# Supporting Interoperable Geospatial Data Fusion by adopting OGC and ISO TC 211 standards

Simone Giannecchini, Francesco Spina
NURC, NATO Undersea Research Centre
La Spezia, Italy.
giannecchini,spina@nurc.nato.int

Bryce Nordgren
USDA Forest Service, Fire Science Lab
Missoula, MT, U.S.A.
bnordgren@fs.fed.us

Martin Desruisseaux
Institut de Recherche pour le Developpement / US-Espace
Nouma, Nouvelle-Caldonie.
martin.desruisseaux@noumea.ird.nc

***Abstract - Geospatial data fusion refers to the ability to process (*fuse*) data from a variety of sources which capture and/or model earth-related phenomena in order to produce added-value information. This paper provides an overview of the latest advancements in standardization for interoperability in the geographical information community achieved by the International Organization for Standardization (ISO) and the Open Geospatial Consortium (OGC). It also provides guidelines and suggestions for designing superior architectures to support geospatial data fusion by employing OGC/ISO specifications.***

**Keywords**: *geospatial data fusion, interoperability, ogc/iso specifications, geospatial services, GIS.*

# 1 Introduction

## 1.1 Interoperable Geospatial Data Fusion

*Geospatial data fusion* refers to the ability to process (*fuse*) data from a variety of sources which capture and/or model earth-related phenomena in order to produce added-value information. It relies on a robust, flexible and interoperable architecture to handle all the issues related to data and metadata management and dissemination in order to focus strictly on the *fusion* process itself.

Let us provide a definition of *interoperability*. Quoting from[1],

**Definition 1** ***Interoperability*** *refers to the capability to communicate, execute programs or transfer data among different functional units in a manner which requires the user to have little or no knowledge of the unique characteristics of those units. The outcome is the seamless exchange of information and procedures between different organizations employing public, well-known standard protocols and interface*[1] *connections.*

*Interoperability* is a concept which focuses on the exchange of information, the protocols ruling such exchange and the interfaces that connect modules providing different capabilities. Common misconceptions include the assumption that being interoperable means converting as many data sources as possible to a proprietary format or that it means having a client and a server developed by the same company, i.e. using a proprietary protocol, running over a network link. Hence this paper will utilize definition 1.

A standard unquestionably inherits its credibility from the proposing entity, therefore it is worthwhile to briefly introduce the two leading organizations which are issuing standards in the geospatial community: ISO Technical Committee 211[2] (TC 211) and Open Geospatial Consortium Inc (OGC)[3].

The OGC describes itself as *"a non-profit, international, voluntary consensus standards organization that is leading the development of standards for geospatial and location based services*[4]*"*.

The goal of this consortium, which comprises most of the relevant private companies and public organizations within the geospatial community, is to gather, standardize, formalize into specifications and then publish state-of-the-art guidelines and application programming interfaces for designing and building interoperable geospatial services. The consortium also organizes and promotes interoperability projects which test and validate specifications.

The scope of TC 211 is *"standardization in the field of digital geographic information, with the aim of establishing a structured set of standards for information concerning objects or phenomena that are directly or indirectly associated with a location relative to the Earth"*.

TC211 and OGC work in conjunction in order to avoid duplication of efforts. This will surely add impetus to a wider adoption of geographic information specifications.

[1] An *interface* is a well-recognizable set of operations characterizing the behaviour of an entity.

[2] *http://www.isotc211.org/*

[3] *http://www.opengeospatial.org/*

[4] A *service* is a distinct part of the functionality that is provided by an entity through interfaces.

## 1.2 Foundation Concepts

It is common practice in the geographic information community to separate geospatial data into two main categories, *raster* data and *vectorial* (or *vector*) data, depending on their nature. Referring to [1][2],

**Definition 2** ***Raster Data*** *deals with real world phenomena that vary continuously over space. It contains a set of values, each associated with one of the elements in a regular array of points or cells. It is usually associated with a method for interpolating values at spatial positions between the points or within the cells.*

**Definition 3** ***Vector Data*** *deals with discrete observable and/or measurable world phenomena whose spatial characteristics responsible for its geolocation are represented by a set of one or more geometric primitives (points, curves, surfaces, or solids). Other characteristics of the phenomenon are recorded as alphanumeric or "thematic" attributes.*

The concepts of discrete and continuous phenomena are not mutually exclusive, hence their representations as vector or raster are not mutually exclusive. The same phenomenon may be represented as either a vector or a raster.

A lake may be viewed as a vector data set returning a single value for each attribute, such as its name, area, estimated age, number of emissaries, etc. The lake may also be represented as a raster data set providing different values for water quality, temperature, visibility for each (x,y,z) position at different time instants.

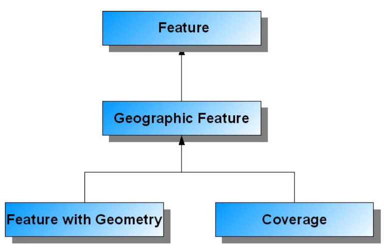


Figure 1: Feature hierarchy.

The above introduced concepts were put through a standardization process by the OGC and became the foundation of an interoperable geospatial information model. The concepts of *feature*, *geographic feature*, *feature with geometry* and *coverage* were introduced. Referring to figure (1) and to [1][2][3][4][5][6][7], we will now provide definitions for such concepts following OGC and ISO standards.

**Definition 4** *A* ***Feature*** *can represent an abstraction of a real world phenomena which is somehow observable and/or measurable. It can also describe an abstract entity.*

**Definition 5** *A* ***Geographic Feature*** *is a feature which describes a phenomenon related to a location on the earth, which means that it carries geolocation information along with it. It is the fundamental unit of geospatial information.*

**Definition 6** *A* ***Feature with Geometry*** *is a geographic feature which has one or more properties having values which are geometries.*

*Features with geometry* usually represent discrete observable phenomena, therefore they can be considered the same as vectorial data. It is common practice, when there is no ambiguity, to refer to them simply as *features*.

A feature may have a number of *properties* which may be operations, associations and attributes. *Attributes*, may be spatial or temporal coordinates, an identifier for a location (e.g. place name) and/or "thematic". Features have both *instances*, which indicate manifest real world or virtual observed phenomena, as well as *feature-types*, which represent the schema of a class of feature instances with common characteristics.

**Definition 7** ***Coverage*** *is a subtype of geographic feature that deals with phenomena which vary continuously over space and time. It can be generally defined as a function mapping from a multidimensional spatiotemporal domain to an heterogeneous multidimensional range*[5].

It is straightforward to point out that *coverage* can be employed to represent raster data[6]. It is a mapping function which associates a point in a spatiotemporal space $(x, y, z, t)$ with a user-defined set of parameters (e.g., temperature, pressure, owner, address, cost, plot name, etc.).

Before proceeding we would like to introduce the important concept of *geospatial metadata*.

**Definition 8** ***Geospatial metadata*** *is a special type of data, which characterizes geospatial data and services in terms of content, quality and condition in order to allow a more efficient exploitation of the described object.*

*Metadata* is vital for organizing, locating and maintaining archives of geospatial data since it provides a means for describing datasets and services.

It is common practice to separate *metadata* into three categories (see [8]), *Discovery Metadata*[7] , *Exploration Metadata*[8] and *Exploitation Metadata*[9].

The above mentioned categories depict different usage phases of a metadata repository. Discovery metadata enables us to find data sources matching our requirements using a minimum amount of metadata. After *discovering* possible data sources, we need more detailed metadata to determine if these sources meet our specific requirements, for instance, to solve a problem or to run an analysis. Finally, we need as much metadata as possible about a selected data source to

[5] A *range* is a record of values of heterogeneous nature both in terms of data types as well as in term of unit of measure.

[6] A *coverage* can represent more than just raster data.

[7] They provide the minimum amount of information to convey to the inquirer the nature and content of the data resource.

[8] They provide sufficient information to enable the inquirer to ascertain that data are applicable for a given purpose.

[9] They provide properties required to access, transfer, load, interpret and exploit the searched data in the final application.

proficiently apply, transfer or manipulate the geospatial data we have found.

First, we discover all the possible sources matching our requirements by using a minimum amount of metadata. Second, we refine our search to ensure that we find what we really need, by using more detailed metadata. Last, we need as much metadata as possible to efficiently utilize the geospatial data we are looking for.

# 2 Support for stand-alone geospatial data fusion

**Definition 9** *With the term* ***stand-alone geospatial data fusion*** *we refer to the geospatial data fusion which is performed mainly on a single machine without reliance on external capabilities.*

Supporting stand-alone geospatial data fusion, from our perspective, means introducing and exploiting *interoperability* at the *data management* or *toolkit level*. The goal is to seamlessly exchange not only data, but also *operations*[10].

The experience of one of the authors illustrates this concept.

Working in an oceanographic research institute the author found himself in a situation where he had to manage data from various organizations in a wide range of proprietary formats. His first approach was to convert the data into a single commonly accepted format but this quickly became cumbersome and inconvenient. Every new data source or every update of the old ones required building a conversion routine (usually with a *loss of information*). After much experimentation this author adopted an alternate approach: leaving the data sources in the original format, and producing tools capable of abstracting from the underlying format. He concluded that *interoperability* does not aim to convert all data to a common format but rather it aims to understand the data sources as they stand.

In addition, the author was often required to employ several toolkits to allow scientists and developers to perform complex operations. It was common to encounter situations where data produced by one toolkit was required as input by another. Connecting software developed by different vendors is a demanding task that is not always possible without adopting well-accepted standards, as *interoperability* suggests.

## 2.1 Reference Model

We have previously introduced and discussed the rationale behind the adoption of *interoperability* for supporting stand-alone data-fusion. We will now propose and describe a reference model for *data and metadata management*, based on OGC and ISO standards.

[10] An *operation* is the specification of a transformation or query that an object may be asked to execute. It usually has a name and a list of input and, where applicable, output parameters.

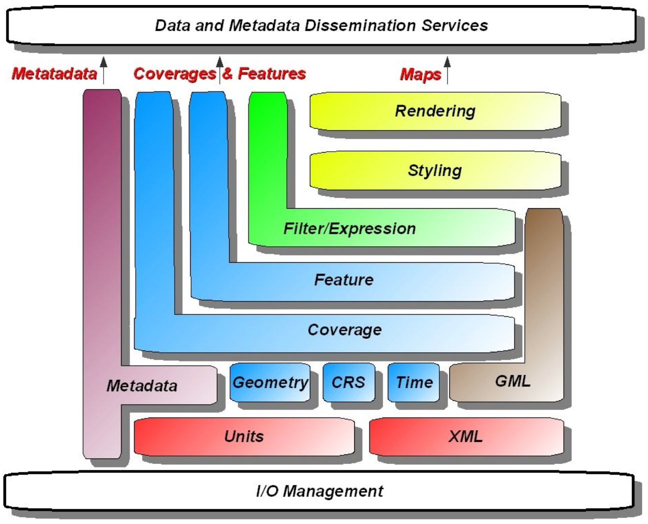


Figure 2: Reference model for supporting stand-alone data fusion.

Referring to figure (2), we can point out the existence of several components which can be grouped together in order to form various homogeneous functional units.

The lowest level of the diagram comprises of two red blocks, namely XML[9] and *Units*, which represent the foundations upon which the reference model is built. We will solely focus on the latter block.

The ability to coherently associate units of measure to quantities is crucial for data fusion purposes. Several specifications aiming at standardizing such topic exist; here we would like to report [10][11] where base SI[11] and non-SI unit of measures are dissected. An XML schema for their encoding as well as extensive support for conversions is provided.

Above the XML and Units level are three different groups of components: Metadata, GML[11] and what we call the *Data Model* (which is comprised of the light cyan modules). These groups represent the core of the reference model since they are responsible for ensuring a sufficient level of *abstraction* from underlying data sources.

*GML*, which stands for Geographic Mark-up Language, is an XML grammar for the modeling, transport, and storage of geographic information. It is primarly a tool for describing stored data (features or coverage) but it can also be used to describe other components, such as coordinate reference systems, coverage and features.

The *Data Model* finds its roots in the *Geometry*, *Time* and in the *Coordinate Reference System (CRS)* modules. The *Geometry* module, based on [3][7], provides standardized support for geographically referenced geometric and topological objects, whereby the term "geometry" we usually refer to set of direct positions[12] in a particular coordinate reference system. This module provides the basic means to describe geolocation of other objects.

[11] International System of Units (SI).

[12] Position described by a single set of coordinates.

The *Time* module, based on [12], provides standardized support for describing temporal characteristics of geographic features and metadata. Specifically, such characteristics can be employed both as thematic feature attributes as well as to provide the independent function for describing variation of earth-related phenomena (i.e. *geographic feature*) over time.

The *Coordinate Reference System (CRS)* module, based on [13], provides a standard method for describing spatial or spatiotemporal coordinate reference systems used to geolocate geospatial datasets.

On top of the *Geometry* and the *CRS* building blocks take place the modules to handle *coverage* and *features*, the foundations of the *data model* for providing *data abstraction* from originating data sources (recall section 1.2).

Where data access is concerned, we want to point out the presence of the *Filter/Expression* module implementing the "Filter Encoding Specification"[14] from OGC. This specification provides a standard way for building complex queries based on GML filters and expressions, both geospatial and non geospatial, which are semantically equivalent to an SQL WHERE statement. This module is crucial for refining queries in order to reduce the amount of data retrieved when requesting subparts of large data sets.

An example of a numeric filter is shown below here.

```
<Filter xmlns:gml="http://www.opengis.net/gml">
<PropertyIsBetween>
<PropertyName>PERSONS</PropertyName>
<LowerBoundary>
<Literal>2000000</Literal>
</LowerBoundary>
<UpperBoundary>
<Literal>4000000</Literal>
</UpperBoundary>
</PropertyIsBetween>
</Filter>
```

The last two important building blocks depicted are the *Rendering* and *Styling* blocks which are responsible for rendering *coverage* and *features* using custom styles[13] to produce georeferenced maps.

These blocks form the *Styled Rendering Model*. Specifically the *Styling* module implements the "Styled Layer Description Specification"[15] which provides a standard way of describing XML-based styles for the rendering phase (see section 2.1.3).

In the following sections we will detail further some of the concepts introduced above.

### 2.1.1 Internal Common Data Format

In section 1.2 as well as in section 2.1 we introduced the *Data Model* which extracts well-defined spatial, temporal and relational meaning from underlying data sources accomplishing superior *data abstraction*. This goal is achieved by using the concepts of *feature* and *coverage*, which shield the higher tiers from data management related issues. This not only enhances *interoperability* but it also greatly improves architectural *robustness* since the design of the higher tiers does not have to deal with data source related issues, like format transcoding. It also enforces net layer separation (*multi-tiers architecture*) and allows for the design of *independent services*.

The outcome of such an approach is the introduction of what we call *Internal Common Data Format*. Regardless of the format of the geospatial data sources we encounter, we can always treat them as if they were either *features* or *coverage*. A *Data Management Layer* exploiting this abstraction is able to present to the higher tiers of the architecture, responsible for data and metadata dissemination, only two different logical data types, *coverage* and *feature*. This enhances the overall robustness and provides the foundations for further extensibility (see figure (2)).

### 2.1.2 Coordinate Reference System Management.

*Coordinate Reference System (CRS) Management* refers to the capability to precisely refer geospatial locations to the earth and to convert such locations to different coordinate reference systems. Strong support for such capability is provided in [13].

**Definition 10** *A* ***Coordinate Reference System*** *is a Coordinate System which is related to the earth by a datum. A CRS is a combination of a Coordinate System and a Datum.*

**Definition 11** *A* ***Coordinate System*** *is composed by a set of coordinate axes with known metric.*

**Definition 12** *A* ***Datum*** *specifies the relationship of a coordinate system to the earth, creating a CRS.*

Several types of datums exist which can be used to define different types of CRS. Examples of coordinate reference systems are: Geocentric, Geographic, Projected, Vertical, Temporal, Engineering and Image.

Moreover specification [13] defines other important concepts which allows for the creation of new instances of coordinate reference systems besides the basic types listed above. By mean of these capabilities, geographic data referring to different coordinate reference systems can be merged together for coherent data fusion.

A *Coordinate Conversion* is a change of coordinates from a CRS to another one based on the same datum (an example is a *map projection*[14]). A *Coordinate Transformation* is a mathematical transformation between two CRS which requires a change of datum (an example is the orthorectification process which maps from an Image CRS to a Geographic CRS.).

### 2.1.3 Styled Rendering

Rendering capabilities are a central requirement of geospatial data fusion. The OGC has standardized

[13] A *Style* is a set of rules defining either how to render topological objects as lines and polygons (features) or how to apply palettes to numerical ranges (coverage).

[14] *Coordinate conversions* from a geodetic coordinate reference system to a planar surface.

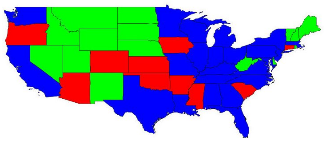

Figure 3: Styled rendering example.

them through specification [15] which introduces and describes *SLD*, a map-styling language based on XML which allows the user to state how to render certain topological objects, for *features*, or numerical ranges, for *coverage*, in a standardized way. SLD specification heavily relies on the "Filter Implementation Specification"[14] in order to provide the user with the ability to build complex rules for including or excluding collections of features from the rendering and for differentiating drawing options (filling color, line color, line thickness, etc.) depending on the feature's thematic alphanumeric attributes values or on the feature's geographical positions.

In figure (3) we can observe a rendering of the USA produced by employing a style where the color of each state depends on its population. An excerpt of the style used is shown below.

```
<Rule>
<Filter xmlns:gml="http://www.opengis.net/gml">
<PropertyIsGreaterThan>
<PropertyName>PERSONS</PropertyName>
<Literal>4000000</Literal>
</PropertyIsGreaterThan>
</Filter>
<PolygonSymbolizer>
<Fill>
<CssParameter name="fill">#0000FF</CssParameter>
</Fill>
</PolygonSymbolizer>
</Rule>
```

### 2.2 Final Considerations

In the previous section we briefly discussed various aspects regarding support for *stand-alone geospatial data fusion* in the OGC and ISO context. Standard support for custom rendering and coordinate reference system management was mentioned. Provision of an adequate level of *data abstraction* was introduced and validated.

As a final annotation we would like to point out that OGC and ISO specifications also provide application programming interfaces. This fact allows us not only to interoperate with anyone who implements the same interfaces, achieving the above mentioned *interoperability* at the *toolkit level*, but also to check whether the specifications are inadequate in order to propose new solutions and improvements. For further information refer to the GeoAPI project[15].

[15] *http://geoapi.sourceforge.net.*

## 3 Support for distributed geospatial data fusion

**Definition 13** ***Distributed geospatial data fusion*** *refers to the geospatial data fusion performed in a distributed, network-centric manner exploiting standard service and communication protocols in order to exchange data and information as well as commands for the execution of remote procedures.*

From our perspective such an architecture can be viewed as *client-server architecture* where the server applications are responsible for data and metadata management and dissemination (i.e. ingestion, storage and retrieval) and client applications are responsible for performing real data fusion.

In figure (4) the core multi-tier architecture for supporting distributed geospatial data fusion is depicted. In this section we aim to describe the *Data and Metadata Disseminations Tiers* which are responsible for the dissemination of geospatial data as well as the associated metadata. Based on concepts and considerations stated in the previous section it is assumed that we are working upon an underlying *Data Management Tier* which provides us with the needed level of *data abstraction*. Hence we need to deal only with the concepts of *coverage*, *feature* and *metadata*, regardless of the nature of the originating data sources.

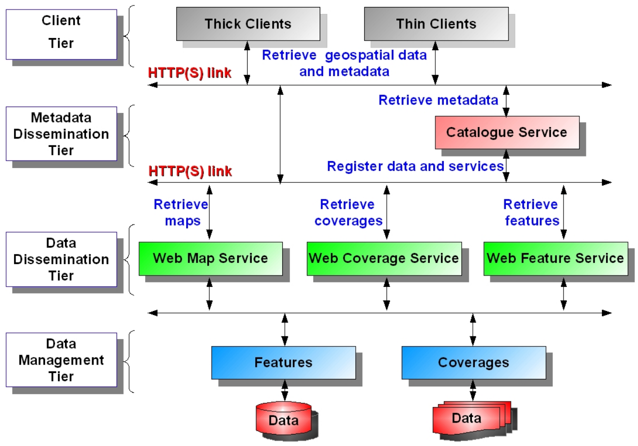


Figure 4: Core geospatial services for supporting distributed geospatial data fusion.

Referring to figure (4), we note that interactions between clients and servers take place on an HTTP or HTTPS link, by employing XML based protocols (as described below) in order to leverage technologies and protocols which are freely available and widely supported. This allows us to focus attention on the definitions of protocols and architectures which strictly deal with geospatial data and metadata dissemination issues.

Before proceeding, it is worthwhile to introduce the concepts of *thin* and *thick* clients.

**Definition 14** ***Thin clients*** *rely on invoking the services of other components (servers, middlewares) for most of the computation they need. They also rely* *on other components to manage most of the data and metadata. Thin client is often used as a synonym for browser-based client.*

**Definition 15** ***Thick clients*** *handle much of the necessary computation and data/metadata management themselves, rather than relying on external capabilities. They are usually built upon a robust and complex low-level data-access tier which handles basic input and output operations.*

A *thin* client usually is a *lightweight* component if compared to a *Thick* client since it requires fewer hardware and software resources. On the other hand it usually provides the user with less powerful capabilities.

## 3.1 Data Dissemination Tier

This section focuses on describing with more detail the functionalities provided by the OGC services present in the *Data Dissemination Tier* in figure (4), namely WMS[16], WFS[17] and WCS[18].

**Definition 16** *A* ***Web Mapping Service*** *exposes an XML interface to provide clients with a georeferenced map obtained by superimposition of several layers, coverage and/or features, employing rules specified by the user using styled layer descriptor files (recall section 2.1.3).*

A *Web Mapping Service* standardizes the way through which clients can specify requests for maps, in terms of output format, resolution, coordinate reference system, geographic envelope and style. A WMS service interface provides three basic operations: *GetCapabilities*, which retrieves a collection of metadata (using GML[11]) describing service's information content and accepted request parameters; *GetMap*, which retrieves the requested map; *GetFeatureInfo*, which allows the retrieval of information on a particular feature which was rendered on a returned map.

**Definition 17** *A* ***Web Feature Service*** *exposes an XML interface to let the clients extract geospatial features using GML (other equivalent formats are allowed). A transactional WFS system might also support data manipulation such as updates, insertions and deletions by providing support for long term transactions.*

A *Web Feature Service* can be a *basic* or *read-only WFS*, in such case only *GetCapabilities*, *DescribeFeatureType* and *GetFeature* are supported, or it can be a *Transactional WFS* which provides additional support for the *Transaction* operation and also, optionally, for the *LockFeature* operation.

*DescribeFeatureType* provides detailed information about the schema of a feature in order to allow the user to build refined queries using geospatial and alphanumeric filters. *GetFeature* allows the user to request collections of features for further processing, by means of complex queries based on filters and expressions as seen in section 2.1. Finally a transactional WFS allows the user to lock features for further client-side manipulation in order to ensure coherency and avoiding concurrency issues. Commit and Rollback capabilities are provided to make the performed modifications permanent or to reject them.

**Definition 18** *A* ***Web Coverage Service*** *exposes an XML interface which supports the exchange of coverage. It provides access to intact (unrendered) geospatial information for client-side rendering, or input to numeric processing (e.g. meteorological models).*

A *Web Coverage Service* consists of three operations; *GetCapablities*, which returns a description of the service itself in terms of supported parameters and service metadata as well as a basic description of the served coverage; *DescribeCoverage*, which provides a detailed explanation of the structure of a *coverage* to allow refined requests; the *GetCoverage* operation, which returns coverage, or subparts of them, in the requested format, with the requested interpolation and coordinate reference system.

It is worthwhile mentioning that all the services share the same base structure, built upon three main operations, *GetCapabilities*, *Describe*XXX and *Get*XXX. OGC Web Services aims to be self-describing by supporting the *GetCapabilities* operation. The role of the *GetCapabilities* operation is to provide a *capabilities document* or *service profile* to give a high-level description of a service instance and of its contents. Information provided consists of a human readable description of the service, a specification of the functionalities provided and a set of base metadata for the served content.

## 3.2 Metadata Dissemination Tier

The metadata dissemination tier comprises of the *Web Catalogue Service (WCaS)*[19].

**Definition 19** *A* ***Web Catalogue Service*** *is a container of resource descriptions, called catalogue entries, which are essentially summary information about a particular geospatial resource. It may be used to discover both geospatial data and geospatial services which can be referred as geo-resources.*

Catalogue services are a key technology for locating, managing and maintaining distributed geo-resources (i.e. geospatial data, applications and services). With catalogue services, client applications are capable of searching for geo-resources in a standardized way (i.e. through standardized interfaces and operations). Ideally, they are based on a well-known information model, which includes spatial references and further descriptive (thematic) information that enables client applications to search for geo-resources in an efficient manner.

We introduced a distinction between *Data Dissemination Tier* and *Metadata Dissemination Tier* in order to clearly illustrate the semantic and operational differences between them. Services like WCS, WMS and WFS disseminate mainly geospatial data (*coverage*, *features* and maps), while the dissemination of *metadata* is an accessory task whose objective is to enable the users to refine their queries. Moreover this accessory task is focused mainly on service-level metadata which describes rules for accessing the above mentioned services.

A *Catalogue Service* has a different goal. It supports the ability to *publish* and *search* collections of descriptive information (*metadata*) for data, services, and related information. Metadata in catalogues represents resources' characteristics which can be queried and presented for evaluation and further processing by both humans and software.

*Catalogue Services* are required to support the *discovery* of and the *binding* to registered information resources within an information community, hence their main activity is the *mediation* between clients and data dissemination services. It is correct to state that *Catalogue Services* treat other services (including other catalogues) along with the data and metadata they contain as their data sources.

A typical interaction between a client and a *Catalogue Service* should comprise of three steps. First, the client (human or software) sends a keyword-based search phrase to a *Catalogue Service* describing in detail what it is looking for. Second, the target *Catalogue Service* searches in its data base for matching information or alternatively forwards the search to other catalogues. Third, if matching information is found, the target catalogue can retrieve it back to the client (*Brokered Service*) or provide the client with the links to the services containing the requested data (*Direct Service*).

### 3.3 Final Considerations

Many people think that OGC and ISO standards can not be proficiently employed in real *smart*[16] clients where the geospatial component of the handled information is just a single aspect of the whole and sometimes not even the most important. The objection is that a real *smart* client should not be aware of the service protocols described above because introducing such technologies would be very difficult and inefficient.

Such objection is mostly misplaced. It is unquestionable that a *smart* client, especially when main concern is not on geospatial data, should focus predominantly on implementing the requested business logic. However leveraging the above mentioned services architecture would be of tremendous help for achieving such goal since all the issues related to the geospatial data management (ingestion, transcoding, querying, etc...) would be confined to the remote geo-services and the low-level part of the client itself. This would result in increased *encapsulation*, *robustness*, *extensibility* and *efficiency*.

We would finally like to remark on the benefits of the architecture introduced here. Net separation between different layers of the architecture (*Multi-Tiers Architecture*) ensure *robustness* and *extensibility* by means of progressive *data abstraction*. Dissemination of geospatial data and metadata through standard protocols and interfaces is supported. Complex refined queries can be issued in order to reduce the amount of data transferred, improving bandwidth efficiency. Support for *thin* and *thick* clients is provided. *Modularity* is assured by the division of roles and responsibilities between different services which can work independently or in an integrated architecture (clients are responsible for the fusion of data and metadata coming from complementary services). Finally superior design can be easily achieved as such standards are nothing more than state-of-the-art guidelines for building interoperable geospatial services.

[16]An Internet-connected software that allows the user to implement a complex business logic by interacting both with local data, other applications and with a variety of web services based on heterogeneous protocols.

## 4 Future Directions and Conclusions

### 4.1 Future Directions

Both the *Data* and *Metadata Dissemination Tiers* provides no means for supporting *remote geospatial data fusion*, where the term *remote* refers to the ability to request, by mean of a standard protocol, the execution of computations procedures on a remote server. However, this is a crucial capability, especially when performing computationally-intensive procedures on very large datasets where data transfer could prove cumbersome and counterproductive (as an instance, in terms of bandwidth consumption).

A major discussion is taking place in the OGC in order to address the mentioned shortcoming by providing an ad-hoc specification which leverages and finalizes ISO wide experience in this area. Most recents results are captured in discussion paper [20], where the *Web Processing Service (WPS)* is proposed.

Referring to [20], "a *Web Processing Service would provide access to pre-programmmed calculation and/or computation models that operate on spatially referenced data which can be delivered across a network, or available at the server*".

Likewise the services discussed above, the WPS interface comprises of three different operations: *GetCapabilities*, which provides the user with brief service-related metadata describing service's abilities and processes offered; *DescribeProcess*, which describes in detail one or more offered processes along with its input and output parameters; *Execute*, which allows the user to trigger the execution of specified remote processes by provision of the requested input values.

*Synchronous* and *asynchronous* executions are addressed. Moreover such services may support the *store* option in order to make the complex results of executed processes available through the web. *Status* information, recommended for procedures with a long run time, can be provided in order to track progress of the execution of asynchronous processes.

Provision of standard self-containing formats for storing raster data along with their metadata, with emphasis on support for tiling, multiple resolutions and compression, is under discussion. Partial results of this standardization are summarized in specifications [21][22][23].

Particular attention should to be drawn to [21] and [23]. The first specification aims to exploit the JPEG2000[17] standard as a storage format for coverage, since it provides the ability to store both metadata, using XML dialects as GML, and data. In addition, it provides support for wavelet compression and tiling which is critical for acceptable performances on large datasets. The second specification aims to extend preexisting standards with regard to support for coverage's metadata. This issue is becoming more and more important in order to proficiently exploit the huge amount of data being released, mostly at no costs, from various organizations (see EROS[18] at USGS).

### 4.2 Conclusions

This paper presented an overview of the capabilities of standards issued by the Open Geospatial Consortium and the Technical Committee 211 of ISO. Special focus was placed upon support for *geospatial data fusion*, both *stand-alone* and *distributed*.

Base definitions for fundamental concepts have been provided in section 1.2. Particular emphasis has been placed on grouping together capabilities from different specifications in order to organized them in tiers with respect to the issues they address and the level of abstraction they achieve. The various modules composing the *Data Management Tier* (see section 2) as well as the services exposed by the *Data and Metadata Dissemination Tier* (see section 3) were discussed. An explanation of each module and service, as well as a description of main benefits and capabilities has been provided. Field experience has been used to validate the presented approach.

Finally an overview of the recent advancements of the ongoing standardization process has been provided in section 4.1 in order to allow the reader to place the proposed architecture in a more general context which includes future extensions.

[17] *http://www.jpeg.org/jpeg2000/index.html*
[18] http://seamless.usgs.gov/

## References


[1] Percival, G., "OGC Reference Model, Version 0.1.3", Document 03-040.

[2] ISO/TC 211 Geographic Information/Geomatics, "Geographic information – Schema for coverage geometry and functions.".

[3] Herring, J., "Topic 1: Feature Geometry (ISO 19107 Spatial Schema), Version 5", Document 01-101.

[4] Kottman, C., "Topic 5 - Features , Version 4", Document 99-105r2.

[5] "Topic 6: The Coverage Type and its Subtypes, Version 4", Document 00-106.

[6] Kottman, C. "Topic 10: Feature Collections, Version 4", Document 99-110.

[7] Ryden, K., "Simple Features For SQL Implementation Specification (version 1.1)", Document 99-049.

[8] Nebert, D. (ed.) "Developing Spatial Data Infrastructures: The SDI Cookbook".

[9] Bray, T., Paoli, J., Sperberg-McQueen, C. M., Maler, E. Yergeau, F., "Extensible Markup Language (XML) 1.0 , Third Edition".

[10] Bobbitt, J. "OpenGIS Recommendation Paper-Unit of Measure Use and Definition Recommendations Version 1.0".

[11] Cox, S., Daisey, P., Lake, R., Portele, C., Whiteside, A. (eds.) "Geography Markup Language (GML) Implementation Specification (version 3.1.0)"', Document 02-023-r4.

[12] ISO/TC 211 Geographic Information/Geomatics, "Geographic information – Temporal Schema".

[13] Lott, R. "Abstract Specification Topic 2: Spatial referencing by coordinates, ", Document 04-046r3.

[14] Panagiotis A. Vretanos, "Filter Encoding Implementation Specification (version 1.0.0)", Document 02-59.

[15] Lalonde, B, "Styled Layer Descriptor Implementation Specification (version 1.0.0)", Document 02-070.

[16] Jeff de la Beaujardiere, W. (ed.) "Web Map Service Implementation Specification (version 1.3.0)" OGC Document 01-109r1.

[17] Vretanos, P. (ed.) "Web Feature Service Implementation Specification (version 1.0.0)", OGC Document 02-058.

[18] Evans, J. (ed.) "Web Coverage Service Implementation Specification (version 1.0)", OGC Document 03-065r6.

[19] Nebert, D., Whiteside, A. "Catalogue Services Implementation Specification (version 2.0.0)", OGC Document 04-021r2.

[20] Schut, P., Whiteside, A., "OpenGIS Web Processing Service, Version 0.4", OGC Document 05-007r4.

[21] Kyle, M., Burggraf, D., Forde, S., Lake, R., "OpenGIS GML in JPEG 200 for Geographic Imagery, Version 1.0.0", OGC Document 05-047r3.

[22] Schlesinger, B., "Imagery Metadata, Version 1.0", OGC Document 05-015.

[23] ISO/TC 211 Geographic Information/Geomatics, "Geographic information – Metadata – Part 2: Extensions for imagery and gridded data".