# Data and metadata management automation for an effective approach to sharing environmental data

F. D'Amore[1,2], S. Cinnirella[1] and N. Pirrone[1]

[1] CNR-Institute of Atmospheric Pollution Research, c/o: UNICAL-Polifunzionale, 87036 Rende ITALY, damore@iia.cnr.it
[2] University of Calabria, Dept. of Electronic Informatics and System (DEIS), Rende(CS), ITALY, fdamore@deis.unici.it

**Abstract.** In the market of geospatial systems there are several applications that handle data. Some of these components are oriented towards managing and storing data, map visualizations, and data streams, while other components are oriented towards data description and the creation of metadata. Neither of these applications offers users an overview of the process leading to the data creation up to its export to the end user. In the existing literature there are attempts to automate metadating - the term given to descriptions of data. Some companies are also trying an approach that allows, by means of workflow systems, the automation of the creation of a geospatial dataset. As a result, users of geospatial data are increasingly looking for a more structured process of managing geospatial data and metadata, and any tool that handles this process is likely to find reasonable success within this community. Public entities, mainly local ones, are often called on to deal with regulations that require web systems that reveal parameters of any given territory to users/citizens. Likewise, research organizations, especially those dealing with the environment, increasingly find themselves analyzing spatial data. Difficulties that arise when handling such data can be overcome using the approach we are proposing, which involves using a single tool that handles all the necessary steps in exporting spatial data. In this paper we present several methodologies used to manage geospatial data and metadata by means of GeoInt, a middleware tool developed at CNR-IIA that manages geospatial data produced in different research projects. GeoInt is a middleware program offering basic services that permit users to define both data and metadata. It allows users to manage map servers, and allows them to control the download and sharing processes. This research illustrates the ways in which GeoInt has been improved to minimize metadata editing by its users.

**Key words:** GeoInt, SDI, Gi-Cat, GIS, Metadata

## Introduction

To coordinate cross-disciplinary systems for promoting GEOSS and contribute to the GEOSS Common Infrastructure (GCI), CNR-IIA required its researchers to participate in an e-infrastructure to develop a Spatial Data Infrastructure (SDI). Based on the principle of interoperability, this e-infrastructure aims to store, mine and visualize information, and it was based on open-source tools that are compliant with widely-used standards. In order to simplify the interaction between users and the SDI, we developed a middleware that manages the necessary processes for data integration (D'Amore et al., 2012). That middleware, named GeoInt, is currently used for many national and international projects, facilitating several data management processes and exporting geospatial services, in order to build the SDI.

GeoInt is more than a set of services, it is a system that is continually under development, that is oriented towards defining a methodology for geospatial data management. In this paper we present the latest
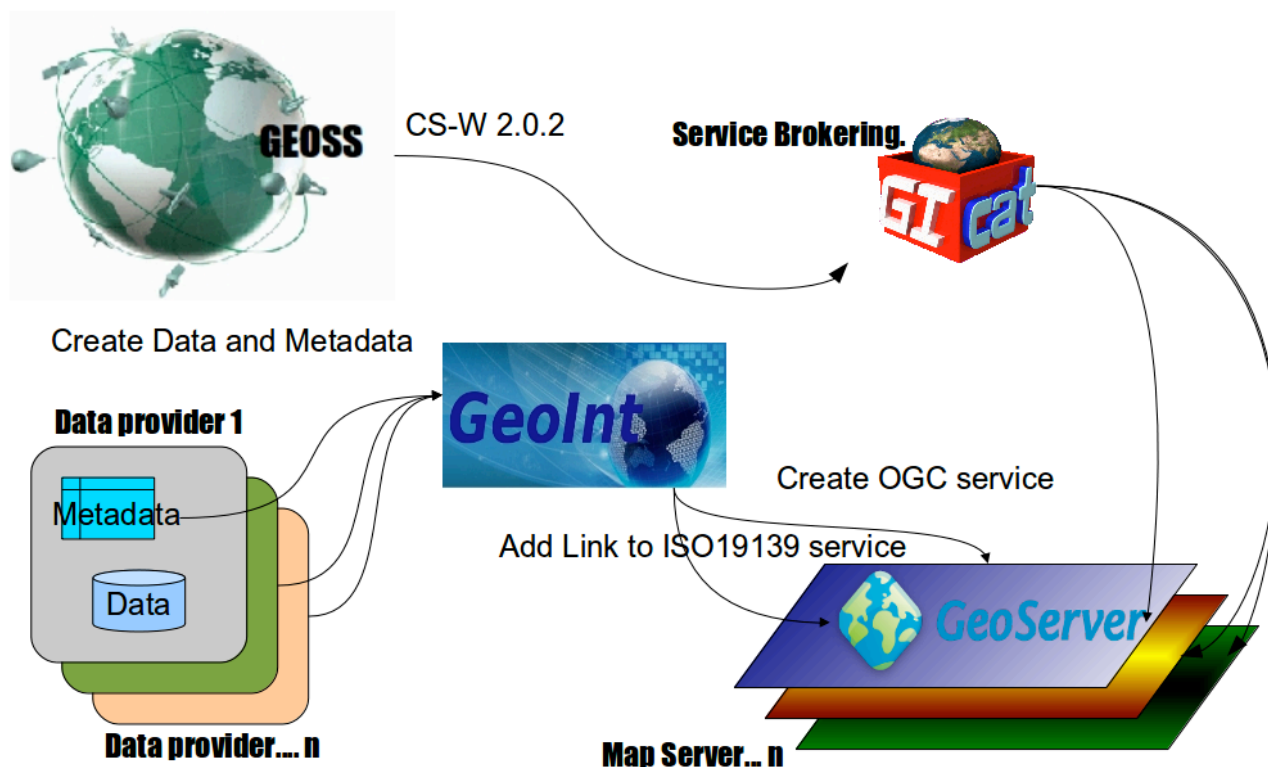
**Fig. 1**. The process schema demonstrating how data and metadata are uploaded in GeoInt, managed by Geoserver, and Gi-cat and then integrated into GEOSS

improvements associated with minimizing metadata editing.

Using Gi-Cat (Nativi et al. 2009), a Service Broker developed at the ESSI-LAB, metadata is collected directly from Geospatial services, as described in Fig. 1.

In the geospatial systems community, there are several applications that handle various aspects of creating and defining data. Some of these components are oriented to managing and storing data, others define thematic maps or data streams, and still others are oriented to describing the process of metadata creation. Yet, none of these applications offer users the possibility to have an overview of the processes leading from data creation all the way to its export to the end user. In the existing literature, there exist attempts to automate metadata creation. Some companies are beginning to try an approach that allows, by means of workflow systems, for the automation of the creation of maps. Hence, a more structured process of preparing geospatial data has become more important. In this paper, we describe GeoInt as a prototype of one such tool, capable of defining global processes with regard to geo data management.

GeoInt is a tool developed as a Web Application and distributed as a SaaS (Software as a Service).The Software Architecture that we adopted permits us to plug new features into the system, and to integrate GeoInt with other components (D'Amore et al., 2012).

**Results and Discussion**

Fig. 1 shows the schema we designed to improve GeoInt as a Geo Service Provider.

GeoInt was introduced as a middleware in order to match the requirements of the GIIDA project (Nativi, 2009; D'Amore et al., 2012). In the first version, GeoInt interacts with Geonetwork in order to provide users with a metadata system and a C-SW 2.0.2catalog. Unfortunately, this approach made metadata creation and management unfriendly for its users.

Our solution was to try to minimize the numbers of components required by the SDI, using GeoInt for both data and metadata. Users can upload data and metadata in GeoInt by means of services provided by the middleware and its Web User Interface. GeoInt permits its users to upload many data formats as CSV (Plain text) or in Shapefile. For any dataset uploaded, it is possible to upload the related metadata, across different formats. In this way users can use their own metadata editor to create the requisite metadata. In addition, for particular projects managed by GeoInt, the system can also provide an application form to complete this task. In both cases, GeoInt stores data and metadata, exposing the metadata trough a REST link. GeoInt also manages GeoServer Instances, an open source map server capable of exporting OWS (OGC Web Service). Using Geoint, users can choose which Geoserver instances to use to export data as a OWS service. OWS services permit

users to include a link for Metadata (where by GeoInt will include a REST link, created before the metadata, in each OWS service. Gi-cat, as the service Broker, will collect any OWS, using the metadata link in order to create a set of information about the data. The metadata collected will then be used by Gi-Cat in order to create a Web Geo Portal, or to include information in a complex System of Systems, such as GEOSS. Gi-Cat is the cornerstone of the metadata creation process, as it can collect metadata information from many types of data formats and export it through different data protocols.

The design shown in Fig. 1 describes a prototype, and to-date only some of its features are implemented. In creating this tool, we aim to create a complex middleware that is capable of minimizing the efforts of its users. To do this, we needed a workflow system (as showed in Fig. 1), where various components interoperate together.

## Conclusion

This paper focused on improving GeoInt, a middleware tool used by the CNR-IIA to process spatial data and metadata. Currently, developers are working to complete the schema presented in this paper, and to define other services to integrate into the tool. Additional areas of focus include improving GeoInt's support for raster data.

The successful integration of GeoInt, Geoserver and Gi-cat gives users an effective tool for managing geospatial data, from its creation to its final user exploitation.

## References

D'Amore F., Cinnirella S., Pirrone N., 2011. Architecture solution for air quality management through a dedicated SDI. Geophysical Research Abstracts, 13, EGU2011-7113.

D'Amore F., Cinnirella S., Pirrone N., 2012. Ict methodologies and spatial data infrastructure for air quality information management. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (in press).

Nativi, S.; Bigagli, L.; Mazzetti, P.; Boldrini, E.; Papeschi, F.; 2009. GI-Cat: A Mediation Solution for Building a Clearinghouse Catalog Service. International Conference on Advanced Geographic Information Systems & Web Services. GEOWS '09, : 68-74.

Nativi, S.; 2009. The GIIDA (Management of the CNR Environmental Data for Interoperability) project. EGU General Assembly 2009, held 19-24 April, 2009 in Vienna, Austria http://meetings. copernicus.org/egu2009, p. 3425.