Connecting REANA and the CERN-VRE: a JupyterLab extension middleware

Rubén Pérez Mercado

Enrique García García, Giovanni Guerrieri

IT-GOV-ENG August 2024

CERN openlab

Table of Contents

01 Introduction

02 Motivation

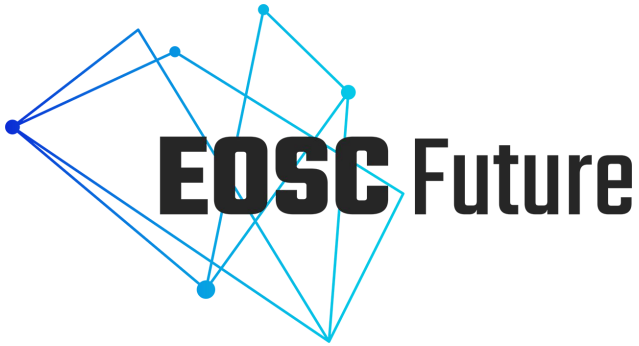
03 Reana JupyterLab extension

04 Conclusions

Introduction

The Virtual Research Environment (VRE):

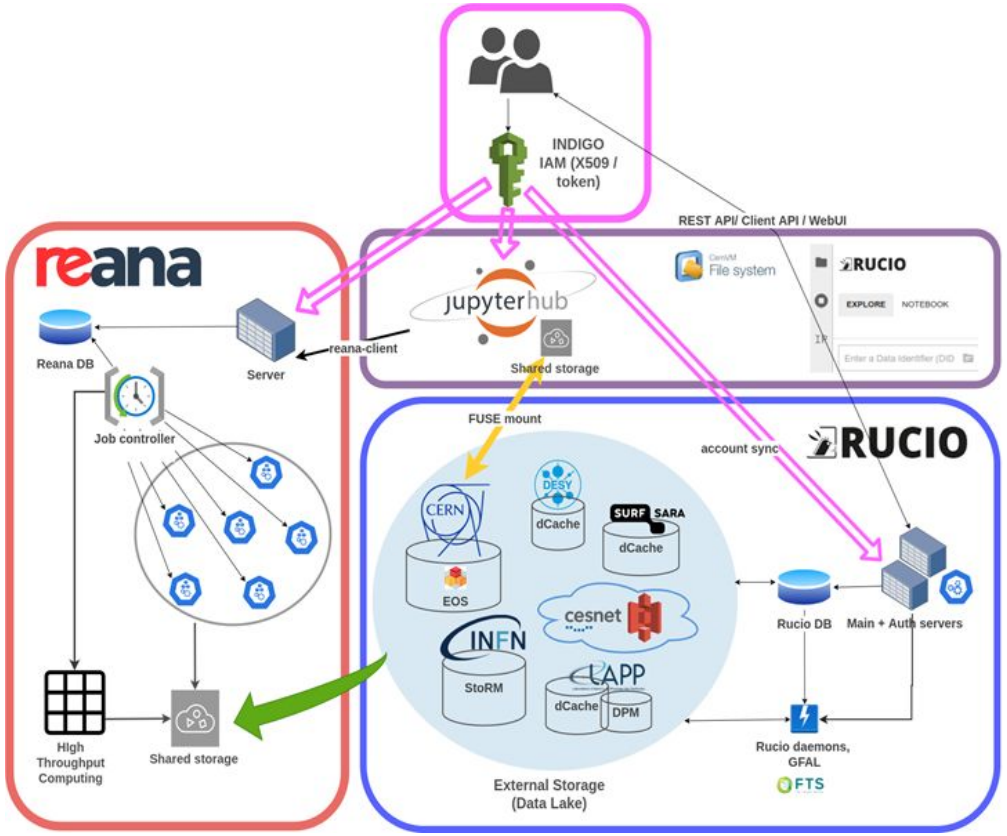
- Collaborative analysis platform where researchers can develop and share end-to-end workflows.
- Adheres to FAIR principles to enhance the usability and impact of scientific data (**F**indable, **A**ccessible, **I**nteroperable, **R**eusable).
- Developed at CERN within the ESCAPE and EOSC (European Open Science Cloud) Future projects.
- Showcases how disciplines ranging from HEP to Astrophysics could benefit from the usage of common technologies.



Introduction

The Virtual Research Environment (VRE):

- **Main components:**
  - A federated Authentication and Authorization layer.
  - A federated distributed storage solution (the Data Lake) with a Data Management framework (Rucio).
  - An enhanced **notebook interface**.
  - A computing cluster with a re-analysis software **(Reana)**

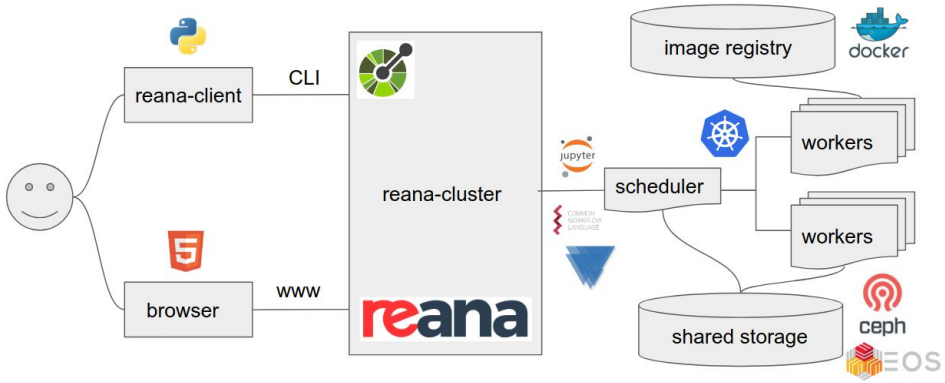


VRE-hub: https://vre-hub.github.io

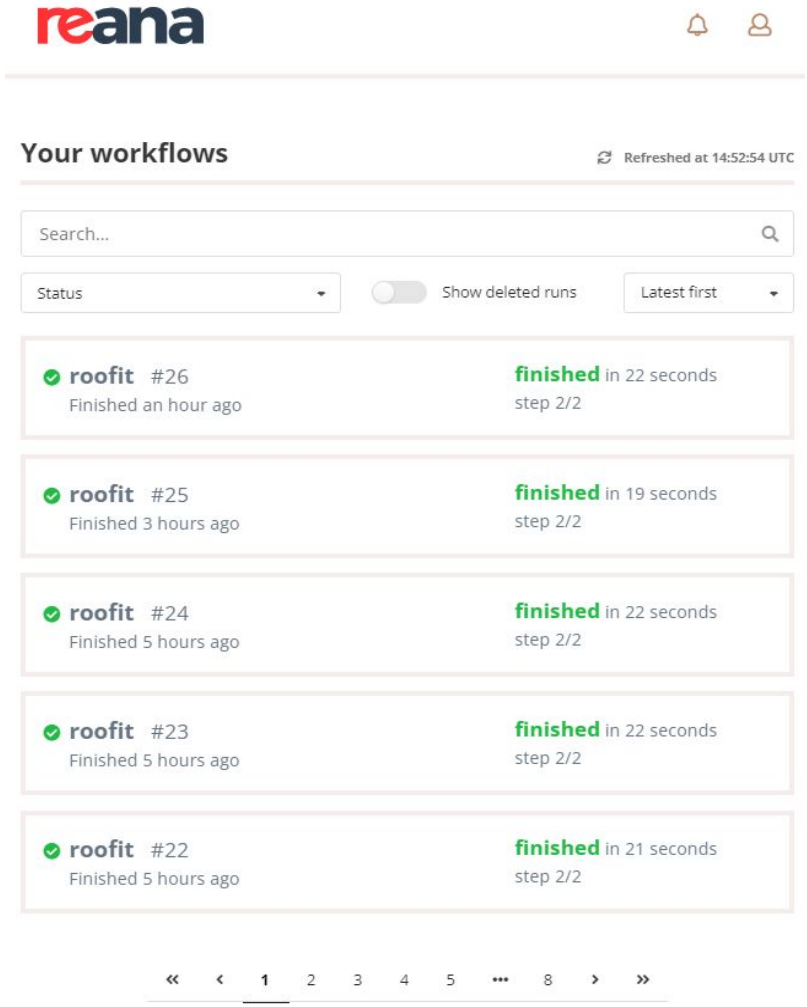
Introduction

REANA:

- REANA is a **re**producible **ana**lysis platform that allows scientists to run containerised data analysis pipelines on remote compute clouds.
- Analyses are defined in YAML files, describing the inputs, workflow steps and the outputs.



Reana: https://docs.reana.io/

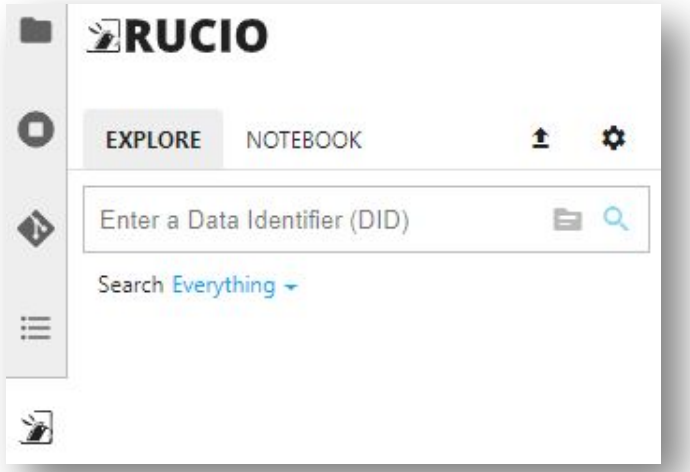


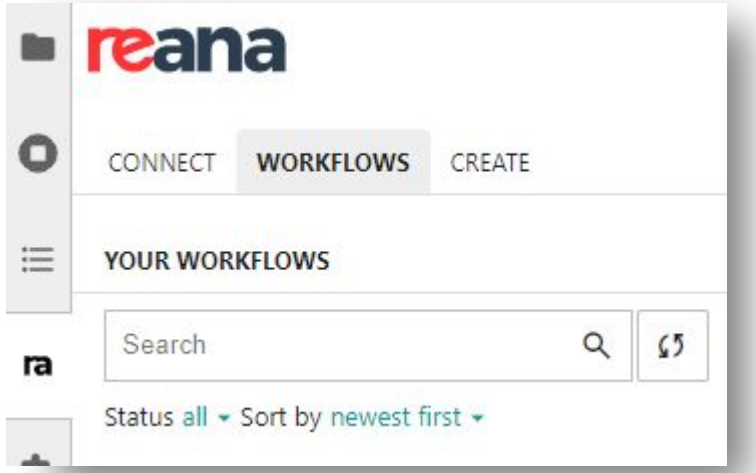
Reana-UI: https://github.com/reanahub/reana-ui

Motivation

- In order to use Reana in our virtual environment we need to know how to use **`reana-client`**. If we want to use **`reana-ui`**, we have to open a new browser tab, separated from the JupyterLab interface.

- As we did for Rucio, we aim to integrate REANA and its functionalities into the JupyterLab environment.



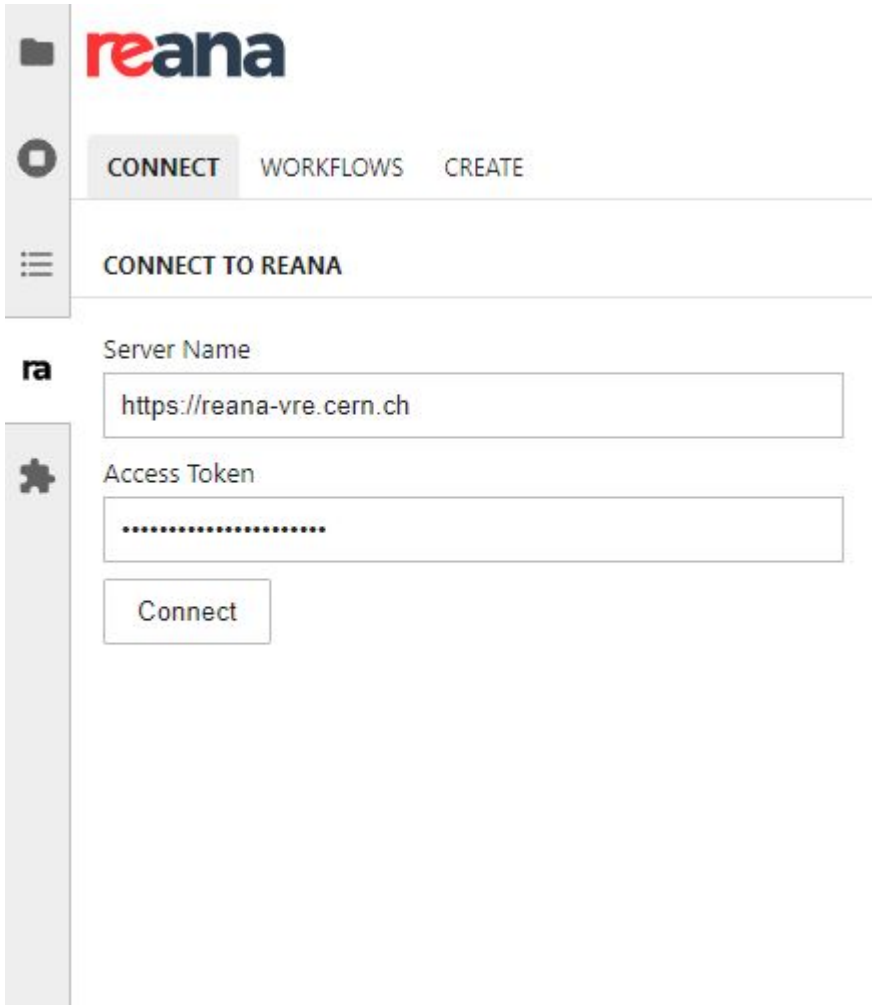


Reana JupyterLab extension

- **Main features**
  - Connect to any Reana server
  - List workflows
  - Interact with a workflow
  - Create workflow from the environment
- **UI Requirements**
  - Responsive
  - User-friendly

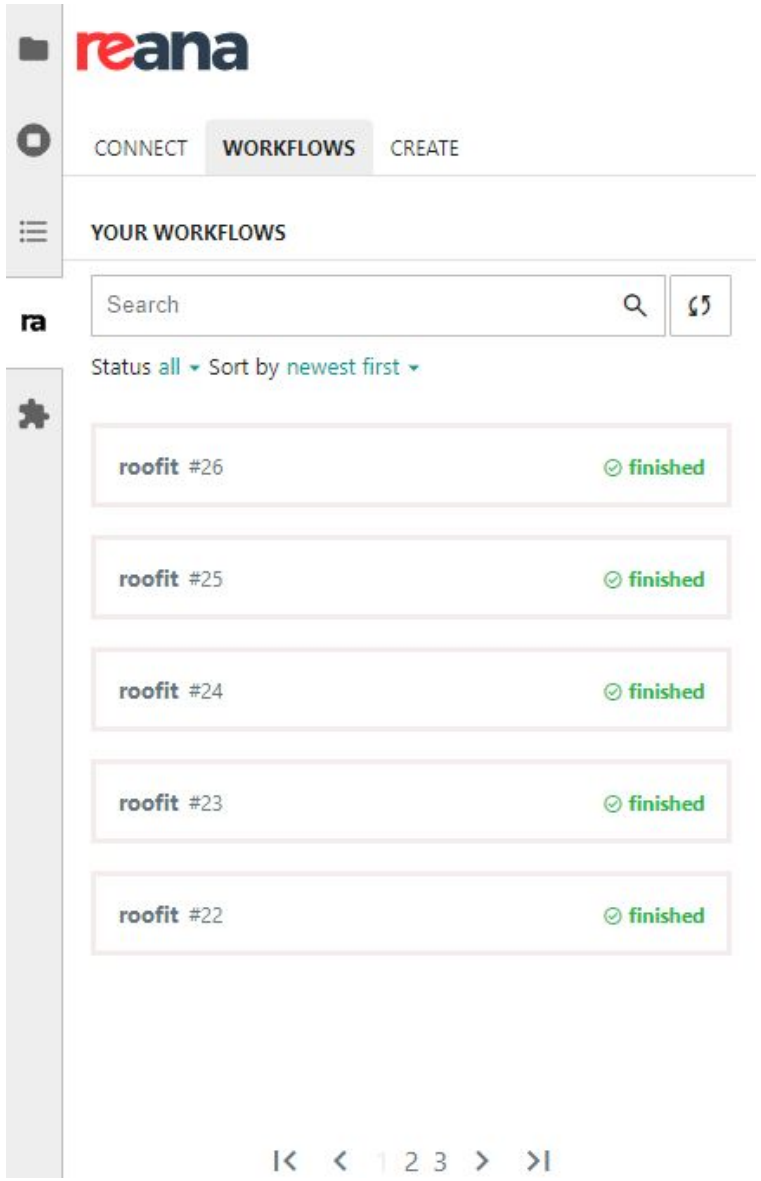
Reana JupyterLab extension

- **Main features**
  - Connect to any Reana server ✓
  - List workflows
  - Interact with a workflow
  - Create workflow from the environment
- **UI Requirements**
  - Responsive
  - User-friendly



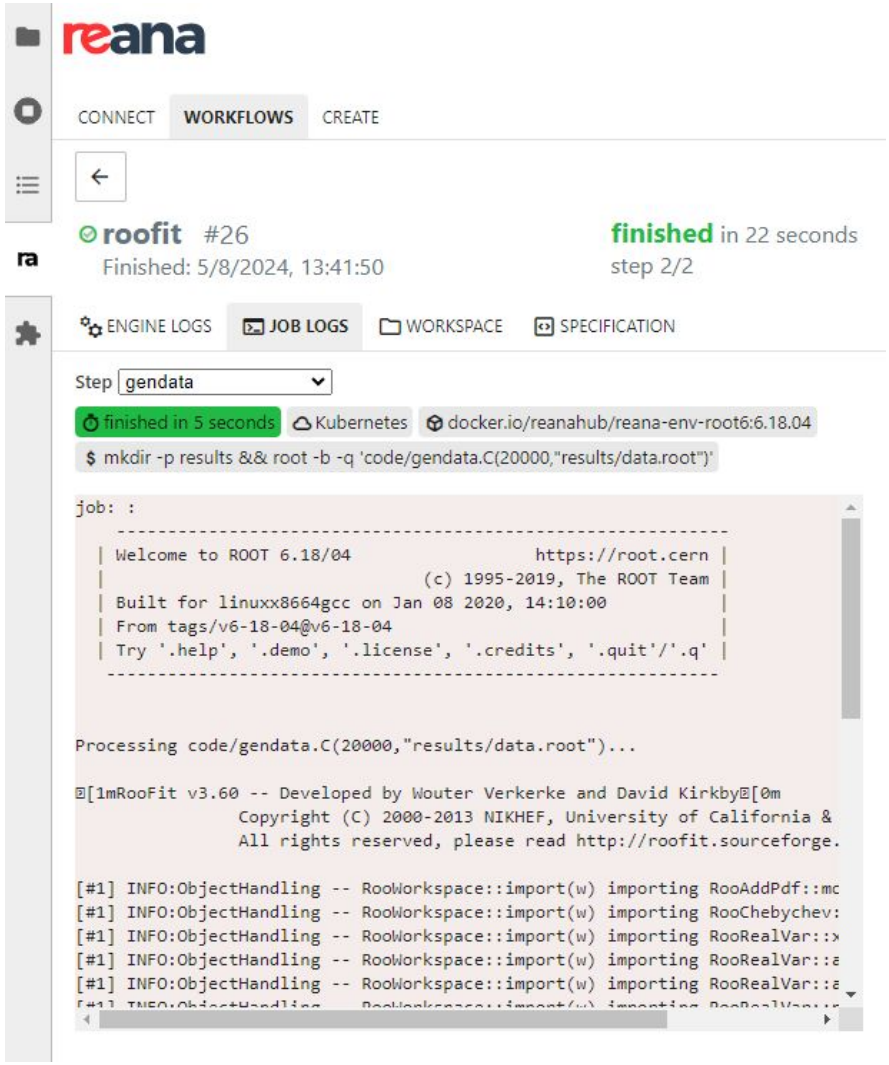
Reana JupyterLab extension

- **Main features**
  - Connect to any Reana server ✓
  - List workflows ✓
  - Interact with a workflow
  - Create workflow from the environment
- **UI Requirements**
  - Responsive
  - User-friendly



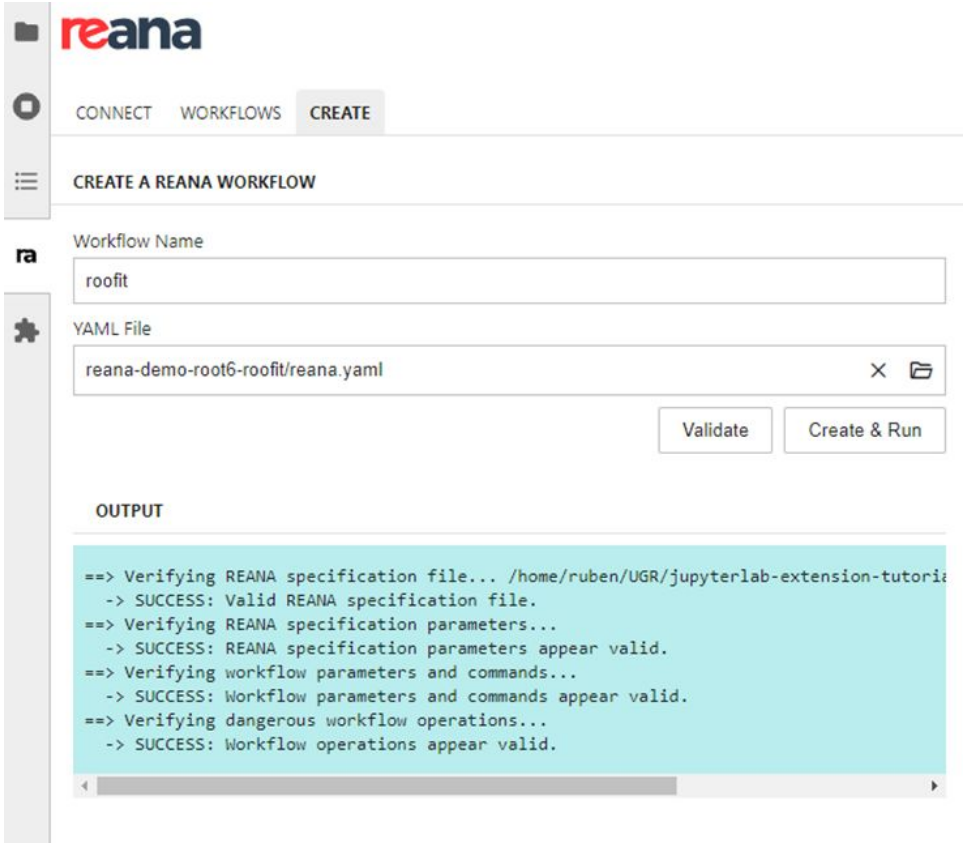
Reana JupyterLab extension

- **Main features**
  - Connect to any Reana server ✓
  - List workflows ✓
  - Interact with a workflow ✓
  - Create workflow from the environment
- **UI Requirements**
  - Responsive
  - User-friendly



Reana JupyterLab extension

- **Main features**
  - Connect to any Reana server ✓
  - List workflows ✓
  - Interact with a workflow ✓
  - Create workflow from the environment ✓
- **UI Requirements**
  - Responsive
  - User-friendly



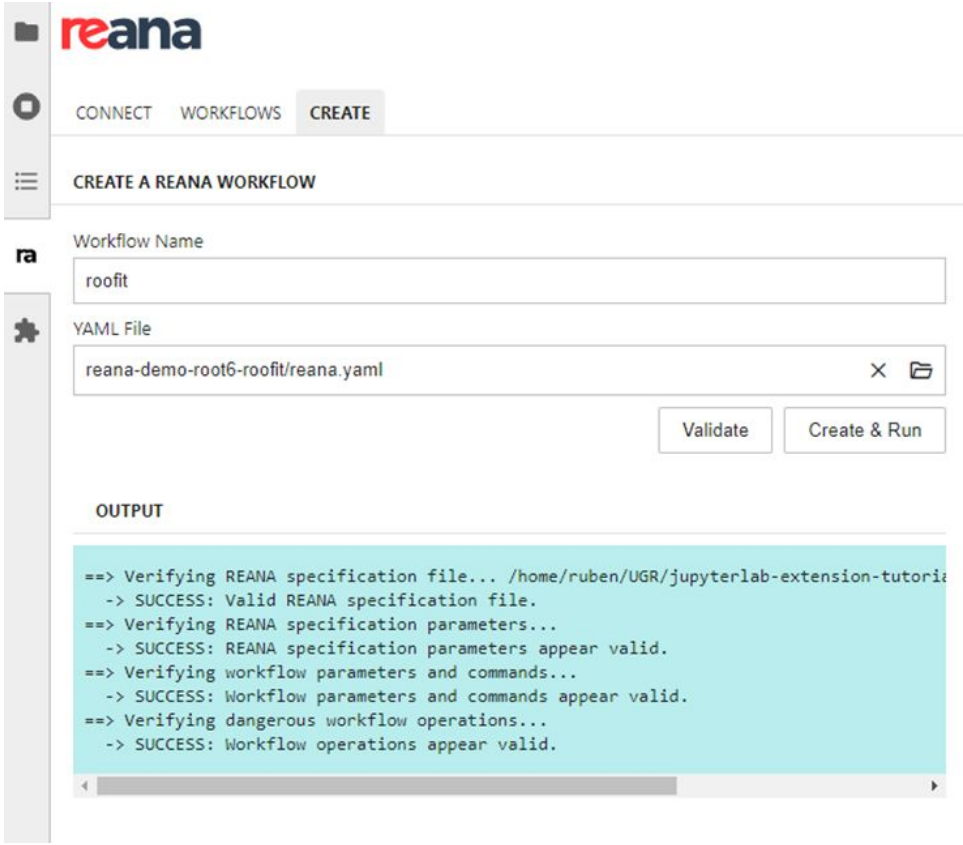
Reana JupyterLab extension

Main features

- Connect to any Reana server ✓
- List workflows ✓
- Interact with a workflow ✓
- Create workflow from the environment ✓

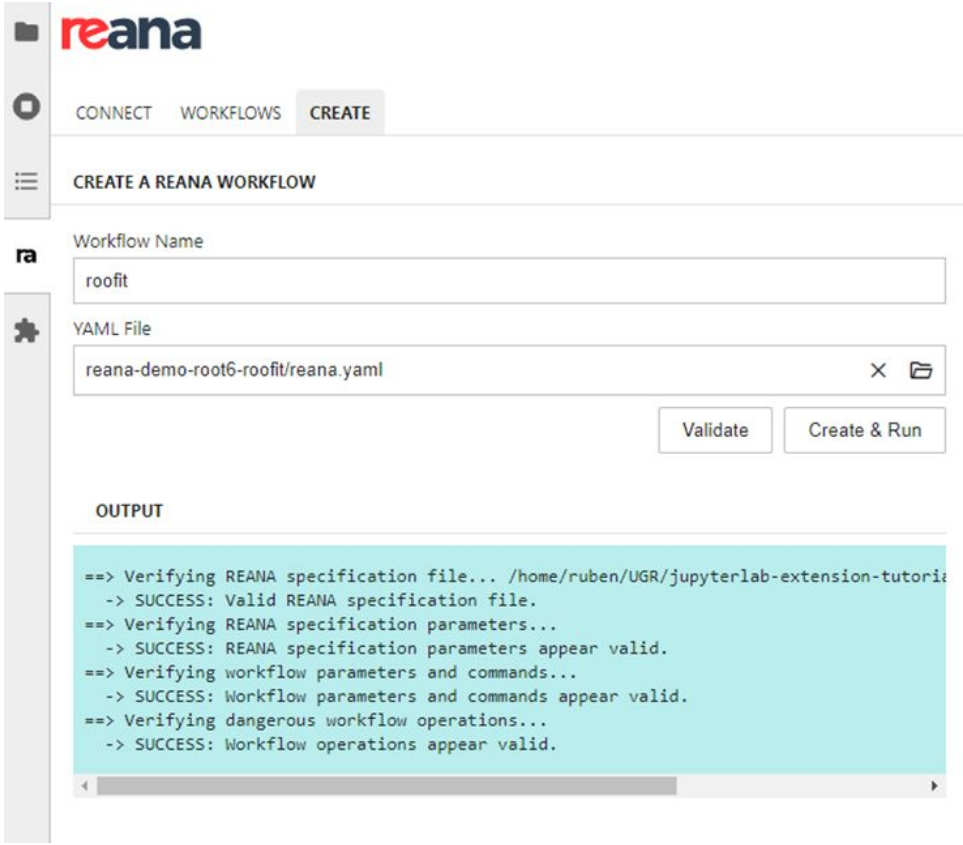
UI Requirements

- Responsive ✓
- User-friendly



Reana JupyterLab extension

- **Main features**
  - Connect to any Reana server ✓
  - List workflows ✓
  - Interact with a workflow ✓
  - Create workflow from the environment ✓
- **UI Requirements**
  - Responsive ✓
  - User-friendly ✓



Reana JupyterLab extension

- **Frontend:**
  - Jupyter lab extension
  - Based on jupyterlab/extension-template
  - React, Typescript, CSS, HTML
- **Backend:**
  - Jupyter server extension
  - Python
- **Connection to REANA API:**
  - `reana-client` in Python
  - `reana-client` in CLI
  - Reana REST API

Docker image and Python package available!

```
docker pull ghcr.io/vre-hub/reana-jupyterlab-extension:latest
pip install reana-jupyterlab
```

Reana JupyterLab extension

- **Frontend:**
  - Jupyter lab extension
  - Based on jupyterlab/extension-template
  - React, Typescript, CSS, HTML
- **Backend:**
  - Jupyter server extension
  - Python
- **Connection to REANA API:**
  - ~~`reana-client` in Python~~
  - `reana-client` in CLI
  - Reana REST API

Docker image and Python package available!

```
docker pull ghcr.io/vre-hub/reana-jupyterlab-extension:latest
pip install reana-jupyterlab
```

Conclusions

- **`reana-jupyterlab-extension`** integrates all the basic functionalities of **`reana-client`** and **`reana-ui`** in the notebook interface.
- Having access to the extension in the VRE will make the researchers' work easier.

- **Future work:**
  - Use the Python methods provided by **`reana-client`** to make the requests to the server.

DEMO

Thank you!

Questions? Suggestions?

ruben.perez.mercado@cern.ch

rubenpermerc@gmail.com

linkedin.com/in/rubenperezmercado

github.com/vre-hub/reana-jupyterlab-extension

CERN openlab

Bibliography

1. Gazzarrini, E., Garcia, E., Gosein, D., Moya, A. V., Kounelis, A., & Espinal, X. (2023). The Virtual Research Environment: towards a comprehensive analysis platform. *arXiv preprint arXiv:2305.10166*.

2. Gazzarrini, E., Garcia, E. G., Gosein, D., & Espinal, X. (2024). The Virtual Research Environment: A multi-science analysis platform. In EPJ Web of Conferences (Vol. 295, p. 08023). EDP Sciences.

3. Šimko, T., Heinrich, L., Hirvonsalo, H., Kousidis, D., & Rodríguez, D. (2019). REANA: A system for reusable research data analyses. In EPJ Web of Conferences (Vol. 214, p. 06034). EDP Sciences.

BACKUP SLIDES

Interaction with REANA

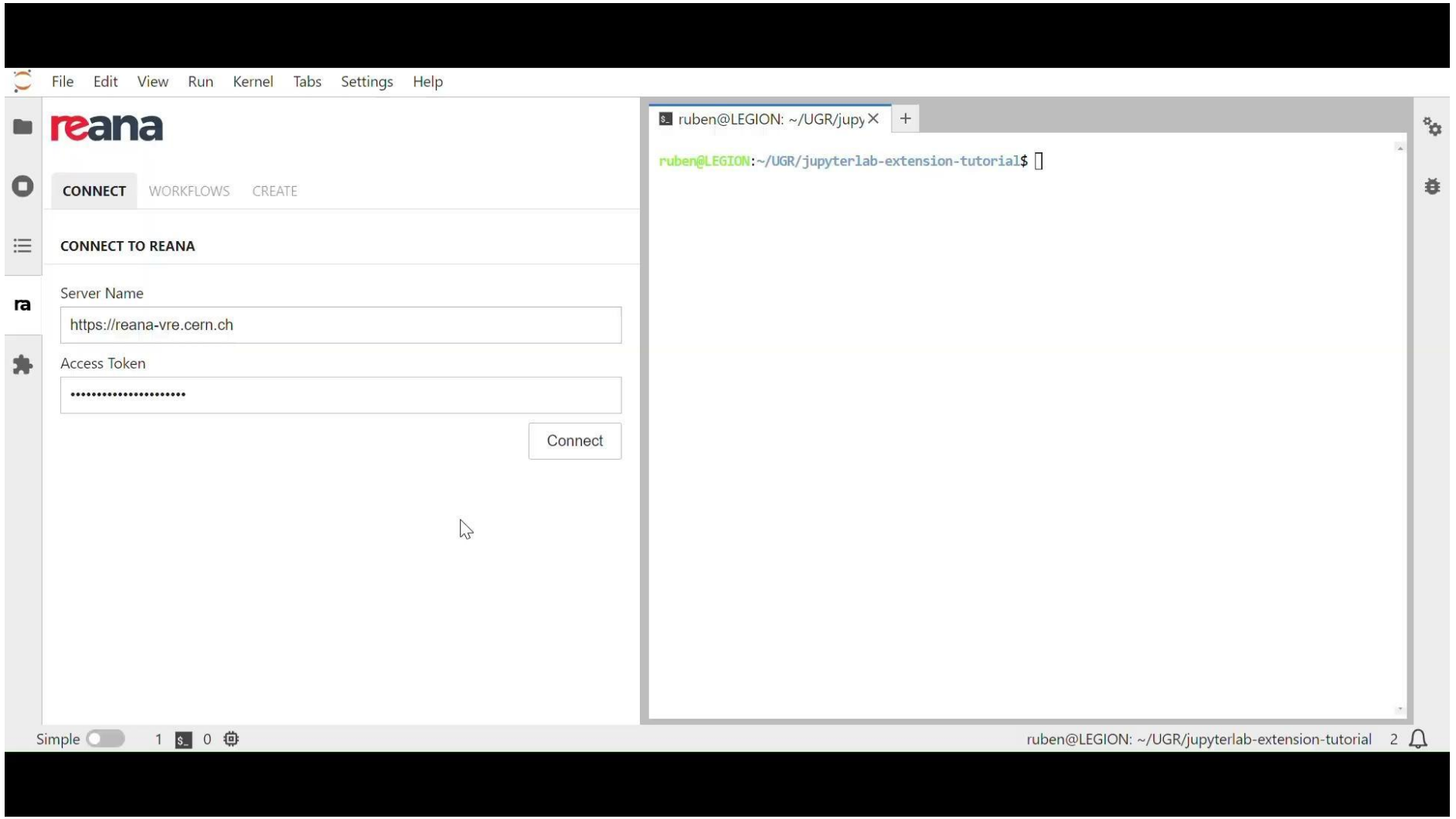
- We started using **`reana-client`** on this project, but we faced a few problems along the way.
  Some of them:
  - **`jsonschema`** incompatibility (https://github.com/reanahub/reana-commons/issues/461).
  - Server URL issue (https://github.com/reanahub/reana-client/issues/722).
  - No *sort* parameter in **`get_workflows`**.
  - No information about navigation in **`get_workflows`** (needed to paginate the results).
  - No **`run`** shortcut for creating the workflow, uploading the files and starting the workflow.

- We are using the **`reana-server REST API`** instead.

- Problem: Workflow creation is complex. We need to use the commands in **`reana-client`**.

```
reana-client run
reana-client validate
```

How Reana is being used in the VRE without the extension?

Elena Gazzarini. ATLAS Dark Matter Dilepton Resonance on the Virtual Research Environment.

How Reana will be used in the VRE with the extension?