

# GENESI-DR

**INFRA-2007-1.2.1 : Scientific Digital Repositories**  
**Duration : January 1, 2008 – April 30, 2010**  
**Total EC funding : 4.4 M€**

## ***GENESI-DR... and its future***

**R. Cossu / L. Fusco - ESA**



EC Grant Agreement no. 212073

**GEO Alliances and Harmonization Workshop**  
**Wash DC, November 11-12, 2009**



- Goals and Achievements of GENESI-DR
- Details:
  - Use of web portal to discover data
  - A more complex data access/ processing scenario
  - GENESI-DR metadata model and search
- From GENESI-DR to GENESI-DEC....

# What is unique in GENESI-DR

a **LARGE DISTRIBUTED DATA INFRASTRUCTURE (DRs)**

with **centralized discovery service**  
for seamless and **controlled**  
access to **heterogeneous**  
sources of data...



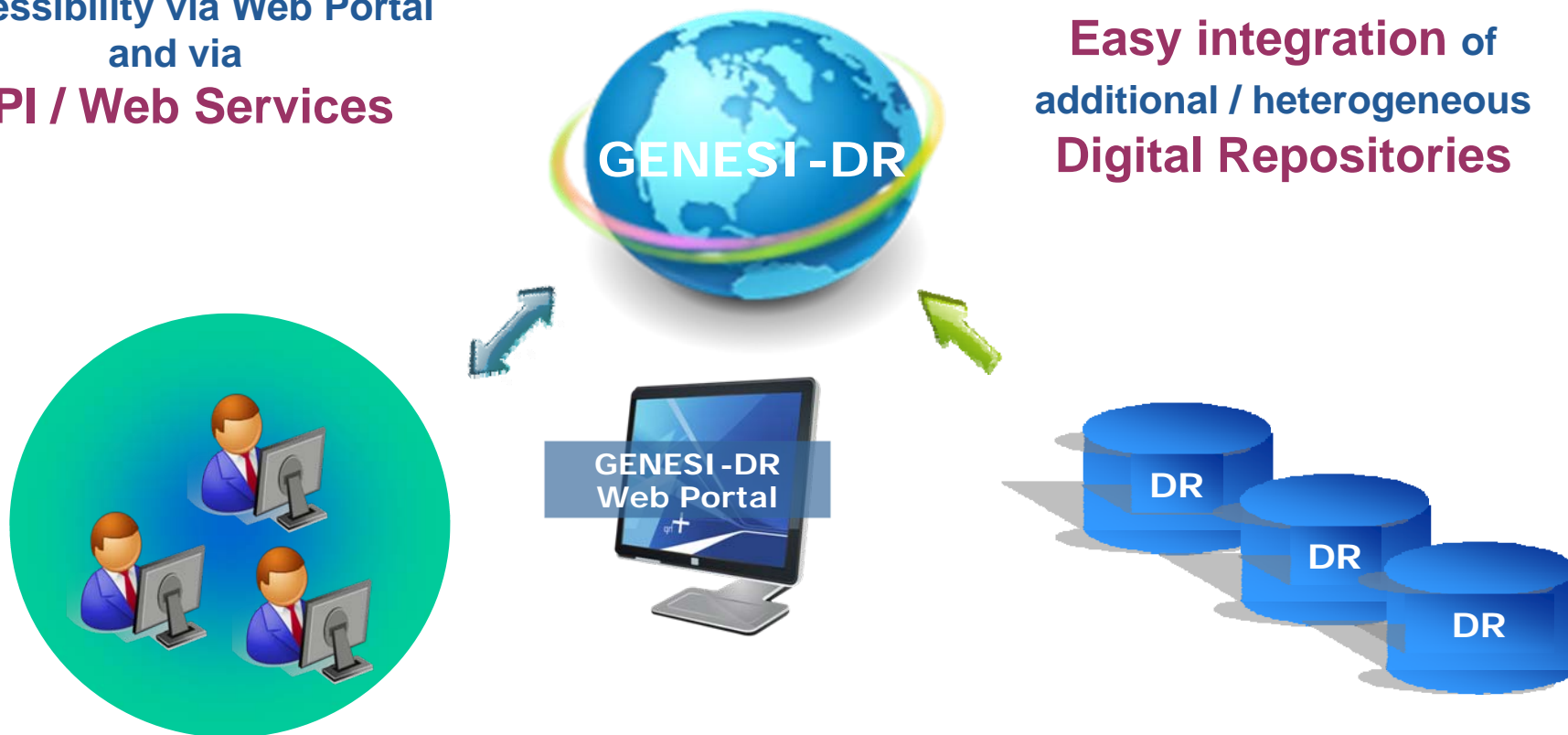
...and **processing resources**,  
to meet **world-wide community** needs

# What is unique in GENESI-DR

**Standardisation,**  
**flexibility** and **scalability** of the architecture

Accessibility via Web Portal  
and via  
**API / Web Services**

**Easy integration** of  
additional / heterogeneous  
**Digital Repositories**



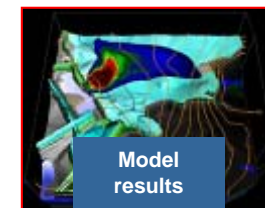
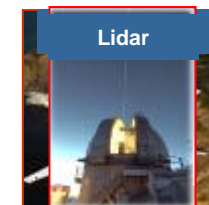
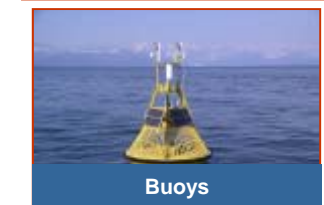
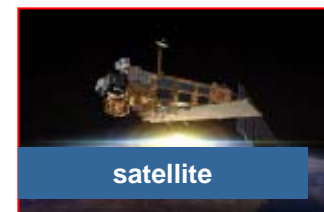
## GENESI-DR has successfully

- demonstrated a community **data e-infrastructure** that provides **discovery** and **access** services for
  - huge amounts of historical and near-real-time data **owned by operational data providers**, and
  - **on-demand (grid-)processing** resources
- Implemented an **open-standard** interface to heterogeneous data holdings
  - Compliant with EC INSPIRE directives
  - Using OGC approved OpenSearch (with geo extension) **for easy and fast integration** of DRs in the federation.
  - Emerging protocols for access (e.g. **BitTorrent**) and Curation / **preservation...**

## At present:

12+ sites in Europe are “GENESI-DR compliant”

50+ different dataset series (Satellite data, models, in situ time series data, results, ...)





**GENESI-DR** has successfully set up many collaborations with projects and organizations:

**GRID and Research Infrastructures**, e.g. EGEE, SEADATANET, Cyclops, SEE-GRID-SCI, Metafor..

**EC GEO/GMES** - Research science community, e.g. EnviroGRIDS...

**International environmental programmes**, e.g. Charter for Disaster Management, GEO/GEOSS, CEOS, International Society for Digital Earth...

...Many more relations established in Europe and world-wide...

# Standardisation efforts in GENESI-DR to support GEOSS

GENESI-DR will operate, validate and optimise the integrated access and use available data, information, products and knowledge originating from space, airborne and in-situ sensors from many repositories dispersed all over the world.

GENESI-DR will demonstrate how Europe can best respond to the emerging global needs relating to the state of the Earth, a demand that is unsatisfied so far.

GENESI-DR adopts OpenSearch that is now considered by OASIS and OGC a complementing way to provide a machine accessible search interface to a data repository.

GENESI-DR, short for Ground European Network for Earth Science Interoperations - Digital Repositories, with the overall goal to meet the great ambitions of GEO, has the challenge of establishing open and seamless access to Earth Observation and Earth science digital repositories for European and world-wide science users.

OpenSearch



GROUP ON  
EARTH OBSERVATIONS

THE GLOBAL EARTH OBSERVATION  
SYSTEM OF SYSTEMS



[www.genesi-dr.eu](http://www.genesi-dr.eu)

*Presented to the EC GEO workshop 8-9 Oct 2009 in Turkey*

# Discovery and access services

e.g., ESA CAT-1 registration

MERIS FRS Level 1 Eu (systematic processing)

MERIS FRS Level 2 Eu (systematic processing)

RA-2 IGDR, GDR and M (off-line processing at F)

SCIAMACHY level 1 co (off-line processing at D)

DORIS Precise Orbit (F)

GOME level 1 products

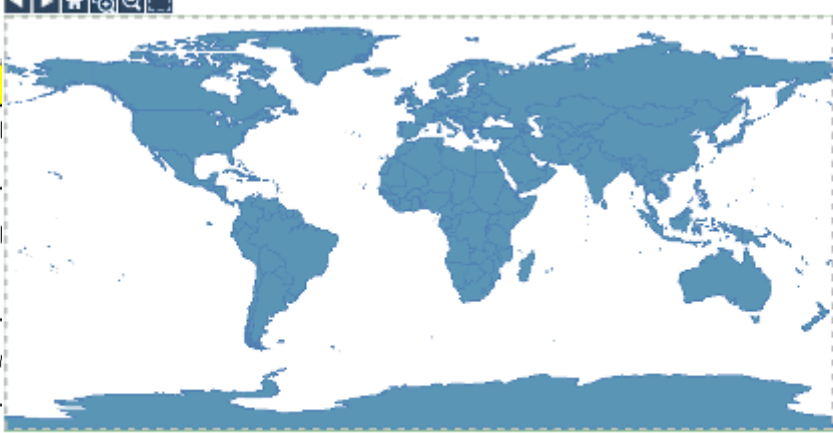
GOME level 2 products

PROBA CHRIS and HR (CHRIS and HRC direct)

Kompsat-1 city dataset

MERIS RR reprocessed

AATSR MERCI reproce



Free text search: info:search/terms/

Query

## GENESI-DR → Single access point VO management

ID	URL	Product	Start Time	End Time	Action
<input type="checkbox"/>	<a href="https://...">https://</a>	MER_RR_1P	2008-06-09T21:39:09.00	2008-06-09T22:23:12.00	show
<input type="checkbox"/>	<a href="https://...">https://</a>	MER_RR_1P	2008-06-09T23:19:45.00	2008-06-10T00:03:49.00	show
<input type="checkbox"/>	<a href="https://...">https://</a>	.._ASA_APP_1P	2008-06-09T01:55:56.580000	2008-06-09T01:55:51.523000	show
<input type="checkbox"/>	<a href="https://...">https://</a>	.._ASA_APP_1P	2008-06-09T01:55:51.744000	2008-06-09T01:56:06.421000	show
<input type="checkbox"/>	<a href="https://...">https://</a>	.._ASA_APP_1P	2008-05-10T11:11:30.641000	2008-05-10T11:11:45.252000	show
<input type="checkbox"/>	<a href="https://...">https://</a>	.._ASA_APP_1P	2008-05-11T19:57:49.549000	2008-05-11T19:58:03.561000	show
<input type="checkbox"/>	<a href="https://...">https://</a>	.._ASA_APP_1P	2008-05-13T19:34:57.351000	2008-05-13T19:35:11.255000	show
<input type="checkbox"/>	<a href="https://...">https://</a>	.._ASA_IMP_1P	2008-06-09T20:00:45.597000	2008-06-09T20:00:59.064000	show
<input type="checkbox"/>	<a href="https://...">https://</a>	.._ASA_IMP_1P	2008-06-10T19:29:17.048000	2008-06-10T19:29:31.376000	show
<input type="checkbox"/>	<a href="https://...">https://</a>	.._ASA_IMP_1P	2008-06-12T02:01:21.296000	2008-06-12T02:01:36.665000	show
<input type="checkbox"/>	<a href="https://...">https://</a>	.._ASA_IMP_1P	2008-06-12T02:01:36.665000	2008-06-12T02:01:56.560000	show

Address	Username
a.int/ra	merfrsusr
a.int/ra	merfrs2
a.int/ra	ra2_data
<a href="https://s-fp.eo.esa.int">s-fp.eo.esa.int</a>	scia1usr
<a href="https://s-dp.eo.esa.int">s-dp.eo.esa.int</a>	dorisusr
<a href="https://s-dp.eo.esa.int">s-dp.eo.esa.int</a>	gomeers*
<a href="https://s-dp.eo.esa.int">s-dp.eo.esa.int</a>	gomeers*
<a href="https://es.eo.esa.int/ra">es.eo.esa.int/ra</a>	probausr
<a href="https://OLISA">OLISA</a>	ITCR3312
<a href="https://es.eo.esa.int/ra">es.eo.esa.int/ra</a>	kompsatusr
<a href="https://y.eo.esa.int/merci/">y.eo.esa.int/merci/</a>	ITCR3312
<a href="https://ci-uk.eo.esa.int">ci-uk.eo.esa.int</a>	merciuser05

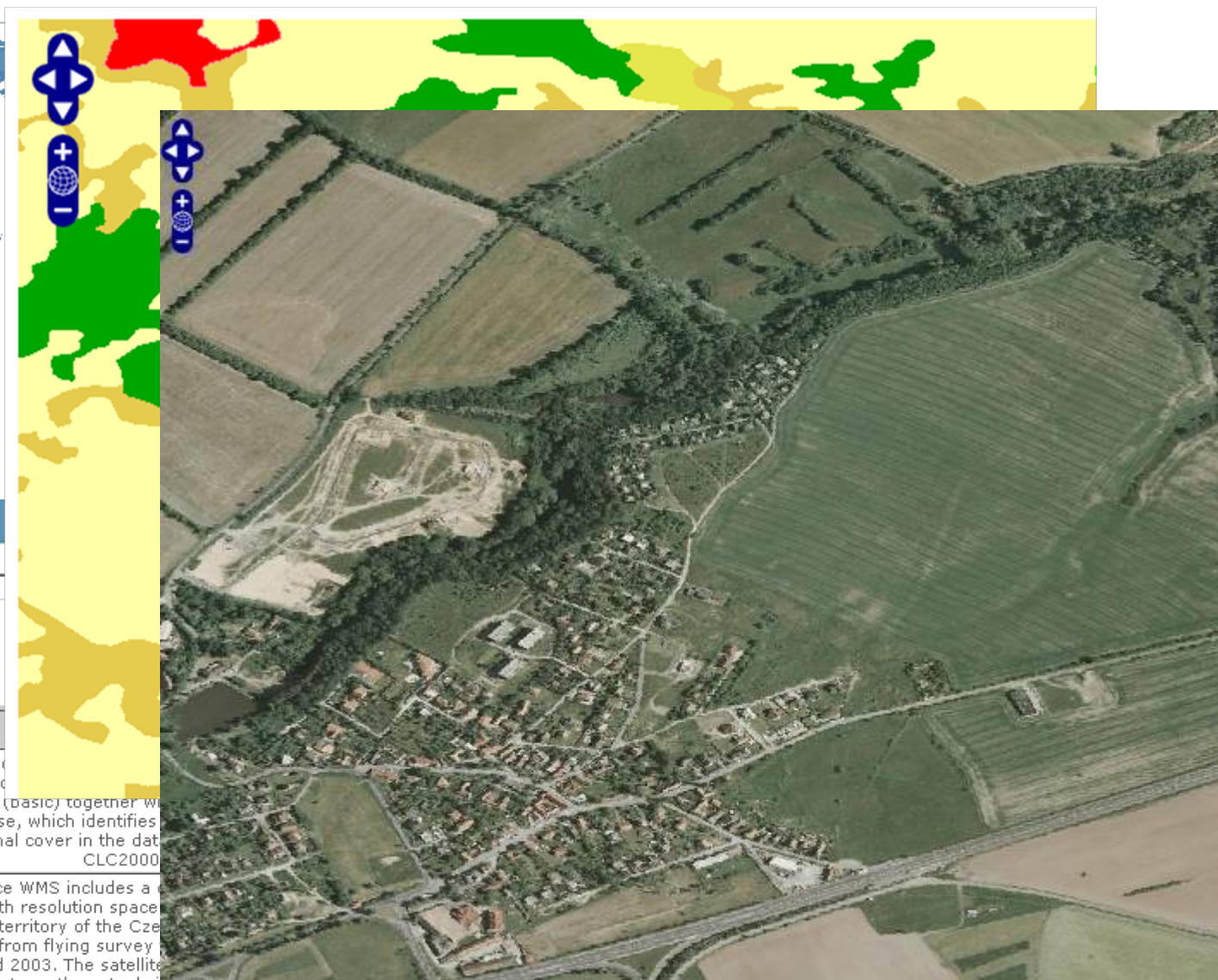




Free text search : q{searchTerms}

czech

Query



ID	Title	
<input type="checkbox"/>	<b>CORINE Land Cover 1990, 2000</b>	The CORINE Land Cover (CLC) is a European Union database, which identifies regional cover in the data. The CLC2000 levels (basic) together with the CLC1990 levels, which identifies regional cover in the data.
<input type="checkbox"/>	<b>Coloured Satellite Photo of the Czech Rep. (1m)</b>	This service WMS includes a map with resolution space of 1m, covering the whole territory of the Czech Republic, executed from flying survey in 2002 and 2003. The satellite photo reflects the actual situation on the ground, but it makes it possible to compare vector data with reality....

User can discover via the **GENESI-DR Web portal data and services** on the base of **free text string** and **metadata...**



such as  
geographical  
area, temporal  
range..

...and other **specific parameters** as applicable

[Click here to watch the movie](#)

# A test case: Ocean data/products discovery, access, and processing

**Aim:** to discover, access, process and visualize (off-line) different typologies of Ocean related data and products

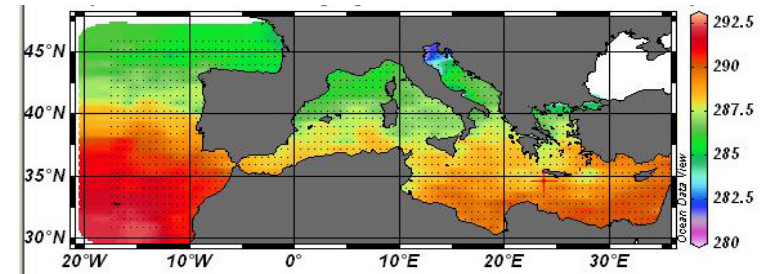
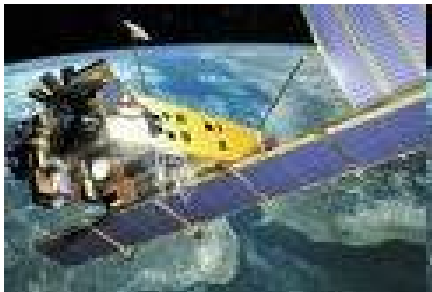
## Data and services:

Sea Surface Temperature maps generated and archived at CNR

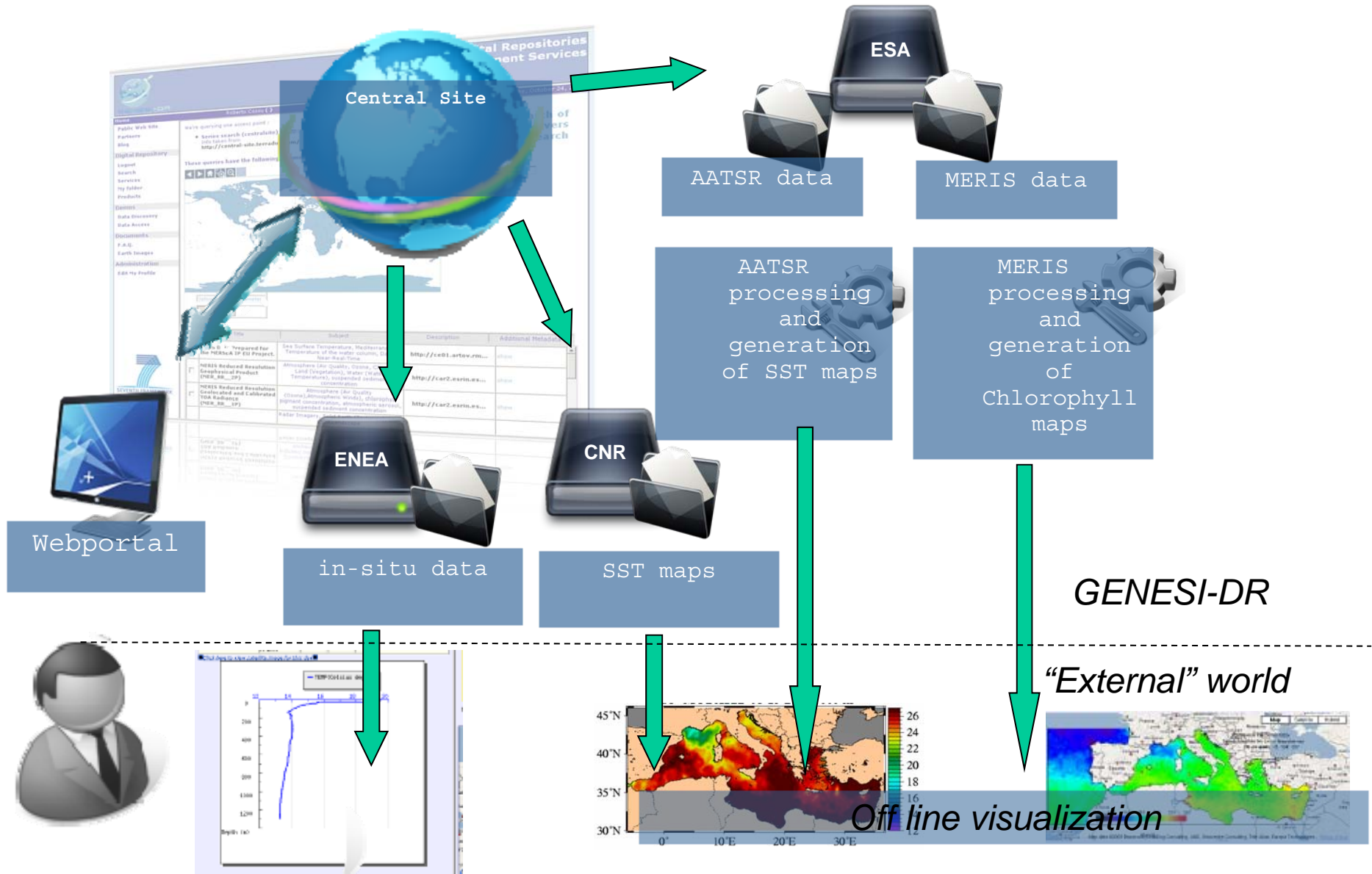
Voluntary Observing Ship (VOS) data archived at ENEA

L3 products (algal bloom, SST) generated on-the-fly from Envisat data (MERIS and AATSR) archived at ESA-ESRIN

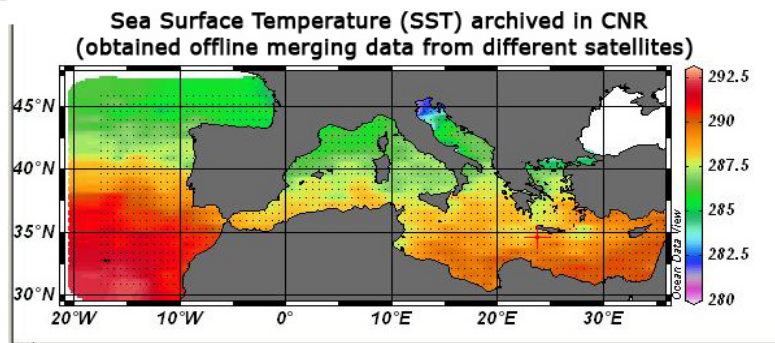
The scenario includes in-situ and satellite data, already generated products, and products generated on the fly from satellite data



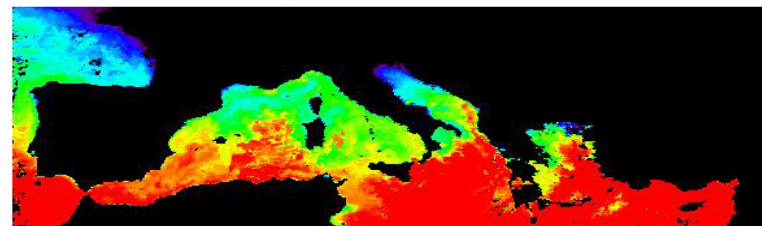
# Ocean data/products discovery, access, and processing



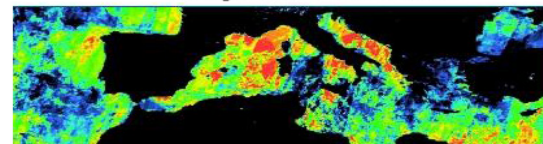
# Ocean data/products discovery, access, and processing



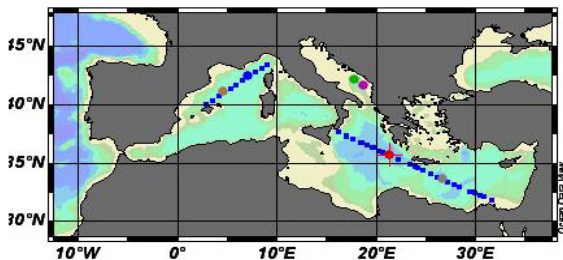
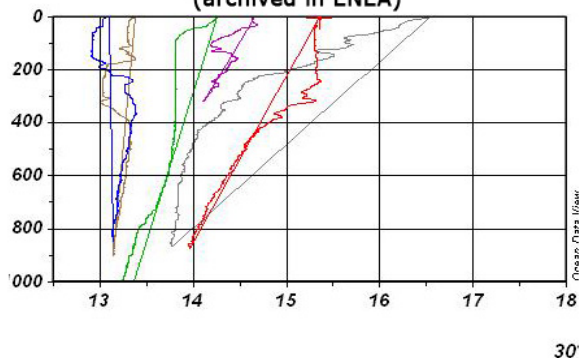
SST generated on-the-fly by GENESI-DR processing service using Envisat AATSR data



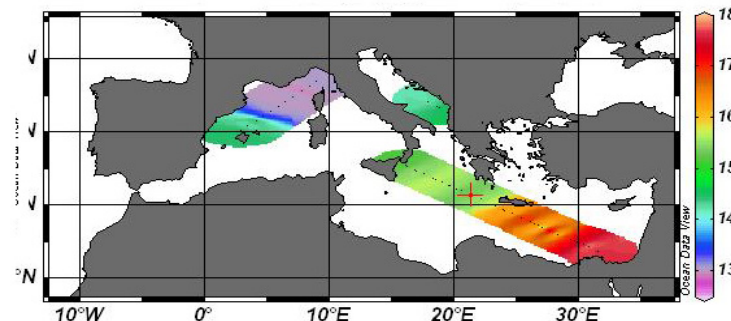
Chlorophyll map generated on-the-fly by GENESI-DR processing service using Envisat MERIS data



Temperature profiles measured by buoys  
(archived in ENEA)



Surface Temperature measured by buoys  
(archived in ENEA)



- Metadata core properties based on the **INSPIRE Implementing Rules for Metadata** as minimal set following the criteria of their usefulness to data discovery and operational setting.
- Design for practical use as a **Dublin Core Application profile** ([dclite4g.xmlns.com](http://dclite4g.xmlns.com)) in a RDF model and reuses known and common vocabularies
- Concern on the elements most useful for data sharing and the building of **search and discovery services**
- Focus on any type of data with spatial-temporal characteristics

- **OpenSearch** was adopted and promoted in GENESI-DR as the minimal compliance level to develop discovery mechanisms in heterogeneous sites
- The **geospatial extension** allow to formulate geospatial requests e.g. point-plus-radius, a bounding box, or a polygon
- Together with the **Time extension**, OpenSearch can specify time start, finish, and slices for searching data.

`start={time:start?}&stop={time:end?}`

- OpenSearch is now considered by **Open Geospatial Consortium** a complementing way to provide a machine-accessible search interface to a data repository (OGC 09-084)

- natural language capabilities on the portal search
  - **map of chlorophyll in the Mediterranean Sea last month**
- Last Month = 2009-05-01/2009-05-31T23:59:59
  - Also valid are date intervals (e.g. from sep 2004 until 2004-10-03, from last week until yesterday)
- Mediterranean Sea = -4,30,26,45
  - Includes global regions (e.g. Caucasus, east tropical Africa), seas and oceans (e.g. Sea of Marmara), countries and regions (cities will be next)
  - Allows the selection of north/south/east/west of area (e.g. North of Portugal)
- Returns direct links and metadata of the products in HMA/EOP, RDF/XML or even in MetaLink (for common download managers)
- This feature is in final testing phase and is being integrating in these days

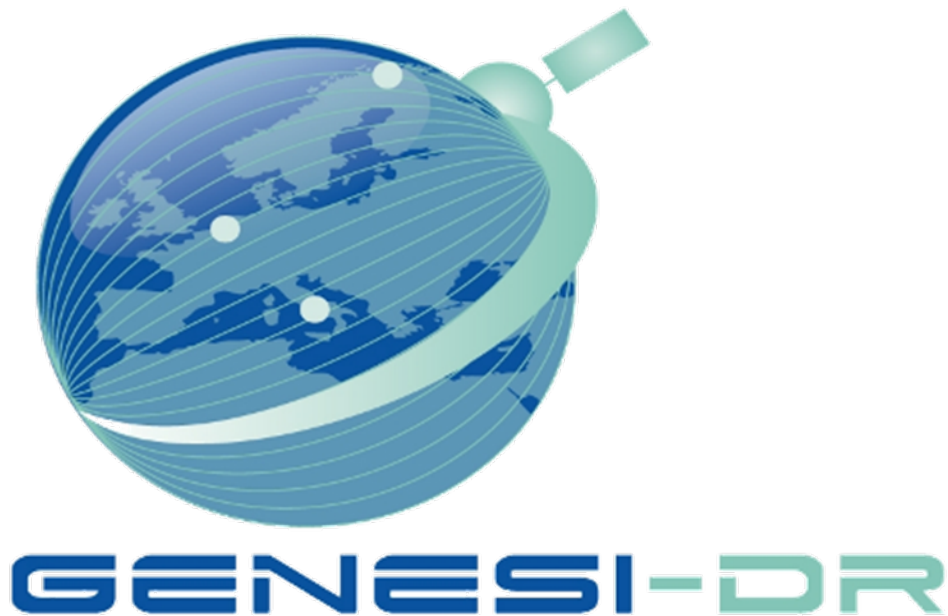


- All collaborations demonstrated the value of the federated approach and **express interest** in GENESI-DR infrastructure to be **evolved and moved in operations**
- GENESI-DR is extendible to the **wide and multidisciplinary Earth Science user communities**, including research, related institutional, industry & operational actors
- Communities still need **customised services** in their specific community environment while sharing data for collaboration

# Ideas for a GENESI-DR follow-up project

## Digital Earth Communities

Virtual Research



INFRA-2007-1.2.1 : Scientific Digital Repositories  
Duration : January 1, 2008 – April 30, 2010  
Total EC funding : 4.4 M€

*To be presented to EC Call  
INFRA-2010-1.2.3: Virtual  
Research Communities  
(deadline 24 Nov 09)*

**If successful ... to be  
started in continuation with  
GENESI-DR**



EC Grant Agreement no. 212073

# Embracing Digital Earth Communities: an International scenario on ...

- **ESFRI Communities:** SeaDataNet, ...
  - New interfaces: EMSO, GEO-SEAS...
- **EC DG Research GEO:** Envirogrid...
- **European DG Regional Development programme ESPON**
  - The **E**uropean **S**patial **P**lanning **O**bservation **N**etwork supports policy development to the aim of territorial cohesion involving geo-economic applied research.
- **European Spatial Data Infrastructure:**
  - Italian Ministry of Environment Cartographic Portal
- **GMES and GEO/GEOSS**
  - International Society for Digital Earth: e.g. CEODE (China), Asian Pacific Network for Climate Change
  - Ready to interface this alliance DA-09-02



## INFRA-2010-1.2.3

- Enable an increasing number of users from all disciplines to access, share and use e-Infrastructures
- Integrate and link regional e-Infrastructures
- Addressing e-Infrastructures for the ESFRI projects and communities is encouraged

## Objectives

## New technical Services & Research activities:

- Provide tools for building multidisciplinary services and customised Digital Earth facilities, e.g:
  - Collaborative services for the mash-ups, execution and management of interdisciplinary experiments
  - Handling of VOs in federated infrastructures
  - Alert services when new relevant information is accessible in the infrastructure
  - Visualization and data manipulation tools of complex results
- Investigate technological paradigm for making
  - Scalable data access, distribution and processing: cloud computing
  - easier and transparent information discovery: ontology and geo-semantics

### **INFRA-2010-1.2.3 Objectives**

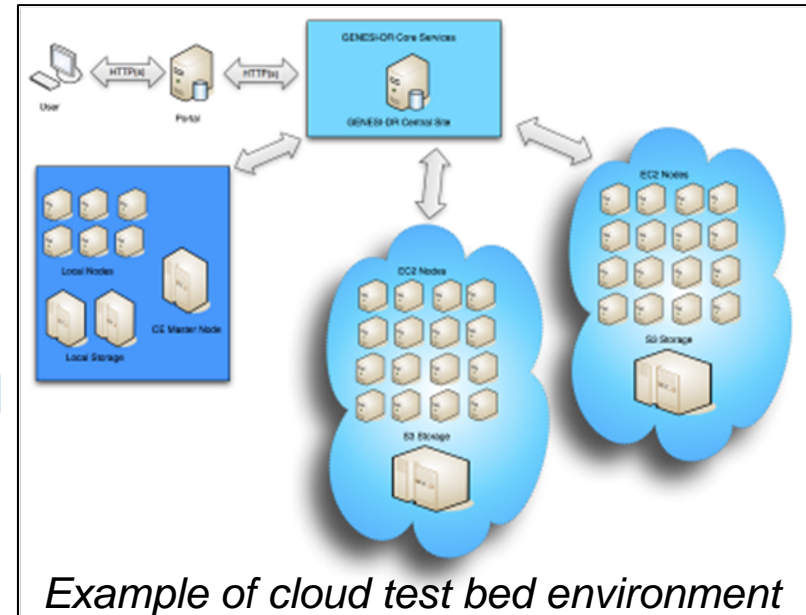
- Deployment of e-Infrastructures in research communities to enable multi-disciplinary collaboration
- Deployment of end-to-end e-infrastructure services and tools for integrating and increasing research capacities
- Build user-configured virtual research facilities and test-beds from collection of diverse resources

**Cloud Infrastructure-as-a-Service** will be experimented in GENESI-DR 2 to offer “infinite” infrastructure on-demand in a complex multidisciplinary/distributed environment:

GENESI-DR 2 should be able to run virtually any software application or configuration while paying only for the capacity and services it consumes

GENESI-DR 2 data could be partially stored and made accessible in Cloud Computing storages with virtually “infinite” capacity

Alternative Cloud infrastructures (existing and under development) will be analysed to assess performance, costs, readiness for operation...



Next GENESI-DR meeting on Data Policy  
ESAC - Madrid 4<sup>th</sup> Dec 2009

Attached to PV 2009 Conference

<http://www.sciops.esa.int/index.php?project=CONFERENCE&page=pv2009>

**Thank You!**

<http://www.genesi-dr.eu> <http://portal.genesi-dr.eu>