Abstract

openBIS^[1] is an extensible, open source software framework for constructing user-friendly, scalable and powerful information systems for data and metadata acquired in biological experiments. It enables users to collect, integrate, share, publish data and to connect to data processing pipelines.

The **Single Cell Unit** is a centralized scientific facility at the D-BSSE^[2] that provides equipment in the areas of flow cytometry, advanced automated microscopy and laboratory automation and maintains a department-wide openBIS installation for the management of all data generated on the facility's acquisition stations.

The openBIS Importer Toolset is a tightly integrated collection of tools that allows for the semiautomated, semi-unsupervised registration of annotated datasets into openBIS directly from the acquisition stations.

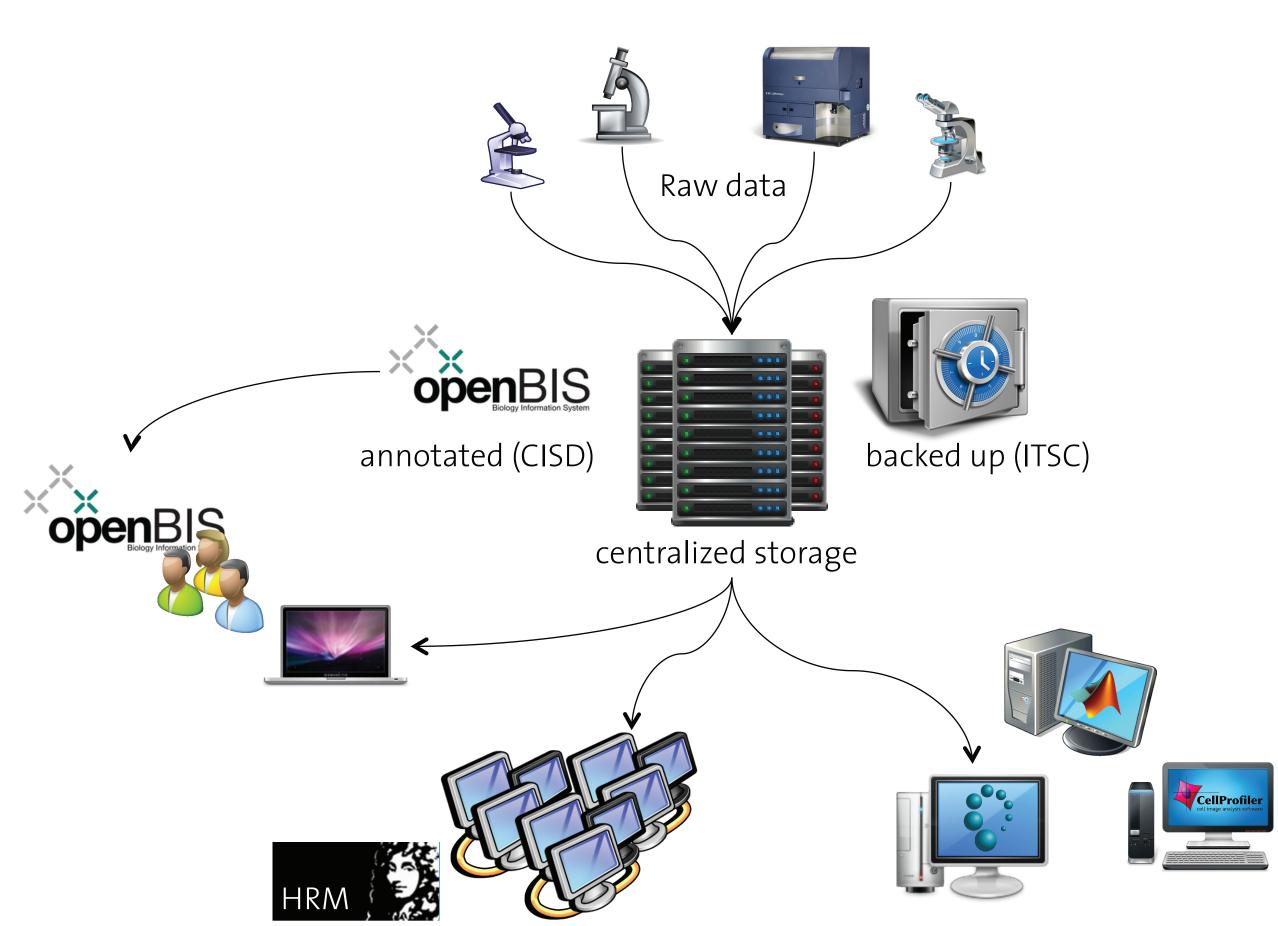


Figure 1: The Single Cell Unit in collaboration with the IT Services and Consulting group (ITSC^[3]) and the Center for Information Sciences and Databases (CISD^[4]) provides centralized, secured storage for all the raw data acquired in the facility. Associated metadata is stored in a department-wide installation of openBIS. Data registered in openBIS will be easily accessible for further analysis and processing work. The SCU provides high-end workstations and dedicated processing servers (e.g. the Huygens Remote Manager). Alternatively, data can be accessed from users' machines. Group instances of openBIS can access data in the SCU openBIS instance via linked datasets. This allows group to organize their openBIS instance as desired and without any data duplication.

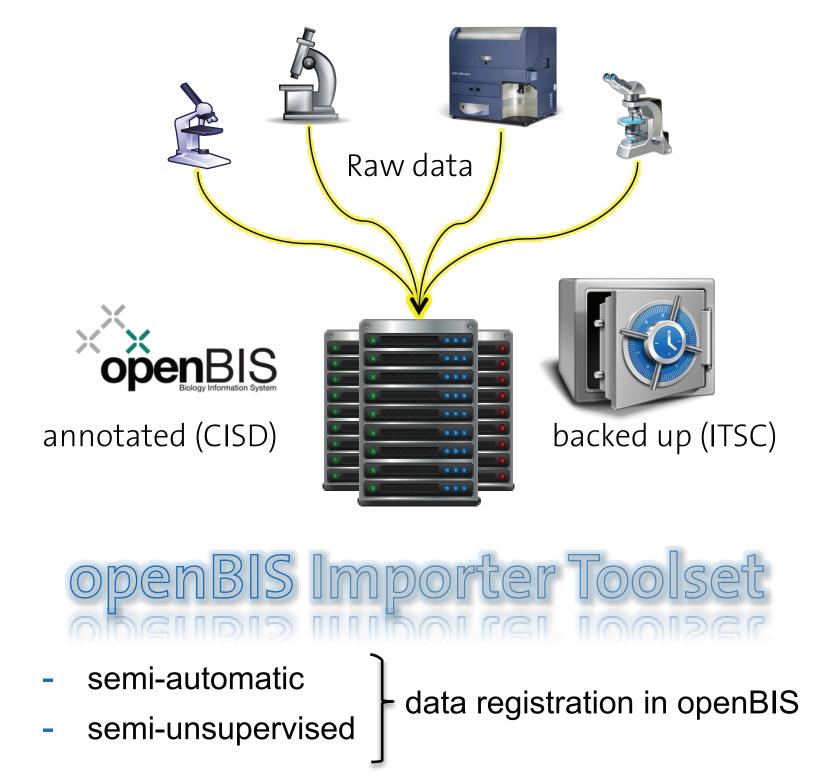


Figure 2: The openBIS Importer Toolset provides tools for the semi-automatic, semi-unsupervised registration of data and associated metadata into openBIS (upper portion of Figure 1). It also extends openBIS by providing dedicated views (as full-fledged HTML5 web apps).

openBIS Importer Toolset

Aaron Ponti, Single Cell Unit, Department of BioSystems Science and Engineering (D-BSSE), ETH Zurich (Basel) aaron.ponti@bsse.ethz.ch

eter) from BD BioSciences as an example.

Workflow

BD LSRFortessa cell analyzer Acquisition station 2 Scan 🛯 🛛 Export **Datamover** as a service Python openBIS

 Image: Comparison of Compar Single Cell Unit openBIS Please login to start your session User: Password: Click here for help.

Figure 3: The openBIS Importer toolset is composed of several tools and applications: Annotation Tool, Datamover as a service, Python dropbox scripts, and custom openBIS views (HTML5 apps).

Annotation Tool

The **Annotation Tool** guides the process of data registration into openBIS.

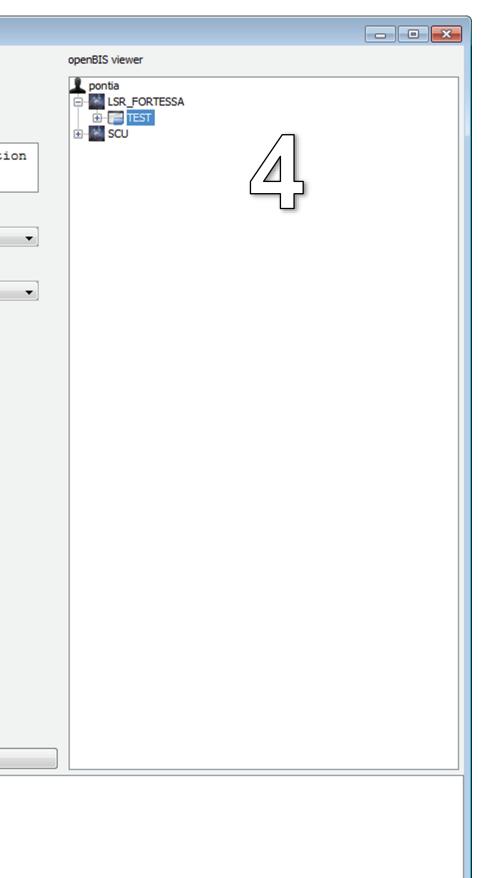
Data viewer	Metadata editor
<pre>/pontia /pontia Kymb 090512 FCS EXport % Kymb 090512 % Symb 090512 % Symb 090512 % Quality Control &Peak Quality_Control_MD % Symb 090512 % Symb 090512 % Symb 090512 % 488 and 405 excitation % 488 and 405 excitation % Only 405 excitation % Only 405 excitation % Tube_014 % Tube_015 % Tube_016</pre>	<pre> Quality_Control_MD Kymb 090512 Description Yeast mtFP1 and Citrin. 405 and 445-nm excitate comparison. Plate geometry 96_WELLS_8X12 Target openBIS project /LSR_FORTESSA/TEST 55 55 55 55 55 55 55 55 55</pre>
Scan Metadata viewer Name Value \$BEGINANALYSIS 0 \$BEGINDATA 3029 \$BEGINSTEXT 0 \$BTIM 15:02:38 \$BYTEORD 4,3,2,1 \$CYT LSRII \$DATATYPE F \$DATE 04-APR-2012 \$ENDDATA 6420.39 Invalid datasets Same Same Same Same Same Same Same Same	The second secon
18-06-2013 11:06:23: Ready. 18-06-2013 11:06:31: Logging in to openBIS 18-06-2013 11:06:46: Scanning user data folder 18-06-2013 11:06:47: Successfully processed folder "Kymb 090512 FCS EXpor 18-06-2013 11:06:47: Successfully processed folder "Mark" 18-06-2013 11:06:47: Successfully processed folder "Quality Control 8Peak" 18-06-2013 11:06:47: Successfully processed folder "Quality_Control_MD"	

Figure 4: The Annotation Tool scans the Datamover incoming folder and displays the experiments in the data viewer in the most familiar form for the user (1). Each element in the data viewer tree has metadata associated, that is displayed in the metadata viewer (2) upon selection. Here, the metadata stored in the selected FCS file is shown. The Annotation Tool also performs validation on the content of the incoming folder, and lists all invalid datasets in (3); these must be fixed or removed before registration is possible. The openBIS viewer (4) displays all entities the user is permitted to access (currently, read-only operations are allowed). In the metadata editor (5), additional metadata is associated to the various experiments, and the target openBIS project is chosen for registration. We are now defining a minimum required set of experiment metadata (not present in the datasets) that should be provided upon registration. When metadata is ready, the user clicks on upload to openBIS (6) and all information is saved as an XML file along with the experiments in the incoming folder. The XML file signals Datamover that the dataset is ready to be transferred to the openBIS Data Store Server.

The following figure depicts the registration workflow with the LSRFortessa cell analyzer^[5] (flow cytom-







Datamover as a service

Datamover^[5] is a program that takes care of moving typically large amount of data to a (usually) remote central storage. Running in the background, it checks for new, incoming data periodically and starts the moving process - optionally under the control of scripts (hand-shaking mechanism). In particular, the presence of a special XML file generated by the Annotation Tool triggers the operation of Datamover. Datamover was extended to run as a Windows service.

Python dropboxes

The python (or, more precisely, jython) dropbox feature makes it possible for a script written in the python language to control the data set registration process of the openBIS Data Store Server. A script can modify the files in the dropbox and register data sets, samples, and experiments as part of its processing. The framework provides tools to track file operations and, if necessary, revert them, ensuring that the incoming file or directory is returned to its original state in the event of an error. Machine/file format specific dropbox scripts are provided that make use of the XML file generated by the Annotation Tool to register the datasets and all associated metadata information into openBIS. In the future, **attachments** will be supported as well.

openBIS custom views (HTML5)

openBIS supports the integration of customized views in the form of HTML 5 web applications. We make use of this functionality to provide the users with a consistent view into the registered datasets and to integrate additional visualizations and analysis tools.

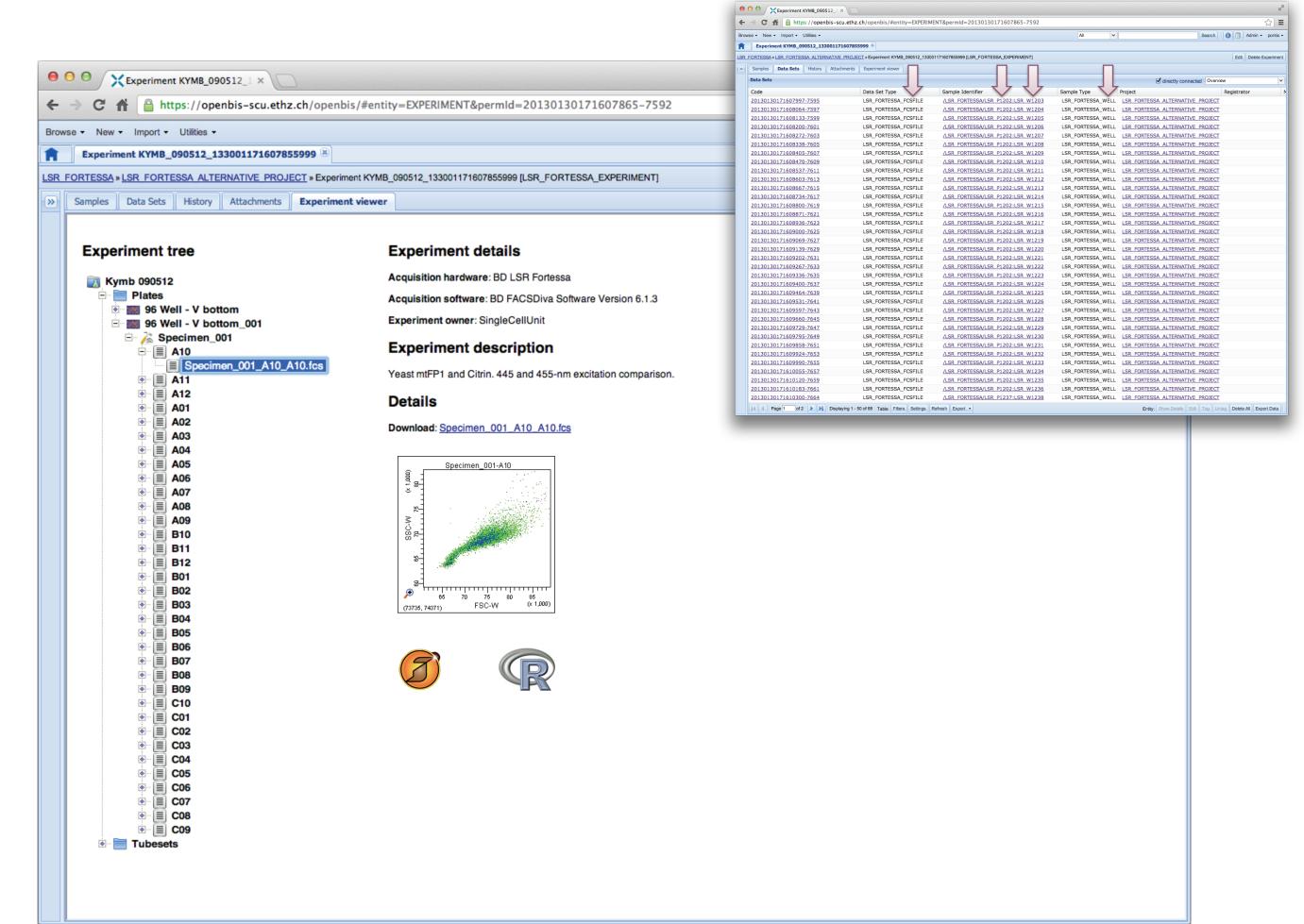


Figure 5: Custom openBIS views recreate a familiar data organization for the user (in this case the tree structure that is also found in the Annotation Tool replicates the data visualization found in the BD FACS DIVA software that drives the acquisition on the LSR Fortssa flow cytometer), in contrast to the standard openBIS view shown in the inset on the topright corner of the figure. Specific analysis pipelines will in future be started directly from within openBIS.

Summary and outlook

In summary, we are developing tools to streamline registration of acquired data and metadata into openBIS and to extend openBIS via custom views and apps for working with data. Some of the future steps will be to (1) build a plug-in framework for hardware specialization, (2) implement tools for easier user management, (3) add functionality to create/delete openBIS entities in the Annotation Tool, (4) integrate more and more (interactive) analysis functionalities, and (5) investigate synergies with **OMERO**.

References

- [1] http://www.cisd.ethz.ch/software/openBIS
- [2] http://www.bsse.ethz.ch/
- [3] http://www.bsse.ethz.ch/administration/it ITSC