

# Towards automating and publishing workflow analyses in Galaxy

Andrea Bagnacani, Markus Wolfien, Martin Scharm, Olaf Wolkenhauer

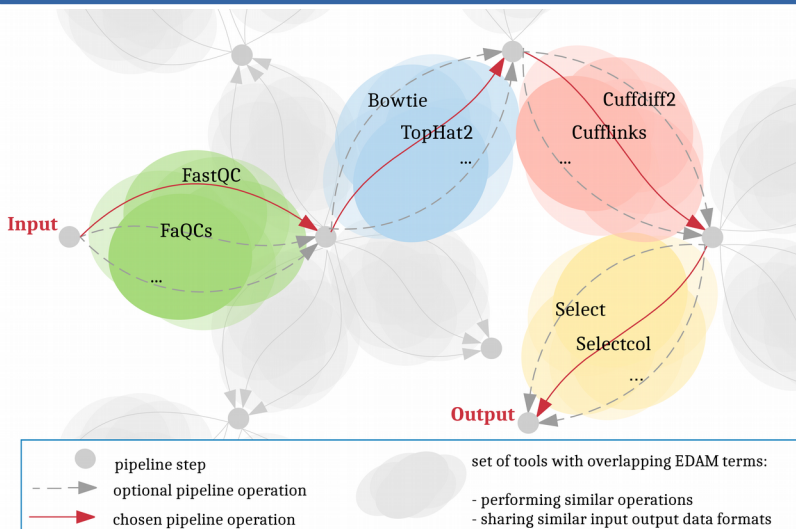
## Research gaps and objectives for the present work

The Galaxy community is promoting RNA-Seq protocols and best practices through the reuse of existing tools, and the consolidation of a Training Network to provide guidance to researchers through example datasets, tutorials, and interactive tours. However, the more tools and techniques are showcased, the more complex the options for tool chaining and parametrization become. And as more material is gathered to illustrate approaches on the analysis of Life Science data, the higher the risk to overwhelm users. To assist researchers in carrying out their investigations, we propose to integrate the Galaxy framework with an interactive recommendation system that leverages on community consolidated best practices as well as EDAM-annotated tools. Such a system would promote the adoption of well established pipelines, while at the same time allow room for testing further experimental tools, easing reuse, while consolidating protocols and reproducibility.

## References

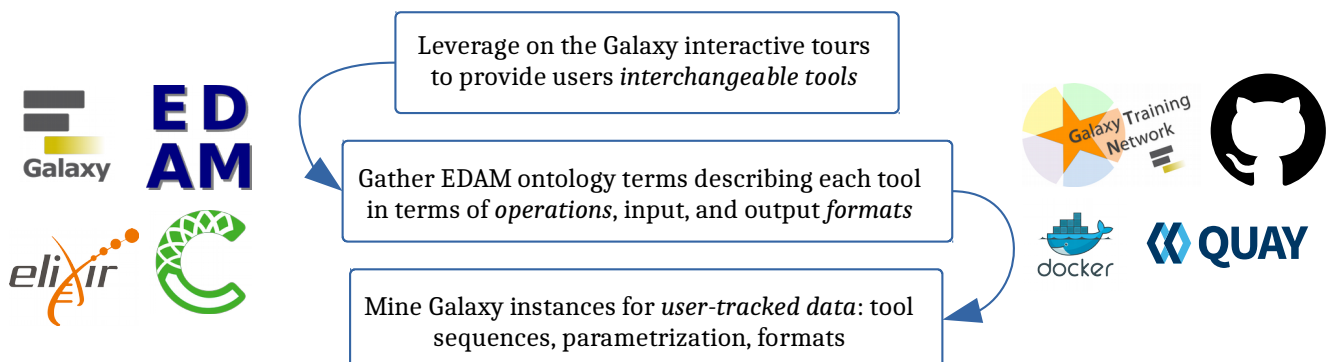
- Wolfien M**, Rimmbach R, Schmitz U, Jung JJ, Krebs S, Steinhoff G, David R, **Wolkenhauer O**. TRAPLINE: A standardized and automated pipeline For RNA sequencing data analysis, evaluation and annotation. BMC Bioinformatics, 2016. [10.1186/s12859-015-0873-9](https://doi.org/10.1186/s12859-015-0873-9)
- Lott SC, **Wolfien M**, Riege K, **Bagnacani A**, **Wolkenhauer O**, Hoffmann S, Hess WR. Customized workflow development and data modularization concepts for RNA-Sequencing and metatranscriptome experiments. Journal of Biotechnology, 2017. [10.1016/j.jbiotec.2017.06.1203](https://doi.org/10.1016/j.jbiotec.2017.06.1203)
- Grüning BA, Fallmann J, Yusuf D, Will S, Erxleben A, Eggenhofer F, Houwaart T, Batut B, Videm P, **Bagnacani A**, **Wolfien M**, Lott SC, Hoogstrate Y, Hess WR, **Wolkenhauer O**, Hoffmann S, Akalin A, Ohler U, Stadler PF, Backofen R. The RNA workbench: best practices for RNA And high-throughput sequencing bioinformatics in Galaxy. Nucleic Acids Research, 2017. [10.1093/nar/gkx409](https://doi.org/10.1093/nar/gkx409)

## Bridging the gap



**The de.STAIR interactive recommendation system.** The recommendation system traces a path from the beginning of a Life Science data analysis to its end. Here, researchers can decide which path to walk towards the completion of the desired analysis by providing an input dataset and an analysis goal. Tools are recommended step by step, based on both tool's pertinence within the scope of the requested goal, and existing best practices within the current analysis step. A best practice approach is shared among the scientific community through the Galaxy platform as reusable workflows, while a tool pertinence is inferred from the provided EDAM ontology terms associated to its operations and input/output data formats.

## Towards the automation



Andrea Bagnacani  
andrea.bagnacani@uni-rostock.de

Systems Biology and Bioinformatics  
University of Rostock  
[www.sbi.uni-rostock.de](http://www.sbi.uni-rostock.de)