



EOSC-Life: Building a digital space for the life sciences

D2.2 — A common environment that can run cross-RIs workflows

WP2 – Tools Collaboratory: Deployment of life-science data integration and analysis workflows in EOSC

Lead Beneficiary: CRMMP and EMBL

WP leader: Antonio Rosato and Jean-Karim Hériché

Contributing partner(s): EMBL, CIRMMP, ALU-FR, INRA, SU, Fraunhofer, KNAW, USMI, CNR, UNIMAN, VLIZ, INFRAFRONTIER, CRS4, VIB, ECRIN, FZJ, Lygature, ERINHA

Authors of this deliverable: **Antonio Rosato, Jean-Karim Hériché**

Contractual delivery date: **31 Aug 2022**

Actual delivery date: **17 October 2022**

H2020-INFRAEOSC-2018-2

Grant agreement no. 824087

Horizon 2020

Type of action: RIA

Table of Contents

Executive Summary	3
Project Objectives.....	3
Detailed Report on the Deliverable	3
1. Background	3
2. Description of Work	4
3. Next Steps	8
References	10
Abbreviations.....	11
Delivery and Schedule	11
Adjustments.....	11
Appendices	12
Appendix A: List of community-specific Galaxy instances	12
Appendix B: Non-exhaustive list of publications with WP2 contribution to tools and workflows deployment	12



Executive Summary

Research in the communities behind the BMS RIs relies on a large and diverse set of software tools and computational workflows that are often not easily reusable outside their domain or community of origin. The core purpose of WP2 is to make these tools and workflows Findable, Accessible, Interoperable and Reusable (FAIR) in the European Open Science Cloud (EOSC), so that users across the BMS RIs, as well as the broad scientific community, can take full advantage of the EOSC through a common set of resources and environment. This deliverable reports on the implementation of the WP2 roadmap to create such an environment supporting the deployment of tools and workflows from different domains of the life sciences in the EOSC.

Project Objectives

This deliverable has contributed to the following project objectives:

- a. Make BMS RI tools ready for deployment in the EOSC following FAIR principles;
- b. Drive implementation of workflows that cross disciplines and RI boundaries;
- c. Promote and develop the use of the EOSC by fostering tool-focused collaborations and sharing between communities via cloud implementation and sharing of workflows that integrate the data and data analysis methods of different BMS RI.

Detailed Report on the Deliverable

1. Background

WP2 is concerned with running computational tools and workflows from the different life science domains in the EOSC. A workflow describes a set of computational tasks and their relationships. A tool is a piece of software used by researchers to carry out a computational task. In WP2, workflows are specified in terms of tools and the flow of data between them.

BMS RIs produce a large and diverse array of tools and workflows for managing, integrating and analysing a variety of data types. In addition, access to computational resources and skills is also highly variable. As a consequence, cloud deployment of tools and workflows is very heterogeneous across BMS RIs communities.

In order to overcome challenges posed by this diversity and heterogeneity, WP2 identified relevant technologies that were already in use in some BMS RIs communities (see milestone M2.1) to form the basis of the WP2 roadmap for the development of a common computational environment for the life sciences that will foster interoperability and a tighter integration of data across BMS RI communities in the EOSC.



2. Description of Work

In order to promote a cross-RI EOSC environment, WP2 has worked in two areas:

1. A tools and workflows **deployment** framework
2. A tools and workflows **FAIRification** framework

2.1. A framework for deploying tools and workflows in EOSC

In order to address data access, tools interoperability and cloud portability, WP2 focuses on the parts of the software stack that are required to implement workflows, namely tool packaging, containerisation, workflow management systems and other relevant platforms, such as computational notebooks. These technologies, which form the basis of the deployment framework of the WP2 roadmap (Figure 1), are listed below.

- The Conda package manager¹ is used for tool packaging. It standardises software installation and is already used by many researchers: the Bioconda channel, which collects a majority of tools relevant to the life sciences, now contains more than 8000 packages.
- Singularity² and Docker³ are used for “containerisation” to provide isolated execution environments and ensure their portability. Over 10,500 life sciences tools are now available in containerized form in the BioContainers registry⁴.
- WP2 promotes the adoption of workflow engines that make use of the above technologies, such as Nextflow⁵ and Snakemake⁶, for running workflows on the command line, and the Galaxy platform⁷, as a web-based user-friendly interface for building and running data analysis workflows. In particular, Galaxy makes tools and workflows available to non-expert users across many different communities through community instances (see list in Appendix A), and also through national instances (e.g., usegalaxy.fr⁸, usegalaxy.be⁹, usegalaxy.es¹⁰).

¹ <https://docs.conda.io/en/latest/>

² <https://github.com/sylabs/singularity>

³ <https://www.docker.com/products/container-runtime/>

⁴ <https://biocontainers.pro/>

⁵ <https://www.nextflow.io/>

⁶ <https://snakemake.readthedocs.io/en/stable/>

⁷ <https://galaxyproject.org/>

⁸ <https://usegalaxy.fr/>

⁹ <https://usegalaxy.be/>

¹⁰ <https://usegalaxy.es/>



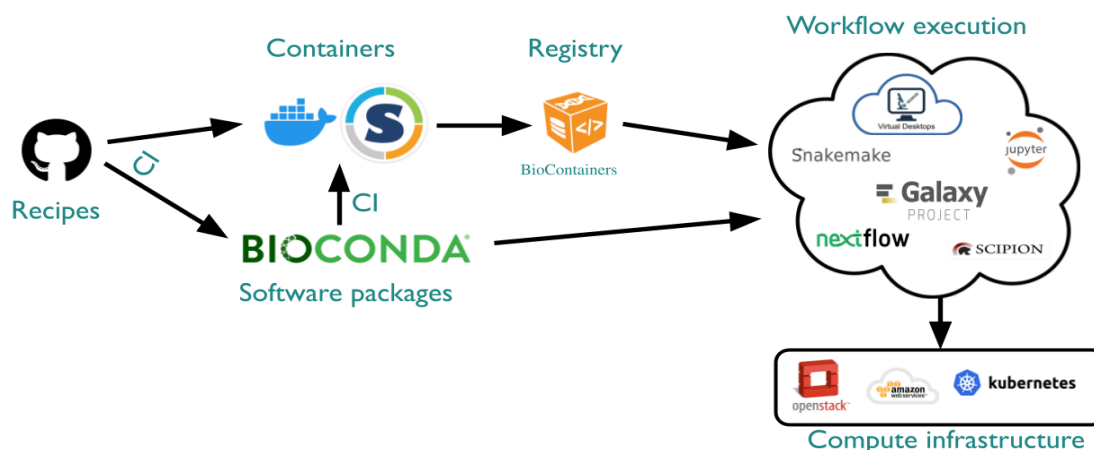


Figure 1: EOSC-Life WP2 tools and workflows deployment framework. Bioconda packages and Docker and Singularity containers provide portable software that can be deployed through registries across a variety of cloud execution environments, from workflow management systems to computational notebooks and virtual desktops, running on a variety of compute infrastructures.

During the COVID-19 pandemic, particular efforts and resources were devoted to supporting and deploying SARS-CoV-2 related workflows^{11,12}. See also ELIXIR's newsletter "Hacking the pandemic"¹³.

WP2 promotes the adoption of its roadmap and of the associated technologies for the deployment of tools and workflows by:

- contributing to training (e.g., workshop on tool packaging¹⁴, workshop on workflow platforms¹⁵),
- contributing to community events (e.g., Galaxy Community Conferences, European Galaxy days, meetings of scientific societies),
- contributing to EOSC-Life open call¹⁶ community projects,
- through publications (see Appendix B),
- raising awareness within each RI using their communication networks (e.g., Euro-Biolmaging virtual pubs, ELIXIR All Hands meetings).

It is noteworthy that the WP2 roadmap is finding adoption beyond the life sciences, for example through EOSC-Life collaboration with EOSC-Nordic¹⁷ on the Climate Science Workbench used to build climate models.

¹¹ <https://covid19.workflowhub.eu/>

¹² <https://galaxyproject.org/projects/covid19/>

¹³ <https://elixir-europe.org/news/hacking-pandemic>

¹⁴ <https://docs.google.com/document/d/1uowLhMxMpaJc7kleDZe8KjCxwej4gKLpbUXruaCFqel/edit>

¹⁵ <https://docs.google.com/document/d/1Od0JzTduih7DIalaoYS-MnfFU1UVmy6DprqhBlqAV8/edit>

¹⁶ <https://www.eosc-life.eu/calls/>

¹⁷ <https://nordicesmhub.github.io/EOSC/>



2.2. A framework for the FAIRification of tools and workflows

Existing mechanisms for making computational tools FAIR were selected early in the project (see milestone M2.1¹⁸) and were put on the roadmap (e.g., bio.tools, BioContainers). Due to the aforementioned heterogeneity in practices and segregation of communities, tools from some communities were under-represented or absent from the framework. WP2 participated in the increase in the number of tools available with some effort devoted to increasing domain coverage, for example by contributing tools from the bioimaging community and by raising awareness of the existence of the registries.

However, there was no equivalent for making computational workflows FAIR [Goble 2020¹⁹] and a major effort in WP2 was then devoted to building the necessary resources. This led to the creation of the WorkflowHub²⁰, a registry for describing, sharing and publishing computational workflows, which is now in beta. Over 250 workflows are already registered in the WorkflowHub, covering multiple BMS RIs (e.g. ELIXIR, MIRRI, Euro-BioImaging, INSTRUMENT, INFRAFRONTIER, BBMRI) across over 100 research groups and projects²¹.

The WorkflowHub enhances workflows FAIRness in many ways. It provides identifiers (including DOIs²²) and rich metadata. It also provides added value services, such as organising workflows in collections and supporting the life cycle of a workflow with versioning, snapshotting, sub-workflows and test data. Moreover, it provides links to testing and execution services.

Recently, the Research Data Alliance (RDA) working group “FAIR for Research Software²³” has also published an extension of FAIR Principles to cover Research software [Chue Hong 2022²⁴], which WP2 has contributed to. This work on how FAIR can be realised for software will continue within the new FAIR-IMPACT²⁵ project where UNIMAN, EMBL-EBI, CSC and CNR are partners, with publishing workflows in WorkflowHub being one of the use cases.

To facilitate discovery and re-use of workflows, the WorkflowHub leverages a number of standards to which WP2 also contributed, such as Workflow RO-Crate²⁶, a standard for packaging executable workflows together with auxiliary files and enriched metadata, and Workflow Testing RO-Crate²⁷, an extension of Workflow RO-Crate that adds a formalism to specify metadata related to workflow testing.

To support this work, WP2 also contributed tools to enable working with the standards, such as ro-crate-py²⁸, a Python library and CLI to create and consume RO-Crates, and repo2rocrate²⁹, a tool to generate Workflow (Testing) RO-Crates out of workflow repositories following community standards. repo2rocrate generalises part of the code from an earlier tool developed by WP2 to

¹⁸ https://docs.google.com/document/d/1lck8E6pC5s9U5dWNDuz0gZJKG7wEf10Dof_e_0CcE3g/edit#heading=h.n68mymbc5lev

¹⁹ <https://direct.mit.edu/dint/article/2/1-2/108/10003/FAIR-Computational-Workflows>

²⁰ <https://workflowhub.eu/>

²¹ <https://workflowhub.eu/projects>

²² <https://about.workflowhub.eu/docs/citable/>

²³ <https://www.rd-alliance.org/groups/fair-research-software-fair4rs-wg>

²⁴ <https://zenodo.org/record/6623556#.Y0P8yezMITV>

²⁵ <https://fair-impact.eu/>

²⁶ <https://about.workflowhub.eu/Workflow-RO-Crate/>

²⁷ https://crs4.github.io/life_monitor/workflow_testing_ro_crate

²⁸ <https://github.com/ResearchObject/ro-crate-py>

²⁹ <https://github.com/crs4/repo2rocrate>



automatically generate Workflow Testing RO-Crate metadata for the highly curated Galaxy Intergalactic Workflow Commission (IWC) workflow collection³⁰.

Information going stale is a major risk for registries, since it can lead to workflows that stop working due to software collapse [Hinsen 2019³¹]. Extending the useful life of registered workflows requires them to be regularly tested in various contexts. Periodic testing reveals problems when they arise and helps identify issues so they can be rectified. To facilitate workflow maintenance, and in particular to help with the automation and monitoring of workflow tests, WP2 is developing the LifeMonitor³² workflow testing service.

In addition to tools that facilitate the automatic execution of tests on continuous integration services (e.g., GitHub Actions³³), LifeMonitor integrates with WorkflowHub, includes a web application reporting on the status³⁴ of all tested workflows, and can notify people who manifested their interest when workflows show problems.

Interoperability between WorkflowHub and LifeMonitor is ensured by the adoption of Workflow (Testing) RO-Crate as a common data exchange format. LifeMonitor also includes a GitHub app, currently in pre-release testing, that leverages repo2rocrate to interactively guide workflow maintainers through the steps required to generate a “standardised usage” workflow repository following community best practices.

Such best practices include actions like specifying a licence, including test setups and adding detailed (or at least minimum required) metadata (author, organisation, version, etc.), which considerably enhance the workflow’s reusability and findability. The app then automatically converts the repository to a Workflow Testing RO-Crate for registration to WorkflowHub and LifeMonitor.

An example of a workflow following community standards is the FAIR CRCC image conversion workflow³⁵, developed with contributions from BBMRI and EuBI and interaction with WP1 and WP2, that aims to make the colorectal cancer cohort³⁶ dataset more accessible. The dataset contains thousands of digital pathology images. The workflow supports converting and maintaining these as images in open, rather than vendor-specific, formats. The workflow also implements an envelope encryption strategy based on Crypt4GH [Senf 2021³⁷] to support secure sharing of this sensitive data collection.

The companion workflow FAIR CRCC send data³⁸ implements a secure data transfer procedure for the encrypted CRCC data. It adds the authorised receiver’s decryption key to the enveloped data and sends it to a selected destination storage location. Here, only the receiver – the only one who has the decryption key – can access it. These workflows apply the recommendations of the WP2 roadmap, so their components are also fully containerised and, as a result, they show excellent portability: in fact, they have already been executed on different computing infrastructures

³⁰ <https://github.com/galaxyproject/iwc>

³¹ <https://ieeexplore.ieee.org/document/8701540>

³² <https://www.lifemonitor.eu/>

³³ <https://github.com/features/actions>

³⁴ <https://app.lifemonitor.eu/dashboard>

³⁵ <https://github.com/crs4/fair-crcc-img-convert>

³⁶ <https://www.bbMRI-eric.eu/scientific-collaboration/colorectal-cancer-cohort/>

³⁷ <https://doi.org/10.1093/bioinformatics/btab087>

³⁸ <https://github.com/crs4/fair-crcc-send-data/>



without any changes. The workflow repositories, which adopt the standard Snakemake layout³⁹, were converted to a Workflow Testing RO-Crate with repo2rocrate and registered to WorkflowHub and LifeMonitor.

A full report on the workflow FAIRification framework (figure 2) will be available in deliverable D2.3: *Implementation of a comprehensive workflow registry solution for publishing and sharing workflows across instances of the Tools Collaboratory.*

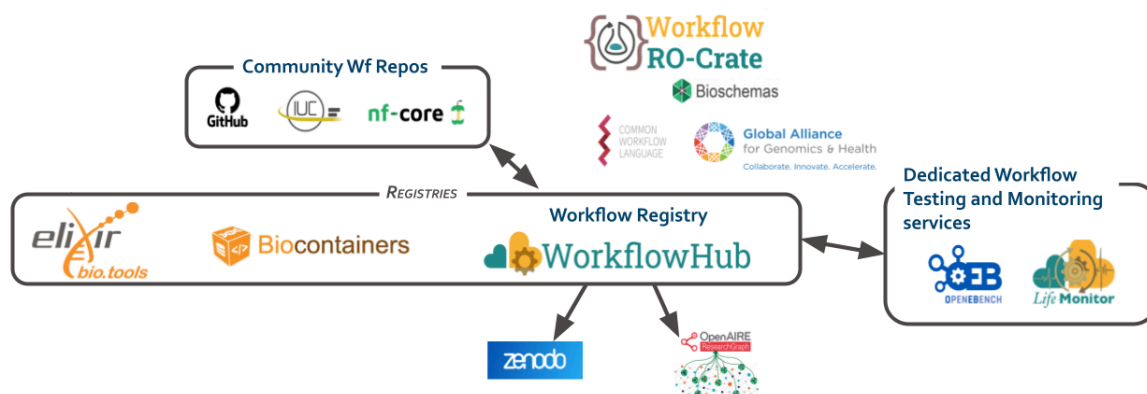


Figure 2: EOSC-Life WP2 tools and workflows FAIRification framework. FAIRification of tools and workflows from repositories such as GitHub, the Intergalactic Utilities Commission's toolshed and nf-core is supported by tools registries bio.tools and biocontainers and the workflow registry workflowhub through the development and use of standards (CWL, workflow-RO crate, Bioschemas, GA4GH) and monitoring systems (OpenEbench, Life-Monitor).

3. Next Steps

3.1. Access to data

As the success of the Galaxy platform demonstrates, researchers are keen to use cloud computing services. However, connecting computational workflows running in a cloud environment to input data is emerging as a major hurdle for EOSC users. While structured public data repositories usually provide well-defined data access mechanisms, the diversity of these mechanisms can be confusing, and changes to them can lead to incompatibilities with existing tools and workflows making the data inaccessible to them.

In addition, researchers are generally interested in running workflows on data that have been freshly generated for their research project and are not publicly available yet. However, users often have no clear means of connecting locally generated data to cloud services. A workshop on

³⁹ <https://snakemake.readthedocs.io/en/stable/snakefiles/deployment.html#distribution-and-reproducibility>



data access, movement, and management⁴⁰ organised by ELIXIR with participation of EOSC-Life WP2 members already touched upon some of these issues.

To bring this a step further, a working group involving EOSC-Life WP1 and WP2 is looking into how to address these issues with the goal of coming up with a set of best practices.

3.2. Sustainability of services and technologies

Further European-wide developments around the Galaxy platform will continue under the new EOSC project EuroScienceGateway⁴¹, which started in September 2022. EuroScienceGateway will leverage the Galaxy platform, the Pulsar Network⁴² (a wide distributed job execution system to scale computing power over heterogeneous resources) and FAIR workflow services pioneered in the EOSC-Life⁴³ cluster (WorkflowHub⁴⁴, Workflow RO-Crate⁴⁵ and metadata standards like schema.org). According to its proposal, EuroScienceGateway aims to lift Pulsar from TRL-7 to TRL-9, and WorkflowHub from TRL-7 to TRL-9.

In order to standardise the software installation processes, ensure reproducibility of workflows and allow automation of software deployments in cloud environments, EOSC-Life WP2 relies on a now well-established technology stack.

Crucially, this stack depends on a number of freely and publicly accessible services and technologies that are currently provided free-of-charge by commercial enterprises. These services and technologies have been key enablers of the development of cloud computing services that make research more efficient, reproducible and collaborative. Some of these free services and technologies have now taken such a central role that they are now considered essential.

However, most of the time, researchers do not have formal contracts and agreements for the use of these services and technologies, but rely on the tacit assumption that conditions under which they will be provided will remain compatible with the research community's interests, e.g. free of charge, freely accessible, with suitable performance characteristics, etc.

For example, at the end of 2020, Docker Inc. which provides Docker Hub, the largest repository of container images used by default as source of containers by many systems, imposed a limit to the number of container images that can be downloaded for free, which forced some organisations to implement mitigating measures such as mirroring locally parts of the repository.

As this example illustrates, the unexpected introduction of usage restrictions to some of the free services and even the disappearance of others as the providers search for profitability highlight the fragility of the current model. In light of this issue, WP2 plans to evaluate in more detail the risks to the sustainability of the tools included in its roadmap.

⁴⁰ <https://docs.google.com/document/d/1hliYWbur3t6fAYsiitDMxGqE-1VF3m44mGzkdj7SYw/edit>

⁴¹ <https://esciencelab.org.uk/projects/eurosciencegateway/>

⁴² <https://pulsar-network.readthedocs.io/en/latest/>

⁴³ <https://www.eosc-life.eu/>

⁴⁴ <https://workflowhub.eu/>

⁴⁵ <https://about.workflowhub.eu/Workflow-RO-Crate/>



References

[Chue Hong 2022] Neil P. Chue Hong, Daniel S. Katz, Michelle Barker; Anna-Lena Lamprecht, Carlos Martinez, Fotis E. Psomopoulos, Jen Harrow, Leyla Jael Castro, Morane Gruenpeter, Paula Andrea Martinez, Tom Honeyman; Alexander Struck, Allen Lee, Axel Loewe, Ben van Werkhoven, Catherine Jones, Daniel Garijo, Esther Plomp, Françoise Genova, Hugh Shanahan, Joanna Leng, Maggie Hellström, Malin Sandström, Manodeep Sinha, Mateusz Kuzak, Patricia Herterich, Qian Zhang, Sharif Islam, Susanna-Assunta Sansone, Tom Pollard, Udayanto Dwi Atmojo; Alan Williams, Andreas Czerniak, Anna Niehues, Anne Claire Fouilloux, Bala Desinghu, Carole Goble, Céline Richard, Charles Gray, Chris Erdmann, Daniel Nüst, Daniele Tartarini, Elena Ranguelova, Hartwig Anzt, Ilian Todorov, James McNally, Javier Moldon, Jessica Burnett, Julián Garrido-Sánchez, Khalid Belhajjame, Laurents Sesink, Lorraine Hwang, Marcos Roberto Tovani-Palone, Mark D. Wilkinson, Mathieu Servillat, Matthias Liffers, Merc Fox, Nadica Miljković, Nick Lynch, Paula Martinez Lavanchy, Sandra Gesing, Sarah Stevens, Sergio Martinez Cuesta, Silvio Peroni, Stian Soiland-Reyes, Tom Bakker, Tovo Rabemanantsoa, Vanessa Sochat, Yo Yehudi, FAIR4RS WG (2022): **FAIR Principles for Research Software version 1.0 (FAIR4RS Principles v1.0)**.

Research Data Alliance

<https://doi.org/10.15497/RDA00068>

[Crusoe 2022] Michael R. Crusoe, Sanne Abeln, Alexandru Iosup, Peter Amstutz, John Chilton, Nebojša Tijanić, Hervé Ménager, Stian Soiland-Reyes, Bogdan Gavrilović, Carole Goble, The CWL Community (2022):

Methods Included: Standardizing Computational Reuse and Portability with the Common Workflow Language. *Communications of the ACM* 65(6)

<https://doi.org/10.1145/3486897>

[Goble 2020] Carole Goble, Sarah Cohen-Boulakia, Stian Soiland-Reyes, Daniel Garijo, Yolanda Gil, Michael R. Crusoe, Kristian Peters, Daniel Schober (2020):

FAIR Computational Workflows. *Data Intelligence* 2(1):108–121

https://doi.org/10.1162/dint_a_00033

[Senf 2021] Alexander Senf, Robert Davies, Frédéric Haziza, John Marshall, Juan Troncoso-Pastoriza, Oliver Hofmann, Thomas M. Keane (2021):

Crypt4GH: a file format standard enabling native access to encrypted data, *Bioinformatics*, Volume 37, Issue 17, Pages 2753–2754, <https://doi.org/10.1093/bioinformatics/btab087>

[Soiland-Reyes 2022] Stian Soiland-Reyes, Peter Sefton, Mercè Crosas, Leyla Jael Castro, Frederik Coppens, José M. Fernández, Daniel Garijo, Björn Grüning, Marco La Rosa, Simone Leo, Eoghan Ó Carragáin, Marc Portier, Ana Trisovic, RO-Crate Community, Paul Groth, Carole Goble (2022):

Packaging research artefacts with RO-Crate. *Data Science* 5(2)

<https://doi.org/10.3233/DS-210053>

[Wilkinson 2016] Mark Wilkinson, et al. (2016):

The FAIR Guiding Principles for scientific data management and stewardship.

Scientific Data 3

<https://doi.org/10.1038/sdata.2016.18>



[Hinsen 2019] K. Hinsen, "Dealing With Software Collapse," in *Computing in Science & Engineering*, vol. 21, no. 3, pp. 104-108, 1 May-June 2019, <https://doi.org/10.1109/MCSE.2019.2900945>.

Abbreviations

- **BMS RIs** - Biological and Medical Sciences Research Infrastructures
- **CWL** – Common Workflow Language [[Crusoe 2022](#)] <https://www.commonwl.org/>
- **FAIR** – Findable, Accessible, Interoperable and Reusable [[Wilkinson 2016](#)] <https://www.go-fair.org/fair-principles/>
- **IWC** – Galaxy’s *Intergalactic Workflow Commission* <https://github.com/galaxyproject/iwc>
- **RO-Crate** – Research Object Crate [[Soiland-Reyes 2022](#)] <https://w3id.org/ro/crate>
- **TRL** – technology readiness level

Delivery and Schedule

The delivery is delayed: Yes

The delivery is slightly delayed due to some staff absence over the summer break.

Adjustments

Adjustments made:

None



Appendices

Appendix A: List of community-specific Galaxy instances

- rna.usegalaxy.eu
- clipseq.usegalaxy.eu
- metagenomics.usegalaxy.eu
- hicexplorer.usegalaxy.eu
- cheminformatics.usegalaxy.eu
- proteomics.usegalaxy.eu
- imaging.usegalaxy.eu
- metabolomics.usegalaxy.eu
- ecology.usegalaxy.eu
- nanopore.usegalaxy.eu
- singlecellomics.usegalaxy.eu
- humancellatlas.usegalaxy.eu
- virology.usegalaxy.eu
- streetscience.usegalaxy.eu

Appendix B: Non-exhaustive list of publications with WP2 contribution to tools and workflows deployment

- Paul Brack, Peter Crowther, Stian Soiland-Reyes, Stuart Owen, Douglas Lowe, Alan R Williams, Quentin Groom, Mathias Dillen, Frederik Coppens, Björn Grüning, Ignacio Eguinoa, Philip Ewels, Carole Goble (2022):
Ten Simple Rules for making a software tool workflow-ready. *PLOS Computational Biology* **18**(3):e1009823 <https://doi.org/10.1371/journal.pcbi.1009823>
- Simon A. Bray, Xavier Lucas, Anup Kumar & Björn A. Grüning (2020):
The ChemicalToolbox: reproducible, user-friendly cheminformatics analysis on the Galaxy platform. (2020) *Journal of Cheminformatics* **12**, 40. <https://doi.org/10.1186/s13321-020-00442-7>
- Michael R. Crusoe, Sanne Abeln, Alexandru Iosup, Peter Amstutz, John Chilton, Nebojša Tijanić, Hervé Ménager, Stian Soiland-Reyes, Bogdan Gavrilović, Carole Goble, The CWL Community (2022):
Methods Included: Standardizing Computational Reuse and Portability with the Common Workflow Language. *Communications of the ACM* **65**(6) <https://doi.org/10.1145/3486897>
- Rafael Ferreira da Silva, Henri Casanova, Kyle Chard, Ilkay Altintas, Rosa M Badia, Bartosz Balis, Tainã Coleman, Frederik Coppens, Frank Di Natale, Bjoern Enders, Thomas Fahringer, Rosa Filgueira, Grigori Fursin, Daniel Garijo, Carole Goble, Dorran Howell, Shantenu Jha, Daniel S. Katz, Daniel Laney, Ulf Leser, Maciej Malawski, Kshitij Mehta, Loïc Pottier, Jonathan Ozik, J. Luc Peterson, Lavanya Ramakrishnan, Stian Soiland-Reyes, Douglas Thain, Matthew Wolf (2021):



- A Community Roadmap for Scientific Workflows Research and Development.** *2021 IEEE Workshop on Workflows in Support of Large-Scale Science (WORKS)*, pp 81–90.
<https://doi.org/10.1109/WORKS54523.2021.00016> arXiv:2110.02168 [cs.DC]
- Carole Goble, Sarah Cohen-Boulakia, Stian Soiland-Reyes, Daniel Garijo, Yolanda Gil, Michael R. Crusoe, Kristian Peters, Daniel Schober (2020):
FAIR Computational Workflows. *Data Intelligence* **2**(1):108–121
https://doi.org/10.1162/dint_a_00033
 - Alex Hardisty, Paul Brack, Carole Goble, Laurence Livermore, Ben Scott, Quentin Groom, Stuart Owen, Stian Soiland-Reyes (2022):
The Specimen Data Refinery: A canonical workflow framework and FAIR Digital Object approach to speeding up digital mobilisation of natural history collections. *Data Intelligence* **4**(2) https://doi.org/10.1162/dint_a_00134
 - Annika Jacobsen, Ricardo de Miranda Azevedo, Nick Juty, Dominique Batista, Simon Coles, Ronald Cornet, Mélanie Courtot, Mercè Crosas, Michel Dumontier, Chris T. Evelo, Carole Goble, Giancarlo Guizzardi, Karsten Kryger Hansen, Ali Hasnain, Kristina Hettne, Jaap Heringa, Rob W.W. Hooft, Melanie Imming, Keith G. Jeffery, Rajaram Kaliyaperumal, Martijn G. Kersloot, Christine R. Kirkpatrick, Tobias Kuhn, Ignasi Labastida, Barbara Magagna, Peter McQuilton, Natalie Meyers, Annalisa Montesanti, Mirjam van Reisen, Philippe Rocca-Serra, Robert Pergl, Susanna-Assunta Sansone, Luiz Olavo Bonino da Silva Santos, Juliane Schneider, George Strawn, Mark Thompson, Andra Waagmeester, Tobias Weigel, Mark D. Wilkinson, Egon Willighagen, Peter Wittenburg, Marco Roos, Barend Mons, Erik Schultes (2020):
FAIR Principles: Interpretations and Implementation Considerations. *Data Intelligence* **2**(1):10–29 https://doi.org/10.1162/dint_r_00024
 - Anna-Lena Lamprecht, Magnus Palmblad, Jon Ison, Veit Schwämmle, Mohammad Sadnan Al Manir, Ilkay Altintas, Christopher J. O. Baker, Ammar Ben Hadj Amor, Salvador Capella-Gutierrez, Paulos Charonyktakis, Michael R. Crusoe, Yolanda Gil, Carole Goble, Timothy J. Griffin, Paul Groth, Hans Ienasescu, Pratik Jagtap, Matúš Kalaš, Vedran Kasalica, Alireza Khanteymooi, Tobias Kuhn, Hailiang Mei, Hervé Ménager, Steffen Möller, Robin A. Richardson, Vincent Robert, Stian Soiland-Reyes, Robert Stevens, Szoke Szaniszló, Suzan Verberne, Aswin Verhoeven, Katherine Wolstencroft (2021):
Perspectives on automated composition of workflows in the life sciences. [version 1; peer review: 2 approved] *F1000Research* **10**:897
<https://doi.org/10.12688/f1000research.54159.1>
 - Vincenzo Laveglia, Andrea Giachetti, Linda Cerofolini, Kevin Haubrich, Marco Fragai, Alessio Ciulli, Antonio Rosato (2021):
Automated determination of nuclear magnetic resonance chemical shift perturbations in ligand screening experiments: The PICASSO web server. *Journal of Chemical Information and Modeling* **61**(12), 5726. <https://doi.org/10.1021/acs.jcim.1c00871>
 - Paul-Gilloteaux, P.; Tosi, S.; Hériché, J.-K. *et al.*
Bioimage analysis workflows: community resources to navigate through a complex ecosystem [version 1; peer review: 2 approved]. *F1000Research* 2021, **10**:320
<https://doi.org/10.12688/f1000research.52569.1>
 - Strelak, D.; Jiménez-Moreno, A.; José, L. *et al.* **Advances in Xmipp for cryo-electron microscopy: from Xmipp to Scipion.** (2021) *Molecules* **26**(20), 6224.
<https://doi.org/10.3390/molecules26206224>
 - Stian Soiland-Reyes, Peter Sefton, Mercè Crosas, Leyla Jael Castro, Frederik Coppens, José M. Fernández, Daniel Garijo, Björn Grüning, Marco La Rosa, Simone Leo, Eoghan Ó Carragáin,



- Marc Portier, Ana Trisovic, RO-Crate Community, Paul Groth, Carole Goble (2022): **Packaging research artefacts with RO-Crate**. *Data Science* 5(2) <https://doi.org/10.3233/DS-210053>
- Stian Soiland-Reyes, Genís Bayarri, Pau Andrio, Robin Long, Douglas Lowe, Ania Niewielska, Adam Hospital, Paul Groth (2022): **Making Canonical Workflow Building Blocks Interoperable across Workflow Languages**. *Data Intelligence*, 4(2), 342. https://doi.org/10.1162/dint_a_00135
 - Tangaro, M.; Mandreoli, P.; Chiara, M. *et al.* (2021): **Laniakea@ReCaS: exploring the potential of customisable Galaxy on-demand instances as a cloud-based service**. *BMC Bioinformatics* 22, 544. <https://doi.org/10.1186/s12859-021-04401-3>
 - Peter Wittenburg, Ivonne Anders, Christophe Blanchi, Merret Buurman, Carole Goble, Jonas Grieb, Alex Hardisty, Sharif Islam, Thomas Jejkal, Tibor Kálmán, Christine Kirkpatrick, Laurence Lannom, Thomas Lauer, Giridhar Manepalli, Karsten Peters-von Gehlen, Andreas Pfeil, Robert Quick, Mark van de Sanden, Ulrich Schwardmann, Stian Soiland-Reyes, Rainer Stotzka, Zachary Trautt, Dieter Van Uytvanck, Claus Weiland, Philipp Wieder (2022): **FAIR Digital Object Demonstrators 2021**. Report, FAIR Digital Objects Forum. *Zenodo* <https://doi.org/10.5281/zenodo.5872645>
 - Zullino, S.; Paglialonga, A.; Dastrù, W. *et al.* (2022): **XNAT-PIC: extending XNAT to preclinical imaging centers**. *Journal of Digital Imaging*. <https://doi.org/10.1007/s10278-022-00612-z>
 - The Galaxy Community (2022) **The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update**, *Nucleic Acids Research*, Volume 50, Issue W1, Pages W345–W351, <https://doi.org/10.1093/nar/gkac247>
 - Beatriz Serrano-Solano, Anne Fouilloux, Ignacio Eguinoa, Matúš Kalaš, Björn Grüning, Frederik Coppens (2022) **Galaxy: A Decade of Realising CWFR Concepts**. *Data Intelligence* 4 (2): 358–371. doi: https://doi.org/10.1162/dint_a_00136

