



Galaxy EU

Helena Rasche

About

Community

Subdomains

Partners

Users

Training

Materials

TIaaS

Testing

Scalability

Infrastructure

Community

Migration

Cloud

Openness

Intersection

State of the European Galaxy Project

Helena Rasche

2018-11-20



Community — Users — Training — Infrastructure

Galaxy EU

Helena Rasche

About

Community

Subdomains

Partners

Users

Training

Materials

TIaaS

Testing

Scalability

Infrastructure

Community

Migration

Cloud

Openness

Intersection



About Me

Galaxy EU

Helena Rasche

About

Community

Subdomains

Partners

Users

Training

Materials

TIaaS

Testing

Scalability

Infrastructure

Community

Migration

Cloud

Openness

Intersection

- Joined team 16 months ago
- UseGalaxy.eu System Administrator
- de.NBI Cloud Freiburg Administrator
(3.5k cores, 15TB ram)
- Developed Galaxy Circos/JBrowse tools, RStudio GIE



Community

Galaxy EU

Helena Rasche

About

Community

Subdomains
Partners

Users

Training
Materials
TlaaS
Testing
Scalability

Infrastructure

Community
Migration
Cloud
Openness

Intersection

Community is the most important part of what
we do

Galaxy / Metagenomics

Analyze Data Workflow Visualize Shared Data Help Login or Register

You are using the new UseGalaxy.eu backend server, let us know if you encounter any issues!

Tools

- search tools
- bioinformatics
- Quality Control
- Assembly
- Metagenomic analyses
- Oligo
- Mothur
- Chimera_vsearch
- Chimera_pintail
- Coccorense
- SPLIT_ABUND
- Get_atus
- Chimera_slayer
- Filter_sharable

Welcome to Galaxy Metagenomics (ASaIM) – a webserver to process, analyse and visualize Metagenomic and Microbiota data in general.

1. Get started
2. Tools
3. Tutorials
4. Workflows
5. References

Get started

Are you new to Galaxy, or returning after a long time, and looking for help to get started? Take a [guided tour](#) through Galaxy's user interface.

Want to learn about metagenomics analyses? Check our [tutorials](#) or take one of our guided tour:

- Introduction of amplicon data analyses using mothur tool suite
- Introduction to shotgun metagenomics data analyses
- 16S Microbial Analysis with Mothur MiSeq SOP

Check also the standard but customizable [workflows](#) available there.

Galaxy / HiCExplorer

Analyze Data Workflow Visualize Shared Data Help Login or Register

You are using the new UseGalaxy.eu backend server, let us know if you encounter any issues!

Tools

- search tools
- deeeTools
- HiCExplorer
- hicEndTADs
- hicPlotTADs
- hicMergeMatrixBins
- hicPlotMatrix
- hicBuildMatrix
- hicPlotViewpoint
- hicTransform
- hicCorrelate

Tools to process and visualize chromosome conformation

HiCExplorer

Fidel Ramirez, Joachim Wolf, Björn Grönig, Vivien Shereik, Gauder Richard, Gina Penzner, Rolf Backofen, Ralf Gieck, Thomas Meine, Stephan Niedrig, Devan Ryan

<https://github.com/deeeTools/HiCExplorer>

Max Planck Institute of Immunobiology and Epigenetics

Galaxy deeNBi PTI D-SP

Sekundärmetabolit und -verarbeitung

LUDWIG WIRTSCHAFTS UNIVERSITÄT MÜNCHEN

Get started with Galaxy HiCExplorer

Are you new to Galaxy, or returning after a long time, and looking for help to get started? Take a [guided tour](#) through Galaxy's user interface.

Take a [guided tour](#) for an introduction to Galaxy HiCExplorer and Hi-C data analysis. This tour is guides you through the Hi-C tutorial on the [Galaxy Training Network](#) where you can analyse Hi-C data of Drosophila melanogaster. Follow the tutorial to understand the analysis steps better or as a help which parameters are useful.

Galaxy / Proteomics

Analyze Data Workflow Visualize Shared Data Help Login or Register

You are using the new UseGalaxy.eu backend server, let us know if you encounter any issues!

Tools

- search tools
- PROTEOMICS, METABOLOMICS, CHEMISTRY
- Proteomics
- MALDIquant peak detection
- Peak detection, binning and filtering for mass-spectrometry imaging data
- MSI mzImage
- mass spectrometry imaging mz heatmaps
- MSI filtering
- filtering tool for mass spectrometry imaging data
- MALDIquant preprocessing
- Preprocessing of mass-spectrometry imaging data
- MSI nist spectra
- mass spectrometry imaging mass spectra plots
- MSI segmentation
- mass spectrometry imaging spatial clustering
- MSI Qualitycontrol
- mass spectrometry imaging QC
- MSI preprocessing
- mass spectrometry imaging

Welcome to Galaxy Proteomics

This European Galaxy branches off from the main [usegalaxy.eu](#) server and focuses on one specific domain: Proteomics.

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.

Feel free to check out another server maintained by the GalaxyP team: [Proteogenomics Gateway](#).

Get started

Are you new to Galaxy, or returning after a long time, and looking for help to get started? Take a [guided tour](#) through Galaxy's user interface.

Data Driven Research for Everyone

Galaxy is an open, web-based platform for accessible, reproducible, and transparent computational biological research.

Galaxy / Ecology

Analyze Data Workflow Visualize Shared Data Help Login or Register

You are using the new UseGalaxy.eu backend server, let us know if you encounter any issues!

Tools

- search tools
- FILE AND META TOOLS
- Get Data
- Convert Formats
- Collection Operations
- GENERAL TEXT TOOLS
- Text Manipulation
- Filter and Sort
- Join, Subtract and Group
- GENOMICS, NGS
- Extract Features
- BED Tools
- Fetch Alignments
- Operate on Genomic Intervals
- FASTA/FASTQ manipulation
- Multiple Alignments
- FASTA/FASTQ manipulation
- Picard
- Quality Control
- Assembly
- Mapping
- Variant Calling
- Genome editing

Welcome to Galaxy for Ecology – a web platform to get, process, analyze and visualize ecological data

Getting started

Are you new to Galaxy, or returning after a long time, and looking for help to get started? Take a [guided tour](#) through Galaxy's user interface.

Want to learn about ecology analyses? Check our [tutorials](#) or take one of our guided tour:

- Introduction to species phenology study and abundance index as trends computation.

Tools

Almost 140 tools are proposed in this custom Galaxy instance. They were chosen for their use in exploitation of ecology data:

- General tools
 - Text Manipulation: JQ process JSON, Replace Text, cast, melt, Subtract, Complement, Cluster, Replace text in a specific column, Replace parts of text, text reformatting, Text transformation, Unfold columns, Replace column, Add input name as column, Create text files, Sort a row, reverse a file, Compress an expression,

Data Driven Research for Everyone

Galaxy is an open, web-based platform for accessible, reproducible, and transparent computational biological research.

- Accessible:** Users can easily run tools without writing code or using the CLI; via a user-friendly web interface.
- Reproducible:** Galaxy captures all the metadata from an analysis, making it completely reproducible.
- Transparent:** Users share and publish analyses via interactive web-based documents that can enhance analyses with user-supplied annotation.
- Scalable:** Galaxy can run anything, from a laptop, to large clusters, to the cloud.


[Use Galaxy.eu](#)

News

Nov 7, 2018
Upcoming Server Migration

Nov 5, 2018
New Paper on "Uncontrolled Diabetes Mellitus Has No Major Influence on the Platelet Transcriptome"

Nov 2, 2018
New Paper on "The Transcription Factor ETV1 Induces Atrial Remodeling and Arrhythmia"

Oct 31, 2018
Training Infrastructure Feedback from Maria Doyle

Oct 30, 2018
New article 'Selbst ist der Forcher: Bio-Data-Analysen'

Events

Feb 25, 2019 - Mar 1, 2019
Galaxy workshop on HTS data analysis

Jan 28, 2019 - Feb 1, 2019
2019 Galaxy Admin Training

Dec 6, 2018 - Dec 7, 2018
Galaxy RNA-seq data analysis workshop in Freiburg

Nov 19, 2018 - Nov 20, 2018
European Galaxy Days

Nov 15, 2018
GTN CoFest on Training Material

News

Nov 1, 2018
Enthusiastic Response to the First ELIXIR-NL Galaxy Training Course in the Netherlands

May 3, 2018
Nanopolish tools for signal-level analysis of Oxford Nanopore sequencing data is now available in Galaxy

Events

Nov 12, 2018 - Nov 16, 2018
Elixir BioHackathon 2018

Nov 8, 2018
NGS Introduction using Galaxy

Oct 17, 2018 - Oct 19, 2018
Bioinformatics for Translational Medicine using Galaxy: it's cool, it's teach it!

Aug 27, 2018 - Aug 29, 2018
Next Generation Sequencing data analysis, LUMC

Aug 17, 2018
GTN CoFest on Metagenomics Training Material



How to use Galaxy?

Tools

In the left panel of this page you will find the list of available tools, you can see this page at any point by clicking 'Analyze Data' in the top menu.

History

The history of your current data analysis is shown in the right panel. A good practice when running multiple analysis in parallel is to create several histories and give each one unique name. You can click on 'View all histories' button (upper right corner in histories panel) to get an overview and switch between the current histories.

Input Data

Every analysis starts with getting the input data into your current history (right panel). To do this you can upload your own input files or use shared datasets. To upload files to your computer or install Galaxy to download files from the web you have to use of the upload tool: Get data (tool panel on the left) → upload file. Please DO NOT UPLOAD FILES (>GB), but contact the administrator to create a central data repository! To use available shared data you have to click on Shared data (top menu) → Data Libraries, then browse the available libraries and select the file/s you want to use. By clicking 'History' button and choosing the desired History name you will import these files and it'll then available to use as input for future analysis.

Execution

To run a job select the tool from tool panel on the left, then the corresponding interface will be loaded and you will be able to select corresponding input data and (re)define parameters. Please read carefully the labels and help text next to the input fields. After clicking Execute you will be able to see entries for each in your history. The color of a

describes the current status of the underlying job.

Color State

Gray The job is being evaluated to run (new dataset) or is queued. Allow this to complete.

Yellow The job is executing.

Green Successful processing.

Red The job has failed.

Blue The job is paused. This indicates either a problem with an input (a previous step in the workflow may have failed) or that you have exceeded disk quota set by the administrator of the Galaxy instance you are working on.

How to use Galaxy

- Learn how to use Galaxy
- History Introduction** - A detailed introduction to the Galaxy History
- Galaxy UI** - A gentle introduction to the Galaxy User Interface
- Workflow - Introduction** - An introduction on how to display multiple datasets and visualizations next to each other.
- Pasteur Galaxy Initiation slides are available here.

For pasteur users:

- Instructions on how to Import big data here.
- Galaxy Initiation slides are available here.
- If you need a new tool or a new genome index for an existing tool (bwa, blast, star...), send an email to [Pasteur_galaxy_support](#).

News

Mar 14, 2018
Updated PhageTerm

Jan 29, 2018
Updated Salmonella CRISPR Typing

Nov 30, 2017
Updated multqc

Nov 24, 2017

How to cite Galaxy

Citing Galaxy

If you use Galaxy@Pasteur please cite
doi:10.7490/1000/research.114334.1

Cluster Load



Past Year on UseGalaxy.eu

Galaxy EU

Helena Rasche

About

Community

Subdomains

Partners

Users

Training

Materials

TIaaS

Testing

Scalability

Infrastructure

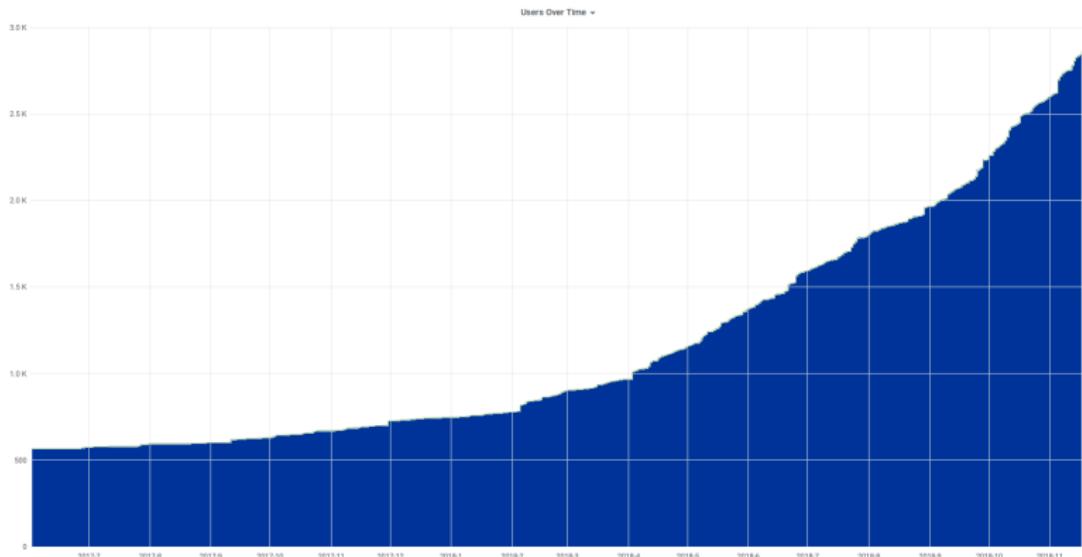
Community

Migration

Cloud

Openness

Intersection





Users

Galaxy EU

Helena Rasche

About

Community
Subdomains
Partners

Users

Training
Materials
TlaaS
Testing
Scalability

Infrastructure
Community
Migration
Cloud
Openness

Intersection

Galaxy Users

2863

Workflows

8274

Galaxy Histories

59650

Workflow Invocations

45836

HDAs

8167027

Jobs

4552868

2017	60 active users/month	10k jobs/month
2018	580 active users/month	100k jobs/month



Training

Galaxy EU

Helena Rasche

About

Community

Subdomains

Partners

Users

Training

Materials

TIaaS

Testing

Scalability

Infrastructure

Community

Migration

Cloud

Openness

Intersection

We love training!

- Training Materials
- Training Infrastructure
- Training Testing



Community + Training

Galaxy EU

Helena Rasche

About

Community

Subdomains

Partners

Users

Training

Materials

TIaaS

Testing

Scalability

Infrastructure

Community

Migration

Cloud

Openness

Intersection

- Amazing “User Conference”
- Successful BioHackathon
- Improvements to training materials / infrastructure
- 20 pull requests
 - Translation Support
 - Search
 - Training handbook
- 6 contributors
- GTN Meeting
- GTN CoFest for Genome Annotation



Training Infrastructure as a Service | TlaaS

Galaxy EU

Helena Rasche

About

Community

Subdomains

Partners

Users

Training

Materials

TlaaS

Testing

Scalability

Infrastructure

Community

Migration

Cloud

Openness

Intersection

What is it?

- Free Training Infrastructure for any Galaxy training.
- Private Queue
- No Administration
- No Maintenance
- Have someone to blame when the server burns down

galaxyproject.eu/tiaas

Enthusiastic Response to the First ELIXIR-NL Galaxy Training Course in the Netherlands

«

»

Nov 1, 2018

Erasmus MC

training

tiaas

An ELIXIR-NL training course entitled [Bioinformatics for Translational Medicine using Galaxy: see it, do it, teach it!](#) was held on October 17 - 19, 2018 in Amsterdam. Participants were enthusiastic about the course contents, and particularly interested in the link of using Galaxy workflows for training and teaching. For this workshop we used the [European Galaxy server](#) via their wonderful [Training Infrastructure as a Service \(TaaS\)](#) setup.

The workshop consisted of two parts, which were separately targeting two groups of people: 1) Biomedical & translational medicine researchers/students who are keen on bioinformatics and 2) educators in the "bio" domain who would like to start using Galaxy in their courses as an introduction to Bioinformatics.

In this workshop, a series of bioinformatics topics were taught in the Galaxy environment and the application of Galaxy into teaching has triggered interest and good response from the attendees. The follow-up feedback from the attendees is very positive.

The full list of topics covered in this training course can be found at [the course webpage](#) and all the training materials are available on the [Galaxy training materials website](#) (also suitable for self-study!)

The course was organized by members of the [DTL Galaxy Interest Group](#), more specifically by Chao (Cico) Zhang (VU), Sanne Abeln (VU), Anton Feenstra (VU), Saskia Hiltemann (ErasmusMC), Mateusz Kuzak (DTL) and Celia van Gelder (DTL).

A big *thank you!* to the GalaxyEU team for providing the TaaS service: it saved us a lot of time in



TIaaS Usage

Galaxy EU

Helena Rasche

About

Community

Subdomains

Partners

Users

Training

Materials

TIaaS

Testing

Scalability

Infrastructure

Community

Migration

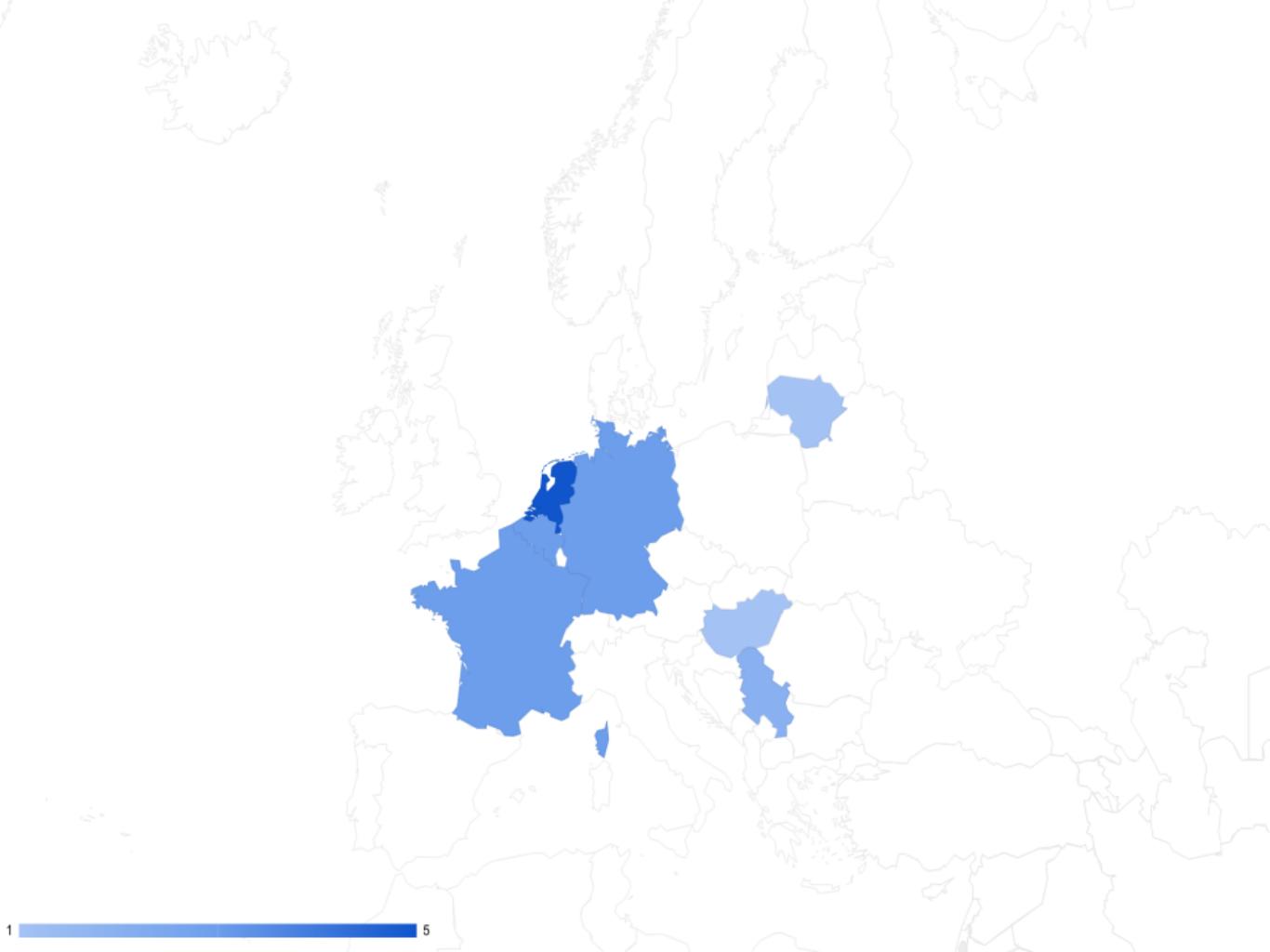
Cloud

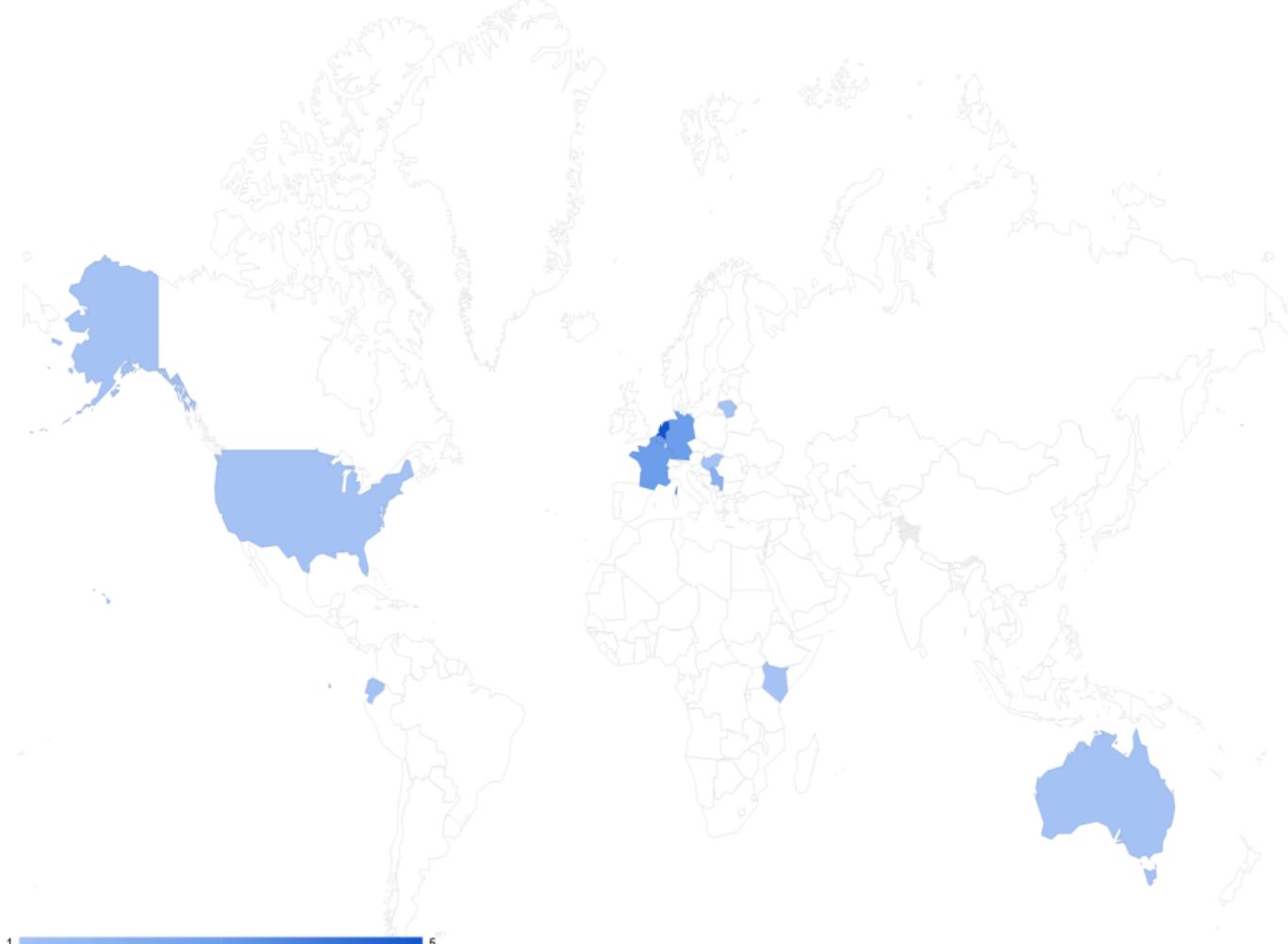
Openness

Intersection

In past 6 months:

- 14 Completed Trainings
- 5 Upcoming Trainings
- >500 Trainees





Training + Infra: Workflow Testing

Galaxy EU

Helena Rasche

About

Community

Subdomains

Partners

Users

Training

Materials

TIaaS

Testing

Scalability

Infrastructure

Community

Migration

Cloud

Openness

Intersection

github.com/usegalaxy-eu/workflow-testing

Workflow	Status
example1/wf3-shed-tools.ga	build passing
example2/wf4-shed-tools.ga	build passing
training/epigenetics/hicexplorer/hicexplorer.ga	build passing
blockclust/blockclust_workflow.ga	build passing
training/variant-analysis/microbial-variants/microbial_variant_calling.ga	build passing
training/sequence-analysis/quality-control/quality_control.ga	build passing
training/sequence-analysis/mapping/mapping.ga	build passing
training/chip-seq/formation_of_super_structures_on_xi/formation_of_super_structures_on_xi.ga	build passing
training/assembly/general-introduction/assembly-general-introduction.ga	build passing
training/assembly/unicycler-assembly/unicycler.ga	build passing
training/metagenomics/general-tutorial/amplicon.ga	build passing
training/transcriptomics/ref-based/ref_based.ga	build passing
GraphClust2/GC-lite.ga	build passing
sklearn/ard/ard.ga	build passing
sklearn/adaboost/adaboost.ga	build passing
training/variant-analysis/mapping-by-sequencing/mapping_by_sequencing.ga	build passing
training/proteomics/protein-id-sg-ps/protein-id-sg-ps.ga	build running
training/proteomics/protein_quant_sil/protein_quant_sil.ga	build passing



Scalability

Galaxy EU

Helena Rasche

About

Community

Subdomains

Partners

Users

Training

Materials

TIaaS

Testing

Scalability

Infrastructure

Community

Migration

Cloud

Openness

Intersection

- Australia broadcast “hybrid” training across entire country
- Some hypothetical plans for doing this in Europe
- Dependent on finding grants



Busy year at Galaxy EU

Galaxy EU

Helena Rasche

About

Community

Subdomains

Partners

Users

Training

Materials

TIaaS

Testing

Scalability

Infrastructure

Community

Migration

Cloud

Openness

Intersection

Overhaul of Infrastructure

- New services: CI / Monitoring / HAProxy / CVMFS
- Everything in Ansible
- Everything Open
- Cloud Infrastructure
- Automation: Ansible + Terraform



Community

Galaxy EU

Helena Rasche

About

Community

Subdomains

Partners

Users

Training

Materials

TIaaS

Testing

Scalability

Infrastructure

Community

Migration

Cloud

Openness

Intersection

Remote Compute

- VGCN github.com/usegalaxy-eu/vgcn
- Czech group + others interested
- Compute power dedicated to UseGalaxy.eu

Why?

- No maintenance
- No administration
- Run our VM, give us your compute power :)

Migration

Galaxy EU

Helena Rasche

About

Community

Subdomains

Partners

Users

Training

Materials

TaaS

Testing

Scalability

Infrastructure

Community

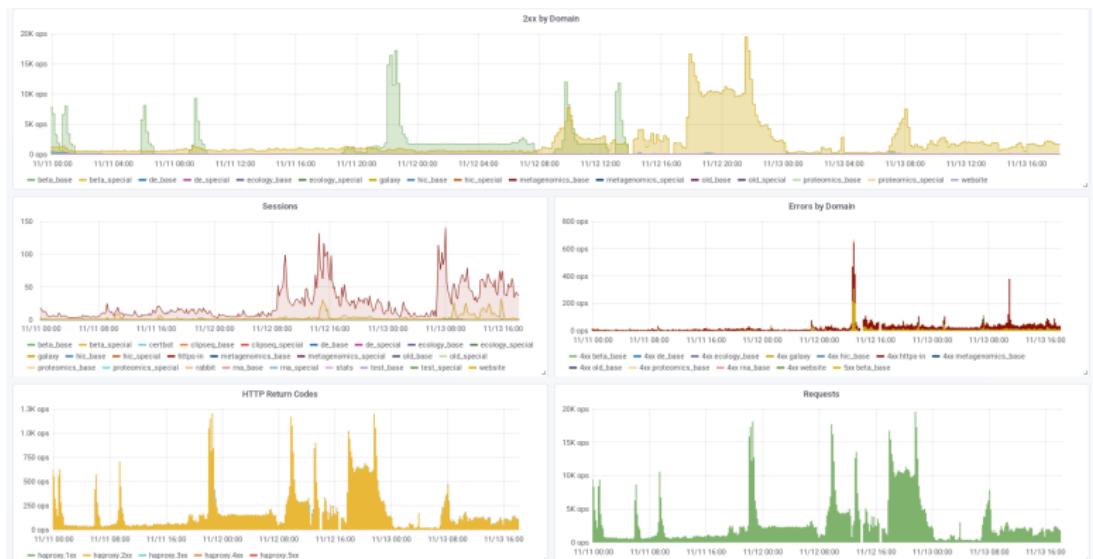
Migration

Cloud

Openness

Intersection

- Old head node was dying — migrated to a new one
- Live! (Mostly) No one noticed





Cloud

Galaxy EU

Helena Rasche

About

Community

Subdomains

Partners

Users

Training

Materials

TlaaS

Testing

Scalability

Infrastructure

Community

Migration

Cloud

Openness

Intersection

- Cloud Migration, moved entire infrastructure from old OpenStack to a new one.
- Terraform!
- ≈ 100 VMs in Cloud
- ≈ 1,3k Cores
- ≈ 8 TB Ram



Open Development + Contributions

Galaxy EU

Helena Rasche

About

Community

Subdomains

Partners

Users

Training

Materials

TlaaS

Testing

Scalability

Infrastructure

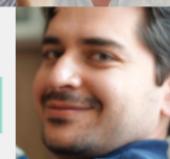
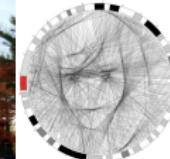
Community

Migration

Cloud

Openness

Intersection



Community + Users + Training + Infrastructure

Galaxy EU

Helena Rasche

About

Community

Subdomains
Partners

Users

Training

Materials
TlaaS
Testing
Scalability

Infrastructure

Community
Migration
Cloud
Openness

Intersection

