UNIVERSITÀ DELLA CALABRIA

Dipartimento di ELETTRONICA, INFORMATICA E SISTEMISTICA

UNIVERSITÀ DELLA CALABRIA

Dipartimento di Elettronica,
Informatica e Sistemistica

Dottorato di Ricerca in
Ingegneria dei Sistemi e Informatica
XXIVciclo

Tesi di Dottorato

A Spatial Data Infrastructure for Geo Data Management

Francesco D'Amore

DEIS- DIPARTIMENTO DI ELETTRONICA, INFORMATICA E SISTEMISTICA
Novembre

Settore Scientifico Disciplinare: ING-INF/05

Università della Calabria

Dipartimento di Elettronica,
Informatica e Sistemistica

Dottorato di Ricerca in
Ingegneria dei Sistemi e Informatica
XXIV ciclo

Tesi di Dottorato

A Spatial Data Infrastructure for Geo Data Management

Francesco D'Amore

Coordinatore
Prof. Luigi Palopoli

Supervisore
Prof. Domenico Talia

Dott. Sergio Cinnirella

DEIS

DEIS- Dipartimento di Elettronica, Informatica e Sistemistica

Novembre
Settore Scientifico Disciplinare: ING-INF/05

A mia madre

Preface

The research presented in this dissertation was made possible by the Institute of Atmospheric Pollution Research - Division of Rende, of the Italian National Research Council (IIA-CNR). Additional scientific teachings during the first period of my PhD studies were provided by the Department of Informatics and System (DEIS), University of Calabria, Rende. I am sincerely grateful for the support provided by the following various groups throughout the time I spent as a PhD student.

Particularly, I would like to acknowledge my supervisor and chair, Prof. Domenico Talia, that has given to me the opportunity to "touch" the fascinating world of scientific research and that I am honoured to have as a guide.

I would like to thank especially my outstanding mentor, Dr. Sergio Cinnirella, for the insightful comments and thoughtful conversations that helped to enhance the quality of this research. His guidance supported me in all the time of research and writing of this thesis.

I am also grateful to my colleagues for their friendship and for the valuable help in all aspect of my work.

In particular, I want to thank Mariantonia Bencardino, Laura Fantozzi, Giovanni Manca and Lynne Gratz for their support over the years in the office at CNR-IIA.

Lastly, I would like to thank my family for all their love and encouragement. For my parents who raised me with a love of science and supported me in all my pursuits.

Francesco D'Amore
Rende,

Contents

Part I A Spatial Data Infrastructure and Geo Data Management

Part II The Global Mercury Observation System

Part I

A Spatial Data Infrastructure and Geo Data Management

1

An Air Quality Spatial Data Infrastructure

1.1 Introduction

Health diseases caused by environmental pollution are a growing concern worldwide, which in recent years addressed more studies on causal links between environmental pollution and health at regional and global levels [1] [2]. A greater understanding is thus needed of the links between exposure to pollution and its effect on health, as well as the long term impact on health of chemical substances, biological organisms and physical changes in the environment. Monitoring and modelling are, therefore, a crucial activity for identifying the key pressures on the environment, the condition or state of the environment, and the level of environmental quality being achieved by society [3]. Monitoring and modelling are inevitably challenging, not only from a technical point of view but also due to the complexity of the problems being addressed, originating in the interaction of multiple parameters at various levels of organization (anthropogenic or biological, individual or population) and scale (from global to local). Using the information collected through monitoring provides big challenges in integrating and connecting the various information sources used and the technologies implemented. A number of initiatives already exist to link different types of environmental data for various purposes, including efforts from the EEA (EIONET), the WMO (IGACO), the OECD (ECOSANTE), UNECE (EMEP), AC (AMAP), the EU and ESA (GMES) and the EU information system (INSPIRE). These monitoring actions involve several monitoring networks, a wide range of observational platforms and the use of techniques for data harmonization (also called data assimilation), interconnectivity and linkage. Integrated health assessment is not the main purpose of these monitoring networks, nevertheless it requires both routinely available and by-project collected dataset, combined in multiple ways.

Data and model simulations are, therefore, crucial to support policy makers and public participation within any environmental decision making process as well as for a broad understanding of the environment. However, these data

are not always available to the public and are not usually in a format that is understood by all the different stakeholders [4] [5].
Also, monitoring systems show relevant discrepancies in terms of spatial and temporal trends, as they do not cover appropriate regions, are discontinue along years and are often application-oriented. To fulfil the gap Spatial Data Infrastructures (SDIs) have been developed. An SDI is a framework of policies, institutional arrangements, technologies, data and people which enables the sharing and effective usage of geographic information [6]. For example, the Canadian Geospatial Data Infrastructure (CGDI) was implemented as an easy-to-use, advanced, online information resource for offering valuable benefits to decision makers in four priority areas [7], which are:

- public safety: to share maps of roads, bridges, electrical grids, water systems, buildings, and the like, to better plan for and respond to emergencies and disasters;
- public health community: to share location-based information securely to track pandemics, analyse trends, and monitor population health;
- local population community: to connect people and communities, map the future, and realize opportunities; and
- environment and sustainable development: to better manage land and water, assess the environment, and monitor ecosystems.

In order to overcoming inconsistencies in spatial data collection, lacking of documentation, incompatibility of spatial data sets, barriers to share and re-use of existing spatial data and incompatibility of geographic information initiatives, the EU adopted the INSPIRE Directive, which was developed to enable effective sharing of geographical data locally, and across borders [8]. The overall objective of INSPIRE is to make harmonized and quality spatial information readily available to support environmental policies and policies or activities which may have a direct or indirect impact on the environment in Europe. The INSPIRE proposal lays down general rules for the establishment of an SDI in Europe based on infrastructures for spatial information established and operated by the Member States. The component elements of those infrastructures include: metadata; key spatial data themes and spatial data services; network services and technologies; agreements on sharing, access and use; co-ordination and monitoring mechanisms; process and procedures.

In addition, the Group on Earth Observations(GEO) launched a program to coordinate efforts to build a Global Earth Observation System of Systems (GEOSS) aimed to obtain benefit in a broad range of societal benefits ranging from human health to biodiversity loss throughout resource's management, climate prediction, weather forecast and ecosystems protection. GEO provided a web-based interface (GEO Portal) for searching and accessing the data, information, imagery, services and applications available through the Global Earth Observation System of Systems (GEOSS) [10].

At Italian National level, some regions have built SDI projects, with different platforms, different costs and different data (e.g. Lombardia, Sardegna, Piemonte) or are developing projects concerning spatial data and integrated information services (Abruzzo, Puglia, Calabria, Sicilia, Molise, Sardegna, Campania). Also the Department of Civil Protection and the Department of Environment of the Italian Environment Ministry have portals respectively oriented to risk prediction/mitigation and collection of a vast Remote Sensing Plan for multi spectral data. Also the Comando Carabinieri per la Tutela dellAmbiente (Environmental Protection Command of Carabinieri Corp) has developed an important project for the control of crimes against the environment, with a wide collection of spatial data, including many iperspectral coverage as well as the Department of Demanio Marittimo has rebuilt and updated cadastral maps of all the Italian coasts. Spatial datasets are, therefore, a basic part of the information systems of the public administrations and the aforementioned projects are connected with the mission of the public administration, but they are hardly ever aimed at achieving an SDI. Most of the projects use standards for interoperable data and network services and often have a web geo-portal, but there is no uniform SDI that covers all of Italy [11].

In this work we present methodologies used in two projects oriented to geo data management in environmental fields: GIIDA and GMOS.

To coordinate national earth and cross-disciplinary systems for promoting GEOSS and to support INSPIRE implementation, the Italian National Research Council (CNR) promoted the **GIIDA** project (Integrated and Interoperable Management of Environmental Data) [12]. GIIDA aims to "implement the Spatial Information Infrastructure (SII) of CNR for Environmental and Earth Observation data". It was also aimed to design and develop a multidisciplinary cyber infrastructure for the management, processing and evaluation of Earth and environmental data. This infrastructure will contribute to the Italian presence in international projects and initiatives, such as: INSPIRE, GMES, GEOSS and SEIS.

GIIDA was divided in seven main thematic areas/domains: Biodiversity, Climate Changes, Air Quality, Soil and Water Quality, Risks, Infrastructures for Research and Public Administrations, Sea and Marine resources following the main research areas of the CNR. CNR-Institute of Atmospheric Pollution Research (CNR-IIA) leads the Working Group on Air Quality, which developed i) a specific Web Portal; ii) a thematic catalog service; iii) a thematic thesaurus service; iv) a thematic Wiki; v) standard access and view services for thematic resources such as: datasets, models, and processing services; vi) a couple of significant use scenarios to be demonstrated.

The **Global Mercury Observation System** (GMOS) project was established in support of the Group on Earth Observations (GEO) 2012-2015 work plan and the task HE-02-C1,"Global Mercury Observation System". This task supports the achievement of the goals of GEOSS and other ongoing international programs, including the UNEP Mercury Program, as well as major international conventions, such as the UNECE-LRTAP TF HTAP. The overall goal of GMOS is to develop a coordinated global observation system for mercury, including ground-based stations at high altitude and sea level locations, ad-hoc oceanographic cruises over the Pacific, the Atlantic and the Mediterranean, and free tropospheric mercury measurements. This will provide high quality data for the validation and application of regional and global scale atmospheric models, to provide concrete findings that can be used to support future policy development and implementation.

In the first part, we present Software Architectures and methods for geo data management. Next, we present GMOS project as use case of such methodologies.

1.2 Air Quality and Spatial data Infrastructure

1.2.1 Introduction to Air Quality Management

The goal of air quality management is to protect and enhance air quality for preserving human health and ecosystems. To accomplish the goal numerous regulations and standards, a broad suite of management tools, and several monitoring networks to track progress have been established (e.g. AERONET, EMEP, GAW). All of these components depend on robust and up-to-date scientific and technical input, which includes an understanding of relationships between air pollutant levels and impacts on human health, ecosystems, atmosphere composition and materials. At this level a SDI is fundamental as repository but also as interface between science and management as well as policy.

1.2.2 Role of SDI in Air Quality Management

A SDI oriented to Air Quality Management is a framework that should collect air quality information at both dataset and metadata level, store them in a database and make them accessible through a web-interface. The idea is to connect heterogeneous databases at different locations and access them from a simple web interface such that they appear as one, virtual database for the end user. Through a data catalogue (metadata) the users requests can be directed towards the appropriate database server(s). The results of such a query can then be downloaded from a central server or emailed to the customer. Large collections of air quality observation and simulation data can be made available through the cooperation of as many data centres as possible, and the physical location of these datasets becomes irrelevant. The only real problem to be overcome is the different data policies, but this is changing owing to a European Directive (Directive 2003/4/EC) on public access to environmental information [9].

1.2.3 Metadata and Interoperability in Air Quality Management

Storing and organizing data needs to contextualize data themselves in order to give information on collection methodology,lineage, spatial and temporal domain, copyright, context of use, etc. Such documentation is called metadata. Metadata helps organise and maintain data and provides information about an organisations data holdings in catalogue form; avoids duplication of effort by ensuring the organisation is aware of the existence of data sets; helps users to locate all available geospatial and associated data relevant to an area of interest; builds upon and enhances the data management procedures of the geospatial community; promotes the availability of geospatial data beyond the traditional geospatial community; advertise and promote the availability of their data and potentially link to on line services [6].
Though air quality data are important inputs to simulation models and decision-making systems, they are highly distributed and heterogeneous, and thus difficult to access in a coordinated manner. The first step to make them interoperable is to produce a spatial and temporal metadata registry that enables the science community to more easily use and exchange data and services. Metadata are very important for interoperability and systems integration, because with metadata is possible to define a common standard between systems. Use metadata is the main step for building complex geo-system where interoperability is a common target. The second step is to produce data in common formats following internationally accepted content standards by government agencies, academic institutions, and commercial companies around the world. The central management and storage of air quality data and metadata enable the interoperable data access at all levels of this framework by integrating in a central database data retrieved in most common format, including netCDF, ESRI©shapefile and Microsoft©Excel.

1.2.4 Users of SDI in Air Quality Management

From the perspective of an End User, a SDI should organize information and resources and distributes them, providing their services via a single access point, often online. The resources managed by a SDI should be Geo Portals or catalogs of Geographic Web Services oriented to users. In both cases, through the access points to the SDI, the end user should find the resources searching among metadata.
Users accessing an air quality SDI can be a human operator or a Service Client, often a Web Service Client. In the first case, the SDI is used by operators that can be data readers or users that create data and metadata. Among them, the decision makers in public administration and researchers interested in the Air

Quality are fundamental. In order to match the requirements for this kind of users is important create tools to search geographic data, necessarily based on metadata, allowing the detection of geographical data based on keywords and the geographic location. The data visualization during the search phase has no a fundamental impact on the user (reference), but a minimal visualization tool for performing research based on geographic localization is needed.
After this phase, visualization is a critical requirements in Air Quality management, and the visualized products should be different as result of a different use of data. Resulting metadata must indicates access points related to data use.
Data producer has different needs from data reader, as he should create data and metadata. It is important to use a unique standard and a single access point to metadata, in order to manage standard and access policy. Often the access point for metadata is a Web Application, with metadata editor and an end point for metadata catalog service.
Data and metadata entered must be reached even by non-human clients, typically Web Service Clients. This is an important requirement to ensure the scalability of the system towards systems of systems such as GEOSS or GIIDA.
In the first part of this work will be presented methodologies used in order to develop frameworks and process for Spatial Data Infrastructure. In the second part such methodologies will be used in order to build a Data Sharing Framework for the Global Mercury Observation System (GMOS) network. GMOS is an Observing System aimed to collect and share data about mercury pollution. Data will be collected by different provider by means of sharing architectures oriented to geo data management.

1.3 Development of Spatial Data Infrastucture for Air Quality Theme under GIIDA

1.3.1 Data Analysis and state of the art

Data and state of the art analysis is crucial to understand what kind of SDI to create. In our case, data are related to pollutants holding geographical information (i.e. latitude, longitude, elevation, pollutant concentration or emission). Data are gathered from monitoring systems and numerical simulations running on a parallel computing cluster. Simulations outputs are often in netCDF format, which is a binary format used for its flexibility and widely used as a container for scientific information (reference). Both measurements and simulations must be integrated into the SDI by considering work methodologies. Some data are or have been collected in geographical containers (e.g. ESRI©shapefile or NetCDF) and come from previous studies that have produced geographic data. They represent an essential background for the startup the Air Quality SDI.

1.3.2 SDI Architecture and used technologies

The Architecture of the Air Quality SDI aims to provide geographic services for integration into a Service Oriented Architecture (SOA) as GIIDA and GEOSS. SOA is the main orientation of further development within the environmental sector. The core of the system is represented by the PostGIS geographic database, which holds vector information (Figure 1.1). PostGIS, an extension of PostgreSQL DBMS, is a de facto standard in the panorama of open source software for storage of vectorial geographic data, but its current limit concern raster data storage in comparison with not-free competitors (e.g. Oracle).

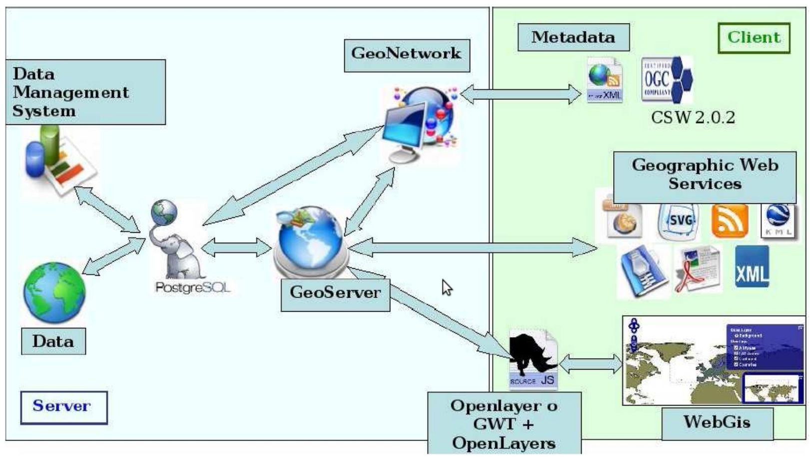

Fig. 1.1. The Architecture of the Air Quality Spatial Data Infrastructure

Data are stored into the database with a dedicated tool, the Data Management System (DMS), developed at CNR-IIA. The DMS receives as input the results of numerical models or measurement data, reads the content and persists them in the central database. Data input is in the format of XML meta-language, netCDF or shapefile. The DMS was developed with a modular architecture that allows to change the functionality to meet futures requirements. The goal is to use the DMS as an information platform for the management of vector data coming from measurements and numerical models.
As shown in Figure 1.2, the DMS enables the persistence of vector data using Hibernate as Object relational mapping (ORM) directly connected to the PostGis database. The system has a Web interface that allows interaction with the end-user. We are studying the way to integrate the DMS with used models (e.g. Weather Research and Forecasting Model, WRF) in order to minimize human interactions and enhance integration and interoperability.

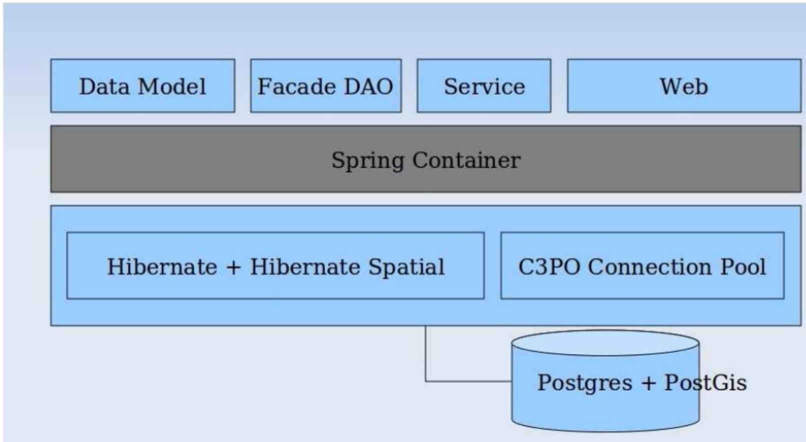

Fig. 1.2. The Data Management System (DMS) Architecture, developed at CNR-IIA

ESRI©shapefiles are de facto a standard for spatial data and represent the product of previous works but they are often distributed in several personal computers making data redundancy and not interoperable. As consequence, these vector files have been converted and stored in the database as our goal is to store all geographic data in the central DBMS. A problem remains for raster data that currently are not supported by PostGis. In the SDI they are handled directly by Geoserver and stored in the File System. Nevertheless, we are studying ways to store raster data on the PostGIS database by using mechanisms similar to those used for vectorial data (WKTRaster).
Geoserver is a map server that exports the data by creating Geographic Web Services, complaint with OGC (WMS, WFS, WCS).
These services can be used directly by end-users in complex systems SOA compliant, or geoportals built with Web technologies. The OGC services are also used by GeoNetwork for integrate metadata with maps and other geographic data. GeoNetwork is a tool used by Air Quality SDI for the creation and management of metadata based on INSPIRE directive (http://cartoserver.iia.cnr.it/). The metadata is stored in central database, as illustrated in Figure 1.2, which thus represents the central unit of storage both for data and metadata. The metadata managed by GeoNetwork are exported via the CSW 2.0.2. This service is the basis for integration of Air Quality SDI in complex systems such as GIIDA, as explained in 1.3.5.

1.3.3 Metadata Editor: Standard and Metodologies

In the Air Quality SDI the Geonetwork tool is successfully employed to create and manage metadata for geographic data. This tool allows users to create metadata compliant with European and international standard, such as ISO

19139 and INSPIRE Directive. The tool allows validation and check of created metadata and links with Geographical Services such as WMS, correlation with differents, and even not structured, data sources. Metadata managed by Geonetowrk are stored in a Postgres DBMS as showed in 1.1.
To improve interoperability, Geonetowrk exports metadata with the protocol CS-W. This feature permits the creation of complex Systems of Systems by federating different nodes. In our Air Quality SDI this feature is exploited by Gi-cat in order to integrate our SDI with GIIDA as explained in 1.3.5.

1.3.4 Metadata Editor: Standard and Metodologies

In Air Quality SDIs, are important three types of clients:

1. Web based clients, oriented to specific products and projects related to Air Quality theme;
2. DektopGIS oriented to researcher and operators in data elaboration;
3. Web Services, compliant with standard OGC and designed for system integration and interoperability.

The DesktopGIS systems can access information in two ways: by connecting directly to the geographic database or at Geoserver services, using WMS. Connecting directly to the database, the user of DesktopGIS that has the credentials to do so, can modify and work on the data. These changes will be made available directly to all services that use the data, given the centrality of the database in our architecture.

The web-based systems are used by end-users by means of geoportals, in which geographic data are used to create high level services and specific projects and are related to particular areas of interest such as monitoring the ozone precursors emission from facilities, or simply view a list of information layers contained in the Air Quality SDI. The web based systems are therefore the ideal tool to realize information systems that reach a wide range of users.

The technologies used for the realization of these systems are based mostly on javascript frameworks as Openlayers, which allow to develop GIS applications with AJAX support. Openlayers is a very effective tool that allows a connecttion directly to a WMS service in order to develop WebGis effective and easy to use. These methodologies currently represent the most used technique in the area of opesource WebGis products and offer excellent results. However, new methodologies for creating Web- based applications are available in the marketplace that allow to overcome the development approach based on Javascript, as building, reusing, and maintaining large JavaScript code bases and AJAX components can be difficult and fragile. For our purposes, GWT a technology integrated with Openlayers, which directly uses

Java or similar Software Engineering techniques, was adopted.

1.3.5 Integration of Air Quality SDI in GIIDA

Interoperability and systems integration has become important in geographic systems and SOA architectures become popular even in geo-systems area. The Air Quality SDI was created having in mind the more high level system, GIIDA. The integration of Air Quality SDI with GIIDA occurs primarily through the use of OGC CSW protocol that supports the ability to publish and search collections of metadata about geospatial data, services and related resources. It is therefore an integration based on common use of metadata standards according to the INSPIRE directive. The Air Quality SDI exports metadata entered through Geonetwork, using the protocol CSW. The end point for CSW is the catalog Gi-cat that collects data from Geonetowork. The information carried from CSW are linked to GIIDA which can acts as a collector of geographic information. Similarly, other participants to GIIDA can incorporate information coming from their SDIs using CSW protocol, creating thus a complex system of systems that offers integrated services to users.

2

ICT Methodologies and Spatial Data Infrastructure

2.1 The SDIs role in Air Quality information Management

The goal of air quality information management is to protect and enhance air quality to preserve human health and ecosystems [15]. To accomplish the goal, several regulations and standards have been established, and these have utilized a broad suite of management tools and monitoring networks to track progress (e.g. AERONET, EMEP, GAW, GMOS). All of these components depend on robust and up-to-date scientific and technical inputs, which include the most advanced knowledge of relationships between air pollutant levels and their impacts on human health, ecosystems, atmosphere composition, and cultural heritage sites. At this level an SDI is not only a repository, but an intermediary mechanism that provides raw data for science, management, and policy.

An SDI oriented to air quality information management is a framework that should collect air quality information at both dataset and metadata levels, store it in a database, and make it accessible through a web-interface. The SDI can also integrate a component that connects external and heterogeneous databases, enabling access to a single user-friendly web interface that makes both appear as one, virtual database to the end-user. Therefore, external databases can be exported as a single data warehouse by using a data catalog component, making these databases available within the SDI. By incorporating this component, user's requests can be forwarded to appropriate dataset(s). Results of such a query can then be downloaded from a central server or emailed to the customer. Large collections of air quality observation and simulation data can be made available through the cooperation of as many data centres as required, and the physical location of these datasets becomes irrelevant. The only real problem to be overcome is the differing data policies established by data owners, but even this problem is changing as new approaches to public access regarded environmental information are adopted (e.g. the European Directive 2003/4/EC) [9].

2.1.1 ICT methodologies to construct the SDI

In an SDI, additional ICT tools and frameworks are more often required. These tools manage Web Sensors, Data Notification, Processing, and complex data Visualization. Furthermore, an SDI may manage a huge amount of data and processes like atmospheric chemical models that are coupled with meteorological models, for which the integration of a complex computational framework is required.
The importance of integration between different tools is also supported by the Open Geographic Consortium (OGC), which released standards like Sensor Web Enablement [13] or Web Processing Services [14] in order to integrate results coming from different processes.
The integration of such technologies requires a different approach in developing an SDI, for example by using ICT methodologies often adopted in e-business. This flexible Information Infrastructure solves integration issues among SDI components by hiding the complexity of geographic systems to final users [16], who are often scientists or decision makers but often not IT experts. Moreover, they cannot always rely on having access to the full-time technical staff that is administrating these complex IT infrastructures, and hence, using an SDI allows them to remain focused on their area of expertise.

2.1.2 Type of data to be handled in the SDI

The SDI can handle information on air quality that is diverse yet unique to any particular project. In this case, the included data ranged from concentration of contaminants emitted into the atmosphere (or released to water and/or soil), to concentration of contaminants measured at permanent sites (as well as along dedicated monitoring campaigns). These meteorological parameters are the input for transport models aimed to draw chemical weather.
Outputs from models can be included in the SDI, and hence they can be used by policymakers to evaluate mid and long-term scenarios.

2.2 The SDI Architecture and Implementation

With the above parameters in mind, we developed and implemented an SDI as the cornerstone of a project with an integrated architecture, so that additional components could be plugged in as necessary [21]. Open-source components were used in building the Air Quality SDI. Postgis, Geoserver and Geonetwork were utilized for geographic data storage, export geographic web services, and managing metadata, respectively, whereas Javascript libraries embedded in OpenLayers [26] were used to display geographic information. These tools effectively made the SDI a pluggable system (a system built through components plugged together), and required additional effort to integrate the different components. To this end, both an effective client application and a

flexible software middleware were fundamental in crafting a common interface for the components involved in the SDI. Services and processes provided are now controlled through an Information Infrastructure (GeoInt), designed and developed at CNR-IIA, which wraps most useful processes in order to provide high-level services (e.g. data integration and management) to final users. GeoInt, described in section 2.3, acts as middleware between users and the SDI in order to provide a more friendly interface to the data and processes managed by the Air Quality SDI.
The architecture of the Air Quality SDI was designed to provide geographic services for integration into a Service Oriented Architecture (SOA) like GI-IDA and GEOSS. SOA is now the primary orientation for development within the environmental sector.
Hereafter, the SDI Architecture is described through different viewpoints by using the terminology of the Reference Model of Open Distributed Processing (RM-ODP) (ITU-T Rec. X.901-X.904 — ISO/IEC 10746) [16] [27]. RM-ODP is a model used to describe complex ICT systems. The following RM-ODP viewpoints have been considered:

- Enterprise
- Information
- Computational
- Engineering
- Technology

2.2.1 Enterprise Viewpoint

The enterprise viewpoint focuses on the purpose, scope, and policies of a system. It provides the context and the overall environment within which the system will be built, and therefore indicates constraints and obligations that must be applied to all other viewpoints, therefore representing the global requirements that the SDI must respect. Through the Enterprise Viewpoint, the SDI quality is described from a stakeholder's point of view.

2.2.1.1 Actor and Users

From the perspective of an end-user, an SDI should both organize information and resources and distribute them by providing their services via a single access point, often on line. Resources managed by an SDI should be geoportals or catalogs of geographic Web Services oriented to users. In both cases the end-user should find resources through the access point by searching metadata.
Users accessing an SDI can be human operators or they can be Service Clients, namely Web Service Clients. With human operators, the SDI is used by data readers or data and metadata creators. In order to match requirements for this user category it is important to create tools to search geographic data

(necessarily based on metadata), allowing for detection based on keywords and geographic location.
In summary, three types of clients have been considered for the Air Quality SDI:

- Web based, oriented to specific products and projects related to the air quality theme or dedicated to decision makers in public organizations.
- Desktop GIS, oriented to researchers, operators and data providers.
- Web Service, compliant with a standard and designed for system integration and interoperability.

2.2.1.2 Functional requirements

In air quality information management data are the cornerstone for simulation models and decision-making systems. They are both widely distributed and heterogeneous, and so they are difficult to access in a coordinated way.
The first step in making systems truly interoperable is to adopt agreed standards for data storage, searching and distribution. Metadata will also enable the community to get information on the data within the SDI.
The second step is to produce data in a commonly recognized format. Hence, data storage can be virtual or even on a physical device, but it must be re-engineered and harmonized in order for more services to be produced from it. In many organizations, environmental information is based on proprietary formats and stored in local computers without any kind of metadata. This information must be harmonized in order to foster multidisciplinary approaches and advanced analysis. An essential requirement of any SDI is a database that holds harmonized data and related metadata.
From the users point of view, creating metadata can be a tedious task, and the management of large XML files with verbose descriptions may discourage some operators. Organizations that manage large datasets need some tools that assist with the creation of metadata and its storage in a database. An additional point that discourages operators is the necessity of following standards that are often updated or changed.
Metadata stored in the database will be used to search data. After this initial phase, visualization is the next critical requirement and visualized products should differ as a result of a different use of data. Here, metadata must indicate the access point related to data use. For example, web-based systems are browsed by end-users by means of geo-portals, in which geographic data are used to create high-level services in specific projects. These geo-portals are related to particular areas of interest, such as the monitoring of ozone precursor emissions from facilities, or simply viewing a list of information layers contained in the SDI. Therefore, web-based systems are the preferred tools for realizing information systems that reach a wide range of users. Users need to interact with an SDI in all phases of data management: data and metadata creation, process management, data access, data searching, and, finally, visualization. To these ends, many tools have been developed that show data in

a user-friendly way (e.g. Openlayers). Still, an additional Information Infrastructure that hides the complexity of the SDI and helps the data provider can be useful. Middleware, like GeoInt, help users add data and metadata without any direct interaction with the complex interface of the SDI.

2.2.1.3 Non-Functional requirements

In order to match new requirements that arise from new and changing scenarios in e-Science and geomatic (Sensor integration, 3D supports), the SDI has crafted a modular approach, where each component implements a single feature and then talks with the others in order to perform the required services. With this approach a change in a functional component has a minimal impact on the system as a whole, because components are only loosely coupled. This architectural design allows for an easier system upgrade, and it adds scalability as new modules can be added to balance the load.

2.2.1.4 Integration Constraints

Interoperability and systems integration have become important in geographic systems, and SOA architectures have become popular even in geo-system areas. The Air Quality SDI was created with the higher level GIIDA domain in mind.
The integration of the Air Quality SDI with GIIDA occurs primarily through the use of the OGC CS-W protocol, a protocol that allows for the ability to publish collections of metadata, geospatial data, and services and related resources. Thus, it is an integration based on the common use of metadata standards according to the INSPIRE directive. Interoperability enables data harmonization by linking more systems in a complex and robust distributed system, like in GIIDA or GEOSS.

2.2.2 Information Viewpoint

The information viewpoint is focused on the information semantics and information processing that is performed. It describes the information managed by the system and the structure and type of content of the supporting data. It then describes the system in terms of data managed.

2.2.2.1 Data analysis and state of the art

Data analysis and infrastructure state of the art are crucial in understanding which kind of SDI to create. In the case of the CNR-IIAs project, data are related to contaminants holding geographical information (i.e. latitude, longitude, elevation, depth, chemical concentration), which are gathered from monitoring systems and numerical simulations running on parallel computing

clusters or existing datasets. Measurements, simulations and datasets must be integrated into the SDI by considering working methodologies. For example, some data have been collected in geographical containers (e.g. ESRI Shapefile). Data coming for past work represent essential background information for the planning and implementation of the Air Quality SDI.
GeoInt wraps any data source and provides an interface for final users in order to upload data and metadata in SDI databases. Data Sources can be managed by user (data file) or automatically managed in the data acquisition processes. To this end, it is strategic to provide integration between hardware and software components within the SDI, to minimize human actions in the data acquisition. This kind of integration is possible only with a flexible software middleware, integrated directly into the SDI, where any source can be wrapped with a specific Software Layer.

2.2.2.2 Metadata: The knowledge description

Description of data stored in an SDI is as fundamental as the data themselves. Storing and organizing data requires it to be contextualized in order to give information on any collection methodology, lineage, spatial and temporal domain, copyright, context of use, etc. Such documentation is called metadata. Metadata helps to organize and maintain data in order to:

- provide information on an organizations data hold in a catalogue form;
- avoid duplication of efforts by ensuring the organization is aware of the existence of data sets;
- help users to locate all available geospatial and associated data relevant to an area of interest;
- build upon and enhances the data management procedures of the geospatial community;
- promote the availability of geospatial data beyond the traditional geospatial community;
- advertise and promote the availability of their data and potentially link them to on-line services [6].

2.2.2.3 Thesaurus Integration

The Air Quality SDI was constructed with a thesaurus developed at CNR-IIA [18]. Indexing and searching by thesaurus to provide access to geo-information resources is still a quality issue in the Internet world. Nothing can be more frustrating than searching by subject when no indexing strategy is present: each searching session has to face all the uncertainties of natural language (synonymy, polysemy, homonymy) combined with all the uncertainties of a full text search (no relevance control on the retrieved occurrences). Moreover, most people realize that an indexing strategy itself is not enough, especially when we have to face either broad classification systems or indexing by using natural language. With regard to broad classification systems, navigation in wide virtual 'containers' is often a time-wasting operation; whereas in natural language indexing synonyms, polysemes, or homonyms limit and/or delay efficient and effective information retrieval. These problems grow exponentially when large numbers of documents and multilingual contexts are involved. Adopting a thesaurus:

- guarantees the effective control of the indexing language, covering each selected concept with a preferred term (in each language, in the case of multilingual thesauri) and ensuring inter-language equivalence among these descriptors;
- provides a systematic display of the descriptors, making navigation through the terminology easier;
- allows indexing and searching by combining several descriptors ex post, in order to refine and personalize both the semantic description and the information retrieval.

In this way, adopting a thesaurus creates results that balance the number of retrieved documents with their relevance. This was the reason behind the implementation of EARTh [18], a thesaurus oriented towards environmental issues that included a specific section for GIS terminology.

2.2.3 Computational Viewpoint

The computational viewpoint enables distribution through functional system decomposition into objects that interact by interface. It describes the architecture of an SDI by means of its components. This SDI has a classical three-layer architecture (shown in Figure 2.1) with a Data Storage Layer (DSL) connected to a Business Logic Layer (BLL), hosted in a Tomcat Server that is linked to the Application Layer (AL).
The core of the system is represented by the DBMS, which holds vector information and functional data. The DBMS represents the DSL in the SDI architecture, and stores metadata and geographic data in separate databases to maintain a different logic structure for each type of data. Additional databases are dedicated to functional tasks and to information coming from web sensors. In detail, the following databases are contained in the DSL:

- g_container: used to store geographic information;
- g_metadata: used by Geonetwork for metadata stored in g_container;
- geoint: used by GeoInt to perform functional tasks;
- sos_database: which contains data coming from web sensors.

The latter was a database designed by 52North [23] in order to match Sensor Web Enablement (SWE) specifications, which is connected to 52north SOS components deployed in the BLL, and is used to export sensor information. In the BLL, server components are plugged in to perform the system

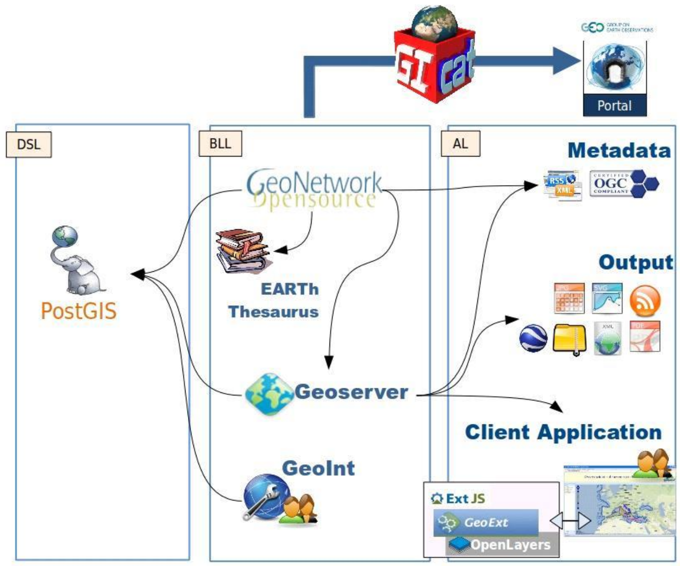

Fig. 2.1. SDI Architecture showing the Data Storage Layer (DSL), the Business Logic Layer (BLL) and the Application Layer

functionalities like metadata management, data and map creation, and data dissemination. The components related to this layer are:

- GeoInt: the data producer and Information Infrastructure, oriented to the Geo Infrastructure;
- Geoserver: the map server;
- Geonetwork: the metadata manager and catalog;
- EARTh: the Thesaurus used by the metadata editor.
- GiCat: the Service Broker. [20]

The AL represents a set of tools oriented towards final users. This container holds Desktop GISs, Web GISs, Metadata viewers and data downloaders. Each component is described in more detail in the Engineering Viewpoint, within section 2.2.4.

2.2.4 Engineering Viewpoint

The Engineering viewpoint focuses on interactions between the distributed components of the system. Each component installed in the BLL and AL is described hereafter. The communication channel between components is realized through web links and REpresentational State Transfer services (REST) [22]. Until now, only the Thesaurus has been physically installed in Geonetwork as the Simple Knowledge Organization System (SKOS) services has not

been implemented.
GeoInt (http://sdi.iia.cnr.it/geoint/) is the Information Infrastructure developed at CNR-IIA that wraps the SDI in order to provide facilities related to data and metadata creation and harmonization to the final-user. With GeoInt, users can work with a set of tools that make using the SDI components easier, particularly with regard to data and metadata creation, sensors integration and process management. This tool allows users to create metadata that is compliant with European and international standards, like those detailed in the ISO 19139 and the INSPIRE Directive.
Geoserver (http://sdi.iia.cnr.it/geoserver) is a map server that exports data by creating Geographic Web Services, compliant with OGC (WMS, WFS, WCS). These services can be used directly by end-users in complex SOA systems or geo-portals built with Web technologies. OGC services are also used by Geonetwork to integrate metadata with maps and other geographic data.
Geonetwork (http://sdi.iia.cnr.it/geonetwork) is a tool used to manage metadata. The metadata are exported via the CS-W 2.0.2 protocol, which is the basis for integrating the Air Quality SDI into complex systems. The tool allows links with Geographical Services such as WMS and correlation with different, even non-structured, data sources.
EARTh (http://ekolab.iia.cnr.it/earth.htm) is the Thesaurus linked within Geonetwork in order to support the metadata editing process.
GiCat (http://sdi.iia.cnr.it/gicat) [20] is a Service Broker used as a collector for Geographic Web Services. It supports several protocols like WMS, WFS, THREDDS, CS-W and accepts a wide variety of inputs, extracts information, and exports that information in a standardized protocol like CS-W. It is linked directly with Geonetwork to export metadata and to integrate the Air Quality SDI into a more complex System of Systems like those constructed within GIIDA or GEOSS. GiCat is a Service Broker developed at the ESSI-Lab (http://essi-lab.eu).
Some WebGISs (http://www.webgis.iia.cnr.it) were developed using services generated by Geoserver and are released to end-users by means of a geo-portal, where geographic data are used to create high-level services in specific projects related to particular areas of interest. The web-based systems are therefore ideal tools to realize information systems that reach a wide range of users.

2.2.5 Technology Viewpoint

The Technology Viewpoint is focused on the systems technology selection. It describes the technologies selected to provide processing, functionality and presentation of information. It examines products available in the marketplace that are used to build the SDI. The SDI is based on the adoption of international standards. These are mainly OGC for data representation, and ISO19115 and INSPIRE, for metadata creation. Moreover, standards like WMS 1.1.1 and WFS 1.1.0 were adopted for data, and CS-W 2.0.2 was adopted for metadata dissemination through Geonetwork and GiCAT.

As outlined in the ISO 19139 standard, metadata are presented in an XML document. ISO 19139 defines Geographic MetaData XML (gmd) encoding, an XML Schema implementation derived from ISO 19115. XML is a mark-up language, widely used in ICT systems, aimed to describe any type of information regardless of the means by which this information will be used or transmitted.

Geonetwork acts as a component plugged into the Air Quality SDI architecture, so that metadata can be managed at all levels (metadata creation, storage, and management) and exported through OGC CS-W, the standard protocol used to describe it.

Data served by the SDI are often used in WebGISs. In order to develop such applications, javascript frameworks were used.

In our case we adopted Openlayers, which allowed us to develop applications with AJAX support. Openlayers is a very effective tool that can establish a connection directly to WMS services in order to develop effective, easy-to-use applications. These methodologies currently represent the most-used technique in the area of open-source WebGIS products and they offer excellent results. However, new methodologies for creating Web-based applications are available in the marketplace that can overcome a development approach based on Javascript.

2.3 GeoInt

GeoInt is the Information Infrastructure that was developed at CNR-IIA to provide data creation and harmonization. The architecture in Figure 2.2 shows interactions between the database, OGC services exported via Geoserver, and final users. GeoInt is an Enterprise Application, developed with Java and deployed in a Serverlet Container (Tomcat). The scope of this framework provides an infrastructure that centralizes data persistence, processes, services creation, processes control over geographic information, and geographic data storage management. It provides an AJAX web user interface in order to access features provided by the framework. The web interface of GeoInt was developed with ZK [24] in order to improve user friendliness and usability. ZK enables a declarative approach for web application and it can be very useful in order to develop complex web systems oriented towards data visualization. ZK is a Web Framework that allows to develop an AJAX interface without the direct use of Javascript. Figure 2.4 shows a screenshot of GeoInt during the selection of a DataSource.

In the following section we describe each component used in building GeoInt.

2.3.1 GeoInt Data Storage Management

As we described above, geographic data are stored in a Postgis database (see section 2.2). This database is managed through GeoInt by means of Data

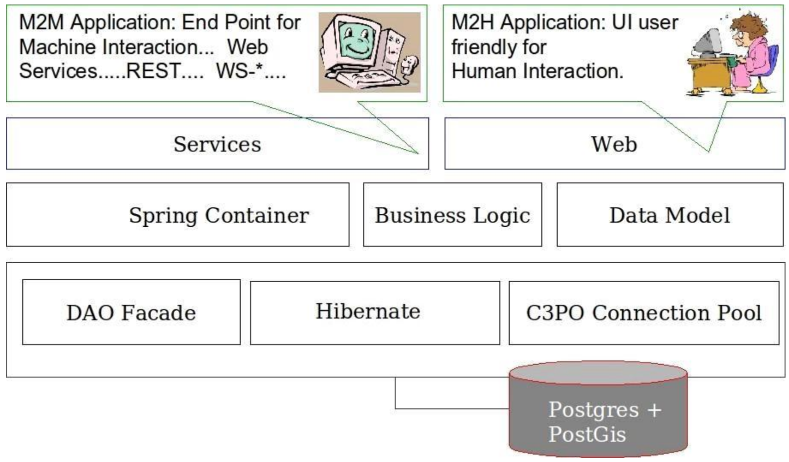

Fig. 2.2. GeoInt Components Overview

Access Objects (DAOs). DAOs are patterns used in Software Engineering to decouple storage systems from software layers. In our case, there are DAOs objects that manage any given instance of data storage.
The DAO interface designed in GeoInt is the GContainerDAO. Few methods of the DAO interface are listed in the UML Class Diagram in Figure 2.3
GContainerDAO exports the CRUD (Create, Retrieve, Upload, Delete) operation on geographic objects stored in a vectorial storage device. DAO for raster data are in the development phase. In the next version of GeoInt, DAO interfaces for raster will be tested and deployed. GeoInt, without a DAO for raster, cannot handle this data directly. To date, raster data are stored in the File System and handled by Geoserver.
In the case of vectorial data, Postgis is the *de facto* open-source standard for geographical DBMS.
By using DAOs in GeoInt, we allow for the possibility of a switch from Postgis to a different storage system. In GeoInt, each vectorial data storage system can be activated by the implementation of a corresponding GContainerDAO. A user/developer that wants to enable a different storage strategy must provide the specific implementation of GContainerDAO to obtain any selected strategy.
In addition, in GeoInt a specific DAO is used to perform user management and functional tasks, namely GeoIntDAO. GeoIntDAO defines a set of methods for user-notification, users roles, credentials on datasets and other functional tasks.
The design of the DAO system is described using UML (Figure 2.3). In each instance, the Postgis database, which is used and managed by GeoInt, is wrapped by an instance of PostgisContainerDAOImpl that implements the GContainerDAO in order to plug different Postgis Databases into GeoInt.

Developers with different user requirements will provide specific implementations of GContainerDAO, such as OracleContainerDAOImpl, MySqlContainerDAOImpl, and others, depending on the storage devices being used.
The GContainerDAO and GeoIntDAO are collected in a Façade. Often used in Software Engineering, The Façade pattern is a very useful strategy for offering a direct interface into a complex system. In this case, components in the second layer of Figure 2.2 can manage any object by means of the method provided by the Façade. This approach demonstrates a more robust Software Architecture and a more flexible use of the underlying resources.

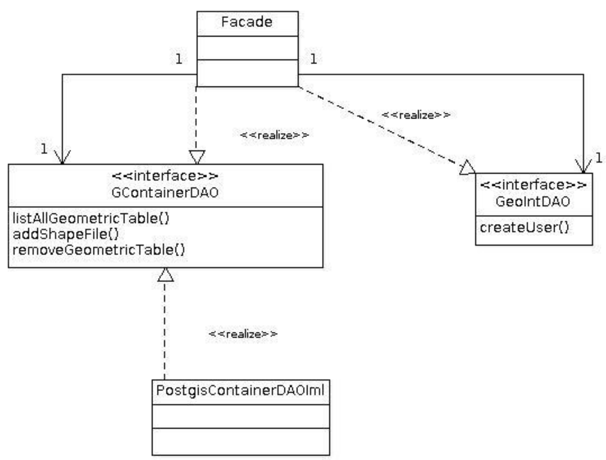

Fig. 2.3. Façade and DAO in GeoInt

For example, to switch from one Postgis database to another, a user must provide the Façade with the appropriate instance of PostgisContainerDAOImpl, even at run time. This is possible through the GeoInt web interface (Figure 2.4), once the current user has the appropriate credentials to use the selected database. The Business Logic Component and other Software Layers that use the Façade can use the method without gaining awareness of the type and location of the new data storage. Furthermore, the Façade is used by components hosted in the higher-level layer of the GeoInt architecture, in order to manage different data sources.

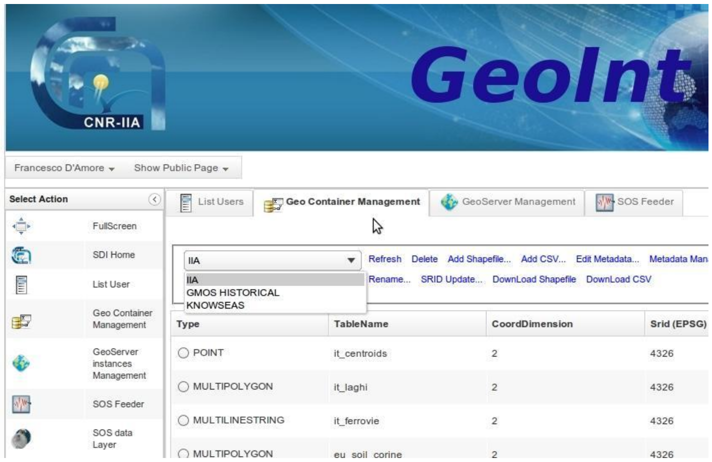

Fig. 2.4. GeoInt. Select a DataSource

Spring [23] is used as IoC framework in GeoInt. IoC is a pattern that enables a flexible setup of complex enterprise applications, which are used as Application Context in order to setup any component (and the relationship between components). Spring is used to manage not only DAO components, but also Connections Pools (C3PO), JDBC Data Sources, Logging components (Log4j), and services like mail systems and application utilities. All components are linked together by means of the Spring framework, at startup in GeoInt.

2.3.2 OGC Service Creation with GeoInt: Geoserver Integration

Figure 2.2 shows the Architecture wherein there is a layer that contains the business logic to perform data process and services. This layer, where appropriate, is stored and by the Façade. This BLL can also contain services used to integrate GeoInt with additional software components.
The communication component between GeoInt and Geoserver is fundamental to exporting data stored in the geographic databases through OGC Web Services (OWS). For this we used the RESTfull endpoint provided by Geoserver, which helps to manage the map server thereby making it possible to add a new WMS/WFS service programmatically by means of a RESTfull operation. A UML sequence diagram showing the process is reported in Figure 2.5. Through GeoInt, users can select a Geodatabase and finally a georesource. They can create a OWS Service by means of GeoInt, which will send a command to Geoserver and get a return end point of the service created. This end point is managed directly by GeoInt.
This simple scenario shows how GeoInt can hide the complexity of resource

management to the final users of an SDI, as GeoInt can manage the creation of a resource and the creation of Geo Services compliant with OGC standards through a single interface.

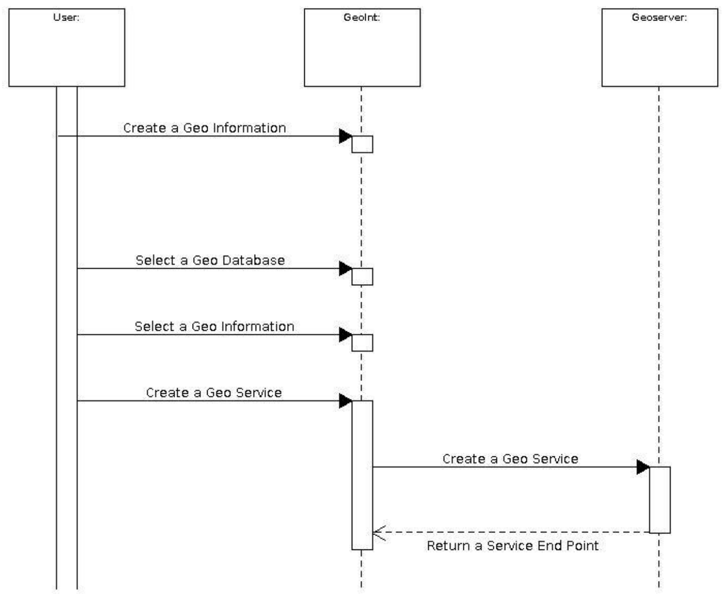

Fig. 2.5. Service Creation with GeoInt

2.4 System Integration

The Air Quality SDI is federated in the GIIDA domain that acts as System of Systems (SoS). The Air Quality SDI, like all GIIDA SDIs, exports information by means of CS-W protocol or other standard protocols by means of a brokering catalog, namely GiCat.

The central node of GIIDA acts as a Brokering Service by utilizing GiCat. Data coming from different partners are then collected and processed in the central node. This Brokering Service allows final users to exploit data and services provided by multiple nodes as if they were coming from a unique data source. Therefore, the top-level GIIDA domain offers integrated services to end users or to a more complex SoS (E.g. GEOSS), as if it were a collection of data owned by GIIDA partners.

Fig. 6 shows the integration schema adopted by the Air Quality SDI within the GIIDA framework. Concurrent partners of GIIDA can implement different SDI architectures, but all of them interact with the central node by means of well-known data protocols managed by a GiCat instance installed in the central node.

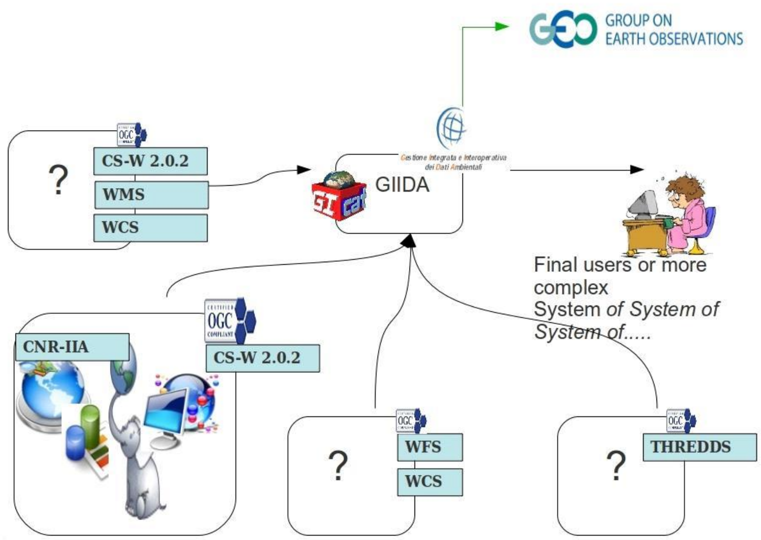

Fig. 2.6. Integration of the Air Quality SDI within GIIDA domain

This service interface standardization enables the decoupling of the SDI implementation from the external world: the local architecture can be modified in some component without affecting interaction with final users. To this end, final users do not see internal changes and their approach to the SDI remains unaffected in terms of complexity. This allows for a high level of flexibility, which is essential when the SDI must be integrated into the SoS.

2.5 Geo Services Factory

Fig. 1 shows the schema designed in order to improve GeoInt as Geo Service Provider. The idea is to minimize the numbers of component required by the SDI in order to match users requirements. Users can upload data and metadata in GeoInt by means of services provided by the middleware and its Web User Interface. GeoInt permits to upload many data format as CSV (Plain text) or Shapefile. For any dataset uploaded it is possibile to upload also a metadata related, using different formats. In this way users can use their own metadata editor in order to create a metadata. For particular projects managed by GeoInt, the system can provide also an application form in order to create metadata. In both case, GeoInt store data and metadata, exposing metadata trough a REST link. GeoInt manages also GeoServer Instances, an open source map server capable to export OWS (OGC Web Service). Users can decide, by means of GeoInt, the Geoserver instances to use in order to export data as OWS service. OWS permit to include a link for Metadata and GeoInt will include in each OWS service, the REST link created before for metadata. Gi-cat, as service Broker, will collect any OWS, using the metadata link in order to create a set of information about data. The metadata collected will use by Gi-Cat in order to create a Web Geo Portal or to include information in complex System of System, such as GEOSS. Gi-cat is core stone of metadata process as it can collect metadata information from many type of data format and export trough different data protocol.

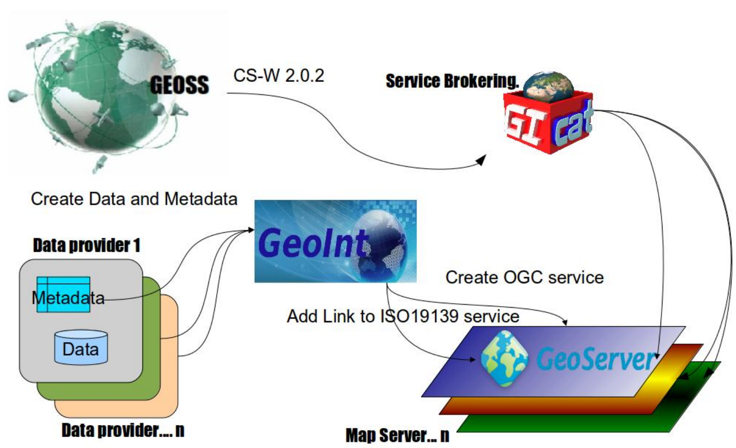

Fig. 2.7. The process schema that shows how data and metadata are uploaded in GeoInt, managed by Geoserver and Gi-cat and integrated in GEOSS

The design showed in Fig. 1 is a draft and currently only some of its features are implemented. We aim to create a complex middleware capable to minimize the efforts of users. In order to match this requirements, a workflows system as showed in Figure 2.7 is need, where few components interoperate toghether.

Part II

The Global Mercury Observation System

3

The GMOS Project

Mercury is emitted into the atmosphere from a variety of anthropogenic and natural sources. Of the anthropogenic sources, among the most important are fossil fuel combustion, smelting, cement production and waste incineration, while the oceans are the largest natural source of Hg to the atmosphere, volcanism also makes an important contribution [15]. Most mercury emissions are of elemental Hg but anthropogenic emissions also include Hg in different chemical and physical forms. Mercury cycling between different environmental compartments depends on the dynamic of chemical and physical processes (oxidation, dry deposition, wet scavenging) which influence its transport and residence time in the global environment. It has been suggested that due to intensified anthropogenic emission of mercury into the atmosphere since the beginning of industrialization this global Hg pool has increased in the past 150 years.

Evidence of long-term changes in the atmospheric mercury burden have been derived from chemical analysis of lake sediments, ice cores, peat deposits and firn air records [15]. These studies identify a peak in the atmospheric mercury concentration during the 70s in the Northern Hemisphere. A growing number of these records from both hemispheres demonstrate approximately a threefold increase of mercury deposition since pre-industrial times. In principle, an increase in the global atmospheric pool should also be reflected in the background concentration. However, since the first reliable measurement data were published just 30 years ago, it is extremely difficult to determine a global trend from the existing spatially and temporally inchoate air concentration data sets. For example, Asian mercury emissions are believed to have rapidly increased over the last decade, however, this is not reflected neither in the long-term measurement of Total Gaseous Mercury (TGM) at Mace Head, Ireland covering the period between 1996 to 2006, nor in the precipitation data of the North American Mercury Deposition Network (MDN).

Currently, there is no coordinated global observational network for mercury that could be used by the modelling community or to establish recommendations to protect human health and ecosystems on a global scale. Current

national/regional monitoring networks are inadequate: they lack, (1) observations of all forms of Hg in the ambient air and in both wet and dry deposition samples; (2) long-term measurements of Hg and other air pollutants; (3) comprehensive monitoring sites at altitudes which permit the sampling of free tropospheric air; and (4) measurement sites that allow a careful investigation of inter-hemispheric transport and trends of background concentrations.
Recognizing that TGM and Hg in wet deposition are spatially heterogeneous, several studies have aimed to set up monitoring networks in order to compare trends between sites in the same region, between regions, and to determine the influence of local and regional emission sources. There is also interest in understanding the processes that contribute to Hg variability on a diurnal, weekly, seasonal, and annual basis. In 1995, Fitzgerald argued for and defined the basic requirements of an Atmospheric Mercury Network (AMNet). This has partly been accomplished on a regional scale within the Canadian Atmospheric Mercury Network (CAMNet), which may be considered as seminal in this respect. Part of the CAMNet contributes also to the Arctic Monitoring and Assessment Program (AMAP) under the Ministry Council of Circum Polar Countries and where there is also contribution from the Scandinavian countries, e.g. on Greenland. Nevertheless, although the number of atmospheric Hg monitoring stations has increased, the database is sparse, especially in remote locations.
The **Global Mercury Observation System** (GMOS) project was established in support of the Group on Earth Observations (GEO) 2012-2015 work plan and the task HE-02-C1,"Global Mercury Observation System". This task supports the achievement of the goals of GEOSS and other on-going international programs, including the UNEP Mercury Program, as well as major international conventions, such as the UNECE-LRTAP TF HTAP. The overall goal of GMOS is to develop a coordinated global observation system for mercury, including ground-based stations at high altitude and sea level locations, ad-hoc oceanographic cruises over the Pacific, the Atlantic and the Mediterranean, and free tropospheric mercury measurements. This will provide high quality data for the validation and application of regional and global scale atmospheric models, to provide concrete findings that can be used to support future policy development and implementation.
The specific objectives of GMOS are:

1. To establish a global observation system for mercury able to provide ambient concentrations and deposition fluxes of mercury species around the world, by combining observations from permanent ground-based stations, and from oceanographic and tropospheric measurement campaigns.
2. To validate regional and global scale atmospheric mercury modelling systems able to predict the temporal variations and spatial distributions of ambient concentrations of atmospheric mercury, and Hg fluxes to and from terrestrial and aquatic receptors.

3. To evaluate and identify source-receptor relationships and their temporal trends for current and projected scenarios of mercury emissions from anthropogenic and natural sources.
4. To develop interoperable tools to allow the sharing of observational and models output data produced by GMOS, for the purposes of research and policy development and implementation as well as for enabling societal benefits of Earth observations, including advances in scientific understanding in the nine Societal Benefit Areas (SBA) established in GEOSS.

In order to achieve these goals, a coordinated international scientific effort is needed to enable users throughout the world to openly share and put to use vast quantities of global mercury data, thereby advancing scientific research in many disciplines, promoting technological and economical evaluation of abatement technologies implementation, facilitating impact assessment on environment and human health, establishing benefits in enhancing the quality of life of members of society. The importance of making mercury data openly available to all countries and individuals is underscored by various international agreements. Data-sharing in GMOS is an important way to increase the ability of researchers, scientists and policy-makers to analyse and translate data into meaningful assessments of mercury impact on environment and human health, reports and knowledge. Data sharing also discourages duplication of effort in data collection reducing use of resources in data collection; encourages diverse thinking, as other researchers are able to use the data to answer questions that the initial data collectors may not have considered; encourages accountability and transparency, enabling to validate one another's findings; and allows for comparisons that cross national and regional boundaries.
In the next sections will be described some components of the SDI developed for GMOS project and Air Quality Management [19].

3.1 GMOS Interoperable System

Within GMOS project will be developed tools (i.e. databases, catalogs, services) to collect GMOS datasets, harvest mercury databases, offer services like search, view and download spatial datasets from the GMOS portal (www.gmos.eu). The system will be developed under the framework of the Infrastructure for Spatial Information in the European Community (INSPIRE) Directive and the Directive 2003/4/EC on public access to environmental information, which aim to make available relevant, harmonised and high quality geographic information to support formulation, implementation, monitoring and evaluation of policies and activities which have a direct or indirect impact on the environment. It is proposed to establish three data bases (emissions, field data and model results) which will be equipped with state-of-the-art and opensource software to allow highest performances. Web-based user interfaces

and prototype applications will be developed to demonstrate the potential of blending different data sets from different servers for environmental assessment studies.
Several services (i.e. catalog browsers, WMS and WCS services, web GIS services) will be developed to facilitate data integration, data re-use, and data exchange within and beyond the GMOS project. Different types of measurement and model datasets provided by projects partners and from other sources will be integrated in PostgreSQL-PostGIS, harmonized by creating metadata INSPIRE compliant and make it available to a larger community of stakeholders, policy makers, scientists, NGOs as well as other public and private institutions, as dictated by the Directive 2003/4/EC. Since interoperability is a central concept of the Global Earth Observation System of Systems (GEOSS), the Global Monitoring for Environmental and Security (GMES) and the INSPIRE Directive, the guidelines developed in these frameworks will be adopted. The use of standards is a key concern along the encoding process. International standards for data and spatial schemas (ISO19107, ISO14825), for metadata (ISO19115:2003, ISO/DTS19139:2005, ISO15836) and for services (WMS 1.1.1, WFS 1.0, SLD 1.0, GML 3.1) will be used. On the other side for data exchange XML together with SOAP, XSD, J2EE (for applications development) and W3C for standard interfaces will be the reference standards.
With specific reference to GMES, the global database on mercury monitoring and the model output resulting from GMOS will be made available in terms of series of monitoring, forecast and reanalysis services. The GMOS operational services will hopefully contribute to the Monitoring Atmospheric Composition and Climate (MACC) project by providing access to atmospheric environmental services.
In order to define a framework for data sharing in GMOS, the Work package 9 (WP9) will be structured in five Tasks:

- Task 9.1: System architecture definition
  The goal of this task will be the definition of the overall system architecture. It will have an overview of existing architectural approaches in the framework of INSPIRE. The system architecture has to consider multi-level data communication and processing. The system architecture will integrate methods for data collection at the lower end, methods for data processing and data fusion on the middle layer, and methods for data presentation and portal solutions on the top layer. All layers will have distinct requirements regarding communication models, time constraints, data formats and sizes, and security and access control as well as open-source software selection.
- Task 9.2 Data Collection tools development
  The goal of this task is to develop the data collection tools and to set up of the communication and networking between the collection tools at the front end (monitoring sites and model outputs). This task has tight

relation to Task 9.3 which bridges the developed tools to the front end. This task is also related to Task 9.5, where all the elements are going to be integrated to one system. The full performance of the data collection tools will be tested.

- Task 9.3 Data fusion mediation system development
  In this task it is our objective to develop tools which will take all collected data from different data sources (Task 9.2) and harmonize them into one comprehensive and complex database. This database will be consequently used as input for application developed in Task 9.4.

- Task 9.4 Development of the GMOS portal (Web based Applications, web GIS services, WMS services, Decision Support System)
  The aim of this task is to develop the GMOS portal and a set of Web based applications, web GIS services, WMS and WFS services. This task has tight relation to WP 10. The system will be designed with security, redundancy and scalability in mind so that further functionalities can easily be included afterwards. This task is largely dependent on the design carried out during the definition phase along Task 9.1, as the developed applications will have to adapt to the architecture defined in aforementioned tasks. Along implementation this task will interact with Task 9.1 in a continuous basis especially in the database definition and access. Specifically, synthetic and clear data, as well as project results will be made freely available (open access) according the Clauses 29 and 39 of the Grant Agreement.
- Task 9.5 Harvesting of other data services
  This task will collect a list of related metadata, databases and services (the so called harvesting process) and connect them to the GMOS portal. As primary aspect it will look into opportunity to link GMOS portal to GEO Portal and as second aspect will look into opportunity to link the portal to the WMO information system and to the MACC project. The process will contribute to discussions on the standardisation of data formats and names related to atmospheric mercury models and observations.

Finally the WP-9 activity will ensure the adoption of the GEOSS data sharing principles and register the data component (SDI) in the GEOSS community.
This report on system architecture (Deliverable 9.1) gives an overview of how the GMOS Information and Communication Technology (ICT) system is structured.

4

Architecture definition and development of data exchange system

4.1 ICT techniques in the SDI design

In our view, the GMOS ICT system is designed as a complex middleware because user's requirements range from scientific purposes to policy, from simple assessments dedicated to citizens to information made available to modelers.
The central node of the infrastructure is a Spatial Data Infrastructure (SDI) useful to manage Web Sensors, Data Notification, Processing and complex data Visualization. Furthermore, a SDI may manage a huge amount of data and processes like atmospheric chemical models coupled with meteorological models, for which the integration of a complex computational framework is required.
The importance of integration between different tools of the SDI can be supported by the Open Geographic Consortium (OGC) standards (e.g. Sensor Web Enablement or Web Processing Services) in order to integrate results coming from different processes.
The integration of such technologies requires a different approach to develop a SDI, by using Information Communication Technology (ICT) methodologies as often adopted in e-business. A flexible Information Infrastructure, which solves integration issues between SDI components by hiding the complexity of geographic systems to final users, may be useful in our case.

4.2 The SDI architecture and implementation

With the above in mind, we developed an SDI architecture, which will be the corner stone of an integration architecture, where any other component can be plugged in. Services and processes provided by the SDI are controlled through an Information Infrastructure that wraps most useful processes of the proposed SDI in order to provide to final users high level services (e.g. data integration). This Information Infrastructure as middleware between users and

the SDI to provide a more friendly façade to the SDI. Open source components can be used to build the GMOS SDI.
We suggest PostGis, Geoserver and Geonetwork for geographic data storage, export geographic web services and manage metadata, respectively; whereas some Javascript libraries embedded in OpenLayers can be used to display geographic information. This approach enabled the construction of a SDI as pluggable system. This scenario needs additional efforts to solve the integration process between the different components. Information Infrastructure can plays an important role, exporting an uniform and unique interface for the components involved in the SDI.
The architecture of the GMOS SDI is aimed to provide geographic services for integration into a Service Oriented Architecture (SOA) like GEOSS. SOA is the main orientation of further development within the environmental sector. The SDI Architecture is hereafter defined trough the description of different viewpoints of the systems, using the Reference Model of Open Distributed Processing (RM-ODP) terminology (Naumenko 2001). RM-ODP is a model view used to describe complex ICT systems. In our case, the GMOS ICT system is described through viewpoints of the system itself and its environment. A part of RM-ODP is herein used in an informal way, as the formal specification of ICT system is out of the scope of this document. In the next section the following RM-ODP viewpoints will be described:

- Information Viewpoint
- Enterprise Viewpoint
- Computational Viewpoint
- Technology Viewpoint

4.3 Information Viewpoint

The information viewpoint is focused on the semantics of the information and the performed information processing. It describes the information management by the system and the structure and content type of the supporting data.

4.3.1 Data Analysis and state of the art

Data analysis and infrastructure state of the art are crucial to understand what kind of SDI to create. In the case of GMOS, data are related to mercury holding geographical information (i.e. latitude, longitude, elevation, depth, concentration or emission), which are gathered from monitoring systems and numerical simulations, running on a parallel computing clusters. Both measurements and simulations must be integrated into the SDI by considering working methodologies. For example, some data are or have been collected in

geographical containers (e.g. ESRI Shapefile, netCDF). Data coming for previous monitoring campaigns or model runs represent an essential background for the startup the SDI.
The proposed Information Infrastructure must wrap any data source and provide an interface to the final user in order to upload data and metadata. The SDI can be the entry-point for the collection of data coming from different sources, providing information to be shared between operators working in the same area of interest. Meanwhile it is very important to provide integration between hardware and software components within the SDI, to minimize human action along data acquisition. This kind of integration is possible only with a flexible software architecture, where any source can be wrapped with a specific Software Layer.

4.3.2 Metadata: The knowledge description

The description of the data stored in a SDI is fundamental as data themselves. Storing and organizing data needs to contextualize data themselves in order to give information on collection methodology, lineage, spatial and temporal domain, copyright, context of use, etc. Such documentation is called metadata. Metadata help organize and maintain data and provide information about an organizations data holdings in catalogue form; avoid duplication of effort by ensuring the organization is aware of the existence of data sets; help users to locate all available geospatial and associated data relevant to an area of interest; build upon and enhances the data management procedures of the geospatial community; promote the availability of geospatial data beyond the traditional geospatial community; advertise and promote the availability of their data and potentially link to on line services.
A detailed description of metadata an standards adopted in GMOS is the object of the Report on definition of metadata, standards with respect to INSPIRE Directive (Deliverable 9.2). To our knowledge it is now fundamental to say the metadata are stored as eXtensible Markup Language (XML), which is is a markup language for documents containing structured information. Structured information contains both content (words, pictures, etc.) and some indication of what role that content plays (for example, content in a section heading has a different meaning from content in a footnote, which means something different than content in a figure caption or content in a database table, etc.). Almost all documents have some structure. A markup language is a mechanism to identify structures in a document. The XML specification defines a standard way to add markup to documents (www.xml.com).
For the management of XMLs a tool is required, which leverage metadata creation and store it in a database. Among the open-source tools we selected GeoNetwork, which is well suitable for metadata management providing a metadata template compliant with INSPIRE. Geonetwork acts as a component plugged in the GMOS SDI Architecture, in order to manage metadata at all levels (metadata storage, metadata management and creation) and export

these metadata through OGC CS-W, a standard protocol used to describe metadata.

4.3.3 Thesaurus integration

The GMOS SDI will be integrated with a Thesaurus developed at CNR-IIA. Indexing and searching by a thesaurus to provide access to geo-information resources is still a quality issue in the Internet world. Often is waste of time when searching by subject where no indexing strategy is present: each searching session has to face all the uncertainties of natural language (synonymy, polysemy, homonymy), combined with all uncertainties of a full text search (no relevance control on the retrieved occurrences). Moreover, we can easily realize that even an indexing strategy itself is not enough, when we have to face either broad classification systems or indexing by using natural language: in the first case navigation in wide virtual 'containers' is often a time wasting operation; in the second case synonyms, polysemes or homonyms limit and/or delay efficient and effective information retrieval. These problems rapidly multiply when a large number of documents and a multilingual context are involved.
On the contrary, a thesaurus: i) guarantees the effective control of the indexing language, covering each selected concept with a preferred term (in each language in the case of multilingual thesauri) and ensuring inter-language equivalence among these descriptors; ii) provides a systematic display of the descriptors, making navigation through the terminology easier; iii) allows indexing and searching by combining several descriptors ex post, in order to refine and personalize both the semantic description and the information retrieval. In this way a good balance between the numbers of retrieved documents and their relevance is assured. These have been the main reasons for using the thesaurus EARTh (Plini 2004), which is oriented to environmental area of interest with a specific section for GIS terminology.
The Simple Knowledge Organization Systems (SKOS) will be used to integrate EARTh within the SDI.

4.4 Enterprise Viewpoint

The enterprise viewpoint focuses on purpose, scope, and policies of a system. It provides the context and overall environment within which the system will be built, and therefore constraints, and obligations that must apply in all other viewpoints. It represents the global requirements that the SDI must respect.

4.4.1 Actors and Users

From the perspective of an end-user, a SDI should organize information and resources and distribute them, providing their services via a single access point,

often online. Resources managed by a SDI should be geo-portals or catalogs of geographic Web Services oriented to users. In both cases, through the access point, the end-user should find resources by searching among metadata.
Users accessing the GMOS SDI can be a users or a Service Client, often a Web Service Client. In the former case, the SDI is used by operators that can be data readers or users that create data and metadata. Among them, decision makers in public administration and researchers interested in information are fundamental. In order to match the requirements for this kind of users it is important to create tools for searching geographic data, necessarily based on metadata, allowing the detection based on keywords and the geographic location.
Summarizing, three types of clients have been considered in the GMOS SDI:

- Web based, oriented to specific products and projects related to mercury in the environment or dedicated to decision makers in public organizations.
- Desktop GIS, oriented to researchers, operators and data providers.
- Web Service, compliant with OGC standard and designed for system integration and interoperability.

4.4.2 Functional Requirements

In environmental management, data are important inputs to simulation models and decision-making systems. They are widely distributed and heterogeneous and so difficult to access in a coordinated way. The first step to make systems interoperable, is to produce a spatial and temporal metadata registry that enables the community to more easily use and exchange data and services. Metadata are very important for interoperability and systems integration, because with metadata is possible to define a common standard between systems. The use of metadata is the main step for building complex geo-system where interoperability is a common target. It is important for GMOS SDI to manage metadata of data to improve the system interoperability.
The second step is to produce data in commonly recognized formats. A central management and storage of air quality data and metadata may enable the interoperable data access at all levels, by integrating in a central database data retrieved in most common format, including netCDF, ESRI Shapefile and Microsoft Excel and Access. Data storage can be virtual or even a physical device, but data centralization is important because with this approach it is possible to harmonize the knowledge in order to produce more services based on data stored. In many organizations working with geographic data, centralization and a harmonization of the knowledge are missing, making difficult multidisciplinary approach and advanced analysis. For this reason, the use of central data storage and a description trough metadata are a very strong requirements for our SDI.
From a user point of view, the creation of metadata can be a tedious task,

as the management of large XML with verbose descriptions may discourage some operator. As GMOS will manage large datasets some tools that leverage metadata creation and store it in a database is required. An additional point which discourage operators is the necessity to follow standards that are often updated or integrated. Metadata stored in database will support data search. The metadata visualization during the search phase has not a fundamental impact on the user, but a minimal visualization tool to perform researches based on geographic localization is required. After this phase, visualization is a critical requirements and the visualized products should be different as result of a different use of data. Resulting metadata must indicate access point related to data use. For example, web-based systems are browsed by end-users by means of geo-portals, in which geographic data are used to create high level services in specific projects and are related to particular areas of interest such as monitoring the ozone precursors emission from facilities, or simply view a list of information layers contained in the GMOS SDI. Web-based systems are therefore ideal tools to realize information systems that reach a wide range of users.

4.4.3 Non-functional requirements

In order to match new requirements, the SDI has been designed through a modular approach where each component implements some feature and talks with each others in order to perform required services. Through this approach, the change of a functional component has minimal impact on the system because each component is loosely coupled with each others. This architectural design supports positively the system upgrade and the scalability as new modules can be added to balance the load.
Other global requirements for the SDI are related to the friendliness of the user endpoint. Users need to interoperate with the SDI in all phase of the data management: data and metadata creation, process management, data access, data searching and, finally, visualization. Many tools have been developed to show data in a user-friendly manner (e.g. OpenLayers) but, in order to hide the complexity of SDI, some Information Infrastructure that helps data provider is required: a middleware that helps users to add data and metadata without any direct action on the SDI.

4.4.4 Integration Constraints

Interoperability and systems integration has become important in geographic systems and SOA architectures become popular even in geo-systems area . The SDI will be created having in mind the more high level system, GEOSS. The integration of GMOS SDI with GEOSS will occur primarily through the use of OGC CS-W protocol that supports the ability to publish collections of metadata of geospatial data, services and related resources. It is therefore an integration based on the common use of metadata standards according to the

INSPIRE directive. Interoperability enable data harmonization, through federation of more systems in complex and robust distributed system, as GEOSS.

4.5 Computational Viewpoint

The computational viewpoint enables the distribution through a functional decomposition of the system into objects which interact as interfaces. It describes the functionality provided by the system and its functional decomposition.
This SDI has a classical three layer architecture showed in Figure 4.1. A Data Storage Layer, connected to the Business Logic Layer that can be hosted in a Tomcat Server which can be linked to the Application Layer.

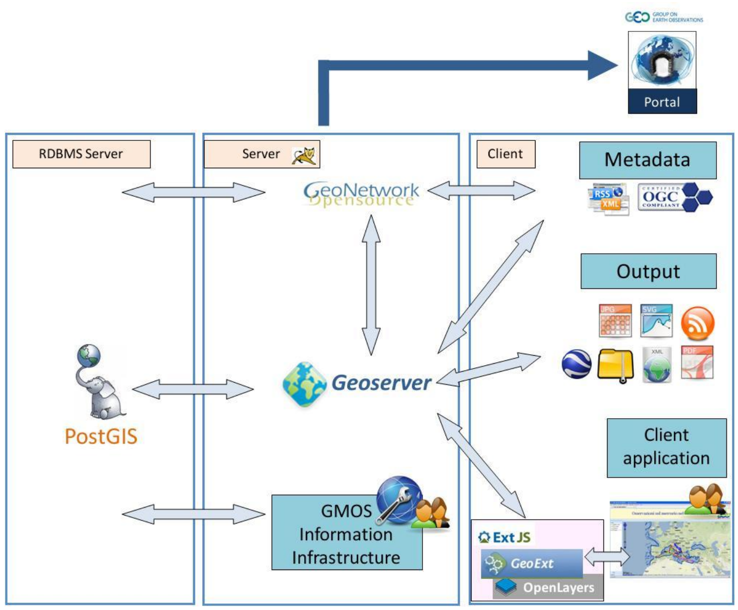

Fig. 4.1. GMOS SDI Architecture

The core of the system will be represented by the Database Management System (DMBS), which holds vector information and functional data. The DBMS will represent the Data Storage Layer in the SDI architecture, and will store metadata and geographic data in separate databases to maintain a different logic structure for each type of data. Additional databases dedicated to functional tasks and to information coming from web sensors can be implemented. In detail, the following databases can be contained in the Data Storage:

- d_container: used to store geographic information;
- g_container: used by Geonetwork to store metadata of data stored in the d_container;
- it_container: used by the Information Technology to perform functional tasks;
- sos_container: which contains data coming from web sensors.

The latter is a database designed by 52North (52north.org) in order to match the Sensor Web Enablement (SWE) specifications. sos_container can be connected to 52north SOS components deployed in the Business Logic Layer, and used to export sensor information.
Server components that perform the system functionality like metadata management, data creation, map creation and data dissemination pertain to the Business Logic Layer. These components are related to the Business Layer:

- Information Infrastructure: which will be oriented to the Geo Infrastructure;
- GeoServer: the map server;
- GeoNetwork: the metadata manager and cataloger;
- GiCat: the service broker.

The Information Infrastructure will wraps the SDI in order to provide to final user facilities related to data and metadata creation and data harmonization. Users can work with a set of tools that enable a more easy use of the SDI components that concern data and metadata creation, sensors integration and process management. This tool will allows users to create metadata compliant with European and international standard, such as the ISO 19139 and the INSPIRE Directive (see Deliverable 9.2).
Geoserver (geoserver.org) is a map server that exports data by creating geographic Web Services, compliant with OGC (WMS, WFS, WCS). These services can be used directly by end-users in complex SOA systems or geo-portals built with Web technologies. OGC services can also be used by GeoNetwork to integrate metadata with maps and other geographic data.
GeoNetwork (geonetwork-opensource.org) is a tool used to manage metadata. The metadata can be exported via the CSW 2.0.2 protocol, which is the basis for integration of GMOS SDI in GEOSS. The tool allows links with Geographical Services such as WMS, correlation with different, and even not structured, data sources.
GiCat (essi-lab.eu/cgi-bin/twiki/view/GIcat) is a Service Broker used as a collector for Geographic Web Services. It supports several protocols like WMS, WFS, THREDDS, CS-W. GiCat accepts many inputs, extracts information and exports them in standardized protocol as CS-W. It is linked directly with GeoNetwork to export metadata and to integrate the Air Quality SDI in more complex Systems of Systems like GIIDA or GEOSS.
The Application Layer represents a set of tools oriented to final users. This

container holds Desktop GISs and Web GISs. Desktop GIS systems can access information by a direct connection to the geographic database or through WMS services. The direct connection to the database enable the data modification to those that have credentials. Given the centrality of the database in our architecture, changes are made available to all services that use the data. Some Web GIS can be developed using services generated by Geoserver and are released to end-users by means of geo-portals, in which geographic data are used to create high level services in specific projects and are related to particular areas of interest. The web based systems are therefore ideal tools to realize information systems that reach a wide range of users.

4.6 Technology Viewpoint

The Technology Viewpoint is focused on the selection of technology of the system. It describes the technologies chosen to provide the processing, functionality and presentation of information.
The GMOS SDI can be based on the adoption of international standard, mainly OGC for data representation while ISO19138 and INSPIRE for metadata creation. In detail, standards like WMS 1.1.1 and WFS 1.1.0 must be adopted for data whereas CS-W 2.0.2 will be adopted for metadata dissemination trough GeoNetwork and GI-CAT.
WebGIS will be developed through javascript frameworks, as OpenLayers which allows to develop applications with AJAX support.
OpenLayers is a very effective tool that allows a connection directly to a WMS services in order to develop effective applications and easy to use. These methodologies currently represent the most used technique in the area of open source WebGIS products and offers excellent results. However, new methodologies for creating Web-based applications are available in the marketplace that allow to overcome the development approach based on Javascript.

4.6.1 Information Infrastructure

The Information Infrastructure will be developed to provide data creation and harmonization. The architecture in Figure 4.2 has been drawn to enable interactions between the database, OGC services exported via Geoserver and final users. Information Infrastructure will be an Enterprise Application, developed with Java and deployed in a Web Server Container (Tomcat). The scope of this framework is to provide to an institution, where geographic data and process are the core information, an infrastructure that centralize data persistence process and service creation, process control over geographic information and geographic data storage management. The Information Infrastructure will provides a AJAX web user interface in order to access to the features provided by the framework.

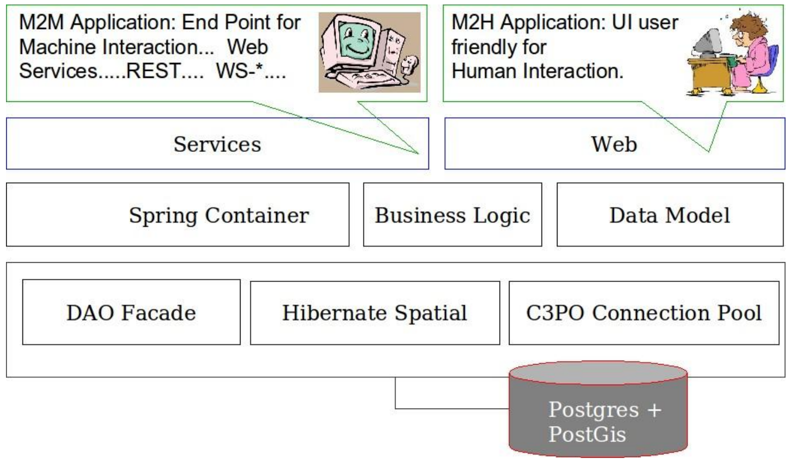

Fig. 4.2. Information Infrastructure components view

The web interface of will be developed with ZK (www.zkoss.org) in order to improve user friendliness and usability. ZK permits declarative approaches for web application and it can be very useful in order to develop complex web systems oriented to data visualization.
In this section each component that concurs to build the Information Infrastructure will be described.

4.6.2 Information Infrastructure Data Storage Management

Geographic data in the GMOS SDI will be stored in a Postgis database. This database is managed through the Information Infrastructure by means of Data Access Objects (DAO). The DAO is a patterns used in Software Engineer in order decoupling the storage systems from other Software Layer. In our case, there are DAOs that manage any instance of data storages.
At present, only relational databases are used by means of the DAO. However, the architecture proposed for the Information Infrastructure enable to switch from relational databases to other storage system. The DAO patterns allows this level of abstraction. Each Data Storage system can be enabled in the Information Infrastructure by mean of the corresponding DAO. For each tyupe of data source, a corresponding DAO will be implemented to activate a different storage strategy, according with the DAO interface provided.
This DAO exports the CRUD (Create, Retrieve, Upload, Delete) operation on the geographic objects stored in some storage device. The Information Infrastructure provides, by default, an implementation for Postgis but due to the DAO abstraction, many others relational databases could be used (i.g Oracle Spatial, MySql).
Postgis is a de facto open source standard as geographical DBMS. It supports

vector data storage but concern remains for raster data that currently are not supported by PostGis. Actually they are handled directly by Geoserver and stored in the File System. This is could be a limitation in the SDI because is very important to provide a single data storage for both vector and raster data. Nevertheless, an add-on to Postgis is expected in the near future to store raster data on the database by using mechanisms similar to those used for vectorial data (i.e. WKTRaster). Others relational DBMS used in GIS already supports raster data, but they arent open source.
In order to enable different databases than Postgis, a different implementation of the DAO could be provided and linked in the Information Infrastructure by means of the IoC (Inversion of Control) framework that manages the whole application. IoC is a pattern that enables a flexible setup of complex enterprise applications.
In the Information Infrastructure, Spring will be used as IoC Framework. In detail, it can be used as Application Context in order to setup any component and the relationship between components. Spring can manage not only DAO components but also Connections Pools (C3PO), JDBC Data Source, Logging component (Log4j), and services like mail systems and application utilities. All components are linked together by means of Spring framework at the Information Infrastructure start up.
Hibernate Spatial is one of the component that can be linked together by Spring framework in order to setup the Information Infrastructure. Hibernate is an Object Relational Mapping (ORM), a system that support the conversion of relational model to Object Model. It is very useful when a Object Oriented (OO) language is used with a relational database as persistence device, and, as the Information Infrastructure is developed in Java as OO language, Hibernate can play an important role where a relational database is used as storage system. In our case Hibernate will be used with a particular extension: Hibernate Spatial. This permits a more robust use of Geometric objects, stored in a relational DBMS.
In the Information Infrastructure a specific DAO will be used to perform user management and functional task: GeoIntDAO. The DAO system design can be described using UML (Figure 4.3). Each instance of PostGis database used and managed by the Information Infrastructure can be wrapped by an instance of PostgisContainerDAOImpl, that implement GContainerDAO. GContainerDAO will be collected in a Facade with GeoIntDAO, used for functional task such as account management.

The Façade pattern is a very useful strategy used in Software Engineer in order to provide a complex system with a more useful and direct interface. In this case, the components in the second layer of Figure 4.2 will manage any object by means of the method provided by the Faade. This approach realize a more robust Software Architecture and a more flexible use of the underlying resources.
For example, in order to switch from a Postgis database to another, the user

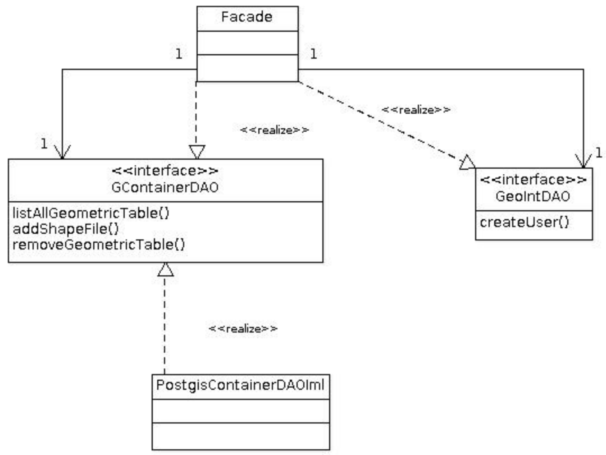

Fig. 4.3. Facade and DAO in the Information Infrastructure

must provide the Facade with the appropriate instance of PostgisContainerDAOImpl, even at run time. The Business Logic Component and others Software Layer that use the Facade will use the method without get awareness about the type and the location of the new Data Storage. Furthermore, the Facade can be used by components hosted in the higher level layer of the Information Infrastructure Architecture, in order to manage different data sources.

4.6.3 OGC Service Creation with the Information Infrastructure: Integration with Geoserver

Figure 4.2 showed the Architecture in which there is a layer that contains the business logic to perform data process and services. In this layer we can host Process and Services to store and manage data by means of the appropriate Façade. This business logic layer can contains also services used to integrate the Information Infrastructure with other software components. The component communication between the Information Infrastructure and Geoserver is fundamental in order to export the data stored in the Geographic databases through geographic Web Services designed by OGC (OWS).
The RESTfull endpoint provided by Geoserver can help to manage an instance of this map server: it is possible to add a new WMS/WFS services programmatically by means of a RESTfull operation enabled in Geoserver.
A use case scenario is reported in Figure 4.4 by means of UML sequence

diagram. Through the Information Infrastructure the user can select a Geo Database and finally a geo resource. The user can create a OWS Service by means of the Information Infrastructure: the Information Infrastructure will send a command to a Geoserver instance a get a return end point of the service created. This end point will be managed directly by the Information Infrastructure.
This simple scenario shows as the Information Infrastructure can hide the complexity of resource management to final user of the SDI, as the Information Infrastructure can manage the creation of a resource and the creation of Geo Services compliant with OGC standard through a single interface.

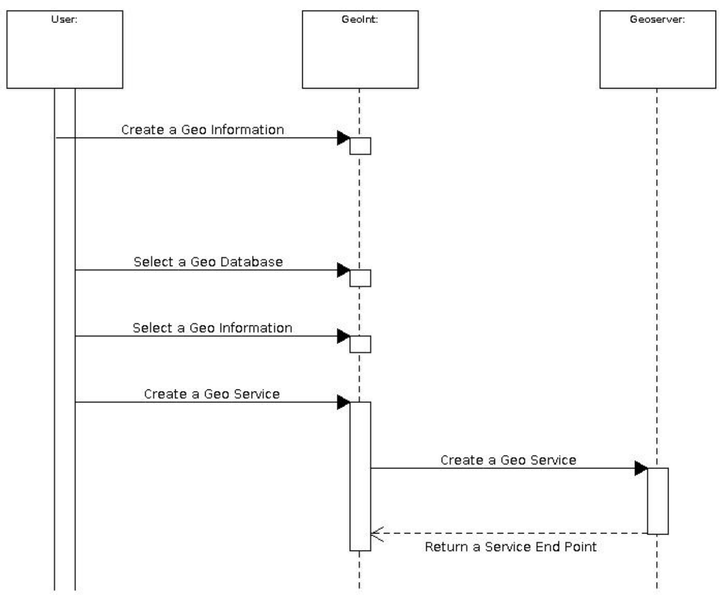

Fig. 4.4. Service Creation with the Information Infrastructure

4.7 Data Integration: upload data in the GMOS ICT System

Users interested to add data to the ICT System described above, can adopt different approaches to upload data (Figure 4.5). The Infrastructure will provide to User a Web Interface where Administrator can manage data and process. Two main method to upload data in the GMOS ICT System have been considered:

1. Using the Web User Interface, partners can upload their data directly. The data will be managed by Cyber Infrastructure and stored in the Data Storage Layer;
2. the Infrastructure gets a connection with patterns ICT framework, if present. It reads information from partners in a defined and configurable way including read frequency, location of data sources, protocol used (FTP, HTTP....), data type and data format. Different patterns of interaction could be implemented about interaction between GMOS ICT System and ICT framework developed by partners: in a pull behaviour, GMOS ICT periodically call partners about new data, while in a push behaviour partners can notify GMOS system about new data to be integrated in GMOS infrastructure.

Both scenarios described above require information about data sources. Any parters have to provide information about:

1. Data Format of data to be integrated in GMOS System (CSV, XLS, DBF, SHP, etc.)
2. Metadata about data to be integrated. In Example: type of measurements, localization of sensor, quality of measurements, headers.
3. If applicable or available, an example of such measurements for each partners.
4. A description of ICT system developed in partners organizations, if any. (Remote repository of data, Web Services or each information infrastructure that can interacts with GMOS ICT System).
5. Protocols available in partners organizations in order to data transmission (FTP, HTTP)
6. Any type of information about data managed by partners that join GMOS and that participants consider important in order to develop an Interoperable System.

These information will be very useful in order to design GMOS ICT System. The System about data integration will be developed and awareness about partners data structure is required.

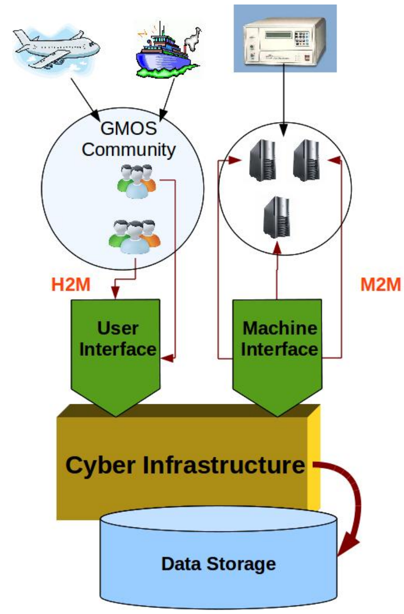

Fig. 4.5. The GMOS ICT System: Data Integration

5

Definition of metadata and standards with respect to INSPIRE

5.1 Metadata and Data Sharing

One of the main task in GMOS project is to enable the interoperability and, where practicable, harmonisation of spatial data sets and services regarding the Global Mercury Observation Network.
It is important to note that interoperability has to go beyond ”mercury particular community”, but take the various cross-community information needs into account as the context of GEOSS is. The interoperability faces with harmonisation of information, often including geospatial features, and description of information.
It is also important to note that ”interoperability” is understood as providing access to spatial data sets through network services in a representation that allows for combining them with other such spatial data sets in a coherent way. In our specific case interoperability means provision of data and related services conformal to pre-defined standards, which make data and services prompt to whatever kind of use under whatever type of system for whatever user.
The use of standards is a key concern to establish an interoperable system. Despite GMOS is developed within an international context involving partner outside the EU (e.g. USA, China, Russia, Japan) it will follow the European INSPIRE directive (Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007) establishing an Infrastructure for Spatial Information in the European Community to support Community environmental policies, and policies or activities which may have an impact on the environment. To this extent, to ensure the GMOS spatial data infrastructures compatibility, the GMOS Interoperable team adopted the INSPIRE rules (which however stands on the ISO standards), which help harmonisation of spatial data sets and services within Europe and the GEOSS community. Common Implementing Rules (IR) in specific areas are reported in Table 5.1.

This Deliverable will address the following components to help GMOS community to adopt the above IR:

Table 5.1. Implementing rules adopted for data, metadata and services

Area	**Implementing rules guidance documents**
Metadata	INSPIRE Metadata Implementing Rules: Technical Guidelines based on EN ISO 19115 and EN ISO 19119 (Version 1.2) 16.06.2010
Data Specifications	INSPIRE Specification on Coordinate Reference Systems - Guidelines v 3.1 03.05.2010
Network Services	Technical Guidance for the implementation of INSPIRE Discovery Services 30.03.2011 Technical Guidance for the implementation of INSPIRE View Services 30.03.2011 Technical Guidance for the INSPIRE Schema Transformation Network Service 15.12.2010 Draft Technical Guidance for INSPIRE Coordinate Transformation Services 15.03.2010 Draft Technical Guidance Download Services (version 2.0) 25.09.2009
Monitoring and Reporting	Not applicable

- rules for application schemas;
- coordinate referencing and units;
- identifier management;
- multi-lingual text and cultural adaptability;
- multiple representations (levels of detail), consistency, and more.

5.2 Rules for application schemas

An application schemas is a formal descriptor (normative rules) of data structure and content. The purpose of this component is to:

- provide a computer-readable data description defining the data structure - enabling automated mechanisms for data management;
- achieve a common and correct understanding of the data, by documenting the data content of the particular theme, thereby making it possible to unambiguously retrieve information from the data.

5.2.1 Metadata

In general, access to information (i.e. data sets) is facilitated if they are catalogued. In the case of digital information, the catalogue stores the identity, meaning, representation, and relationships of concepts or things in the real world as they are managed in digital systems. A catalogue, then, acts as a dictionary for data sets. This descriptive entity is called metadata.
The word metadata shares the same Greek root as the word metamorphosis.

"Meta-" means change and metadata, or "data about data" describe the origins of and track the changes to data. Metadata is the term used to describe the summary information or characteristics of a set of data. This very general definition includes an almost limitless spectrum of possibilities ranging from human-generated textual description of a resource to machine-generated data that may be useful to software applications. More recently, the term metadata has even been applied to services as a description of published service characteristics (SDI Cookbook, 2009).

Metadata helps the:

- Discovery metadata - What data sets hold the sort of data I am interested in? This enable organisations to know and publicise what data holdings they have.
- Exploration metadata - Do the identified data sets contain sufficient information to enable a sensible analysis to be made for my purposes? This is documentation to be provided with the data to ensure that others use the data correctly and wisely.
- Exploitation metadata What is the process of obtaining and using the data that are required? This helps end-users and provider organisations to effectively store, reuse, maintain and archive their data holdings (SDI Cookbook, 2009).

Each of the above purpose, requires a different level of information:

- Discovery metadata "is the minimum amount of information that needs to be provided to convey to the inquirer the nature and content of the data resource".
- Exploration metadata "provides sufficient information enable an inquirer to ascertain that data fit for a given purpose exists, to evaluate its properties, and to reference some point of contact for more information".
- Exploitation metadata "include those properties required to access, transfer, load, interpret, and apply the data in the end application where it is exploited".

Significant benefits derives from metadata creation as it helps:

- Organisation and maintenance of data by providing a catalogue of data;
- Coordination to avoid duplication and aware of the existence of data sets;
- Data location in a specific geographic context;
- Enhancement of the data management procedures of the geospatial community;
- Promoting the geospatial data availability beyond the traditional geospatial community
- Advertising and promoting the availability of own data and potentially link to on line services (e.g. text reports, images, web mapping and e-commerce) that relate to specific data sets.

The most important reference standard for geospatial data are the ISO 19115 (International Standard) and ISO 19139 (Draft Technical Specification).
An ISO standard for standard metadata was published and approved in 2003 . The ISO standard was derived from inputs from the the various national bodies and their implementations of the respective metadata standards assisted by metadata software. Indeed, most of the existing standards already have a great deal in common with each other, and a robust international discussion has ensured that the ISO standard has accommodated most of the various international requirements. ISO 19115 provides an abstract or logical model for the organization of geospatial metadata. It does not provide for rigorous compliance testing as there is no normative guidance on formatting the metadata included in the standard. A companion specification, ISO 19139, standardises the expression of 19115 metadata using the Extensible Markup Language (XML) and includes the logical model (UML) derived from ISO 19115. In North America, work is beginning to create a North American Profile of Metadata based on ISO 19139 for Canada, the United States, and Mexico. This will allow for the compliance testing of metadata files using XML.
Metadata also forms an important part of the OpenGIS Abstract Specification. The OpenGIS Consortium (OGC) http://www.opengis.org is an international membership organisation engaged in a co-operative effort to create open computing specifications in the area of geoprocessing. As part of its draft 'OpenGIS Abstract Specification' OGC has adopted ISO 19115 as the abstract model for metadata management within the consortium. OGC is working closely with FGDC and ISO/TC 211 to develop formal, global spatial metadata standards. At their plenary meeting in Vienna, Austriain March 1999, ISO/TC 211 welcomed the satisfactory completion of the co-operative agreement between the OpenGIS Consortium and ISO/TC 211 and endorsed the terms of reference for an ISO/TC 211 / OGC co-ordination group.
Each of the initiatives is promoting the standards and use of discovery metadata as a foundation of their respective metadata directory initiatives. This discovery metadata provides sufficient information to enable an inquirer to ascertain that existence of data fit for purpose exists and to reference some point of contact for more information. If, after discovery, more detail is needed about individual data sets then more comprehensive and more specific metadata is required. It is possible that organisations may wish to develop metadata at different but complementary levels - at one level discovery metadata for external use and for in-house / internal use more detailed metadata. And to avoid duplication of effort those elements common to both are flagged. These guidelines have been developed with recognition of the importance of more extensive metadata required for data management and each of the organisations is promoting the adoption of ISO Metadata Standard (SDI Cookbook, 2009).
INSPIRE will be based on the infrastructures for spatial information that are created and maintained by the Member States. The components of those in-

frastructures include: metadata, spatial data themes (as described in Annexes I, II, III of the Directive), spatial data services; network services and technologies; agreements on data and service sharing, access and use; coordination and monitoring mechanisms, processes and procedures.
The guiding principles of INSPIRE are that the infrastructures for spatial information in the Member States will be designed to ensure that spatial data are stored, made available and maintained at the most appropriate level; that it is possible to combine spatial data and services from different sources across the Community in a consistent way and share them between several users and applications; that it is possible for spatial data collected at one level of public authority to be shared between all the different levels of public authorities; that spatial data and services are made available under conditions that do not restrict their extensive use; that it is easy to discover available spatial data, to evaluate their fitness for purpose and to know the conditions applicable to their use (Drafting Team "Data Specifications" Guidelines for the encoding of spatial data).

5.2.1.1 Implementing rules

INSPIRE implementing rules are based on international standards established for metadata for spatial data (EN ISO 19115) and spatial data services (EN ISO 19119). INSPIRE also implements extended elements and different mandatory elements than the standard ISO. Tables below compare the core requirements of ISO 19115 to the requirements of INSPIRE for spatial dataset and spatial dataset series (Table 5.2 and Table 5.3) and for services (Table 5.4 and Table 5.5) as defined in the Implementing Rules (INSPIRE Metadata Implementing Rules: Technical Guidelines based on EN ISO 19115 and EN ISO 19119).

The practical implementation of the INSPIRE rules passes through an encoding process following the ISO 19139 XML (eXtensible Markup Language). Once encoded in this way, a metadata record can be validated using:

- XML schema, which can test the essential structure of the XML file
- XSLT Schematron, which can test for rules applied to the content of the XML file.

INSPIRE has an on-line validator available at:
http://www.inspire-geoportal.eu/index.cfm.

5.2.2 Data

Data specification pertains to WP2, WP3, WP4 and WP5 as they will prepare Standard Operating Procedures (SOPs) to collect data from filed instruments

Table 5.2. Requirements for ISO and INSPIRE Core Metadata Elements - Part 1

ISO 19115	**Core INSPIRE**	**Comments**
Dataset title (M)	Part B 1.1 Resource Title	-
Dataset reference date (M)	Part B 5 Temporal Reference	ISO 19115 is more demanding. The metadata shall contain a date of publication, revision or creation of the resource, while in INSPIRE the Temporal Reference can also be expressed through Temporal Extent.
Dataset responsible party (O)	Part B 9 Responsible organisation	INSPIRE is more demanding by mandating both the name of the organisation, and a contact e-mail address.
Geographic location of the dataset (C)	Part B 4.1 Geographic Bounding Box	INSPIRE is more restrictive. A Geographic bounding box is mandated
Dataset language (M)	Part B 1.7 Resource Language	ISO 19115 is more demanding. It mandates the dataset language, even if the resource does not include any textual information. The ISO 19115 Dataset language is defaulted to the Metadata language.
Dataset character set (C)	-	ISO 19115 is more demanding. The dataset character set has to be documented in ISO 19115 when ISO 10646-1 is not used.
Dataset topic category (M)	Part B 2.1 Topic Category	-
Spatial resolution of the dataset (O)	Part B 6.2 Spatial Resolution	-
Abstract describing the dataset (M)	Part B 1.2 Resource abstract	-
Distribution format (O)	-	-
Additional extent information for the dataset (vertical and temporal) (O)	Part B 5.1 Temporal extent	INSPIRE is more demanding. A temporal reference is mandated, and can be expressed as a temporal extent.
Spatial representation type (O)	-	-
Reference system (O)	-	-
Lineage (O)	Part B 6.1 Lineage	INSPIRE is more demanding. A general lineage statement is mandated.

Table 5.3. Requirements for ISO and INSPIRE Core Metadata Elements - Part 2

ISO 19115	**Core INSPIRE**	**Comments**
On-line resource (O)	Part B 1.4 Resource Locator	-
Metadata file identifier (O)	-	-
Metadata standard name (O)	-	-
Metadata standard version (O)	-	-
Metadata language (C)	Part B 10.3 Metadata Language	INSPIRE is more demanding. The metadata language is mandated even if it is defined by the encoding.
Metadata character set (C)	-	ISO 19115 is more demanding. The metadata character set has to be documented in ISO 19115 when ISO 10646-1 is not used.
Metadata point of contact (M)	Part B 10.1 Metadata point of contact	INSPIRE is more demanding by mandating both the name of the organisation, and a contact e-mail address.
Metadata date stamp (M)	Part B 10.2 Metadata Date	ISO is more restrictive because this element shall contain the ”date that the metadata was created” and INSPIRE may contain the date when the metadata record was created or updated.
-	Part B 1.3 Resource Type	INSPIRE is more demanding.
-	Part B 1.5 Unique Resource Identifier	INSPIRE is more demanding.
-	Part B 3 Keyword	INSPIRE is more demanding.
-	Part B 7 Conformity	INSPIRE is more demanding.
-	Part B 8.1 Conditions for access and use	INSPIRE is more demanding.
-	Part B 8.2 Limitations on public access	INSPIRE is more demanding.

Table 5.4. Requirements for ISO and INSPIRE Core Services Elements - Part 1

ISO 19115 Core	**INSPIRE**	**Comments**
Dataset title (M)	Part B 1.1 Resource Title	-
Dataset reference date (M)	Part B 5 Temporal Reference	ISO 19115 is more demanding. Despite its name, this ISO 19115 Core metadata element applies to services. A reference date of the service (date of publication, revision or creation...) is mandated.
Dataset responsible party (O)	Part B 9 Responsible organisation	-
Geographic location of the dataset (C)	-	See INSPIRE Geographic Bounding Box.
-	Part B 4.1 Geographic Bounding Box	The Geographic Bounding Box is handled in ISO 19119 with a different metadata element from the one corresponding to "Geographic location of the dataset".
Dataset language (M)	-	Not applicable to services.
Dataset character set (C)	-	Not applicable to services
Dataset topic category (M)	-	Not applicable to services
Spatial resolution of the dataset (O)	Part B 6.2 Spatial Resolution	In the current version of ISO 19119, it is not possible to express the restriction of a service concerning the spatial resolution
Abstract describing the dataset (M)	Part B 1.2 Resource abstract	-
Distribution format (O)	-	-
Additional extent information for the dataset (O)	-	-
Spatial representation type (O)	-	-
Reference system (O)	-	-

and define data temporary storage. Following collection and quality controls, data will be stored in the SDI within a Postgresql -Postgis database (See at chapter 4, for details).

5.2.3 Services

Data specification pertains to WP2, WP3, WP4 and WP5 as they will prepare Standard Operating Procedures (SOPs) to collect data from filed instruments

Table 5.5. Requirements for ISO and INSPIRE Core Services Elements - Part 2

ISO 19115 Core	**INSPIRE**	**Comments**
Lineage (O)	-	-
On-line resource (O)	Part B 1.4 Resource Locator	-
Metadata file identifier (O)	-	-
Metadata standard name (O)	-	-
Metadata standard version (O)	-	-
Metadata language (C)	Part B 10.3 Metadata Language	INSPIRE is more demanding. The metadata language is mandated.
Metadata character set (C)	-	ISO 19115 is more demanding. The metadata character set has to be documented in ISO 19115 when ISO 10646-1 is not used.
Metadata point of contact (M)	Part B 10.1 Metadata point of contact	-
Metadata date stamp (M)	Part B 10.2 Metadata Date	ISO is more restrictive because this element shall contain the "date that the metadata was created" and INSPIRE may contain the date when the metadata record was created or updated.
-	Part B 1.3 Resource Type	INSPIRE is more demanding.
-	Part B 1.6 Coupled Resource	Optional in INSPIRE.
-	Part B 2.2 Spatial Data Service Type	INSPIRE is more demanding.
-	Part B 3 Keyword	INSPIRE is more demanding.
-	Part B 7 Conformity	INSPIRE is more demanding.
-	Part B 8.1 Conditions for access and use	INSPIRE is more demanding.
-	Part B 8.2 Limitations on public access	INSPIRE is more demanding.

and define data temporary storage. Following collection and quality controls, data will be stored in the SDI within a Postgresql -Postgis database (See chapter 4, for details).
GMOS Discovery Services allow users and computer programs to search for spatial datasets and services based on their metadata records. To ensure a consistent and compatible implementation of services, some requirements and recommendations must be considered.
The base functionality of the GMOS Discovery Service is derived from INSPIRE (Technical Guidance or the implementation of INSPIRE Discovery Services).
Following this Technical Guidance, services will be based on OGC Catalogue Services Specification 2.0.2 - ISO Metadata Application Profile for CSW 2.0, which will implement the following operations:

- Get Discovery Service Metadata: Provides all necessary information about the Discovery Service and describes service capabilities;
- Discover Metadata : Allows requesting INSPIRE metadata elements of spatial data sets and services from a Discovery Service;
- Publish Metadata: Allows editing of INSPIRE metadata elements of resources in the
- Discovery Service (push or pull metadata mechanisms). Editing meaning insert, update and delete;
- Link Discovery Service: Allows the declaration of the availability of a Discovery Service for the discovery of resources through the Member State Discovery Service while maintaining the resource metadata at the owners location.

Searching criteria are required following common *queryable* as defined by OGC. Table 5.6 and Table 5.7 report INSPIRE search criteria to be implemented in GMOS.

5.3 Coordinate reference system, projection and units

5.3.1 Coordinate system

Standardisation of coordinate reference systems and projections for the storage and treatment of geographic databases and map display in GMOS will follow international standards as the inability of the large geodetic systems, such as European Datum (ED50), North American Datum (NAD), and Tokyo Datum (TD), to provide a worldwide geo-data basis, and the final GEOSS context to which will contribute the project.
The spatial referencing is usually referred to selected points of the earth surface. In the case of GMOS, such points are given by permanent monitoring stations or temporary monitoring sites (points) over oceans or in the stratosphere.

Table 5.6. Mandatory search criteria (*queryable*) to be implemented in GMOS - Part 1

INSPIRE *queryable* metadata elements	**INSPIRE Discovery Service *queryable* properties**	**Is mandatory for INSPIRE Discovery Service?**
Keyword	Subject	Yes
Topic category	TopicCategory	Yes, if resources of type *dataset* or *series* are supported by the catalogue service instance
Spatial data service type	ServiceType	Yes, if resources of type *service* are supported by the catalogue service instance.
Lineage	-(not supported)	Yes
Spatial resolution	SpatialResolution	Yes, if resources of type *dataset* or series are supported by the discovery service instance
Specification	-(not supported)	Yes
Degree	-(not supported)	Yes
Geographic bounding box	BoundingBox	Yes, if resources of type *dataset* or *series* are supported by the catalogue service instance
Conditions applying to access and use	-(not supported)	Yes
Limitations on public access	-(not supported)	Yes

It was established to adopt the World Geodetic System (WGS) and natural coordinates to store data holding geographic coordinates. WGS is a standard for use in cartography, geodesy, and navigation, a standard coordinate frame for the Earth, a standard spheroidal reference surface (the datum or reference ellipsoid) for raw altitude data, and a gravitational equipotential surface (the geoid) that defines the nominal sea level.
The latest revision is WGS-84 (dating from 1984 last revised in 2004, which will be valid up to about 2010) is developed, maintained and enhanced by the National Geospatial-Intelligence Agency.

The coordinate origin of WGS-84 is meant to be located at the Earth's center of mass; the error is believed to be less than 2 cm. In WGS 84, the meridian of zero longitude is the IERS Reference Meridian. It lies 5.31 arc seconds east of the Greenwich Prime Meridian, which corresponds to 102.5 metres at the latitude of the Royal Observatory.
As of the latest revision, the WGS-84 datum surface is defined as an oblate spheroid (ellipsoid), with major (transverse) radius a = 6,378,137 m at the equator, semi-minor radius b = 6,356,752.3142, and flattening f = 1/298.257

Table 5.7. Mandatory search criteria (*queryable*) to be implemented in GMOS - Part 2

INSPIRE *queryable* metadata elements	**INSPIRE Discovery Service queryable properties**	**Is mandatory for INSPIRE Discovery Service?**
Responsible party	OrganisationName	Yes
Responsible party	role	Yes
Resource Title	Title	Yes
Resource Abstract	Abstract	Yes
Resource Type	Type	Yes
Unique resource identifier	ResourceIdentifier	Yes
Temporal Reference	TemporalExtent PublicationDate RevisionDate CreationDate	Yes
Metadata language	MetadataLanguage	Yes

223 563 (this is a flattening of 21.384 685 755 km, or 0.335% in relative terms).
Presently WGS 84 uses the 1996 Earth Gravitational Model (EGM96) geoid, revised in 2004. This geoid defines the nominal sea level surface by means of a spherical harmonics series of degree 360 (which provides about 100 km horizontal resolution). The deviations of the EGM96 geoid from the WGS 84 reference ellipsoid range from about -105 m to about +85 m.
The geographic latitude (abbreviation: Lat, , or phi) and longitude (abbreviation: Long, , or lambda) will be stored as coordinates containing degrees (integer), minutes (integer), and seconds (integer, or real number) (DMS) or coordinate containing degrees (integer) and minutes (real number) (MinDec) (41.899546 N; 12.474723 E).

5.3.2 Projection

he representation on plane of objects lying over curved surface needs the projection. Therefore, more generally, a map projection is any method of "flattening" into a plane a continuous surface having curvature in all three spatial dimensions. Map projection distort the surface in some fashion. Depending on the purpose of the map, some distortions are acceptable and others are not; therefore different map projections exist in order to preserve some properties of the sphere-like body at the expense of other properties.
As main map representation activity within GMOS will conceive technical reports to the European Commission, it is suggested to adopt the European Environmental Agency guidelines as much as possible.
The European Terrestrial Reference System 1989 (ETRS89) can be used as geodetic datum for pan-European spatial data representation. Also, following the European Commission recommendation the Lambert Azimuthal Equal

Area (ETRS89-LAEA; 52 N, 10 E, false easting: 4 321 000, false northing: 3 210 000) projection can be selected due to its suitability for map display purposes. Moreover, this is the EEA projection recommended for spatial analysis, e.g. grid analysis, as this is an area-true projection and it is recommended for use when combining layers, measuring areas and distances, and in sampling processes for statistical purposes. For practical reasons the WGS84 datum can be chosen as a secondary coordinate reference system. Although any reference vector data should be stored as WGS84 datum, it is not suitable for map display purposes, nor analysis. Nevertheless, it is easily transformable to ETRS89 and widely used (i.e. in GPS devices, it is the GoogleEarth datum).
In the case of global map representation, the Winkel Tripel projection (Winkel III) is suggested. This is a modified azimuthal map projection, which produces small distance errors, small combinations of Tissot indicatrix ellipticity and area errors, and the smallest skewness of any other projection.
In 1998, the Winkel Tripel projection replaced the Robinson projection as the standard projection for world maps made by the National Geographic Society.

5.4 Identifier management

Unique identification of spatial objects is provided by external object identifiers. Identifiers can be published by the responsible data provider with the intention that they may be used by third parties to reference the spatial object or can be established by data owner. For example a third party data identifier can be the country name (official short names in English) as given in ISO 3166 or NUTS codes for administrative areas.
Unique object identifiers in the context of GMOS are related to names of monitoring sites for which was established to adopt a key for each station. This key will be adopted along any data process systems. For example, for station *Bariloche* the code ALE is adopted. This code is used in order to identifies datasets related to Bariloche site.

5.5 Multi-lingual text and cultural adaptability

In the context of application schema, multilingual aspects are crucial with geographical names and spatial object and attribute classifications (feature catalogues, feature concept dictionaries, code list values). Translation of geographical names shall not be required in GMOS, only exonyms may be used. For example "Rome" will become "Roma" in Italian (exonym).
There shall not be a limitation to the number of names in different languages for one spatial object. For example names of cities (e.g. "Bruxelles", "Brussel", "Brussels", "Brssel", etc.) or other spatial object names in a multilingual area ("Brussel-Zuid", "Bruxelles-Midi").
The solution to multi-lingual issues is not the translation of everything into a

common language (e.g., English). Often, it is sufficient to obtain resources in their original production language, rather than in its translated version. An example may be textual descriptions of a feature. However, in order to find spatial objects, certain service interfaces, portals or client applications should support translation capabilities.

5.6 Data consistency

Data consistency refers to the usability of data and is often taken for granted in the modelling environment. Consistency between data is a very complex subject to deal with. Even if data are harmonised according to very well defined rules, they rarely fit exactly for various reasons. Certain inconsistencies between data sets may be related to the temporal differences between datasets, e.g. update frequencies.

"Inconsistencies" related to temporal differences are strictly speaking not inconsistencies. However, in practise this may cause difficulties in the production of harmonised datasets and may impact modelling activity in GMOS.

The following recommendations for processes introducing and maintaining consistency between data can reduce inconsistency.

The first consistency to check is the conformance to data specifications reported in SOPs (see specific documents) including the data capturing rules. Before checking the consistency between different data sets, each data set shall be verified to conform to the corresponding, in particular complying to the same set of constraints.

The harmonisation of data specifications is the best way to promote consistency. Specific tools related to the use of ontologies (i.e. Thesauri) in data specifications will be used.

Recommended data capture rules should be specified in sufficient detail to indicate the data requirements for a dataset and minimise harmonisation inconsistencies over a longer time frame. Since existing data will be the basis of GMOS, feasibility (as well as the contribute to GEOSS) and cost/benefit will not always allow for a harmonisation of capturing rules (unless for example they can be covered by transformation services).

6

Server and databases

6.1 SDI implementation

The GMOS Spatial Data Infrastructure (SDI) as designed in chapter 4, was installed and tested. This small Deliverable reports main key points as most of them can be experienced directly on the web.
The activity *Data Collection tools development* was here reported. The goal of this task is to develop the data collection tools and to set up of the communication and networking between the collection tools at the front end (monitoring sites and model outputs). This task has tight relation to Task 9.3 which bridges the developed tools to the front end.
The SDI has a three layer architecture as showed in Figure 6.1 with a Data Storage Layer, connected to the Business Logic Layer that can be hosted in a Tomcat Server which can be linked to the Application Layer. A small difference to that reported in chapter 4) is represented by a new component (GI-cat), which enable the SDI to be retrieved by brokering services.

The core of the system is represented by the Data Storage Layer, which holds vector and functional information. The PostgresSQL+PostGIS DBMS represents the Data Storage Layer in the SDI architecture, which store metadata and geographic data in separate databases to maintain a different logic structure for each type of data. Additional databases dedicated to functional tasks and to information coming from web sensors are implemented.

In detail, the following databases are contained in the Data Storage:

- d_container: is used to store geographic information;
- g_container: is used to store metadata of data stored in the d_container;
- it_container: is used by the Information Technology to perform functional tasks;

Server components that perform the system functionality like metadata management, data creation, map creation and data dissemination pertain to the Business Logic Layer. The followings are related to the Business Layer:

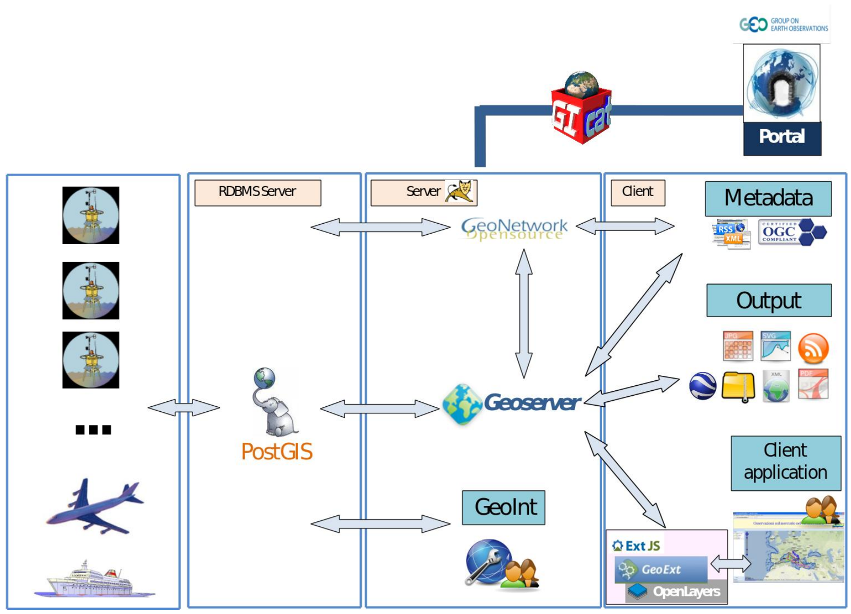

Fig. 6.1. GMOS SDI. Architecture

- Information Infrastructure: which will be oriented to the Geo Infrastructure;
- Geoserver: the map server;
- Geonetwork: the metadata manager and cataloguer;
- Gi-cat: the service broker.

The Information Infrastructure wraps the SDI in order to provide to final user facilities related to data and metadata creation and data harmonization. Users can work with a set of tools that enable a more easy use of the SDI components that concern data and metadata creation, sensors integration and process management. This tool allows users to create metadata compliant with European and international standard, such as the ISO 19139 and the INSPIRE Directive (see chapter 5).
Geoserver (geoserver.org) is a map server that exports data by creating geographic Web Services, compliant with OGC (WMS, WFS, WCS). These services can be used directly by end-users in complex SOA systems or geo-portals built with Web technologies. OGC services can also be used by Geonetwork to integrate metadata with maps and other geographic data.
Geonetwork (geonetwork-opensource.org) is a tool used to manage metadata. The metadata can be exported via the CSW 2.0.2 protocol, which is the basis for integration of GMOS SDI in GEOSS. The tool allows links with Geographical Services such as WMS, correlation with different, and even not structured, data sources.

GiCat (essi-lab.eu/cgi-bin/twiki/view/GIcat) is a Service Broker used as a collector for Geographic Web Services. It supports several protocols like WMS, WFS, THREDDS, CS-W. Gi-cat accepts many inputs, extracts information and exports them in standardized protocol as CS-W. It is linked directly with Geonetwork to export metadata and to integrate the GMOS system in more complex Systems of Systems like MACC or GEOSS.

6.2 Main characteristics of the user-friendly interface

The SDI has a user-friendly interface (Figure 6.2), which is managed by the System Administrator.

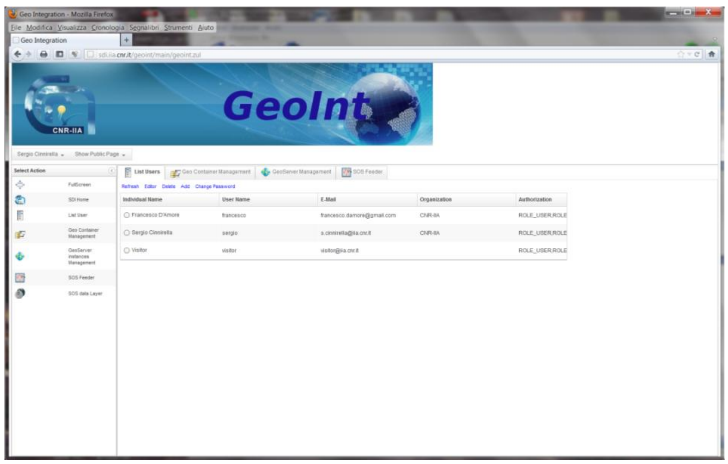

Fig. 6.2. The user-friendly interface for the GMOS SDI manager

The Administrator can set the role for each user. At present four roles have been implemented:

- View only: the user can only view information on established database/s;
- Editor of data: the user can upload and delete data on selected database/s;
- Editor of Metadata: the user can add/delete Metadata, which are used for the download page;
- Download of data: the user can download data.

Through this interface, users can manage GMOS databases by uploading data (as different formats. csv, shp, etc.) delete them, manage naming and view content (Figure 6.3).

Fig. 6.3. The user-friendly interface for the GMOS SDI manager

6.3 Main characteristics of collected information

Data collected by Work Package 2 (WP2), WP3, WP4 and WP5 have been uploaded on a temporary space in order to make quality control, before their final upload on the GMOS database.

The following parameters have been collected for land-based monitoring sites(Table A1, Annex 1):

- METEO parameters
- Total Gaseous Mercury
- Gaseous Elemental Mercury
- Reactive Gaseous Mercury
- Total Mercury in precipitation
- Methilmercury in precipitation
- Mercury deposition

Not all site have the same set of parameters whereas most data suffer lack of information as well as several metadata have not been delivered. In the case of cruise campaigns, the following parameters have been collected.

- METEO parameters
- Dissolved Gaseous Mercury
- Total Gaseous Mercury
- Reactive Gaseous Mercury
- Methilmercury in water
- Dimethilmercury in water
- Mercury fluxes
- Particle bounded Mercury

- Several water phisico-chemical parameters

Also in this case lack of information for both data and metadata is high. Finally, dataset on emissions from anthropogenic source have been collected from W2. They pertain point and areal sources for different mercury species. Emissions have been estimated at different stack level.
In this case, the lack of information pertain the grid distribution for the most recent dataset.

7

Communication and networking for data retrieval

7.1 Configuration of all GMOS ground-based sites

One of the task objective of GMOS network is to implement system and tools which can help data collection and data integration. Data coming from different instruments and collected by GMOS partners, will be integrated (i.e. collected) in a core databases managed by the GMOS-Cyber Infrastructure (GMOS-CI).
This Deliverable shows how the GMOS-CI works in order to collect raw data on mercury observations and ancillary parameters coming from the GMOS network.
GMOS ground-based site have been established in the southern and northern hemisphere as reported in Figure 7.1. Differences between sites pertain to measurement methodology.

Master stations are those where Gaseous Elemental Mercury (GEM), Gaseous Oxidized Mercury (GOM), mercury associated to suspended particulate matter (PBM2.5) and Hg in precipitation are continuously measured. Secondary stations are those where only Total Gaseous Mercury (TGM) and Hg in precipitation are continuously measured. In addition, ancillary parameters are often monitored as for example CO2, SO2, N2O, NO, NO2, NOx, Heavy Metals, VOC, POPs, meteorology, and aerosol scattering coefficient (see chapter 6).

7.2 Raw output files from monitoring instruments

The most common instrument adopted for Hg measurement is the Tekran 2537A. Few stations used the Lumex RA-915. To enable the automatic or manual upload, the structure of output files for each adopted instrument should be known.
Data output template and serial data records of Tekran 2537A contain very detailed information on:

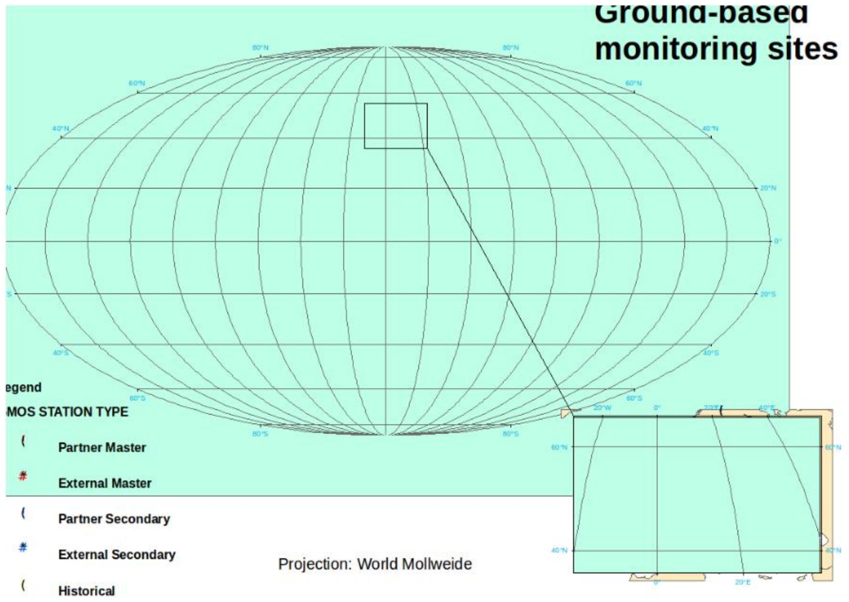

Fig. 7.1. Worldwide distribution of ground-based monitoring sites

- Calibrations
- Instrument settings
- Method parameters
- Serial data output (mercury analysis report)

The Lumex RA-915 raw data are saved in the file "calibration.dat" at the PC embedded in the Lumex mercury monitors. Columns delimiter is Space with the following fields:

- The first column is date: d(d).m(m).year), e.g. 8.10.2012 (day - according UTC)
- The 2-nd column is UTC time: hour:min:sec, e.g., 7:57:4
- The 3-rd column is the UTC time in decimal format: 7.95135 (what is = 7:57:4)
- The 4 6 columns are service codes: -0.726938 1313.45 11.5 (they are not necessary for further data processing)
- The 7-th (the last) column is mercury concentration in ng/m$\hat{3}$, e.g. 1.15766

7.3 Scenarios

GMOS-CI was designed to integrate data coming from different instruments using different network protocols and data formats. In order to comply with

this vision, different strategies are needed to match each GMOS partners requirements.
Each GMOS partner is served by ITC infrastructures with different architectures and capabilities. In some case, monitoring sites are reached by an Internet connections, while in other they are not. GMOS-CI offers two options in order to integrate data:

- **Data Upload** through the GMOS web site, for those stations NOT CONNECTED to Internet (also called Human to Machine, H2M); or
- **Data Integration**, using the System Integration capabilities of GMOS-CI, for those stations CONNECTED to Internet (also called Machine to Machine, M2M).

7.3.1 Data Upload

To upload data a GMOS participant/partner must define a *Contact User* for each station. The Contact User will be provided with credentials to upload data for the selected station.

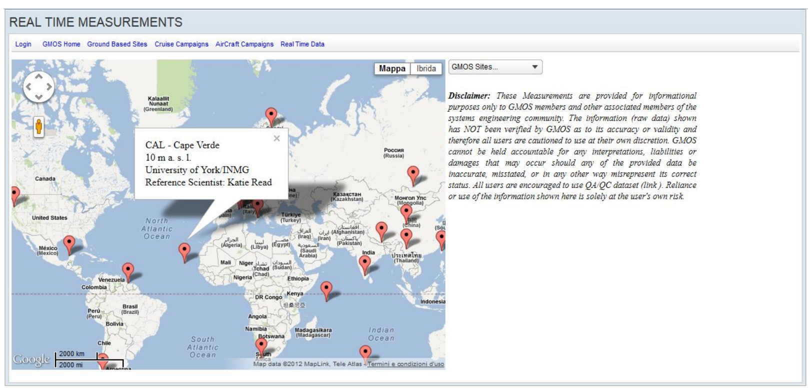

Fig. 7.2. Simple information on GMOS stations

The window REAL TIME MEASUREMENTS shows displacement of GMOS monitoring stations. By clicking on one station very simple metadata information appear (Figure 7.2). To login a user must use the registration e-mail and assigned password (Figure 7.3). The automatic user registration process was not implemented therefore Contact User must send an e-mail to the GMOS-CI administrator to register.

Once logged the Contact User is enabled for the assigned station and the **Upload Data** option appears on the menu (Figure 7.4). A click on this menu voice shows the upload window. The Contact User can upload raw data by using the form showed in the Figure 7.4. A simple upload button opens a selection window by which the Contact User can browse computer folders for

Fig. 7.3. Login window

uploading a single file.
For larger datasets, users can check the asynchronous button for uploading files. The selection of "Async mode" enables the mail box. This e-mail will be used by the system to notify the completion of uploading process to the user. The data format supported by the system is ASCII text file, containing raw data formatted in a CSV format. Data acquisition format is suitable of adaptation following an agreement between GMOS partners and GMOS-CI managers and after compatibility tests.
Data formats and data protocol exchange are described in section 7.4.

Fig. 7.4. The upload mask opened for the Longobucco station

7.3.2 Data Integration

GMOS-CI integrate data coming from IT frameworks provided by partners by means of an Enterprise Service Bus (ESB) Architecture. Stations connected to Internet will use a M2M process. The GMOS-CI will read data from a computer installed in the Contact User side. This computer can be directly connected to an instrument or can be a storage of instrument dataset. All data are directly stored it in the real-time GMOS database.

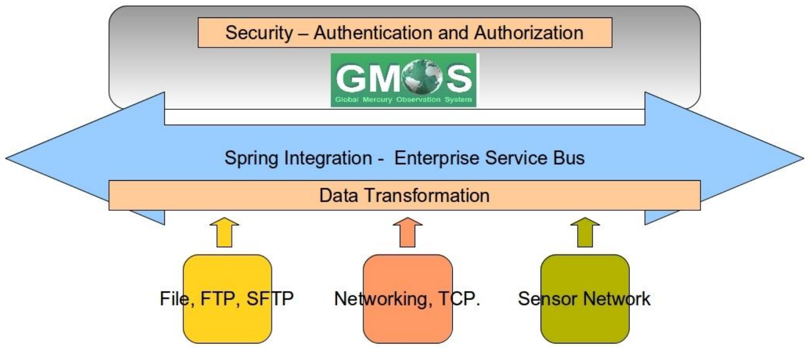

Fig. 7.5. ESB for System Integration

Data sources will be plugged at GMOS-CI by means of the intermediation of ESB (Figure 7.5). Each station can use its own data protocols and data formats. The GMOS-CI will make a data transformation to store data in the managed databases. In order to implements this architecture, partners must provide information on data format and data protocol.
To link a station to the GMOS-CI, protocol and data format should be agreed with GMOS-CI managers. At present the File Transfer Protocol (FTP) has been implemented and Tekran 2537A and Lumex RA-915 output files can be read.
This process needs:

- an FTP server hosted on a computer of the station;
- a folder shared with the GMOS-CI;
- data format of instrument used (Lumex, Tekran..).
- data rate generation.

For security reasons, the computer should have a firewall and a shared folder in read only mode. The GMOC CI will take care of data acquisition, pooling and data storage.

7.4 Data Format and Data Protocol

In order implement both approaches described above, it is important to define both data format and data protocol for Data Integration. GMOS-CI can read

different data formats, in relation to data output for each site. At present, procedures able to acquire data coming from Tekran and Lumex analysers have been developed. As showed in Figure 7.5, the component *Data Transformation* take informations coming from Sensor Network to store them in managed databases. Data Transformation component exports such procedures for Tekran and Lumex. Other procedures can be developed in order to match different requirements of GMOS partners.
GMOS-CI can interact with different data protocols in order to fill the database. FTP, HTTP/Web Services, TCP Sockets and other are the most common protocols that can be adopted.
The Contact User or the Reference Scientist must provide to GMOS-CI managers information on data format and protocol in order to setup the transfer process.

7.5 Data Storage

Datasets collected from GMOS stations are stored in the GMOS-CI managed database. These data are stored as raw and will be treated in the QA/QC to obtain the final database. Raw data are not available to download but the last 200 time series can be visualized if the user is logged in the system.

7.6 Data Visualization

The GMOS web site provides features showing data coming from measurement stations. For each station, raw data are showed by means a web component able to paint time series (Figure 8). These time series are coming from the core databases managed by GMOS-CI, where raw data are stored.
In order to get time series, user must be logged in the GMOS-CI (Figure 7.3) .
The component showing time series, show the last 200 measurements acquired for each stations

7.7 Data Policy

The WP9 in collaboration with WP3 has draft a document on GMOS Data Policy (see chapter 10). The document is based on the specific objectives of GMOS:

- To establish a global observation system for mercury able to provide ambient concentrations and deposition fluxes of mercury species around the world, by combining observations from permanent ground-based stations, and from oceanographic and tropospheric measurement campaigns.

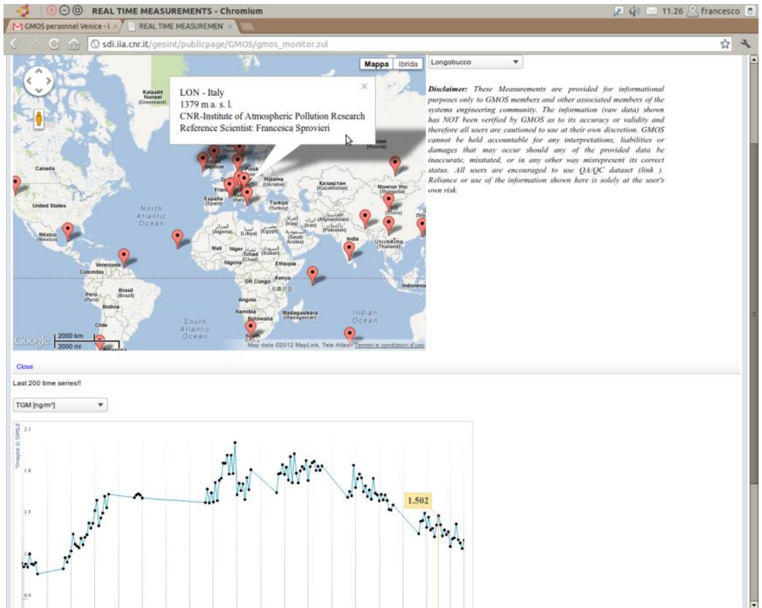

Fig. 7.6. Time series of raw data for Longobucco

- To validate regional and global scale atmospheric mercury modelling systems able to predict the temporal variations and spatial distributions of ambient concentrations of atmospheric mercury, and Hg fluxes to and from terrestrial and aquatic receptors.
- To evaluate and identify source-receptor relationships and their temporal trends for current and projected scenarios of mercury emissions from anthropogenic and natural sources.
- To develop interoperable tools to allow the sharing of observational and models output data produced by GMOS, for the purposes of research and policy development and implementation as well as for enabling societal benefits of Earth observations, including advances in scientific understanding in the nine Societal Benefit Areas (SBA) established in GEOSS.

Data Policy document clearly outline to GMOS partners, data creators, and data users how data will be stored, managed and shared within the framework of the project objectives and beyond. The purpose of the policy is to support the open sharing of new and existing mercury datasets, within an appropriate framework for property rights and due attribution. The Data Policy document is in compliance to GEO Data Sharing Principles.
It is the intention that the resulting GMOS database should:

- be freely shared and distributed following GEOSS Data Sharing Principles, while encouraging co-operation and coherence;
- be global in scale, though implemented nationally and regionally;

- serve to disseminate technological capacity by drawing on and making widely available scientific and technical information; and
- make mercury data universally available, while fully acknowledging the contribution made by those gathering and publishing these data.

Full support of this data policy by GMOS participants and interested parties will support the use of the available datasets for scientific advancement as well as environmental policy development.

8

Implementation of tools for data fusion

8.1 GI-cat clients

GI-cat is an implementation of a broker catalog service made by ESSI-Laboratory of CNR-IIA (http://essi-lab.eu), which allows clients to discover and evaluate geoinformation resources over a federation of data sources. It publishes different catalog interfaces, allowing different clients to use the service.

Fig. 8.1. The GMOS GI-portal EO which allows to query GI-cat by means of the OpenSearch interface

In the GMOS SDI a GI-cat instance was activated and the brokering access to the SDI can be used by means of two lightweight clients: GI-portal is a lightweight client which allows to query GI-cat by means of the OpenSearch interface (http://sdi.iia.cnr.it/gicat/gi-portal/) (Figure 8.1).

GI-go geobrowser is an extended GI-cat client written in Java which allows complex queries and enhanced results presentation (http://essi-lab.eu/gi-go) (Figure 8.2).

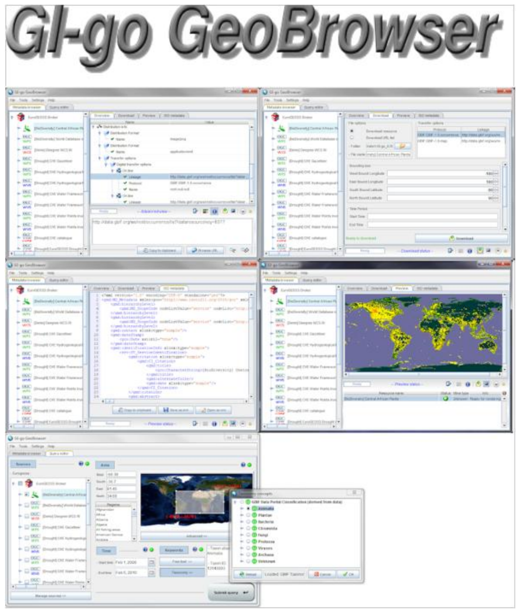

Fig. 8.2. The GMOS GI-go geobrowser, an extended GI-cat client written in Java.

8.1.1 Published interfaces

The following catalog interfaces are available under GI-cat:

- CSW/ISO 2.0.2 interface
- GI-CAT interface
- OPENSEARCH interface

8.1.2 Federated resources

Searches can be performed against the following resources:

- GN@CNR-IIA of type GEONETWORK version 2.4 (CSW/ISO 2.0.2)
- Endpoint: http://sdi.iia.cnr.it/geonetwork/srv/csw?

8.1.3 GI-cat interface

GI-cat features a custom SOAP interface that extends the standard OGC CSW specification and enables the following additional capabilities:

- Configuration: to dynamically configure the GI-cat instance
- Feedback: to retrieve feedbacks status information from the GI-cat instance
- Incremental queries: to obtain incremental query responses from the GI-cat instance

The list of available services and related WSDLs are reported in Table 8.1.

Table 8.1. List of SOAP Services

Available SOAP services	**Methods**
Catalog	GetResource Cancel Close GetContent GetAccessParameters GetData ClearCache Status Query
HTTPProvider	-
Manager	GetBrokerCapabilities RemoveConfiguration InsertConfiguration Harvest ActivateConfiguration Transaction ListConfigurations
OpenSearchService	-
SOAP1_2Provider	-

9

Implementation of GMOS Portal

9.1 Public portal services

The public face of the portal does not require any log-on information. It provides news and information about the GMOS as well as links to services and sites of interest to the research community.
The most interesting public service is the measurement activity viewer, which endpoint is http://www.webgis.iia.cnr.it/gmos/ Measurement programs (ground-based, over-water and aircraft) are synthetically described in the relative sections while geographical displacement of measurement locations can be found in the section Browse measurements (Figure 9.1) through a WebGIS. This very simple viewer must not be confused with the more complex SDI.

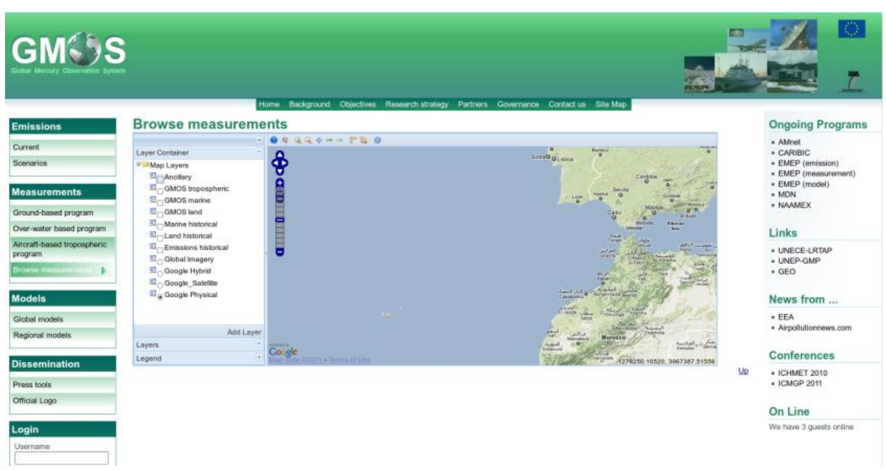

Fig. 9.1. The viewer of GMOS data

The scope of this viewer is to show historical and current displacement of monitoring station and sampling points of oceanographic cruise and atmospheric measurements. Under this page, measurement are not shown, whereas metadata can be viewed to get information on scientists collecting information and to whom require permissions for accessing data.

9.2 Private portal services

The portal provides also a secure business environment where historical and current data are stored.
Measurement programs (ground-based, over-water and aircraft) can be accessed and downloaded after registering (Figure 9.2).

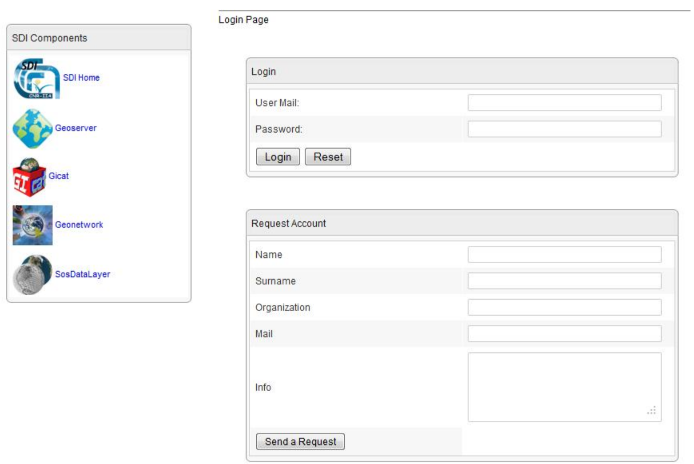

Fig. 9.2. The form to require private access to GMOS protected section

After accessing, a user can browse monitoring stations as well as measurement campaigns (Figure 9.3).

The example shows historical data by year for available stations. A user has several browsing options:

- Directly on the map by clicking on a station;
- By selecting a station through dedicated buttons;
- By selecting a station reported in the table.

Once selected, dataset can be downloaded or plotted (Figure 9.4). Different options can help customization of data plotting.

Where available ancillary information (i.e. meteorological and chemical data) are reported in a different table.

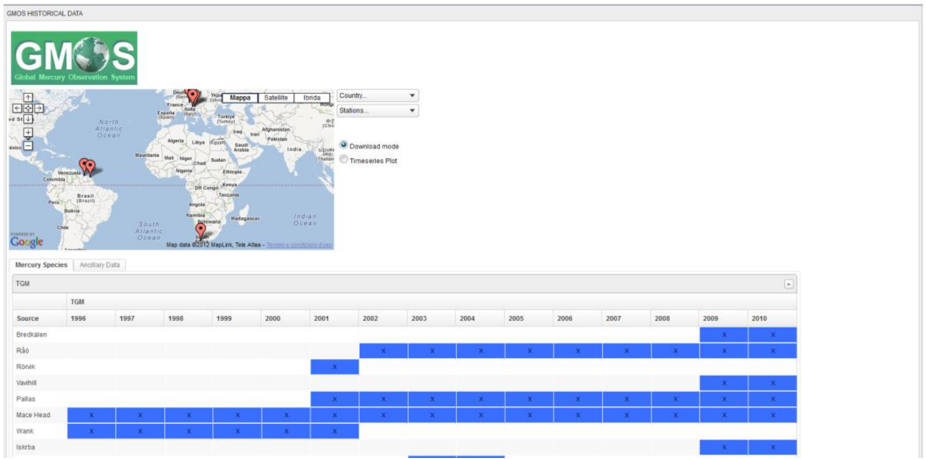

Fig. 9.3. The page for browsing and downloading data

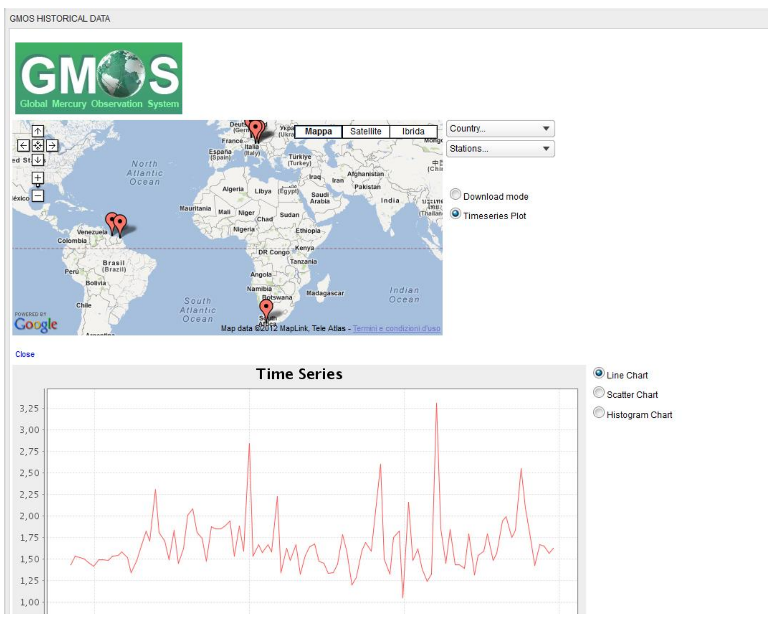

Fig. 9.4. Time-series plot of data

9.3 Download format

Data downloaded follow the schema reported in Table 9.1. The following fields are reported within each file:

- gid: measure number
- datasource: name of the station/campaign
- elev: station/point elevation/depth
- value: measurement value
- uom: unit of measure
- timestamp: time in the format ddmmyy hh.mm
- lat: latitude of the sampling site/point;
- lon: longitude of the sampling station /point.

Table 9.1. Schema of download data

gid	datasource	elev	value	uom	timestamp	lat	lon
1	ZEP	550	-9999	ng/m^3	1-1-00 1.00	78.9	11.88
2	ZEP	550	-9999	ng/m^3	1-1-00 2.00	78.9	11.88
3	ZEP	550	-9999	ng/m^3	1-1-00 3.00	78.9	11.88
4	ZEP	550	-9999	ng/m^3	1-1-00 4.00	78.9	11.88
5	ZEP	550	-9999	ng/m^3	1-1-00 5.00	78.9	11.88
6	ZEP	550	-9999	ng/m^3	1-1-00 6.00	78.9	11.88
7	ZEP	550	-9999	ng/m^3	1-1-00 7.00	78.9	11.88
8	ZEP	550	-9999	ng/m^3	1-1-00 8.00	78.9	11.88

9.4 Additional services

At present Work Package 9 (WP9) is developing additional web services to make available data coming from model outputs.

10

GMOS Data Policy

GMOS has developed the present Data Policy document to clearly outline to GMOS partners, data creators, and data users how data will be stored, managed and shared within the framework of the project objectives and beyond. The purpose of this policy is to support the open sharing of new and existing mercury datasets, within an appropriate framework for property rights and due attribution. This Data Policy document is in compliance to GEO Data Sharing Principles (see section 10.1).

It is the intention that the resulting GMOS database should:

1. be freely shared and distributed following GEOSS Data Sharing Principles, while encouraging co-operation and coherence;
2. be global in scale, though implemented nationally and regionally;
3. serve to disseminate technological capacity by drawing on and making widely available scientific and technical information; and
4. make mercury data universally available, while fully acknowledging the contribution made by those gathering and publishing these data.

Full support of this data policy by GMOS participants and interested parties will support the use of the available datasets for scientific advancement as well as environmental policy development.

10.1 GEOSS Data Sharing Principles

The data sharing methodologies implemented by GMOS have been designed within the framework of the GEOSS Data Sharing Principles. The methodologies and data sharing processes described in this Data Policy document are compliant with the Data Sharing Action Plan and Data-CORE initiative within GEOSS project.
Datasets collected within the GMOS project will be integrated into the GEOSS Common Infrastructure (GCI) as part of GMOS Data-CORE.

The GEOSS Data Sharing Principles use the term "full and open exchange" of data, metadata and products as the guiding strategy. Data, metadata and products will be provided, according to GEOSS Data Sharing principles, with minimal time delay compatible with rules established in section 10.4 and with as few restrictions as possible, on a non-discriminatory basis, at minimum cost for no more than the cost of reproduction and distribution.

10.2 Definitions of Terms

- Ground-based station A fixed-location atmospheric monitoring site where in situ measurements of mercury in ambient air and/or precipitation are collected.
- Oceanographic cruise campaign A period (or multiple periods) of time dedicated to collecting over-water, air-sea exchange, and other related measurements on-board an oceanographic research vessel.
- Aircraft campaign - A period (or multiple periods) of time dedicated to collecting measurements of ambient air mercury and other atmospheric compounds at high altitude (e.g. free troposphere) on-board an air plane.
- Observation data Measurements of environmental or atmospheric parameters taken at a ground-based monitoring site, on-board a research cruise, or on-board a research aircraft.
- Data point An individual measurement of an environmental or atmospheric parameter
- Dataset A group of data points collected over a period of time at a measurement location or campaign.
- GMOS site GMOS monitoring sites are divided into Master and Secondary sites. Master sites are those where measurements of speciated ambient air mercury and mercury in precipitation are collected; Secondary sites are those where measurements of total gaseous mercury in ambient air and mercury in precipitation are collected.
- Quality Assurance and Quality Control (QA/QC) The process by which a dataset is closely and systematically examined to ensure that it was collected and/or generated according to acceptable standard procedures, and that dataset is consistent, within acceptable limits of precision and accuracy, and comparable to related datasets.
- Historical measurement data Observational mercury data from ground-based sites, oceanographic cruise campaigns, or aircraft campaigns collected prior to the official start of the GMOS project. GMOS partners or external partners may offer to contribute historical data for inclusion in the GMOS database.
- Internal GMOS site measurement data Observational data collected from a monitoring site or campaign by a GMOS partner (a researcher or research group who is receiving funding from the GMOS project to support the collection of mercury measurements).

- External GMOS site measurement data (e.g., third-party data) Observational data collected from a monitoring site or campaign by an external GMOS partner (e.g. a research group or network who is not eligible to receive financial support through GMOS but has agreed to share their data with the GMOS project and measurement database)
- Model output Information produced by the regional and global scale models that are being run in support of the GMOS project
- Data creator A researcher or research group who collects measurement data or produces model output, and agrees to provide this data to the GMOS database.
- Data user An interested party who wishes to access current or historical data through the GMOS database and utilize that data for informational or analytical purposes.
- Metadata Information about measurement or model datasets, including the type of data, the location where it was collected, date and time, units of measurement, precision and accuracy of data, measurement or modeling methods, and details about data ownership.
- Database A database is an organized collection of data stored in digital form and logically structured in order to improve data access and data retrieval.
- GMOS SDI - An Spatial Data Infrastructure(SDI) is a coordinated series of agreements on technology standards, institutional arrangements, and policies that enable the discovery and use of geospatial information by users and for purposes other than those it was created for. GMOS SDI is the SDI developed in order to manage geospatial data for the GMOS project.
- Interoperable system A system is interoperable if is able to interact with different systems or services in order to perform common task.
- Cyber Infrastructure A well organized collection of services and processes, often distributed over the internet or other net infrastructures, in order to provide users with a set of features able to match some requirements.
- GEOSS The Global Earth Observation System of Systems, designed to link existing observational systems worldwide and support the development of new systems, with the intention of providing useful content and information in support of policy decision-making.
- GEO The Group on Earth Observations is a voluntary partnership of governments and international organizations seeking to coordinate efforts in support of developing GEOSS.
- Open Geographical Consortium - The Open Geospatial Consortium (OGC) is an international industry consortium of companies, government agencies and universities participating in a consensus process to develop publicly available interface standards. OGC Standards support interoperable solutions that "geo-enable" the Web, wireless and location-based services and mainstream IT.

- Web services A set of operation, often stateless, distributed over the Web, using protocols and methodologies proper of the Web.
- File Transfer Protocol (FTP) The FTP is a standard network protocol used to transfer files from one host or to another host over the Internet.

10.3 The GMOS Spatial Data Infrastructure

The GMOS sites are free to quality assure / quality control (QA/QC) and use their data in their own manner, but all GMOS data should be processed by GMOS project managers in the same way. The most important process related to data management within the GMOS project is the transfer of collected data to a central system. As such, GMOS will provide a cyber Infrastructure to all the partners which will allow the sharing of:

- information and data from historical databases,
- measurements collected at GMOS ground-based sites and measurement campaigns,
- model output,
- Metadata for each dataset.

The development of the Cyber Infrastructure system will consider a range of data formats given that data will be provided from in-situ or mobile sensors, from oceanographic or aircraft measurement campaigns, or from numeric models. The GMOS Cyber Infrastructure serves the GMOS Spatial Data Infrastructure (SDI) in order to:

- Archive data collected from historical measurements/campaigns and from literature;
- Archive data collected from on-going measurements/campaigns;
- Archive data from run of simulation models;
- Archive metadata for archived information;
- Make available data and metadata for download in interoperable formats;
- Make available web services related to search, discovery, access, preview and export.

In the SDI there are different components in order to perform different tasks. Data will be stored in a Data Storage Layer and managed by the Cyber Infrastructure, which will serve as an integration system for data coming from GMOS partners activities.
Data storage will occur:

- By directly uploading dataset. GMOS site administrators or measurement campaign responsible will have an account in the Cyber Infrastructure to manually upload the data (Human to Machine process, H2M);

- By connecting the Cyber Infrastructure to a computer linked to instruments and to internet. The Cyber Infrastructure will download data at least on a daily basis.

Data stored in the Data Storage Layer and managed by the Cyber Infrastructure will be provided to users by means of different devices contained in an Application Layer.
Each device represents a different view of the data managed by the Cyber Infrastructure. The GMOS website can be used as device where data will be provided to users as simple link for data link, or by a Web visualization system to visualize information. Additional devices will be oriented to a machine access, such as Web Services OGC compliant. Access to GMOS Cyber Infrastructure for data search, discovery and visualization can occur also by the GEO web portal.

10.4 Policy for Data Usage

10.4.1 General conditions

This section of the GMOS data policy provides information to those who wish to access and use data that is held within the GMOS database. GMOS will make data available freely, without charge for the data itself, to all interested parties. Data will also be provided without restriction or discrimination against any individual or research group. Metadata will be provided freely without any conditions. Access to data and other products will require registration, acceptance of conditions, and acceptance of the license agreement.

10.4.2 Restrictions to access

In general, all data provided by the GMOS Cyber Infrastructure can be accessed by any interested party. However, the following conditions apply:

- Embargo period: Given the need for researchers to collect, analyze, and publish their own research before their dataset can be made publicly available, it is necessary to restrict access to data until the GMOS participants have had time to do so. GMOS will impose a two year embargo period from the time data has been collected before it can be publicly available through the GMOS Cyber Infrastructure. The duration of the embargo period can be made shorter if the researcher or group responsible for the data agrees to this condition.
- Ongoing long-term data sets: An exception to the embargo period is ongoing, long-term measurement datasets. Earlier data may be made publicly available, whereas more recent data may still be covered under the embargo period. The individual researcher or group responsible for the data may provide specific requests for such datasets in the interest of allowing them

to publish findings on long-term measurement trends before making data publicly available.
- Special cases: According to the GEOSS Data Sharing Implementation guide, low resolution data, images, and maps will be provided to the end-user as they have been uploaded in the GMOS database for dissemination purposes, research and education.

10.4.3 Community institutions access to data

GMOS information may be relevant for the INSPIRE themes of "Human health and safety" and "Environmental monitoring facilities". National and international institutions and bodies may be licensed to the GMOS high resolution data when they are relevant for human health and safety and when they are related to emergency management. The European Guidance on the 'Regulation on access to spatial data sets and services of the Member States by Community institutions and bodies under harmonised conditions' is adopted to access the GMOS database and services.

10.4.4 License agreement

All data provided through the GMOS Cyber Infrastructure will be accompanied by a user license agreement. This is an agreement between the data user and GMOS, outlining any limitations on how the data may be used, how the user must acknowledge the source of the data, and the limits of liability for the data that GMOS provides. GMOS does not accept liability for the users interpretation of data. Correct and scientifically sound interpretation of data is solely the responsibility of the user.
In the license agreement, users will be requested to not supply GMOS data to third-party users. Those interested users should contact GMOS directly to access data contained within the GMOS database. This is important for ensuring that the most recent versions of datasets are being utilized by interested users, and that all interested users can have direct contact with the GMOS database managers as well as the creators of the data.
If a user encounters any problems or errors in a GMOS-provided dataset, they will be requested to inform GMOS in a timely manner so that any necessary corrections can be applied to the dataset for future use. These collaboration is important for ensuring the production and use of high-quality datasets.

10.4.5 Acknowledgement

Anyone who uses data supplied by GMOS must acknowledge in any publication or report the contribution of the data creators. The GEO Data Citation standards will be disseminated through the GMOS SDI to help end user to proper cite and acknowledge data providers/creators.

10.4.6 How to access data

Access to GMOS data will be made available to:

- Authenticated users that can download/visualize data for a specific dataset and visualize time series by means of their personal account.
- Anonymous users that can access only metadata and visualize the coordinates of measurement sites.

Data provided includes historical and current: emission estimates, ground-based sites measurements; oceanographic and atmospheric measurements.
Data can be downloaded as Comma-Separated Values (CSV) format, which is an ASCII file, human readable format. Additional OGC compliant data format will be available through OGC Web Services (OWS), namely Web Feature Service (WFS).
Metadata can be downloaded as eXtensible Markup Language (XML) format, using the ISO19139 data format.

10.5 Policy for Data Contributions

10.5.1 Data sources

Data sources that can be made available through the GMOS Cyber Infrastructure may include:

- Historical or recent measurements from ground-based monitoring sites
- Historical or recent measurements from oceanographic cruise campaigns
- Historical or recent measurements from aircraft campaigns
- Metadata from monitoring sites and/or campaigns
- Model output

Data sources submitted to the GMOS Cyber Infrastructure should clearly specify any restrictions that should be placed on the distribution or availability of data.

10.5.2 Data submission

Submission of data to the GMOS portal can be accomplished in several ways depending on the type of data and/or monitoring location. Historical data and data from individual monitoring campaigns (e.g. oceanographic cruises, aircraft measurements) can be sent by e-mail or uploaded by FTP to the Cyber Infrastructure. Continuous monitoring data from ground-based sites can be supplied in a variety of ways. If the site has a continuous internet connection, then the GMOS Cyber Infrastructure can obtain data directly from the monitoring site. Alternatively, data can also be submitted by means

of GMOS Cyber Infrastructure. Each GMOS site manager has an account to upload data. This data will be managed as coming from monitoring site.
Model output can be submitted by sending an e-mail or uploading by FTP to the Cyber Infrastructure. In order to share data information across to GMOS web site, metadata information for each dataset must be sent by e-mail or uploaded by FTP to the Cyber Infrastructure.
No data will be accepted without providing the companion metadata.

10.5.3 Data formats

Measurement data can be submitted using a simple CSV format, where the first row defines the header information for each column. Each dataset, or group of datasets, will be integrated with the metadata information using the ISO19139 format.
In certain cases, it may not be possible to use supply data in CSV format. For example, this is the case for certain types of model output. In these special cases, that data provider should contact the GMOS data manager in order to define an acceptable data format for submission. In all cases, metadata related to the uploaded dataset is always required.

10.5.4 Data storage

Data are stored in Relational Data Base Management System (RDBMS). All databases used to store data are collected in a logical Data Storage Layer, hosted in the GMOS Cyber Infrastructure. The servers in the Cyber Infrastructure are firewalled with the exception for local access by components inside the Application Layer in the GMOS Cyber Infrastructure. The GMOS data is accessible only trough the GMOS Cyber Infrastructure.

10.5.5 Data availability

Data submitted to the GMOS Cyber Infrastructure are the intellectual property of the data creator. Digital data will be made available up-on request of GMOS website users and after an agreement with the data creator. Real-time data will not be accessible for a user to download, but may be graphically displayed from the past 24 hours for the purposes of displaying the interesting work going on within GMOS.

10.5.6 QA/QC

All researchers are free to QA/QC their own data before publication of their own results and findings. Additionally, within the GMOS network, a unique and separate QA/QC methodology must be applied to the GMOS datasets in order to achieve the goals of data consistency and comparability. For this reason, when submitting data to the GMOS portal, it is requested that data be

submitted as close to the raw form as possible so that the GMOS data managers can apply a consistent QA/QC procedure to all submitted data. The GMOS-defined QA/QC procedures will be applied to all datasets produced within the GMOS project. For each type of on-going measurement within GMOS (ground-based, cruise and aircraft campaign measurements), ad-hoc QA/QC procedures will be developed. After submission, the datasets will be stored in two different databases: one consisting of original raw data without any post-processing, and the other one with data resulting from GMOS QA/QC procedures.
To apply the appropriate QA/QC procedure to GMOS measurements, GMOS site operators and GMOS campaign participants are requested to compile and upload the corresponding site or campaign logbook. Specifically for the GMOS ground-based sites, a site log form will be developed and distributed for the site operators to fill out during each visit.
Historical data, produced before the GMOS project and without the adoption of the same monitoring criteria (i.e., following GMOS Standard Operating Procedures), will not undergo the QA/QC process by GMOS data managers. In this case, when submitting historical data to the GMOS portal, a prior validation from the data owner it is requested along with indications about the specific QA/QC methodology that was applied. Additionally, as an added value, the GMOS data managers will conduct a necessary data harmonization in such a way that all historical mercury datasets will be available within the GMOS portal in a standardized and consistent format.

10.5.7 Data cancellation policy

Users can update or remove their own previously uploaded datasets only by sending an e-mail request to the GMOS Cyber Infrastructure manager.

10.5.8 Embargo periods

Before making data publicly available, GMOS will allow individual researchers and project partners adequate time to assemble, analyze, and publish their own findings. GMOS considers two years to be a reasonable embargo period for researchers to accomplish these tasks, after which point the data should be made publicly available through the GMOS portal. However, in certain cases, a longer or shorter embargo period may be necessary as long as it can be justified by the research coordinator for the project. Changes to the embargo period should be agreed upon in advance with the GMOS project coordinator. During the embargo period, data submitted to the GMOS portal will remain restricted and will not be publicly available to anyone. During this period, as researchers are compiling and analyzing datasets, updates or modifications can be made to the dataset submitted to the portal.
The embargo period begins from the end date of data collection (such as the end of a monitoring campaign, or the point at which data has been gathered

from an instrument). Once the embargo period has passed, the data will be made available through the GMOS portal to anyone who requests it and signs the corresponding license agreement unless there is a legitimate reason for denying access to a user. An exception to the embargo period is continuous measurements from ground-based monitoring sites. In these cases, the most recent two years of data shall remain under the embargo unless a different duration of the embargo period has been approved by the GMOS project coordinator.

10.5.9 Intellectual property rights

Nothing in this document should be read to alter the scope and application of Intellectual Property Rights and benefit sharing agreements as determined under relevant laws, regulations and international agreements of the Participants. Intellectual Property Rights in data that a researcher generates depends on where that researcher works and their contract of employment. The rights owner for any data submitted to the GMOS portal will be required to grant GMOS a non-exclusive licence to allow GMOS to mange and supply the data for use.
To the greatest extent possible, the GMOS database is an open-access facility, and all users should have equal access to data in databases affiliated with or developed by GMOS. As such, GMOS will promote the free dissemination of mercury data and, in particular:

- will not assert any proprietary rights to the data in databases that are developed by other organisations and that subsequently become affiliated with GMOS;
- will seek, to the greatest extent possible, to make freely and openly available, with the least possible restrictions on reuse, any data commissioned, created or developed directly by GMOS; and
- will respect conditions set by data publishers that affiliate their databases to the GMOS portal. GMOS will seeks to ensure that the publisher/holder of data is acknowledged and requests that such attribution be maintained in any subsequent use of the data.

It will be a condition of access to and use of the accessible data that users acknowledge that the validity of the data in any databases affiliated with the GMOS database cannot be assured. GMOS does not claim responsibility for the accuracy and reliability of the data as well as for the suitability of its application for any particular purpose. GMOS may claim appropriate Intellectual Property Rights available within applicable national jurisdictions over any tools, such as search engines or other software products that are developed by GMOS while carrying out the GMOS Work Programme.

10.5.10 Acknowledgement

All data supplied by GMOS to a user will include a data license agreement, specifying that data users must acknowledge the creator of the data in any publication or report.

11

Conclusion

During the last decade, air pollution in the Mediterranean region has received considerable attention due to the high level of industrialization and dense population around the coastal zone of the Mediterranean Basin. Several research studies have shown that the Mediterranean atmosphere is influenced by air masses originating from Central-Europe and North Africa, but also from Eastern Europe and sometimes from south-east Asia. For this reason the Mediterranean has been described as an air pollution crossroad.
This example, with others around the world, are very effective in order to describe the importance of geo data management and the importance of tools able to analyse and export environmental data to public decisors and researchers.
When developing tools and scenarios for air quality information management, some of the most limiting factors are discrepancies among monitoring systems, particularly in terms of spatial and temporal differences and data availability. Fostered by European directives (e.g. INSPIRE, Air Quality Directive) and international programs (i.e. GEOSS) that have pushed the scientific community toward the development of advanced interoperable systems, CNR-IIA created a framework based on open-source tools compliant with standards and oriented towards an integrated system that facilitates data management. These interoperable systems can be helpful in assuring real time data analysis and dissemination within the scientific community, as well as offering the information to other stakeholders and policymakers.
This work has suggested a solution to design and implement a complex Spatial Data Infrastructure based on open-source components. It describes the development of an Information Infrastructure to provide end-users with a friendly interface for data creation and harmonization. This framework manages the SDI data through DAOs and enables service-creation through RESTfull endpoints.
The Air Quality SDI holds information on concentration of contaminants measured at permanent sites, as well as in dedicated monitoring campaigns. It also incorporates meteorological parameters and outputs of meteorological mod-

els. This SDI can support both modelling activities and environmental assessments for different case studies that aim to evaluate the impact of atmospheric pollution on terrestrial and aquatic ecosystems, and on human health.
The framework developed, GeoInt, is a middleware between Data Storage and Services Layer. GeoInt exports common services to the applications layers and upon this framework is possible to plug-in different project and application oriented to geo data management.
The tools developed de-couple storage systems from the software layers, helping to manage both vector and raster information, though the latter is still under development and will be provided in the next GeoInt version.
Additional work will focus on improving GeoInt processes and data management. GeoInt will then be able to export an ISO19139 document by means of CS-W 2.0.2 without any interaction with others tools. This will allow us to remove a complex component from the SDI (and it will address a related issue regarding upgrading and maintenance). Additional areas of focus will include improving raster data.
GeoInt is defined as a Geo Data Factory for SDI. This middleware is currently used in the Global Mercury Observation System (GMOS) in order to build the sharing infrastructure about data providers and data users around the world. The SDI holds the global observation system for mercury, including ground-based stations at high altitude and sea level locations, ad-hoc oceanographic cruises over the Pacific, the Atlantic and the Mediterranean, and free tropospheric mercury measurements. This work done along the project, will finally support the Gourp on Earth Observations (GEO) 2012-2015 work plan and the task HE-02-C1, "Global Mercury Observation System".

References

1. K. Katsouyanni, D. Zmirou, C. Spix, J. Sunyer, J.P. Schouten, A. Ponka, H.R. Anderson, Y. Le Moullec, B. Wojtyniak, M.A. Vigotti, Short-term effects of air pollution on health: a European approach using epidemiological time-series data. The APHEA protocol. J Epidemiol Community Health. 1996 April; 50(Suppl 1): S12-S18W.-K. Chen, Linear Networks and Systems (Book style). Belmont, CA: Wadsworth, 1993, pp. 123135.
2. A. Seaton, D. Godden, W. MacNee, K. Donaldson. Particulate air pollution and acute health effects. Lancet, 345, 176-178, 1995
3. J. Roca, E. Padilla, M. Farr, V. Galletto. Economic growth and atmospheric pollution in Spain: discussing the environmental Kuznets curve hypothesis. Ecological Economics. 39 pp 8599.2001
4. O. Gunther. Environmental information systems. ACM SIGMOD Record, Volume 26 Issue 1, March 1997 ACM New York, NY, USA New York.
5. C. Gouveia, A. Fonseca, A. Cmara, F. Ferreira.Promoting the use of environmental data collected by concerned citizens through information and communication technologies. Journal of Environmental Management. 2004 Jun;71(2):135-54.
6. GSDI.Developing Spatial Data Infrastructures: the SDI Cookbook. Editor: Douglas D. Nebert, Technical Working Group Chair, 2004.
7. CGDI.Canadian Geospatial Data Infrastructure. Online at: http://www.geoconnections.org/en/aboutcgdi.html. 2009.
8. EC. Directive 2007/2/EC establishing an Infrastructure for Spatial Information in the European Community (INSPIRE). Online at: http://eur-lex.europa.eu/en/index.htm. 2007.
9. EC,. Directive 2003/4/EC on public access to environmental information. Online at: http://eur-lex.europa.eu/en/index.htm. 2003.
10. GEOSS. The Global Earth Observation System of Systems (GEOSS) 10-Year Implementation Plan. Online at: http://www.earthobservations.org. 2009.
11. D. Vandenbroucke, P. Beusen. Spatial Data Infrastructures in Italy: State of play 2007. SPATIAL APPLICATIONS DIVISION, K.U.LEUVEN RESEARCH & DEVELOPMENT, Celestijnenlaan 200 E, B-3001 LEUVEN. 2007
12. S. Nativi. 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.

13. M. Botts, G. Percivall, C. Reed. J. Davidson. OGC Sensor Web Enablement: Overview And High Level Architecture.OGC White Paper, pp 07-165. 2007.
14. P. Schut, A. Whiteside. OpenGIS Web Processing Service, OGC 05-007r4. 2005.
15. N. Pirrone, K.R. Mahaffey. Where We Stand on Mercury Pollution and its Health Effects on Regional and Global Scales.In: Dynamics of Mercury Pollution on Regional and Global Scale (Pirrone N. And Mahaffey K. Eds), Springer Verlag Publischers, Norwell, MA, USA, Chapter 1, pp 1-21. 2005.
16. Z. Nedovi-Budi, N.R. Budhathoki. Technological and Institutional Interdependences and SDI The Bermuda Square? International Journal of Spatial Data Infrastructures Research. 1: 36-50. 2006.
17. A. Naumenko, A. Wegmann, G. Genilloud, W.F. Frank. Proposal for a formal foundation of RM-ODP concepts. Proceedings of ICEIS 2001, WOODPECKER2001, J. Cordeiro, H. Kilov (Eds.), Setbal, Portugal, July 2001.
18. P. Plini, F. Mazzocchi. EARTh Environmental Applications Reference Thesaurus. A proposal for a new environmental thesaurus model, UNEP Thesaurus/Terminology Workshop. Ginevra, 14-15-IV-2004
19. D'Amore S. Cinnirella, N. Pirrone. Architecture solution for air quality management through a dedicated SDI. Geophysical Research Abstracts, 13, EGU2011-7113
20. R.T. Fielding,. Architectural Styles and the Design of Network-based Software Architectures. PhD dissertation in Information and Computer Science, University of California, Irvine. Online at: http://www.ics.uci.edu/ fielding/pubs/dissertation/top.htm. 2000
21. 52North. SOS Interface implementation. Online at: http://52north.org/communities/sensorweb/sos/index.html. 2012
22. SPRING. Application development framework for enterprise Java. Online at: http://www.springsource.org/. 2012.
23. SKOS, Common data model for sharing and linking knowledge organization systems via the Semantic Web. Online at: http://www.w3.org/TR/skos-reference/. 2012.
24. ZK. Web Framework, Online at: http://www.zkoss.org/. 2012.
25. GWT. Web Framework, Online at: http://code.google.com/webtoolkit/. 2012.
26. OL. Web Framework, Online at: http://openlayers.org/. 2012
27. RM-ODP. Recommendations and international standards defines essential concepts necessary to specify open distributed processing systems. Online at: http://www.rm-odp.net/. 2012
28. Ponka, A., Anderson, H., Le Moullec, Y., Wojtyniak, B., Vigotti, M. and et, a., 1995. Short-term effects of air pollution on health: a european approach using epidemiological time-series data. the aphea project: background, objectives, design. Eur Respir J 8(6), pp. 10301038.
29. 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, 5 (6).
30. Cinnirella S., D'Amore F., Pirrone N., 2010. An Air Quality Spatial Data Infrastructure for Environmental Management. In Proceedings of the International Symposium on Geo-information for Disaster Management (Gi4DM). Torino, February 2-4.

31. Cinnirella S., D'Amore F., Mazzetti P., Nativi S., Pirrone N., 2012. A Spatial Data Infrastructure for the Global Mercury Observation System. In Proceedings of The 16th International Conference on Heavy Metals in the Environment (ICHMET). Rome, Italy 23-27 September.
32. D'Amore F., Cinnirella S., Pirrone N., 2012. Data and metadata management automation for an effective approach to sharing environmental data. In Proceedings of The 16th International Conference on Heavy Metals in the Environment (ICHMET). Rome, Italy 23-27 September.
33. Cinnirella S., D'Amore F., Pirrone N. 2009. GIIDA-Air Quality Thematic Area, Term of Reference report. Technical report CNR/IIA/2009/XX, Rende, Italy: 4 pp
34. Cinnirella S., D'Amore F.,D9.1 - GMOS ICT System. Architecture and Data Integration . On Line: www.gmos.eu
35. Cinnirella S., D'Amore F.,D9.2 - Report on definition of metadata, standards with respect to INSPIRE. On Line: www.gmos.eu
36. Cinnirella S., D'Amore F., D9.3 - Servers installed and databases for observations, emissions and model results set-up. On Line: www.gmos.eu
37. Cinnirella S., D'Amore F., D9.4 - Implementation of communication and networking for data retrievall. On Line: www.gmos.eu
38. Cinnirella S., D'Amore F., D9.5 - Implementation of tools for data fusion. On Line: www.gmos.eu
39. Cinnirella S., D'Amore F., D9.6 - Implementation of GMOS portal. On Line: www.gmos.eu