

openBIS and Galaxy

Manuel Kohler

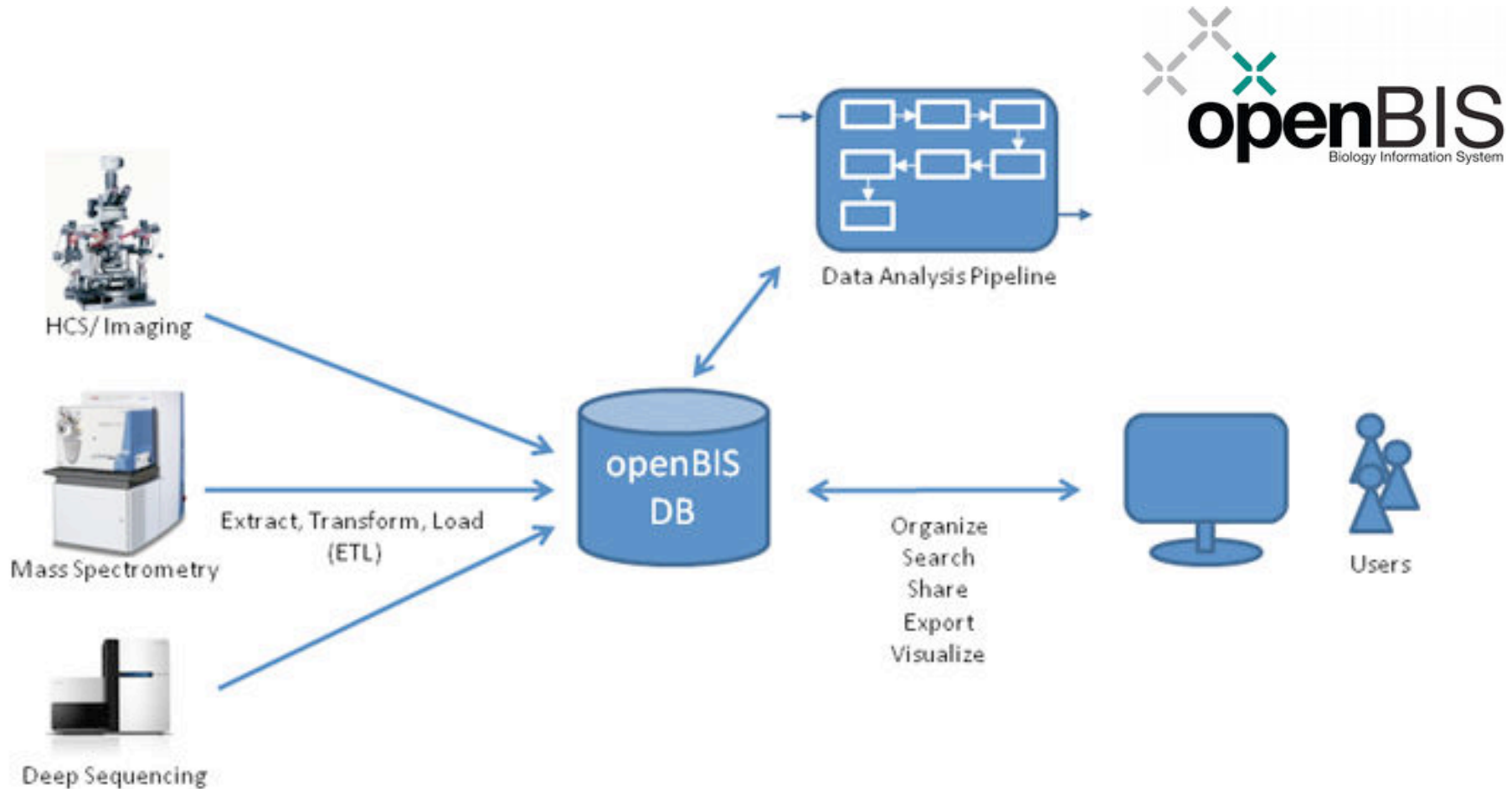
Center for Information Sciences and Databases (C-ISD)



OUTLINE

- openBIS
- Goal
- Integration (Proof-of-Concept)
- Problems

What is openBIS?



Hierarchy: Space → Project → Experiment → Sample → Data Set → File

BSSE_QGF_TEST » QGF_TEST » 2012.09 » Sample 120917_SN792_0158_AC125KACXX:4 [ILLUMINA_FLOW_LANE]

Sample Properties

Sample	/BSSE_QGF_TEST /120917_SN792_0158_AC125KACXX:4
PermID	20120917113811524-60404917
Sample Type	ILLUMINA_FLOW_LANE
Registrar	Server, BSSE ETL
Registration Date	2012-09-17 11:38:11
Project	/BSSE_QGF_TEST/QGF_TEST
Experiment	/BSSE_QGF_TEST/QGF_TEST/2012.09
Parents	97
Container	/BSSE_FLOWCELLS /120917_SN792_0158_AC125KACXX [ILLUMINA_FLOW_CELL]
Data transferred	2012-09-21 12:14:47 +0200
Common Parent Properties:	
Name of Contact Person	Ina Nissen
Email of Contact Person	ina.nissen@bsse.ethz.ch
Name of Principal Investigator	Christian Beisel
Email of Principal Investigator	christian.beisel@bsse.ethz.ch
Sample Kind	GENOMIC_DNA
Sequencing Application	GENOMIC_DNA_SEQ
Library preparation kit	NEXTERA XT DNA SAMPLE PREPARATION KIT(ILLUMINA)
Unique Barcodes in Pool?	OK
/BSSE_QGF_TEST/BSSE-QGF-8341 [ILLUMINA_SEQUENCING]:	
Affiliation	Other
External Sample Name	275
Organism (NCBI Taxonomy)	Erwinia amylovora
Index 1	Index1 (i7) N707 CTCTCTAC
Index 2	Index2 (i5) S506 ACTGCATA
Calculated DNA concentration of library (nM)	0
Calculated ul DNA for 2nM stock	0
Calculated ul EB for 2nM stock	0
Notes	Isolation: Stammheim Position on Plate: F7 15 Swiss old
Number of Attachments	0
Barcode_Complexity_Checker	No Index
/BSSE_QGF_TEST/BSSE-QGF-8342 [ILLUMINA_SEQUENCING]:	

Contained Children Parents **Data Sets** History Attachments Queries

Data Sets						<input checked="" type="checkbox"/> directly co
Code	Data Set Type	Registration Date	Index 1	Sample	Pro	
20120921200631770-60405080	QUALITY_PDFS	2012-09-21 20:06:33 GMT+02:00		120917_SN792_0158_AC125KACXX:4	QGF	
20120921191159478-60405063	FASTQC	2012-09-21 19:12:01 GMT+02:00		120917_SN792_0158_AC125KACXX:4	QGF	
20120921121444499-60405021	FASTQ_GZ	2012-09-21 12:15:14 GMT+02:00	Index1 (i7) N705 C	120917_SN792_0158_AC125KACXX:4	QGF	
20120921121445729-60405022	FASTQ_GZ	2012-09-21 12:15:14 GMT+02:00	Index1 (i7) N712 C	120917_SN792_0158_AC125KACXX:4	QGF	
20120921121445889-60405023	FASTQ_GZ	2012-09-21 12:15:14 GMT+02:00	Index1 (i7) N704 T	120917_SN792_0158_AC125KACXX:4	QGF	
20120921121446060-60405024	FASTQ_GZ	2012-09-21 12:15:14 GMT+02:00	Index1 (i7) N701 T	120917_SN792_0158_AC125KACXX:4	QGF	
20120921121446226-60405025	FASTQ_GZ	2012-09-21 12:15:14 GMT+02:00	Index1 (i7) N702 C	120917_SN792_0158_AC125KACXX:4	QGF	
20120921121446396-60405026	FASTQ_GZ	2012-09-21 12:15:14 GMT+02:00	No Index	120917_SN792_0158_AC125KACXX:4	QGF	
20120921121446483-60405027	FASTQ_GZ	2012-09-21 12:15:14 GMT+02:00	Index1 (i7) N706 T	120917_SN792_0158_AC125KACXX:4	QGF	
20120921121446616-60405028	FASTQ_GZ	2012-09-21 12:15:14 GMT+02:00	Index1 (i7) N707 C	120917_SN792_0158_AC125KACXX:4	QGF	
20120921121446735-60405029	FASTQ_GZ	2012-09-21 12:15:14 GMT+02:00	Index1 (i7) N703 F	120917_SN792_0158_AC125KACXX:4	QGF	
20120921121446856-60405030	FASTQ_GZ	2012-09-21 12:15:14 GMT+02:00	Index1 (i7) N711 F	120917_SN792_0158_AC125KACXX:4	QGF	
20120921121446994-60405031	FASTQ_GZ	2012-09-21 12:15:14 GMT+02:00	Index1 (i7) N706 T	120917_SN792_0158_AC125KACXX:4	QGF	
20120921121447208-60405032	FASTQ_GZ	2012-09-21 12:15:14 GMT+02:00	Index1 (i7) N708 C	120917_SN792_0158_AC125KACXX:4	QGF	
20120921121447368-60405033	FASTQ_GZ	2012-09-21 12:15:14 GMT+02:00	Index1 (i7) N710 C	120917_SN792_0158_AC125KACXX:4	QGF	
20120921121447516-60405034	FASTQ_GZ	2012-09-21 12:15:14 GMT+02:00	Index1 (i7) N709 C	120917_SN792_0158_AC125KACXX:4	QGF	
20120920210309276-60405014	RUNINFO	2012-09-20 21:04:32 GMT+02:00		120917_SN792_0158_AC125KACXX:4	QGF	

Goal

A. Make openBIS a data source in Galaxy

Get Data

- Upload File from your computer
- openBIS : gets data set

B. And ideally make openBIS a data sink in Galaxy as well

C. Finally upload to the Tool Shed

Experiment

openBIS (step 1 of 4)

Experiment:

select Experiment

Next step

What it does

Gets data set from openBIS

Sample

openBIS (step 2 of 4)

Sample:

120420_SN792_0109_BC0P8LACXX

select Sample

Next step

What it does

Gets data set from openBIS

Data Set

openBIS (step 3 of 4)

Data Set:

20120807161244695-4183

select Data Set

Next step

What it does

Gets data set from openBIS

FASTQ File

openBIS (step 4 of 4)

Files:

select Files

Execute

What it does

Gets data set from openBIS

```
<tool id="openbis_get_data_set" name="openBIS" version="1.0">
  <description>: gets data set</description>
  <command interpreter="python">getFileFromOpenBIS.py $dataset $file $userEmail $output</command>
<inputs>
  <page>
    <param name="email" type="hidden" value="$__user_email__" />

    <param name="experiment" type="select" label="Experiment" help="select Experiment"
      refresh_on_change="true" dynamic_options="getExperiments()"/>
  </page>
  <page>
    <param name="sample" type="select" label="Sample" help="select Sample"
      refresh_on_change="true" dynamic_options="getSamplesForExperiments(experiment)"/>
  </page>
  <page>
    <param name="dataset" type="select" label="Data Set" help="select Data Set"
      refresh_on_change="true" dynamic_options="getDataSetforSamples(sample, experiment)"/>
  </page>
  <page>
    <param name="file" type="select" label="Files" help="select Files"
      refresh_on_change="true" dynamic_options="getFiles(dataset)"/>
  </page>
</inputs>
<outputs>
  <data format="fastqsanger" name="output" label="more foo"/>
</outputs>
<options refresh="True"/>
<code file="openbis_get_dataset.py" />
<help>
  **What it does**
```

Not Good

Use of the deprecated `<code>` tag:

“**Deprecated** do not use this unless absolutely necessary. [...] These hooks are being replaced by new tool config features and methods in the `~/lib/galaxy/tools/__init__.py` code file.”

- So what to use instead?

Critical Blocker:

How to make the openBIS authorization work in Galaxy?

```
<param name="email" type="hidden" value=${__user_email__} />
```

```
<param name="experiment" type="select" label="Experiment"  
help="select Experiment" refresh_on_change="true"  
dynamic_options="getExperiments(email)"/>
```

This does not work. Ideally I would like to have something like this:

```
dynamic_options="getExperiments(${__user_email__})"
```

No Answer from the mailing list...

Example workflow for not well documented tags



Meta Data in Galaxy

Add more meta data fields to a data set?

External Sample Name	Input_R3(IDX1)
Cycles	50
Name of Contact Person	Ina Nissen
Email of Contact Person	ina.nissen@bsse.ethz.ch
Name of Principal Investigator	Christian Beisel
Email of Principal Investigator	christian.beisel@bsse.ethz.ch
Sample Kind	GENOMIC_DNA
Affiliation	Other
External Sample Name	275
Organism (NCBI Taxonomy)	Erwinia amylovora
Index 1	Index1 (i7) N707 CTCTCTAC
Index 2	Index2 (i5) S506 ACTGCATA



Edit Attributes

Name:

Stam_HEK293_CTCF_1_SRR299276.fastq

Info:

uploaded fastq sanger file

Annotation / Notes:

None

Add an annotation or notes to a dataset; an

Database/Build:

Human Feb. 2009 (GRCh37/hg19) (hg19)

Save

Auto-detect

This will inspect the dataset and attempt to

Summary

- openBIS and Galaxy can be integrated as soon as the missing piece of authorization is solved
- Documentation for Galaxy could be better
- Integration of custom code should be clear
- Meta data should be supported better

