SDIGER: Experiences and identification of problems on the creation of a transnational SDI

J. Nogueras-Iso, M.Á. Latre, R. Béjar, P.R. Muro-Medrano F.J. Zarazaga-Soria,

Depto. de Informática e Ingeniería de Sistemas
Universidad de Zaragoza (España)
jnog@unizar.es, latre@unizar.es, rbejar@unizar.es, prmuro@unizar.es, javy@unizar.es

Abstract: *SDIGER is a pilot project on the implementation of the Infrastructure for Spatial Information in Europe (INSPIRE), funded by the Statistical Office of The European Communities (EUROSTAT), which aims at demonstrating the feasibility and advantages of the solutions for sharing spatial data and services proposed by the INSPIRE position papers and to estimate the costs and to find the problems and obstacles of implementing interoperability-based solutions on the basis of real cases. SDIGER consists in the development of a Spatial Data Infrastructure (SDI) to support access to geographic information resources concerned with the Water Framework Directive (WFD) within an inter-administration and cross-border scenario that involves: two countries, France and Spain; and, the two main river basin districts at both sides of the border, the Adour-Garonne basin district and the Ebro river basin district. The objective of this paper is to present the project and its objectives, making special emphasis on the feedback provided by the development of this project and the problems associated with the creation of a transnational Spatial Data Infrastructure.*

INTRODUCTION

SDIGER is a pilot project on the implementation of the Infrastructure for Spatial Information in Europe (INSPIRE) [1]. This project has been funded by the European Commission through the Statistical Office of the European Communities (Eurostat), contract number "2004 742 00004" for the supply of informatics services in the various domains of the Community Statistical Programme. The objectives fixed by Eurostat for this project are three fold. Firstly, it will serve to test and demonstrate the feasibility and advantages of solutions for sharing spatial data and services, observing the principles and standards proposed by the INSPIRE position papers in 2002 and their interoperability-based approach. Secondly, it is useful to acquire experience in implementing interoperable solutions and develop processes able to be reused when INSPIRE is put into operation. And thirdly, it can help to estimate the costs of implementing interoperability-based solutions on the basis of real cases, together with the problems, obstacles which might be encountered during the subsequent large-scale implementation of INSPIRE.

The "call for tender" for this project required the cross-border application to be focused on an environmental subject. The SDIGER project that was then proposed consists in the development of a Spatial Data Infrastructure (SDI) to support access to geographic information resources concerned with the Water Framework Directive (WFD) [2] within an inter-administration and cross-border scenario that involves: two countries, France and Spain; and, the two main river basin districts at both sides of the border, the Adour-Garonne basin district, managed by the Water Agency for the Adour-Garonne River Basins (*L'Agence de l'Eau Adour-Garonne*) and the Ebro river basin district, managed by the Ebro River Basin Authority (*Confederación Hidrográfica del Ebro*).

This project is being developed by a consortium consisting of the following entities: IGN France International (*Institut Géographique National France International*), the National Geographic Institute of France (*Institut Géographique National*), the National Centre for Geographic Information of Spain (*Centro Nacional de Información Geográfica*), and the University of Zaragoza (*Universidad de Zaragoza*), together with experts from University Jaume I. Additionally, this consortium counts on the help of the following collaboration entities: the National Geographic Institute of Spain (*Instituto Geográfico Nacional*), the Water Agency of Adour-Garonne (*Agence de l'Eau Adour-Garonne*), the Ebro River Basin Authority (*Confederación Hidrográfica del Ebro*), the Regional Direction of the Ministry of Environment for the Midi-Pyrenees region, and the GIS-ECOBAG association. As it can be observed, these entities (most of them public institutions) are the main providers of the topographic and hydrographic data in the cross-border area.

The paper is structured as follows. Next section provides a more detailed explanation of the objectives of the pilot project, organized by activities. Then, the following section explains the problems found in the development of the different activities. These problems are classified in 5 subsections: problems found in the definition of a useful application scenario, problems found in metadata related activities, problems related with data specifications, problems related with the set-up of services, problems found in the definition of the geoportal, and general management problems. Finally, current state of the project and the next steps to be taken are defined in the Conclusions section.

ACTIVITIES OF THE PROJECT

SDIGER is a two-year project, divided by Eurostat in the "call for tender" in a set of activities orientated to face the problems that may arise in the large-scale implementation of INSPIRE. The following subsections detail these activities. All of them, except for the last one, correspond to the first year of the project.

Definition of a cross-border scenario

The SDIGER project consists in the development of a Spatial Data Infrastructure (SDI) (see figure 1) to support access to geographic information resources concerned with the Water Framework Directive (WFD) within an inter-administration and cross-border scenario that involves: two countries, France and Spain; and, the two main river basin districts at both sides of the border, the Adour-Garonne basin district, managed by the Water Agency for the Adour-Garonne River Basins (*L'Agence de l'Eau Adour-Garonne*) and the Ebro river basin district, managed by the Ebro River Basin Authority (*Confederación Hidrográfica del Ebro*).

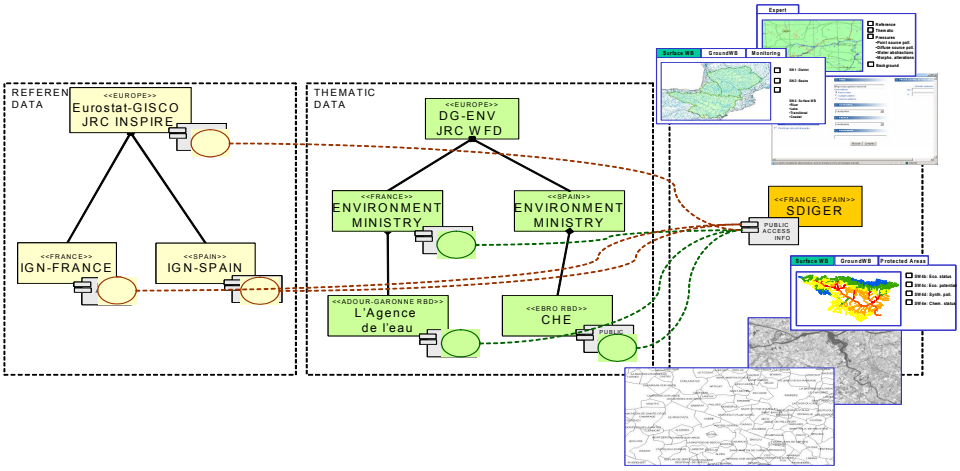

Figure 1. Architecture of the SDIGER SDI.

The area covered by this SDI project is particularly interesting because although most of the Adour and Garonne river basins lay in French territory and Ebro river basin lay in Spanish territory, some stream and river headwaters are located in the other country. This is the case, for instance, of the Garonne river source, which is located in Spain and managed by the Ebro River Basin Authority, and of the Irati river headwaters, an Ebro river tributary which, on the contrary, is located in France and managed by the Water Agency for the Adour-Garonne River Basins. Cross-border information is, thus, of great importance for each of the Basin Authorities in order to assure that the Water Framework Directive requirements are fulfilled in each of the river basin districts. Additionally, this cross-border area includes several protected areas included within Natura 2000, the network of protected areas in the European Union.
Within this scenario, two applications are going to be developed following the INSPIRE principles and proposed architecture:

- WFD Reporting. The WFD introduces a new approach to data and information collection and reporting. This use case proposes to use INSPIRE principles for fulfilling the reporting requirements of the member states to the European Commission. In particular data required by articles 3 and 5 of the WFD are taken as an example to implement the reporting mechanisms in an INSPIRE compliant way, i.e. the required data and information will be directly accessed within a spatial data infrastructure.
- Water abstraction request use case. In France and Spain, the use of both surface and groundwaters for private purposes requires an authorisation given by an authority. The administrative process for a water abstraction request requires users to provide an application form specifying the characteristics of the water abstraction point, water use and water discharge point. The objective of this use case is to provide users with some guidance, data and documentation needed to follow the administrative process of the water abstraction request. Stakeholders will be provided with a web application that will enable them to specify the request details (including location of the geographical elements related to the request). Users will be presented a report indicating whether the request is possible according to the legislation of the affected country, tabular information needed for the application form and that depends on the spatial data provided by the user, a printed map with the spatial data provided by the user to be included with the request, and an orientative, not legally binding report about the request compliance with WFD criteria.

Details of both applications can be found at the "Application Scenario" document available at the SDIGER portal (http://sdiger.unizar.es).

Metadata related activities

In this area we can distinguish three main tasks: the definition of metadata profiles customized to the type of resources to be described in the project, the development of a metadata edition tool, and the own task of creating metadata contents.

Definition of metadata profiles

Within the SDIGER project, three metadata profiles have been developed: a metadata profile for geographical data mining, a generic metadata profile for INSPIRE for assessing and using geographical data and a metadata profile for the Water Framework Directive. The standards ISO19115 [3] and Dublin Core [4] have been the basis for the development of these profiles. In general, a metadata application profile should have in their objectives to facilitate the metadata creation process by: providing an specification of the subset of the metadata standard terms (choosing the ones that are more relevant for a specific domain); offering guidelines for filling in the fields of the metadata according with the specific domain; and providing specific keyword controlled lists, thesaurus and ontologies for the context where the application profile is used. For the development of these metadata profiles, several standards and initiatives in the context of spatial data infrastructures have been taken into account additionally:

- The proposal for the INSPIRE (Infrastructure for SPatial Information in Europe) [1]. Chapter II of the proposal makes explicit references to the information that metadata should contain to describe spatial resources.
- The Draft Technical Report on "Geographic Information – Metadata – European core metadata set" developed by the Working Group 5 (Spatial Data Infrastructures) of CEN/TC287 [5].
- The draft spatial application profile of Dublin Core proposed by the European Standardization Committee (CEN) [6].
- The guidelines for metadata included within the document "Guidance Document on Implementing the GIS Elements of the Water Framework Directive" [7]. The first application domain tackled by the INSPIRE proposal is environmental data and WFD data is directly related with environmental aspects. Additionally, the WFD metadata profile should observe and take into consideration this guidance document.
- The core metadata recommendations for the Spanish Spatial Data Infrastructure [8].

From a conceptual point of view, the metadata profiles should be organised hierarchically putting Dublin Core in the top level because it is the most general metadata standard. This metadata could be specialised with the application profile that has been proposed within the scope of this project for geographical data mining. More details should be included in the INSPIRE application profile (it ought to take into account the effect of the multicultural and multilingual heterogeneities in the creation of metadata and provide guidelines to avoid this heterogeneity) and this one should be refined with the WFD application profile (specialised in water resources).
The "Dublin Core Metadata Application Profile for geographical data mining" is based mainly in the Dublin Core Spatial Application Profile [6], though several modifications have been done, such as including new elements (*provenance* and *rightsholder*), new refinements (*distance* and *equivalentScale.denominator* to the element *spatialResolution*) and changes in the encoding scheme (drop of TGN and inclusion of EUROVOC, AGROVOC and INSPIRE_SpatialThemes). The "Generic metadata profile for INSPIRE for assessing and using geographical data" is based on an application profile of ISO19115 [3] with the objective of describing the spatial resources in accordance with the proposal for the INSPIRE directive, focusing on the Chapter II, which makes explicit references to the information that metadata should contain to describe spatial resources. Finally, the "Metadata profile for the Water Framework Directive" is an ISO19115 profile mainly based on the guidelines for metadata included within the document "Guidance Document on Implementing the GIS Elements of the Water Framework Directive" [7], making special emphasis on the description of data quality.

Metadata edition tool

Within the SDIGER project, an open source metadata management tool with support to the aforementioned metadata profiles has been provided. This application is based in the CatMDEdit tool [9] that was previously developed by the authors. This tool, multiplatform and multilingual (Spanish, English, French, Polish and Czech) , includes as main functionalities:

- Support for edition and visualisation of metadata entries according to different ISO19115 and Dublin Core profiles.
- A thesaurus management tool, allowing the management of thesauri supported by a thesauri database. The main functions of this tool are: creation/deletion/modification of thesauri; edition/visualisation of terms in a hierarchical and alphabetical structure; and import/export from/to text files in different formats. For the SDIGER project, and new multilingual thesauri support (in Spanish, French and English) has been included:

UNESCO (about 5.000 terms), AGROVOC (about 11.600 terms), EUROVOC (about 7.200 terms) and GEMET.

- An XML Import/Export tool, enabling the exchange of metadata records in XML format conforming to different standards such as CSDGM, ISO19115 and Dublin Core. Besides, the tool also facilitates more readable presentations of metadata records in HTML format. The import process has two possibilities: it can either create a new metadata record, or update the content of a metadata record previously selected from the metadata repository.
- A Metadata Generation tool, enabling the semi-automatic generation of metadata for several types of resources. For instance, this tool is able to obtain descriptive information from ESRI *shapefiles*. Additionally, the tool can also extract metadata corresponding to the relational structure of tabular sources (e.g. Excel, Access, Oracle…).
- A Contact Management Tool, allowing reusing contact information (e.g. name, address, telephone…), which is needed in several metadata fields. Thanks to this tool, the contact information about a person is only inserted once and used whenever it is required.

Metadata creation

The metadata catalogue was performed for core topographic data provided by National Mapping Agencies, thematic data provided by Water Agencies and some of the European data provided by EuroGeographics, JRC and Eurostat. Behind providing the metadata to be integrated into the geoportal for description of the data, this procedure served to test and provide the comments concerning both the profiles and the CatMDEdit tool. The results of this survey will be present in the study report to be provided to the Eurostat by the end of this year and the metadata records (see table below) are actually available on the geoportal of the project.

Organization (Data Provider)	**Count**
IGN Spain	23700
CHE	35
IGN France	3
AEAG	18
DIREN	8
SMEAG	12
Eurostat (NUTS)	1
JRC (IMAGE2000)	2000
EEA (NATURA2000)	1
EuroGeographics	2
Total	**25780**

Table 1. Count of records in the metadata catalog

Multilingual access portal to data and services.

The Geoportal of the SDIGER project is already accessible at http://sdiger.unizar.es and providing access to data and services produced and served by the institutions being partner or collaborator of the SDIGER consortium. This portal is structured in four main sections:

- General information about the project. This section provides details about the project (objectives, partners, results …), useful links and any other kind of information that could be interesting for the project audience.
- Generic services. This sections offers access to the three basic services considered: geodata catalog search application, gazetteer application and geographic information visualization application.
- Use case applications. This section provides the applications which implement the two use cases described in the application scenario deliverable.
- Private area. This section provides access (with login and password) to a restricted area where the documents and the deliverables are stored.

Additionally, this portal offers these capabilities with no language restrictions (Spanish, French and English). In order to develop such a multilingual portal, two issues must be solved: the GUI internationalization and a cross-language information retrieval model. With respect to the first issue, the GUI components (labels, buttons, value lists, …) must be displayed in the language specified by the user. For these requirements, Java internationalization techniques and XML technologies (including XSLT) have been used to dynamically internationalize the software components, load web pages contents stored as XML documents, and apply the appropriate style sheets to display the required portal style and with the appropriate language for text labels. And as regards the second issue, a cross-language information retrieval model will be proposed. There are a lot of geographic information resources that are catalogued using only one language, but users that make their queries in one language may be interested in existent resources that have been

described in another language. The user is more interested in the resource (map, image or multimedia resource in general) rather than in the metadata describing it. Thus, catalogs must provide users with the mechanisms facilitating the multilingual search without forcing cataloguing organizations to describe their resources in all the possible languages. Next activity explains the multilingual resources used to facilitate this cross-language retrieval.

Multilingual aspects of the application

French and Spanish are the official languages of the two countries directly involved in the project. Besides offering data and services in these two languages, an English version of the geoportal will be also available to facilitate accessibility to users not familiar with these other two languages. Therefore, multilingual resources like multilingual thesauri (GEMET [10], UNESCO [11], EUROVOC[1] and AGROVOC[2]) and multilingual gazetteers will be used to facilitate the creation of metadata and the development of ergonomic search interfaces for data and service catalogs [12].
Additionally, although the multilingual thesauri facilitate the cross-language information retrieval, it is also important to help the user understand the content of metadata records that may have been written in a language different from the user query language. In that case, it would be desirable to translate on-line the records obtained as a result of the query by means of a machine translation service. For that purpose, SYSTRANLinks from SystranSoft (http://www.systransoft.com) has been selected as the machine translation service used in this project.

Creation of a common object-oriented data model for the data used in the application

As the SDIGER project is especially focused on thematic water data, this activity has given priority to the harmonisation of data models related with the water resources, in particular the ones required by the WFD.
As a common schema for interoperable access to national data was required for the application scenario use cases, the national layers in both countries have been analysed for modelling harmonisation. That is the reason why it was decided to make the data model for the thematic data concerned by the application scenario, then add other core data also involved into the application and make a study concerning task for harmonisation of the all the data, which are integrated into the geoportal.

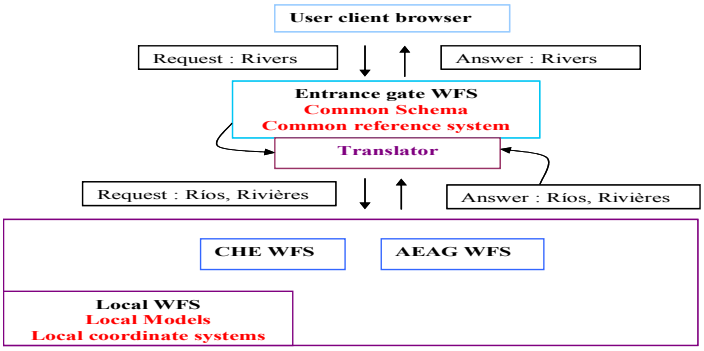

Figure 2: GiMoDig approach for on-the-fly conversions

The process followed to fulfil this harmonisation of data models has been the following:

- Analysing the data inventoried and identifying matching data
- Taking into account other data models experiences: GISCO, Eurosion, data model proposed by GIS Working Group for the WFD Common Implementation Strategy
- Developing of a common data object oriented model (using UML)
- Establishing mappings between original data structures at each institution and the common data model to guarantee a seamless data integration and use.

Finally, we are studying now the approaches to allow the conversion of data on the fly. One of these approaches is the analysis of the use of an on-the-fly schema translation tool from the GIMODIG project, which is based on the use of XSLT documents (see *Figure 2*). Additionally, other more pragmatic approaches are also taken into account, such as the creation of database views to transform national data into data compliant with the harmonized models.

Configuration of servers

This activity consists in the configuration of the servers for accessing the data and services covered by the application according to the ISO and Open Geospatial Consortium standards. The types of services that are offered are basically three: discovery of geographic data through the use of geographic data catalogs; map on-line visualization on the web

[1] http://europa.eu.int/celex/eurovoc
[2] http://www.fao.org/agrovoc

by means of the use of Web Map Servers (WMS); and a selected access to data through Web Feature Servers (WFS), mainly for the web application described in the next activity.

Internet application

This activity, nowadays in progress, consists in the development of web applications fulfilling the use cases proposed by the first activity and making profit of data and services established in the previous activities.

Business plan for implantation of INSPIRE at the European level

This activity, still under development, will produce a report based on the previous activities will be written to provide elements to identify problems, solutions and the costs of using configurations commensurate with the European scale of INSPIRE. Thanks to the analysis of costs, problems and successful experiences in the development of previous activities, we will be able to define an accurate SDI reference model, which will include solutions to problems found at the pilot project and fit to the features of INSPIRE objectives at the European level. We will provide a business plan making special emphasis on the estimation of costs, which will be classified as follows:

- Creation of the data model. Based on the costs of developing the models developed within this pilot project (mostly focused on water reference data and hydrological resources), we will estimate the extension of these models to other application domains.
- Configuration of the catalogue. We will estimate the costs involved in the configuration of a catalog services server for prototypical institution that decides to form part of INSPIRE network of nodes.
- Configuration of a server according to OGC standards. Apart from a catalog services server, we will also make an estimation of the costs for installing the main services to provide on-line visualization and access to data, i.e. the cost of set-up of a WMS, WFS, WCS and a gazetteer service.
- Costs of multilingualism management. We will include the costs of acquiring and integrating multilingual resources to facilitate multilingual capabilities of geoportals and services.
- Costs of developing the application excluding the aforementioned costs. Here, we will estimate the budget of developing customized applications built upon integrated access to the data and services offered by an SDI.
- Costs of re-engineering the data if necessary. Given the experience of harmonizing the data provided by the institutions involved in this project, we will estimate the average cost that must be addressed to adapt data to INSPIRE standards and common models.
- Costs of licences for software and data; cost of hardware. The costs due to software licenses and hardware needed for this project will be similar to the costs needed for the configuration of servers in a node participating in the European SDI aimed by INSPIRE.
- Costs for maintaining servers and availability of services on the Internet. Given the experience of this project and previous projects developed by the partners of the SDIGER consortium, maintenance costs and issues concerned with 24/7 availability will be estimated.
- Other costs not covered by the above list. Thanks to the experience of developing this pilot project, new aspects and problems not included initially in the tender specification will be taken into account.

Maintenance for the second year period

The main goal of the second year project will be the assurance of the services and applications functionality, ensuring that average up-time of all servers allows a correct use of the infrastructure. In order to fulfil this, it will be necessary to plan security mechanisms and to design contingency plans. The work will be focused in four main areas: version control in each node, specification of a contingency plan, definition of backup procedures, and to establishment of a personal plan to guarantee response time. The node version control should take into account and to inventory: the characteristics of the hardware that supports each node services (processor, memory, hard disks,…), operating system version, base software versions (database management system, web server, application server,…), application software versions and instances (web applications, portals, services) and data. Additionally, it should be necessary to have a dependency map between data, services and applications and to have an installation map that shows the distribution of the whole contents. The main objective of the node version control is to provide the information necessary to “clone” a node from the backups as soon as possible.

PROBLEMS FOUND

This section provides details of the problems found during the development of the SDIGER project. Identifying these problems is one of the main objectives of this project. In some cases, it could be possible to provide solutions to be exported to other contexts. In most cases, solutions implemented are just “ad hoc” solutions that should be redefined

over new rules and recommendations provided by the INSPIRE initiative. The problems found have been classified into five sections: problems found in the definition of a useful application scenario, problems found in metadata related activities, problems related with data specifications, problems related with the set-up of services, problems found in the definition of the geoportal, and general management problems.

Problems related with the definition of a useful application scenario

The first task of this project has been the specification of an application scenario. Maybe the most relevant problem has been to find a useful use-case. The status of most SDIs is that they have just created just geoportals with quite attractive map viewers and a search service for data holdings based on metadata. However, project supervisors realize that this is only a very first step that did not fulfil their expectations. It is needed to prove that SDIs solve real problems in an easier way than developing stand-alone applications from the scratch. Our first intention was to set-up a set of servers able to display and make searches on data required for the Water Framework Directive. However, this was not enough. European Commission wanted to verify that forcing member states to create WFD data, real useful applications could be derived. That is the reason why we finally proposed the "water abstraction use case" that, on the one hand, solves a real need and improves an administrative process, and on the other hand makes profit of the data required by the WFD. In addition, the "WFD reporting use case" will show how to combine two new requests to the member states imposed by the European Commission: reporting described in articles 3 and 5 of the WFD and INSPIRE requirements.

Problems related with metadata activities

The project will provide the three specific application profiles detailed previously and a tool that will be able to facilitate the creation of metadata according with these profiles. We would like to detail two of the difficulties found in the definition of the metadata profiles:

- ISO19115 is a complex metadata standard. Although trying to define simple and customised profiles for INSPIRE and WFD force you to create also complex profiles.
- Although establishing the correspondence between ISO19115 and Dublin Core, these standard are still quite far from each other. We have proposed a Dublin Core Application Profile for Geographical Data Mining as a first step for description of resources. But we need also to define a complex crosswalk to transform this first step Dublin Core metadata into ISO19115 for full description.

As a consequence of these difficulties, the use of the metadata management tool has shown two main problems:

- Creation of guidelines must be defined. Despite that the understanding of the semantics of a metadata element is sill quite subjective.
- Automatic metadata creation tools are needed as metadata creation is usually done after data creation.

Once the tools for creating and maintaining metadata are available, next step is to create them. This process produces a very interesting and complex to solve subset of problems. Maybe the most relevant could be the following:

- There is still no culture about metadata creation. Metadata creation is still a project-driven approach in public institutions. That is to say, metadata is created only when public institutions commit themselves to an SDI-like project. Quite the opposite, they should encourage metadata creation as another task performed together with data creation.
- The detail of metadata is quite heterogeneous in different institutions, together with the contents of the items filled. It is necessary to provide standard ontologies to be used by the creators in order to be able to use the same concepts in the metadata creation processes.
- The metadata creator does not usually take into account that this metadata will be later searched through a catalog. Examples like errors in the name of data providers may derive in the fact that you can not found.

Problems related with data specifications

At the beginning, the tender of the project was too ambitious as it was aiming data harmonization for too many layers: all national and European layers were liable to be harmonized. Several European projects have been created for that purpose and most of them have not been very successful. Finally, the scope was narrowed to the layers related with hydrology and the Water Framework Directive. However, although narrowing the scope, this directive does not provide a common model suitable for every Competent Authority and several aspects related to the WFD are not present (such as pressure and impact data). Thus, each Water Agency or member state defines its own model based on the WFD GIS Working Group recommendations. Additionally, these national and local models are not fully stable as the WFD is just in the initial phase of implementation. On the other hand, several problems have been found in the harmonization of data models. Maybe the most relevant are:

- The techniques for data harmonization are not mature enough.
- When appropriate metadata about the data is not available, the understanding of the data semantics of data sources to be harmonised is very difficult and suggestions for harmonization can be erroneous, particularly

with thematic data (where some expertise is needed for their understanding) and data in different languages (where some knowledge of the languages involved is needed).
- Where data has not been created following a predefined model (such as pressure data for the WFD), differences among datasets to be harmonised are so big that the harmonisation is hard and can only be performed at a very high level.
- Conversion of data on the fly is not very efficient.

Problems related with services

Definition of portrayal services

During the development of this project, it has been necessary to install and put on line several Web Mapping Services. In addition, these ones and another ones provided by the project partners have been include in the possibilities offered by WMS clients. The most important problems found during this work could be the next ones:

- Existence of security restrictions (firewalls). It has been necessary to develop and install WMS-Proxies in order to allow the access to services provided by some public institutions that have not their WMS available at Internet level. In addition, it will be necessary to access to all the WMS using a relay system in order to be able to control the use of them (server systems offered by third parties can not be configured for proving log information).
- Different Spatial Reference Systems (SRS): Multiple WMS can only interoperate if they share at least one SRS as a common denominator. Lack of support of specific or common projection systems may prevent the visualisation of more than one WMS at a time.
- Multilingualism: How to treat multilingualism, e.g. in legends derived from different WMS or in the textual information given on a selected feature in response to a GetFeatureInfo request.
- Coherent use of scale hints: The WMS specification gives a general rule but no explicit reference on how to treat different scales. There is no provision of information concerning the scale of a specific dataset and this, may result in unpredictable results. Thus maps of very different scales may be combined in senseless ways or, on the contrary, it may not be possible to combine maps with almost similar scales.
- Consistent principles of cartographic styling: As most of the current WMS do not fully support the Styled Layer Descriptor specifications, a user defined cartographic styling of the advertised map layers is not yet feasible, thus leading to colourful patchworks of adjacent map layers served from different WMS or resulting in visually hardly to interpret overlays composed out of the selected layers coming from various WMS.

Definition of a distributed catalog

The initial objective of the project was to be able to access to the catalogs provided by the partners of the project (IGN Spain, IGN France, CHE and AEAG) within their SDIs. The solutions could be to accessing them in an on-line distributed catalog manner, or by using harvesting technologies. Finally, the project provides only one own catalog where the metadata provided by the four institutions have been loaded. The reason for choosing this alternative is based in the following problems:

- Only one of the four institutions had its catalog on-line. The others had some metadata created, or even no metadata.
- The definition of distributed catalog that not has matured specifications. There are two work lines in this area: the real on-line distributed catalog, and the use of harvesting techniques. In both cases, there are many technical problems that mast to be solved.
- In addition, there is a lack of real implementations compliant with well-established specifications. To be able to develop systems with the capacity for connecting with on-line catalogs that satisfy current standards is a hard work. OGC has provided a good job trying to accord a standard for catalogs. Nonetheless, and even though great progress has been achieved due to these initiatives, the fact is that at present there are far too many catalogue services' implementation profiles and standards available, even just inside OGC: Z39.50, CORBA-IIOP, SRW or CSW (SOAP, XML-Post and KVP).

Definition of a transnational gazetteer

The situation at the beginning of the project was the following: IGN Spain provided the gazetteer that could be adapted to the OGC recommendations (but with an extra effort), and the inventory in France has shown that there is no real gazetteer (but there is a toponyms database suitable to be transformed into one). This work was performed by IGN France staff especially for the SDIGER project. The resulting gazetteer containing more than 94,000 items for French

Midi-Pyrénées region and complaint with specifications provided by Spanish partners is integrated into the geoportal. The following table shows the number of geographic names accessible through the Gazetteer service.

Institution	Content	Nr
IGN -Spain	- geographic names - administrative names - hydronyms	361581
CHE Ebro	- water points	52902
IGN-France	- hydronyms - populated places - non-populated places - oronyms - other	94107
Total		**508590**

Table 2. Count of records in the Gazetteer

In addition, the initial objective of the project was to be able to offer a distributed gazetteer service. This initial objective found the same problems than the distributed catalog. Furthermore, to be able to offer a centralised gazetteer it was necessary to solve two additional problems:

- There are no approved specifications for gazetteer, with the consequent problems for the exchange of data.
- Problems in the different typology of features. The IGN-F provides 74 different types, the IGN-S provides 50 different types, and the CHE provides 20 different types. These make a set of 144 different types. After studied them, the current gazetteer has 129 different types.

Problems related with the definition Geoportal

We have distinguished two main kind of problems related with the development of the GeoPortal. The first one is related with the web application that will be provided. These problems can be separated into the ones related with it specification (see problems with the application scenario) and the problems of performance. In this way, it is necessary to mention the use of Web Feature Servers. They seem to be the appropriate specification for the exchange of feature data on the fly, but in practices they result to be quite inefficient.
The other main set of problems is related with the multilingualism of the portal. This set is integrated with the following main difficulties:

- Difficulties for the development of Web portal infrastructure with capabilities for internationalization.
- Difficulties for the translation of contents of the projects in 3 different languages.
- Difficulties for the internationalization of legends in Web Mapping viewers. The Web Mapping Service does not take into account the management of names of layers in different languages.
- Difficulties in the presentation of metadata into the 3 different languages. Metadata has been created using one specific language but should be able to be presented in the rest of languages used for the project. An automatic translation tool has been used (Systran) but we are not sure about the results provided by this tool.
- Difficulties in the data modelling and harmonization, in particular with thematic terms and data.

General problems

In addition, a set of important difficulties related with the conception and nature of the project have been found during its development. They have been classified into the following three main sets.

The dream of reusing the infrastructure created for other SDIs

The initial planning of SDIGER as an SDI built upon existing SDIs is still a dream. We have faced that the SDIs developed by the public institutions involved in the project are still in a very initial state, and in many cases with a no clear future and/or objectives. Several basic services have been created specifically for this project.

The management of a transnational/transinstitutional project

This project involves a big set of institutions and administrations. The nature of the contractor (EUROSTAT from the European Commission) and the importance and the repercussion of the results of this project should give them an incentive to get involved with extreme effort and interest. Unfortunately, the development of this project is suffering difficulties to have real contribution from collaborator entities not really engaged in it. In addition, many problems related with the communication among the partners involved in the project have been identified.

Data policies

Finally, another set of problems identified are the ones related with the policies of the institutions. This kind of problems is especially relevant in the case of data policies. Some partner institutions are not allowed to provide public access to their data, not even for display. This has force the development of special services and the use of restricted areas.

CONCLUSIONS

This paper has presented SDIGER, a two-year pilot project on the implementation of INSPIRE, funded by Eurostat, that aims to test, estimate the costs and identify the problems of applying the solutions for sharing spatial data and services proposed by the INSPIRE position papers in 2002. The base for achieving these objectives is by developing a SDI to support access to geographic information resources concerned with the WFD in a cross-border scenario that involves the Adour-Garonne and Ebro river basin districts. The work done till now has shown a set of problems that have been classified into five categories: problems found in the definition of a useful application scenario, problems found in metadata related activities, problems related with data specifications, problems related with the set-up of services, problems found in the definition of the geoportal, and general management problems. At the end of the year 2005, the SDI with the additional applications will be available after solving these problems. However, the main objective of this project, as has been mentioned before, is not to have this functionality available. The objective is identify areas that should be clarified by the INSPIRE recommendations and rules, and to provide a useful tool for future project development in order to be able to provide better plans for them.

ACKNOWLEDGMENTS

This work is funded by the European Commission through the Statistical Office of the European Communities (Eurostat), contract number “2004 742 00004” for the supply of informatics services in the various domains of the Community Statistical Programme. Additionally, this work has been partially supported by the Spanish Ministry of Education and Science through the project TIC2003-09365-C02-01 from the National Plan for Scientific Research, Development and Technology Innovation.

BIBLIOGRAPHY

1. Commission of the European Communities (CEC), 2004. Proposal for a Directive of the European Parliament and of the Council establishing an infrastructure for spatial information in the Community (INSPIRE). COM(2004) 516 final, 2004/0175 (COD).
2. Oficial Journal (OJ) of the European Union, 2000. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. The EU Water Framework Directive - integrated river basin management for Europe. L 327, 22/12/2000 pp. 0001-0073 (2000)
3. Geographic information - Metadata. ISO 19115:2003, International Organization for Standardization, 2003.
4. Information and documentation - The Dublin Core metadata element set. ISO 15836:2003, International Organization for Standardization, 2003.
5. Draft Technical Report on “Geographic Information – Metadata – European core metadata set”. Document TC 287 WI 00287XXX , WG5, 2005.
6. Zarazaga-Soria, F. J., Nogueras-Iso, J., Ford, M. “Dublin Core Spatial Application Profile”. CWA 14858, CEN/ISSS Workshop - Metadata for Multimedia Information - Dublin Core, September 2003. Accesible at: http://www.cenorm.be/cenorm/businessdomains/businessdomains/isss/cwa/cwa14858.asp
7. Vogt, J. (ed.), 2002. “Guidance Document on Implementing the GIS Elements of the Water Framework Directive”, European Communities.
8. “SGTNEM_2005_01: Núcleo Español de Metadatos”. Infraestructura de Datos Espaciales Española, Consejo Superior Geográfico (Ministerio de Fomento), 2005.
9. Zarazaga-Soria, F.J., Lacasta, J., Nogueras-Iso, J., Torres, M.P., Muro-Medrano, P.R., (2003). A Java Tool for Creating ISO/FGDC Geographic Metadata, *Geodaten- und Geodienste-Infrastrukturen – von der Forschung zur praktischen*, pp. 105-118.
10. European Environment Agency (EEA), 2001. GEneral Multilingual Environmental Thesaurus (GEMET), version 2.0. European Topic Centre on Catalogue of Data Sources (ETC/CDS), http://www.mu.niedersachsen.de/cds/.
11. UNESCO, 1995. UNESCO Thesaurus: A Structured List of Descriptors for Indexing and Retrieving Literature in the Fields of Education, Science, Social and Human Science, Culture, Communication and Information. United Nations Educational, Scientific and Cultural Organization (UNESCO) Publishing, Paris. http://www.ulcc.ac.uk/unesco/.
12. Nogueras-Iso, J., Zarazaga-Soria, F.J., Lacasta, J., Tolosana, R., Muro-Medrano, P.R., 2004. Improving multilingual catalog search services by means of multilingual thesaurus disambiguation. In 10th European Commission GI&GIS Workshop, ESDI: The State of the Art, Warsaw, Poland.