



Contribution ID: 33

Type: **Tutorial or Skill-Up**

A short introduction to Nextflow: Bring your data science pipelines to the next level

Wednesday 26 February 2025 11:00 (1h 30m)

Analysing data typically consists of several steps with dedicated tools, chained one after another. In theory, all this can be achieved with well-written scripts. However, workflow managers help developers increase the reproducibility of their pipelines and results by providing features for workflow and data provenance, portability, readability, and fast prototyping.

Nextflow is a reproducible, versatile and powerful workflow management system designed to simplify the development and execution of data-driven computational pipelines. It is predominantly used in bioinformatics, but its flexibility makes it suitable for other scientific domains and data science. It's designed to be environment-agnostic, meaning that workflows can be executed across a variety of computing platforms, including high performance clusters and cloud services, without modification. It comes with close integration with software containers (e.g., Docker, Singularity) and Conda and Spack environments.

Nextflow has an active community, including nf-core: a community-driven movement that aims to establish best practices, tools, and guidelines for developing, testing, and curating Nextflow pipelines. The nf-test test framework enables you to test all components of your data science pipeline, from end-to-end testing of the entire pipeline to specific tests of processes or custom functions.

In this workshop, we will briefly introduce Nextflow and valuable nf-core resources: We will cover the fundamental components of a Nextflow script and work on a small hands-on example workflow. Our goal is to lay the foundation with Nextflow, empowering you to optimize your pipeline development and enhance the efficiency and scalability of your scientific research.

Training material and prerequisites:

- Bring your own laptop ☑
 - Basic knowledge how to work on a terminal
 - It is highly recommended to use Gitpod
 - If this is not option for you, please checkout the Environment Setup
 - Training material: Hello Nexflow
-

I want to participate in the youngRSE prize

no

Primary authors: LATARETU, Marie; WOLK, Paul (Robert Koch-Institut)

Presenters: LATARETU, Marie; WOLK, Paul (Robert Koch-Institut)

Track Classification: Data and Software Management: computational reproducibility