Database Merging status

Solange Delageniere

ESRF | The European Synchrotron

Database merging: why ?

ISPyB meeting : Database merging status | January 18th 2016 I Solange Delageniere
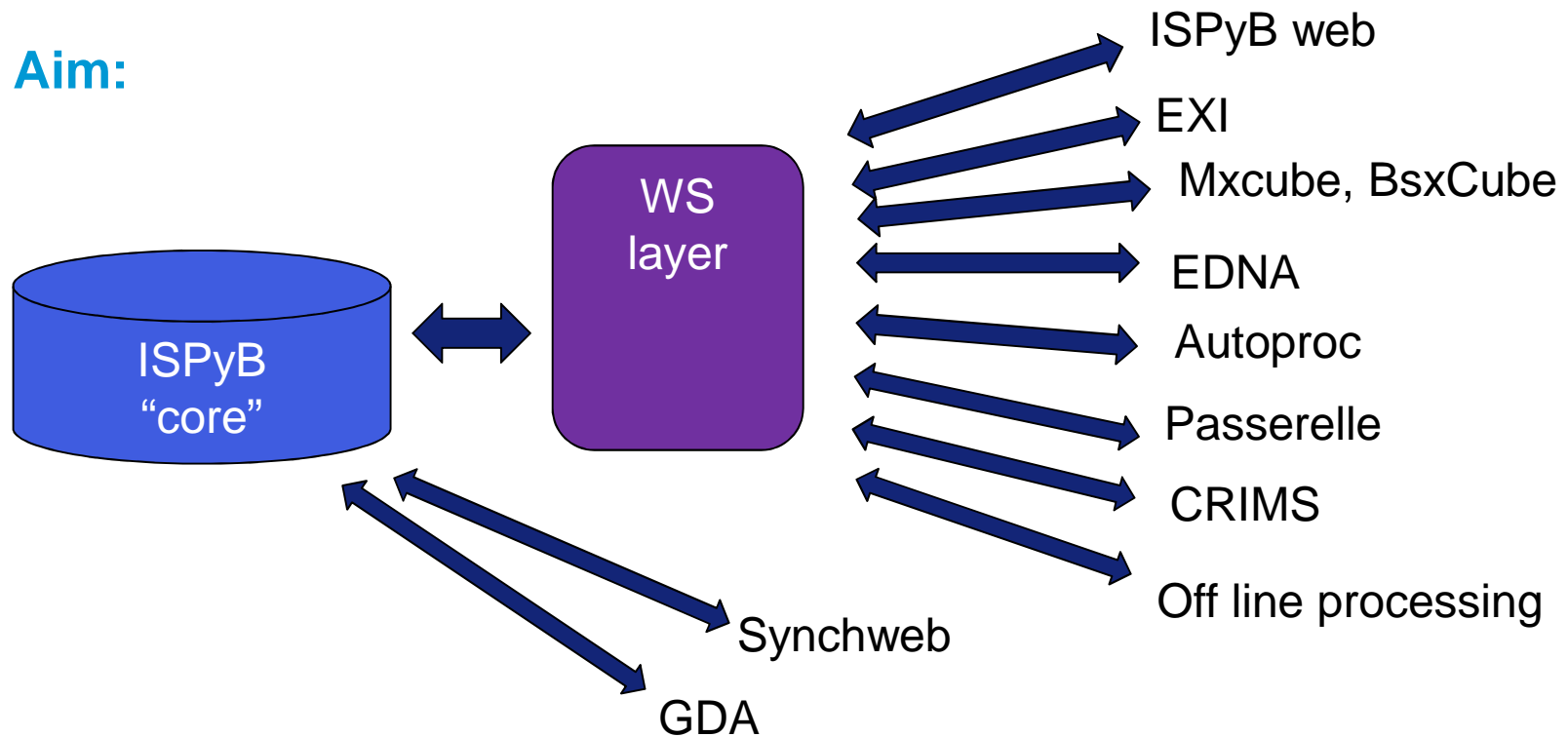
The European Synchrotron | ESRF
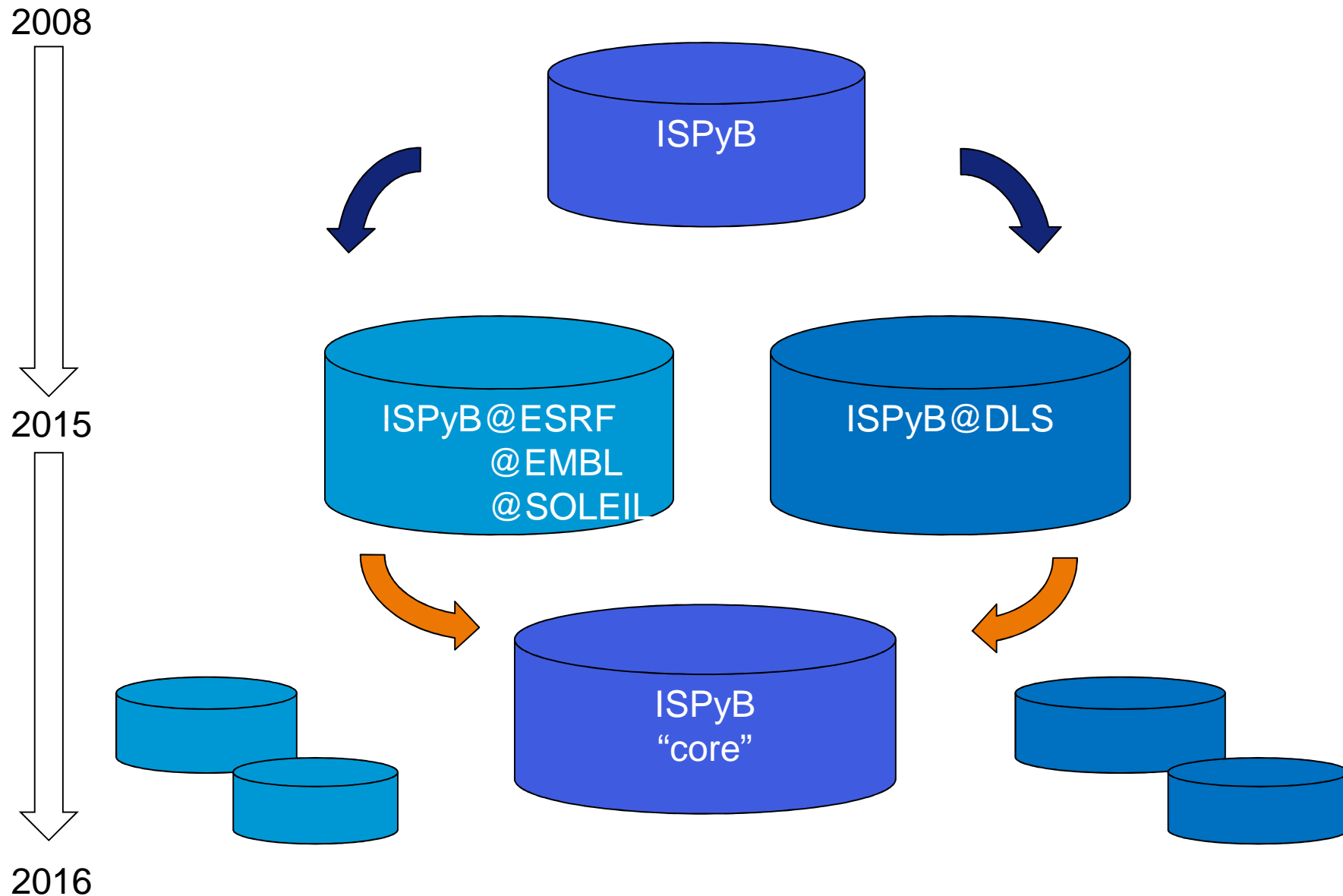
Database merging: why ?

Scope: MX centered → may evolve ?

Let the tables linked to other techniques as they are

Saxs, Em, Tomo

Aim:

ISPyB "core" ↔ WS layer

- ISPyB web
- EXI
- Mxcube, BsxCube
- EDNA
- Autoproc
- Passerelle
- CRIMS
- Off line processing
- Synchweb
- GDA

ISPyB meeting : Database merging status | January 18th 2016 I Solange Delageniere

The European Synchrotron | **ESRF**

Database merging: how ?

Compare the schemas

Common tables

1 by 1 : field by field

To keep or not to keep ?

To add or not to add ?

Compare

Think

Drink

Discuss

Eat cookies

Think
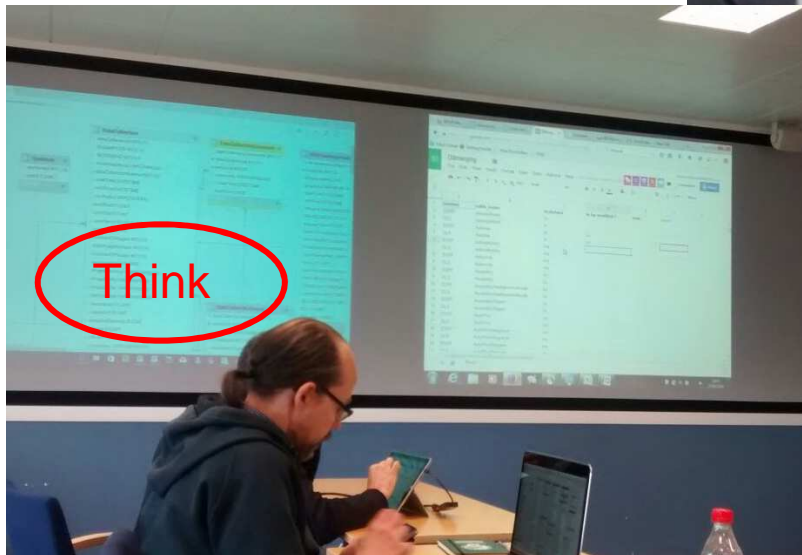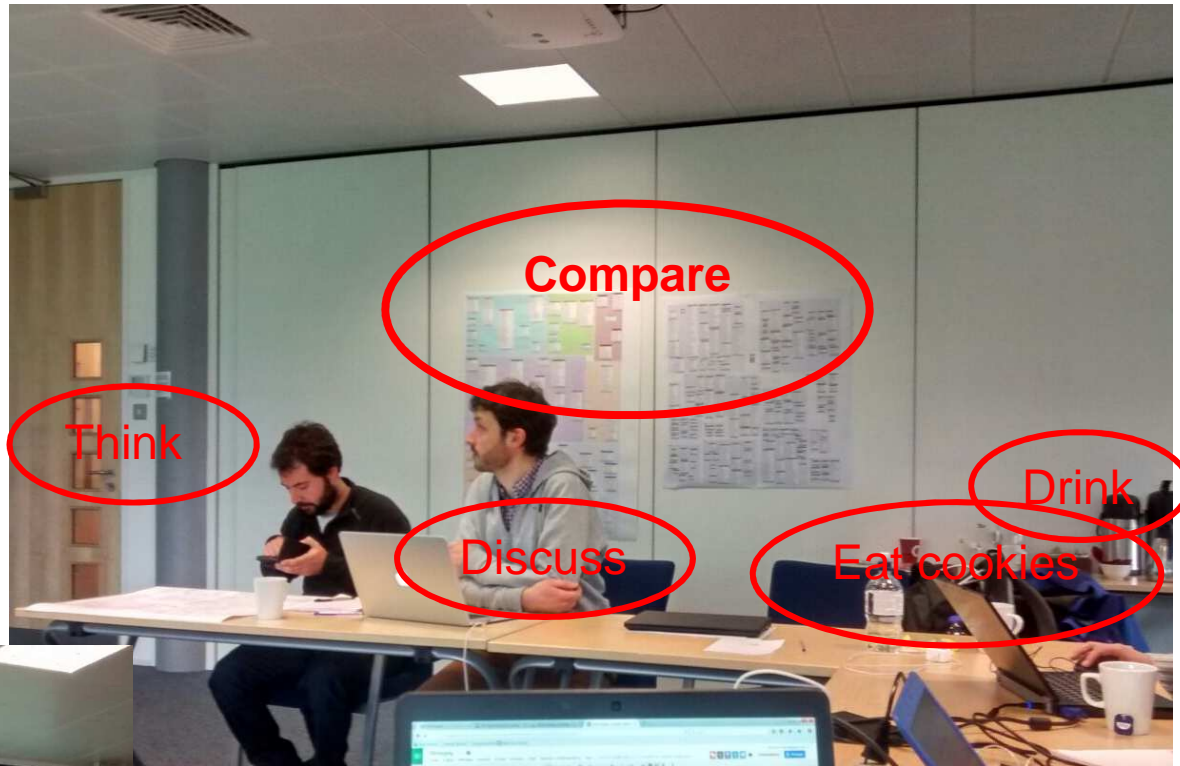
Actions

Missing explanations

Write scripts

Run scripts

The European Synchrotron | ESRF

Database merging: plan

technique	table_name	comments	to add ?	todo: May 20th	todo: June 6th	todo: June 12th
	Aperture		no		will be dropped and aperture size put in DataCollection table	
	BeamApertures	no data no part of official schema	no			
	BeamCentres	no data no part of official schema	no			
	BeamlineStats	no data no part of official schema	no			
	BF_xxx	Beamline faults tables:log book	no			
	BLSampleImage		?	DLS Provides documentation	ESRF might need it adapt them to CRIMS, but the table need documentation	
	BLSampleImageScore		?	DLS Provides documentation	ESRF might need it	ESRF might need it
	BLSampleType_has_component	DLS: why need of Many2Many ?	?	ESRF not sure to use it has to be discussed		BLsampleType will replace Crystal and Component will replace Protein
	BLSession_has_SCposition		yes	DLS provides script	ESRF will add it	ESRF will add it
	CalendarHash	DLS specific	no			
	Component	DLS will later change the Protein table name to Component, and add automatically entries in componentType type for old data	?			ESRF not sure to use it has to be discussed
	Component_has_subType	not convinced MX datamodel shall be merged this more generic model	?			not convinced MX datamodel shall be merged this more generic model
	ComponentSubType	ESRF not sure to use it has to be discussed	?			ESRF not sure to use it has to be discussed
	ComponentType	describes the component for example protein, small molecule,...	?			ESRF not sure to use it has to be discussed
	Containerinspection		no			
	DataCollectionComment		?			
biosaxs?	DataReductionStatus		no			
	DewarRegistry		no			
	DHLTermsAccepted	users have to accept terms - I promise not to ship other things from DLS account -	yes		DLS To change it to CourierTermsAccepted and provide the script	ESRF will integrate it
EM	EMMicroscope		no			
	Imager	formulatrix tray hotel - for plates	no			
plates	Inspection Type		no			
for MR	MXMRRun		yes	DLS to provide a script	- ESRF to add	
for MR	MXMRRunBlob		yes	DLS to provide a script	- ESRF to add	
	Permission	role table linked to UserGroup_has_Permission,	no			
	PhPsession	should be outside the data model	no	to be kept separatly		
	Project tables	not used a lot	yes	DLS will make study to see which ones are used	and then simplify the project model accordingly, provide the creation scripts	
	Protein_has_pdb		yes	DLS provides the script	ESRF will add it	
	RobotAction		yes	DLS provides the script	ESRF will add it	
	Protein_has_pdb		yes	DLS provides the script	ESRF will add it	
	RobotAction		yes	DLS provides the script	ESRF will add it	
	Schedule	equivalent to CRIMS	no			
	ScheduleComponent	equivalent to CRIMS	no			
	Screen	equivalent to CRIMS	no			
	ScreenComponent	equivalent to CRIMS	no			
	SessionType		?			not sure ESRF needs it , all info is stored in SMIS and avoid duplicating data. To be discussed with collaboration.

Database merging: in progress…

Initial status

Total nb of tables : > 100

Common tables: 48, but several missing fields

Quite some work to be done …

How ?

Write scripts

Run scripts

Exchange emails

Discuss

Database merging: discussions

Ex: "We would like some new tables/fields"

OK to add the "easy" ones:

add new fields in existing tables, when these new fields can be null as long as we have the update scripts:
- externalId columns

OK in some cases

When the new field is a foreign pointing to an existing table, and agreement.

To be discussed later → meeting

When the new fields are foreign keys pointing to new tables.

Some new tables where we think we need to have a discussion about them.

ISPyB meeting : Database merging status | January 18th 2016 I Solange Delageniere

The European Synchrotron | ESRF

Database merging: merged items

December 2016

Total nb of tables : 125 / 133

Common tables: 79 very small discrepancies

DLS extra tables: 54, 13 for xi, 4 for perm, 12 for BF, 11 for project

ESRF extra tables: 45 for BioSaxs, 1 for admin

Remaining points:

DataCollection table

Links to tables outside the "core" -> foreign keys to external tables

The European Synchrotron | ESRF

Database merging: links to tables outside the "core"

Solution 1:

Do not add foreign keys fields, not add extra tables
pros: very clean on the "core" side, no empty tables
cons: the schemas are not really merged, and if we share the API we have to be careful so that it is consistent, have a API for the "core"and a API with the extra field/table

Solution2:

Add the fields but not as foreign keys
pros: no need to add empty tables, but extra empty fields
cons: scripts to be run to redefine the foreign keys to use the site specific tables

Solution3:

Add all foreign keys with the extra tables linked
pros: schema merged
cons: the "core" may be huge with a lot of empty tables, and we have to check that foreign keys pointing to empty tables do not raise errors.

The European Synchrotron | ESRF

Database merging: remaining items

Remaining points:

- DataCollection table

- Links to tables outside the "core" -> foreign keys to external tables

- Scope: Keep all or reduce and have a core ? How to define the core ?

  -> How to indicate tables from the core, and extra tables ?   prefixes ? **yes**

**to be discussed in the STC
and dev. meeting**

ISPyB meeting : Database merging status | January 18th 2016 I Solange Delageniere

The European Synchrotron   **ESRF**

Database merging: conclusion

Almost there !

    ->Track carefully the DB modifications

Thanks to:

- **Alejandro DeMaria,**
- **Ivars Karpics,**
- **Karl Levik,**
- **Stuart Fisher**

The European Synchrotron | **ESRF**