Searching the New Grail: Inter-Disciplinary Interoperability

S. Nativi[(1)], M. Craglia[(2)], L. Vaccari[(2)], M. Santoro[(1)]

[(1)] National Research Council of Italy - IMAA, C.da S. Loja, 85050 Tito Scalo, Italy
[(2)]Spatial data Infrastructures Unit, Institue for Envirnment and Sustainability, Joint Research Centre of the European Commission, Via Fermi, 20127 Ispra, Italy
{nativi, santoro}@imaa.cnr.it, {massimo.craglia, lorenzino.vaccari}@jrc.ec.europa.eu

ABSTRACT

The implementation of the INSPIRE Directive in Europe and similar efforts around the globe to develop spatial data infrastructures and global systems of systems have been focusing largely on the adoption of agreed technologies, standards, and specifications to meet the (systems) interoperability challenge. Addressing the key scientific challenges of humanity in the 21st century requires however a much increased multi-disciplinary effort, which in turn makes more complex demands on the type of systems and arrangements needed to support it. This paper analyses the challenges for interdisciplinary interoperability using the experience of the EuroGEOSS research project. It argues that interdisciplinary requires mutual understanding of requirements, methods, theoretical underpinning and tacit knowledge, and this in turn demands for a flexible approach to interoperability based on mediation and brokering. The paper demonstrates the implications of adopting this approach and charts the trajectory for the evolution of current spatial data infrastructures.

1 INTRODUCTION

One of the most fundamental challenges facing humanity at the beginning of the 21st century is to respond effectively to the global changes that are increasing pressure on the environment and on human society. This priority is articulated by the International Council for Science (ICSU) as follows:

> *"Over the next decade the global scientific community must take on the challenge of delivering to society the knowledge and information necessary to assess the risks humanity is facing from global change and to understand how society can effectively mitigate dangerous changes and cope with the change that we cannot manage. We refer to this field as 'global sustainability research'.[1]"*

ICSU identified five scientific priorities, or Grand Challenges, in global sustainability research (see*Figure 1*) through a broad consultation involving over 1000 scientists from 85 countries in 2009-2010. These Grand Challenges include:

1. Developing the **observation** systems needed to manage global and regional environmental change.
2. Improving the usefulness of **forecasts** of future environmental conditions and their consequences for people.
3. Recognizing key **thresholds** or non-linear changes to improve our ability to anticipate, recognize, avoid and adapt to abrupt global environmental change.
4. Determine what institutional, economic and behavioural **responses** can enable effective steps toward global sustainability.
5. Encouraging **innovation** (coupled with sound mechanisms for evaluation) in developing technological, policy, and social responses to achieve global sustainability.

The increasing importance of linking the scientific effort necessary to underpin the sustainability agenda with innovation and sustainable economic growth is also at the heart of the European Union's Europe 2020 strategy[2], focusing on smart, sustainable, and inclusive growth.

---

[1] http://www.icsu-visioning.org/wp-content/uploads/GrandChallenges_Pre-publication.pdf
[2]COM(2010)2020

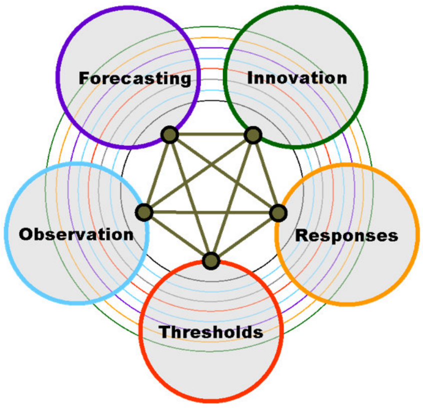

Figure 1:The five ICSU Grand Challenges in Global Sustainability Research

The Global Earth Observation System of Systems (GEOSS[3]), envisioned by the group of eight most industrialized countries (G-8) in 2003 and currently half way in its 10-year implementation plan provides the indispensable framework to integrate the earth observation efforts of the 84 GEO-members and 58 participating organisations. A major role of GEOSS is to promote scientific connections and interactions between the observation systems that constitute the system of systems, and address some of the scientific challenges identified by ICSU with a particular focus on nine societal benefit areas[4]. Such interactions also promote the introduction of innovative scientific techniques and technologies in the component observing systems. In this respect therefore the development of GEOSS can make a strategic contribution in delivering the objectives of the Europe 2020 strategy.

For these reasons the European Commission plays a very active role in developing GEOSS. This includes participating in and co-chairing GEOSS Committees and the Data Sharing Task Force, and implementing important initiatives to collect and share environmental information for the benefit of the global society: the Infrastructure for Spatial Information in Europe (INSPIRE Directive), the Global Monitoring for Environment and Security (GMES) initiative, and the Shared Environmental Information System (SEIS). The European Commission also contributes to the implementation of the GEOSS Work Programme through research projects like EuroGEOSS[5], which are funded from its Framework Programme for Research & Development.

2 PROGRESS AND MAIN RESULTS TO DATE

The concept of inter-disciplinary interoperability and the need for it in managing societal issues is central to the addressing the challenges of sustainability research identified by ICSU. With this in mind, EuroGEOSS was launched on May 1st 2009 for a three year period with the aim to demonstrate the added value to the scientific community and society of making existing earth observing systems and applications interoperable and used within the GEOSS and INSPIRE frameworks. The project builds an initial operating capability (IOC) in the three strategic areas of Drought, Forestry and Biodiversity, and undertakes the research necessary to develop this further into an advanced operating capability (AOC) that provides access not just to data but also to analytical models made understandable and useable by scientists from different disciplinary domains. To achieve the AOC requires research in advanced modelling from multi-scale heterogeneous data sources, expressing models as workflows of geo-processing components reusable by other communities, and ability to use natural language to interface with the models. The extension of INSPIRE and GEOSS components with concepts emerging in the Web 2.0 communities in respect to user interactions and resource discovery, also supports the increased dialogue between science and society, which is crucial for building consensus on the collective action necessary to address global environmental challenges.

[3] http://www.earthobservations.org/geoss.shtml
[4] Disasters, Health, Energy, Climate, Agriculture, Ecosystems, Biodiversity, Water and Weather
[5] http://www.eurogeoss.eu/

EuroGEOSS has completed the first half of its activities. During these first 18 months of the project, the key objectives were:

1. Achieving an IOC, i.e. the development of the services necessary to make it possible to discover view, and access the information resources made available by the partners of the project in the thematic areas of biodiversity, drought, and forestry.
2. Registering these resources as GEOSS components.
3. Developing the framework for assessing the added value of the project and of GEOSS to the communities of users.

All of these objectives have been achieved: the IOC in the fields of biodiversity, drought, and forestry has been established, it has been registered with GEOSS, and a multi-layered framework of surveys and models to assess the longitudinal impact of the project and the benefits of GEOSS have been put in place.

The Forestry IOC has been achieved giving priority to the development of federated metadata[6] catalogues and a map viewer, which are then integrated into the EuroGEOSS brokering framework. These priorities were expressed in an analysis of forestry users' requirements (Figueiredo et al., 2009). The IOC Metadata Catalogue was developed based on the open source package GeoNetwork v2.4.3 and populated with spatial and non-spatial metadata from the European Forest Data Centre at JRC. Metadata adjustments have been made to fit Dublin Core[7] elements and ensure compliance with INSPIRE and relevant ISO 19115 (ISO, 2003), ISO 19119 (ISO, 2005) and ISO 19139 (ISO, 2007) standards. The Metadata Catalogue functionalities and interface have been adjusted to meet the specific forestry theme requirements. As a result the IOC Catalogue provides search, discovery and preview facilities of spatial and non-spatial metadata. The catalogue successfully harvests metadata from national and local forestry catalogues such as those of the national Spanish spatial data infrastructure (IDEE), and is federated in the EuroGEOSS brokering framework so that its resources are globally accessible and viewable by the GEOSS community.

The biodiversity IOC has been achieved based on the analysis of user requirements (O'Tuama et al., 2009) by developing a series of metadata catalogues and services at the partners' institutions, and integrating them into the EuroGEOSS brokering framework. A key milestone has been the development of the metadata catalogue for the Global Biodiversity Information Facility (GBIF[8]) with a specialised profile using the Ecological Metadata Language to support better community needs, especially for species names datasets and natural history collections, and for multiple natural languages. A metadata sharing service has been established based on the Open Archive Initiative, harvesting metadata from the participating GBIF catalogues and integrating them into the EuroGEOSS brokering framework.

In parallel to this and related developments at other partners' institutions, significant work has taken place to develop a Digital Observatory for Protected Areas (DOPA[9]) a facility with initial focus on Africa but with a global reach as part of the GEOBON observation network (GEO, 2008). DOPA will be developed in an iterative way, starting with an information system able to visualise and interact through a single graphical user interface with key datasets hosted by the partners, namely boundaries of protected areas (United Nations Environment Programme - World Conservation Monitoring Centre, UNEP-WCMC), species occurrences (GBIF) and maps of bird distributions (Birdlife International and Royal Society for the Protection of Birds, RSPB). During the execution of the EuroGEOSS project, these developments will become more and more web based allowing the integration of information made available in the other thematic areas. The initial phase of the project has focused on the setting up of a prototype of DOPA that includes a specialised database, an advanced web client, and the preparation of unique datasets regarding bird distributions that will become available in the form of species occurrences via GBIF and in the form of species distribution maps directly through the DOPA.

---

[6]Metadata is a description of an information resource, including key elements such as what it is, who is responsible for it, where can it be found, and how it can be accessed.
[7] http://dublincore.org/
[8] http://www.gbif.org/
[9] http://dopa.jrc.ec.europa.eu/

The Drought IOC has been achieved developing a series of web services to discover, view, and access drought data providers at the European level (EC Joint Research Centre, JRC), regional level (Observatory for South East Europe), and national/regional levels (Spanish Drought Observatory, and observatory for the Ebro river basin). The goal of connecting drought data providers on the three scale levels (continental, national/international, regional/local) was one of the key priorities expressed by users (Hofer, Niemeyer et al., 2010), and its achievement is an important proof of concept of a nested multi-scale system of systems. All the partners have in place an infrastructure for providing web map services (Open Geospatial Consortium - Web Map Server, OGC WMS) (OGC, 2006) and update their services regularly. Some partners (EC-JRC and University of Lubjana) provide also web map services of time series (WMS-T) for accessing data sets of a chosen date or period.

The integration of services from different partners in a common viewer, i.e. the map viewer of the European Drought Observatory (EDO), allows the linkage to services from the other thematic areas (e.g. forest) and opens new options for drought data analysis. These options will be further explored in the second half of the project. In addition to the European perspective, an interoperable EDO contributes to a future Global Drought Early Warning System under consideration by the World Meteorological Organization (WMO), and GEO/GEOSS. To this end, a prototype Global Drought Monitor has been established as a first building block of the Global Drought Early Warning System in partnership with the North American GEO/GEOSS community, the U.S. National Integrated Drought Information System (NIDIS) and the Princeton African Drought Monitor prototype. A first demonstration pilot of such Global Drought Monitoring System has been achieved and was demonstrated at the GEO Beijing Summit in November 2010[10].

Central to the multidisciplinary IOC is the EuroGEOSS discovery broker, which is a component able to read and mediate among the many standards and specifications used by different scientific communities. By building bridges among the practices of these communities, the broker makes it possible to search, and discover the resources available from heterogeneous sources. During this initial phase, the EuroGEOSS discovery broker gives access to over 400 datasets and 26 services, including multiple catalogue services in the three thematic areas. By registering the broker as a GEOSS component, all of the thematic resources of the project are also accessible to the global research community. The following Section discusses briefly the key achievements and challenges of the brokering approach developed for EuroGEOSS.

3 THE BROKERING APPROACH

The EuroGEOSS multidisciplinary IOC was built on the comparative analysis of the thematic user requirements (Vaccari et al., 2010) and is developed applying several of the principles/requirements that characterize the System of Systems (SoS) approach and the Internet of Services (IoS) philosophy:

1. Keep the existing capacities as autonomous as possible by interconnecting and mediating standard and non-standard capacities.
2. Supplement but not supplant systems mandates and governance arrangements.
3. Assure a low entry barrier for both resource users and producers
4. Be flexible enough to accommodate existing and future information systems as well.
5. Build incrementally on existing infrastructures (information systems) and incorporate heterogeneous resources by introducing distribution and mediation functionalities to interconnect heterogeneous resources.
6. Specify interoperability arrangements focusing on the composability of inter-disciplinary concepts rather than just the technical interoperability of systems.

[10] seehttp://www.ogcnetwork.net/pub/ogcnetwork/GEOSS/AIP3/pages/Demo.html both Drought European and Drought Global

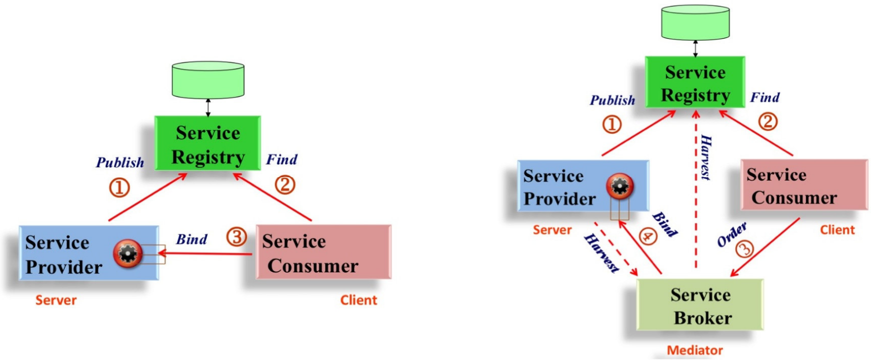

Figure 2: SOA and SOA-brokering Approach, source (Nativi and Bigagli, 2009)

The key features of the EuroGEOSS multidisciplinary IOC are the brokering and mediation capabilities that allow discovering and accessing autonomous and heterogeneous resources from the three thematic domains of the project. This is achieved by applying a *Brokering* approach. This approach extends the traditional SOA archetype introducing an "expert" component: the Broker (see*Figure 2*). It provides the necessary mediation and distribution functionalities which, in contrast, are allocated to the service consumer for the traditional SOA approach.Mediation addresses the communication between one client and one server of different types (heterogeneity issue). The main task of an ideal "mediation component" is to integrate heterogeneous servers by adapting their technological (protocol), logical (data model) and semantic (concepts and behavior) model (Nativi and Bigagli, 2009). Distribution addresses the communication between one client and multiple servers of the same kind, which are presented as if they were a unique instance.The brokering approach allows client applications to:*(i)* bind to heterogeneous service providers in a transparent way (mediation), and *(ii)* interact with them using a single and well-known end point (distribution).

Such a solution addresses significant shortcomings characterizing the present SOA implementations for global frameworks. In fact, the SOA pattern is usually implemented in systems where service providers and client applications share both a communication protocol and a data and metadata model. In complex, large, and heterogeneous infrastructures, like GEOSS, these pre-conditions cannot be ensured and seem unrealistic. Demonstrating the added value of this brokering approach is therefore one of the main contributions of EuroGEOSS to the development of GEOSS as anIoS.

The EuroGEOSS Discovery Broker provides the IOC with harmonized discovery functionalities by applyng the brokering approach.Bymediating and distributing user queries against tens of services –presently registered in the EuroGEOSS capability as either catalogs or inventory servers which propagate the query to many other resources – the EuroGEOSSDiscovery Broker makes it possible to select among a list of well-adopted traditionalinterfaces and emerging Web 2.0 services, and easily utilize them. This list includes the service interfaces that comply with INSPIRE and/or OGC, service interfaces which are specific to the three thematic areas, and service interfaces which are well-used by other communities (e.g. Thematic Realtime Environmental Distributed Data Services, THREDDS[11] and Open-source Project for a Network Data Access Protocol, OPeNDAP[12]) or projects (e.g. Ground European Network for Earth Science Interoperations - Digital Repositories, GENESIS-DR[13] and SeaDataNet[14]). Building these bridges to different communities makes it possible to serve the multi-disciplinary needs of scientific research without assuming that everyone mustconverge on one selected standard – hence, lowering the "entry barrier" for data providers.

In turn, the EuroGEOSS Discovery Broker exposes a set of international standard and well-known service interfaces: OGC CSW Core (OGC, 2007), OGC CSW ISO Application Profile (OGC, 2007a), OGC CSW ebRIM CIM and EO Extension packages (OGC, 2007c; OGC, 2008), OpenSearch[15] (with Geo and Time extensions), and OAI-PMH[16].

[11] http://www.unidata.ucar.edu/projects/THREDDS/
[12] http://www.opendap.org/
[13] http://portal.genesi-dr.eu/
[14] http://www.seadatanet.org/
[15]http://www.opensearch.org/Home
[16]http://www.openarchives.org/pmh/

The EuroGEOSSDiscovery Broker is based on the GI-cat technology(Santoro et al., 2010; Nativi and Bigagli, 2009).

Figure 3 depicts the role played by the Discovery Broker in bringing together the capabilities provided by the three thematic areas and those shared by other Communities –e.g. Climatology, Meteorology, Oceanography and Hydrology. A partial list of the supported service interfaces is showed.

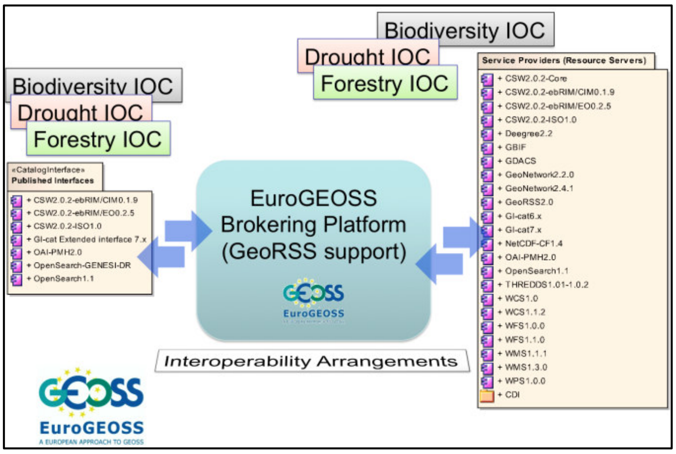

Figure 3: Broker supporting multiple practices

Applying the same principles and technology, recently, a broker to facilitate multi-disciplinary data access was introduced and is part of the EuroGEOSS IOC. The EuroGEOSS Access Broker makes it possible for users to access and use the datasets resulting from their queries which are returned to them based on a common grid environment they have, previously, specified by selecting the following common features: Coordinate Reference System (CRS), spatial resolution, spatial extent (e.g. subset), data encoding format.

In keeping with the SoS principles, the EuroGEOSS Data Access Broker carries out this task by supplementing, but not supplanting, the access services providing the datasets requested. That is achieved by brokering the necessary transformation requests (those that the access services are not able to accomplish) to external processing services. Following the IoS and Web 2.0 principles, the broker publishes Web applications allowing users to: *(i)* select a default business logic (i.e. algorithms) implementing dedicate processing like CRS transformation and space resolution resampling;*(ii)* upload their own business logic (i.e. processing schemes) and set it as default; *(iii)* select the order of the processing steps. The EuroGEOSS Data Access Broker also publishes an interface which realizes the INSPIRE transformation service abstract specification (Howard et al., 2010).

When developing multidisciplinary infrastructures, an important challenge is semantic interoperability(Smits et al, 2007; Klien et al., 2004; Lemmens et al, 2006). To address this, EuroGEOSS, in collaboration with FP7 GENESIS [17] project, prototyped a semantic Discovery Augmentation Component (DAC) (Santoro et al., 2010a) which harnesses the Discovery Broker capacity. It implements a "third-party discovery augmentation approach": enhancing discovery capabilities of infrastructures by developing new components that leverage on existing systems and resources to automatically enrich available geospatial resource description with semantic meta-information. In fact, the EuroGEOSS DAC is able to use existing discovery (e.g. catalogs and discovery brokers) and semantic services (e.g. controlled vocabularies, ontologies, and gazetteers). DAC provides users with semantics enabled query capabilities –contributing to bridge a gap which is important for multi-disciplinary SOA infrastructures.

The EuroGEOSS DAC federates both multilingual controlled vocabularies providing semantics (i.e. Simple Knowledge Organization System*,* SKOS, repositories) and ISO-compliant geospatial catalogue services. The DAC can be queried using common geospatial constraints (i.e. what, where,

[17] http://www.genesis-fp7.eu/

when, etc.). Currently, two different augmented discovery styles are supported: *(i)* automatic query expansion; *(ii)* user-assisted query expansion.

*Figure 4*shows a snapshot of a Web 2.0 application (i.e. a client of the DAC) which was specifically developed allowing EuroGEOSS users to perform semantically augmented searches.

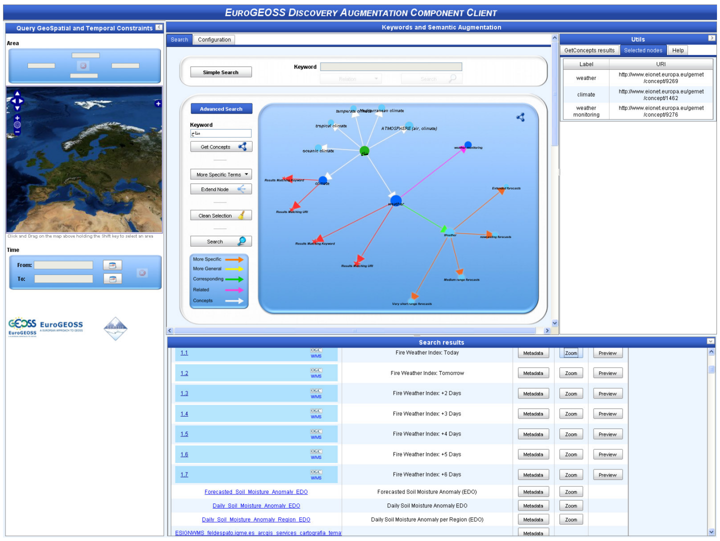

Figure 4: SemanticallyAugmentedSearch

4 DISCUSSION AND NEXT STEPS

There exist clear challenges on using and integrating multi-disciplinary resources to develop cross-disciplinary applications.They include:

- High Entry Barrier: users need to "learn" and develop many (sometimes, immature) information technologies.
- Limited functionalities: international community has mainly focused on discovery functionality implementation; while, cross-domain evaluation, transfer and use functionalities are still lacking.
- Limited semantic interoperability: interoperability for heterogeneous disciplinary resources and different domain semantics are still main issues.
- Limited sustainability: as for scalability, a flat approach to interconnect resources is not sustainable in presence of hundreds of thousands of (heterogeneous) entries and hundreds of registered standards; as for flexibility, future systems and specifications must be easily added, as well.

EuroGEOSS experimented a brokering framework to address these challenges. In fact, this solution can provide a homogeneous discovery, evaluation, and access framework to heterogeneous resources in a seamless way for users –lowering the entry barrier. It is able to implement conceptual composability (not just technical interoperability) allowing a major flexibility and scalability. It can enable semantic query by making use of existing semantic engines –developed by the diverse communities, preserving their autonomy and replacing them, if necessary.

This is achieved by extending the SOA approach and advancing it through the use of "expert" components. Brokers proved to be effective components on which the emerging IoS approach can be realized.

During the next 18 months the project will build its AOC, so that it is possible to access and use not just data across multiple thematic areas but also models and analytical process expressed in workflows and implemented through web-based chains of services. The main expected impact of this development is to make the EuroGEOSS resources accessible and usable not only from specialists in the individual fields, but also from scientists from multiple disciplines that will be able to have a clear picture of how the resources available can be used to address specific questions and how they may be adapted for their specific needs. In addition, the work already started in the project on natural language interfaces and lessons to be learned from Web 2.0 social networks offer the opportunity to expand the use of the EuroGEOSS infrastructure to a much wider audience that transcend scientific disciplines.

BIBLIOGRAPHY

Figueiredo, C., Gaigalas, G., Achard, F. and Iglesias, J.M., 2009, Report on User Requirements for the EuroGEOSS Forestry Operating Capacity. EuroGEOSS D3.1, http://www.eurogeoss.eu/Documents/EuroGEOSS_D3_1.pdf, Last date accessed 1/2011.

GEO, 2008, GEO Biodiversity Observation Network Concept Document, GEO-V document 20, Geneva, Switzerland,http://www.earthobservations.org/documents/cop/bi_geobon/200811_geobon_concept_document.pdf, Last date accessed 1/2011.

Hofer, B., Niemeyer, S., Ceglar, A., et al., 2010, Report on User Requirements for the EuroGEOSS Drought Operating Capacity. EuroGEOSS D5.1,http://www.eurogeoss.eu/Documents/EuroGEOSS_D_5_1.pdf, Last date accessed 1/2011.

Howard, M., Payne, S. and Sunderland, R., 2010,Technical Guidance for the INSPIRE Schema Transformation Network Service, http://inspire.jrc.ec.europa.eu/documents/Network_Services/JRC_INSPIRE-TransformService_TG_v3-0.pdf, Last date accessed 1/2011.

ISO, 2003,Geographic Information –Metadata, ISO/IEC_19115:2003.

ISO, 2005, Geographic Information – Services, ISO_19119:2005.

ISO, 2007, Geographic Information – Metadata – XML Schema Implementation, ISO/TS_19139:2007.

Klien, E., Einspanier, U., Lutz, M., &Hubner, S., 2004, An Architecture for Ontology-based Discovery and Retrieval of Geographic Information, in: Toppen, F. &Prastacos, P. (Eds.): 7th Conference on Geographic Information Science (AGILE 2004): 179-188.

Lemmens, R., Wytzisk, A., de By, R., Granell, C., Gould, M., van Oosterom P., 2006, Integrating Semantic and Syntactic Descriptions to Chain Geographic Services, *IEEE Computer Society*.

Nativi,S. and Bigagli, L., 2009, Discovery, Mediation, and Access Services for Earth Observation Data, Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of Volume 2, No 4, Page(s):233– 240.

Open Geospatial Consortium, 2007, OpenGIS® Catalogue Services Specification 2.0.2, OGC 07-006r1.

Open Geospatial Consortium, 2007a, OpenGIS® Catalogue Services Specification 2.0.2 - ISO Metadata Application Profile, OGC 07-045.

Open Geospatial Consortium, 2007c, OGC® Cataloguing of ISO Metadata (CIM) Using the ebRIM profile of CS-W (0.1.8), OGC 07-038r1.

Open Geospatial Consortium, 2008, EO Products Extension Package for ebRIM (ISO/TS 15000-3) Profile of CSW 2.0 (0.1.9), OGC 06-131r4.

Open Geospatial Consortium,2006,Web Map Server (WMS) Implementation Specification ver. 1.3.0, OGC_06-042.

O' Tuama, E., Dubois., G., Cottam, A., May, I. and Fisher., I., 2009, Report on User Requirements for the EuroGEOSS Biodiversity Operating capacity. EuroGEOSS D 4.1, http://www.eurogeoss.eu/Documents/EuroGEOSS_D4_1.pdf, Last date accessed 1/2011.

Santoro, M., Boldrini, E., Papeschi, F., Baldini, A., Nativi, S., (ed.s), (2010) Specification of EuroGEOSS Initial Operating Capacity.EuroGEOSS Deliverable D.2.2.2.available at:http://www.eurogeoss.eu/Documents/EuroGEOSS_D_2_2_2.pdf, Last date accessed 1/2011.

Santoro, M., Mazzetti, P., Fugazza, C., Nativi, S.andCraglia, M., 2010a, Semantics Enabled Queries in EuroGEOSS: a Discovery Augmentation Approach, American Geophysical Union, Fall Meeting 2010.

Smits, P., Friis-Christensen, A., 2007, Resource Discovery in a European Spatial Data Infrastructure, *IEEE Transactions on Knowledge and Data Engineering*, Vol. 19, No. 1.

Vaccari, L.,Nativi, S. and Santoro, M. (2010) Report on Requirements for Multidisciplinary Interoperability. EuroGEOSS D2.1.1. http://www.eurogeoss.eu/Documents/EuroGEOSS_D_2_1_1.pdf, Last date accessed 1/2011.