Accepted Manuscript

Managing user generated information in geospatial cyberinfrastructures

Laura Díaz, Carlos Granell, Michael Gould, Joaquín Huerta

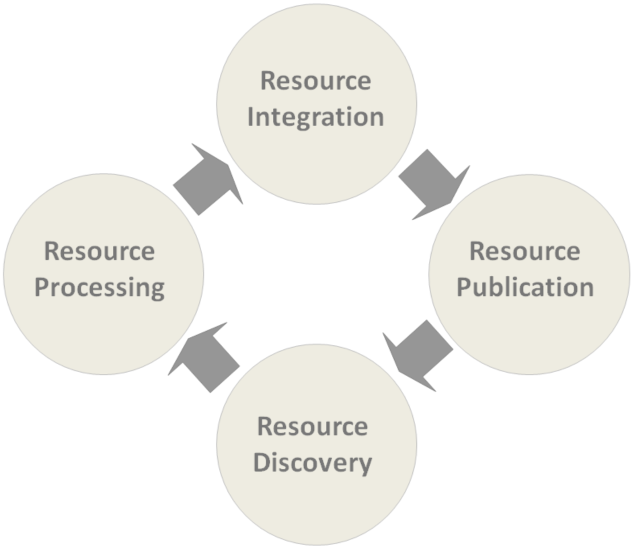
PII: S0167-739X(10)00172-X
DOI: 10.1016/j.future.2010.09.002
Reference: FUTURE 1936

To appear in: *Future Generation Computer Systems*

Received date: 25 January 2010
Revised date: 23 August 2010
Accepted date: 1 September 2010

Please cite this article as: L. Díaz, C. Granell, M. Gould, J. Huerta, Managing user generated information in geospatial cyberinfrastructures, *Future Generation Computer Systems* (2010), doi:10.1016/j.future.2010.09.002

Graphical Abstract



Research highlights

- User-driven methodologies to maintain information infrastructures
- Assist users in these steps during the resource life cycle.
- Concealing technology to integrate user content generating standard components.
- New middleware components in SOA to help users to participate actively.

Managing user generated information in geospatial cyberinfrastructures

Laura Díaz*, Carlos Granell, Michael Gould, Joaquín Huerta

Institute of New Imaging Technologies (INIT)

Universitat Jaume I, Castellón (Spain)

{laura.diaz, carlos.granell, gould, huerta}@uji.es

*Corresponding author

Abstract

Information systems built using standards-based distributed services have become the default computing paradigm adopted by the geospatial community for building information infrastructures also known as Spatial Data Infrastructures (SDI). Government mandates such as the INSPIRE European Directive recommend standards for sharing resources (e.g. data and processes) with the goal of improving environmental (and related) decision-making.
Although SDI present benefits to data providers in terms of data sharing and management, most of geospatial infrastructures have been built following a top-down approach where official providers (most commonly mapping agencies) are permitted to deploy and maintain resources. Because the mechanisms to deploy resources in these infrastructures are technologically complex, there has been limited participation from users, resulting in a scarcity of deployed resources.
To address these limitations, we present a distributed architecture based on INSPIRE principles and extended with a Service Framework component. This component improves ad hoc integration and deployment of geospatial data resources within geospatial information infrastructures. The Service Framework addresses the need to improve the availability of geospatial data resources by providing mechanisms to assist users in wrapping resources to generate INSPIRE-based services.

Keywords: *User content, Processing Services, Collaborative Services, Service Oriented Architecture; SOA; Spatial Data Infrastructure; SDI; Geospatial Cyberinfrastructures; GCI; Hybrid SDI building, Assisted Deployment and Publication*

*Corresponding author contact information

laura.diaz@uji.es

Institute of New Imaging Technologies (INIT)

Universitat Jaume I

Av. Vicente Sos Baynat s/n

12071, Castellón (Spain)

Tlf +34 964 72 90 78

Second review paper to FGCS, special issue on "Data management for eScience"

Managing user generated information in geospatial cyberinfrastructures

Abstract

Information systems built using standards-based distributed services have become the default computing paradigm adopted by the geospatial community for building information infrastructures also known as Spatial Data Infrastructures (SDI). Government mandates such as the INSPIRE European Directive recommend standards for sharing resources (e.g. data and processes) with the goal of improving environmental (and related) decision-making.

Although SDI present benefits to data providers in terms of data sharing and management, most of geospatial infrastructures have been built following a top-down approach where official providers (most commonly mapping agencies) are permitted to deploy and maintain resources. Because the mechanisms to deploy resources in these infrastructures are technologically complex, participation from users has been limited, resulting in a scarcity of deployed resources.

To address these limitations, we present a distributed architecture based on INSPIRE principles and extended with a Service Framework component. This component improves ad hoc integration and deployment of geospatial data resources within geospatial information infrastructures. The Service Framework addresses the need to improve the availability of geospatial data resources by providing mechanisms to assist users in wrapping resources to generate INSPIRE-based services.

Keywords: *User content, Processing Services, Collaborative Services, Service Oriented Architecture; SOA; Spatial Data Infrastructure; SDI; Geospatial Cyberinfrastructures; GCI; Hybrid SDI building, Assisted Deployment and Publication*

1. Introduction

Multidisciplinary research teams require the support of suitable cyberinfrastructures that allows them share distributed data and computing capabilities to perform collaborative science more efficiently (Hey and Trefethen, 2005). In the environmental and geospatial domains this approach has been identified as a solution to the challenge of generic interoperability and data integration (Denzer, 2005; Yang et al., 2010). Indeed, Geospatial Cyberinfrastructures (GCI) are regarded as cyberinfrastructures that focus particularly on geospatial resources to support processing capabilities for end users such as geospatial analysis, environmental modelling, and decision making. The integration and use of accurate, up-to-date geospatial data through GCI leverages efficient management, sharing and exploitation of geospatial resources scattered among numerous agencies, a critical aspect recognized in most environmental applications (Mansourian et al., 2005; Rocha et al., 2005; Nayak and Zlatanova, 2008).

Spatial Data Infrastructure (SDI) is one of the enabling technologies that contribute in the development of GCI. The capability of discovering, accessing and sharing a diversity of geospatial resources, among a wide range of stakeholders, is being addressed by interconnected SDI nodes at different scales to build a global information infrastructure (Masser et al., 2008; Rajabifard et al., 2002). This network of operational SDI nodes exposes large volumes of geospatial resources through standards-based geospatial services so that scientists can combine them into customized applications. This approach has been proven to be effective in multiple application domains like urban planning (Bishop et al., 2000), forestry (Davis et al., 2009), risk management (Friis-Christensen et al., 2009) and hydrology (Granell et al., 2010).

Scientists collaborate on large-scale projects by accessing available computing and data resources to process and generate new information. In the geospatial community, however, sophisticated capabilities are still required to automatically collect, interpret and integrate these new resulting information into the sharing knowledge of the research teams. Furthermore extending SDI nodes to make geospatial data readily available to scientists, policy makers, and users in general poses at least the following two challenges:

- The ability to manage and trace the outputs generated by workflows of services is a desirable aspect to compare, share and replicate certain experiments among different teams of scientists (Deelman et al., 2009; Fook et al., 2009). For instance, some forestry models compute fire assessment maps used exclusively during a given experiment. Yet, this behaviour, where valuable results are often hidden and isolated, limits other scientists to potentially harness such maps (resources) for validation and calibration in their own experiments.

- Current operational SDI nodes rely on geospatial web services exposed by standards-based interfaces (Friis-Christensen et al., 2007; Kiehle, 2006; Krishnan & Bhatia, 2009), though, such information infrastructures still require advanced technological capabilities (skilled personnel, geospatial middleware technologies and tools, etc.) both for regular refinement and maintenance due to their dynamic nature (Ackland, 2009) to effectively manage the increasing rates of user generated information.

The lack of these capabilities impedes sharing crucial information rapidly with other interested stakeholders avoiding the availability and visibility of geospatial resources in a short time (Buyya et al., 2009). In this line, Yang et al. (2010) have identified the need of making efforts in the development of geospatial middleware and tools to better support citizen contributions, advanced services, models, and decision support analysis needed by stakeholders and the general public.

In this paper we address precisely these issues related to the capability of generating and integrating new knowledge and user generated information into the mainstream SDI, and hence in GCI, to be efficiently reused and shared. Our contributions thus are two-fold:

- Semi-automatic mechanisms to augment the availability of user generated information by assisting users to wrap resources as standard web services. This greatly would increase the rates of sharing and availability of geospatial resources by massively adding them in SDIs in forms of standards-based geospatial services.
- Semi-automatic mechanisms to improve the visibility of geospatial resource by generating metadata descriptions and publishing them in open catalogues. As pointed in the literature (Craglia et al., 2007; Nogueras et al., 2005), metadata and catalogue services are key in SDI in order for data and service resources to be properly discovered.

Both contributions are materialized in the geospatial middleware toolset called *Service Framework,* which will be the focus of this paper. To illustrate our proposed solution, we describe a use case within the European-funded project EuroGEOSS[i] (European approach to GEOSS). EuroGEOSS pursues the improvement and establishment of interconnections among systems and resources to take benefit from multidisciplinary data and tools available at global, national and regional levels. EuroGEOSS is broaching geospatial thematic areas like forestry. In this context, scientists are looking for suitable and efficient ways to prevent and assess fire impacts. Forest fires analysis requires the development and management of information systems, datasets and services which could interact and communicate with each other. We stick to this forestry scenario through the paper, where we illustrate a protected areas damage assessment use case as a running example.

The rest of the paper is organized as follows. Section 2 highlights the role of users as a driving force in capacity building in GCI. In Section 3, we discuss the Service Framework architecture, various software components and services, and how we have addressed the presented challenges. In Section 4 we evaluate the Service Framework against the protected areas damage assessment use case. We discuss related work in Section 5. Finally, we present our conclusions in Section 6.

2. Extending geospatial infrastructures with user-driven capacity building

Traditionally SDI building follows a top-down approach. This scenario leads to the provider-consumer paradigm, where only official providers like National Mapping Agencies (NMA) and other environmental agencies, centrally, manage and deploy resources according to institutional policies (Béjar et al., 2009) while end-users are limited to the pure consumer role.

Nevertheless, recent advancements in web technologies have enabled new ways of participation on the web. It is not surprising to see how the web has changed the way we communicate, how we do our daily routines, and even our social behaviour (Castells, 2001). Citizens, experts and non-experts alike, are increasingly participating in the process of generating continuous information and collaborating with others in problem-solving tasks. This highlights the transition of the role of users from just mere data consumers to active participants and providers. Consequently end users now interact, use, and access information infrastructures in a different manner (Ackland, 2009).

The shift in the role of users has also been reflected in the geospatial domain. Volunteered Geographic Information (VGI), coined by Goodchild (2007), highlights the fact that users are becoming active producers of geographic information rather than only passive recipients. The sharing and availability of user-generated geographic information within the mainstream SDI may substantially improve traditional geospatial analysis and decision support tasks (Flanagin and Metzger 2008; Pultar et al., 2009). Indeed, authors (Budhathoki et al., 2008; Omran and van Etten, 2007) have even suggested a new SDI generation largely influenced by user needs.

In order to extend SDI and GCI with user-driven capacity building, we draw on mechanisms to assist users to integrate and share their resources. Such mechanisms consist of wrapping geospatial resources as target standard services to generate automatically standardized components that guarantee the interoperability across underlying information infrastructures. As geospatial web services come with standards-based interfaces for describing, discovering and accessing resources, they provide a common alphabet (see Section 3). Yet a true lingua franca also requires agreement on protocols, data formats, and ultimately semantics (Foster et al., 2005). At this point is where common and general frameworks for a SDI like the Infrastructure for Spatial Information in Europe directive (INSPIRE; INSPIRE, 2007) come into the scene, providing the standards, protocols and common patterns used in our approach to reach interoperability at different levels.

The benefits of user-driven capacity building can be seen in multiple domains and user communities. Taking GEOSS (Global Earth Observation System of Systems[ii]), as a real example of how geospatial infrastructures are used, we explain next how GEOSS may benefit from user participation to support the sharing and utilization of geospatial resources.

GEOSS is based on existing systems, standard specifications and components whose primary aim is the organization of the underlying geospatial resources by means of easy discovery and interoperable access. Essentially, GEOSS demonstrates the added value to the scientific community and society of making existing systems and applications interoperable within SDI and GCI infrastructures.

Figure 1 shows several of the general use cases (blue circles) identified in the GEOSS Architecture Implementation Pilot phase 2 (AIP-2) (Butterfield et al., 2008) in order to illustrate how the capacity building of the systems in GEOSS follows a top-down methodology. These use cases clearly differentiate two main roles or actors. The *GEOSS Resource Provider* who is the only primary actor able to deploy and register resources (data and processes), and, the *GEOSS end user* who can search and consume resources but can not create and deploy new resources limiting thus its participation in the building process of GEOSS.

Figure 1. AIP-2 GEOSS Transverse Technology (simplified) Use Cases

What we would like to remark in Figure 1 is that we have added two new use cases (green circles) that lay out our contribution. We extend the initial scenario, following the same direction of the web evolution, to allow users to participate in deploying and publishing resources. Our approach takes a hybrid approach to put these new use cases in place. We follow a top-down building methodology which relies on standard specifications and service interfaces, and a bottom-up methodology since we provide mechanisms to overcome the complex deployment mechanisms of the top-down building methodology to let *GEOSS end users* participate in the building of SDI and GCI. Our running example using the *Service Framework* toolset will demonstrate these new use cases to enable user participation.

3. The Service Framework anatomy

GCI users are continuously generating new resources, such as information and models, which should be stored and persisted in GCI to guarantee their accessibility to the rest of users. However, the complex deployment mechanisms and the traditional top-down building methodologies impede end users to participate in the expansion of such GCI by deploying and publishing these new resources.

Regarding the persistence of models in GCI, in previous works we described concrete aspects in the implementation and deployment of OGC web processing services in INSPIRE-compliant infrastructures (Granell et al, 2010; Díaz et al, 2008). The Open Geospatial Consortium (OGC[iii]) is an international, consensus-based organization whose aim is to foster interoperability between geospatial data, clients and services by promoting the definition of well-established interfaces to a wide range of geospatial web services. One of the conclusions of our previous works is that deploying resources (data and tools) as services, described using these standard interfaces, in distributed environments notably increases the reusability of these resources in multiple contexts.

Based on these principles, our approach has its roots in a generic architecture according to INSPIRE guidelines (INSPIRE, 2007), as shown in Figure 2 (left). Basically, INSPIRE-based SDIs are built on three layers, where all geospatial resources (data, processes and metadata) reside in the resource layer. The middle or service layer contains all the geospatial services that are deployed and registered into the GCI according to international standards and agreements. These services are classified according to INSPIRE service types into services that expose discovery, view, download and processing capabilities. Finally, the application layer holds the business logic, workflow engines and other modules needed to execute and combine services according to the needs of client applications. The traditional top-down approach is exemplified in the left side of Figure 2 by means of up arrows, since users through the application layer connect to available services to discover, access, and consume geospatial resources.

Figure 2. Service Framework and its components

The right side of Figure 2 shows the *Service Framework*, the proposed extension to SDI generic architecture to permit the persistence of user resources. It represents the strategy and methodology to manage and deploy user-generated resources in GCI. This framework provides the mechanisms to assists users in managing geospatial resources, not only in terms of accessing distributed resources, but also in deploying and publishing them in target SDIs. Since the scope of this work focuses on data resources, the main goal is to describe the *Service Framework* as a real implementation example of the hybrid capacity building approach, where the down arrows represent the provider role taken by end users: they now contribute to the growing of GCI.

The use of web services, described with OGC standard interfaces, have been proven to be successful in order to improve interoperability (at syntactical level) when accessing and sharing

geospatial resources in GCI. The Service Framework plays a vital role in the service provision paradigm because it acts as a service generator. It gets a data resource and returns an updated standard-based service deployed into the GCI with new content that corresponds to this resource.
As shown in Figure 2 (right), the Service Framework is organized in a 3-tier structure, where the *Service Deployer* is the core component to deploy resources in the geospatial services instances placed in the Service Layer. In the following we describe the main components contained in the Service Framework that are involved in the management of user-generated information.

3.1 Service Connector

Users may interact with geospatial services using various mechanisms and client applications such as geoportals, mashups and desktop GIS. As we can see in the Figure 2, the *Service Connector* module contains the components that provide the functionality to access these (standard and interoperable) services.

The Service Connector component offers a simple but convenient API to deal with OGC-based services from the client perspective. It includes the OGC Web Mapping Service (WMS; De La Beaujardiere, 2004) connector to access portrayal capacities, the OGC Web Feature Service (WFS; Vretanos, 2005) and OGC Web Coverage Service (WCS, Whiteside & Evans, 2008) connectors for data downloading and the OGC Catalogue Service (CS-W, Nebert et al., 2007) connector for registration and discovery purposes. One of the most relevant connectors is the *Web Processing Service connector* (*WPS connector*) that exposes easy-to-use methods to query and execute geoprocessing services interfaced by the OGC Web Processing Service specification (WPS, Schut, 2007). The WPS connector queries and executes WPS services independently of the granularity and the process model of the WPS and their underlying implementation. The *WPS connector* has been proven to be efficient to communicate with OGC WPS-based services in real applications (Granell et al., 2010). Therefore, this component deals with accessing and performing distributed geoprocessing functions or models to modify or create new information.

3.2 Service Deployer

The Service Deployer is the component that deals with the deployment and persistence of the new generated information into the GCI, both from source resources injected directly by users and resources resulting of executions of WPS-based services. The Service Deployer assists users in managing and integrating data resources in the underlying infrastructure by means of two components: the *Data Wrapper* and the *Service Publisher*. The former assists users in wrapping data as standard services according with INSPIRE implementing rules (availability). The latter aims at creating metadata descriptions to publish them in a catalogue service (visibility).

3.2.1 Data Wrapper

The Data Wrapper component deals with the deployment of data resources by wrapping them as standard data services implemented according to INSPIRE implementing rules[iv]. The Data Wrapper generates data services of two INSPIRE service types: *View* and *Download* services (see Figure 2). The View services are implemented using the service interface described by WMS specification (INSPIRE, 2008). For downloading purposes, the Data Wrapper generates Download services using either the WFS specification for vector data or the WCS specification for raster data (INSPIRE, 2009). It is worth mentioning that the Data Wrapper feeds these INSPIRE Services with any dataset independent of the data model. At its current status one of the known limitations is the spatial data format of the dataset, which is limited to the formats supported by the service implementation.

Among the existing OGC-compliant implementations, Geoserver[v] is the implementation of choice in our work to instantiate the OGC-based standard services. Geoserver implements the three service interfaces relevant for our purpose (WMS, WFS, and WCS). In particular, the addition of new geospatial datasets to a given Geoserver instance is supported by the Geoserver RESTful API[vi], which lets the Data Wrapper to configure and access Geoserver instances programmatically via HTTP methods as the application protocol (GET, POST and PUT). Consequently new data resources are deployed and made available through the standard services in the service layer as shown in the right side of Figure 2.

3.2.2 Service Publisher

The Service Publisher component generates metadata descriptions to be published in catalogue services (Nogueras-Iso et al., 2005) implemented according to the INSPIRE implementing rules for the *Discovery* service type. It contains two components: the *Metadata Generator* and the *Resource Publisher.* The Metadata Generator utilizes the Service Connector to query the OGC services, where the data resources have been previously deployed. In response, the Metadata Generator collects the service descriptions that contain the functional and non-functional descriptors needed to generate a minimum set of metadata elements.

The Resource Publisher utilizes the Service Connector to connect to the service catalogue implementing the CS-W specification (INSPIRE, 2009b) and publish the metadata description with compliance to the INSPIRE recommendation for geospatial metadata (ISO, 2003). The task of publishing metadata documents in existing catalogue services requires transformations between the source OGC service description and the target ISO metadata description. The Resource Publisher performs these transformations either by using XSLT technology or by sending the OGC service descriptions directly to a catalogue service, which, if supported, internally performs such transformations. For the first prototype we have chosen Geonetwork[vii], as implementation of the OGC CS-W, so that both transformation alternatives are possible.

4. Service Framework Evaluation

In this section we proceed to evaluate the usefulness of the Service Framework in a running example. We describe the user-driven strategy for adding data resources in GCI, that is, the sequence of actors and steps involved in assisting users in deploying data resources. Finally, we elaborate on the Service Framework execution applied to the protected areas damage assessment running use case, being developed within the context of the EuroGEOSS project.

4.1 User-driven data integration strategy

The sequence diagram in Figure 3 illustrates the steps and components involved in the Service Framework strategy to assist users in deploying data resources in GCI as interoperable services. The main actors involved in our contribution have been highlighted in purple.

The first step is to select the source of the data resource to be deployed. It can be a local or remote dataset. The sequence diagram illustrates the case in which a user deploys processed information derived from a WPS-based service execution. Therefore the first steps illustrate the fact of generating new information by performing distributed processing. When the user selects a WPS service through the Service Framework interface the communication with the target WPS is managed by the WPS connector, a component within the Service Connector (see Section 3.1). The WPS connector conceals interactions with the target process to the end users, such as the *getCapabilities* request to the WPS to inspect the available processes. After the available processes are shown the user selects one of them. In the next step the WPS connector sends a *DescribeProcess* request to the WPS to examine the description of the required inputs and outputs parameters to execute the process. At this point, the user must properly select the input data resources, which can be retrieved from local storage or accessed through remote, standard download services (e.g., WFS, WCS). In the latter case, the Service Connector is used to request this remote data. Once the input data resources are in place, the user can execute the process assisted by the Service Framework interface. Once again the WPS connector sends an *execute* request to the WPS services and retrieves the resulting dataset (information) that will be portrayed in a map viewer integrated in the Service Framework interface.

Once user information resources are obtained, the Data Wrapper component within the Service Deployer (See Section 3.2) is in charge of generating the corresponding standard data service. This component uses the Geoserver RESTful API (via PUT method) to add this new dataset in the Geoserver instance than implements the View and Download Service types. After that, the Data Wrapper component configures a new service layer according to the data nature. Note that the Data Wrapper yields a couple of deployed geospatial services from the same data resource. One is a view service interfaced by the OGC WMS specification; the other is a download service using either the OGC WFS or the OGC WCS, depending on the nature of the resource: vector or raster respectively.

Figure 3. Sequence diagram to deploy user information through the Service Framework.

This process has addressed the challenges mentioned in the introductory section. The proposed mechanisms deal with the management of the outputs generated by workflows to be later shared and compared to replicate experiments among different teams of scientists. For that we have assist users to participate in building and maintaining the GCI to efficiently manage the increasing rates of user generated information turning it into an increased amount of interoperable resources available.

4.2 Use case: Protected areas damage assessment

To illustrate with a practical approach we describe the sequence diagram seen in Figure 3 with a use case as a running example. The interactions between the user and the Service Framework components are illustrated in Figures 4-7. These figures show a prototyped user interface to access the functionalities of the Service Framework, which at this current status are (1) upload local datasets as we can see in the *File Upload* tab in the figures, (2) Execute distributed WPS-based services to generate new information as shown in the *Web Processing Service* tab, (3) visualize local and processed datasets in the viewer included in the interface, and (4) deploy these datasets as INSPIRE-based services as shown in the *Deploy* tab.

The use case describes a protected area's damage assessment where a local technician estimates the damage of the protected areas after the forest fires in the summer of 2005 in the Valencia region in the east coast of Spain. At a local level the *Conselleria de Medi Ambient, Aigua, Urbanisme i Habitatge*[viii] (Environmental, Water, Urban planning and living regional government) generates datasets on burned and protected areas on a bigger scale than its counterpart's datasets that are being generated at the national or European level (Giovando et al, 2010). The differences between data models across different administrative levels are notable, though, this issue is not considered here but needs further discussions in future developments. Nevertheless, the key aspect described here is the fact that the Service Framework would assist the local technician in deploying and sharing the results of the assessment. We will see how the Service Framework, independently of the data model, assists the local technician to upload a local dataset and the assessment results into the infrastructure. This information will be deployed as a standard data service. At this moment, the deployed service can be shared with other technicians at different administrative levels, for instance to compare the assessment at other scales.

Figure 4 shows how the protected areas and burned areas datasets are uploaded by the user and deployed in the GCI. In the *Deploy* tab the technician can deploy these resources (behind the scene via the Data Wrapper) and publish them in a catalogue service (via the Service Publisher), so that the resulting layers may be shared through standard data services. As a result, the Service Framework component returns two URLs (Figure 4): one to an OGC WMS service, recommended by INSPIRE to be used as a view service, while second URL refers to a OGC WFS service, a standard also recommended to be use to download services for vector data. Both URLs can be used to fulfill the metadata elements of the deployed dataset that will be registered in a catalogue service to improve its discovery.

Moreover, the Service Framework permits displaying the deployed datasets on a map viewer; however, note the map viewer's displays a couple of layers, one for protected areas in green and the other burned areas in red, not locally, but accessing directly through the corresponding view services.

Figure 4. After deploying data, service layers are visualized on the viewer.

Besides deploying local data sets, users can process data and extract information through the part of the user interface for accessing geoprocessing services. Although the process model is out of scope here, it is worth mentioning that the developed OGC WPS for this particular use case implements the intersection algorithm, among other topology processes. The valid inputs for this process are described as vector data in OGC GML encodings (Cox et al, 2002). This means that the process is independent of the data model as far as the input data is given in GML format.

Figures 5 and 6 show the steps we described previously in Figure 3, referring to a geoprocessing service execution. When the user selects a certain process in the *Web Processing Service* tab (Figure 5), this process is then invoked through the WPS connector component to retrieve its available processes. These processes are displayed on the interface. The user chooses the intersection process to calculate the protected areas damaged by the forest fires. By doing so, the user is requested to provide the two input parameters, in particular, a couple of geometries (Figure 6). As such, both input parameters are a reference (URL) to the download data service where the user had deployed the previous datasets (Figure 4).

After executing the intersection process, the result of the process (a new dataset with the damaged areas) can be visualized and what is more important, it can be deployed as a standard data service to be shared. In the *Deploy* tab in Figure 6, the user selects the resulting dataset and the Service Deployer (concretely the Data Wrapper) component generates the corresponding geospatial services. Now, forestry technicians are able to share intermediate results in terms of services from the local to the global level. Again, the Service Framework component returns two URLs that point to the generated geospatial services, as in the case of Figure 4.

Figure 5. Geoprocessing process selection for execution.

Figure 6. Process results are deployed as standard services and visualized on the viewer.

The user generated information deployed as standard services have been generically integrated in the GCI. At this moment, other stakeholders can access these data through standardized services for further analysis from other systems and software tools, like web applications on top of the GCI and desktop GIS applications in compliance with OGC specifications (see Figure 7).

Figure 7. Deployed data retrieved by gvSIG (left) and UDig (right) desktop GIS tools. Both implement OGC standards

5. Related Work

This section describes relevant related works and other projects and initiatives that give a notion of the context and support the motivation of this work.

As we previously introduced the purpose of the GEOSS initiative is to achieve improved monitoring of the state of the Earth by expressing interface interoperability as standard service definitions. In order to meet these challenges, it is necessary to establish a common framework within which various systems can communicate and share resources in an interoperable manner. This interoperable framework defines the GEOSS common infrastructure which promotes the use of common principles, rules, techniques, and standards for all GEOSS systems (Butterfield et al., 2008). GEOSS relies on existing infrastructures like SDI and GCI as institutional and technical precedents. As a result, the GEOSS common architecture interoperability is pursued by means of standardized services, mostly implementing OGC interfaces.

In this same context INSPIRE sets out a legal framework for the European SDI, with regard to policies and activities having environmental impact. INSPIRE is actually based on GCIs which have already been set up and are managed by each member state, thereby creating an infrastructure of SDI nodes that are operational at a national, sub-national and thematic level for sharing and accessing data in multidisciplinary and cross-border projects. SDI initiatives as a whole contribute to GEOSS by providing a portfolio of standards, protocols, and interfaces to allow geospatial data to be accessed and exchanged. This set of specifications and standards promoted by INSPIRE considerably enhances interoperability between the services and components provided by SDI nodes.

Therefore and based on this context, the main focus of this work is to extend SDI and GCI, being respectful to current standards, protocol and specifications to maximize interoperability, with new geospatial tools to enable the integration and management of new resources as interoperable geospatial services.

Regarding deployment mechanisms, the technique of wrapping resources as standard web services is described in the research performed by (Fileto, 2001) and (Díaz et al., 2009). Both describe several approaches for geospatial data integration over the web like common wrapper techniques (Roth and Schwarz, 1997) and standards for exchanging geographical data among heterogeneous systems (Albrecht, 1999).

Regarding the availability and reuse of user generated information in environmental applications Rocha et al., (2005) report how the MEDSI tool uses OGC web services to provide and manage geospatial information. Abdalla et al. (2007) present a case study to demonstrate the utility of interoperable web services for disaster management discussing the strengths and weaknesses of leveraging geospatial data interoperability. In contrast, our approach improves these scenarios by the integration of information, not in the form of documents, but through standard services that can be consumed in other scenarios and by any interoperable tool to complete and enrich future analysis.

Brunner et al. (2009) described a system to provide distributed geospatial processing to support collaborative and rapid emergency response, and also a system for storing results in public databases. They also proposed as future work the implementation of relevant standards as OGC WPS.

Fook et al. (2009) presented a Geospatial service architecture that supports the sharing of modelling results that also enables researchers to perform new modelling experiments. They presented the Web Biodiversity Collaborative Modelling Services (WBCMS). The Access Processor context presented in their work supports queries and displays model instances. In the developed prototype within the OpenModeller Project, users are allowed to add metadata and reuse the results of the models. Although they have used OGC interfaces for web data services, they do not mention how processing capabilities can be accessed from outside their architecture. In addition, they proposed to reuse processing results by storing them as files described in catalogues services.

GMES[ix]-funded projects like AWARE[x] or ORCHESTRA[xi] argued for the use of standard services deployed in SDI for effective geospatial information in the disaster management context. In this line Kiehle (2006), Friis-Christensen et al. (2007), Foerster and Schäffer (2007), Díaz et al. (2008), Friis-Christensen et al. (2009), and Granell et al. (2010) propose similar approaches to run geoprocessing on top in SDI using OGC standards to generate information.

Regarding the use of SDI standards like OGC in cyberinfrastructures, (Lee and Percival,2008) define the OGC WPS as a first step of OGC towards providing grid computing capability in GCI. We have demonstrated how WPS services can be used to generate user information and how the Service Framework can assist users in integrating and deploying WPS execution results in the information system as a standardized service.

Using WPS as standard for developing sophisticated processes is described by (Foerster et al., 2010) where the authors implemented and deployed generalization processes in open architectures, sending the results of the processes to be served by OGC WMS. They concluded that these results can be accessed from any WMS client and increase syntactical interoperability.

Regarding the WPS client interface, we found applications like the uDig desktop client extended with WPS client (Foerster and Schäffer, 2007) to access different WPS instances. The *WPS connector* implemented as a Java library is an independent library implementing different versions of the WPS specification that can be integrated in multiple applications.

6. Conclusions and future work

SDI and GCI provide the infrastructure in which spatial wrappers and mediators play a facilitating role. OGC service specifications are the standard *the facto* used to wrap data resources, abstracting data from their machine representation, and becoming accessible to diverse users uniformly. The adoption of these OGC interfaces and standards makes spatial data integration possible in a distributed environment where semantic differences are not too great. Similarly, as mentioned, the use of OGC WPS specification offers abstraction at a syntactic level to offer interoperable operations in GCI. Global initiatives like GEOSS and framework directives like INSPIRE rely on these standards and deployed systems to build global geospatial information infrastructures.

Scientists, technicians and other users generate huge volumes of new information in their daily work. GCI provide them with standard components to discover and access distributed resources in an interoperable fashion. However, these infrastructures do not consider the needed mechanisms to facilitate user collaboration and participation in the building and maintenance of these infrastructures. This provokes a scarcity of scientific resources deployed in standard services and available for sharing in these infrastructures.

We have described mechanisms to assist users in automating the deployment and integration of information as standard services in GCI. This combines two main points for data sharing: first concealing technology from scientists so that they can massively add new information resources to geospatial infrastructures; and second, to wrap this information as OGC services, guarantying the interoperability of the brand-new deployed resources. Furthermore we demonstrate how we can support GEOSS assisting users to deploy resources as standardized services following the European directive INSPIRE’s technical approach to fulfil the requirements of maximizing the interoperability.

In the use case we have demonstrated how GCI users at different levels could benefit from these mechanisms. The rapid and automatic deployment of assessment results in the GCI allows these users to share these results to compare models efficiency using for instance different scales datasets.

Our proposal assists in the back end of the infrastructure, by potentially increasing the number of available interoperable resources. Complementary mechanisms and languages for service workflows descriptions like BPEL could be used jointly with the presented approach. The Service Framework offers proper mechanisms to deploy the results of a process or a workflow execution (described in BPEL) into the GCI infrastructure, with the added value that such resources are deployed as a standard services according to INSPIRE and GEOSS initiatives. The combination of both –workflow engines and the Service Framework– can be used to deploy workflow results in common infrastructures and then improve the rates of sharing and reusing of resources to build comprehensive domain specific applications.

Furthermore, future research on semantic interoperability is needed to reach generally acceptable levels of ad hoc spatial data and process integration and orchestration (Lemmens et al., 2006) and a shared semantic enablement layer for SDI (Janowicz et al., 2010).

Ongoing work focuses on WPS profiles as an attempt to provide a semantic layer on the available processes and their results in order to generate more elaborated descriptions and metadata of the processes execution. To address some of the current limitations, ongoing work is also focusing on improving the automatic deployment of the data resources. Currently we are limited to the data formats supported by Geoserver because we deploy resources using its API. We are investigating the existing transactional interfaces of the different OGC specifications in order to use them to feed services independently of their implementation.

Acknowledgements

This work has been partially supported by the European project EuroGEOSS and by the CENIT España Virtual project through the Instituto Geográfico Nacional (IGN). We would like to thank Jorge Suarez (Conselleria de Medi Ambient, Aigua, Urbanisme i Habitatge), who kindly granted the forest fire data for experimentation purpose. Finally, the authors also wish to thank Dori Apanewicz for proof-reading the manuscript and the anonymous reviewers for their valuable comments.

References

Abdalla, R. , Tao, V., Li, J., 2007. Challenges for the application of GIS Interoperability in Emergency Management. In: Lecture Notes in Geoinformation and Cartography. Editors Jonathan Li, Sisi Zlatanova, Andrea Fabbri City: New York Publisher: Springer-Verlag Pages: 389–405

Ackland, R. (2009). Social Network Services as Data Sources and Platforms for e-Researching Social Networks. *Social Science Computer Review*, 27:481-492.

Albrecht, J. 1999. Geospatial information standards a comparative study of approaches in the standardisation of geospatial information. Computers & Geosciences, 25(1) 9–24.

Brunner, D., Lemoine, G. Bruzzone, L. 2009. Distributed geospatial data processing functionality to support collborative and rapid emergency response. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2(1): 33-46.

De La Beaujardiere, J. (ed.) 2004. Web Mapping Service Implementation Specification, Version 1.3.0. Open Geospatial Consortium.

Béjar, R., Latre, M. Á., Nogueras-Iso, J., Muro-Medrano, P. R., Zarazaga-Soria, F. J. 2009. An architectural style for spatial data infrastructures. International *Journal of Geographical Information Science*, 23(3), 271-294.

Bishop, I.D. et al,. 2000. Spatial data infrastructures for cities in developing countries. Cities, 17(2), 85-96

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

Butterfield, M.L., Pearlman, J. S., Vickroy, S. C. 2008. A System-of-Systems Engineering GEOSS: Architectural approach. *IEEE Systems Journal* 2(3) 321-332.

Buyya, R., Yeo, C. S., Venugopal, S., Broberg, J., Brandic, I. 2009. Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility. *Future Generation Computer Systems* 25(6) 599-616.

Castells, M. 2001. The Internet Galaxy: Reflections on the Internet, Business and Society. Oxford: Oxford University Press.

Cox, S., Cuthbert, A., Lake, R., Martell, R. (Eds). 2002. OpenGIS® Geography Markup Language (GML) Implementation Specification, version 2.1.2. Open- GIS® Implementation Specification, 17 September 2002. OpenGIS Project Document Number 02-069

Craglia, M, Kanellopoulos, I., Smits, P.2007. Metadata: where we are now, and where we should be going. Proceedings of 10th AGILE International Conference on Geographic Information Science 2007. Aalborg University, Denmark

Davis, C.A. Jr., Fonseca, F.T., Câmara, G.2009. Beyond SDI: Integrating Science and Communities to Create Environmental Policies for the Sustainability of the Amazon. International Journal of Spatial Data Infrastructures Research, 2009, Vol. 4, 156-174

Deelman, E., Gannon, D., Shields, M., Taylor, I. 2009. Workflows and e-Science: An overview of workflow system features and capabilities. *Future Generation Computer Systems* 25 (5) 528-540.

Denzer, R., 2005. Generic integration of environment decision support systems – state-of-the-art. *Environmental Modelling & Software* 20 (10) 1217-1223.

Díaz, L., Granell, C., & Gould, M. 2008. Case Study: Geospatial processing services for web-based hydrological applications. In J.T. Sample, K. Shaw, S. Tu, M. Abdelguerfi (Eds.): Geospatial Services and Applications for the Internet. Springer. ISBN 978-0-387-74673-9.

Díaz, L., Granell, C., Gould, M. 2009. Spatial Data Integration over the Web. In V. E. Ferraggine, J.H. Doorn, L.C. Rivero (Eds.) Handbook of Research on Innovations in Database Technologies and Applications: Current and Future Trends, Information Science Reference (Hershey, . 2009), Volume II, pp. 325-333.

Fileto, R. 2001. Issues on Interoperability and Integration of Heterogeneous Geographical Data. Proceedings of 3rd Brazilian Symposium on GeoInformatics. 2001.

Flanagin A. J., Metzger, M. J. 2008. The credibility of volunteered geographic information. *GeoJournal* 72 137-148

Foerster, T., Shäffer, B. 2007. A Client for Distributed Geo-processing on the Web. In Proceedings of International Symposium on Web and Wireless Geographical Information Systems (W2GIS 2007), Cardiff (UK). Berlin: Springer LNCS 4857, (2007) pp. 252-263.

Foerster, T., Lehto, L., Sarjakoski, T., Sarjakoski, L.T., Stoter, J. 2010. Map generalization and schema transformation of geospatial data combined in a Web Service context. Computers, Environment and Urban Systems 34(1): 79–88

Fook, K. D., Monteiro, A. M. V., Câmara, G., Casanova, M. A., Amaral, S. 2009. Geoweb Services for Sharing Modelling Results in Biodiversity Networks. *Transactions in GIS* 13(4) 379-399.

Foster, I. 2005. Service-oriented science. Science 308 (5723) 814-817.

Friis-Christensen, A., Ostländer, N., Lutz, M., Bernard, L., 2007. Designing Service Architectures for Distributed Geoprocessing: Challenges and Future Directions. Transactions in GIS 11 (6) 799-818

Friis-Christensen, A., Lucchi, R., Lutz, M., Ostländer, N., 2009. Service chaining architectures for applications implementing distributed geographic information processing. International Journal of Geographical Information Science, 23 (5) 561-580.

Giovando, C., Whitmore, C., Camia, A., San Miguel, J., Boca, R., Lucera, J. 2010. Geospatial support to forest fire crisis managements at the European level. Proceedings of Gi4DM 2010 Conference on Geomatics for Crisis Management. Torino, Italy, Feb 2010.

Goodchild, M.F. (2007). Citizens as voluntary sensors: Spatial data infrastructure in the world of Web 2.0. *International Journal of Spatial Data Infrastructures Research*, 2: 24-32.

Granell, C., Díaz, L., Gould, M. 2010. Service-oriented applications for environmental models: Reusable geospatial services (in press). Environmental Modelling and Software, 25(2): 182-198, 2010. ISSN: 1364-8152

Hey, T., Trefethen, A. E., 2005. Cyberinfrastructure for e-Science. *Science* 308 817-821.

INSPIRE EU Directive. 2007. Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE). Official Journal of the European Union, L 108/1, Volume 50, 25 April 2007. http://eur-lex.europa.eu/JOHtml.do?uri=OJ:L:2007:108:SOM:EN:HTML.

INSPIRE Network Services Drafting Team. 2008. Draft Technical Guidance to implement INSPIRE View Services

INSPIRE Network Services Drafting Team. 2009. Draft Technical Guidance for INSPIRE Download Services

INSPIRE Network Services Drafting Team. 2009b. Technical Guidance for INSPIRE Discovery Services

ISO, 2003 ISO, 2003. Geographic Information—Metadata. ISO 19115:2003, International Organization for Standardization (ISO).

Janowicz, K., Schade, S., Bröring, A., Keßler, C., Maué, P., Stasch, C. 2010. Semantic Enablement for Spatial Data Infrastructures. Transactions in GIS, 14(2): 111-129

Kiehle, C. 2006. Business logic for geoprocessing of distributed geodata. Computers & Geosciences 32 (10) (2006), 1746-1757.

Krishnan, S., Bhatia, K. 2009. SOAs for scientific applications: experiences and challenges. *Future Generation Computer Systems* 25(4) 466-473.

Lee, C., Percivall, G. 2008. Standards-Based Computing Capabilities for Distributed Geospatial Applications. Computer 41(11): 50-57.

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., Rajabifard, A., Williamson, I..2008. Spatially enabling governments through SDI implementation. International Journal of Geographical Information Science. Vol. 22, No. 1, (2008) 5–20

Mansourian, A., Rajabifard, A., Zoej, M.V. 2005. Development of a Web-Based GIS Using SDI for Disaster Management, in: *Geo-information for Disaster* Management. 2005, pp. 599-608.

Nayak, S.; Zlatanova, S. (Eds.). Remote Sensing and GIS Technologies for Monitoring and Prediction of Disasters. 2008, X, 272 p, Springer.

Nebert, D., Whiteside, A., Vretanos, P. (eds.) 2007. OpenGIS Catalogue Services implementation, Version 2.0.2. Open Geospatial Consortium.

Nogueras-Iso, J., F.J. Zarazaga-Soria, R. Béjar, P.J. Álvarez, P.R. Muro-Medrano, 2005. OGC Catalog Services: a Key element for the development of Spatial Data Infrastructures. *Computers and Geosciences*, 31(2) 199-209.

Omran, E.L.E., van Etten, J. (2007). Spatial-data sharing: Applying social-network analysis to study individual and collective behaviour. *International Journal of Geographical Information Science*, 21(6), 699-714.

Pultar, E., Raubal, M. Cova, T. J., Goodchild, M.F. 2009. Dynamic GIS Case Studies: Wildfire Evacuation and Volunteered Geographic Information. Transaction in GIS 13(s1): 85-104.

Rajabifard, A., Feeney, M-E. F., Williamson, I. P. 2002. Future directions for SDI development. International Journal of Applied Earth Observation and Geoinformation 4 (2002) 11–22

Rocha, B. Cestnik, M.A. Olivera, Interoperable Geographic Information Services to Support Crisis Management in J. Li and C.Vangenot (Eds): 5th International Workshop on Web and Wireless Geographical Information Systems, LNCS 3822 Springer-Verlag. Laussane, Switzerland, 2005, pp 246-255

Roth, M.T., Schwarz, P. Don’t Scrap It, Wrap It! A wrapper Architecture for Legacy Data Sources. Proceedings of the International Conference on Very Large DataBases, (1997), 266–275.

Schäffer, B., Foerster, T. 2008. A client for distributed geo-processing and workflow design. Journal of Location Based Services, 2(3): 194-210.

Schut, P., (eds.), OGC Web Processing Service (WPS) version 1.0.0. OGC Standard Document, Open Geospatial Consortium. 2007

Vretanos, P. 2005. OGC Web Feature Service Implementation Specification, Version 1.1.0. Open Geospatial Consortium. 2005

Whiteside, A., Evans, J.D.(eds).2008. Web Coverage Service (WCS) Implementation Standard. Open Geospatial Consortium. 2008.

Yang, C., Raskin, R., Goodchild. M., Gahegan, M. 2010. Geospatial Cyberinfrastructure: Past, present and future. *Computers, Environment and Urban Systems* 34(4) 263-358.

[i] http://www.eurogeoss.eu

[ii] http://www.earthobservations.org/

[iii] http://www.opengeospatial.org

[iv] http://www.inspire-geoportal.eu/index.cfm/newsid/4204

[v] http://geoserver.org/display/GEOS/Welcome

[vi] http://geoserver.org/display/GEOSDOC/RESTful+Configuration+API

[vii] http://geonetwork-opensource.org

[viii] http://www.cma.gva.es/web/indice.aspx

[ix] http://www.gmes.info

[x] http://www.aware-eu.info

[xi] http://www.eu-orchestra.org

Fig1

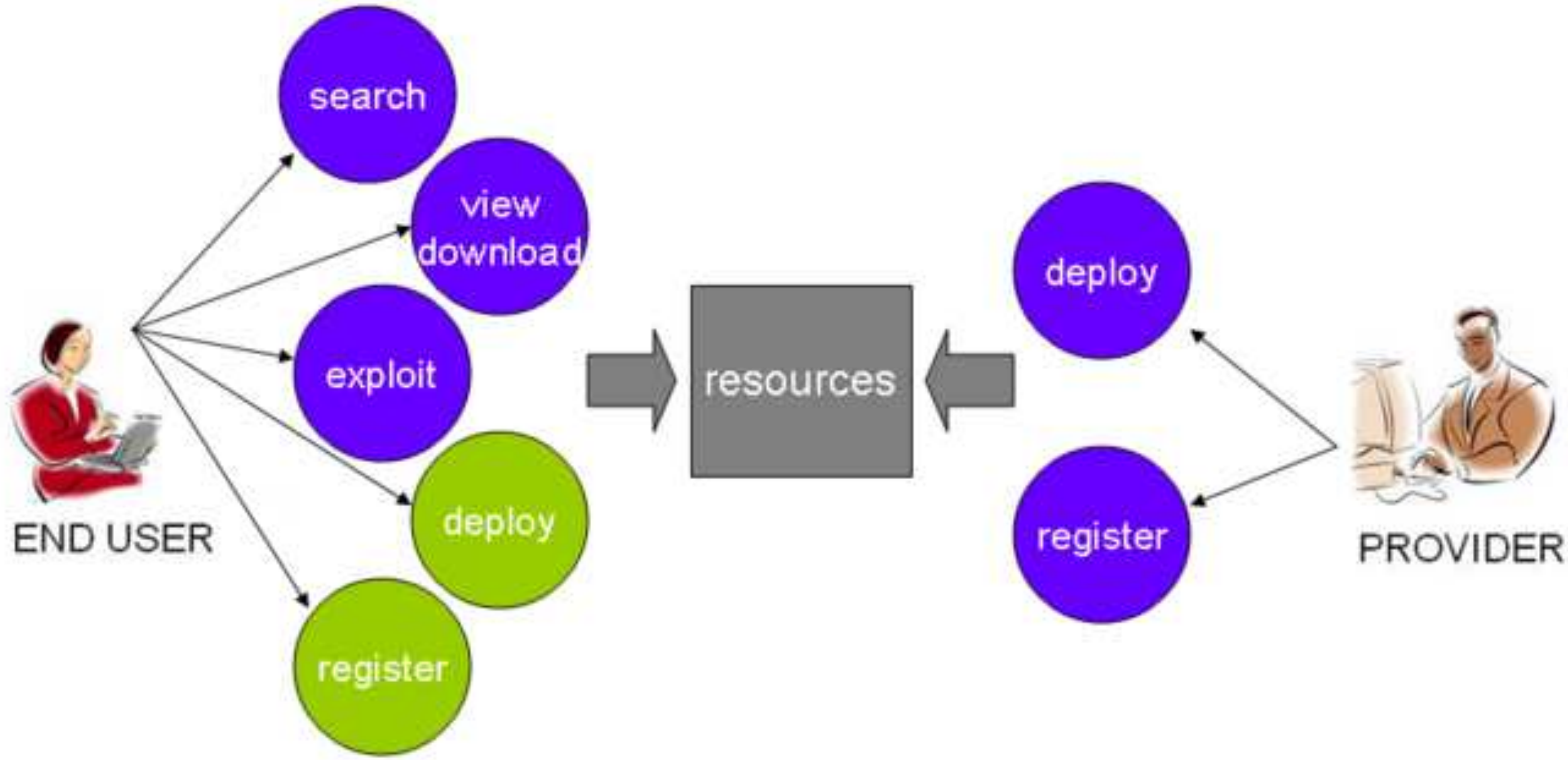

search
view
download
exploit
deploy
register
END USER
resources
deploy
register
PROVIDER

Fig2
Click here to download high resolution image

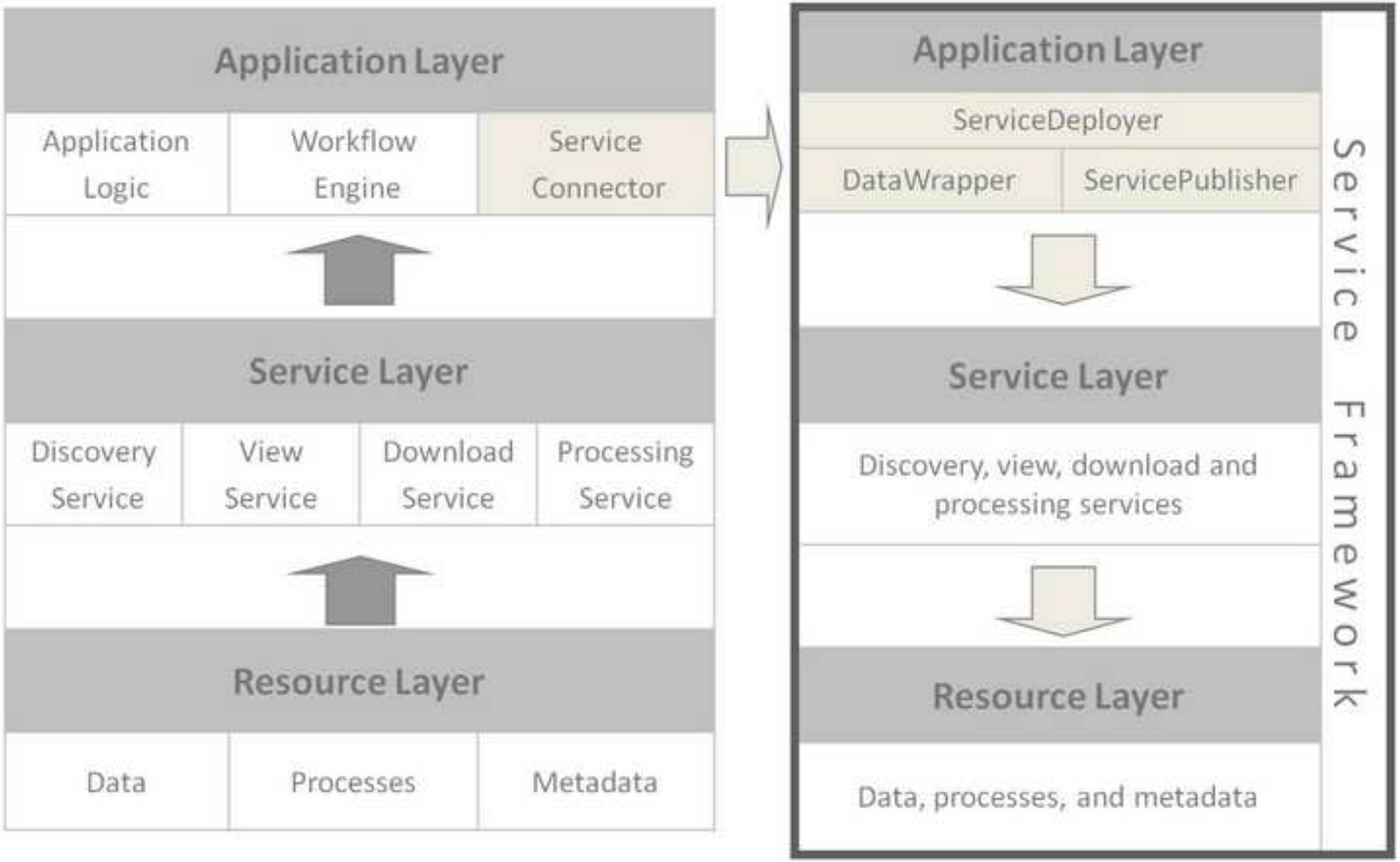


Fig3

Click here to download high resolution image

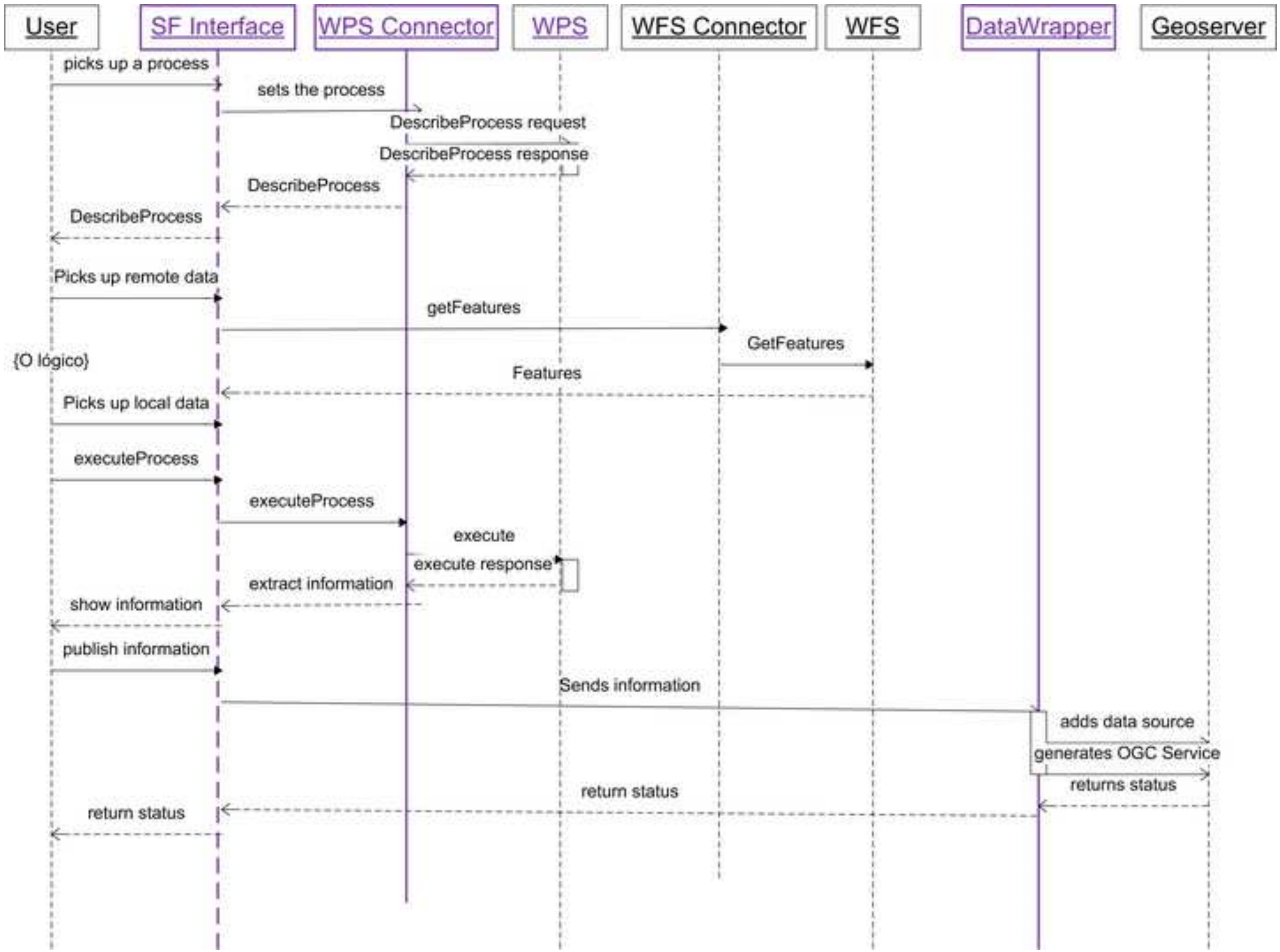


Fig4
Click here to download high resolution image

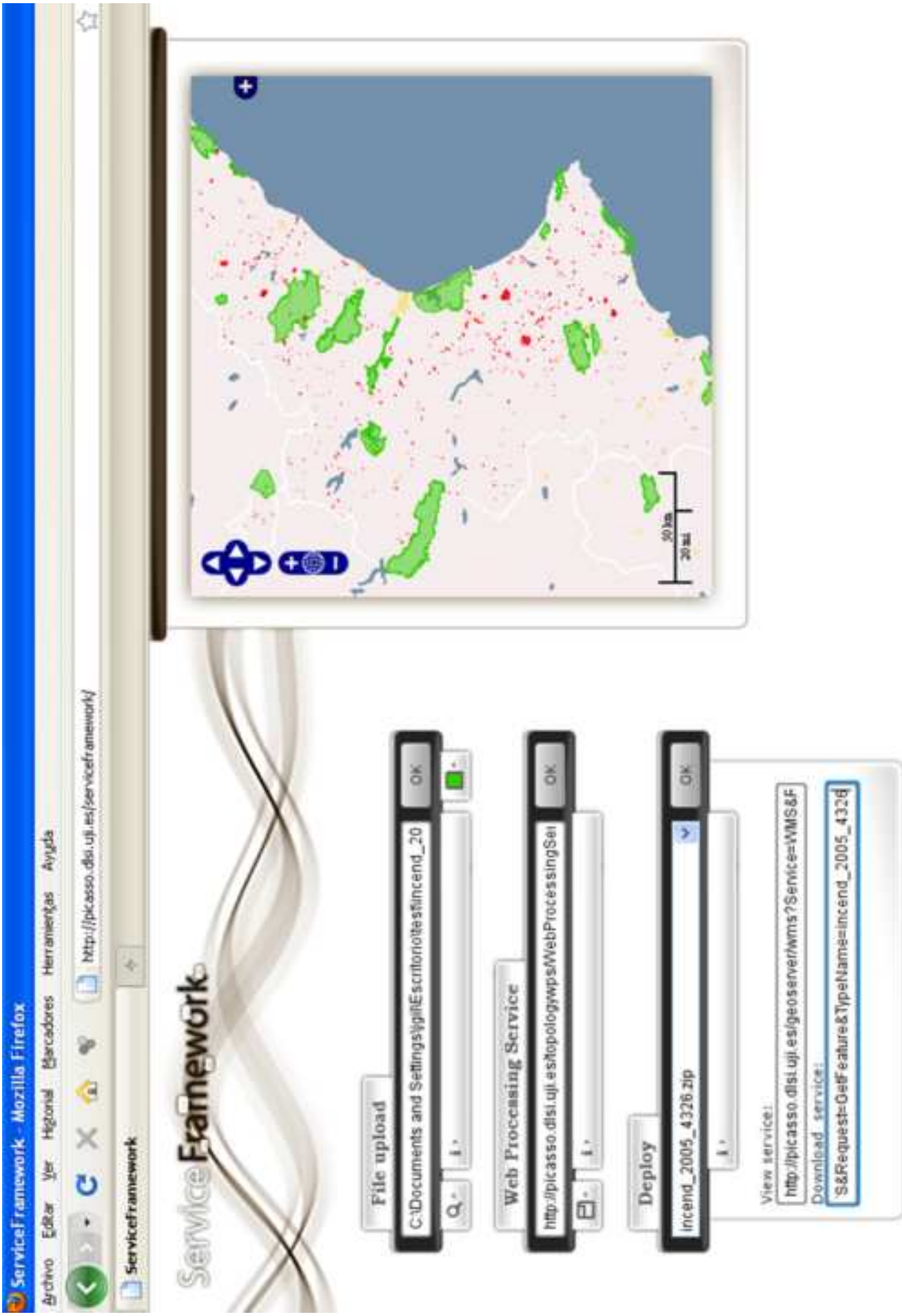
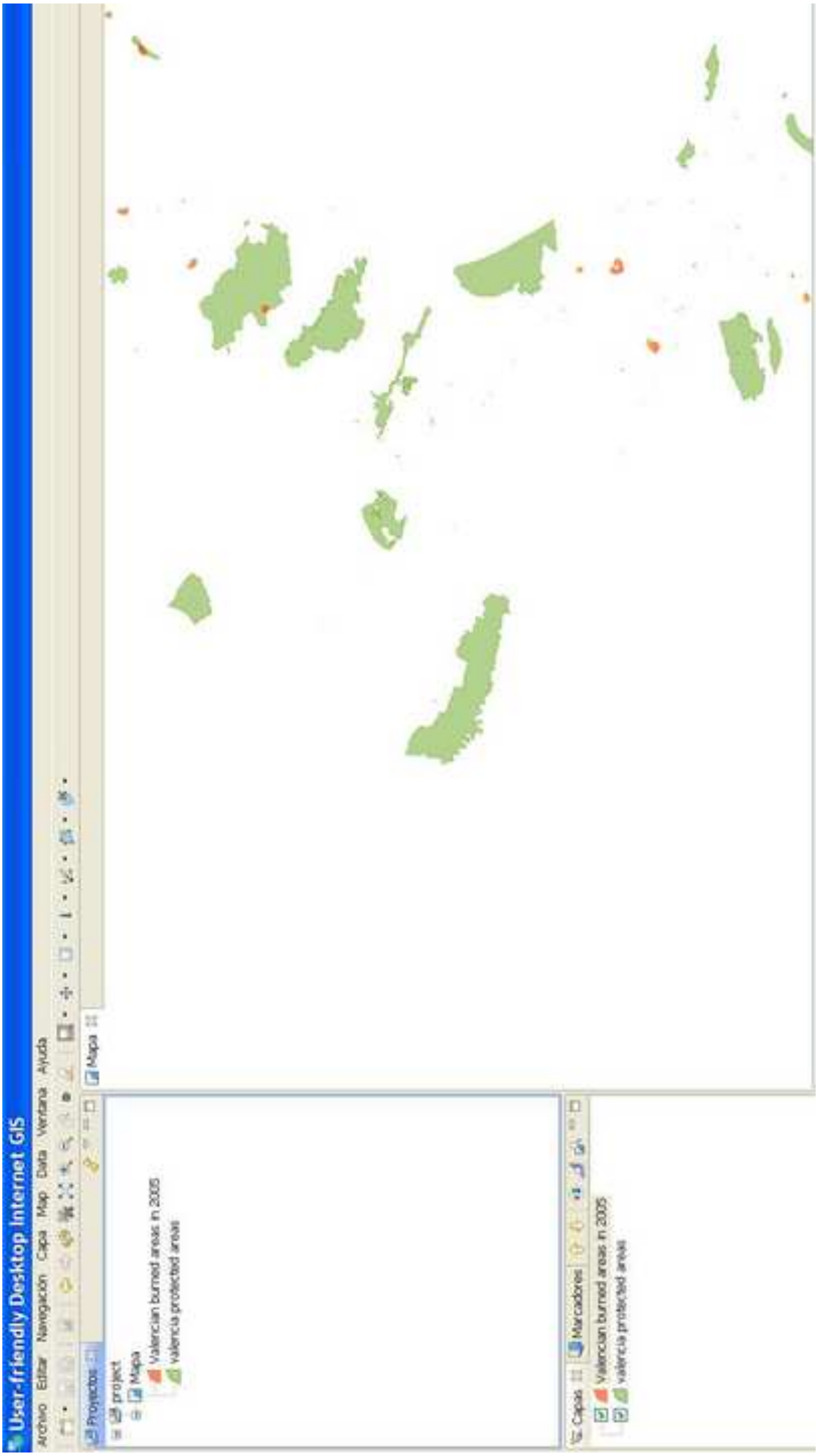
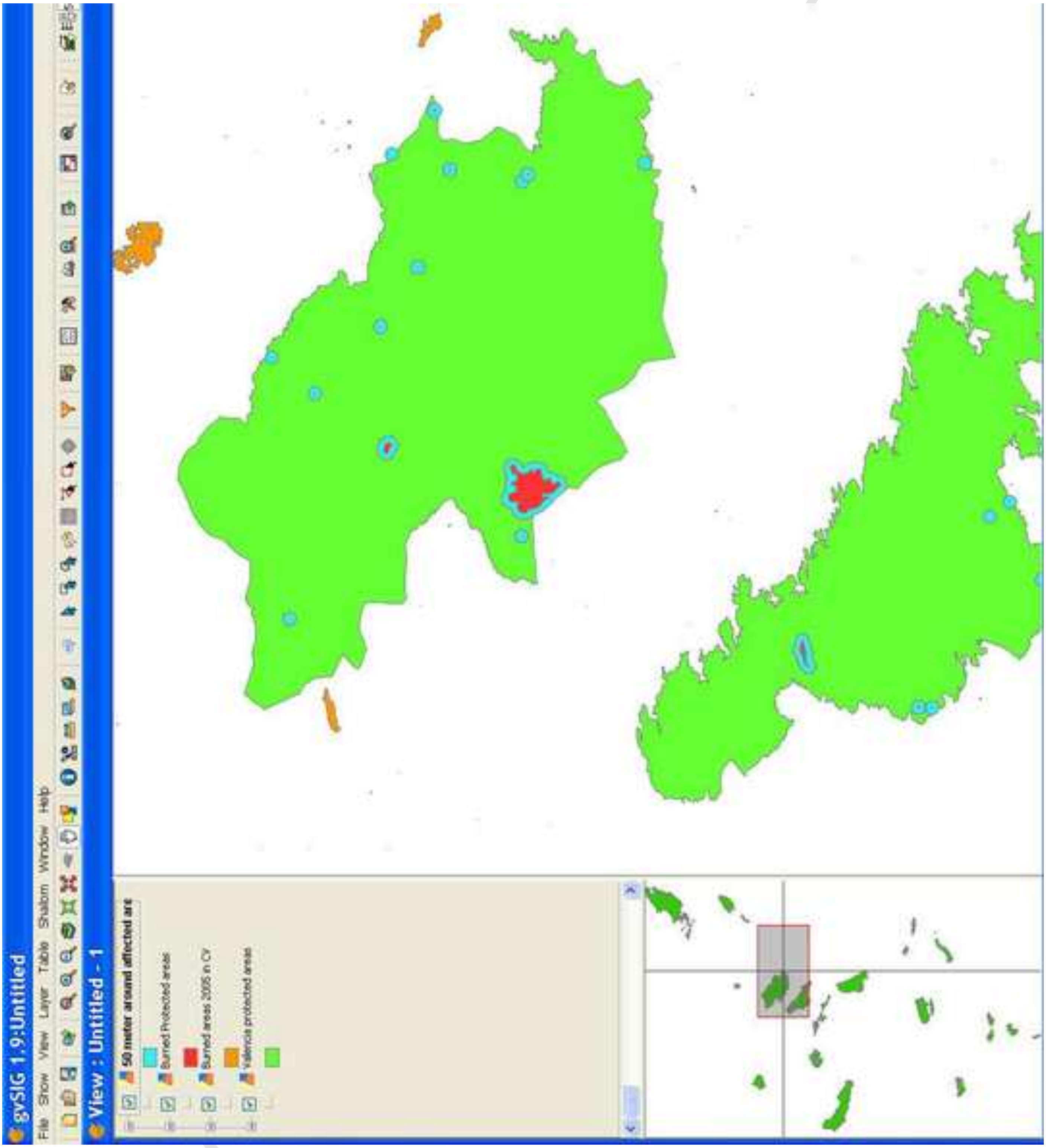


Fig5
Click here to download high resolution image

Fig6


User-friendly Desktop Internet GIS
Archivo Editar Navegación Capa Map Data Ventana Ayuda
Proyectos
project
Mapa
Valencian burned areas in 2005
valencia protected areas
Capas
Marcadores
Valencian burned areas in 2005
valencia protected areas


gvSIG 1.9:Untitled
View : Untitled - 1
50 meter around affected are
Burned Protected areas
Burned areas 2005 in CV
Valencia protected areas

Dr. Laura Díaz graduated in Computer Engineering (2000) from the Universitat de València. She obtained her Master in Intelligent Systems (2008) and received her doctorate in Geospatial Science (2010) at the Universitat Jaume I of Castellón where she is currently a postdoctoral researcher. Her research lines are novel approaches for improving geoservices descriptions and deployment, distributed geoprocessing and service interoperability. She has participate in R&D projects, working for the Institute of Robotics of the Universitat de València and specialist companies in Geospatial Information Systems such as Geodan (The Netherlands) and Iver TI (Spain). ESMI, AWARE, and currently EuroGEOSS are among the most important international projects she has taken part in.

Dr. Carlos Granell is a postdoctoral researcher at the UJI, where he graduated in Computer Engineering (2000) and received his doctorate in Computer Science (2006). His lines of research are centred on Geographic Information Systems (GIS), Spatial Data Infrastructures (SDI), interoperability, geoprocessing services, workflow, and the composition and reuse of geographic services. He has taken part in several public funded research projects, both at a Spanish and a European level, the most important of which are the European projects ACE-GIS and AWARE. He has carried out research stages at SINTEF (Norway) and at the ITC (the Netherlands).

Dr. Gould earned his Ph.D. in geography (GIS) from the National Center for Geographic Information and Analysis (NCGIA), University at Buffalo. Most recently, he was an interdisciplinary professor of information systems and co-creator of an Erasmus Mundus-funded geospatial technologies master's program at Universitat Jaume I, Castellón, Spain. Gould's research expertise includes creating Spatial Data Infrastructures (SDI); researching standards-based interoperability with the Open Geospatial Consortium, Inc. (OGC); and developing geospatial Web services.

Prof. Joaquín Huerta is an associate professor in the Department of Information Systems at UJI, where he teaches GIS and Internet Technologies. He holds a PhD in Computer Science from Jaume I University. He is currently leading two EU projects at UJI: EUROGEOSS (FP7) and eSDI-NET+ (FP6) and a Spanish project "España Virtual" funded by Cenit Programme. His current research interest are Geospatial Technologies and Computer Graphics. He is Director of the Master in Geospatial Technologies funded by Erasmus Mundus program. In addition to academic activities Dr Huerta is founding board member of an Internet service provider and, thus, possesses considerable business experience as well as experience in systems integration

Diaz

Granell

Huerta
Click here to download high resolution image