



ENVRI
FAIR



EOSC ENVRI DevOps framework -- Progress report

Zhiming Zhao

University of Amsterdam

ENVRI-FAIR WP7 leader



ENVRI-FAIR has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824068

ENVRI: a cluster of research infrastructures in environmental and earth sciences

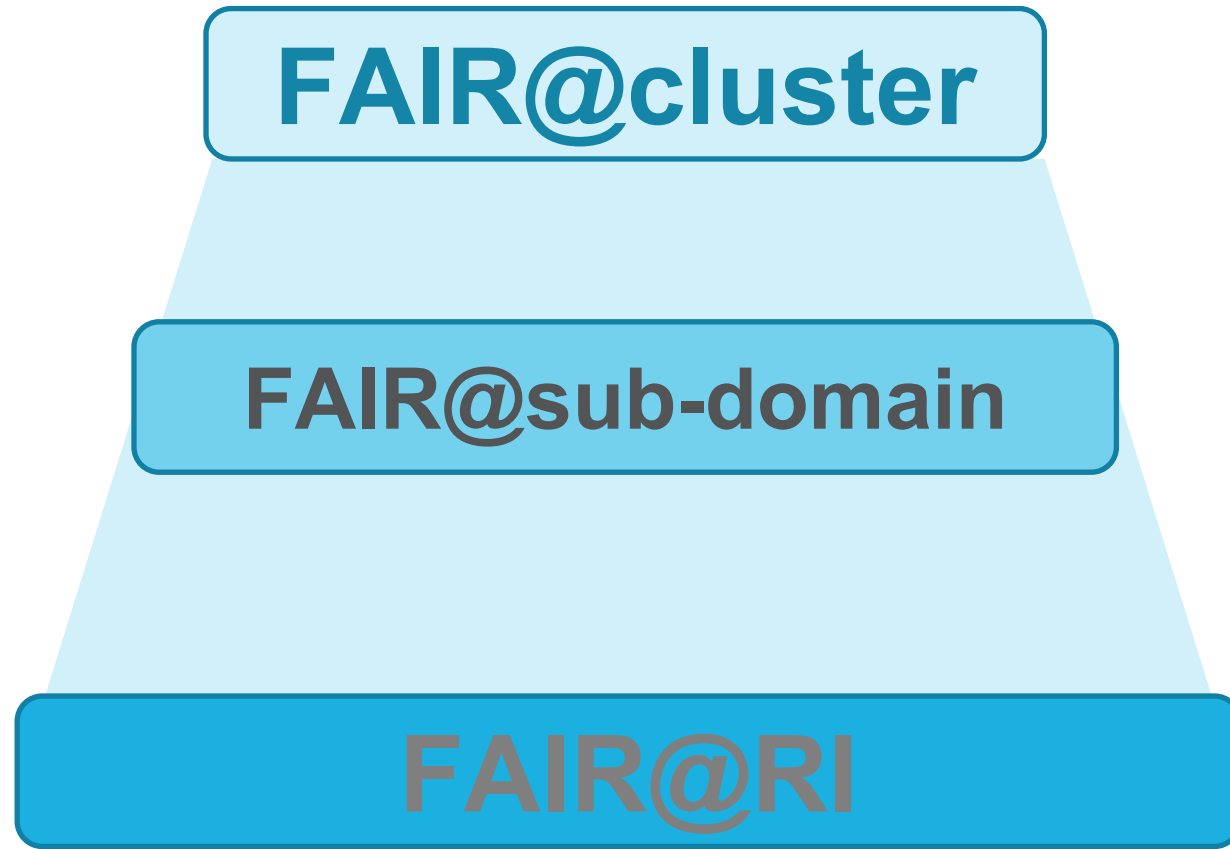


www.envri.eu





ENVRI-FAIR development activities





Time line

- Use case 1: workflow automation
- Use case 2: Jupyter notebook for data intensive science
- Use case 3: DOSC DevOps framework

<p>Phase 1 Describe the key use case: scenario, scope, KPI, steps etc. Get familiar with the EOSC services, following training and practices from the other projects etc. Get the requested resource provisioned Setup OLA with resource providers and agreement with Jelastic.</p>	<p>Phase 2 DevOps pipeline configured, including Git, automated testing, integration and deployment demonstrate in at least via two service development</p>	<p>Phase 3 Demonstrate the initial version of the workflow from ENVRI-FAIR, with automated workflow execution in Cloud; Demonstrate the other common data services identified in the ENVRI communities (optional)</p>	<p>Phase 4 Exploit the results to the development activities in ENVRI-FAIR sub domain Sustain the development by finding new opportunities, e.g., new EOSC projects etc.</p>
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------



Agreements and confluence pages

1. EGI OLAs agreed with the resource providers (CESGA and INFN-CATANIA)
2. VO vo.envri-fair.eu setup and SLA agreed
 - a. <https://operations-portal.egi.eu/vo/view/voname/vo.envri-fair.eu>
 - b. <https://documents.egi.eu/public/ShowDocument?docid=3595>
3. Signed agreement with Jelastec for the installation of the platform
4. EAP confluence page updated:
 - a. <https://confluence.egi.eu/display/EOSC/EOSC+DevOps+framework+and+virtual+in+frastructure+for+ENVRI-FAIR+common+FAIR+data+services>
5. CRM DB entry added:
 - a. <https://confluence.egi.eu/display/EOSC/ENVRI-FAIR>



Internal management

- WP7 members from ENVRI-FAIR project
- Weekly call of WP7
- Current status:
 - Develop a demonstrator for IaaS
 - Provided a cloud introduction training
 - Integrating the Jupyter development
 - Collecting workflow and test cases
 - Learning Jelastic

The screenshot shows a Trello board for the 'ENVRI-FAIR-EOSC-EAP' project. The board is organized into three columns: 'Todoist', 'doing', and 'done'. The 'doing' column contains five cards with the following titles: 'Provisioning resources', 'Create a guideline for each case', 'Experiments of the workflow automation', 'Collect potential experiments for the EOSC EAP testbed', and 'Study the DevOps (Jelastic) tool'. Each card has a status indicator (eye icon) and a user profile picture. The 'done' column contains two cards: 'Add document to the EOSC wiki' and 'Create the ENVRI-FAIR EOSC EAP team in Trello'. The board also features a 'Burndown Chart' and an 'Invite' button.

How to access EOSC services

This document provides instructions to access the resources in the [ENVRI-FAIR EOSC early adopter program \(EAP\) project](#).

If you have questions related to this document, please contact Zhiming Zhao (z.zhao@uva.nl), or Andrea Manzi (andrea.manzi@egi.eu).

1. Background

[ENVRI-FAIR EOSC early adopter program \(EAP\) project](#) [1] is a 12 months mini-project supported by EU EOSC-HUB project (April 2020-Dec 2020, with possibly three-month extension in 2021).

The objective of the ENVRI-FAIR EOSC EAP is to

- Gain practices for using the current EOSC services, for infrastructure (IaaS), for integrating/deploying services (Jelastic), and for scientific experiments (e.g., Jupyter notebook);
- Use the IaaS (provided by EOSC) as a Common testbed for testing and integrating data management services developed by ENVRI RIs/sub-domains;
- Preparing for the integration between ENVRI results and emerging EOSC ecosystem.

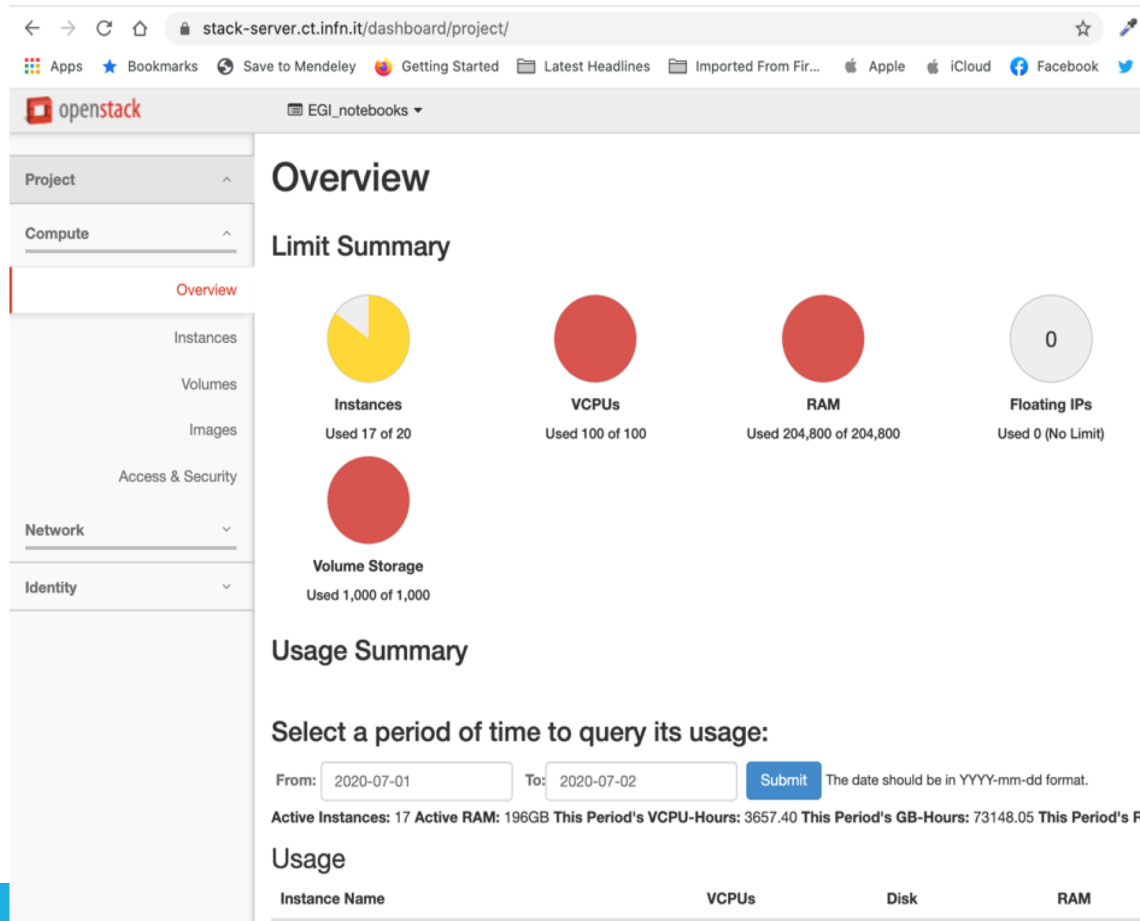
The EAP is structured into three scenarios:

- Automated virtual infrastructure provisioning and software deployment for scientific workflows
- Continuously software testing, integration and deployment for data management services

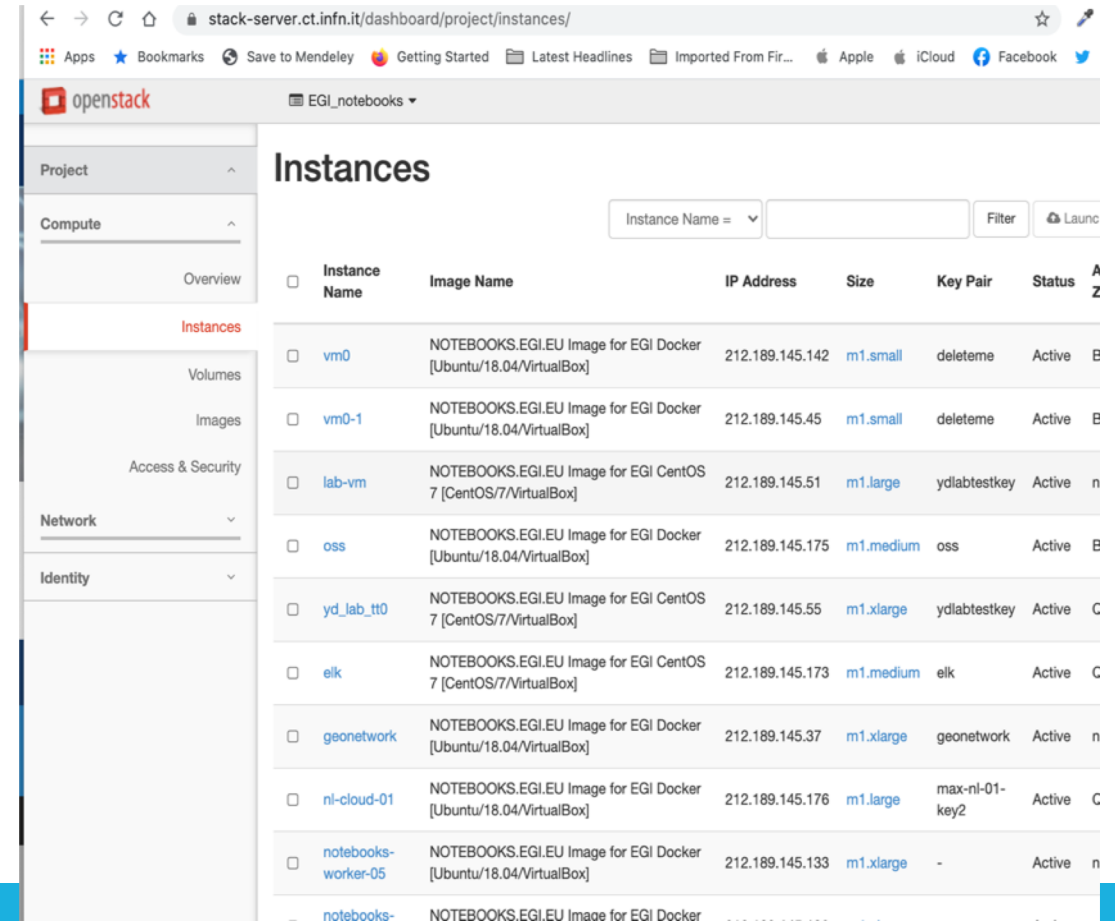


Use case 1: automated workflow using IaaS

1. Get familiar with the OpenStack dashboard



The screenshot shows the OpenStack dashboard for the 'EGI_notebooks' project. The 'Overview' section displays a 'Limit Summary' with four circular gauges: Instances (Used 17 of 20), VCPUs (Used 100 of 100), RAM (Used 204,800 of 204,800), and Floating IPs (Used 0 of No Limit). Below this is a 'Usage Summary' section with a date range selector (From: 2020-07-01, To: 2020-07-02) and a 'Submit' button. The usage summary text reads: 'Active Instances: 17 Active RAM: 196GB This Period's VCPU-Hours: 3657.40 This Period's GB-Hours: 73148.05 This Period's RA'. A table titled 'Usage' is partially visible at the bottom, with columns for Instance Name, VCPUs, Disk, and RAM.



The screenshot shows the OpenStack dashboard for the 'EGI_notebooks' project, specifically the 'Instances' page. It features a table listing various instances with columns for Instance Name, Image Name, IP Address, Size, Key Pair, and Status. The instances listed include vm0, vm0-1, lab-vm, oss, yd_lab_tt0, elk, geonetwork, nl-cloud-01, notebooks-worker-05, and notebooks-.

Instance Name	Image Name	IP Address	Size	Key Pair	Status
vm0	NOTEBOOKS.EGI.EU Image for EGI Docker [Ubuntu/18.04/VirtualBox]	212.189.145.142	m1.small	delete-me	Active
vm0-1	NOTEBOOKS.EGI.EU Image for EGI Docker [Ubuntu/18.04/VirtualBox]	212.189.145.45	m1.small	delete-me	Active
lab-vm	NOTEBOOKS.EGI.EU Image for EGI CentOS 7 [CentOS/7/VirtualBox]	212.189.145.51	m1.large	ydlabtestkey	Active
oss	NOTEBOOKS.EGI.EU Image for EGI Docker [Ubuntu/18.04/VirtualBox]	212.189.145.175	m1.medium	oss	Active
yd_lab_tt0	NOTEBOOKS.EGI.EU Image for EGI CentOS 7 [CentOS/7/VirtualBox]	212.189.145.55	m1.xlarge	ydlabtestkey	Active
elk	NOTEBOOKS.EGI.EU Image for EGI CentOS 7 [CentOS/7/VirtualBox]	212.189.145.173	m1.medium	elk	Active
geonetwork	NOTEBOOKS.EGI.EU Image for EGI Docker [Ubuntu/18.04/VirtualBox]	212.189.145.37	m1.xlarge	geonetwork	Active
nl-cloud-01	NOTEBOOKS.EGI.EU Image for EGI Docker [Ubuntu/18.04/VirtualBox]	212.189.145.176	m1.large	max-nl-01-key2	Active
notebooks-worker-05	NOTEBOOKS.EGI.EU Image for EGI Docker [Ubuntu/18.04/VirtualBox]	212.189.145.133	m1.xlarge	-	Active
notebooks-	NOTEBOOKS.EGI.EU Image for EGI Docker				



Use case 1: automated workflow using IaaS

1. Get familiar with the OpenStack dashboard
2. Notebook of the automation tool, working on the code for the EAP IaaS

https://github.com/QCDIS/CONF/blob/develop/jupyter_notebooks/vre_demo.ipynb

The screenshot shows a Jupyter Notebook interface with the following content:

```
File Edit View Insert Cell Kernel Help Not Trusted Python 3
```

```
run /opt/conda/lib/python3.7/site-packages/urllib3/connectionpool.py:986: InsecureRequestWarning: Unverified HTTPS request is being made to host 'lifewatch.lab.uvalight.net'. Adding certificate verification is strongly advised. See: https://urllib3.readthedocs.io/en/latest/advanced-usage.html#ssl-warnings
InsecureRequestWarning,

In [42]: toasca = get_tosca(provisioned_tosca_id)

tosca_dict = yaml.safe_load(tosca)
graph = build_graph(tosca_dict['topology_template']['node_templates'])
nx.draw(graph, with_labels=True)

for node_name in toasca_dict['topology_template']['node_templates']:
    if toasca_dict['topology_template']['node_templates'][node_name]['type'] == 'tosca.nodes.QC.VM.Compute':
        print(node_name+' '+tosca_dict['topology_template']['node_templates'][node_name]['attributes']['public_ip'])

/opt/conda/lib/python3.7/site-packages/urllib3/connectionpool.py:986: InsecureRequestWarning: Unverified HTTPS request is being made to host 'lifewatch.lab.uvalight.net'. Adding certificate verification is strongly advised. See: https://urllib3.readthedocs.io/en/latest/advanced-usage.html#ssl-warnings
InsecureRequestWarning,

compute: 3.120.209.252
compute_1: 3.126.139.56
```

Below the code, a network graph is displayed with nodes and edges:

```
graph TD
    ws_pema --> kubernetes
    kubernetes --> topology
    topology --> compute
    topology --> compute_1
```




Use case 1: automated workflow using IaaS

1. Get familiar with the OpenStack dashboard
2. Notebook of the automation tool, working on the code for the EAP IaaS
3. Develop the infrastructure planner GUI for workflow

IaaS Planner

Parameter optimizer

Planner

- 1 Select application type
 - Regular Workflow
 - Regular Workflow**
 - Time-Constrained Workflow
 - Microservices
 - IOT
- 2
- 3
- 4 Configure QoS demands

Optimizer

Robin

Select the workflow of your application

Insert workflow file in cwl format

compile1.cwl (1.7 kB)

0 files (0 B in total)

Select the performance models that you wish to compare

Insert performance models in yaml format

5 files (731 B in total)

COMPARE CLEAR

Optimizer

Select the workflow of your application

Insert workflow file in cwl format

compile1.cwl (1.7 kB)

1 files (1.7 kB in total)

Select the performance models that you wish to compare

Insert performance models in yaml format

5 files (731 B in total)

COMPARE CLEAR

Lowest cost	Lowest makespan	Lowest total costs + makespan
File name: input_pcp_-_Copy.yaml	File name: input_pcp_-_Copy.yaml	File name: input_pcp_-_Copy.yaml
Total costs: 27	Total costs: 27	Total costs: 27
Makespan: 27	Makespan: 27	Makespan: 27



Use case 1: automated workflow using IaaS

1. Get familiar with the OpenStack dashboard
2. Notebook of the automation tool, working on the code for the EAP IaaS
3. Develop the infrastructure planner for workflow (screen snapshot)
4. Delivered a training

Programme and speakers

• July 13th 2020

- 09.30 - 10.00 (CEST): Welcome and general presentation
- 10:00 - 12:00 (CEST): **Cloud computing and application development for research infrastructures**
In this webinar, we will discuss the basic concepts of cloud computing, including virtualization, containerization, service models, and cloud application development. We will also discuss how clouds can support data management and scientific workflows in the research infrastructures via examples from ENVRIplus and ENVRI-FAIR projects.

Speaker: Zhiming Zhao

Dr Zhiming Zhao is an assistant professor at University of Amsterdam (UvA). He leads the "Quality Critical Distributed Computing" research team in the group of Multiscale Networked Systems (MNS) at the System and Networking Lab (SNE). His research interests include big data management, Cloud and edge computing, software engineering, and blockchain. He leads the development support WP in ENVRI-FAIR and the VRE development in the LifeWatch-ERIC Dutch Virtual Laboratory Innovation Center. He is also the UvA PI in SWITCH, ENVRIplus, ARTICONF and several other projects.



After the course: <http://tiny.cc/phbasz>

Tutorial for the Webinar "An introduction to Cloud computing"

Speaker: dr. [Zhiming Zhao](#), Support: dr. [Spiros Koulouzis](#),
University of Amsterdam, Amsterdam, NL
LifeWatch ERIC, vLab & Innovation Center, Amsterdam, NL

The tutorial is part of the webinar "[an introduction to Cloud computing](#)", in the [ENVRI community winter school 2020](#). In this tutorial, you will learn how to define a simple REST service using OpenAPI. You will also learn how to use Ansible and Kubernetes, a.k.a K8s to deploy the RESTful Web Service on a VM in Cloud environments.

We sincerely thank dr. Giuseppe Larocca and dr. Andrea Manzi from EGI to provide the testbed via the EGI training platform. The tutorial is supported by the [EOSC early adopter program](#) via [ENVRI-FAIR](#) project, and [LifeWatch-ERIC](#). The testbed will be accessible after the webinar for 10 days; during those days we will also provide support for all technical questions.

0. Before you Begin

Install Ansible on local machine (laptop)

You will need Ansible for the assignment. Please install it on your local computer based on the following instructions:

What you need:

1. A laptop
2. Follow instruction to get a VM (from EGI)
3. Follow the tutorials.

The VM will be available for 10 days.





Use case 2: notebook for data sciences

- Jupyter extension development
 - Jupyter hub
 - Data hub
 - They will be connected soon



Use case 2: notebook for data sciences

- Jupyter extension development
 - Jupyter hub
 - Data hub
 - They will be connected soon
- FAIR-Cells: Customize Jupyter environment

FAIR-Cells

About

FAIR-Cells is developed by the research team for Quality Critical Distributed Computing (QCDCS) in the Multi Scale Networking System Group (MNS) of the University of Amsterdam. The development of FAIR-Cells is coordinated by Dr. Zhiming Zhao: z.zhao@uva.nl

Developers:

- W. Kruijerwilco@uva.nl
- S. Koulouzis@uva.nl

MNS UvA

FAIR-Cells

Build

Image name: nb_helper

Base image: Data Science Notebook

Cell: Cell 5

Preview:

line = [0.25, 1.5]
mean_line = 0.875

Conda environment.yml:

```
# Autogenerated by FAIR-Cells, please adjust.
channels:
- conda-forge
dependencies:
- pip:
- matplotlib
```

Variable	Query	Post	Disabled
sum	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
len	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
line	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
mean_line	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
plt	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>



Use case 3: DevOps

- Get Jelastic provisioned
- Learning the technology

The screenshot displays the Jelastic dashboard. At the top, there are navigation buttons for 'NEW ENVIRONMENT', 'IMPORT', and 'MARKETPLACE'. Below this, the 'Env Groups' section is visible, featuring a table with columns for 'Name', 'Status', and 'Tags'. The table is currently empty. At the bottom of the interface, there are buttons for 'Archive', 'Git / SVN', 'Upload', and 'Delete'. Below these buttons is a table with the following data:

Name	Comment	Size
HelloWorld.zip	Sample package which you can deploy to your environm...	488 KB



Summary

Phase 1

Describe the key use case: scenario, scope, KPI, steps etc.

Get familiar with the EOSC services, following training and practices from the other projects etc.

Get the requested resource provisioned
Setup OLA with resource providers and agreement with Jelastic.

Phase 2

DevOps pipeline configured, including Git, automated testing, integration and **deployment demonstrate** in at least via two service development

Phase 3

Demonstrate the initial version of the workflow from ENVRI-FAIR, with automated workflow execution in Cloud;
Demonstrate the other common data services identified in the ENVRI communities (optional)

Phase 4

Exploit the results to the development activities in ENVRI-FAIR sub domain
Sustain the development by finding new opportunities, e.g., new EOSC projects etc.



Next phase

- Use case 1:
 - Demonstrator: infrastructure planning + automation
- Use case 2:
 - Demonstrator 2.1: from Jupyter to service flow
 - Demonstrator 2.2: demos from community
- Use case 3:
 - Demonstrator 3.1: demonstrate the automation of testing, integration, deployment



Sustainability

1. Exploitation of the results to the ENVRI communities
2. Getting support from other ongoing relevant projects, e.g. ARTICONF, BlueCloud and CLARIFY
3. Other opportunities



ENVRI
FAIR

EOSC Earlier Adopter Program

www.envri.eu



envri.eu/envri-fair



[@ENVRIcommunity](https://twitter.com/ENVRIcommunity)



[ENVRI community](https://www.linkedin.com/company/ENVRIcommunity)



facebook.com/ENVRIcommunity