

## FAIR and scalable education

### The Galaxy training network (GTN) and a Training Infrastructure as a Service (TlaaS)

Anika Erxleben-Eggenhofer<sup>1</sup>[\[https://orcid.org/0000-0002-7427-6478\]](https://orcid.org/0000-0002-7427-6478), Teresa Müller<sup>2</sup>[\[https://orcid.org/0000-0003-1252-9684\]](https://orcid.org/0000-0003-1252-9684), and Bérénice Batut<sup>1</sup>[\[https://orcid.org/0000-0001-9852-1987\]](https://orcid.org/0000-0001-9852-1987)

<sup>1</sup> Freiburg Galaxy Team, Bioinformatics Group, Department of Computer Science, University of Freiburg, Germany

<sup>2</sup> Bioinformatics Group, Department of Computer Science, Albert-Ludwigs-University Freiburg, Germany

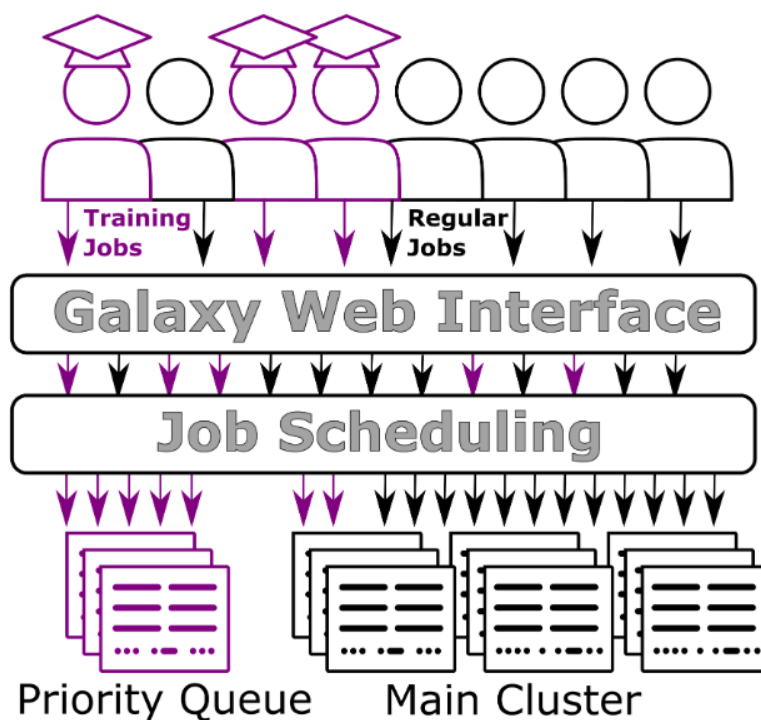
**Abstract.** The Galaxy Project [1][2] is a widely-used open-source platform for data-driven research that offers an extensive suite of tools and services for analyzing and visualizing large-scale data. Besides Galaxy as data analysis framework, the Galaxy Training Network (GTN; [3][4]) provides a huge collection of training material for researchers, developers and trainers. Among the various features offered by the Galaxy platform, the GTN and the Training Infrastructure as a Service (TlaaS; [5]) stands out as an innovative solution for delivering scalable and flexible training to researchers and educators. TlaaS is built on top of the Galaxy platforms which enable users to request dedicated compute resources for training purposes. This infrastructure, in combination with GTN, provides an excellent solution for delivering high-quality training to researchers and educators around the world, regardless of their geographical location or hardware limitations. One of the key benefits of the GTN is its ability to enable users to create and share training materials for a variety of data analysis workflows, including e.g. genomics, proteomics, and metabolomics but also for imaging and climate data. These training materials can be delivered in various formats, including interactive tutorials, videos, and webinars, and can be accessed from any location with an internet connection. Moreover, the platform offers a wide range of community-contributed training materials that cover a vast array of topics in biomedical research and beyond, making it an invaluable resource for researchers and educators. To use TlaaS, instructors simply need to register their training event in a simple form to get access to the dedicated compute resources.

TlaaS was created by the implementation of two components: a web service, and a default set of Galaxy job scheduling rules, which function together to present a private queue for users in specific Galaxy user groups. The web service enables registering requests for resources for the training event specifying time, topic with tools to be used and number of trainees. Additionally then a training group is created in Galaxy by adding members to those groups in a GDPR compliant way, as needed. TlaaS coordinators or system administrators review these requests, using information about the class size, the tools used in the training materials, as well as the resource allocations of those tools on the infrastructure, to estimate the required compute resources.

Training participants access a specific training URL at the start of the training event, after which they are automatically registered in the TlaaS system without further user interaction and without instructors needing to manually manage group membership. The job scheduler, once aware of the training group, will place any job run by someone in that training onto the dedicated training nodes (Fig. 1).

In this talk we will highlight how Galaxy and TiaaS in combination with well curated training material can scale to thousands of simultaneous trainees and enable Massive Open Online Course (MOOC) events that we have been executing the last 3 years during the pandemic.

**Keywords:** Galaxy, Galaxy Training Network, training, education, TiaaS



**Figure 1.** Scheme of the TiaaS queuing system. Jobs are processed by the same Galaxy server, but when those jobs come from users in the training group, they receive special handling.

## Competing interests

The authors declare that they have no competing interests.

## References

1. The Galaxy Community - Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update; Nucleic Acids Research, Volume 50, Issue W1, 5 July 2022, Pages W345–W351, <https://doi.org/10.1093/nar/gkac247>
2. <https://galaxyproject.org> (accessed Apr. 24, 2023).
3. Hiltmann & Rasche et al. - Galaxy Training: A powerful framework for teaching! PLoS Comput Biol 19(1): e1010752. <https://doi.org/10.1371/journal.pcbi.1010752>
4. <https://training.galaxyproject.org> (accessed Apr. 24, 2023).
5. <https://usegalaxy-eu.github.io/tiaas.html> (accessed Apr. 24, 2023).