Article under Review for the International Journal of Spatial Data Infrastructures Research, submitted 2013-02-27

Assisted Generation and Publication of Geospatial Data and Metadata*

Sergio Trilles[1], Laura Díaz[1], José Gil[1], Joaquín Huerta[1]

[1]NIT, Universitat Jaume I;
strilles@uji.es; laura.diaz@uji.es; jose.gil@uji.es; huerta@uji.es

Abstract

Today, the existence of metadata is one of the most important keys for an effective discovery of geospatial data published in Spatial Data Infrastructures (SDI). However, due to lack of efficient mechanisms, integrated in user daily workflow, to assist users in generating metadata, little metadata is produced. This paper presents a mechanism for generating and publishing metadata with a Publication Service. This mechanism is provided as a web service implemented with a standard-base interface, the Web Processing Service (WPS) specification to improve its interoperability with other SDI components. This paper extends previous research when designing a Publication Service in the framework of the European Directive INSPIRE as a solution to assist users in publishing geospatial data and metadata automatically in order to improve, among others, SDI maintenance and usability.

Keywords: metadata, automated assistance, geoprocessing, discovery, SDI, INSPIRE

1. INTRODUCTION

The current trend is to deploy and organize this information in Geospatial Information Infrastructures (GIIs) also known as Spatial Data Infrastructures (SDIs) (Masser, 2005). To increase the efficiency and interoperability of GII many

---

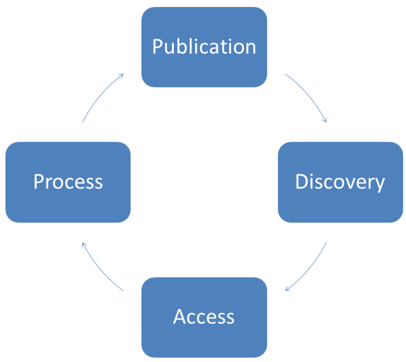
regional and global initiatives work in the establishment of open standards and agreements.

To increase the efficiency and interoperability of GII many regional and global initiatives work in the establishment of open standards and agreements. A GII on European scale has been legally mandated: the Infrastructure for Spatial

Information in Europe (INSPIRE). It should provide environmental data related to 34 themes, including transport networks, land cover and hydrography, INSPIRE provides important parts of the European contribution to a Global Earth Observation System of Systems (GEOSS).

Content is managed by means of regulated and standardized service types. This imposes a distinct life cycle of geospatial content in distributed environments which can be described in four steps as illustrated in Figure 1. First, content must be made available to a distributed system, i.e., content must be published in standard services like discovery and access services. Second, users need to discover content which will be finally accessed by using these services (third step). Finally, users process the content and generate new content, which should be integrated and published in the distributed system closing this cycle.

Figure 1: Content Life Cycle in GII



The GEOSS Service Factory (GSF) is our proposal to develop a generic publication service to assist users, both, experts of environmental domains or casual users, in content publication on certain system. To illustrate with a practical example we show the development and deployment of the GSF in a forestry fires system, where users benefit from this new channel to provide information.

Compared to existing Geographic Information System (GIS) tools, which publishing geospatial data and maps using proprietary software, the proposed solution is implemented as a standard interfaced web service and can be generally applied a wide range of content types, service standards and information systems.

Because of the current status of the SDI and the mechanisms deployed for discovery, metadata (necessary to describe the resources) and Catalogue Services are the key elements for the discovery and resource fusion possibilities (Nogueras-Iso et al 2004) (Díaz et al 2007). In this context, the INSPIRE Directive mandates the creation and maintenance of metadata and related Discovery Services (Craglia et al 2007). According to the INSPIRE Directive (INSPIRE EU Directive 2007), it is necessary to generate and publish metadata for each data published in infrastructure. This metadata must be validated (meet standards) and accurately describe the data being referred (Nogueras-Iso et al 2009).

Currently there are a lot of geographic data published online but still there is a lack of is associated metadata published online, making difficult its discovery (Nogueras-Iso et al 2005). One of the reasons could be the implicit difficulty and arduous process of creating them manually and the lack of mechanisms to do this combined with the creation of the data which would be the best moment to collect all of the information about the data (Zarazaga-Soria et al 2003).

To this day there is no trivial automated mechanism that does not require technicians to assist in the process (Díaz et al 2011a). This is because for an effective generation of metadata, the generation and publication of data and metadata have to be done together to increase the information about the data when generating the metadata.

This work aims to bridge the gap in the generation and publication process of geographical data together with its metadata in an interoperable and scalable component. We present a publication service focusing on describing its capability of automatically generates and publishes the associated metadata's (Trilles et al 2011). We provide this functionality implementing an SDI standard-base service that facilitates the creation and publication of metadata.

There are two ideas that drive our motivation throughout the project. First, to improve the sharing of geospatial data by assisting in the publication of data as standard-base services and second by facilitating the discovery of this geographic data.

Section 2 lists the different works in the literature for the generation and publication of metadata. We present the GSF as part of GII architectures in Section 3, before detailing the involved components for the data publication (Section 4) and metadata generation and publication (Section 5), the experiments

(Section 6) and concluding the paper with a discussion and an outlook to future work in Section 7.

2. RELATED WORK

Metadata are created by data providers and stored in catalogs according to the standards dictated by international organizations such as ISO* they are adopted as standards by the community. The purpose of this process is for users to find specific data in distributed environments.

Today the creation of metadata is not combined with the creation and publication of data, because the existence of difficulty of metadata creation. In addition there are not realized functions of maintenance of the metadata (Beltran et al 2009).

Metadata creation is mainly a manual process normally separate from the process of data generation, this provokes metadata creation to be a time consuming task that in many cases lacks critical information such as where the data is published online and how can it be accessed (Nogueras-Iso et 2005).

In the literature, there is research that addresses this issue by proposing automatic tools for metadata generation, which is a complex task that can be categorized according to the generation techniques (Bulterman et 2004) (Manso et al 2004) (Beltran et al 2010a). These are manual generation, related works describe automatic and semi-automatic techniques (Campbel et al 2004) (Currier et al 2004) (Jorum 2004). (Beard 1996) describes five categories: manually, extending the data stored with values obtained by consulting, using automatic measurements and observations, using data extracted and calculated and finally inferred from other elements. While other authors categorize the automatic generation into two classes: metadata extraction and inference (Beltran et al, 2009).

Tools such as CatMDEdi (Zarazaga-Soria et al 2003) are used for automatic extraction of metadata in different formats, as reflected in the work (Taussi 2007); the amount of information that can be extracted fundamentally depends on the representation model used and its file format. Thus, there are elements that can only be extracted from certain types of data and files, while others, such as the size of the data, can be obtained in all circumstances.

Other works describe how to integrate the generation of metadata within GIS used to generate data as well. These tools already provide the functionality to be able to read from the data formats which facilitates the extraction of metadata In

* International Organization for Standardization, http://www.iso.org

this works we can highlight the few existing GIS tools that offer an automatic deduction of metadata for raster and vector containers that are based on the analysis of these specific formats, and application of ad-hoc mechanisms that process the data to extract information is used in the metadata construction (Mansó et al 2004) and the gvSIG metadata editor (Beltran et al 2009b).

Besides generation techniques we can describe the nature of the tools which provide metadata generation functionality. Normally this functionality is included in Catalogues Services where metadata will be later stored, this is, for example, in the case of Geonetwork[†] and ESRI ArcCatalog, that is used for the automatic generation of metadata for geographical data. This tool enables loading basic fields automatically and updating data synchronization and metadata. In both cases, to improve these tools we can add extensions to transform metadata from one metadata standard to another. ARCCatalogue provides a Metadata Editor from the profile *Nucleo Español de Metadatos* (NEM). It is integrated into the ArcCatalog.

In order to provide a component for metadata generation and publication using extraction techniques we wrap this functionality to be provided as an interoperable service that can be combined with other services for data generation. We aim at having the data and metadata generation and publication in the same workflow as a standard services in SDI. This facilitates metadata generation so as to know more about the data itself, like when has been generated, where is it available, so we increase the linkage between data and metadata in order to facilitate further data visibility and accessibility.

In our approach is a Publication Service called GSF (Díaz et al 2011b) in this work we extend previous research work by adding the ability to generate metadata and publish it to this Publication service, this process when executed after data generation increases the amount of data descriptions thereby generating more useful metadata.

In order to implement our service interface we choose the standard Open Geospatial Consortium (OGC) Web Processing Service (WPS) (Schut 2007). WPS allows the deployment of any functionality on the Internet. This standard provides standards to describe any computation (process) and how to make service requests and responses.

[†] Geonetwork OpenSource, http://geonetwork-opensource.org

3. CONTEXT: EXTENDING SDI WITH A PUBLICATION SERVICE

The work described is part of the GSF (Díaz et al 2011a) (Díaz et al 2011b), a publication service that facilitates the provision of content and maintenance in the realm of SDI by all participants. The GSF provides, as an interoperable service implementing the OGC WPS specification, the functionality to publish geospatial data as a standard data service deployed in an SDI according to the European directive INSPIRE. GSF offers a solution for publishing geospatial content aimed at increasing data accessibility for improving data sharing, enabling automatic publication of them and providing a collaboration tool for the end-user.

The life cycle of content available in an information system, consists of four steps: publishing, discovery, access and processing (Díaz et al 2011a). GSF improves the publication stage of content, by simplifying and automating the publishing process. The publication is part of the natural workflow for processing information. After the discovery, access and/or data processing, the online publication of new content needs to be created to be reused by others interested stakeholders in the SDI.

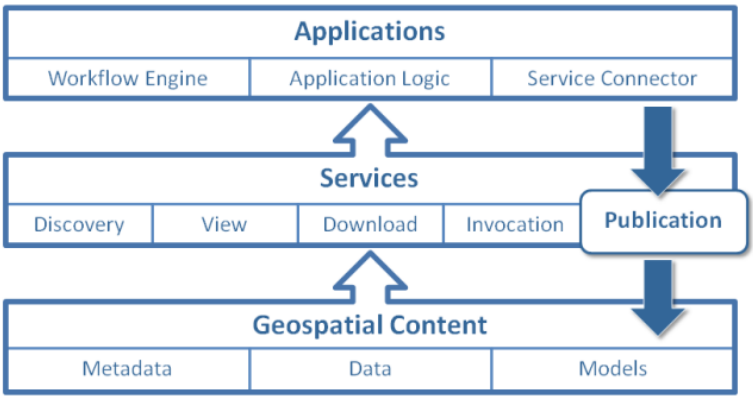
Within the geospatial domain, interoperability is ensured by standardization efforts most prominently by the OGC. OGC has proposed a number of standards, which promote syntactic interoperability through the use of services (Percival, 2008). The existing specifications have been proven to help in setting up GII; these include: OGC Catalogue Services (CS-W), OGC Web Map Service (WMS), OGC Web Feature Service (WFS), OGC Sensor Observation Service (SOS), and OGC Sensor Event Service (SES) interface specifications as examples to mediate geospatial content. Other specifications such as OGC WPS provide an interface for accessing processing functionality as distributed Web Services. INSPIRE specifically recommends the use of OGC Services to implement its service types. An overview of OGC Service Interface Specifications together with superseding INSPIRE Service Types is given in Table 1.

Table 1: Geospatial Service Types and Common Standards

Description	OGC Specification	INSPIRE Service Type
Vector Data Download Service	WFS	Download
Raster Data Download Service	WCS	Download
Catalog and Discovery Service	CS-W	Discovery
Sensor Data Download Service	SOS	Download

In the previous section we described the fact that existing applications to create and publish metadata are mainly desktop applications. In contrast, we provide the functionality of generation and publishing metadata as an interoperable, scalable service that can be added to other service calls to be chained. Figure 2 shows where the GSF is deployed in an INSPIRE base SDI architecture, where all the functionality is provided as a service in the service layer. The 'Application' layer includes end user applications, ranging from complex Environmental Decision Support Systems (EDSS) to simple clients on mobile devices. Client applications access geospatial content stored in repositories through services in the middle layer. *Geoportals* are a special type of applications. They provide the entry point to domain-specific GIIs (Bernard et al, 2005). Such portals are deployed in most of the recent GII implementations.

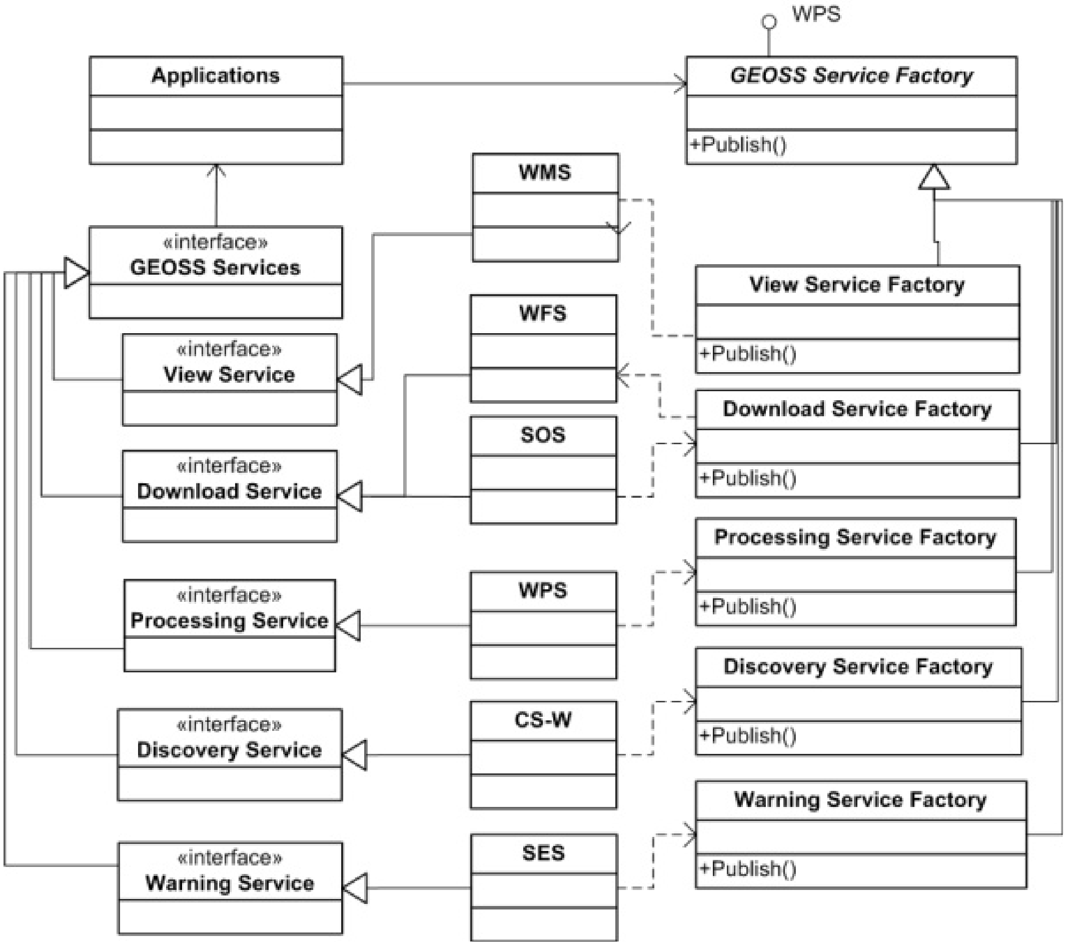
Figure 2: Proposed Extended GII Architecture



GSF is implemented with an OGC standard interface specification, particularly the GSF is an OGC WPS. GSF internal design uses the software design pattern: the factory pattern (Vlissides et al 1995). The factory pattern is a creational pattern that provides a scalable mechanism to create new entities according to particular criteria. In this case, the factory pattern is used to design within the service a battery of factories each of them dealing with the creation of new entries of a certain type in existing SDI services (Díaz et al 2011b). This allows the GSF to publish new content (data and metadata) by adding new entries to different services types (according to factory type) View, Download and Discovery.

Figure 3 shows the Abstract Factory diagram adapted to our scenario. We use inheritance to derive the most specific OGC standards from the more general service types defined by INSPIRE and GEOSS (central and left part of the Figure 3). View and Download Services are some of the INSPIRE Service Types adopted in our approach to generate GEOSS Services and to deploy them in a certain GII (see also Table 1). The individual services implement a particular OGC standard specification following INSPIRE guidelines to implement INSPIRE service types. For example, a WMS is used to implement a *View Service*, a WFS to implement a *Download Service*.

Figure 3: Abstract Factory Pattern Applied to (GEOSS) Service Factory



In this work the scope is limited to publish content exclusively in OGC services; due to their specifications we differentiate two possibilities of communication between the factories and OGC service instances:

1. Service interface with transactional operations: Some OGC standards (such as WFS, SOS, and CS-W) include transactional operations in their specification. In this case, the factory can publish content by implementing the client side of the interface. In this case, the factory implementation is independent of service implementations. It will be able to deploy content in any instance that implements the chosen (transactional) standard service interface.
2. Service interface without transactional operations: Other OGC specifications do not provide transactional operations. In these cases, the publication has to

be supported by other means. This is, for example, the case for the WMS, a common standard to implement View Services.

The drawback in this case is that the factory depends on the implementation technology of the service instance to publish content. At this point, it is worth to remark the special behavior of the DiscoveryFactory. This factory can be invoked when the content nature is a metadata, i.e. data about data or services, but it could also be invoked anytime after publishing any other content. We envision de DiscoveryFactory to contain the intelligence to be able to request other service types where new content has just been published in order to extract some metadata and generate automatically a small set of metadata elements to be published in a Discovery Service.

According to INSPIRE implementation rules for these types of services that should be used for implementation are ISO 19128:2005 WMS for visualization services, WFS for download and CS-W for discovery.

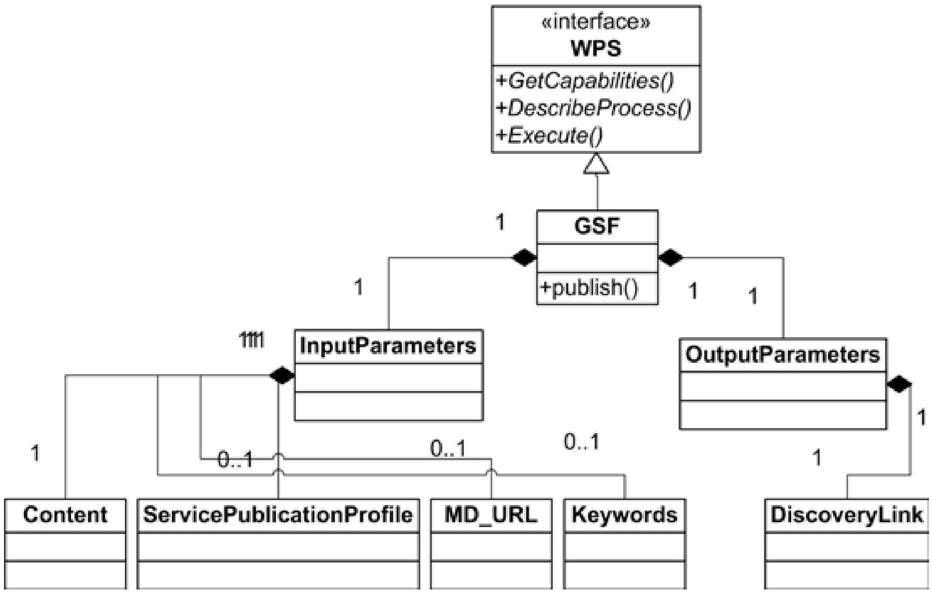
The result of executing the GSF are metadata Uniform Resource Locator (URLs) describing the content just published, this metadata contains information about the download, visualization and discovery services where the data is available.

The GSF is associated with a profile or Service Publication Profile (SPP), this contains the necessary information for the publication and metadata generation. The SPP is an Extensible Markup Language (XML) document that can be sent as an input.

4. GENERATION AND PUBLICATION OF GEOSPATIAL DATA

The GSF is designed to be implemented as a service component with a standard interface to be re-used in different scenarios. Since the OGC WPS (Schut, 2007) is used to reach processing interoperability, it is our standard of choice to implement the GSF as a service. Figure 4 shows the UML class diagram with a simplification of the GSF interface and the signature of the publish process regarding input and output parameters.

Figure 4: Proposed Extended GII Architecture



We offer a single process, called *Publish* (see Figure 4). WPS specifies that inputs and outputs can be encoded in many alternative ways. The Publish process considers the following parameters:

- Content: Only this input parameter is mandatory. This content can be passed by value or by reference, where a Uniform Resource Locator (URL) to the content can be used. It can vary from a vector or raster data set, a workflow description, or a metadata document.
- ServicePublicationProfile: XML encoded parameter that describes the publication policy. This parameter includes information regarding where each data type should be published within this GII.
- MD_URL: This parameter indicates that this content is already published in the GII and there are available metadata that should be reused when updating it. This parameter is optional.
- Keywords: The optional ‘keywords’ parameter provides an initial capability for metadata creation.
- DiscoveryLink: This is the only output parameter. This parameter contains the information needed to discover the content published in the system. In the case of the GII, where content is registered in Catalogue services, this parameter contains the end point to the metadata available in the Catalogue Service that contains the description of the content just being published. This contains information about the data services end points serving the content. This parameter is optional.

Each application system deployed in a GII can have its own publication policy. This policy establish rules, for example, which content type is published in each service type that implement a particular specification and that is located in a concrete end point. At technical level we describe this policy as a Service SPP. SPP is set for a GSF deployed in a GII and it will configure the GSF to decide where each factory publishes the content. For example, the SPP determines which content types are published for visualization and download and also publishes their metadata for discovery purposes.

The use of the SPP allows the GSF be more scalable and flexible. When invoking GSF, the SPP can describe which content is published for different purposes in different service types. For example, a single geospatial data set may be published in an OGC WMS and in an OGC WFS at the same time. Otherwise the SPP can determine that the GSF only publishes this content for one service type. We explain the SPP in more detail in the next section, where we introduce a prototype implementation of the GFS in the context of forest fire management.

GSF is able to publish raster or vector data. If we talk about vector data the GSF is able to publish Shapefile in the WFS service. In addition, we have added the ability insert OM (Observations and Measurements) data from sensors in the SOS services that we describe in the SPP. If we speak of raster data, the GSF is capable of inserting data such as GEOTIFF, Keyhole Markup Language (KML) and Geography Markup Language (GML) in the WMS service.

5. GENERATION AND PUBLICATION OF METADATA

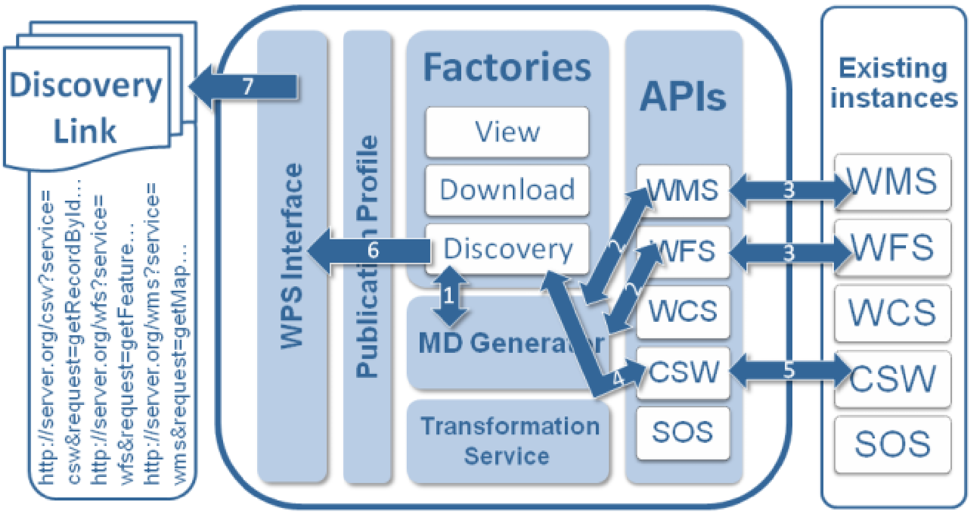
The present work aims to define a module that is able to generate and publish metadata within GSF. To generate the metadata information the data and the site where it is published must be specified. Part of this information can be extracted from the services where the data has been published. These data services are mainly available in SDIs.

These services conform to the INSPIRE implementing standards are as OGC services WMS, WFS, Web Coverage Service (WCS) and SOS. Other information that is more difficult to identify are the keywords, so we designed GSF with input parameters such as keywords, which can be indicated by the user indicated by the user.

In general, the metadata generator will have to run when the data is published, if not stated otherwise. The factory responsible for working with metadata generator is the Discovery Factory.

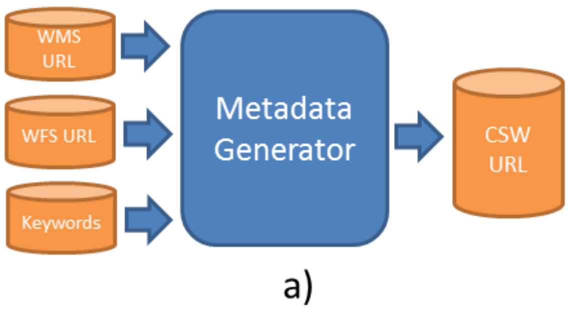
The first step is to consult WMS and WFS services to get the information needed to create the metadata. This is done through appropriate requests (GetCapabilities). After obtaining the necessary information the process applies metadata generation process and proceeds to their publication in the catalog (CSW). Finally, the WPS will return the URL result to the publication. This operation can be seen in Figure 5.

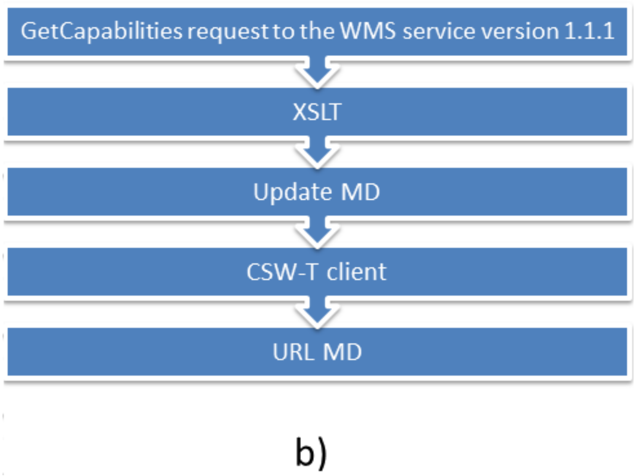
Figure 5: Steps for the Publication and Generation Metadata in the GSF



After publication of content, if the WPS has not received a URL with the metadata (MD_URL) metadata is then called metadata generator. This requires the URLs to public services to generate the requests, in addition to input keywords. The only output parameter of the GSF has a URL to access the metadata catalog service where it has been published. To further explain the process of metadata generation, we have the following image (Figure 6) which shows the inputs and outputs of the Metadata Generator module. Figure 6b shows the steps involved.

Figure 6: A) Input and Output Parameters Metadata Generator B) Steps Metadata Generator





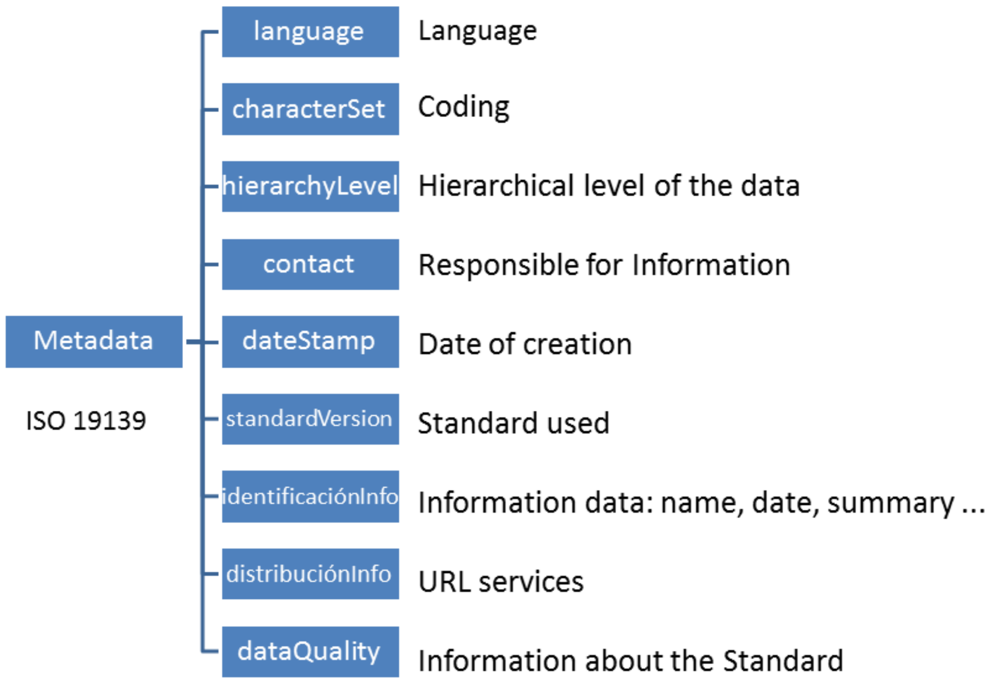
In the first step the GetCapabilities request is performed to services of information where the previous content has been published. For example, if the content has a vector, this will be published for visualization in WMS and for download in WFS. This request returns a XML with the characteristics of the information that will be necessary for the metadata creation. Secondly, the same step is done by the WFS, to provide more information to the metadata.

The following step must apply the Extensible Stylesheet Language Transformations (XSLT) to generate metadata. The transformation will take the XML entry obtained from the previous request. The result obtained after the transformation is another document XML, which is the metadata with the ISO 19139 standards and INSPIRE.

The third step consists of parsing the XML metadata for fields that have not been filled in the transformation, such as keywords or URLs to services. After this step, the metadata will be considered completed.

Then the metadata will be published in the catalogue that is defined in the SPP. This step is performed similarly to the publication of data but in this case to the catalog. In our work we use transactional profile (CSW-T) according to the OpenGIS Catalogue Services specification. In this way, the system will have to implement all operations allowing the CSW-T protocol. These are: GetCapabilities, GetRecords, GetRecordById, GetDomain, DescribeRecord and the transactional.

Figure 7: Some Metadata Fields



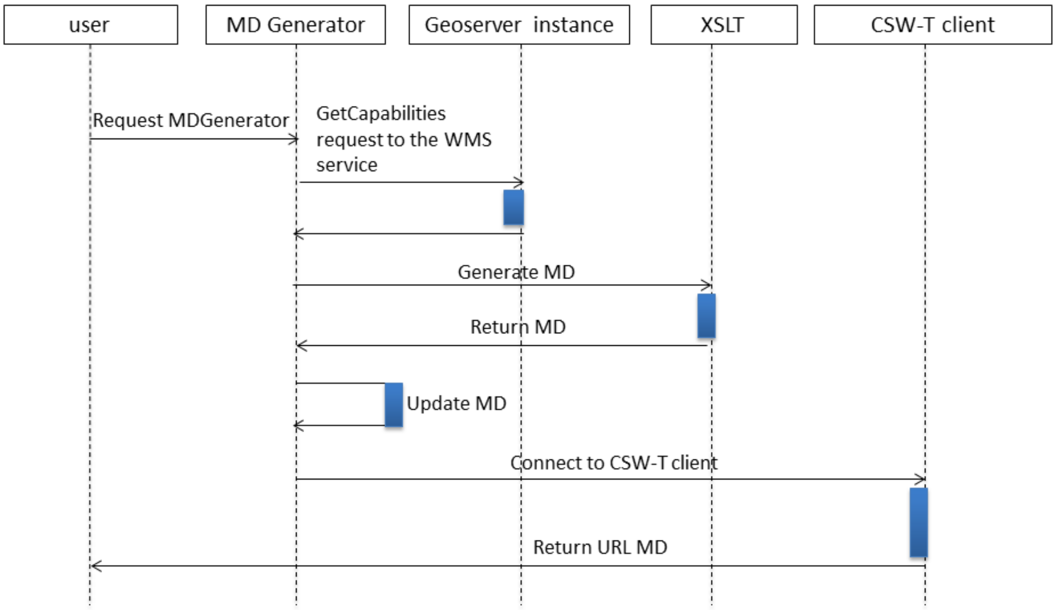
The last step is to get the URL that identifies the metadata posted on the server. After obtaining the metadata URL, the generator metadata is considered as finished.

For the generation of metadata has been chosen to create the template using XSLT. This technology is a standard from the World Wide Web Consortium (W3C) organization that allows us a way of transforming XML documents or other types of documents. As we have already stated metadata is nothing more than an XML following a standard so that we can apply this type of transformations.

We should mention that the metadata that is generated does not have all the fields in the specification and some cannot be obtained automatically. We have chosen the fields you can see in the figure below (Figure 7). It shows the main fields that contain the metadata generated.

The following figure (Figure 8) shows the sequence diagram of the execution flow for the metadata generation and publication.

Figure 8: Sequence Diagram



6. EXPERIMENTATION

Currently, GSF is able to publish different raster and vector formats, to perform the experiment, we use a Shapefile as an example for vector data and Geotiff file as an example for raster format. The data files, the content, is sent to the GSF as an input parameter in the WPS Execute request (as a mandatory parameter).

In particular, figures show the publication process of the vector data describing fire-burned areas in the province of Castellón, in Spain in the year 2005. Having this data available as an interoperable SDI services, improves its accessibility for further reuse by other stakeholders to carry out analysis of environmental impact caused by such events. Furthermore, in order to improve its visibility, it is necessary that it is published not only in a data service, for visualization or download, but also that its metadata is published in a metadata catalog, offering information about what is it, how is it and where it can be accessed.

One of the (open source) implementations of the OGC-based SDI services is Geoserver, which offers implementation for OGC WMS, WFS and WCS. We have chosen Geoserver to test the data publishing. Moreover, the metadata catalog GeoNetwork has been chosen for the metadata publication.

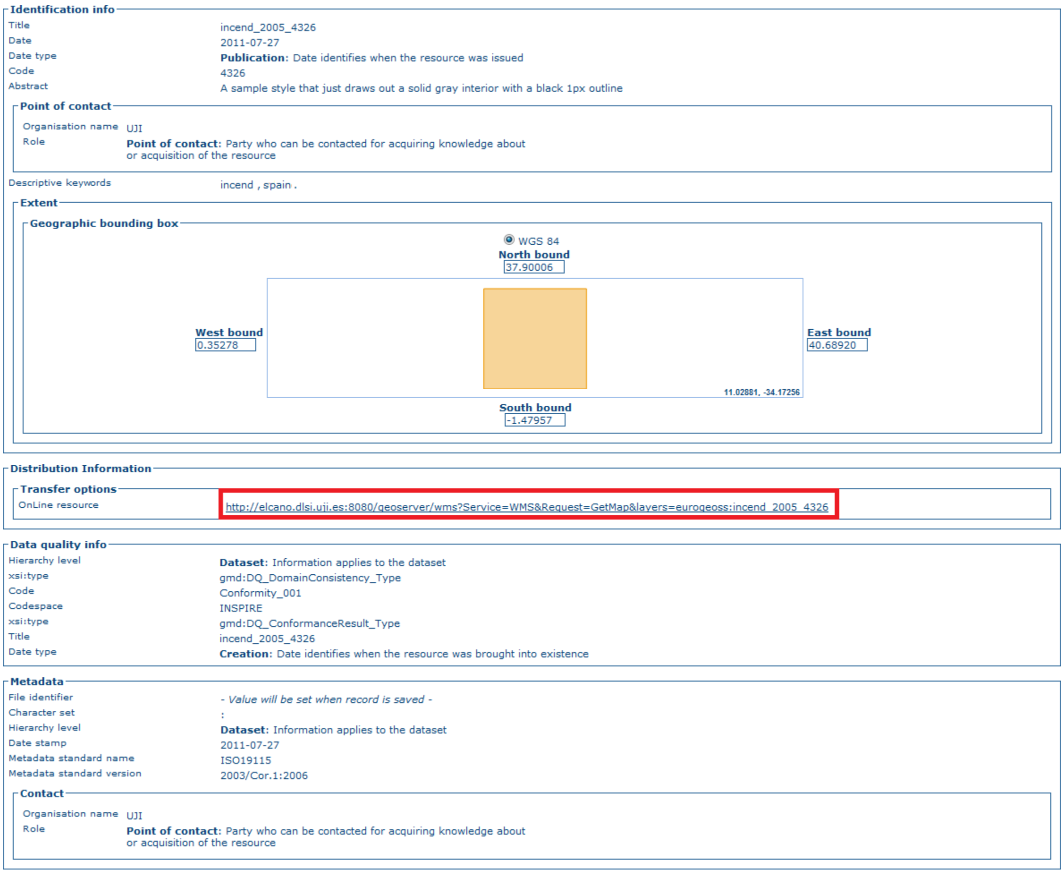
Figure 9: Example of a GetCapabilities Layer

```xml
<Layer queryable="1">
  <Name>eurogeoss:incend_2005_4326</Name>
  <Title>incend_2005_4326</Title>
  <Abstract/>
  <KeywordList/>
  <SRS>EPSG:4326</SRS>
  <!--WKT definition of this CRS: GEOGCS["WGS 84",
  DATUM["World Geodetic System 1984",SPHEROID["WGS
  84", 6378137.0, 298.257223563, AUTHORITY["EPSG",
  "7030"]], AUTHORITY["EPSG","6326"]], PRIMEM["Gre
  enwich", 0.0, AUTHORITY["EPSG","8901"]], UNIT["d
  egree", 0.017453292519943295], AXIS["Geodetic lo
  ngitude", EAST], AXIS["Geodetic latitude", NORTH
  ], AUTHORITY["EPSG","4326"]]-->
  <LatLonBoundingBox minx="-1.4795661150193993"
  miny="37.90006305852903" maxx="0.3527841387963444"
  maxy="40.68920370363359"/>
  <BoundingBox SRS="EPSG:4326" minx="-1.47956611501
  93993" miny="37.90006305852903" maxx="0.352784138
  7963444" maxy="40.68920370363359"/>
  <Style>
    <Name>polygon</Name>
    <Title>Default polygon style</Title>
    <Abstract>A sample style that just draws out
    a solid gray interior with a black 1px outline
    </Abstract>
    <LegendURL width="20" height="20">
      <Format>image/png</Format>
      <OnlineResource xmlns:xlink="http://www.w3.org
      /1999/xlink" xlink:type="simple" xlink:href=
      "http://elcano.dlsi.uji.es:8080/
      geoserver/wms?request=GetLegendGraphic&amp;
      format=image%2Fpng&amp;width=20&amp;height=20
      &amp;layer=incend_2005_4326"/>
    </LegendURL>
  </Style>
</Layer>
```

After the execution we see how the data has been published to the geographic data server Geoserver and the getCapabilities is accessible to WMS and WFS services. Figure 9 shows the layer information published.

On the other hand, we also have published in the metadata server, GeoNetwork. In Figure 10 we see the metadata information that is visual. The online resource to the data is inside the red box.

Figure 10: GeoNetwork Capture



The full execution of WPS, publication of data and metadata generation and publication, has a time cost of 18 seconds. This delay is caused by the XSLT templates is costly. Services and client are on the same server‡.

Finally, we have used the various operations to the server SOS, register a sensor and register an observation. We have recorded different sensors and published observations of some sensors.

‡ Intel(R) Xeon(R) CPU 5160 @ 3.00GHz, RAM Memory 16 GB DDR2 FB-DIMM Synchronous 667 MHz

We have published a sensor, first using the SensorRegister operation for the publication of a new sensor in the SOS server. Figure 11 shows an example of sensor registration.

Figure 11: Example of Register Sensor

```xml
<sml:SensorML xmlns:sml="http://www.opengis.net/sensorML/1.0.1"
xmlns="http://www.opengis.net/sensorML/1.0.1"
xmlns:gml="http://www.opengis.net/gml"
xmlns:swe="http://www.opengis.net/swe/1.0.1"
xmlns:xlink="http://www.w3.org/1999/xlink"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
version="1.0.1"
xsi:schemaLocation="http://www.opengis.net/sensorML/1.0.1 http://schemas.opengis.net/sensorML/1.0.1/sensorML.xsd">
<sml:member>
     <sml:System xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"   >
         <!--sml:identification element must contain the ID of the sensor-->
          <sml:identification>
               <sml:IdentifierList>
                    <sml:identifier>
                         <sml:Term definition="urn:ogc:def:identifier:OGC:uniqueID">
                              <sml:value>urn:ogc:object:feature:Sensor:ESPVA0300000003009B_Direccion</sml:value>
                         </sml:Term>
                    </sml:identifier>
               </sml:IdentifierList>
          </sml:identification>
          <!-- sml:capabilities element has to contain status and mobility information -->
          <sml:capabilities>
                 <swe:SimpleDataRecord>
                     <!-- status indicates, whether sensor is collecting data at the moment (true) or not (false) -->
                     <swe:field name="status">
                           <swe:Boolean>
                              <swe:value>true</swe:value>
                            </swe:Boolean>
                        </swe:field>
                        <!-- status indicates, whether sensor is mobile (true) or fixed (false) -->
                        <swe:field name="mobile">
                            <swe:Boolean>
                              <swe:value>false</swe:value>
                            </swe:Boolean>
                         </swe:field>
                    </swe:SimpleDataRecord>
          </sml:capabilities>
         ...
     </sml:System>
     </sml:member>
     </sml:SensorML>
```

We have also implemented the function (InsertObservation) to publish measurements. Figure 12 shows an example of publication of measurements.

Figure 12: Example of Measurement


```
<om:Measurement xmlns:sml="http://schemas.opengis.net/om/2.0"
xmlns="http://schemas.opengis.net/om/2.0"
xmlns:gml="http://www.opengis.net/gml"
xmlns:swe="http://www.opengis.net/swe/1.0.1"
xmlns:xlink="http://www.w3.org/1999/xlink"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
version="1.0.1"
xsi:schemaLocation="http://schemas.opengis.net/om/2.0 http://schemas.opengis.net/om/2.0/observation.xsd">
   <om:samplingTime>
   <gml:TimeInstant>
   <gml:timePosition>2012-02-06T17:44:15+00:00</gml:timePosition>
   </gml:TimeInstant>
        </om:samplingTime>
          <om:procedure xlink:href="urn:ogc:object:feature:Sensor:ESPVA0300000003009B_Direccion"/>
          <om:observedProperty xlink:href="urn:ogc:def:phenomenon:OGC:1.0.30:_Direccion"/>
          <om:featureOfInterest>
               <sa:SamplingPoint gml:id="foi_ESPVA0300000003009B">
                   <gml:name>foi_ESPVA0300000003009B</gml:name>
                   <sa:sampledFeature xlink:href=""/>
                   <sa:position>
                     <gml:Point>
                       <gml:pos srsName="urn:ogc:def:crs:EPSG::4326">38.4  -0.5</gml:pos>
             </gml:Point>
                   </sa:position>
                 </sa:SamplingPoint>
          </om:featureOfInterest>
          <om:result uom=" ">8.5</om:result>
     </om:Measurement>
```


The execution of WPS for the publication of sensors or observations has a time cost of 2 seconds on the same server[§].

7. CONCLUSIONS

We argued that content publication to GIIs causes the central bottleneck in environmental information sharing. Bottom-up approaches are needed in the context of GIIs to assist users in populating these infrastructures with content. Following the work of international activities, the GSF was promoted as a solution. We proposed the GSF as a publication service of GII architectures. In order to ensure platform independence, it is provided as a separate component that becomes accessible through standard interface such as OGC WPS from the Geoportal frontend of a GII. In future a possible specialization of this interface can be developed using profiling, to offer a common profile for content publication in GII. In this work a core operation called ‘Publish’ and its mandatory and optional parameters have been defined. We recommended the deployment of INSPIRE Service Types, because INSPIRE provides a formal framework to GEOSS while at the same time providing an abstraction layer on top of OGC standards. A prototype has been depicted as proof of concept.

[§] Intel[(R)] Xeon[(R)] CPU 5160 @ 3.00GHz, RAM Memory 16 GB DDR2 FB-DIMM Synchronous 667 MHz

Our proposal alters the role of GII users, being either professionals or casual users. They turn from rather passive consumers into active participants playing a more interactive role and providing new content (Budhathoki et al, 2008). Now, users can participate in the maintenance and updating of the GII. This means that users, besides searching, accessing and analyzing data, could massively publish newly generated content as interoperable components. This would improve the availability of interoperable content in global, regional and local services related to domain specific scenarios and could increase the effectiveness of GIIs.

In this work there has developed a mechanism of generation and publication automatic of data and metadata, extending the previous work of publication of geographical information in SDIs (Díaz et al 2011a). We also added the ability to publish observations SOS and sensors SOS.

The suggested flexible implementation of the GSF offers possibilities for resource plug-and-play. New factories can be added using class inheritance. This may be extended in order to publish other content types, such as environmental simulations and Volunteered geographic information (VGI), including geospatial data models extended with uncertainty information (Williams et al. 2008) or nongeospatial content, into GIIs. The suggested approach also provides flexibility in terms of functional extensions, such as content validation, security, and automated reasoning capabilities.

The global component addresses the improvement of data availability in the environment of the SDI as well as its maintenance, the particular generation and publication of metadata facilitates where data is found, because it will generate and publish metadata which defines and facilitates its location. There has been a possible solution to the problem of availability of data and metadata, and the most important the link between the information and the metadata, so that every time we publish data and/or metadata, it's to improve not only the visibility and discovery of the data but also its access.

Different branches of expansion have been defined as future work. The first is continuing the current work to generate a client which will facilitate the implementation of the WPS, both in fixed devices such as a phone. Another line of research is to automatically increase the elements of the metadata generated by using ontologies.

With the presented approach, we move closer to real usage of GIIs, because end users become involved in content provision. Once the barrier of motivation has been overcome, we will be able to benefit from GIIs for effective and efficient information sharing, one of the main goals of GEOSS.

ACKNOWLEDGMENTS

This work was supported in part by the European FP7 Project 226487 EuroGEOSS, the GEOCLOUD project reference IPT-430000-2010-11, subprogram Innpacto 2010, Ministerio de ciencia e innovación and Generalitat Valenciana (Grant ACIF/2012/112).

REFERENCES

Beard, M. K. (1996). “A Structure for Organizing Metadata Collection.” *In Proceedings 3rd International Conference on Integrating GIS and Environmental Modeling. Sante Fe, NM. January 21-26 1996. Santa Barbara, Canada.*

Beltrán, A., Díaz, L., Granell, C. and J. Huerta (2009a). Geocrawler y gvsig: un tándem para la generación automática de metadatos. *In 5th gvSIG User Conference. November 2009, Valencia, Spain.*

Beltrán, A., Díaz, L., Granell, C., Huerta, J. and C. Abargues (2009b). Description and Publication of Geospatial Information, Geoscience and Remote Sensing, Pei-Gee Peter Ho (Ed.), *InTech*.

Beltrán, A., and C. Martín (2010). Generando descripciones de recursos para gvsig desde geocrawler. *In 6th gvSIG User Conference. November 2011, Valencia, Spain.*

Bernard, L., Kanellopoulos, I., Annoni, A., Smits, P. (2005). The European geoportal- one step towards the establishment of a European Spatial Data Infrastructure.Computers, *Environment and Urban Systems*, vol 29, issue 1, pp.15-31.

Budhathoki, N., Bertram, B., Nedovic-Budic, Z. (2008). Reconceptualizing the role of the user of spatial data infrastructure”. *GeoJournal*, vol, 72, pp. 149-160.

Bulterman, D.C.A. (2004). Is it time for a moratorium on metadata? *MultiMedia, IEEE* , vol.11, no.4, pp. 10- 17, Oct.-Dec. 2004

Campbell. T. (2004) Fostering a culture of metadata production. *In GSDI10: Tenth International Conference for Spatial Data Infrastructure,* St. Augustine, Trinidad*.*

Craglia, M., Kanellopoulos, I., Smits, P. (2007). Metadata: Where We Are Now, and Where We Should Be Going - *In: Proceedings of the 10th AGILE International Conference on Geographic Information Science, Aalborg (Denmark), pages 1-7, University of Aalborg (Publ.), Aalborg*

Currier, S., Barton, J., O'Beirne, R. and B. Ryan (2004). Quality assurance for digital learning object repositories: issues for the metadata creation process. *Association for Learning Technology Journal*, 12 (1). pp. 5-20.

Díaz, L., Granell, C., Gould, M. and J. Huerta (2011a). Managing user-generated information in geospatial cyberinfrastructures. *Future Gener. Comput. Syst*. 27(3): 204-314 March 2011

Diaz, L., Martin, C., Gould, M., Granell, C., and M.A. Manso (2007). Semi-automatic metadata extraction from imagery and cartographic data. *In Geoscience and Remote Sensing Symposium 2007. IGARSS 2007. IEEE Internacional, 3051-3052*

Díaz, L. and S. Schade (2011b). GEOSS Service Factory: Assisted Publication of Geospatial Content. *Proceedings of the 14th AGILE International Conference on Geographic Information Science - Advancing Geoinformation Science for a Changing World, Utrecht, 2011. AGILE. Utrecht, The Netherlands, April 2011*

INSPIRE EU Directive (2007). Directive 2007/2/ec of the european parliament and of the council of 14 march 2007 establishing an infrastructure for spatial information in the european community (inspire). *Technical report, Official Journal of the European Union, L 108/1, Volume 50, 2007.*

Jorum (2004). The jisc online repository for [learning and teaching] materials: Jorum scoping and technical appraisal study. *In JORUM. Volume V: Metadata, 2004.*

Masser, I. (2005). GIS Worlds: Creating Spatial Data Infrastructures. *Redlands: ESRI Press.*

Manso, M.A., Nogueras-Iso, J., Bernabe, M.A. and F.J. Zaragasa-Soria (2004). Automatic metadata extraction from geographic information. *In 7th Conference on Geographic Information Science (AGILE 2004), Heraklion, Greece, 2004.*

Nogueras-Iso, J., Barrera, J., Rodríguez, A.F., Recio, R., Laborda, C. and , F.J. Zaragaza-Soria (2009). *Development and deployment of a services catalog in compliance with the INSPIRE metadata implementing rules. SDI Convergence: Research, Emerging Trends, and Critical Assessment. The Netherlands Geodetic Commission (NGC), 2009, pp 21-34.*

Nogueras-Iso, J., Zarazaga-Soria, F.J., Béjar, R., Álvarez, P.J. and P.R. Muro-Medrano (2004). OGC Catalog Services: a key element for the development of Spatial Data Infrastructures, *Computers & Geosciences*, Volume 31, Issue 2, March 2005, pp 199-209, Nogueras-Iso, J., Zarazaga-Soria, F.J., and P.R. Muro-Medrano (2005). *Geographic information metadata for spatial data infrastructures: resources, interoperability and information retrieval*. Springer.

Schut, P. (2007). Opengis web processing service version 1.0.0. OpenGeospatial Consortium., 2008.

Taussi, M. (2007). Automatic production of metadata out of geographic datasets. *PhD thesis, master thesis, Department of Surveying, Helsinki University of Technology, 2007.*

Trilles, S., Gil, J., Díaz, L. and J. Huerta (2011). Un geoprocesamiento para la publicación de metadatos geográficos. *II Jornadas Ibéricas de las Infraestructuras de Datos Espaciales (JIDEE), 2011*

Vlissides J., Helm, R., Johnson, R. and E. Gamma (1995). *Design patterns: Elements of reusable object-oriented software.* Addison-Wesley.

Williams, M., Cornford, D., Bastin, L., Pebesma, E. (2008) Uncertainty Markup Language (UncertML). *OGC Discussion Paper*.

Zarazaga-Soria, F.J., Lacasta, J., Nogueras-Iso, J., Torres, M.P. and P.R. Muro-Medrano (2003). A java tool for creating ISO/FGDC geographic metadata. *Geodaten-und Geodienste-Infrastrukturen-von der Forschung zur praktischen Anwendung. Beiträge zu den Mänsteraner GI-Tagen, 26:27, 2003.*