THE GENESI-DR INFRASTRUCTURE: AN OPPORTUNITY FOR THE OCEAN SCIENCE COMMUNITY

Andrea Kaiser-Weiss [1], Stefano Migliorini [1], Giuseppe Manzella [2], Roberto Cossu [3], Eliana Li Santi [3], Steven Hosford [4], Luigi Fusco [3]

[1] *University of Reading, Department of Meteorology, Earley Gate, PO Box 243, Reading, RG6 6BB, UK, Email: a.k.kaiserweiss@reading.ac.uk, s.migliorini@reading.ac.uk*

[2] *ENEA CLIM, P.O. Box 224, I-19100 La Spezia, Italy, Email: giuseppe.manzella@enea.it*

[3] *ESA/ESRIN via Galileo Galilei - Casella Postale 64 - 00044 Frascati, Italy, Email: roberto.cossu@esa.int, eliana.li.santi@esa.int, luigi.fusco@esa.int*

[4] *CNES 18, avenue Edouard Belin, 31401 Toulouse cedex 9, France, Email: steven.hosford@cnes.fr*

ABSTRACT

The Ground European Network for Earth Science Interoperations – Digital Repositories (GENESI-DR) is establishing a reliable, easy, and long-term access to Earth science data. GENESI-DR offers a single access point to the heterogeneous data (from space, airborne and in-situ sensors) located at various individual repositories.

Via the GENESI-DR Central Site, secure access is provided to the data stored at the federated archives. The connection of further data repositories and computing resources is encouraged and can readily be accomplished. The GENESI-DR e-infrastructure can inter-operate with other infrastructures (e.g., SeaDataNet), as demonstrated in an ocean-related application.

The GENESI-DR capabilities can be accessed directly by users via the Web Portal, or by automated external applications.

Self-explanatory training material is available on-line (www.genesi-dr.eu).

1. THE GENESI-DR PROJECT

Ground European Network for Earth Science Interoperations - Digital Repositories (GENESI-DR) (www.genesi-dr.eu) is an ESA-led, European Commission funded two-year project, aimed at providing reliable, easy, long-term access to historical and recently acquired Earth science data from space, airborne and in-situ sensors archived in large distributed repositories.

The specific strength of GENESI-DR lies in the concept of offering a single access point to the petabytes of heterogeneous data located at a variety of different European Earth Science Digital Repositories through the same interface in a transparent and homogeneous way.

2. THE GENESI-DR TECHNICAL CONCEPT

All the GENESI-DR capabilities can be accessed by users via web interface, the Web Portal, or by external applications via open standardized interfaces exposed by the system. Automated services can be set up allowing the user to directly retrieve data products without any ordering stage.

Firstly, the user has to register. After being authenticated, users obtain security credentials to be delegated to services acting on their behalf without transferring private keys over the network. Credentials can be delegated to Grid portals. A Grid portal is a web server that provides an interface to Grid services, allowing users to submit compute jobs, transfer files, and query Grid information services from a standard web browser. Users wanting to access the GENESI-DR infrastructure can obtain a credential through the GENESI-DR Portal. Credentials are also generated for hosts on which GENESI-DR services will run and are used by the GENESI-DR services running on those hosts to securely interact with other GENESI-DR services.

GENESI-DR users are organised in groups, each identifying a different set of data and services to which users belonging to that group are allowed to access. VOMS (Virtual Organization Membership Service) serves as a central repository for user authorization information, providing support for arranging users into a general group hierarchy, and keeping track of their roles.

The administrator is managing users, data and security. Monitoring of the overall infrastructure is performed by means of Nagios, which is a popular open source computer system and network monitoring software application. Questions or troubles reported are managed with a ticket system.

Each time a data new series is added in a digital repository, in order to make it available to GENESI-DR, it has to be published to the Central Site. Metadata have to be extracted, mapped to RDF (Resource Description Framework) format and loaded in the underlying database. The generation of the RDF files can be done either manually or by means of DR-specific tools or by using the GENESI-DR Catalogue Generator.

A series management interface dedicated to each DR allows the DR operator to register the new series to the Central Site. Note that the series management page is only available for the DRs currently registered into GENESI-DR. Interested DR operators shall contact the Central Site administrator.

3. A GENESI-DR APPLICATION EXAMPLE

The infrastructure already deployed in September 2009 comprises ESA, CNES, DLR, KSAT, ASI, NILU, Infoterra, JRC, ENEA, and linked digital repositories of ISPL/EGEE and CENIA (see Fig. 1). The data hold by these repositories can be discovered, accessed, and even processed (respecting the individual data policies of the repositories) from a single access point: the GENESI-DR Central Site.

Figure 1. A registered user has direct access to the federated digital repositories (DR's) through the GENESI-DR Central Site. Note the data remain at the federated digital repositories and are not copied to the Central Site, and the individual data DR's policies being respected.

The GENESI-DR e-Infrastructure can inter-operate with other infrastructures (e.g., SeaDataNet) and can be validated against a related application (in this case, the ocean-related application of ENEA and CNR ISAC subset of SeaDataNET distributed data bases).

The user has access to the federated archives and services (e.g., map generation). For instance, users can generate on-the fly sea surface temperature (SST) maps from the Advanced Along Track Scanning Radiometer (AATSR) using the ESA facilities and can choose their own user parameters and can discover, at the same time and from the same portal, the daily SST data generated offline by the Italian National Council of Research CNR and vertical profiles of sea temperature measured by Volunteer Opportunity Ships (VOS) and archived at ENEA (see Fig. 2).
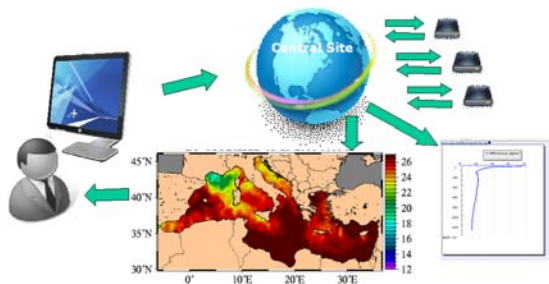

Figure 2. GENESI-DR inter-operates with other infrastructures to provide products and services, for example to provide archived data, or to produce on-the-fly higher level data from satellite measurements, or to generate maps as specified by the end user.

4. TRAINING FOR NEW DATA PROVIDERS AND DATA USERS

GENESI-DR provides training for data access and processing (see *Training* at www.genesi-dr.eu). A short introduction and overview is provided in [1], a full user manual in [2]. Detailed DR-owner tutorials also available from the website are aimed at data repository holders who would like to "genesi-fy" their data, i.e., to link their own data (or data repositories) to GENESI-DR, see also [3]. Training events are regularly organized (please check the project web page for recent events). Please contact the GENESI-DR team for further information.

5. REFERENCES

1. Li Santi, E. & Kaiser-Weiss A. (2009) GENESI-DR Training WorkshopVenice, 24 Sept. 2009. http://www.genesi-dr.eu/index.php?menu=training&type=training_material

2. Li Santi, E. (2008) GENESI-DR User Manual. GENESI-DR Project Report DSA1.10, ESA. http://www.genesi-dr.eu/documents/GENESI-DR-SA1-DEL-DSA1.10-v1.0.pdf

3. Li Santi, E., Goncalves, P., Parrini , A., Sottilaro, M., Pacini, F. (2009) GENESI-DR Operations Manual. GENESI-DR Project Report DSA1.7, ESA. http://www.genesi-dr.eu/documents/GENESI-DR-SA1-DEL-DSA1.7-v1.0.pdf