

Galaxy and UC13

Cyclone usecases Hackathon, 15 nov 2016.

Samuel Blanck, EA 2694, Université Lille 2 Droit et Santé

What is Galaxy ?



- “Online bioinformatics analysis for everyone”
- Open, web-based platform for data intensive biomedical research
- Academic Free License version 3.0
- Penn State, John Hopkins University
- www.galaxyproject.org

What is Galaxy ?

The screenshot displays the Galaxy web interface. On the left is the 'Tools' panel with a search bar and a list of tool categories including Get Data, Send Data, Lift-Over, Text Manipulation, Datamash, Convert Formats, Filter and Sort, Join, Subtract and Group, Fetch Alignments/Sequences, NGS: QC and manipulation, NGS: DeepTools, NGS: Mapping, NGS: RNA Analysis, NGS: SAMtools, NGS: BamTools, NGS: Picard, NGS: VCF Manipulation, NGS: Peak Calling, NGS: Variant Analysis, NGS: RNA Structure, NGS: Du Novo, NGS: Gemini, NGS: Assembly, Operate on Genomic Intervals, and Statistics.

The central workspace contains a text block stating: "Galaxy is an open source, web-based platform for data intensive biomedical research. If you are new to Galaxy [start here](#) or consult our [help resources](#). You can install your own Galaxy by following the [tutorial](#) and choose from thousands of tools from the [Tool Shed](#)." Below this is a tweet from the Galaxy Project (@galaxyproject) dated 28-30 Nov, mentioning an advanced sequence analysis workshop in Bordeaux. The tweet includes logos for GCC 2017 Montpellier and cgfb Bioinformatique. At the bottom of the workspace are logos for Penn State, Johns Hopkins, TACC, and CyVerse.

On the right is the 'History' panel, showing a search bar and a list of datasets under 'Unnamed history'. The datasets are numbered 1 through 10, with names like 'galaxy_demo.txt', 'Change Case on data 1', 'Examples liste_sondes.csv', 'Convert on data 3', 'Compute on data 5', 'Sort on data 6', and 'Sort on data 6'. Each entry has icons for viewing, editing, and deleting.

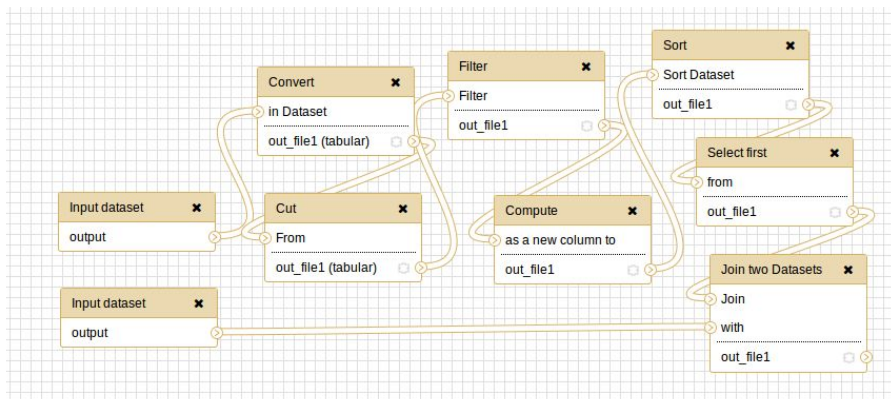
What is Galaxy ?

- Database
 - Sqlite (by default)
 - Possibility to switch to MySQL or PostgreSQL
- Web server
 - Built-in HTTP server, written in Python

Galaxy also supports TORQUE PBS, PBS Pro, Platform LSF, and Sun Grid Engine clusters

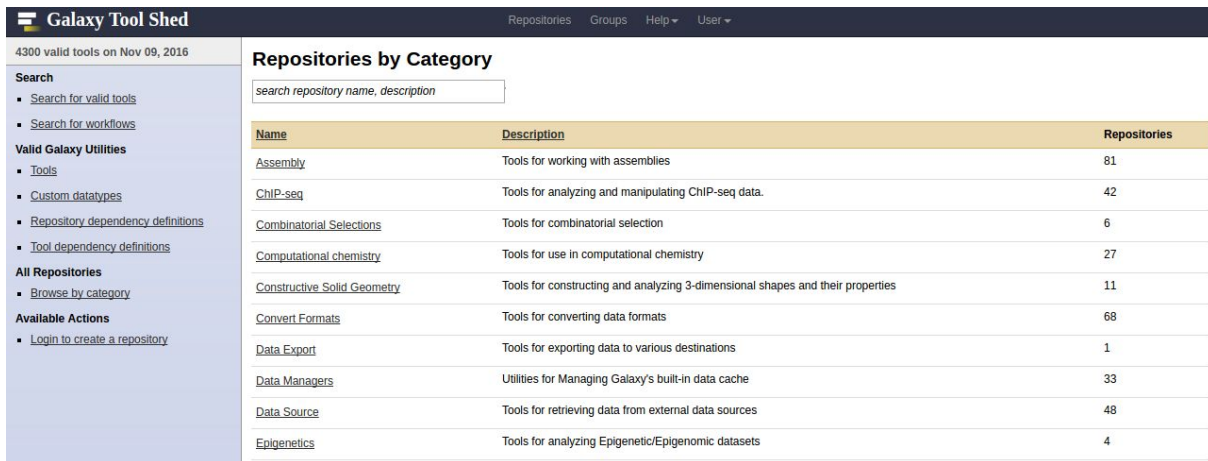
What is Galaxy ?

- “Push-button” interface, easier than command line
- Keep track of analysis
- Possibility to create workflows
- Share results, histories and workflows
- Greater reproducibility of the analysis



What is Galaxy ?

- Galaxy is customizable
- Possibility to integrate custom tools thanks to a “XML wrapper” framework
- Galaxy toolshed is the Galaxy appstore
- Thousands of tools available

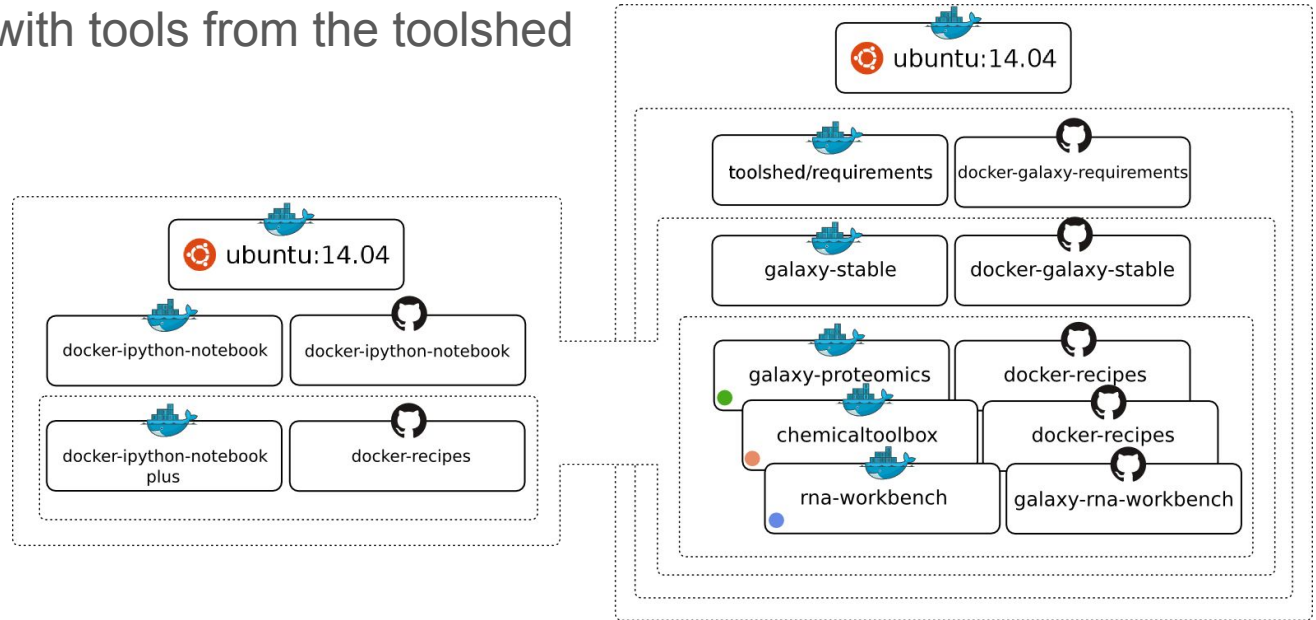


The screenshot displays the Galaxy Tool Shed web interface. At the top, a dark navigation bar contains the 'Galaxy Tool Shed' logo and links for 'Repositories', 'Groups', 'Help', and 'User'. Below the navigation bar, a light blue sidebar on the left provides navigation options, including a search bar, links for 'Valid Galaxy Utilities' (Tools, Custom datatypes, Repository dependency definitions, Tool dependency definitions), 'All Repositories' (Browse by category), and 'Available Actions' (Login to create a repository). The main content area, titled 'Repositories by Category', features a search input field and a table listing various tool categories and their associated repository counts.

Name	Description	Repositories
Assembly	Tools for working with assemblies	81
ChIP-seq	Tools for analyzing and manipulating ChIP-seq data.	42
Combinatorial Selections	Tools for combinatorial selection	6
Computational chemistry	Tools for use in computational chemistry	27
Constructive Solid Geometry	Tools for constructing and analyzing 3-dimensional shapes and their properties	11
Convert Formats	Tools for converting data formats	68
Data Export	Tools for exporting data to various destinations	1
Data Managers	Utilities for Managing Galaxy's built-in data cache	33
Data Source	Tools for retrieving data from external data sources	48
Epigenetics	Tools for analyzing Epigenetic/Epigenomic datasets	4

Galaxy + Docker

- Several “flavours” of dockerized Galaxy
- Easy to deploy
- Easy to populate with tools from the toolshed



Where to find Galaxy ?

Many public servers available

- usegalaxy.org (“Galaxy team” server)
- In France : ABiMS (Roskoff), GalaxEast (Strasbourg), Genouest (Rennes),

In the clouds

- AMIs on AWS
- IFB cloud

Galaxy and IFB

Several Galaxy VM are available on IFB-core stratuslab cloud :

- Galaxy FROGS
- Galaxy STACKS
- Galaxy MODAL

bilille (IFB Platform of Lille)

- Test Galaxy VM on Openstack

But it's not possible to share data between users

UC13 workflow

