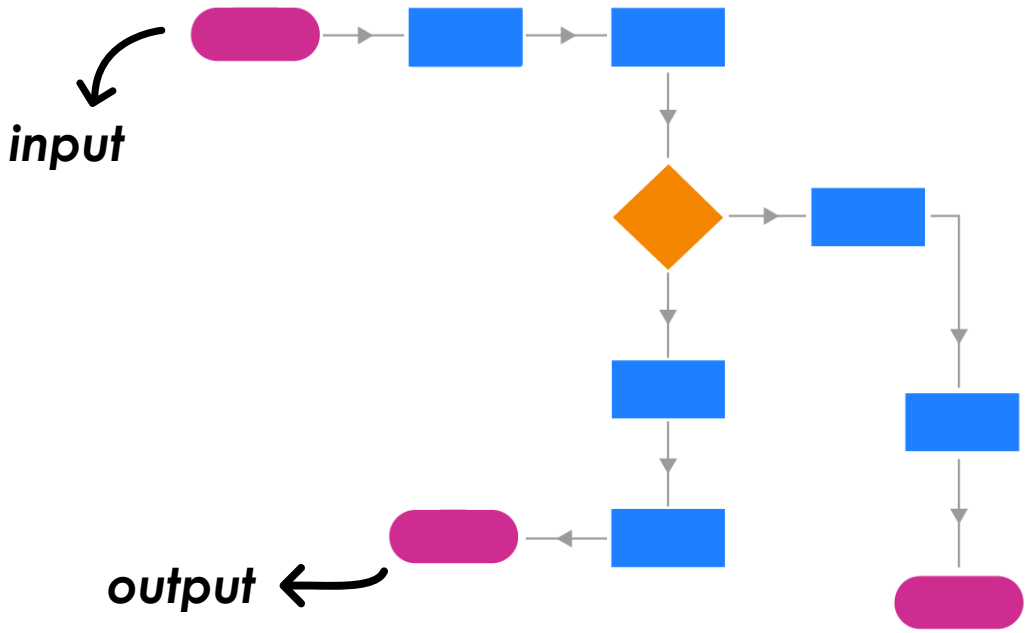
COMMON
WORKFLOW
LANGUAGE

Introduction to workflow languages

Common **W**orkflow **L**anguage (**CWL**)

What is a data-analysis workflow?

Workflow the process that moves a scientific investigation from raw data to coherent research question to insightful contribution[1]



1. Accepts one or more inputs
2. follows a logical sequence of steps
3. Produces one or more outputs

[1] https://doi.org/10.1371/journal.pcbi.1008770

COMMON WORKFLOW LANGUAGE
elixir GREECE
ATHENA Research & Innovation Information Technologies
INAB INSTITUTE OF APPLIED BIOSCIENCES
CERTH CENTRE FOR RESEARCH & TECHNOLOGY HELLAS

What is CWL?

CWL stands for **C**ommon **W**orkflow **L**anguage and it is a standard for describing computational data-analysis workflows.

CWL has applications mainly in computational sciences and the research areas of:

- Bioinformatics
- Medical Imaging
- Astronomy
- Physics
- Chemistry

They key goal is to produce **workflows** that are **reproducible in different computational environments**

How are CWL workflows structured?

Tools ⚙

- They are **wrappers** that describe to the CWL engine how a command-line tool works
- They **list its inputs and outputs**, and the command to produce one from the other
- A lot of them are **already available for commonly used tools**, and can be used directly into workflows

Workflows

- Explain **how tools are connected to each other** and in what order
- Are generally **project-specific**
- Can be nested inside each other*

*This is not going to be explained in this workshop. Available resources can be found here

CWL resources

[CWL website]

[CWL user guide]

[CWL GitHub]