

The GEOSS Common Infrastructure for the heavy metal pollution community applications

P. Mazzetti¹, S. Nativi¹, M. Santoro¹ and E. Boldrini¹

¹ Institute of Atmospheric Pollution Research, National Research Council of Italy, Rome, ITALY, {paolo.mazzetti;stefano.nativi;mattia.santoro;enrico.boldrini}@cnr.it

Key words: GEOSS, GCI, multidisciplinary interoperability, data sharing

Introduction

The Group on Earth Observations (GEO) is coordinating efforts to build a Global Earth Observation System of Systems (GEOSS) to provide decision-support tools to a wide variety of users. The GEOSS is conceived as a 'system of systems' linking together existing and planned observing systems around the world and supporting the development of new systems where gaps currently exist. This is made technically possible through a set of software components (portals, registries, specialized services) that constitute the so-called GEOSS Common Infrastructure (GCI). The GCI is in charge of addressing all of the issues to make possible the interaction of the users with the systems that are connected to GEOSS. Due to the need of keeping existing systems autonomous in terms of governance and technical evolution, interoperability (i.e. the ability of two or more systems or components to exchange information and to use the information that has been exchanged) is a major issue in building the GEOSS.

The current GCI is based on a few principles:

- GEOSS Data Sharing principles
 - full and open exchange of data (not necessarily free of charge);
 - Brokering framework.
- Supplement existing systems without supplanting them
 - Lower entry barrier for both users and resource providers.

A cyber-infrastructure for the Heavy Metal Community

Heavy metal pollution can compromise human health and broadly affect the environment. The integration of information about heavy metal pollution and other environmental parameters (weather, land-use, etc.) may support prevention, early warning, research, health-care

planning and delivery, and timely public alerts.

Therefore it is important that the heavy metal pollution community develops a community cyber-infrastructure which is compliant with the existing and future multidisciplinary infrastructure such as GEOSS and others (e.g. EarthCube, EyeonEarth, etc.).

To achieve this objective several interoperability challenges must be addressed:

- organizational;
- semantic;
- technological;
- legal, data policy;
- economical challenges.

This paper is going to focus on semantic and technological challenges.

Technological Challenges

Disciplinary level

Concerning technological challenges at the disciplinary level (e.g. biodiversity, human health, ecosystems, etc.), interoperability is achieved by adopting (a set of) common standards as far as:

- metadata and data model;
- encoding format and language;
- Web-based protocols to share resources (i.e. data and metadata).

Metadata are data about data (i.e. a description of data); standardization is essential to allow data "discovery" and "evaluation". While data model and encoding format/language standardization is necessary to allow data "use". Eventually, Web protocols standardization is essential to guarantee digital systems interoperability -e.g. interoperability among the data discovery, access, and

visualization services managed by different organizations.

These set of disciplinary standards concur (along with the analogous semantic specifications, a common data policy, and other best practices to form a disciplinary "service bus".

Multidisciplinary Level

At the multidisciplinary level, the different disciplinary "service buses", must be interconnected by adopting mediation and proxy solutions to lower the entry barriers for scientists.

The recent adoption of the "brokering approach" in the GCI implements such interconnections, making possible to provide harmonized discovery and access of datasets from heterogeneous sources.

Scientists are able to discover datasets from catalogues and inventories belonging to different cyber-infrastructures in a seamless way. Moreover, they can access selected datasets on a Common Environment (same Coordinate Reference System, resolution, format). On-going activities include the access and chaining of environmental models.

A Heavy Metal Pollution cyberinfrastructure can be developed to provide specific discovery and access services and then can be connected to the GCI (and other multidisciplinary systems) through the brokering approach and technological solutions.

Semantic Challenges

Disciplinary Level

As for semantic challenges at the disciplinary level (e.g. biodiversity, human health, ecosystems, etc.), interoperability is achieved by adopting (a set of) common standards as far as:

- controlled vocabularies/thesauri and ontologies;
- gazzeters;
- Web-based protocols to share resources (i.e. ontologies and gazzeters).

Controlled vocabularies are essential to specify a list of well-known and recognized terms to be used and shared in a given Community. Gazzeters are important to share a set of well-known (temporal) events, spatial

phenomena/features, etc. Eventually, Web protocols standardization is essential to guarantee digital systems interoperability -e.g. interoperability among the discovery and access services to ontologies and gazzeters managed by different organizations.

These specifications are comprised by a Community "service bus".

Multidisciplinary Level

At the multidisciplinary level, the GCI is able to provide mediation between different knowledge services including multilingual thesauri, and ontologies to enrich data discovery with multilinguality, suggestions, and query expansion (e.g. towards broader and narrower concepts). In this way specific community thesauri can be developed and made accessible through the GCI for multidisciplinary applications.

Results and Discussion

The multidisciplinary capabilities of the GCI have been successfully demonstrated at the last GEO Plenary in Istanbul (November 2011).

Presently, the GCI is brokering millions of resources managed by many different networks and systems dealing with all the nine Societal benefit Areas of GEOSS (Disaster, Health, Energy, Biodiversity, Ecosystems, Water, Climate, Weather, Agriculture).

Conclusion

The Heavy Metal Community is asked to develop a cyber-infrastructure to share data and services and contribute to international initiatives, such as GEOSS, Eye-on-Earth, and EarthCube.

The Community must face technological and semantic challenges to build a Community "service bus". Once ready, Brokering solutions (like the GCI) will interconnect the Heavy Metal capacity to the other disciplinary ones allowing multidisciplinary applications and systems.

Acknowledgements

The authors thank the financial supports provided by the European Commission through the project EuroGEOSS and GMOS of the FP7.