Report on Training Activities

Work package	**WP 2 Community building and dissemination**	
Task	Task 2.3 Training	
Author (s)	Fulvio Marelli	T2
	Pedro Goncalves	T2
	Stefano Salvi	INGV
Reviewer (s)	Rosemarie Leone	ESA
Approver (s)	Cristiano Silvagni	ESA
Authorizer	Mirko Albani	ESA
Document Identifier	EVER-EST DEL WP2-D2.5	
Dissemination Level	Public	
Status	Approved by the EC	
Issue	1.1	
Date	14 December 2018	

Document Log

Date	Author	Changes	Version	Status
20/10/2017	Fulvio Marelli	TOC	0.1	Draft
05/11/2017	Fulvio Marelli	Contributions from partners integrated	0.2	Draft
12/12/2017	Fulvio Marelli	Materials from AGU integrated	0.3	Draft
11/01/2017	Fulvio Marelli	Final draft, ready for internal review	0.4	Draft
09/07/2017	Helen Glaves, Cristiano Silvagni	Major corrections	0.5	Draft
13/02/2017	Fulvio Marelli	Comments merged. Ready for delivery	1.0	Draft to be approved by the EC
14/12/2018	Cristiano Silvagni	Document status	1.1	Approved by the EC

Table of contents

List of Tables

List of Figures

Definitions and Acronyms

Acronym	Description
AHM	All Hands Meeting
CNR-ISMAR	Consiglio Nazionale delle Ricerche – Istituto di Scienze Marine
COTS	Commercial Off the Shelf
ES	Earth Science
EO	Earth Observation
ESA	European Space Agency
EVER-EST	European Virtual Environment for Research - Earth Science Themes
GUI	Graphical User Interface
RDA	Research Data Alliance
RDF	Resource Description Framework
RDFS	RDF Schema
RO	Research Objects
SATCEN	European Union Satellite Center
VRC	Virtual Research Community
VRE	Virtual Research Environment
W3C	World Wide Web Consortium

Applicable Documents

Document ID	Document Title
EVER-EST DEL WP2-D2.4	Training Strategy and Plan
EVER-EST DEL WP3-D3.1	Virtual Research Environment use case detailed definition
EVER-EST DEL WP4-D4.1	Workflows and Research Objects in Earth Science
EVER-EST DEL WP5-D5.1	VRE architecture and interfaces definition

Reference Documents

EVER-EST DEL WP2-D1.1	Project Management Plan
EVER-EST DEL WP2-D2.1	Dissemination strategy and plan
EVER-EST DEL WP2-D2.4	Training Strategy and Plan

1 Introduction

The purpose of this document is to provide a report related to the training activities carried out during the first two years of the EVER-EST project. The logic and roadmap for the EVER-EST training activities has been designed and documented in the deliverable D2.4 Training Strategy and Plan. This introductory chapter provides a short summary of D2.4 to recall and focus on its main guidelines.

1.1 Document structure

The document is organized as follows:

- **Chapter 1 - Introduction**: provides a brief overview of the deliverable D2.4 Training Strategy and Plan and its logical connection to this report.
- **Chapter 2 - Internal Training**: the core element of this document is to provide information about all the internal training activities conducted during the first two years of the project. It is divided into three main components:
    - Training on Research Objects (RO): Introducing RO's in the Earth Science context for the first time required mutual understanding about their meaning and practical usage;
    - Technical training: from the e-infrastructure developers to the user communities to describe the services provided;
    - User actions training: this activity was not foreseen at the beginning of the project however it resulted necessary to align terminology and constraints in order to understand the detail of a research workflow and how the Virtual Research Environment (VRE) could help.

- **Chapter 3 – External training**: provides an overview of the training activities carried out with stakeholders outside the EVER-EST project which took place during the first 24 months of the project. This training focused on three main objectives/activities:
    - Virtual Research Community (VRC) building: enlarging the VRC communities based on the work carried out by CNR and INGV within their organisations (i.e. other research departments)
    - External training at AGU 2017 based on INGV work
    - Early adopters: based on joint work carried out with the University of Tor Vergata

- **Annexes**: a selection of the most relevant internal training materials produced during the first two years of the project. The rationale for the selection of these particular training resources will be elucidated throughout this document. It should be noted that these annexes represent only a small part of the training material that, although not included in this document, are available on the project repository and can be accessed through the VRE platform in the Help section (https://vre.ever-est.eu/)

1.2 From training strategy to implementation

The training strategy outlined in deliverable D2.4 Training Strategy and Plan was formulated during the first few months of the project to support the following EVER-EST's main objectives:

1. Creation of a Virtual Research Environment for the Earth Sciences;
2. Validate the infrastructure by demonstrating the benefit of the VRE for the four pre-selected Virtual Research Communities;

3. Apply the concept of Research Objects and demonstrate the potential benefits for the Earth Science domain.

Based on the project roadmap and key milestones, delivery of the final version of the EVER-EST VRE is planned to occur in M24. The associated training activities have therefore been divided into two main activities:

- **Internal training for the EVER-EST VRCs**: this Internal training will be provided throughout the EVER-EST project, with a particular effort being made during the first two years of the project to allow the development of a mutual understanding between the VRCs and the technical development team, identify potential gaps and establish a common understanding throughout the project team. This activity will also address the definition of the training material that shall be made available to early adopters or to new Earth Science communities.

- **External training for associated communities:** external training for communities which are close to the four VRCs that are represented in the project (e.g. Volcanic community or another INGV department), and early adopters coming from new Earth Science research sectors shall commence at the end of the first two years of the project.

Note: this document reports on the training activities undertaken during the first 24 months and therefore focuses largely on internal training. Some external training activities have commenced after being engaged by new stakeholders while presenting the project in public events. Those training sessions were not foreseen at the beginning of the project but have been added also to consolidate the training material and overall training approach.

Since the very beginning of the project, internal training sessions turned out to be a critical activity as they were the means to facilitate a shared and common understanding of how Virtual Research Environment can enhance the way researchers share their work within and across teams and institutions. The four VRC's involved in the project had a different perception and background, both at theoretical and at practical level, of how a VRE and related services could be adopted and integrated. This clearly included technical aspects, but also IPR's, security issues, barriers to data/research results sharing, etc.

As expected, one of the most difficult aspects that has been encountered by the EVER-EST team, both by the VRCs and technical partners, has been to agree, adopt and develop the full potential of Research Objects for the Earth Science domain. As this is the first time that the integration of the RO concept into an Earth Science e-infrastructure has been attempted a constant dialogue between RO specialists, technical teams and the VRCs was initiated. This resulted in a mutual understanding of the overall approach which led to the adaptation of the Earth Science technologies and the VRC user requirements to the new paradigm. The first section of the document therefore describes the main events, challenges, ideas and feedback taken into consideration when developing the training activities (**Error! Reference source not found.**).

The RO training phase commenced once the EVER-EST team had a more detailed understanding of RO's potential which led to the consolidation and creation of new advanced VRE user requirements based on the adoption of Research Objects. On one side the VRC's were capable to envisage new workflows within and across their teams and on the other side the technical partners were able to engineer and integrate their Earth Science services into the new context and correctly link them to the RO technologies.

Another key milestone for the internal training is represented by the User Actions document that allowed the RO experts and technical partners to gain an in depth view of the behaviour and tasks of a single researcher and how they share information. This activity enabled the technical teams to identify each of the actions performed by the VRCs which in turn facilitated the integration of the services within the VRE.

Demos and hands-on session were held in accordance with the original plan and while the Earth Science services where being developed and reaching a concrete level of maturity allowing the VRC to adopt the services, first as standalone services but progressively as platform integrated services, a series of meetings were organised by the technical partners with the aim to inform the VRCs about each new service release, implemented technology, user procedures and work instructions.

As anticipated, external training activities were performed during the first 24 months of the project. Although the VRE was not available to external users and still under development, the level of maturity of some modules and the logic of the overall e-infrastructure was demonstrated to the external audience which led to training events for new user communities.

2 Internal Training

The internal training described in the current report cover phase 1 and 2 that were defined in D2.4:

- **Internal Training Phase 1 - The basics (M1 – M12):** the goal was to start the dialogue on the future usage of the platform between software developers and Earth Science researchers. The training focuses on explaining to the community how to familiarize themselves with the most innovative technologies that the project will deliver.
- **Internal Training Phase 2: Hands on the EVER-EST services (M12 – M18):** the training program continues with interactive activities using the first software modules which will be delivered by the technical team through ad hoc meetings and via the training environment. The VRC's will be trained on how to use these modules and connect to them through workflows.

2.1 Research Objects training

The methodology adopted for the analysis of research objects in the Earth Science communities and the elicitation of requirements has three main pillars:

1. A **Research Object Requirements Questionnaire** that contained questions related to the goals, content, metadata and user requirements of ROs. The questionnaire was given to each of the four VRCs and the latter's answers were analysed in order to derive requirements. The questions were based on existing RO requirements from other scientific communities and the goal was to establish a) which of these requirements are applicable in Earth Science as well and b) what additional requirements may be needed.
2. A **Research Object Survey** that is addressed to the broader Earth Science community (i.e. not merely the four VRCs of this project) and aims to a more comprehensive understanding of its needs with respect to ROs and scientific knowledge sharing and preservation.
3. A series of **Research Object Hackathons** where users from the four VRCs received comprehensive training on research objects and had the chance to create their own ones using existing tools and models. This exercise helped clarify and distil the requirements derived from the questionnaire and helped identify potential issues and challenges with respect to their implementation.

As widely described in the introductory section, training on Research Objects was a critical activity aimed to facilitate their integration into an Earth Science platform and to ensure their adoption by the communities. Two main events were organised during the first year of the project to introduce and discuss in detail how Research Objects can change the way researchers work together. Although the main audience were the Virtual Research Communities both training events were also attended by the majority of the technical partners. More specifically, the RO hackathons allowed:

- **Having hands-on workshops with the EVER-EST VRCs** where members of the scientific communities can actually use the available tooling provided by previous EU projects in order to create their own research objects. Note that the Research Objects created this way respond to specific challenges of the VRCs' daily work, either focused on the scientific aspects involved (Sea Monitoring, Supersites) or from a more operational point of view (Land Monitoring and Natural Hazards).
- **Collecting first-hand impressions from the VRCs** on the utility of the Research Object concept as well as identifying the main conceptual gaps, technological limitations, extensions, and the customization that these communities require in order to leverage the whole potential of the research object paradigm.
- **Providing the VRCs with the necessary training to model their Golden Exemplar ROs (GE)** as compelling instruments for communication with the wider communities. The existence of such GEs is key in order to raise awareness within the communities, showing the benefits of research objects and the feasibility of their creation by actual VRC members (not software or knowledge engineers) with just some initial limited training. The hackathons provided the VRCs with the necessary knowledge to continue the GE modeling

work offline, back at their laboratories. The fact that the VRCs have incorporated research objects and the supporting technology gives evidence of the uptake of this paradigm obtained so far in the project.

These training sessions were crucial for the evolution of the project as the idea of putting Research Objects at the centre of the paradigm, first time in Earth Science, could have induced a resistance by final users unwilling to change the way they worked. It was therefore fundamental to have a first hackathon before the completion of the user requirement definition (D3.1) in order to ensure that some of the use cases were already reflecting a basic understanding and tentative adoption of RO's technology and concepts. The details of both events including the training objectives, material, agendas, feedbacks and discussions are reported below. It must be underlined that these face to face meetings were not the only occasion to discuss the adoption of RO's; a continuous discussion via email, virtual conferences and sharing of documents via the Alfresco internal repository took place for almost 18 months, until it was clear that all VRC's were fully capable to create, manage, modify, share Research Objects of different types, within and across their communities and working environments.

2.1.1 First Research Objects hackathon

Date and venue: January 2016 at CNR ISMAR – Bologna (IT).

The first hackathon, held in January 2016, started with a tutorial on scientific workflows and Research Objects, followed by a demonstration and tutorial of the Taverna scientific workflow management system. The participants were introduced to the Research Object lifecycle management platform ROHub.org and learned how they could use it to build Research Objects on top of the resources comprised in their own research.

Workshop objectives:

As previously mentioned, the entire consortium attended the first hackathon. Therefore, objectives were set for each of the different typologies of attendants. For Research Object experts, the main goal was to understand how the VRCs worked and to identify the gap between the RO tools which were available and the actual needs of the VRCs. For the VRCs the crucial point of the workshop was to understand what Research Objects can do for them, and how they can be used to improve the way they work. The approach used was "learn by doing", putting VRC's to work on some concrete examples to understand the potential of RO's.

Preliminary material:

To elicit a first set of requirements about Research Objects for the four project Virtual Research Communities, a comprehensive questionnaire of 14 questions (Appendix A – Introduction to RO's in Earth Science - Survey) was created, containing concrete questions regarding:

- The goals and objectives for the adoption of the research object paradigm by the community.
- The expected content of a research object and the kind of meta-information that these should have.
- The different user roles with respect to the creation, management and exploitation of ROs.
- The required system functionality per role.

More in detail, each VRC was requested to provide the description of a potential Research Object they could envisage being used in their working environment. As a minimum, the following input was requested:

- Main RO purpose, e.g.:
    - Share and access scientific data.
    - Share scientific methods and their implementations as e.g. software, data processing capabilities, executable simulation models.
    - Validate scientific findings by reproducing the implemented methods and observations.
    - Support novel forms of scholarly communications, e.g. executable papers.

- Explicit content of the RO, like specific datasets, software, documents.
- Metadata, e.g. creation date, author, dependencies between components, access rights, citations to related work.

Main topics discussed – taken from the hackathon agenda:

1. Introduction to RO's main concepts
2. Scientific workflows and research objects - Motivation and main concepts (ESI) 20'
3. Looking into a specific workflow management system: Taverna
    a. 101 tutorial (ESI) 30'
    b. Data & Service customization (ESI) 30'
4. Overview of ROHUB, main research object management functionalities (PSNC) 30'
5. Development with ROHUB of VRC's Research Object Supersites
    a. Use case presentation (All partners)
    b. RO construction (joint session, led by PSNC) 60'
6. Development with ROHUB of each VRC's selected Research Object

Training material and hands on exercises

The first hackathon training material was based on PowerPoint presentations, prepared by ESI and PSNC, covering RO concepts and technologies. All RO training documentation is available on the project Content Management System and has been regularly updated in order to reflect the evolution of the discussion between RO representatives and the Virtual Research Communities. These documents constitute the basis of the Research Objects training material which was also used during the second hackathon that took place in Bologna, July 2016. The training document is attached in Appendix B – Research Objects Training Material.

Feedback from first hackathon

The first hackathon was a key milestone for the internal training programme as the RO experts were exposed for the first time to the Earth Scientist, and vice-versa. This enabled both parties to learn from each another giving the opportunity to further understand the different working environments. The Virtual Research Communities demonstrated a significant level of interest in the RO technology and although it was challenging to fully grasp the RO concept after just one introductory session they had a clear picture of the potentiality that the RO concepts and technology represented.

2.1.2 Second Research Objects hackathon

Date and venue: July 2016 at CNR ISMAR – Bologna (IT).

The second hackathon that took place in Bologna relied on a more solid understanding and advanced adoption of ROs by the communities. The VRCs had the opportunity to work together on their first Golden RO (see 2.1.2.1), wrapping their scientific workflows into Research Objects to be inserted, executed and shared via the VRE.

Workshop objectives:

- Finalize the VRC's Golden Exemplar ROs
- Agree on a sustainable protocol for large-scale RO generation
- Discuss the three fundamental use cases for RO utilization by VRCs and the tools to support it
- Search and recommendation
- Long-term preservation

- Publication
- Introduce new supporting GUIs and tools

2.1.2.1 VRC golden exemplar ROs

The first main objective of the second EVER-EST hackathon was the definition of the so-called Research Object Golden Exemplar (GE) which has proven to be powerful instruments for adoption and dissemination both within the VRCs and amongst the wider community. This core activity demonstrated the degree of understanding that the VRC developed with respect to the potentials of ROs in the Earth Science domain. The following GE ROs were developed by the four VRCs:

Sea Monitoring Virtual Research Community (CNR-ISMAR)

- Deep Sea Habitat **Sustainability** Model**:** In this research object the MSFD indicator 1.5 was derived (Habitat area) to assess the biological diversity descriptor. To do this in deep sea environment, the scientist (user) needs to implement a habitat suitability model.
- The Citizen science and jellyfish distribution**:** A crowdsourcing app sponsored by Italian magazine and other different media provides scientific data to study jellyfish. CNR-ISMAR wants to fully exploit within the EVER-EST initiative the potential of the app to generate meaningful indicators in MSF perspective.
- Trend in the evolution of invasive jellyfish distribution**:** Starting from Jellyfish sightings, data is elaborated to produce explicit geographical information concerning trends about the evolution and distribution of alien species according with MSF directive descriptors.

Natural Hazards Virtual Research Communities (NERC)

- Hazard Impact Model Development**:** Research object to facilitate development of surface water flooding early warning systems and their impacts within the UK.

Land Monitoring Virtual Research Community (SatCen)

- Land monitoring Golden Exemplar**:** Research object for the ingestion of satellite images acquired on land areas (with the support of information coming from social sensing sources and other geotagged information) to automatically detect changes and send an alarm to the user.

GeoHazards Supersites Virtual Research Community (INGV)

- Volcano Source Modelling (VSM) - Application to Campi Flegrei (Italy)**:** In this workflow-centric research object the workflow is used to model ground deformation from satellite InSAR (Interferometric Synthetic Aperture Radar) data observed at Campi Flegrei (Italy) during 2011-2013 (up to 15 cm uplift), in order to derive the magmatic source characteristics. The research object is based on the Volcano Source Modelling (VSM) software tool by INGV.
- IPWV map generation**:** This research object contains the workflow which allows obtaining an integrated map of the perceptible water content over the Etna supersite, by using satellite and GPS data. Moreover, the research object produces ascii files containing IPWV values at a resolution of 0.2 deg.
- 2013 Mount Etna Eruption (bibliographic Search**):** This is a bibliographic research object, which supports search of bibliographic repositories, e.g. Google Scholar, ISI Web of Science and Earth-Prints, for publications about the 2013 eruption on Mount Etna. In 2013, 19 paroxysmal, mainly strombolian, eruptions occurred causing several environmental effects, which are the subject of ongoing scientific investigations. This research object can be reused to automatically update the list of scientific literatures on this eruptive phase.

2.1.2.2 Training material and hands on exercises

The material used during the second hackathon related to RO creation and management via ROHub was based upon the input collected during the first RO hackathon. Since then the material is available for external training, directly accessible from the VRC help guide.

The RO training document provides a general guide for creating new Research Objects from scratch in ROHub. It also provides a series of questions at each step of the RO creation procedure and at the end of the document, to collect the feedback from the VRCs and tune the future developments of the ROHub. The entire document is available in Appendix B – Research Objects Training Material.

2.2 EVER-EST technical training

All those training activities that are not focused on Research Objects but related to Earth Science specific e-infrastructure and e-research service fall within the technical training category. Two distinct training session related to Earth Science services and technologies were provided to all the communities and project members, the first in November 2016 and the last one in February 2017. At that stage, the EVER-EST Earth Science services were gaining a level of maturity and usability that allowed the technical teams to organise demos and training sessions to facilitate their adoption by the communities. As for the Research Objects training sessions, a series of virtual one to one or general meetings where held before and after these main consortium training events to focus on specific aspects, or to assess the comprehension of services and functionalities by the different communities. Informal training-like discussions were also held during two of the project All Hands Meetings (Edinburgh and Poznan) to first introduce and then discuss with the communities the new EVER-EST services.

2.2.1 Common Services training

These services are quite low-level and don't have much direct impact on the VRC operational activities. In this sense, descriptive slides on the technologies involved in all Common Services were prepared and discussed with the communities, covering the following topics:

Identity Server

- Description of the authentication mechanisms
- Description of User Management + How To subscribe, authenticate, manage user account
- Description of security management

Enterprise Service Bus

- Description of Communications Adaptation
- Description of Information Transformation Enhancement (Role of Middleware in general)

Data Analytics

- Description of Data logging in the light of KPI measurement

Data Discovery and Access

- Description of discovery functionality within a VRE (what to discover and why) + How To Discover and Access EO data
- How To Discover RO (TBC, probably already covered by RO-Services training).

All training material used for these different slots is available on the project's Content Management System and used, when deemed necessary, for dedicated and public events.

2.2.2 Web Processing Service training

In EVER-EST, the WPS (Web Processing Service) functionality allows invoking a remote geospatial processing service as a Web Service. This geospatial processing service is provided by T2's Big Data processing infrastructure and it is invoked by the EVER-EST VRE user interface, passing the inputs to the processing chain and returning an output.

WPS was firstly adopted for the SatCen use case, where the processing service is based on the Change Detection process, which allows to select a pair of images from the Sentinel-1 archive, within a user-defined timeframe, and to identify clustered changes. The service launches a set of chained processing modules based on SNAP[1] (i.e. Subset, Orbit Correction, Thermal Noise Removal, Calibration and Terrain Correction). Successively, an in-house Change Detection algorithm identifies the areas where possible changes occurred. The output of the Change Detection is a raster product containing the pixels where possible changes have been detected.

It is worth mentioning that the CNR community has developed a cross-fertilization study in junction with SatCEN starting from a discussion during this training session.

WPS – Hands on training example

The required input for the processing are:

- An Area Of Interest (AoI), called *Polygon 1*;
- Two Sentinel 1 images, called *Master* and *Slave*.

The AoI is defined using the "Draw Option/Draw Polygon" icon and the images have been previously selected with the "Data Discovery" service. It is advisable to select two images acquired under the same interferometric conditions (i.e. same acquisition parameters).

To launch the service in the main interface, the user clicks the icon "WPS Manager" on the sidebar (Fig.1).

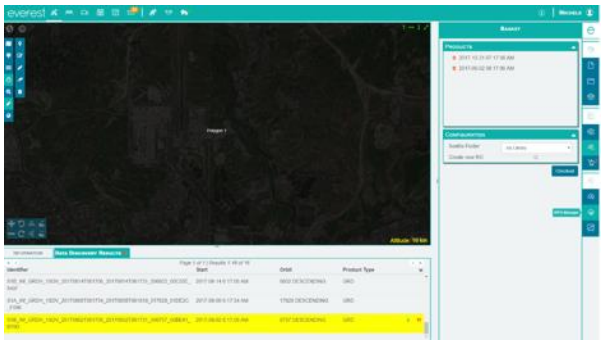

Figure 1 WPS Manager launch

Once the WPS Manager panel is opened (Figure 2), the user should select the input images in the *Master* and *Slave*[2] sections of the *Inputs* window (the images on which Change Detection will be computed). In the *Polygon* window, the AoI is automatically set.

[1] http://step.esa.int/main/toolboxes/snap/

[2] In the current implementation, the order of the two images is not relevant.

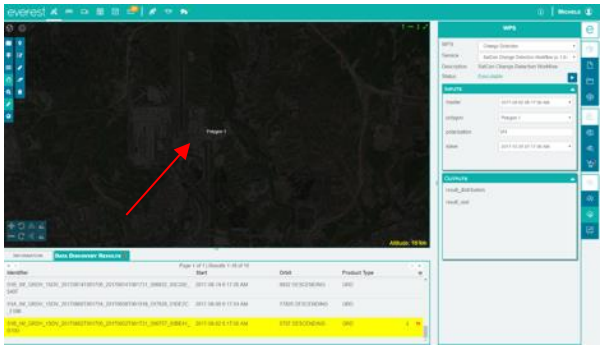

Figure 2 WPS Menu

When the input images are set, the user can launch the WPS by clicking on the play button next to “Executable”. Information on the processing status (percentage of completed steps) is shown to the user on the WPS Manager Icon. When the WPS execution ends, the interface directly zooms on the results (to better visualize the results, it is advisable to uncheck some layers from the “Manage Layer” icon on the left menu (Figure 3).

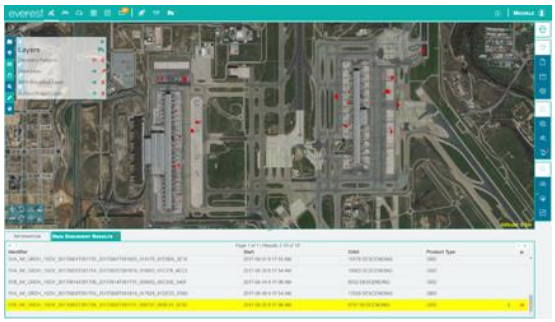

Figure 3 WPS Results Visualization

The change detection output, is visualized on the map, with changed areas marked in red. The maps are also saved as *. PNGW, and together with selected input metadata (e.g. original name of the images, period) they can be stored in a RO.

Processing Services: The Sandbox

Steps for the Sandbox training (taken from D2.4)

1. Introduction to the cloud platform: what is the Terradue cloud platform, its mission, main services, portal and dashboard.
2. Building applications on the cloud platform: hello world example, what is the workflow involved, working with a developer cloud sandbox, the workflow descriptor, running an executable.
3. Delivering processes to users: testing the workflow, sharing processing jobs, starting the activities.
4. How to use the cloud platform provided by EVER-EST: inserting a job into an individual workflow management system.

A specific set of on-line exercises are available to the communities for hands-on training.

These exercises include:

- Hands-on Exercise 1: building a basic workflow
- Hands-on Exercise 2: make a robust workflow and debug it
- Hands-on Exercise 3: staging data
- Hands-on Exercise 4: using a toolbox
- Hands-on Exercise 5: using parameters
- Hands-on Exercise 6: a multi-node workflow
- Hands-on Exercise 7: debug a multi-node workflow
- Hands-on Exercise 8: browse published results
- Hands-on Exercise 9: using an OpenSearch catalogue
- Hands-on Exercise 10: prepare an OGC processing service

EVER-EST User Interface

It is difficult to define a training activity on User Interface design. GUI layout, logic and functionalities are the result of a continuous joint activity of development, deployment, discussion and adaptation, based on Agile developing standards, involving all partners in the project. Therefore, no specific internal training has been performed for this specific topic as the GUI is the result of a wider shared approach.

Nevertheless, following the comments provided by the communities during GUI early adoption, an online help describing the user interface functionalities has been directly integrated to guide users during the first steps and ensure proper understanding of the platform and its sub components. The “Take a Tour” popup functionality is available for all users and guests. A couple of steps of the Take a Tour are shown in the following figures.

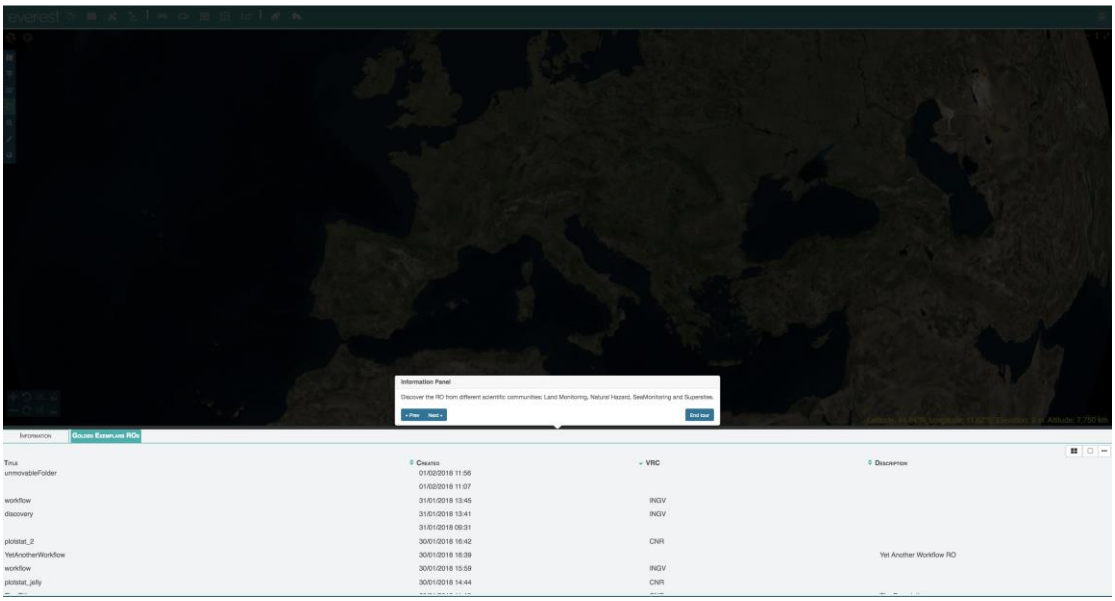

Figure 4 A “Take a tour” step to describe GUI functionalities

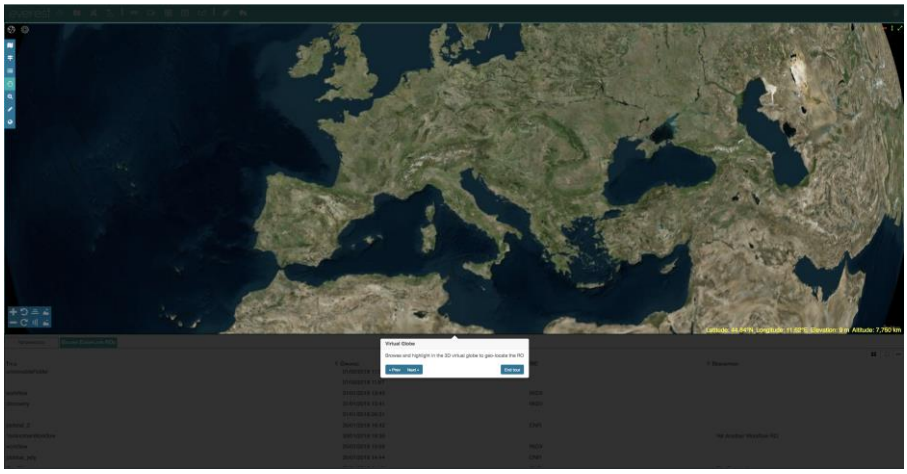

Figure 5 A "Take a tour" step to describe GUI functionalities

2.3 User Actions Training

The User Actions training can be considered as one of the project's training pillars; It has contributed to the creation of the EVER-EST community and is at the heart of the mutual understanding that has developed between the Research Object experts, the Earth Science technical team and the Virtual Research Communities.

Based on the template provided as an example in Appendix D (Appendix D – User Actions Template Document), during the All Hands-On meeting in Venice each VRC was requested to bring an example scenario where their daily working practices and use case were divided into small steps. They were requested to clearly identify each member of the team involved and their role, the goal of the scenario, the detailed sequence of micro-steps- called User Actions that are foreseen by the scenario and the compliance of each User Action with the requirements described in D3.1.

The goal of this training exercise was to enter into the everyday routine of each VRC with a practical working example and understand together how the EVER-EST technology could enhance and improve it.

The key elements of the discussions can be summarised as follow:

- Determine whether a processing step within a scenario, obtained by using COTS, could be fragmented into smaller steps to be completely handled by researchers via workflows;
- Understand which parts of the working scenario can be fully automatized via workflows and where the presence of a person taking decisions is mandatory;
- Identify the type of decisions that researchers normally take with respect to research intermediate results;
- Understand team dynamics to facilitate the adoption of Live Research Objects to support the team collaborative working;
- Identify which are the Common and E-Collaboration services which are de-facto the priority within each scenario.

Having the sequence of user actions turned out to be a sort of "Rosetta stone" within the consortium as it allowed a shared understanding of the same actions starting from two different perspectives; the VRC team necessity on one side and technological solution and improvement by software developers on the other. It is a crucial element for training activities as it solves issues such as different dictionaries and action interpretation, that could create conceptual gaps during future training events. This is why, as already mentioned, after the Venice meeting each VRC was asked to deploy all its user requirements listed in D3.1 into user actions documents, that were discussed during on-line training sessions during the next three months.

The User Actions document constitutes a central element for future training activities as it can dramatically facilitate the dialogue between the EVER-EST consortium team members and the external communities.

It is worth reminding that an additional list of “Generic Earth Science User Actions” was built on top of this exercise and was a fundamental element for the discussion of the User Interface design as it highlighted those functionalities which are needed by any Earth Science community and must therefore be available and easily identified during the User Experience.

3 External Training

No major activities for external training were planned during the first two years of the project. All training activities were focused around the four EVER-EST VRCs however with the introduction of additional stakeholders within the current VRCs new training opportunities were identified. This is the case of CNR ISMAR who created internal training events with the support of the EVER-EST team for UNISALENTO Lecce (also part of CNR ISMAR). Similarly, INGV performed dedicated training sessions for the University of Iceland in the context of the Supersites initiative. The result of these activities, and the main asset used after M24 for external training is the EVER-EST user manual, which is attached at the end of the document (see Appendix C – EVER-EST User Manual).

Nevertheless, additional external training sessions took place during the reported period. A master thesis on Data Management with the University of Tor Vergata opened up the opportunity for a training session with early adopters and INGV performed training sessions to a completely new Earth Science community.

3.1 Master with University of Tor Vergata – Activities and related training

The University of Tor Vergata (UTV) organizes every year a Master Program in Data Science. In this context, a thesis has been proposed in joint collaboration with the European Space Agency entitled "Data Scientist with EVER-EST platform".

The main objectives of the thesis were

Propose and consolidate a collaboration between ESA and University of Tor Vergata, based on the EVER-EST infrastructure as a reference platform to be used during the overall Master Programme in Data Science.

Use the EVER-EST platform to demonstrate the added value for data scientists of the Research Object paradigm, which represents the logical kernel of the e-infrastructure. Conduct a study on disciplinary cross-fertilization possibilities between Research Objects created during the master and Research Objects which have been previously inserted in the VRE by one of the VRC.

Collaboration between ESA and University

The EVER-EST infrastructure allows access to different Earth Science catalogues with a specific focus on Earth Observation data and provides computing processing capabilities. Following this approach, a new application can be integrated using "virtual machines" that have their own specifications (disk size, processor speed, operating system etc.) and run on shared private (physical deployment over local hardware) or commercial cloud infrastructures. This kind of approach is pivotal in allowing the adaptation and customization which are specifically needed for the Master activities. Currently, master students face the need to complete their work despite the lack of adequate technologies, infrastructures and applications. Therefore EVER-EST can be configured as a series of Virtual Machines responding to Data Science requirements in terms of installed toolkit and requested performances.

Training activities: steps in the creation of the Research Object for the Master

A Research Object has been created using Earth Observation data which is already available in EVER-EST. Datasets can be accessed via the platform via catalogue but also as a Datacube allowing to choose a pixel or a polygon on a map with the result to obtain a relevant space data time series. The case study, i.e. the Research Object under construction, is related to the "Mediterranean Sea Anomalies detection". This included information on sea surface temperatures, chlorophyll, dissolved and detrital organic matter, particulate backscattering coefficient and sea surface height. The initial analysis on the associated measures and variables is carried out in R or Tableau tools for data visualization, which can be deployed on the platform when needed. The obtained statistics and charts can be loaded and executed via the Taverna workbench, another tool available in EVER-EST, and can shared within the community. This information can therefore be reused and reproduced in different working environments by other users. The resulting Research Object is validated in terms of content of attributes compiled and a DOI for the

citation is applied. The final representation of the results of the resulting Research Object is the map of the detected anomalies on the Mediterranean Sea.

Further steps: Cross-fertilization between UTV and EVER-EST VRC

One of the main feature of EVER-EST is to provide Data Scientists the possibility to use already validated RO and/or create a cross validation between ROs and between Virtual Research Environments. These characteristics have been used for a cross-fertilization study in synergy with CNR and UniSalento biological researchers group. The group is specialized on the quantification of deterministic and stochastic components of environmental change that lead to outbreaks of maritime species, in this specific case, the jellyfish. The Research Objects created by UniSalento have been cross-fertilized with the RO on "Mediterranean Sea Anomalies detection" developed during the Master. This can be considered as a good example of joint work between two communities, Earth Observation researchers and Maritime Biologist, which could be not necessarily strictly linked in their everyday activities and that was de facto facilitated by the common use of RO's and the adoption of the EVER-EST infrastructure as working environment. The analysis led to the successful identification of correlations between the two phenomena over specific areas of the Adriatic Sea. The final result has been graphically represented using an EVER-EST GIS tool overlapping all information produced by both studies.

3.2 Activities at INGV

3.2.1 Iceland community training

Training event agenda (11th October 2017)

Time	Owner	Topic
09:00 – 09:10	S. Salvi	Intro to EVER-EST
09:10 – 09:30	E. Trasatti	VRE and Research Object concept
09:30 – 10:30	E. Trasatti V. Romaniello	Live DEMO Volcano Source Modeling (VSM) Water Vapor estimation (IPWV)
10:30-11:15	S. Corradini D. Stelitano	Live DEMO Ash and SO_2 retrieval by Volcanic Plume Removal (VPR)
Coffee Break		
11:30-12:15	All	Plan of the activities – All the audience or in separate groups
Lunch		
13:15-14:00	All	Plan of the activities, continued

Table 1: Iceland community training agenda

3.3 Activities at CNR

Training framework: case study of the Lagoon of Venice (CNR-ISMAR Venice)

The Lagoon of Venice is the biggest lagoon in the Mediterranean Sea with a surface of about 550 km2 and with an average depth of about 1 m. In the last century, the morphological and ecological properties of the lagoon changed dramatically, the surface of the salt marshes was reduced by 60% and some parts of the lagoon are deepening with a net sediment flux exiting from the inlets.

To understand and monitor the future evolution of the Lagoon of Venice, ISMAR within the project RITMARE (a National Research Programme funded by the Italian Ministry of University and Research) carried out an extensive survey, involving a team of more than 25 scientists, to collect high resolution (0.5 m) bathymetry of key study areas such as the tidal inlets and channels. These data are stored in a public repository and are freely available to be used with a CC-BY license agreement (Madricardo et al., 2017. High resolution multibeam and hydrodynamic datasets of tidal channels and inlets of the Venice Lagoon. Sci. Data 4:170121 doi: 10.1038/sdata.2017.121 https://www.nature.com/articles/sdata2017121). Following a broad multidisciplinary approach, bathymetric and backscatter intensity data are now employed for geomorphologic studies, habitat mapping and modeling of evolution trends of this highly dynamical and complex transitional environment.

Bathymetric data for physical habitat mapping – Geomorphometry

ISMAR scientists developed a number of new methodologies to extract geomorphometric properties from the high-resolution bathymetry to be used to classify different physical habitat type within the channel of the lagoon of Venice: They also planned new field experiments to understand the evolution of the highly dynamical channels of the lagoon of Venice.

To this aim, they need to re-use raw bathymetric data and ADCP data, to correctly preserve the processed data (e.g. High resolution DTM), to ensure the reproducibility of the new methodologies and experiments in different shallow water areas in order to compare results, to share data, methods and documents, to be cited and acknowledged to other scientists. In their daily work, they also need big storage and computing capabilities.

The role of the EVER-EST VRE and related trainings

The ISMAR Venice research group was trained to access the Sea monitoring VRE with personal accounts and produce the following workflow ROs using all EVER-EST services and capabilities:

1) RO-ROUGHNESS: Evaluation of sea-floor roughness from high-resolution multibeam data and in parallel collected ADCD data using Virtual Machine to implement algorithms and to encapsulate the workflows into Taverna Workbench and EVER-EST Workflow runner;

2) RO-DUNES: Evaluation of subaqueous dune fields characteristics and migration velocities analyzing seafloor high-resolution DEMs through ARCGiS and Matlab routines included into Taverna Workbench and executable trough the WF runner service (FFT, automatic extraction of crests and troughs).

3) RO-VIRTUAL FIELD EXPERIMENT LIDO: in September 2014, a field experiment was carried out in the Lido Inlet, Venice, Italy in order to estimate the bedload transport in the inlet: we collected MBES data for 22 repeated surveys and, at the same time. ADCP data, water samples, sand traps, CTD casts. The RO will include the entire experiment and make the collected data and the full analysis procedures open and available to be reproducible and re-used through the EVER-EST VRE.

3.4 External training at AGU (by INGV)

Demonstration of the EVER-EST Virtual Research Environment at the American Geophysical Union Fall meeting, New Orleans. The demonstration was carried out from 5:00 to 6.30 pm on Thursday, December 14 at the Hilton Riverside hotel, Warwick room.

INGV invited the community of the Geohazard Supersite network, including over 350 scientists from all over the world. Over 20 people had registered online for the event during the previous month, but 12 of them actually attended. This was later justified by conflicts with simultaneous scientific sessions. There were three presenters from INGV: E. Trasatti, S. Salvi and C. Tolomei.

This is the list of attendees:

Name	**Organization**
Mike Poland	USGS – CVO
Eric Fielding	NASA – JPL – ARIA Project
Marco Moro	INGV – CNT
Giuseppe Pezzo	INGV – CNT
Tim Stough	NASA Applied Science
Giuseppe Solaro	IREA – CNR
Francesco Casu	IREA – CNR
Claudio De Luca	IREA – CNR
Gerald Bawden	NASA – HQ
Charles Meertens	UNAVCO
Nikos Svigkas	INGV – CNT
Christian Bignami	INGV – CNT
Enrico Serpelloni	INGV – Bologna

Table 2: AGU training session attendees

The session started with a presentation by S. Salvi and E. Trasatti (Research Object available at: http://sandbox.rohub.org/rodl/ROs/EVER_EST_demonstration_AGU_Fall_meeting_2017/), in which a description of the context was given, introducing the Research Object concept, its benefits for the community, and the concept of Virtual Research Environment.

Then two use cases were demonstrated: 1) a Geodetic data inversion to model a magmatic source on the Etna volcano, and 2) the InSAR processing of Sentinel-1 data for an earthquake in Iran. Use case one was demonstrated by E. Trasatti, while use case two was demonstrated by C. Tolomei, while S. Salvi added comments during the demonstration and replied to questions from the audience.

The re-use of an existing RO generated by another scientist was demonstrated in the first use case, while the second use case demonstrated the capacities of the VRE to access Sentinel 1 data and process them remotely in a powerful processing environment. Following these use cases, it was also shown how to run a workflow RO using the Taverna Server application embedded in the VRE.

During the session, there were several questions from the audience, mostly concerning the RO concept and their capabilities. The capacity of ROs to document research work in a structured way was much appreciated, as well as the possibility to share scientific results while protecting the IPRs of the researcher by obtaining a DOI for the Research Object. Some attendees asked if it was possible to use the VRE and how. The INGV team explained that as soon as the user manual was finished they were to be contacted and authorized to register on the VRE, then they could be given access to the Virtual Machines of their choice.

Appendix A – Introduction to RO's in Earth Science - Survey

Q1: What is the main objective which, in your opinion, a Virtual Research Environment based on research objects should contribute to in your community? Which one is the primary one?		
Objective	**Applicable? (Yes/No)**	**Primary? (Yes/No)**
Share and access scientific data and knowledge		
Share and use computational resources (software, workflows, etc.) and data processing capabilities		
Validate scientific findings by reproducing in-silico experiments and their executions		
Contribute to new forms of scholarly communications, e.g. executable papers		
Other	(Please describe)	

Table 3: Research object goals questions

Q2: If in your community you publish/share data processing methods, either individually or as part of experiments, what are the typical forms these methods take?		
Method implementation form	**Applicable? (Yes/No)**	**Comments**
Source code / script (Python, R, etc)		
Executable system / Web service		
Scientific workflow (Taverna, Pegasus, Wings, Kepler, Triana, Galaxy, etc.)		
Spreadsheets (Excel)		
Other	(Please describe)	

Table 4: Research object computational resources questions

Q3: For a given research object (data, method, etc.) that will be available on the VRE, which of the information below you want to know:		
	Applicable? (Yes/No)	**Comments**
When it was created		
Where it was created		

Who created it		
Who was the last to modify it		
What did those modifications imply		
Which purpose(s) it has been conceived and used for		
What it contains		
How is its content structured		
Who can access it and for what purpose		
What has been its evolution history		
To what extent its content can be trusted as being valid		
To whom each of its constituent parts belongs, e.g. data, web services, etc.		
It is complete or, on the contrary, constantly revised		
Whether it still works or is it broken/deprecated (for methods and experiments)		
Other	(Please describe)	

Table 5: Research object metadata questions

RO Management and Usage Requirements

Q4: Which of the following roles are applicable for the potential users of your community		
Role	**Applicable? (Yes/No)**	**Comments**
Creator: Scientist conducting an investigation who wishes to collect together resources as a research object that can then be reused		
Contributor: Scientist who provides materials/methods/data that may be used within a research object, but who is not necessarily creating a complete research object		

Collaborator: A collaborator is a scientist who provides materials/methods/data that may be used within a research object, but who may not even be aware of the fact that she is actually contributing content to a research object. Collaborators may be naïve in terms of their understanding or experience of the research object approach		
Curator: Scientist or support staff who makes sure the research object remains healthy and up to date. If it contains a workflow or other software implementing a scientific method, ensures it is executable. If datasets are involved, looks after their accessibility and availability		
Reader: The reader is looking for related works, state of the art, in her field of research, skimming the titles, keywords, themes and abstracts of the published research objects		
Comparator: Similar to reader, a researcher who compares their work, or planned work, with that of other researchers		
(Re-)User: Scientist who is looking for research objects as a basis for implementing her own methods in order to answer her research question		
Evaluator/Reviewer: a researcher who needs to evaluate/validate the work of a peer		
Publisher: The publisher wants her work and group to be known among the community, being the main author of the published research object. Alternatively, the person in charge of a scientific publication, receiving, processing, and delivering submissions from and amongst the community as scientific papers		
Librarian: Person in charge of custodying scientific knowledge, ensuring that the research objects remain accessible and looking after their preservation and conservation		

Other	(Please describe)

Table 6: Research object user roles Questions

Q5: As a Creator of a research object, which of the following actions would you like to perform:		
Action	**Applicable? (Yes/No)**	**Comments**
Create resources and aggregate them into research objects		
Describe the relationships between aggregated resources		
Reference external data		
Annotate resources and ROs using semantic models		
Assign a persistent id to a research object		
Be recognised as the creator of the RO		
Other	(Please describe)	

Table 7: Research object creator requirements questions

Q6: As a Contributor of a research object, which of the following actions would you like to perform:		
Action	**Applicable? (Yes/No)**	**Comments**
Provide content for a research object		
Modify the content of a research object		
Be credited for your contributions		
Annotate resources and ROs using semantic models		
Assign a persistent id to a research object		
Be recognised as the creator of the RO		
Have access to ROs being created by other users		
Other	(Please describe)	

Table 8: Research object contributor requirements questions

Q7: As a Reader of a research object, which of the following actions would you like to perform:		
Action	**Applicable? (Yes/No)**	**Comments**
Browse ROs		
Search for ROs		
Find ROs according to their reputation		
Find ROs by purpose		
Access the data and metadata of a research object		
Examine the relationships between resources		
Follow the steps taken in an experiment or process		
Other	(Please describe)	

Table 9: Research object reader requirements questions

Q8: As a Reviewer/Evaluator of a research object, which of the following actions would you like to perform:		
Action	**Applicable? (Yes/No)**	**Comments**
Access/examine data, metadata, content and resources		
Examine the relationships between resources		
Check if the external data has changed		
Follow the steps taken in an experiment or process		
Rate content		
Express opinion about the quality of external data sources		
Other	(Please describe)	

Table 10: Research object reviewer/evaluator requirements questions

Q9: As a Re-User of a research object, which of the following actions would you like to perform:		
Action	**Applicable? (Yes/No)**	**Comments**
Build a new RO based on an existing one		
Build a research object using components/parts of another one		
Run an existing workflow/process with new data		
Rerun parts of a workflow		
Use results or data from an existing experiment as input to a new one		
See versions of a research object		
Extract content		
Other	(Please describe)	

Table 11: Research object re-user requirements questions

Q10: As a Publisher of a research object, which of the following actions would you like to perform:		
Action	**Applicable? (Yes/No)**	**Comments**
Publish ROs		
Provide references to a research object		
Advertise a research object		
Restrict access to parts of a research object		
Other	(Please describe)	

Table 12: Research object publisher requirements questions

Q11: As a Collaborator of a research object, which of the following actions would you like to perform:		
Action	**Applicable? (Yes/No)**	**Comments**

Provide content		
Other	(Please describe)	

Table 13: Research object collaborator requirements questions

Q12: As a Curator of a research object, which of the following actions would you like to perform:		
Action	**Applicable? (Yes/No)**	**Comments**
Check and annotate data and methods submitted by scientists		
Help develop tools to monitor and ensure the quality of the research objects and their content		
Make sure experiments encapsulated in research objects remain executable, that data are accessible, etc.		
Other	(Please describe)	

Table 14: Research object curator requirements questions

Q13: As a Comparator of research objects, which of the following actions would you like to perform:		
Action	**Applicable? (Yes/No)**	**Comments**
Compare a research object with others		
Select similar exemplars		
Look for peers working on similar topics		
Other	(Please describe)	

Table 15: Research object comparator requirements questions

Q14: As a Librarian of research objects, which of the following actions would you like to perform:		
Action	**Applicable? (Yes/No)**	**Comments**
Cataloguing research objects, annotating them with information useful for scientists accessing them, classification and categorization in verticals, etc.		

Distinguish between scholarly (i.e. made by scholars, reputable, peer-reviewed, cited, credited) and not-scholarly material on the web and in the repository		
Find scholarly material in the research object repository		
Make sure research objects are linked to and accessible from the main thematic (trans) national repositories		
Other	(Please describe)	

Table 16: Research object librarian requirements questions

Appendix B – Research Objects Training Material[3]

ROHub prototype live demo is available at:

http://sandbox.wf4ever-project.org/portal (also available via http://www.rohub.org/)

Pre-requisite:

- To have an OpenID account with a provider using version 1.x.
  - If you don't have, you can create an account in one of the sites supporting this version, e.g.,:
    - http://www.livejournal.com/ (e.g., rapw3k.livejournal.com) nice identifier (same as account)
    - https://www.yahoo.com/ (e.g., me.yahoo.com) long cryptic identifier (e.g., https://me.yahoo.com/a/cPNpaOo40eTDI9vLPxnedsPtiA--#26c9b)
  - The resources associated to the RO have been gathered

Tasks

1) Log-in to ROHub

a) Click "sign-in" on the top right corner (see Figure 6)

b) Enter your OpenID (see Figure 7) - Note: the google link is not working as google currently uses v2)

c) Question:

i) Would you recommend something for improving the user experience for this task?

[3] Please note all pictures in the manual refer to the old RO-HUB user interface: at M24 a new version of RO-HUB has been released and the user manual is under upgrade to reflect the new graphics.

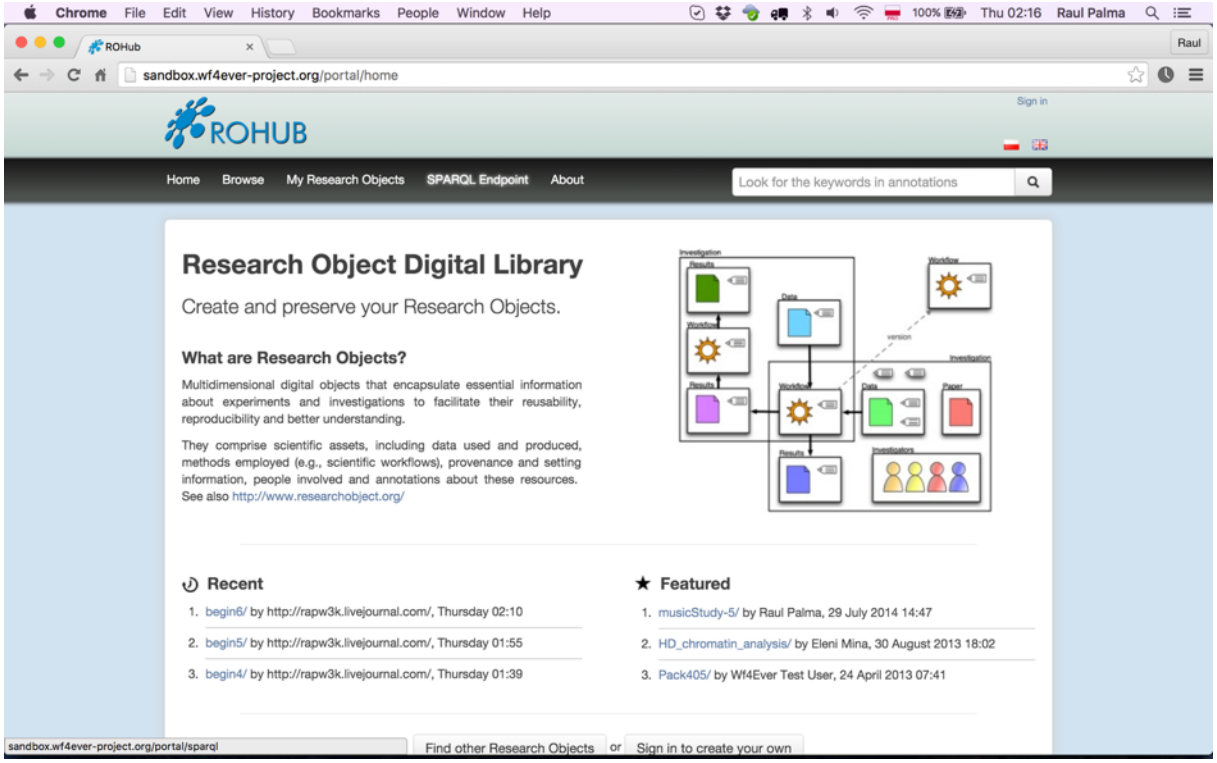

Figure 6 ROHub main page

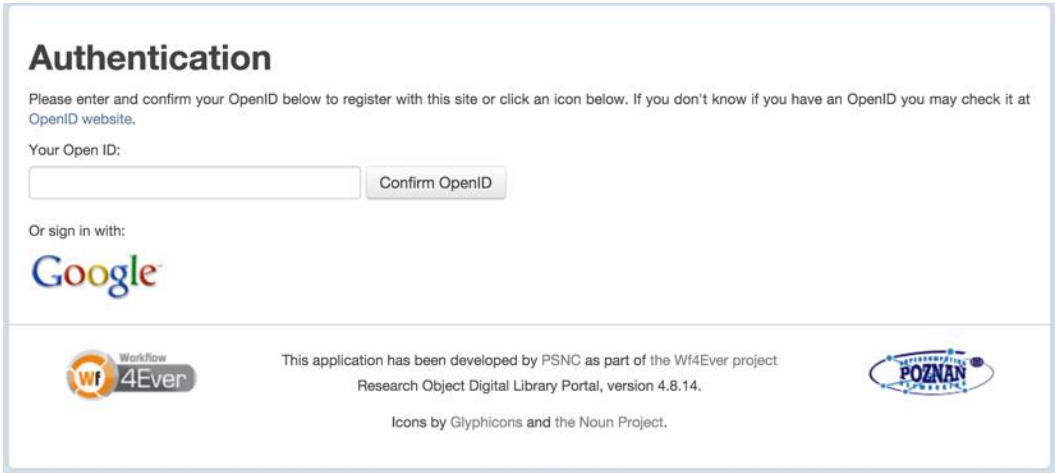

Figure 7 log-in page

2) Create a new RO from scratch
   a) Click "My Research Objects" tab (see Figure 6)
   b) Click "+Create" button (see Figure 8)
   c) In the new RO window (see Figure 9), provide:
      i) RO identifier: some good identifier with no spaces, case sensitive
      ii) Do not use templates[4] at the moment (see step 3)

[4] Templates enable to create a predefined structure of folders in the research object, so that it will be easier to organize resources

iii) (optionally) RO title: a short title of the RO (spaces are allowed) like the purpose of the RO (e.g., supersite validation RO). This can be added later at any time.
iv) (optionally) RO description: longer description of the RO. This can be added later at any time.

d) Question:
   i) Would you request other metadata at the creation time?
   ii) Would you recommend something for improving the user experience for this task?

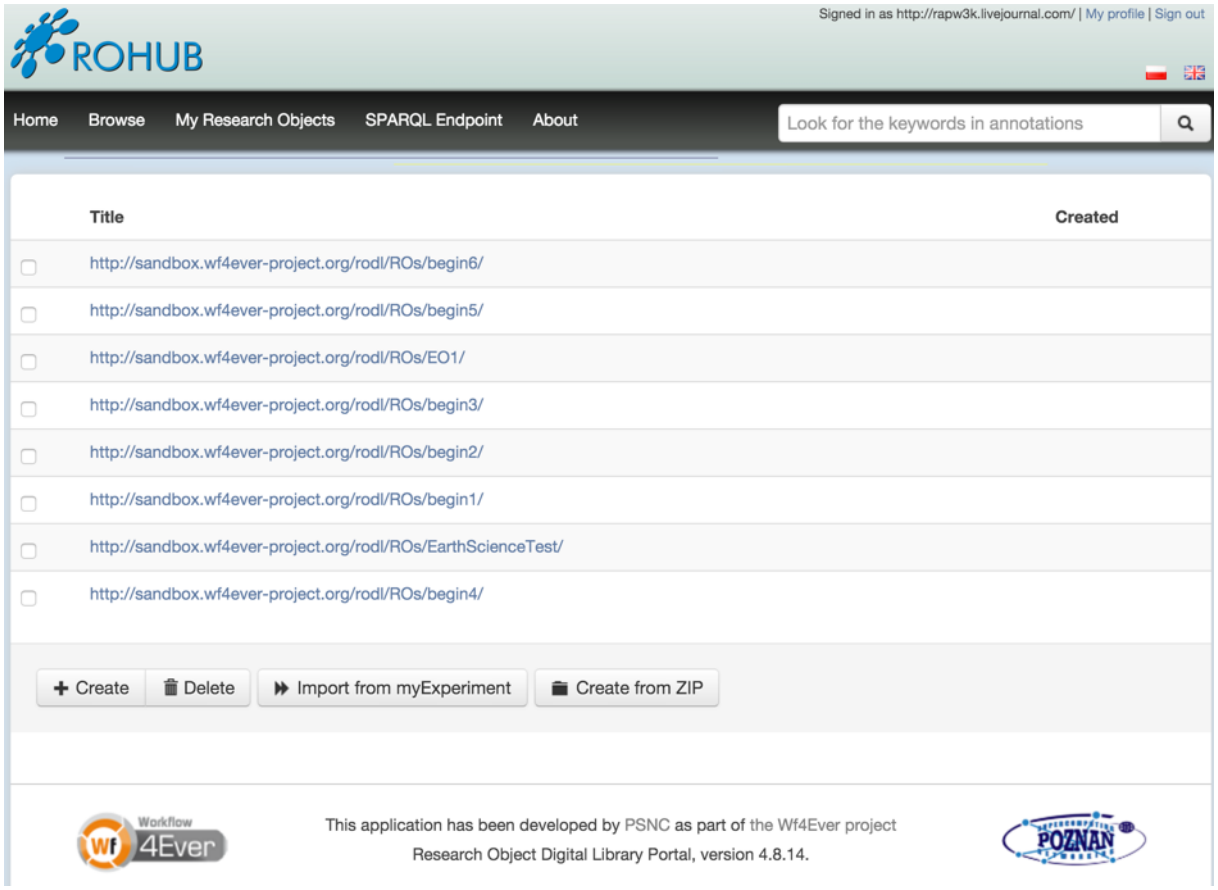

Figure 8 My Research Objects tab

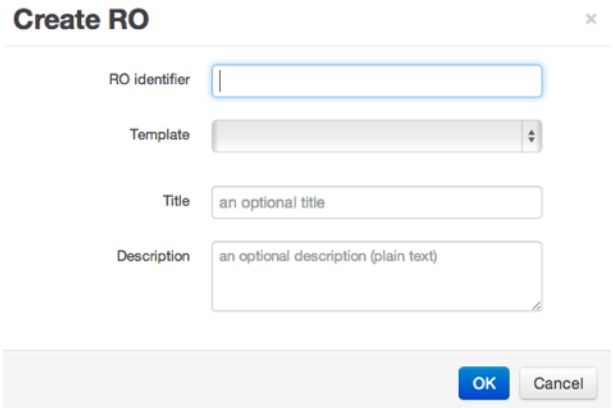

Figure 9 New RO window

3) Open an RO (your new RO)

a) Question:
   i) How much time in seconds is acceptable for opening a RO?

4) Create the RO structure
   a) Click the “Content” tab and click “+Folder” button to create the folder structure you like (see Figure 10)
   b) Enter the Folder name. Tip: as in your file system, try to use short, descriptive names
   c) Question:
      i) Is a folder structure useful for you? Why?
      ii) Would you recommend something for improving the user experience for this task?

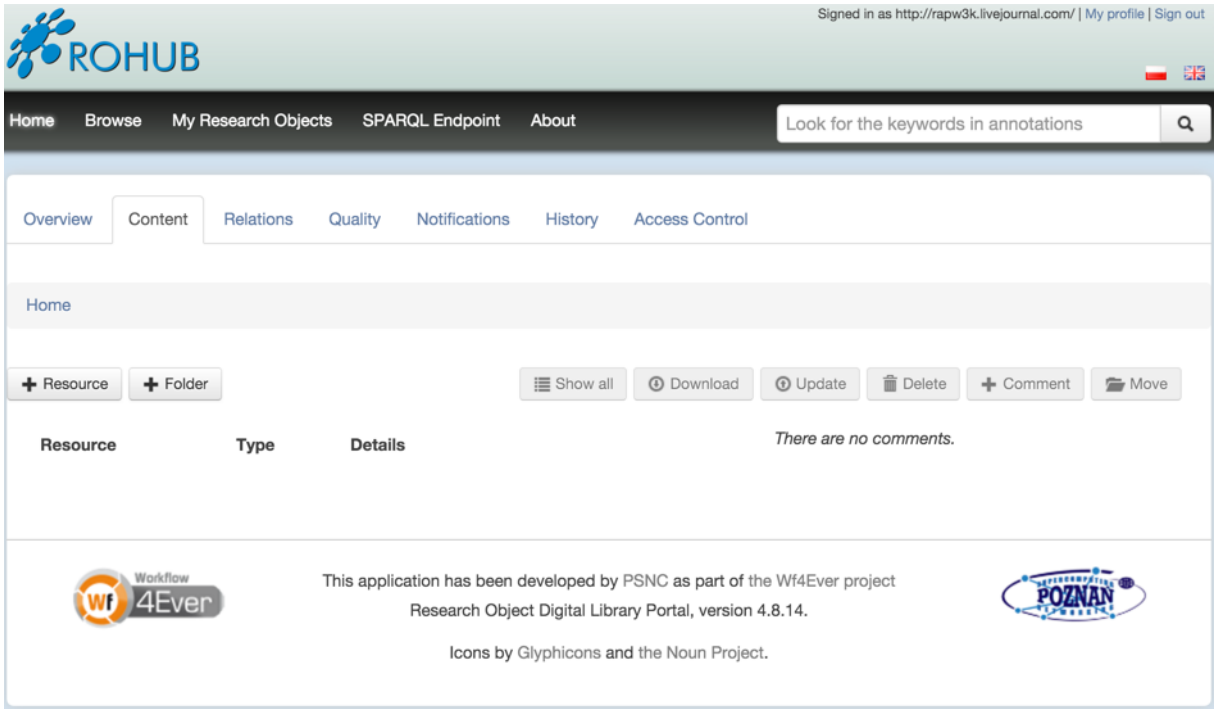

Figure 10 RO Content tab

5) Populate the RO
   a) Click the “Content” tab and click “+Resource” to add resources in the desired folder (see Figure 10).
   b) In the “Upload a resource” window (see Figure 11) provide:
      i) Resource kind-> Resource to be added can be
         (1) Local – to upload the physical resource in rohub
         (2) Web resource – to aggregate an external resource, by adding its URL
      ii) (Optionally) select the resource type - this can be added later at any time
      iii) Question:
         (1) Are there missing some important resource types? (list them)
         (2) Would you recommend something for improving the user experience for this task?

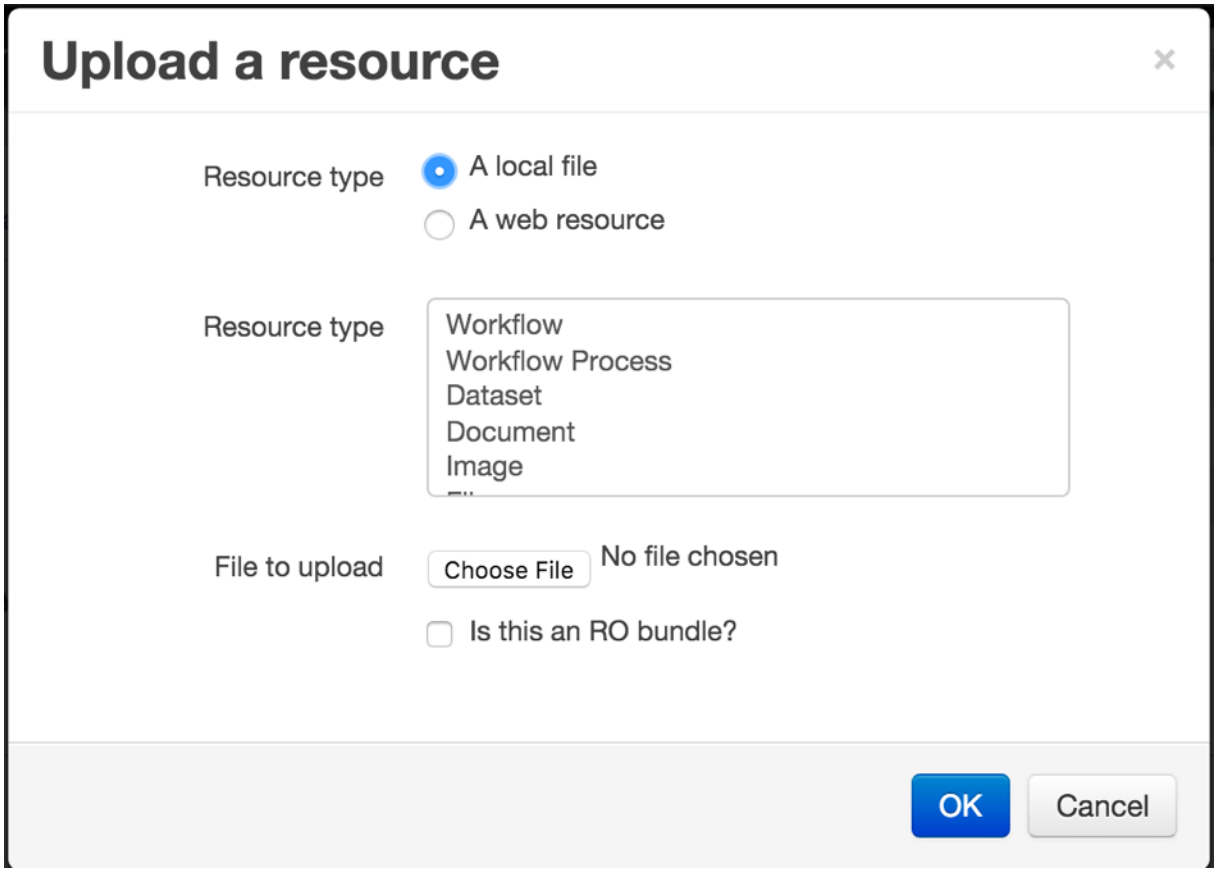

Figure 11 Upload a resource window

6) Annotate the RO and its aggregated resources (attach metadata)

a) For the RO

i) click the "Overview" tab

ii) (optionally) add/edit title

iii) (optionally) add/edit description

iv) click "Advanced view" to enter advanced mode and see all the annotations associated to the RO (see Figure 12)

(1) click "Annotate" and potentially add some metadata (e.g., license and add URL of license (e.g., https://opensource.org/licenses/MIT))

v) Question:

(1) Would you recommend something for improving the user experience for this task?

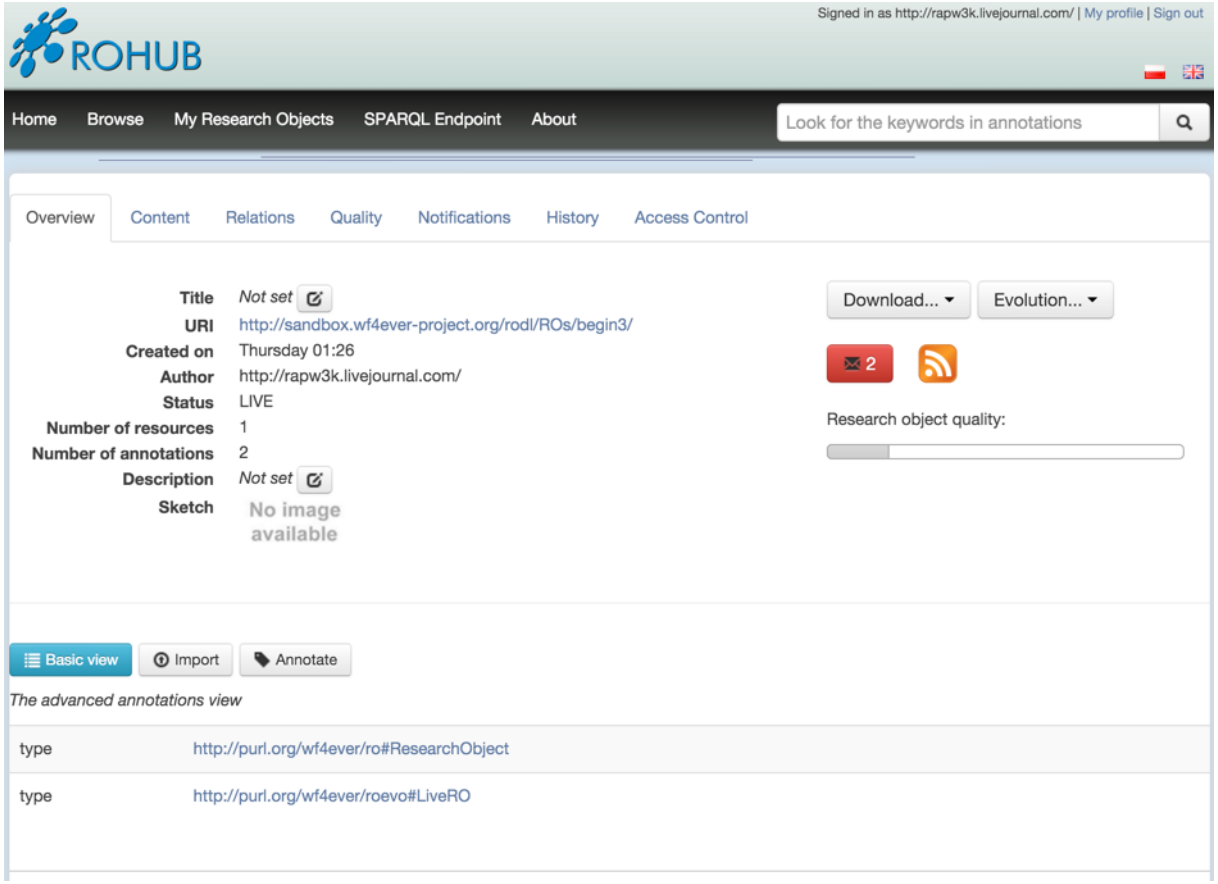

Figure 12 RO annotations - advanced view

b) For an aggregated resource
   i) click the “Content” tab
   ii) select an aggregated resource (see Figure 13)
   iii) (optionally) add/edit title
   iv) (optionally) add/edit description
   v) (optionally) add/edit type
   vi) click “Advanced view” to enter advanced mode and see all the annotations associated to the resource.
      (1) click “Annotate” and potentially add some metadata (e.g., license and add URL of license (e.g., https://opensource.org/licenses/MIT))
   vii) Question:
      (1) Would you recommend something for improving the user experience for this task

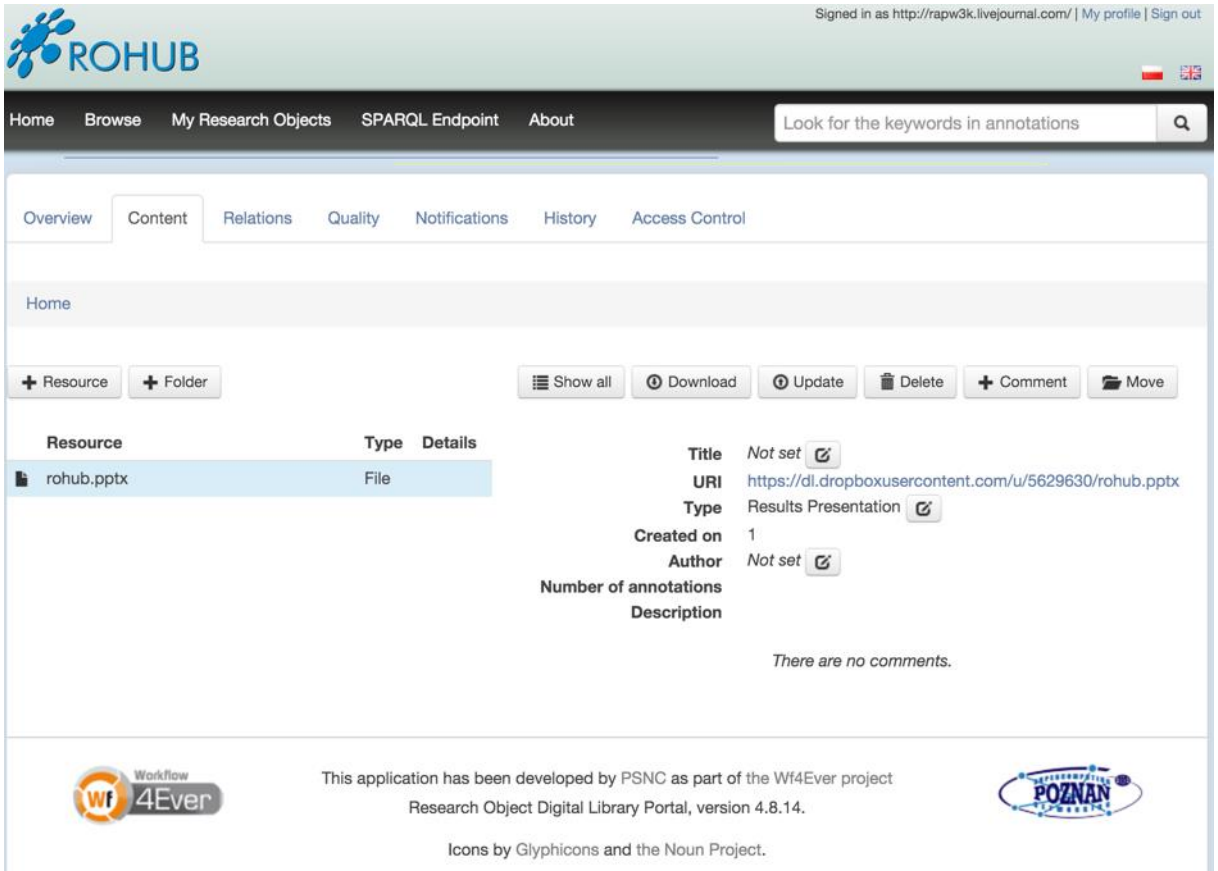

Figure 13 RO content - aggregated resource selection

7) Add relations in the RO
   a) Click “Relations” tab (see Figure 14)
   b) Select the subject of the relation (RO or aggregated resource)
   c) Select a property from the pre-defined list
   d) Select the object of the relation (RO or aggregated resource or URI of external resource)
   e) Question:
      i) Are there relevant properties missing? (list them)
      ii) Would you recommend something for improving the user experience for this task?

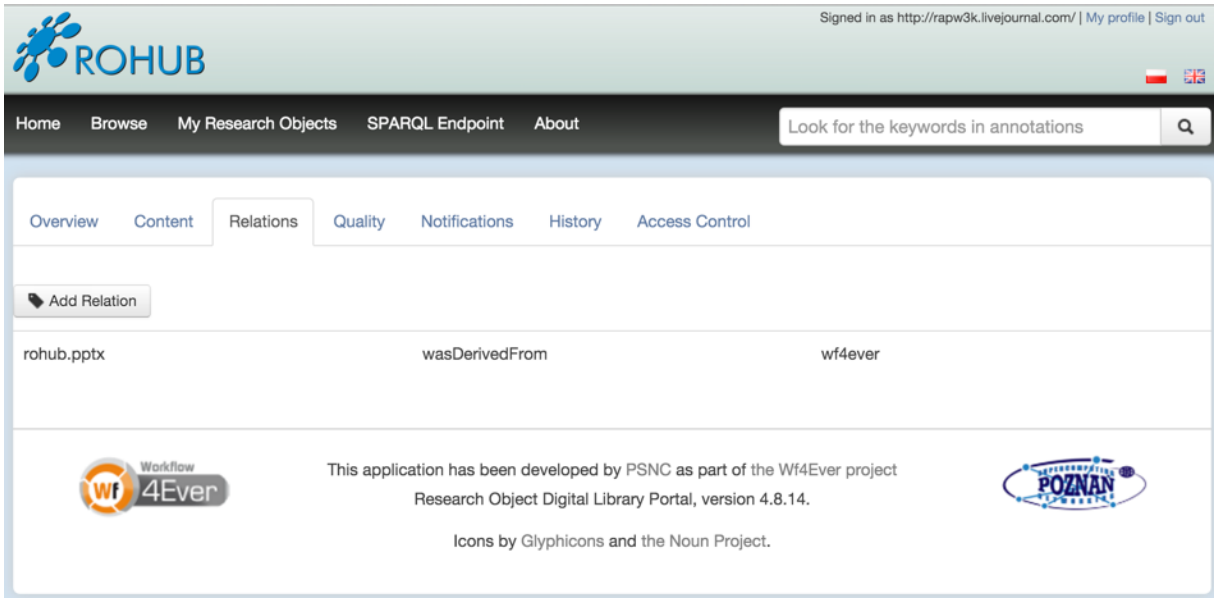

Figure 14 RO relations tab

8) Comment the RO and its aggregated resources

a) For the RO

i) click the “Overview” tab (see Figure 15)

ii) click the “+Comment” button (if not visible return to basic view)

iii) Add comment text

iv) Question

(1) Would you recommend something for improving the user experience for this task?

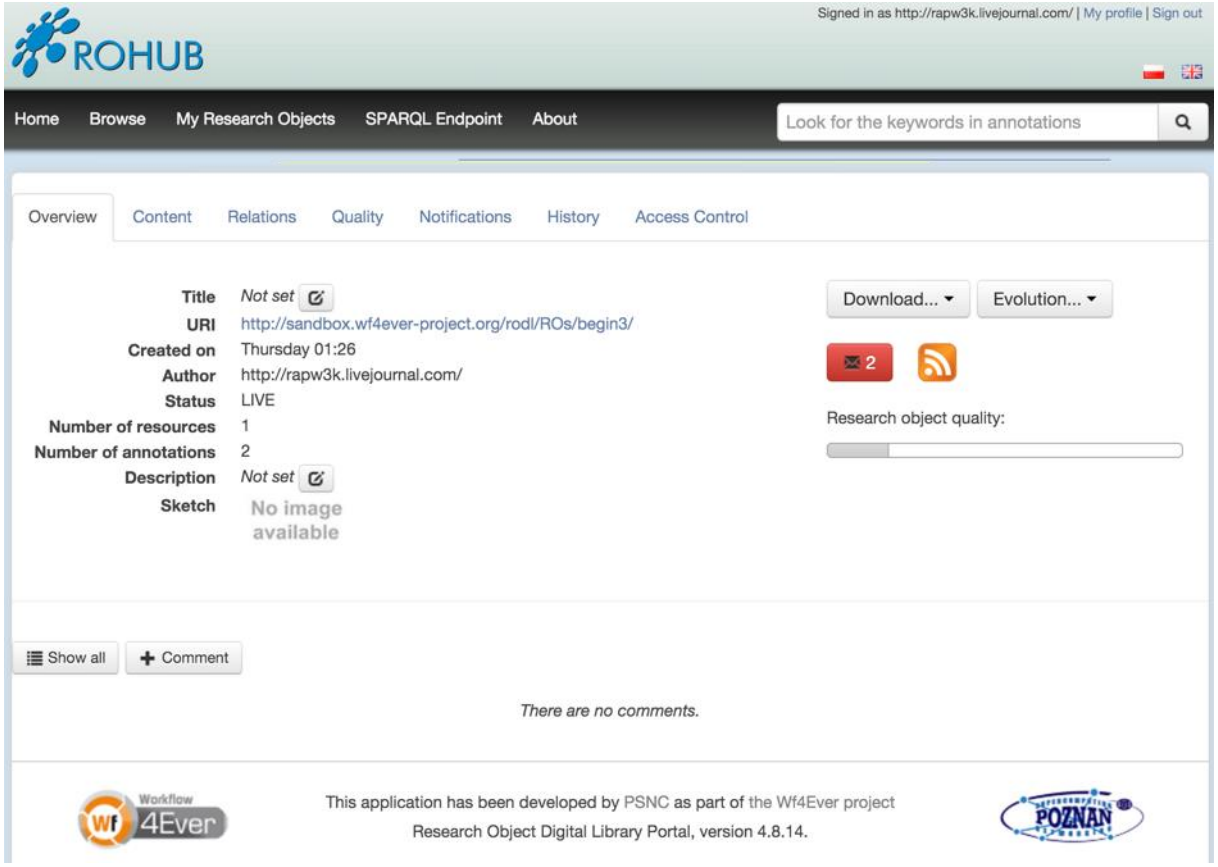

Figure 15 RO Overview tab

b) For the aggregated resource
   i) click the “Content” tab
   ii) select a resource (see Figure 16)
   iii) click the “+Comment” button (if not visible return to basic view)
   iv) Add comment text
   v) Question
      (1) Would you recommend something for improving the user experience for this task?

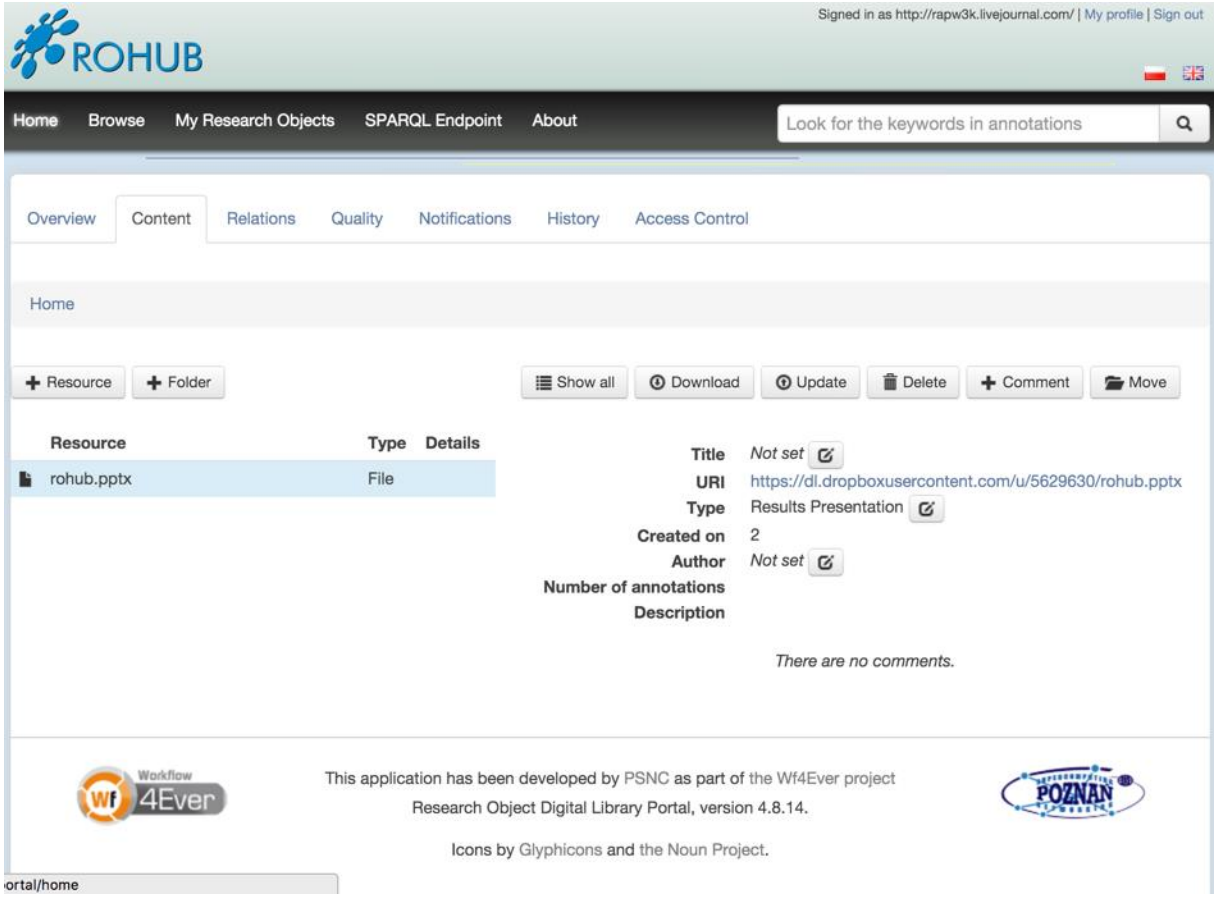

Figure 16 RO content tab - aggregated resource selected

9) Visualize & navigate an RO (e.g., your new RO)
- a) Open an RO
- b) Click “Overview” tab
- c) Click “Content” tab
- d) Click “Relations” tab
- e) Question:
  - i) Would you recommend something for improving the user experience for this task?

10) Check the RO quality
- a) Open an RO (e.g., your new RO)
- b) Click “Overview” tab
  - i) Click “Research Object quality” bar and analyse (see Figure 17)
- c) Click “Quality” tab (see Figure 18)
  - i) Select any “checklist template” and analyse
- d) Open an RO with some quality evaluation history, e.g., http://sandbox.wf4ever-project.org/rodl/ROs/musicStudy-5/ and in the “Quality” tab, click “See quality history with RO Monitoring Tool” link
  - i) Visualise and analyse the graphic (see Figure 19)
- e) Question:
  - i) Would a checklist template be useful for your VRC? If yes, could you provide such (draft) checklist?

ii) Is the decay monitoring tool interesting/useful for your VRC? Why?
iii) Would you recommend something for improving the user experience for this task?

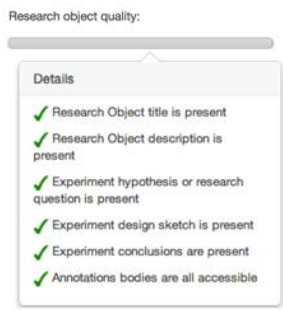

Figure 17 RO quality bar

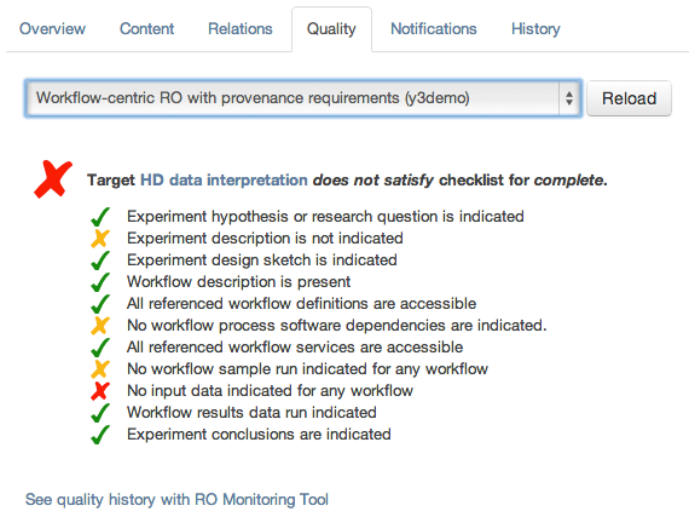

Figure 18 RO quality tab

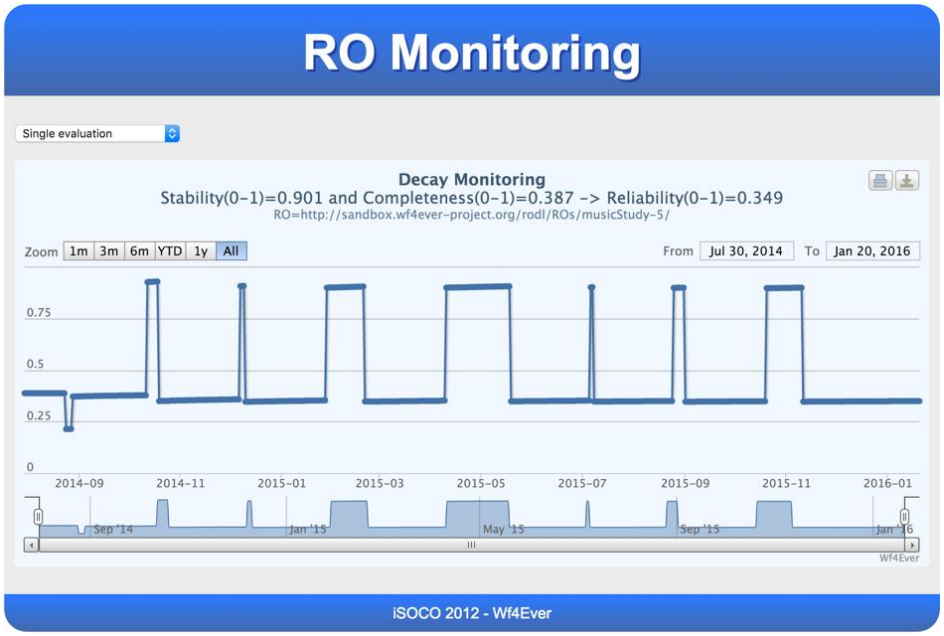

Figure 19 RO decay monitoring tool

11) Manage RO evolution
- a) Open your new RO
- b) Click the "Overview" tab
- c) Click the "Evolution" button & select "Snapshot"[5] (see Figure 20). Note: this task takes some time (in the background)
  - i) This task will create a copy of the RO at the current time, so you can continue working on the main (live) RO and share the snapshot with other colleagues/reviewers.
- d) Visualise the RO evolution
  - i) Click the "History" tab (see Figure 21)
  - ii) Visualise and analyse the graphic
  - iii) Click the snapshot RO (this will open the RO)
  - iv) Return to your RO
- e) Question
  - i) What other evolution tasks or RO stages do you envision? E.g., Fork to create new (live) RO from the existing one, and move to another direction
  - ii) How can we improve the user experience for this task?

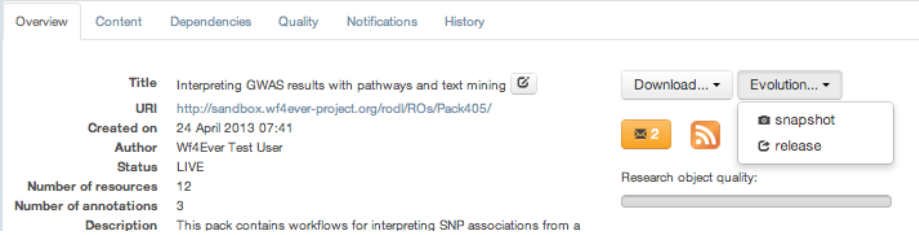

Figure 20 RO evolution tasks

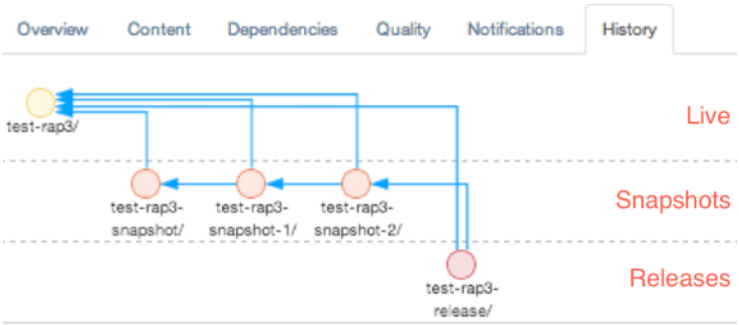

Figure 21 RO history tab

12) Set RO access control
- a) Open your new RO
- b) Click the "Access Control" tab (see Figure 22)
- c) Set the access mode to your RO
  - i) Open – everybody can read/edit
  - ii) Public – everybody can read, only owner and editors can edit
  - iii) Private – only owner, readers and editors can read, and only owner and editors can edit
    - (1) To add readers/editors, you need the user uri, e.g., http://rapw3k.livejournal.com/

[5] Evolution stages are: live (for the working copy of the RO); snpashot (a read-only copy of the RO at a certain point in time for sharing with colleagues/reviewers); release (a read-only copy of the RO at the final stage when the research has been concluded, for preservation purposes).

d) Question
   i) Do you envision other roles or access modes? (list them)
   ii) How can we improve the user experience for this task?

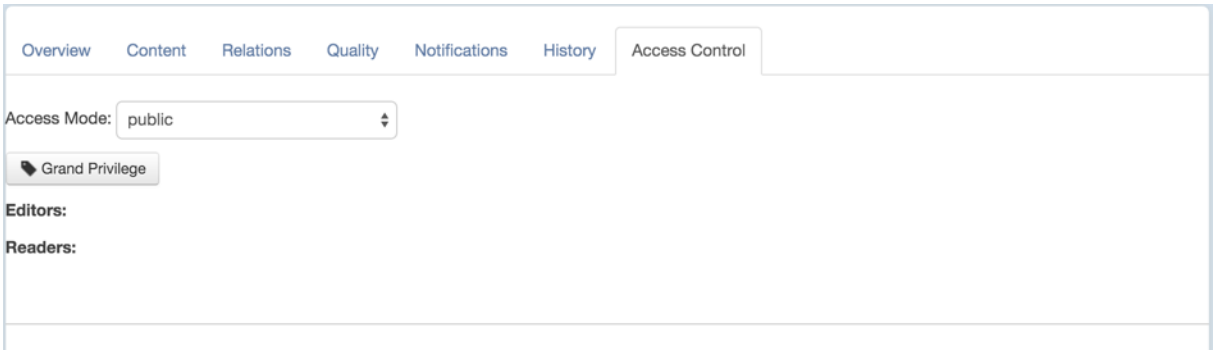

Figure 22 RO access control tab

13) Check the RO notifications
   a) Open your new RO
   b) Click the yellow box with message symbol in the "Overview" tab (see Figure 23), or click directly the "Notifications" tab (see Figure 24)
   c) Read some notifications
   d) Optionally add feed to your applications like outlook, web browser, etc.
      i) Click the "feed" button in the "Overview" tab
      ii) Copy the URL
      iii) Paste the URL to your application
   e) Question
      i) How can we improve the user experience for this task?

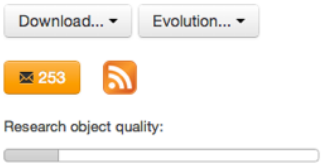

Figure 23 RO notifications links

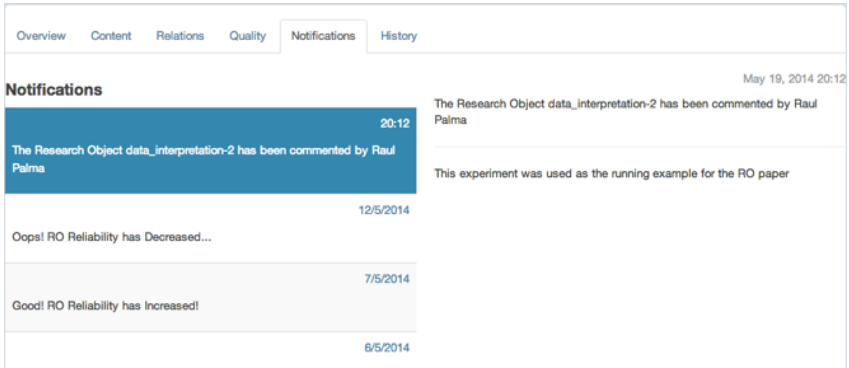

Figure 24 RO notifications tab

14) Discover, access & download ROs in ROHub consumption

a) Open ROHub
b) Click “Browse” tab (see Figure 25)
   i) Play for a while with the filter of the faceted search
   ii) Try also the sorting options
c) Enter some text in the search textbox at the top to find your new RO
d) Open the discovered RO
e) Download the RO (as zip)
f) Question:
   i) Would you recommend something for improving the user experience for this task?

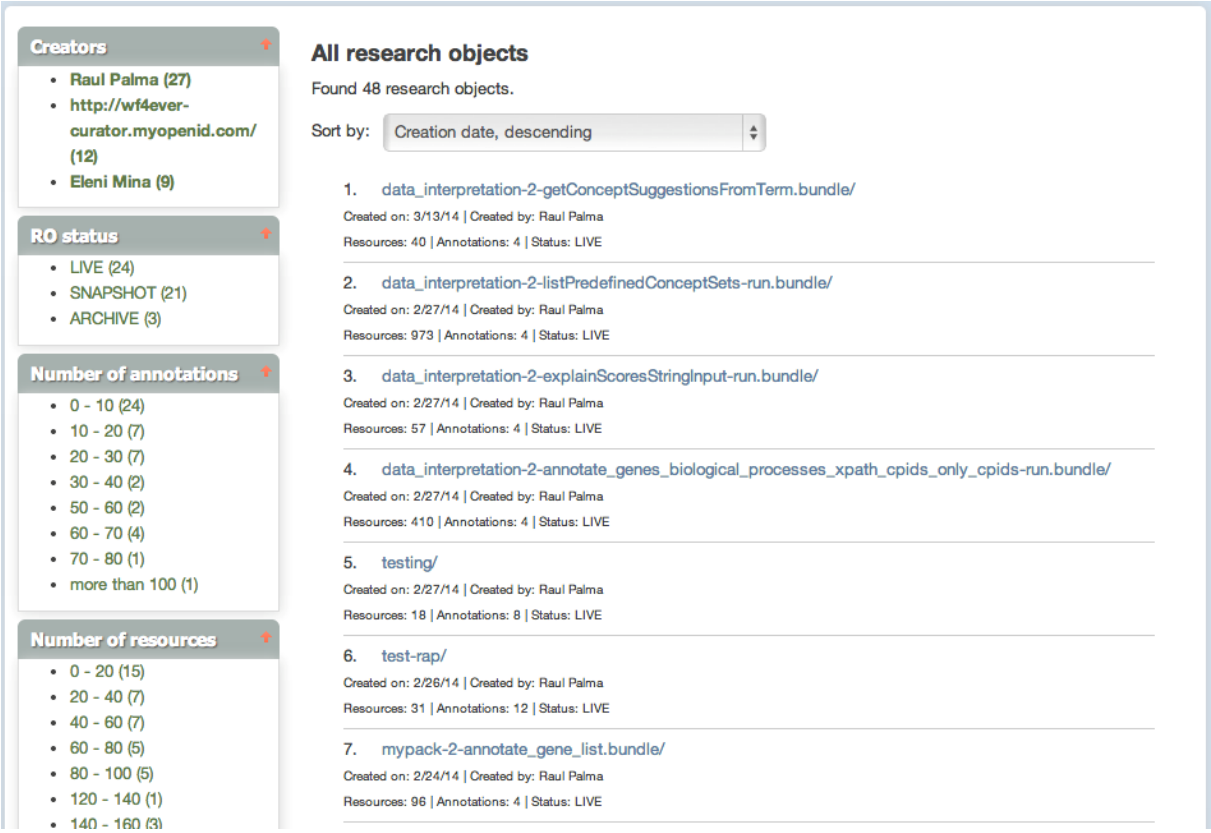

Figure 25 ROHub browse tab

Final Questions

15) What are best features of rohub
16) What are the weak features of rohub
17) What are the missing features in rohub
18) Is reproducibility important for your VRC - why?
19) What is the current state of reproducibility in your VRC, what is the goal (for you)?
20) Any other important aspect to be considered in rohub?

Appendix C – EVER-EST User Manual

1. Introduction to the use of the VRE and the Research Objects

Virtual Research Environments, also called Virtual Laboratories or Science Gateways, are community-oriented digital systems designed to help scientists collaborate and pursue common scientific goals. As such they are one of the enabling pillars of the Open Science paradigm.

The EVER-EST VRE has been developed to support (among others) the scientific community gathered around the GEO Geohazard Supersites and Natural Laboratories (GSNL) initiative (www.geo-gsnl.org). It provides a set of digital services which allow scientists to share data, results, knowledge and workflows, to use high performance processing capacities, to document all the components of their research in a structured way exploiting the Research Object (RO) concept. It also allows to assign Digital Object Identifiers (DOI) to proprietary research work, providing one of the preconditions needed to establish a trustworthy collaboration within a globally dispersed scientific community.

Data discovery

The EVER-EST VRE provides an interface to access satellite Earth Observation data from a variety of sources (Sentinel-1 and MODIS for our community), and store them in the cloud repository (based on Seafile file-hosting software system) in a personal area (which can be shared for collaboration purposes) for subsequent analysis or processing.

Data analysis

The EVER-EST VRE provides the user access to a Linux or Windows Virtual Machine (VM), already configured with some software tools needed to carry out the most common EO data analyses used by the GSNL community. The user can install and run his/her own codes in the VM, taking also advantage of the Taverna workflow management system to make the code easily reusable by others. The user can also upload, modify and execute existing codes (workflows) which have been saved as Research Objects (ROs).

Knowledge sharing and attribution

The EVER-EST VRE provides interfaces to access and create Research Objects as a way to facilitate knowledge sharing and proper attribution. The concept of Research Object was developed by Bechhofer et al. (2010), and successively expanded through demonstration and practical use in various scientific disciplines (see www.researchobject.org for a description of the RO concept and an updated bibliography; www.rohub.org provides access to ROs).

In short, ROs are aggregations of resources that bring together data, methods, results and people to document scientific investigations. Their goal is to encapsulate digital knowledge and provide a structured mechanism for sharing and reusing it. The concept of the RO was first proposed as a new digital-era tool for sharing research investigation results (Bechhofer et al., 2010). Later, the role of ROs in facilitating not only the sharing but also the reuse, reproducibility and preservation of research work was envisioned by developing the concept of workflow-centric ROs (Bechhofer et al., 2013).

The content of a RO can be for instance:

- a set of scientific articles and/or grey literature reports/bulletins concerning a specific subject, event and/or area (e.g. all material describing a specific eruption in one of the volcano Supersites);
- an aggregation of different datasets used for a specific experiment (e.g. to facilitate data sharing and reuse);
- a complete representation of a processing workflow, including for instance reference to the input data, the executable code, the output results, information on the workflow procedure, information on attribution and provenance of the research (using e.g. a DOI and the researcher's ORCID ID).

In the EVER-EST VRE users have an interface to create one of the RO types above, to reuse an existing workflow-centric RO, and in general to display the content of any RO. To manage the RO metadata (e.g. to complete the information required to make the RO fully documented and reusable) the user then needs to employ the ROHUB

portal (www.rohub.org), which provides access to the RO database and to the complete set of RO management functions. The full exploitation of the RO potential by the scientific community is one important step towards the implementation of the Open Science paradigm, in the GEO-GSNL initiative and elsewhere.

Manual structure

This manual is organized as follows: Section 2 provides a description of the main components of the VRE, Sections 3, 4, and 5 provide three different examples of use of the VRE during a scientific investigation.

To a new user, we suggest to start from the description of the use cases, and to use the component section to detail unclear aspects.

Please provide your feedback and suggestions for improvement to: email GSNL

2. High level description of the VRE Components

The **EVER-EST VRE** for the Supersite is https://vre.ever-est.eu/supersites/. The use of Chrome is strongly suggested to avoid mis-functionalities.

The VRE provides access to distributed services. Two of the systems providing such services to EVER-EST are ROHub and the Collaboration Sphere. Their user interfaces can be accessed also externally to EVER-EST through proprietary GUIs .

2.1 *Registration and authentication*

The user needs to sign in at the **Supersite VRE** site, https://vre.ever-est.eu/supersites/ by clicking on the icon at the top right corner:

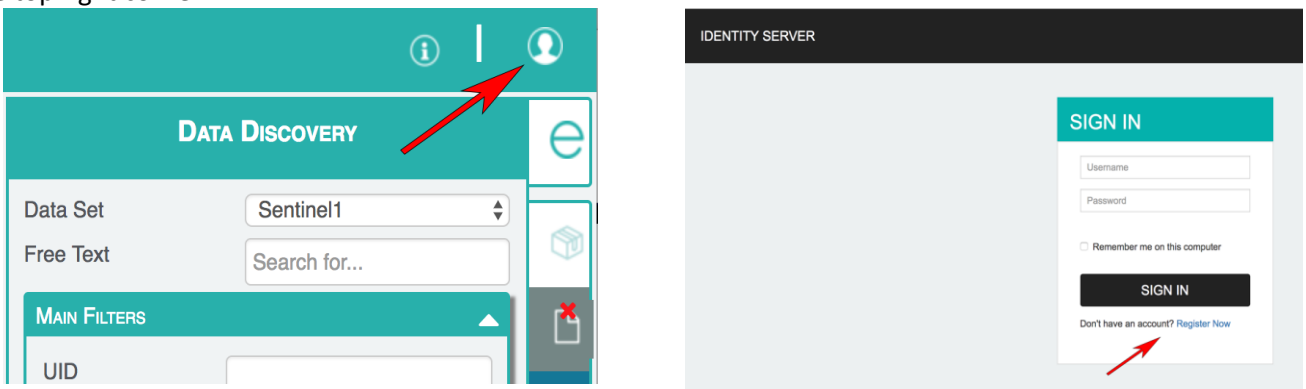

Figure 26 Left: Sign in in the GUI. Right: complete sign in

The icon lands to the identity server as above (right). In case of first registration, the new user can create an account from there.

Once authorized, the user holds an **EVER-EST id** and can log in from the same icon.

The VRE uses the Single Sign On technology, and the registered EVER-EST authentication can be used to access also the other VRE components: Seafile, Rohub, the Collaboration Sphere and the cloud controller platform.

Figure 27 Same user authentication, the EVER-EST id, for all the VRE components

2.2 *User storage area*

The EVER_EST VRE provides a storage area to the users, based on the Seafile enterprise file-hosting platform (similar to DropBox or GoogleDrive). Personal files can be put in a specific server, can be synced and shared across

different devices and may be accessed as a virtual disk. The user can access the https://box.everest.psnc.pl/ server via web with the EVER-EST credentials. “Shibboleth” log in or “Single Sign On” must be chosen.

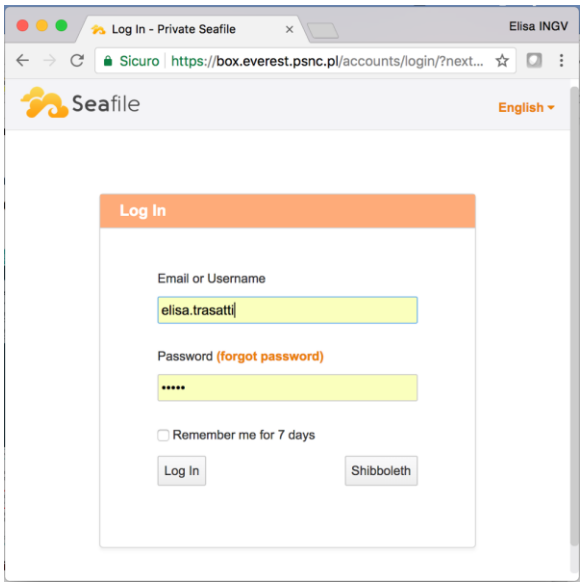

Figure 28 Log in the Seafile server at https://box.everest.psnc.pl/.

Seafile can be also used directly from the main VRE interface, or through a client in a Virtual Machine (VM) or in your local device. In the latter case the user needs to download the client from https://www.seafile.com/en/download/. Configure the client to use the https://box.everest.psnc.pl/ server, then select the Single Sign On option (see Figure below - missing). User and password are the EVER-EST credentials. From the client app, the local folders can be synced with their correspondents on Seafile, and viceversa. The owner can choice which folders to sync, while the whole cloud is accessible via web as described above.
The Seafile personal storage area is also accessible from the GUI. Click on the icon and sign. After a successful authentication, the Seafile icon turns from a red cross to a green approval.

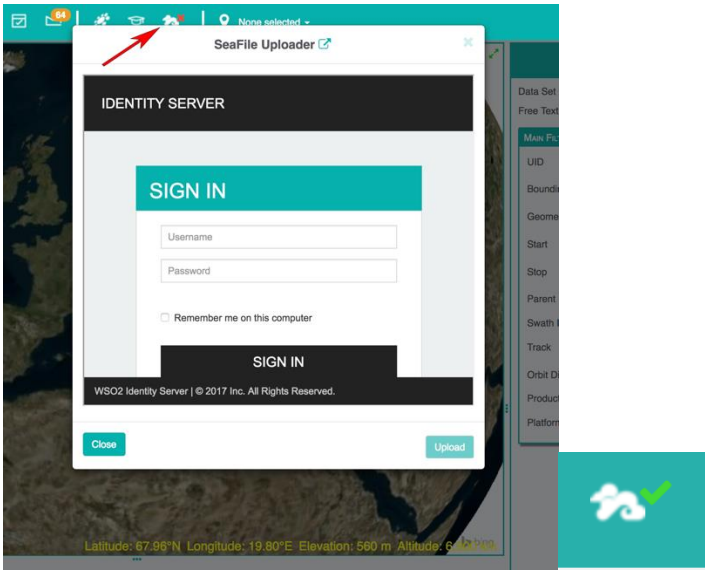

Figure 29 Seafile personal storage cloud is accessible from the GUI. If connected, the icon is with a green check

2.3 *Rohub*

Rohub is a research object management platform supporting the preservation and lifecycle management of scientific investigations, research campaigns and operational processes. It makes available and discoverable new knowledge through the research object paradigm. As the only existing platform implementing natively the full research object model and paradigm, resources associated to a particular experiment are aggregated in a single digital entity (research object), and metadata relevant to understand and interpret the content is represented as semantic annotations that are user and machine readable. It is accessible at http://www.rohub.org. User's own research objects may be created and managed by accessing the system with the EVER-EST id. There is a direct connection to Rohub.com from the GUI, since research objects may be created and managed in the GUI, and may be further managed in the Rohub web site.

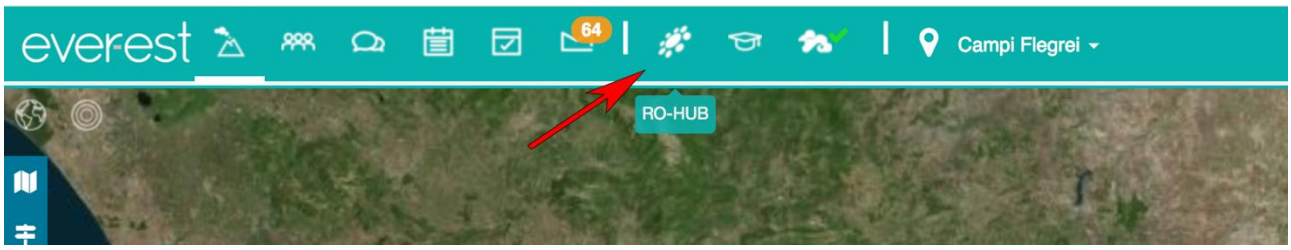

Figure 30 Rohub is accessible from the GUI

2.4 *Collaboration Sphere*

The **Collaboration Sphere** finds more relevant ROs and structure them through semantic intelligence. If you ever felt overwhelmed by large collections of documents and it was even hard to think of the query you would need to submit, the Collaboration Spheres will enable you to search by example, providing recommendations based on a selection of ROs or researchers from your social network. Just drag and drop to see related research objects obtained through the automatic analysis of the selected content. More information can be found at http://everest.expertsystemlab.com/home/. The site for the semantic search is http://everest.expertsystemlab.com/spheres/index.html, and the user can log in at the top right corner.

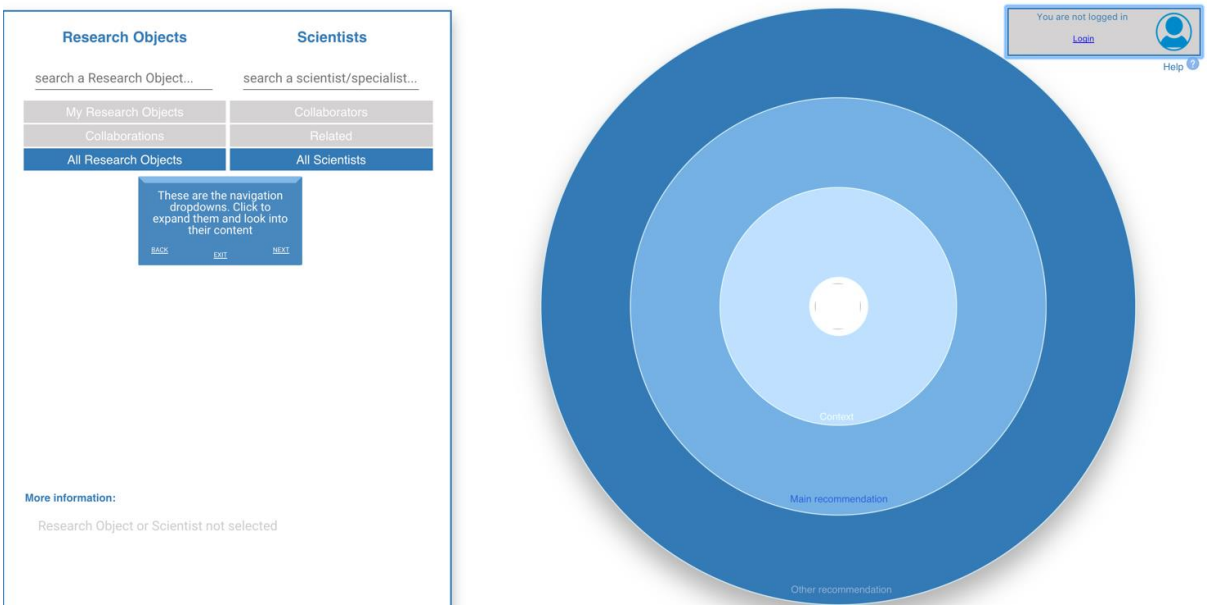

Figure 31 The collaboration sphere creates a structured collection of resources (RO and scientists) related with each other with different degrees of affinity

Once logged in, the user is in the center of the sphere, and can drag & drop a research object, a scientist and other resources in order to find resources related. In the following, a research object is chosen and dropped in the center, so that further research objects related to the topic appear. More information can be found here http://everest.expertsystemlab.com/home/.

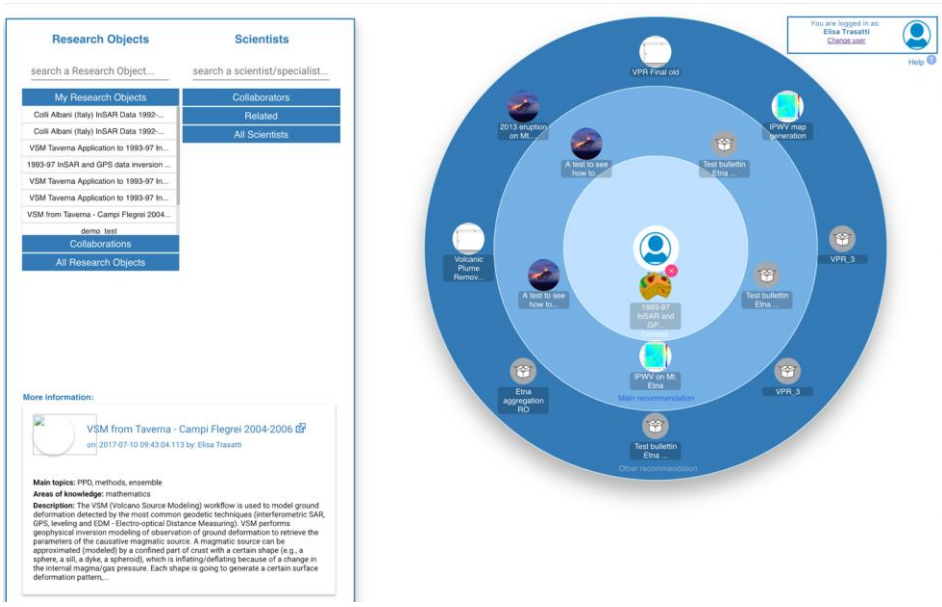

Figure 32 Example of the use of the collaboration sphere

The Collaboration Spheres are integrated with the EVER-EST VRE through a single sign on mechanism, and are accessible in the GUI at the top left corner of the virtual globe.

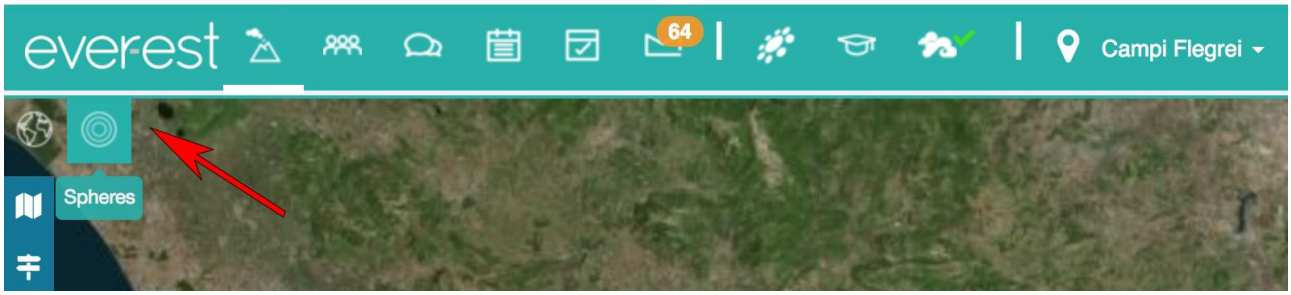

Figure 33 The collaboration sphere is accessible from the GUI

2.5 *Workflow management system*

EVER-EST supports the **Taverna** open source Workflow Management System - a suite of tools used to design and execute scientific workflows. The Taverna suite is written in Java and includes the Taverna Engine (used for enacting workflows) that powers both **Taverna Workbench** (the desktop client application) and **Taverna Server** (which executes remote workflows). Taverna is also available as a Command-line Tool for faster execution of workflows from a terminal without the overhead of a GUI. Taverna automates experimental methods through the use of a number of different (local or remote) services from a diverse set of domains – earth sciences, biology, chemistry and medicine to music, meteorology and social sciences. Effectively, Taverna enables a scientist who has a limited background in computing, limited technical resources and support, to construct highly complex analyses over data and computational resources that are both public and private, all from a standard PC, UNIX box or Apple computer. It can be downloaded from https://taverna.incubator.apache.org/.

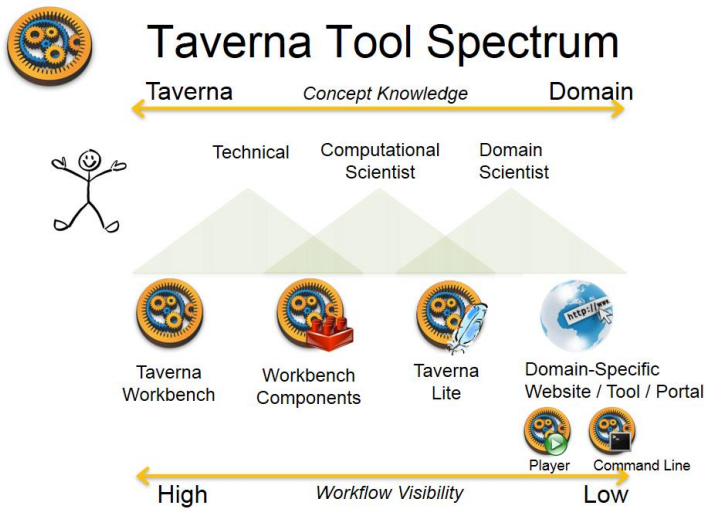

Figure 34 Scheme of the Taverna Tools

In EVER-EST the use of Taverna is suggested to easily and rationally organize the input/output of personal workflows, to make them reusable, and to easily create Research Objects from a workflow execution.

2.6 *Processing Services*

The EVER-EST VRE offers two types of processing services. It offers virtual machine resources, either Linux or Windows, and the Taverna server. The first option means that the user is given an IP to connect (explain better). In the new environment, the user is able to run his personal codes, to re-use others codes (e.g., from the reuse of a research object). The Linux VM is equipped with the following installed by default: Seafile, Taverna workbench, Fortran (gfortran 90), Matlab. The windows VM is equipped with Seafile, Taverna workbench, ENVI and Sarscape.

The second option for processing server is the execution of a workflow within the GUI without the need of local execution or connection to a virtual machine. The Taverna server, available in the GUI, can be used to run workflows created previously in Taverna (with t2flow extension). These are then executed in the EVER-EST VRE without the need to connect to external computational resources (see example in section 4). This requires that the inputs files are accessed by global URLs.

2.7 *EVER-EST GUI description*

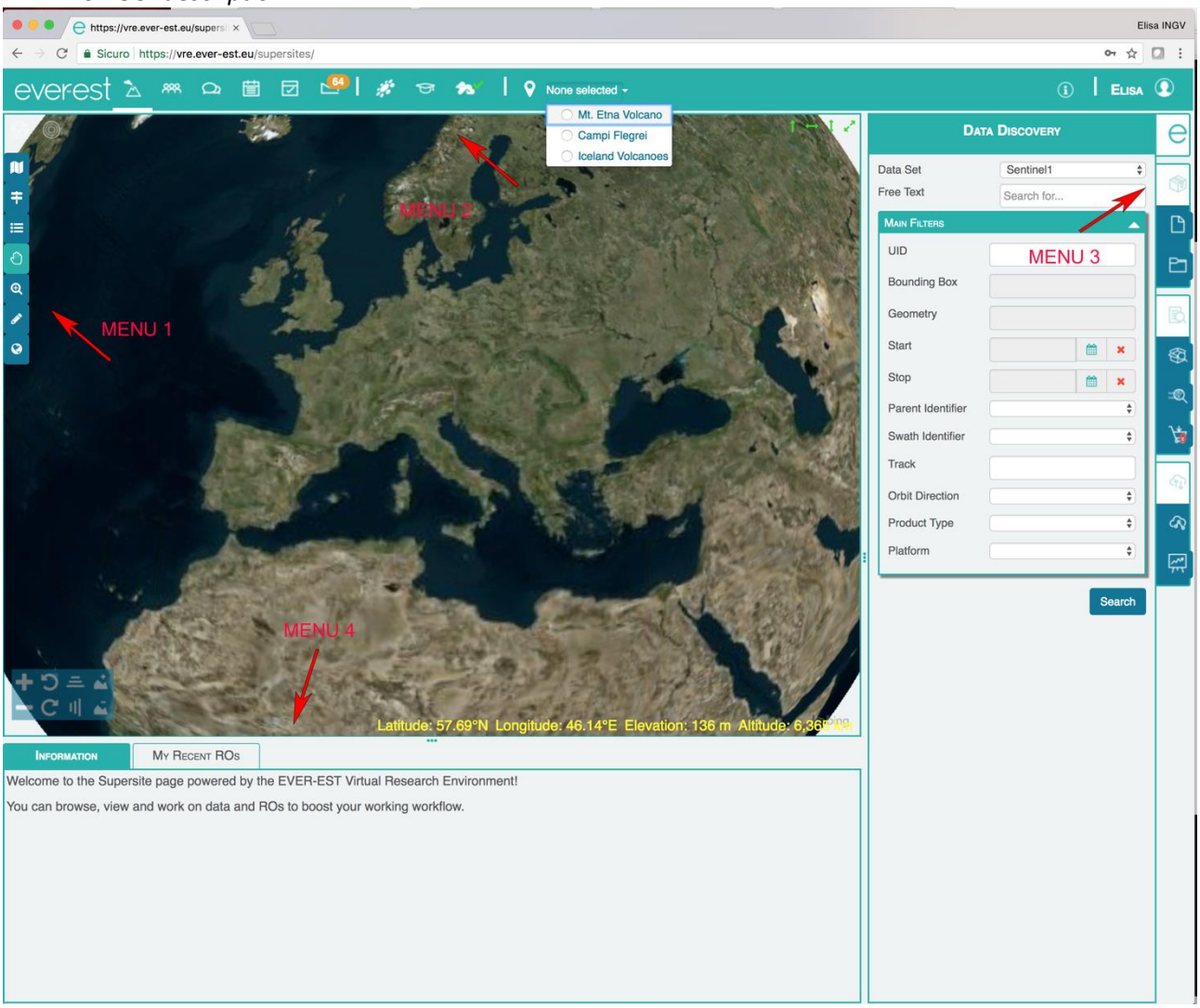

Figure 35 The EVER-EST VRE GUI for the Supersite community

Menu 1 collects a set of functionalities to operate within the virtual globe.

From top to bottom:  map projection (few lines of explanation of each button),

- search location…
- manage layers (layers such as .shp, .png can be added and visualized in the virtual globe),
- pan…
- zoom box…
- draw options (to draw areas and polygons)…
- change map (to shift to satellite and road)…

Also here, put in separate lines with short explanation Menu 2 collects a group of user-specific functionalities, such as (in order from left to right): instant messaging (chat), forum, tasks, my validation requests, mailing (as the users are mailed when new activity is performed by the user group). Then after the vertical bar, link to Rohub (see 2.3), e-learning, and the link to upload to Seafile (must have green check). After the second vertical bar, the direct selection of one of the three volcanic European Supersites, i.e., Mt. Etna, Campi Flegrei and the Icelandic volcanoes.

Same as above Menu 3 is the operational menu with the heart of the VRE functionalities. It has three main sub-menus, with white icons, that are the research object, data discovery and cloud controller. From the research object menu, it can be created a new RO or opened a recent one. From the data discovery, the user can search research objects or search data (Sentinel-1 and MODIS). The last icon of this submenu is the basket, used when discovering data. The chosen image, indeed, are stored in the basket until check out, when they are downloaded to the personal Seafile storage area. The last submenu lists the available services. The first choice is the cloud controller. It connects the user to the Terradue cloud platform, where the user must log in with the EVER-EST id (there is a specific icon). In case of the following result after sign in (bottom left): THIS MUST NOT HAPPEN

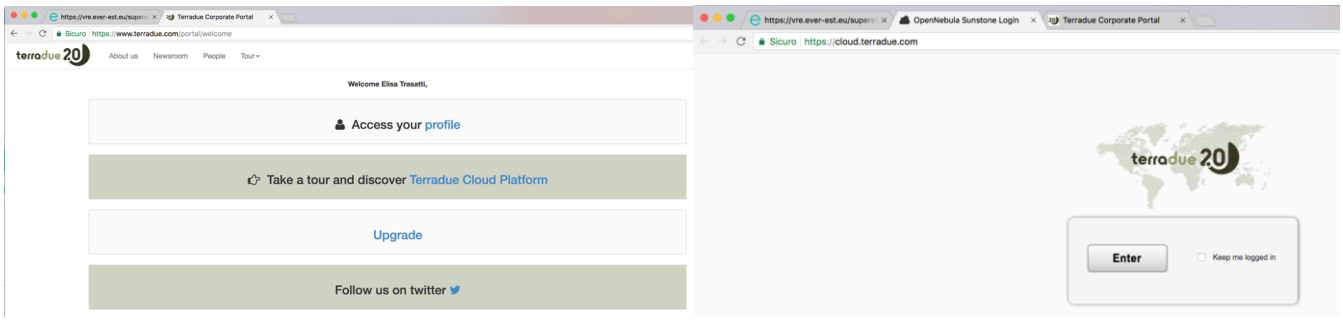

Figure 36 Access to the Terradue cloud platform. Left: first connection. Right: correct connection

the user is required to go back to the VRE and click on the cloud controller again, obtaining the result above(right). Once in, there are virtual machines (Linux and Windows) that can be chosen and cloned, simply by using the given IP and connecting from a remote desktop software (Microsoft Remote Desktop is suggested for Mac). The Linux VM is also available via ssh.

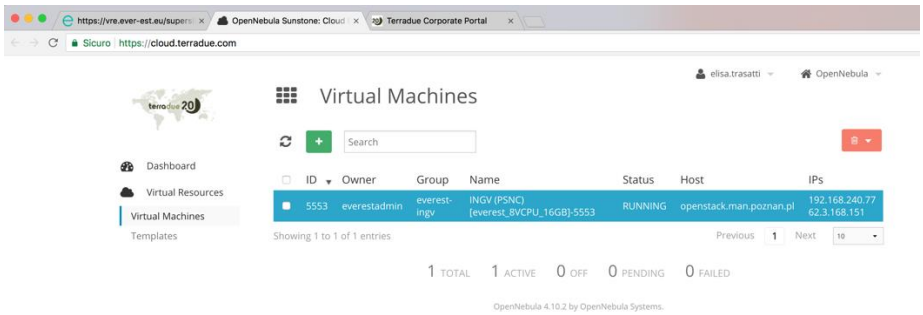

Figure 37 Virtual machines as golden images available at the Terradue cloud platform

The second and last icon of the services submenu is the workflow server, where .t2flow files (native Taverna files) can be run within the GUI. See related example below.
This is ok Menu 4 is the info area. It contains information regarding the selected Supersite, and in additional tabs it may contain the results of data query, or research object query, or the list of recent research object created, etc.

3. Use case 1: execution of a Taverna WF in a Linux Virtual Machine

3.1 Description of the use case

During eruptions, volcanoes emit large quantities of particles (ash, water vapor, ice) and gases into the atmosphere. The Volcanic Plume Removal (VPR) procedure (Pugnaghi et al., 2013; Guerrieri et al., 2015; Pugnaghi et al., 2016) has the capability to retrieve, simultaneously and in real time, the volcanic ash and SO2 clouds physical parameters from multispectral MODIS data in the TIR spectral range. The only input parameters the procedure needs, once the coefficients have been computed for a specific area and volcano, are the cloud altitude and the atmospheric profiles.

Using EVER-EST VRE we'll download the MODIS image (**section 3.2**) we want to process directly on Seafile repository and the Research Object of another researcher containing the VPR procedure. Then we use the Linux Virtual Machine where Taverna and Matlab software are installed. Having all the inputs, we'll run VPR code within Taverna to compute the volcanic cloud ash Mass, Effective radius and Aerosol Optical Depth and the SO2 columnar content. Results are stored in Seafile.

3.2 Access to and management of the Linux Virtual machine

After the creation of a temporary Virtual Machine and the access as described above (subsection 2.7) we could find in “Applications” menu, Seafile Client (“Internet” section, please login by your account) and Taverna software (“Others” section).

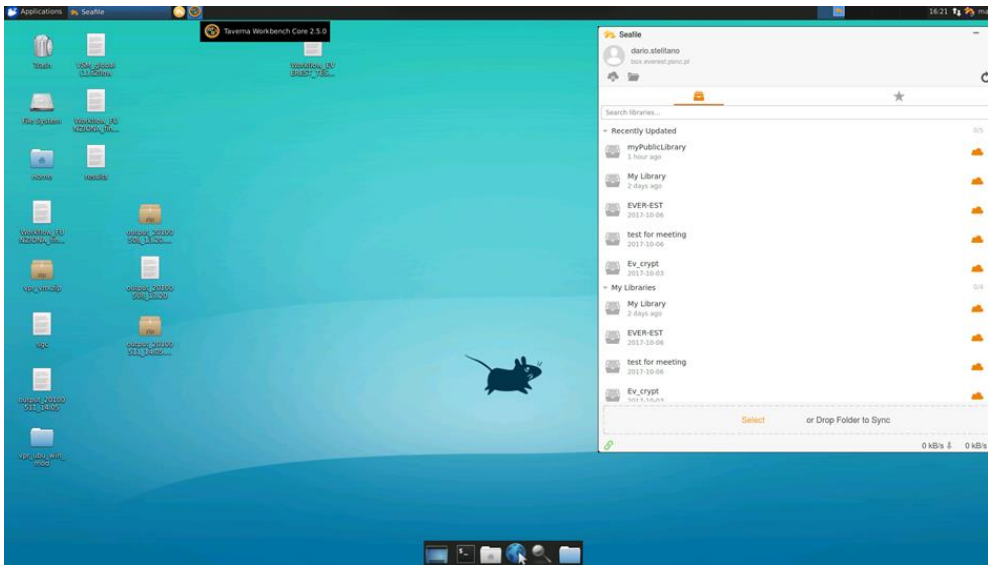

Figure 38 Linux Virtual Machine desktop. The Seafile client is running

3.3 Data discovery and access

3.3.1 VPR Algorithm data discovery

In this case we are looking for the algorithm uploaded by Manzo (creator field, menu 3). In case your RO was created into the VRE website and not in Rohub flag “only VRC ROs” (menu 3). We could see the result on menu 4 (“Volcanic Plume Removal 717”).

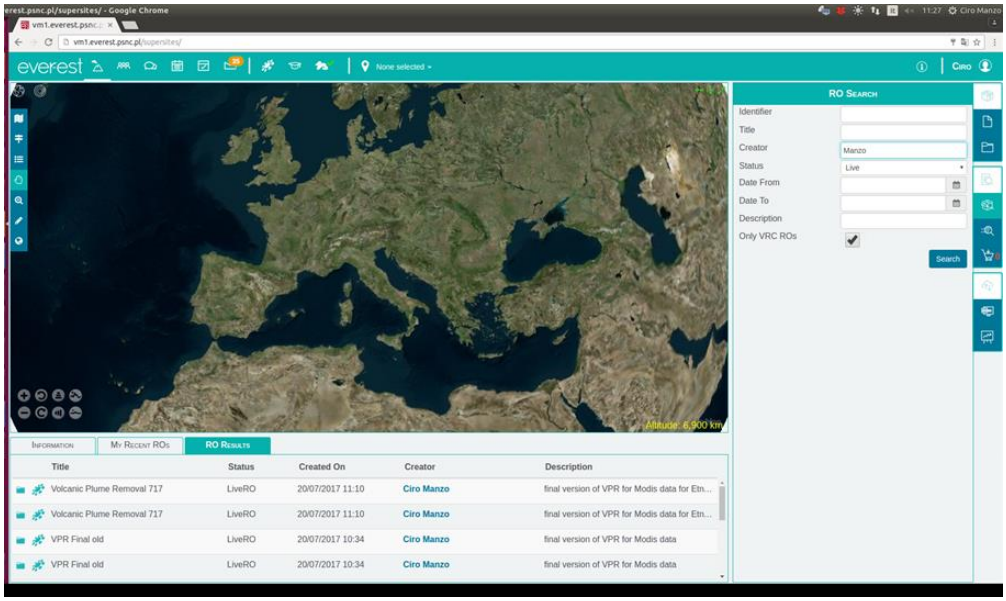

Figure 39 RO search and discovery in the GUI

To download the workflow to run the algorithm you must click on the second icon of RO selected which open the Rohub environment in the webpage http://www.rohub.org/portal/ro?ro=http://sandbox.rohub.org/rodl/ROs/VolcanicPlumeRemoval717/

By the "Overview" tab you can move to "Content" in order to check all data and documents stored into the RO.

Figure 40 The content of the selected RO is deeply examined in the Rohub web site

Following the link "Content→Workflows".
Select the Taverna file ("Workflow_Funziona_finale.t2flow"), the download button is active.
Download the file on your computer.
This t2flow file works only with Linux machine and with Matlab 14 installed.

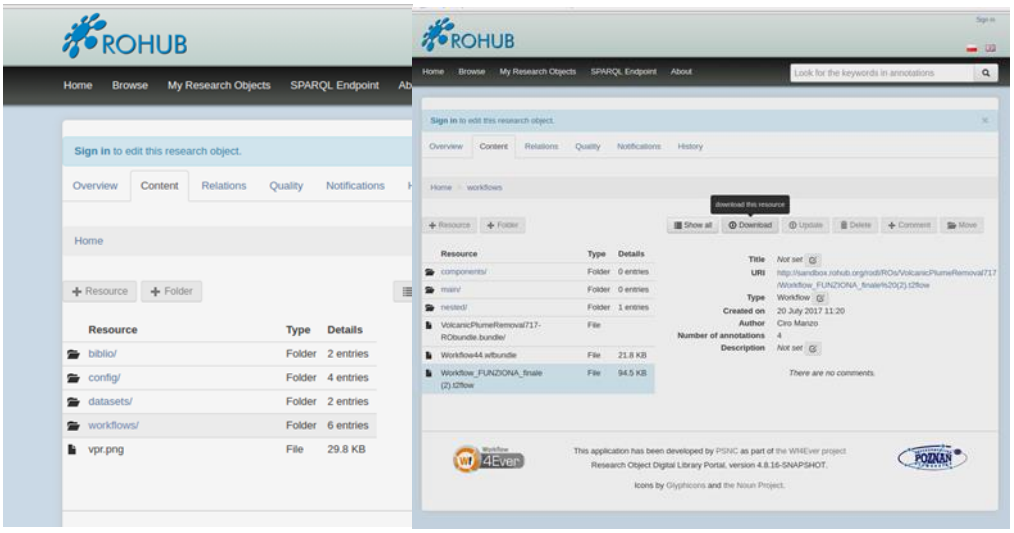

Figure 41 Download of the workflow from the RO stored in Rohub

3.2.2 Data discovery of new MODIS data

Select your area of interest on EVER-EST site (after login, please see previous paragraphs).

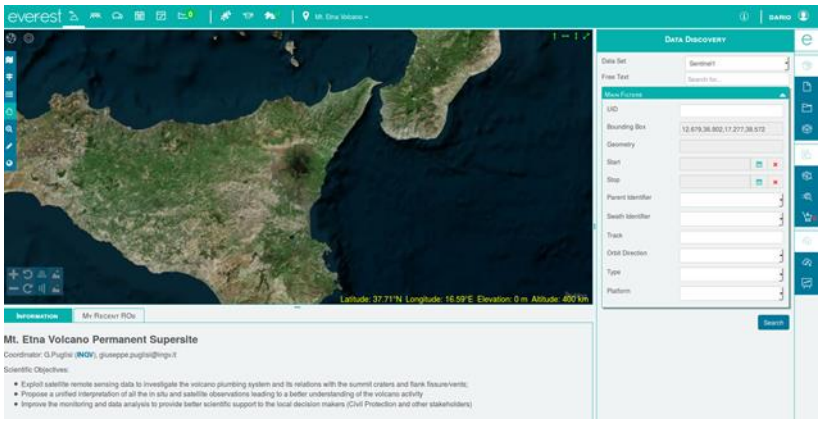

Figure 42 Mt Etna volcano Supersite selected from Menu 2

Select draw polygon on menu 1, draw your area of interest and select MxD03 in Dataset field and your temporal slot in menu 3.

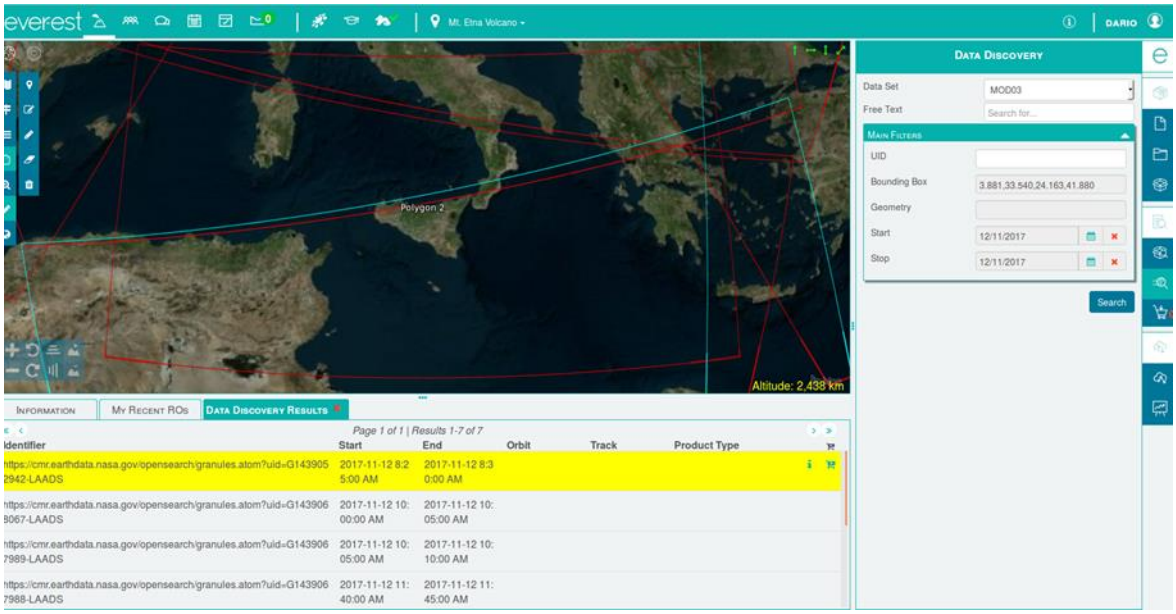

Figure 43 Data discovery and selection

Add the image to your basket (red icon on right of menu 4 row). After you could decide in which Seafile folder you want to download it.

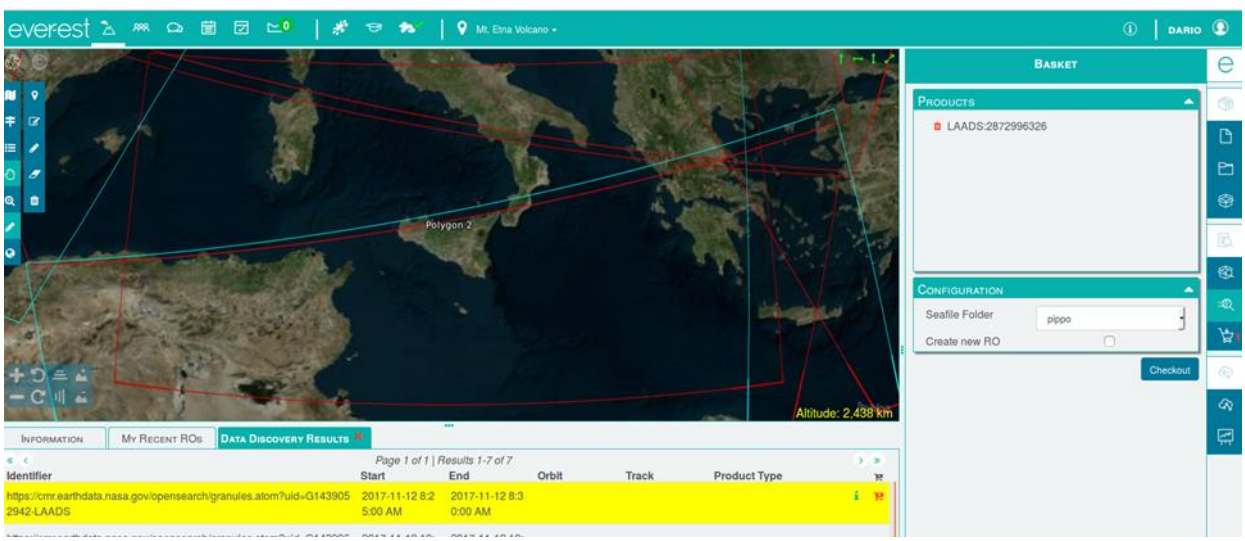

Figure 44 The selected data are added to the basket and downloaded in Seafile

Repeat this procedure also for MxD21 MODIS Data.

3.4 Data processing

Once into the virtual machine it is possible to access Taverna software, by icon in the top bar:

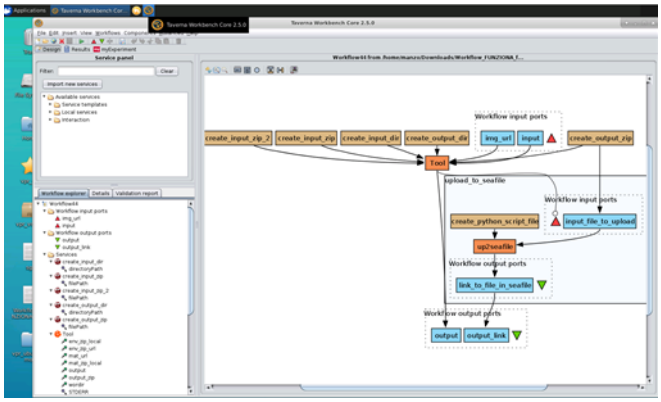

Figure 45 Use of the workflow in Taverna from the Linux VM

The VPR algorithm nested into the Taverna flow requires a series of input that are provided in the test case of RO. These inputs must be:
- MODIS images MxD03 and MxD21, at specific time.
- Plume mask per each plume component (Ash named by suffix *Vol or And*, Sulfide Anidrid by suffix *SO2*…). For any additional detail, related to the format check bibliography and the file into the example or contact the authors.
- The atmospheric temperature file must be named according the format of WMO files (for any additional detail, related to the format contact the authors and check the file into the example).
To download go to http://weather.uwyo.edu/surface/meteorogram/europe.shtml

Must firstly upload zipfile with all required input (please see above) on Seafile folder, then generate the link of the zip file link and then download on the virtual machine.

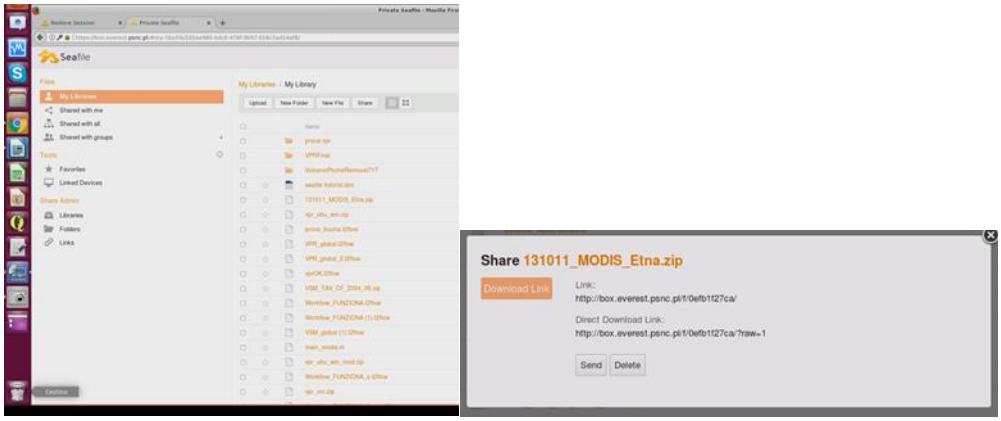

Figure 46 Data uploaded in Seafile and generation of the unique download link

Copy this link and paste into the form related to “img_url” tab. (figure below) and run the workflow (green triangle at bottom of page).

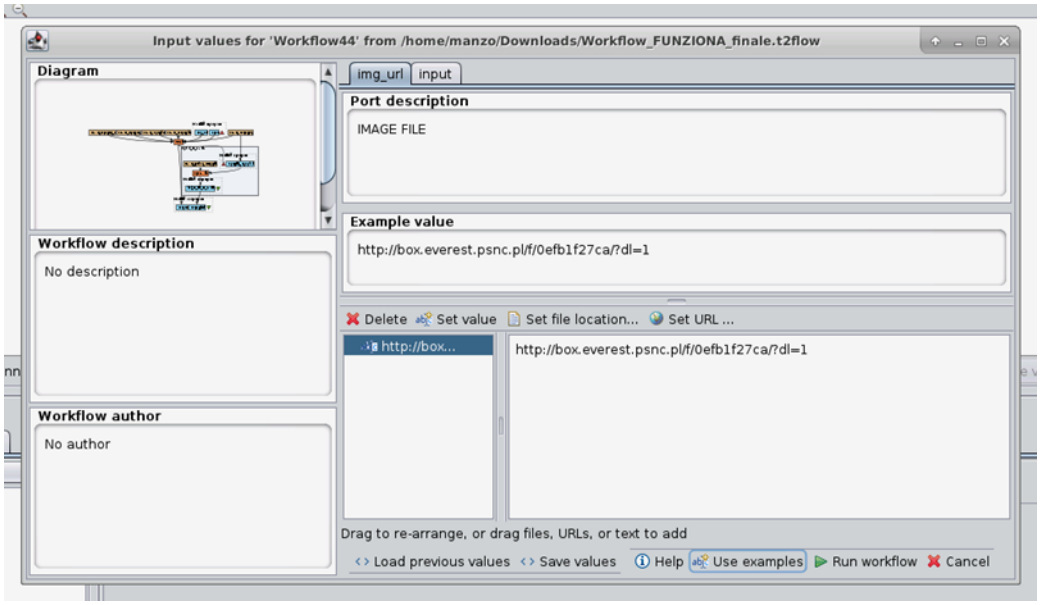

Figure 47 The input is the url of the image file. The url is the download link previously generated

During VPR workflow run the procedure will ask you:
the image file

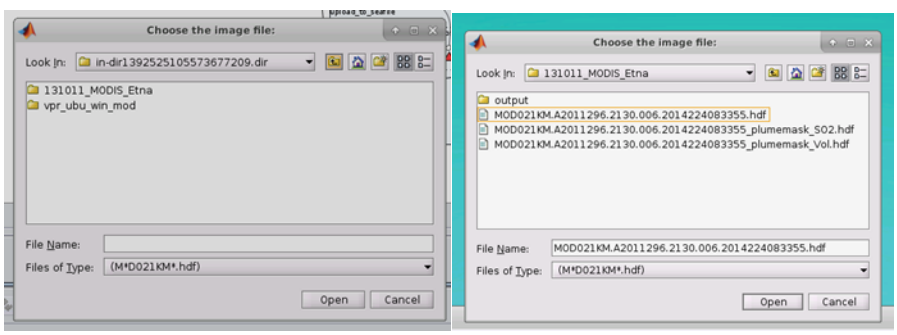

Figure 48 Input during VPR run

The volcano

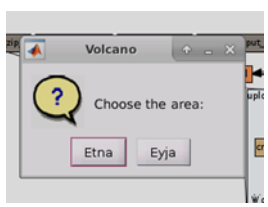

Figure 49 Input during VPR run

And the height, both for SO2 and Ash plumes

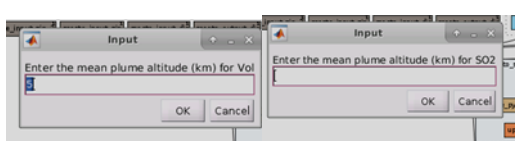

Figure 50 Input during VPR run

3.5 Management of output results

After the VPR running the flow upload the results on a Seafile folder, to check them out it is possible click on the output_link tab and copy and paste the corresponding link.

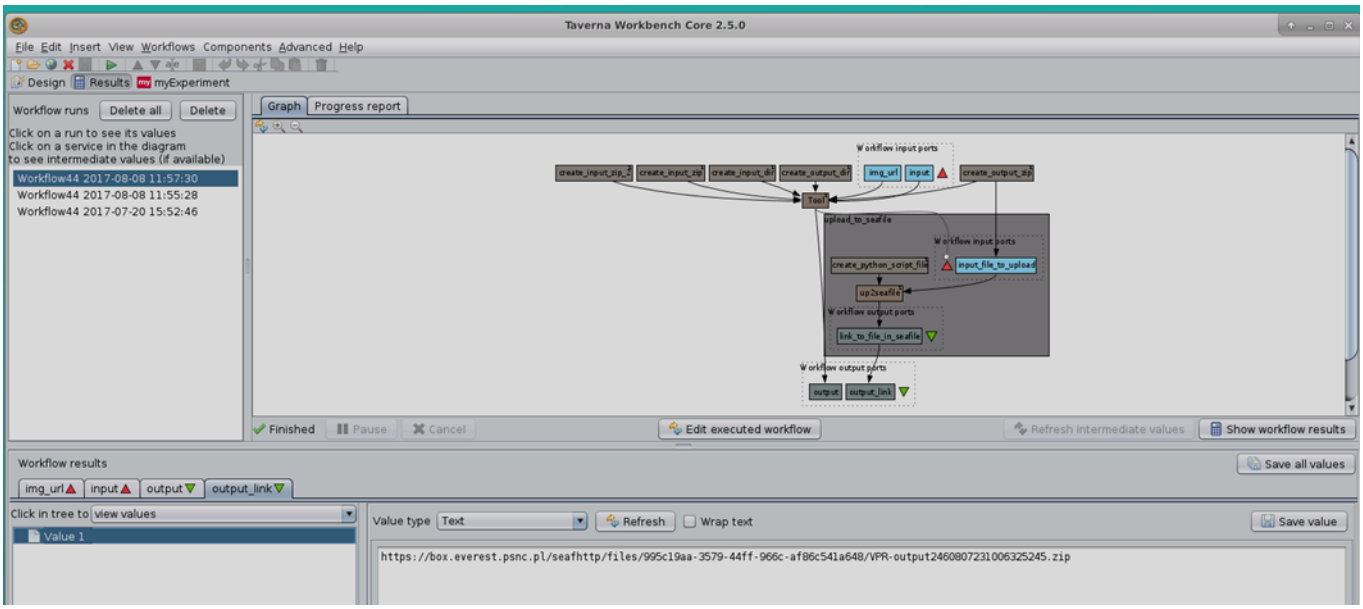

Figure 51 After the VPR run, Taverna manages a second step in which the results are stored in a zip file and uploaded in Seafile The result is the url to that zip file

To create a RO of results, click on "Save all values" button and then into the new pop up window on "save provenance bundle". It takes some moments to create the bundle zip file. When the process finishes create new RO using the procedure in section 6. How do I use the bundle? First store bundle in seafile and then create the RO?

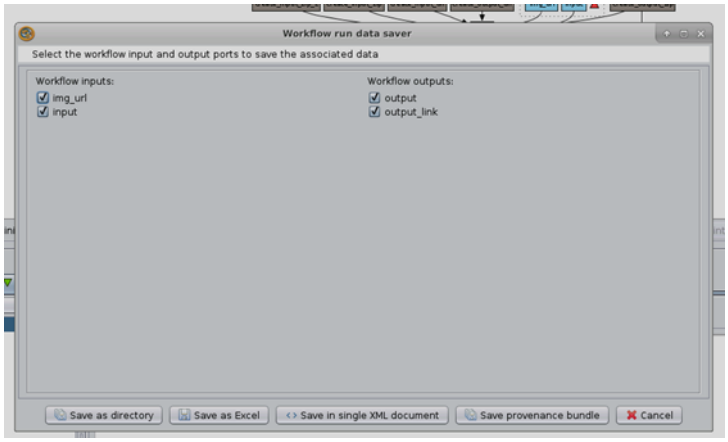

Figure 52 All the resources used and created during the run in Taverna can be stored in a zip file to be saved and exported in a RO

4. Use case 2: execution of a WF within the VRE

4.1. Description of the use case

This use case shows a re-use of a RO within the GUI. Let's consider there is an existing workflow-type RO, that includes a workflow that can be run in Taverna, together with descriptions. The RO may contain resources of a complete run, e.g., input and output data, but it may be consisting in the workflow alone. Consider also the user has data that can be the input of the tool referenced in the RO. He can re-use this resource with his data by

downloading the RO and run the workflow locally or using the virtual machine. The EVER-EST VRE offers the processing service of running the workflow within the GUI by using the Taverna server. This may be suitable in case of serial runs using the same algorithm embedded in the workflow.
In this particular case, there is RO containing the results of the Volcanic Source Modelling (VSM) tool. VSM is a tool to model geodetic data of volcanic areas, and considering the deformation as due to magma input at depth, it performs inversion to estimate the position, the depth and the shape of the magma chamber of magma storage. The RO collects InSAR data of 2004-2006 at Campi Flegrei (Italy), showing uplift of about 6 cm. The data is modelled supposing the action of a spherical magma chamber. The RO is a workflow type and archived with a doi. The user has new data from Campi Flegrei, regarding 2011-2013 uplift of about 16 cm in Pozzuoli, and want to test this data against a magma chamber recharging hypotesis. The user can re-use the existing RO by loading the workflow containing the VSM within the VRE. The run can be done in the VRE using the WF runner without downloading the RO resources, instead of locally or connecting to the VM.

4.2. RO discovery and access

The user signs in the VRE-Supersite.
In this use case satellite data have been previously processed, and it is stored in the personal Seafile storage area. Data can be visualized in the virtual globe as a layer, being .png files.

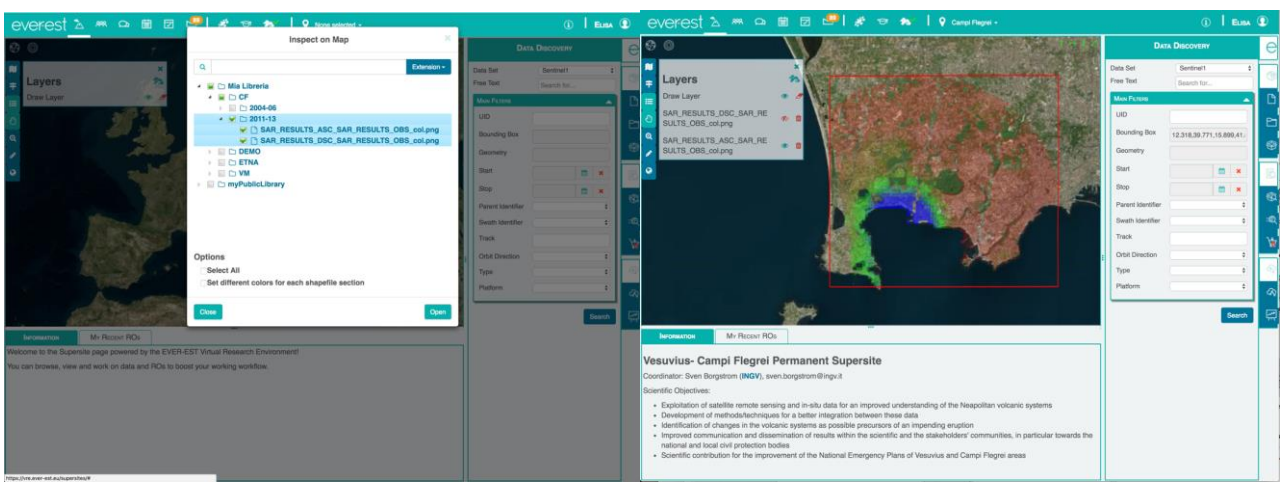

Figure 53 Data visualization within the virtual globe. Blue areas evidence uplift in the Pozzuoli bay

The InSAR data evidence the active deformation of uplift (together with horizontal displacement). The user looks for a RO containing a tool useful to study the deformation, by searching for the “VSM” keyword.

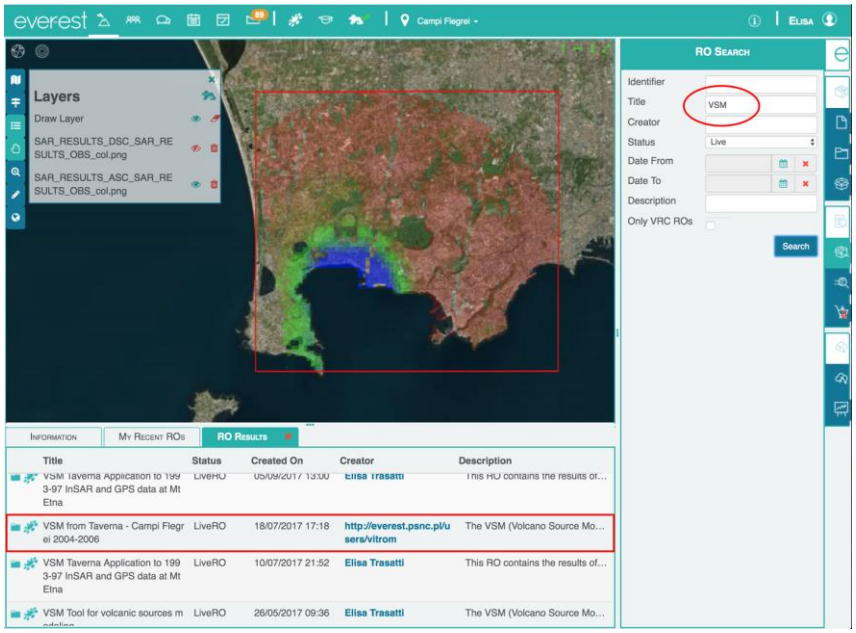

Figure 54 The user clicks on the RO search and searches for "VSM" in order to find ROs using this tool

From the results, the user chooses a RO that seems more related to his data (the one regarding Campi Flegrei). The use can discover the content within the VRE. The RO comes with the workflow containing the VSM tool, and documentation useful to learn how to use it.

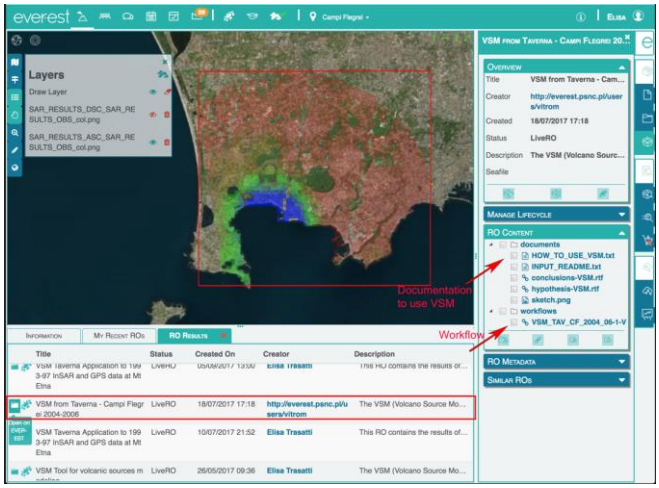

Figure 55 The user chooses one of the RO and can open it in the VRE to inspect the contents

The user is able to format the input data and the modelling set up as required by VSM by following the information contained in the documentation. The input file is a zip file containing the all the input files (two ascii files, one for each satellite orbit in this case, plus an ascii file containing GPS data and two ascii files specific for the inversion set up). The input file is stored in the personal area in Seafile and a download link has been generated (as described above, see figure below).

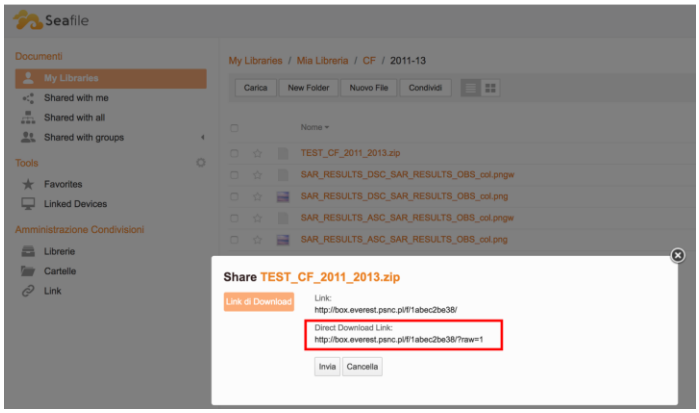

Figure 56 Storage in Seafile of the input data files in a single zip file, and generation of a unique download link

4.3. Taverna runner

With the RO still open in the VRE, it is possible to click on the bottom icon of Menu 3, the "Workflow Runner". It implicitly recognize the presence of a Taverna workflow within the open RO (.t2flow file). The workflow can be executed by entering the required input urls. In this specific case, the inputs are the url to the zip file containing data and inversion settings, and a second url is referred to the Linux or iOS compiled Fortran code. This second url must be chosen among the two suggested in the workflow documentation (contained in the RO). This happens because the RO can be downloaded and the workflow executed locally or in a VM with different operative system. In the case of the Taverna server, the OS is Linux and the correct link to the Fortran executable is http://box.everest.psnc.pl/f/11ef876776/?raw=1. Also the choice of the Seafile folder in which to store the results is required before starting the execution.

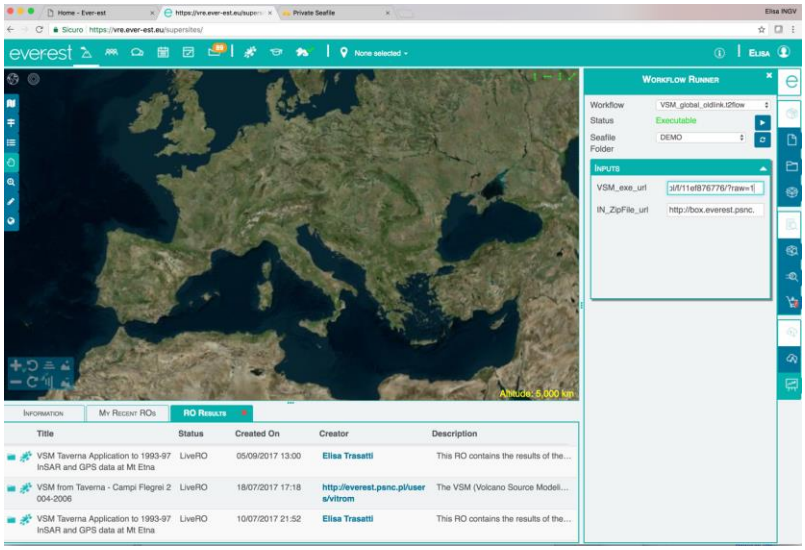

Figure 57 Workflow execution with the Workflow runner

After launching the execution, a green message alerts that the execution started and an orange working icon appears in correspondence of the Workflow Runner icon. When a successful run ends, the icon has a green check.

Figure 58 Icons indicating the running state of the workflow. Left, running; right, successful execution

Use case 3: execution of SARscape on a Windows Virtual Machine

Description of the use case

This use case shows the possibility to retrieve surface displacement due to an earthquake or the ground mean velolcity occurred in a certain temporal interval exploiting SAR (Synthetic Aperture Radar) data by means of the differential SAR interferometry (DInSAR) methodologies.

SAR interferometric techniques combine complex images recorded by antennas at different locations or at different time to form *interferograms* which permit the determination of minute differences in the range (distance) to corresponding points of an image pair, and with a centimeter-to-millimeter accuracy.

In case of an earthquake occurrence considering two acquisitions, one before and one after the event, and forming the corresponding interferogram, then performing some by now consolidated processing steps the surface field displacement can be easily retrieved.

An effective way to study the temporal behavior of the detected phenomena is obviously the generation of time-series that allow us to follow the evolution of the monitored deformations; that is implemented by the known multi-temporal InSAR algorithms. These techniques use a large number of radar acquisitions, considering a lot of the possible interferometric pairs. The information available from each data pair must be properly related to those included in the other acquisitions generating an appropriate sequence of DIFSAR interferograms. Such interferogram stack has to be inverted in order to solve the system of equations; that inversion is obtained via the application of the singular value decomposition (SVD) method.

Data discovery and access

The user signs in the VRE-Supersite. In this use case, user wants to study the effects derived by the occurrence of an earthquake.

To do this, user has to select the two SAR data by means of the dedicated GUI and then import and processing them using the dedicated InSAR software, that is SARscape©.

Firstly, the user defines the Area Of Interest by means of the drawing tools shown in the left side of the display. Then, user has to set the different search parameters in the 'Data Discovery' panel and run the query through the '*Search*' Button.

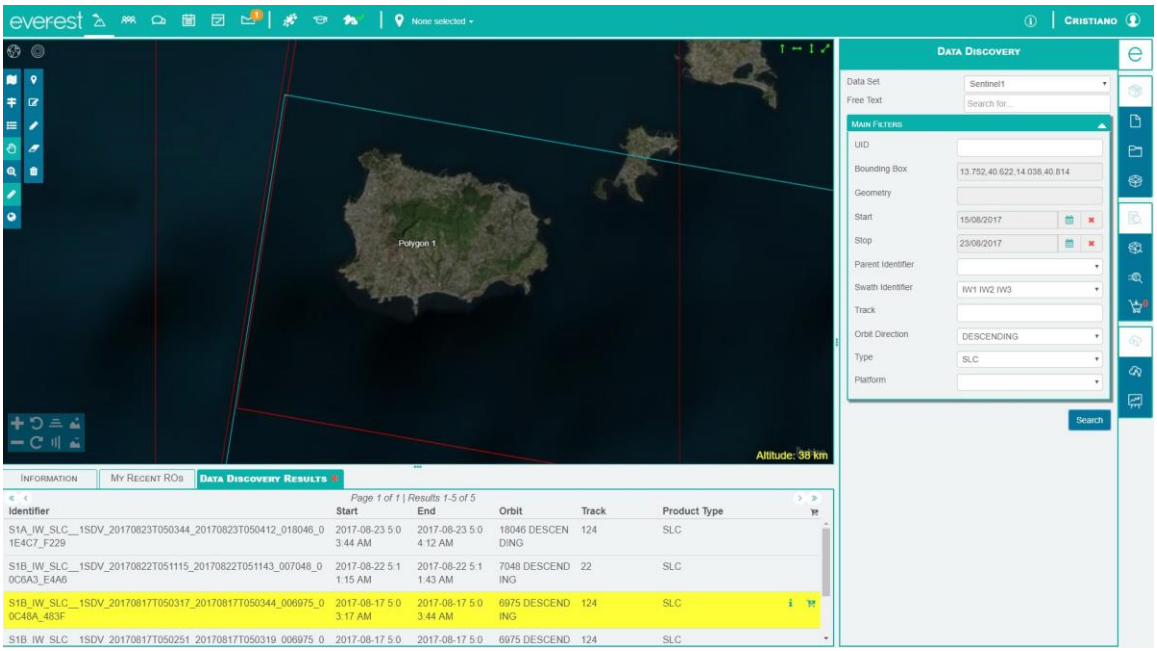

Figure 59 Data Discovery

The results will be displayed in the '*Data Discovery Results*' panel in the below part of the interface. Then selecting the desired items, it is possible to add them to the basket.

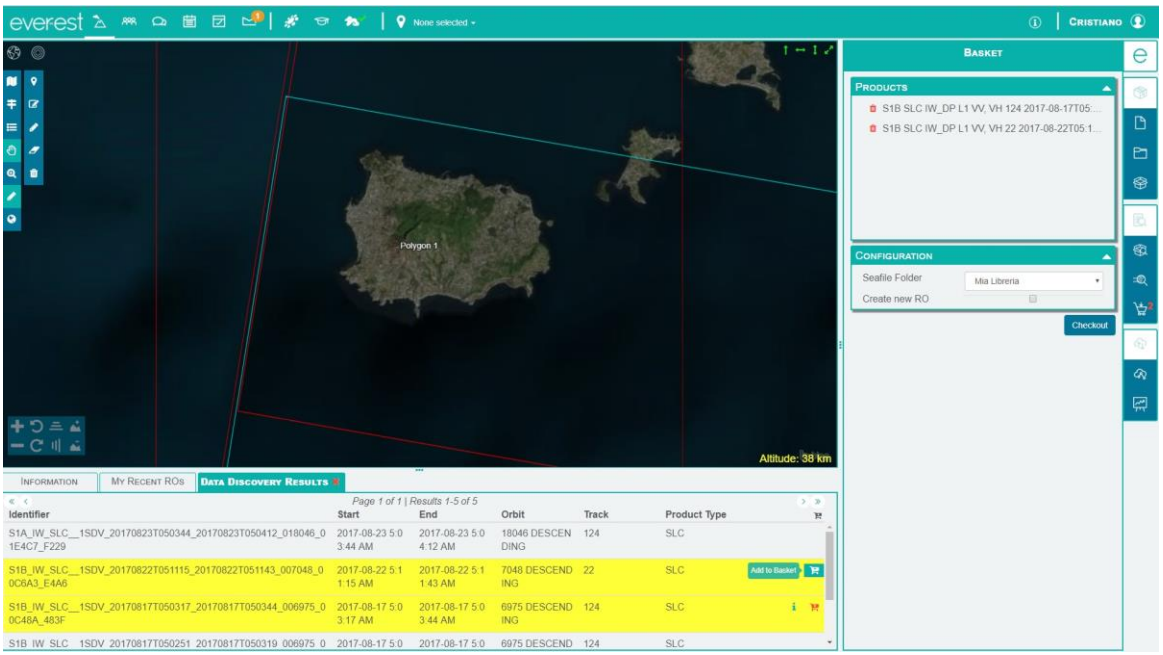

Figure 60 Data Selection

It is also possible to set the output folder on the Seafile repository where data will be stored ('*checkout'* button).

5.1. Access to and management of the Virtual environment

To retrieve the desired products (coseismic displacement or ground mean velocity map) the user takes advantage of a virtual machine (VM) working on the Windows operative system. The user will access to it via remote desktop connection using the dedicated credentials.

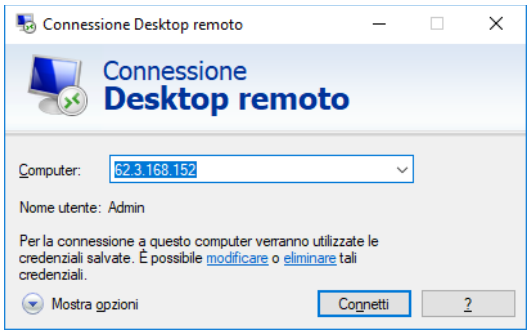

Figure 61 Remote desktop connection

The user will have the availability of the SARscape software to process the SAR data now available in the selected folder (see previous section).

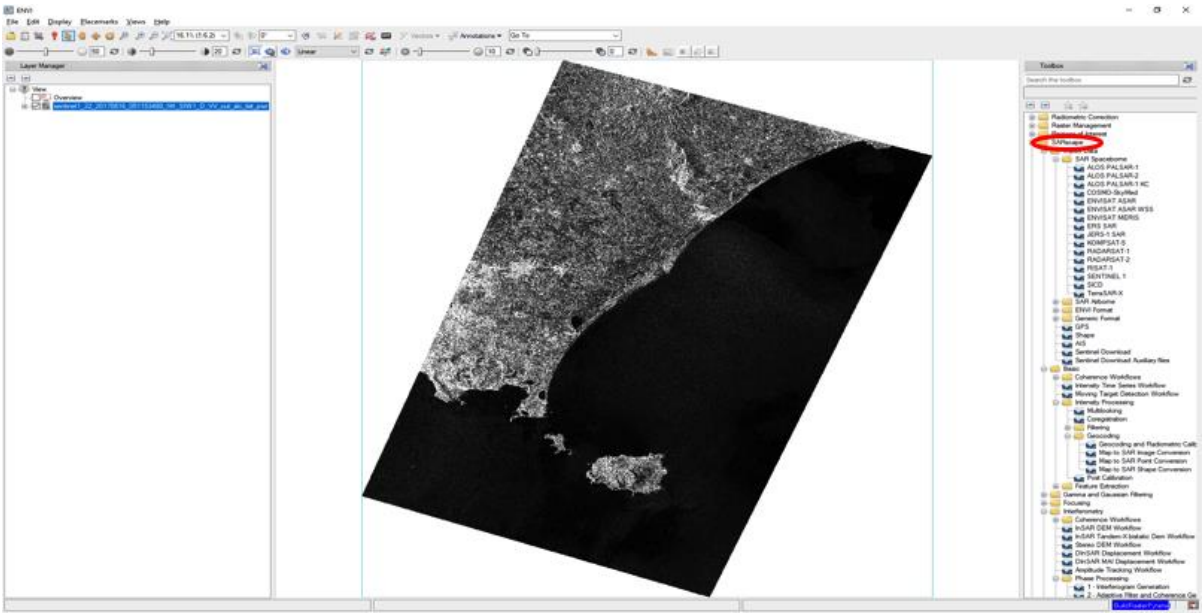

Figure 62 Sarscape Software

5.2. Data processing

Once accessed the VM, the user run the SARscape software and process the data according with the desired methodology (standard interferometry or multitemporal stacking) following the foreseen steps, each with its different settings.

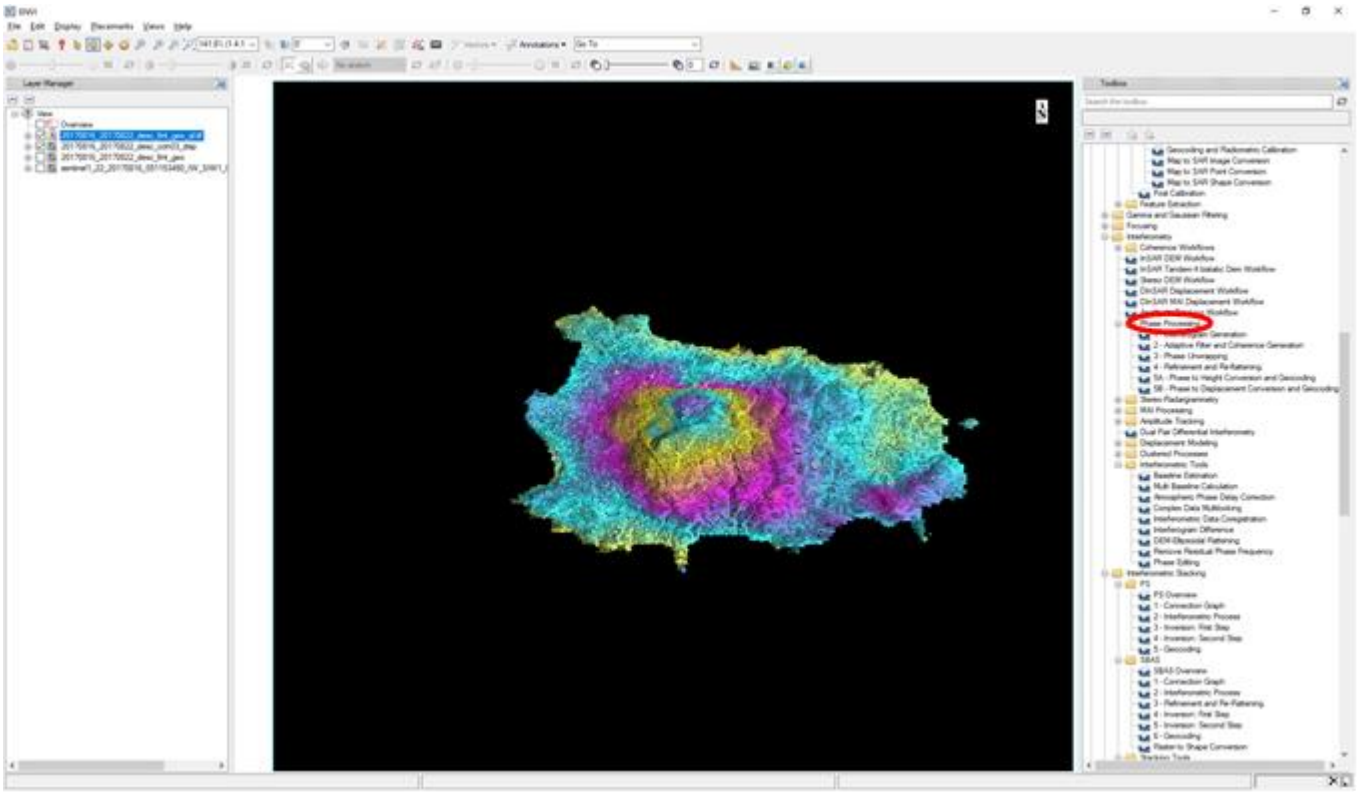

Figure 63 Data Processing

Finally, the output files will be stored in the personal area into Seafile and a download link will be also generated.

5.3. *Management of output results*

Once retrieved the desired results, it is possible to generate a results RO by means of the dedicated panel in the GUI (toolbar on the right side).

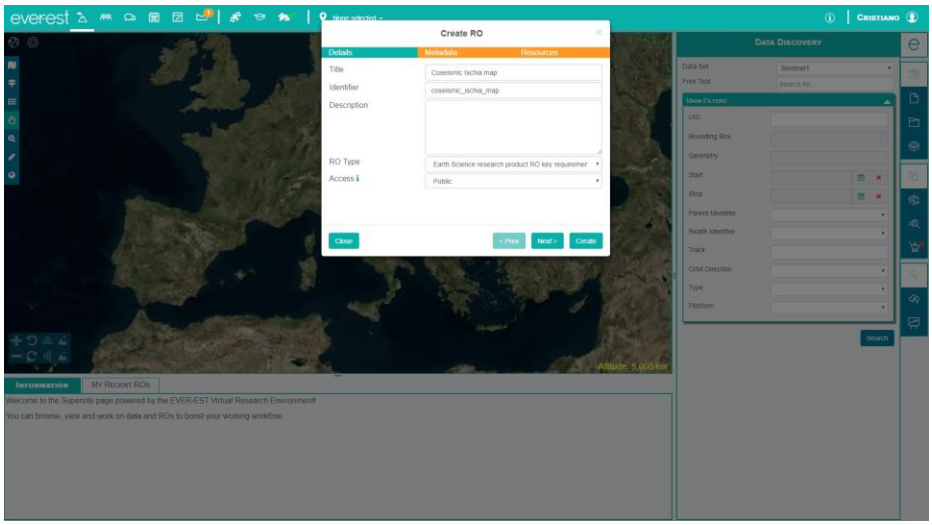

Figure 64 Output Results Management

At the beginning of the RO creation, user must fill the title for the RO, its type and the description. After clicking on the '*Next*' button, user can fulfills the metadata tab. Finally, in the '*Resources*' tab, user can add the obtained results to the RO, using the previously generated link to Seafile.

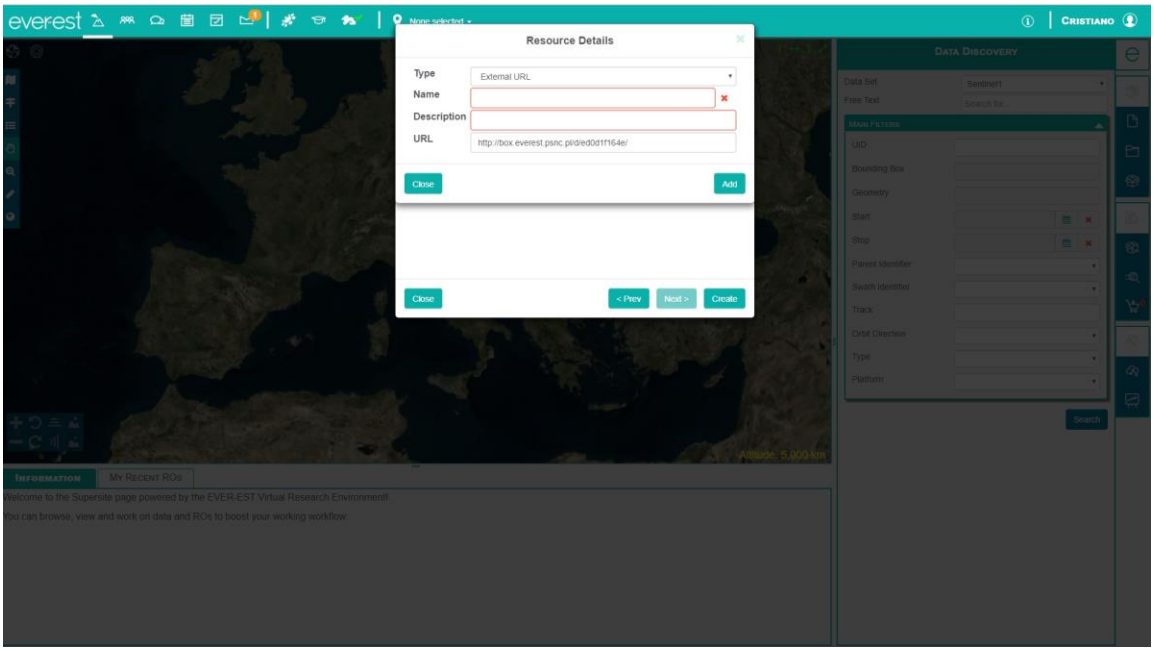

Figure 65 RO Creation

5. Creation of new RO's from the VRE

This chapter is dedicated to the description of a new RO creation from WF outputs. You can use the menu on the right panel (MENU 3) and press the button "New RO".

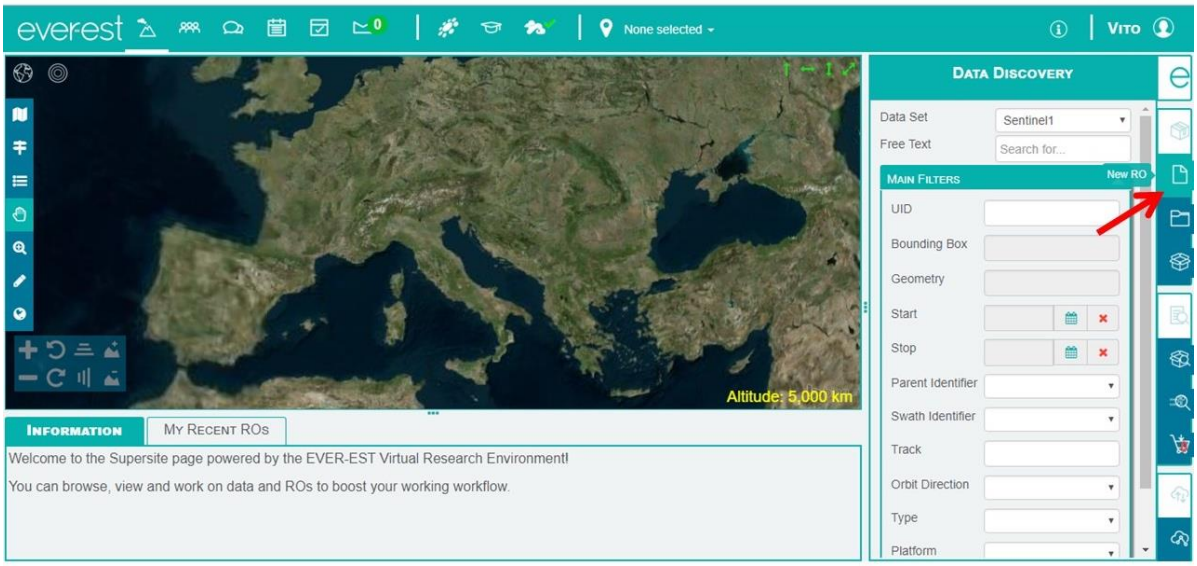

Figure 66 New RO

So, a new window appears with the possibility to create the RO. The first field to compile is the Title that is mandatory and represents the key sentence to discover the new RO on the ROhub portal.

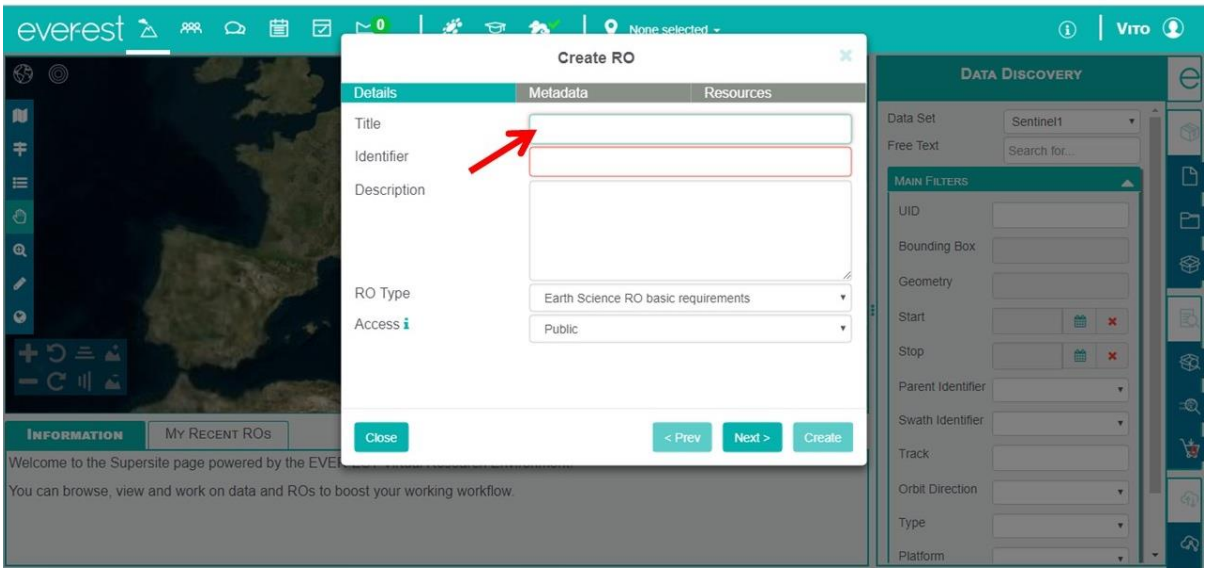

Figure 67 New RO Metadata

In such window, in addition to the title, you can set the Identifier, write a Description, select the RO type and the Access mode (Public or Private). This window consists of three different tabs: Details, Metadata and Resources.

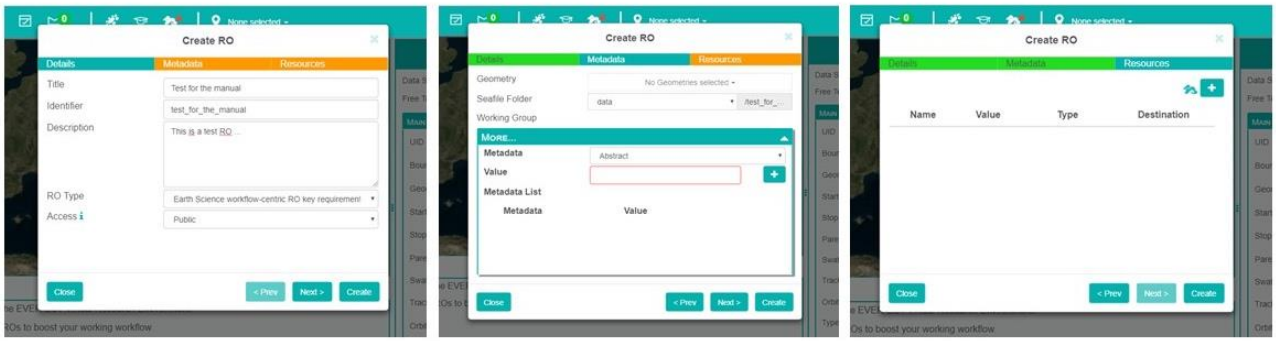

Figure 68 New RO Fields to be filled for metadata generation

You can use the "Next" and "Back" buttons to move among these three tabs. Once set fields in the Details tab, you can go on the Metadata tab to add several types of metadata associated to the RO (i.e. date, format, source etc.). In the Resources tab it is possible to add resources from local files and/or from an External URL (i.e. dataset, documents, files etc.). Now, you can create the new RO by pressing the "Create" button.

References

- Bechhofer, Sean & De Roure, David & Gamble, Matthew & Goble, Carole & Buchan, Iain. (2010). *Research Objects: Towards Exchange and Reuse of Digital Knowledge.* Nature Precedings. doi: 10.1038/npre.2010.4626.1.
- Bechhofer S, Buchan I, De Roure D, Missier P, Ainsworth J, Bhagat J, Couch P, Cruickshank D, Delderfield M, Dunlop I, Gamble M, Michaelides D, Owen S, Newman D, Sufi S, Goble C. (2013). *Why linked data is not enough for scientists.* Futur. Gener. Comput. Syst. 29:599–611. doi: 10.1016/j.future.2011.08.004.
- Guerrieri, L., Merucci, L., Corradini, S., and Pugnaghi, S., (2015). *Evolution of the 2011 Mt. Etna ash and SO2 lava fountain episodes using SEVIRI data and VPR retrieval approach*. JVGR, 291, 63-71.
- Pugnaghi, S., Guerrieri, L., Corradini, S., Merucci, L., Arvani, B.. (2013) *A new simplified approach for simultaneous retrieval of SO2 and ash content of tropospheric volcanic clouds: an application to the Mt Etna volcano.* Atmos. Meas. Tech. 6 (5), 1315–1327.
- Pugnaghi S., Guerrieri L., Corradini S., Merucci L. (2016). *Real time retrieval of volcanic cloud particles and SO2 by satellite using an improved simplified approach*. Atmos. Meas. Tech., 9. doi:10.5194/amt-9-1-2016.

Appendix D – User Actions Template Document

The purpose of the document is to provide a detailed view of the actions performed within each VRC participating to the project. The use cases described in D3.1 have first to be divided into a certain number of Scenarios.

A scenario is an end-to-end story describing a specific activity/workflow performed within the VRC[6].

In this template we describe a couple of scenarios related to the CNR Use Case. Each scenario has to be summarized in max 50 words. Once the scenarios have been identified, for each Scenario, the VRC will have to compile (eventually helped by ambassadors TBC) the precise list of user actions that are describing the workflow step-by-step.

Legenda

In the document, each of the user actions description will have the following layout

Action ID	Action description	Dependencies	Challenge
US_CNR_01			

Action ID: progressive number to list all actions (e.g. US_01, US_02)

Action description: To be written in as → A <**user**> performs an <**action**> to <**achieve a goal**> (optional)

Example: User1 wants to upload data on the server to share it with others

- The main messages are "keep it simple" and iterative specifications. Start from the basics and then release future versions going each time more in detail on some specific aspects. Discuss the document with the technical partners!

- Note: "to achieve a goal" section could be sometimes useless (E.g. of course I am processing data to obtain results). So it can be omitted in most of the cases. Nevertheless in some cases the goal could be revealing a peculiar willing of the user. (E.g. I want to immediately see the results to avoid some information goes public before I can provide my judgement). In this case the goal can be very helpful for the technical team to setup new solutions. So: use it only if you really believe this is adding a new piece of information, feel free to ignore it in all other cases.

Dependencies: a pointer to all the possible dependencies and links related to an action. These include links to Data, Software, Other Infrastructure, etc. whose descriptive information can be inserted in the proper tables which are available in the last section of the document.

[6] **Note**: In an ideal situation the sum of all scenarios shall reflect the entire VRC use case described in D3.1.

Example: User1 wants to upload data on the server to share it with others

Dependencies: Data_01 and Infrastructure_01

"where"

Data_01: Bathimetry data, in ASCII format, contained on a USB stick, total 5GB

Infrastructure_01: Coconet catalogue, accessible via Open Search, contact person Federica Foglini

The goal is to keep the Action Description clean and resolve details by pointing at the proper tables in the following sections.

Note 1: Following the discussion in Venice a specific table on "**Human Control Point**" has been added to allow the writer to specify when, where and how the set of actions has a mandatory need for a human intervention and can't be automatized in any way. What is the action that the person is going to perform (e.g. validation), what are the options (e.g. Approved/Not Approved) and what are the actions that derive (e.g. proceed or go back to User Action N°).

Note 2: The dependency "User Action – GUI requirements" does not refer to the whole GUI concept (which is developed in a parallel thread) but to specific behaviors that are needed in relation to a specific action. I don't have to write as a "User Action - GUI requirement" that I must be able to select a ROI for a search, as it is a general behavior of the interface itself. But I could require a special popup window (e.g.) when another colleague has approved my work with a special message.

Challenge N°: points to the proper challenge described in D3.1 which is addressed by this action. (See the pdf example showing a couple of challenges). It now refers to the general challenges (e.g. EVER-COL_10 see D3.1), potentially divided in sub-challenges for better clarity (e.g. EVER-COL-101, 102, etc. TBD): in the first iteration this column can be left blank and filled in a second moment.

Personas

Personas are used to better focus on the characteristics of the people using the VRE. Use at least one persona for each group of users.

Personas	Representing group	Description
Federica	CNR – Ismar Bologna	Expert user. Creates RO's. Works on data with Fantina
Fantina	CNR – Ismar Venice	Works on data with Federica
Marco	Arpa Bologna	Provides data to Federica, gets results to publish
Roberto	Journalist	Guest user: gets information from Marco and publishes it on the newspapers
Ferdinando	University of Lecce	Works with the Jellyfish app
Francesco	Citizens	Use the Jellyfish app to inform about the presence of jellyfish

User actions examples

Scenarios	Free text description (max 50 words)
1. The ARPA – CNR investigation	CNR and Arpa work together on Sea Environmental monitoring. Every six months ARPA produces a new dataset. CNR processes it and returns to ARPA final result for publication.
2. The Jellyfish app	A crowdsourcing app sponsored by Italian magazine Focus provides scientific data to study jellyfish. CNR wants to fully exploit within the EVER-EST initiative the app DB potential.
3. Third Scenario	

Scenario 1: The ARPA-CNR investigation

Action ID	Action description	Dependencies	Challenge
US_CNR_01	Marco has just finished a new acquisition campaign. He wants to upload the new dataset on the VRE *to obtain the processing (optional)*	Data: DS_2	EVER-COM-10
US_ CNR_02	Federica wants to be automatically notified each time Marco uploads a new dataset *to optimise her work (optional)*		EVER-GEN-10
US_ CNR_03	Federica wants to move the data from Marco in a working directory provided by the VRE		
US_ CNR_04	Federica needs to perform a statistical analysis using Maxent.	Software: SW_01	
US_ CNR_05	Fantina receives the results from Federica and wants to check the results are correct	HI_01	EVER-COL-10

Scenario 2: The Jellyfish app

Action ID	Action description	Dependencies	Challenge
US_ CNR_06	Francesco provides input data using a proper template in the app about presence absence and outbreaks of jellyfish along the Italian coasts on a Virtual Portal (e.g. Meteo Meduse).	Data: DS_03	EVER-GEN-10
US_ CNR_07	Federica wants to be automatically notified each time Francesco uploads new information		
US_ CNR_08	Federica checks and validates the new data	HI_02, Infra_1	
US_ CNR_09	Federica process the data to produce a daily map of distribution for each species sighted	Software: SW_01	
US_ CNR_10	Federica uses the diff function to check the variation in jelly fish distribution	Software: SW_01	EVER-COL-10

<u>Dependencies</u>

Datasets/ Products

N°	Name	Source	Volume	Format
DS_ CNR_1	(e.g.) Bathymetry	(e.g. Infra_1/ DVD/)	1TB	Raster and Ascii

				file
DS_ CNR_2	Arpa Data	USB	10GB	
DS_ CNR_3	Jelly fish occurrences	Information from citizens using the app		

Infrastructures/catalogues

Infrastructure/ Catalogue Name	Infrastructure/ Catalogue Description	Access policy	Com. Protocol	Contacts
Infra_ CNR_1	Coconet	Open access	OpenSearch	Foglini
Infra_ CNR_2	Emodnet	Private Catalogue		

Software requirements

N°	Name	Description	Open source/ IPR	Contacts
SW_ CNR_01	Maxent			Foglini
SW_ CNR_02	ArcGIS			Foglini

Algorithms

N°	Name	Description	Sw Used/Code type	Interaction
Algo_ CNR_01				Automatic/ Manual
Algo_ CNR_02				

User stories-actions GUI requirements

Specific GUI Requirement	Free text description/Image
GUI_ CNR_01	The user interface shall show me a popup window once the processing is over. I shall be able to check the processed image and press “ok” or “re-do” in case the image is correct or not.
GUI_ CNR_02	User interface to populate the Infra_1 with new data using a proper template

Control points

N°	Description	Conditions (YES/NO)	If YES/NO - Go to
HI_ CNR_01	User needs to check if the final map has been correctly plotted Uses GUI_01	YES: User checks the image and confirm it is correct NO: The image is too much dark and provides no information	YES: Press OK button and proceed to the following step in the user stories (go to User Stories N+1) NO: Press “Re-DO” button and Re-Run the processing (go to User Story N-2)
HI_ CNR_02	User need to validate data from the App (are the information reliable?)	Yes: user confirm the data are reliable and correct No: user evaluate the data are not correct	Yes: Press ok and use GUI_02 to populate Infra_1 No User rejects the data