Contribution ID: **89**   Type: **Workshop/Training**

Onedata hands-on workshop: Workflow Processing, Dataset Management and Archiving

Friday, 23 June 2023 09:00 (1h 30m)

Onedata [1] is a high-performance data management system with a distributed, global infrastructure that enables users to access storage resources worldwide. It supports various use cases ranging from personal data management to data-intensive scientific computations. Onedata has a fully distributed architecture that facilitates the creation of a hybrid-cloud infrastructure with private and commercial cloud resources. Users can collaborate, share, and publish data, as well as perform high-performance computations on distributed data using POSIX-compliant data access applications. The latest Onedata release introduces the integration of a powerful workflow execution engine, which is powered by OpenFaas [2]. This integration enables the creation of complex data processing pipelines that can leverage transparent access to organizationally distributed data. In addition, the new software version offers several new features and improvements that enhance its capabilities in managing distributed datasets throughout their lifecycle.

This hands-on workshop will focus on the latest Onedata release version, 21.02.1. Participants will explore its features through interactive exercises, with a special focus on data processing using automation workflows, distributed dataset management, and archive preservation. Other covered topics will include directory size statistics and the Space Marketplace. The training materials will correspond to the scenarios from the Onedata demonstration, presented during another session. The workshop will be conducted on the Onedata services at EGI DataHub, with the intention of easy reproducibility by EGI users.

Keywords: workflows, distributed dataset, data lifecycle, data access, distributed systems, file sharing.

[1] Onedata project website. https://onedata.org.
[2] OpenFaaS - Serverless Functions Made Simple. https://www.openfaas.com/.

Other key topic

Key Topic

Data Spaces

Primary authors: OPIOLA, Lukasz (CYFRONET); ORZECHOWSKI, Michal; KRYZA, Bartosz; DUTKA, Lukasz

Presenter: OPIOLA, Lukasz (CYFRONET)