H2020-EINFRA-2015-1

VI-SEEM

VRE for regional Interdisciplinary communities in Southeast Europe and the Eastern Mediterranean

Deliverable D5.4

Report on integrated services and the VRE platform

Author(s): Demetris Antoniades (editor)

Status –Version: Final – d

Date: November 30, 2016

Distribution - Type: Public

Abstract: Deliverable D5.3 – "Report on the integrated services and the VRE Platform" provides a detailed description of the initial domain-specific services and their integration for creation of the VRE platform.

The VI-SEEM Consortium consists of:

GRNET	Coordinating Contractor	Greece
CYI	Contractor	Cyprus
IICT-BAS	Contractor	Bulgaria
IPB	Contractor	Serbia
NIIF	Contractor	Hungary
UVT	Contractor	Romania
UPT	Contractor	Albania
UNI BL	Contractor	Bosnia-Herzegovina
UKIM	Contractor	FYR of Macedonia
UOM	Contractor	Montenegro
RENAM	Contractor	Moldova (Republic of)

IIAP-NAS-RA	Contractor	Armenia
GRENA	Contractor	Georgia
BA	Contractor	Egypt
IUCC	Contractor	Israel
SESAME	Contractor	Jordan

The VI-SEEM project is funded by the European Commission under the Horizon 2020 e-Infrastructures grant agreement no. 675121.

Document Revision History

Date	Issue	Author/Editor/Contributor	Summary of main changes
09-11-2016	a	Demetris Antoniades	Initial version of ToC
21-11-2016	b	Theodoros Christoudias, Michael Alvanos, Stelios Erwtokritou, Panagiotis Charalambous, Georgios Artopoulos, Demetris Antoniades	1st draft
25-11-2016	c	Dusan Vudragovic, Tamas Maray, Stelios Erwtokritou, Demetris Antoniades	2nd draft
30-11-2016	d	Ioannis Liabotis, Demetris Antoniades, Ognjen Prnjat, Evangelia Athanasaki	Final edits

Preface

In the last decade, a number of initiatives were crucial for enabling high-quality research - by providing e-Infrastructure resources, application support and training - in both South East Europe (SEE) and Eastern Mediterranean (EM). They helped reduce the digital divide and brain drain in Europe, by ensuring access to regional e-Infrastructures to new member states, states on path to ascension, and states in European Neighbourhood Policy area – in total 14 countries in SEE and 6 in EM.

This VI-SEEM project brings together these e-Infrastructures to build capacity and better utilize synergies, for an improved service provision within a unified Virtual Research Environment (VRE) for the inter-disciplinary scientific user communities in the combined SEE and EM regions (SEEM). The overall objective is to provide user-friendly integrated e-Infrastructure platform for regional cross-border Scientific Communities in Climatology, Life Sciences, and Cultural Heritage for the SEEM region; by linking compute, data, and visualization resources, as well as services, models, software and tools. This VRE aspires to provide the scientists and researchers with the support in full lifecycle of collaborative research: accessing and sharing relevant research data, using it with provided codes and tools to carry out new experiments and simulations on large-scale e-Infrastructures, and producing new knowledge and data - which can be stored and shared in the same VRE. Climatology and Life Science communities are directly relevant for Societal Challenges.

The driving ambition of this proposal is to maintain leadership in enabling e-Infrastructure based research and innovation in the region for the 3 strategic regional user communities: supporting multidisciplinary solutions, advancing their research, and bridging the development gap with the rest of Europe. The VI-SEEM consortium brings together e-Infrastructure operators and Scientific Communities in a common endeavor.

The overall objective is to provide user-friendly integrated e-Infrastructure platform for Scientific Communities in Climatology, Life Sciences, and Cultural Heritage for the SEEM region; by linking compute, data, and visualization resources, as well as services, software and tools.

The detailed objectives of the VI-SEEM project are:

1. Provide scientists with access to state of the art e-Infrastructure - computing, storage and connectivity resources - available in the region; and promote additional resources across the region.
2. Integrate the underlying e-Infrastructure layers with generic/standardised as well as domain-specific services for the region. The latter are leveraging on existing tools (including visualization) with additional features being co-developed and co-operated by the Scientific Communities and the e-Infrastructure providers, thus proving integrated VRE environments.
3. Promote capacity building in the region and foster interdisciplinary approaches.
4. Provide functions allowing for data management for the selected Scientific Communities, engage the full data management lifecycle, link data across the region, provide data interoperability across disciplines.
5. Provide adequate user support and training programmes for the user communities in the SEEM region.

6. Bring high level expertise in e-Infrastructure utilization to enable research activities of international standing in the selected fields of Climatology, Life Sciences and Cultural Heritage.

The VI-SEEM project kicked-off in October 2015 and is planned to be completed by September 2018. It is coordinated by GRNET with 15 contractors from Cyprus, Bulgaria, Serbia, Hungary, Romania, Albania, Bosnia-Herzegovina, FYR of Macedonia, Montenegro, Moldova (Republic of), Armenia, Georgia, Egypt, Israel, Jordan. The total budget is 3.300.000 €. The project is funded by the European Commission's Horizon 2020 Programme for Excellence in Science, e-Infrastructure.

The project plans to issue the following deliverables:

Del. no.	Deliverable name	Nature	Security	Planned Delivery
D1.1	Project management information system and "grant agreement" relationships	R	CO	M01
D1.2	3-Monthly progress report	R	CO	M03n *
D1.3a	First period progress reports	R	CO	M18
D1.3b	Final period progress reports	R	CO	M36
D2.1	Internal and external communication platform, docs repository and mailing lists	DEC	PU	M02
D2.2	Promotional package	DEC	PU	M04
D2.3	Dissemination and marketing plan	R	PU	M05
D2.4	Training plan	R	PU	M06
D2.5	Promotional package with updates	R	PU	M16
D2.6	1st Dissemination, training and marketing report	DEC	PU	M18
D2.7	2nd Dissemination, training and marketing report	R	PU	M35
D3.1	Infrastructure and services deployment plan	R	PU	M04
D3.2	Service registry, operational and service level monitoring	R	PU	M12
D3.3	Infrastructure overview, assessment and refinement plan	R	PU	M18
D3.4	VRE AAI Model and compatibility with other eInfrastructures	R	PU	M27
D3.5	Final infrastructure overview and assessment report	R	PU	M36
D4.1	Data sources and services deployment plan	R	PU	M06
D4.2	Description of the initial deployed data services	R	PU	M11

D4.3	Description of the final data platform available to VRE users	R	PU	M23
D4.4	Final report on data, services, availability and usage	R	PU	M35
D5.1	Detailed technical implementation plan for VRE services and tools	R	PU	M04
D5.2	Data management plans	R	PU	M06
D5.3	User-oriented documentation and training material for VRE services	R	PU	M13
D5.4	Report on integrated services and the VRE platform	R	PU	M14
D5.5	Final report on integrated services and the VRE platform	R	PU	M36
D6.1	Framework for VRE resource and service provision	R	PU	M09
D6.2	1st Report of open calls and integration support	R	PU	M20
D6.3	Sustainability and business model	R	PU	M24
D6.4	2nd Report of open calls and integration support	R	PU	M36

Legend: R = Document, report, DEC = Websites, patent fillings, videos, etc., PU = Public, CO = Confidential, only for members of the consortium (including the Commission Services).

**n=1,2,3,…12*

Table of contents

References

[1] Project VI-SEEM-675121 – Annex I – Description of the action

[2] VI-SEEM Deliverable D3.1: Infrastructure and services deployment plan

[3] VI-SEEM Deliverable D4.2: Description of the initial deployed data services

[4] VI-SEEM Deliverable D5.1: Detailed technical implementation plan for VRE services and tools

[5] VI-SEEM Deliverable D5.2: Data management plans

[6] VI-SEEM Deliverable D5.3: User-oriented documentation and training material for VRE services

[7] Joomla framework
https://framework.joomla.org/

[8] Bootstrap library
http://getbootstrap.com

[9] VI-SEEM code repository
https://code.vi-seem.eu

[10] Live Access Server
http://las.vi-seem.eu/las

[11] Clower
http://dchrepo.vi-seem.eu

[12] ChemBioServer
http://bioserver-3.bioacademy.gr/Bioserver/ChemBioServer/

[13] VI-SEEM training portal
https://training.vi-seem.eu/

[14] Docker
https://www.docker.com/

[15] Google Analytics
https://analytics.google.com/analytics/web/

[16] VI-SEEM data repository
https://repo.vi-seem.eu/

[17] VI-SEEM Service Catalogue and Portfolio
https://services.vi-seem.eu/ui/catalogue/services/

[18] VI-SEEM Wiki
http://wiki.vi-seem.eu/index.php/Main_Page

[19] VI-SEEM Login Service
https://services.vi-seem.eu/ui/catalogue/services/#VI-SEEM_Login

List of Figures

List of Tables

Glossary

AMBER	Assisted Model Building with Energy Refinement molecular simulation programs
CBIR	Content-Based Image Retrieval
COSMO	Consortium for Small-scale Modeling
CPU	Central Processing Unit
DICOM	Digital Imaging and Communications in Medicine
DREAM	The Dust REgional Atmosphere Model
ECHAM	Global Climate Model developed by the Max Planck Institute for Meteorology
EMAC	ECHAM/MESSy Atmospheric Chemistry
ERT	Electrical Resistivity Tomography
FERRET	Interactive computer visualization and analysis environment
FFTW	Fastest Fourier Transform in the West, library for computing the
FIRFLY	Ab initio and density functional theory chemistry program
GAMESS	General Atomic and Molecular Electronic Structure System is a general ab initio quantum chemistry package
GATK	Genome Analysis Toolkit
GIS	Geographic Information System
GrADS	Grid Analysis and Display System
GROMACS	Molecular Dynamics Software Toolkit
GUI	Graphical User Interface
HPC	High Performance Computing
IDL	Interactive Data Language, a programming language used for data Analysis
LAS	Live Access Server
MEDICI	A multimedia content management system
MESSY	Modular Earth Sub-model System
MM5	The PSU/NCAR mesoscale model
NAMD	Scalable Molecular Dynamics Toolkit
NCL	NCAR Command Language
NetCDF	Network Common Data Form
NWCHEM	High Performance Computational Chemistry Software
OPENFOAM	Open source Field Operation And Manipulation toolbox for continuum Mechanics

PIDs	Persistent Identifiers
RegCM	The Regional Climate Model system
SEEM	South East Europe and Eastern Mediterranean region
SOL	Soft Ontology Layer
VI-SEEM	VRE for regional Interdisciplinary communities in Southeast and the Eastern Mediterranean
VRE	Virtual Research Environment
WEST	Wind Energy Simulation Toolkit
WRF	Weather Research and Forecasting Model

Executive summary

What is the focus of this Deliverable?

The focus of this deliverable is to describe the status of the VI-SEEM integrated services, the design of the VRE platform and to showcase how the integrated services are included in the VRE platform.

What is next in the process to deliver the VI-SEEM results?

The contents of this deliverable will form a basis for the other work packages to make efficient use of the integrated service and the VRE platform. The VRE platform whose design and implementation is presented in this deliverable, provide access to both the common services (WP3 for e-Infrastructure and WP4 for data services) that are used by all three scientific communities, as well as the application-specific services that are deployed in WP5.

In particular, the content of this deliverable will be used in following VI-SEEM activities:

- WP2.2 – Dissemination and marketing
- WP2.3 – Training
- WP2.4 – Innovation management
- WP3.3 - Operations and resource management of e-Infrastructure communities
- WP3.4 – Authentication, authorization and access management
- WP4.3 – Data access, preservation and reuse
- WP4.4 – Data analysis
- WP5.1 – Refinement of service requirements & tech assessment for integration
- WP5.4 – Overall integration of services
- WP5.5 – Scientific support
- WP6.2 - Implementation of open calls for access to VRE resources
- WP6.3 - Integration of the new applications to the VRE environment
- WP6.4 - Sustainability and business model

What are the deliverable contents?

The deliverable contents include an overview of the main structure of the VRE platform, and how it integrates the various types of VI-SEEM services - starting with the generic VI-SEEM services. It also provides information about the domain-specific services that are dealing with the integration of domain-specific workflows, software tools and application codes. Furthermore, it provides a list of the integrated application-level services. For each service available in the VRE, information on how a user can access the existing service and contribute to updating existing service is given; as well as instructions for including new services. Finally, a detailed list of expected upgrades and new services is given. In detail, the deliverable includes:

- Description of the VI-SEEM Virtual Research Environment Platform (chapter 2)
- Description of the VI-SEEM scientific application environment as a domain specific service addressing the needs of each of the three project scientific communities and their applications (chapter 3)

- Description of the VI-SEEM workflows, software tools and codes repository as a domain-specific service addressing the needs of each of the three project scientific communities (chapter 4)
- Description of the VI-SEEM regional community datasets as a domain-specific service addressing the needs of the three project scientific communities and their data (chapter 5)
- Description of the VI-SEEM application-level services (chapter 6)
- Plan for the update of existing and integration of new services (chapter 7)

Conclusions and recommendations

The deliverable is the main output of T5.3 “Development of the VRE platform” and T5.4 “Overall integration of services” and describes the design behind the VRE platform and how this reflects the integration of the generic and domain-specific VI-SEEM services. The deliverable describes the different sections of the VRE platform, illustrating how each section manages to integrate the VI-SEEM offered services. Details on how the user can easily get information and use the VI-SEEM provided services are given along with the current services’ integration status. Detailed instructions on how the user can integrate hers/his own applications and services to the VI-SEEM Virtual Research Environment are given in each section.

The deliverable and the VRE platform in general will act as a guideline for further integration of applications and services. This will allow application developers participating in the VI-SEEM integration phases to make their tools, datasets and services available to the wider community.

1 Introduction

VI-SEEM brings together service providers, service enablers/experts and scientists from the three relevant to the project scientific communities, to develop the platform for the provision of a Virtual Research Environment (VRE). The VRE provides the framework and the entry point for the scientific community users, where data and analysis tools can be promoted and accessed. The data, application and services are relevant to the three scientific communities: climate, digital cultural heritage and life sciences. While initially non-existent or scattered around the different communities' local infrastructures, the VI-SEEM VRE integrates them in a stable and secure regional infrastructure providing a consolidated access point for them to the users.

This document describes the integration of the generic and domain specific VI-SEEM services and the VRE platform. The VRE platform is user-tailored portal that connects the unified infrastructure developed in WP3, the generic data services deployed in WP4 and the domain specific services, tools and APIs integrated through WP5.

We first describe the structure and the design of the VRE platform in chapter 2. In this chapter we provide details on how the user can get access to the generic infrastructure and data services and how she/he can get information for the three scientific communities supported by the project.

The following chapter provides more details on the domain-specific services offered by the project and the current integration status as made possible through the VI-SEEM integration phase described by D5.1 [5]. Chapter 3 describes the VI-SEEM scientific application environment where the user can get information on optimized applications and tools available and supported in the e-Infrastructure. Such applications are available in the HPC or Grid sites provided by the project partners. Additionally, the service provides ready-to-use Virtual Machine images that the user can instantiate to familiarize herself and run scientific community specific tools in the regional Cloud infrastructure. Chapter 4 describes the workflow and software tools provisioning services where the user can get access to manual or automated scientific workflows and scientific codes provided by the VI-SEEM communities. Chapter 5 describes the regional community datasets domain specific service where the user can access available scientific datasets. Chapter 6 describes a number of services grouped together under the VI-SEEM application-level services umbrella. These services form a range of web-based or visualization services providing easy access to underlying workflows, applications and resources for each regional community. For each of the available services details on how a user can access the available material through the VRE platform are provided as well as details on how a user can contribute to the VRE by providing new software modules, workflows and datasets and new application-level services. Finally, in chapter 7 we describe the forthcoming updates to the VI-SEEM domain specific services, as these are planned through the WP5 integration phases. Additionally, we provide a list of new application-level services that are expected to be deployed during the 2nd year of the project.

2 VI-SEEM Virtual Research Environment platform

In this section, we provide a description of the VRE platform made available in M12 of the project. We describe the design of the VRE platform and how it integrates the VI-SEEM domain specific services.

2.1 Scope

The VRE portal provides access to the VRE services and applications. The portal connects the unified infrastructure developed in WP3, the generic data services deployed in WP4 and hosts the user-tailored front-end interface. This integrated platform enables easy communication and data sharing between the various user groups. The VI-SEEM VRE portal is available at http://vre.vi-seem.eu/.

2.2 VRE platform structure

The VRE homepage (Figure 1) provides quick access to the VI-SEEM Virtual Communities information, access to VI-SEEM resources description as well as access to the VI-SEEM services, through clearly annotated images and buttons which direct the user to the relevant sections of the portal or other VI-SEEM websites (such as the VI-SEEM training portal). Through the VRE homepage the user can navigate to all available information provided by the VI-SEEM Virtual Research Environment.

The “Access to resources” section provides information about the available compute and data resources, listing available optimized applications and libraries in the infrastructure, and information on how the user can acquire access. Additionally it provides a link to the VI-SEEM training portal [14] where the user can find training and documentation material for the applications, libraries, infrastructure usage and datasets available. “Access to compute resources” section integrates the VI-SEEM unified infrastructure generic service, developed by WP3, by providing information for the available HPC, Grid and Cloud infrastructures [3]. “Access to data resources” integrates the VI-SEEM generic data services deployed by WP4 and listed in Deliverable 4.2 [4]. The training portal evolves with the VI-SEEM integration phases and VI-SEEM training workshop, constantly being updated with more information, howtos and documentation material for relevant tools, workflows and datasets. A list of the currently available information in the VI-SEEM training portal is available in D5.3 [7].

The “Contribute to the Virtual Research Environment” section provides direct links to step-by-step guidelines on how a user can contribute to the VI-SEEM scientific application environment, making software tools, workflows, datasets and application-level services available to the community.

Finally, the “Virtual communities” section provides direct access to the available codes, workflows, datasets and application-level services for each of the VI-SEEM scientific communities.

Figure 1 - The VRE portal homepage

The VRE menu (depicted at the top of Figure 1) is structured in a way that provides instant access to different information available to the VRE users. Each of the items corresponds to a VI-SEEM domain specific service, with sub-categories corresponding to the different tools, software and datasets available to the user and/or the supported scientific communities. The menu has the following four main sections:

- Scientific application environment: this section provides access to several modules relevant for the work of the regional scientific communities of interest – such as optimized applications and libraries, virtual machine images and developed use cases. The following submenu items are provided:
    - How to contribute: providing detailed information on how the user can contribute to the environment, making software, use cases and Virtual Machines available to the community
    - Optimized application and libraries: listing the application and libraries available to the VI-SEEM integrated infrastructure services, along with documentation and training material
    - Virtual machine images: Providing access to virtual machine images tailor made for specific scientific communities
    - Developed use cases: providing information on how different communities utilize the VI-SEEM integrated environment for developing specific scientific use cases
- Workflow and software tools: this section provides access to several modules such as documents containing best practice procedures and workflows for the production of scientific results relevant to the application categories identified in the region. As this service is being implemented mainly via the use of the VI-SEEM code repository it is also integrated with the VI-SEEM Login service [19] that provided the basic AAI infrastructure for the project. The following submenu items are provided:
    - How to contribute: providing detailed information on how the user can contribute her own workflows and software tools and make them available to the community
    - Scientific workflows: providing access, documentation and training material for the available workflows, developed through the VI-SEEM integration phases and made available to the community
    - Code repository: providing access to the available software tools developed through the VI-SEEM integration phases and made available to the community.
- Regional community datasets: this section provides datasets (which can take the form of scientific data, publications and simplified data) of regional importance for the scientific communities. Given the VRE is a public website, these communities can be VI-SEEM collaborators, worldwide scientific communities or any member of the general public if the associated license of the data sets permits. To secure the VI-SEEM datasets that are available in the VI-SEEM repository the service is integrated with the VI-SEEM Login service that provides the basic AAI mechanism covering the users of the region. The following submenu items are provided:
    - How to contribute: providing detailed information on how the user can make its datasets available to the community, listing all available steps and information needed in order for the dataset to be fully utilized and lead to new and reproducible research
    - Climate datasets: lists the available datasets for the climate scientific community
    - DCH datasets: lists the available datasets for the digital cultural heritage scientific community

    - Life sciences datasets: lists the available datasets for the life sciences scientific community
- Application-level services: In this section users can find a range of web-based or visualization services providing easy access to underlying workflows, applications and resources for each regional community. The following submenu items are provided:
    - How to contribute: providing detailed information on how the user can make its application-level service available to the community, listing all available steps and information needed
    - Climate: lists the application-level services related to the climate scientific community
    - Digital cultural heritage: lists the application-level services related to the DCH scientific community
    - Life sciences: lists the application-level services related to the life sciences scientific community

Each section, along with the current state of the integration process, is further described in the following sections.

2.3 Technical details

The VRE is hosted by The Cyprus Institute upon a Docker [15] container which runs in a virtual machine and since its first deployment has had about 98% uptime. The VRE website is built using the Joomla content management system [8] enhanced with the Bootstrap front-end framework [9]. The service is deployed using an Apache webserver and a MySQL Database. An SSL certificate is also installed to allow for secure connections between guests of the VRE and the web server.

Google analytics [16] have also been enabled to monitor the visits to the VRE, allowing detailed access analytics for the different services provided by the VI-SEEM VRE in order to optimize and improve. Figure 2 illustrates the geographical access patterns and visitor statistics for the VRE platform over the last month.

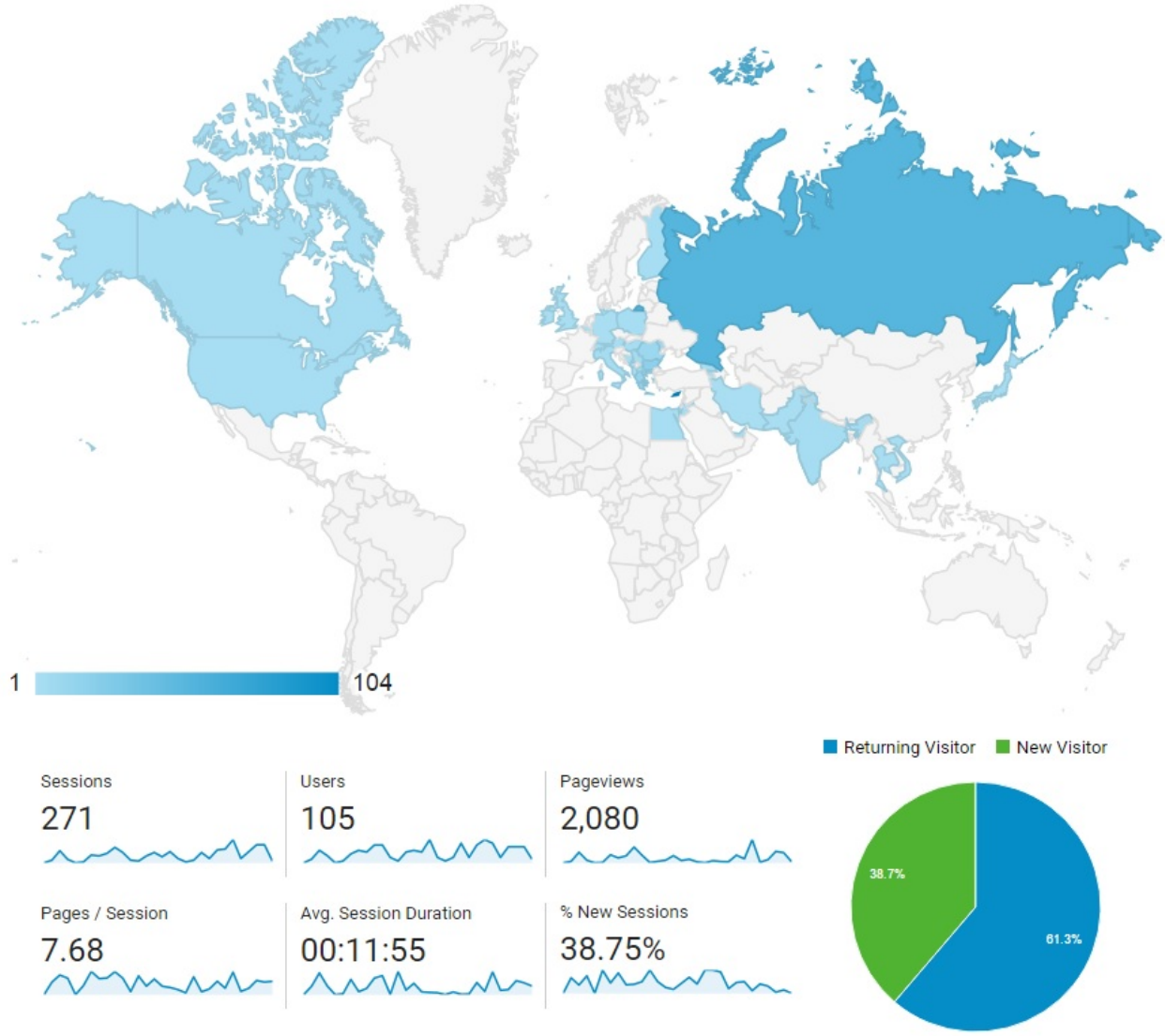


Figure 2 - Google analytics data for the VRE October 26 - November 23 2016

3 VI-SEEM scientific application environment domain specific service

3.1 Scope and service description

The VI-SEEM scientific application environment is a domain specific service provides the VI-SEEM user community with access and information to several modules including optimized applications and libraries, Virtual Machine (VM) images and a full list of available codes, relevant for the work of the three regional scientific communities of interest. The modules that are readily available, and additional modules expected to be provided in the future by the VI-SEEM community for each support scientific community are listed in the following sections. More details can be found in D5.1 [5].

The services provided through the VRE scientific application environment service include:

- Optimized applications and libraries
- Virtual Machine (VM) images, and
- Developed use cases from the three scientific communities.

The applications libraries include short descriptions and access/documentation details for the specific scientific communities service modules provided by the VI-SEEM scientific application environment as well as the e-Infrastructure resources which support them. The VM section provides access and documentation details for the available VM images provided by the VI-SEEM scientific application environment and the computational resources which support them. Finally, the service provides a list of specific developed scientific applications available through the computational facilities of the VI-SEEM scientific community.

The service is available at https://vre.vi-seem.eu/index.php/scientific-application-environment

3.2 Services integration

The VI-SEEM scientific application environment consists of a collection of scientific service modules provided as a ready-to-use production environment in a number of VI-SEEM partner sites. These scientific modules are made available through the VI-SEEM integration phases described in D5.1 [5], requested from the specific use cases participating in the integration phases. As a part of WP3, continuous updates of the VI-SEEM infrastructure are performed in order to make the software modules available in as many infrastructure sites as possible. Table 1 lists the optimized applications and libraries that are currently available in the VI-SEEM scientific application environment.

Climate	CCLM	DREAM	EMAC	FERRET	GIS
	GRADS	IDL	MM5	NCL	OPENFOAM
	Paraview	R	RegCM	WRF	WRF-CHEM
Digital Cultural Heritage	3DINV	AutoGR	CH-CBIR	CLOWDER	Soft Ontology Layer (SOL)
Life Sciences	AMBER	Abogardo	BioCORE	CellProfiller	CRYSTAL
	Desmond	FFTW	FIREFLY	GAMESS	GROMACS
	ImageJ	Image-Pro	JVM	LabVie	Maestro
	Molekel	NAMD	OpenCV	PDB	Pymol
	R bbb	Rasmol	TCL	Vaa3D	VMD
	WinCoot				

Table 1 - Available optimized applications and libraries

Virtual machine images have also been provided by the scientific communities. Currently two virtual machine images are available:

1. The Live Access Server VM: a ready to use VM instance of the Live Access Server Climate Scientific Community Application-level service
2. Clowder: pre-configured and ready to use Docker [15] components for installing and using the Clowder digital cultural heritage research data system.

3.3 Access through the VRE platform

3.3.1 Accessing the scientific application environment

The main page of the VI-SEEM scientific application environment (Figure 3) provides to the user general information about the scientific environment and easy access to the different categories of software modules and tools hosted. Each section then contains specific information for the three different scientific communities, namely climate, digital cultural heritage and life sciences.

In the optimized applications and libraries category, the service provides a list of available software for the three scientific communities. The climate and weather forecast models are large community codes that were developed over long time periods by large international user communities. The cultural heritage user community employs a number of codes and models, which are summarized in the table below. The life sciences research community in the SEEM region is using a variety of codes and tools to simulate the processes under investigation. For each scientific community a list of the available software modules is available as illustrated in Figure 4. For each module additional

information is provided in a separate page. This information includes general description of the module, links to documentation and training material and a list of the VI-SEEM infrastructures that are currently supporting the module. An example of such page for the weather research and forecast model is illustrated in Figure 5. More material is uploaded as the integration phase progress and training and documentation material is made available.

Figure 3 - VI-SEEM scientific application environment service

Climate

Climate and weather forecast models are large community codes that were developed over long periods of time by large international user communities. All models below are free to use for research purposes and some are available open-source. EMAC requires a license, which is free of charge. A number of more specialized models are used in some communities.

Climate Supported Software Modules				
CCLM	DREAM	EMAC	FERRET	GIS
GRADS	IDL	MM5	NCL	OPENFOAM
Paraview	R	RegCM	WRF	WRF-CHEM

Climate and Weather forecasting models are clearly the computational focus of this user community, however,other software related to data analysis and visualization is also used. Examples for popular free and/or open source visualization software are:

- The Grid Analysis and Display System (GrADS) is an interactive desktop tool that is used for easy access, manipulation, and visualization of earth science data.
- ParaView is a multi-platform data analysis and visualization application. ParaView users can quickly build visualizations to analyse their data using qualitative and quantitative techniques.
- FERRET is an interactive computer visualization and analysis environment designed to meet the needs of oceanographers and meteorologists analysing large and complex gridded data sets.

Cultural Heritage

The user community employs a number of codes and models which are summarized in the table below. These codes and tools are integrated into the VRE providing user friendly access to the groups in the SEEM region.

Cultural Heritage Supported Software Modules				
3DINV	AutoGR	CH-CBIR	CLOWDER	Soft Ontology Layer (SOL)

In addition to the list given this table, the project partner SESAME is supporting a number of software and applications for data analysis for current and future needs, associated with Cultural Heritage, such as OMNIC, Unscrambler, Igor, PyMCA, IFEFFIT, FEFF, WINXAS, Matlab, PEAKFI.

Life Sciences

The Life Sciences Research Community in the SEEM region is using a variety of codes and tools to simulate the processes under investigation.

Life Sciences Supported Software Modules				
AMBER	Avogadro	BioCORE	CellProfiller	CRYSTAL
Desmond	FFTW	FIREFLY	GAMESS	GROMACS
ImageJ	Image-Pro	JVM	LabView	Maestro
Molekel	NAMD	OpenCV	PDB	Pymol
R bbb	Rasmol	TCL	Vaa3D	VMD
WinCoot				

Figure 4 - List of scientific modules and tools available in the VI-SEEM infrastructure listed by scientific community

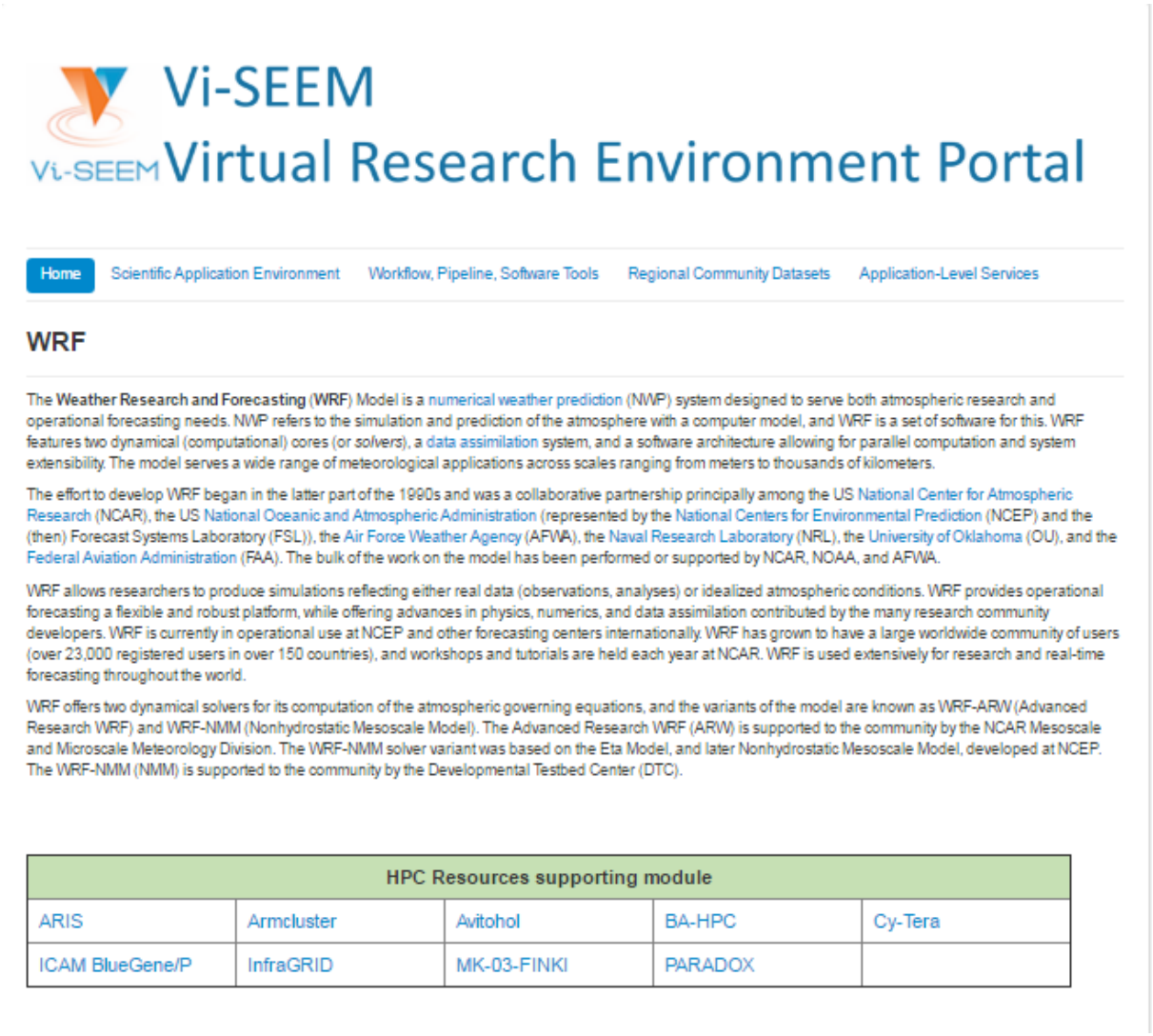


Figure 5 - Example of software module specific information for the WRF model

Regarding virtual machines (VMs), the VRE platform provides description and access details for the available VMs for the scientific communities. The VM information includes the operating system version, the software packages installed, the running services, and access instructions. Figure 6 provides an example of how the VRE platform presents and provide access to the available virtual machines. The example illustrates how the user can get information and the required Docker components for setting up and running a version of the Clowder research data system.

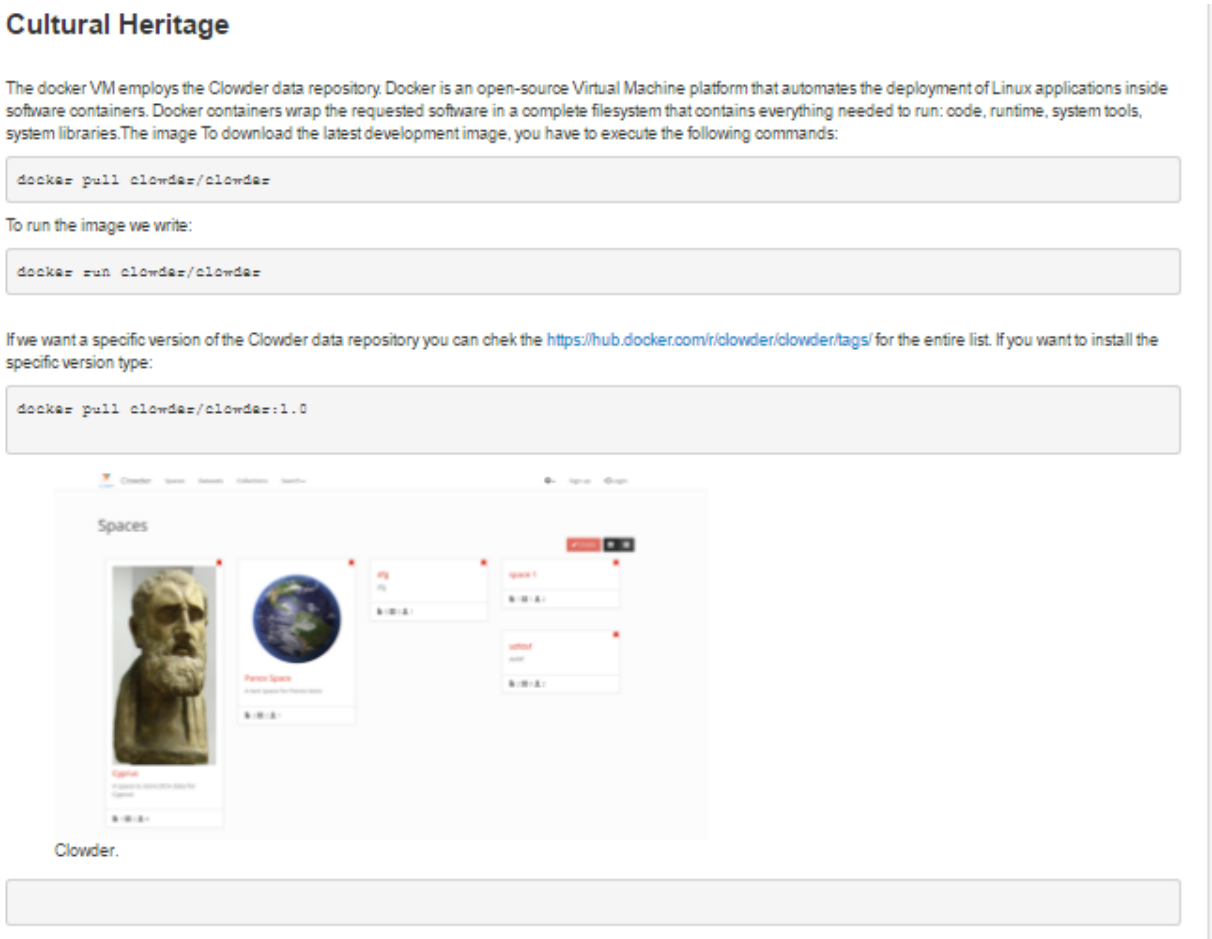


Figure 6 - Example of a Cultural Heritage available virtual machine through the VI-SEEM scientific application environment

Finally, the VI-SEEM scientific application environment provides information regarding the use cases developed through the VI-SEEM integration phases. This page lists scientific specific use cases of development or usage of software modules among scientific partners. In each phase of the project, the scientific partners verify the integration, applicability and readiness for use of a number of optimized applications. For each scientific community a summary table with the list of the developed cases, the integration phase, completion date, country, institute, and the scientific contact is given. An example for the case of the life sciences scientific community is illustrated in Figure 7 The interested user can the follow the link to the specific use case of interest for more information, which includes a small description regarding the scope of the module, developers, and contact people (e.g. in Figure 8). Thus, if any user interested in using of any of these developed cases, she is able to find relevant contact points via the VRE portal.

Life Sciences Scientific Community

	Application Acronym	Integration Phase	Integration Completion	Country	Responsible Partner	Institute	Scientific Contact
Life Sciences	MD-Sim	1st	M12	Greece	GRNET	Biomedical Research Foundation Academy of Athens	Zoe Cournia
	PSOMI	1st	M12	Montenegro	UOM	Faculty of Natural Sciences, University of Montenegro	Miljan
	CCC	2nd	M18	Jordan	SESAME	Computer and Information Technology, Tafila Technical University	Mohammad Alfraheed
	SEMaCD	2nd	M18	Serbia	IPB	Institute of molecular genetics and genetic engineering Belgrade	Djordje Francuski
	THERMOGENOME	2nd	M18	Bulgaria	IICT-BAS	Institute of Molecular Biology Bulgarian Academy of Sciences	Stoyno
	MS4DD	2nd	M18	Bulgaria	IICT-BAS	Institute of Molecular Biology Bulgarian Academy of Sciences	Nicolay Dodoff
	DICOMNetwork	2nd	M18	Moldova	RENAM	Institute of Emergency Medicine,	Diana Zagadailo Natalia Golubev
	CNCADD	2nd	M18	FYR of Macedonia	UKIM	Faculty of pharmacy, UKIM	Marija Glavash Dodov
	SQP-IRS	3rd	M28	Jordan	SESAME	SESAME	Gihan Kamel
	BioMoFS	3rd	M28	Armenia	IIAP-NAS-RA	Yerevan State University, Armenia	Yevgeni Mamasakhlisov
	MDSMS	3rd	M28	Armenia	IIAP-NAS-RA	The International Scientific-Educational Centre of NAS RA, Bioinformatics Department	Armen Poghosyan
	NGS1	3rd	M28	Cyprus	CVI	Cyprus Institute of Neurology and Genetics	Athina Theodosiou
	NGS2	3rd	M28	Cyprus	CVI	Cyprus Institute of Neurology and Genetics	Petros Mina

Figure 7 - List of the Life Sciences specific use cases developed over the VI-SEEM scientific application environment

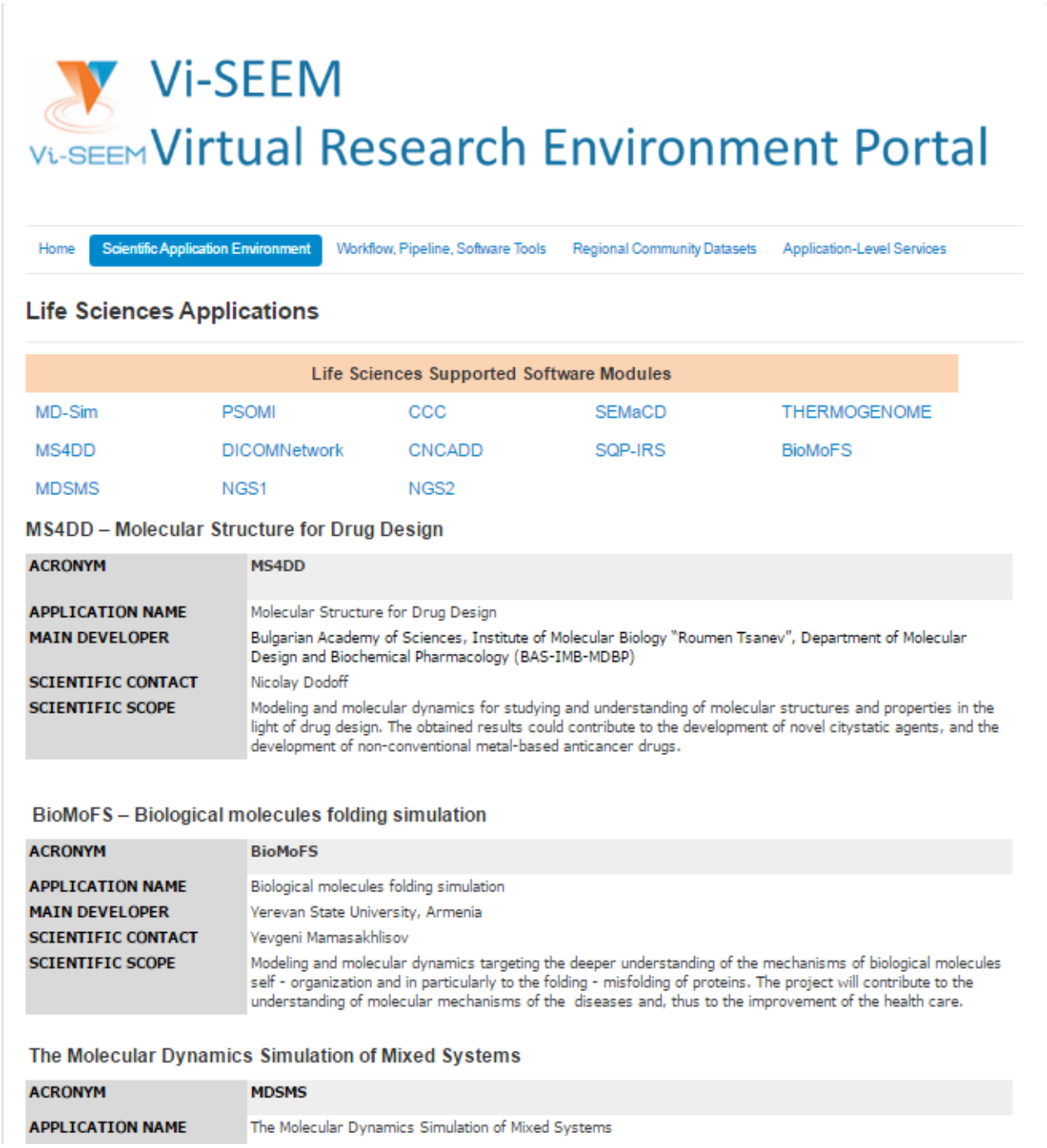


Figure 8 - A detailed view of the information provided for the life sciences use cases developed over the VI-SEEM scientific application environment

3.3.2 Contributing to the scientific application environment

To streamline the procedure of contributing to the VI-SEEM scientific environment, each project partner assigns a service enabler who is responsible to coordinate and assist researchers from their own country during the process of the service integration. In particular, each service enabler is responsible to:

- Explain the integration procedure to researchers and ensure that the integration follows the agreed timelines as set in the integration plan.
- Assist researcher's access to the VI-SEEM infrastructure and ensure the smooth initiation of service integration.
- Provide technical support and address any problems encountered during the integration possibly with the support of other experts from the project.
- Ensure that researchers receive all required support from the partners that provide the computing resources.

The researcher(s) should provide a description of each service module and its use cases on the VRE portal, with links to accompanying documentation, training material, and contact details of the responsible scientist.

Technical documentation must be made available in the VI-SEEM wiki [19], and must entail instructions on how to access, compile/deploy and/or use each service module. Furthermore, the training material will cover scientific aspects of the application, and must be hosted on the VI-SEEM training portal. Finally, the researcher(s) should provide the contact details of the person(s) responsible for providing scientific support to future users of the VRE. The responsible scientist for each service module will then need to make sure that each service module, as defined in the VRE integration plan, is set up correctly, and that it works and is accessible to users as expected.

The integration process and specific checkpoints to be followed when contributing to the scientific environment are available at the VRE platform at https://vre.vi-seem.eu/index.php/scientific-application-environment/how-to-contribute and illustrated in Figure 9.

Figure 9 - Guidelines for contributing to VI-SEEM scientific application environment

4 VI-SEEM workflows and software tools domain specific service

4.1 Scope and service description

This service provides access and information to several modules such as documents containing best practice procedures and workflows for the production of scientific results relevant to the application categories identified in the SEEM region. The service describes and documents scientific workflows provided by the VI-SEEM platform. Similar details are also provided for the list of codes that are available and can be downloaded from the VI-SEEM code repository. Details of how researchers can contribute code are given. Details for contributing workflows – which can take the form of scripts, which automate data generation and processing, interfaces that ease the use of tools and software applications and documents that describe scientific processes, are also provided.

Scientific workflows are made available either as ready-to-use tools that automate the process or as documents describing a specific process. The former category is made available through the VI-SEEM code repository [10], while the latter is uploaded in the VI-SEEM data repository [17]. Codes produce by the VI-SEEM scientific communities are made available through the VI-SEEM code repository. Both the VI-SEEM code repository and the VI-SEEM data repository are integrated with the VI-SEEM Login service, the AAI infrastructure of the VRE that provides the authorization and authentication framework for all users of the region. This service is available at https://vre.vi-seem.eu/index.php/workflow-pipeline-and-software-tools-repository

4.2 Services integration

During the first year of the project and with the end of the 1st integration phase several workflows and codes have been made available to the community. As the integration phases progress more workflows and codes are planned to be made available through the VRE platform. Table 2 lists the currently available workflows and codes providing the specific links to the VRE platform. Workflows and codes are listed along the specific use case developed through the integration phase.

	Application acronym	Codes	Workflows
Climate	EMAC	EMAC code already uploaded to VISEEM repository available to the VI-SEEM community	
LS	MD-Sim	ChemBioServer (pending Material Transfer Agreement - MTA)	Workflows for simulating a protein with NAMD and GROMACS

Table 2 - Workflows and codes available through the VRE platform

4.3 Access through the VRE platform

4.3.1 Accessing workflows and software tools

The main page of the “wokflows and software tools service” is depicted in Figure 10. The service initially provides some general information for the service. Following, specific links are provided for accessing available scientific workflows (left) and software tools (right). A direct link for accessing available workflows and codes for each specific scientific community is also provided.

The “scientific workflows” page provides a list with the available scientific workflows for each scientific community. Figure 11 shows an example of how the scientific workflows are listed for the life sciences community. The table provides a link to the specific workflows along with a short description and contact information for the responsible scientists.

In a similar manner, the “code repository” page (Figure 12) provides a list with the available codes for each scientific community. Additional, generic codes are also listed in this section. These codes are specific software modules build by the VI-SEEM community during the integration of the generic services, and are openly available for use.

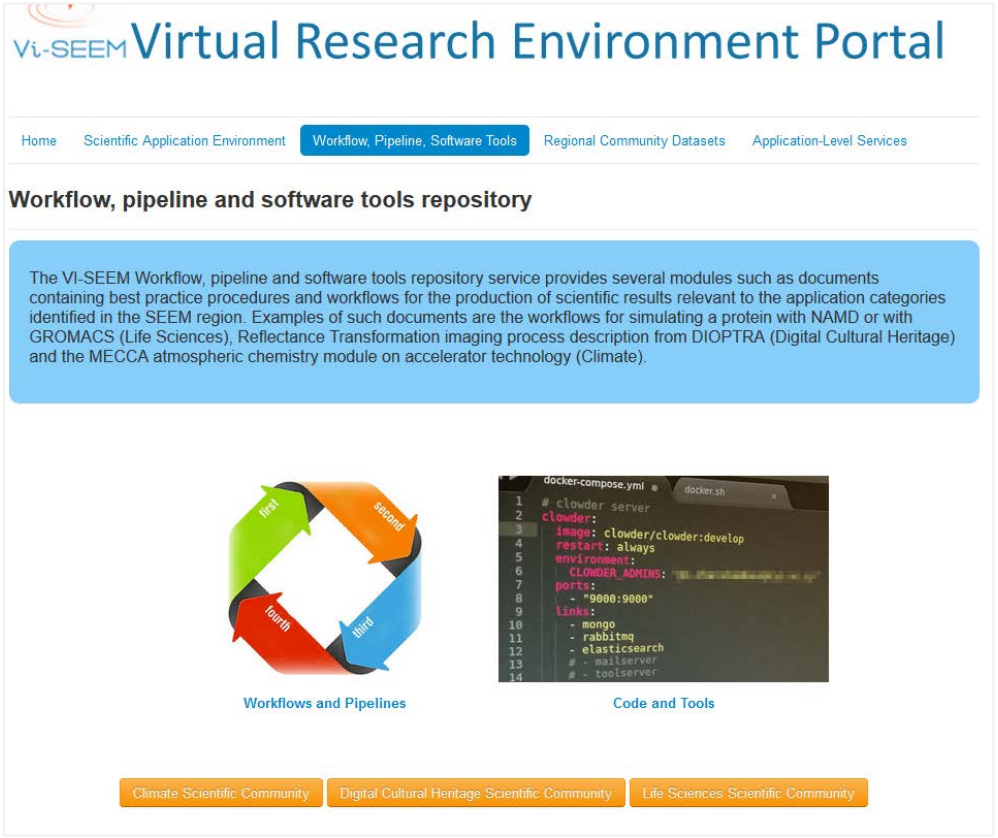
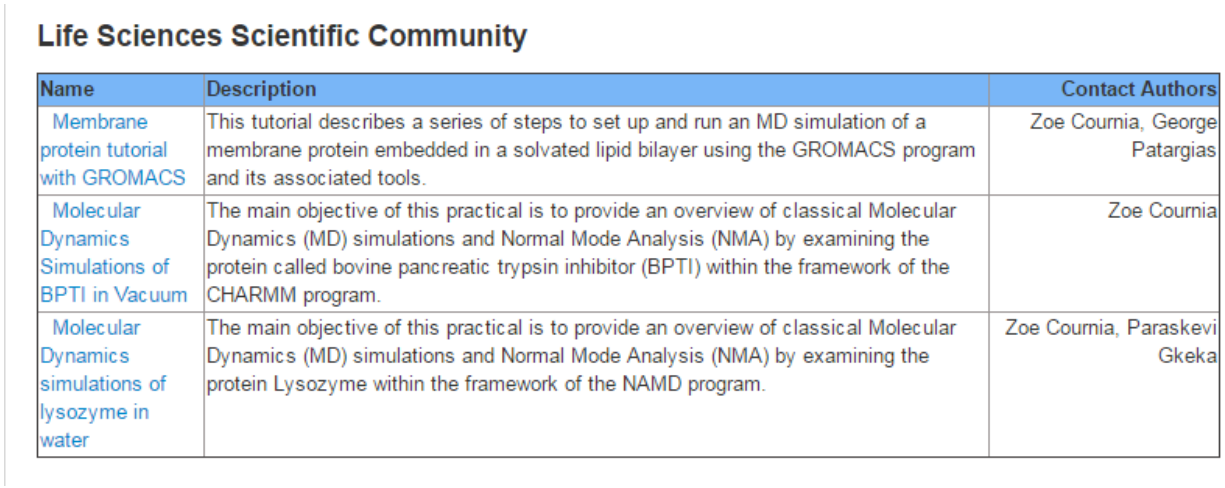


Figure 10 – Workflow and software tools repository main page



Life Sciences Scientific Community

Name	Description	Contact Authors
Membrane protein tutorial with GROMACS	This tutorial describes a series of steps to set up and run an MD simulation of a membrane protein embedded in a solvated lipid bilayer using the GROMACS program and its associated tools.	Zoe Cournia, George Patargias
Molecular Dynamics Simulations of BPTI in Vacuum	The main objective of this practical is to provide an overview of classical Molecular Dynamics (MD) simulations and Normal Mode Analysis (NMA) by examining the protein called bovine pancreatic trypsin inhibitor (BPTI) within the framework of the CHARMM program.	Zoe Cournia
Molecular Dynamics simulations of lysozyme in water	The main objective of this practical is to provide an overview of classical Molecular Dynamics (MD) simulations and Normal Mode Analysis (NMA) by examining the protein Lysozyme within the framework of the NAMD program.	Zoe Cournia, Paraskevi Gkeka

Figure 11 - List of available scientific workflows for the life sciences community as described in the VRE platform

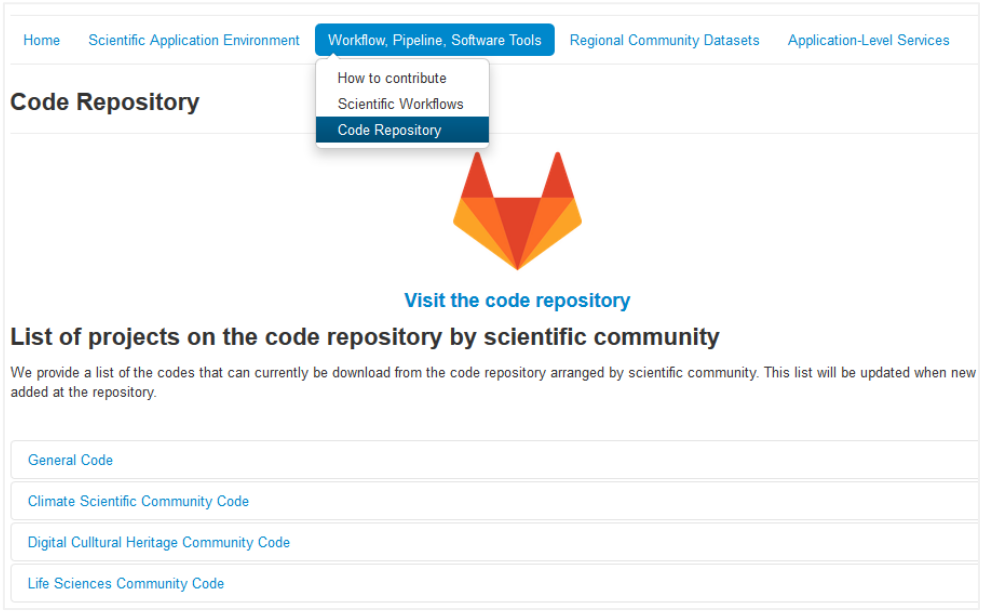


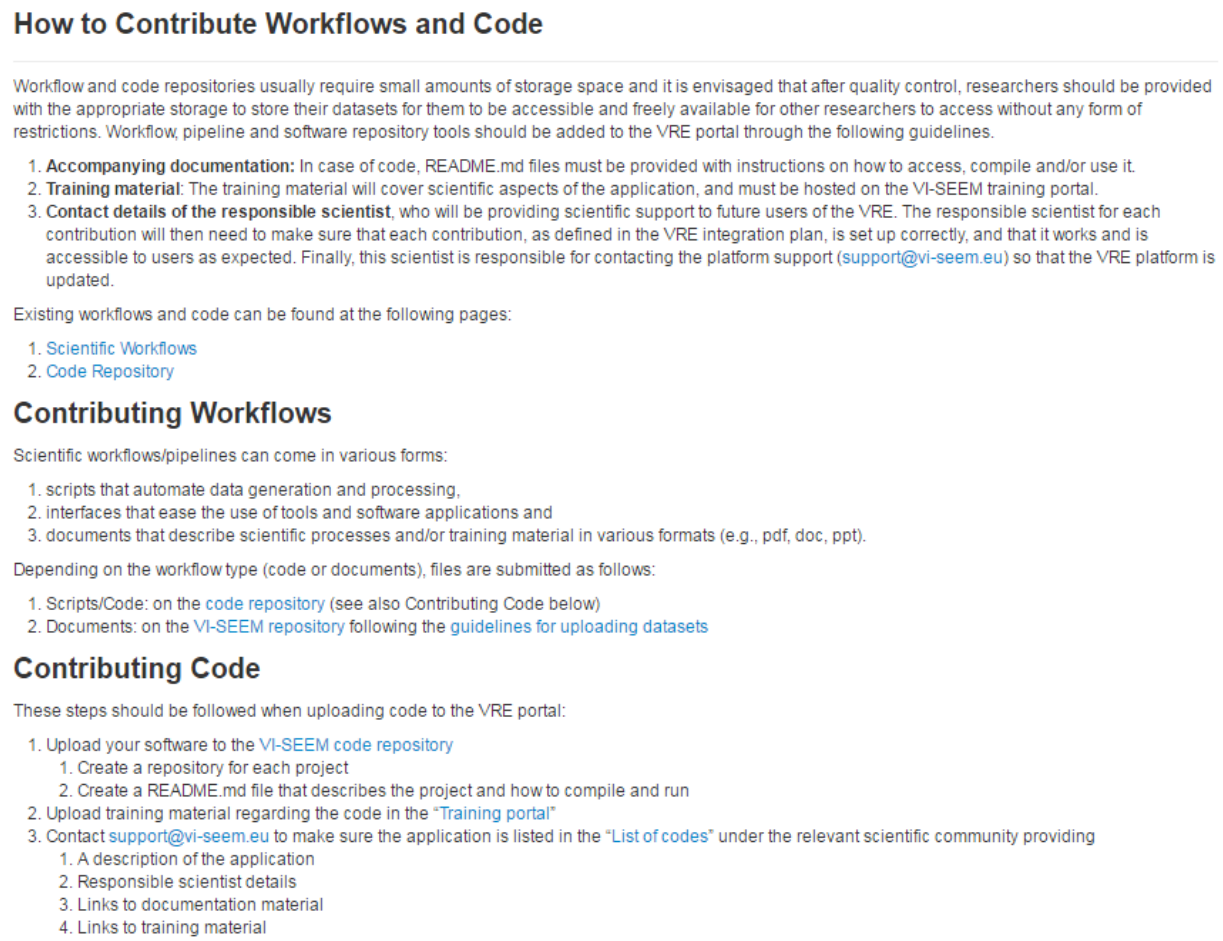
Figure 12 - VRE platform code repository page

4.3.2 Making workflows and software tools available to the VRE platform

A specific “how to contribute” page is available to the interested researchers to ease the process of making workflows and codes available to the VI-SEEM communities. The page lists in an easy to follow manner the needed steps and information researchers need to follow and provide in order for the workflows/codes to be made available (Figure 13). Workflow and software tools should be added to the VRE portal accompanied with the following information:

1. Documentation: in case of code, README.md files must be provided with instructions on how to access, compile and/or use them.
2. Training material: the training material will cover scientific aspects of the application, and must be hosted on the VI-SEEM training portal.
3. Contact details of the responsible scientist, who will be providing scientific support to future users of the VRE. The responsible scientist for each contribution will then need to make sure that each contribution, as defined in the VRE integration plan, is set up correctly, and that it works and is accessible to users as expected. Finally, this scientist is responsible for contacting the platform support (support@vi-seem.eu) so that the VRE platform is updated to reflect the contribution.
4. License: Licensing information on the use of the material uploaded in the core repository must be also provided.

The complete guidelines are available at https://vre.vi-seem.eu/index.php/workflow-pipeline-and-software-tools-repository/how-to-contribute



How to Contribute Workflows and Code

Workflow and code repositories usually require small amounts of storage space and it is envisaged that after quality control, researchers should be provided with the appropriate storage to store their datasets for them to be accessible and freely available for other researchers to access without any form of restrictions. Workflow, pipeline and software repository tools should be added to the VRE portal through the following guidelines.

1. **Accompanying documentation:** In case of code, README.md files must be provided with instructions on how to access, compile and/or use it.
2. **Training material**: The training material will cover scientific aspects of the application, and must be hosted on the VI-SEEM training portal.
3. **Contact details of the responsible scientist**, who will be providing scientific support to future users of the VRE. The responsible scientist for each contribution will then need to make sure that each contribution, as defined in the VRE integration plan, is set up correctly, and that it works and is accessible to users as expected. Finally, this scientist is responsible for contacting the platform support (support@vi-seem.eu) so that the VRE platform is updated.

Existing workflows and code can be found at the following pages:

1. Scientific Workflows
2. Code Repository

Contributing Workflows

Scientific workflows/pipelines can come in various forms:

1. scripts that automate data generation and processing,
2. interfaces that ease the use of tools and software applications and
3. documents that describe scientific processes and/or training material in various formats (e.g., pdf, doc, ppt).

Depending on the workflow type (code or documents), files are submitted as follows:

1. Scripts/Code: on the code repository (see also Contributing Code below)
2. Documents: on the VI-SEEM repository following the guidelines for uploading datasets

Contributing Code

These steps should be followed when uploading code to the VRE portal:

1. Upload your software to the VI-SEEM code repository
    1. Create a repository for each project
    2. Create a README.md file that describes the project and how to compile and run
2. Upload training material regarding the code in the "Training portal"
3. Contact support@vi-seem.eu to make sure the application is listed in the "List of codes" under the relevant scientific community providing
    1. A description of the application
    2. Responsible scientist details
    3. Links to documentation material
    4. Links to training material

Figure 13 - Step-by-step guidelines for making workflows and software codes available to the VRE platform

5 VI-SEEM regional community datasets domain specific service

5.1 Scope and service description

This service provides access and information regarding datasets of regional importance for the scientific communities of interest. The service integrates the generic data services provided by WP4 and links them to the scientific communities [4].

Datasets can be made available to the scientific communities through three different methods:

1. Uploaded to the VI-SEEM dataset repository service [19]
2. Uploaded to an application-level service as those described in chapter 6
3. Linked through an external source

The first two methods provide to the data owned the storage infrastructure and services for making the dataset available. The latter case covers data that are already hosted in an external infrastructure. In all cases the datasets should follow the VI-SEEM data management plan as described in D5.2 [6].

5.2 Services integration

During the first year of the project and with the end of the 1st integration phase several datasets have been made available to the community. As the integration phases progress more datasets are planned to be made available through the VRE platform.

Table 3 lists the currently available datasets providing the specific links to the VRE platform. Datasets are listed along the specific use case developed through the integration phase. In the VI-SEEM repository the user can view all datasets available organized by the different scientific communities that the datasets belong to, as shown in Figure 14 . Selecting the specific scientific community the user can then get access to the available datasets together with all relevant metadata that accompany the dataset.

	Application Acronym	Datasets
Climate	WRF-ARW	Current atmospheric and weather predictions available from the Cyprus Department of Meteorology: • Current Weather Predictions DSM Values • Current Weather Predictions Point Values • Current Atmospheric Predictions (RAOB)
DCH	BVL	A set of rare books that represent documentation sources for the culture and civilization on the Banat region (Banatica collection). A subset of 200 digitized books have been made available in UVT's GridFTP. • Banatica Dataset

	ELKA	The Electronic Corpus of Karamanlidika (ELKA) offers access to a number of manually digitized Karamandlidika texts available for testing and improving OCR methods. • Karamaldidika texts
Life Sciences	MD-Sim	MD trajectories of oncogenic proteins with mutations relevant to the SEEM area: • MD simulation Trajectory for the wild-type PI3Ka protein solvated in water (5 replicas of 100 ns each and 2 replicas of 800 ns each) • MD simulation Trajectory for the oncogenic PI3Ka mutant H1047R protein solvated in water (5 replicas of 100 ns each) • MD simulation Trajectory for the oncogenic PI3Ka mutant E545K protein solvated in water (4 replicas of 800 ns each) • MD simulation Trajectory for the wild-type RXRa protein (1 simulation of 800 ns each) • MD simulation Trajectory for the oncogenic RXRa mutant S427F RXRa (2 replicas of 800 ns each)

Table 3 - Regional community datasets available through the VRE platform

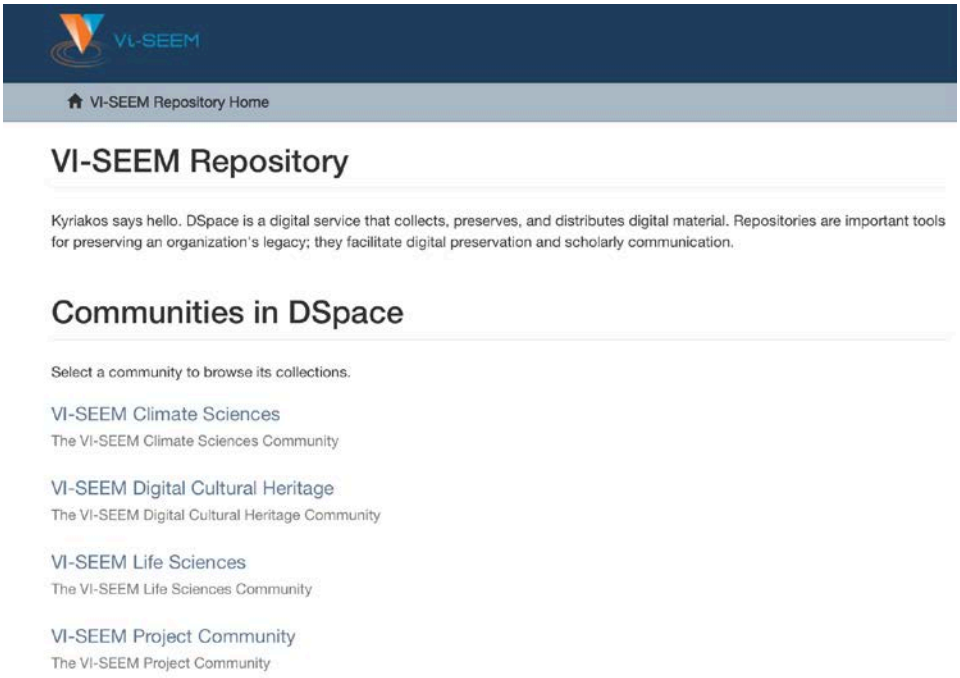


Figure 14: The home page of the VI-SEEM repository

5.3 Access through the VRE platform

5.3.1 Accessing datasets

The main page of the "regional community datasets" service is depicted in Figure 15. The service initially provides some general information for the types of data available through VI-SEEM and the users categories the datasets are targeting. For data producers interested in making their dataset available to the VI-SEEM communities a direct link to the guidelines for uploading datasets is provided. Finally, links to the available datasets for the three scientific communities are provided.

Figure 15 - Regional community datasets service main page

For each scientific community the available datasets are listed in a different page in the VRE platform. Figure 16 shows an example of the page using the climate scientific community. The page first explains the types of datasets available from the scientific community and also the structure of the metadata accompanying each dataset. For each of the dataset types a list of the available datasets is given.

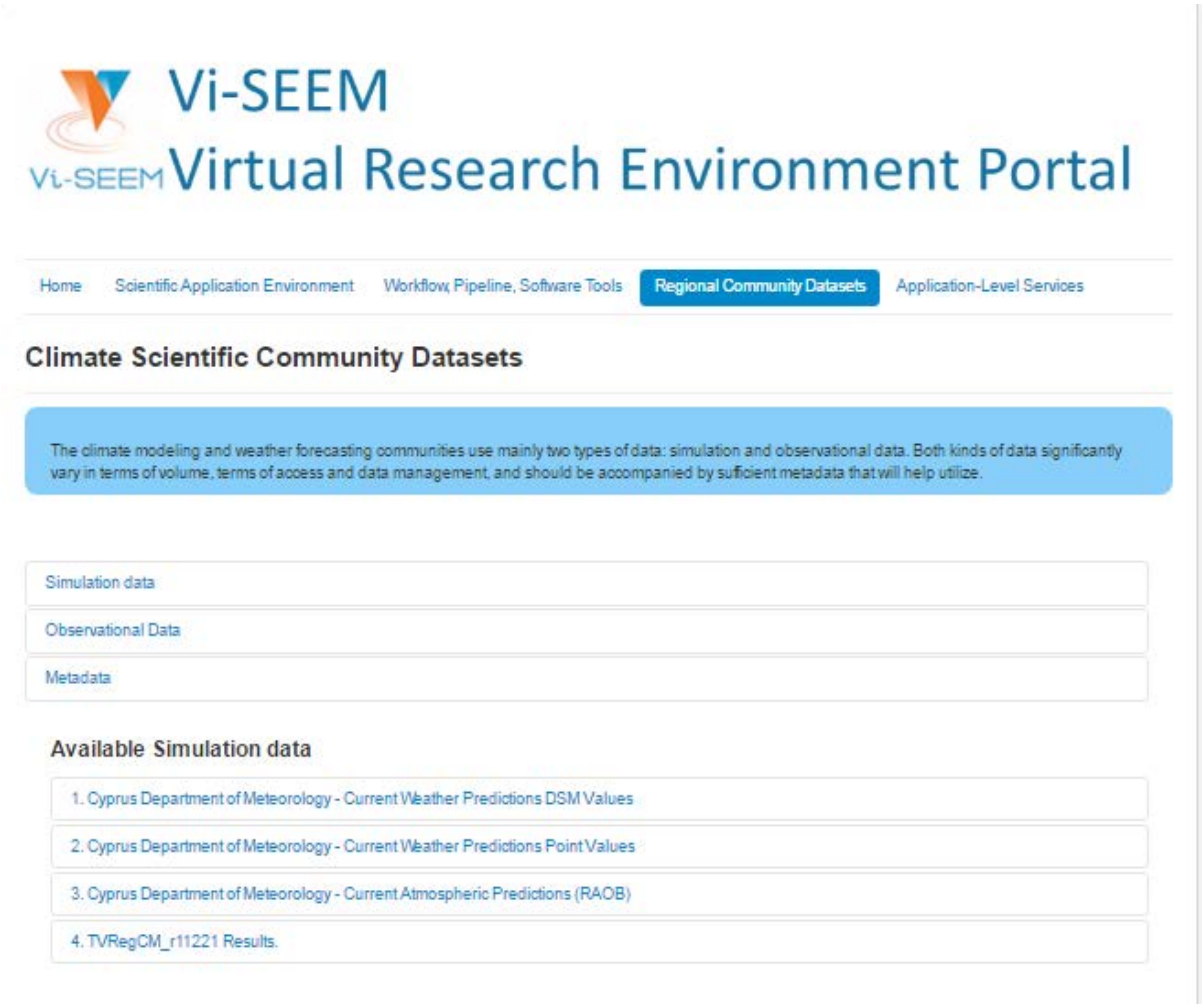


Figure 16 - Climate scientific community datasets main page.

Continuing the example for the climate scientific community Figure 17 shows the description for the climate simulation data; figure also depicts how available datasets are shown to the user. A direct link to the list of available datasets is given to the user through the yellow button. For each available dataset a small description is given to aid the user understand whether the dataset is of use to him. Additionally, metadata regarding the openness, preservation, pre-processing and responsible scientists are given for the dataset.

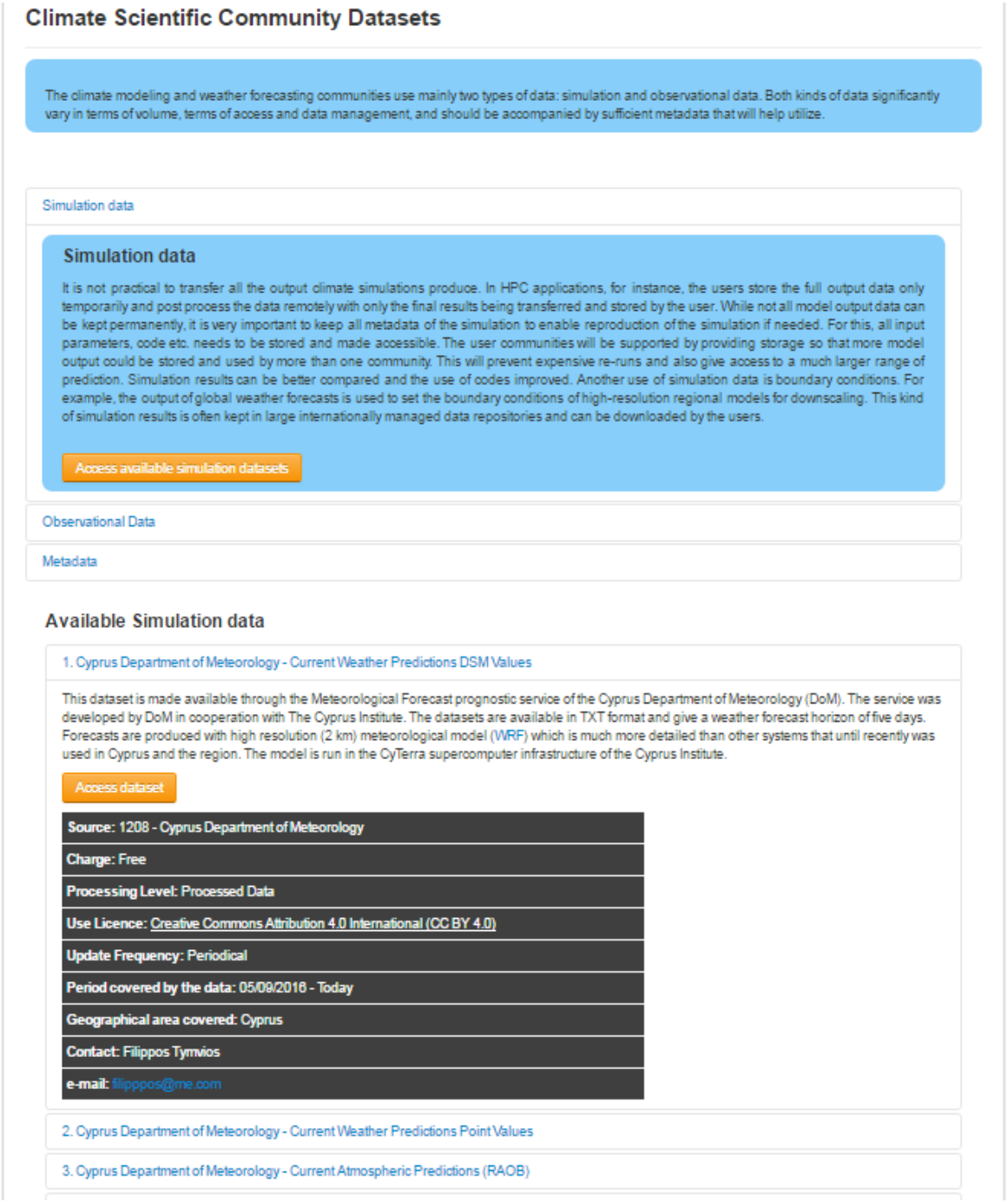


Figure 17 - Climate scientific community dataset page showing a description of the simulation data for climate research

5.3.2 Making datasets available to the VRE platform

A specific "how to contribute" page is available to the interested data producer to ease the process of making datasets available to the VI-SEEM communities. The page lists in an easy to follow manner the needed steps and information the data producer need to follow and provide in order for the dataset to be made available (Figure 18). The complete guidelines are available at https://vre.vi-seem.eu/index.php/regional-community-datasets/how-to-contribute. Specific instruction on how to use the different services to upload datasets given. For example, Figure 19 shows the detailed guidelines for adding a dataset in the VI-SEEM repository service. The guidelines are available at https://vre.vi-seem.eu/index.php/regional-community-datasets/how-to-contribute?id=96

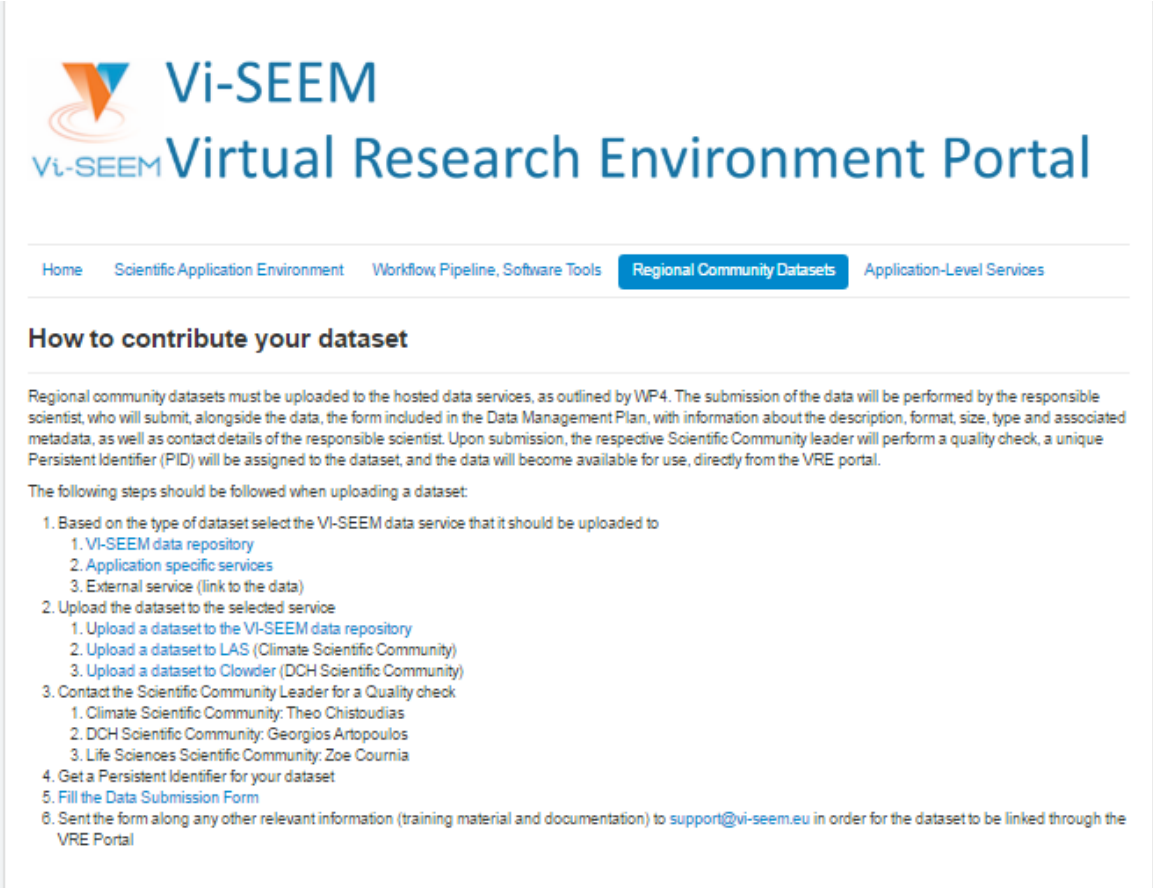


Figure 18 - Guidelines for listing datasets through the VRE platform

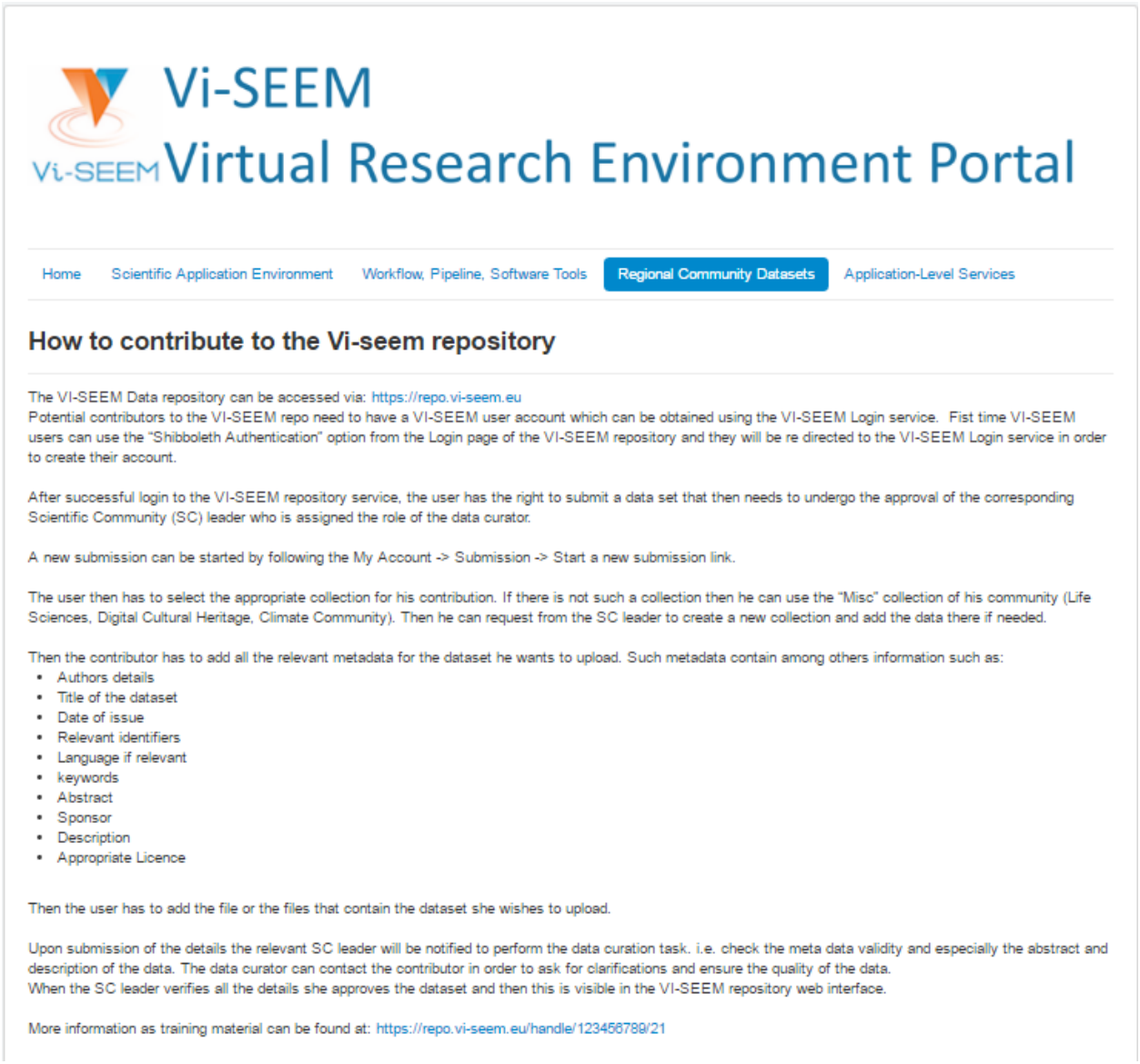


Figure 19 – Step-by-step guidelines for adding a dataset in the VI-SEEM repository

6 VI-SEEM application-level services

6.1 Scope and service description

This category contains web-based or visualization services providing easy access to underlying workflows, applications, and resources. Currently there are three available application-level services, one for each scientific community. These are standalone services that are provided the necessary hosting and storage infrastructure for their operation through the VI-SEEM generic services. The service allow for uploading and pre-processing datasets that are then made available to the scientific communities.

The listing of the application-level services is available at https://vre.vi-seem.eu/index.php/application-level-services-for-the-regional-communities

6.2 Service integration

Currently 3 application-level services are available through the VRE platform. These are:

1 Live Access Server (LAS), from the climate scientific community. LAS is a highly configurable server designed to provide flexible access to geo-referenced scientific data. It can present distributed data sets as a unified virtual data base through the use of DODS networking. Ferret is the default visualization application used by LAS, though other applications (Matlab, IDL, GrADS etc) can also be used. LAS enables web user to:

- visualize data with on-the-fly graphics
- request custom subsets of variables in a choice of file formats
- access background reference material about the data (metadata)
- compare (difference) variables from distributed locations

The main page of VI-SEEM LAS service is illustrated in Figure 20.

2 Clowder, from the digital cultural heritage scientific community. Clowder is a research data management system designed to support any data format and multiple research domains. It contains three major extension points: pre-processing, processing and previewing. When new data is added to the system, pre-processing is off-loaded to extraction services for extracting appropriate data and metadata. The extraction services attempt to extract information and run pre-processing steps based on the type of the data, for example to create previews. This raw metadata is presented to the user in the Clowder web interface. Users can upload, download, search, visualize and get various information about these data. Data in the case of VI-SEEM and more specifically in the field of digital cultural heritage can be of very diverse types and formats. More specifically users can upload massively or individual files of:

- 3D Models: extractors clean up and prepare for visualization on platform itself.
- Scanned books and their metadata: OCR algorithms will be used to extract the text in the documents so that users can find books using both metadata information and the book's contents.

- Image, text and sound files and their metadata, organised in collections.
- Advanced documentation data, such as Reflectance Transformation Imaging, and analysis of material properties of structures, works of art and artefacts.

The main page of VI-SEEM Clowder service is illustrated in Figure 21.

3 ChemBioServer from the life sciences scientific community. ChemBioServer is a web-application for effectively mining and filtering chemical compounds used in drug discovery. ChemBioServer allows for pre-processing of compounds prior to an in silico screen, as well as for post-processing of top-ranked molecules resulting from a docking exercise with the aim to increase the efficiency and the quality of compound selection that will pass to the experimental test phase. It provides researchers with the ability to:

- browse and visualize compounds along with their properties
- filter chemical compounds for variety of properties such as steric clashes and toxicity
- apply perfect match substructure search
- cluster compounds according to their physicochemical properties providing representative compounds for each cluster
- build custom compound mining pipelines
- quantify through property graphs the top-ranking compounds in drug discovery procedures.

The main page of VI-SEEM ChemBioServer service is illustrated in Figure 22.

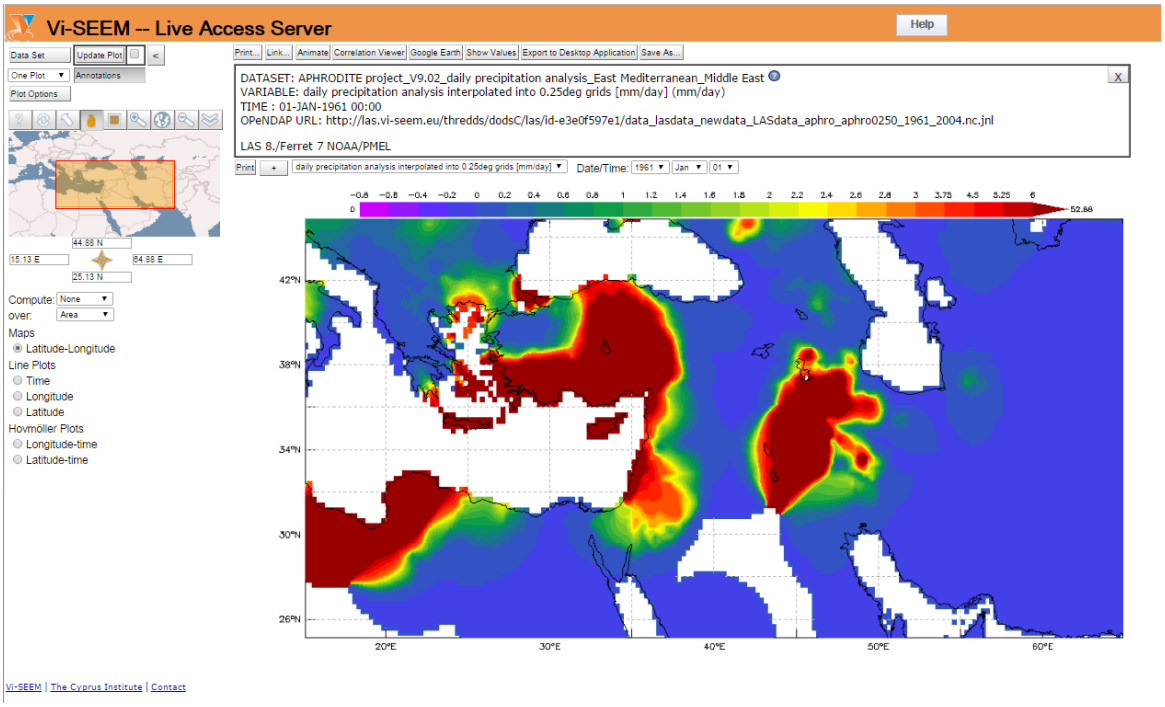


Figure 20 - VI-SEEM Life Access Server service for climate scientific community

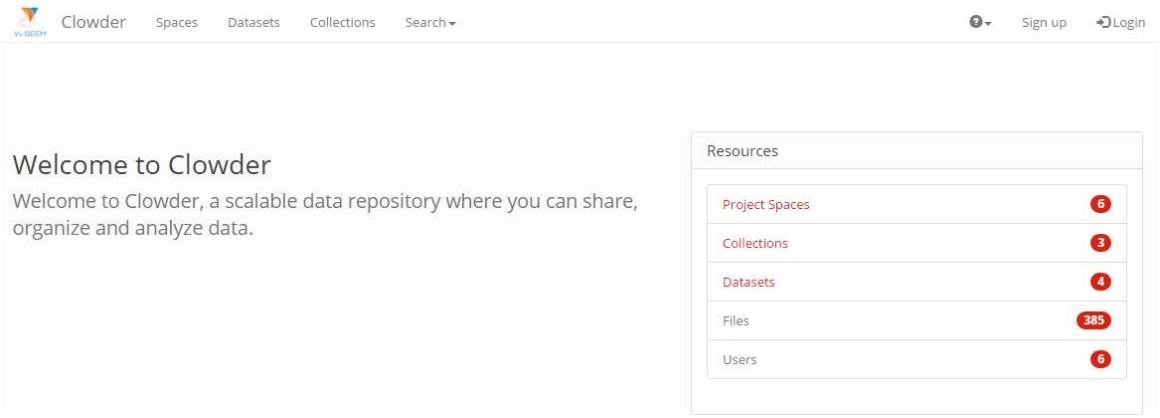


Figure 21 - VI-SEEM Clowder service for DCH scientific community

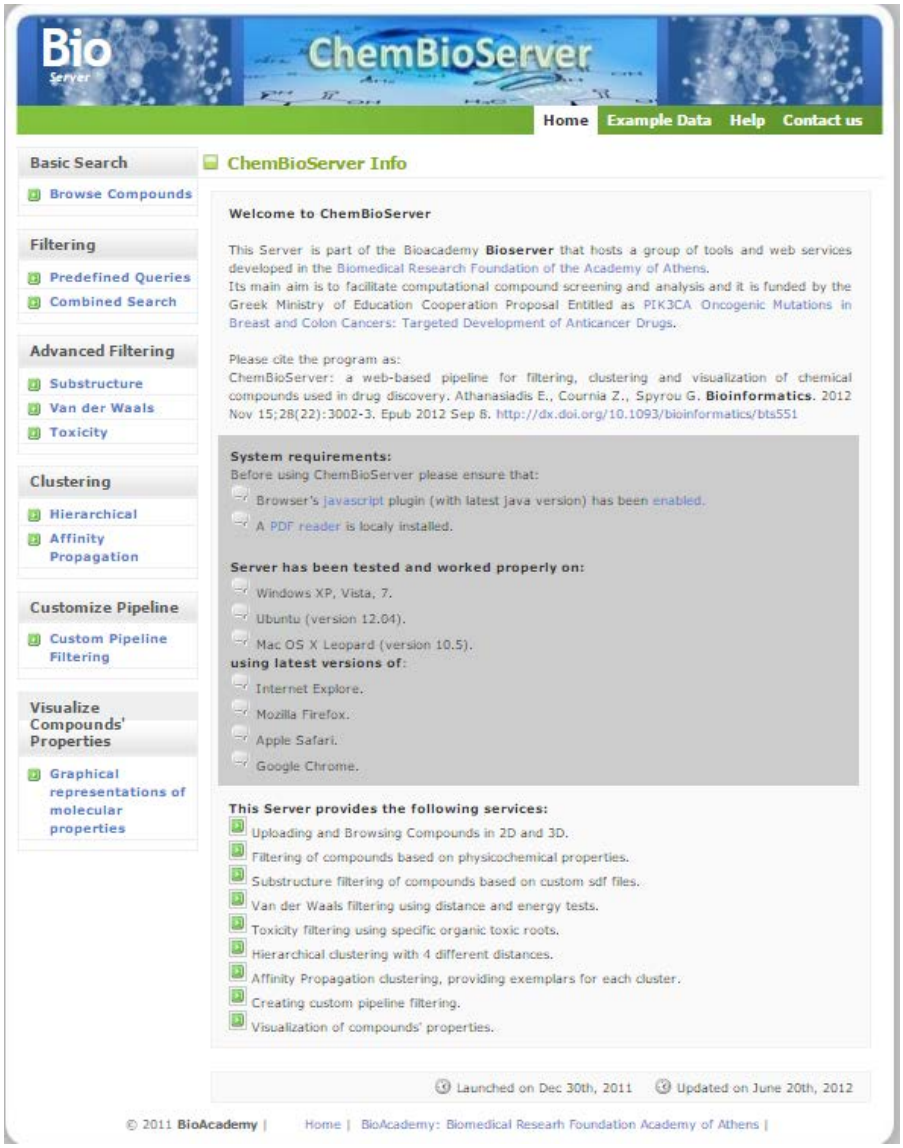


Figure 22 - VI-SEEM ChemBioServer service for life sciences community

6.3 Access through the VRE platform

6.3.1 Accessing available services

The main page of the "application-level services" category, as depicted in Figure 23, provides a short introduction on the services and easy one-click access to the available services. The page also provides a link to specific information for making new application-level services available through the VRE platform. Following the links and through the menu, the user can get more information and access to the available application-level services. The available services in each community are listed along a short description. Direct access is provided to the service. Each service is also registered in the VI-SEEM service catalogue [18] offering more information about the service current status, including version and support information. Figure 24, Figure 25 and Figure 26 show the information provided to the VI-SEEM platform user for the available application-level services in each scientific community.

Figure 23 - VRE platform main page for application-level services

Figure 24 - Available climate application-level services

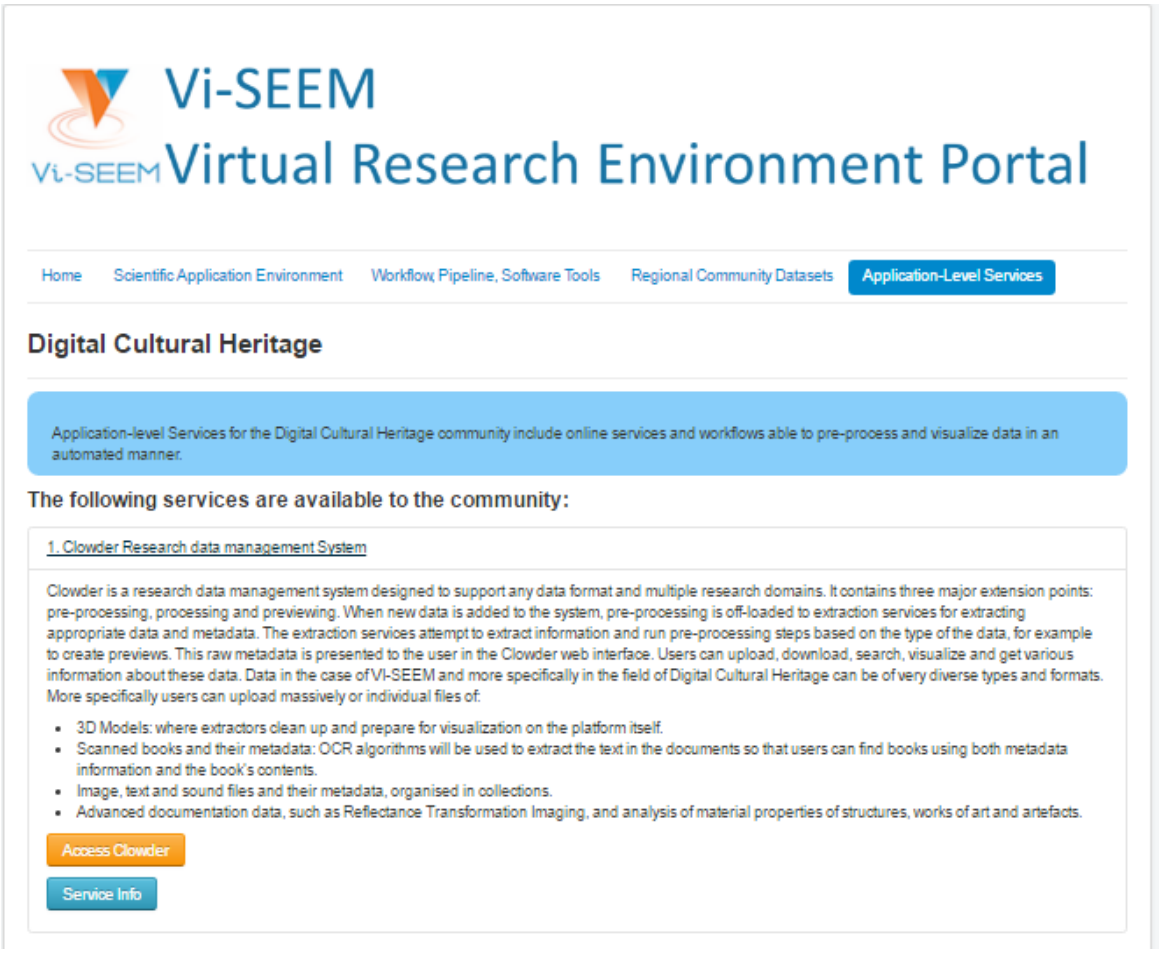


Figure 25 - Available DCH application-level services

Figure 26 - Available life sciences application-level services

6.3.2 Contributing new application-level services

A specific "how to contribute" page is available to the interested application-level service owner to ease the process of listing a service to the VI-SEEM communities. The page lists in an easy to follow manner the needed steps and information the service owner needs to follow and provide in order for the application-level service to be made available (Figure 27). The detailed guidelines can be found at https://vre.vi-seem.eu/index.php/application-level-services-for-the-regional-communities/how-to-contribute

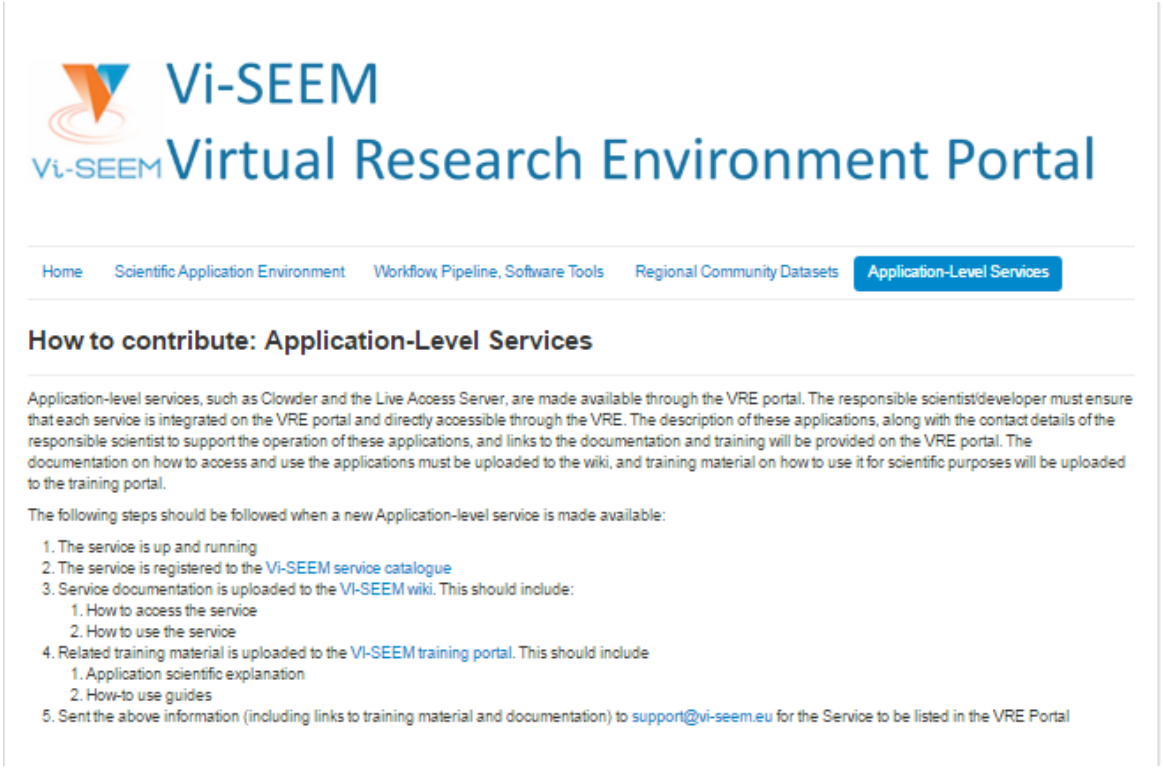


Figure 27 - Guidelines for listing an application-level service to the VRE platform

7 Update of existing and integration of new services

As the VI-SEEM project and the integration phases progress, new services are expected to be made available and the existing services will be significantly updated including new software tools, workflows, codes and datasets. In this section, we describe the current time plan for integration of new services and upgrade of existing ones. The time plan was derived after communication with the application developers currently in the 1st and 2nd integration phases and also discussions with additional application developers from the community. Further datasets, software codes and workflows are expected from the use cases participating in the forthcoming integration phases as those are described and planned in D5.1 [5].

7.1 Update of scientific application environment service

As part of the work performed in WP3 the unified e-Infrastructure is constantly updated with tools and software needed by the application in the running integration phases. This effort will continue and the plan is for all available infrastructures to install and support all needed software modules and optimized libraries. Additionally, use cases in the integration phases are expected to provide the ready to use virtual machine images listed in Table 4.

	Application Acronym	VMs
	VINE	A Virtual Machine for running the Regional Climate Model Evaluation System (RCMES)
	ClimStudyArmenia	Virtual machine including all software needed for running Accurate Prediction and Investigation of Weather and Climate in Armenia and South Caucasus
DCH	Manuscript	A virtual machine image that includes the Manuscript system for annotation of historical manuscript and a prototype of Interactive Page Segmentation
Life Sciences	DICOMNetwork	A Virtual machine including all components of the DICOMNetwork
	PSOMI	A Virtual Machine include workflows for simulations of protein interactions using Gromacs and Namd.

Table 4 - List of VM images expected to be available through integration phases

7.2 Update of workflows and software tools service

The use cases currently and the 1st and 2nd integration phases of the project are expected to provide the workflows and software tools listed in Table 5.

	Application Acronym	Codes	Workflows
Climate	ClimStudy AM		Software as a Service using persistent identifiers (PID)
	HIRECLIM S	CCLM (Climate version of high resolution, limited-area model CLM) and associated scripts and procedures.	Visualisation and processing procedures written in GRADS and IDL
	VINE	• WRF-Chem (end of 2016) • RegCM4 (May 2017) • WRF-Climate (Oct 2017) • RCMES: Regional Climate Model Evaluation System (March 2018)	Data processing and visualisation tools of output data from: RegCM4 (June 2017); WRF-Climate (Nov 2017)
	WRF-CHEM	• WRF-Chem model, version 3.7.1 • NCL scripts for post-processing WRF-Chem model output	N/A
	DREAMCLI MATE	• DREAM model code (open for the VI-SEEM community)	An on-line environment for data access providing plotting utilities for model output variables.
	ACIQLife	• Scripts used in ACIQLife (end of 2016)	Workflows and Software tools used: Climlib; IOAPI
DCH	Manuscript	• Software for annotation of historical manuscript • A backbone library • Interactive Page Segmentation software	
Life Sciences	MD-Sim	• Automated Force Matching Method (AFMM) for small molecule parameterization	
	PSOMI		Workflow for analysis of interaction between the amino acid residues of the active sites of the membrane protein(s) Workflow for energy analysis of the interaction, length of newly formed interaction bonds and their strength Workflow for the synthesis of the test molecule in the laboratory Workflow for analysis of molecular dyn.

Table 5 - Scientific workflows and codes expected from the use cases participating in the 1st and 2nd integration phases

7.3 Update of the regional community dataset service

The use cases currently and the 1st and 2nd integration phases of the project are expected to provide the datasets listed in Table 6.

	Application Acronym	Datasets
Climate	ClimStudyArmenia	WRF model Outputs and Inputs
	HIRECLIMS	Data needed for CCLM model run
	VINE	RegCM4 long range prediction (1959-2100) output. RCMES results, (Nov 2017)
	WRF-CHEM	• Gridded datasets of meteorological variables for Europe & North Africa (2N – 60N, 20W – 62E), for the 1989-2009 period (end of 2016) • Gridded datasets of the ECMWF ERA-Interim surface and upper-air reanalyses, for the period 1989-2009. (end of 2016) • (3) Gridded datasets of dust aerosols concentrations for Europe and North Africa (2N – 60N, 20W – 62E), for the 1989-2009 period.
	DREAMCLIMATE	• Datasets from test runs, describing transport of Saharan dust in GRIB1 format, containing 3D maps of dust concentration in 8 size bins, dust load, aerosol optical thickness are available. • Longer period (several years) datasets with 3h temporal resolution
	ACIQLife	NCEP Global Analysis and UBA data. Until end of 2016
DCH	Manuscript	A dataset for writer identification and word spotting (end of 2016)
	PSOMI	• (2R, 3R, 4R, 5S) -5-hydroxy-5 - ((S) -methylene-2-oxo-1,3-dioxolan-4-yl) pentan-1,1,3,4-tetraacetate Data format: PDB, GRO

Table 6 - Regional community datasets expected from the use cases participating in the 1st and 2nd integration phases

7.4 New application-level services

During the first year of the project, the following application-level services have been identified from the scientific communities and are expected to be deployed and made available to the VI-SEEM community in the forthcoming year.

- ARCHES: an open-source, geospatially-enabled software platform for cultural heritage inventory and management, developed jointly by the Getty Conservation Institute and World Monuments Fund. The system is freely available for organizations worldwide to download, install, and configure in accordance with their individual

needs and without restrictions. Arches is not one single repository; however, an organization could set it up as its own central repository if desired.

Arches has been purpose-built for the international cultural heritage field, and the off-the-shelf or default version is configured to inventory and document all types of immovable heritage, including buildings, cultural landscapes, heritage ensembles or districts, and archaeological sites.

  - Arches helps organizations achieve a number of objectives that promote the understanding, appreciation, and management of heritage places. These include:
  - identification and inventory
  - research and analysis
  - monitoring and risk mapping
  - determining needs and priorities for investigation, research, conservation and management
  - planning for investigation, conservation, and management activities
  - raising awareness and promoting understanding among the public, as well as governmental authorities and decision makers

- CH-CBIR: Content-Based Image Retrieval and Classification in Cultural Heritage Applications. Development of discriminative and compact image representations for Cultural Heritage applications. Effective and efficient algorithms for image retrieval and classification in large collections.
- DICOMNetwork: an application for distributing medical image sets. It is expected that "DICOM Network" will help doctors and medical personnel with processing and comparing investigation results. As a result, research community will have access to a large number of investigation archive databases that will offer for making decisions based on experience database. The system structure comprises the following data servers and modules:
  - Data from equipment are collecting at the “DICOM Server” modules that can be installed in the same location with the used medical equipment or can be distributed through other institutions and even countries.
  - All the investigations (DICOM Images) are archiving on DICOM Servers, but the information about investigation is stored in DICOM Portal (like www.dicom.md) database. Many DICOM Services can be connected to one DICOM Portal.
  - DICOM Portal stores all data like user’s info, access info, system settings, DICOM Server settings and some other, but not DICOM images it salves. Each Institution can deploy DICOM Portal internally on own server.
  - DICOM DATA Interface collects information about users and investigations from all DICOM Portals and provides functionality for data exchange and unification.

  This architectural approach provides a flexible information exchange that can be adjusted by requirements of individual customer or specific institution.

8 Conclusions

The deliverable is the main output of T5.3 “Development of the VRE platform” and T5.4 “Overall integration of services” and describes the design behind the VRE platform and how this reflects the integration of the generic and domain specific VI-SEEM services. The deliverable describes the different sections of the VRE platform, illustrating how each section manages to integrate the VI-SEEM offered services. Details on how the user can easily get information and use the VI-SEEM provided services are given, together with the current services integration status. Detailed instructions on how the user can integrate hers/his own applications and services to the VI-SEEM Virtual Research Environment are given in each section.

The deliverable and the VRE platform in general will act as a guideline for the further integration of applications and services. This will allow application developers participating in the VI-SEEM integration phases to make their tools, datasets and services available to the scientific communities.

