(Draft) SDI Convergence: Research, Emerging Trends, and Critical Assessment.
The Netherlands Geodetic Commission (NGC), 2009, p. 21-34

Development and deployment of a services catalog in compliance with the INSPIRE metadata implementing rules

J. Nogueras-Iso[1], J. Barrera[1], A.F. Rodríguez[2], R. Recio[1], C. Laborda[3]
, F.J. Zarazaga-Soria[1]

[1]Computer Science and Systems Engineering Department, University of Zaragoza
María de Luna 1, 50018-Zaragoza (Spain)
{jnog,jesusb,rociorm,javy}@unizar.es

[2]National Geographic Institute of Spain
General Ibáñez de Ibero, 3, 28003-Madrid (Spain)
afrodriguez@fomento.es

[3]GeoSpatiumLab
Carlos Marx, 4, local der., 50015-Zaragoza (Spain)
claborda@geoslab.com

Abstract

In order to facilitate the availability of and access to spatial data, Spatial Data Infrastructures must set up a series of services to be reused by their community of users in the construction of different applications and value-added services. One of the key elements to exploit the resources provided by these infrastructures is to facilitate a catalog of services describing the features of the services offered to their users. This article presents the development and deployment of a services catalog within the Spatial Data Infrastructure of Spain, a catalog in compliance with the INSPIRE implementing rules.

Keywords: Metadata, Services, Spatial Data Infrastructures, Services Catalog

1. INTRODUCTION

According to the Global Spatial Data Infrastructure Association Cookbook (Nebert, 2004), “the term Spatial Data Infrastructure (SDI) is often used to denote the relevant base collection of technologies, policies and institutional arrangements that facilitate the availability of and access to spatial data”. SDIs provide the framework for the optimisation of the creation, maintenance and distribution of geographic information at different organisation levels (e.g., regional, national, or global level) and involving both public and private institutions.

From a technical point of view, in order to facilitate the access to and exploitation of spatial data, SDIs must set up a series of services to be reused by their community of users in the construction of different applications and value-added services. In addition, these services must be accessible on the Internet through standardized transport protocols and interfaces established by organisations and standardisation bodies such us OGC (Open Geospatial Consortium), ISO/TC211 (International Organization for Standardization, Technical for Committee Geographic Information/Geomatics) or W3C (World Wide Web Consortium).

Therefore, one of the key elements to exploit the reusability of the resources provided by an SDI is to set up a services catalog that provides metadata describing the features of the services offered to the community of users. The objective of a services catalog is to help users in the task of finding the most appropriate service according to their needs and requirements.

Up to now and as far as implementation of SDI initiatives is concerned, most of the efforts have been devoted to the development of spatial data catalogs. However, catalogs for describing services have received little attention. The objective of this article is to present the development of a services catalog, which has been deployed within the Spatial Data Infrastructure of Spain (Infraestructura de Datos Espaciales de España - IDEE). The Web client application of this services catalog to be integrated within the IDEE geoportal provides an easy and intuitive interface to browse and evaluate the features of services accessible through this SDI. Additionally, it must be noted that this services catalog complies with the most relevant standards and specifications defined at international level, making a special emphasis on the compliance with the metadata implementing rules of the INSPIRE Directive (EC, 2007), a Directive for establishing an Infrastructure for Spatial Information in the European Community.

The rest of this article is structured as follows. Section 2 describes the state of the art in metadata schemas for service description. Section 3 presents the design of the services catalog, including a description of its architecture, the adoption of INSPIRE metadata implementing rules for metadata modelling, and the automatic method proposed to derive metadata from the capabilities information provided by OGC services. Section 4 describes the deployment of the services catalog in the Spain SDI. Finally, this article ends with some conclusions and proposals for future work.

2. STATE OF THE ART IN APPROACHES TO SERVICE DESCRIPTION

Within the context of SDIs, the first approach for service description arose with the definition of the first specifications for OGC services more than ten years ago. For descriptive purposes, all OGC service specifications support a *getCapabilities* operation to obtain service-level metadata (also named as capabilities) describing the content and acceptable request parameters of an OGC service. The XML responses to *getCapabilities* requests include information about service identification (general metadata for discovery such as information as title, abstract, or keywords), service provider, and available operations.

In 2005, ISO/TC211 defined an extension to the ISO 19115 (ISO, 2003a) geographic metadata standard for the description of the specific features of services. This extension has been defined within the context of ISO 19119 standard (ISO, 2005). This standard aims at establishing the foundations of geographic information services. It provides a taxonomy of geographic information services, and it extends the ISO 19115 comprehensive model with specific elements and data types for service descriptors. This proposal has been also adopted by OGC in the definition of one of the profiles of the CSW (Catalogue Services for the Web) protocol binding for catalog services specifications. This profile, called “ISO Metadata Application Profile” (Voges and Senkler, 2007), proposes a combination of the models in ISO 19115 and ISO 19119 as the information model followed by the metadata records to be managed through catalogs.

Recently, the European Commission started a procedure for the adoption of measures to implement the INSPIRE directive as regards metadata (EC, 2008a). These implementing measures define at an abstract level those descriptors that are essential for the discovery of data and services. Besides, these measures are accompanied with non-binding guidelines to establish the mapping between this set of abstract descriptors and the most important metadata standards such as ISO 19115 or Dublin Core (also adopted by ISO as ISO15836 (ISO, 2003b)). In particular, the guideline containing the mapping to ISO 19115 (EC, 2008b) also includes the mapping to the specific service descriptors defined in ISO 19119.

However, the above proposals seem insufficient to satisfy the current needs for building applications following a Service Oriented Architectures (SOA) (Lieberman, 2003), where the syntactic and semantic description of services is highly relevant. Nowadays, the development of the services offered by spatial data infrastructures, and in general the development of services in any type of networked infrastructures, is usually guided by the Web Services Architecture proposed by W3C (Booth et al., 2004). This architecture aims at providing a standard means of interoperating between different software applications (the Web services), running on a variety of platforms and/or frameworks. According to the W3C Web Services Glossary (Haas and Brown, 2004), a Web Service is a software system designed to support interoperable machine-to-machine interaction over a network. Web Services are self-contained, self-describing, modular applications that can be published, located, and invoked across the Web (Doyle and Reed, 2001). The great impact of Web Services has increased the importance of metadata that describe the processing capabilities of services. The details of a Web Service can be published in a catalog, so that a client's (or another service's) request for such a service can lead to the client invoking that service.

The leading and most accepted standard for service syntactic metadata is WSDL (Web Services Description Language). WSDL (http://www.w3.org/TR/wsdl) is a means of describing a service connection (operation signature or binding) for software to connect to it. Service Directory specifications like UDDI (Universal Discovery, Description and Integration of Web Services protocol, http://www.uddi.org/) can use WSDL to express the machine-readable connect to a service. The main disadvantage of WSDL is that it does not have the ability to characterize the semantic capabilities of services. OWL (Web Ontology Language, http://www.w3.org/TR/owl-features/) is one of the options to formalize the semantic description of services, and in particular OWL-S (http://www.w3.org/Submission/OWL-S/), which is an ontology based on OWL to model the features of Web services. OWL-S is used to annotate services with semantic descriptions. Other options for semantic annotation are WSMO (Web Services Modeling Ontology) or WSDL-S (Web Service Semantics). On the one hand, WSMO (http://www.wsmo.org/) provides a conceptual model to describe the semantic aspects of services in order to automate their discovery, invocation and composition. The WSMO infrastructure has defined WSML (Web Services Modeling Language) as representation language. On the other hand, WSDL-S (http://www.w3.org/Submission/WSDL-S/) defines a mechanism to associate semantic annotations to Web services already described with WSDL.

As a conclusion of this state-of-the-art discussion, it can be stated that the current initiatives in the SDI context still tend to use metadata models (e.g., ISO 19119 extension of ISO 19115) that are more focused on the description of services for mere discovery purposes than in providing the means for the automatic invocation and composition of services. However, it can be noticed that there is an increasing consciousness about the need to move towards more interoperable descriptions of services. For instance, OGC has defined an application profile of the CSW protocol (Martell, 2007) that is compatible with the ebXML (Electronic Business using eXtensible Markup Language, http://www.ebxml.org/) infrastructure, which consists in a modular suite of specifications to exchange business messages, conduct trading relationships, communicate data in common terms and define and register business processes. It must be noted that one of the requirements for the services catalog proposed in this article is to be compliant with the INSPIRE implementing rules. Therefore, the metadata model proposed will follow the guidelines for mapping the metadata implementing rules to ISO 19115/19119 (see section 3.2).

3. DESIGN OF THE SERVICES CATALOG

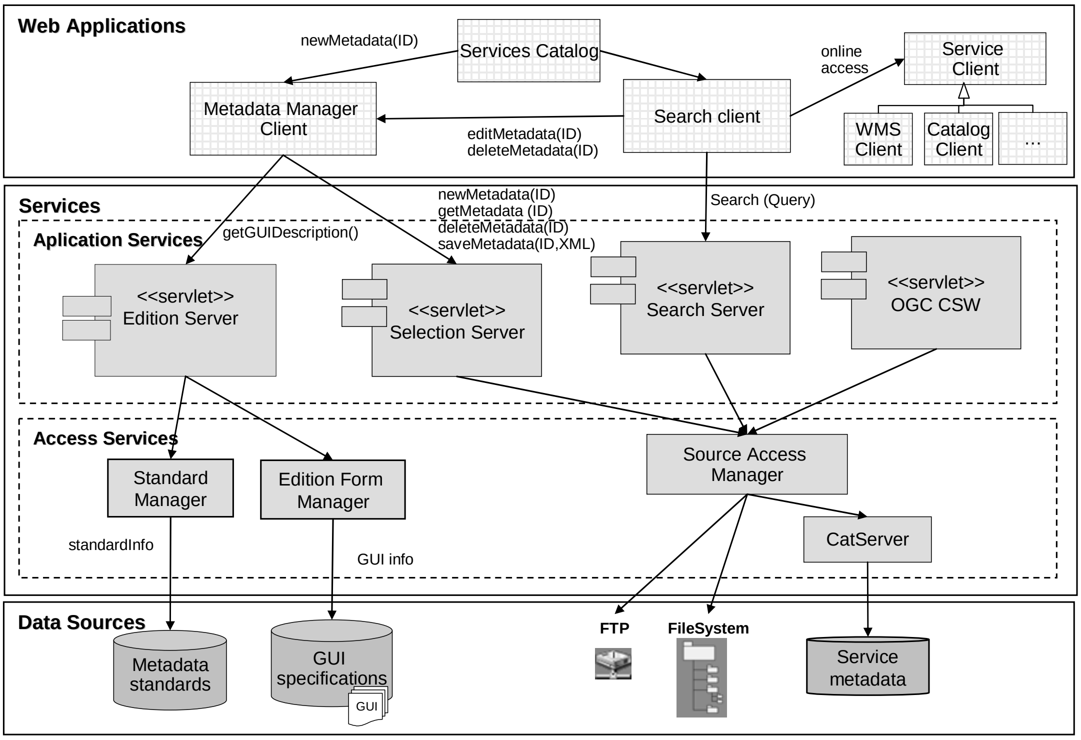
This section describes the architecture of the services catalog. Section 3.1 presents the architecture. Then section 3.2 presents the metadata model followed by the metadata records managed through the services catalog. Since one of the main requirements of the services catalog presented in this article is the compliance with the INSPIRE implementing rules, this section describes the implications in the modelling and encoding of service metadata. Finally, section 3.3 describes the process followed to derive service metadata from the capabilities information obtained through OGC services.

3.1. Architecture

The services catalog has been structured following a multi-layer architecture model, where the different components have been grouped in different levels according to their functionality with respect to data access, processing, or interaction with the final user. In particular, three architectural layers have been distinguished (see Figure 1): (1) the *Data Sources* layer includes the different storage repositories used by the services catalog, (2) the *Services* layer integrates the components in charge of the access to data (*Access Services*) and their processing (*Application Services),* and *,* (3) the *Web Applications* layer consists of the components that interact with the end user, either receiving its requests, or providing the results generated in the lower layer.

As it can be observed in Figure 1, the main component in the top level of the architecture (*Web Application* layer) is the *Services Catalog* application, which is the application that a final user can access through a Web browser. This application provides its functionality thanks to the use of two components called *Metadata Manager Client* and *Search Client*. Despite the fact that they are embedded in a single application, they could be the base of independent applications. Whereas the aim of the first component is the creation, elimination and modification of services metadata; the second component is in charge of querying the catalog and showing the results to the user. This *Services Catalog* application also enables online connections with the services returned by the catalog thanks to the integration of generic clients compliant with the most common OGC specifications. In order to facilitate the integration of components in this layer, they have been developed using the *Google Web Toolkit (GWT)* technology. This technology, sponsored by Google, provides a set of free software tools to build web applications with AJAX using Java as programming language.

Figure 1: Services Catalog general architecture



With respect to the *Services* level, we can distinguish two categories of components: an *Application Services* category including the components that carry out tasks of data processing; and an *Access Services* category integrating the components that deal with data and information retrieval. Within the *Application Services* category we can find four components using the Java servlets technology: *Edition Server*, *Selection Server*, *Search Server*, and *OGC CSW*. The first two servlets are designed to give support to the *Metadata Manager Client* application. *Edition Server* provides a machine-readable definition (in XML format) for the Graphic User Interface (GUI) of the service metadata edition forms to be displayed by *Metadata Manager Client*. *Selection Server* provides management operations (i.e. insert, update and delete operations) to update the contents of the metadata repository. The *Search Server* servlet, invoked by the *Search Client* application, provides query and present operations. It processes the restrictions found in client queries and returns a list of results satisfying these restrictions. The list of results may be optionally grouped and sorted according to different criteria. And finally, the *OGC CSW* component offers a standardized interface to the services catalog according to the OGC specifications (Nebert et al., 2007). In particular, this component implements the CSW (Catalog Services for the Web) protocol binding to allow the communication between catalog clients and servers over HTTP.

The second category of the *Services* layer includes three main components: *Source Access Manager, Standard Manager*, and *Edition Form Manager*. *Source Access Manager* is the component in charge of retrieving metadata from the final storage device. It provides an abstraction layer over the different types of metadata sources facilitating a uniform access mechanism to the components in the higher layers of the architecture. That is to say, thanks to *Source Access Manager* other components do not need to worry about the storage device, which may be either a file system accessed via FTP, or something more complex such as a XML metadata database implemented on top of a relational database (for instance, see the *CatServer* system described by Tolosana-Calasanz et al., 2005). This component is accessed by three components in the *Application Services* category: *Selection Server* invokes it to perform management operations on the metadata repository; *Search Server* uses this component to find the metadata records satisfying the user queries; and *OGC CSW* provides a standardized wrapper to access the operations offered by this component.

In order to understand the functionality of the other two components that belong to the *Access Services* category, it is necessary to know the mechanism used for metadata edition. This mechanism is the one used in version 4.0 of the *CatMDEdit* desktop application (Zarazaga-Soria et al., 2003; Nogueras-Iso et al., 2008). Using a machine readable definition of metadata standards and a set of rules for GUI layout, this mechanism generates dynamically the edition forms to modify the contents of metadata records in conformance with the correspondent metadata standard. With this purpose in mind, *Edition Server* (in the *Application Services* layer) invokes the *Standard Manager* component to recognize and recover the definition of the metadata standard followed by the metadata record(s) to be updated. Then, *Edition Server* invokes the *Edition Form Manager* component to generate the GUI description of the forms to be displayed by the *Metadata Manager Client*. The *Edition Form Manager* creates the GUI description as a result of applying GUI production rules to the elements contained in the metadata standard. Each GUI production rule specifies the most appropriate GUI component (e.g., text field, list, text area, choice) for the data type of each metadata element. In the case of the *Services Catalog*, metadata conforms to the INSPIRE metadata implementing rules and their mapping to ISO 19115 and ISO 19119 metadata standards. But this automatic mechanism for metadata edition could be applied to develop other Web metadata editors in conformance to other metadata standards.

Finally, the bottom layer of the architecture (*Data Sources* layer) consists of three different data sources: the data source where the information about metadata standards is stored; the data source with the rules for GUI layout; and the data source for storing service metadata. All these data are encoded in XML format and, although all the data sources used in the IDEE Services Catalog are file systems, the final storage device for service metadata could have been accessed through more complex mechanisms such as FTP or CatServer (accessing an XML database on top of a relational database).

3.2. Adoption of the INSPIRE metadata implementing rules for metadata modelling

As mentioned in the introduction, one of the main requirements of this services catalog is to be compliant with the INSPIRE implementing rules, which will dictate the development of national SDIs in Europe in the following years. Therefore, the metadata model proposed will follow the guidelines for mapping the metadata implementing rules to ISO 19115/19119 (EC, 2008b), which is the more mature existent guideline to translate the INSPIRE implementing measures into a particular metadata standard.

As an example of this metadata model, Figure 2 shows the ISO 19119 elements for service identification that must be included in the metadata model to comply with the INSPIRE implementing rules (the correspondent INSPIRE descriptors are shown on the right side). Additionally, it must be noted that we need to include in this metadata model all the mandatory elements of ISO 19119 despite the fact that some of them have no equivalent in the set of abstract descriptors contained in the INSPIRE metadata implementing rules.

Figure 2: ISO 19119 elements for service identification in compliance with INSPIRE metadata implementing rules, extracted from (EC, 2008b)

+ citation [1] : CI_Citation	
+ title [1] : CharacterString	Resource title (See 2.2.1)
+ date [0..*] : CI_Date	See note 1
+ date [1] : Date	Date of publication (See 2.6.2)
+ dateType [1] : CI_DateTypeCode	publication
+ date [0..1] : CI_Date	See notes 1 and 2
+ date [1] : Date	Date of last revision (See 2.6.3)
+ dateType [1] : CI_DateTypeCode	revision
+ date [0..1] : CI_Date	See Note 1 and 3
+ date [1] : Date	Date of creation (See 2.6.4)
+ dateType [1] : CI_DateTypeCode	creation
+ abstract [1] : CharacterString	Resource abstract (See 2.2.2)
+ pointOfContact [1..*] : CI_ResponsibleParty	See 3.5.1
+ descriptiveKeywords [1..*] : MD_Keywords	
+ keyword [1..*] : CharacterString	Keyword value (See 2.4.1)
+ thesaurusName [0..1] : CI_Citation	Originating controlled vocabulary (See 2.4.2)
+ resourceConstraints [1..*] : MD_Constraints	See 3.6
+ serviceType [1] : GenericName	Spatial data service type(See 2.3.2)
+ couplingType [1] : SV_CouplingType	Mandated by ISO 19119. See Note 8
+ containsOperations [1..*] : SV_OperationMetadata	Mandated by ISO 19119
+ operationName [1] : CharacterString	Mandated by ISO 19119. Default value is unknown
+ DCP [1..*] : DCPList	Mandated by ISO 19119. Default value is WebServices
+ connectPoint [1..*] : CI_OnlineResource	Mandated by ISO 19119.
+ linkage [1] : URL	Mandated by ISO 19119. See Note 5
+ extent [1..*] : EX_Extent	See Note 4
+ geographicElement [1..*] : EX_GeographicBoundingBox	Geographic bounding box (See 2.5.1)
+ westBoundLongitude [1] : Decimal	
+ eastBoundLongitude [1] : Decimal	
+ southBoundLatitude [1] : Decimal	
+ northBoundLatitude [1] : Decimal	
+ temporalElement [0..*] : EX_TemporalExtent	See Note 6
+ extent [1] : TM_Primitive	Temporal extent (See 2.6.1)
+ operatesOn [0..*] : MD_DataIdentification	Coupled resource (See 2.2.6 and Note 7)

With respect to the encoding of service metadata in XML format, it must be noticed that we have followed the guidelines established by the technical specification ISO/TS 19139 (ISO, 2007). This technical specification defines the way to translate the UML models proposed in ISO 19115 (and other related standards) into an XML syntax. In particular, the syntax that has been used is the one proposed in the XML-Schemas (http://schemas.opengis.net/iso/19139/20060504/srv/) accompanying the "ISO Metadata Application" profile (Voges et al., 2007) of the OGC catalog services specification, which adopts ISO 19115/19119 as metadata information model.

3.3. Automatic generation of metadata from capabilities

Metadata are the key element to allow the discovery and reusability of services provided by an SDI, however manual metadata creation is a hard and tedious task. In order to make this process easier, this services catalog includes an automatic method to derive metadata from the capabilities information returned by the services that comply with OGC specifications. The responses to *getCapabilities* operations are usually implemented as an XML file that contains three main sections (the name of these sections may vary according to the service type or the capabilities specification version):

- *ServiceIdentification* (or Service in WMS specification): General metadata including identification information such as the title, the abstract or the keywords.
- *ServiceProvider* (or *ContactInformation* in WMS specification): Metadata about the organisation that provides the specific service instance.
- *OperationsMetadata* (or *Capability* in WMS specification): Metadata about the operations and related abilities specified by the service.

The automated method proposed is able to obtain metadata from services compliant with five OGC service specifications: WMS (Web Map Service), CSW (Catalog Services for the Web), WFS (Web Feature Service), WCS (Web Coverage Service) and WPS (Web Processing Service). This automated method is based on a mapping between the capabilities defined in these specifications and the set of descriptors defined in the INSPIRE metadata implementing rules. This mapping can is shown in Table 1.

As shown in Table 1, the *getCapabilities* response of a WMS has a different structure from the capabilities in contrast to other OGC services. However, there is an easy mapping between this specific schema and the more uniform schema used in other OGC specifications, which is defined in (Whiteside, 2007).

Table 1: Mapping between OGC Capabilities Specifications and INSPIRE metadata profile

INSPIRE metadata	WMS 1.3.0	CSW 2.0.2	WFS 1.1.0	WCS 1.1.0	WPS 1.0.0
Resource Title	Service > Title	ServiceIdentification> Title	ServiceIdentification> Title	ServiceIdentification> Title	ServiceIdentification> Title
Resource Abstract	Service > Abstract	ServiceIdentification > Abstract	ServiceIdentification > Abstract	ServiceIdentification > Abstract	ServiceIdentification > Abstract
Responsible organisation	Service > ContactInformation	ServiceProvider > ServiceContact	ServiceProvider > ServiceContact	ServiceProvider > ServiceContact	ServiceProvider > ServiceContact
Keyword	Service > KeywordList	ServiceIdentification > Keywords	ServiceIdentification > Keywords	ServiceIdentification > Keywords	ServiceIdentification > Keywords
	WMS	*CSW*	*WFS*	*WCS*	*WPS*
	infoMapAccessService	*humanCatalogueViewer*	*infoFeatureAccessService*	*infoCoverageAccess Service*	
Conditions for access and use	Service > AccessConstraints	ServiceIdentification> AccessConstraints	ServiceIdentification > AccessConstraints	ServiceIdentification > AccessConstraints	ServiceIdentification > AccessConstraints
Spatial data service type	*view*	*discovery*	*download*	*download*	*transformation*
Resource locator	Created from the getCapabilities url	Created from the getCapabilities url	Created from the getCapabilities url	Created from the getCapabilities url	Created from the getCapabilities url
containsOperations *	Capability > Request	OperationsMetadata > Operation	OperationsMetadata > Operation	OperationsMetadata > Operation	OperationsMetadata > Operation

*: Only in INSPIRE metadata implementing rules based on ISO 19115 and ISO 19119

Additionally, it must be noted that, apart from extracting some information from a *getCapabilities* response, this automatic method can also infer some information which is not directly present in a capabilities response. It adds three elements for the classification of services: the spatial data service type according to the service types in annex D.3 of the INSPIRE metadata implementing rules; a keyword to identify the service type according to the ISO 19119 classification of geographic information services; and another keyword corresponding to the acronym of the OGC specification. For instance, in the case of a WMS, this method will fill: the *spatial data service type* element with the value *"view"*; and two keywords with the values *"infoMapAccessService"* (according to the ISO 19119 classification) and *"WMS"* (according to OGC classification).

4. DEPLOYMENT OF THE SERVICES CATALOG IN THE IDEE

According to the architecture described in the previous section a first prototype of the IDEE Services Catalog has been developed (http://www.idee.es/IDEE-ServicesSearch/ServicesSearch.html?locale=en). This prototype allows search and access to the description of services subscribed to the IDEE. Before this prototype, this description of services was updated manually within the IDEE services directory (http://www.idee.es/show.do?to=pideep_catalogo.EN), a set of static Web pages to report the services offered by the IDEE member organisations. Additionally this prototype facilitates the online connection with OGC Web Map Services (WMS).

Figure 3: Search criteria, result presentation and online connection with the services

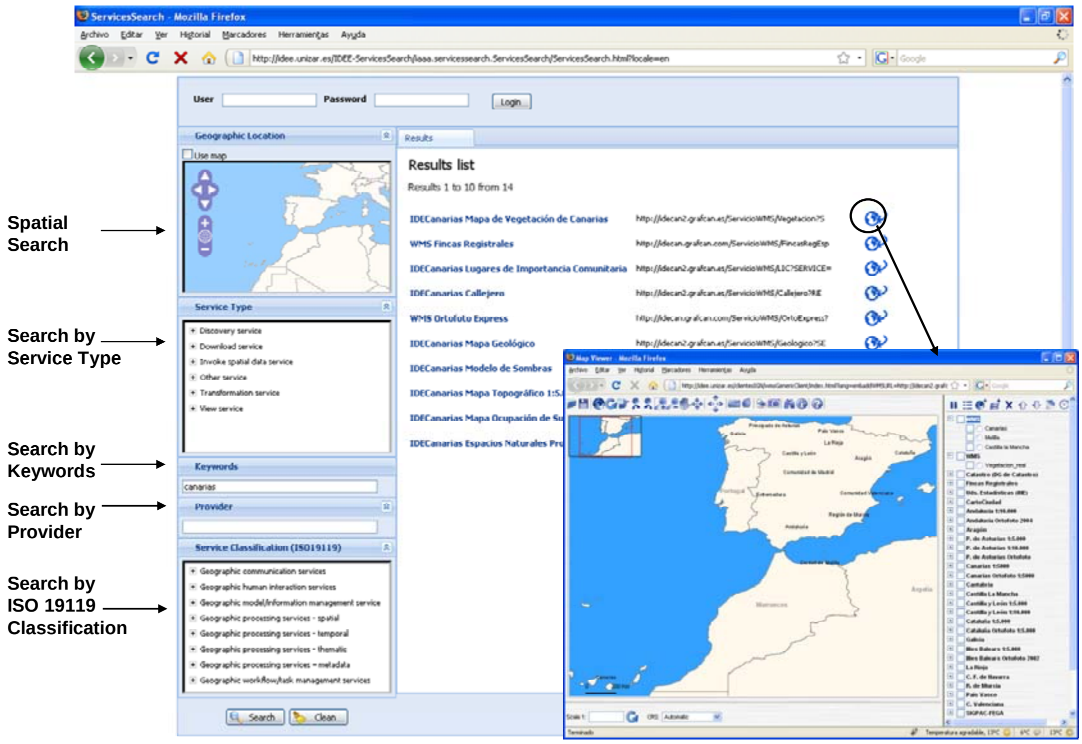
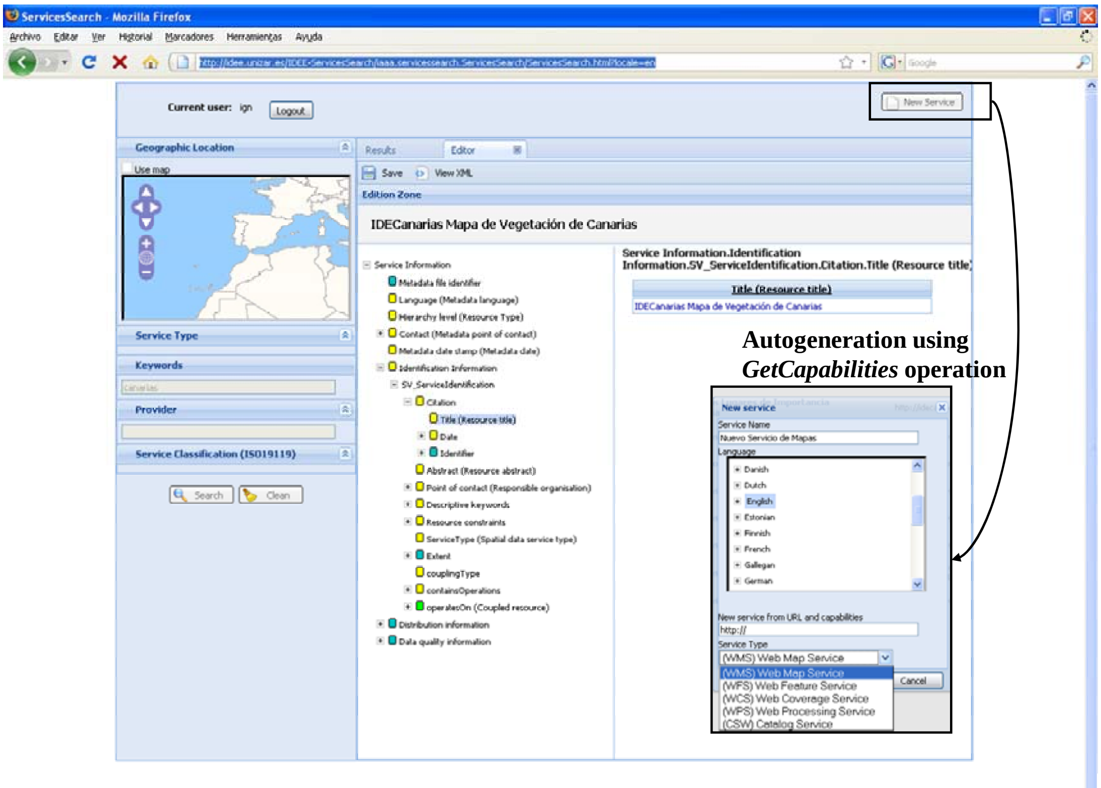


Figure 3 shows a screenshot of the prototype. With respect to the construction of search requests, the query specification form (see left part of Figure 3) allows to establish different restriction criteria such as the geographic extension of the data provided by the service, the keywords, the service type, or the services provider. Once the results have been presented, the application enables the online edition of metadata items (see Figure 4). Besides, an important feature of the application is that it allows the automated generation of metadata for services compliant with OGC specifications by means of their c*apabilities* information. This functionality is provided when editing the information of a new service (see *New Service* button in Figure 4). Finally, it must be also noted that the online connection with the services returned by the catalog can be activated through the map icon in the list of results. For instance, Figure 3 (right part) shows how an OGC WMS Client is launched to add the new layers served by the WMS discovered through the catalog.

Figure 4: Online metadata edition



As regards the contents that are accessible through this prototype, we must mention that we have established a process to compile all the service URLs contained in the original static directory (335 services altogether) and apply the automatic method described in section 3.3 to convert the *getCapabilities* response into a metadata record compliant with INSPIRE and ISO 19115/19119 metadata models. Table 2 shows some statistics derived from this upload process.

Table 2: Statistics about the upload of contents to the services catalog (Nr = number)

<table>
<tr><th colspan="2"></th><th colspan="24">Subscribed services</th></tr>
<tr><th rowspan="3">Service category</th><th rowspan="3">Service</th><th colspan="2" rowspan="3">Nr_ss</th><th colspan="2" rowspan="3">% (=Nr_ss/ tot_ss)</th><th colspan="16">Automatic creation</th><th colspan="4">Manual creation</th><th colspan="4">Duplicated cases</th></tr>
<tr><th colspan="2" rowspan="2">Nr_ac</th><th colspan="2" rowspan="2">% (=Nr_ac/ Nr_ss)</th><th colspan="4">Correct</th><th colspan="4">Manual review</th><th colspan="2" rowspan="2">Nr_mc</th><th colspan="2" rowspan="2">% (=Nr_mc /Nr_ss)</th><th colspan="2" rowspan="2">Nr_dc</th><th colspan="2" rowspan="2">% (=Nr_dc/ Nr_ss)</th></tr>
<tr><th colspan="2">Nr_acc</th><th colspan="2">% (=Nr_acc /Nr_ac)</th><th colspan="2">Nr_acr</th><th colspan="2">% (=Nr_acr /Nr_ac)</th></tr>
<tr><td>Discovery</td><td>CSW</td><td colspan="2">5</td><td colspan="2">1,5</td><td colspan="2">1</td><td colspan="2">20,0</td><td colspan="2">0</td><td colspan="2">0,0</td><td colspan="2">1</td><td colspan="2">100,0</td><td colspan="2">4</td><td colspan="2">80,0</td><td colspan="2">0</td><td colspan="2">0,0</td></tr>
<tr><td>View</td><td>WMS</td><td colspan="2">300</td><td colspan="2">89,6</td><td colspan="2">238</td><td colspan="2">79,3</td><td colspan="2">200</td><td colspan="2">84,0</td><td colspan="2">38</td><td colspan="2">16,0</td><td colspan="2">40</td><td colspan="2">13,3</td><td colspan="2">22</td><td colspan="2">7,3</td></tr>
<tr><td rowspan="2">Download</td><td>WCS</td><td rowspan="2">29</td><td>15</td><td rowspan="2">8,7</td><td>4,5</td><td rowspan="2">15</td><td>3</td><td rowspan="2">51,7</td><td>20,0</td><td rowspan="2">7</td><td>0</td><td rowspan="2">46,7</td><td>0,0</td><td rowspan="2">8</td><td>3</td><td rowspan="2">53,3</td><td>100,0</td><td rowspan="2">14</td><td>12</td><td rowspan="2">48,3</td><td>80,0</td><td rowspan="2">0</td><td>0</td><td rowspan="2">0,0</td><td>0,0</td></tr>
<tr><td>WFS</td><td>14</td><td>4,2</td><td>12</td><td>85,7</td><td>7</td><td>58,3</td><td>5</td><td>41,7</td><td>2</td><td>14,3</td><td>0</td><td>0,0</td></tr>
<tr><td>Transfor-mation</td><td>WPS</td><td colspan="2">1</td><td colspan="2">0,3</td><td colspan="2">0</td><td colspan="2">0,0</td><td colspan="2">0</td><td colspan="2">0,0</td><td colspan="2">0</td><td colspan="2">0,0</td><td colspan="2">1</td><td colspan="2">100,0</td><td colspan="2">0</td><td colspan="2">0,0</td></tr>
<tr><td>Invoke</td><td></td><td colspan="2">0</td><td colspan="2">0,0</td><td colspan="2">0</td><td colspan="2">0,0</td><td colspan="2">0</td><td colspan="2">0,0</td><td colspan="2">0</td><td colspan="2">0,0</td><td colspan="2">0</td><td colspan="2">0,0</td><td colspan="2">0</td><td colspan="2">0,0</td></tr>
<tr><td colspan="2" rowspan="2">TOTAL auto./manual/dupl=</td><td colspan="8">Correct/manual=</td><td colspan="2">207</td><td colspan="2">81,5</td><td colspan="2">47</td><td colspan="2">18,5</td><td colspan="8"></td></tr>
<tr><td colspan="2">335 (=tot_ss)</td><td colspan="2">100,0</td><td colspan="2">254</td><td colspan="2">75,8</td><td colspan="8"></td><td colspan="2">59</td><td colspan="2">17,61</td><td colspan="2">22</td><td colspan="2">6,57</td></tr>
</table>

The services subscribed to the IDEE have been grouped in Table 2 according to the spatial data service type defined in annex D.3 of the INSPIRE metadata implementing rules (EC, 2008a): *Discovery*, *View*, *Download*, *Transformation* and *Invoke*. And within a category, there is specific row for each OGC compliant service in this category: services compliant with CSW (Catalogue Service for the Web) protocol binding of the OGC Catalogue Service Specifications within the *Discovery* category; services compliant with the OGC WMS (Web Map Service) specification within the *View* category; services compliant with the OGC WCS (Web Coverage Service) and WFS (Web Feature Service) specifications within the *Download* category; and services compliant with the OGC WPS (Web Processing Service) specification within the *Transformation* category. For each category and OGC specification the table shows statistics about the number and percentage of the services subscribed to the catalog including details about the metadata creation process:

- Column *Nr_ac* (and its associated percentage) under the *"Automatic creation"* header counts the number of records that were automatically created from the *capabilities* responses. And within this category it is possible to distinguish two subcategories: number of records that were generated without any problem (see columns under "*Correct"* header); and number of records that required some kind of manual review (see columns under *"Manual review"* header) due to old versions or capabilities or problems in XML files (incomplete or invalid files). Although not included in the table, it can be derived that 61.8% of records were automatically created without manual review (207 records out of 335 services).
- Column *Nr_mc* (and its associated percentage) under the *"Manual creation"* header counts the number of records that were manually created. This column represents the number of service whose URL was invalid. In this case, metadata records were filled manually with the information stored in the static directory.
- Column *Nr_dc* (and its associated percentage) under the *"Duplicate cases"* header counts the number of records that had been introduced twice.

As a conclusion from the analysis of the figures in this table, we have detected some typical problems to solve while working with services catalogs. Firstly, information becomes rapidly out-of-date: 17.61% of service URLs included in the service directory was not valid anymore. Secondly, interoperability problems between different versions of OGC services persist: 18.5 % of automatically created metadata records were reviewed manually because their capabilities response could not be parsed. Thirdly, there is no control about duplicates: 6.57% of original service descriptions had been introduced twice (but with slight differences). The advantage of using a services catalog with respect to a static directory is that it is easier to establish an automatic mechanism to report these problems.

5. CONCLUSIONS AND FUTURE WORK

This work has presented the development and deployment of a first prototype of Services Catalog in the SDI of Spain, replacing the static content offered until now by IDEE under the name of services directory. Thanks to this new prototype, it is possible to provide online services such as metadata creation, dynamic indexing and metadata search, as well as the connection with generic clients, which allow verifying the correct operation of services.

Thanks to the upload of metadata in the services catalog using the content of the original static service directory, we have detected some problems that could be common to other SDI initiatives: information becomes rapidly out-of-date (it is difficult to maintain service URLs up-to-date); interoperability problems between different versions of OGC services persist (capabilities responses are still very heterogeneous); and there is little control about duplicates. Although the use of a services catalog does not solve these problems directly, it is easier to establish an automatic mechanism to report these problems and mark the records that are not valid at the moment. Additionally, we must remark the low quality, in general, of the metadata that are usually found as a response to the *getCapabilities* request of OGC-compliant services. For instance, most of these metadata do not even contain a citation to the organisation that has set up the service on the internet.

As further lines of this work, we plan to integrate more generic clients to facilitate the online connection with other service types apart from WMS services. Currently, the connection with OGC-compliant catalogs through generic CSW clients is being developed. With respect to the automatic metadata generation, we also plan to infer more metadata elements from the information contained in a capabilities response. For instance, we expect to infer the extent of a WMS from the bounding box of the map layers provided by this service. Last, new functionalities of the application have been proposed to facilitate the monitoring of the services registered in the catalog.

To sum up, we think that Services Catalogs will constitute a key element for SDI development. They are essential to facilitate the reusability of services, and in the future, they will probably receive more requests than the data catalogs. We must not forget that if SDIs are based on a Service Oriented Architecture, the services are the basic concept around which an SDI will be conceived and structured. Therefore, the Services Catalogs will play the main role in the exploitation of SDIs.

ACKNOWLEDGEMENTS

This work has been partially supported by the Spanish government through the projects “España Virtual” (ref. CENIT 2008-1030) and TIN2007-65341; the National Geographic Institute of Spain (IGN); and GeoSpatiumLab S.L.

REFERENCES

Booth D., Haas H., McCabe F., Newcomer E., Champion M., Ferris C. and D. Orchard (eds.) (2004). *Web Services Architecture. W3C, W3C Working Group Note 11 February 2004*. http://www.w3.org/TR/2004/NOTE-ws-arch-20040211/.

Doyle A. and C. Reed (ed.) (2001). *Introduction to OGC Web Services*. An OGC® White Paper, Open Geospatial Consortium Inc. http://portal.opengeospatial.org/files/?artifact_id=14973

EC (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)*. European Commission (EC), Official Journal of the European Union. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2007:108:0001:0014:EN:PDF

EC (2008a). *Commission Regulation (EC) No 1205/2008 of 3 December 2008 implementing Directive 2007/2/EC of the European Parliament and of the Council as regards metadata (Text with EEA relevance)*. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:326:0012:0030:EN:PDF

EC (2008b). *INSPIRE Metadata Implementing Rules: Technical Guidelines based on EN ISO 19115 and EN ISO 19119, v. 1.0*. European Commission (EC), Joint

Research Centre. http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/metadata/MD_IR_and_ISO_20081219.pdf

Haas H. and A. Brown (eds.) (2004). *Web Services Glossary. W3C, W3C Working Group Note 11 February 2004*. http://www.w3.org/TR/2004/NOTE-ws-gloss-20040211/

ISO (2003a). *Geographic information - Metadata. ISO 19115:2003*. International Organization for Standardization (ISO).

ISO (2003b). *Information and documentation - The Dublin Core metadata element set. ISO 15836:2003.* International Organization for Standardization (ISO).

ISO (2005). *Geographic information - Services. ISO 19119:2005.* International Organization for Standardization (ISO).

ISO (2007). *Geographic information - Metadata - XML schema implementation. ISO/TS 19139:2007*. International Organization for Standardization (ISO).

Lieberman J. (2003). *OpenGIS Web Services Architecture. OpenGIS Discussion Paper OGC 03-025*. Open Geospatial Consortium Inc. http://portal.opengeospatial.org/files/?artifact_id=1320

Martell R. (ed.) (2007). *CSW-ebRIM Registry Service - Part 1: ebRIM profile of CSW (Version: 1.0.0)*. OGC 07-110r2, Open Geospatial Consortium Inc. http://portal.opengeospatial.org/files/?artifact_id=27092

Nebert D. (ed.) (2004). *Developing Spatial Data Infrastructures: The SDI Cookbook* v.2.0. Global Spatial Data Infrastructure. http://www.gsdi.org

Nebert D., Whiteside A. and P. Vretanos (eds.) (2007). *OpenGIS - Catalogue Services Specification (version: 2.0.2)*. OGC 07-006r1, Open Geospatial Consortium Inc. http://portal.opengeospatial.org/files/?artifact_id=20555

Nogueras-Iso J., Barrera J., Gracia-Crespo F., Laiglesia S. and P.R. Muro-Medrano (2008). "Integrating catalog and GIS tools: access to resources from CatMDEdit thanks to gvSIG". *Proceedings 4th International gvSIG conference*. 3-5 December 2008, Valencia. http://www.jornadasgvsig.gva.es/fileadmin/conselleria/Documentacion/4asJornadas/articulos/Sesion3/CatMDEdit_gvSIG.pdf

Tolosana-Calasanz R., Portolés-Rodríguez D., Nogueras-Iso J., Muro-Medrano P.R. and F.J. Zarazaga-Soria (2005). "CatServer: a server of GATOS". *Proceedings of AGILE Conference 2005, Estoril, Portugal,* pp. 359–366 ISBN 972-8093-13-6.

Voges U. and K. Senkler (2007). *OpenGIS Catalogue Services Specification 2.0.2 - ISO Metadata Application Profile*. OGC 07-045, Open Geospatial Consortium Inc. http://portal.opengeospatial.org/files/?artifact_id=21460

Whiteside A. (ed.) (2007). *OGC Web Services Common Specification (Version: 1.1.0 with Corrigendum 1).* OGC 06-121r3. Open Geospatial Consortium Inc. http://portal.opengeospatial.org/files/?artifact_id=20040

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". In: *Geodaten- und Geodienste-Infrastrukturen - von der Forschung zur praktischen Anwendung. Beiträge zu den Münsteraner GI-Tagen* 26./27. Juni 2003. Volume 18., Münster, Germany, IFGIprints, pp. 17–30.