Analysis Reproducibility with REANA on Kubernetes
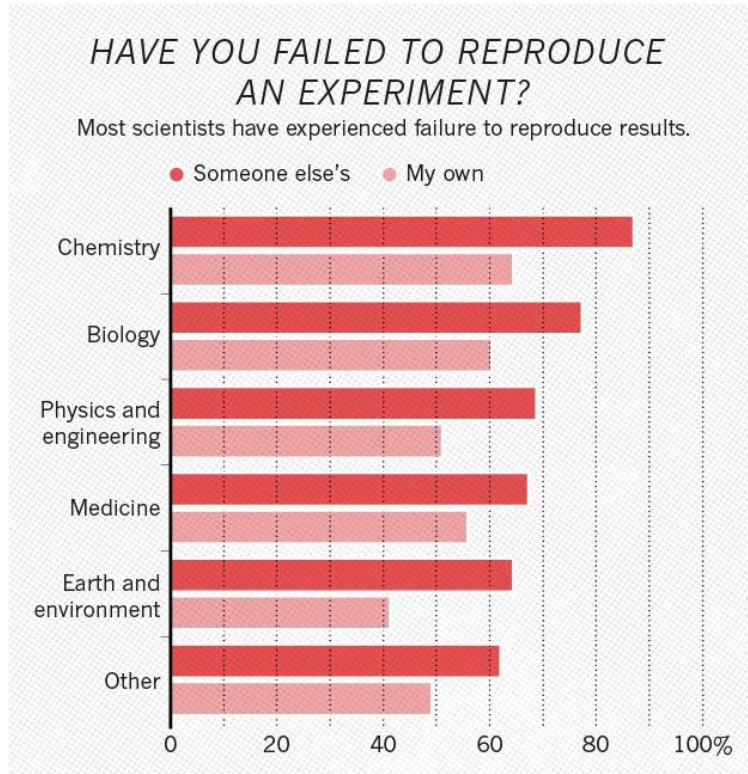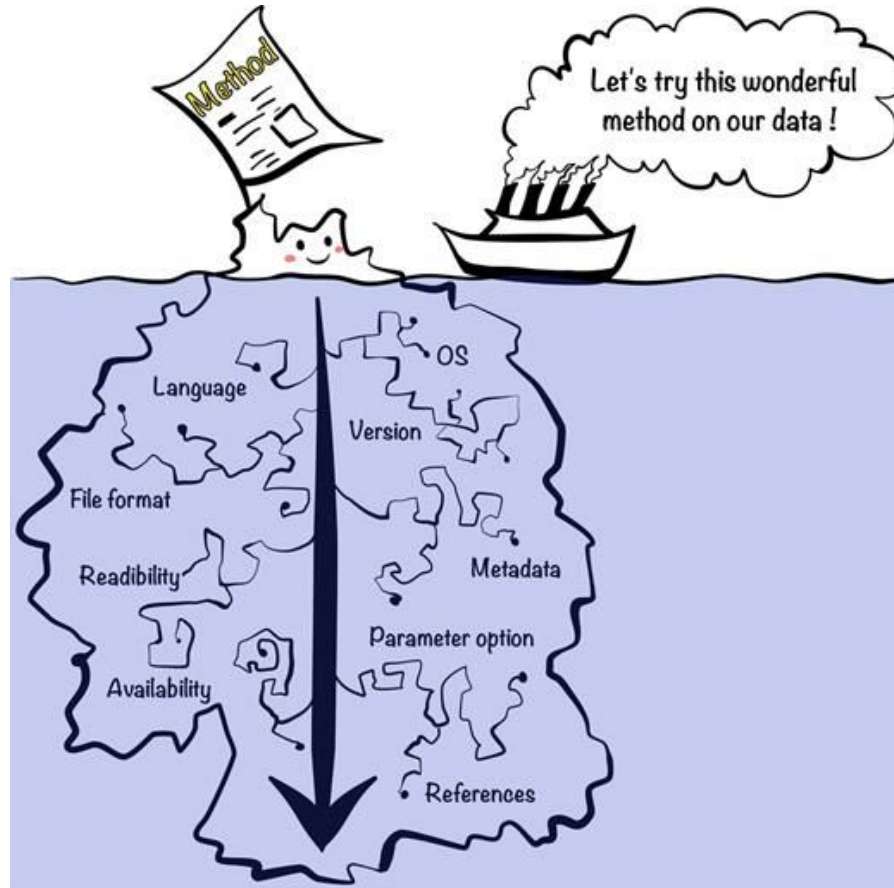
Diego Rodriguez
on behalf of the REANA team

Reproducibility?

HAVE YOU FAILED TO REPRODUCE
AN EXPERIMENT?
Most scientists have experienced failure to reproduce results.

● Someone else's   ● My own

Long-term value of knowledge?

It is harder than you think

The four questions

where is the **data**?
EOS, local hard drive…
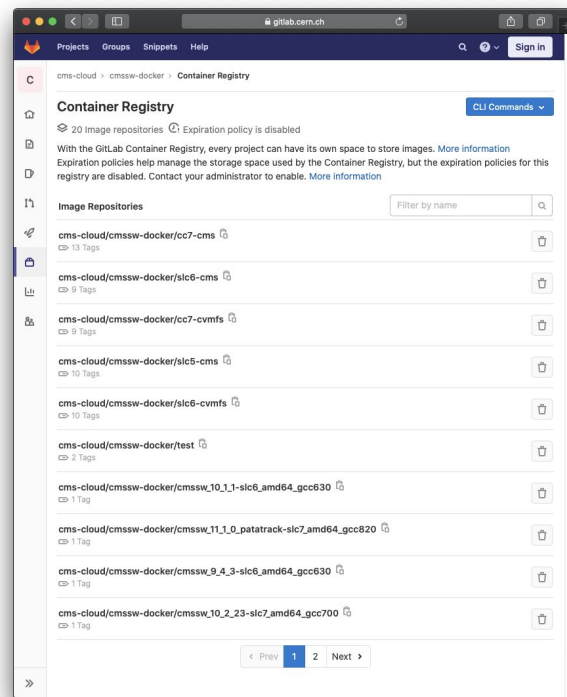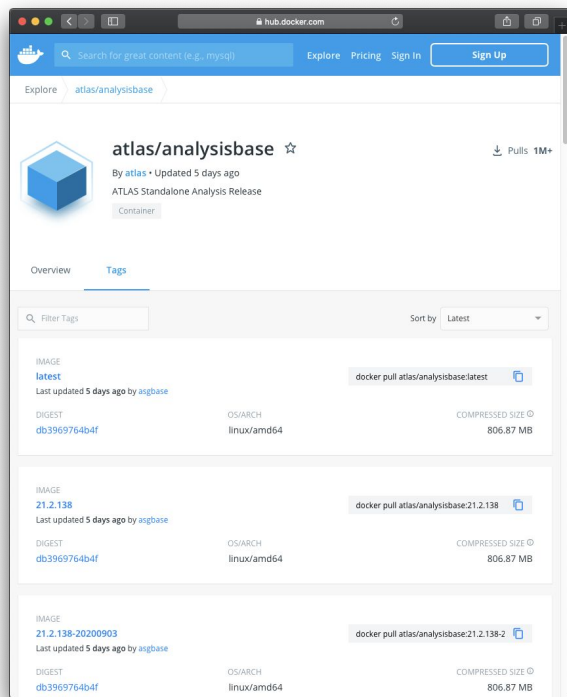
where is the **code**?
GitLab, local copy...

where is the **environment**?
my own computer, lxplus...

where is the **workflow**?
I remember it, bash script, readme...

reana

ATLAS and CMS containerised environments

A recipe for reproducibility
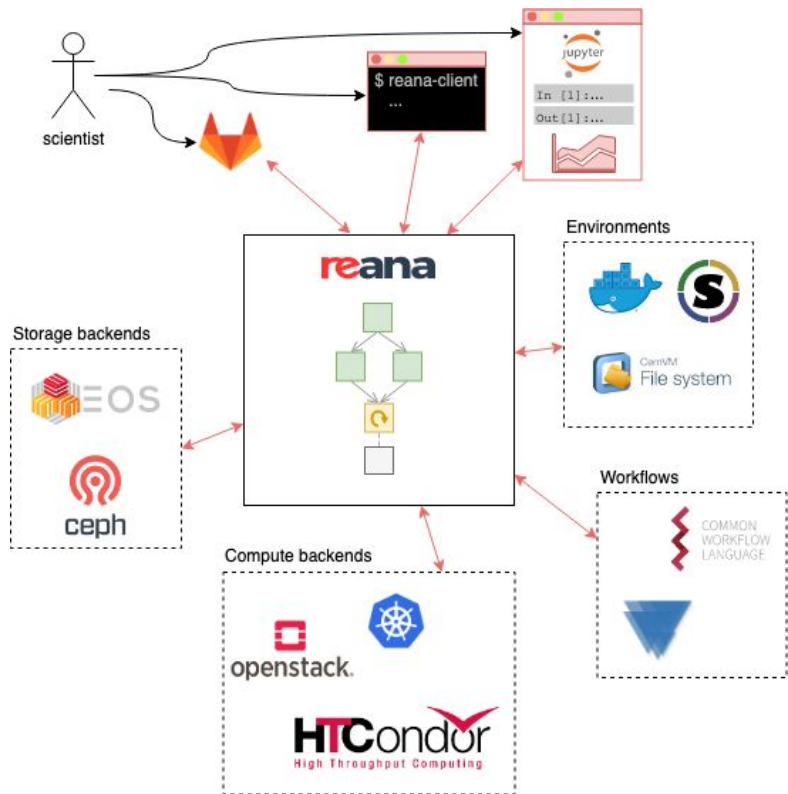
🥑 + 🔪 + 🥘 + 👩‍🍳👨‍🍳 = 😄

`data + code + environment + workflow =` ♻️

Overview

Different **storage backends**

Multiple **compute backends**:
Kubernetes, HTCondor, Slurm

Different **workflow languages** (from
simple pipelines to complex
map-reduce operations)

Integrations with other systems

Private **GitLab** images

Possibility to access private **EOS**
resources with **Kerberos**

Support for **VOMS**-proxy certificates

Demo

Demonstrate REANA UI and CLI

Walk-through simple physics model

analysis

GitLab integration

Live architecture overview

Kubernetes cluster/admin view


REANA CLI
-
end-user view

Kubernetes/OpenStack
-
admin view

A declarative approach

```
version: 0.6.0
inputs:
 files:
    - code/gendata.C
    - code/fitdata.C
 parameters:
    events: 20000
    data: results/data.root
    plot: results/plot.png
workflow:
 type: serial
 specification:
    steps:
      - name: gendata
        environment: 'reanahub/reana-env-root6:6.18.04'
        compute_backend: htcondorcern
        commands:
        - mkdir -p results && root -b -q
'code/gendata.C(${events},"${data}")'
      - name: fitdata
        environment: 'reanahub/reana-env-root6:6.18.04'
        compute_backend: htcondorcern
        commands:
        - root -b -q 'code/fitdata.C("${data}","${plot}")'
outputs:
 files:
    - results/plot.png
```
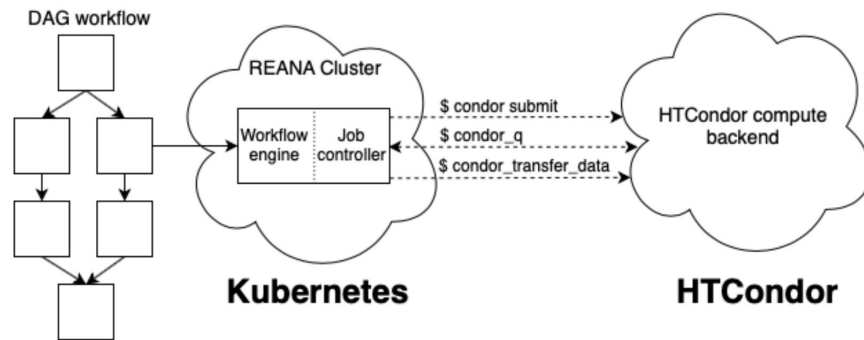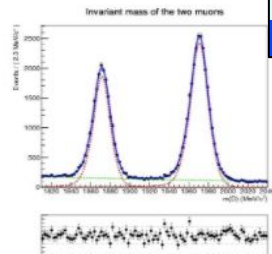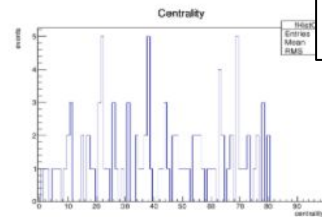

reana

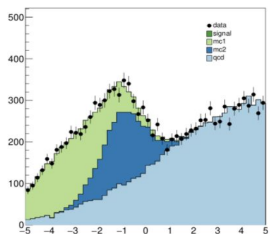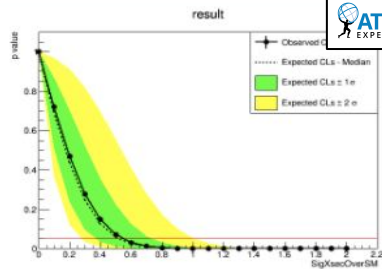# Scaling out: Beyond Standard Model searches
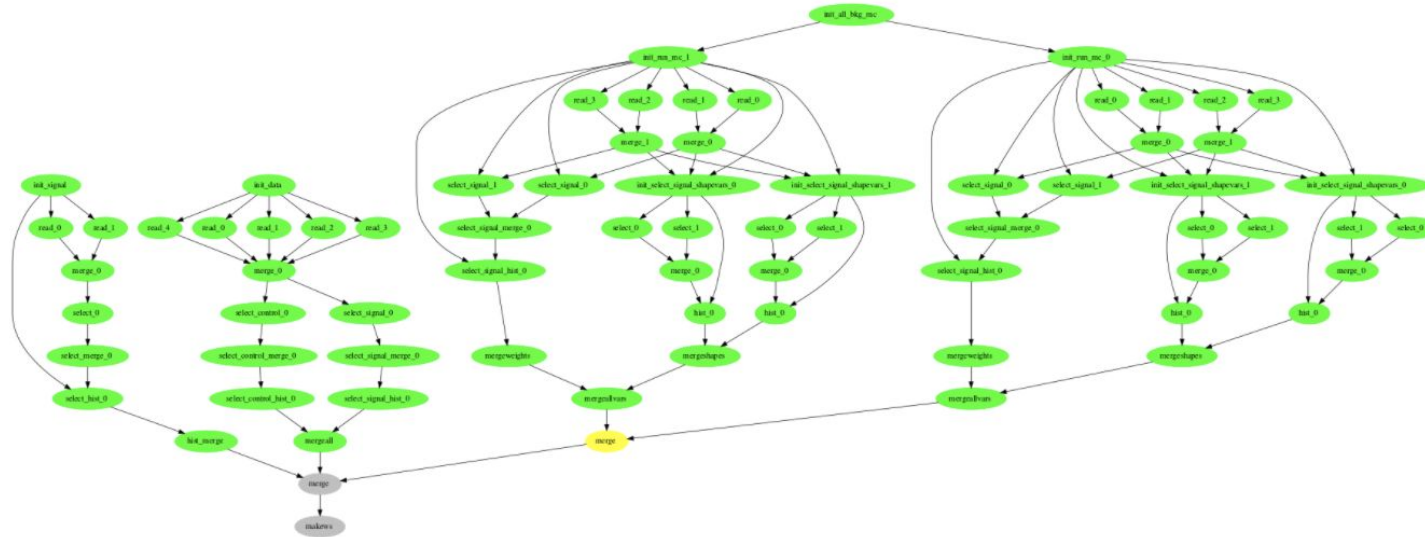
Pilot examples
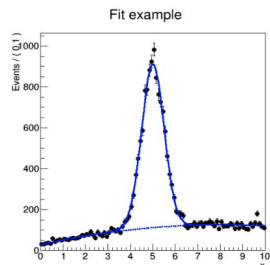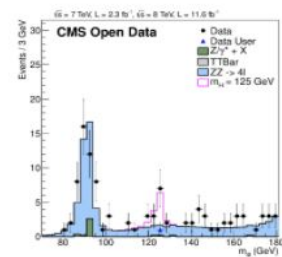
ROOT/RooFit physics analysis

LHCb rare charm decay search

ALICE LEGO train test run

ATLAS BSM search

ATLAS RECAST

CMS Higgs-to-four-leptons

reana@CERN

Install on premises
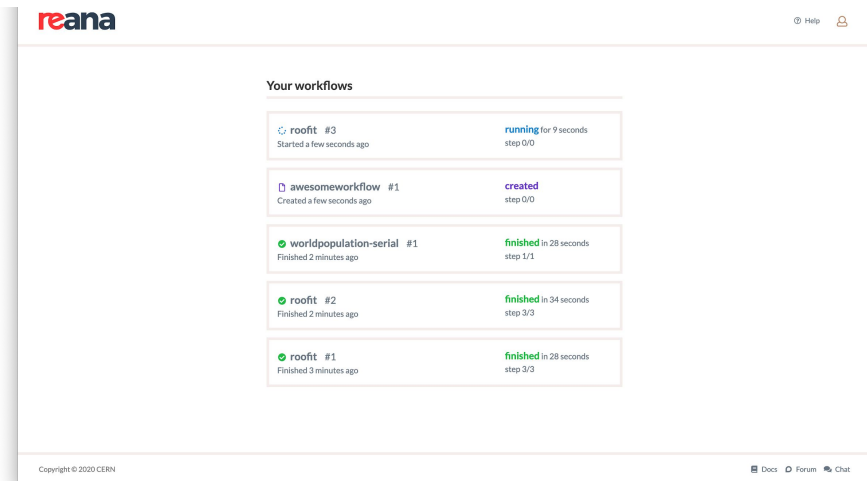
https://reana.cern.ch/

```
$ helm repo add reanahub \
    https://reanahub.github.io/reana

$ helm repo update

$ helm install --devel reana reanahub/reana
```

Helm repository at http://reanahub.github.io/reana, more docs

at http://docs.reana.io/development/deploying-at-scale/

Get in touch!

reana@CERN → https://reana.cern.ch/

Documentation → http://docs.reana.io/

Forum https://forum.reana.io/

Mattermost@CERN https://mattermost.web.cern.ch/it-dep/channels/reana

REANA@Gitter https://gitter.im/reanahub/reana

REANA@Github https://github.com/reanahub/reana

Q&A