With the support of the Erasmus+ Programme of the European Union Sector Skills Alliances N° 591991-EPP-1-2017-1-IT-EPPKA2-SSA-B

D 3.3 – Revised and improved platform and tools

Author(s)/Organisation(s):

- Jeroen Dries (VITO)
- Hande Erdem (VITO)
- Sven Casteleyn (UJI)

Work package / Task:

WP3 - Developing an integrated platform of collaborative tools

T3.2 – Revising, fine-tuning and improving existing platforms and tools

Short Description:

This document describes the EO tools that have been built or improved for EO4GEO. The publicly accessible tools are referenced in this document.

Keywords:

Dissemination Level		
PU	Public	x
RE	Restricted to other programme participants (including Commission services and project reviewers)	
CO	Confidential, only for members of the consortium (including EACEA and Commission services and project reviewers)	

With the support of the Erasmus+ Programme of the European Union Sector Skills Alliances N° 591991-EPP-1-2017-1-IT-EPPKA2-SSA-B

Revision History:

Revision	Date	Author(s)	Status	Description
1.0	16/12/2019	Jeroen Dries	Draft	First version
2.0	18/12/2019	Sven Casteleyn	Draft	Second version
3.0	15/01/2020	Hande Erdem	Final	Quality evaluators' comments have been incorporated into the document

Table of Contents

Acronyms

BoK	Body of Knowledge
VET	Vocational Education and Training
MEP	Mission Exploitation Platform
LMS	Learning Management System
NDVI	Normalized Difference Vegetation Index
LAI	Leaf Area Index
FAPAR	Fraction of Absorbed Photosynthetically Active Radiation
FCOVER	Fraction of Vegetation Cover

Glossary

eduGAIN is an international interfederation service interconnecting research and education identity federations. It enables the secure exchange of information related to identity, authentication and authorisation between participating federations. https://edugain.org

H5P is a free and open-source content collaboration framework based on JavaScript. H5P is an abbreviation for HTML5 Package, and aims to make it easy for everyone to create, share and reuse interactive HTML5 content. Interactive videos, interactive presentations, quizzes, interactive timelines and more can be developed and shared using H5P. https://h5p.org/

Jupyter Notebook is a web-based interactive computational environment for creating Jupyter notebook documents. The "notebook" term can colloquially make reference to many different entities, mainly the Jupyter web application, Jupyter Python web server, or Jupyter document format depending on context. https://jupyter.org

OpenID is an open standard and decentralized authentication protocol. https://openid.net

QGIS (previously known as Quantum GIS) is a free and open-source cross-platform desktop geographic information system (GIS) application that supports viewing, editing, and analysis of geospatial data. https://qgis.org/en/site/

1. Introduction

1.1. EO4GEO project

EO4GEO is an **Erasmus+ Sector Skills Alliance** gathering **26 partners from 13 EU countries**, most of which are part of the **Copernicus Academy Network**. Be they from academia, public or private sector, they are all active in the education and training fields of the space / geospatial sector. The project is also supported by a strong group of Associated Partners mostly consisting of associations or networks active in space/geospatial domain. The project started on January 1st, 2018, upon approval by the EU Education, Audiovisual and Culture Executive Agency (EACEA) and runs over four years.

EO4GEO **aims to help bridging the skills gap in the space/geospatial sector** by creating a strong alliance of players from the sector/community reinforcing the existing ecosystem and **fostering the uptake and integration of space/geospatial data and services**. EO4GEO will work in a **multi- and interdisciplinary** way and apply innovative solutions for its education and training actions including: case-based and collaborative learning scenarios; learning-while-doing in a living lab environment; on-the-job training; co-creation of knowledge, skills and competencies; etc.

EO4GEO defines a long-term and sustainable strategy to fill the gap between supply of and demand for space/geospatial education and training taking into account the current and expected technological and non-technological developments in the space/geospatial and related sectors (e.g. ICT). The strategy is implemented by: creating and maintaining an ontology-based Body of Knowledge for the space/geospatial sector based on previous efforts; developing and integrating a dynamic collaborative platform with associated tools; designing and developing a series of curricula and a rich portfolio of training modules directly usable in the context of Copernicus and other relevant programmes and conducting a series of training actions for a selected set of scenario's in three sub-sectors - integrated applications, smart cities and climate change to test and validate the approach. Finally, a long-term Action Plan is developed and endorsed to roll-out and sustain the proposed solutions. For more information on the project please visit http://www.eo4geo.eu/about-eo4geo/.

1.2. Objectives of the work package

Work package 3 has following specific objectives:

- To identify, assess and select existing platforms and components that can be used/reused and (eventually) integrated;
- To identify the minimal functional and non-functional user requirements the platform(s) should match;
- To discuss and get feedback on the proposed architecture and technical solutions proposed;
- To test and exploit existing platforms and tools, analyse their use/usability and collect input for potential improvement;
- To describe, plan and document the changes that will be made to existing platforms and tools;
- To describe new tools that will be developed;
- To modify and extend coding of existing platforms and tools and to develop new tools;
- To document and provide guidelines for users and to organize acceptance tests.

1.3. Objectives of the task – update with respected to DoW

Work Package 3 essentially foresees the selection, customization, fine-tuning and development of tools relevant for the project. The objective of Task 3.2 as foreseen in the Document of Work is to revise, fine-tune and improve existing platforms and tools. On the other hand, the objective of Task 3.3 as foreseen in the Document of Work is to develop new innovative tools for the platform. While this distinction between existing and new tools made sense at the time of the proposal, due to various reasons detailed below, it was decided during the execution of the project to re-organize the work into managerially and logically connected/related tools. Concretely, it was decided to group all tools related to Geographic Information (GI) and Earth Observation (EO), together with the learning platform, in Task 3.2; on the other hand, all tools related to the Body of Knowledge were grouped in Task 3.3.

The motivation for this was fourfold: (i) unforeseen circumstances and decisions made throughout the project caused some tools that were originally foreseen to be re-used to be re-developed. Concretely, due to unsustainability of the existing solution, the BoKWIKI platform was per

executive decision replaced by the Living Textbook (developed/extended by partner UTC) with the restful services exposing data of the BoK (newly developed by partner UJI). Therefore, tools automatically transitioned from Task 3.2 to Task 3.3. (ii) due to the architectural decision to develop the BoK-related tools as much as possible as re-usable and interacting components, it was logically better to group functionality related to the BoK-related tools under one umbrella. For example, a re-usable component to visualize and browse through the BoK was partly extracted from the Curriculum Design Tool (originally T3.2), made available as a new independent tool (originally Task 3.3) and re-used both in the renewed Curriculum Design Tool (originally T3.2), the Occupational Profile and Job Offer tools (both originally T3.3). As illustrated by the example, the original classification between existing and new tools was no longer tenable. (iii) the learning platform, originally foreseen to be reported in D3.4 (Task 3.3), is closely connected to the PROBA-V Mission Exploitation Platform, as it interacts with it to provide new, interactive learning material, relevant for Task 4.3, and the case-based scenarios (WP5). It was thus preferable to group both under a single task. (iv) both for the GI&EO related tools and for the BoK-related tools, there is one project partner that is the main developer of the tools. Therefore, from a project management and software development point of view, if makes sense to group the respective tools under the responsibility of the respected partners, i.e., VITO for GI&EO related tool (now T3.2) and UJI for BoK-related tools (now T3.3).

Despite the re-organization of tools among Task 3.2 and Task 3.3, it needs to be stressed that all foreseen work has been / is carried out.

As such, currently under Task 3.2 (Deliverable D3.3) the following tools fall:

- The PROBA-V Mission Exploitation Platform, to be extended with more functionalities to allow easier use in training context, some additional apps for use within the case-based learning actions.
- The learning platform (e.g., Moodle, EOCollege), for exchanging learning materials and connect to the PROBA-V Mission Exploitation Platform to develop practical learning material and support use cases.

The BoKWiKi (replaced by the Living Textbook), the RESTful services to expose BoK data and the Curriculum Design tool, will be reported along with other BoK-related tools in Deliverable D3.4 (Task 3.3).

1.4. Purpose of the document

In task 3.2, the requirements and use cases of the EO4GEO project (gathered in D3.1) were used as input to improve and extend EO tools available in the PROBA-V MEP platform. As the PROBA-V MEP platform is now also extending into other Copernicus data (e.g. Sentinel 1,2,3,5), the platform is rebranded as 'Terrascope', the name of the Belgian collaborative ground segment. On the other hand, a learning platform foreseen for EO4GEO was installed and connected with Terrascope content.

Main purpose of this document is to describe the EO tools and the setup of the learning platform that make up deliverable 3.3.

1.5. Structure of the document

After a brief look at the requirements, we will go into details of each tool that is provided as part of the deliverable. These are:

- E-learning platform
- EO Platform
    - EO data viewer
    - Jupyter notebooks
    - User virtual machines

2. Requirements analysis

An analysis of the requirements and use cases described in D3.1 allowed us to identify the main areas of work:

2.1. E-learning platform

The requirements on the e-learning platform were compatible with most mature LMS systems available today. Some features do depend on the specific configuration or setup of the system. So the selection of an actual specifically configured instance became relevant.

Based on the feedback from project partners, and the steering committee, it became clear that no single learning management system (LMS) could be selected for use within the project. The main reason for this is that different (educational) institutions usually already have an LMS in place, so the focus shifted to ensuring interoperability with various LMS systems.

The requirement to reuse and import modules (513, 515) led to specific integration testing with H5P content. Various content types were tested to see if they could be used across systems, and in the context of developing training material for EO use cases.

2.2. EO-platform

The following requirements lead to the most important improvements of the platform:

606 – Single sign on: Support for openID connect was implemented throughout components in the system. This was then used to integrate the EduGain identity provider. This provider groups a large number of educational/research institutions in Europe, allowing for instance university students to log in with the account from their institutions. This was certainly not the easiest choice to integrate with, but we believed that logging in with a verified account is important if the tools are to be used in the setting of a course.

709 – Learning platform integration: integration approaches with the learning management systems were explored and verified, so that they can be used in the development of new training material.

703 – Scalability and 705 – automation and testing: these non-functional requirements led to an overhaul in the deployment of the Jupyter notebooks service adding for instance the capability to automatically clean up old instances improved the scalability. Also addition of automated testing of notebooks increased the reliability.

There are also requirements that required a lot of work which was not implemented with the resources of the EO4GEO project. We'd like to mention some noteworthy requests here:

- Addition of new Copernicus derived products (Coherence, Sentinel 5p,…)
- Numerous improvements to the Terrascope viewer, enabling data discovery
- Full upgrade of the cloud environment that hosts the user virtual machines
- Upgrades of EO software such as QGis and the Sentinel Toolbox

3. Provided tools

3.1. E-learning platform

For interoperability tests, a Moodle instance was provided (https://eo4geo.planetek.it/). Several types of test accounts were created to test the various roles (student, teacher). This setup will be used further during the testing phase, when validating the case-based learning scenarios (WP3), and also explored when Improving existing and developing new training material (T4.3).

Likewise, the EO-college (https://eo-college.org) platform can be used in a similar manner. Having two platforms, built on differing technologies (Moodle/Wordpress), should sufficiently ensure the portability of learning material developed in further work packages.

To validate the platform a proof of concept was done on the Proba-V MEP Drupal based setup and on the Moodle instance hosted by planetek. The proof-of-concept on the MEP can be found here: https://proba-v-mep.esa.int/node/182. Note that it focuses on using various technologies, not on the educational/pedagogic aspects. Nevertheless, the learning objectives of this particular H5P exercise within LMS can be summarised as:

1. Being able to visualize and compare different images using sliders or other easy user controls
2. Being able to interpret the differences in various vegetation indices (LAI, in this case)
3. Answer quiz questions to get immediate feedback

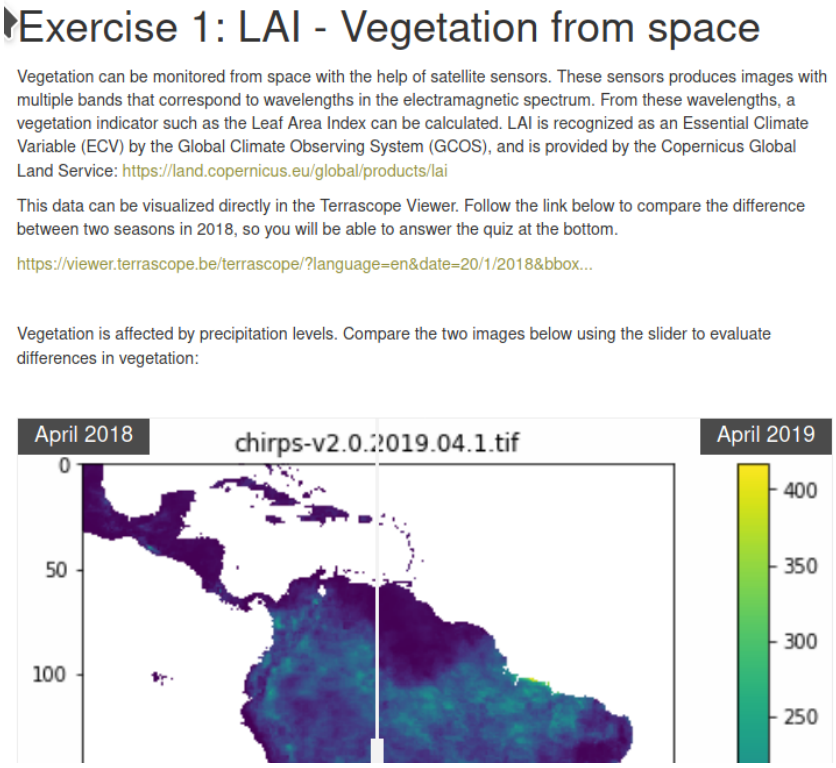

Figure 1Excercise using H5P widget integrated in a Drupal portal

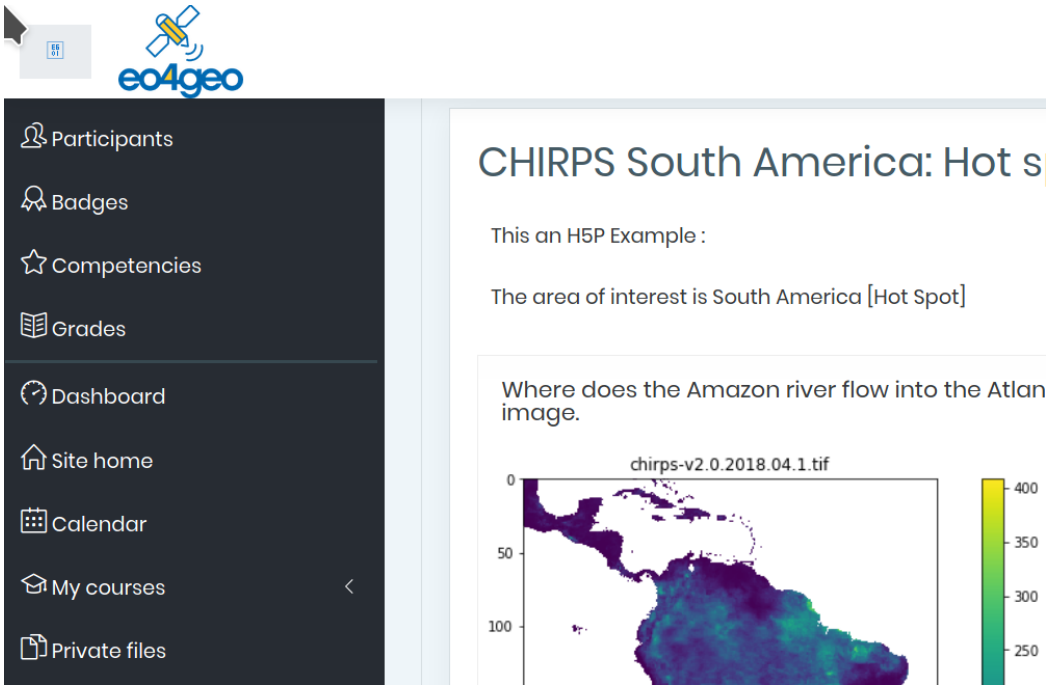

Figure 2 Moodle LMS showing EO4GEO H5P test

3.2. EO Platform

This section gives an overview of the tools that make up the EO platform. These tools are also listed on a (temporary) landing page: https://proba-v-mep.esa.int/eo4geo.

A main requirement for the platform was to integrate with a single-sign-on service. Therefore, integration with the EduGain system has been implemented, but is still being rolled out across the production components. This needs to be done in a controlled and gradual manner, to ensure continued operation/accessibility of the platform for existing users.

3.2.1. EO data viewer

The EO data viewer (https://viewer.terrascope.be) allows users to browse and explore available datasets. To support the use cases, we provide a number of products:

- Sentinel-2 L2A over Europe
- Sentinel-2 derived vegetation indices (NDVI, LAI, FAPAR, FCOVER)
- Sentinel-1 SLC, Sigma0 and Coherence to support SAR use cases
- Copernicus Global Land timeseries provide global timeseries starting in 1998. Can be used for instance in the climate case based learning scenario.

The viewer also includes some additional features:

- Comparison of images for different dates
- Generating animations
- Extracting aggregated timeseries over an area of interest

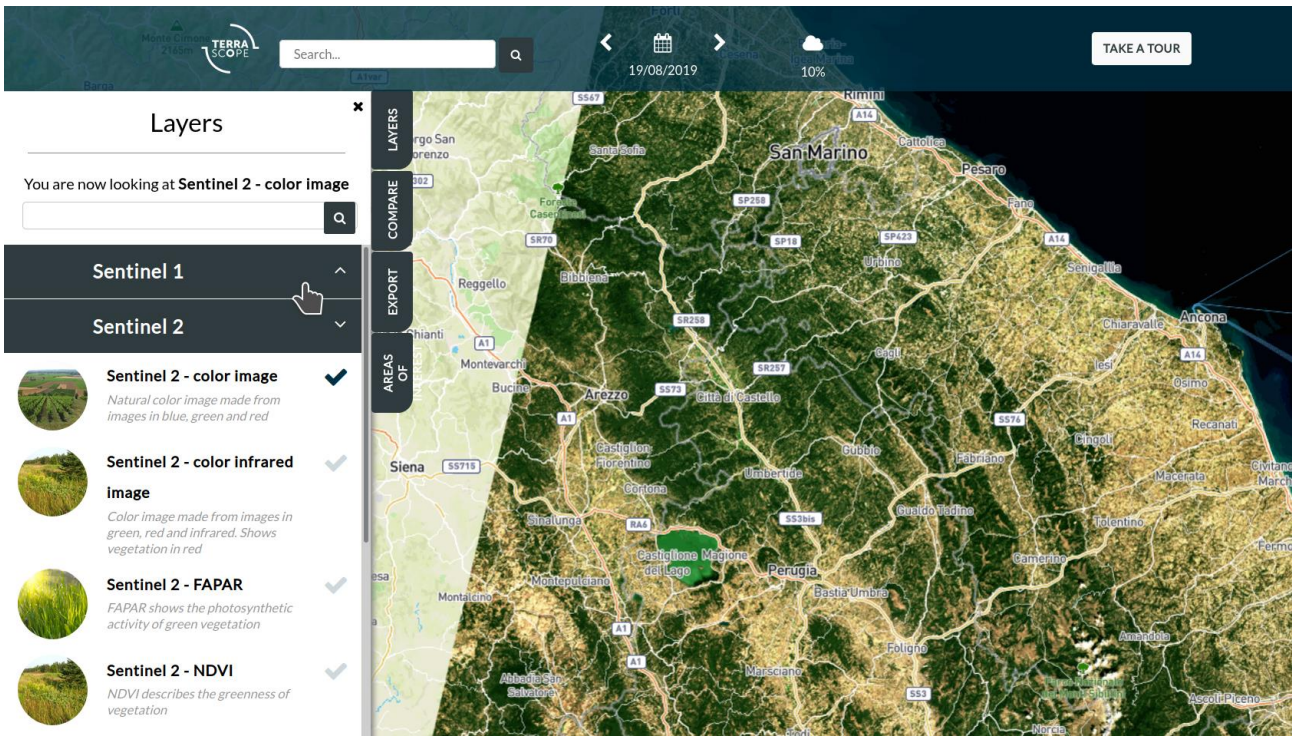

Figure 3Terrascope viewer enabling data discovery

3.2.2. Jupyter notebooks

The Jupyter notebooks service (https://notebooks.terrascope.be) allows users to create notebooks in an environment that has direct access to the data. It is also preconfigured with a number of useful libraries, so that students/teachers do not need to spend time on installing tools. It also ensures that results are reproducible, as they always run in the same environment.

A system of shared folders allows students and/or teachers to exchange data directly on the platform.

A set of example notebooks is provided as training material for the platform. These notebooks are tested automatically to ensure that they continue to run as the underlying platform is upgraded.

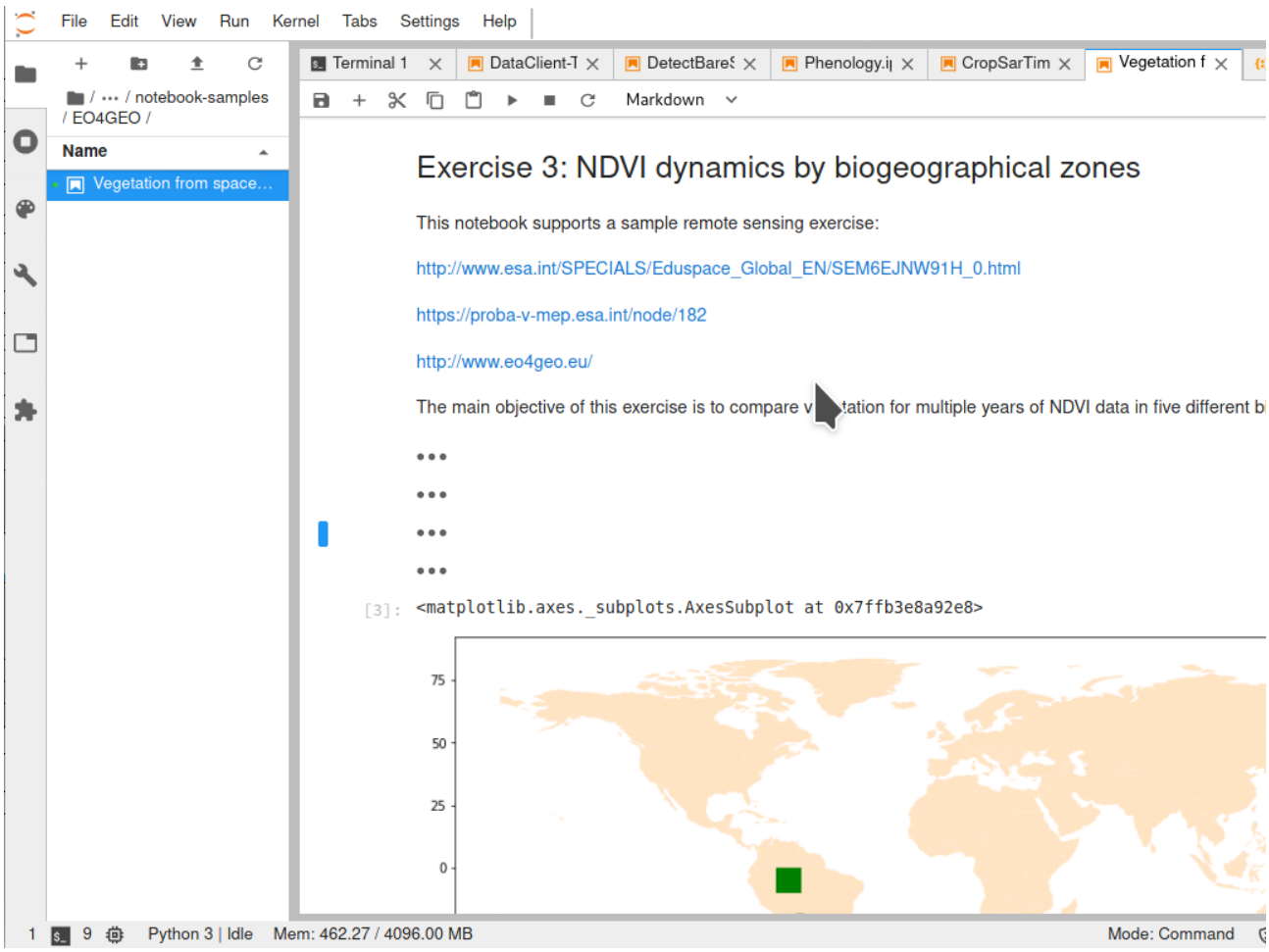

Figure 4 Jupyter notebook service showing EO4GEO excercise

3.2.3. User virtual machines

https://proba-v-mep.esa.int/vre

This tool allows users to request a virtual machine with a number of EO tools installed and direct access to an archive of earth observation data. Users can also install their own tools as needed.

From this machine, a powerful processing cluster can also be accessed for more advanced use cases such as implementing a processing workflow.

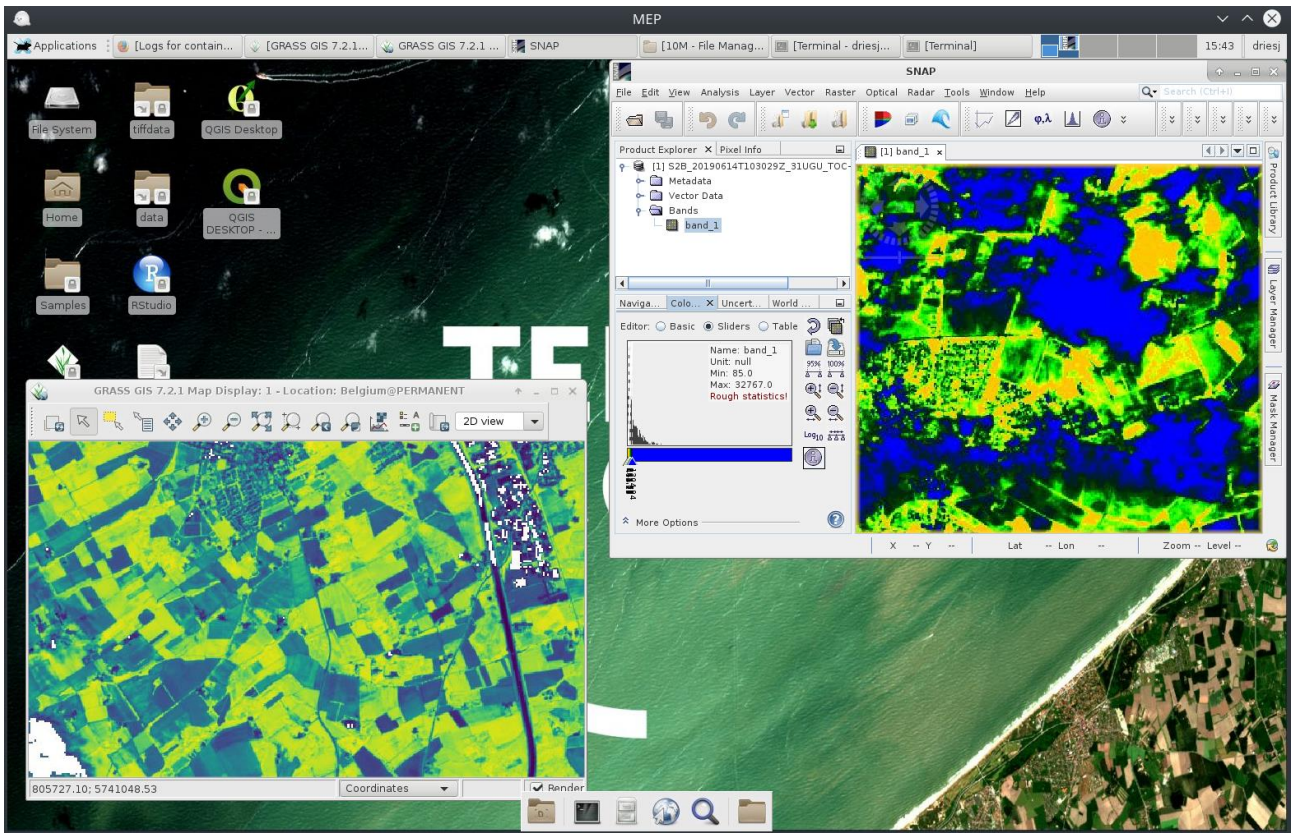

Figure 5 Desktop interface of user VM

4. Next steps

The presented tools will now have to be used in work package 4 and 5, to support the creation of innovative learning material and validate the case-based learning scenarios. We anticipate that this will result in new requests for the tools, or improvements to the implemented functionality. As the partners of this task (3.2) are also involved in the relevant tasks of work package 4 and 5, we foresee that these requests can be evaluated case by case, and implemented where needed. For this reason, and in agreement with WP3 leader and EO4GEO project coordinators, some resources have been kept, in order to support further develop the EO tools and learning platform if required by WP4 and/or WP5. If this is the case, an addendum to this deliverable will be filed, explaining the additional work carried out.

5. Conclusions

Both the e-learning platform and the EO platform (including its sub-tools) have been delivered as part of Task 3.2. To be more precise:

- During the task, the focus was on the content types rather than the LMS since the need or choice of the LMS will be related to the results of WP4 and WP5.
- EO Platform that is being used for EO4GEO project is based on the MEP PROBA-V EO Platform, also known as Terrascope EO Platform. This allows EO4GEO project to benefit from an extensive set of EO platform functionality that has been developed over the years. By being the platform of choice for EO4GEO project, MEP PROBA-V EO Platform benefits from extending its reach to new users.

Moreover, EO platform had additional features, improvements and upgrades done outside the resources of the EO4GEO but nevertheless benefits EO4GEO project by aligning with the objectives of providing extensive tools on a collaborative platform.

6. Annex

This section describes the general approach that is being utilized in validating PROBA-V MEP EO Platform which is used in EO4GEO project. The aim of this section is to give an overview of the test breakdown and test results. Therefore, this document does not go into details of test cases and scenarios. As this document covers the complete MEP platform, some test cases may not be directly applicable to EO4GEO project.

PROBA-V MEP EO Platform is a software product developed using agile methodology and continuous integration over several years. As a result of using this methodology, majority of the testing is done in an automated fashion using build pipelines. Tests are executed whenever there is a code change or whenever there is an explicit schedule to run them. This brings significant advantage and raises the quality of the product considerably.

Graphical user interface tests and performance (load/stress) tests are executed manually for different reasons. It is costly to write and maintain the stability of user interface tests, therefore, we currently execute them manually.

Load/Stress tests are also done manually using special tools. The summary result of those tests can be found in this document. We must add that MEP EO Platform is an operational platform and performance of various components are continuously monitored and measured by a specialized software operations team using state of the art tools. What we describe in this document is only a subset of those efforts.

6.1. Test Results

All relevant tests (both functional and performance) have passed successfully. Few functional tests that have failed (3 out of 2367) are out of the scope of EO4GEO project and do not have any impact on the platform functionality.

6.1.1. Functional Tests

Test Category	Success #	Failed #	Total #
geo.atcconf_generator	5	0	5
geo.CatalogClient	10	0	10
geo.dias-s2chain	50	0	50
geo.ecmwf-meteo-processor	1	0	1
geo.es5-gateway	11	0	11
geo.GeotrellisTimeseries	296	0	296
geo.gwc-geotrellis	1	0	1
geo.JobControl-Callback	47	0	47
geo.JobControl-Callback-Python	19	0	19
geo.JobControl-dashboard	59	0	59
geo.openeo-python-helpers	1	0	1
geo.oscars-mgmnt	38	0	38
geo.PDFDataStore	9	0	9
geo.pid-java	60	0	60
geo.POD-HMA-Subscription	72	0	72
geo.POD-REST	108	0	108
geo.SeedingScripts	8	3	11
geo.sentinel2-biopar	15	0	15
geo.sentinel2-biopar-integrationtests	7	0	7
geo.snap-spark	9	0	9
geo.TileSeeding-pipeline	16	0	16
geo.Timeseries_Analysis_Service	8	0	8
openEO » openeo-geopyspark-driver	63	0	63
openEO » openeo-geopyspark-integrationtests	19	0	19
openEO » openeo-geotrellis-extensions	35	0	35
openEO » openeo-python-client	139	0	139

openEO » openeo-python-driver » feature/processes-gitsubmodule	48	0	48
openEO » openeo-python-driver	133	0	133
OSCARS	123	0	123
PROBAV-Composite	78	0	78
Processing Information Datastore Python client	18	0	18
ProcessingJobs	778	0	778
S1 SLC Coherence	1	0	1
UI Tests	72	0	72
Notebook Tests	5	0	5
Single sign on (eduGain) Test	2	0	2
Total	**2364**	**3**	**2367**

6.1.2. Load and Stress Tests

Test Category	Passed	Failed	Total
Timeseries execution @ 250 requests – Long range	1	0	1
Timeseries execution @ 250 requests – Short range	1	0	1
Timeseries execution @ 300 requests – Long range	1	0	1
Timeseries execution @ 300 requests – Short range	1	0	1
Timeseries execution @ 400 requests – Long range	1	0	1
Timeseries execution @ 400 requests – Short range	1	0	1
Timeseries execution @ 500 requests – Long range	1	0	1
Timeseries execution @ 500 requests – Short range	1	0	1
Mapcache load test @ 500 requests	1	0	1
Mapcache load test @ 1000 requests	1	0	1
Mapcache load test @ 1200 requests	1	0	1
WMTS load test 186000 request by 100 users	1	0	1
WMTS load test 186000 request by 1000 users	1	0	1
WMTS load test 217000 request by 1000 users	1	0	1
WMTS load test 403000 request by 100 users	1	0	1
Total	**15**	**0**	**15**