REANA ↔ GitLab integration

Leticia Farias Wanderley

Supervisor: Diego Rodríguez Rodríguez

IT-CDA-DR
Reproducibility Crisis

HAVE YOU FAILED TO REPRODUCE AN EXPERIMENT?

Most scientists have experienced failure to reproduce results.

- Chemistry
- Biology
- Physics and engineering
- Medicine
- Earth and environment
- Other

https://www.nature.com/news/1-500-scientists-lift-the-lid-on-reproducibility-1.19970
High turnover of researchers
Pillars of a reproducible research analysis

Data
Where is your input data?

Code
Where is your code?

Environment
Which OS and libraries did you use?

Workflow
Which steps did you follow?
Reproducible research data analysis platform
Currently, manual interaction
Can we do better?
Designed two solutions depending on user needs
Physicists

#9z2n

#9z2n

reana
Physicists

#9z2n

×

#9z2n

×

#9z2n

reana
Physicists

#s6d7

reana

#s6d7
Physicists

#s6d7

Leticia Farias Wanderley
reana:
  image: "python:3"
  script:
    # install reana-client
    - pip install reana-client
    
    # create the workflow
    - reana-client create -n $CI_PROJECT_NAME
    
    # upload code and input files
    - reana-client upload ./data ./code
    
    # run the workflow
    - reana-client start --follow
    
    # download output files
    - reana-client download results/greetings.txt
  
  artifacts:
    paths:
      - results/greetings.txt
    expire_in: 1 day
    when: on_success
```yaml
reana:
  image: "python:3"
  script:
    - pip install reana-client
    - reana-client create -n $CI_PROJECT_NAME
  artifacts:
    paths:
      - results/greetings.txt
    expire_in: 1 day
    when: on_success
```

```
# install reana-client
- pip install reana-client

# create the workflow
- reana-client create -n $CI_PROJECT_NAME

# upload code and input files
- reana-client upload ./data ./code

# run the workflow
- reana-client start --follow

# download output files
- reana-client download results/greetings.txt

Designed for short execution time
Flexible manual configuration
```
Authorize **REANA** to use your account?

An application called **REANA** is requesting access to your GitLab account. This application was created by team@reana.io. Please note that this application is not provided by GitLab and you should verify its authenticity before allowing access.

This application will be able to:

- **Access the authenticated user's API**
  Grants complete read/write access to the API, including all groups and projects.

[Deny] [Authorize]
Authorize REANA to use your account?

An application called REANA is requesting access to your GitLab account. This application was created by team@reana.io. Please note that this application is not provided by GitLab and you should verify its authenticity before allowing access.

This application will be able to:

- Access the authenticated user’s API
  Grants complete read/write access to the API, including all groups and projects.

Zero configuration

Unlimited execution time
reana-demo-helloworld | Pipeline #75598780 has succeeded for master | 352a646f

---

<table>
<thead>
<tr>
<th>Project</th>
<th>Leticia Wanderley / reana-demo-helloworld</th>
</tr>
</thead>
<tbody>
<tr>
<td>Branch</td>
<td>master</td>
</tr>
<tr>
<td>Commit</td>
<td>352a646f</td>
</tr>
<tr>
<td>Commit Author</td>
<td>Leticia Wanderley</td>
</tr>
</tbody>
</table>

Pipeline #75598780 triggered by Leticia Wanderley successfully completed 4 jobs in 1 stage.
+ reana = ❤
Impact

Implemented and integrated ✅

1st CERN GitLab application ✅

"The integration of REANA with GitLab is a great step towards running reproducible CMS analysis workflows. This will also significantly facilitate our analysis preservation efforts."

Clemens Lange (CMS)

Future work

User configurability
Thank you!

reana.io

leticiafwanderley@gmail.com

@lfwanderley
@leticiawanderley

Special thanks to Alex, Daniel, and the whole GitLab team @ CERN!