

**CERN**  
**Analysis Preservation**

# CERN Analysis Preservation and REusable ANALYSIS

Sünje Dallmeier-Tiessen  
(for many other collaborators)  
CERN

HSF/WLCG, Naples, March 27th 2018

CERN Analysis Preservation



<http://analysispreservation.cern.ch/>



<http://github.com/cernanalysispreservation>

REANA



<http://reanahub.io/>



<http://github.com/reanahub>



@reanahub

**CERN IT** H. Hirvonsalo, D. Rodríguez, T. Šimko **CERN SIS** S. Dallmeier-Tiessen, R. Dasler, S. Feger, P. Fokianos, A. Lavasa, I. Tsanaktsidis, A. Trzcinska **ALICE** M. Gheata, C. Grigoras, M. Zimmermann **ATLAS** K. Cranmer, L. Heinrich, A. Sanchez Pineda, D. Rousseau, F. Socher **CMS** A. Calderon, A. Geiser, A. Huffman, K. Lassila-Perini, T. McCauley, A. Rao, A. Rodriguez Marrero **LHCb** S. Amerio, B. Couturier, S. Neubert, A. Trisovic **CERN CernVM** J. Blomer **CERN Kubernetes** R. Rocha **CERN EOS** L. Mascetti **DASPOS** M. Hildreth, H. Meng, D. Thain, A. Vyushkov **DPHEP** J. Shiers

# Outline

Concepts

CERN Analysis Preservation

REANA

Challenges

# CERN Analysis Preservation

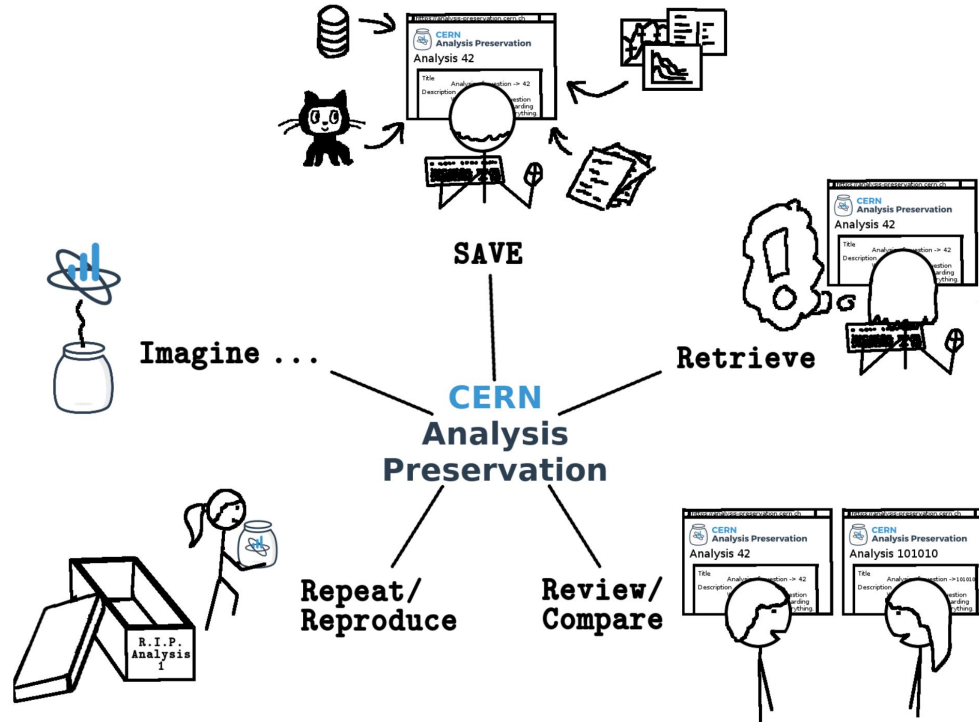


**CERN**  
**Analysis Preservation**

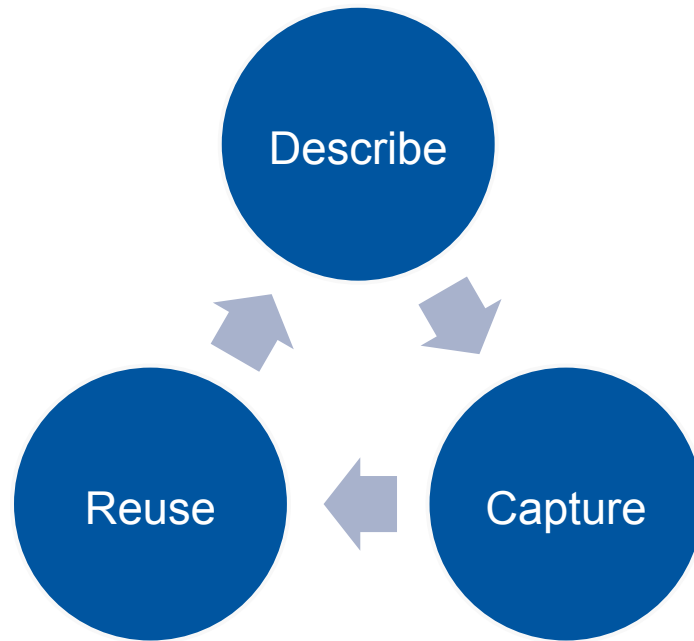
# Use cases



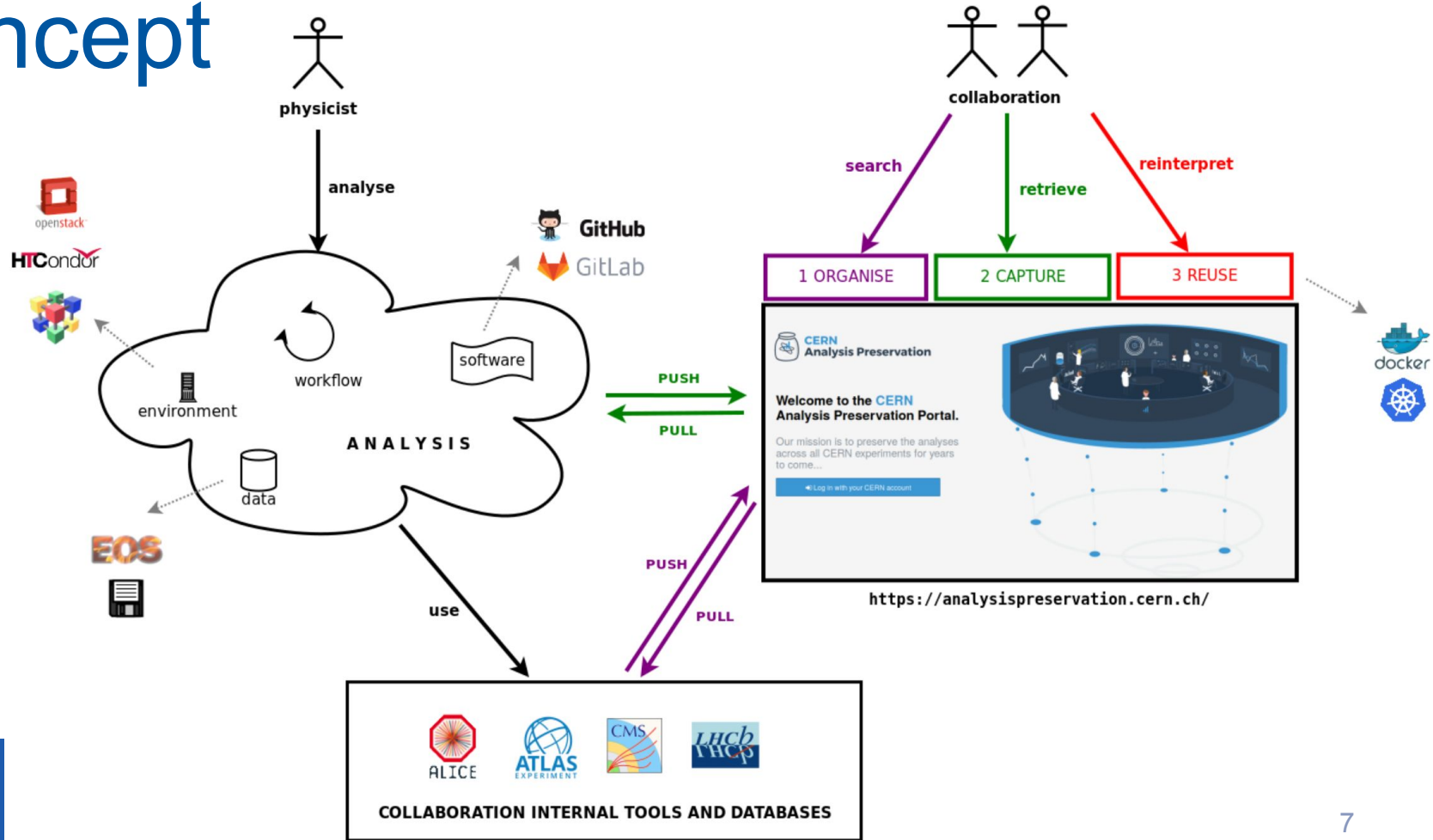
## CERN Analysis Preservation



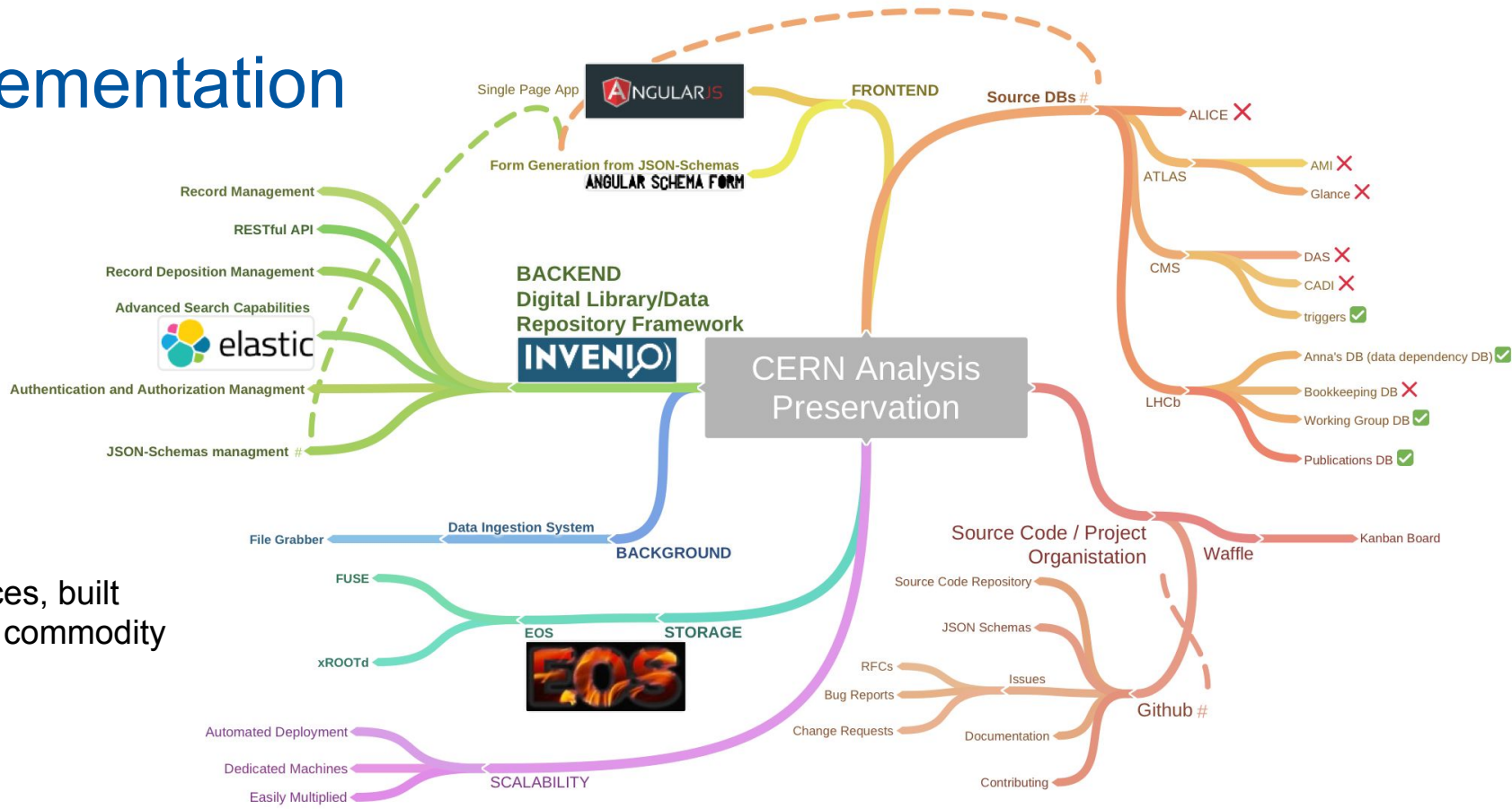
# Three pillars



# Concept



# Implementation



Many pieces, built mostly on commodity solutions





## CERN Analysis Preservation

### Welcome to the CERN Analysis Preservation Portal.

Our mission is to preserve the analyses  
across all CERN experiments for years  
to come...

[Log in with your CERN account](#)



<http://analysispreservation.cern.ch/>

# Current state: CERN Analysis Preservation

Command-line client to ease submission through REST API

Import software from GitLab

Connections to collaboration databases to profit from existing information

Running user tests

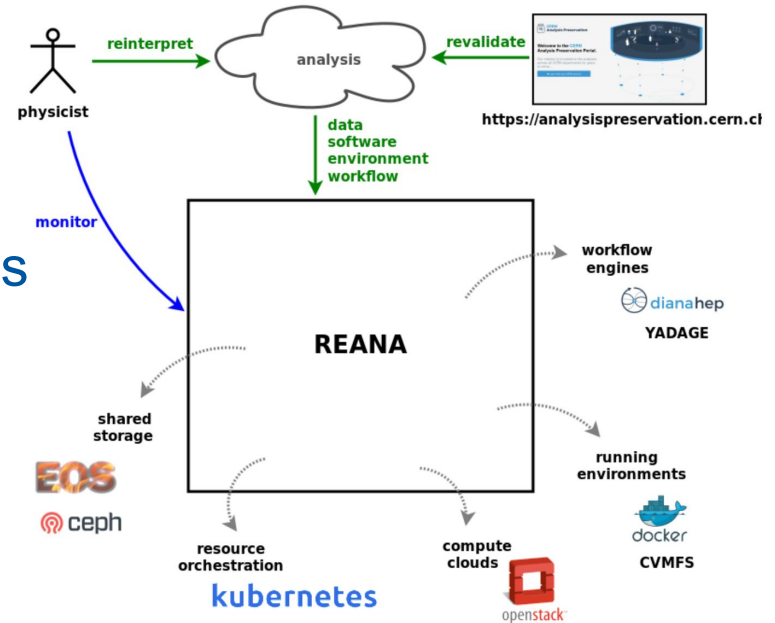
<http://analysispreservation.cern.ch/>

The screenshot displays the CERN Analysis Preservation web interface. The top navigation bar includes the CERN logo, the text "Analysis Preservation", and the LHCb logo. A search bar is present on the right, along with "Create", user profile, and help icons. The main content area is titled "LHCb Analysis 17/05/2017, 08:19:57" and features a "Save" button. The left sidebar contains a navigation menu with sections: "Basic Information" (8/3 req), "DST selection" (2), "Code" (3), "Analysis Steps" (1 items), and "Additional Resources" (4). The "Basic Information" section lists: Analysis Name, Measurement, Proponents, Status, Reviewers, Review eGroup, Working Group, and Keywords. The "Code" section lists: Application, Platform, User code (0 items), Production information (2), Collision Data (0 items), and MC Data (0 items). The main form area is titled "PRODUCTION INFORMATION" and contains a "COLLISION DATA" section with an "Add New" button. Below this, there are input fields for "Bookkeeping path" (with an "Autofill" button), "Processing Pass", "RECONSTRUCTION SOFTWARE" (with a "Name" field), and "Version".

REANA = REusable ANAlyses

# REANA

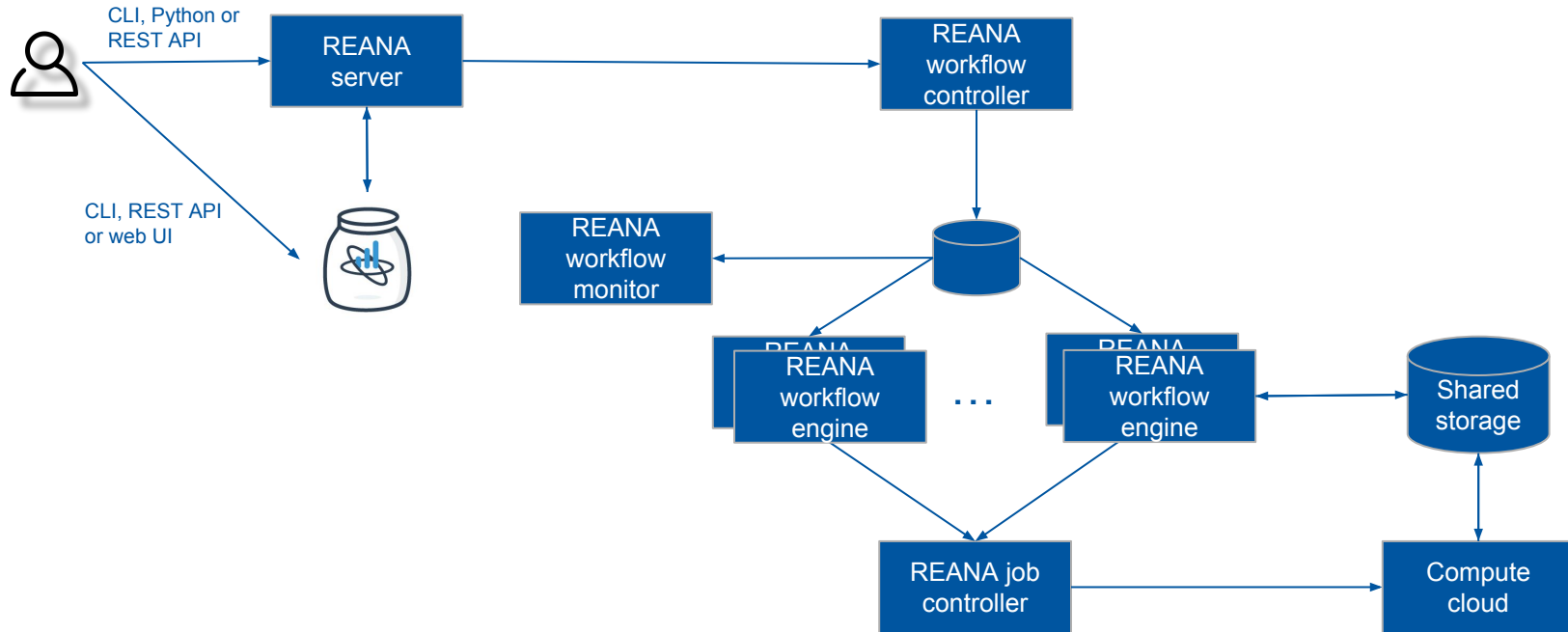
- Building system to instantiate preserved analysis on the cloud  
<https://reanahub.io>
- Cloud native
- Aiming to support multiple scenarios
  - compute clouds
  - distributed storage systems
  - workflow engines
  - container technologies



# To consider for reproducible research

- **Data?** (root files, csv files etc...)
- **Software?** (analysis code)
- **Environment?** (OS, libraries, CPU, memory...)
- **Workflow?** (complex graphs)

# REANA architecture

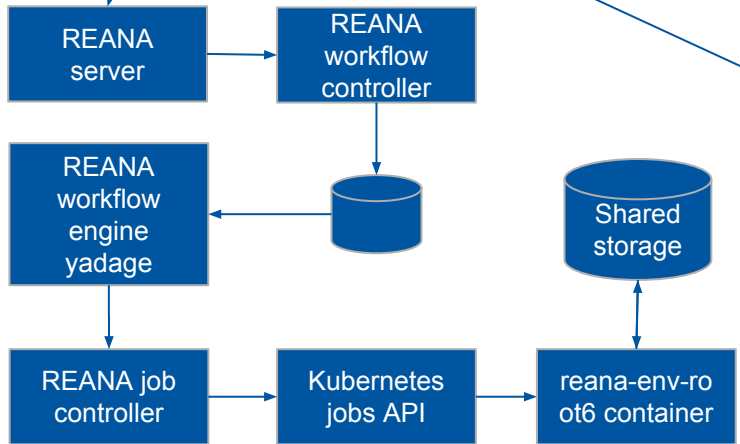


# REANA example

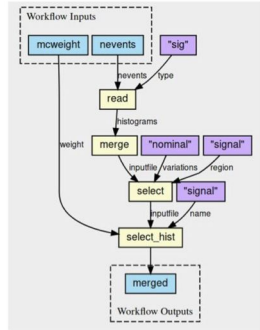
```

$ pip install reana-client
$ export REANA_SERVER_URL=https://reana.cern.ch
$ reana-client run workflow.yml
[INFO] Starting reana-demo-root6-roofit analysis...
[...]
[INFO] Done. You can see the results in the 'output/' directory.
    
```

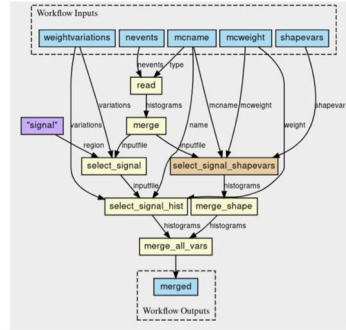
*reana-client  
not released yet!*



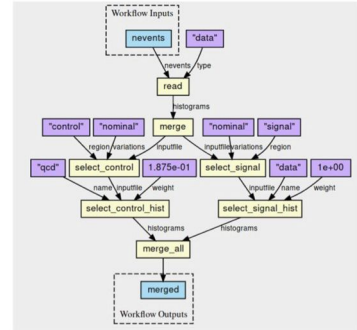
more at <https://github.com/reanahub/reana-demo-root6-roofit>



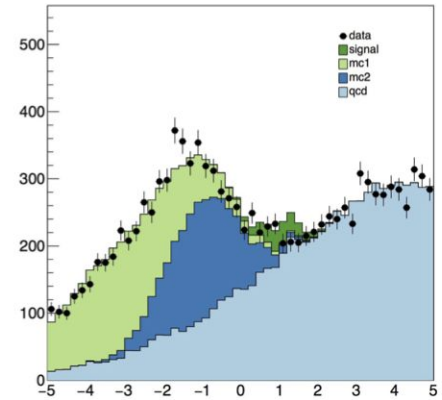
sig



mc



data



Atlas full full chain example

Workflows represented in yml files.

Common Workflow Language now supported

# What's next

- Extend CAP beta testing to more experiments
- Build bridge between CAP and REANA
- Test REANA more widely

## Challenge

- Cope with ever increasing data size and ever changing requirements (changing way of doing analysis)



CERN Analysis Preservation



<http://analysispreservation.cern.ch/>



<http://github.com/cernanalysispreservation>

REANA



<http://reanahub.io/>



<http://github.com/reanahub>

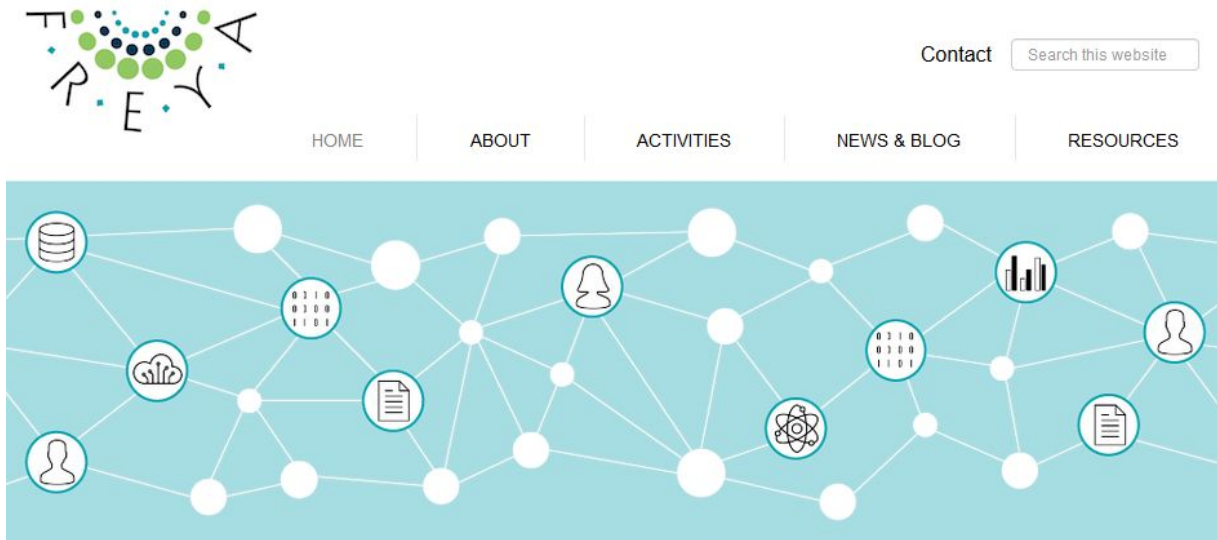


@reanahub

**CERN IT** H. Hirvonsalo, D. Rodríguez, T. Šimko **CERN SIS** S. Dallmeier-Tiessen, R. Dasler, S. Feger, P. Fokianos, A. Lavasa, I. Tsanaktsidis, A. Trzcinska **ALICE** M. Gheata, C. Grigoras, M. Zimmermann **ATLAS** K. Cranmer, L. Heinrich, A. Sanchez Pineda, D. Rousseau, F. Socher **CMS** A. Calderon, A. Geiser, A. Huffman, K. Lassila-Perini, T. McCauley, A. Rao, A. Rodriguez Marrero **LHCb** S. Amerio, B. Couturier, S. Neubert, A. Trisovic **CERN CernVM** J. Blomer **CERN Kubernetes** R. Rocha **CERN EOS** L. Mascetti **DASPOS** M. Hildreth, H. Meng, D. Thain, A. Vyushkov **DPHEP** J. Shiers

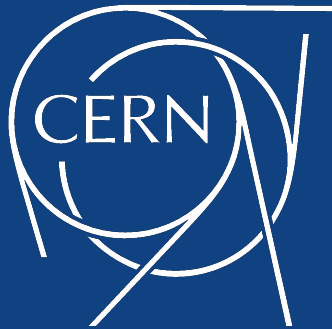
Supported by

[www.project-freya.eu](http://www.project-freya.eu)



## Welcome to FREYA

Connected Open Identifiers for Discovery, Access and Use of Research Resources



Questions?