

IBERGRID

2024 28-30 OCT
UNIVERSITY
OF PORTO

better
software
for
better
science

13TH IBERIAN GRID CONFERENCE



SQAaaS platform as the quality gate for Digital Twins

Pablo Orviz

orviz@ifca.unican.es
IFCA-CSIC



LABORATÓRIO DE INSTRUMENTAÇÃO
E FÍSICA EXPERIMENTAL DE PARTÍCULAS



Infraestructura
Nacional de
Computación
Distribuida



CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS



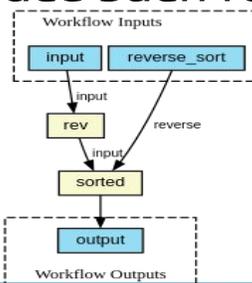
SQAaaS: the QA building block for Digital Twins

📌 SQAaaS features:

- Programmatic QA assessment (RESTful API) and awarding of digital research content
 - SQA characteristics on code and (web) services (custom)
 - FAIR principles on data (RDA maturity model)
- Adhoc composition of CI/CD pipelines
..besides providing clients (web portal, CLI) to streamline usage

📌 Digital Twins orchestrate and operate **computational end-to-end data analysis workflows**

- Workflow steps consume/execute/produce *such research content*:
 - Input and output data
 - Software/code and/or services



POST	/pipeline/assessment	Creates a pipeline for assessment (QAA module).
GET	/pipeline/assessment/{pipeline_id}/output	Get the assessment output (QAA module)
DELETE	/pipeline/{pipeline_id}	Delete pipeline by ID
GET	/pipeline/{pipeline_id}	Find pipeline by ID
PUT	/pipeline/{pipeline_id}	Update pipeline by ID
GET	/pipeline/{pipeline_id}/badge	Gets badge data associated with the given pipeline

Bronze badge	Silver badge	Gold badge
Code Accessibility	Versioning	Security
Licensing	Code metadata	Code Style
Documentation		



[Verify](#) [Go to Badgr's award page](#)

Share your badge in popular code and data repository platforms using Markdown

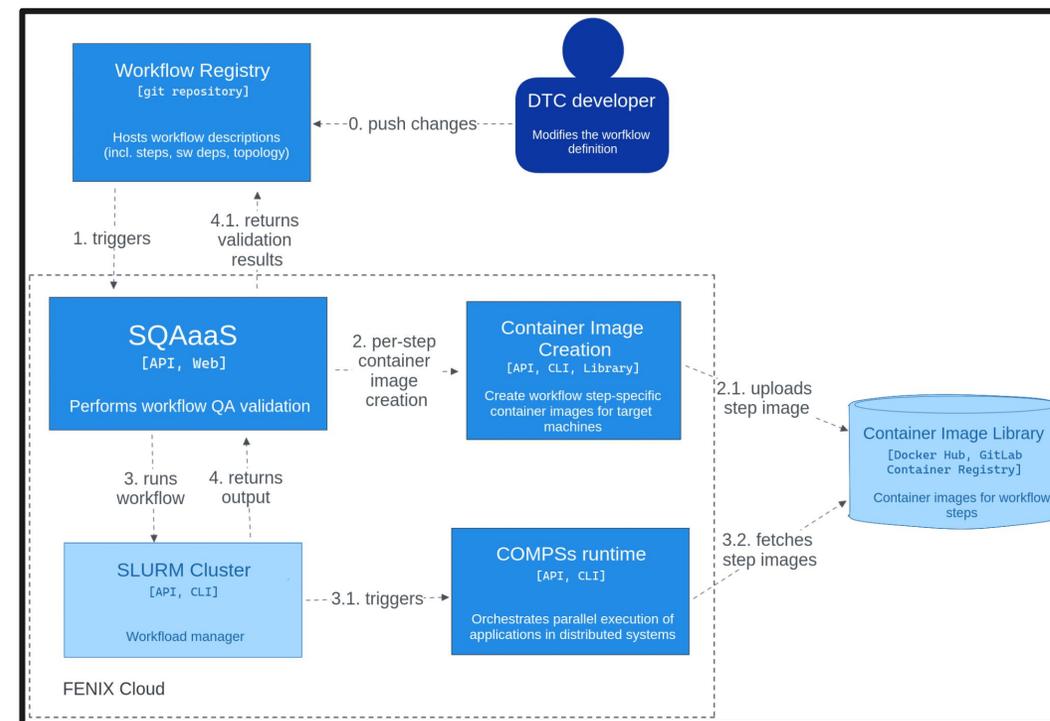
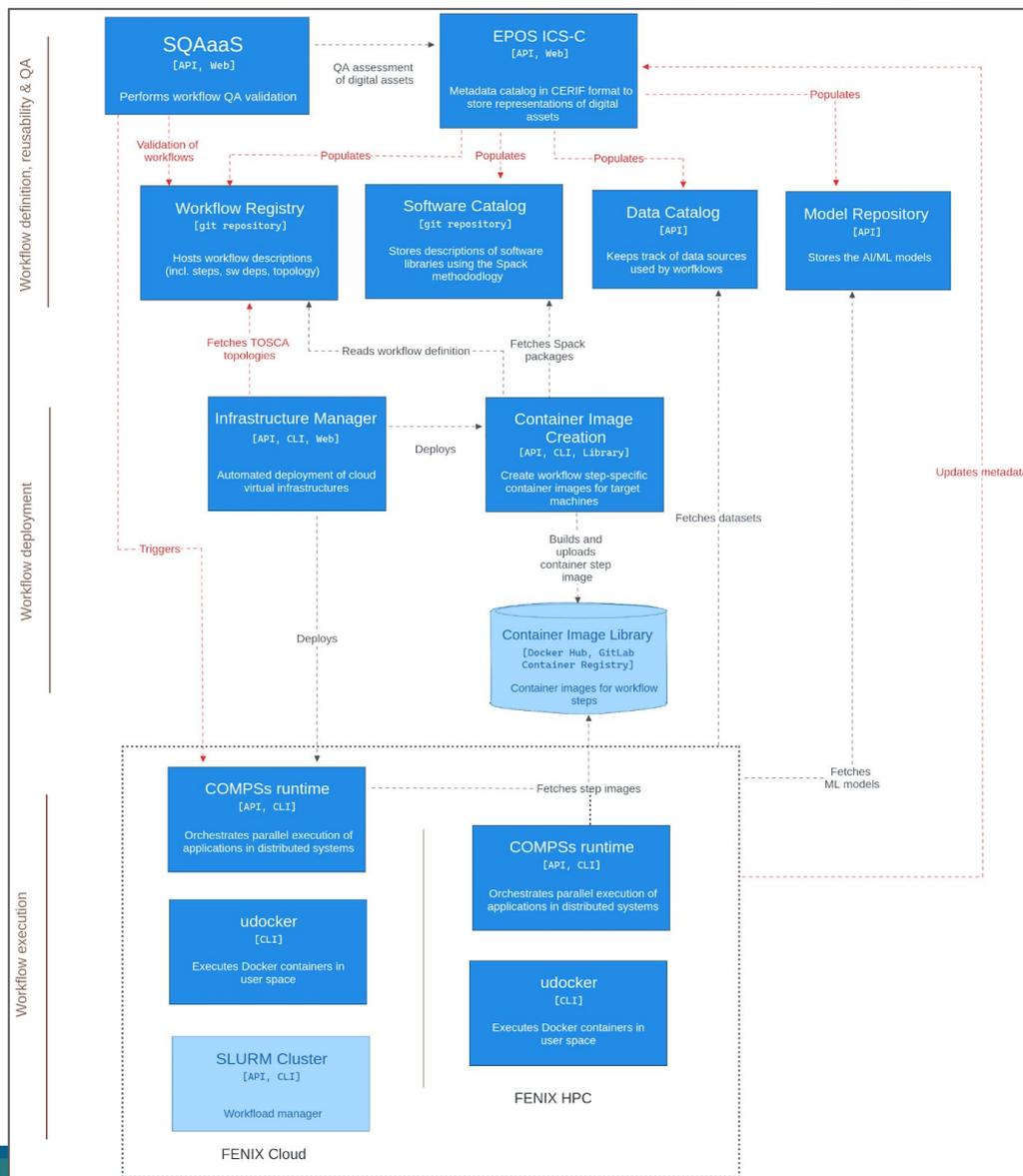
[Get Badge Image](#) [Get Badge Shield](#)

SQAaaS role in Digital Twin projects (EC-funded)

- SQAaaS fits in as the **QA building block during the (i) development and (ii) operation phases of a Digital Twin**
 - QA+FAIR assessment of individual digital assets (development life cycle)
 - Integration with Workflow Management Systems (operation)
 - Integration with Scientific Data Management technologies (operation)
- SQAaaS is already an architectural component within projects developing Digital Twin solutions:
 - InterTwin: developing an engine (DTE) for building, validating and deploying multidisciplinary digital twins
 - DT-GEO: building a multi-component prototype digital twin on geophysical extremes



Role of SQAaaS in Digital Twin projects: DT-GEO



Integrations & features: Code hosting platforms

- **Suitable for:** QA work done at development time, in the platform where the code is maintained
- **Target/s:** source code (incl. workflow code)
- **Automation:** reacts to events triggered by the hosting platform (push, change request, releases, etc.)
- **Platform/s:** GitHub and GitLab

GitHub



Features:

- #1. Leverages already existent CI technologies: **GitHub Actions & GitLab CI**
- #2. **Reporting summary** for each build
- #3. **Dynamic badging**, updated as SQAaaS assessment progresses

Integrations & features: Code hosting platforms

Feature #1: Leverages already existent CI technologies: **GitHub Actions & GitLab CI**

[Marketplace](#) / [Actions](#) / SQAaaS source code assessment

GitHub Action
SQAaaS source code assessment
 2.4.1 Latest version
 SQAaaS | source code | building | License: GPLV3 | REUSE compliant | code style: black

Baseline evaluation

```
uses: eosc-synergy/sqaaas-assessment-action@v2
```

GitHub

```
uses: eosc-synergy/sqaaas-step-action@v1
with:
  name: model_validation_step
  tool: pytest
  test-path: ./tests
```

A 2nd action ([sqaaas-step-action](#)) defines a series of steps to be executed as part of the assessment in the SQAaaS platform

Extended evaluation

```
uses: eosc-synergy/sqaaas-assessment-action@v2
with:
  qc_uni_steps: model_validation_step
```

SQAaaS tutorial at Ibergrid 2024 conference @
Mon 28

Integrations & features: Code hosting platforms

Feature #1: Leverages already existent CI technologies: **GitHub Actions & GitLab CI**

Baseline evaluation



SQAaaS CI Component 
Collection of pipeline templates

Following the approach of GitHub, a 2nd CI component has been implemented to define an additional step in the SQAaaS assessment

SQAaaS tutorial at Ibergrid 2024 conference @
Mon 28

```
include:
  - component: gitlab.com/dtgeo/sqa-assessment-ci-template/SQA.gitlab-ci@1.0.0
```

Extended evaluation

```
variables:
  INPUT_NAME: pycompss-step
  INPUT_TOOL: commands
  INPUT_COMMANDS: echo 'Here we will execute PyCOMPSS'
  INPUT_QC_UNI_STEPS: pycompss-step

stages:
  - SQA

include:
  - component: $CI_SERVER_FQDN/dtgeo/metadata/sqa-assessment-ci-template/SQA-step.gitlab-ci@1.0.0
  - component: gitlab.com/dtgeo/sqa-assessment-ci-template/SQA.gitlab-ci@1.0.0
```

Integrations & features: Code hosting platforms

Feature #2: Reporting summary for each build

Summary

Jobs

- Job that triggers SQAaaS platfo...

Run details

- Usage
- Workflow file

Job that triggers SQAaaS platform summary

SQAaaS summary

Result	Assertion	Subcriterion ID	Criterion ID
✓	Source code uses Git for version control	QC.Acc01	QC.Acc
✓	A README file is present in the code repository	QC.Doc06.1	QC.Doc
✓	A CODE_OF_CONDUCT file is present in the code repository	QC.Doc06.3	QC.Doc
✓	A CONTRIBUTING file is present in the code repository	QC.Doc06.2	QC.Doc
✓	Documentation resides in the same repository as code	QC.Doc01.1	QC.Doc
●	Docs are not fully compliant with markdownlint standard	QC.Doc02.X	QC.Doc
✓	An Open Source license found in the code repository: GPL-3.0	QC.Lic01	QC.Lic
✓	LICENSE file is visible at the root path of the code repository: LICENSE	QC.Lic01.1	QC.Lic
✓	License GPL-3.0 is approved by the Open Source Initiative	QC.Lic02	QC.Lic
✓	License GPL-3.0 is listed under the Open Source Initiative popular category	QC.Lic02.1	QC.Lic
✓	JSON files are compliant with jsonlint standard	QC.Sty01	QC.Sty
✓	The code repository uses tags for releasing new software versions	QC.Ver01.0	QC.Ver
✓	Latest release tag 2.9.1 is SemVer compliant	QC.Ver01	QC.Ver
✓	All release tags are SemVer compliant	QC.Ver02	QC.Ver
□	No matching files found for language <i>CodeMeta</i> in repository searching by extensions or filenames No matching files found for language <i>Citation File Format</i> in repository searching by extensions or filenames	QC.Met01	QC.Met
●	No matching files found for language <i>Python</i> in repository searching by extensions or filenames No matching files found for language <i>Go</i> in repository searching by extensions or filenames	QC.Sec02	QC.Sec

View full report at https://raw.githubusercontent.com/eosc-synergy/sqaas-api-spec/assess.sqaas/main/report/assessment_output.json

Job summary generated at run-time

- Displayed in GitHub Actions' job view
- Provides link to full report in SQAaaS platform

Integrations & features: Code hosting platforms

Feature #3: **Dynamic badging** (powered by shields.io), updated as the SQAaaS assessment progresses

Intermediate status



Final status

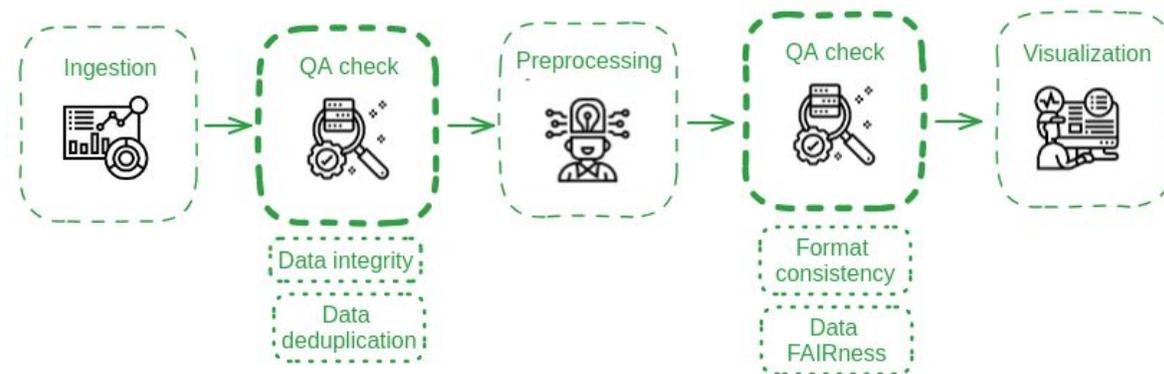


More on SQAaaS badging
docs.sqaaaS.eosc-synergy.eu

Further Integrations & features (in the making)

- **Workflow Management Systems (WfMSs)**

- **Suitable for:** individual QA work during the execution of a workflow
- **Target/s:** components of a workflow
- **Automation:** not through events, coded in the workflow specification
- **Tool/s:** PyOphidia & CWL (InterTwin)



- **Data Management Systems**

- **Suitable for:**
 - QA work on data (FAIR, quality dimensions)
 - React to CRUD operations on data repositories/lakes
- **Target/s:** data
- **Automation:** notifications sent by the data management system on CRUD operations
- **Tool/s (potential):** Rucio (InterTwin)

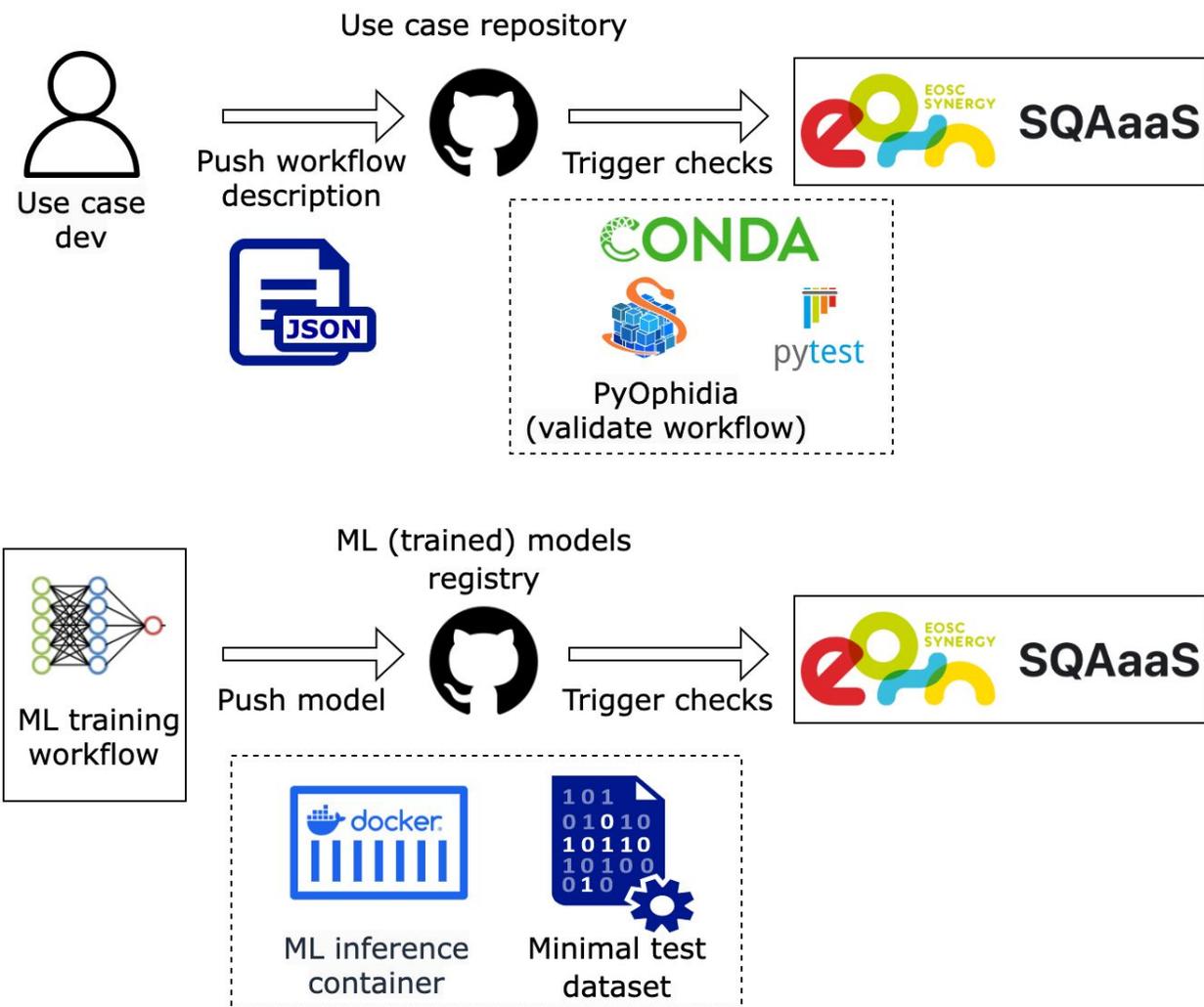
Use cases: InterTwin's Climate Change (PyOphidia+SQaaS)

↘ Main integration leveraging SQAaaS GitHub actions:

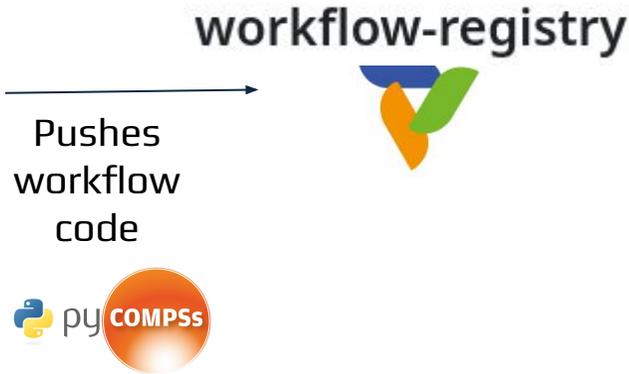
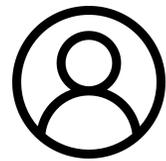
1. **Validation of (PyOphidia) workflow description** document
2. **Test trained models statically** with small test set

↘ Further potential integrations (ad hoc support into Workflow and Data Management Systems): **Stronger scientific validation of the model** with statistically relevant (larger) test set

- **Data quality and ML model FAIRness** (to be better explored)



Use cases: DTC development in DT-GEO



#1 Builds Docker images



#2 Triggers SQAaaS assessment



#3 Triggers workflow execution



- Offloads work to GitLab CI
- SQAaaS GitLab CI component performs the workflow validation (PyCOMPSSs execution) in an HPC-like ecosystem in the Cloud

Potential integrations with:

- PyCOMPSSs WfMS: define QA work within workflow steps (e.g. data FAIRness)
- Software and data registries: react to changes

Further uses of SQAaaS platform: QA tracker



Target	Platform services, DTE core, DTE thematic modules, DT applications
Frequency	Upon product release
Means	Manual
Reporting	Tracked through JIRA issues (release-based)
Managed by	Software Release Management team (WP3, Mario & Joao from LIP bearing the burden)

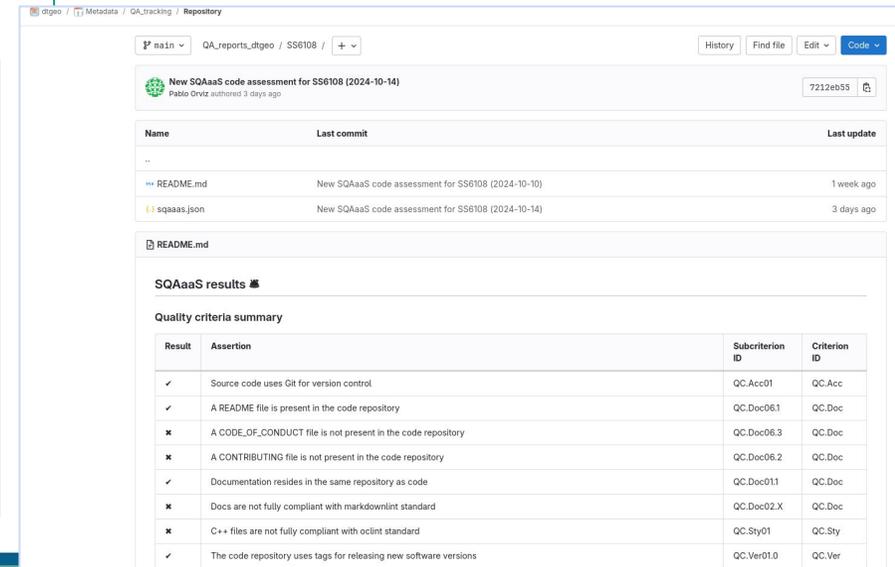


Target	DT applications
Frequency	Periodically / cron-based
Means	Automated / Pipeline schedules / SQAaaS GitLab CI component
Reporting	Git-powered, summaries in READMEs
Managed by	Digital asset quality tracking (WP4's T4.2)

Software Release Management Kanban
<https://jira.egi.eu/secure/RapidBoard.jspa?rapidView=81#>

Software developed or modified within the project

SW	Developers (Name and Contact)	In Jira?	So																						
INDIGO PaaS Orchestrator Dashboard	@ Marica Antonacci (main)	<table border="1"> <thead> <tr> <th>T</th> <th>Key</th> <th>Summary</th> <th>Assignee</th> <th>Reporter</th> <th>P</th> <th>Status</th> <th>Resolution</th> <th>Created</th> <th>Updated</th> <th>Due</th> </tr> </thead> <tbody> <tr> <td>+</td> <td>ITRM-6</td> <td>INDIGO PaaS Orchestrator Dashboard 3.2.1</td> <td>Marica Antonacci</td> <td>João António Tomásio Pina</td> <td></td> <td>RELEASE VALIDATED</td> <td>Unresolved</td> <td>2024 Jan 08</td> <td>2024 Oct 15</td> <td></td> </tr> </tbody> </table> <p>1 issue Refresh</p>	T	Key	Summary	Assignee	Reporter	P	Status	Resolution	Created	Updated	Due	+	ITRM-6	INDIGO PaaS Orchestrator Dashboard 3.2.1	Marica Antonacci	João António Tomásio Pina		RELEASE VALIDATED	Unresolved	2024 Jan 08	2024 Oct 15		https://github.com
T	Key	Summary	Assignee	Reporter	P	Status	Resolution	Created	Updated	Due															
+	ITRM-6	INDIGO PaaS Orchestrator Dashboard 3.2.1	Marica Antonacci	João António Tomásio Pina		RELEASE VALIDATED	Unresolved	2024 Jan 08	2024 Oct 15																
itwnaai	@ Matteo Bunino @ Alexander Zoebbauer @ Rakesh Sama	<table border="1"> <thead> <tr> <th>Key</th> <th>Summary</th> <th>T</th> <th>Created</th> <th>Updated</th> <th>Due</th> <th>Assignee</th> <th>Reporter</th> <th>P</th> <th>Status</th> <th>Resolution</th> </tr> </thead> <tbody> <tr> <td>ITRM-22</td> <td>Itwinai 0.2.1</td> <td>+</td> <td>2024 Feb 01</td> <td>2024 Jun 12</td> <td></td> <td>Matteo Bunino</td> <td>João António Tomásio Pina</td> <td></td> <td>RELEASE VALIDATED</td> <td>Unresolved</td> </tr> </tbody> </table> <p>1 issue Refresh</p>	Key	Summary	T	Created	Updated	Due	Assignee	Reporter	P	Status	Resolution	ITRM-22	Itwinai 0.2.1	+	2024 Feb 01	2024 Jun 12		Matteo Bunino	João António Tomásio Pina		RELEASE VALIDATED	Unresolved	https://github.com
Key	Summary	T	Created	Updated	Due	Assignee	Reporter	P	Status	Resolution															
ITRM-22	Itwinai 0.2.1	+	2024 Feb 01	2024 Jun 12		Matteo Bunino	João António Tomásio Pina		RELEASE VALIDATED	Unresolved															



dtgeo / Metadata / QA_tracking / Repository

QA_reports_dtgeo / SS6108 / +

New SQAaaS code assessment for SS6108 (2024-10-14)
Pablo Orvez authored 3 days ago 7212eb55

Name	Last commit	Last update
...		
README.md	New SQAaaS code assessment for SS6108 (2024-10-10)	1 week ago
sqaaaS.json	New SQAaaS code assessment for SS6108 (2024-10-14)	3 days ago

README.md

SQAaaS results

Quality criteria summary

Result	Assertion	Subriterion ID	Criterion ID
✓	Source code uses Git for version control	QC.Acc01	QC.Acc
✓	A README file is present in the code repository	QC.Doc06.1	QC.Doc
✗	A CODE_OF_CONDUCT file is not present in the code repository	QC.Doc06.3	QC.Doc
✗	A CONTRIBUTING file is not present in the code repository	QC.Doc06.2	QC.Doc
✓	Documentation resides in the same repository as code	QC.Doc01.1	QC.Doc
✗	Docs are not fully compliant with markdownlint standard	QC.Doc02.X	QC.Doc
✗	C++ files are not fully compliant with oclint standard	QC.Sty01	QC.Sty
✓	The code repository uses tags for releasing new software versions	QC.Ver01.0	QC.Ver

IBERGRID

2024

28-30 OCT
UNIVERSITY
OF PORTO

better
software
for
better
science

13TH IBERIAN GRID CONFERENCE



Q&A

Thanks for your attention

orviz@ifca.unican.es
IFCA-CSIC



LABORATÓRIO DE INSTRUMENTAÇÃO
E FÍSICA EXPERIMENTAL DE PARTÍCULAS



Infraestructura
Nacional de
Computación
Distribuida



CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

