

VI-SEEM project

1st Call For Proposals For Projects Accessing VI-SEEM Services

<https://vi-seem.eu>

Todor Gurov, IICT-BAS
gurov@bas.bg



- ❑ The project proposals should address open research topics in specific fields of **Life Sciences, Climate research, and Digital Cultural Heritage**.
- ❑ The call is addressed to scientists and researchers that work in academic and research institutions in the region of South Eastern Europe and the Eastern Mediterranean. More specifically these are (in alphabetical order):
- ❑ **Albania, Armenia, Azerbaijan, Bosnia and Herzegovina, Bulgaria, Cyprus, Egypt, FYR of Macedonia, Georgia, Greece, Hungary, Israel, Jordan, Lebanon, Moldova, Montenegro, Romania, Serbia and Turkey.**



- ❑ In the field of Life Sciences
 - ❑ LS Area A: Modeling and Molecular Dynamics (MD) study of important drug targets
 - ❑ LS Area B: Computer-aided drug design
 - ❑ LS Area C: Analysis of Next Generation DNA sequencing data
 - ❑ LS Area D: Synchrotron data analysis
 - ❑ LS Area E: Image processing for biological applications



- ❑ In the field of Climate Research
 - ❑ CR Area A: Regional climate modelling to better understand and predict climate change and impacts, and phenomena such as dust storms
 - ❑ CR Area B: Air quality modelling, including atmospheric chemistry and air pollution transport
 - ❑ CR Area C: Weather forecast and extreme weather prediction, model development, application



Scope and criteria of access (3/3)

- ❑ In the field of Digital Cultural Heritage
 - ❑ Area A: Online services and access to repositories in order to enable studies of the cultural heritage assets in the region (e.g., searchable digital libraries; with support of meta-data and OCR for Latin characters).
 - ❑ Area B: Online visualization tools and data management systems to drive breakthrough contributions to art historical problems (e.g., interactive visualization viewer of RTI files and 3D models with digital libraries integration).
 - ❑ Area C: Unsupervised feature learning in photogrammetric techniques, data processing for image classification; semantic referencing; and geo-referencing.



Computational and Storage Services Available (1/4)



- ❑ **HPC Resources:** The HPC resources of the project consist of clusters with low-latency interconnection or supercomputers. Most of the systems are based on CPUs with x86_64 instruction set, some of them equipped with accelerators (GPUs and Intel Xeon Phi coprocessors), but also there are BlueGene/P systems, as well as one based on the Cell processor (PS3 cluster IMAN1-Booster/King).
- ❑ **In total 19 million CPU core hours, 370 million GPU core hours and 16 million Phi core hours will be provided in this call.**

HPC Resources:



Country	Site Name	Details	Dedicated Cores	Project				Project Hours			
				CPU	GPU	Phi	Cell	CPU	GPU	Phi	Cell
Greece	ARIS	RAM: 64 GB/server, 3.2 GB/core; FDR-14 Infiniband	5%	426	0	0	0	3 000 000	0	0	0
Cyprus	Cy-Tera	RAM: 48 GB/server, 3 GB/core; QDR Infiniband	15%	209	2 419	0	0	1 829 088	21 192 192	0	0
Bulgaria	Avitohol	RAM: 64 GB/server, 4 GB/core; FDR Infiniband	10%	240	0	1 830	0	2 102 400	0	16 030 800	0
Serbia	PARADOX	RAM: 32 GB/server, 2 GB/core; QDR Infiniband	5%	85	5 427	0	0	742 848	47 542 272	0	0
Hungary	NIIFI SC		5%	38	0	0	0	421 882	0	0	0
Hungary	Leo		5%	67	31 450	0	0	588 672	275 498 496	0	0
Romania	ICAM BlueGene/P	RAM: 4 GB/server, 1 GB/core; IBM 3D Torus Interconnect	20%	819	0	0	0	7 176 192	0	0	0
Romania	InfraGRID	RAM: 10 GB/server, 1.25 GB/core; QDR Infiniband	20%	91	627	0	0	798 912	5 494 272	0	0
Albania	UPT-HPC	RAM: 8GB/server, 0.5 GB/core; Ethernet	10%	14	0	0	0	126 144	0	0	0
FYROM	MK-03-FINKI	RAM: 24 GB/server, 2 GB/core; QDR Infiniband	5%	38	0	0	0	336 384	0	0	0
Armenia	Armcluster	RAM: 2 GB/server, 1 GB/core; Myrinet2000	10%	13	0	0	0	112 128	0	0	0
Egypt	BA-HPC	RAM 64 GB/server, 8 GB/core; DDR Infiniband	20%	208	0	0	0	1 822 080	0	0	0
Jordan	Gamma	RAM: 32 GB/server, 4 GB/core	100%	8	2 496	0	0	70 080	21 864 960	0	0
Jordan	Zaina	RAM: 6 GB/server, 0.75 GB/core; Ethernet	30%	17	0	0	0	147 168	0	0	0
Jordan	Booster/King	RAM: 256 MB	30%	77	0	0	614	672 768	0	0	5 382 144
				2 257	42 419	1 830	0	19 126 810	371 592 192	16 030 800	0

Computational and Storage Services Available(2/4)



- Cloud Resources:** The Cloud resources available in the call can be used in two ways. Those clouds that provide the ability to launch VMs with public IPs give the possibility to deploy VRE services for their main or backup/fail-over instance. VMs that possess only private IPs can be used for distributed data processing where necessary.
- In total around 300 VM cores are to be provided in this call.

No.	Country	Site Name	Details	VM Cores	Dedicated VM Cores	Project VM Cores	Project VM-hours
1	Greece	Okeanos	CPU: 1-8 cores; RAM: 1-8 GB	10 000		50	438 000
2	Cyprus	Cyl Cloud Facility	CPU: 1-8 cores; RAM: 1-96 GB	176		18	157 680
3	Bulgaria	Avitohol	CPU: 1-16 cores; RAM: 1-64 GB	2 400	5%	120	1 051 200
4	Romania	InfraGrid	CPU: 1-8 cores; RAM: 1-8 GB	400		46	402 960
5	Albania	UPT-Cloud	CPU: 1-8 cores; RAM: 1-8 GB	12	50%	6	52 560
6	Bosnia	ETFBL-CC01	CPU: 1-8 cores; RAM: 1-4 GB	60		13	113 880
7	FYROM	MK-04-FINKI_CLOUD	CPU: 1-12 cores; RAM: 1-12 GB	436		24	210 240
8	Moldova	MD-Cloud	CPU: 1-4 cores; RAM: 1-4 GB	12		3	26 280
9	Armenia	IIAP Cloud	CPU: 1-8 cores; RAM: 1-16 GB	96		10	87 600
10	Israel	IUCC InfinityCloud	CPU: 1-8 cores; RAM: 1-16 GB	560	5%	28	245 280
				14 092		305	2 671 800

Computational and Storage Services Available(3/4)

- Grid Resources:** The Grid resources, available in this call, are provided mostly from smaller clusters. Grid resources for the VI-SEEM VRE are provided from the following countries: Bulgaria, FYR of Macedonia, Georgia, Greece, Moldova, Montenegro and Serbia.

No.	Country	Site Name	Cores		Dedicated Cores	Project Cores		Project Hours CPU
			CPU	GPU		CPU	GPU	
1	Greece	Hellas Grid	864		5%	43	0	378 432
2	Bulgaria	BG01-IPP	640		5%	32	0	280 320
3	Serbia	AEGIS01-IPB-SCL	704		5%	35	0	308 352
4	FYROM	MK-03-FINKI	768		5%	38	0	336 384
5	Montenegro	MREN01CIS	32		50%	16	0	140 160
6	Moldova	MD-GRID	40			12	0	105 120
7	Georgia	GE-01-GRENA	128		10%	13	0	112 128
			2 536	0		158	0	1 380 576

Computational and Storage Services Available(4/4)



- ❑ **VI-SEEM Simple Storage Service:** (<https://ipbbox.ipb.ac.rs/index.php>) The VI-SEEM Simple Storage Service (VSS) is a secure data storage service provided to VI-SEEM users for storing and sharing research data as well as keeping it synchronized across different computers. Data sharing will be possible with other registered VI-SEEM users or with anyone else by using public links which can be protected with passwords if needed. Each user will be provided with 50 GB of storage for up to two years from the beginning of their project.
- ❑ **VI-SEEM Repository Service:** (<https://repo.vi-seem.eu/>) The main storage service that will allow the users of the VI-SEEM VRE to deposit and share data is the VI-SEEM Repository Service (VRS). This is VI-SEEM main repository for hosting the Regional Community Datasets. It can also be used to host publications and their associated data, as well as software or references to software and workflows used to generate such data and publications. The VRS is also the service designated for storing simplified data formats such as images, videos or others formats suitable also for the general public. Each project is eligible for up to 10 TB of storage for up to two years from the start of the project.

Computational and Storage Services Available(4/4)



- ❑ **VI-SEEM Archival Service:** Data archiving is the practice of moving data that is no longer being used or are being used in a less frequent fashion into a separate storage service. It is a single set or a collection of historical records specifically selected for longer term retention and future reference. Additionally, data archives contain data that are important for future reference or it is important to preserve them for regulatory and audit purposes. In science, archived data are important for future reference and reproducibility of scientific simulations. Each project will be eligible for storing up to 10 TB in the archival service of the project for at up to two years from the start..
- ❑ **VI-SEEM work storage space / local storage and data staging:** This service refers to storage space available by the computational resource providers to store temporary data for the purposes of processing them, or for storing results of computations. The service will be available for 12 months from the start of the project. The maximum capacity depends on the service provider.

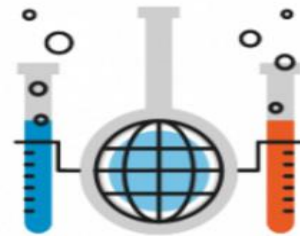
VRE Portal is available <http://vre.vi-seem.eu/> Continuously populated with new content



Virtual Communities

Climate

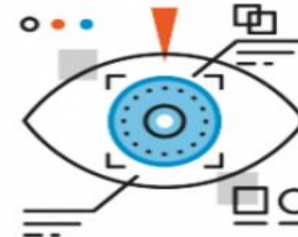
- Climate Applications
- MM5
- OPENFOAM
- WRF
- WRF-CHEM



Virtual Communities

Life Sciences

- Life Sciences Applications
- VMD
- NAMD
- JVM
- GROMACS



Virtual Communities

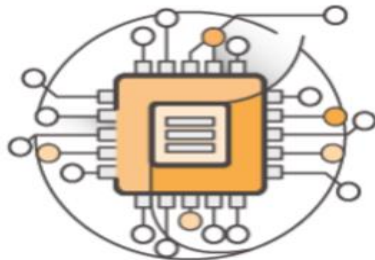
Digital Cultural Heritage

- Digital Cultural Heritage Applications
- SOL
- Clowder
- CH-CBIR
- AutoGR

Quick Links

[How to upload code](#)

[How to upload a dataset](#)



Access to Compute Resources



Access to Data Resources



Access to VI-SEEM Training Portal

The Live Access Server (LAS)

- ❑ The Live Access Server (LAS) is a highly configurable server designed to provide flexible access to geo-referenced scientific data. LAS enables the 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

Clowder

- ❑ 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 (zipped) or individual files of:
 - ❑ 3D Models: where extractors clean up and prepare for visualization on the 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

ChemBioServer

- ❑ 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 a 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.

Criteria for the evaluation of projects

- ❑ **Project proposals must be of high scientific and social value**
 - ❑ Scientific excellence
 - ❑ Scientific and/or social impact of the proposed research
 - ❑ The need for usage of the selected services and resources
 - ❑ The ability to provide project results (mainly data sets but also services and software) as services for other future VRE users
 - ❑ Maturity and experience of the principal investigator and his/her team in the research field as well as in using the selected resources and services
 - ❑ Feasibility of the project based on the technical evaluation and the availability of resources
 - ❑ Potential for the collaboration among scientists in more than one eligible country for this call

VI-SEEM aims at a balanced provision of resources to the whole spectrum of scientific fields between the three target communities

Important Dates



- ❑ Opening date: **14th October 2016**
- ❑ Closing Date: **14th November 2016**
- ❑ Clarification provided by applicants if needed: **18th November 2016**
- ❑
- ❑ Allocation decision: **January 2017**
- ❑ Allocation Start Date of awarded proposals: **January 2017**
- ❑ Allocation end date of award: **January 2018** for computational services, **January 2019** for some data services.
- ❑ Final report from successful projects: **May 2018**

Application process

- ❑ All proposals should be submitted electronically via the following link:
- ❑ <http://survey.ipb.ac.rs/index.php/112349?lang=en>
- ❑ The application form is also available in pdf format in order for applicants to have the full list of questions available. Please note that you have to fill in the on line form for your application to be taken into account.
- ❑ VI-SEEM Access Team will be available to answer questions while the call is open. You can contact the access team by sending e-mail to: service-access@lists.vi-seem.eu
- ❑ **Deadline for submission of proposals to the 1st call is Monday the 14th of November 2016 at 18:00 CEST**