Active participants and contact person:

ESRF	demariaa@esrf.fr
DESY	clemente.borges@desy.de
SOLEIL	idrissou.chado@synchrotron-soleil.fr
ALBA	marmenter@cells.es;blorenzo@cells.es;acampsm@cells.es;ecenteno@cells.es
HZB	michael.hellmig@helmholtz-berlin.de; alexander.dillmann@helmholtz-berlin.de
GLOBALPHASING	rhfogh@globalphasing.com
MAXIV	alberto.nardella@maxiv.lu.se
DLS	karl.levik@diamond.ac.uk;james.p.hall@diamond.ac.uk
EMBL	dvonstetten@embl-hamburg.de

- Creation of working groups in order to focus in well defined areas
  - Dedicated meetings
  - Output is shared in the weekly coordination meetings for approval

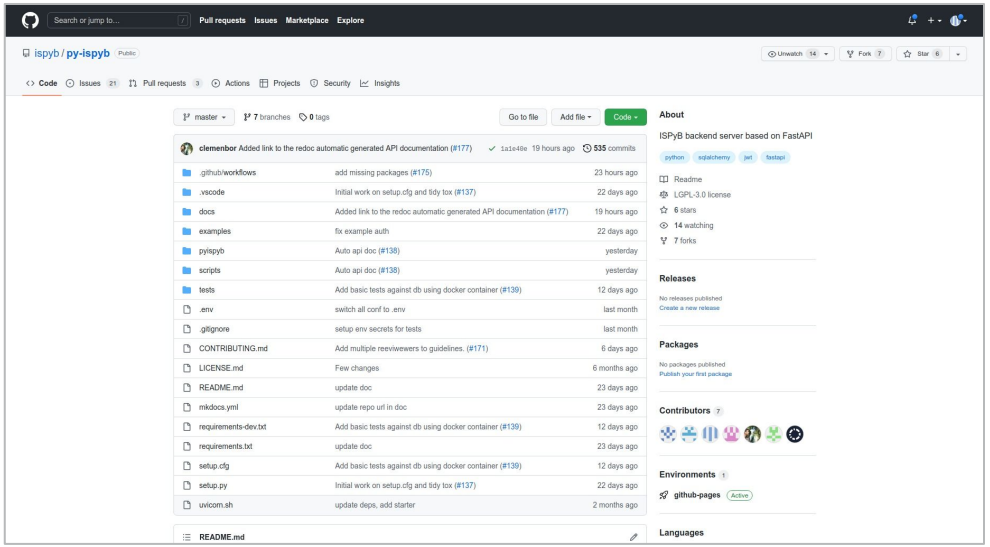
Working Groups

Tasks	Brief description	Responsible(s)	Participant(s)
Framework architecture	Organize the project structure, choose the libraries and define the best good practices to be adopted (automatic testing and documentation), deployments, etc...	ESRF	All
Authentication	Develop the authentication/authorization mechanism(s)	ESRF	Soleil, DESY, DLS
User Portal Sync	Development of a fairly generic mechanism to synchronize the data from the UP. It includes entities like proposals, sessions, proteins and samples, etc...	DESY	ESRF, ALBA, DLS
Shipping	Implementation of the sample tracking system	DESY, ESRF, DLS	SOLEIL, GP
EM	Development of cryo-electron microscope (cryoEM) for single particle experiments	ESRF	DLS
MX	Development of MX	GP, SOLEIL	EMBL, DESY, ESRF, MAXIV, ALBA, DLS
BioSAXS	Development of BioSAXS		
SSX	Development of synchrotron serial crystallography (SSX) experiments.	ESRF	EMBL, DESY, DLS, GP
Others techniques?			
Documentation	Ensure coherent and up to date documentation. User, Developres and Application Developers	GP	ESRF, SOLEIL
X-ray imaging		EMBL	

Framework/Architecture Group

Goal:
Organize the project structure, choose the libraries and define the best good practices to be adopted (automatic testing and documentation), deployments, etc...

Activity:
Github repository (https://github.com/ispyb/py-ispyb)
FASTAPI
Automatic API Documentation
Testing
Discussions on best practices



Authentication/Authorization Group

Goal:
Develop the authentication/authorization mechanism(s)

Authentication mechanism survey

Site	Keycloak	LDAP	Custom	Comments
ESRF	X	X		We want/need to support both user and proposals accounts
HZB		X	X	In future we would need both user and proposal authentication.
SOLEIL	X	X	X	Only user acconts are supported at SOLEIL either for BAG proposal or any type of proposals Possibility to add authentication from ORCID would be appreciated. Note: custom refers to authentication agains DB for the historical ISPyB
ALBA				
DLS				
EMBL				
MAXIV	X	X		Only user accounts are supported. We actually use LDAP but we want to use keycloak in the future
DESY			X	We use a REST API to authenticate users. So far only user accounts are foreseen

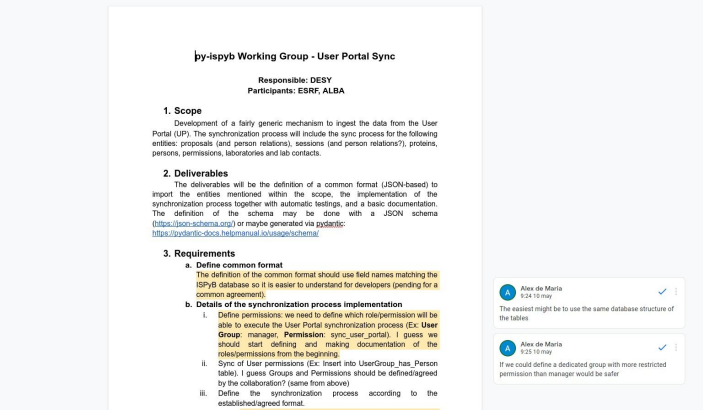
Implementation and documentation

- Flask implementation has been ported to FastAPI
- Documentation can be found on https://ispyb.github.io/py-ispyb/auth/

User Portal Sync Group

Goal:
Development of a fairly generic mechanism to synchronize the data from the UP. It includes entities like proposals, sessions, proteins and samples, etc...

Requirement analysis



py-ispyb Working Group - User Portal Sync

Responsible: DESY
Participants: ESRF, ALBA

1. Scope

Development of a fairly generic mechanism to ingest the data from the User Portal (UP). The synchronization process will include the sync process for the following entities: proposals (and person relations), sessions (and person relations?), proteins, persons, permissions, laboratories and lab contacts.

2. Deliverables

The deliverables will be the definition of a common format (JSON-based) to import the entities mentioned within the scope, the implementation of the synchronization process together with automatic testings, and a basic documentation. The definition of the schema may be done with a JSON schema (https://json-schema.org/) or maybe generated via pydantic: https://pydantic-docs.helpmanual.io/usage/schema/

3. Requirements

a. Define common format
The definition of the common format should use field names matching the ISPyB database so it is easier to understand for developers (pending for a common agreement).

b. Details of the synchronization process implementation

i. Define permissions: we need to define which role/permission will be able to execute the User Portal synchronization process (Ex: User Group: manager, Permission: sync_user_portal). I guess we should start defining and making documentation of the roles/permissions from the beginning.

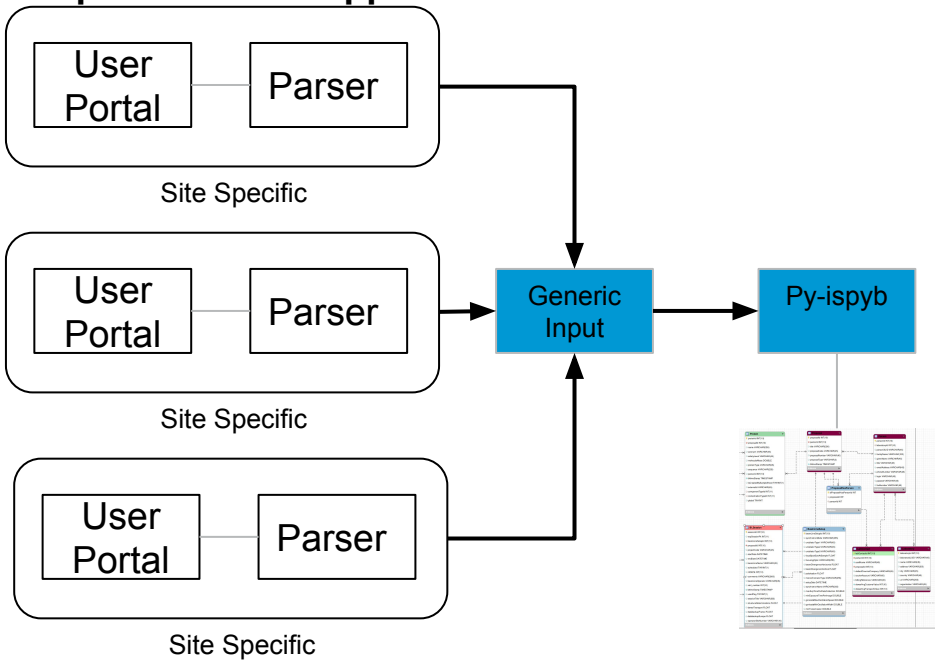
ii. Sync of User permissions (Ex: Insert into UserGroup_has_Person table). I guess Groups and Permissions should be defined/agreed by the collaboration? (same from above)

iii. Define the synchronization process according to the established/agreed format.

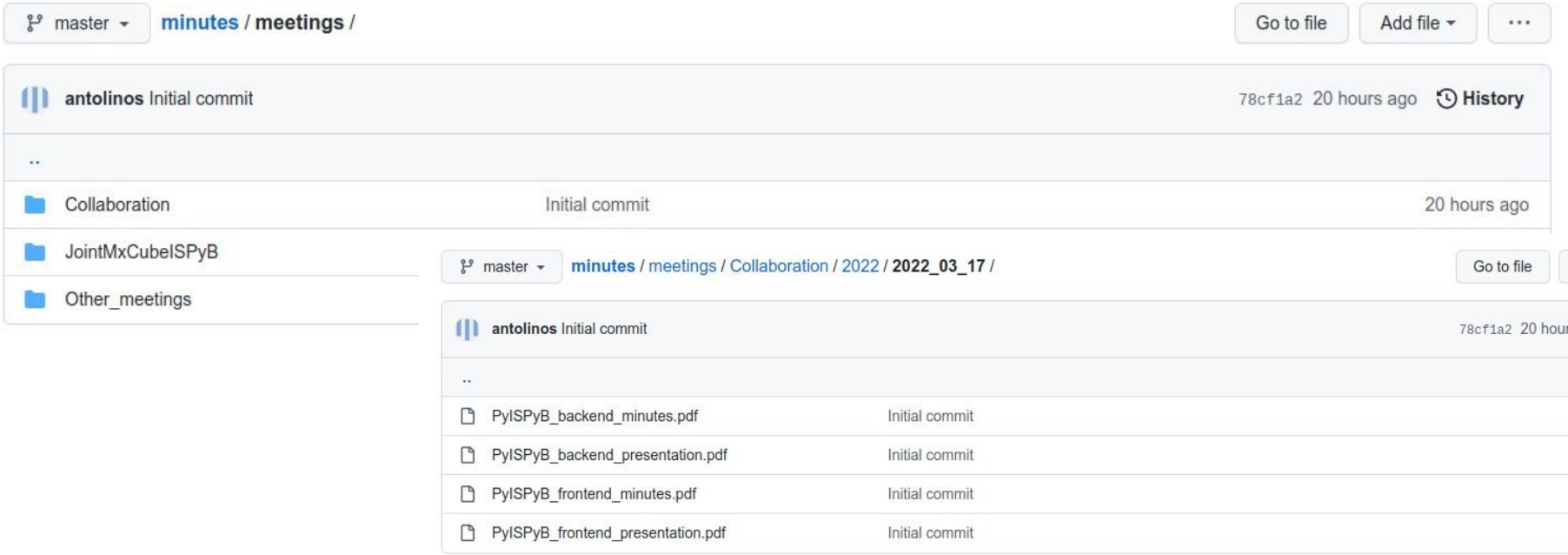
Alex de Maria 9:24 10 may
The easiest might be to use the same database structure of the tables

Alex de Maria 9:25 10 may
If we could define a dedicated group with more restricted permission than manager would be safer

Implementation Approach



Minutes



https://github.com/ispyb/minutes

Conclusions

- ISPyB collaboration has well defined roadmap and a nice momentum
- Enough technical skills and experience to improve the current implementations
- More and more people is actively collaborating
- We still need to encourage all facilities to participate depending on their resources
- Collect feedback from all the actors: users, scientists, industrials, etc…
- Adoption of ISPyB for new facilities
- The implementation of SSX will evaluate the effectiveness of the collaboration

Thanks!