

GEOSPATIAL WEB SERVICE INTEGRATION AND MASHUPS FOR WATER RESOURCE APPLICATIONS

C. Granell^a *, L. Díaz^a, M. Gould^a

^aCenter for Interactive Visualization, Department of Information Systems, Universitat Jaume I, E-12071 Castellón, Spain - (carlos.granell, laura.diaz, gould)@uji.es

KEY WORDS: Web Services, Hydrological Applications, Web Processing Services, Earth Imagery Exploitation, Internet/Web, Interoperability

ABSTRACT:

Water resources management is critical due to its importance for human lives. In this context, we have developed a web-based system for hydrological applications that integrates geospatial processing services and mapping mashups to provide an interactive and user-friendly geoportals interface to the configuration and execution of scientific models. The server side provides hydrological model logic through a library of distributed geospatial processing services that prepares and presents all geospatial data –satellite imagery, cartography, digital elevation models, and sensor measurements– necessary for running two popular river runoff models. The geoportals’s client side facilitates catalogue searching for appropriate geospatial data, interacts with the geoprocessing services according to the hydrological model parameters, and displays the results as a web mapping mashup using the Google Maps API and data retrieved by different geospatial web services to provide quick feedback about the status and behaviour of the hydrological model, in order to assist and improve decision making.

1. INTRODUCTION

Hydrologists traditionally have been consumers and collectors of huge amounts of heterogeneous data needed in order to run sophisticated hydrological models providing a physically based representation of the hydrological processes occurring in a given watershed. These models require a wide variety of datasets such as measurements of ground temperature and precipitation, stream gauge measures, satellite imagery representing the snow coverage area of the watershed, digital elevation models, locations of the weather stations, the geographic boundary of the watershed, and many other variables and parameters related to the physical characteristics of the watershed. Hydrologists need then to bring together all these disparate datasets to then try to interpret the resulting runoff predictions and ultimately to improve environmental decision-making. In many cases scientists still work individually on concrete datasets using a chain of different applications such as desktop GIS, spreadsheets and remote-sensing tools, and manually exchanging intermediate results among them (via emails, etc.). From a user’s perspective this chain of tasks is tedious, time consuming and inefficient, and often leads to incomplete or inaccurate results.

Recent advances in web services technology allow these scientists to improve workflow via indirect access to distributed hydrological models and their related datasets (instead of requiring huge data downloads) and also to process them remotely. The goal pursued here is one of facilitating the connection of hydrologists and other scientists to their data, models, and support services and to provide a distributed, scalable yet less complex approach to their workflow. Our solution relies on creating and integrating geospatial services for data processing combined with mapping mashup technology for data visualization.

Most geospatial applications that have migrated from desktop to web are based on a distributed model composed of independent and specialized geospatial web services, linked by international standards-based interfaces (Alameh, 2003). Geospatial web services have evolved to become an efficient alternative for interoperable, modular and distributed GIS applications, such as in this specific case of water resource applications. Furthermore geospatial web services are beginning to exploit Spatial Data Infrastructures (SDI) and the recent possibilities for service chaining in order to build customized and scalable web applications supporting the most common requirements and needs of spatial information users (Lemmens et al., 2006). It is expected that as the number of online web services increases so too will sources of spatial and non-spatial dataset repositories accessed via traditional web service interfaces and protocols. Adding further to the offer of web services is the cadre of mapping mashup APIs and tools that offer increased possibilities for user customization.

We have designed and developed a concrete prototype for water resource management that integrates geospatial processing services, heterogeneous datasets, and the Google Maps API (Jones, 2007) for mapping mashups, among other open source technologies, to provide interactive visual synthesis and exploration of scientific data generated using distributed geospatial processing services.

The remainder of this paper is structured as follows. The next section introduces some basic concepts needed for the remaining sections. Section 3 describes the system architecture and its main components. We demonstrate our architecture in a water resource application scenario in Section 4. Finally, section 5 concludes the paper.

* Corresponding author.

2. CONTEXT

2.1 Spatial Data Infrastructure and Geospatial Web Services

Spatial Data Infrastructure (SDI) comprises a set of policy and standards activities promoting creation of a geospatial information infrastructure to assist diverse user communities to collect, share, access and exploit geo-referenced information resources (Masser, 2005). Traditional discovery and visualization-based SDI is evolving to a more intermediate service-based vision in which geospatial services are used not only to access geospatial data, but also to transform them and process them, often in service chains. Many geospatial services implement interfaces defined by the Open Geospatial Consortium (OGC), such as Web Map Service (WMS), Web Feature Service (WFS), and Web Coverage Service (WCS), etc.

The recent OGC Web Processing Service (WPS) specification (Schut, 2007) provides an interface for accessing more complex services by exposing functionality to, or by wrapping existing off-line services as web services. The WPS interface provides three methods to provide the functionality of a certain geoprocessing service. First the *getCapabilities* method, common in other OWS services (WMS, WFS, etc.), returns the nature of the processes offered. A WPS defines input and output parameters of a process in a very detailed way by providing a *describeProcess* method. Finally, the *execute* method invokes one of the processes of the geoprocessing service. The basic operational unit of the OGC WPS is the notion of process: a spatial operation, with inputs and outputs of a defined type. This means that a WPS instance may offer one or various operations or processes just as normal web services do (Papazoglou, 2008). In our opinion, OGC WPS can and should go beyond providing only geoprocessing routines because it can be an interesting specification within which to wrap both spatial and non-spatial processing services, leading to increased interoperability between OWS services and general purpose web services.

2.2 Mapping Mashups

The term mashup is normally used to designate web applications that are a collage of services with a common user interface that integrates diverse data sources, and that provide solutions to very specific and narrow problems. Mapping or geographical mashups are becoming a viable alternative to developing geospatial visualization tools from scratch for which the required flexibility is difficult and time-consuming to achieve (Wood et al., 2007). *Mashup* is considered to be one of the new types of web content today - a successful product of the Web 2.0 phenomenon. Mashups increasingly rely on web applications that allow web pages to be updated in real time (using AJAX) without the need to refresh the entire page on every update. Mapping mashups require a client-side map viewer and access to some reliable source for the type of data or information that the mashup will display.

3. SYSTEM MODEL

Figure 1 illustrates our basic system architecture, based on the general INSPIRE (European SDI) technical architecture (INSPIRE Drafting Teams, 2007) and implemented with the specific components and technologies described in following sections. Its goal is to identify the services and components

needed to solve user requirements, to facilitate the composition, integration and reuse of concrete instances of each kind of requested service, and to permit the rapid visualization and exploration of data by means of mapping mashups.

The architecture is composed of loosely-coupled layers. Thus this approach cleanly separates the user-interface part of the web application (geoportal and mashup), from the integration logic and the instances of each kind of geospatial web service. Scientists and other decision makers interact with the Presentation layer that includes components to support user interface and interaction, and data visualization and exploration. The Horizontal layer is concerned with integration, instantiation, and invocation of services instances. This layer also comprises the description and implementation of other components needed both for the business logic (in our case hydrological model logic), and for processing visualization data that will be available to clients through the mapping mashup. The Service layer combines a set of service instances grouped in categories such as discovery, view (maps), download (accessing geospatial data) and processing services. Spatial data sets and metadata reside in the Data layer, which is neither illustrated in Figure 1 nor described in the following sections, because it is out of the scope of this paper.

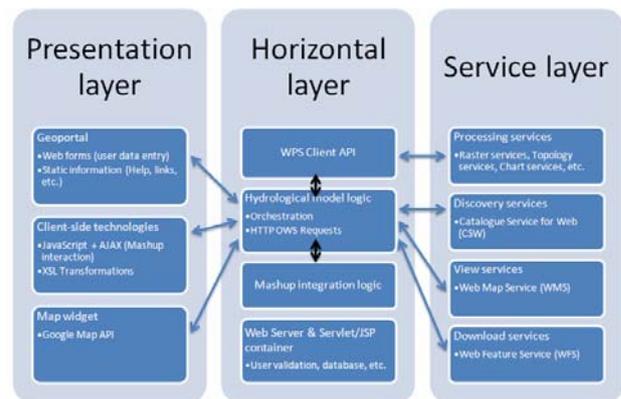


Figure 1. Layered architecture and components.

3.1 Presentation layer – Geoportal

The Presentation layer provides the entry point for users and decision makers to access the data and services provided by our web application. In this sense, our web application plays the role of a one-stop geoportal (Bernard et al., 2005) that offers expert users all of the capabilities necessary to configure and run hydrological models for a particular watershed of study.

The user interface of the Geoportal is composed of multiple web forms (wizard-like interfaces) that guide users in the sequence of steps involved in configuring hydrological models. Each and all web forms (i.e. model steps) offer a consistent user interface through the whole process, as illustrated in Figures 2-4. The left side contains HTML forms for data collection and displaying useful information generated as the model steps are being executed. The right side contains the map widget to visualize geodata.

It is important to mention that multiple technologies have been interrelated (mashed-up) to compose the user interface of our Geoportal. First we have used server-side technologies (e.g. JSP) for retrieving information and generating content dynamically

for HTML forms. These technologies are closely related with the service and data integration component in the Horizontal layer. For data visualization, we have used the Google Maps API (Jones, 2007) for building the user-interface part of the mapping mashup in our Geoportal. Client-side technologies such as JavaScript and AJAX (Asynchronous JavaScript and XML) are also common in building mashups for enabling user interaction, as well as other technologies focused on the specification of data and visual encodings for feeding mapping mashups (e.g. KML), which actually are processed in the Horizontal layer.

3.2 Horizontal layer – Service Integration and Mashup Integration Logic

Service and data integration is at the core of the Horizontal layer to enable the communication between the Presentation and Service layer. For this we have developed a set of components (the first three components in Figure 1) to deal with integration, composition, instantiation, and invocation of services instances in the Service layer.

The first component, WPS Client API component, enables the communication between the user and web processing services, facilitating connection to and access of WPS. For example, if some geospatial web service is necessary while a web form is being completed (e.g. introduction of parameter values), the WPS client API can invoke the corresponding service by transforming user requests into OGC WPS requests. Once results are returned, the WPS also extracts results (output parameters) to be properly processed by the remaining components involved in service and data integration.

The WPS Client API component may communicate with the other two components in the Horizontal layer: the hydrological model logic and the mashup integration logic. The former implements auxiliary functions, concrete data constraints, and other specific restrictions subject to the logic of the hydrological model. For example, certain tasks are actually complex workflows which demand several executions of geospatial services in a certain composition structure. The hydrological model logic component then dictates the service orchestration (e.g. iterations, sequence) of the different geospatial processing services involved in the workflow as a chain of WPS calls via the WPS Client API component.

Once the hydrological logic model obtains the output results for a given workflow, the mashup iteration logic component streams these results through the mapping mashup (Presentation layer) by converting mainly processing data (workflow results) from Geographic Markup Language (GML) (Portele, 2007) format to visualization data encoded in KML, ready to be displayed by Google Maps. (Alternatively one could visualize the GML using a Web Feature Service (WFS)).

Basically, these three main components communicate with each other in order to manage all aspects of service management. Rather than providing a unique component for service and data integration, we have intentionally decoupled it, creating three components in order to more easily accommodate other hydrological models to the Geoportal. At this moment, two hydrological models are supported therefore we created two specialized components of the hydrological model logic, each dedicated to one model. The addition of a new hydrological model requires implementation of a specialized component of the hydrological model logic.

3.3 Service layer – Geospatial Web Services

The Service layer groups service instances in basic functional categories. Although our work has been focused mainly on processing services, our Geoportal application also accesses other INSPIRE network services types, in order to facilitate access to not only geospatial data processing but also to discovery, viewing and download of data.

3.3.1 Processing Services: Processing services can perform complex computations on geospatial data. They are frequently described or packaged using the OGC WPS specification, as introduced in Section 2. Our processing service design methodology has been based on wrapping. In order to provide useful geospatial processing services which suit the concrete requirements of hydrological models, we have identified atomic functions shared among the model tasks. An atomic function performs a basic operation that should be easily tested and be domain-independent enough to be applicable to other contexts. An atomic function may be either already provided by existing GIS libraries and projects, or may require more elaborate and customized capabilities. For the former, we have wrapped existing GIS functions as WPS services. For the latter we have created new functions and exposed them as WPS services. In this case, such functions may be built both from scratch or combining existing atomic WPS service as chains of geoprocessing services. The ultimate goal is to create a library of geospatial processing services in which customized and elaborated functions rely on other much more simple, atomic and well-tested functions. In this way, the reuse of geospatial processing services is fostered because the process of creating new complex geospatial processing services is made possible by reusing already available geospatial processing services from shared libraries. Using Free and Open Source Software (FOSS) projects we have implemented more than 20 geospatial processes grouped in four WPS. Users can exploit these services to mine and to analyze data from any OGC-compliant online data sources.

3.3.2 Discovery Services: These are a special kind of service that offers end users a common mechanism to register, search, and access discoverable geospatial data and services. In our case, we use a catalogue service which contains metadata records of satellite imagery. As users progress through the web forms, the Geoportal is able to collect useful data that are used to perform automatic queries against the service catalogue. The left side of Figure 2 shows the list of metadata records that matched the user criteria. At this moment, the user only needs to select the proper metadata records according to her expertise. The target data (described by the metadata), in this case satellite imagery, is retrieved automatically by means of download services.

3.3.3 Download Services: These are the basic geospatial services that serve geospatial data to a client application or to other services internally. The most prominent example of download services is the Web Feature Service (WFS), which filters and retrieves vector format representations of geospatial features and feature collections encoded in GML. From our perspective, the Geoportal uses a WFS service to retrieve the GML data associated to the metadata records selected by the user.

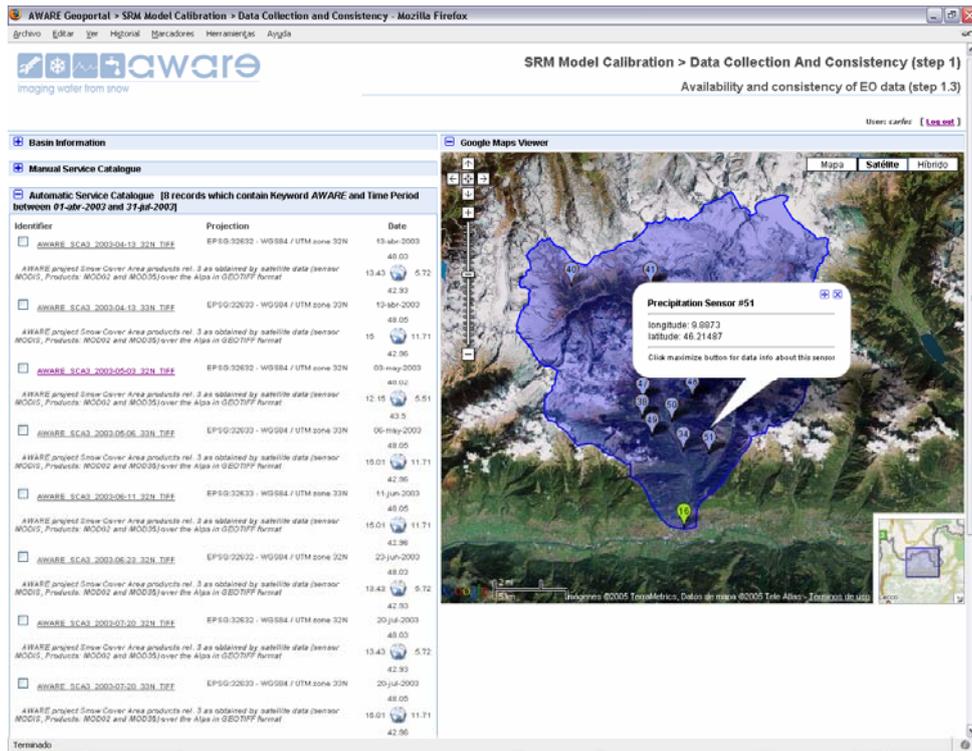


Figure 2. Integrating a catalogue service client in a mapping mashup.

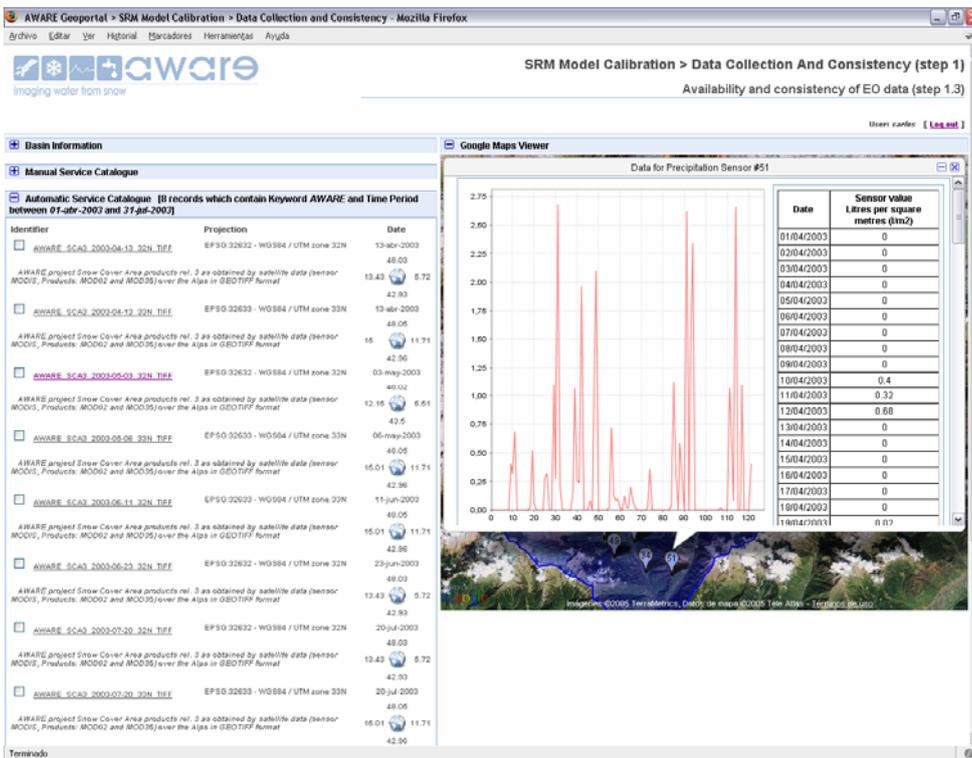


Figure 3. Detail of data visualization and exploration.

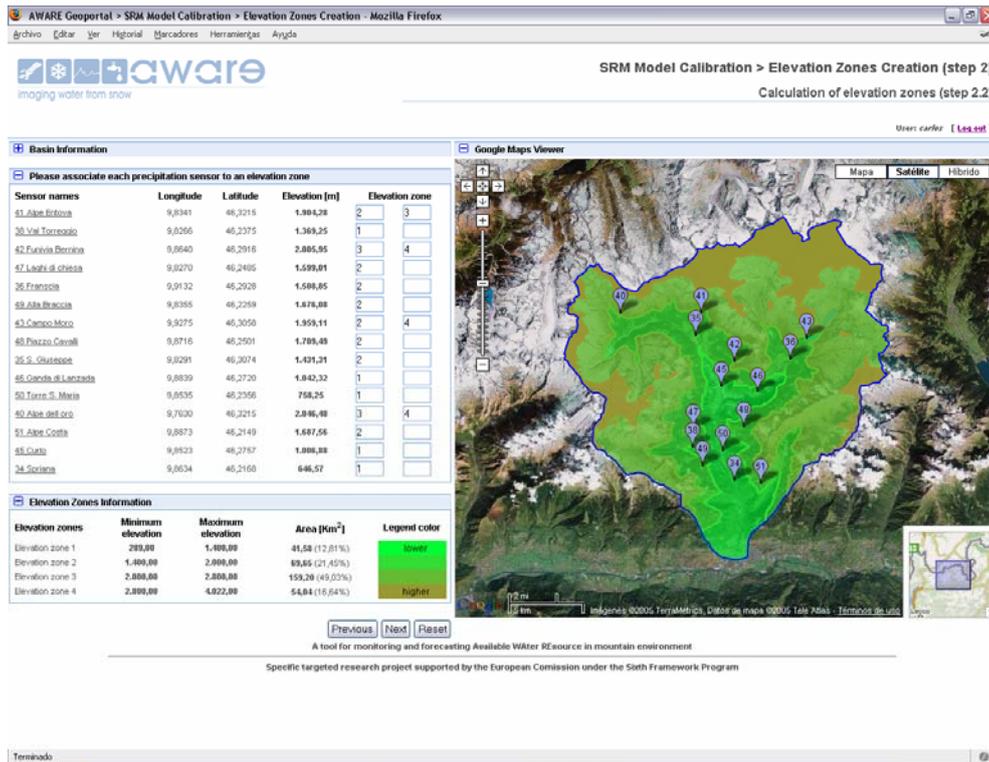


Figure 4. Calculation and visualization of the elevation zones of a watershed.

3.3.4 .View Services: View or portrayal Services may be also considered a specialized download service that produce rendered data such as portrayed maps, perspective views of terrain, annotated images, and so on. Examples are the Web Map Service (WMS) that dynamically produces spatially referenced maps of client-specified criteria from one or more geographic datasets, returning the map views in well-known image or graphics formats. Again, from the metadata records, the Geoportal is capable of retrieving image maps and the legend for the satellite imagery selected by users. Figure 2 shows the WMS link (the globe icon under “date”) associated to each metadata record retrieved from the catalogue service.

4. IMPLEMENTATION

As described in the Introduction section, many datasets are available in regions such as North America and Europe to provide scientists with a more complete view and context of hydrologic processes: stream gauge datasets; temperature and precipitation data (collected from weather stations at different locations); the location of all gauges, weather stations, the geographic boundary of the watershed (drainage basin); the Digital Elevation Model (DEM) of the watershed; and satellite imagery capturing the snow coverage area of interest. Our prototype for water resource management facilitates the collection, validation and integration of spatial data for analyzing a given watershed, and provides access only to those specific processes that are needed for any given model, reducing wasted software implementation, maintenance and cost.

Figure 2 shows an example of our prototype, in which metadata records are retrieved from the catalogue service automatically. For each record, multiple fields are displayed, such as the title, projection, date, bounding box, etc. The right side of Figure 2

shows the map viewer, the user-interface part of the mapping mashup, displaying the basin boundary together with the location of meteorological sensors. Each sensor is identified according to fill colour (red-temperature; blue-precipitation, and green-stream gauge). As explained in section 3, the map shown in Figure 2 is rendered in the Presentation layer but the input data has been processed and integrated by the components set devoted to data and service integration in the Horizontal layer. Figure 3 shows an example of the possibilities offered by the prototype for data visualization and exploration. Users may click on a sensor icon to obtain more information about the sensor data, as in the case of Figure 3. This action is not executed on the client side, but rather on the server side via requests to the corresponding components and services in the Middleware. In this sense, we prefer to keep the browser simple (used mostly for user interaction and for data entry) rather than becoming the integration platform for running interactive mapping mashups as typically happens in software mashups (Daniel et al., 2007). Furthermore, both the plot and raw table displayed in Figure 3 are processed in the server side, forwarding in this case the HTML code that can be easily rendered by the browser.

Figure 4 shows one of the main steps contained in one of the tasks within the major hydrological model supported in our prototype (SRM). It is dedicated to the calculation and visualization of elevation zones for a watershed and involves multiple input data and processing services. The workflow of this step involves several processing service such as the raster analysis process, spatial intersections, and coordinate transformation, all wrapped as WPS services (Díaz et al., 2008). Again, the resulting elevation zones are processed in GML but transformed into KML by the mashup integration logic component to be visualised in the mapping mashup.

5. CONCLUSIONES

A first observation derived from the scenario experience, which coincides with conclusions in previously referenced work, is that the approach based on distributed geoprocessing services leads to a collection of reusable geoprocessing services, which may be useful for other users providing they are well-documented and registered in open catalogues. This is possible in principle because WPS-based geoprocessing services do not work with pre-established datasets but rather they preserve a loosely-coupled relationship between data and processing capabilities (algorithms), making it possible to chain them to other geospatial web services such WMS and WCS. We have demonstrated that wrapping and exposing existing algorithms as web processing services makes available and reusable certain functionality that was previously only accessible by acquiring multiple desktop software packages. These services need not be maintained or executed locally avoiding the computation time required for this. Furthermore for implementing certain complex scientific models we only need a model engine that performs the orchestration of this web processing services. An interesting research problem to be addressed is the case where a geoprocessing service requires transport and calculation of large amounts of Earth Observation (EO) data, and so this is a future goal for the Geoportal application.

REFERENCES

- Alameh, N., 2003. Chaining Geographic Information Web Services. *IEEE Internet Computing* 7(5): 22-29.
- 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*, 29(1), pp. 15-31.
- Portele, C., 2007. OpenGIS Geographic Markup Language (GML) Encoding Standard. <http://www.opengeospatial.org/standards/gml> (accessed 30 Apr. 2008).
- Daniel, F., Matera, M., Yu, J., Benatallah, B., Saint-Paul, R., Casati, F. (2007). Understanding UI integration: A survey of problems, technologies, and opportunities. *IEEE Internet Computing* 11(3), pp. 59-66.
- Díaz, L., Granell, C., Gould, M., 2008. Case Study: Geospatial Processing Services for Web-based Hydrological Applications. In *Geospatial Services and Applications for the Internet Era*. Springer, New York.
- INSPIRE Drafting Teams, 2007. INSPIRE Technical Architecture. http://www.ec-gis.org/inspire/reports/ImplementingRules/network/INSPIRETechnicalArchitectureOverview_v1.2.pdf (accessed 24 Apr. 2008).
- Jones, M. T., 2007. Google's Geospatial Organizing Principle. *IEEE Computer Graphics and Applications* 27(4): 8-13.
- Lemmens, R., Wytzisk, A., de By, R., Granell, C., Gould, M., van Oosterom, P., 2006. Integrating Semantic and Syntactic Descriptions to Chain Geographic Services. *IEEE Internet Computing* 10(5): 42-52.
- Masser, I., 2005. *GIS Worlds: Creating Spatial Data Infrastructures*. Esri Press, Redlands.
- Papazoglou, M.P., 2008. *Web Services: Principles & Technology*. Pearson, Essex.
- Schut, P. (ed), 2007. OpenGIS Web Processing Service 1.0.0. Open Geospatial Consortium. <http://www.opengeospatial.org/standards/wps> (accessed 30 Apr. 2008).
- Wood, J., Dykes, Slingsby, A. Clarke, K., 2007. Interactive Visual Exploration of a Large Spatio-Temporal Dataset: Reflections on a Geovisualization Mashup. *IEEE Transactions on Visualization and Computer Graphics* 13(6): 1176-1183.

ACKNOWLEDGEMENTS

This work has been partially supported by the AWARE project SST4-2004-012257 co-funded by the EU and the GMES initiative. Institut Cartogràfic de Catalunya (ICC, Spain) assisted in the design of the Geoportal, and Istituto per il Rilevamento Elettromagnetico dell'Ambiente (IREA, Italy) has set up some data services and the catalogue service detailed here; both are partners of the AWARE project.