EnviroGRIDS interoperability guideline

Giuliani, Gregory; Ray, Nicolas; Lehmann, Anthony; Charvat, Karel

D2.1: EnviroGRIDS interoperability guideline

Title	:	EnviroGRIDS interoperability guideline
Creator	:	Gregory Giuliani (UniGE/GRID)
Creation date	:	14.08.2009
Date of last revision	:	26.10.2009
Subject	:	Interoperability, Spatial Data Infrastructure, Data sharing, Standards
Status	:	Validated
Type	:	Word document
Description	:	Guideline presenting key concepts about Spatial Data Infrastructure, Interoperability, OGC/ISO standards and related tools, allowing partners to share their data in an OGC compliant way
Contributor(s)	:	Gregory Giuliani, Nicolas Ray, Anthony Lehmann, Karel Charvat
Rights	:	Public
Identifier	:	enviroGRIDS_D21.doc
Language	:	English
Relation	:	EnviroGRIDS data storage guideline (D2.2), EnviroGRIDS sensor data use and integration guideline (D2.3), EnviroGRIDS remote sensing data use and integration guideline (D2.4), Grid infrastructure sustainability for environmental sciences (D2.5)

EXECUTIVE SUMMARY

The aim of the EU FP7 EnviroGRIDS project is to build capacities in the Black Sea region to use new international standards to gather, store, distribute, analyze, visualize and disseminate crucial information on past, present and future states of this region, in order to assess its sustainability and vulnerability. To achieve its objectives, EnviroGRIDS will build a Grid-enabled Spatial Data Infrastructure (GSDI) becoming one of the integral systems in the Global Earth Observation System of Systems (GEOSS), and compatible with the new EU directive on Infrastructure for Spatial Information in the European Union (INSPIRE), as well as UNSDI developments.

The scientific aim of the EnviroGRIDS @ Black Sea Catchment project is to start building an Observation System that will address several GEO Societal Benefit Areas within a changing climate framework. This system will incorporate a shared information system that operates on the boundary of scientific/technical partners, stakeholders and the public.

It will contain an early warning system able to inform in advance decision-makers and the public about risks to human health, biodiversity and ecosystems integrity, agriculture production or energy supply caused by climatic, demographic and land cover changes on a 50-year time horizon.

To achieve and support the EnviroGRIDS' vision and objectives, the work package 2 will focus its attention on the creation of a Grid enabled Spatial Data Infrastructure (GSDI) so that the data necessary for the assessment of GEO Societal Benefit Areas, as well as the data produced within the project can be gathered and stored in an organized form and accessible in an interoperable way on the Grid infrastructure and distributed across the Grid in order to provide a high performance and reliable access through standardized interfaces.

The present deliverable tries to give a good overview as complete as possible of the basic concepts concerning Spatial Data Infrastructures and related topics, with a special focus on interoperability and standards that will be used in the context of EnviroGRIDS. The aim of the document is to give the necessary knowledge and to assist the partners of the project to share and integrate geographical data and information in a more efficient way.

One of the challenges we are facing today is to make sense of the vast amount of data and information we continuously receive and access in order to turn them into understandable information. Most of these data and information products have a reference or are linked to a position on the Earth' surface, and we say that they are spatially referenced or georeferenced.
The Agenda 21 fosters the importance of geospatial information to support decision-making and management on the degradation and threats that are affecting the environment and therefore the need for availability and access to appropriate information, development of databases and exchange information are the conditions for creating the basis for a sustainable development.
Experiences from the developed countries show that more than two-third of human decisions are affected by georeferenced information ans thus geospatial data are indispensable to make sound decisions at all levels, from the local to the global.
Despite the fact that administrations and governments are recognizing that spatial information needs to be efficiently coordinated and managed for the interest of all the citizens, this huge amount of geospatial data is stored in “electronic silos” in different places, by different organizations, and in different format. The vast majority of those data are not being used as effectively as they should because the collected data are often incompatible. This inevitably leads to inefficiencies and duplication effort. Moreover, geospatial information is an expensive resource and time consuming to produce. For this reason it is of high importance to improve the access and availability of data and promote its reuse because many of the decisions that different organizations have to take and want to achieve can only be made if good and consistent georeferenced data is available and readily accessible.

In consequence, it is essential to make these data easily available and accessible in order to give the opportunity to the users to turn them into understandable information. This will participate to building a better informed society because achieving the goal of a sustainable development involves the integration of a large number of different types of data from different sources. Through agreed common standards and a clear will to share data and information, these data can be interchanged and integrated in an interoperable way and leading to a collaborative approach to decision-making.

The growing recognition that once a geodata has been created it will be used as well in public and private sectors as in business and that there is a clear need to store data into databases and made accessible to everyone for others envisioned use leads to the concept that geodata could be a shared resource, which will be maintained continuously.

The advantages of having geographical data in a digital form are:

- easy storage,
- easy dissemination,
- facilitation of data exchange/sharing,
- faster and easier updates and corrections,
- ability to integrate data from multiple sources,
- customization of products and services.

As a result of the previous considerations the concept of the Spatial Data Infrastructure (SDI) was developed in order to facilitate and coordinate the exchange and sharing of geospatial data. SDIs encompass the data sources, systems, network linkages, standards and institutional issues involved in delivering geodata and information from many different sources to the widest possible group of potential users.

The overall objective of an SDI is to maximize the reuse of geospatial data and information avoiding the duplication efforts and expenses and thus enabling users to save resources, time and effort when trying to acquire or maintain datasets. This objective cannot be achieved without good coordination, good communication, good collaboration, adherence to known and accepted standards and procedures, institutional arrangements and policies, adequate human and technical resources. Moreover, the existence of geospatial data and information does not alone ensure that it is used for decision-making, and different other factors are important to consider in order to ensure that data and information will be used and reused:

- to be used, people need to know that the data exist, and where to obtain it.
- they need to be authorized to access and use the data.
- they need to know the history of the data capture, in order to interpret it correctly, trust it and be able to integrate it meaningfully with data coming from other sources.
- to know if the data depends on other data sets, in order to make sense of it.

In summary, SDIs intend to create an environment that foster activities for using, managing and producing data and in which all stakeholders can co-operate with each other and interact with technology, in order to better achieve their objectives at different political/institutional levels.

The driving forces behind SDIs are:

- promoting economic development,
- stimulating better government,
- fostering environmental sustainability,
- modernization of government,
- environmental management.

Today's effort on the technical development of SDI components clearly focus on the exchange of geodata in an interoperable way through web-based services that allow efficient access to spatially distributed databases and foster the need for an interoperable environment based on reusability and standardized components allowing the exchange of data, share of tasks and automate processes.

The interoperability is no more than the ability of two or more components to communicate and exchange information through a common system and use the information that has been exchanged. This give users the ability to:

- find what they needs,
- access it,
- understand and employ it,
- have good and services responsive to their needs.

As of today, in a climate of economic constraint, interoperability and standardization have never been so important because a non-interoperable system impedes the sharing of data, information and computing resources, leading organizations to spend much more than necessary on data, software and hardware. Moreover, being non-interoperable increases the risk for a system or an infrastructure to fail delivering its expected benefit and in consequence to lead to a user disappointment and system failure.

As a response to the need of standards to support interoperability in the geospatial communities, the Open Geospatial Consortium (OGC) aims to tackle the non-interoperability caused by the diversity of systems creating, storing, retrieving, processing and displaying geospatial data in different formats. In addition to this, software vendors usually do not communicate among them to agree on how data should be structured and stored and how systems must exchange information. This leads inevitably to a non-interoperable environment, isolating geospatial data in “electronic silos” and resulting in expensive duplication of data and difficulty in sharing and integrating information.

The OGC has pointed out the general user needs:

- need to share and reuse data in order to decrease costs (avoid redundancy collection), obtain additional or better information, and increase the value of data holdings;
- need to choose the best tool for the job and the related need to reduce technology and procurement risks (avoid being locked in to one vendor);
- need to leverage investment in software and data, enabling more people to benefit from using geospatial data across applications without the need for additional training;

The OGC believes that responding to user needs of interoperability will have a profound and positive impact in the public and private sectors, opening the doors to new business opportunities and new human activities.

In other words, interoperability enhances: communication, efficiency and quality for the benefit of all citizens allowing them to access data in a good, consistent and transparent way. To enable effective interoperability, it is necessary to commit to it, to follow standards, to foster collaboration, to commit to the collection and maintenance of metadata, and to commit to training and education.

All these commitments show that being interoperable is not only a matter of technology but also requires a change of philosophy, a spirit of mind to go "open". Only at this condition the following benefits of interoperability will be leveraged:

1. Allow to share and reuse of data: gives access to distributed and heterogeneous sources of data.
2. Avoid data duplication: data are collected and maintained at the most appropriate place.
3. Reduce costs: of maintenance, of operations and of production.
4. Integration: multi-source data integration could only be achieved with an effective interoperability.
5. Reduce the complexity: through common knowledge, standards offer a set of rules that every data provider can follow, understand and become familiar with. Moreover, when a user shares data in a standardized way, another person will be immediately able to use it.
6. Increase efficiency.
7. Vendor neutral: avoid to be locked in to one vendor.
8. Improve decision-making: offering standardized access to a vast amount of data and information and to used them as effectively as they should.
9. New opportunities and knowledge: open the doors to new activities and relations unforeseen before.

To support the implementation and the creation of Spatial Data Infrastructures and the use of open standards like OGC and ISO, several initiatives have emerged at the regional or global levels (e.g. INSPIRE or GEOSS). They are all concerned with data access, harmonization, standardization, interoperability, seamless integration and services. They coordinate actions that promote awareness and implementation of complementary policies, common standards and effective mechanisms for the development and availability of interoperable digital geographic data and technologies to support decision making at all scales and for multiple purposes. These actions encompass the policies, organizational remits, data, technologies, standards, delivery mechanisms, and financial and human resources necessary to ensure that those working at the global and regional scale are not impeded in meeting their objectives.

In conclusion, due to the diversity of the partners involved in the enviroGRIDS project and the associated heterogeneity (both in term of technical capabilities and knowledge) of SDI and interoperability matters, we envision the following technical and collaborative challenges:

- Assessing the current status of the geospatial sector in the region, meaning that we have to identify what is available in term of data and GIS capabilities, what is the level of implementation of OGC standards and the thematic coverage of data related to the GEOSS Societal Benefits Area and INSPIRE themes. This assessment will help us to identify gaps in data, thematic coverages, and GIS capabilities.
- Facilitating access to the data sets either by direct agreements with data providers (in case data are not freely available) or by having one or several metadata catalogues of good quality.
- Making them interoperable, in order to allow easy access, retrieval and use.
- Promote and building capacities on OGC and open standards if the identified data are not already interoperable, in order to assist data providers to share their data.

- Providing appropriate levels of network service access: could be an issue, if the quality of network communications avoid or at least impede good web-based communications.
- Pushing for data harmonization, using the INSPIRE data specifications.
- Integrating and connecting heterogeneous data sources.

In order to achieve the goal of a "sharing spirit", it is of high importance to promote collaboration and cooperation among partners. Good communication and good organization are fundamental for sharing not only data but also information, knowledge and capacities. In consequence, high priority must be given to the creation of a well understood and accepted governance structure in order to develop a clear strategy for the deployment of the infrastructure. This will allow partners to share the same vision and to feel a common sense of ownership of the future Black Sea Observation System. This will certainly create a spirit of commitment around the project and thus would greatly facilitate the endorsement of a Spatial Data Infrastructure and its related concepts.

As presented in this deliverable, sharing fosters the notion of re-use of data, but also re-use of technical capabilities and skills developed, highlighting the importance of capacity-building, the necessity to learn from others and to share also knowledge among the different partners of the project.

If collaboration is achieved with the help of good communication and organization, the agreement on the use of new standards, the development of guidelines, good practices and policies will greatly enhance the "open and shared spirit".

Finally, the key challenge of the EnviroGRIDS project is to bridge the technological gap between the world of Spatial Data Infrastructure and Grid computer infrastructure, creating a Grid-enabled SDI.
With the ever increasing spatial and temporal resolution of geospatial data causing a tremendous increase in term of size of these data, we progressively reach the limits of the processing capacities of the traditional GIS/SDI technology. With the uptake of Grid technology and the progressive deployment of large infrastructure project like the EGEE many more scientific domains and technology-related activities have access to sizable computing resources. However, several differences between OGC-compliant SDIs and Grid infrastructures exist concerning the description of services, their interfaces, their state, the way they implement security and the way they catalogue metadata. All these differences must be taken into account in order to extend the analytical capabilities of traditional SDIs.
Based on the foreseen challenges, we propose a set of both general and specific recommendations to successfully implement interoperability among the partners of the project and beyond, and build capacities around the different topics of Spatial Data Infrastructure. Participants will act as a key drivers and pioneers, to promote sharing of data supported by SDI initiatives and serving and receiving data to and from systems like INSPIRE and GEOSS.

General recommendations:

- **GR1: Assessing the status and environment of geospatial sector in the region.**
- **GR2: Defining a core-geodatasets needed for the project.**
- **GR3: Creating appropriate governance structure.**
- **GR4: Developing a clear strategy and vision.**
- **GR5: Facilitating access to data.**
- **GR6: Making data interoperable.**
- **GR7: Promoting the use of open standards.**
- **GR8: Endorsing the GEOSS and INSPIRE principles.**
- **GR9: Building capacities.**
- **GR10: Obtaining commitment**
- **GR11: Having good communication and organization**
- **GR12: Taking care of the differences between SDIs and Grid Infrastructure.**

- **GR13: Proposing an envisioned architecture for a Grid-enabled SDI.**

Specific recommendations:

- **CB1: Building on existing efforts and best practices,**
- **CB2: Focusing on user needs,**
- **CB3: Foster collaboration and partnerships,**
- **CB4: Fostering information exchange and knowledge development,**
- **CB5: Facilitating the exchange of ideas and best practices,**
- **CB6: Facilitating the sharing of human and technical resources,**
- **CB7: Facilitating education and training**
- **CB8: Promoting the sharing of data, information, reports, articles and guidelines.**
- **CB9: Identifying experts in each specific areas,**
- **CB10: Identifying, coordinating and building synergies between existing and future capacity building efforts and SDIs initiatives.**
- **CB11: Enhancing the sustainability of existing and future SDIs initiatives and capacity building efforts by building awareness amongst the different stakeholders.**

To achieve and meet these recommendations it is of high importance to clearly define:

– **who are and who could be the end users of the grid-enabled SDI?**

At this stage the two main end-users will be the Black Sea Commission (BSC) and the International Commission for the Protection of the Danube River (ICPDR). The infrastructure must help them to increase their capacities and improve their every day work, especially regarding data and information management, state of the environment reporting, assessments, water quality studies, early-warning systems and modeling.

It would be good also to identify and reach other potential users of the infrastructure (other than the partners of the project that are obvious users) within the region in order to promote the use of SDI and Grid infrastructure.

– **what is the design of the architecture?**

In the deliverables 2.1 and 2.2, we present general architectures for the Spatial Data Infrastructure (fig.16) as well as the Grid Infrastructure. Now it is important to define and agree on a clear and complete architecture connecting the two types of infrastructures allowing to discover and retrieve data through the SDI and analyze them on the Grid.

For that purpose, we recommend to organize a specific workshop at the beginning of year 2010 (but before the General Assembly), with the relevant partners of the project. The output of the project will be a document presenting the EnviroGRIDS grid-enabled SDI specifications, a list of tasks and a roadmap to develop the overall infrastructure, a list of proposed tools to be used, and finally an efficient and practical implementation strategy allowing our infrastructure to be scalable and to adapt it on the coming and future developed solutions (like G-OWS, 52north, ...).

– **what is the long-term plan for the infrastructure?**

As we are trying to connect two types of infrastructures (SDI and Grid) and due to the fact that participating to the Grid involves some financial investments, it is important to already think and clarify what will be the sustainability plan for our infrastructure especially after the end of the project. It is essential to guarantee a long-term access to the Grid. Otherwise the grid-enabled SDI will lose its analytical and processing component and it will be highly damageable for the end-users of the infrastructure, as well as the EnviroGRIDS consortium.

This issue will be covered in deliverable D2.5.

– **who's go what?**

To identify the different data sources it is necessary that metadata be immediately available. If partners and/or data providers within the Black Sea catchment have already metadata catalogue we have to know them so that we can query their catalogue directly from the EnviroGRIDS URM (http://www.envirogrids.cz). In the case that metadata is unavailable because either lack of capacities to publish them using CSW or even worst because data are not documented, the EnviroGRIDS project should help them to learn at least to write metadata into the catalogue system of our URM, but it would be better to build capacities and help them to learn to install and use tools like GeoNetwork.

– **Data policy?**

In order to give a clear vision of the data exchange to the data providers and to allow data sharing at a large extent, it is of primary importance to have a good and precise data policy. It must be clear enough so that data providers would be interested in sharing their data and cover items of primary interest to them, like: data sharing principles, intellectual property rights, access rights, etc...

In the annexes of this document a draft proposition is already made but must be clearly improved. We recommend circulating this data policy among partners and especially the BSC and ICPDR to have their inputs. The aim is to finalize it for the next General Assembly in spring 2010 where all partners should endorse it.

– **Data quality?**

Once we have access to the vast amount of data that is available in the region, we need to assess its quality, because not all data will be useful for the end-users of the infrastructure. Therefore, we recommend to define a methodology to assess data quality.

– **build on best practices**

The last two points highlight the need to build our knowledge on best practices that are already available on the GEO/GEOSS and INSIPIRE websites.

– **build capacities**

As we already mention previously, the main goal of the project is to build capacities in the Black Sea region especially regarding data sharing and related technologies (web services, metadata catalogues, etc.) allowing data providers to register their data into GEOSS. The success of the project could and/or will be measured on how many new contributors of the region will register their services into the GEOSS Common Infrastructure.

In order to help them to get the necessary knowledge and use the GEOSS system, we will organize:

- High level workshops (for the different stakeholders of the region) aiming to build awareness on the benefits of sharing data, to present GEO/GEOSS and its added values;
- Technical trainings (focusing on GIS/IT people) to teach them how to install and use softwares like GeoServer, GeoNetwork and/or ESRI products such as ArcGIS Server.

These two types of workshops could be replicated at least in the different countries around the Black Sea and thsi learning material could be translated and distributed in the different regional languages.

Our hope is to plant and grow some seeds of a new vision based on the benefits of data sharing and offering standardized access to a vast amount of data and information. The challenge will then become to use this data as effectively as they should in order to allow the decision-makers within the Black Sea catchment to take better informed decisions.

Tables of Content

1.1 Presentation of the EnviroGRIDS project

The Black Sea Catchment is internationally known as one of ecologically unsustainable development and inadequate resource management, which has led to severe environmental, social and economic problems. The EnviroGRIDS @ Black Sea Catchment project addresses these issues by bringing several emerging information technologies that are revolutionizing the way we observe our planet. The Group on Earth Observation Systems of Systems (GEOSS) is building a data-driven view of our planet that feeds into models and scenarios to explore our past, present and future. EnviroGRIDS aims at building the capacity of scientist to assemble such a system in the Black Sea Catchment, the capacity of decision-makers to use it, and the capacity of the general public to understand the important environmental, social and economic issues at stake.

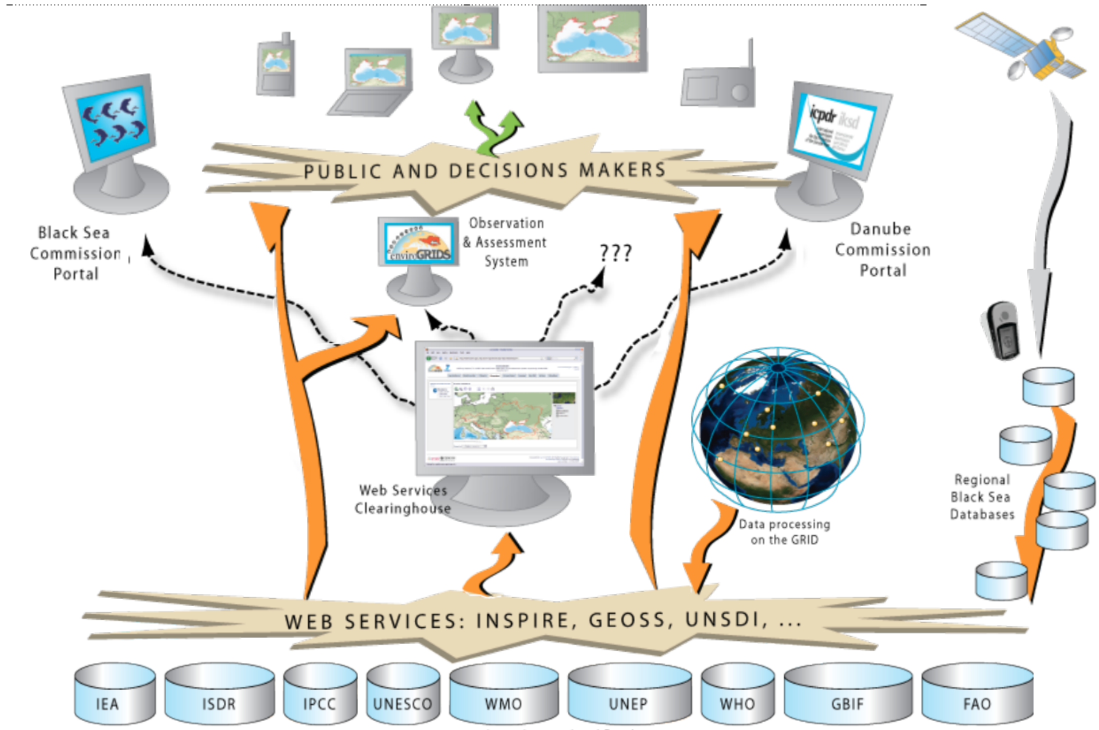


Fig.1: Project Vision and Goals.

EnviroGRIDS aims at building capacities in the Black Sea region to use new international standards to gather, store, distribute, analyze, visualize and disseminate crucial information on past, present and future states of this region, in order to assess its sustainability and vulnerability. To achieve its objectives, EnviroGRIDS will build a Grid-enabled Spatial Data Infrastructure (GSDI) becoming one of the integral systems in the Global Earth Observation System of Systems (GEOSS), and compatible with the new EU directive on Infrastructure for Spatial Information in the European Union (INSPIRE), as well as UNSDI developments.
The scientific aim of the EnviroGRIDS @ Black Sea Catchment project is to start building an Observation System that will address several GEO Societal Benefit Areas within a changing climate framework. This system will incorporate a shared information system that operates on the boundary of scientific/technical partners, stakeholders and the public.
It will contain an early warning system able to inform in advance decision-makers and the public about risks to human health, biodiversity and ecosystems integrity, agriculture production or energy supply caused by climatic, demographic and land cover changes on a 50-year time horizon (Lehmann, 2009).

1.2 Presentation of the Work package 2: Spatial Data Infrastructure

The aim of this WP is to create a Grid enabled Spatial Data Infrastructure (GSDI) so that the data that are necessary for the assessment of GEO Societal Benefit Areas, as well as the data produced within the project to fulfill EnviroGRIDS specific objectives. This data can be gathered and stored in an organized form on the Grid infrastructure and distributed across the Grid in order to provide a high performance and reliable access through standardized interfaces. Using the standardized technologies of the Grid we can provide a Single Information Space for environmental data in the Black Sea Catchment.

First a gap analysis is undertaken by all partners to analyse the state of development of EO & SDI in the different countries within the Black Sea Catchment under the supervision of BSC PS and ICPDR partners. The recommendations derived from the gap analysis should aim at complementing the existing geographical information systems of the ICPDR and BSC. It will serve also to bring new partners from the Black Sea countries into the project in order to fill some thematic or technical gaps.

The proposed Grid-enabled system will require the creation of a Spatial Direct interface to load and download spatial data in different format and projections. Partners will therefore be able to make available existing sources of data more or less publicly available (e.g. historical climate data or geographic information), or newly collected data from sensors and satellites. This system will be fully compatible with developments made on interoperability standards such as INSPIRE, GEOSS, UNSDI, OGC, SensorML or TML. Grid services can be used to replicate and distribute the data from source sites to other data centres to improve availability but also access performance.

We intend to base our work firmly on the experience that the EGEE project (and in particular our partner CERN) has acquired in similar projects, but with non-geographic data (e.g. biomedical data). CERN also collaborated with UNOSAT to store satellite images and geographic metadata on the Grid. In particular, CERN has been providing the AMGA metadata service as part of the gLite middleware of the EGEE project that allows access on the Grid to databases storing GIS information. We intend to adapt this catalogue and its extensive replication and federation features to provide a unified view of all available metadata. EnviroGRIDS GSDI will also allow distributing intensive calculations such as those needed for hydrological modelling and calibrations.

1.3 Scope and Purpose of the Task 2.2: Interoperability and data storage

This task is separated into two subtasks, one regarding interoperability and one regarding storage. The concern of the present deliverable (D2.1) is on the first subtask, the part concerning storage is covered by another complementary deliverable (D2.2).

The societal benefits of Earth observations cannot be achieved without data sharing. EnviroGRIDS will therefore apply the following GEOSS data sharing principles:

- There will be full and open exchange of data, metadata, and products shared within GEOSS, recognizing relevant international instruments and national policies and legislation.
- All shared data, metadata, and products will be made available with minimum time delay and at minimum cost.
- All shared data, metadata, and products free of charge or no more than cost of reproduction will be encouraged for research and education.

In order to develop observation systems it is crucial to agree on relevant standards to describe, merge as well as disseminate distributed information. This is what is proposed here through the exploitation of geospatial standards like OGC and ISO.

EnviroGRIDS metadata and data standards will also be coordinated and aligned with GEOSS standards, which are currently under development. Note that EnviroGRIDS is now an official task named "Vulnerability of Sea Basin" (EC-09-02c) in the GEO work plan 2009-2011 under the Area "Ecosystems" and the overarching task "Ecosystem vulnerability to Global Change".

The European directive on "Infrastructure for Spatial Information in the European Community" (INSPIRE) is a major milestone for the use of Geographical Information in Europe, and is a central contribution to environmental policy and sustainable development. The directive is a legal agreement that encourages European countries to have: up-to-date metadata using identified themes lists and numbers of attributes, interoperability for dataset and services, facilitation of network access and sharing of data. EnviroGRIDS will therefore be built around the INSPIRE directive to guarantee full compatibility among European and International Cooperation Partner Countries (ICPC).

The main standards that will be used are OGC and ISO 19115/19139. The Open Geospatial Consortium (OGC) is a non-profit, international, voluntary consensus standards organization that is leading the development of standards for geospatial and location based services. The consortium is constituted of 350 companies, government agencies and universities participating in a consensus process to develop publicly available interface specifications. OpenGIS Specifications support interoperable solutions that "geo-enable" the Web, wireless and location-based services, and mainstream IT. The specifications empower technology developers to make complex spatial information and services accessible and useful with all kinds of applications.

Published in May 2003, ISO 19115 is the most widely used international standard for describing geographic information and services. It provides information about the identification, the extent, the quality, the spatial and temporal schema, spatial reference, and distribution of digital geographic data. It is applicable to cataloguing of datasets and clearinghouse activities. As well as geographic datasets, dataset series, individual geographic features and feature properties, the standard can be extended to fit specialized needs (description of statistics, documents e.g.). ISO 19139 is the standard that aims to define an XML encoding (XML schema implementation) for the metadata elements defined in ISO 19115.

In the context of INSPIRE, a survey done in 2006 showed that a majority of the European countries were using ISO standards (19139, 19115). The use of this standard is growing year after year. For instance, the current world leader of GIS software ESRI proposes in its ArcGIS package a module called ArcCatalog that allows (since 2006) the management of ISO 19115 compliant metadata. Tailored Visual basic metadata editor using ArcCatalog dll can also be used or developed. The INSPIRE survey analysis showed that OGC specifications were applied to 39% of all metadata holdings. Like ISO19115, OGC specifications are widely and internationally used. They facilitate the interaction, sharing and visualisation of geospatial dataset. The use of these specifications allow direct use of number of existing resources using OGC webservices WMS (Web Map Service), WMF (Web Map Features) or WCS (Web Coverage Services) and a better dissemination and use inside the EnviroGRIDS SDI infrastructure as well as for external partners. The Web Feature Service (WFS) is certainly one of the most valuable specifications of the OGC. It provides a generic way for accessing raw geographic data over the web. To the general user, this can potentially provide a wealth of information embedded in the map being viewed. Parts of the WMS specification tried to implement this functionality, but using WFS gives much more control over how to actually access the data. A WFS leads to greater transparency and openness in mapping applications.
Instead of merely being able to look at a picture of the map, as the provider wants the user to see, the power is in the hands of the user to determine how to visualize the raw geographic and associated data. The data can be downloaded, analysed and combined with other data, or it can be chained with other web services to do even more interesting things on the web.

These standards will be implemented in EnviroGRIDS through the metadata guidelines production and integration, as well as through data geoservices and data geoprocessing services that will be designed around OGC specifications.

1.4 Scope and purpose of Deliverable 2.1

This document aims to provide a technical overview and basic concepts governing Spatial Data Infrastructures and associate topics in order to share this necessary knowledge with the partners involved in the EnviroGRIDS project and beyond to assist them to share and integrate spatial data sets more effectively.

The different chapters of the deliverable cover the following points:

- Spatial Data Infrastructure (chapter 3)
- Interoperability and standards (chapter 4)
- Global and regional SDIs initiatives (chapter 5)
- Organizations involved in developing standards for the GIS/SDI industry (chapter 6)
- Description of the relevant standards for the project (chapter 7)
- Description of the tools that implement those standards (chapter 8)

We conclude with some important challenges (chapter 9) and recommendations (chapter 10).

This report is part of a set of guidelines and are linked with:

- D2.2: EnviroGRIDS data storage guideline,
- D2.3: EnviroGRIDS sensor data use and integration guideline,
- D2.4: EnviroGRIDS remote sensing data use and integration guideline.

Altogether, these deliverables form the base set of knowledge for the project

1.5 Contributors to the deliverable

Gregory Giuliani:

After obtaining a degree in Earth Sciences, he went on to complete a master in Environmental Sciences, specializing in remote sensing and GIS. He previously worked as a GIS Consultant for the World Health Organization, as a University tutor in remote sensing and GIS and as a GIS developer in a local Swiss GIS company. He works at UNEP/GRID-Europe since 2001 and is the focal point for Spatial Data Infrastructure (SDI). In 2008, he also started to collaborate closely with the enviroSPACE laboratory where he begins a Ph.D thesis and works also for the FP7 ACQWA project. In EnviroGRIDS, he is involved as WP2 leader where his objective is to coordinate SDI and Grid technology researches.

Nicolas Ray:

Dr. Nicolas Ray is the EnviroGRIDS project manager. He holds a Master in Environmental Sciences, a PhD in Biology from University of Geneva in the field of Human Population Genetics, and a postgraduate certificate in Computer Sciences. Nicolas' research activities focused on the modeling of animal movement and habitat, with the development of several spatial and statistical analysis tools to integrate various data types (genetic, environmental, demographic). Through his research and various consultancy works, Nicolas acquired solid competencies in GIS, spatial analysis, software development, and grid computing integration. After six years of postdoctoral research in Australia, UK and University of Bern, he recently joined the Institute of Environmental Science at University of Geneva and UNEP/GRID-Europe to dedicate his time to the enviroGRIDS project.
In addition to managing the daily activities of the project and ensuring high quality deliverables, he will take an active part in WP2 with all tasks related to the integration of grid computing and SDI.

Anthony Lehmann:

Dr. Anthony Lehmann is the EnviroGRIDS project initiator and coordinator. He holds a Masters Degree and a PhD in Aquatic Biology from the University of Geneva, and a Postgraduate Master in Statistics

from the University of Neuchâtel. He specialized during his career in combining GIS analyses with statistical models. At the University of Geneva he is in charge of the enviroSPACE laboratory exploring Spatial Predictions and Analyses in Complex Environments.
He is sharing is working time at a 50% rate with the United Nations Environment Programme (UNEP) Global Resource Information Database (GRID) under a special agreement between the University of Geneva and UNEP. At GRID, Dr. Lehmann is responsible for organizing research activities by leading the "environmental monitoring and modelling" unit. With the EnviroGRIDS project, his personal objective is to motivate all the partners to give their best in order to create a great observation system for the Black Sea Catchment. He will be especially involved in WP3 and WP5.

Karel Charvat:
Dr. Karel Charvat has a PhD in theoretical cybernetics from Charles University in Prague and currently works for Czech Centre for Science and Society.

2. Responding to our changing environment, the need for data sharing

Today we are living in a globalized world where everything is changing rapidly (growing population, environmental deterioration,...) and where communication means have taken a remarkable place in our life. Every day we access an enormous and continuous flow of information and much of it refer to a position or a specific place on the surface of our planet. This information is therefore, and by definition, georeferenced.

In the last 30 years, the amount of georeferenced data available has grown dramatically following the evolution of the communication means and due to the rapid development of spatial data capture technologies such as Global Positioning System (GPS), remote sensing images, sensors, etc (Philips *et al.*, 1999). Over the last ten years, with the advent of applications like Google Earth, we have seen that geoinformation has been incorporated and routinely embedded into business and workflows of agencies at all levels of government, as well as in the private sector (Booz *et al.*, 2005).

Despite the fact that administrations and governments are recognizing that spatial information is important and must be part of the basic information infrastructure that need to be efficiently coordinated and managed for the interest of all citizens (Ryttersgaard, 2001), this huge amount of geospatial data is stored in different places, by different organizations and the vast majority of the data are not being used as effectively as they should.

Moreover at the United Nations Conference on Environment and Development in Rio de Janeiro in 1992, the so-called Agenda 21 resolution fosters the importance of georeferenced information to support decision-making and management on the degradation and threats that are affecting the environment (GSDI, 2004). This means that there is a strong need for availability and access to appropriate information. The development of databases and exchange of information are the conditions for creating the basis for a sustainable development and to support the information management needs for implementing and monitoring sustainable development policies and goals like the UN Millennium Development Goals (UNGIWG, 2007).

Thus, geospatial information is a critical element underpinning decision making for many disciplines (Rajabifard and Williamson, 2001) and is indispensable to make sound decisions at all levels, from local to global. Experiences from developed countries show that more than two-third of human decision-making are affected by georeferenced information (Ryttersgaard, 2001).

However, geospatial information is an expensive resource, it is time consuming to produce, and for this reason it is of high importance to improve the access and availability of data, and promote its reuse. Many of the decisions that different organization need to make depend on good and consistent georeferenced data, available and readily accessible (Rajabifard and Williamson, 2001).

In 1998, the former vice-president of the United States, Al Gore, presented its visionary concept of a Digital Earth (Gore, 1998), "*a multi-resolution, three-dimensional representation of the planet, into which we can embed vast quantities of georeferenced data"*. As of today this vision is clearly not fully realized, but gives us an interesting support to our purpose as it is still actual.

Talking about geospatial data, Al Gore (1998) mentioned that the difficult part of taking advantage of this vast amount of information will be "*making sense of it, turning raw data into understandable information"* because at the moment we have more information than we can handle and it is stored in *"electronic silos of data"*, remaining mostly unused. He envisioned applications where *"information can be seamlessly fused with the digital map or terrain data"* allowing the user to move through space and time, but of course, to achieve this vision, a collaborative effort (from government, industry, academia and citizens) is needed.

All the technologies and capabilities required to realize this vision and to build a Digital Earth are already available:

- computational science: even a simple desktop computer can process complex models and simulations. With the potential that technologies such as the Grid are offering new insights into the data are possible, giving us the ability to simulate phenomena that are impossible to observe.
- mass storage: storing tera-bytes of data on a desktop computer is not a problem anymore.
- satellite imagery: a lot of satellites are continuously observing the Earth offering high spatial and temporal observations.
- broadband networks: are already a reality giving the ability to connect different databases together.
- interoperability: this is a key point to allow communication and integration of distributed data, allowing the geospatial data generated by one software to be read by another.
- Metadata: are important as they describe the data, allowing a user to evaluate and discover the data before using them.

Even if all technologies are ready, organizations and agencies around the world are still spending billions of dollars every year to produce, manage and use geographical data but without having the information they need to answer the challenges our world is facing (Rajabifard and Williamson, 2001). These authors also highlight the facts that most organizations and/or agencies need more data than they can afford, they often need data outside their jurisdictions, and the data collected by different organizations are often incompatible. This inevitably leads to inefficiencies and duplication of effort, and thus it is evident that countries can benefit both economically and environmentally from a better management of their data (UNGIWG, 2007; GSDI, 2004).

In consequence, it is now essential to make these data easily available and accessible in order to give the opportunity to the user to turn them into understandable information with a clear and broad benefits for the society and the economy because *"working together, we can help and solve many of the most pressing problems facing our society…"* (Gore, 1998).

It is clear that there are a lot of challenges to face, both tangible and intangible, when we start sharing data but we have to overcome them in order to improve our knowledge, share our experiences and try to build a better informed society. Achieving the goal of a sustainable development requires the integration of a large number of different types of data from different sources. Through agreed common standards and a clear political will, these data can be interchanged and integrated in an interoperable way, leading to a new collaborative approach to decision-making.

In conclusion, for Arzberger *et al.* (2004), ensuring that data are easily accessible, so that they can be used as often and as widely as possible is a matter of sound stewardship of public resources. Availability should be restricted only in certain specific cases like national security. These authors argue that *"publicly funded research data should be openly available to the maximum extent possible"*, because publicly funded data are a public good, produced in the public interest.

It seems to be the right time to really work together and to share our data in order to provide a state-of-the-art grid-enabled spatial data infrastructure for the Black Sea catchment.

3. Spatial Data Infrastructure

3.1 Definition, concepts and rationale

The term Spatial Data Infrastructure (SDI) if often used to describe the mechanisms or the enabling environment that support easy access to, and utilization of, geographical data and information (UNECA, 2005). This definition is quite reductive as it gives the idea that SDIs are essentially technical. The primary objective of SDIs is to provide a basis for geospatial data discovery, evaluation, and application for users and providers within all levels of government, commercial and the non-profit sectors, academia and citizens (GSDI, 2004).

This means that SDIs are more than just data repositories. SDIs store data and their attributes, and their related documentation (metadata), offering a mean to discover, visualize, and evaluate their fitness to different purpose, and finally provide access to the data themselves. In addition to these basic services, there are often additional services or software supporting the use of the data. Finally, to make an SDI work efficiently, it is necessary to include all the organizational agreements needed to coordinate and administer it.

In consequence, following Masser (2005) and GSDI (2004), we can give a more complete definition of what are SDIs:

"A spatial data infrastructure supports ready access to geographic information. This is achieved through the coordinated actions of nations and organizations that promote awareness and implementation of complimentary policies, common standards and effective mechanisms for the development and availability of interoperable digital geographic data and technologies to support decision making at all scales for multiple purposes. These actions encompass the policies, organizational remits, data, technologies, standards, delivery mechanisms, and financial and human resources necessary to ensure that those working at the national and regional scale are not impeded in meeting their objectives".

Before going further in details, we have to explain the concepts underlying the rational of SDI, in particular geospatial data and information (also named geodata or georeferenced data). A geodata describes a location on Earth, giving through its attributes a comprehensive picture of the physical world both in term of spatial and/or temporal extent. Geodata are extremely valuable as users can build spatial relationships between features and data. For example, just after a flood event, one can overlay remote sensing images with existing georeferenced data of settlements to evaluate the extent of the damage and then focus humanitarian assistance. In consequence, geodata has a key role to play in our knowledge-based economy affecting directly or indirectly different sectors like forestry, urban planning, security, telecommunication, environmental protection, etc.

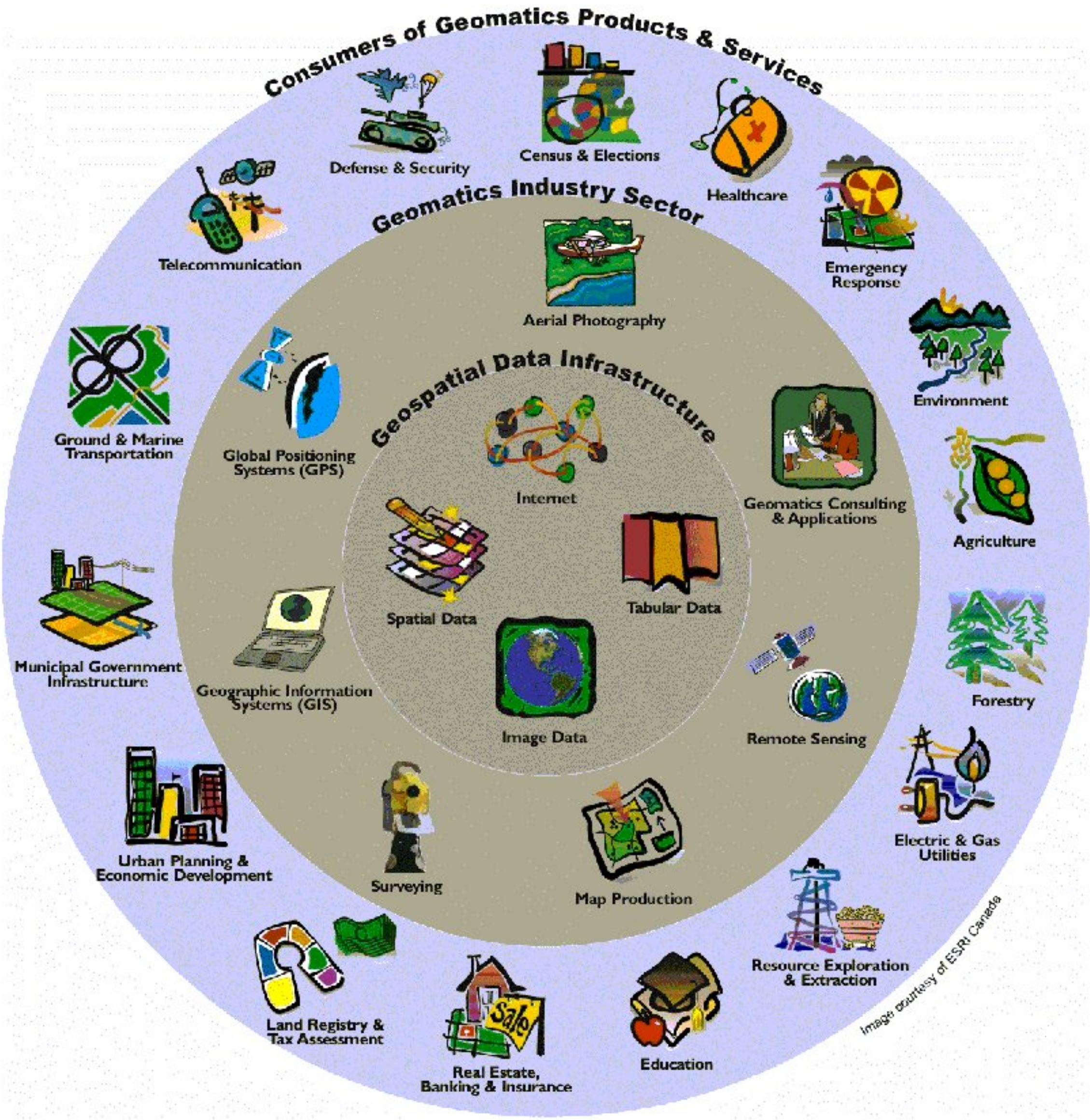


Fig.2: GIS and economy (Source: Geoconnections).

If, previously, geographical information was mostly presented in the form of paper maps, with the increasing means to capture information in digital formats, geospatial data are nowadays used and viewed within a Geographical Information System (GIS). This computer system is capable of assembling, storing, manipulating, and displaying geographically referenced information (UNECA, 2005).

A GIS gives the ability to merge different existing information from different sources facilitating collaboration in creating and analyzing data. Due to these new possibilities of reusing existing data and working on collaboratively greater scale, new challenges arise. When someone wishes to create a new information layer based on different data sets or different formats, with different terminology, and perhaps different projection, it is quite difficult to bring them together. Harmonizing geodata is a complex, costly and time-consuming task, but could be achieved by agreeing among data capturers before the work begins.

The growing recognition that once a geodata set has been created it could be used for public and private sectors (Ryttersgaard, 2001), reinforces the need to store data into databases that are made accessible for different purposes (Philips *et al.*, 1999). This leads to the concept that geodata could be a shared resource, which will be maintained continuously.

The advantage of having geographical data in a digital form (UNECA, 2005; UNGIWG, 2007) are:

- easy storage,
- easy dissemination,
- facilitation of data exchange/sharing,
- faster and easier updates and corrections,
- ability to integrate data from multiple sources,
- customization of products and services.

As a result of the previous considerations, the concept of the SDI was developed in order to facilitate and coordinate the exchange and sharing of geospatial data (Rajabifard and Williamson, 2001), encompassing the data sources, systems, network linkages, standards and institutional issues involved in delivering geodata and information from many different sources to the widest possible group of potential users (Coleman et al., 1997).

The vision of an SDI incorporates different databases, ranging from the local to the national, into an integrated information highway and constitutes a framework, needed by a community, in order to make effective use of geospatial data (UNECA, 2005).

3.2 Objectives

Following Masser's definition (2005) and the different considerations highlighted in the previous section we can list different objectives underpinning SDIs:

- The overall objective of an SDI is to maximize the reuse of geospatial data and information.
- SDIs cannot be realized without coordination (especially by governments).
- SDIs must be user driven, supporting decision-making for many different purposes.
- SDIs implementation involves a wide range of activities, including not only technical topics such as data, standards, interoperability, and delivery mechanisms, but also institutional arrangements, policies, financial and human resources.
- The term infrastructure is used to promote the idea of a reliable and supporting environment, analogous to a road or a telecommunication network, facilitating the access to geoinformation using a minimum set of common practices, protocols, and specifications (GSDI, 2004). This allows the movement of spatial information instead of goods.
- SDIs are all about (UNGIWG, 2007):
  - re-use: of data, technical capabilities, skills developed, invested effort and capital.
  - sharing: "sharing-not-wearing" the costs of data, people, technology,... helping to realize more rapid returns on investment.
  - learning from others: avoiding the pitfalls experienced by others.
- Avoid duplication efforts and expenses and enables users to save resources, time and effort when trying to acquire or maintain datasets (Rajabifard and Williamson, 2001).
- SDIs are "about working smarter, not harder" (UNGIWG, 2007).
- Implies to scale from specific and monolithic (data-centric) towards independent and modular (service-oriented) information systems.
- Integrate these systems together into an information highway which both links together environmental, socio-economic and institutional databases and provides a movement of information from local to national and global levels.
- Encompass the sources, systems, network linkages, standards and institutional issues involved in delivering spatially-related information from many different sources to the widest possible group of potential users.

Altogether these objectives intend to create an environment that foster activities (fig.3) for using, managing and producing geospatial data and in which all stakeholders can co-operate with each other and interact with technology, to better achieve their objectives at different political/institutional levels (Rajabifard and Williamson, 2004).

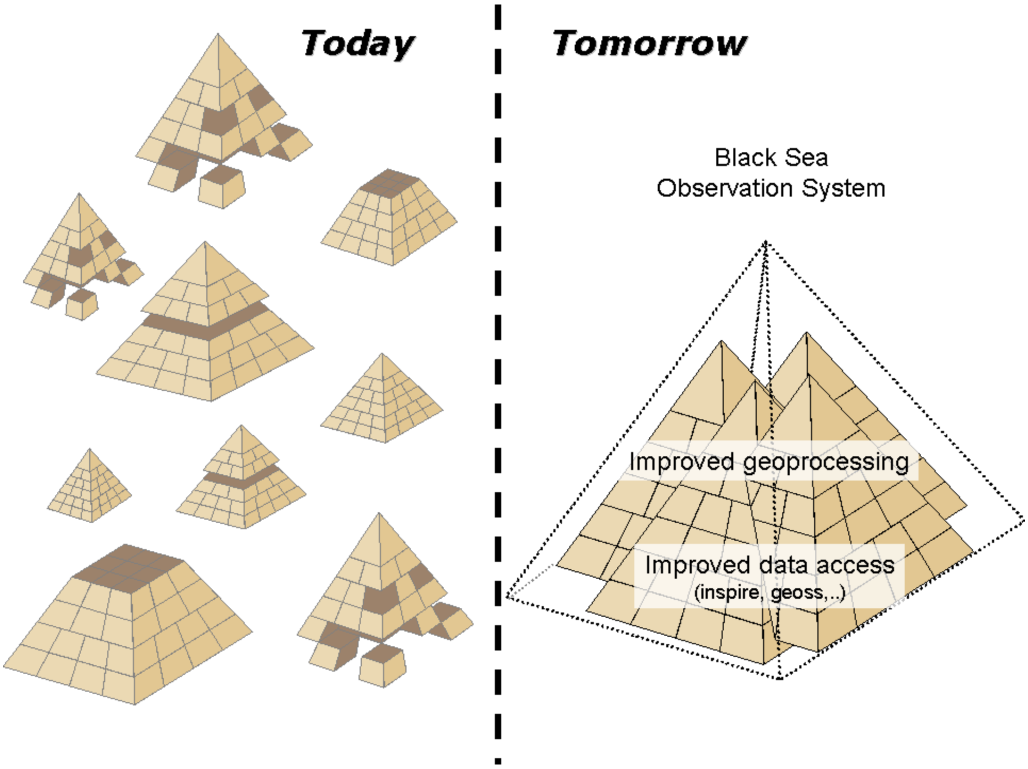


Fig.3: EnviroGRIDS aims

3.3 Components

Masser (2005) identifies the most important stakeholders with special interests in geoinformation/SDIs matters and shows the diversity both in terms of size and resource of the large numbers of players involved:

- Central government organizations,
- Local government organizations,
- Commercial sector,
- Non-for-profit/non-governmental organizations,
- Academics,
- Individuals.

Thus the temptation will be to create a centralized “one-size-fits-all” spatial database, in order to provide all the information needed by a country or a specific community of common interest. But as reported by UNECA (2005), UNGIWG (2007) and GSDI (2004) the existence of geodata and information does not alone ensure that it is used for decision-making. Different other factors are important to consider in order to ensure that information will be effectively used and reused:

- To be used, people need to know that the data exist, and where to obtain it.
- They need to be authorized to access and use the data.
- They need to know the history of the data capture, in order to interpret it correctly, trust it and be able to integrate it meaningfully with data coming from other sources.
- They need to know if the data depends on other data sets, in order to make sense of data.

Consequently, to leverage the full potential of geospatial data, an SDI must be made of different components to allow users to find, discover, evaluate, access and use these data, namely:

- A clearly defined core of spatial data.
- The adherence to known and accepted standards and procedures.
- Databases to store data and accessible documentation about the data, the so-called metadata.
- Policies and practices that promote the exchange and reuse of information.
- Adequate human and technical resources to collect, maintain, manipulate and distribute geospatial data.
- Good communication channels between people/organizations concerned with geodata, allowing the establishments of partnerships and share knowledge.
- The technology for acquiring and disseminating data through networks.
- Institutional arrangements to collaborate, co-operate and coordinate actions.

But as stated by Rajabifard and Williamson (2001), there is an important additional component represented by people. This includes the users of geospatial data but also the providers and any other data custodians. For these authors, people are the key to transaction processing and decision-making. Facilitating the role of people and data in governance that appropriately supports decision-making and sustainable development objectives is central to the concept of SDI.

In order to meet the requirements of all stakeholders involved, an SDI must (Coleman *et al.*, 1997):

- be widely available,
- be easy to use,
- be flexible,
- form the foundation for other activities.

In summary, Rajabifard and Williamson (2001) suggest that an SDI cannot be seen only as composed of geospatial data, services and users but instead involves other issues regarding interoperability, policies and networks.

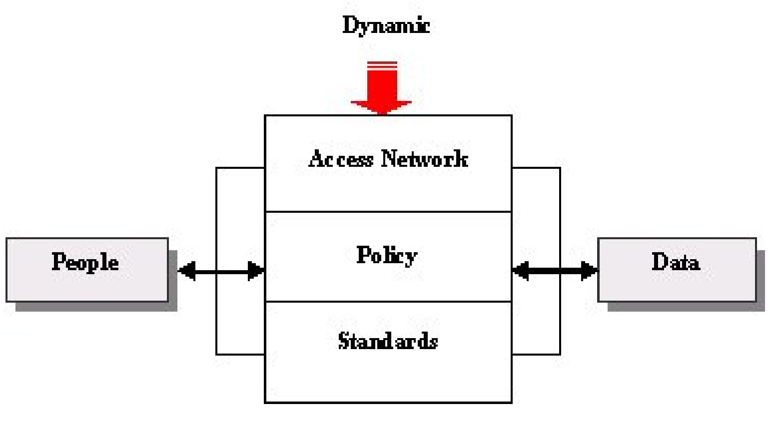


Fig.4: Nature and relations between SDI components (Source: GISCafe).

This shows that an SDI is by nature really dynamic as people who want to access data must interact with technological components (fig.4).

3.4 SDI hierarchy

As a result of the fact that SDI initiatives range from local to national and regional levels (Crompvoets, 2003; Masser 2007) and they all aim to promote economic development, to stimulate better government and to foster environmental sustainability (Masser, 2005), Rajabifard (2002) proposed a model of SDI hierarchy that is made of inter-connected SDIs developed at different levels (from local to global). Each SDI of a higher level is primarily formed by the integration of geospatial datasets developed and made available by the lower level (fig.5).

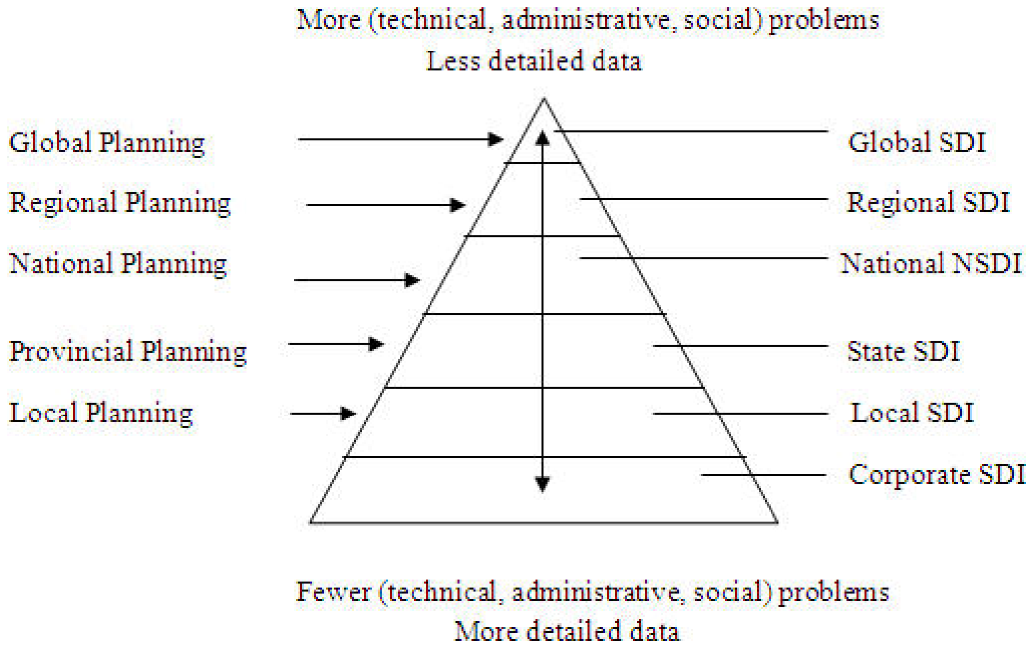


Fig.5: SDI hierarchy. (Source: GISCafe)

Such a hierarchy has two views: in one hand it is an umbrella in which the SDI at a higher level encompasses all the components of SDIs at levels below. On the other hand, it can be seen as the building block supporting the access of geospatial data needed by SDIs at higher levels. The SDI hierarchy allows to create an environment in which users working at any level can rely on data from other levels and integrate geospatial data from different sources (Mohammadi *et al.* 2007). Such a hierarchy is envisioned by regional and global initiatives like INSPIRE and GEOSS that will be further discussed.

For Masser (2006), the SDI hierarchy poses the challenge of a multistakeholder participation in SDI implementation because the bottom-up vision differs a lot from the top-down approach that is implicit in most of the SDI literature. The top-down approach emphasizes the need for standardization and uniformity while the bottom-up stresses the importance of diversity and heterogeneity due to the different aspirations of the various stakeholders. In consequence, it is necessary to find a consensus to ensure some measure of standardization and uniformity while recognizing the diversity and the heterogeneity of the different stakeholders performing different tasks at different levels.

3.5 SDI evolution and (emerging) trends

Different authors (Masser, 2005; Crompvoets, 2003) have studied the diffusion and evolution of SDI around the world and show that driving forces behind SDI initiatives are generally similar:

- promoting economic development,
- stimulating better government,
- fostering environmental sustainability,
- modernization of government,
- environmental management.

They all agree on the fact that, as of today, a critical mass of SDI users has been reached as a result of the diffusion of SDI concepts during the last ten to fifteen years. This provides a basic network of people and organizations that is essential for future development of SDIs.

Rajabifard *et al.* (2001) find that the first generation of SDIs, based on a product model, gave away to a second generation at the beginning of the years 2000, the latest being characterized by a process model. Indeed, the first generation of SDIs were product-based, aiming to link existing and future databases while the second generation aim to define a framework to facilitate the management of information assets allowing reuse of collected data by a wide range of people and/or organizations for a great diversity of purposes at various times. For Masser (2005) this evolution emphasis the shift from the concerns of data producers to those of data users and the shift from centralized structures to decentralized and distributed networks like the web.

The process-based model emphasizes the communication channel of knowledge infrastructure and capacity building towards the creation of an SDI facilitating cooperation and exchange of data and knowledge (Rajabifard and Williamson, 2001). They also highlight the fact that the characteristics of the social system strongly influence the approach taken to implement and develop a Spatial Data Infrastructure. They propose key issues and strategies to be considered for the design process:

- development of a strategic vision and associated implementation strategy,
- recognition that SDI is not an end in itself,
- key institutional strategy is to have all coordinating processes administered by one group.

Today's effort on the technical development of SDI components clearly focus on the exchange of geodata in an interoperable way (Bernard and Craglia, 2005) through services that allow efficient access to spatially distributed data. The shift towards an infrastructure offering services to answer requests rather than a “simple” network allowing to find, view and exchange geodata is highlighted by the concept of web services and the related Service Oriented Architecture (SOA).

Web services are a “new paradigm” (Comert, 2004) where different systems or providers offer some services for certain user groups, allowing an easy access to distributed geographic data and geoprocessing applications. The web services emphasize the necessity that systems involved could talk to each other and the provision of this talk should be easy and cost-effective for businesses to profit. In other words, web services relay on interoperability.

Web services enable the possibility to construct web-based application using any platform, object model and programming language. A service is no more than a collection of operations that allows users to invoke a service, which could be as simple as requesting to create a map or complicated as processing a remote sensing image.

In summary, web services are for application-to-application communication over internet and are based, in general, on open standards like XML (Comert, 2004). SOA is the basic principle concept supporting web services development. It promotes loose coupling between software components so that they can be reused (Sahin and Gumusay, 2008). In a SOA, the key component is services. They are well defined set of actions, self contained, stateless, and does not depend on the state of other services.

There are three components on the web services architecture: service provider, service requester and service broker and three operations: publish, find and bind. A SOA relates the three components to the three operations to allow automatic discovery and use of services.

In a traditional scenario, a service provider hosts a web service and “publishes” a service description to a service broker. The service requester uses a “find” operation to retrieve the service description and uses it to “bind” with the service provider and invoke the web service itself (fig.6).

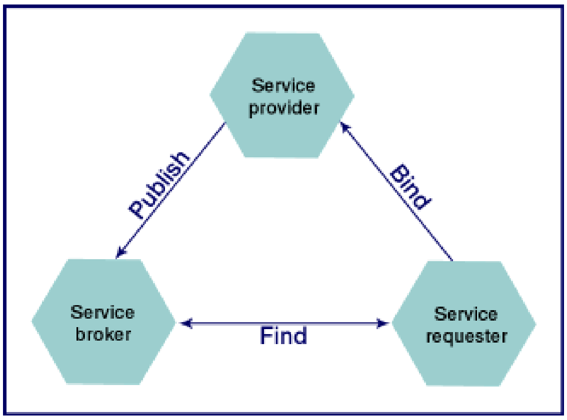


Fig.6: Basic operations in SOA (Source: IBM).

SOA is the underlying concept for an interoperable environment based on reusability and standardized components and thus is of high importance for SDIs allowing applications and related components to exchange data, share tasks, and automate processes over the Internet (OGC, 2004).

The OGC web services are, by far, the most important and relevant web services for our purpose and they will be discussed in details in an upcoming section of this deliverable (paragraph 6.1) .

With the advent of web services into the SDI community new trends/opportunities could be foreseen:

- Actual SDIs are lacking of analysis capabilities, an essential task to turn data into understandable information. This means that processing geospatial data is done in general on desktop computers and thus limit the analytical capacities caused by the huge execution time that geoprocessing tasks require to process a vast amount of data. With the recently introduced Web Processing Service and the promises of high storage and computing capacities offered by Grid infrastructures, new opportunities are emerging within geosciences communities (Padberg and Greve, 2009)
- Semantic web developed vocabularies (called ontologies) for geospatial data with the goal to increase understanding of such data by machines, allowing automated process through web services (Boes and Pavlova, 2008; Vaccari *et al.*, 2009).
- Web services are one aspect of the Web 2.0 revolution. The web 2.0 refers to a second generation of internet based services that allow people to collaborate and share information (Boes and Pavlova, 2007). GIS is also taking advantage of the web 2.0 revolution, with a good example being Google Maps that opened some of the more straightforward capabilities of GIS to the general public (Goodchild, 2007) and allowing, with other tools, the general public to create and generate news sources of data and information. This phenomenon is also know as Volunteered Geographic Information (VGI) (Boes and Pavlova, 2007; Coleman *et al.*, 2009, Craglia, 2007). VGI offers new opportunities and perhaps will influence the development of SDIs and the production of data in a near future.

Finally, it is necessary to mention that building an efficient SDI is almost impossible without partnership because a single agency is unlikely to have all resources, skills or knowledge to undertake the development of all aspects of SDI (UNECA, 2005; UNGIWG, 2007). This is why different authors (Williamson *et al.*, 2003, Rajabifard and Williamson, 2004) stress the importance of the capacity building component in the SDI implementation process.

3.6 Benefits

To conclude, we can highlight some of the (expected) benefits that SDIs can offer:

- universal (anywhere and anytime) access to geodata and related information,
- services and applications to discover and access distributed data sources,
- integration of different geospatial information to provide seamless visualization,

- seamless combination (chaining) of data, services and related applications,
- geospatial data update and maintenance made easy,
- sharing and reuse capabilities,
- collaborative activities,
- wide-scale interoperability, agreeing on open and common standards,
- development of partnerships, collaboration between different stakeholders.

The Canadian Geospatial Data Infrastructure (CGDI) claimed that developing applications using such an infrastructure allows to:

- Reduce costs: Applications can be built by reusing existing services.
- Reduce complexity: Service interfaces hide the underlying complexity.
- Permits less costly integration and interoperability: Standard interfaces simplify interconnection and integration.
- Allow direct access to current, authoritative source data.

Finally an effective and working SDI leads to:

- Inform decision-making: easy access to current information, knowledge and expertise.
- Increase efficiency: standards and specifications, as well as access to services, reduce duplication of effort.
- Enhance usability: providing reliable access to geospatial information to all levels, from the citizens to governments.
- Push for economic growth.

We would not depict SDIs as a "perfect tool" that can solve all problems. It is evident that SDIs represent a great opportunity and a framework with great perspective but as Masser reminds us (2005), SDIs can facilitate access to data to a wide range of users only if profound changes in "sharing spirit" take place. He also mentions the fact that building an SDI is a long term process. In order to be fully operational such a process depends on sustainability and commitment.

In addition, there are several others issues that could limit the implementation of SDI concepts such as: collaboration, funding, political stability, legislation, priorities, awareness, copyrights, privacy, licensing, capacity building and cultural issues. These are all clear challenges that EnviroGRIDS will have to face.

Technical	**Institutional**	**Policy**	**Legal**	**Social**
Computational heterogeneity	Collaboration models	Political stability	Rights, Restrictions	Cultural
Semantic	Funding model	Legislation	Copyright,	Capacity builiding
Reference sys.	Linkage between data units	Priorities/ Sustainable dev.	Intellectual Property rights	Equity
Data quality		Awarness of data existence	Data access and prices	
Metadata			Privacy	
Format			Licensing	

Table 1: Integration issues (Williamson et al., 2006)

To conclude this section, it is important to keep in mind that data sharing and related SDI developments rely first on individuals (Craig, 2005) that have many interests in common. First, their idealism, their sense that better data will lead to better decisions. Second, their self-interest: by sharing, they got something in return even if it is intangible, they are viewed as cooperative partners, and finally they are involved in a professional culture that honors serving society and cooperating with others.

4. Interoperability and standards

4.1 Definition and concepts

Previously, in section 2, we have seen that we are living in a world that is changing rapidly with communication means that are taking an increasing place in our everyday life. This communication revolution is mostly based on the Internet, whose successes are due to interoperability. Interoperability is *"the ability of a system or a product to work with other systems or products without special effort on the part of the customer."* (OGC, 2004). This means that two or more systems or components are able to transmit or exchange information through a common system and to use the information that has been exchanged.

When systems are interoperable, it gives the user the ability to:

- find what he needs,
- access it,
- understand and employ it,
- have goods and services responsive to his need.

As of today, in a climate of economic constraint, interoperability and standardization have never been so important because a non-interoperable system impedes the sharing of data, information and computing resources (OGC, 2004), leading organizations to spend much more than necessary on data, software and hardware. Moreover being non-interoperable increases the risk for a system or an infrastructure to not deliver its expected benefit and in consequence to lead to a user disappointment and system failure.

In order to achieve interoperability, there are two approaches:

- adhering to standards
- making use of a "broker" of services that can convert one product's interface into another product's interface, "on the fly".

One good example of the first approach is the Web, where standards like HTTP, TCP/IP or HTML have been developed by organizations that wish to create standards to *"meet everyone's needs without favoring any single company or organization"* (OGC, 2005).

The great advantage of interoperability, and that is why it is an essential building block for the GIS and SDI industry, is that it describes the ability of locally managing and distributing heterogeneous systems to exchange data and information in real time to provide a service (OGC, 2004). This allows the users to maximize the value of past and future investments in geoprocessing systems and data.

As a response to the need of GIS standards to support interoperability, the OGC aims to tackle the non-interoperability caused by the diversity of systems creating, storing, retrieving, processing and displaying geospatial data in different formats. In addition to this, software vendors often did not communicate among themselves to agree on how data should be structured and stored and how systems must exchange information, leading inevitably to a non-interoperable environment, isolating geospatial data in "electronic silos" and resulting in expensive duplication of data and difficulty in sharing and integrating information (OGC, 2004).

The OGC (2005) has pointed out the general user needs:

- Need to share and reuse data in order to decrease costs (avoid redundancy collection), obtain additional or better information, and increase the value of data holdings.
- Need to choose the best tool for the job and the related need to reduce technology and procurement risks (avoid being locked in to one vendor).
- Need to leverage investment in software and data, enabling more people to benefit from using geospatial data across applications without the need for additional training.

The OGC believes that responding to user needs of interoperability will have a profound and positive impact in the public and private sectors, opening the doors of new business opportunities and new human activities.

In summary, interoperability enhances: communication, efficiency and quality for the benefit of all citizens allowing them to access data in a good, consistent and transparent way.

4.2 Types of interoperability

There are two types of interoperability (OGC, 2004):

- syntactic (or technical): when two or more systems are capable of communicating and exchanging data, they are exhibiting syntactic interoperability. Specified data formats and communication protocols are fundamental. In general, XML or SQL standards provide syntactic interoperability. Syntactical interoperability is required for any attempts of further interoperability.
- Semantic: Beyond the ability of two or more computer systems to exchange information, semantic interoperability is the ability to automatically interpret the information exchanged meaningfully and accurately in order to produce useful results as defined by the end users of both systems. To achieve semantic interoperability, both sides must defer to a common information exchange reference model. The content of the information exchange requests are unambiguously defined: what is sent is the same as what is understood (eg. explaining why INSPIRE is producing data specifications).

Different types of geoprocessing systems (vector, raster, CAD, etc.) producing different types of data, different vendors geoprocessing systems using internal data formats and producing proprietary formats, different vendors systems using proprietary libraries and interfaces and reducing the possibilities of communication between systems... are all causes of syntactic non-interoperability while different data producers using different metadata schemas and/or different naming conventions lead to semantic non-interoperability.

The World Wide Web and its associated technologies offer a great opportunity to overcome both syntactic and semantic non-interoperability because it is an almost universal platform for distributed computing and it provides facilities to semantically process structured text. The web is thus a key enabler for interoperability, by increasing access to geospatial data and processing resources, which in consequence increases the value of those resources (OGC, 2004).

To ensure effective interoperability, it is not only a matter of technology but also and often it requires a change of philosophy, of spirit to go “open”. This is classified under human or legal/policy on the following table summarizing the different types of interoperability.

Technical	**Semantic**	**Human**	**Legal/Policy**
Machine to machine connections	Common understanding concepts, terms, ...	Cooperation	Digital rights, ownerships
Software interaction	Inter-disciplinary vocabularies	Training	Responsibility
APIs			
Formats			

Table 2: Different types of interoperability.

As expressed by the OGC (2004) *“if an organization does not fully embrace the tenets of interoperability and interoperable architectures, then long-term success in integrating geospatial processes into an organization's overall business processes may be problematic”.*

In consequence, organizations will need:

- commitment to interoperability and geospatial standards,

- commitment to collaboration,
- commitment to define a geospatial interoperability and information framework that meets the requirements of the organization,
- commitment to the collection and maintenance of geospatial metadata,
- commitment to training and education

Through all these commitments, an organization will be truly interoperable, maximizing the value and reuse of data and information under its control and will be able to exchange these data and information with other interoperable systems, allowing new knowledge to emerge from relationships that were not envisioned previously.

4.3 Interoperability enablers

To enable effective interoperability, we have already seen that the Internet and standards are probably the most important components at a technical level but here are a lot of other possible enablers, both human and technical, that could help an organization to promote its commitment to interoperability:

- web and networks,
- standards,
- infrastructure,
- metadata,
- support for multiple: languages, views, data formats, projections, datums,...
- sharing of best practices,
- cooperation and collaboration,
- business models,
- business agreements,
- policy framework,
- copy and access rights,
- authorization,
- ...

Altogether they will contribute in a way or another to a successful implementation of the geospatial interoperability by reaching a consensus between the users' need for compatibility with the autonomy and heterogeneity of the inter-operating systems (OGC, 2005).

4.4 Standards

Standards are documented agreements, used in public contracts or international trade, containing technical specifications or other precise criteria to used consistently as rules, guidelines, or definitions of characteristics, to ensure that materials, products, processes and services are fit for their purpose (Ostensen, 2001). In other words, standardization means agreeing on a common system (OGC, 2005).

The existence of non-harmonized standards for similar technologies contribute to the so-called “technical barriers to trade”, avoiding a user to share data, information or services.

Although developing standards is a long and complex process, involving many organizations, based on a consultative approach and aiming to find a consensus between all the parties involved (UNECA, 2005), organizations and agencies are increasingly recognizing that standards are essential for improving productivity, market competitiveness, export capabilities (GSDI, 2004), and lowering maintenance and operation costs over time (Booz *et al.*, 2005; Craglia and Nowak, 2006; Almirall *et al.*, 2008).

We can summarize the functions of standards as follow:

- help to ensure interoperability,
- promote innovation, competition, commerce and free trade,
- increase efficiency,

- make things work,
- affect every aspect of our life (widespread use of standards).

4.5 Benefits

After discussing what are standards and interoperability, we can give an overview of the (expected) benefits of a truly interoperable architecture.

1. Allow sharing and reusing of data: gives access to distributed and heterogeneous sources of data.
2. Avoid data duplication: data are collected and maintained at the most appropriate place.
3. Reduce the costs: of maintenance, of operations and of course of production.
4. Integration: As Mohammadi *et al.* (2007) shows multi-source data integration could only be achieved with an effective interoperability.
5. Reduce the complexity: through common knowledge, standards offer a set of rules that every data provider can follow, understand and become familiar with. Moreover when a user shares a data in a standardized way, another will be immediately able to use it.
6. Increase efficiency.
7. Vendor neutral: avoid the fact to be locked in to one vendor.
8. Improve decision-making: offering standardized access to a vast amount of data and information and to used them as effectively as they should.
9. New opportunities and knowledge: open the doors to new activities and relations that were not foreseen before.

Finally, as OGC (2005) stated:

"Changing internal systems and practices to make them interoperable is a far from simple task. But the benefits for the organization and for those who make use of information it publishes are incalculable".

5. Initiatives

Different initiatives at the regional and global level are influencing and promoting the creation of Spatial Data Infrastructures and the use of open standards. They all are concerned about data access, harmonization, standardization, interoperability, seamless integration and services. They coordinate actions that promote awareness and implementation of complimentary policies, common standards and effective mechanisms for the development and availability of interoperable digital geographic data and technologies to support decision making at all scales for multiple purposes. These actions encompass the policies, organizational structures, data, technologies, standards, delivery mechanisms, and financial and human resources necessary to ensure that those working at the global and regional scale are not impeded in meeting their objectives.

5.1 Infrastructure for Spatial Information in the European Community (INSPIRE)

Website: http://inspire.jrc.ec.europa.eu/

The Infrastructure for Spatial Information in the European Community, namely INSPIRE, is of particular interest for the EnviroGRIDS project. INSPIRE is a European Directive (entered into force in May 2007 and fully operational by 2019) that aims to create a European Union Spatial Data Infrastructure. This will enable the sharing of environmental spatial information among public sector organizations and better facilitate public access to spatial information across Europe (EU, 2007). When fully implemented, it will, theoretically enable data from one Member State to be seamlessly combined with data from all other States. This is particularly important for activities relating to the environment.

The main purpose of INSPIRE is to support the formulation, implementation, monitoring, and evaluation of Community environmental policies (EU, 2008). Therefore the spatial information considered under the directive is extensive and includes a great variety of topical and technical themes and will be based on Spatial Data Infrastructures established and operated by the Member States.

This initiative wishes to overcome the barriers affecting data access and exchange in Europe, including (EU, 2008):

- Inconsistencies in collection of georeferenced data: geodata are often missing and/or incomplete, or are collected twice by different organizations.
- Lacking of documentation, description (metadata) of the data.
- Geodata are often incompatible and thus cannot be combined.
- Infrastructures used to find, access and use geodata often function in isolation and are incompatible.
- Barriers to sharing: cultural, linguistic, institutional, financial and legal.

In order to overcome these barriers, it has been recognized that it would be necessary to develop a legislative framework asking the Member States to coordinate their activities and to agree on a set of requirements, common standards and processes. In consequence, INSPIRE is unique in the sense that it is an important collaborative and participative process to formulate the directive, create implementing rules and develop relative specifications and services.

INSPIRE seeks to create a European SDI and the INSPIRE Directive defines it: “infrastructure for spatial information means metadata, spatial data sets and spatial data services; network services and technologies; agreements on sharing, access and use; and coordination and monitoring mechanisms, processes and procedures, established, operated or made available in accordance with this Directive”. (EU, 2004)

The end users of INSPIRE include policymakers, planners and managers at the local, national and regional levels, and the citizens and their organizations.

INSPIRE is based on common principles (EU, 2007):

1. Data should be collected only once and kept where it can be maintained most effectively.
2. It should be possible to combine seamless spatial information from different sources across Europe and share it with many users and applications.
3. It should be possible for information collected at one level/scale to be shared with all levels/scales; detailed for thorough investigations, general for strategic purposes.
4. Geographic information needed for good governance at all levels should be readily and transparently available.
5. Easy to find what geographic information is available, how it can be used to meet a particular need, and under which conditions it can be acquired and used.

A step by step approach is used to implement and develop the infrastructure because such an initiative cannot be built from one day to another and is asking Member States to drastically change their existing infrastructure. Thus the implementation of services has been stated just after the adoption of the Directive, whereas the harmonization of INSPIRE data themes will be made in three phases up to 2013.

The European Commission Joint Research Center (JRC) plays a major role in this initiative as it has supported the development of the proposal and now endorses the responsibility of the overall technical coordination of the Directive, providing support to the preparation of the technical rules on implementation, data harmonization, documentation and the required services to discover, view and download data.

The Directive provides five sets of Implementing Rules (IR) that set out how the various elements of the system (metadata, data sharing, data specification, network services, monitoring and reporting) will operate and to ensure that the spatial data infrastructures of the Member States are compatible and usable in a Community and transboundary context. The Drafting Teams now working on these

IRs are composed of international experts and the process includes open consultation – particularly with Spatial Data Interest Communities (SDIC) and Legally Mandated Organizations (LMO).

The Directive specifically states that no new data will need to be collected. However it does require that two years after adoption of the Implementing Rules for data sets and their related services each Member State will have to ensure that all newly collected spatial data sets are available in conformity with the IR. Other data sets must conform to the Rules within 7 years of their adoption. Implementing Rules will be adopted in a phased manner between 2008 and 2012 with compliance required between 2010 and 2019 (EU, 2008).

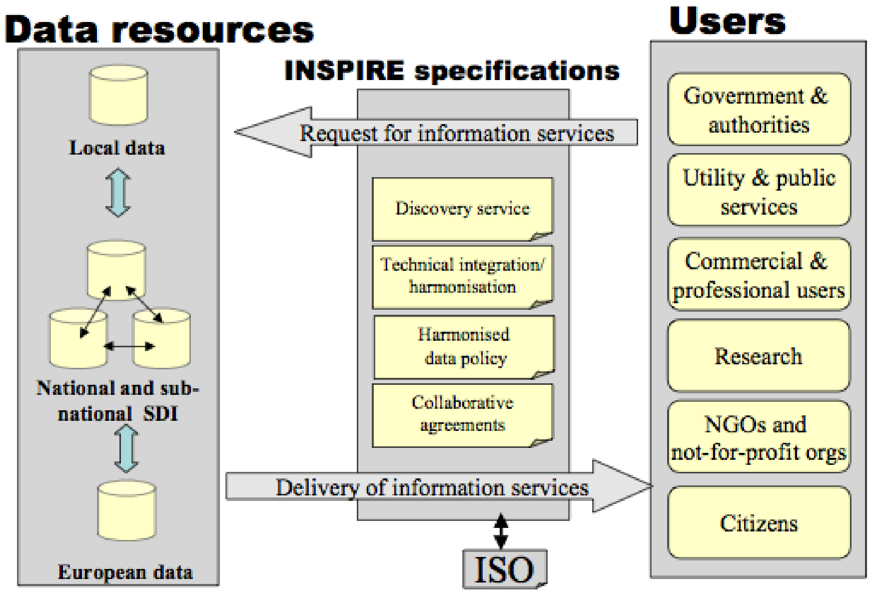


Fig.7: Data and information flow within the INSPIRE framework (Source: INSPIRE).

The envisioned interoperability in INSPIRE is a possibility offered to the user to combine geospatial data and services from different sources across the European Community in a consistent way without involving specific efforts of humans or computers (fig.7). Thus users will spend less time and efforts to integrate data delivered within the INSPIRE framework.

The Directive (EU, 2007) defines 34 “spatial data themes” that have been defined in three Annexes sorted in order of priority. Annex 1 datasets cover the ‘basic’ spatial building blocks such as spatial referencing systems, geographic names, addresses, transport networks, hydrography and land parcels. Because of the range of data types involved, the impact of INSPIRE is comprehensive. Annex 1 datasets have to be prepared and made available from 2011, with the other Annexes at later dates. In order to enable full system interoperability across the EU, each spatial data theme is described in a data specification. As mentioned on the INSPIRE website *“The process for developing harmonized data specifications is designed to maximize the re-use of existing requirements and specifications, in order to minimize the burden for Member States’ organizations at the time of implementation. The consequence of this is that the process of developing Implementing Rules for interoperability of spatial datasets and services may be perceived as being complex: it involves a large number of stakeholders, with many interactions and consultations”*.

Finally, all the data, information and services shared within INSPIRE would be accessible through the INSPIRE Community Geoportal. For Luraschi *et al.* (2009) because the geoportal does not store or maintain data and metadata, it could be seen as a gateway aggregating a number of instances of specific geospatial information services distributed across the Europe and maintained by the organization responsible for the data.

According to the INSPIRE network architecture (EU, 2008), Member States shall establish, operate and provide access to the following network services (fig.8):

- discovery services: support discovery of data, evaluation and use of spatial data and services through their metadata properties
- view services: as a minimum, display, navigate, zoom in/out, pan, or overlay spatial data sets and display legend information and any relevant content of metadata.
- download services: enabling copies of complete spatial data sets, or parts of such sets, to be downloaded.
- transformation services: enabling spatial data sets to be transformed (projection and harmonization).
- invoke spatial data services: enabling data services to be invoked.

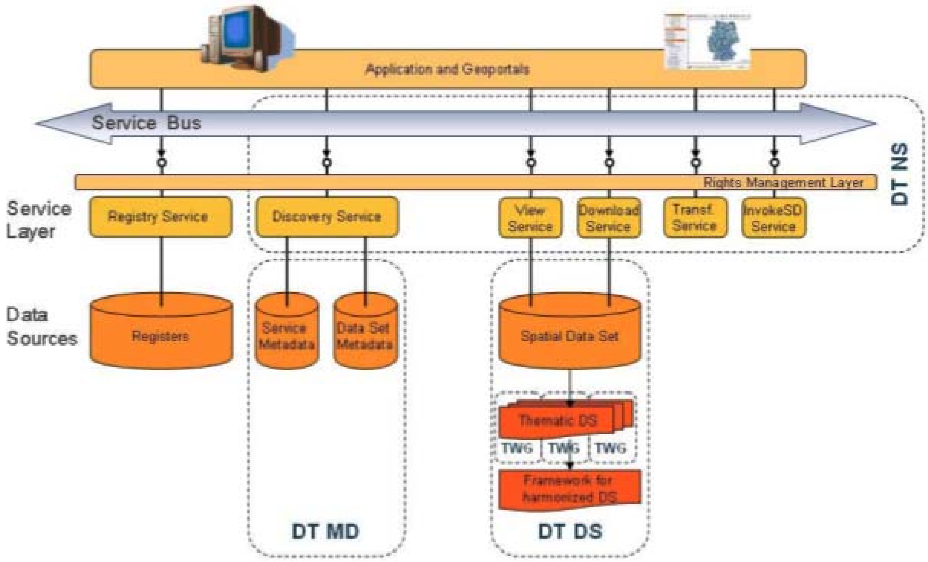


Fig.8: INSPIRE network architecture (Source: INSPIRE).

5.2 Global Earth Observation System of Systems (GEOSS)

Website: http://www.earthobservations.org

The GEOSS is being established by the intergovernmental Group on Earth Observations (GEO) and is a worldwide effort to build a system of systems on the basis of a 10-Year Implementation Plan for the period 2005 to 2015 (GEO, 2005). GEO is voluntary partnership of governments and international organizations where membership and participation is contingent upon formal endorsement of the Implementation Plan mentioned above.

GEOSS is an effort to connect already existing SDIs and Earth Observations infrastructures and thus will not create and/or store data but rather works with and build upon existing systems. GEOSS, through its developing GEOportal, is foreseen to act as a gateway between the producers of environmental data and the end users, with the aim of enhancing the relevance of Earth observations for the global issues and to offer a public access to comprehensive, near-real time data, information and analyses on the environment (GEO, 2007).

GEOSS aims to provide a broad range of so-called Societal Benefits Areas (GEO, 2005):

- Reducing loss of life and property from natural and human-induced disasters,
- Understanding environmental factors affecting human health and well-being,
- Improving the management of energy resources,

- Understanding, assessing, predicting, mitigating, and adapting to climate variability and change,
- Improving water resource management through better understanding of the water cycle,
- Improving weather information, forecasting and warning,
- Improving the management and protection of terrestrial, coastal and marine ecosystems,
- Supporting sustainable agriculture and combating desertification, and
- Understanding, monitoring and conserving biodiversity.

The mechanisms for data and information sharing and dissemination are presented and described in the 10-Year Implementation Plan Reference Document (GEO, 2005) where information providers must accept and implement *"a set of interoperability arrangements, including technical specifications for collecting, processing, storing , and dissemination shared data, metadata and products. GEOSS interoperability will be based on non-proprietary standards, with preference to formal international standards. Interoperability will be focused on interfaces, defining only how system components interface with each other and thereby minimizing any impact on affected systems".* GEOSS is based on existing technologies using internet-based services.

Moreover members must fully endorse the following data sharing principles:

1. There will be full and open exchange of data, metadata, and products shared within GEOSS, recognizing relevant international instruments and national policies and legislation.
2. All shared data, metadata, and products will be made available with minimum time delay and at minimum cost.
3. All shared data, metadata, and products being free of charge or no more than cost of reproduction will be encouraged for research and education.

These principles push data owners to go "open" and to share their data using standards and thus becoming interoperable.

5.3 United Nations Spatial Data Infrastructure (UNSDI)

Website: http://www.ungiwg.org/unsdi.htm

The United Nations Spatial Data Infrastructure in an initiative conducted by the United Nations Geographic Working Group (UNGIWG) that aims at building an institutional and technical mechanism to establish a coherent system to exchange data and services concerning geospatial data and information within the United Nations, and also supporting SDI development activities in the Member Countries.

As stated in the UNSDI Compendium (UNGIWG, 2007), *"Historically, the production and use of geospatial data have been accomplished within the United Nations by its component organizations, in accordance with their individual needs and expertise. But concordant with the recent, rapid increase in the use of geospatial data for UN activities is the need for greater coherence in its management system-wide".*

This initiative aims to contribute to the general mission of the United Nations to maintain peace and security, to address humanitarian emergencies, to assist sustainable development and support achievement of the UN Millennium Development Goals. The hope is to facilitate efficient access, exchange and utilization of georeferenced information in order to make the UN system more effective, increase the system coherence and support its "Delivering as One" policies.

The UNSDI provides an institutional and technical foundation of policies, interoperable standards procedures and guidelines that enable organizations and technologies to interact in a way that facilitates spatial discovery, evaluation and applications (UNGIWG, 2007).

5.4 Global Monitoring for the Environment and Security (GMES)

Websites: http://www.gmes.info/ and http://ec.europa.eu/gmes/

The GMES is a European programme, coordinated by the European Commission and European Space Agency, for the implementation of a European capacity for Earth observation with the objective to monitor and better understand the environment and thus contribute to the security of every citizen. This initiative aims at providing decision-makers and other users who rely on strategic information with regard to environmental and security issues an autonomous, independent and permanent access to timely, reliable and accurate data and services.

The objective is to integrate data on atmosphere, oceans and continental/land processes giving an overview of the state of health of our planet and to deliver information through five thematic areas (served by different services) covering:

- land,
- marine,
- emergency,
- atmosphere,
- security.

allowing policy and decision-makers to prepare legislation (at different level) on environmental topics and to monitor the implementation of such laws.

To gather data and information on Earth Observation, GMES proposes to build and infrastructure around four components:

- space: environmental satellites
- in-situ measurements: ground-based and airborne sensors
- data harmonization and standardization
- services to users.

Like various other data sources, Earth-observation-based services already exist in Europe but are dispersed and fragmented at national and regional level avoiding a sustainable observation capacity (meaning that long-term availability of information is not guaranteed). In consequence, GMES is the answer of the European Commission to develop a reliable and sustainable Earth Observation system and also contributing the GEOSS initiative.

GMES stated that *"By securing the sustainability of an information infrastructure necessary to produce output information in the form of maps, datasets, reports, targeted alerts, etc…, GMES helps people and organisations to take action, make appropriate policy decisions and decide on necessary investments. GMES also represents a great potential for businesses in the services market, which will be able to make use of the data and information it provides according to a full and open access principle."* (GMES, 2009).

5.5 Global Spatial Data Infrastructure (GSDI)

Website: http://www.gsdi.org

The mission of the GSDI Association, a world-wide inclusive body of organizations, agencies, firms and people, is to "*promote international cooperation and collaboration in support of local, national, and international spatial data infrastructure developments that will allow nations to better address social, economic and environmental issues of pressing importance"* (GSDI, 2004).

Its purpose is to focus on communication, education, scientific, research and partnership efforts to support all societal needs for access to and use of spatial data.

This is an association, guided by a board and funded by membership fees, to:

- serve as a point of contact and effective voice for those in the global community involved in developing, implementing and advancing spatial data infrastructure concepts,
- foster spatial data infrastructures that support sustainable social, economic, and environmental systems integrated from local to global scales, and

- promote the informed and responsible use of geographic information and spatial technologies for the benefit of society.

The GSDI community aims to truly develop and achieve the goal of a Global Spatial Data Infrastructure relying on international and open standards, guidelines and policies to enhance data management and access, and support global economic growth, and associated social and environmental objectives (UNGWIG, 2007), through interoperable standards-based services, systems, software and products that operate in a web-based environment.

This vision is guided by five goals (Stevens, 2005):

- Continue to promote and develop awareness and exchanges on infrastructure issues for all relevant levels from local to global.
- Promote and facilitate standards-based data access/discovery through the Internet.
- Promote, encourage, support, and conduct capacity building.
- Promote and conduct SDI development research.
- Collaborate with others to accomplish its Vision and Goals.

To support this vision, the GSDI association acts as a platform and offers a vast choice of publications, conferences, workshops, projects and programs allowing people interested in SDI to learn, exchange, share their knowledge and expertise, because capacity building is one of the key points of SDIs.

6. Standards organizations relevant for GIS/SDI

In the field of geomatics there are several organizations involved in publishing standards to effectively achieve the goal of interoperability. Such standards are increasingly important in the geospatial community allowing the increase of interoperability between systems and data and thus to "geo-enable" the Web.

6.1 Open Geospatial Consortium (OGC)

Website: http://www.opengeospatial.org

The Open Geospatial Consortium (OGC) is a non-profit, international, voluntary consortium of more than 380 companies, government agencies and universities that is leading the development of standards for the geospatial community. Its approach is based on a member-driven consensus process to develop open and publicly available standards and software application programming interface for the geospatial community (UNGIWG, 2007). These standards offer to the developers the possibility to create complex georeferenced information and services accessible to a wide variety of applications and share data in a standardized and interoperable way.

The OGC standards are based on a generalized architecture presented in the Abstract Specification and Reference Model (OGC, 2007). On top of the Abstract Specification, there is a set of standards that have been developed and/or proposed to serve specific needs of the Geographical Information community in order to be interoperable.

These standards are mostly based upon the use of the http protocol to interact through messages over the Internet. In the last two years, the OGC members have been looking with interest to a more common approach used in the Service Oriented Architecture using SOAP protocol and WSDL bindings. There is also work in progress around the Representational State Transfer (REST) protocol for web services.

The OGC is also closely working with the International Organization for Standardization (ISO) through a partnership with the ISO Technical Committee 211 (TC211) to promote and endorse the OGC standards to a higher level of standardization becoming part of the ISO 19100 series. For example, the WMS or the GML are now ISO standards.

The OGC vision is the realization of full societal, economic and scientific benefits of integrating electronic location resources into commercial and institutional processes worldwide. Its mission is to serve as a global forum for the collaboration of developers and users of spatial data products and services, and to advance the development of international standards for geospatial interoperability.

More specifically the OGC aims to (http://www.opengeospatial.org/ogc/vision):

- Provide free and openly available standards to the market, tangible value to Members, and measurable benefits to users.
- Lead the creation and establishment of standards that allow geospatial content and services to be seamlessly integrated into business and civic processes, the spatial web and enterprise computing.
- Facilitate the adoption of open, spatially enabled reference architectures in enterprise environments worldwide.
- Advance standards in support of the formation of new and innovative markets and applications for geospatial technologies.
- Accelerate market assimilation of interoperability research through collaborative consortium processes.

It must be noticed that the OGC and its members want to help users and developers to make usage of OGC's standards offering them different resources (e.g. technical documents, training, best practices) through the OGC Network (http://www.ogcnetwork.net/).

6.2 International Organization for Standardization (ISO)

Website: http://www.iso.org

The International Organization for Standardization (ISO), the world's largest developer, publisher and promoter of international standards, is a non-governmental organization made of a network of more than 160 countries representatives (one per country) with a central secretariat based in Geneva (Switzerland) that coordinates the system.

Even if the main focus of ISO is the development of technical standards, they have an important impact on the economy and the society because many members are coming from a governmental structure or from the private sector. Therefore, ISO is an ideal place to build consensus and solutions that meet the needs and requirements from both the economy and the society.

Within the ISO system there is a Technical committee (http://www.isotc211.org/) whose main area of interest are the Geographical information and the Geomatics aiming to establish a structured set of standards for information concerning objects or phenomena that are directly or indirectly associated with a location relative to the Earth.

At the present day, they have more than 55 standards in the field of Geographical information specifying methods, tools and services for data management (including definition and description), acquiring, processing, analyzing, accessing, presenting and transferring such data in digital/electronic form between different users, systems and locations.

6.3 The World Wide Web Consortium (W3C)

Website: http://www.w3c.org

The World Wide Web Consortium (W3C) is an international consortium where members, organizations, staff and the public work together to create and develop Web standards, protocols and guidelines. Since 1994, the W3C has published more that 110 Recommendations (the W3C standards) aiming to " *lead the World Wide Web to its full potential by developing protocols and guidelines that ensure long-term growth for the Web",* accommodating the growing diversity of people, hardware and software and ensuring the core principles and components of these standards would be supported by everyone.

For the W3C it is crucial to reach the goal of the "web interoperability" allowing the Web to reach its full potential by using technologies that must be compatible with one another and allowing any hardware and/or software to access the Web. By publishing open and non-proprietary standards, the W3C seeks to avoid market fragmentation and thus Web fragmentation.

Therefore the W3C engages in education and outreach, develops software and interoperable technologies that support this mission and acts as an open and vendor-neutral forum for discussion about the Web.

6.4 Organization for the Advancement of Structured Information Standards (OASIS)

Website: http://www.oasis-open.org/

The Organization for the Advancement of Structures Information Standards is a non-profit, global consortium (with 5000 members coming from more than 600 organizations in 100 countries) that drives the development, convergence and adoption of open standards for the global information society, the so-called e-business. OASIS produces different web standards concerning the following categories: Web Services, e-Commerce, Security, Law & Government, Supply Chain, Computing Management, Application Focus, Document-Centric, XML Processing, Conformance/Interop, and Industry Domains.

7. Standards description

ISO and OGC are providing a lot of different specifications regarding geographical data but in the context of the EnviroGRIDS project we propose to concentrate on those that are mostly used in the geospatial community. The aim is to place the first building blocks of a regional SDI for the Black Sea catchment.

The general aim of these standards is to abstract data delivery mechanisms from physical storage formats and offer services that could be consumed by applications through different interfaces.

The OGC defines a general architecture for the geoportal (OGC, 2004) called The Geospatial Portal Reference Architecture. It provides the basis for an open, vendor-neutral portal that is intended to be a first point of discovery for geospatial content in the context of designing and implementing the Spatial Data Infrastructures. The Geospatial Portal Reference Architecture is founded on the tenants of a Service Oriented Architecture (SOA). An SOA is an architecture that represents software functionality as discoverable services on a network yielding the following benefits:

- Easier extension of legacy logic to work with new business functionality
- Greater flexibility to change without the need to constantly re-architect for growth
- Cost savings by providing straight-forward integration.

The Geospatial Portal Reference Architecture specifies also four services that are needed for creating a interoperable and standardized geoportal (fig.9):

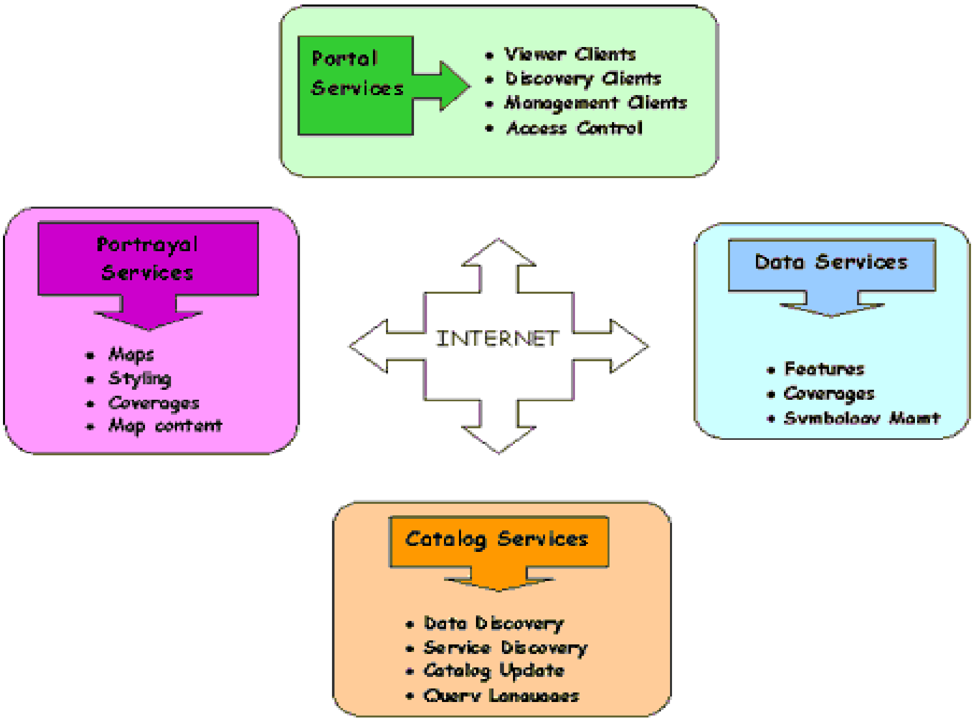


Fig.9: Geospatial Portal Reference Architecture (Source: GeoNetwork).

- Portal Services: provide the single point access to the geospatial information on the portal. In addition, these services provide the management and administration of the portal.
- Catalog Services: used to locate geospatial services and information wherever it is located and provide information on the services and information if finds to the user.
- Portrayal Services: used to process the geospatial information and prepare it for presentation to the user.
- Data Services: used to provide geospatial content and data processing.

To implement and deploy these different service classes, the OGC propose to use web services that are applications running on a computer connecting to a remote web service via a URL allowing access to distributed data and services. As stated by the CGDI "*Web service architectures provide a distributed environment in which you can deploy and invoke services using standard Internet protocols. In this context, a service is a collection of operations, accessible through one or more interfaces, that allows you to evoke a behavior of value to you.*" (http://www.geoconnections.org/publications/Technical_Manual/html_e/s4_ch10.html#10.1)

Using such a Service Oriented Architecture (SOA) provides a distributed computing platform over a network, typically the Internet, allowing to publish standardized services no matter how it is implemented or on which platform it is executed. This is leveraging the full potential of the interoperability and thus web services to be seamlessly coupled, reusable and available for a variety of applications.

A traditional open web service must have the ability to describe its capabilities and provide a standard way to communicate with it, enabling applications and other web services to communicate and interoperate. Through OGC standards, different GIS softwares and/or components can interoperate, work together and exchange information over a network by means of agreed standards.

For example, when two softwares implement the same OGC standard, they can immediately work together without the necessity to develop new components to translate from one file format (used in one software) to another file format (used in a second software). This means that in a SOA environment that implement OGC standards, a user can access in a transparent way the data stored in different databases, with different formats, and running on different Operating Systems.

Without interoperability and standardization, accessing and integrating different data sources is really difficult or in the worst case impossible. This leads to a fragmentation of geospatial data sources and limit organizations to work only within a single software package.

7.1 Catalogue Service for the Web (CSW)

OGC Catalogue Service Specification: http://www.opengeospatial.org/standards/cat

The Catalogue Service defines an interface to publish, discover, search and query metadata about georeferenced data, services and related resources. CSW uses queryable properties, which enable clients to search for geospatial resources by subject, title, abstract , data format, data type, geographic extent, coordinate reference system, originator, publisher, purpose,...

Fig.10: GeoNetwork, a catalogue system using CSW.

7.2 Web Map Service (WMS)

OGC Web Map Service Specification: http://www.opengeospatial.org/standards/wms

The Web Map Service defines an interface that allows a client to retrieve maps of georeferenced data. In WMS context, a map means a graphical representation (jpeg ,gif or png files) of a geospatial data meaning that a WMS service does not give access to the data itself. It is used for mapping purposes and can be combined with other WMS services.

A traditional WMS interface, invoked by a URL, consists of the following operations:

- GetCapabilites: answer to a client telling him what kind of layers are available and which one are queryable.
- GetMap: produce a map as a picture showing selected layers,
- GetFeatureInfo: answer simple queries about the content of the map

As seen on the following examples, invoking a WMS service need to specify different parameters (mandatory or optional) in the URL. For the purpose of this guideline we will focus our attention on the basic operations of the service that provides map layers in predefined styles (made available by the data provider) thus we will not discuss the Styled Layer Descriptor (SLD) capabilities.

- Example of a WMS URL with a GetCapabilites request:

```
http://metafunctions.grid.unep.ch:8080/geoserver/ows?service=WMS&request=Ge
tCapabilities&version=1.3.0
```

The GetCapabilites operation returns to the user an XML document describing the service and the data sets available from which either desktop and/or web clients may request maps. This operation is common for all the OWS and is presented in details in the OpenGIS Web Service Common Implementation Specification (OGC, 2007). To invoke the operation, the user has only to define the service and the request parameters.


```
- <wfs:WFS_Capabilities version="1.1.0" xsi:schemaLocation="http://www.opengis.net/wfs http://preview.grid.unep.ch:8080/geoserver/schemas
  /wfs/1.1.0/wfs.xsd" updateSequence="262">
  - <ows:ServiceIdentification>
      <ows:Title>enviroSDI Web Feature Service</ows:Title>
    - <ows:Abstract>
        enviroSDI is the Spatial Data Infrastructure of the UNEP/DEWA/GRID-Europe (http://www.grid.unep.ch). This is the reference
        implementation of WFS 1.0.0 and WFS 1.1.0, supports all WFS operations including Transaction.
      </ows:Abstract>
    - <ows:Keywords>
        <ows:Keyword>enviroSDI</ows:Keyword>
        <ows:Keyword>UNEP</ows:Keyword>
        <ows:Keyword>GRID</ows:Keyword>
        <ows:Keyword>EUROPE</ows:Keyword>
        <ows:Keyword>WFS</ows:Keyword>
        <ows:Keyword>GEOSERVER</ows:Keyword>
      </ows:Keywords>
      <ows:ServiceType>WFS</ows:ServiceType>
      <ows:ServiceTypeVersion>1.1.0</ows:ServiceTypeVersion>
      <ows:Fees>NONE</ows:Fees>
      <ows:AccessConstraints>NONE</ows:AccessConstraints>
    </ows:ServiceIdentification>
  - <ows:ServiceProvider>
      <ows:ProviderName>UNEP/DEWA/GRID-Europe</ows:ProviderName>
    - <ows:ServiceContact>
        <ows:IndividualName>Gregory Giuliani</ows:IndividualName>
        <ows:PositionName>enviroSDI coordinator</ows:PositionName>
      - <ows:ContactInfo>
        - <ows:Phone>
            <ows:Voice>+41 22 917 84 17</ows:Voice>
            <ows:Facsimile/>
```


Fig.11: Example of the XML file returned after a GetCapabilites request.

- Example of a WMS URL with a GeMap request:

```
http://preview.grid.unep.ch:8080/geoserver/wms?bbox=84.95293,19.82194,-
74.13126,23.19403&styles=&Format=image/png&request=GetMap &version=1.1.1
&layers=preview:cy_buffers&width=640&height=309 &srs=EPSG:4326
```

The GetMap operation is the most important of the three basic operations of the WMS interface as it returns to a client request a map of selected geospatial layers.

In comparison of the GetCapabilities request that needs only two parameters, we can see on the above example that GetMap operation needs several parameters (also mandatory or optional) that we describe hereafter:

Mandatory parameters for the GetMap operation:

- BBOX: coordinates of the bounding box following minx,miny,maxx,maxy,
- STYLES: list of style names separated by a comma. It's necessary to have an exact correspondence between the number of layers and the number of styles. If this parameter has a empty value, the default style provided by the data custodian will be applied.
- FORMAT: graphical format of the returned map (eg: image/png, image/gif, image/jpeg).
- REQUEST: value "GetMap", this the request itself to invoke the specific operation.
- VERSION: the version of the specification.
- LAYERS: list of selected layers separated by a comma.
- WIDTH: specify the width of the returned map (in pixels).
- HEIGHT: specify the height of the returned map (in pixels).
- SRS: identifier of the Spatial Reference System.

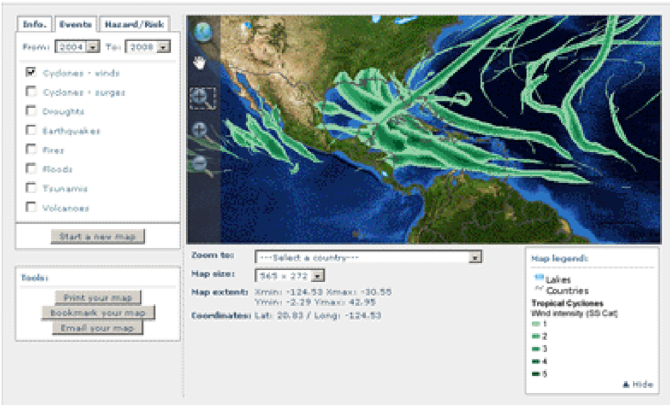


Fig.12: Returned image after a WMS request.

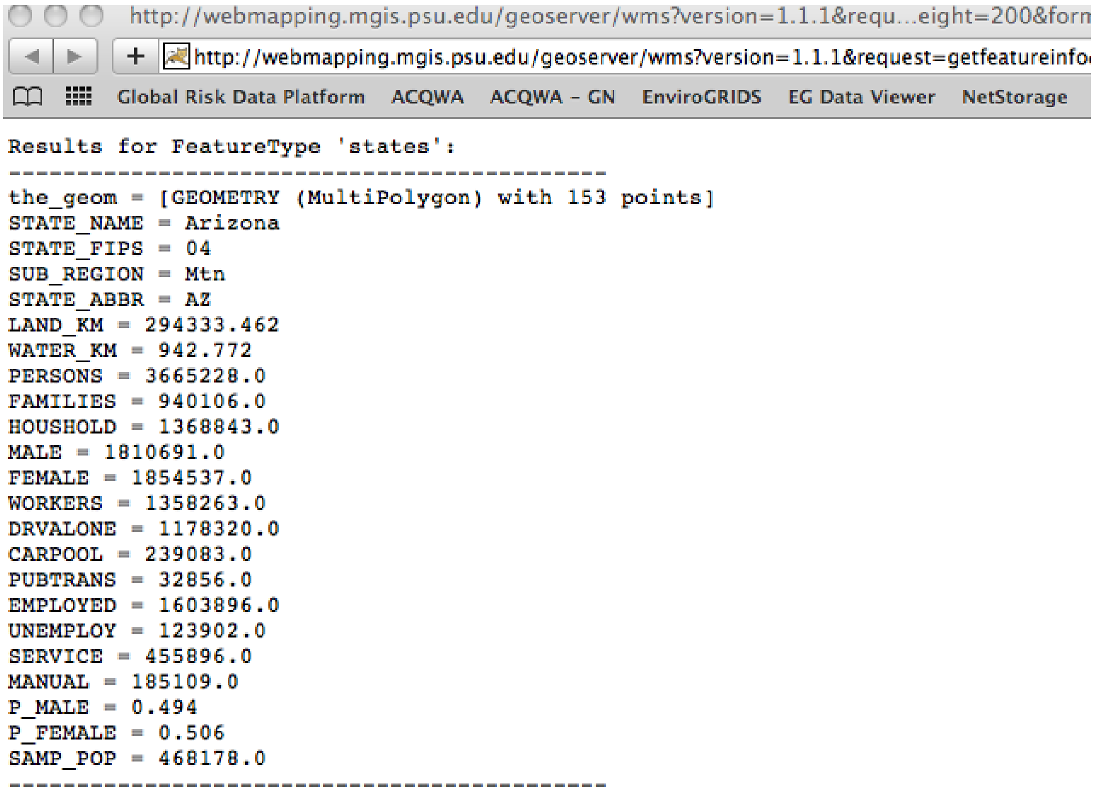
- Example of a WMS URL with a GetFeatureInfo request:

```
http://webmapping.mgis.psu.edu/geoserver/wms?version=1.1.1&request=getfeatu
reinfo&layers=topp:states&styles=population&SRS=EPSG:4326&bbox=-125,24,-
67,50&width=400&height=200&format=text/html&X=100&y=100
&query_layers=topp:states
```

The GetFeatureInfo operation is used to query the attribute table of a selected layer and get information on a specific feature. For example, a user can click on point of a map (retrieved by a GetMap request) and he obtains more information.

Mandatory parameters for the GetFeatureInfo operation:

- VERSION: the version of the specification.
- REQUEST: value "GetFeatureInfo", this the request itself to invoke the specific operation.
- LAYERS: list of selected layers separated by a comma.
- STYLES: list of style names separated by a comma. It's necessary to have an exact correspondence between the number of layers and the number of styles. If this parameter has a empty value, the default style provided by the data custodian will be applied.
- SRS: identifier of the Spatial Reference System.
- BBOX: coordinates of the bounding box following minx,miny,maxx,maxy.
- FORMAT: the format of the returned information (text/xml, text/html, text/plain)
- X,Y: coordinates of the clicked point on the map (in pixels). The origin is at the upper left corner.
- QUERY_LAYERS: list of selected layers to query separated by a comma.




```
Results for FeatureType 'states':
--------------------------------------------
the_geom = [GEOMETRY (MultiPolygon) with 153 points]
STATE_NAME = Arizona
STATE_FIPS = 04
SUB_REGION = Mtn
STATE_ABBR = AZ
LAND_KM = 294333.462
WATER_KM = 942.772
PERSONS = 3665228.0
FAMILIES = 940106.0
HOUSHOLD = 1368843.0
MALE = 1810691.0
FEMALE = 1854537.0
WORKERS = 1358263.0
DRVALONE = 1178320.0
CARPOOL = 239083.0
PUBTRANS = 32856.0
EMPLOYED = 1603896.0
UNEMPLOY = 123902.0
SERVICE = 455896.0
MANUAL = 185109.0
P_MALE = 0.494
P_FEMALE = 0.506
SAMP_POP = 468178.0
--------------------------------------------
```

Fig.13: Result of a GetFeatureInfo query.

7.3 Web Feature Service (WFS)

OGC Web Feature Service specification: http://www.opengeospatial.org/standards/wfs

The Web Feature Service defines an interface that allows a client to retrieve and update features of georeferenced data encoded in Geography Markup Language (GML). The main difference between WMS and WFS is that WFS gives direct access to the geometry and the attributes of a selected geospatial data, meaning that a user can work with a dataset provided by WFS. In brief, the WFS is the specification to access vector datasets.

Similar to the WMS, a WFS interface is invoked by a URL and can perform a certain number of operations allowing a client to manipulate the data. Following the type of operations needed to manipulate the data, we can define two classes of WFS services:

- Basic WFS: a client can retrieve and/or query features,
- Transactional WFS: a client can create, delete or update a feature.

A transaction is defined as one or more data manipulation operations that form a logical unit.

The concept of a geographic feature is described in the OGC Abstract Specification (OGC, 2009) and the retrieved or created data are encoded in the Geographic Markup Language (OGC, 2007).

- Example of a basic WFS URL:

```
http://preview.grid.unep.ch:8080/geoserver/wfs?bbox=-84.95293,19.82194,-74.13126,23.19403 &styles=&request=GetFeature&version=1.0.0 &typename=preview:cy_buffers &srs=EPSG:4326
```

Like the WMS, WFS service is supported by a set of defined operations:

- GetCapabilites: answer to a client describing its capabilities. It tells the client which kind of features are available and what operations are supported on each feature.

- DescribeFeatureType: describe the structure of a selected feature (point, line, polygon).
- GetFeature: retrieve a selected feature encoded in GML. The client can constrain the query both spatially and non-spatially and also specify the feature properties to fetch.
- Transaction: this type of request is made of operations that allow a client to modify features: create, delete and/or update operations. In addition, a client can invoke the LockFeature, in order to be sure that only one user is updating a specific feature, avoiding the risk of multi-edition at the same time.

7.4 Web Coverage Service (WCS)

OGC Web Coverage Service specification: http://www.opengeospatial.org/standards/wcs

Like the WFS allows a client to access vector datasets, the Web Coverage Service allow a client to access raster datasets. By rasters we mean data that are represented as a matrix of cells in continuous space organized in rows and columns where each cells contains a value. Thus WCS service provide access to different types of gridded data such as Digial Elevation Model (DEM), remote sensing imagery, etc... It must be noted that WCS gives only access to the raw data and does not have transactional capabilities.

Like all the OGC Web Services, a WCS interface consists of different operations:

- GetCapabilities: answer to a client describing its capabilities. It tells the client which kind of raster data (or coverage) are available.
- DescribeCoverage: describe the structure of a selected coverage.
- GetCoverage: retrieve the selected coverage.

- Example of a WCS URL with a GetCapabilites request:

```
http://preview.grid.unep.ch:8080/geoserver/ows?service=WCS&request=GetCapab
ilities
```

The GetCapabilites operation returns to a client an XML document describing the service and the data sets available from which either desktop and/or web clients may request coverages.

To invoke the operation, the user has only to define the service and the request parameters.

- Example of a WCS URL with a DescribeCoverage request:

```
http://preview.grid.unep.ch:8080/geoserver/wcs?service=WCS&request=Describe
Coverage&version=1.0.0&coverage=preview:cy_frequency
```

The DescribeCoverage operation returns to a client an XML document describing selected coverages. The information provided must be sufficient for a client to assess the fitness for use of the data. It gives different useful pieces of information such as the supported raster formats, supported SRS, supported interpolation methods, etc...

Mandatory parameters for the DescribeCoverage request:

- SERVICE: value “WCS”, this is the name of the invoked service.
- REQUEST: value “DescribeCoverage”, this is the request to invoke the specific operation.
- VERSION: the version of the specification.
- COVERAGE: list of selected coverages separated by a comma.

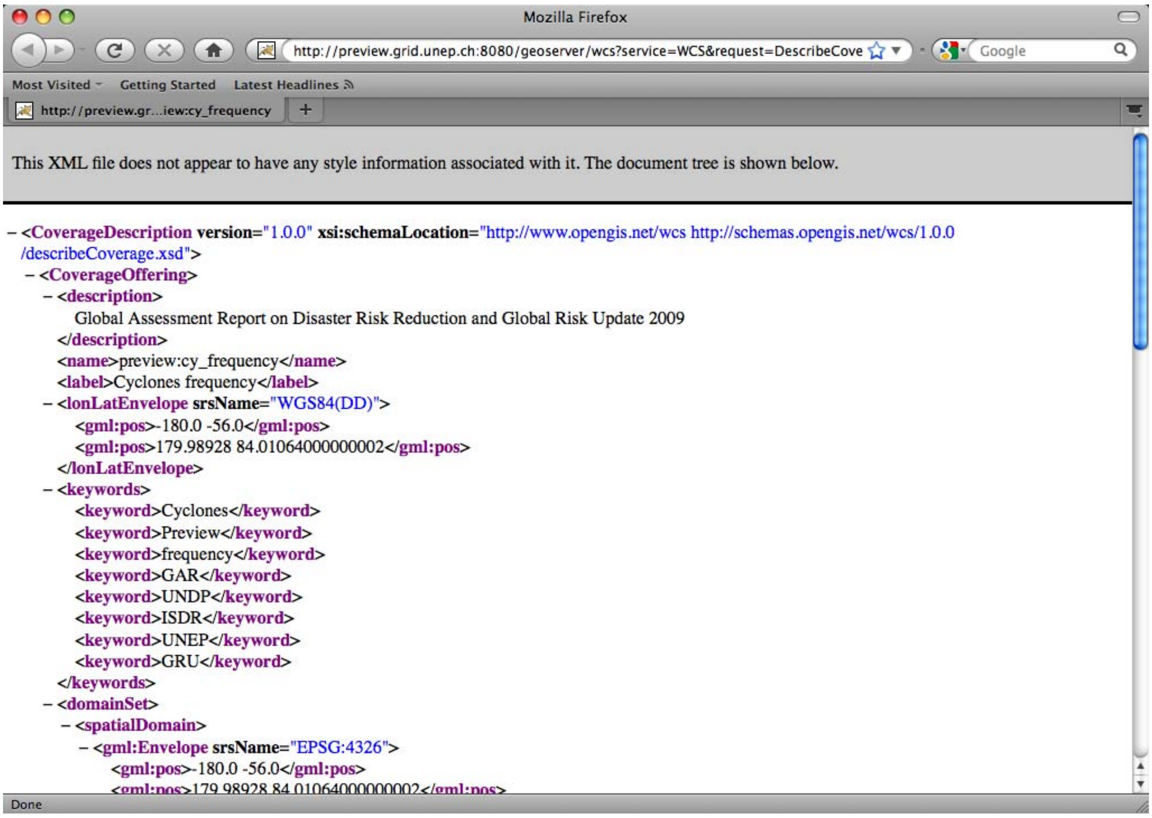


Fig.14: Result of a DescribeCoverage request.

- Example of a WCS URL with a GetCoverage request:

http://preview.grid.unep.ch:8080/geoserver/wcs?bbox=-84.95293,19.82194,-74.13126,23.19403 &service=WCS&styles=&request=GetCoverage&version=1.0.0 &coverage=preview:cy_frequency&width=640&height=309&crs=EPSG:4326 &Format=GeoTiff

The GetCoverage request returns to a client the requested raster data. The syntax and the parameters of the URL are similar to those used in a WMS GetMap resquest.

Mandatory parameters for the GetCoverage request:

- BBOX: coordinates of the bounding box following minx,miny,maxx,maxy
- SERVICE: value “WCS”, this the name of the invoked service.
- STYLES: list of style names separated by a comma. It is necessary to have an exact correspondence between the number of layers and the number of styles. If this parameter has an empty value, the default style provided by the data custodian will be applied.
- REQUEST: value “GetCoverage”, the request to invoke the specific operation.
- VERSION: the version of the specification.
- COVERAGE: name of a single selected coverage.
- WIDTH: specify the width of the returned coverage (in pixels).
- HEIGHT: specify the height of the returned coverage (in pixels).
- CRS: identifier of the Coordinate Reference System.
- FORMAT: the desired format to be used for returning the coverage (eg: GeoTiff, ARCGRID, GTOPO30,...)

If a request is validated then a coverage is extracted (using the BBOX, FORMAT and the different parameters set in the URL) from the selected coverage and sent to the client. If the client is a web browser then the user can download the coverage file. If the request is sent through a Desktop GIS client like ArcGIS then the user gets the coverage directly into it.

7.5 Web Processing Service (WPS)

OGC Web Processing Service specification: http://www.opengeospatial.org/standards/wps

The two previous discussed standards are focusing on data accessibility: WFS allows a client to access vector data while WCS allows a client to retrieve raster data.
Now we need to extend our capabilities in order to process data available using the recently introduced Web Processing Service (OGC, 2007) that provides access to processing and calculations on geospatial data. A WPS service can offer, through a network access, a vast variety of GIS functionalities ranging from a simple calculation to complex models. It acts as a sort of middleware between the client and the process that runs the calculations. It allows users to know which processes are available, to select the required input data and their formats, to create a model and run it, to manage processes (status, storage for the output, …) and to return the output once the computation is completed.

Like the others OWS, the WPS specification includes a set of operations:

- GetCapabilites: answer to a client describing its capabilities. It tells the client which kind of process are available.
- DescribeProcess: describe the parameters a selected process.
- Execute: execute a selected process.

The WPS differs a bit from the others OWS because these operations can be invoked either by SOAP or the traditional http-get and http-post.

- Example of a WPS URL with a GetCapabilites request:

```
http://localhost/wps/wps.py?service=WPS&request=GetCapabilities
```

The GetCapabilites operation returns to a client an XML document describing the service and the processes available for execution.

To invoke the operation, the user has only to define the service and the request parameters.

- Example of a WPS URL with a DescribeProcess request:

```
http://localhost/wps/wps.py?service=WPS&request=DescribeProcess&version=1.0
.0&identifier=soil_process
```

The DescribeProcess operation returns an XML document describing what are the mandatory, optional and default parameters needed for a selected process, as well as the data formats for inputs and outputs.
Mandatory parameters for this operation:

- SERVICE: value “WPS”, this is the name of the invoked service.
- REQUEST: value “DescribeProcess”, this is the request to invoke the specific operation.
- VERSION: the version of the specification.
- IDENTIFIER: the name of the selected process to run.

- Example of a WPS URL with an Execute request:

```
http://localhost/wps/wps.py?version=1.0.0&service=WPS&request=Execute&ident
ifier=soil_process&datainputs=http://localhost/wps/soil_param.xml
```

The Execute operation allows a client to run a selected process using values entered by the client for the required parameters (if needed) and reference the datasets location. Once the process is completed, the result is returned to the client as a new dataset.

7.6 Sensor Observation Service (SOS)

OGC Sensor Observation Service specification: http://www.opengeospatial.org/standards/sos

The OGC Sensor Web Service provides an interface for managing sensors and retrieving data from them. This standard is part of the suite of standards called Sensor Web Enablement (SWE) that are presented in details in the deliverable D2.3.

7.7 ISO 19115/19139

ISO 19115, Geographic information – Metadata:
http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=26020

ISO 19139, Geographic information – Metdata- XML schema implementation:

http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=32557

The ISO 19115 standard defines (with more than 400 metadata elements, 20 core mandatory elements) how to describe a georeferenced information and to provide information about the content, the identification, the quality, the spatial and temporal extent, the access and rights and the spatial reference.

This standard is applicable to:

- the cataloguing of datasets, clearinghouse activities, and the full description of datasets,
- geographic datasets, dataset series, and individual geographic features and features properties.

The main applicability of ISO19115 is for digital data but its principles can be extended and applied to other forms of geospatial data such as maps, charts and textual documents as well as non-geographic data (ISO, 2003).

The ISO 19139 standard complements ISO19115 by defining an XML encoding schema implementation specifying the metadata record format and may be used to describe, validate, and exchange geospatial metadata prepared in XML.

7.8 ISO 19119

ISO 19119, Geographic information – Services:

http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=39890

The ISO 19119 identifies and defines the architecture patterns for service interfaces used for geographic information, defines its relationship to the Open Systems Environment model, presents a geographic services taxonomy and a list of example geographic services placed in the services taxonomy. It also prescribes how to create a platform-neutral service specification, how to derive conform platform-specific service specifications, and provides guidelines for the selection and specification of geographic services from both platform-neutral and platform-specific perspectives. In other words, ISO 19119 specifies the form and content of the XML document that describes the capabilities of a geospatial web service (eg, the answer to the GetCapabilities request).

7.9 Keyhole Markup Language (KML)

OGC Keyhole Markup Language specification: http://www.opengeospatial.org/standards/kml

The KML format is an XML based language schema for describing geographical objects in web-based viewer (the so-called geobrowser). It has been developed and popularized by the Google Earth application and due to its success was turned into an OGC standard.

7.10 Geographic Markup Language (GML)

OGC Geographic Markup Language specification:
http://www.opengeospatial.org/standards/kml

The GML is a very complete XML based language used to describe all geographical features and objects and provides a standard mean for representing geographical features (properties, relationships, geometries, …). GML differs from KML as it is not only used for data visualization (which is the main focus of the KML specification) but serves also as a modeling language as well as an open and interoperable exchange format over the Internet. It is mostly used in the Web Feature Service to send geographical features between servers and clients.

The GML core schema does not contain definitions of features (because it is impossible to describe all features) but could be rather seen as a grammar providing means to define concrete features through GML Application Schemas that are created by users. Although GML is easily readable, it will be mostly used and generated by software or web services answering a specific request and then receiving the result as a GML dataset.

The real advantage of using GML is that XML technologies are available meaning that information stored in a GML file could be easily shared with other information and then specialized application domains could reuse, extend and/or refine GML components in an application schema in order to develop a specific data model.

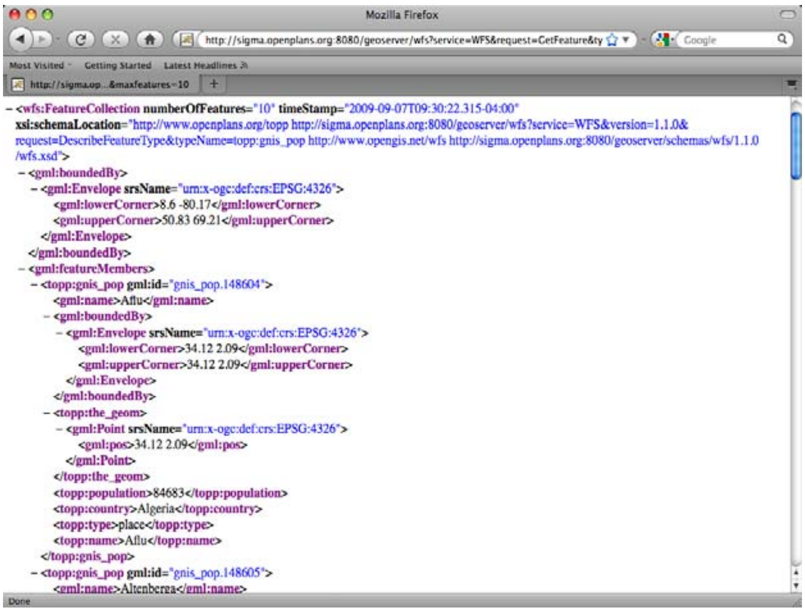


Fig.15: Example of GML returned after a WFS request.

More information available at: http://www.w3.org/Mobile/posdep/GMLIntroduction.html

8. Tools

After reviewing the set of OGC and ISO standards relevant for our purpose it is important to present a selection of tools that implement those standards, allowing the user to produce standardized and interoperable web services. Note that none of these softwares integrate all standards at once. Instead each of them is a building block of a Spatial Data Infrastructure following the OGC Reference Architecture (OGC, 2004; OGC, 2008).

8.1 OGC web services

8.1.1 Mapserver

Website: http://www.mapserver.org

Supported OS: Windows/Linux-Unix/Mac

MapServer is an open source geographic data rendering engine and development environment for building WebGIS applications and sharing data through OGC standards. It can run as a CGI program or via Mapscript which supports several programming languages.

MapServer is now a project of OSGeo and is maintained by a growing number of developers from around the world.

Mapserver main features are:

- Advanced cartographic output: scale, labels, reference map, classification,...
- Support for different scripting and development environments: PHP, Python, Perl, Java and .NET
- Cross-platform support: Linux, Windows, Mac, Solaris, ...
- OGC web services: WMS (client/server), WFS (client/server), WCS, GML, SLD, SOS, ...
- Multitude of raster and vector formats: GeoTiff, shp, PostGIS, ArcSDE, ... via GDAL/OGR
- Map projection support: up to 1000 projections through the Proj.4 librabry.

8.1.2 Geoserver

Website: http://www.geoserver.org

Supported OS: Windows/Linux-Unix/Mac

Geoserver is an open source server designed to publish data from different major data sources using OGC standards and allowing the users to share their data. Unlike Mapserver, Geoserver has no mapping capabilities and is only used for publishing data in an interoperable and standardized way.

Geoserver is a community-driven project sponsored by OSGeo.

Geoserver main features are:

- Java-based.
- Support of WMS, WFS, WCS and KML specifications.
- Various raster and vector formats: PostGIS, Oracle spatial, ArcSDE, DB2, MySQL, shp, GeoTiff, ECW, MrSID and Jpeg2000.
- Production of: KML, GML, shp, GeoRSS, PDF, GeoJSON, JPEG, GIF, SVG and PNG.
- Editing capabilities using WFS-Transactional.
- Includes an OpenLayers client for previewing data layers.

8.1.3 Deegree

Website: http://www.deegree.org/

Supported OS: Windows/Linux-Unix/Mac

Deegree is an open source framework, sponsored by OSGeo, offering the main building blocks of an SDI. Its entire architecture is developed around OGC and ISO standards.

Deegree main features are:

- Java-based.
- Support of WMS, WFS, WCS, WPS and CSW.
- Storage formats: PostGIS, Oracle, shp, GML, jpeg, gif, png, bmp, geotiff.
- Simplified installation and configuration.
- SLD support.
- Envisioned support of Sensor Observation Service (SOS) and Web Terrain Service / Web Perspective and View Service (WTS/WPVS).
- Security implementation using Web Authentication (WAS) and Web Security Service (WSS).
- Built-in web-geoportal.

8.1.4 PyWPS

Website: http://pywps.wald.intevation.org/

Supported OS: Linux-Unix

PyWPS in an implementation of the Web Processing Service specification. The great advantage of PyWPS is that it has been written with a native support of GRASS GIS software, meaning that accessing the GRASS modules via web interface should be really easy. Process can be written using either GRASS or other programs like R, GDAL or PROJ.

PyWPS main features are:

- Support of WPS specification.
- Simple configuration files.
- Method for custom process definition.
- Support for multiple WPS servers.
- Python based

8.1.5 52 north WPS

Website: http://52north.org/maven/project-sites/wps/52n-wps-site/

Supported OS: Windows/Linux-Unix/Mac

52north WPS is another implementation of the WPS specification that aims to create and design an extensible framework (with plug-in mechanism) for providing, orchestrating and executing processes as well as Grid computing on the internet.

52north WPS main feature are:

- Java-based
- Support of WPS specification.
- Pluggable framework for algorithms and XML data handling and processing frameworks
- Build up on robust libraries (JTS, geotools, xmlBeans, servlet API, derby)
- Supports full logging of service activity (exception handling, storing execution results, ..)
- Clients: basic implementation for accessing the WPS & plug-in for uDig and JUMP.
- WPS invocation: synchronous/asynchronous, http-get, SOAP, WSDL
- Supported data types: GeoTiff, ArcGrid, GML.

8.1.6 ArcGIS Server

Website: http://www.esri.com/software/arcgis/arcgisserver/index.html

Supported OS: Windows

ArcGIS Server is part of the ArcGIS family and is the component to provide web-oriented and OGC standardized spatial data services. Since version 9.2, ArcGIS Server also includes the Spatial Data Engine (ArcSDE) that geo-enables databases and allows the user to store their data into popular database system like PostgreSQL or Oracle.

ArcGIS Server main features:

- .NET and Java frameworks.
- ArcGIS Server services can be consumed by web browsers, mobile devices and desktop clients.
- Full implementation of WMS, WFS (basic and transactional), WCS, KML specifications.
- Services: mapping, geocoding, geodata management, geoprocessing, virtual globes, network analysis.
- SOAP and REST API.
- Additional SDKs to build web applications: JavaScript, Flex, Silverlight, ...

8.2 Metadata editor and catalog system

8.2.1 GeoNetwork Open Source

Website: http://geonetwork-opensource.org/

Supported OS: Windows/Linux-Unix/Mac

GeoNetwork is an open source project, sponsored by the UNSDI (UNGIWG, 2007) initiative and supported by several UN agencies (FAO, UNEP, OCHA and WFP) as well as the OSGeo. GeoNetwork implements both the Portal component and the Catalog database of a Spatial Data Infrastructure (SDI) defined in the OGC Reference Architecture (OGC, 2004) allowing a user to search, discover, evaluate, publish, manage and edit metadata on spatial data and related services through the internet.

The main goal of GeoNetwork is to improve the accessibility and thus enhance the data exchange and sharing in a standardized and consistent way between the organizations to avoid duplication, increase the cooperation and coordination of efforts in collecting data and make them available to benefit everybody, saving resources and at the same time preserving data and information ownership.

Main features of GeoNetwork are:

- Instant search on local and distributed geospatial catalogues
- Support of CSW, Z39.50 and OAI protocols.
- Uploading and downloading of data, documents, PDF's and any other content
- An interactive Web map viewer that combines Web Map Services from distributed servers around the world
- Online map layout generation and export in PDF format
- Online editing of metadata with a powerful template system
- Scheduled harvesting and synchronization of metadata between distributed catalogues
- Groups and users management
- Fine grained access control.

8.3 Data storage

8.3.1 PostgreSQL/PostGIS

Website: http://www.postgresql.org/ & http://postgis.refractions.net/

Supported OS: Windows/Linux-Unix/Mac

PostgreSQL is a popular and powerful open-source relational database management system (RDBMS) allowing the user to store data and their relations in the form of tables.

A RDBMS itself cannot store geographical information and thus to geo-enable a database system like PostgreSQL it is necessary to install a middleware to add support for geographic objects into the database. Two softwares are able to work in conjunction with PostgreSQL and add specific tables and functions that extend the capacities of a traditional RDBMS:

- PostGIS: follows the OGC Simple Features Specification for SQL (OGC, 200x).
- ESRI ArcSDE: that implements the powerful geodatabase system.

PostGIS main feature are:

- Geometrytypes for points, linestrings, polygons, multipoints, multilinestrings, multipolygons and geometrycollections.
- Spatial predicates for determining the interactions of geometries
- Spatial operators for determining geospatial measurements like area, distance, length and perimeter.
- Spatial operators for determining geospatial set operations, like union, difference, symmetric difference and buffers.
- Powerful spatial indexes for high speed spatial querying.
- Index selectivity support, to provide high performance of query plans for mixed spatial/non-spatial queries.
- No support (for the moment) of raster data.

8.3.2 PostgreSQL / ArcSDE

Website: http://www.esri.com/software/arcgis/geodatabase/storage-in-an-rdbms.html

Supported OS: Windows/Linux-Unix

ArcSDE is the second software that could “geo-enable” a RDBMS and implements the powerful concept of the geodatabse allowing to store vector and raster data in a central data repository for easy access and data management. It can be leveraged in dekstop, server and mobile applications and is the common data storage and management system of the ArGIS family of softwares products.

In the latest version 9.3 of ArcGIS Server, ArcSDE in now a component of that and offers an integrated environment. Geospatial data is managed as a database accessible by the users using a desktop client and can be easily published on the internet. It allows query, mapping, analysis and editing in a multi-user environment.

ArcSDE main features are:

- Store a rich collection of spatial data in a centralized location.
- Apply sophisticated rules and relationships to the data.
- Define advanced geospatial relational models (e.g., topologies, networks).
- Maintain integrity of spatial data with a consistent, accurate database.
- Work within a multiuser access and editing environment.
- Integrate spatial data with other IT databases.
- Easily scale storage solution.
- Support custom features and behavior.
- Support DB2, Informix, SQL Server, Oracle and PostgreSQL.

8.3.3 File system

The simplest way to store data is probably under a file system arborescence. The inconvenient is that the arborescence must be well structured and self-explainable in order to rapidly find the desired data. We do not recommend to use file system because it is a complex and not efficient system to manage and maintain geographical data. The only advantage is that a user can see a small increase in performance when accessing data but this advantage disappears as soon as he works in a environment where there is a concurrent accesses.

8.4 Web Mapping

8.4.1 Open Layers

Website: http://openlayers.org/

Supported OS: Windows/Linux-Unix/Mac

OpenLayers is and open-source JavaScript API for creating web-mapping application.

Main features are:

- load map data from many sources: WMS, WFS, GeoRSS, ...
- Support for displaying geographic features, with markers and popups
- Easy mouse/keyboard navigation.
- Layers selection.
- Easy build configuration, designed to help build OpenLayers into other applications
- Javascript API to allow full control over OpenLayers-powered map from within Javascript on a web page.

8.4.2 Mapfish

Website: http://www.mapfish.org

Supported OS: Windows/Linux-Unix/Mac

MapFish is JavaScript API and web-mapping framework using the latest web 2.0 technology and integrates different components like OpenLayers, ExtJS and GeoExt.

8.4.3 Google Maps

Website: http://code.google.com/apis/maps/
Supported OS: Windows/Linux-Unix/Mac
Google Maps and its JavaScript API are free services provided by Google and allow developers to embed Google Maps into their web pages using their own data. The API provides a number of functionalities for manipulating maps, adding content and allowing to create simple and robust maps applications.

8.4.4 Mapserver

As already discussed in the section 8.1.1 MapServer can produce OGC web services but has also cartographic capabilities using different scripting languages like PHP, Perl or Python.

9. EnviroGRIDS Challenges

9.1 Technical challenges

This is clear that due to the diversity of the partners involved and the associated heterogeneity both in term of technical capabilities and knowledge on SDI, we foresee the following technical challenges:

- Assessing the current status and environment of the geospatial sector in the region, meaning that we have to identify what is available in term of data and GIS capabilities, what is the level of implementation of OGC standards and the thematic coverage of data in regards of the

GEOSS Societal Benefits Area and INSPIRE themes. This assessment will help us to identify gaps in data, thematic coverages, GIS capabilities.
- Once we identify the relevant data, we have to facilitate the access to them either by direct agreements with the data provider (in case data would not be freely available yet) or by having one or several metadata catalogues of good quality.
- The key point for good and efficient access to data is to make data repositories interoperable, in order to allow easy access, retrieval and use.
- If the identified data are not already interoperable, we have to promote and build capacities on OGC and open standards, in order to assist data providers in sharing their data.
- Provide appropriate levels of network service access: could be an issue, if the quality of network communications avoid or at least impede good web-based communications.
- Promote data harmonization, using the INSPIRE data specifications.
- Finally, integration and connection of heterogeneous data sources.

Part of these different technical challenges will be covered by the task 2.1 on the Gap Analysis that will deliver its assessment on data access and GIS capabilities at the end of the first year of the project. The aim is to have the best possible overview of the current status and environment of the geospatial sector in the Black Sea catchment area, allowing us to develop the best strategy to promote the use of open standards through GEOSS and INSPIRE frameworks.

The greatest foreseen challenges are: to identify the relevant data sources and to define clearly the core of geospatial data needed, and to encourage data providers to go "open" and share their data in an interoperable and OGC-compliant way. At the present, technology is not a problem in itself because different solutions based on different software could be proposed and/or developed depending on the requirements and on the technical capabilities available. The most difficult task will be probably to create an environment allowing a wide agreement on data sharing principles.

We think that these challenges could be overcome principally with a participative approach, proving the benefits of data sharing and in consequence stimulating data providers and partners to come on board.

9.2 Collaborative challenges

In order to achieve the goal of a "sharing spirit", it is of high importance to promote collaboration and cooperation among partners. Good communication and good organization are fundamental for sharing not only data but also information, knowledge and capacities. In consequence, high priority must be given to the creation of a well understood and accepted governance structure in order to develop a clear strategy, milestones to achieve, etc... allowing the partners to share the same vision and to feel a common sense of ownership of the future Black Sea Catchment Observation System. This will certainly create a spirit of commitment around the project and thus would greatly facilitate the endorsement on Spatial Data Infrastructure and related concepts.

As mentioned earlier in the document, sharing fosters the notion of re-use of data but also re-use of technical capabilities and skills developed, highlighting the importance of capacity-building, the necessity to learn from others and to share also knowledge among the different partners of the project.

If collaboration is achieved with the help of good communication and organization, the agreement on the use of new standards, the development of guidelines, good practices and policies will greatly enhance the "open and shared spirit".

9.3 Challenges related to the grid computing integration

One of the key challenges of the EnviroGRIDS project is to bridge the technological gap between the world of Spatial Data Infrastructure and Grid infrastructure, meaning that we wish to create a Grid-enabled SDI.

With the ever increasing spatial and temporal resolution of geospatial data causing a tremendous increase in term of size of these data, we are progressively reaching the limits of the processing capacities of the traditional GIS/SDI technology.
With the uptake of Grid technology and the progressive deployment of large infrastructure project like the EGEE many more scientific domains and technology-related activities have access to sizable computing resources.
Recent studies (Muresan O., 2009; Liping D. *et al.*, 2003) had a successful approach to extend Grid technology to the remote sensing community and to make OGC web services Grid-enabled. They both consider that the Grid has a great potential for the geospatial discipline.
Padeberg and Greve (2009) have identified several differences between OGC-compliant SDIs and Grid infrastructures:

1. Service description: Services in a Grid infrastructure are described using WSDL while the OGC web services provide their own description according to a specific XML schema (response to a GetCapabilites request).
2. Service interface: OGC web services traditionally use XML-RPC (http-get and http-post) to send request and receive response while Grid services are invoked using SOAP. Thus we need to find a solution to use OGC interfaces on the Grid.
3. Statefulness: Grid services store state information for later use (such as intermediate results) while OGC service don't handle state information (with the exception of the WPS).
4. Security: On a Grid, it is essential to be identified and to have mechanisms of encryption, authentication and authorization. OGC does not have specifications to address these issues.

Moreover Liping *et al* .(2003) have identified that the current Grid metadata catalog system is not good enough to answer the needs of the geospatial community, especially the requirements of the ISO 19115 standard.
Data storage issues will be discussed in deliverable D2.2 and the whole development and implementation of the gridded SDI will be presented in deliverable D2.10.

10. Conclusions and Recommendations

Based on the foreseen challenges presented in the previous section, we propose a set of recommendations to successfully implement interoperability among the partners of the project and beyond. Participants will act as a key drivers and pioneers, to promote sharing of data supported by SDI initiatives and serving and receiving data to and from systems like INSPIRE and GEOSS.

General recommendations:

- **GR1: Assessing the status and environment of geospatial sector in the region.**
- **GR2: Define a core-geodatasets needed for the project.**
- **GR3: Create appropriate governance structure.**
- **GR4: Develop a clear strategy and vision.**
- **GR5: Facilitate access to data.**
- **GR6: Making data interoperable.**
- **GR7: Promote the use of open standards.**
- **GR8: Endorse the GEOSS and INSPIRE principles.**
- **GR9: Building capacities.**
- **GR10: Obtain commitment**
- **GR11: Have good communication and organization**
- **GR12: Take care of the differences between SDIs and Grid Infrastructure.**
- **GR13: Propose an envisioned architecture for a Grid-enabled SDI.**

One of the major aims of the EnviroGRIDS project is to build capacities around different themes and especially regarding data sharing, interoperability and open standards. Previously, we highlighted the importance of the concept of re-use, not only of data but also of capacities and knowledge. Following the GEO Capacity Building Strategy (2006), we can distinguish three elements of high importance:

- Human capacity building whose main concern is about education and training of people to be aware of, access, use and develop geospatial data and information.
- Institutional capacity building is focused on developing and fostering an environment for the use of Earth observations to enhance decision making. This includes building policies, programs and organizational structures in governments and organizations aimed at enhancing the understanding of the value of geospatial data and information.
- Infrastructure capacity building refers to the technological components (hardware, software,...) required to access, use and develop geospatial data and information for decision making.

In consequence, we propose different specific recommendations for the project aiming to help building capacities around the Black Sea catchment.

Capacity building recommendations:

- **CB1: Build on existing efforts and best practices,**
- **CB2: Focus on user needs,**
- **CB3: Foster collaboration and partnerships,**
- **CB4: Foster information exchange and knowledge development,**
- **CB5: Facilitate the exchange of ideas and best practices,**
- **CB6: Facilitate the sharing of human and technical resources,**
- **CB7: Facilitate education and training**
- **CB8: Promote the sharing of data, information, reports, articles and guidelines.**
- **CB9: Identify experts in each specific areas,**
- **CB10: Identify, coordinate and build synergies between existing and future capacity building efforts and SDIs initiatives.**
- **CB11: Enhance the sustainability of existing and future SDIs initiatives and capacity building efforts by building awareness amongst the different stakeholders.**

To achieve and meet these recommendations, it is of high importance to clearly define:

- **who are and who could be the end users of the grid-enabled SDI?**

At this stage, the two main end-users will be the Black Sea Commission and the International Commission for the Protection of the Danube River. The infrastructure must help them to increase their capacities and improve their every day work, especially regarding data and information management, state of the environment reporting, assessments, water quality studies, early-warning systems and modeling.

It would be good also to identify or bring other potential users of the infrastructure (other than the partners of the project that are obvious users) within the region in order to promote the use of SDI and Grid infrastructure.

- **what is the design of the architecture?**

In the deliverables 2.1 and 2.2, we present general architectures for the Spatial Data Infrastructure (fig.16) as well as the Grid Infrastructure. The next step is to define and agree on a clear and complete architecture connecting the two types of infrastructures allowing to discover and retrieve data through the SDI and analyze them on the Grid.

For that purpose, we recommend to organize a specific workshop at the beginning of year 2010 (but before the General Assembly), with the relevant partners of the project. The output of the project will be a document presenting the EnviroGRIDS grid-enabled SDI specifications, a list of tasks and a roadmap to develop the overall infrastructure, a list of proposed tools to be used, and finally an efficient and practical implementation strategy allowing our infrastructure to be scalable and to adapt it to the coming and future developed solutions (like G-OWS, 52north, ...).

- **what is the long-term plan for the infrastructure?**

As we are trying to connect two types of infrastructures (SDI and Grid) and due to the fact that participating to the Grid involves some financial investments, it is important to already think and clarify what will be the sustainability plan for our infrastructure, especially after the end of the project. It is essential to guarantee a long-term access to the Grid otherwise the grid-enabled SDI will loose its analytical and processing component and it will be highly damageable for the end-users of the infrastructure, as well as the EnviroGRIDS consortium.

This issue will be covered by deliverable D2.5.

- **who's go what?**

To identify the different data sources it is necessary that metadata be immediately available. If partners and/or data providers within the Black Sea catchment have already metadata catalogue we have to know them so that we can query their catalogue directly from the EnviroGRIDS URM (http://www.envirogrids.cz). If metadata is unavailable because either lacking of capacities to publish them using CSW or because data are not documented, the EnviroGRIDS project should help them to learn at least to write metadata into the catalogue system of our URM. However, it would be better to build capacities and help them to learn to install and use tools like GeoNetwork.

- **Data policy?**

In order to give a clear vision of the data exchange to the data providers and to allow data sharing at a large extent, it is of primary importance to have a good and precise data policy. It must be clear enough so that data providers would be interested to share their data and cover items of primary interest to them, like: data sharing principles, intellectual property rights, access rights, etc...

In the annexes of this document a draft proposition is suggested but it must be clearly improved. We recommend to circulate this data policy among partners and especially the BSC and ICPDR to have their inputs. The aim is to finalize it for the next General Assembly in spring 2010 where all partners should endorse it and publish it on the EnviroGRIDS website.

- **Data quality?**

Once we have access to the vast amount of data that is available in the region, we need to assess also the quality of the data because it is clear that not all data will be useful for our purpose as well as for the end-users of the infrastructure.

Therefore, we recommend to define a methodology to assess data quality.

- **build on best practices**

The last two points highlight the need to build our knowledge on best practices that are already available on the GEO/GEOSS and INSIPIRE websites.

- **build capacities**

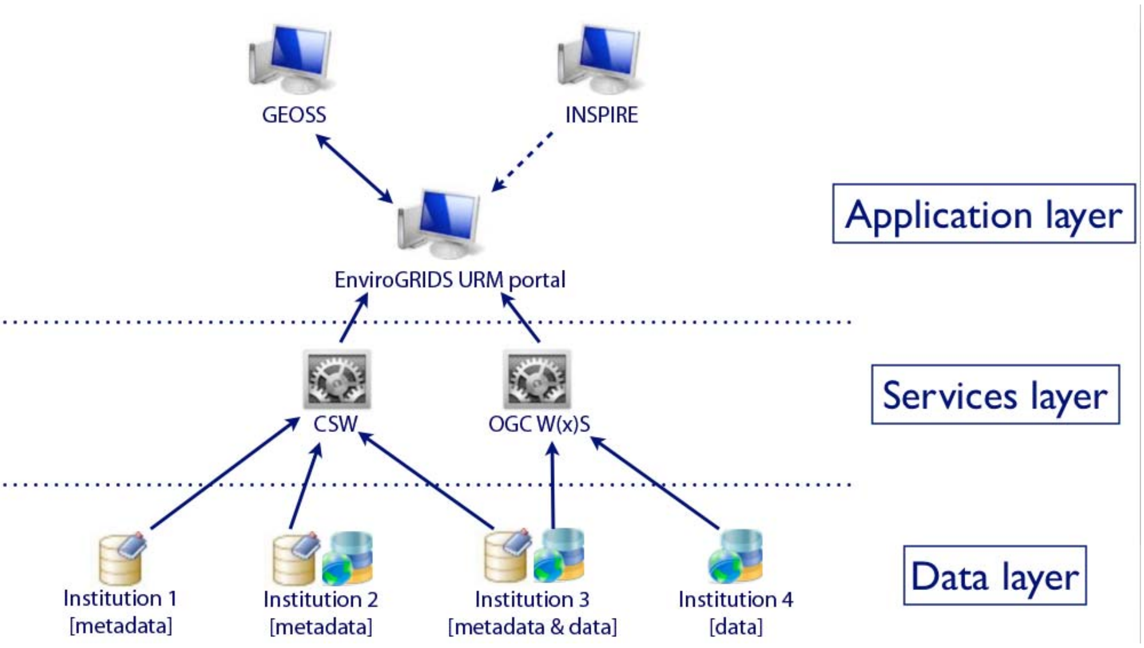
As we already mention previously, the main goal of the project is to build capacities in the Black Sea region especially regarding data sharing and related technologies (web services, metadata catalogues, etc...) allowing data providers to register their data into GEOSS. The success of the project will be measured on how many new contributors of the region will register their services into the GEOSS Common Infrastructure.

In order to help them to get the necessary knowledge and use the GEOSS system, we will organize:

- High level workshops (for the different stakeholders of the region) aiming to build awareness on the benefits of sharing data, to present GEO/GEOSS and its added values.
- Technical trainings (focused on GIS/IT people) teaching them how to install and use softwares like GeoServer, GeoNetwork and/or ESRI products such as ArcGIS Server.

These two types of workshops could be replicated at least in the different countries around the Black Sea and we think that it will be welcome to translate the documentation that will be distributed in the different regional languages.

Our hope is to plant and grow some seeds of a new vision based on the benefits of data sharing and offering standardized access to a vast amount of data and information. The challenge will then become to use this data as effectively as possible in order to allow decision-makers within the Black



Sea catchment to take better informed decisions.

Fig.16: General architecture of the EnvrioGRIDS SDI

The general architecture will be based on a traditional three tier model:

- *Data layer:* will be the level representing the data providers. We envisioned four different cases ranging from sharing only metadata (Institution 1) to sharing only data (Institution 4). The case "Institution 3" will be the best case because data providers want to share metadata and as well as data. The case "Institution 2" represents the fact that probably there will be institutions that have both metadata and data but want to share only metadata. In all cases, depending on the level of knowledge regarding OGC and ISO standards, we aim to build capacities and to help data providers to share their data and/or metadata in an interoperable and standardized way.
- *Services layer:* represents the level where OGC web services will be implemented using softwares like GeNetwork, GeoServer or ArcGIS Server. These services will be implemented and made available in the institutions themselves.
- *Application layer:* is the level of the EnviroGRIDS URM portal that will act as the hub for all metadata and data shared through the services layer. The portal implements functionalities to query the different metadata catalogue as well as showing the different geospatial data. The EnviroGRIDS URM portal will be registered into the GEOSS Common Infrastructure and in consequence all the metadata and data available in the URM portal will also be accessible through the GEOSS geoportal. Finally, the EnviroGRIDS portal will have the possibility to query and access data of GEOSS and INSPIRE geoportals.

EnviroGRIDS developments

EnviroGRIDS partners are already working and developing different tools to facilitate data access and management based on the OGC/ISO standards. In particular, the Uniform Resource Management GeoPortal (from CCSS partner) is an integrated portal using several tools previously presented, allowing effective exchange of data, knowledge and information.

D.1 EnviroGRIDS GeoPortal

The EnviroGRIDS GeoPortal (under development by ArxIT) will help partners to store, describe and manage geospatial data following INSPIRE and GEOSS implementing rules and specifications. This tool aims also to be the entry point for the SWAT users allowing them to extract and prepare data to be processed into the Soil and Water Assessment Tool.

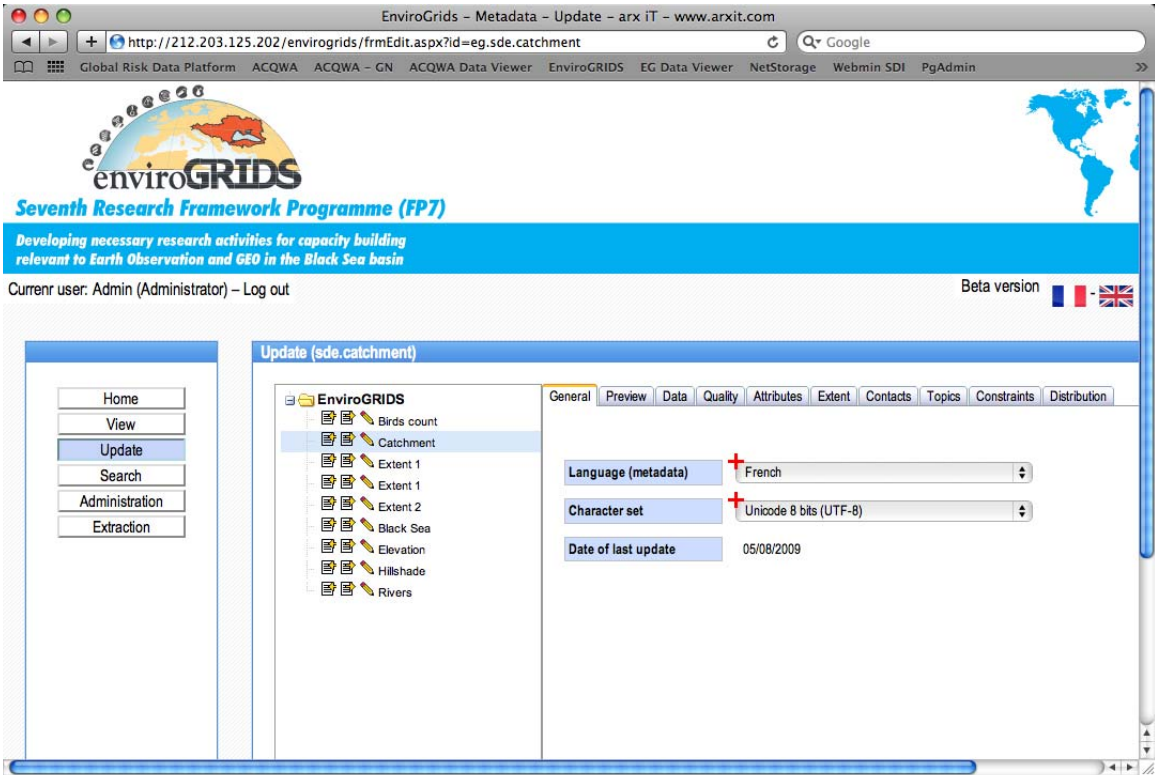


Fig.D1.1: EnviroGRIDS Geoportal

D.2 EnviroGRIDS Clearinghouse

The EnviroGRIDS clearinghouse (under development by UniGE/GRID) will present data and web services used within the project following the nine GEOSS Societal Benefit Areas (SBA). This will act as an easy and simple gateway allowing users to visualize relevant data and information under each SBA.

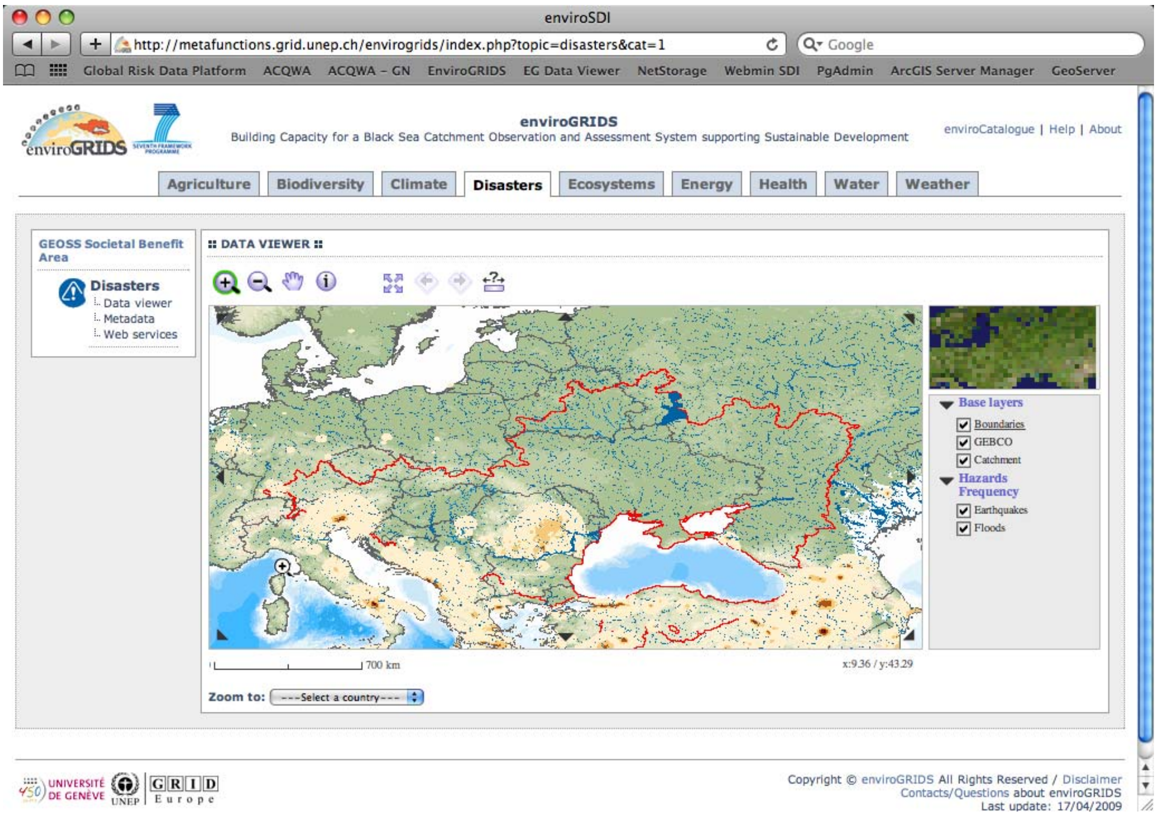


Fig.D2.1: EnviroGRIDS clearinghouse presenting data on disasters shared by the Global Risk Data Platform using WMS.

D.3 Uniform Resource Management (URM) GeoPortal

The URM GeoPortal is a new, integrated solution being designed as combination of previous technologies, Geohosting and new technological development of a visualization client based on Open Layers. The URM GeoPortal is not an integrated solution, but a set of modules and services which are able to communicate trough interoperable services (OGC, W3C). The solution is modular and could be easily modified for different purposes. URM Geoportal is based on Open Source technologies, but it could be integrated with different technologies like MS SQL or ArcSDE. Uniform Resource Management supports validation, discovery and access to heterogeneous information and knowledge. It is based on the utilization of metadata schemes. The URM model currently also integrates different tools, which support sharing of knowledge. The GeoPortal system is developed on the OpenSource platforms (MapServer, GeoServer) and contains common visualization, data sharing, metadata and catalogue functionalities. Additional parts of solution are also tools for management of sensor observations and spatial data transformation and processing.

URM Geoportal is not the product of one company, but it is designed and developed by members of the Czech Living Labs (CLL), which represents a research and development environment in which several research institutions, business companies and regional authorities work together in order to develop projects in the area of new technological concepts. The individual CLL members can be either in a user or researcher position. CLL members assign experts into individual teams for each project. The involvement of the CLL members depends on their interests and possible benefits that they can acquire from each individual project. Users out of Living Lab are contacted in order to get the widest amount of relevant suggestions for solution to any existing task.

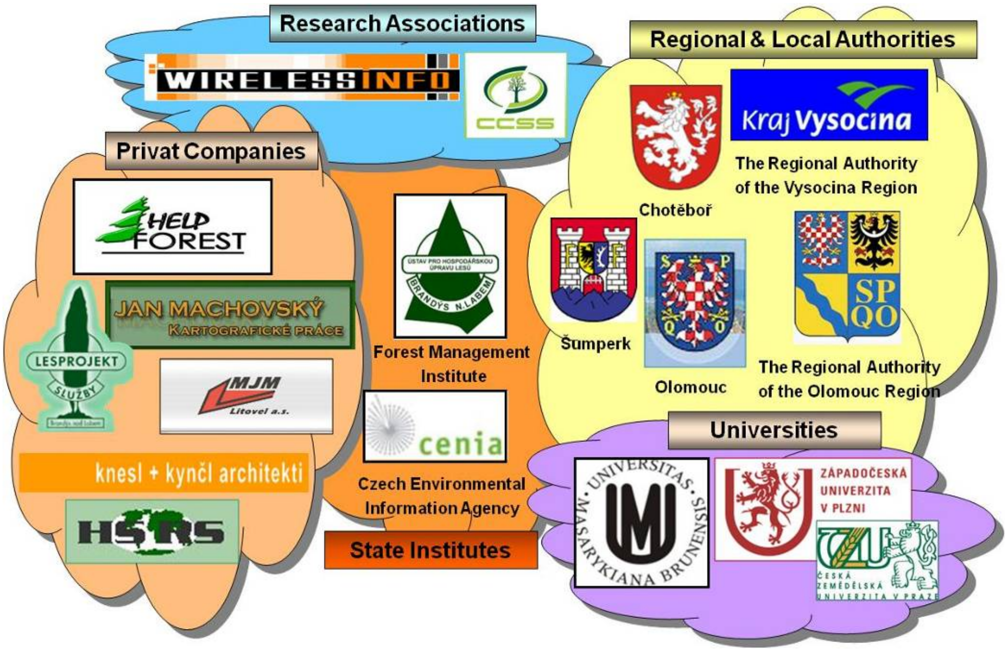


Fig.D3.1: Czech Living Lab.

The CLL is focused mainly on research and development of new technologies around geo-data providing and exchange. The main idea of the CLL is the establishment of new services to change, where possible the physical source data transfer (for example on a CD) into web services that provide on-demand connection to the source data server. The source data remains on the owner's server and can be easily updated. The users are able to connect updated data at any time and use them for their work.

Users inside the CLL define their requirements of innovation and these requirements create the basis for a definition of the target applications. The target applications are formed from collaborative tools – independent components, which have some of the following characteristics - Open Web Services, Collaborative Environment, User Involvement for technical solution, Technological platform independence, Re-using of existing tools, methods and technologies, De-centralized data sources, Open interfaces.

The design and implementation of URM GeoPortal is a long time process and its development is financed from commercial activities of CLL members, but also from different national and European research projects (NaturNet Redime, SpravaDat, MobilDat, Navlog, Earthlook, Impuls, c@r, Winsoc, Humboldt, Plan4all).

Parts of the integrated solution are currently available as commercial tools, while others are Open Source published under GPL license (or dual license). IN 2009, additional tools will be available as Open Source.

Component scheme of URM GeoPortal

The basic components of URM GeoPortal, which are based on INSPIRE principles, are presented in Figure D2.3.

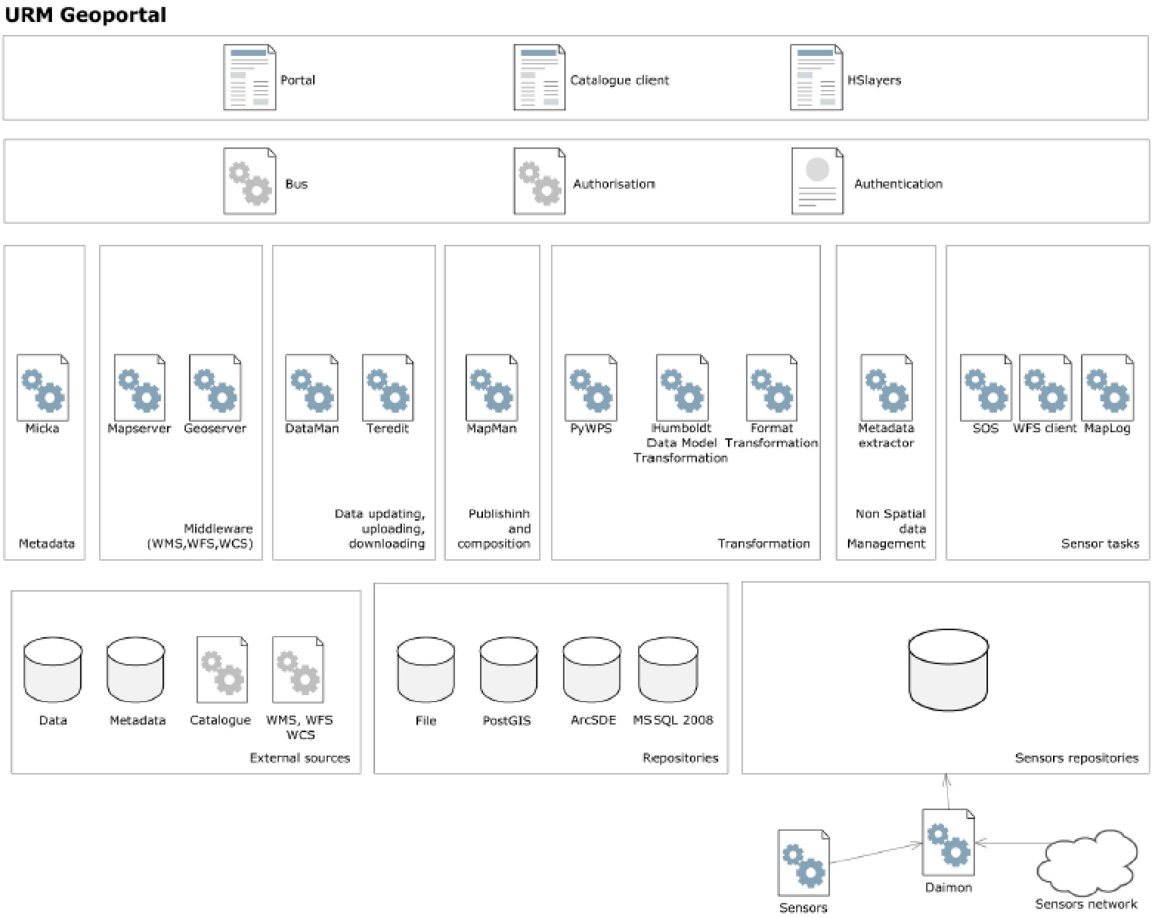


Fig.D3.2: URM GeoPortal components.

External information resources

URM GeoPotal supports communication with external information resources through OGC services. For data and services discovery URM GeoPotal supports communication trough CSW2.0.2 catalogue services. The interconnectivity was tested with Conterra Catalogue and GeoNetwork.

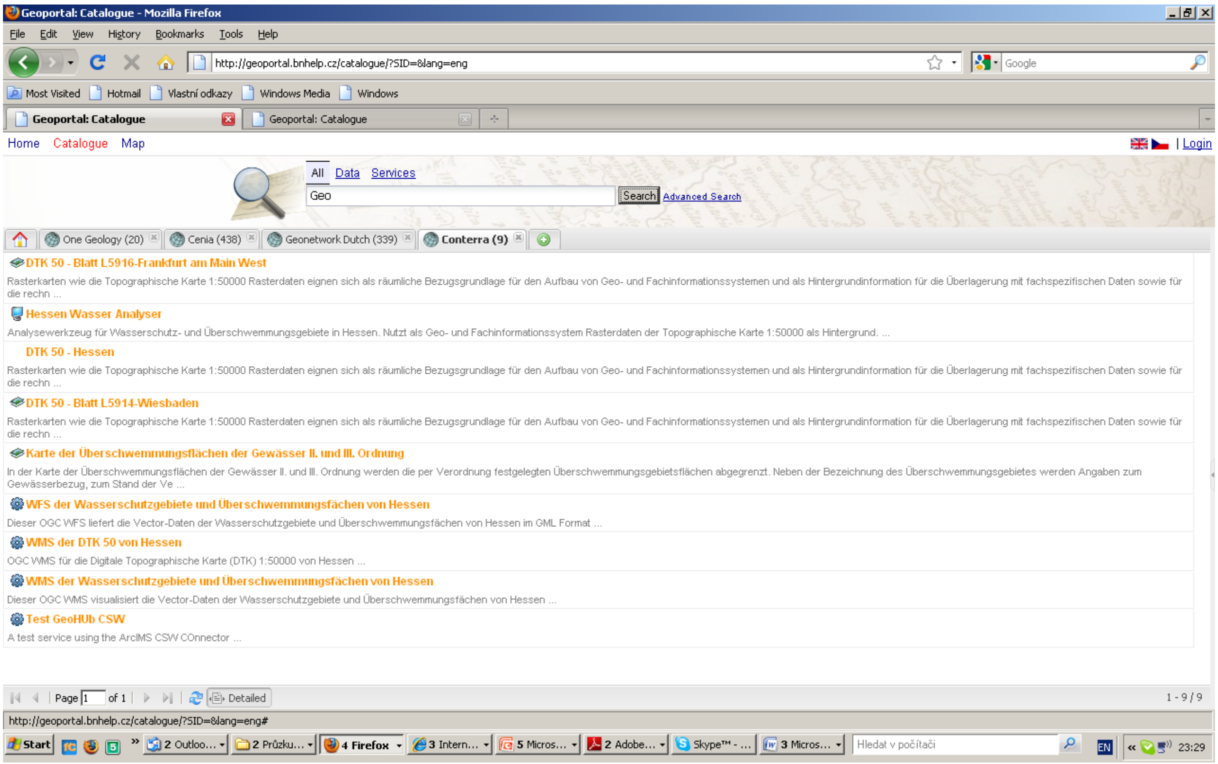


Fig.D3.3: Connectivity with GeoNetwork and Conterra.

For portrayal services is implemented Web Map Service (WMS). For data services are used Web Feature Services (WFS) and Web Coverage Service (WCS) are used.

Internal repository

As its own internal repository, URM GeoPortal uses:

- File system (for shape file, dgn file and raster file like geotiff)
- PostGIS as main data repository

System also offer connectivity to:

- MS SQL 2008 spatial extension
- MS SQL GeoMedia extension
- ArcSDE

Storing of metadata could be done in:

- PostgreSQL
- MySQL
- MS SQL
- Oracle

For better visualization of large data sets GeoPortal use TileCache3

Sensors

Sensors and sensor repositories are discussed in a separate deliverables (D2.3) So this part is not describe here.

Metadata

The URM GeoPortal uses a product of CLL partner Help Service Remote Sensing Micka, MICKA version 3.0.

MIcKa is a complex solution for metadata management and for Spatial Data Infrastructure (SDI) and geoportal building. It contains tools for editing and management of metadata for spatial information, web services and other sources (documents, web sites, etc.). It includes their search on the Internet, portrayal in map or download to local computer.

MIcKA is compatible with the European standards of the INSPIRE directive. Therefore it is ready to be connected with other nodes of prepared network of metadata catalogues (its compatibility with pilot European geoportal is continuously tested).

Functions:

- Functionality with metadata for spatial data (ISO 19115)
- Functionality with metadata for services (ISO 19119)
- Functionality with Dublin Core metadata (ISO 15836)
- Functionality with Feature catalogue (ISO 19110)
- Support of OGC CSW 2.0.2 (catalogue service)
- Functionality with metadata user profiles
- Support of INSPIRE metadata profile
- Web interface for metadata editing
- Multilingual (both user interface and metadata records). The following languages are currently supported: Czech, English, German, French, Latvian, Polish. It is possible to dynamically extend the system for other languages.
- Context help (multilingual)
- Import of the following metadata formats are supported:
  - ESRI ArcCatalog,
  - ISO 19139,
  - MIDAS
  - OGC services (WMS, WFS, WCS, CSW)
  - Feature catalogue XML
- Export – ISO 19139, GeoRSS
- Support of thesauruses and gazetteers.
- Display of changes by using GeoRSS
- User templates for appearance and functionality management.
- Possibility of map client connection for display of on-line map services.
- System requirements:
  - Relational database (ORACLE >= 9, PostgreSQL >= 8.0, MS-SQL >= 2005, or other SQL databases)
  - PHP >= 5.2, support of XSLT
- Independent on Operating system

Method for metadata editing

Metadata are stored in relational database and edited by dynamically generated forms. It is therefore possible to amend other standards or profiles and to switch between profiles while editing. Individual profiles can be distributed into sections. With help of control elements it is possible to double individual items, choose from code list or connect supporting applications. Control of mandatory filled items is enabled while editing.

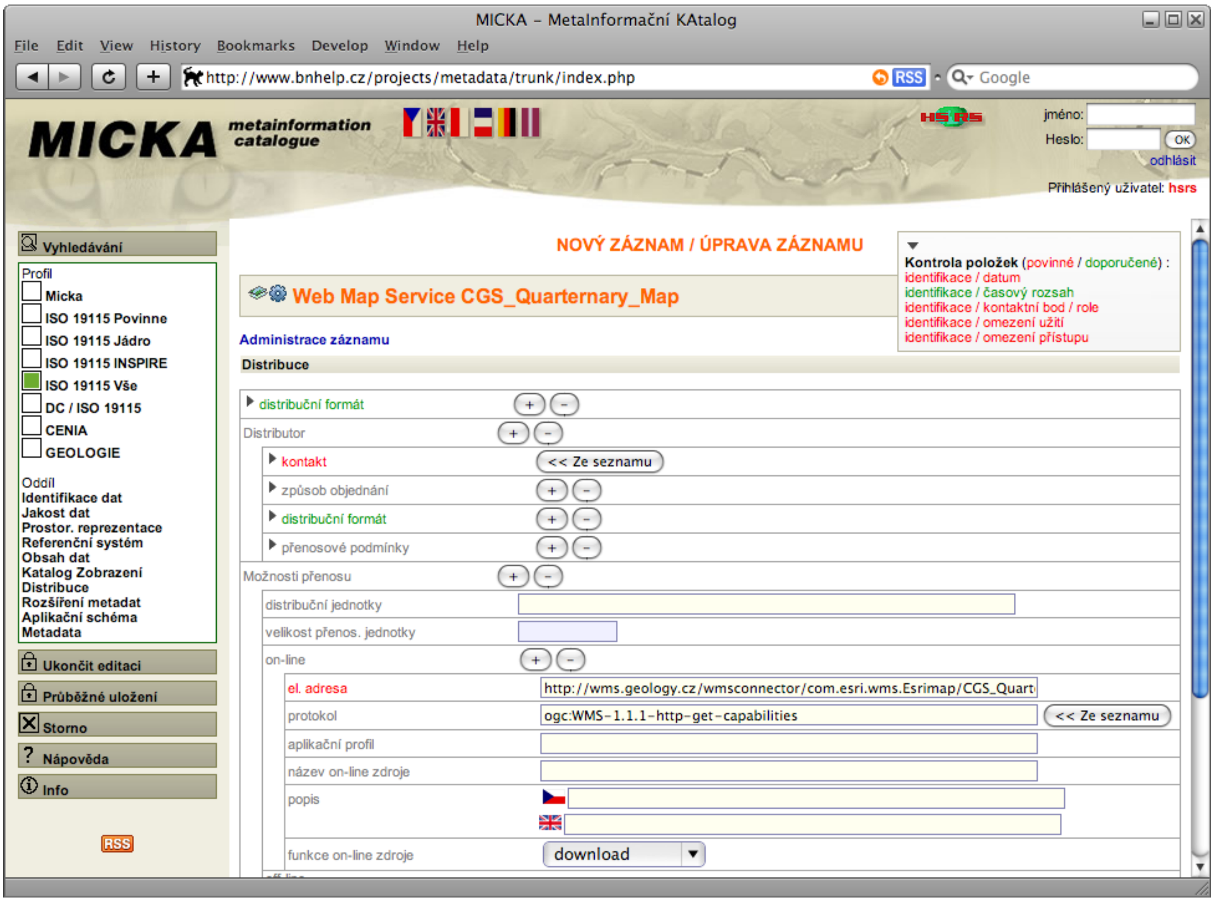


Fig.D3.4: Editing application interface.

Concept of work with spatial data:

MIcKA enables to enter spatial extent of metadata:

- By choosing bounding box directly in map. All maps (for coordinates input or metadata extent display) are realised through connected WMS. Any WMS server can be connected.
- By search in gazetteer. Access to gazetteers is realised through connection to WFS server. By default WFS from HSRS server is connected. It serves administrative division of the Czech Republic and Europe (up to NUTS3 level). The software can connect to any WFS server.

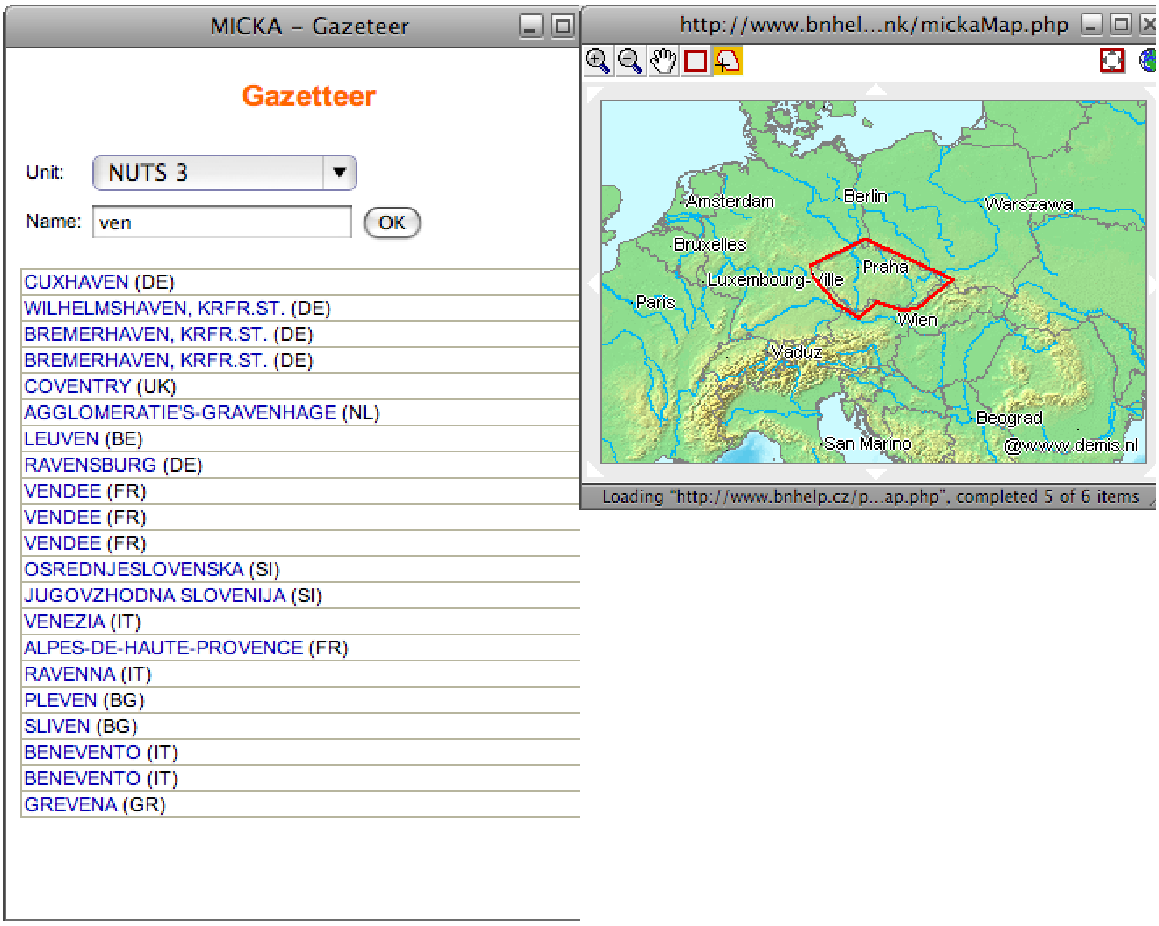


Fig.D3.5: Entering of spatial extent using gazetteer and bounding box.

Concept of work with key names:

Following standards for the classification of records, the system handles:

- Thematic ISO categories (mandatory for datasets)
- Key words:
  - entered by user (arbitrary)
  - chosen from thesaurus

In compliance with INSPIRE requirements, parts of the system are:

- an open source GEMET thesaurus service client - data classification service code list

- an ability to insert key words in order to enable multilingual search.

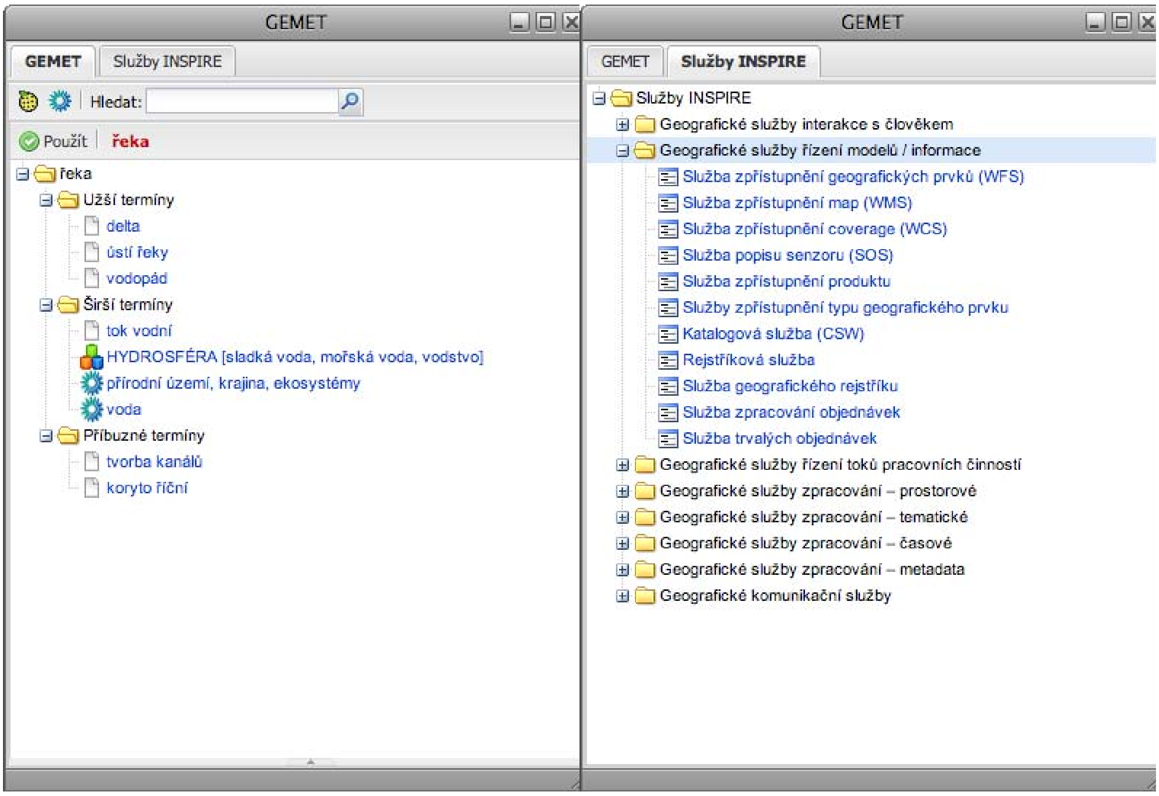


Fig.D3.6: Support of GEMET thesaurus and service classification according to INSPIRE.

Support of the INSPIRE directive:

- INSPIRE metadata profile is part of the system
- Choice from key words from GEMET thesaurus
- Choice from key words from code list of INSPIRE services
- Continuous control of metadata completeness according to the INSPIRE profile
- Batch control of completeness of INSPIRE profile
- Implementation of catalogue service according to OGC CSW 2.0.2 standard

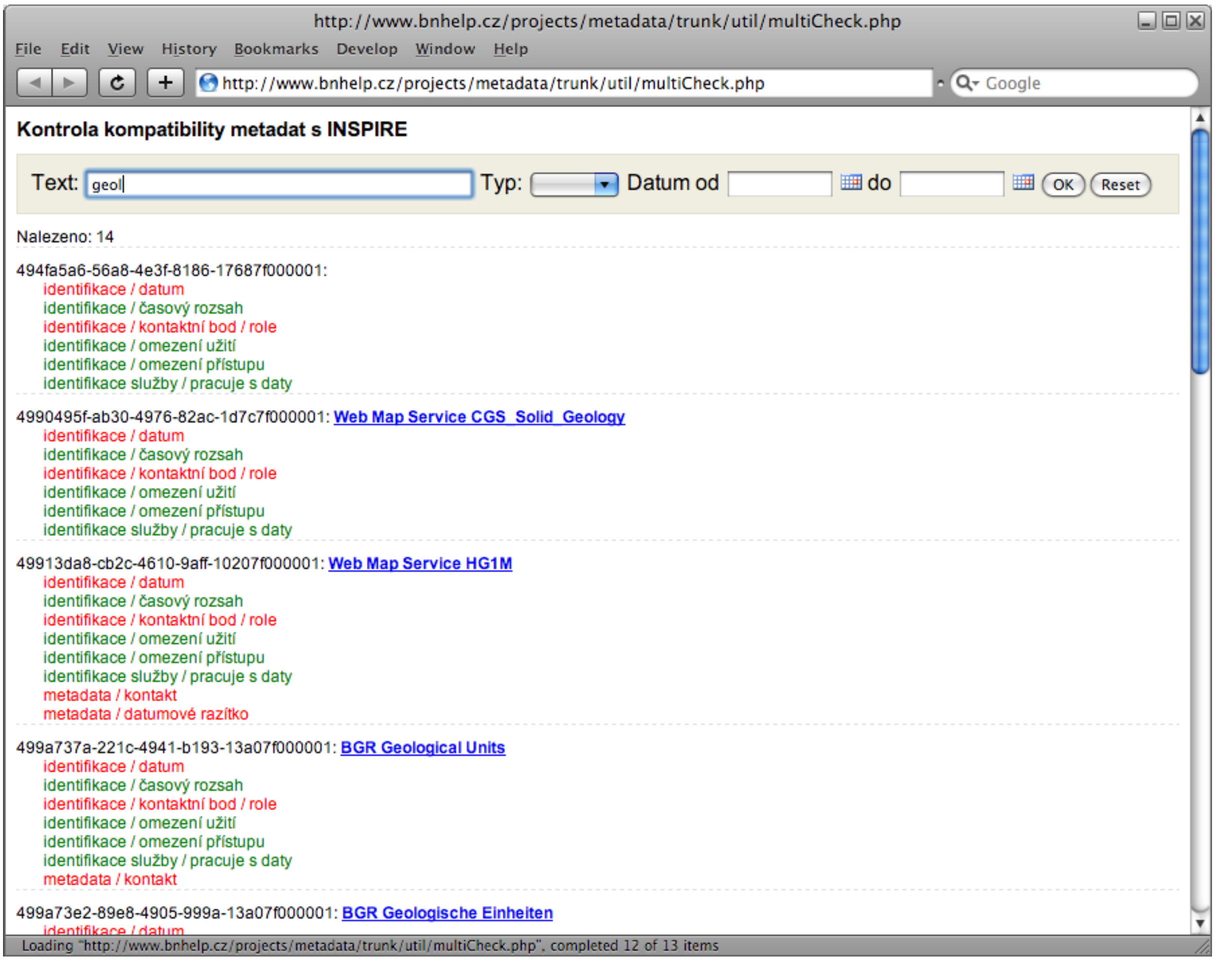


Fig.D3.7: Module for batch control of INSPIRE profile.

Catalogue service:

Catalogue service is part of the system.

It is based on OpenGIS Catalogue Services Specification – profile Catalogue Service for Web (CSW) and OpenGIS Catalogue Services Specification 2.0.2 - ISO Metadata Application Profile standards.

Supported operations:

- Basic: GetCapabilities, DescribeRecord, GetRecords, GetRecordById
- Editing: (CSW-T): Transaction, Harvest
- Inquiry items: according to standards
- Extensions: export to GeoRSS and KML

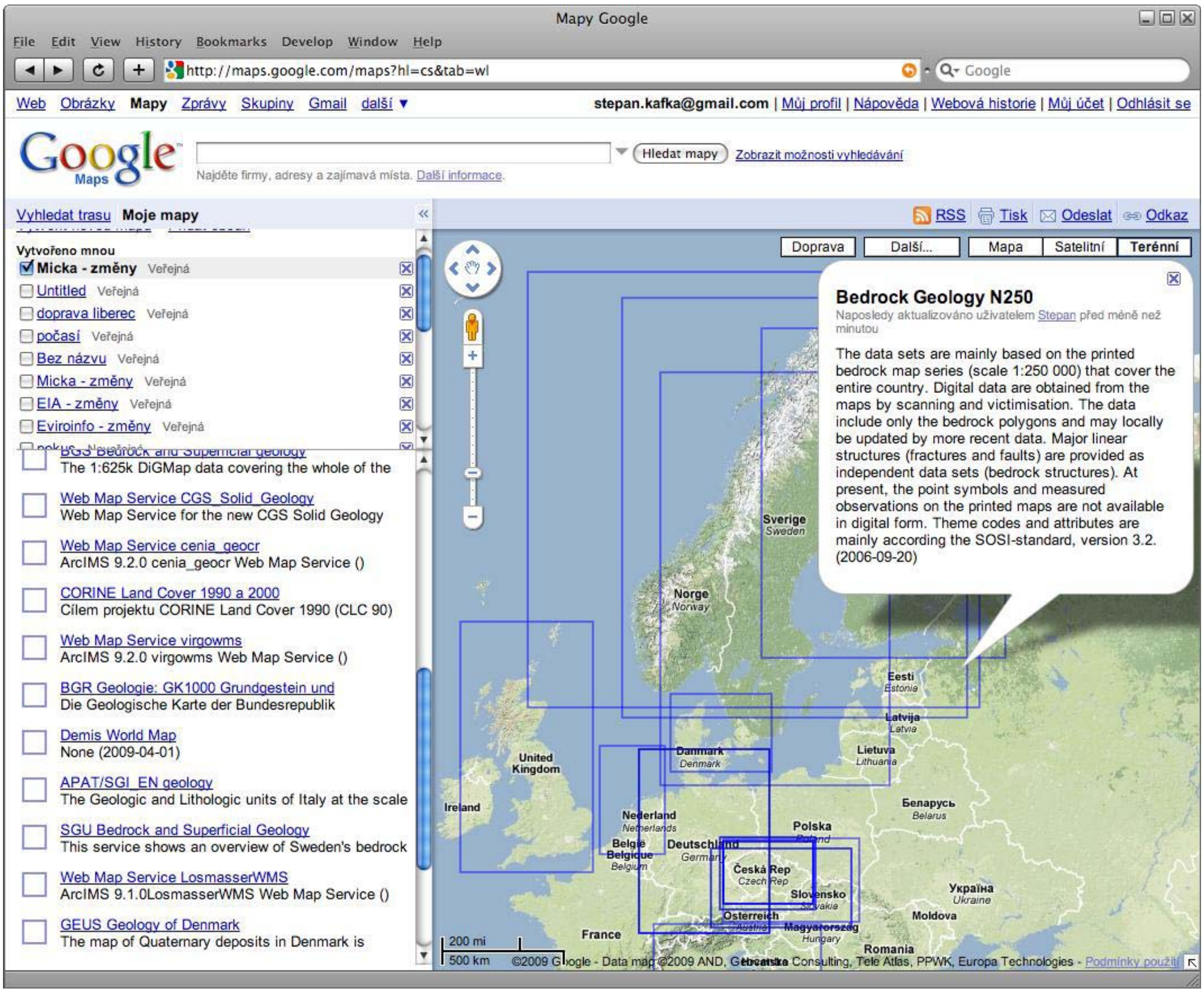


Fig.D3.8: Extent of metadata records as an output of GeoRSS in Google Maps.

Extensions:

The following modules can be connected to the basic MIcKA application:

- WMS Viewer – display of map compositions from available services.
- Micka - Lite – is an application with simple control that contains configuration-enabled form for metadata acquisition in INSPIRE profile and ISO 19139 format. Metadata is possible to store on local disk or send through CSW into MIcKA catalogue.
- Metadata Extractor – enables automatic retrieval of metadata from various sources (texts, images, voice files, web pages, etc.) and their insert into metadata storage using CSW-T.
- Download Manager – batch download of data from data storage for defined territory on the basis of information contained in metadata.

The system MIcKA can be further integrated with other applications. The system is suitable not only for metadata for spatial data management but basically as a tool for central management and evidence of various types of information (documents, data, applications, services, etc.)

Middleware for WMS, WFS and WCS

For implementing of WMS, WFS and WCS, we use two Open Source platforms UMN Mapserver and GeoServer

Data uploading, downloading and updating

For data management, which includes uploading, downloading and updating (include mobile) CLL implement two tools DataMan and Terredit. Both tools will be available as Open Source during late 2009.

DataMan

DataMan is an application for management of spatial data. It supports management of data in databases or files. It supports export and import data, and publishing and updating of related metadata. In database, it is possible to store both vector and raster data, including their attributes. Also for file oriented storage, it supports both vector and raster data. For raster formats, it currently supports TIFF/GeoTIFF, JPEG,GIF, PNG, BMP, ECW. For vector formats ESRI Shapefile, DGN, DWG, GML.

The basic functionality of Dataman are:

- Transfer the spatial file into file repository
- Describe file by metadata (ISO19115)
- Definition of structure in database
- Transfer file file from file system into database (currently only for shapefile)
- Describe record by metadata (19115)

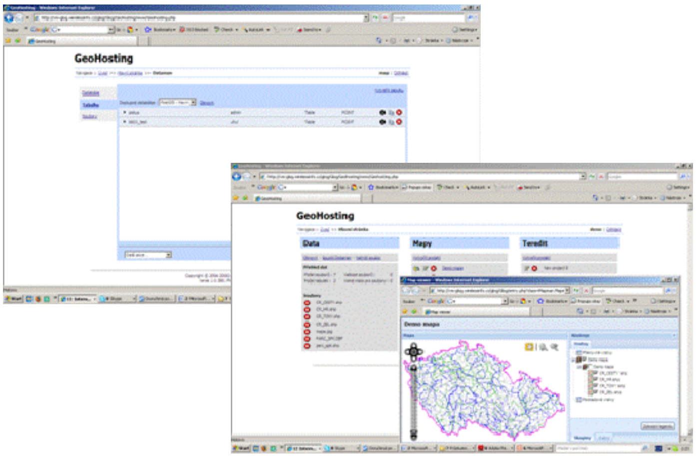


Fig.D3.9: DataMan.

Terredit

Teredit was initially designed as a special tool for data collection in the field, for data transfer from a mobile device (PDA, notebook) to a server and for generating a project for data collection. Currently it is modified to become a general interface between server and any external devices including desktop GISs. Data can be collected and saved onto a server in on-line or off-line mode. Teredit is a global component system where the components are interconnected via a multilevel architecture. Primary data collection provides a "Mobile GIS Editor" – it can be any available data collected software (e.g. ArcPad for PDA) which is up to standard of functionality and communication. The second side is the internal server of the company where the data will be stored. Teredit constitutes a bridge between these two sides.

Teredit sl use web services for data transfer. Web Map Services (WMS) are used to transfer referential layers, Web Features Services (WFS) provide transfer of vector data for editing. The system is proposed with transaction functionality.

"Broker" is a component on the server level of the system and it provides system management and administration. An important task is also to support communication with data, metadata and catalogue servers, checking of batches from PDA or desktop, project management and other.

The next level is consisted of desktop components – project editor, desktop processor and desktop GIS editor. This part is important for field workers who need to check crowd of data collected in the field. The level could be skipped in some project and data can be transferred in this case from a mobile device directly onto the server.

The third level, mobile devices, is necessary for data collection in the field. Teredit on this level contains three main components – Mobile Project Processor, Mobile GIS Editor and External Device Manager. The GIS editor is a SW tool for object editing (e.g. ArcPad). The editor can edit graphic and descriptive (thematic and geographic) attributes. Mobile processor manages data transfer from PDA to a desktop or to a server.

External Device Manager is a component which provides connection with fourth level – with gps, sensors and measurements.

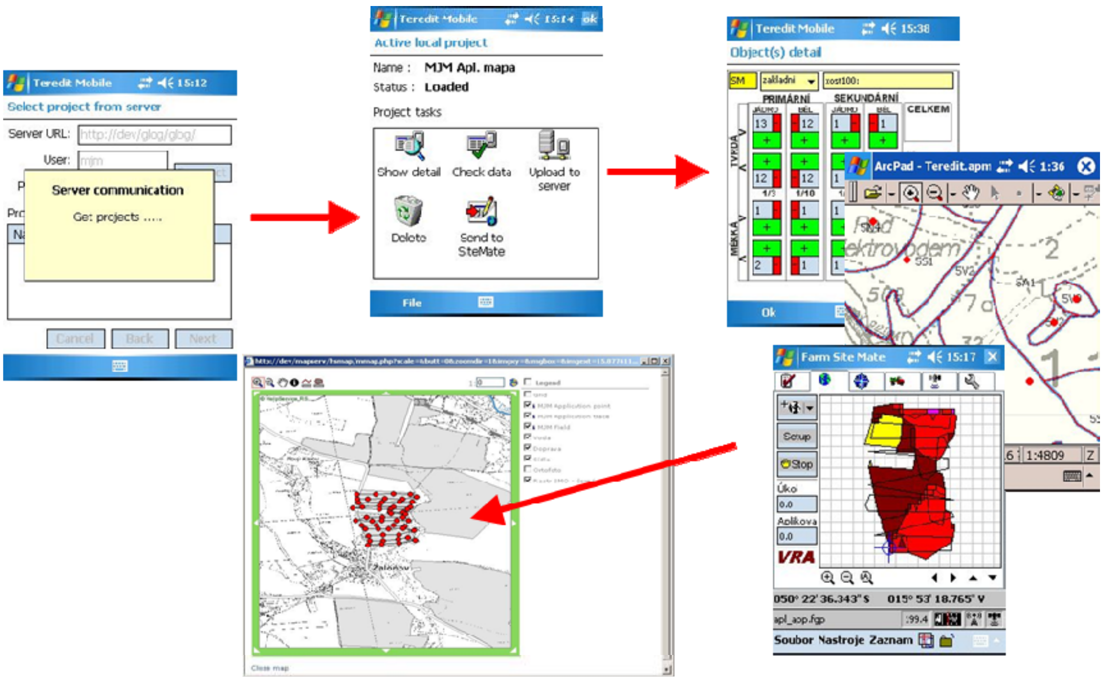


Fig.D3.10: Principle of Teredit data flow server-PDA-server.

The important part of Teredit system is Mobile GIS Editor, which can be presented by common SW utility for object editing. The editor can edit graphic and descriptive attributes.

Data publishing and composition preparing

Another part of the URM GeoPortal solution is the tool that assists publishing local data as WMS and WFS services, and also support preparing and storing visualisation composition from local data and external Web Services. This solution is called MapMan and will be published as Open Source in late 2009.

MapMan

The Map Project Manager (MapMan) is a software tool for users who want to publish local data or create new map projects and compositions from local data and external services. It supports publication of spatial composition from locally stored data (fields or database-stored in DataMan), with external WMS and WFS data services. It supports visualization in a web browser using clients like HSLayers, GoogleMaps, DHTML client, Desktop viewer GoogleEarth, GIS Janitor or publish data as WMS and WFS. All published data are also connected with metadata stored in Micka.

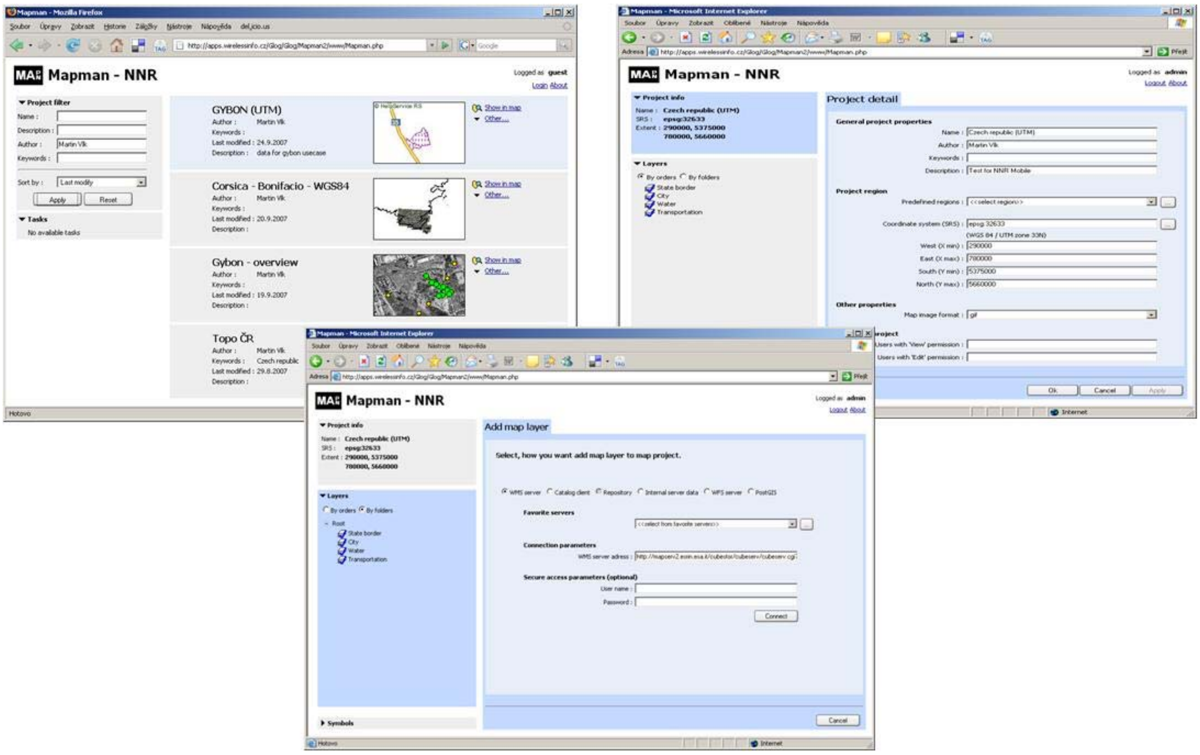


Fig.D3.11: MapMan.

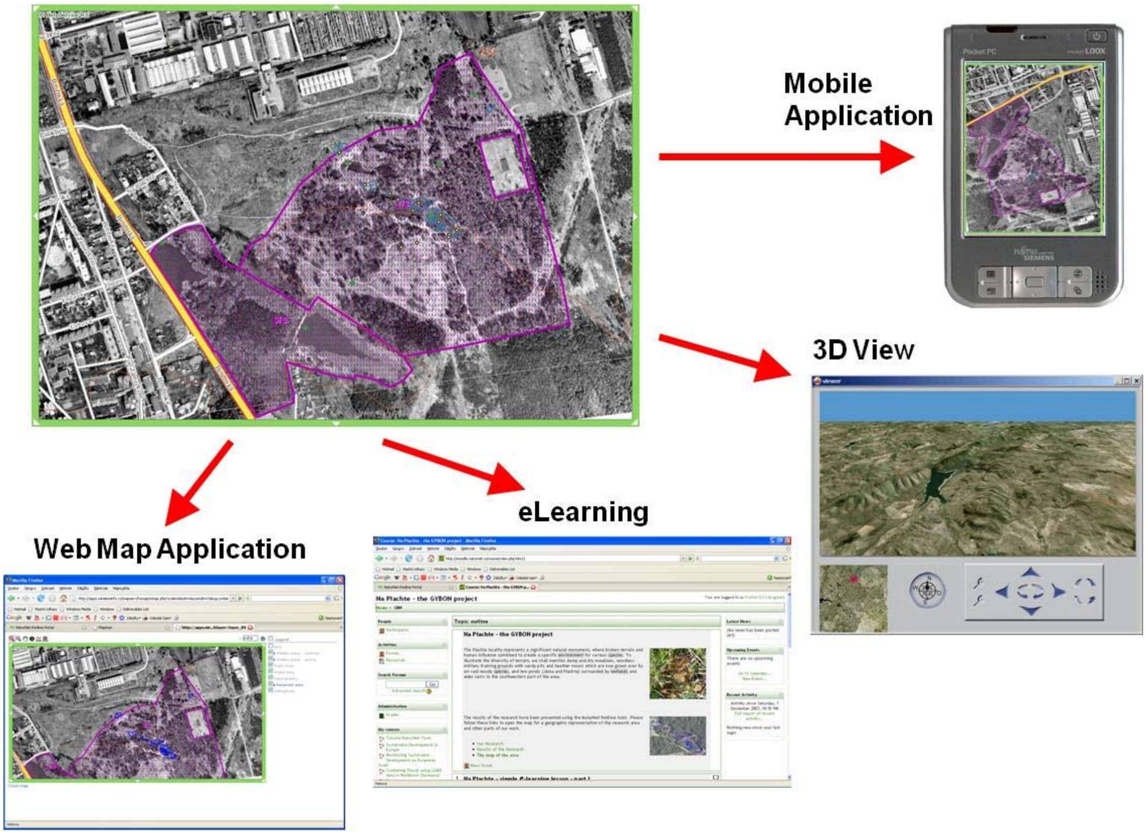


Fig.D3.12: Usage of MapMan composition.

The composition published and used by different applications.

Transformation and processing data

The transformation and processing of data is an important part of any future portal. There is currently one operational tools and two others are under development. The existing tool is PyWPS, an implementation of WPS services. Under development are now Humboldt project transformation services and WPS interface for commercial software system Topol supporting transformation GIS and CAD data.

PyWPS

PyWPS is a project developed since 2006, and which tries to implement OGC WPS standard in its 0.4.0 version. It is written in Python programming language. The main goal of PyWPS is the direct support for GRASS GIS. So, PyWPS can be understood as a kind of translation library, which translates requests complain to WPS standard, overhands them to GRASS GIS or other command line tool (such as GDAL/OGR, PROJ.4 or R statistical package), monitors the calculation progress by informing the user, and returns back the result upon completion.

PyWPS is released under a GNU/GPL license. Currently, version 2.0.0 is available. It is actively maintained by Help Service - Remote Sensing company as one of our projects.

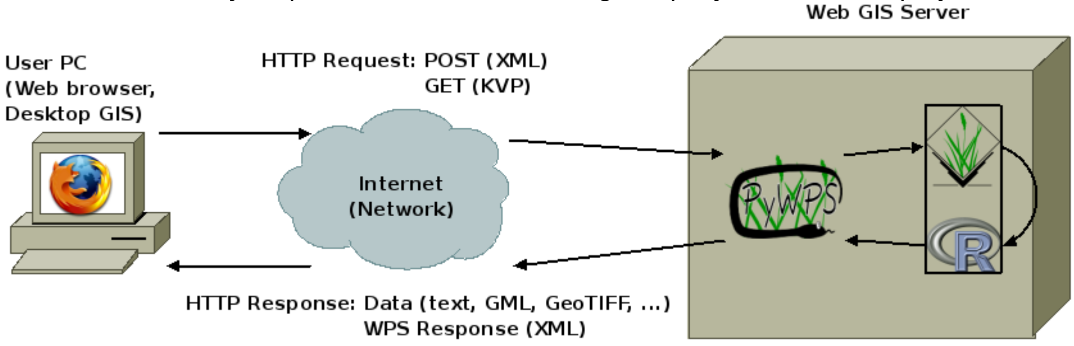


Fig.D3.13: PyWPS implementation.

Humboldt transformation services

The HUMBOLDT framework is concerned with the most important part of the project – to enable organizations to document and harmonize their spatial information, supporting the implementation of a European Spatial Data Infrastructure (ESDI) that integrates the diversity of spatial data available for a multitude of organizations.

The HUMBOLDT framework has three tiers: The HUMBOLDT Harmonisation Toolkit contains applications that can be used by geodata experts to specify conceptual schemas and to define the transformation between them. The HUMBOLDT Service Integration Framework defines both cross-cutting service components that allow the collaborative usage across organizations, countries and domains. The final tier of the HUMBOLDT Framework is formed by individual transformation services, each of which can address a specific harmonisation problem.

One of the major goals of the framework is to be minimally invasive, i.e. not to replace existing systems but rather support and amend them with specific capabilities needed in the data harmonisation process. Therefore, the functionalities of the HUMBOLDT framework are well-isolated from the interfaces by which they are accessed, resulting in components that can be adopted easily for different deployments and process synchronization styles.

First components of the HUMBOLDT framework have been published under the GNU Lesser General Public License version 3 (LGPL v3). They are available for free download at the HUMBOLDT project Web site. This release includes the following software components:

- The HUMOLDT Model Editor, a UML editor that is specifically geared towards the creation of UML application schemas;
- The HUMBOLDT Alignment Editor, a tool that allows to define conceptual schema transformations;
- The Mediator Service, a proxy service that executes transformation chains to provide harmonised geodata;
- The Workflow Service, a service that analyses data sets and decides which processing is required to match a target product description;
- The Context Service, an easy to use-service that can be used to define transformation products;
- Several transformation services exposed as OGC Web Processing Services, such as Coordinate Transformation Service and an Edge Matching Service.

Detailed documentation for all these services and applications, such as full specifications, will be made available under the same license as the software itself. Furthermore, a forum, wiki and bugtracker are prepared for user information and feedback.

Publishing of Non spatial data

Current IT systems require dealing not only with spatial data, but also with non spatial data. Combination of spatial and non spatial data is one of the main objectives of the URM, and has been implemented in Metadata Extractor.

Metadata Extractor

Metadata extractor is a tool to extract available metadata directly from different files (documents, presentation, etc.), edit this metadata and publish metadata and files on URM portal. Other possibility is to extract metadata directly from existing URL addresses and store metadata on URM portal. Access to information is then trough direct URL addresses.

Currently, metadata extractor supports

- publishing documents on the portal – one can select any type of file, extract and edit metadata and published this file on the portal (see Figure D3.14).

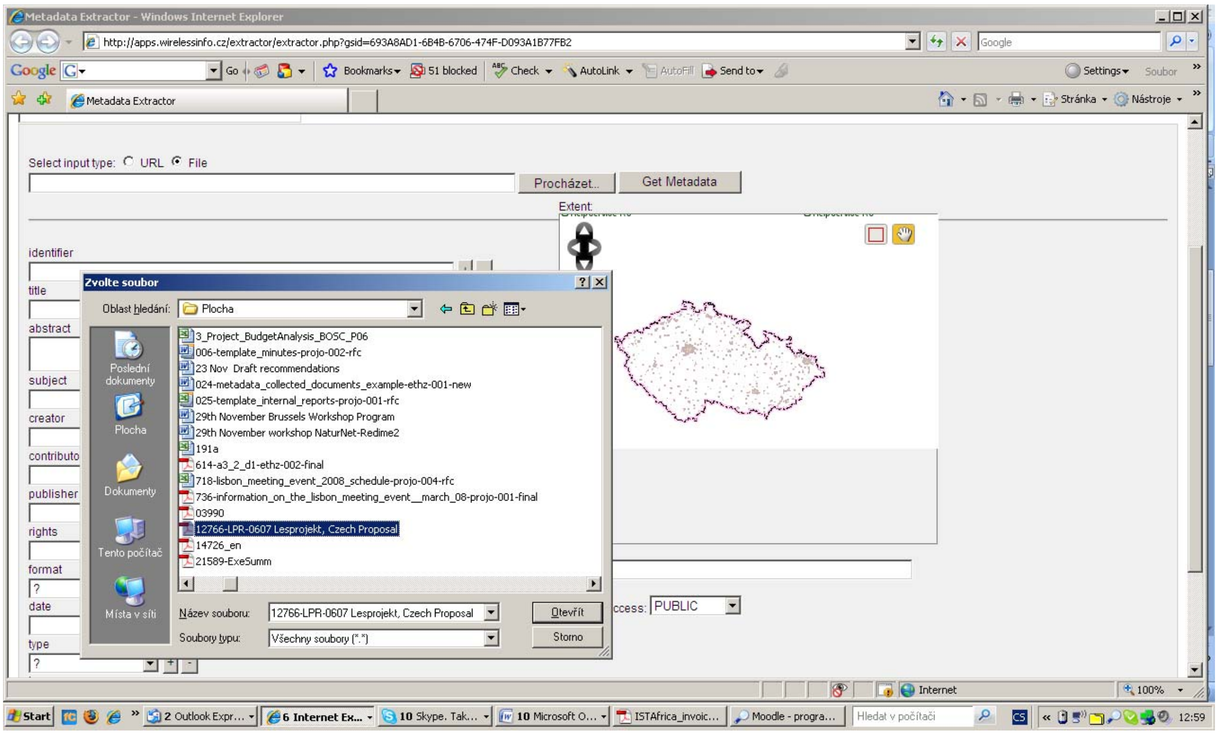


Fig.D3.14: File publishing.

- publishing of links to existing Web pages by putting URL of Web pages to extractor:

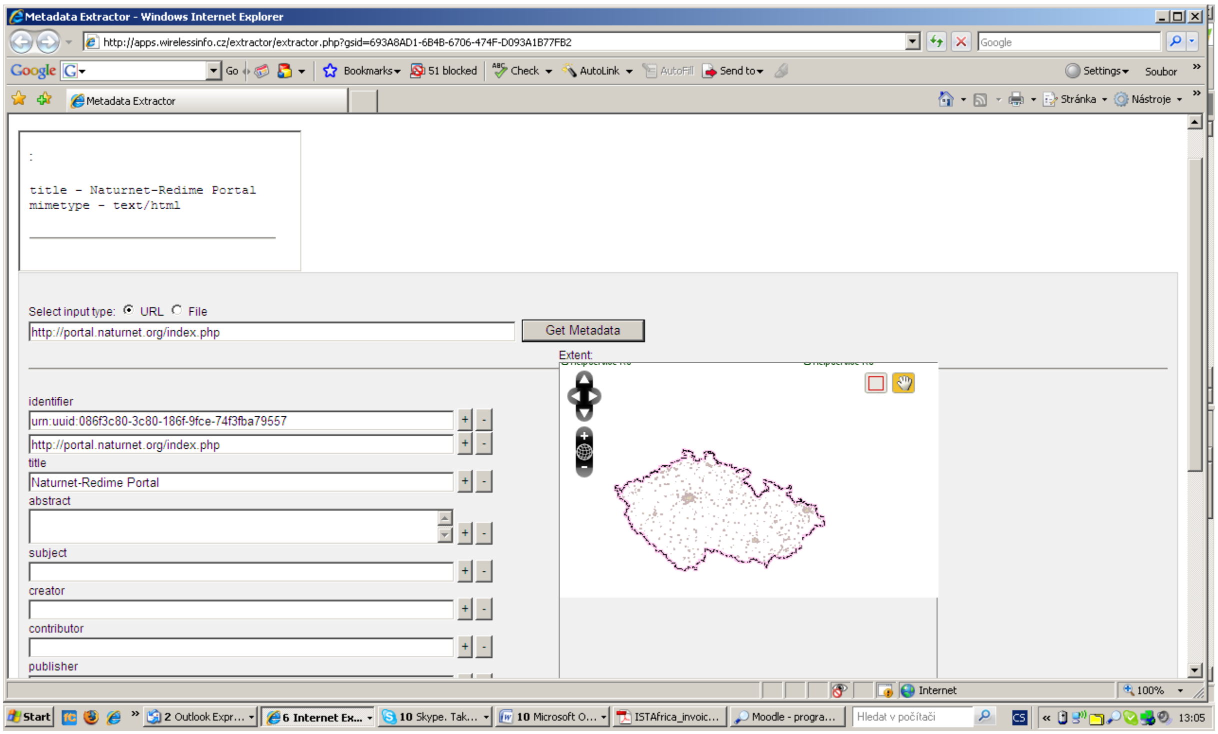


Fig.D3.15:. Publishing of URL.

- or published directly new Web pages stored in Zip file. These Web pages are directly accessible trough URM portal.

BUS

Bus guarantees interconnectivity among all components. Communication is trough SOAP protocol. Security – information confidentiality and security must be assured during the emergency situation. Therefore authorization and authentication services are strongly required and rule based user's attitude for data handling is recommended. Own tools supporting authorization and authentication were implemented. This tools was implemented as own solution and they support one access to all tools and processes.

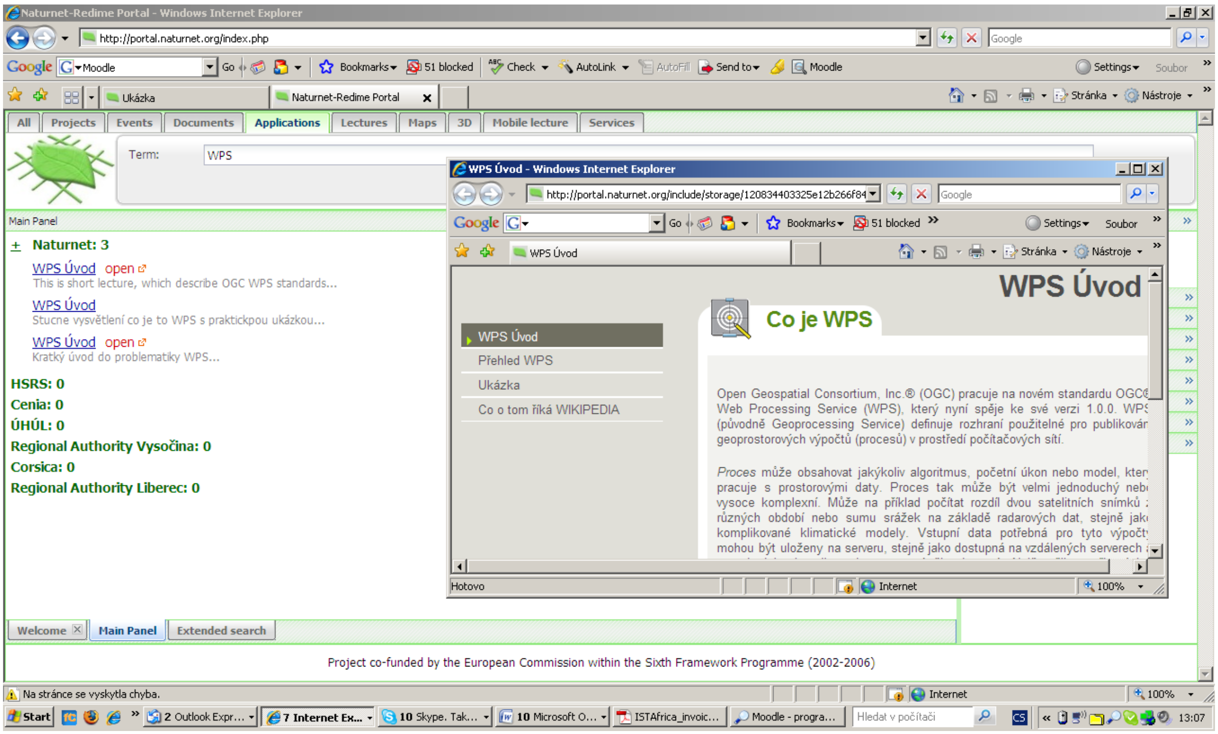


Fig.D3.16: Web pages publishing.

Portal

The portal has two main parts: Catalogue client, which guarantee access to more metadata catalogues and HSLayers as visualisation client.

Portal Catalogue client supporting search in more metadata systems

CatClient – catalogue service client. It enables to search data in configuration-enabled list of connected catalogues. Accessible web services can be displayed directly in the connected map viewer. CatClient can be installed also separately without basic system MIcKA. Current solution on portal is independent on metadata system, is working separately and support search trough mode catalogue systems.

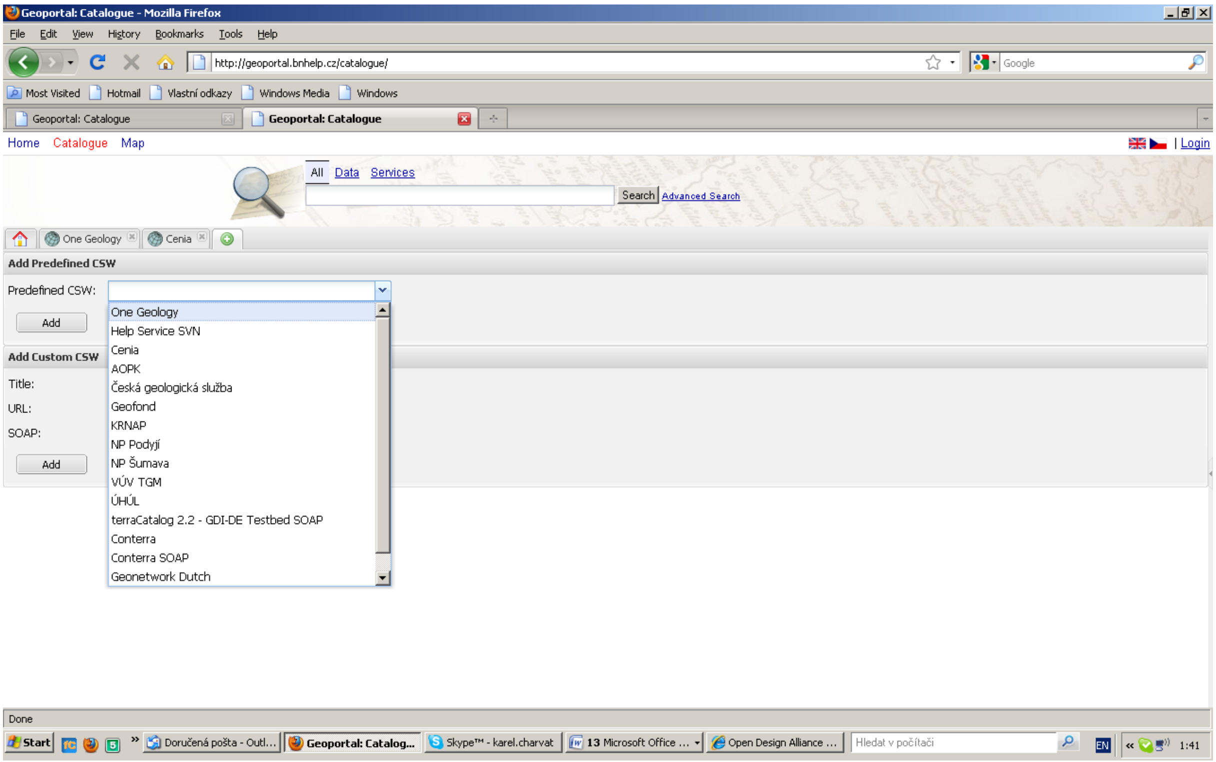


Fig.D3.17: Catalogue client with more servers.

HSlayers

The set of tools HS Layers is used in new generations of our map applications. It is based on JavaScript libraries OpenLayers and ExtJS. Using these tools it is possible to create applications according to user needs. These applications can vary from simple maps accessible through a web site to complicated applications that are similar to desktop GIS tools in their control and functionality. HSLayers are offered in two versions light and portal.

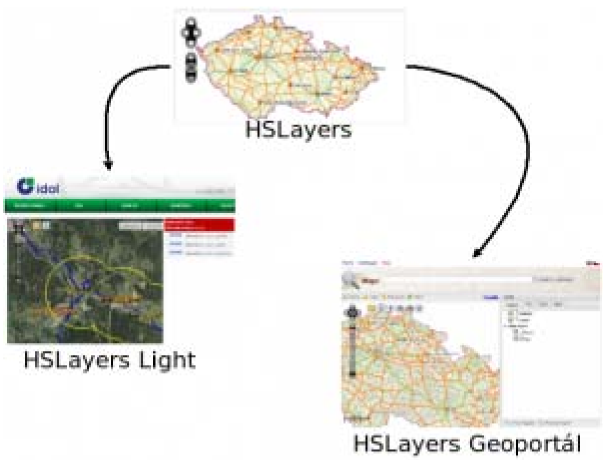


Fig.D3.18: Hslayers.

HSLayers features are coming up from OpenLayers and therefore their characteristics are as follows:

- Portrayal of various types of data:
  - Raster: OGC WMS(-T), Image (PNG, JPEG, GIF), ...

  - Vector: OGC WFS(-T), GML, GeoRSS, KML, GPX, GeoJSON, ...
  - Data sources from commercial servers: Google Maps, Virtual Earth, Yahoo Maps, ...
- The user interface (use control) adheres to current conventions in web map portals.
- Information about queried objects in text bubbles.

HSLayers additional functions:

- Dynamic adding of OGC services into map - clients for WMS and WFS
- Portrayal of independent data sources on the client side. Map composition is composed of the basis of requests to various servers. It is thus not necessary to install a map server.
- Saving of map composition according to WMC (Web Map Context) OGC specification on user computer for repeated future use or for sharing between users.
- Extension of compute functions based on WPS (Web Processing Service) OGC service - according to user needs
- Multilingual environment
- Map requests to various types of data stored on various servers, with automatic processing of results
- Work with micro-formats
- Search on the map
- On-the-fly projection on client side including S-JTSK coordinate system (direct portrayal of e.g. GPS data is therefore possible)
- Connection of the application with catalogue client (OGC CSW) in the geoportal, which enables display of the searched service from catalogue directly on the map.
- Edit function - snapping to chosen layers
- Possibilities for advanced configuration of user requests
- Advanced measuring of length and surfaces
- Print of map compositions - possibility of large print outs (up to A0 format), user configuration of print settings
- Display of description information of cadastre directly from COSMC web site.

Fig.D3.19: Tourism portal Posazavi, A map is a part of the web site.

Light version of HSLayers was made to be inserted into web sites. Its functionalities are limited in comparison with the portal version, especially in user comfort (the light version does not use ExtJS library as the user interface).

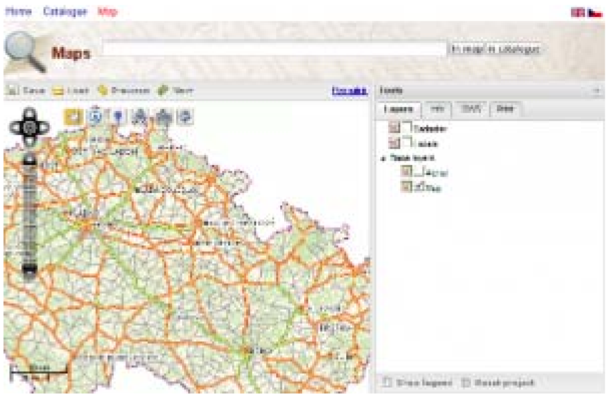


Fig.D3.20: INSPIRE Geoportal.

The Geoportal version is a heavy version of the web based map viewer HSLayers. Beside a map with the necessary control features (layer management, length measurements, map inquiries throughout all displayed layers, print, etc.), catalogue client, MapMan (tool for creation map compositions, supplied by Help Forest), authentication module, metadata editor MIcKA, and other features can be used. All components are interconnected and communicate with each other.

ANNEXES

A.1 INSPIRE themes

Annex 1

1. Coordinate reference systems

Systems for uniquely referencing spatial information in space as a set of coordinates (x, y, z) and/or latitude and longitude and height, based on a geodetic horizontal and vertical datum.

2. Geographical grid systems

Harmonised multi-resolution grid with a common point of origin and standardised location and size of grid cells.

3. Geographical names

Names of areas, regions, localities, cities, suburbs, towns or settlements, or any geographical or topographical feature of public or historical interest.

4. Administrative units

Units of administration, dividing areas where Member States have and/or exercise jurisdictional rights, for local, regional and national governance, separated by administrative boundaries.

5. Addresses

Location of properties based on address identifiers, usually by road name, house number, postal code.

6. Cadastral parcels

Areas defined by cadastral registers or equivalent.

7. Transport networks

Road, rail, air and water transport networks and related infrastructure. Includes links between different networks.

8. Hydrography

Hydrographic elements, including marine areas and all other water bodies and items related to them, including river basins and sub-basins.

9. Protected sites

Area designated or managed within a framework of international, Community and Member States' legislation to achieve specific conservation objectives.

Annex 2

1. Elevation

Digital elevation models for land, ice and ocean surface. Includes terrestrial elevation, bathymetry and shoreline.

2. Land cover

Physical and biological cover of the earth's surface including artificial surfaces, agricultural areas, forests, (semi-)natural areas, wetlands, water bodies.

3. Orthoimagery

Geo-referenced image data of the Earth's surface, from either satellite or airborne sensors.

4. Geology

Geology characterized according to composition and structure. Includes bedrock, aquifers and geomorphology.

Annex 3

1. Statistical units

Units for dissemination or use of statistical information.

2. Buildings

Geographical location of buildings.

3. Soil

Soils and subsoil characterized according to depth, texture, structure and content of par ticles and organic material, stoniness, erosion, where appropriate mean slope and anticipated water storage capacity.

4. Land use

Territory characterized according to its current and future planned functional dimension or socio-economic purpose (e.g. residential, industrial, commercial, agricultural, forestry, recreational).

5. Human health and safety

Geographical distribution of dominance of pathologies (allergies, cancers, respiratory diseases, etc.), information indicating the effect on health (biomarkers, decline of fertility, epidemics) or well-being of humans (fatigue, stress, etc.) linked directly (air pollution, chemicals, depletion of the ozone layer, noise, etc.) or indirectly (food, genetically modified organisms, etc.) to the quality of the environment.

6. Utility and governmental services

Includes utility facilities such as sewage, waste management, energy supply and water supply, administrative and social governmental services such as public administrations, civil protection sites, schools and hospitals.

7. Environmental monitoring facilities

Location and operation of environmental monitoring facilities includes observation and measurement of emissions, of the state of environmental media and of other ecosystem parameters (biodiversity, ecological conditions of vegetation, etc.) by or on behalf of public authorities.

8. Production and industrial facilities

Industrial production sites, including installations covered by Council Directive 96/61/EC of 24 September 1996 concerning integrated pollution prevention and control and water abstraction facilities, mining, storage sites.

9. Agricultural and aquaculture facilities

Farming equipment and production facilities (including irrigation systems, greenhouses and stables).

10. Population distribution — demography

Geographical distribution of people, including population characteristics and activity levels, aggregated by grid, region,
administrative unit or other analytical unit.

11. Area management/restriction/regulation zones and reporting units

Areas managed, regulated or used for reporting at international, European, national, regional and local levels. Includes dumping sites, restricted areas around drinking water sources, nitrate-vulnerable zones, regulated fair ways at sea or large inland waters, areas for the dumping of waste, noise restriction zones, prospecting and mining permit areas, river basin districts, relevant reporting units and coastal zone management areas.

12. Natural risk zones

Vulnerable areas characterized according to natural hazards (all atmospheric, hydrologic, seismic, volcanic and wildfire phenomena that, because of their location, severity, and frequency, have the potential to seriously affect society), e.g. floods, landslides and subsidence, avalanches, forest fires, earthquakes, volcanic eruptions.

13. Atmospheric conditions

Physical conditions in the atmosphere. Includes spatial data based on measurements, on models or on a combination thereof and includes measurement locations.

14. Meteorological geographical features

Weather conditions and their measurements; precipitation, temperature, evapotranspiration, wind speed and direction.

15. Oceanographic geographical features

Physical conditions of oceans (currents, salinity, wave heights, etc.).

16. Sea regions

Physical conditions of seas and saline water bodies divided into regions and sub-regions with common characteristics.

17. Bio-geographical regions

Areas of relatively homogeneous ecological conditions with common characteristics.

18. Habitats and biotopes

Geographical areas characterized by specific ecological conditions, processes, structure, and (life support) functions that physically support the organisms that live there. Includes terrestr ial and aquatic areas distinguished by geographical, abiotic and biotic features, whether entirely natural or semi-natural.

19. Species distribution

Geographical distribution of occurrence of animal and plant species aggregated by grid, region, administrative unit or other analytical unit.

20. Energy resources

Energy resources including hydrocarbons, hydropower, bio-energy, solar, wind, etc., where relevant including depth/height information on the extent of the resource.

21. Mineral resources

Mineral resources including metal ores, industrial minerals, etc., where relevant including depth/height information on the extent of the resource.

A.2 Proposition for a Data Policy for the EnviroGRIDS project

NB: This is a draft proposition for a data policy that must be discussed and completed with all the EnviroGRIDS partners during the next coming 6 month in order to be officially presented and endorsed during the General Assembly in Tulcea (Romania) in April 2010.

1. Scope of this document

This document describes the key issues relating to the supply, custody and use of data in the EnviroGRIDS project through the EnviroGRIDS Spatial Data Infrastructure (SDI).

2. Principles

The overall policy is designed to serve the following aims:

- Timely, easy and free access to the SDI by the EnviroGRIDS community;
- maximum use of the SDI, for data exchange within the EnviroGRIDS community;
- maximum use of the SDI, for presenting the results of EnviroGRIDS various end users;
- easy access to the SDI for users and stakeholders at large, now and in the future.

The policy needs to ensure:

- any existing ownership rights are respected,
- the ownership of each dataset is acknowledged, well referenced;
- the owners are protected from any liability arising from the use of their data;
- groups of users with preferential rights of access to any data are clearly defined;

3. The EnviroGRIDS Spatial Data Infrastructure (EG-SDI) and metadata standards

The EG-SDI will contain the data and metadata stored according to international standards (e.g. ISO 19115) and fed on a continuous basis by the partners tough a web interface.

The SDI contains the metadata and the listing of data owners (EnviroGRIDS partners and their collaborating institutes) who must fill-in their metadata on a mandatory basis. The metadata shall include information on how to access the data or at least identify the person to be contacted to access the datasets.

All WP leaders are to communicate to the SDI management which data and derived products are to be produced and provided to the SDI.

All WP leaders must inform the SDI management on any data from the database that they will use within their planned activities.

Additional information might be sent by the WP/module leaders to the SDI management as deemed necessary.

Access to the SDI for the purpose of adding would be restricted on the basis of password privileges granted only to WP/module leaders.

4. Data access

Data access will vary depending on the type of data and whether the party requesting access is a member of the EnviroGRIDS consortium or not:

The scientific teams and partners will always have free and unlimited access to the metadata of SDI. However they are not permitted to exchange, to sell nor to give away data outside the EnviroGRIDS community without permission of the data owner.

After completion of the project and completion of planned scientific publications, metadata included into the SDI shall be freely available to the scientific community.

Observations collected in the frame of or developed within EnviroGRIDS and the model output remain with the data owners who are encouraged to facilitate the work of EnviroGRIDS by making them available to the scientific teams and partners of EnviroGRIDS.

5. FP7 special clause

Furthermore, the consortium will apply the special clause 29 applicable to the FP7 model grant agreement on access rights to foregrounds (data) for policy purposes and transfer of ownership of foregrounds, which is specific to environment research:

The Project should ensure that protocols and plans for data collection and storage are in line with Community Data Policy.

The Community Institutions and Bodies shall enjoy access rights to foreground for the purpose of developing, implementing and monitoring environmental policies. Such access rights shall be granted by the beneficiary concerned on a royalty-free basis.

Where foreground will no longer be used by the beneficiary nor transferred, the beneficiary concerned will inform the Commission. In such case, the Commission may request the transfer of ownership of such foreground to the Community. Such transfer shall be made free of charge and without restrictions on use and dissemination.

6. GEOSS special clause

The societal benefits of Earth observations cannot be achieved without data sharing. EnviroGRIDS being part of the GEOSS workplan for 2009-2011 it should therefore apply the following GEOSS data sharing principles to present the outputs of the projects to the GEOSS community:

There will be full and open exchange of data, metadata, and products shared within GEOSS, recognizing relevant international instruments and national policies and legislation.

All shared data, metadata, and products will be made available with minimum time delay and at minimum cost.

All shared data, metadata, and products free of charge or no more than cost of reproduction will be encouraged for research and education.

REFERENCES

Almirall G., Bergada M.M, Ros P.Q., Craglia M. (2007) The Socio-Economic Impact of the Spatial Data Infrastructure of Catalonia, JRC Scientific and Technical Reports, 61p.
http://inspire.jrc.ec.europa.eu/reports/Study_reports/catalonia_impact_study_report.pdf

Arzberger P., Schroeder P., Beaulieu A., Bowker G., Casey K., Laaksonen L., Moorman D., Uhlir P., Wouters P. (2004) Promoting Access to Public Research Data for Scientific, Economic and Social Development, Data Science Journal, Vol.3, pp.135-152

Bejar R., Latre, M.A., Nogueras-Iso J., Muro-Medrano P.R., Zarazaga-Soria F.J. (2009) Systems of Systems as a Conceptual Framework for Spatial Data Infrastructures, International Journal of Spatial Data Infrastructure Research, Vol.4, 15p.
http://ijsdir.jrc.ec.europa.eu/index.php/ijsdir/article/view/124/105

Bernard L., Craglia M. (2005) SDI - From Spatial Data Infrastructure to Service Driven Infrastructure, Workshop on Cross-learning on Spatial Data Infrastructures (SDI) and Information Infrastructures (II), 8p.
http://www.ec-gis.org/sdi/ws/crosslearning/papers/PP%20Lars%20Bernard%20-%20Max%20Craglia.pdf

Boes U., Pavlova R. (2007) The Web 2.0 - A New Challenge For GIS And Cartography, Conference on recent problems in Geodesy and related fields with international importance, Sofia, Bulgaria, 10p.
http://www.rgz.sr.gov.yu/DocF/Files/intergeo-east-2007/n11.pdf

Boes U., Pavlova R. (2008) Is there a Future for Spatial Data Infrastructures?, GI Days 2008: Interoperability and spatial processing in GI applications, 10p.
http://www.gi-tage.de/archive/2008/downloads/acceptedPapers/Papers/Boes,Pavlova.pdf

Booz Allen Hamiltion (2005) Geospatial Interoperability Return on Investment Study, NASA, Geospatial Interoperability Office, 47p.
http://www.egy.org/files/ROI_Study.pdf

Coleman, D.J. and J.D. McLaughlin and S.E. Nichols (1997) Building a Spatial Data Infrastructure, Proceedings of the 64th Permananent Congress Meeting of the Fédération Internationale des Géometres (FIG), pp. 89-104, Singapore, May.

Coleman D.J., Georgiadou Y., Labonte J. (2009) Volunteered Geographic Information: the nature and motivation of produsers, GSDI-11 Conference Proceedings, 20p.
http://www.gsdi.org/gsdi11/papers/pdf/279.pdf

Cömert C. (2004) Web Services and National Data Infrastructure (NSDI), ISPRS Conference, 6p.
http://cartesia.org/geodoc/isprs2004/comm4/papers/365.pdf

Craglia M., Novak J. (2006) Assessing the impacts of Spatial Data Infrastructures, Report of International Workshop on Spatial Data Infrastructures' Cost-Benefit / Return on Investment, Ispra, Italy, 12-13 January, 57p.

Craglia M. (2007) Volunteered Geographic Information and Spatial Data Infrastructures: when do parallel lines converge?, VGI Specialist Meeting, Santa Barbara, USA, 3p.
http://www.ncgia.ucsb.edu/projects/vgi/docs/position/Craglia_paper.pdf

Craig W.J. (2005) White Knights of Spatial Data Infrastructure: The Role and Motivation of Key Individuals, URISA Journal, Vol.16, pp.5-13
Crompvoets J., Bregt A. (2003) World Status of National Spatial Data Clearinghouses, URISA Journal, Vol.15, pp.43-50

De Longueville B., Ostländer N, Keskitalo C. (2009) Addressing vagueness in Volunteered Geographic Information (VGI) – A case study, GSDI-11 Conference Proceedings, 21p.
http://www.gsdi.org/gsdi11/papers/pdf/221.pdf

European Commission (2007) Directive 2007/2/EC of the European Parliament and the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE), 14p
http://eur-lex.europa.eu/JOHtml.do?uri=OJ:L:2007:108:SOM:EN:HTML

European Commission (2008) Commission regulation (EC) No 1205/2008 of 3 December 2008 implementing Directive 2007/2/EC of the European Parliament and of the Council as regard metadata, 19p.
http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:326:0012:0030:EN:PDF

European Commission (2008) INSPIRE Generic Conceptual Model, 109p.
http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/DataSpecifications/D2.5_v3.1.pdf

European Commission (2008) INSPIRE Network Services Architecture, 29p.
http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/network/D3_5_INSPIRE_NS_Architecture_v3-0.pdf

European Commission (2007) INSPIRE Technical Architecture - Overview, 12p.
http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/network/INSPIRETechnicalArchitectureOverview_v1.2.pdf

GEO (2005) Global Earth Observation System of Systems 10-Year Implementation Plan Reference Document, 209p
http://www.earthobservations.org/documents/10-Year%20Plan%20Reference%20Document.pdf

GEO (2005) The Global Earth Observation System of Systems (GEOSS) 10-Year Implementation Plan, 11p
http://www.earthobservations.org/documents/10-Year%20Implementation%20Plan.pdf

GEO (2006) GEO Capacity Building Strategy, 13p.
http://www.earthobservations.org/documents/committees/cbc/capacity_building_strategy.pdf

GEO (2007) The Full Picture, 143p.
http://www.earthobservations.org/documents/the_full_picture.pdf

GEO (2008) White Paper on the GEOSS Data Sharing Principles, 93p.
http://www.earthobservations.org/documents/dsp/Draft%20White%20Paper%20for%20GEOSS%20Data%20Sharing%20Policies_27Sept08.pdf

Goodchild M.F. (2007) Citizens as sensors: The World of Volunteered Geography, Workshop on Volunteered Geography Information, Santa Barbara, USA, 15p.
http://www.ncgia.ucsb.edu/projects/vgi/docs/position/Goodchild_VGI2007.pdf

Gore A. (1998) The Digital Earth: Understanding our planet in the 21st Century, 5th international Symposium on Digital Earth, 5p.
http://www.isde5.org/al_gore_speech.htm

Lehmann A. (2009) EnviroGRIDS, Document of Work, 131p.

Liping D., Chen A., Yang W., Zhao P. (2003) The Integration of Grid Technology with the OGC Web Services (OWS) in NWGISS for NASA EOS Data, , 10p.

Masser I. (2006) What's Special about SDI Related Research, International Journal of Spatial Data Infrastructures Research, Vol.1, pp.14-23

Masser I. (2005) The Future of Spatial Data Infrastructures, ISPRS Workshop on Service and Application of Spatial Data Infrastructure, Oct.14-16, Hangzhou, China, 9p.
http://www.commission4.isprs.org/workshop_hangzhou/papers/7-16%20Ian%20Masser-A001.pdf

Masser I. (2007), Building European Spatia Data Infrastructures, ESRI Press, 91p.

Mohammadi H., Rajabifard A. Williamson I.P. (2007) Multi-source Spatial Data Integration within the Context of SDI Initiatives, 18p, Internation Journal of Spatial Data Infrastructures Research, 18p.
http://ijsdir.jrc.ec.europa.eu/index.php/ijsdir/article/view/74/34

Muresan O., Pop F., Gorgan D., Cristea V. (2006) Satellite Image Processing Applications in MedioGRID, Satellite image processing application in mediogrid, pp.253-262

Open Geospatial Consortium (2005) The Importance of Going "Open", OGC White Paper, 8p.
http://portal.opengeospatial.org/files/?artifact_id=6211&version=2&format=pdf

Open Geospatial Consortium (2004) The Havoc of Non-Interoperability, OGC White Paper, 7p.
http://portal.opengeospatial.org/files/?artifact_id=5097&version=3&format=pdf

Open Geospatial Consortium (2004) Integrating Geospatial Standards and Standards Strategies Into Business Process, OGC White Paper, 7p.
http://portal.opengeospatial.org/files/?artifact_id=5098&version=2&format=pdf

Open Geospatial Consortium (2004) The Spatial Web, OGC White Paper, 8p.
http://portal.opengeospatial.org/files/?artifact_id=3859&version=2&format=pdf

Open Geospatial Consortium (2007) OpenGIS Web Service Common Implementation Specification, 153p.
http://www.opengeospatial.org/standards/common

Open Geospatial Consortium (2007) OpenGIS Geography Markup Language (GML) Encoding Standard, 437p.
http://www.opengeospatial.org/standards/gml

Open Geospatial Consortium (2009) The OpenGIS Abstract Specification Topic 5: Features, 50p.
http://www.opengeospatial.org/standards/as

Open Geospatial Consortium (2008) OpenGIS Reference Model, 35p.
http://www.opengeospatial.org/standards/orm

Open Geospatial Consortium (2004) Geospatial Portal Reference Architecture, 23p.
http://portal.opengeospatial.org/files/?artifact_id=6669

Ostensen O. (2001) Global geomatics standards supporting sustainable geospatial data infrastructures, GSDI 5 Conference Proceedings, 6p.
http://gsdidocs.org/gsdiconf/GSDI-5/papers/Olaf%20Ostensen-paper.pdf

Padberg A., Greve K. (2009) Gridification of the OGC Web Processing Service: Challenges and Portential, AGILE Workshop, 5p.

Paluszynski W., Iwaniak A., Sliwinski A., Kubik T. (2007) Building Spatial Data Infrastructure At County Level Using Open Software, SDI-East Africa, 11p.
http://dewa03.unep.org/sdi-ea/system/files/BUILDING_SPATIAL_DATA_INFRASTRUCTURE_AT_C.doc

Phillips A., Williamson I., Ezigbalike C. (1999) Spatial Data Infrastructure Concepts, The Australian Surveyor, Vol.44 No1, p20-28
http://www.sli.unimelb.edu.au/research/publications/IPW/SDIDefinitionsAusSurv.html

Rajabifard A. and Williamson I.P. (2001) Spatial Data Infrastructures: Concept, SDI Hierarchy and Future directions, in Proceedings, of GEOMATICS'80 Conference, Tehran, Iran., 10p.
http://repository.unimelb.edu.au/10187/1247

Rajabifard A. (2002) Diffusion of Regional Spatial Data Infrastructures: with particular reference to Asia and the Pacific, Ph.D thesis, Department of Geomatics, Faculty of Engineering, The University of Melbourne, 213p.

Rajabifard A. and Williamson I.P. (2004) SDI Development and Capacity Building, GSDI-7 Conference on Spatial Data Infrastrcture for a Sustainable Futures, 12p.
http://www.geoinfo.ait.ac.th/download/SCOSA2007/1_DrAbbas/4-SDIDevelopmentCapacityBuilding.pdf

Ryttersgaard J. (2001) Spatial Data Infrastructure, Developing Trends and Challenges, International Conference on Spatial Information for Sustainable Development Proceedings, Nairobi, 8p.
http://www.fig.net/pub/proceedings/nairobi/ryttersgaard-TS1-1.pdf

Sahin K., Gumusay M.U. (2008) Service Oriented Architecture (SOA) based Web Services for Geographic Information Systems, ISPRES Congress 2008, pp.625-630
http://www.isprs.org/congresses/beijing2008/proceedings/2_pdf/5_WG-II-5/03.pdf

Sayar A., Pierce M., Fox G. (2005) OGC Compatible Geographical Information Systems Web Services, Indiana University , Departement of Computer Science, Technical Report, 64p.
http://www.cs.indiana.edu/pub/techreports/TR610.pdf

Stevens A.R, Onsrud J.H, Rao M.(2005) Global Spatial Data Infrastructure (GSDI): Encouraging SDI Development Internationally, ISPRS Conference on Service and Application of Spatial Data Infrastructure, 14-16 October, Hangzhou, China, 10p.
http://www.commission4.isprs.org/workshop_hangzhou/papers/309-312%20Abs&Papr%20GSDI%20China%2014Oct05.pdf

UNECA (2005) SDI Africa: An Implementation Guide, 120p.
http://geoinfo.uneca.org/sdiafrica/default1.htm

UNGIWG (2007) UNSDI Compendium - A UNSDI Vision, Implementation Strategy and Reference Architecture, 150p.
http://www.ungiwg.org/docs/unsdi/UNSDI_Compendium_13_02_2007.pdf

Vaccari L., Shvaiko P., Marchese M. (2008) An emergent semantics approach to semantic intergration of geo-services and geo-metadata in Spatial Data Infrastructures, IJSDIR, 23p.
http://ijsdir.jrc.ec.europa.eu/index.php/ijsdir/article/view/107/113

Williamson I.P, Rajabifard A., Enemark S. (2003) Capacity Building for SDIs, Proceedings 16th UN Regional Cartographic Conference., 14p.
http://repository.unimelb.edu.au/10187/1171

Williamson I., Rajabifard A., Binns A. (2006) Challenged and Issues for SDI Development, International Journal of Spatial Data Infrastructures Research, Vol.1 pp.24-35

Wilson M., Von Hagen C., Howard C. (2009) SDI in East Africa - Leveraging the UN presence, International Journal of Spatial Data Infrastructure Research, Vol.4, pp.1-23
http://ijsdir.jrc.ec.europa.eu/index.php/ijsdir/article/view/109/112

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS & ACRONYMS

BMP:	BitMaP
CERN:	Centre Européen pour la Recherche Nucléaire
CSW:	Catalog Service for the Web
ECW:	ERMapper Compress Wavelets
EGEE:	Enabling Grids for E-scienceE
EO:	Earth Observation
GEO:	Group on Earth Observations
GEOSS:	Global Earth Observation System of Systems
GDAL:	Geospatial Data Abstraction Library
GIF:	Graphics Interchange Format
GIS:	Geographic Information System
GML:	Geographic Markup Language
GPL:	General Public Licence
GPS:	Global Positiong System
GRID:	Global Resource Information Database
GSDI:	Global Spatial Data Infrastructure
INSPIRE:	Infrastructure for Spatial Information in the European Community
ISO:	International Organization for Standardization
IR:	Implementing Rules
IT:	Information Technology
JPEG:	Joint Photographic Experts Group
JSON:	JavaScript Object Notation
KML:	Keyhole Markup Language
OAI:	Open Archive Initiative
OASIS:	Organization for the Advancement of Structured Information Standards
OGC:	Open Geospatial Consortium
OWS:	OGC Web Services
PDF:	Portable Document Format
PNG:	Portable Network Graphics
RDBMS:	Relational DataBase Management System
REST:	Representational State Transfer
RPC:	Remote Procedure Call
RSS:	Really Simple Syndication
SDI:	Spatial Data Infrastructure
SDK:	Software Development Kit
SHP:	Shape File
SLD:	Styled Layer Descriptor
SOA:	Service Oriented Architecture
SOAP:	Simple Object Access Protocol
SOS:	Sensor Observation Service
SQL:	Structured Query Language
SVG:	Scalable Vector Graphics
SWE:	Sensor Web Enablement
TIFF:	Tagged Image File Format
UN:	United Nations
UNEP:	United Nations Environment Programme
UNECA:	United Nations Economic Commission for Africa
UNGIWG:	United Nations Geographical Information Working Group
URM:	Uniform Resource Management
VGI:	Volunteered Geographic Information
WAS:	Web Service Authentication

W3C:	World Wide Web consortium
WMS:	Web Map Service
WFS:	Web Feature Service
WCS:	Web Coverage Service
WSDL:	Web Service Description Language
WSS:	Web Security Service
WPS:	Web Processing Service
WPVS:	Web Perspective and View Service
WTS:	Web Terrain Service
XML:	eXtended Markup Language