Access to ESPON Database by third-party applications

An expertise on the potentialities of web services for accessing the ESPON metadata and data

UNEP/GRID-Geneva
27/06/2013

ABOUT THIS DOCUMENT

A technical expertise

This document is the result of a technical expertise conducted in Spring 2013 by UNEP/GRID-Geneva in the context of the M4D project (Multi-Dimensional Data Base Design and Development – Data Base 2013 phase II, http://www.espon.eu/main/Menu_Projects/Menu_ScientificPlatform/espondatabase2013phaseII.html) funded by the ESPON programme (http://www.espon.eu/main/Menu_Programme/), with the collaboration of the LIG STeamer partner.

Terms of reference of the expertise

The objective of the expertise was to study the access to the ESPON Database by third-party applications, possibly by means of geo-services.

In its first phase (until June 2013) the expertise includes:

1. the conformity assessment of the ESPON Database with OGC standards and the INSPIRE recommendations in particular for discovery and data download services;
2. a review of currently available solutions to query the ESPON Database and their evaluation with respect to solutions based on web services allowing the exploitation of the ESPON metadata and data by third party applications, including geospatial data portals;
3. recommendations for designing a solution based on geographic web services (geo-services), a study of the modalities related to the migration of the current application to this solution.

In its second phase (until December 2013), the expertise will consist in providing guidance and support on technological choices for the development of an application allowing the access to data and metadata of the ESPON Database by means of web services.

Authors

Gregory Giuliani, SDI Coordinator

Hy Dao, Head Metadata & Socio-economics Unit

Jean-Philippe Richard, System Administrator & Database Specialist

Andrea De Bono, GIS Analyst & Data Management Specialist

LIG STeamer contribution (Part I):

Jérôme Gensel

Camille Bernard

Laurent Poulenard

Benoit Le Rubrus

TABLE OF CONTENT

FIGURES

TABLES

1 THE ESPON DATABASE

1.1 THE M4D PROJECT AT A GLANCE

Extract from the following URL:
http://www.espon.eu/main/Menu_Projects/Menu_ScientificPlatform/espondatabase2013phaseII.html

Taking as starting point the achievements of the Project "ESPON Database 2013" (Phase I) and the results and discussions from the ongoing ESPON Project on "Territorial Indicators and Indices" (INTERCO), this project shall maintain, update, develop and further expand the ESPON 2013 Database.

In this framework, the main European activities carried out to promote data comparability and data quality, in particular the INSPIRE Directive and the GMES (Global Monitoring for Environment and Security) initiative will be taken into consideration.

In general, the ESPON 2013 Database shall supply different users (ESPON community and general public), including researchers, policy makers and stakeholders at regional and local level, with data, indicators and (visualisation) tools in the field of European territorial development and cohesion needed for policy formulation, application and monitoring at different geographical levels. The data managed by the project originates from European institutions such as EUROSTAT and EEA, and from all ESPON projects. By doing this, the ESPON 2013 Database shall contribute to better understanding of territorial structures, the current situation and past and future trends of different types of European territories in relation to the various geographical contexts (from local to global) and within a large variety of themes.

1.2 ESPON DB SYSTEM GENERAL ARCHITECTURE

Figure 1 depicts the main logical elements of the ESPON DB System. The ESPON DB System is currently made of two main elements, the Web Application (ESPON Database Portal) and the database itself.

The Web Application has five core functionalities:

1. Resources access: many resources may be accessed through the Web Application. These resources intend to help ESPON TPGs (Transnational Project Group), they may be tools, technical documents or data archives.
2. Tracking tool: this tool allows the delivery of data by the ESPON TPGs. This tool enables a process preceding integration, aiming at ensuring compliance, relevance and overall quality of the data available in the database.
3. Search interface: this interface allows users to retrieve the data, through the use of the metadata associated with it (semantic, spatial, etc.).
4. Downloads: many files or archives are available for download through the Web Application, resources, datasets, etc.
5. Users management: some functionalities or files have restricted access and therefore need authentication.

The official Web Application is available at the following URL: http://database.espon.eu

The database stores all the information related to statistical indicators (metadata, data, spatial information).

The following technologies are used:

- Database Management System (DBMS): PostgreSQL 9.1, with spatial extension PostGIS 1.5.
- The Web Application is mainly written in Java and deployed on a Glassfish Server version 3.1.

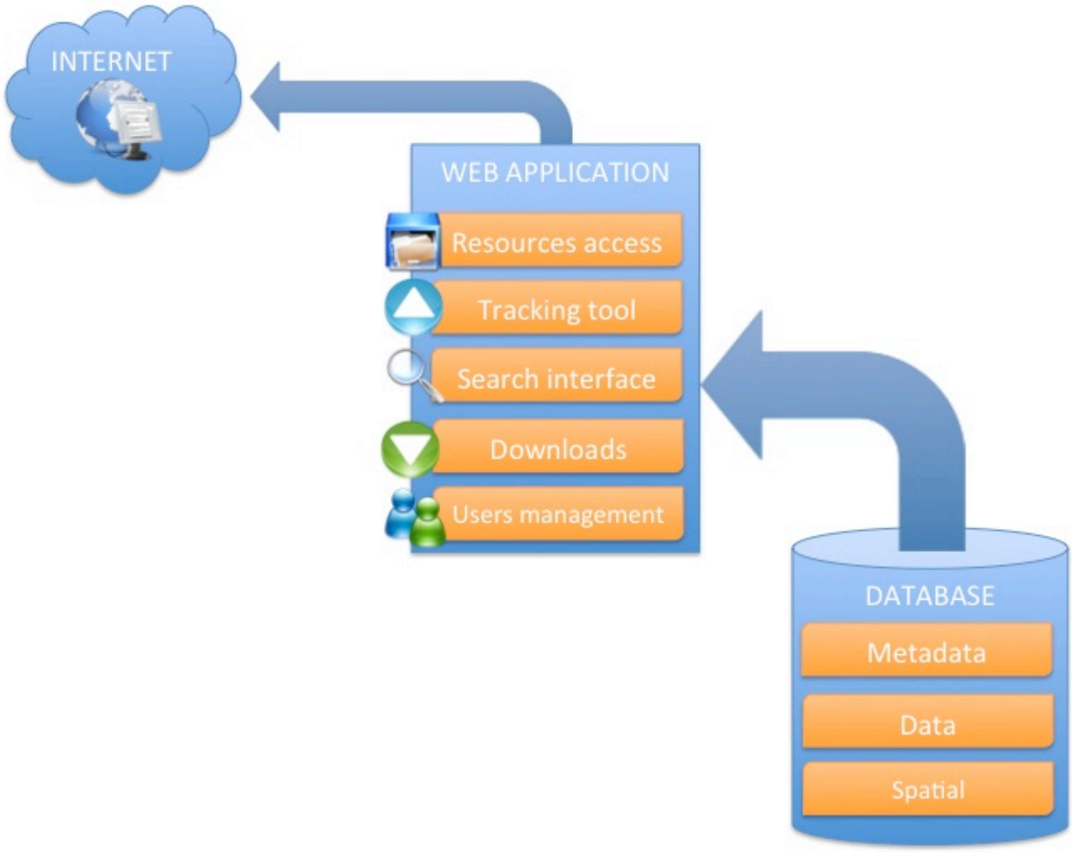


Figure 1. General overview of ESPON DB System

1.3 ESPON DATABASE

Note: in this section of the document, the word "schema" refers to a logical group to organize database objects (analogous to directories at the operating system level). For further information regarding this concept of schema, please consult http://www.postgresql.org/docs/9.1/static/ddl-schemas.html.

Regarding the scope of the web services issue, the three main schemes of the ESPON Database are:

- Metadata: conform with INSPIRE recommendations. The numerous tables of that schema help at describing the data: its nature, methodologies involved, etc.
- Data: the data itself.
- Spatial: spatial objects associated with data, for instance NUTS objects and their geometries.

The following sub-sections respectively describe these three main schemes.

1.3.1 SCHEMA METADATA

Metadata are data about data. They provide information about:

1. Indicator: code, name, dataset…
2. Dataset: name, abstract, source…
3. Source: publication, methodology…
4. Publication: provider…
5. Etc.

The metadata model is a complex system of information. Its elements have distinct roles, and are grouped into hierarchical levels. This model has been designed to be compliant with the INSPIRE directive.

The following diagram displays the main elements of this schema.

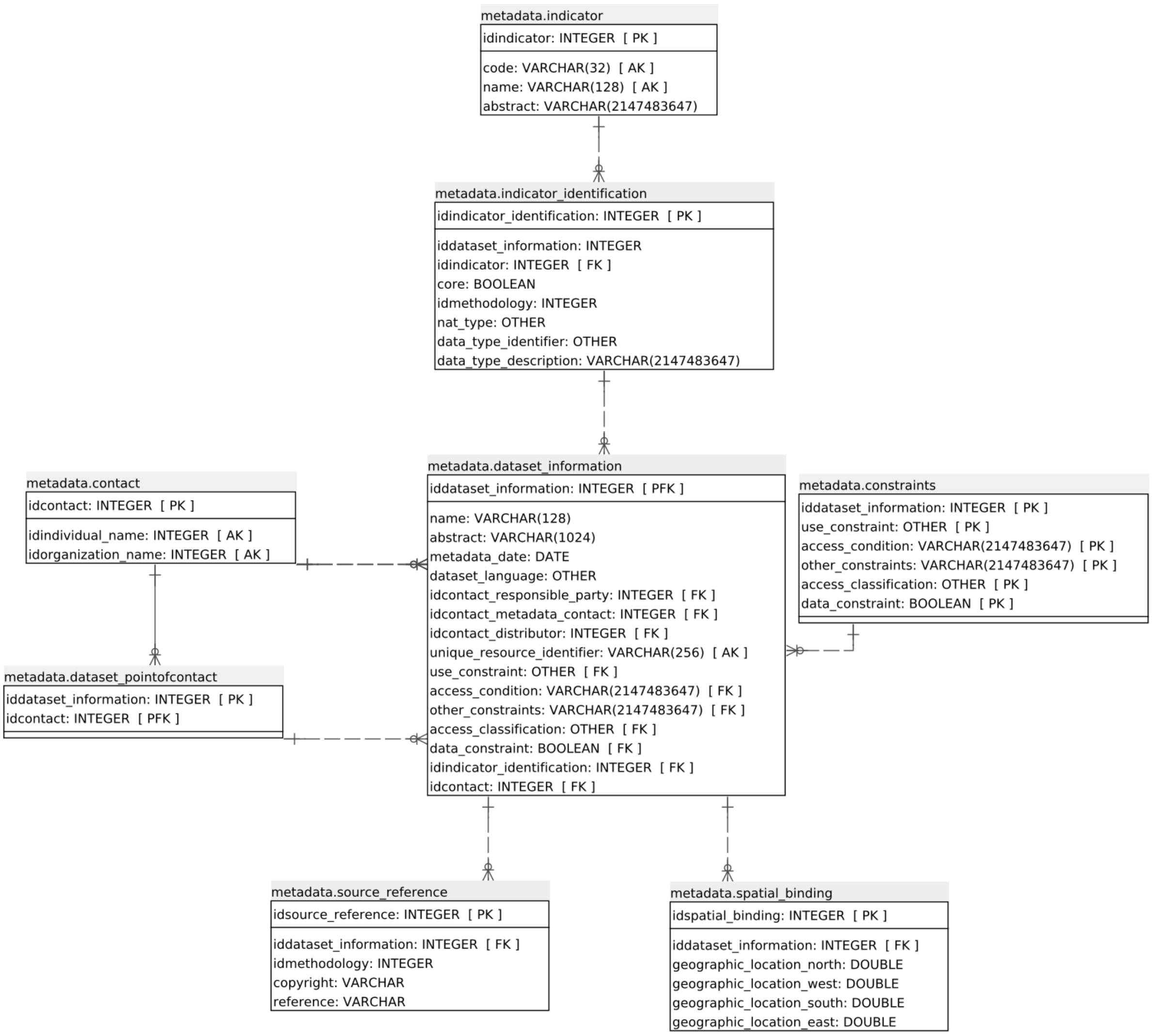


Figure 2 Physical diagram of the database schema METADATA

1.3.2 SCHEMA DATA

Data refer to the values recorded for each statistical unit at a given geographical level. The main entity is "value_registry". It references a spatial object, a statistical indicator and a temporal value. There are eight specializations of "value_registry", depending on the nature of the data (text, integer, enumeration…).

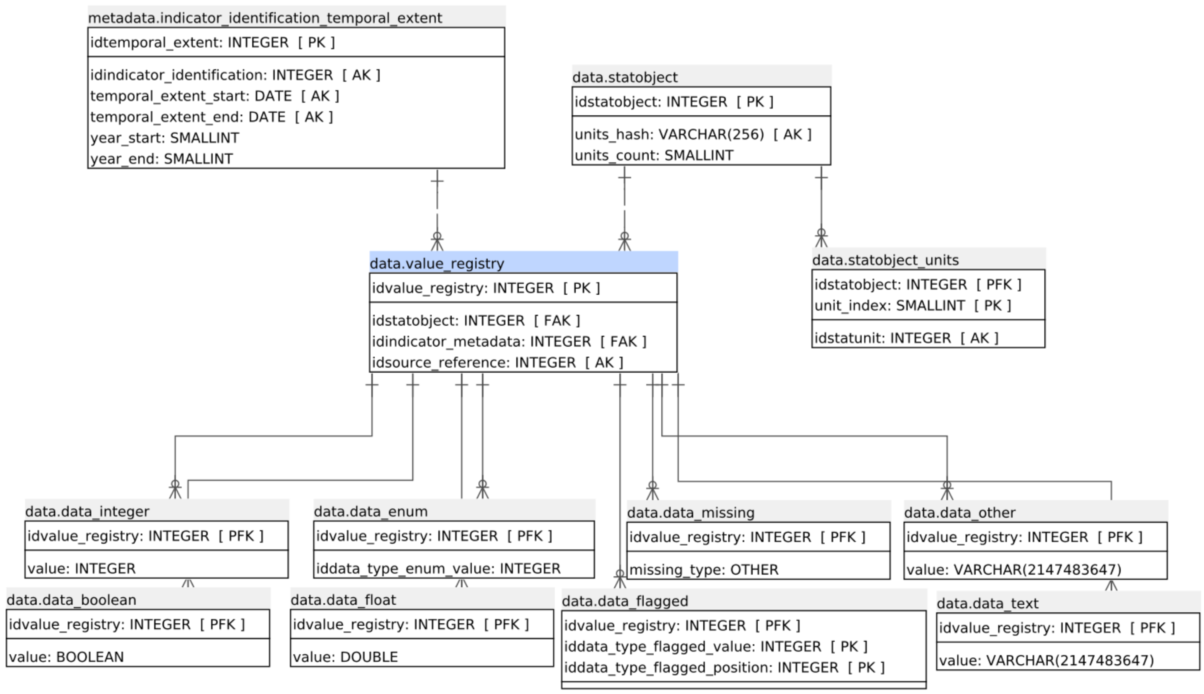


Figure 3. Physical diagram of the database schema DATA

1.3.3 SCHEMA SPATIAL

All the data included in the ESPON database are geographically referenced, by the mean of a link between data value registries and statistical geographic units. These units belong to nomenclatures (NUTS or urban nomenclatures like FUA, MUA, UMZ), and to each of them is associated a geometry. A nomenclature may have more than one version, and is composed of levels, which correspond to different scales (NUTS0: country, NUTS1: big regions, etc.).

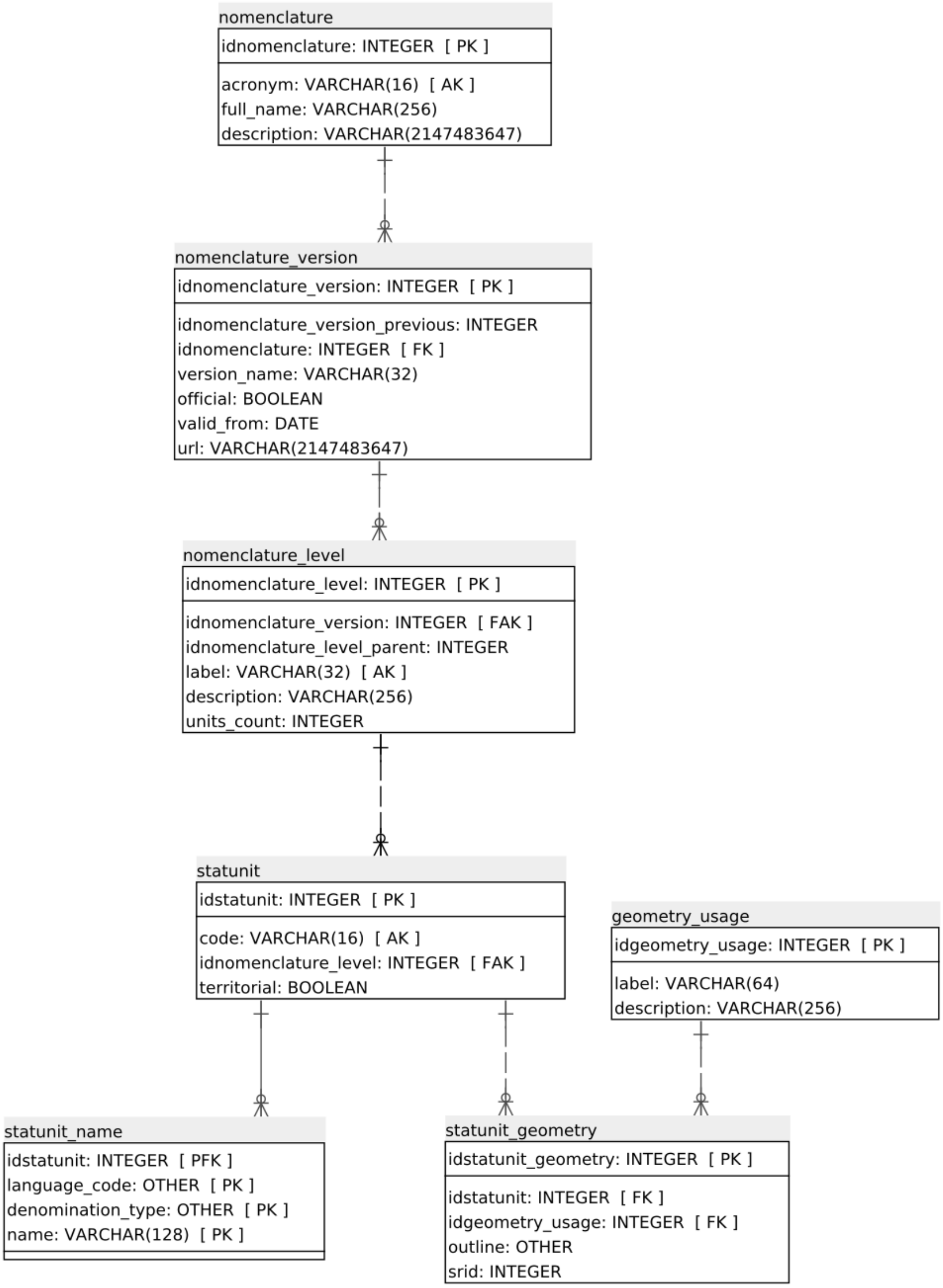


Figure 4. Physical diagram of the database schema SPATIAL

1.4 ENHANCEMENT OF ESPON DB SYSTEM

The actual architecture (database + Web Application) is sufficient to satisfy user's main needs: search, find and get data, but only for a human user, and only through the Web Application.

In order to facilitate the exchange and integration of ESPON DB contents with other (heterogeneous) external data sources, and to increase the use of these data, the LIG seeks to implement these functionalities in an interoperable manner, through the use of web services. This would allow third-party applications to access the content of the database.

The possibility of mapping data (cartographic representation) and processing data will be also tackled but those are of second priority.

Figure 5 shows the ESPON DB System enhanced with web services.

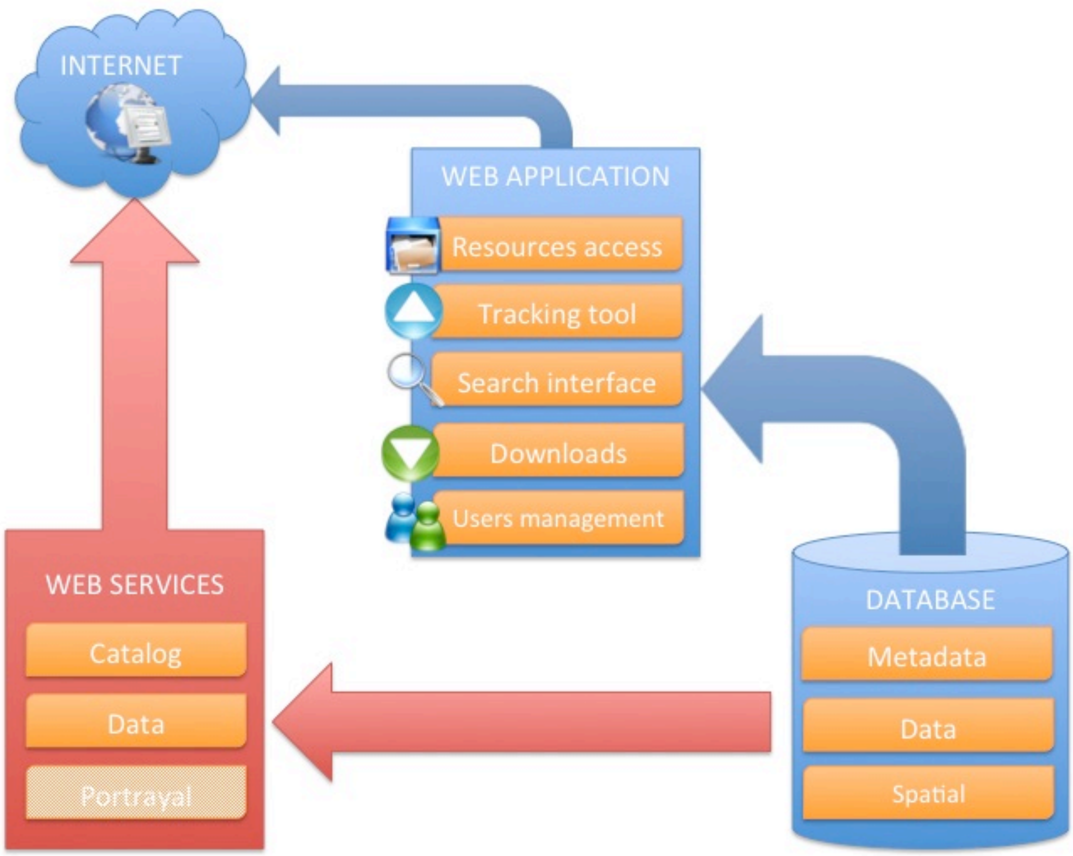


Figure 5. General overview of enhanced ESPON DB System

This would make the contents of the database accessible to human clients through the Web Application, and to third-party applications through the web services. It has to be noted that the web services should only permit read-access to the database, to be consistent with the aim of the tracking process.

2 INTEROPERABILITY

2.1 EXCHANGES BETWEEN SYSTEMS

The key element to enable and ease this exchange and integration of resources is to make the ESPON DB data interoperable. Interoperability is defined as *"the ability of a system or a product to work with other systems or products without special effort on the part of the customer."* (Open Geospatial Consortium 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 users the ability to:

- find what they need,
- access it,
- understand and employ it,
- have goods 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 sharing of data, information and computing resources (Open Geospatial Consortium 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 users to 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"* (Open Geospatial Consortium 2004; Open Geospatial Consortium 2005).

The great advantage of interoperability, and that is why it is an essential building block for the GIS and Spatial Data Infrastructure (SDI) industry, is that it describes the ability of locally managed and distributed heterogeneous systems to exchange data and information in real time to provide a service (Open Geospatial Consortium 2004). This allows 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 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 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 (Open Geospatial Consortium 2004).

The Open Geospatial Consortium (2005) has pointed out 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 public and private sectors, opening 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.

2.2 TYPES OF INTEROPERABILITY

There are two types of interoperability (Open Geospatial Consortium 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 (e.g., explaining why INSPIRE is producing data specifications).

Different types of geoprocessing systems (vector, raster, CAD, etc.) producing different types of data, different vendors of 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 (Open Geospatial Consortium 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 (Table 1).

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 1. Different types of interoperability.

As expressed by the Open Geospatial Consortium (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.

2.3 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 there 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,
- others...

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

2.4 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 (Open Geospatial Consortium 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 (Ezigbalike 2004), organizations and agencies are increasingly recognizing that standards are essential for improving productivity, market competitiveness, export capabilities (Nebert 2005), and lowering maintenance and operation costs over time (Kaufmann and Dorfschmid 2001; Booz, Allen et al. 2005; European Commission 2006).

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).

2.5 BENEFITS OF AN INTEROPERABLE ARCHITECTURE

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

- Allow sharing and reusing of data: gives access to distributed and heterogeneous sources of data.
- Avoid data duplication: data are collected and maintained at the most appropriate place.
- Reduce the costs: of maintenance, of operations and of course of production.
- Integration: as Mohammadi et al. (2010) shows, multi-source data integration can only be achieved with an effective interoperability.

- 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.
- Increase efficiency.
- Vendor neutral: avoid being locked in to one vendor.
- Improve decision-making: offering standardized access to a vast amount of data and information and to used them as effectively as they should.
- New opportunities and knowledge: open the doors to new activities and relations that were not foreseen before.

Finally, as Open Geospatial Consortium (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".

2.6 OGC GEOPORTAL ARCHITECTURE

The Open Geospatial Consortium (OGC) and the International Organization for Standardization (ISO) are the leading institutions that are providing a lot of different specifications regarding geographical information. 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 (Open Geospatial Consortium 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 SDIs. 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 straightforward integration.

The Geospatial Portal Reference Architecture specifies also four services that are needed for creating a interoperable and standardized geoportal (Figure 6):

- *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 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.

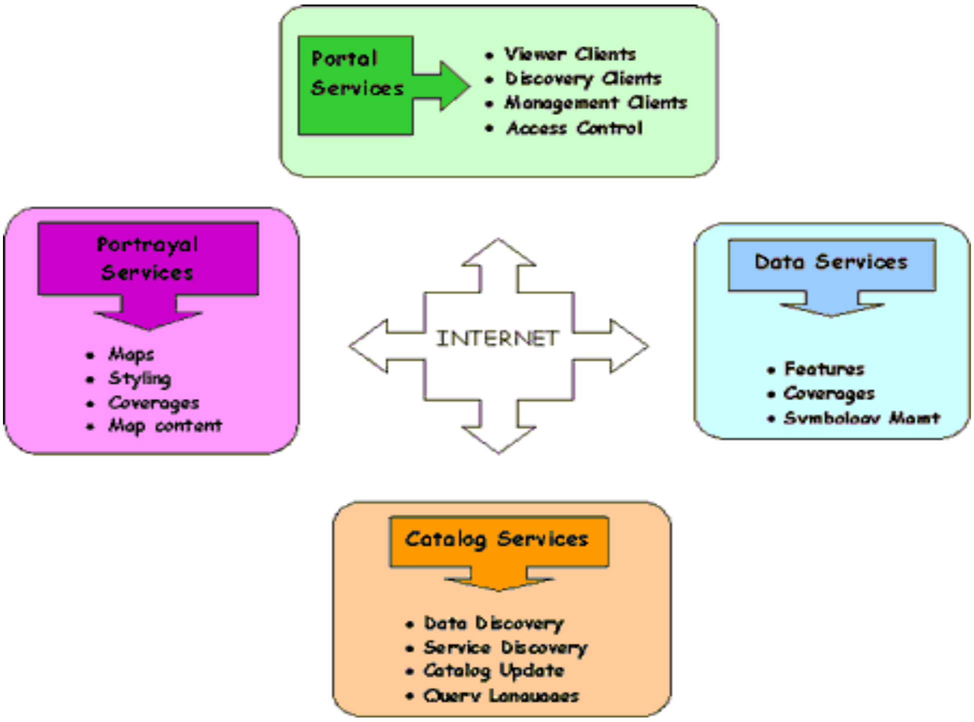


Figure 6. Geospatial Portal Reference Architecture (Source: GeoNetwork)

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 software and/or components can interoperate, work together and exchange information over a network by means of agreed standards.

For example, when two software 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 implements 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.

3 WEB SERVICES IN GENERAL

3.1 ARCHITECTURES AND PROTOCOLS

Today's effort on the technical development clearly focus on the exchange of geospatial and statistical data 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 data is highlighted by the concept of web services and related architectural styles.

Web services are a "new paradigm" (Cömert 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 rely 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 (Cömert 2004).

3.1.1 SERVICE ORIENTED ARCHITECTURE (SOA)

Service Oriented Architecture (SOA) and Resource Oriented Architecture (ROA) are the major architectural styles. They promote loose coupling between 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, and stateless (i.e., do 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 (Figure 7).

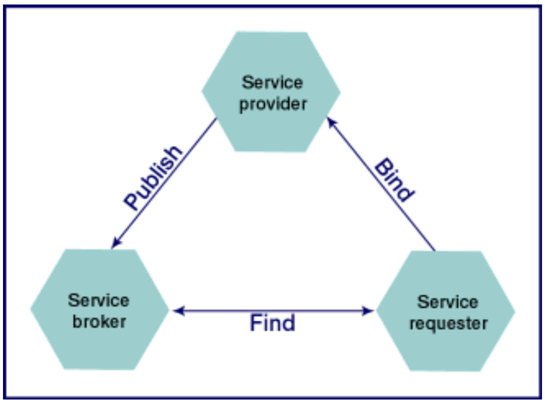


Figure 7. Basic operations in SOA (Source: IBM)

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 (Open Geospatial Consortium 2004).

3.1.2 SOAP (SIMPLE OBJECT ACCESS PROTOCOL)

SOAP (originally Simple Object Access Protocol) is a RPC (Remote Procedure Call) protocol built around XML. It became the underlying layer of a more complex set of web services, based on Web Services Description Language (WSDL). It allows the transmission between distant objects invoking methods. Generally the transfer is done through HTTP protocol but is could use other ones. SOAP protocol is divided in two pieces: an envelope, containing information about the messages and how it is transferred, and a data template.

The advantages of SOAP are the independence of platform and languages, the use of XML standards, the potential extensibility, and the possibility to use HTTP for the transport, which avoids firewalls issues for example. SOAP allows interoperability between heterogeneous architectures.

Nevertheless, the SOAP extensibility potential implied the development of several standards, and finally, this increased the complexity of the method and created differences in implementations. To be able to communicate, applications need to have a description of web services and a data dictionary definition. If services are not well built, implementations could be expensive and could result in inefficient solutions. Today, SOAP web services are often replaced by a more simple and efficient option: REST service (see next chapter). But SOAP remains the best solution when actions performed by the web service are complex or when higher security levels are needed.

3.1.3 REST (REPRESENTATIONAL STATE TRANSFER)

The **REST** architectural style was developed by W3C in parallel with HTTP/1. It is not a protocol, but an architecture that uses HTTP as protocol.

RESTful Web Services" is a web API implemented using HTTP and REST principles. RESTful provides JSON, XML and RDF/XML formats. This technological solution allows exposing a directory structure and querying attributes.

REST-style architectures conventionally consist of clients and servers. Clients initiate requests to servers; servers process requests and return appropriate responses. REST's client–server separation of concerns simplifies component implementation, reduces the complexity of connector semantics, improves the effectiveness of performance tuning, and increases the scalability of pure server components. Layered system constraints allow intermediaries—proxies, gateways, and firewalls—to be introduced at various points in the communication without changing the interfaces between components, thus allowing them to assist in communication translation or improve performance via large-scale, shared caching. REST enables intermediate processing by constraining messages to be self-descriptive: interaction is stateless between requests, standard methods and media types are used to indicate semantics and exchange information, and responses explicitly indicate cacheability.

Such as SOAP web services, REST services are independent from platforms and languages. Compared to SOAP, the most important difference is the simplicity. No need to define a data dictionary, to integrate exchange format in the message (this implies less verbosity and then better performance. It uses the HTTP logic using request method: `GET, POST, PUT, PATCH, DELETE.`

This simplicity implies also some limitations: in terms of security, of possible actions (only Create, Read, Update, Delete – CRUD and even currently some browsers are not implementing correctly Put and Delete). During the conception of web services, if developer needs are focusing on a simple resource oriented solution following the four HTPP actions CRUD, in a not too high security level environment, REST will represent an excellent choice.

3.2 OGC WEB SERVICES

This chapter briefly describes the main characteristics of several OGC web services. The discovery services (CSW) and vector data services (WFS) are those needed for the implementation of the expected functionalities (search and download of statistical data by territorial unit). Three other services - raster data download (WCS), view (WMS) and processing (WPS) - are also presented but will not be implemented in the prototype to be developed by December 2013.

3.2.1 DISCOVERY SERVICES (METADATA) USING OGC CATALOG SERVICE FOR THE WEB (CSW)

The Catalogue Service defines an interface to publish, discover, search and query metadata about geospatial 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, etc.

The CSW interface provides the following operations:

- *GetCapabilities*: allows CSW clients to retrieve service metadata from a server. The response to a GetCapabilities request is an XML document containing service metadata about the server.
- *DescribeRecord*: allows a client to discover elements of the information model supported by the target catalogue service.
- *GetRecordById*: retrieves the default representation of catalogue metadata records using their identifier.
- *GetRecords*: allows querying the catalogue metadata records specifying a query in OCG Filter or CQL languages.
- *Transaction*: defines an interface for creating, modifying and deleting catalogue records.

Example of CSW URL with a GetCapabilities request:

```
http://localhost:8080/geonetwork/srv/en/csw?request=GetCapabilities&service
=CSW&acceptVersions=2.0.2&acceptFormats=application%2Fxml
```

3.2.2 WEB FEATURE SERVICE (WFS) FOR DOWNLOADING VECTOR DATA

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 WFS and the view service WMS (see chapter 3.2.3.2) is that WFS gives direct access to the geometry and attributes of a selected geospatial data, meaning that a user can work with a dataset provided by WFS. In brief, 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 data. Following the type of operations needed to manipulate 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 (Open Geospatial Consortium 2008) and the retrieved or created data are encoded in the Geographic Markup Language (Open Geospatial Consortium 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.

3.2.3 OTHER SERVICES (NOT TO BE IMPLEMENTED IN THE PROPOTYPE)

3.2.3.1 WEB COVERAGE SERVICE (WCS) FOR DOWNLOADING RASTER DATA

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

Like the WFS allows a client to access vector datasets, Web Coverage Service allows a client to access raster datasets. By raster 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 provides 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=GetCapabilities
```

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

To invoke the operation, users have only to define service and 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 (Figure 8). 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.

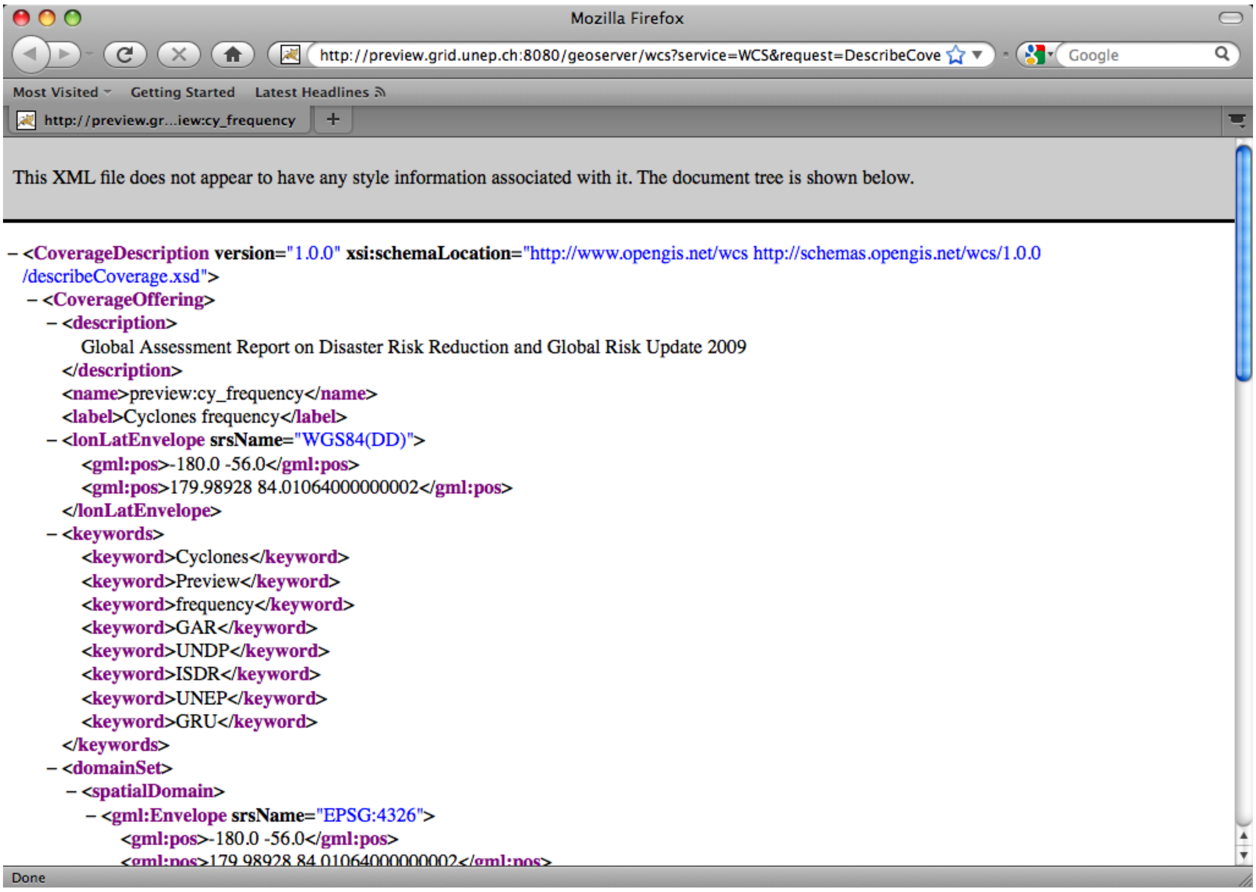


Figure 8. Result of a DescribeCoverage request.

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 is 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 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 users can download the coverage file. If the request is sent through a Desktop GIS client like ArcGIS then users gets the coverage directly into it.

3.2.3.2 VIEW SERVICES (MAPPING) USING OGC 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 (e.g., 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 kinds 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=GetCapabilities&version=1.3.0
```

The GetCapabilites operation (Figure 9) returns to users an XML document describing service and data sets available from which either desktop and/or web clients may request maps. This operation is common for all OWS and is presented in details in the OpenGIS Web Service Common Implementation Specification (Open Geospatial Consortium 2008). To invoke this operation, users have only to define service and 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/>
```


Figure 9. Example of the XML file returned after a GetCapabilites request

Example of a WMS URL with a GetMap 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 returns to a client request a map of selected geospatial layers (Figure 10).

In comparison of a GetCapabilities request that needs only two parameters, we can see on 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 is 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 or CRS (depending on the version of WMS spec): identifier of the Spatial Reference System.

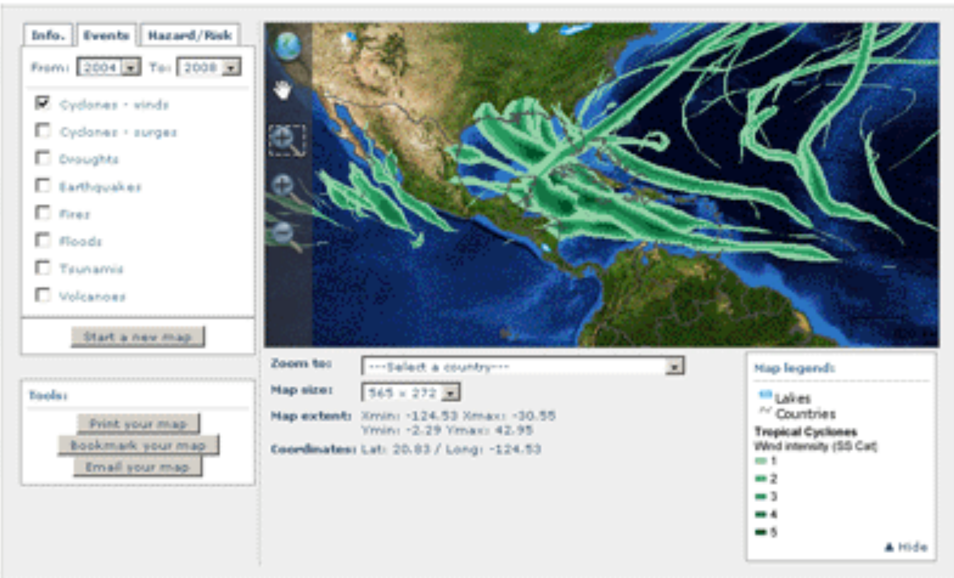


Figure 10. Returned image after a WMS request

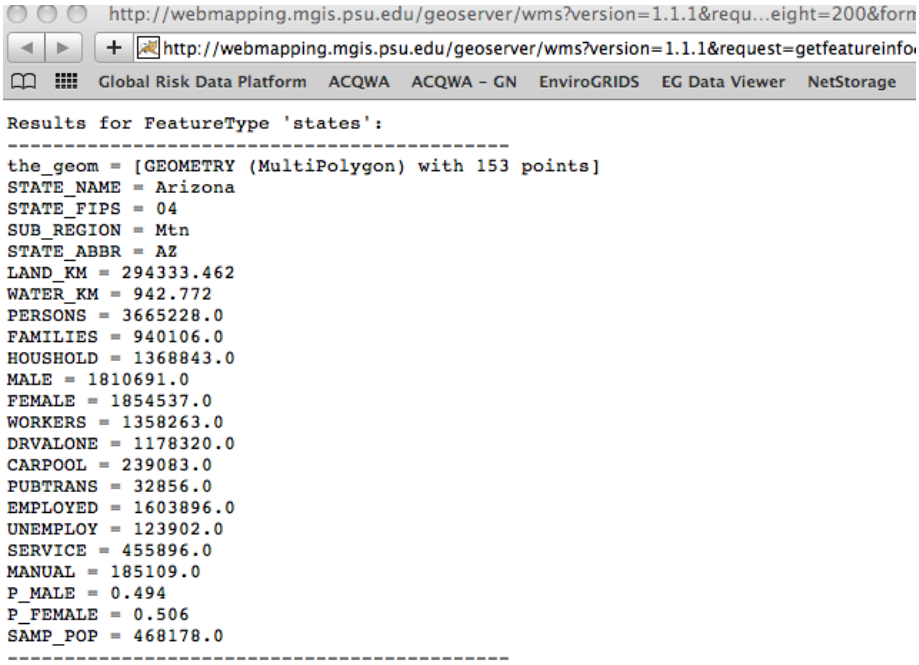
Example of a WMS URL with a GetFeatureInfo request:

```
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 (Figure 11). 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.
- SRS or CRS (depending on the version of WMS spec): 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
--------------------------------------------
```

Figure 11. Result of a GetFeatureInfo query

3.2.3.3 TRANSFORMATION SERVICES (PROCESSING) WITH OGC 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 (Open Geospatial Consortium 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 computation is completed.

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

- *GetCapabilites:* answer to a client describing its capabilities. It tells the client which kinds 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 service and processes available for execution.

To invoke the operation, users have only to define service and 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 mandatory, optional and default parameters needed for a selected process, as well as 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 required parameters (if needed) and references the datasets location. Once the process is completed, result is returned to the client as a new dataset.

3.3 CONFORMITY OF ESPON DATABASE WITH OGC STANDARDS

At first glance, as long as the ESPON data model is INSPIRE compliant and the tabular data are linked to geometries, there is no major compatibility problem with the OGC standards. Both INSPIRE and OGC standards are connected to the GML standard: INSPIRE has developed several GML application schemas (see for instance http://www.ogcnetwork.net/gmlprofiles) and "clients and servers with interfaces that implement the OGC Web Feature Service Interface Standard read and write GML data" (Bray & Ramage 2012).

A complete compliancy checking could be done by LIG in the next phase of the project (prototype implementation), e.g. by mapping the ESPON data model to a GML schema.

4 SWOT ANALYSIS OF WEB SERVICES FOR ESPON DB

This section presents in a synthetic manner the strengths, weaknesses, opportunities and threats (SWOTs) related to the web services previously described in this document. The SWOT characteristics of the web services listed in the below Table 2 and Table 3 are further explored in the next chapter on the steps towards the implementation of a web services prototype.

Services	Strengths	Weaknesses	Opportunities	Threats
CSW	Easy to expose existing metadata with CSW Interoperablity Enhance discoverability Expose to ESPON	Presence of two metadata tools: ESPON and possibly GeoNetwork	Metadata already compliant with INSPIRE Register the CSW interface into GEOSS (http://www.geoportal.org) Use of GeoNetwork additional functionalities http://geonetwork-opensource.org/functionality.html	Maintenance/confusion between the two catalogues Acceptation by the community Services may be temporarily or permanently down
WFS	Easy compliance to INSPIRE IR Interoperability Enhanced accessibility (in particular to updated data – if updated by service providers) Enhanced integration (data can be exposed in various formats and geographic projections) Possibility to download the data	Currently not widely adopted in the statistical community	Make statistical data available as tabular and geospatial data. Demonstrate the benefits of interoperability within the statistical community Link statistical data with other data sources. Use of GeoServer additional functionalities http://geoserver.org/display/GEOS/Features Register the WFS interface into GEOSS (http://www.geoportal.org)	Acceptation by the community Services may be temporarily or permanently down

Table 2. SWOT analysis of CSW and WFS

Services	**Strengths**	**Weaknesses**	**Opportunities**	**Threats**
WCS	Easy compliance to INSPIRE IR Interoperability Enhanced accessibility (in particular to updated data – if updated by service providers) Enhanced integration (data can be exposed in various formats and geographic projections) Possibility to download the data	Currently no data to expose as WCS Currently not widely adopted in the statistical community	Make statistical data available as tabular and geospatial data Demonstrate the benefits of interoperability within the statistical community Link statistical data with other data sources Use of GeoServer additional functionalities http://geoserver.org/display/GEOS/Features Register the WCS interface into GEOSS (http://www.geoportal.org)	Acceptation by the community Services may be temporarily or permanently down
WMS	Easy compliance to INSPIRE IR Interoperability Enhanced accessibility (in particular to updated data – if updated by service providers) Enhanced integration (data can be exposed in various formats and geographic projections)	Currently not widely adopted in the statistical community Data are rendered as images, standard GIS analysis is not feasible using WMS	Make statistical data available as tabular and geospatial data Demonstrate the benefits of interoperability within the statistical community Link statistical data with other data sources WMS can be implemented, if funding available Register the WMS interface into GEOSS (http://www.geoportal.org)	Acceptation by the community Services may be temporarily or permanently down
WPS	Interoperability	Currently not widely adopted in the statistical community	Making available calculations algorithms Demonstrate the benefits of interoperability within the statistical community WPS can be implemented, if funding available	Acceptation by the community Services may be temporarily or permanently down

Table 3. SWOT analysis of WCS, WMS and WPS

5 IMPLEMENTATION (PROOF OF CONCEPT)

5.1 TOWARDS A PROTOTYPE BY DECEMBER 2013

The objective is to expend the capabilities of the ESPON DB and expose its content as interoperable OGC & INSPIRE services. The main idea is to keep the existing ESPON DB as it is because a good community of users already uses it and it answers their needs. Therefore the aim is to provide an added value using OGC-compliant services allowing integrating ESPON DB statistical data with other data repositories (e.g., geospatial). Thanks to these services it will dramatically facilitate the integration and exchange of ESPON DB statistical data with other data sources. Consequently it will allow users to generate more easily new information and find relations between different data sets that were not foreseen before ("this system will be more than the sum of its part"). This will allow different types of community of users to benefit from the content of the ESPON Database. These data sets once published as OGC WFS (and eventually as WMS) services will make data directly available in different formats tabular (XLS, CSV, …) or geospatial (KML, Shapefile, GML, …). Consequently, this will allow to access instantaneously to the thematic, spatial, and temporal dimension of the ESPON DB. Moreover, using such services will facilitate also the dissemination/diffusion of data when updated. Indeed when the ESPON DB will be updated, then thanks to these services this update will be directly available to all users that are already connected to a service. Therefore it will ensure users to always access the most up-to-date data sets. Finally, the data service offered by ESPON DB can be consumed by view and/or processing services developed by the users themselves or by third-parties.

The prototype implementation is aiming at:

1. Exposing the **metadata** stored in the ESPON DB using a catalog system and OGC CSW interface.

- This will enable:
  - efficient discovery mechanism (spatial, temporal, thematical search)
  - Follow INSPIRE rules for discovery
  - Interoperate with other catalogue systems
  - Facilitate the exchange of metadata with other systems
  - Discover and expose ESPON DB metadata through the INSPIRE geoportal (http://inspire-geoportal.ec.europa.eu)
  - Register the CSW interface into the Global Earth Observation System of Systems (GEOSS, http://www.geoportal.org)

  Actions to be implemented:

  1. Export tool to export each metadata record of the ESPON DB as ISO19139 XML file.

  2. Batch import of the exported XML file in GeoNetwork.

  3. Setup the catalog and CSW service in GeoNetwork.

  4. Register the CSW interface in INSPIRE geoportal and GEOSS.

Figure 12 below shows the example of how the metadata of the UNEP's Environmental Data Explorer (a home-made PostgreSQL tool) is currently exported/imported into the Geonetwork application:

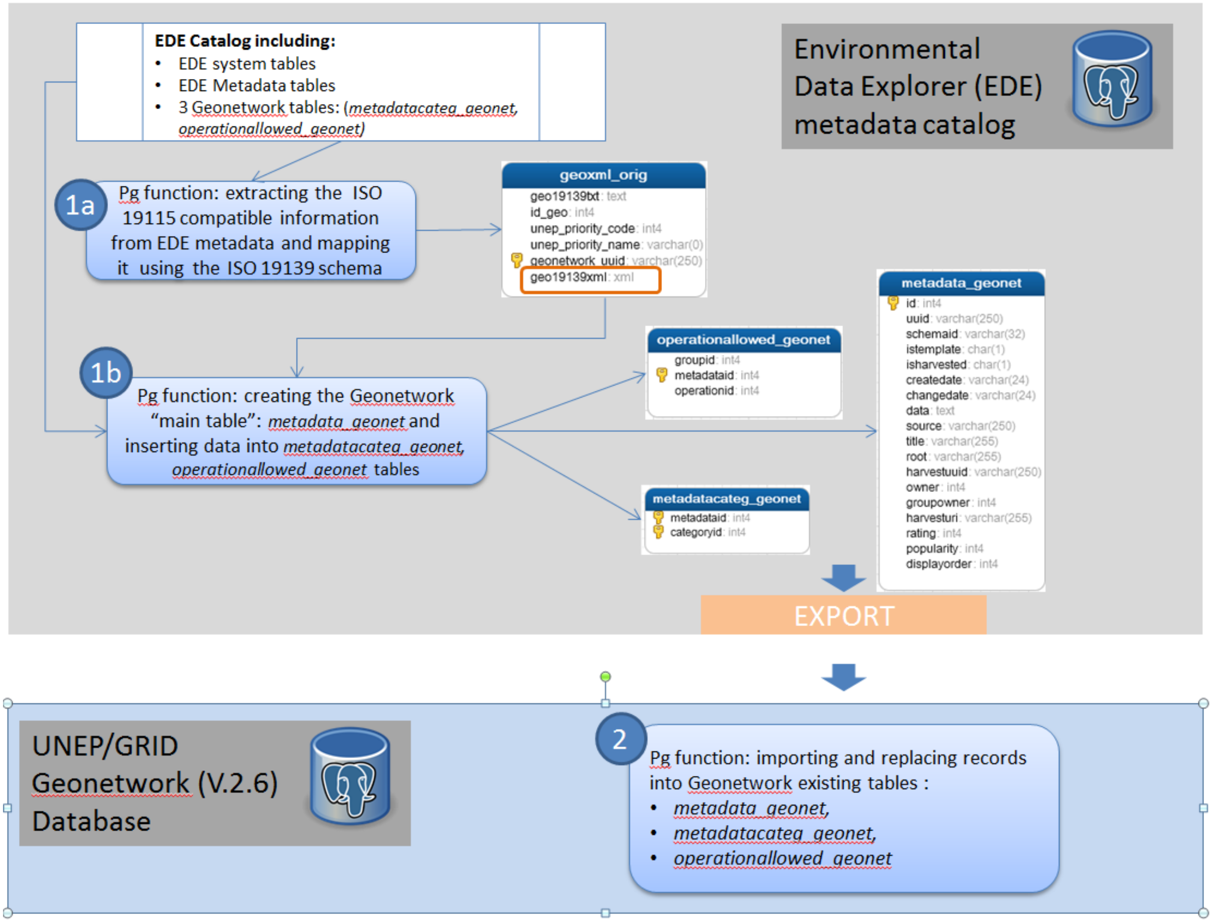


Figure 12. Transferring metadata to Geonetwork (the example of the UNEP/EDE)

Actions 1a and 1b: export metadata as ISO19139 XML file / Action 2 : import metadata into GeoNetwork

2. Exposing the **data** stored in the ESPON DB using OGC WFS services

- This will enable to:
  - Expose ESPON DB data as INSPIRE view and download services
  - Interoperate with other data repositories
  - Give access to data as tables, geospatial data at the same time (no need to change formats, etc.)
  - Facilitate the exchange and integration of data (on-the-fly reprojection, …)

Actions to be implemented:

1. Specify users requirements for download capabilities (formats, etc.).
2. Setup connection between the DBMS and GeoServer .
3. Add the required GeoServer extensions following the output formats needed (CSV, XLS, GeoJSON, KML, GML, etc.).
4. In GeoServer, publish ESPON DB directly as layers if they are stored in PostGIS or using the SQL Views functionality (http://docs.geoserver.org/stable/en/user/data/database/sqlview.html) for example by linking a country layers with ISO codes (geospatial layer) and ESPON DB data (statistical data). This is the proposed solution to make the ESPON DB statistical data also accessible as geospatial data.
5. Setup and configure the WFS and WMS services.

5.2 POSSIBLE FURTHER DEVELOPMENTS BY THE USERS

If desired, users can then develop themselves the following components:

1. Develop **processing** services and expose them as OGC WPS services

- This will enable to:
    - Develop processing algorithms based on ESPON DB data
    - Add interoperable calculations capabilities on top of the ESPON DB
    - Interoperate with other processing services (i.e., building complex analytical workflows based on service chains)
    - Provide both statistical (through R software) and geospatial (through GRASS software) calculations

    Actions to be implemented:

    1. Identify calculations that can be useful or required for users.
    2. Write WPS algorithms (Python code).
    3. Publish these processing services in PyWPS.

2. **Develop a web application** to facilitate the access to all these newly developed services.

- This will enable to:
    - Query and visualize data as tables or maps according to spatial, temporal, and thematic dimensions
    - Extract and download data as tables or geospatial data sets according to spatial, temporal, and thematic dimensions
    - Create graphs online
    - Compute complex calculations online

  An example of such a platform is the Global Risk Data Platform (http://preview.grid.unep.ch) and the Environmental Data Explorer (http://geodata.grid.unep.ch) that are developed and maintained by UNEP/GRID-Geneva.

5.3 SOFTWARE TOOLS

This section describes a number of open source software tools that efficiently implement the previously described web services. They have been successfully tested and used by the authors of the present report.

Catalogue: Geonetwork

GeoNetwork is a catalog application to manage spatially referenced resources. It provides powerful metadata editing and search functions as well as an embedded interactive web map viewer. It is currently used in numerous Spatial Data Infrastructure initiatives across the world.

GeoNetwork has been developed to connect spatial information communities and their data using a modern architecture, which is at the same time powerful and low cost, based on the principles of Free and Open Source Software (FOSS) and International and Open Standards for services and protocols (a.o. from ISO/TC211 and OGC).

The software provides an easy to use web interface to search geospatial data across multiple catalogs, combine distributed map services in the embedded map viewer, publish geospatial data using the online metadata editing tools and optionally the embedded GeoServer map server. Administrators have the option to manage user and group accounts, configure the server through web based and desktop utilities and schedule metadata harvesting from other catalogs.

GeoNetwork is an open source project, sponsored by the UNSDI (Henricksen 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 (Open Geospatial Consortium 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.

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

- Documentation: http://geonetwork-opensource.org/docs.html
- Download: http://geonetwork-opensource.org/downloads.html
- GeoNetwork INSPIREd capabilities: http://geonetwork-opensource.org/manuals/2.6.4/users/search.html?check_keywords=yes&area=default&q=inspire&x=0&y=0
- An excellent report on "INSPIRE support in GeoNetwork": http://www.neogeo-online.net/blog/wp-content/uploads/2011/01/201012_geonetwork_inspire.pdf

Data: GeoServer

GeoServer is an open source software server written in Java that allows users to share and edit geospatial data. Designed for interoperability, it publishes data from any major spatial data source using open standards.

Being a community-driven project, GeoServer is developed, tested, and supported by a diverse group of individuals and organizations from around the world.

GeoServer is the reference implementation of the Open Geospatial Consortium (OGC) Web Feature Service (WFS) and Web Coverage Service (WCS) standards, as well as a high performance certified compliant Web Map Service (WMS). GeoServer forms a core component of the Geospatial Web.

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.

Geoserver natively supports various output formats (GML, GeoJSON, JSON, CSV, Shapeflie) that are listed at:

http://docs.geoserver.org/stable/en/user/services/wfs/outputformats.html

Other formats can be available thanks to the OGR WFS extension (http://docs.geoserver.org/stable/en/user/extensions/ogr.html) and Excel output extension (http://docs.geoserver.org/stable/en/user/extensions/excel.html).

Regarding raster data published as WCS, GeoServer supports natively GeoTiff, GTopo30, ArcGrid (http://docs.geoserver.org/stable/en/user/services/wcs/outputformats.html) and can be extended with GDAL data formats (http://docs.geoserver.org/stable/en/user/data/raster/gdal.html).

Website: http://www.geoserver.org

- Documentation: http://docs.geoserver.org/
- Download: http://geoserver.org/display/GEOS/Download
- INSPIRE extension doc: http://docs.geoserver.org/stable/en/user/community/inspire/index.html
- INSPIRE extension download: http://files.opengeo.org/inspire/
- INSPIRE view service and Geoserver: http://blog.opengeo.org/2011/07/12/inspire-update/
- How to publish INSPIRE view service with GeoServer: http://location.defra.gov.uk/wp-content/uploads/2011/07/Data-Publisher-How-To-Guide-Understand-the-background-to-establishing-an-INSPIRE-View-Service-using-GeoServer.pdf

Processing: PyWPS & GeoServer

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
- SOAP/WSDL

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

It should be noted that GeoServer offers also a set of geoprocessing algorithms exposed as WPS (http://docs.geoserver.org/stable/en/user/extensions/wps/index.html).

5.4 ARCHITECTURE

For security and ease of maintenance reasons, it is recommended to deploy the web services (possibly using GeoNetwork , GeoServer, …) on a dedicated server. This would clearly separate the service layer from the data and the application layers, as shown on Figure 13 below :

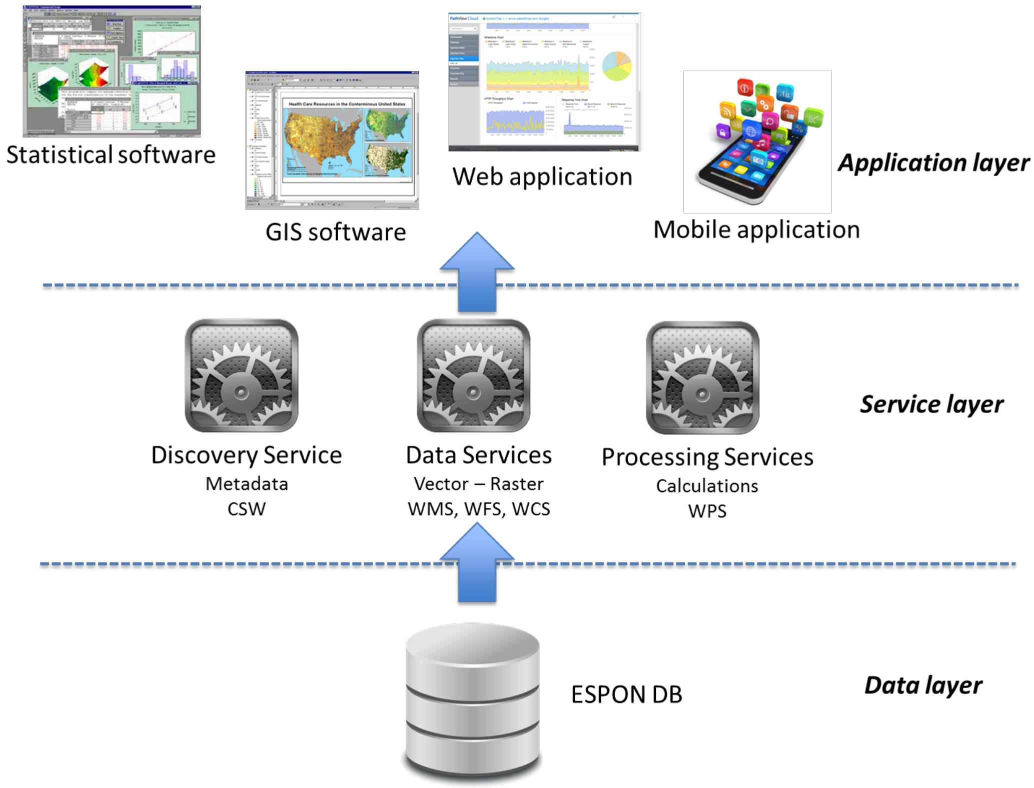


Figure 13. Separation of the application, service and data layers

6 CONCLUSION

The present report shows that the development of web services for searching and downloading the data contained in the ESPON Database is feasible.

The current architecture and the data models of the ESPON Database are suitable for the design and implementation of such services, which can be deployed using available open source software tools.

These web services would add an increased accessibility to the ESPON Database by third-party applications/machines. In particular, these services could be consumed by OGC compliant GIS software and web applications used by other ESPON projects. For instance RIMAP (ESPON Online Mapping Tool) and ETMS (EU Territorial Monitoring System) could take advantage of these web services for their mapping tools. In this respect, these web services perfectly complement the human accessibility already provided by the ESPON Database web application (see Figure 5).

A prototype of web services for data search and download can be achieved by December 2013. Further working time and resources are needed if developments of mapping and processing services by ESPON M4D are envisaged.

7 REFERENCES

7.1 TEXTS

Bernard, L. and Craglia, M. (2005). SDI - From Spatial Data Infrastructure to Service Driven Infrastructure: 8.

Bray, A. and Ramage, S. (2012). Market Report – Open Standards and INSPIRE, OGC, http://www.opengeospatial.org/pressroom/marketreport/inspire.

Booz, Allen, et al. (2005). Geospatial Interoperability Return on Investment, NASA: 47.

Cömert, C. (2004). Web services and National Spatial Data Infrastructure (NSDI). XXth ISPRS Congree. Istanbul: 6.

European Commission (2006). Assessing the impacts of Spatial Data Infrastrucutres. J. R. Center. Ispra: 1-61.

Ezigbalike, C. (2004). SDI-Africa: An implementation guide, UN Economic Commission for Africa.

Fielding, R. T. (2000). Architectural Styles and the Design of Network-based Software Architectures. Doctoral dissertation. University of California, Irvine.

Kaufmann, J. and J. Dorfschmid (2001). Reflections on the benefits and potential economies of geographic data standards: 25.

Mohammadi, Hossein, et al. (2010). Development of an interoperable tool to facilitate spatial data integration in the context of SDI, International Journal of Geographical Information Science : 24/4, 487-505.

Nebert, D. D. (2005). Developing Spatial Data Infrastructure: The SDI Cookbook, GSDI (current on-line version available at http://www.gsdidocs.org/GSDIWiki/).

Open Geospatial Consortium (2004). The Havoc of Non-Interoperability: 7.

Open Geospatial Consortium (2005). Server Architecture Models for the National Spatial Data Infrastructures (NSDI). OGC White Paper: 41.

Open Geospatial Consortium (2007). OpenGIS Geography Markup Language (GML) Encoding Standard, OGC: 1-437.

Open Geospatial Consortium (2007). OpenGIS Web Processing Service: 87.

Open Geospatial Consortium (2008). OGC Reference Model: 35.

Ostensen, O. (2001). Global geomatics standards supporting sustainable geospatial data infrastructures. GSDI-5. Cartagena De Indias, Colombia: 6.

Sahin, K. and M. U. Gumusay (2008). Service oriented architecture (SOA) based web services for geographic information systems. XXIst ISPRS Congress. Beijing: 625-630.

7.2 WEB LINKS

7.2.1 M4D

ESPON DB2 web application https://espondb.liglab.fr/e2/

7.2.2 UNEP/GRID-GENÈVE

PREVIEW portal http://preview.grid.unep.ch

Preview web services http://preview.grid.unep.ch/index.php?preview=services&lang=eng

EDE portal http://ede.grid.unep.ch

EDE web services http://ede.grid.unep.ch/webservices/

EnviroGRIDS FP7 project http://envirogrids.grid.unep.ch/

Envirogrids geoportal http://www.envirogrids.cz/

ANNEX 1. THE INSPIRE SERVICES

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

INSPIRE Geoportal: http://www.inspire-geoportal.eu/ INSPIRE Forum: http://inspire-forum.jrc.ec.europa.eu/

The Infrastructure for Spatial Information in the European Community, namely INSPIRE, is of particular interest for the ESPON 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. 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. 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:

- Inconsistencies in collection of geospatial data: geospatial data are often missing and/or incomplete, or are collected twice by different organizations.
- Lacking of documentation, description (metadata) of the data.
- Geospatial data are often incompatible and thus cannot be combined.
- Infrastructures used to find, access and use geospatial data 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"*.

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:

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.

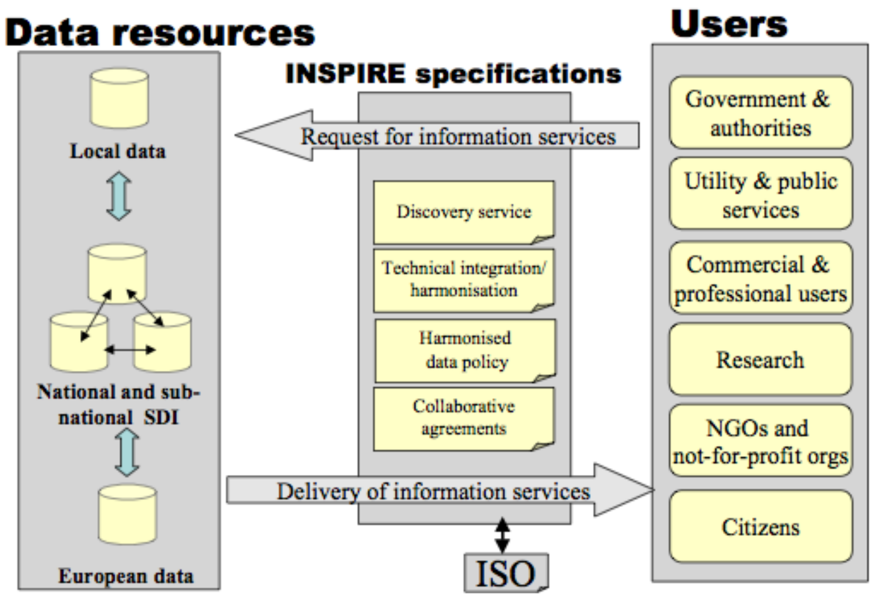
A step-by-step approach is used to implement and develop the infrastructure because such an initiative cannot be built from one day to another and is asking Member States to drastically change their existing infrastructure. Thus the implementation of services has been stated just after the adoption of 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.

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. Thus users will spend less time and efforts to integrate data delivered within the INSPIRE framework.



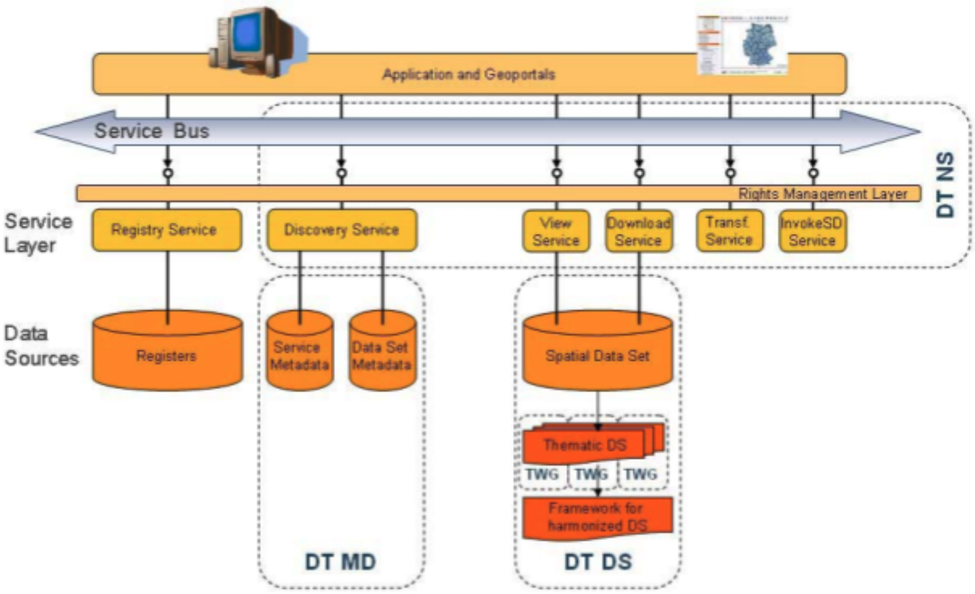
Data and information flow within the INSPIRE framework (Source: INSPIRE)

The Directive 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. This geoportal will not store or maintain data and metadata. Instead, 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, Member States shall establish, operate and provide access to the following network services:

- *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.



INSPIRE network architecture (Source: INSPIRE)

The INSPIRE Directive addresses 34 spatial data themes needed for environmental applications. These themes are subdivided in the three annexes of the directive.

Annex I	**Annex III**
1 Coordinate reference systems	1 Statistical units
2 Geographical grid systems	2 Buildings
3 Geographical names	3 Soil
4 Administrative units	4 Land use
5 Addresses	5 Human health and safety
6 Cadastral parcels	6 Utility and governmental services
7 Transport networks	7 Environmental monitoring Facilities
8 Hydrography	8 Production and industrial facilities
9 Protected sites	9 Agricultural and aquaculture facilities
Annex II	10 Population distribution and demography
1 Elevation	11 Area management/restriction/regulation zones & reporting units
2 Land cover	12 Natural risk zones
3 Orthoimagery	13 Atmospheric conditions
4 Geology	14 Meteorological geographical features
	15 Oceanographic geographical features
	16 Sea regions
	17 Bio-geographical regions
	18 Habitats and biotopes
	19 Species distribution
	20 Energy Resources
	21 Mineral Resources

ANNEX 2. INSPIRE IMPLEMENTING RULES (LINKS)

Implementing Rules: http://inspire.jrc.ec.europa.eu/index.cfm/pageid/47

Metadata: http://inspire.jrc.ec.europa.eu/index.cfm/pageid/101

Data Specifications: http://inspire.jrc.ec.europa.eu/index.cfm/pageid/2

Network Services: http://inspire.jrc.ec.europa.eu/index.cfm/pageid/5

Data and Service Sharing: http://inspire.jrc.ec.europa.eu/index.cfm/pageid/62

Monitoring and Reporting: http://inspire.jrc.ec.europa.eu/index.cfm/pageid/182

ANNEX 3 : INSPIRE SERVICES REQUIREMENTS

3.1 Discovery services

- must support:
    - ISO19115/19139
    - ISO19119
    - additional fields (Resource title, Resource abstract, Resource type, Unique Resource Identifier, Temporal reference)

Minimum search criteria	INSPIRE metadata elements
keywords	Keyword
classification of spatial data and services; (For spatial data sets and spatial data set series)	Topic category
classification of spatial data and services (For spatial data services)	Spatial data service type
the quality and validity of spatial data sets	Lineage
the quality and validity of spatial data sets	Spatial resolution
degree of conformity with the implementing rules provided for in Article 7(1) of Directive 2007/2/EC	Specification
degree of conformity with the implementing rules provided for in Article 7(1) of Directive 2007/2/EC	Degree
geographical location	Geographic bounding box
conditions applying to the access to and use of spatial data sets and services	Conditions applying to access and use
conditions applying to the access to and use of spatial data sets and services	Limitations on public access
the public authorities responsible for the establishment, management, maintenance and distribution of spatial data sets and services	Responsible party
the public authorities responsible for the establishment, management, maintenance and distribution of spatial data sets and services	Responsible party role

Minimum search criteria and corresponding metadata elements

Operation	Role
Get Discovery Service Metadata	Provides all necessary information about the metadata discovery service and describes service capabilities.
Discover Metadata	The Discover Metadata operation allows requesting INSPIRE metadata elements of resources based on a query statement to be retrieved from the target Discovery Service.
Publish Metadata	The Publish Metadata operation allows to edit INSPIRE metadata elements of resources in the Discovery Service (push or pull metadata mechanisms). Editing meaning insert, update and delete
Link Discovery Service	The Link Discovery Service function allows the declaration of the availability of a Discovery Service for the Discovery of resources through the Member State Discovery Service while maintaining the resource metadata at the owner location

Operations

3.2 View services

- must support:
  - WMS 1.3.0,
  - PNG & GIF formats,
  - ETRS89 projection,
  - SLD/SE for styling.

Metadata element	Description
Resource Title	The title of the layer, used for human communication, for presentation of the layer e.g. in a menu
Resource Abstract	Layer abstract
Keyword	Additional Keywords
Geographic Bounding Box	The minimum bounding rectangle in all supported Coordinate Reference Systems of the area covered by the Layer.
Unique Resource Identifier	The Unique Resource Identifier of the resource used to create the layer.
Name	Harmonized name of the layer
Coordinate Reference Systems	List of Coordinate Reference Systems in which the layer is available.
Styles	List of the rendering styles available for the layer. A Style shall be composed of a title and a Unique Identifier.
Legend URL	Location of the legend for each style, language and dimension pairs.
Dimension Pairs	Indicates the supported two dimensional axis pairs for multi dimensional spatial data sets and spatial data sets series

Metadata elements for View Services

Operation	Role
Get View Service Metadata	Provides all necessary information about the service and describes service capabilities.
Get Map	Returns a map containing the geographic and thematic information coming from the available spatial datasets. This map is an image spatially referenced.
Link View Service	Allows a Public Authority or a Third Party to declare a view Service for the viewing of its resources through the Member State View Service while maintaining the viewing capability at the Public Authority or the Third party location.

Operations

3.3 Download services

- Must support :
  - WFS 2.0 for vector
  - WCS 2.0 for raster
  - ETRS89 for projections

Function	Description	non-direct access	direct access
Get Download Service Metadata	Provides metadata about the service and data sets offered by the service to a user and describes service capabilities. Shall at least contain the INSPIRE metadata elements defined for spatial data services as described by the Metadata Implementing Rule.	M	M
Get Spatial Objects	The Get Spatial Objects operation allows spatial object instances to be retrieved. In the case of non-direct access, the operation will retrieve a predefined data set or a pre-defined part of a data set. In the case of download service of a pre-defined data set or predefined part of data set, the operation shall return spatial objects in at least one of the CRSs defined by the Implementing Rule the Annex 1 theme coordinate reference systems. In the case of direct access, the retrieval can be based on an optional query defined by the Define Query operation. The operation shall support user requested CRS belonging to the INSPIRE defined CRSs.	M	M
Describe Spatial Object Types	The Describe Spatial object Type operation generates a description of the spatial object types that the service offers. In the case of download service of a pre-defined data set or predefined part of data set, the function shall return the description of the complete set of spatial object types contained in the data set or part of data set. In the case of a direct access download service, the function can have as parameter a set of named spatial object types for which the description is requested.	O	M
Define Query	Defines a query to be used in the Get Spatial Objects operation. This function is applicable only in the case of direct access download service. The predicates shall express selection criteria based upon the model of the data sets as defined by an INSPIRE Implementing Rule on the interoperability of spatial data sets and services. The capability to define a query is mandatory, but a query can be omitted in a concrete Get Spatial Objects request.	n.a.	M
Link Download Service	Allows the declaration of a Download Service for downloading of its resources through the Member State Download Service while maintaining the downloading capability at the Public Authority or the Third party location.	M	M

M – mandatory; O – optional; n.a. – not applicable

Operations