Contribution ID: **29**   Type: **not specified**

Data Analysis Techniques using SWAN and REANA (part 2 of 3)

Thursday 7 November 2024 09:00 (1 hour)

In the second session of this series, we shall present REANA reusable and reproducible analysis platform. REANA allows researchers to structure their data analyses by means of declarative workflow languages (CWL, Snakemake, Yadage) and run containerised data analysis pipelines on remote compute clouds (Kubernetes, HTCondor, Slurm).

In the first part of this session, we shall discuss the notions of computational reproducibility and reusability, underlying the importance of encapsulating the original computing environments by means of containers and documenting the steps necessary to arrive at results. We shall provide a brief introduction to declarative workflow languages and discuss its pros and cons when compared to imperative analysis code programming.

In the second part of this session, the participants will familiarise themselves with the REANA platform by means of running a simple analysis example. We shall use the https://reana.cern.ch instance at CERN to run a RooFit demo example.

Presenters: DONADONI, Marco (CERN); SIMKO, Tibor (CERN)