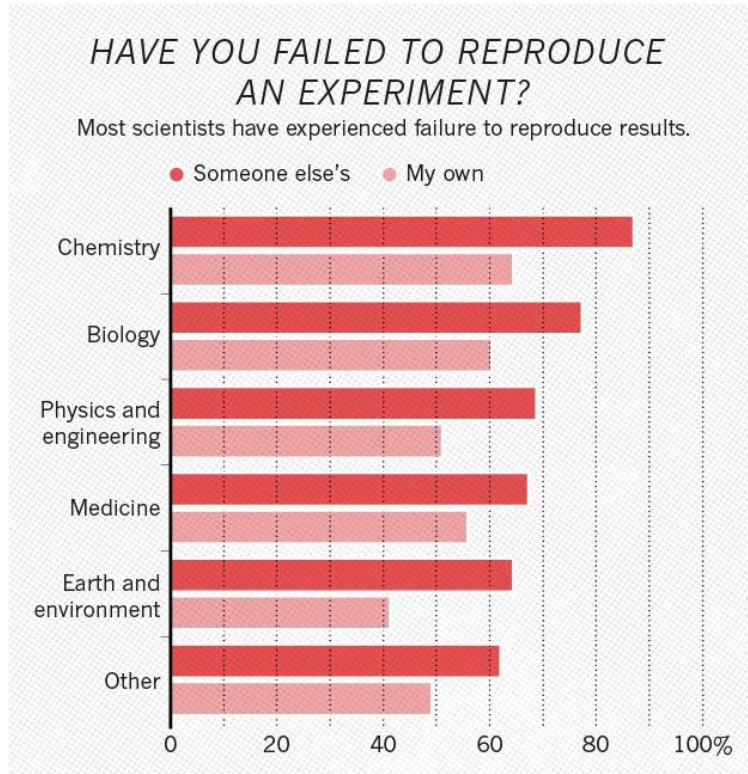


Analysis Reproducibility with REANA on Kubernetes

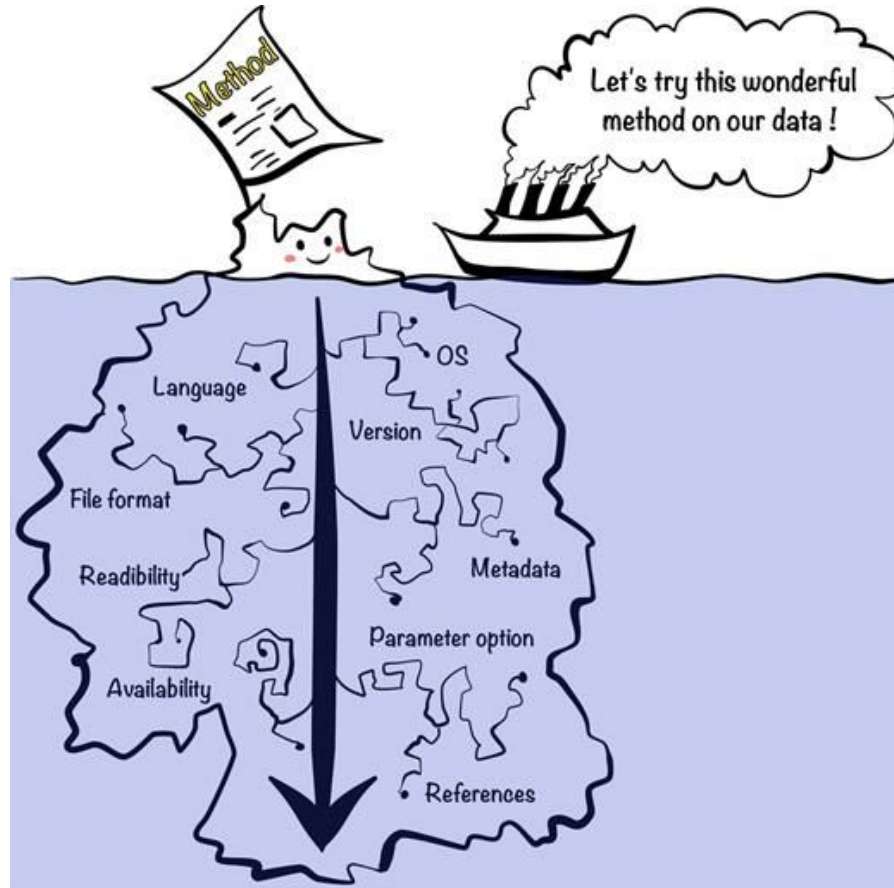
*Diego Rodriguez
on behalf of the REANA team*

Reproducibility?



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...

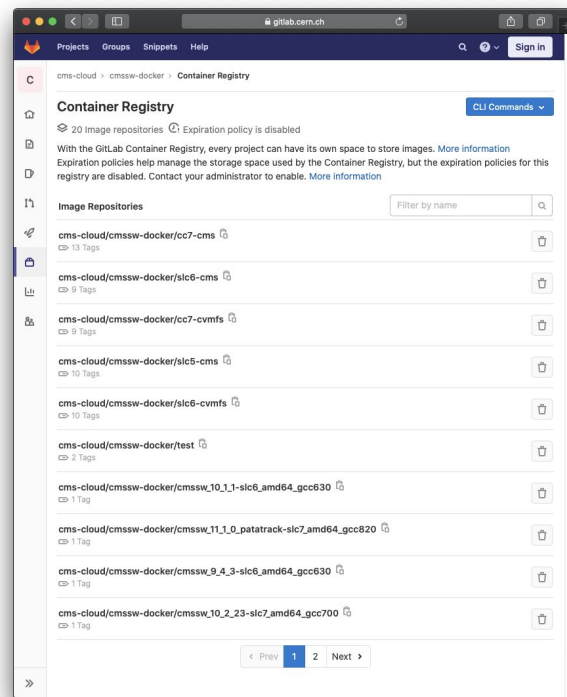
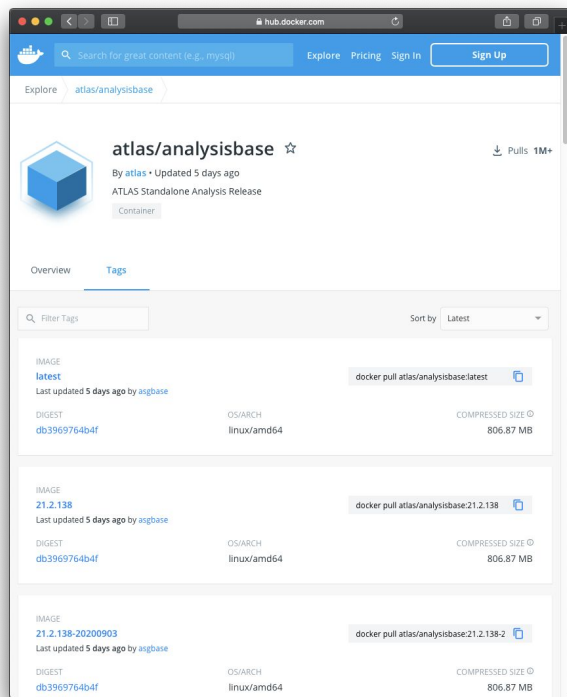
where is the **environment**?

my own computer, lxplus...

where is the **workflow**?

I remember it, bash script, readme...

ATLAS and CMS containerised environments



A recipe for reproducibility



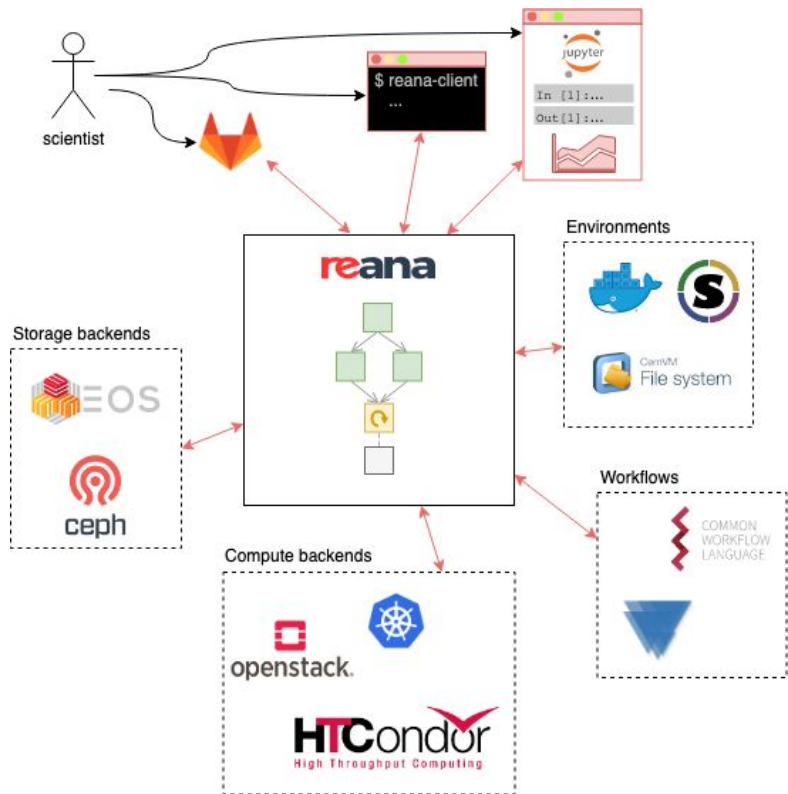
data + code + environment + workflow = 

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-qa) [rodrigdi@lxplus765 reana-demo-root6-roofit]$ reana-client list
NAME                RUN_NUMBER  CREATED          STARTED          ENDED           STATUS
Roofit-HTCondor    1            2020-09-09T07:05:12  2020-09-09T07:05:13  2020-09-09T07:19:35  finished
bss-search          1            2020-09-09T07:03:02  2020-09-09T07:03:06  2020-09-09T07:58:54  finished
REANA example - ROOT6 and Roofit 1            2020-09-08T15:42:27  2020-09-08T15:42:38  2020-09-08T15:54:15  finished
myanalysis          1            2020-09-08T13:59:26  2020-09-08T14:08:14  2020-09-08T14:01:03  finished
(reana-qa) [rodrigdi@lxplus765 reana-demo-root6-roofit]$

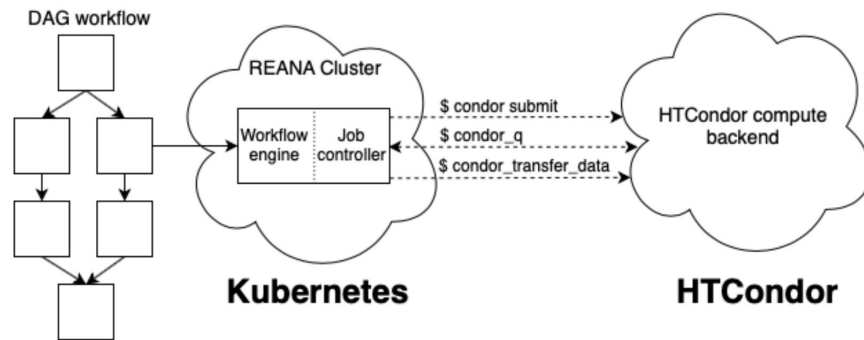
[rodrigdi@lxplus765 reana]$ kubectl get pods -w
NAME                READY   STATUS    RESTARTS   AGE
reana-cache-798568899-9924s    1/2     Running   0           5d18h
reana-message-broker-847844b6f9-mqppr    1/2     Running   0           5d18h
reana-server-b4d596c0b-kg77       2/2     Running   0           5d17h
reana-system-status-1599631208-0722c    0/2     Completed 0           156m
reana-ui-579798cfc-reece          1/1     Running   0           5d18h
reana-workflow-controller-5bf548f7fc-cw4f 2/2     Running   1           5d17h
```

REANA CLI
-
end-user view

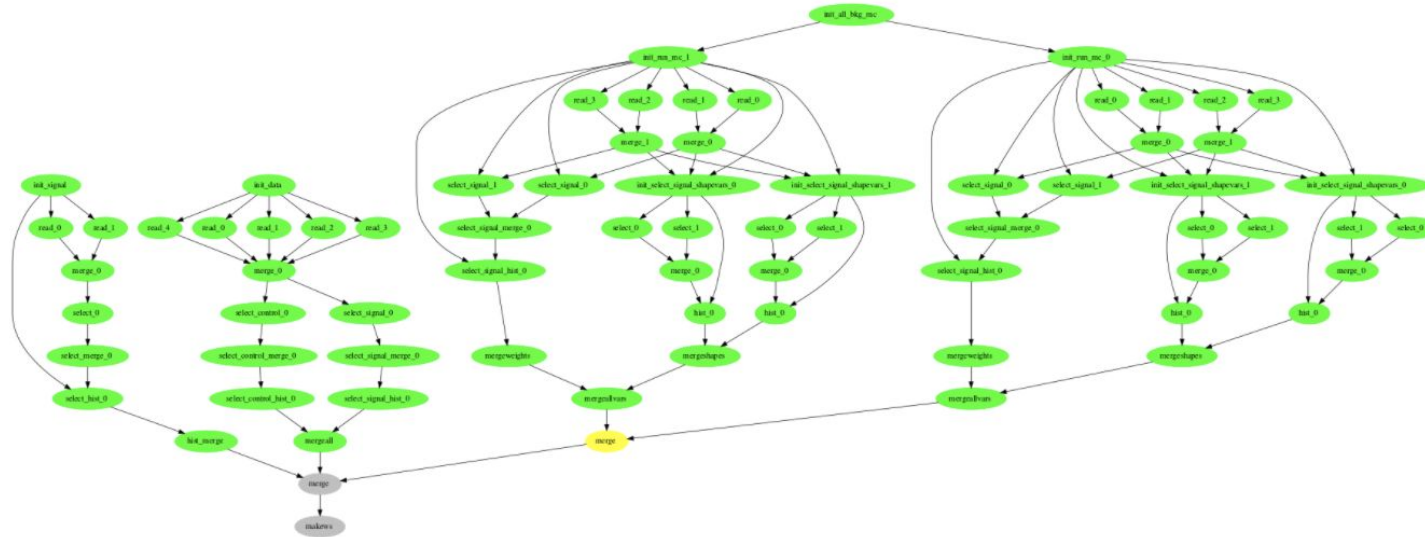
Kubernetes/OpenStack
-
admin view

A declarative approach

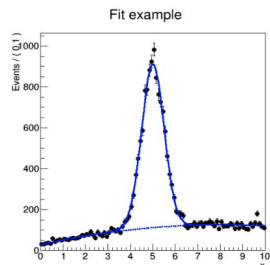
```
version: 0.6.0
inputs:
files:
  - code/gendata.C
  - code/fitdata.C
parameters:
  events: 20000
  data: results/data.root
  plot: results/plot.png
workflow:
type: serial
specification:
steps:
  - name: gendata
    environment: 'reanahub/reana-env-root6:6.18.04'
    compute_backend: htcondorcern
    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
```



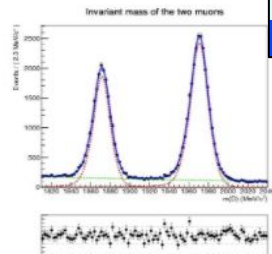
Scaling out: Beyond Standard Model searches



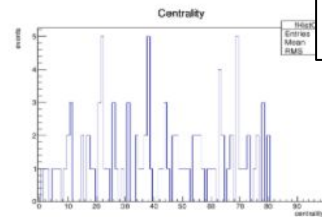
Pilot examples



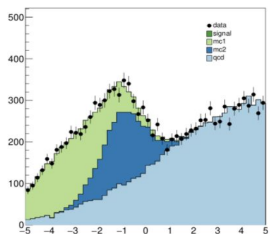
ROOT/RooFit physics analysis



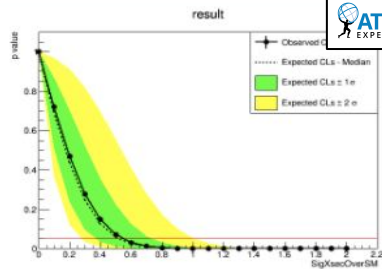
LHCb rare charm decay search



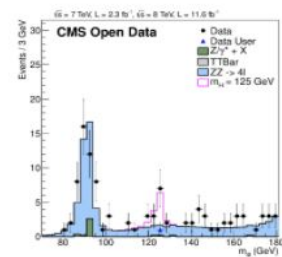
ALICE LEGO train test run



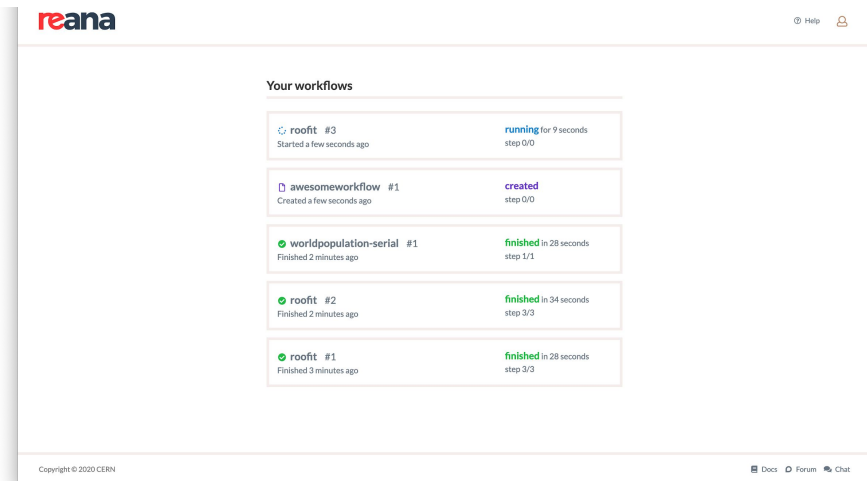
ATLAS BSM search



ATLAS RECAST



CMS Higgs-to-four-leptons



<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