

VI-SEEM

Concertation with International Cooperation Projects

www.vi-seem.eu

Dr. Ognjen Prnjat
Project Coordinator
GRNET
oprnjat at grnet dot gr



Admin details



- VRE for regional Interdisciplinary communities in Southeast Europe and the Eastern Mediterranean
- Start date 01/10/2015
- Duration 36 months
- Total funded effort, PMs: 715
- EC contribution 3.3ME

Geographical and historical context



- VRE for regional Interdisciplinary communities in SEE and EM
- Merging of SEE and EM regions
- SEE: network SEEREN1-2, Grid SEE-GRID-1-2-SCI, HPC HP-SEE
- EM: HPC LinkSCEEM1-2

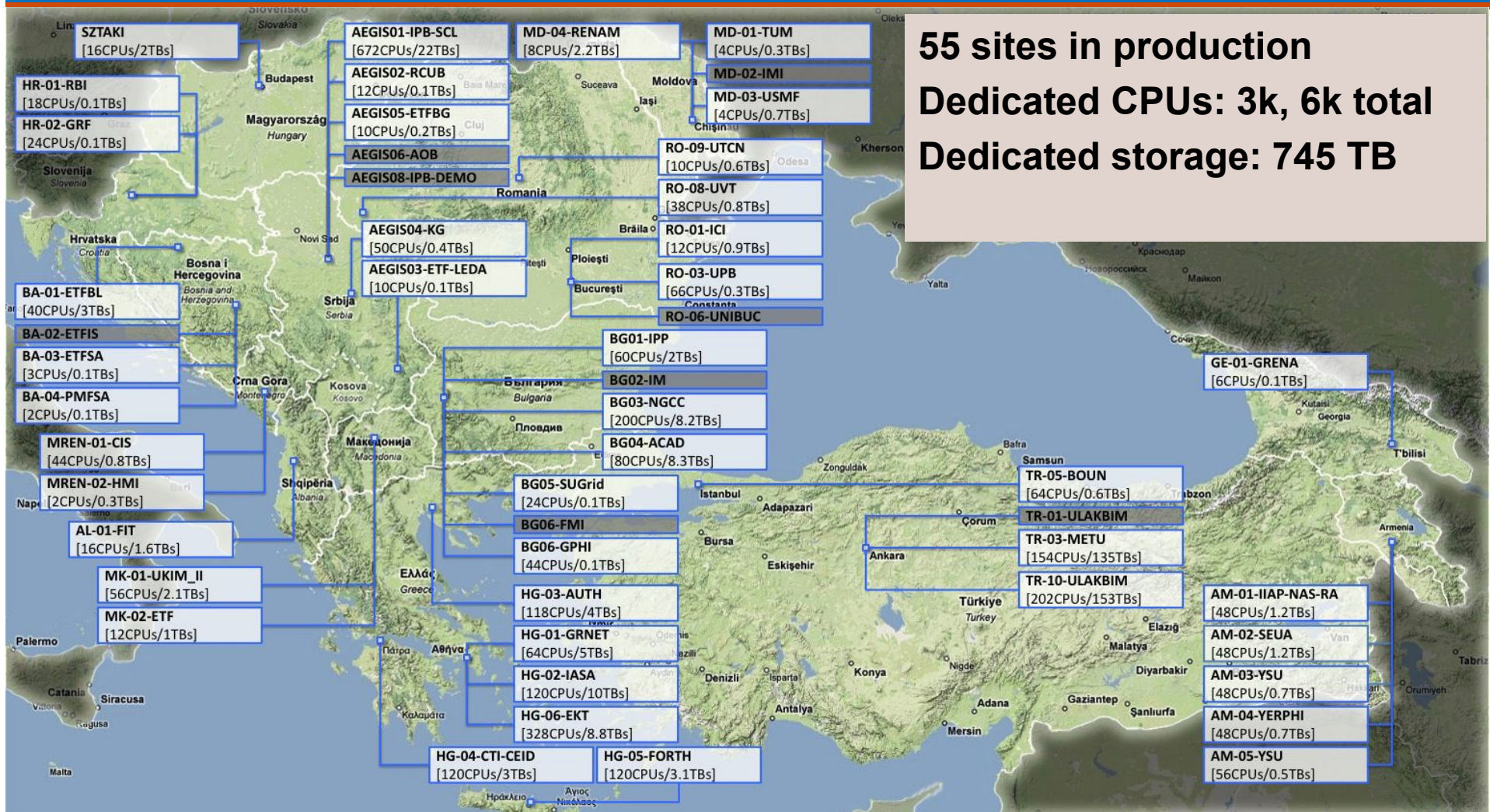


Participant no.	Participant organisation name	Part. short name	Country
1 (Coord)	GREEK RESEARCH AND TECHNOLOGY NETWORK S.A.	GRNET	Greece
2	THE CYPRUS INSTITUTE	Cyl	Cyprus
3	INSTITUTE OF INFORMATION AND COMMUNICATION TECHNOLOGIES – BULGARIAN ACADEMY OF SCIENCES	IICT-BAS	Bulgaria
4	INSTITUT ZA FIZIKU	IPB	Serbia
5	NEMZETI INFORMACIOS INFRASTRUKTURA FEJLESZTESI IRODA	NIIF	Hungary
6	UNIVERSITATEA DE VEST DIN TIMISOARA	UVT	Romania
7	UNIVERSITETI POLITEKNIK I TIRANES	UPT	Albania
8	UNIVERZITET U BANJOJ LUCI	UNI BL	Bosnia and Herzegovina
9	SS CYRIL AND METHODIUS UNIVERSITY OF SKOPJE	UKIM	FYR of Macedonia
10	JAVNA USTANOVA UNIVERZITET CRNE GORE PODGORICA	UOM	Montenegro
11	RESEARCH AND EDUCATIONAL NETWORKING ASSOCIATION OF MOLDOVA	RENAM	Moldova
12	INSTITUTE FOR INFORMATICS AND AUTOMATION PROBLEMS OF THE NATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF ARMENIA	IIAP-NAS-RA	Armenia
13	GEORGIAN RESEARCH AND EDUCATIONAL NETWORKING ASSOCIATION	GRENA	Georgia
14	BIBLIOTHECA ALEXANDRINA	BA	Egypt
15	INTER UNIVERSITY COMPUTATION CENTER	IUCC	Israel
16	SYNCHROTRON-LIGHT FOR EXPERIMENTAL SCIENCE AND APPLICATION IN THE MIDDLE EAST	SESAME	Jordan

Grid infrastructures



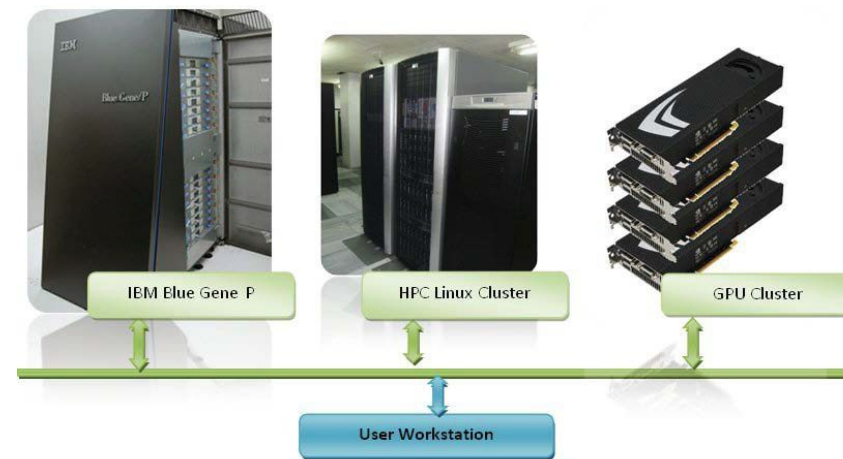
55 sites in production
Dedicated CPUs: 3k, 6k total
Dedicated storage: 745 TB



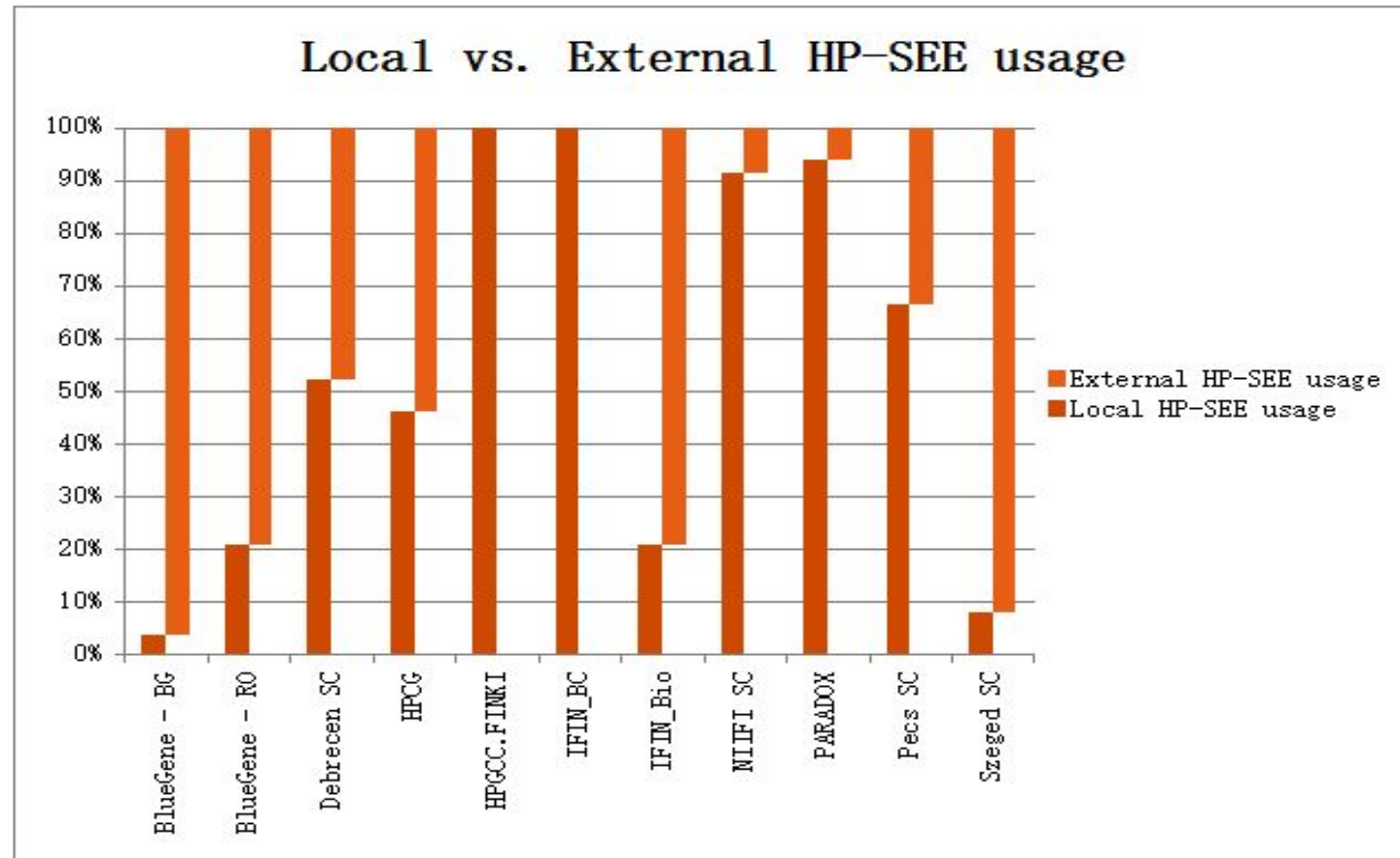
HPC infrastructures



- ❑ 3 top 500 supercomputers (Greece, Bulgaria, Hungary)
- ❑ Heterogeneous infrastructure - 3 Blue Gene/P SCs, several HPC clusters with advanced interconnects. Advantage to users
- ❑ Substantial number of libraries, toolkits and application software deployed, tested and optimised



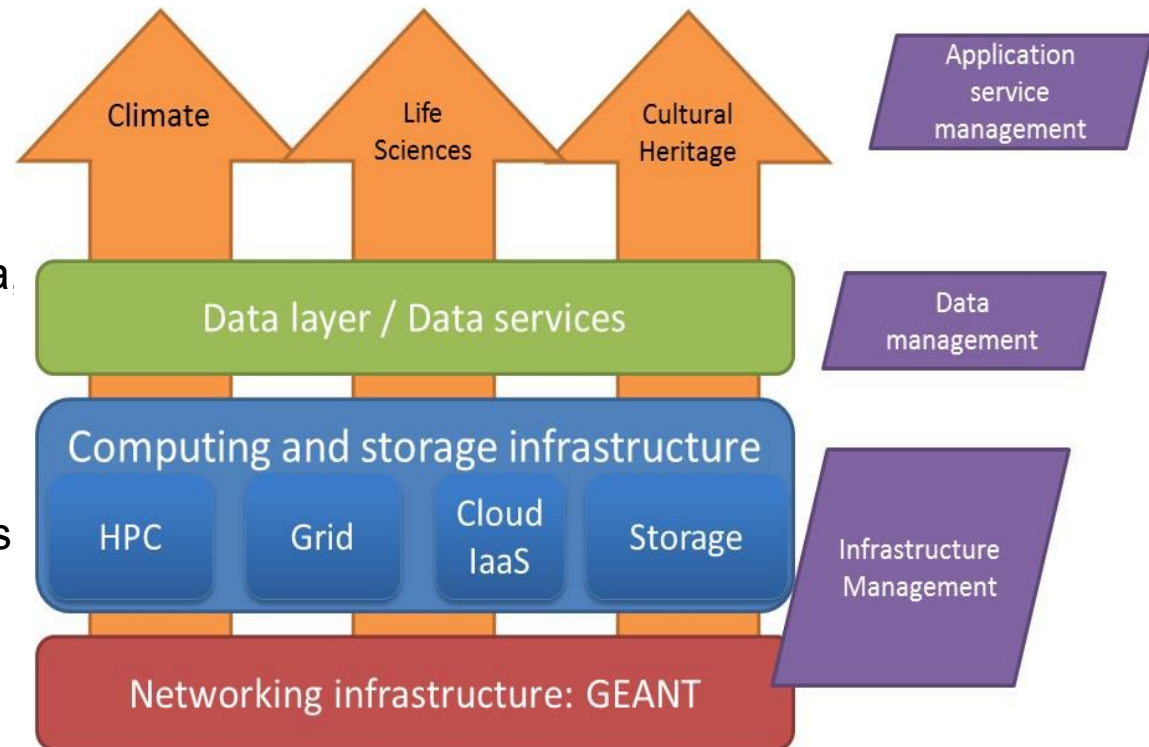
HPC: Trans-national usage and sharing



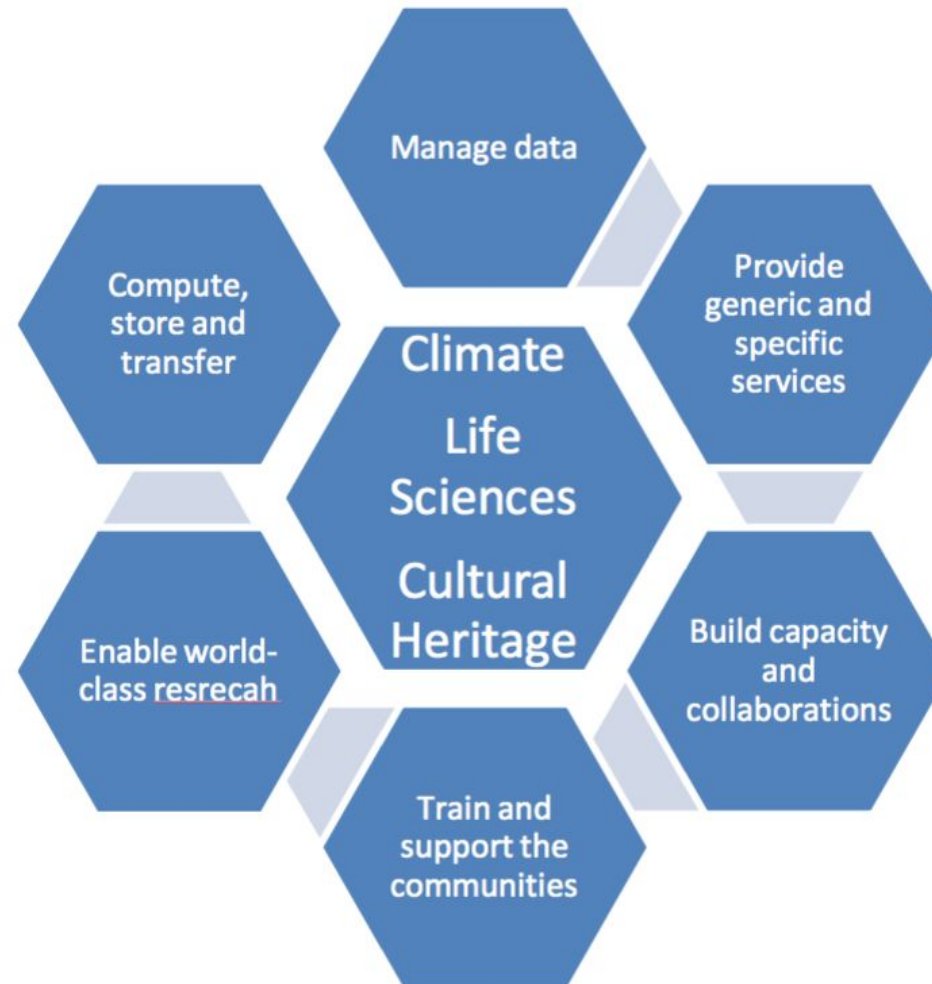
Technology context



- Overall objective: 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.
- Diverse computing technologies
- Advent of big data
- Service orientation
- Transversal services: AAI, service catalogues, common data and computing services in some cases

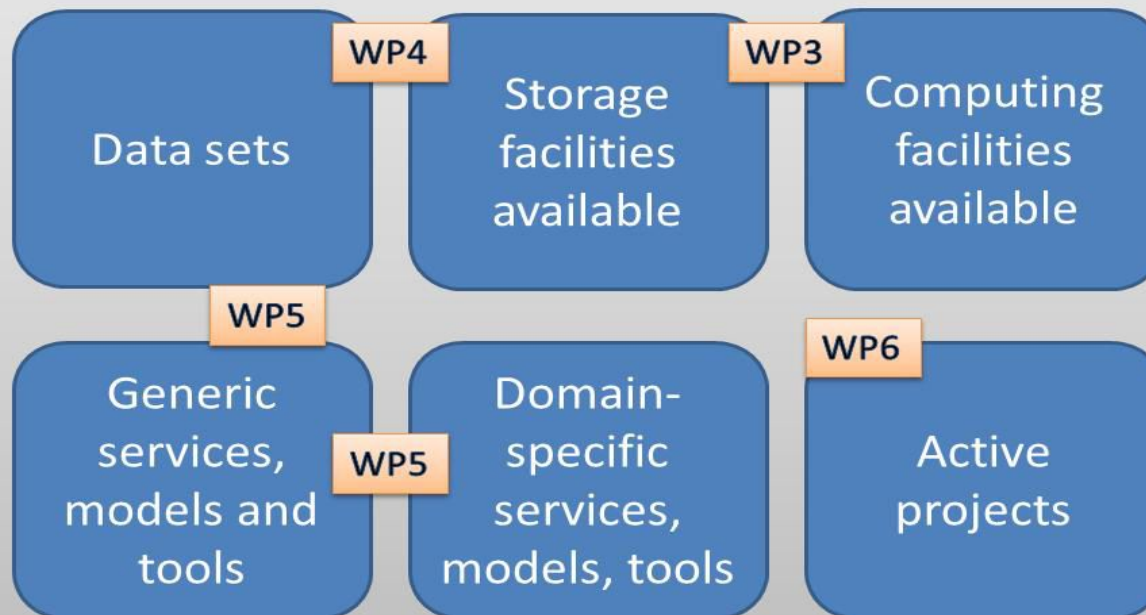


Detail objectives



User-centric approach

Virtual Research Environment: Users' viewpoint



Key Performance Indicators



- e-Infrastructure: 21500 CPU cores, 325000 GP-GPU cores and 18500 Intel Xeon Phi cores of HPC, 2900 grid cores, 10500 cloud VM cores, 11 PBytes of storage (of which dedicated 5-15% , 10-15%, 5% and 10% respectively)
- > 10 specific services, 30 data sets and 25 codes
- 39 applications, 45 research teams taking part. 100 publications, 50 presentations
- 17 dissemination events, 12 training events

Types of applications



Climate	Life Sciences	Cultural Heritage
Regional Climate Modelling	Modeling and Molecular Dynamics (MD) study of important drug targets	Digital Libraries
Global Climate Modelling	Computer-aided drug design	Interactive Visualization Tools
Weather Forecasting	Analysis of Next Generation DNA sequencing data and RNA profiling data	Semantic Referencing
Air Pollution/Quality	Data mining to identify prevalent diseases/mutations in the SEEM region	Image Classification
Model Development	Image processing for biological applications	Modelling of Built Environments and Advanced Representation Techniques
Visualization, Datasets, etc.	Computational simulation of DNA and RNA	Scientific simulation of materiality and systems' properties
Other	Synchrotron data analysis	Geo-referencing Tools
	Other	Bioarchaeology

VI-SEEM Data sets: climate



- Simulation and observation data: global and local sets

Source	Description	URL
Earth System Grid Federation	Data portal for downloading CORDEX and other data (i.e. CMIP5)	http://pcmdi9.llnl.gov/esgf-web-fe/
Climatic Research Unit (CRU)	gridded observations of surface meteorological data and climatic indices	http://www.cru.uea.ac.uk/data
E-OBS	gridded observations of surface meteorological parameters for Europe	http://eca.knmi.nl/download/ensembles/download.php
ECA&D	meteorological station observations for Europe and part of MENA	http://eca.knmi.nl/dailydata/index.php
APHRODITE	Gridded precipitation data for Asia and Middle East	http://www.chikyu.ac.jp/precip/index.html
Global Precipitation Climatology Centre (GPCC)	Gridded global precipitation observations	ftp://ftp-anon.dwd.de/pub/data/gpcc/html/gpcc_normals_down

	Source	Description
European Centre of Medium-Range Weather Forecasts Data Portal (ECMWF)	Mainly for TAU, Israel	Meteorological and hydrological data for Israel, environmental data, such as PM10 and PM2.5 measurements in various locations in Israel
	federation METEO-RO, Romania	a gridded dataset from observations consisting of 9 essential climate variables with daily values (1961-2013) on a spatial resolution of 10 km for the Romanian territory (ROCADA)
AERONET	MODIS (or M key instrum	data recorded by meteorological stations and posts of the State Hydro-meteorological Service
MODIS	NIGGG, Bulgaria	NCEP Global Analysis Data with 1° resolution, available since 2000
ISA-MEIDA		TNO high resolution emission inventory for Europe
TOMS aerosol index	index that det	Data of the pollution levels, measured by the Bulgarian National Network for Air Quality Control, available since 2000
	IGEWE, Albania	data collected from the Hydrological, Meteorological and Seismological networks (such as water flow, temperature, humidity, precipitation etc., and a range of seismic data)

VI-SEEM Data sets: Life Sciences



- Simulation and observation data; global and local sets

Country	Available data	Access Policy (Private, Regional, Public)
Greece	Virtual screening results	Data that is published is public and can be shared
Serbia	SNP database of variations in genes involved in Cystic Fibrosis, Thrombosis, Alpha-1-Antitripsine deficiency, Colorectal cancer, Thyroid cancer, Lung cancer, Endometrium. The database contains data on Serbian population.	Data that is published is public and can be shared
Moldova	Investigation Analytics, Research Results, Pathologies case studies	Investigation Analytics and Research Results are shared information. Pathologies case studies (no personal patient info) – available for research community by creating agreement with institution.

VI-SEEM Data sets: CH



Source	Description	URL
Ancient Cypriot Literature Digital Corpus (AKGDC)	Searchable digital library of Ancient Cypriot Literature. Covers the ancient Cypriot literary production from 7th century BC to 5th-6th century AD and examines it through its wide range of literary genres.	Not public yet.
STARC repository	Repository for digital CH objects created over the last 5 years at Cyl. Including novel interactive methods for accessing and exploring information.	http://public.cyi.ac.cy/starcRepo/
Electronic Corpus of Karamanlidika	Digitalization and transcription into Latin characters of a corpus of Karamanlidika printed material from different periods and genres with access to graphic, phonetic and morphemic varieties search.	Not public yet
The digital collection of Bibliotheca Alexandrina	As a library of the 21 st Century, the Bibliotheca Alexandrina (BA), along with its affiliated academic and cultural centres, is committed to digitization as a means of preserving, managing, and disseminating information and knowledge. The BA sets out to share them with a worldwide audience via the internet, thus promoting greater understanding and tolerance between cultures. To achieve this goal, the BA has formed partnerships with various cultural, academic, governmental, and corporate organizations for the creation of many significant and compelling digital projects. These projects cover a wide array of cultural and educational themes.	http://www.bibalex.org/Libraries/Presentation/Static/12600.aspx?d=0

Data-related activities



- ❑ Functions allowing for data management for selected Scientific Communities, engage the full data management lifecycle
- ❑ Select and integrate data management services tailored to the project's Scientific Communities
- ❑ Deploy and operate data management services that address the full data lifecycle
- ❑ Collect the actual VRE-specific data sets and make them available via common interfaces and services
- ❑ Data storage (live, dropbox-like), data archiving, data manipulation, collaborative access, domain specific interfaces to storage, data annotation and citation, metadata, PIDs, etc.

Data-related Services example: PIDs

- ❑ GRNET member of EPIC consortium
- ❑ Also provided as EUDAT service

Persistent Identifiers

by GRNET

[Browse the documentation](#)

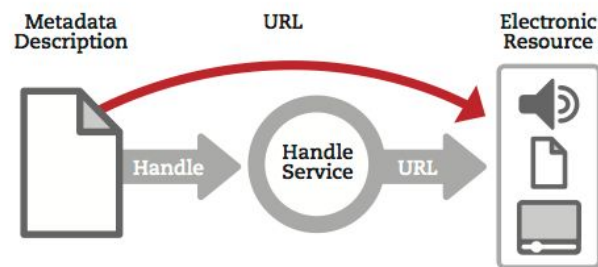


Image from Clarin Project (<http://www.clarin.eu>)



Get Started

Get started with the EPIC API.
[Browse the documentation](#)



API

Enable the use of persistent identifiers in
your application using the EPIC API



Contact

Are you interested in using Persistent Identifiers for your data set? Get
in touch with us to get you started. Just fill out the [form](#).

Concertation with other projects



- ❑ Sharing the application codes, data
- ❑ Possibility of scientific collaborations in certain application domains
 - ❑ Thus also sharing data storage and computing resources
- ❑ Know-how transfer in terms of data services (PIDs etc)
- ❑ Join work on KPIs, service catalogues

Thanks



□ Thanks!