The WorkflowHub: A FAIR Registry for Workflows and Workflow Research Objects

Carole Goble[1], Finn Bacall[1], Stian Soiland-Reyes[1], Stuart Owen[1], Alan Williams[1], Ignacio Eguinoa[2], Bert Droesbeke[2], Hervé Ménager[3], Laura Rodríguez-Navas[4], José Mª Fernández[4], Salvador Capella[4], Michael R. Crusoe[5], Björn Grüning[6], Simone Leo[7], Luca Pireddu[7], Johan Gustafson[8], Phil Ewels[9], WorkflowHub Community[10] and Frederik Coppens[2].

[1] The University of Manchester, UK. Email: carole.goble@manchester.ac.uk

[2] VIB-UGent Center for Plant Systems Biology, Belgium

[3] Institut Pasteur, France

[4] Barcelona Supercomputing Center, Spain

[5] VU Amsterdam, The Netherlands

[6] University of Freiburg, Germany

[7] CRS4, Sardina, Italy

[8] Australian BioCommons, Australia

[9] SciLifeLab National Genomics Infrastructure, Sweden

[10] https://about.workflowhub.eu/acknowledgements/#workflowhub-club

Project Website: https://workflowhub.eu/, https://workflowhub.org

Source Code: https://github.com/seek4science/seek/tree/workflow/

License: (example: **BSD 3-Clause "New" or "Revised" License**)

Main Text of Abstract

The practice of performing computational processes using workflows has taken hold in the life sciences. Like data, workflows should be FAIR, citable, have managed metadata profiles and be openly available for review and analytics.

The WorkflowHub (https://workflowhub.eu) is a new FAIR workflow registry sponsored by the European RI Cluster EOSC-Life and the European Research Infrastructure ELIXIR. It is workflow management system agnostic: workflows may remain in their native repositories in their native forms. As workflows are multi-component objects, including example and test data, they are packaged, registered, downloaded and exchanged as workflow centric Research Objects using the RO-Crate specification, making the Hub an implementation of FAIR Digital Object principles. A

schema.org based Bioschemas profile describes the metadata about a workflow and use of the Common Workflow Language is encouraged, providing a canonical description of the workflow itself. Popular workflow management systems such as Galaxy, Nextflow, and Snakemake are working with the Hub to seamlessly and automatically support object packaging, registration and exchange. The WorkflowHub provides features such as community spaces, collections, versioning and snapshots, and contributor credit. It supports community registry standards and services such as GA4GH TRS and ELIXIR-AAI, and current work integrates with the LifeMonitor workflow testing service.

The WorkflowHub Club open community works together to continuously co-develop the Hub. Beta-released in Sept 2020, the Hub now holds nearly 100 workflows, including 36 curated COVID-19 workflows. It is a listed resource of the European COVID19 Data Portal (https://www.covid19dataportal.org/).

The Hub software is a branch of the open source FAIRDOM-SEEK platform.