

# Galaxy Metabolomics from the Ground Up

Art Eschenlauer  
Mark Esler  
Tim Griffin  
Adrian Hegeman



University of Minnesota - Twin Cities

# Overview

- Organization and analysis of plant metabolomics data
- Why we chose Galaxy
- Why we chose to set up our own Galaxy server
- How well did it work for us?
- Could our solution be applied elsewhere?

# Mass-Spectrometry Based Plant Metabolomics

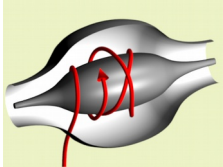
Samples



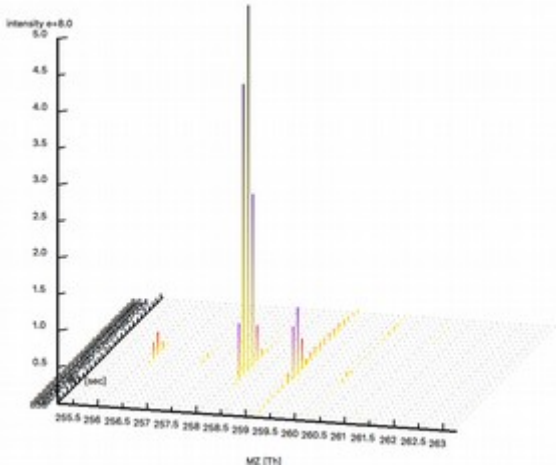
Sample extracts



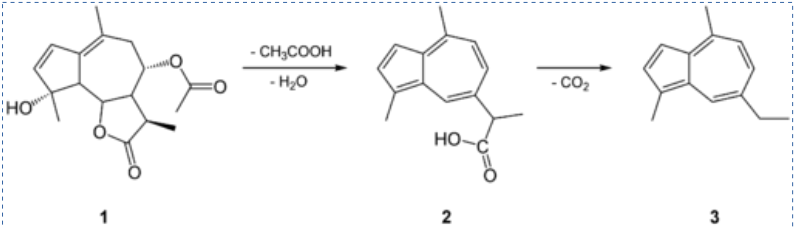
LC-MS  
or  
GC-MS



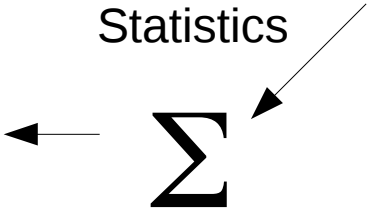
Mass chromatogram



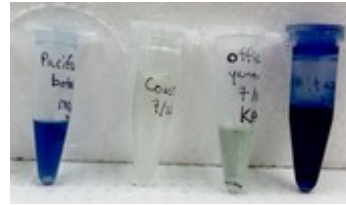
Metabolite abundances



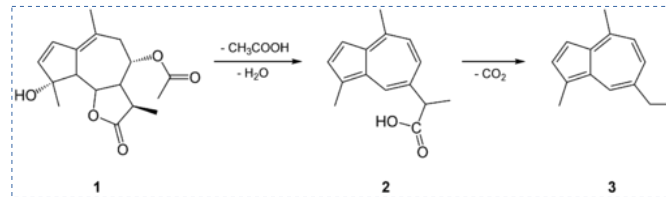
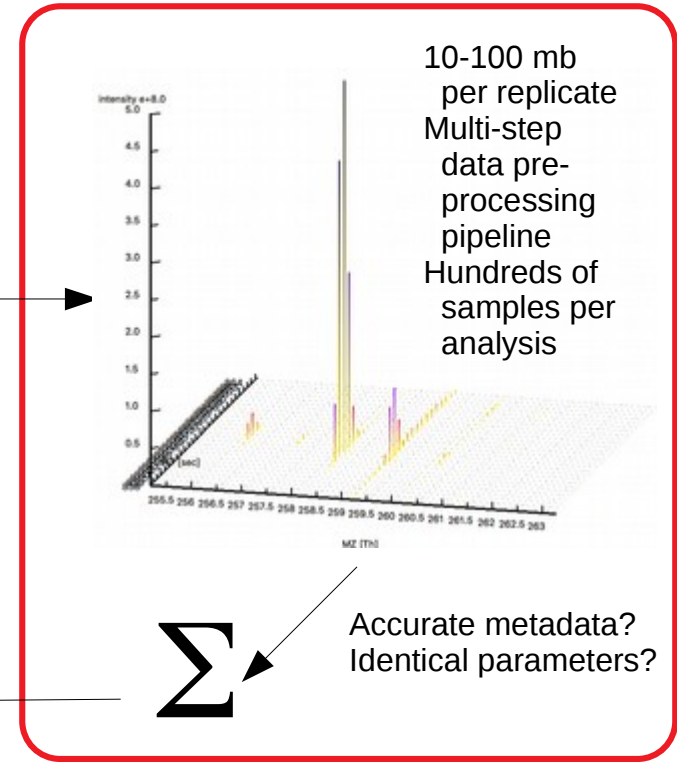
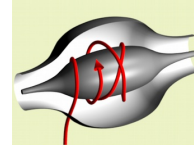
Statistics



# Data-Processing Challenges



Multiple replicates per sample. Blanks. QC Pools.



Accurate metadata?  
Identical parameters?

# Proprietary Solutions

## Value proposition:

- Easy to use for common purposes
- High quality
- Good support

See [J. Proteome Res. 2018, 17, 4, 1335-1339](#)



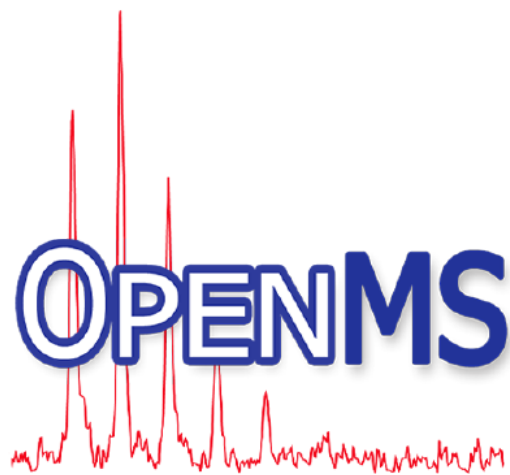
## General issues:

- Cost barriers:
  - To adoption
  - To collaboration
- Slow release cycle:
  - Vendor dependent
  - Staff limited
- Rigid feature set:
  - Monolithic
  - Not customizable
  - Not adaptable
- Not inspectable

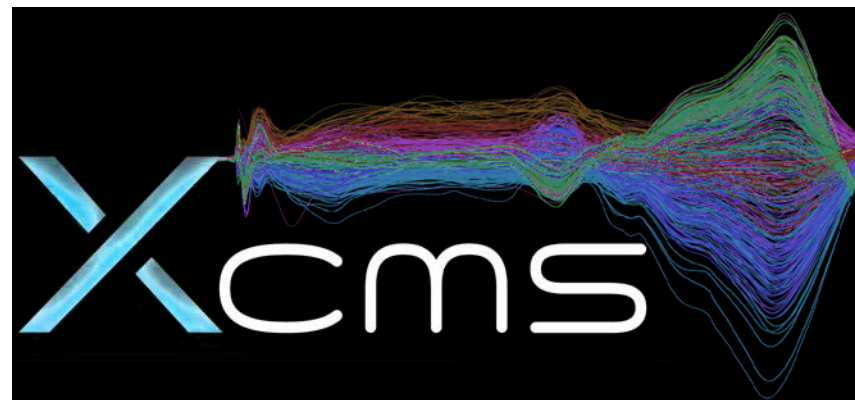
# Popular Open Source Solutions for Metabolomics



<https://mzmine.github.io>



<https://www.openms.de>



<https://github.com/sneumann/xcms>  
<https://xcmsonline.scripps.edu>

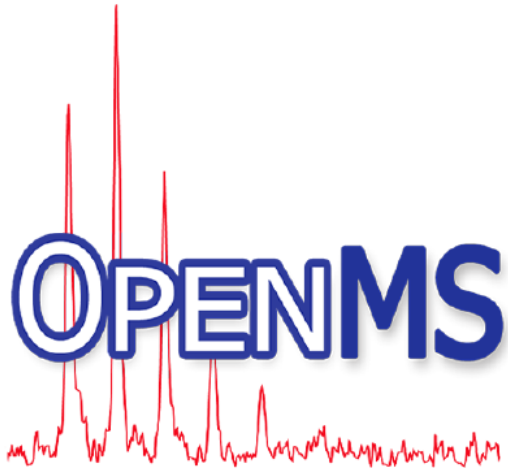
# Advantages of Open Source Solutions

- Communities of engaged users

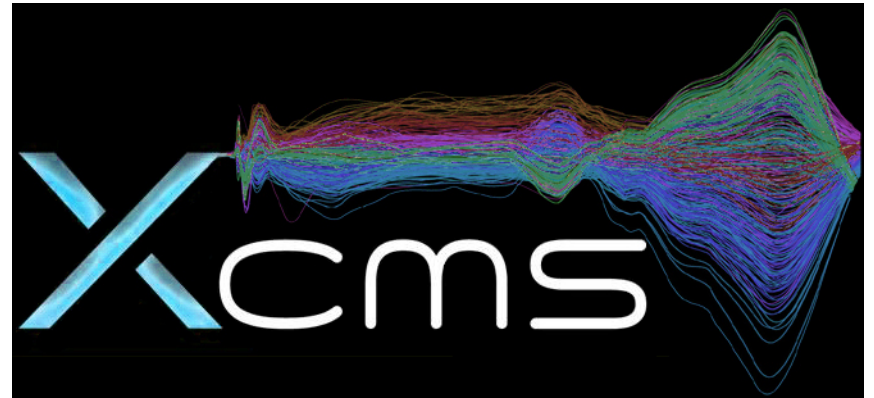


- Inspectable code

- Need-driven, collaborative development



- Transparent issue tracking and resolution



# Tool Software Stands on the Shoulders of Giants



- Your data
  - e.g., mzML
- Software tool
  - e.g., retcor
- Tool package
  - e.g., XCMS
- Software libraries
  - e.g., R programming language
- OS libraries
  - e.g., libquadmath0
- OS distribution
  - e.g., Debian
- OS kernel
  - e.g., Linux



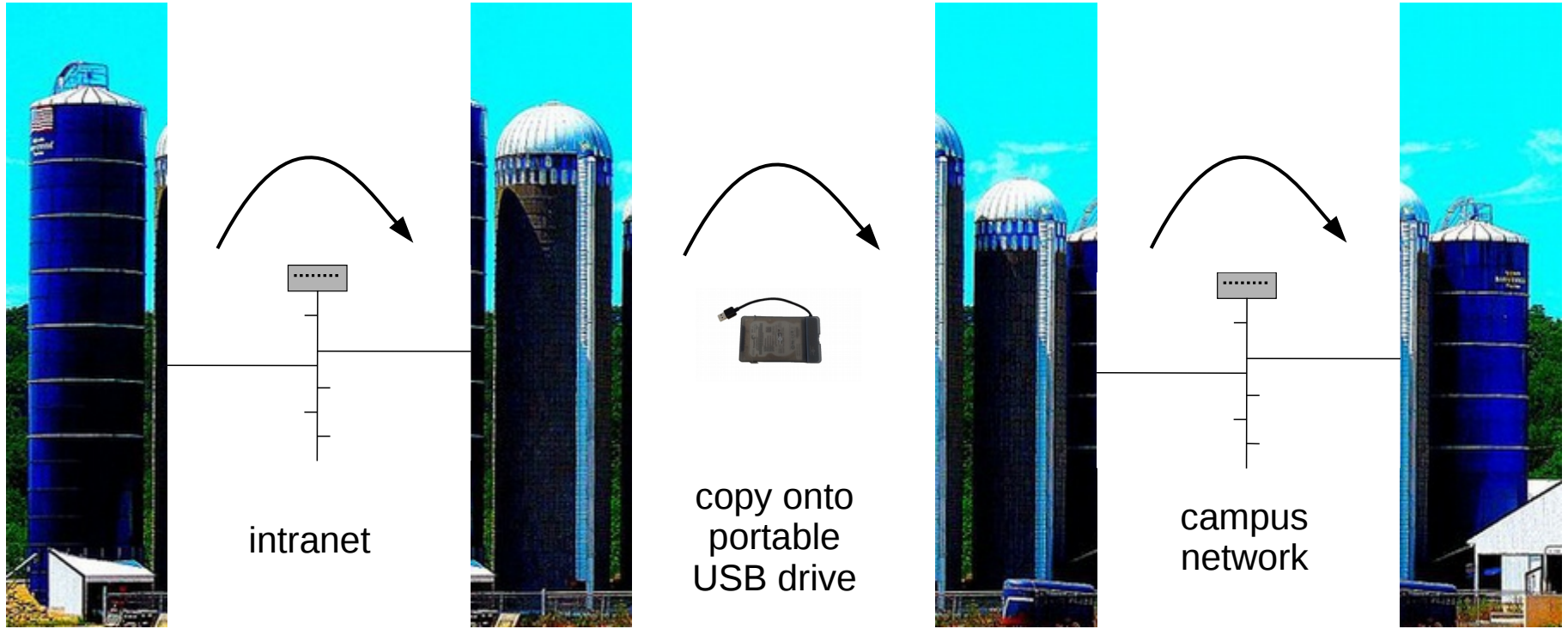
# Different Tools, Different Giants' Shoulders



- Separate machines?
- Separate environments?

# Data Proliferation

with minimal audit trail and challenging reproducibility



Instrument PC

Archive PC

Analysis PC

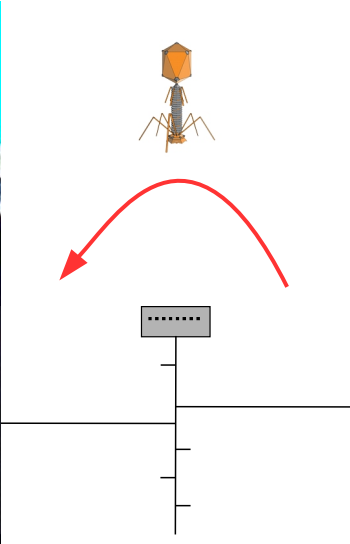
Laptop

# Virus Proliferation

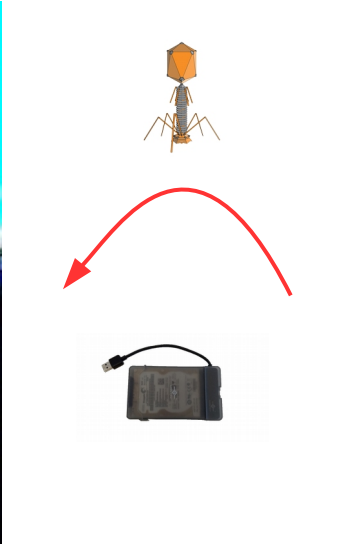
(and worms, malware, etc.)



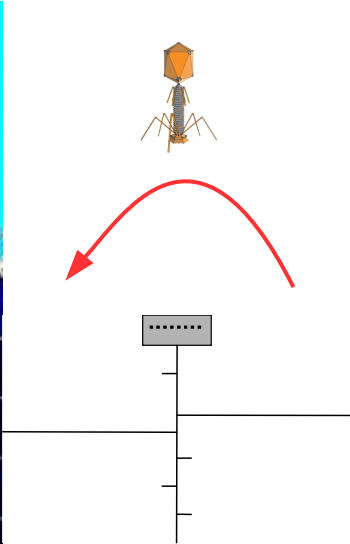
Instrument PC



Archive PC



Analysis PC



Laptop

# Our Wish List for a Solution

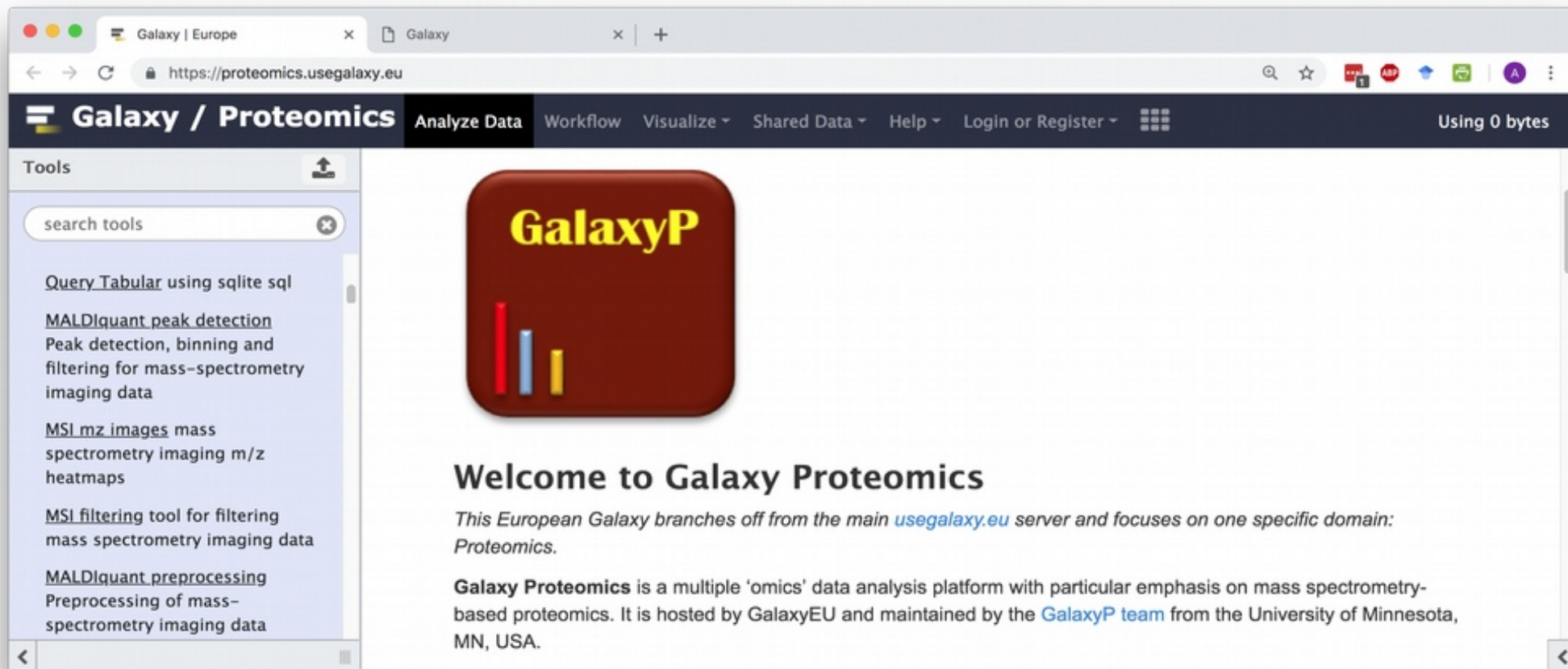
- Easy collaboration
- Easy data import
- Independent operation
- Customizability
- Data integrity/security
- Durable data/histories
- Durable workflows/tools

# Galaxy Tools Execute in their Own Environments



(Tools may run on same or different nodes.)

# Galaxy



The screenshot shows a web browser window with two tabs: "Galaxy | Europe" and "Galaxy". The address bar shows the URL "https://proteomics.usegalaxy.eu". The page header includes the "Galaxy / Proteomics" logo and navigation links for "Analyze Data", "Workflow", "Visualize", "Shared Data", "Help", "Login or Register", and a user profile icon. The top right corner indicates "Using 0 bytes".

The main content area features a large red rounded square icon with the text "GalaxyP" in yellow and a bar chart with three bars (red, blue, yellow). Below the icon, the heading "Welcome to Galaxy Proteomics" is displayed. The text below the heading reads: "This European Galaxy branches off from the main [usegalaxy.eu](https://usegalaxy.eu) server and focuses on one specific domain: Proteomics." and "Galaxy Proteomics is a multiple 'omics' data analysis platform with particular emphasis on mass spectrometry-based proteomics. It is hosted by GalaxyEU and maintained by the [GalaxyP team](#) from the University of Minnesota, MN, USA."

The left sidebar is titled "Tools" and contains a search bar with the text "search tools". Below the search bar, several tool entries are listed:

- [Query Tabular using sqlite sql](#)
- [MALDIquant peak detection](#)  
Peak detection, binning and filtering for mass-spectrometry imaging data
- [MSI mz images mass spectrometry imaging m/z heatmaps](#)
- [MSI filtering tool for filtering mass spectrometry imaging data](#)
- [MALDIquant preprocessing](#)  
Preprocessing of mass-spectrometry imaging data



Can we make this easy to use in our lab?



People will be eager to use a new system only when they perceive that it will make their jobs easier to do.



# Can we make this easy to use in our lab?



- We must process many large files concurrently for each analysis.

# Can we make this easy to use in our lab?



- We must process many large files concurrently for each analysis.
- Researchers are accustomed to such usage idioms as "drag and drop" file management.

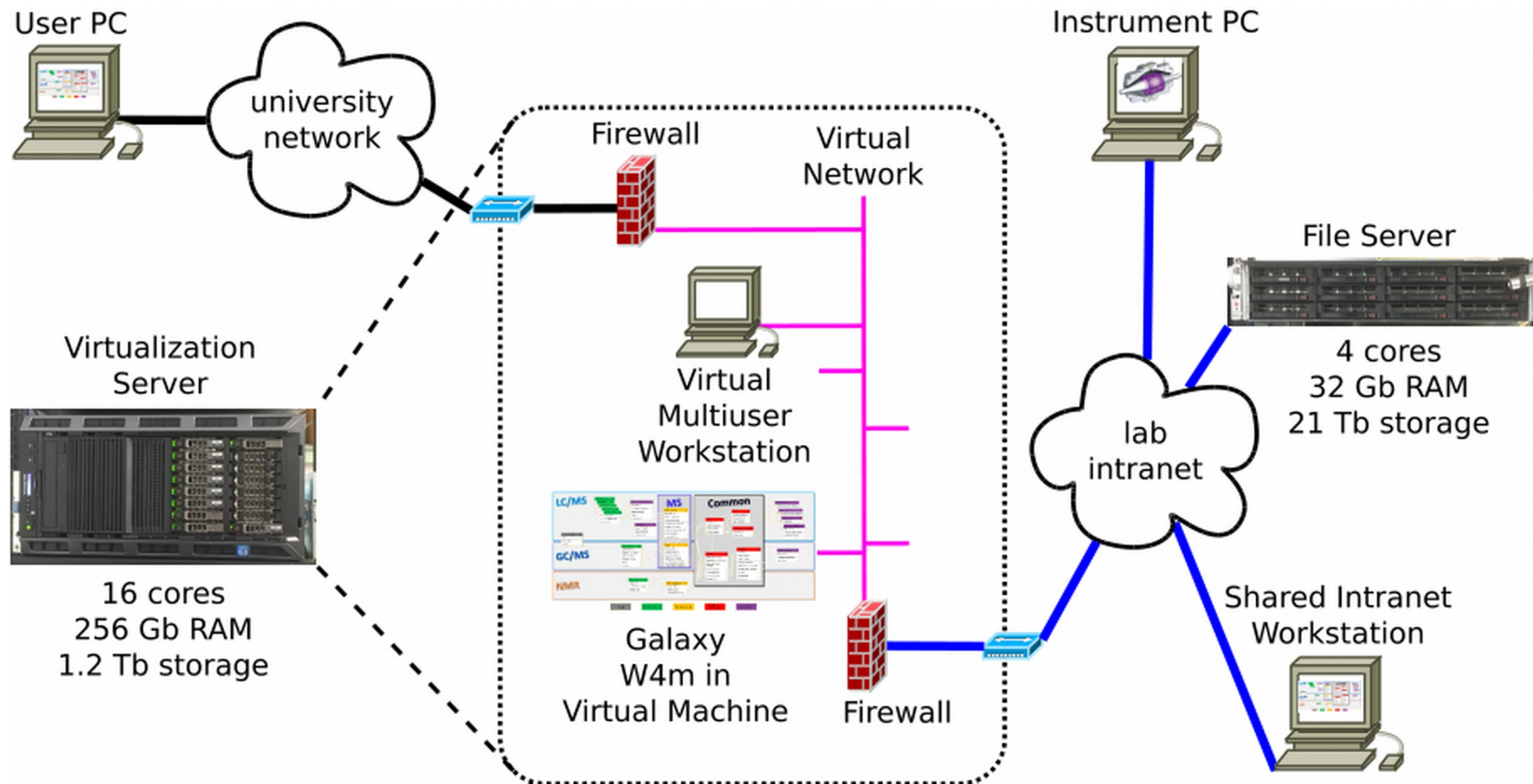
# Can we make this easy to use in our lab?



- We must process many large files concurrently for each analysis.
- Researchers are accustomed to such usage idioms as "drag and drop" file management.

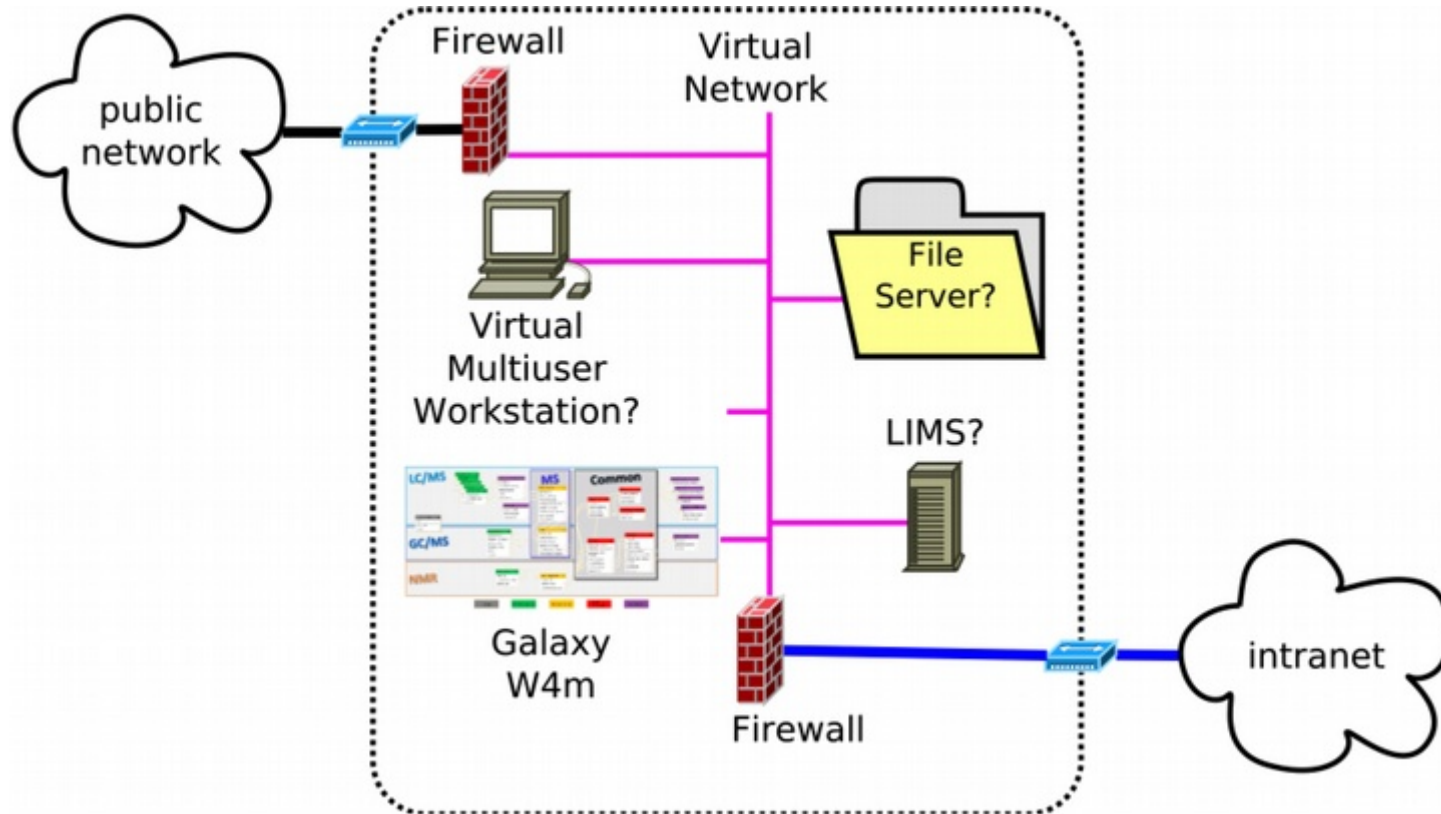
**Let's put Galaxy on our intranet!**

# Our Lab Intranet Solution (Simplified)





# Plan: Package as an Appliance



Perhaps extend BiBBox, <https://bibbox.readthedocs.io>

# Summary

- Galaxy helps us organize, process, and track plant metabolomics data effectively and collaboratively.
- With our own Galaxy we have:
  - created new tools at the direction of end-users;
  - customized and managed our instances independently.
- We plan to apply what we have learned to create an “appliance” for other labs (on our campus and elsewhere).

# Acknowledgements



UNIVERSITY OF MINNESOTA

Driven to Discover<sup>SM</sup>

Soon Li Teh  
Stephen Brockman  
Katherine Sammons  
Katrina Freund  
Yuan Xu  
Renata Pincelli-Sousa  
Kathryn Bushley  
Jerry Cohen Lab  
Gary Gardiner Lab  
Tim Griffin Lab  
Adrian Hegeman Lab



James Johnson



Workflow4metabolomics

Gildas Le Courguillé  
Etienne Thévenot



Steffen Neumann  
Johannes Rainer



Helena Rasche  
Björn Grüning



DBI-1458524  
IOS-1238812



National Institutes  
of Health

1U24CA199347



# Image Attributions

- <http://metacyc.ai.sri.com/META/NEW-IMAGE?type=PATHWAY&object=PWY-6244&detail-level=4>
- Katherine Sammons
- <https://commons.wikimedia.org/wiki/File:Orbitrappe.png>
- [https://commons.wikimedia.org/wiki/File:Ambox\\_octagon\\_stop.svg](https://commons.wikimedia.org/wiki/File:Ambox_octagon_stop.svg)
- [https://commons.wikimedia.org/wiki/File:Fassbender\\_Dairy\\_Farm\\_-\\_panoramio.jpg](https://commons.wikimedia.org/wiki/File:Fassbender_Dairy_Farm_-_panoramio.jpg)
- <https://commons.wikimedia.org/wiki/File:PhageExterior.svg>
- [workflow4metabolomics.org](http://workflow4metabolomics.org)
- <https://commons.wikimedia.org/wiki/File:Man-scratching-head.gif>