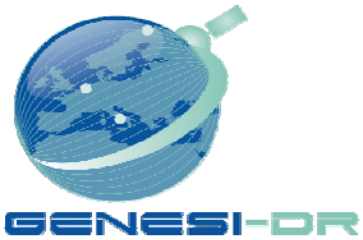

GENESI-DR

Represented by

European Space Agency

EOP-S

Frascati, Italy

Telephone: +390694180530
Facsimile: +390694180532

And with the support of

Terradue srl

Rome, Italy

Telephone: +390680368952

GENESI-DR interoperability with GEOSS-CGI and CGI component evolution and integration analysis in GEOSS Architecture Implementation Pilot – Phase 3 (AIP-3)

Business POC

Luigi Fusco

Luigi.Fusco@esa.int

+390694180530

Technical POC

Roberto Cossu

Roberto.Cossu@esa.int

+390694180607

GENESI-DR Response to the GEOSS AIP-3 CFP

1 Overview

This proposal responds to GEOSS AIP-3 CFP, which seeks organizations to participate in the 3rd Phase of the GEOSS Architecture Implementation Pilot (AIP-3). AIP develops and pilots new process and infrastructure components for the GCI and the broader GEOSS architecture. The main aims of AIP are to reach consensus on Interoperability Arrangements and to register operational components and services that carry forward into persistent operations of GEOSS.

This proposal is presented by two partners of the Ground European Network for Earth Science Interoperations - Digital Repositories (GENESI-DR) project[1], namely the European Space Agency and Terradue, which were significantly involved in the development and deployment of the GENESI-DR infrastructure.

GENESI-DR is a European Commission (EC)-funded two-year project, kicked off in early 2008 which is taking the lead in providing reliable, easy, long-term access to Earth Science data via the Internet. The EC has funded GENESI-DR as a flagship project in Europe to help meet the challenge of facilitating the needs of scientists from different Earth Science disciplines located across Europe to discover, access and use (combining, integrating, processing, …) historical and current Earth-related data from space, airborne and in-situ sensors archived in large distributed repositories.

The GENESI-DR project has already achieved important objectives in providing reliable, easy, long-term access to Earth Science data and associated on-demand processing resources via the Internet[2]. It has designed and implemented a multi-disciplinary platform, providing discovery capabilities for scattered and heterogeneous data, providing easy and fast access to such data, on-demand computing resources, thus making easier the dissemination of newly generated results. Several collaborations with projects and external bodies have been established by GENESI-DR. In particular, the following interfaces with external projects have been set up: EGEE-III NA4 Earth Science Cluster; DG-Research funded SeaDataNet (format conversion, visualisation) and a few others (references can be found on the GENESI-DR web sites).

GENESI-DR achievements have been disseminated in the GEO and GEOSS as an example of effective management of large volumes and diverse types of Earth Observation data. GENESI-DR is playing an active role in Task Force DA09-02.

[1] http://www.genesi-dr.eu

[2] GENESI-DR builds on the operational capabilities developed by predecessors grid infrastructure projects (GPOD - Grid Processing on Demand, http://gpod.eo.esa.int/) which aimed at providing ESA EO data access for on demand processing as needed by Earth Science applications,

With this activity we intend to contribute to GEO-GEOSS by:

- Making GENESI-DR interoperable with GCI so as to provide GEO-GEOSS users with access to additional and heterogeneous data, processing services, and computing resources;
- Analysing how the technical solutions implemented in GENESI-DR can be re-used in GCI to integrate existing components and/or enhance existing ones.

2 Proposed Contributions

2.1 Societal Benefit Area Alignment and Support

Making GENESI-DR interoperable with GCI will provide GEO-GEOSS users with an easy access to heterogeneous data, services, and computational resources. This will be of benefit for Societal Benefit Areas (SBAs). Within DA09-02, GENESI-DR set up cooperation with NASA's Goddard Earth Science Data and Information Services Center (GES-DISC). A scenario has been designed that will make GENESI-DR interoperable with the GES-DISC Interactive Online Visualization ANd aNalysis Infrastructure (Giovanni) tool. The relevant Giovanni instances are the Air Quality and Daily Aerosol Instances, which provide access to data and services for visualizing and analyzing satellite, in-situ, and model data relevant to air quality. However, rather than focusing on a specific SBA, in this pilot project we will focus our attention on technological aspects.

See the following subsection for a brief description of the data that will be available to GEO-GEOSS through GENESI-DR.

2.2 Component and Service Contributions

In this Pilot activity we intend to increase the data and processing services available through the GEOSS component and service.

In detail, we will make GENESI-DR interoperable with GCI in order to increase the number and variety of Earth Science data, processing services, and computing resources available in GEOSS.

Figure 1 gives an overview of the data currently federated to GENESI-DR. All metadata are free and publicly available; data, which are discoverable and accessible through GENESI-DR, remain archived at the Data Provider. Technological solutions have been implemented to guarantee that related usage licenses are fully respected.

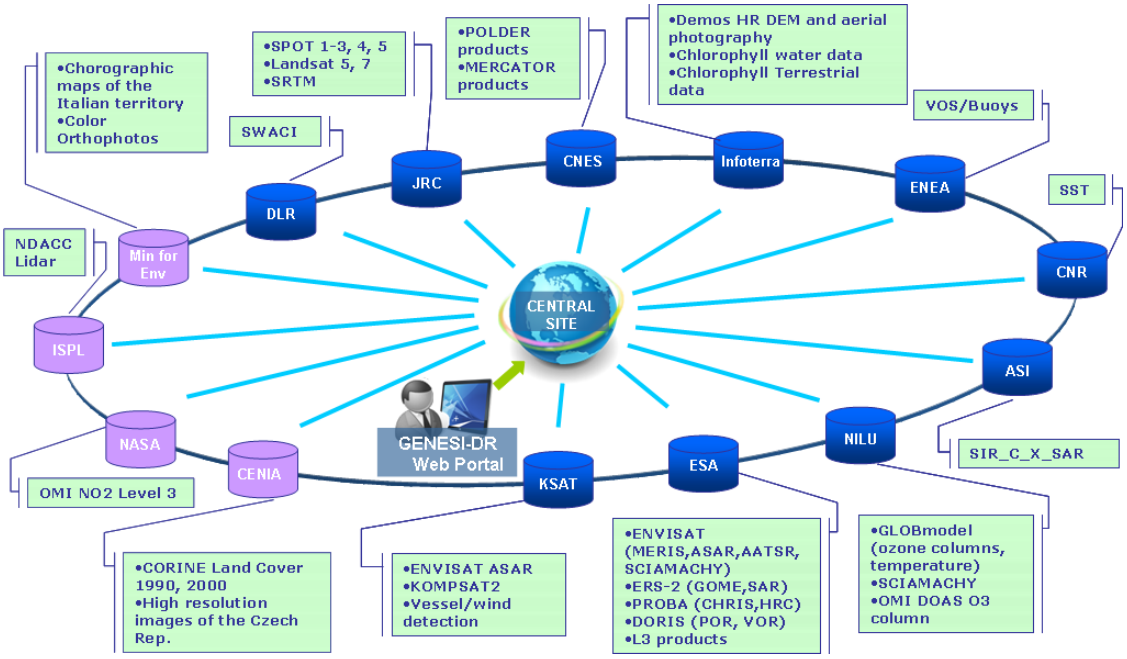

Figure 1 - GENESI-DR federated Digital Resources and available data set series. Controlled access is guaranteed and specific usage licenses are respected.

2.3 Architecture and Interoperability Arrangement Development

During this pilot project, we will analyze how GENESI-DR could contribute to the evolution and integration of GCI components, with the goal of refining the service-oriented architecture and the evolutionary development process.

A description of GENESI-DR architecture is given in the following.

GENESI-DR is responding to application requirements with the design and implementation of a multi-disciplinary platform, providing discovery capabilities of scattered and heterogeneous data (acquired by a multitude of sensors and related to different domains of Earth Science), providing easy and fast access to such data, processing services and computing resources on demand, making the dissemination of newly generated results easier. The GENESI-DR Service Oriented Architecture is implemented using existing and newly developed services, interacting through SOAP and REST interfaces.

In particular, the Central Discovery Service provides the ability to discover and retrieve information regarding data collections and products existing in heterogeneous catalogues at federated DR sites. It identifies the DRs providing products fulfilling the user search criteria and returns the corresponding access points. As a next step, products complying with refined search criteria can be identified and the corresponding metadata are returned to the client which include the product access URL allowing product retrieval. Products never pass through the central site. Only a subset of metadata is stored in the central catalogue to make dataset series discovery possible and to allow GENESI-DR to

automatically forward the refined user query towards the specific DRs. Metadata related to single data products are stored only at the DR owning the data.

Flexibility and performance are taken into consideration by making use of different and efficient data transfer technologies such as HTTPS, GridFTP, and BitTorrent grouped in metalinks as well as specialized geospatial data access services (e.g. the OGC WFS, WCS, and WMS, OPeNDAP). The GENESI-DR architecture provides the DR owners with a mechanism (Catalogue Generator) to produce a metadata catalogue by simply harvesting metadata from their storage systems. The Central Discovery Service communicates with the different DR catalogues through a web service gateway. This provides DR owners the ability to easily make their data and URLs available to a significantly increased audience with no need to duplicate them in a different storage system.

The architecture also embeds services to enable expert users to exploit computational and network resources in order to produce the final desired product. At present, this means passing input data to a processing service (e.g. a Grid Service or an OGC Web Processing Service) available at some sites.

During the proposed activity, detailed plans on how to support the refinement and extension of the currently defined GCI architecture and interoperability arrangements will be produced.

3 Description of Responding Organization

GENESI-DR was initiated in 2008 under ESA coordination and builds on the existing operational European Earth observation infrastructure, involving the key Earth science centres responsible for operational data acquisition, processing, archiving and distribution. Partners include space agencies (DLR, ASI, CNES), space and non-space data providers such as ENEA (IT), Infoterra (UK), K-SAT (NO), NILU (NO), JRC (EU), and industry, such as Elsag-Datamat (IT), CS (FR) and TERRADUE (IT).

This proposal is presented by ESA with the support of TERRADUE, who are finalizing a plan (to be confirmed within the coming two months) to complete the proposed activity in six months.

Contact information:

Business POC	Technical POC
Luigi Fusco	Roberto Cossu
Luigi.Fusco@esa.int	Roberto.Cossu@esa.int
+390694180530	+390694180607