# *SNAKEMAKE BACKEND FOR RECAST: IMPLEMENT A SNAKEMAKE BACKEND FOR RECAST WORKFLOWS*

Fellow: Andrii Povsten

IRIS-HEP Summer Fellowship

Mentors: Matthew Feickert (UWM), Lukas Heinrich (TUM)

RECAST is a framework for extending the impact of existing analyses performed by high-energy physics experiments. This framework allows scientists to access and reuse workflows that have already been performed.

Yadage is a current RECAST backend.

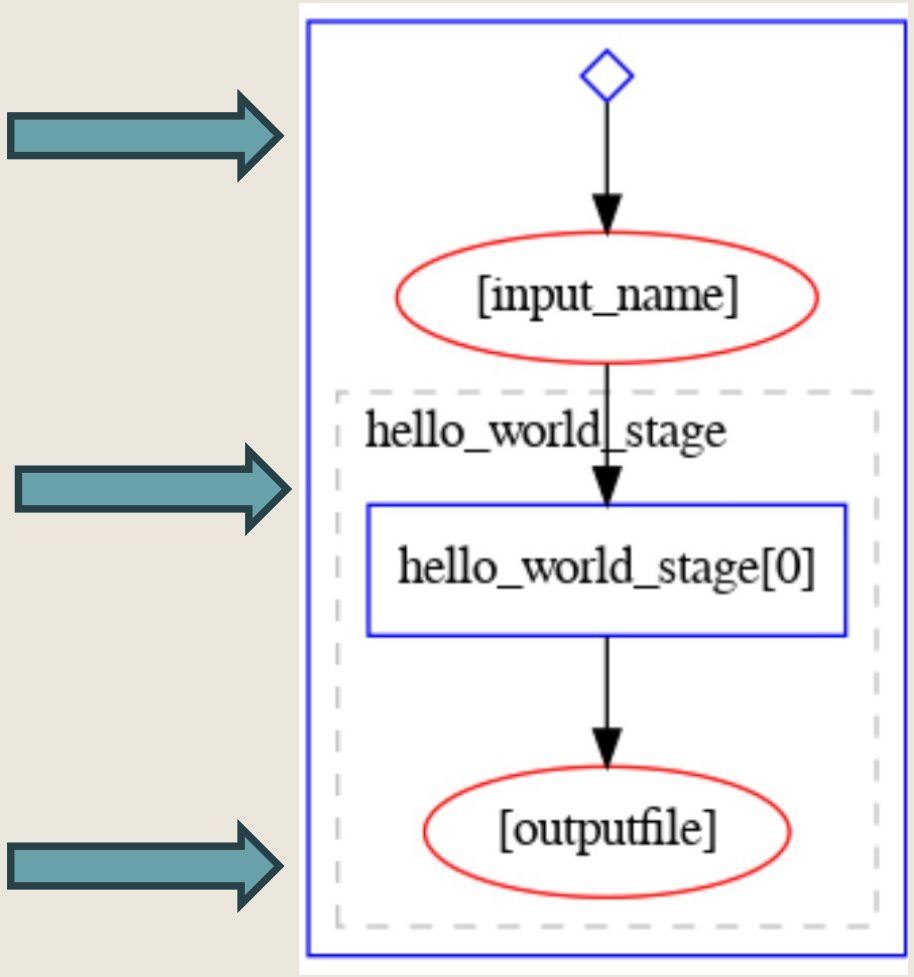
Yadage represents workflow as a DAG.

RECAST can be integrated with REANA cluster, which can collaborate with Snakemake

# Example workflow with Yadage

## Example Workflow

```
cat << 'EOF' > workflow.yml
stages:
- name: hello_world
  dependencies: [init]
  scheduler:
    scheduler_type: singlestep-stage
    parameters:
      name: {step: init, output: name}
      outputfile: '{workdir}/hello_world.txt'
    step:
      process:
        process_type: 'string-interpolated-cmd'
        cmd: 'echo Hello my Name is {name} | tee {outputfile}'
      publisher:
        publisher_type: 'frompar-pub'
        outputmap:
          outputfile: outputfile
      environment:
        environment_type: 'docker-encapsulated'
        image: busybox
EOF
```




## Why Snakemake?

Recreating the same analysis pipeline makes work easily shareable.

Python based workflow management system.

Widely used and robustly supported through contributions from all communities.

Currently supports containers and seems to offer a better supported alternative to Yadage.

Example workflow with Snakemake

```
rule count_words:
    input:
        "input.txt"
    output:
        "output.txt"
    run:
        with open(input[0]) as file:
            text = file.read()
        word_count = len(text.split())
        with open(output[0], "w") as file:
            file.write(str(word_count))
```

```
rule all:
    input:["out/1", "out/2"]
    output:".status"
    shell:"touch"
rule copy:
    input:"in/{file}"
    output:"out/{file}"
    shell:"cp{input}{output}"
```

The key objectives for this project are:

Create a new backend for RECAST with the same yadage functionality.

Submit RECAST Snakemake workflows to the REANA Snakemake workflow engine.

Create and execute new test cases that highlight any differences in strengths between Yadage and Snakemake.

Thank you for your attention!