Interoperability through Web Services: Evaluating OGC Standards in Client Development for Spatial Data Infrastructures

Leonardo Lacerda Alves¹, Clodoveu A. Davis Jr.¹

¹Pontifícia Universidade Católica de Minas Gerais
Instituto de Informática – Mestrado em Informática
leonardo@lacerda.eti.br, clodoveu@pucminas.br

Abstract. Interoperability is one of the most important challenges related to GIS. Through the last years, research on interoperability has evolved from the simple off-line exchange of standardized-format files, through the establishment of spatial data clearinghouses, and to the first initiatives in the treatment of semantic aspects of data. The Open Geospatial Consortium (OGC) has proposed a number of standards to that respect, with the intention of promoting interoperability through the use of services. However, issues regarding fault tolerance, server-independent implementation, delayed-time transactions, privacy, and others reflect the need for further study and discussion. This paper discusses the current status of service-oriented architectures as applied to interoperable GIS, or, more specifically, to the implementation of local spatial data infrastructures (LSDI). A prototype designed to test the services abstract model by simulating a real-world use case is presented. Our conclusions indicate that some improvements may be helpful in the development of GIS/LSDI clients, as well as in the communication between thin clients and servers, supporting delayed-time transactions through asynchronous communication, and enabling clients to support fault-tolerant mechanisms without provider-dependent solutions.

1. Introduction

Interoperability is one of the most important challenges related to GIS. Through the last years, research on interoperability has evolved from the simple off-line exchange of standardized-format files, through the establishment of spatial data clearinghouses, and to the first initiatives in the treatment of semantic aspects of data. Practical interoperability, however, is still hampered by the need to agree on standards, and to develop appropriate tools and methods.

The Open Geospatial Consortium (OGC) has proposed a number of standards to that respect, with the intention of promoting interoperability through the use of services [Percivall 2003]. However, OGC’s definition of Web-based services for spatial information predates the World-Wide Web Consortium’s definition of the Web service architecture [World Wide Web Consortium 2004]. The necessary adjustments between OGC’s and W3C’s proposals are still under way.

Meanwhile, we observe that a number of difficulties arise when someone tries to effectively implement the interoperability-through-services approach. Issues regarding fault tolerance, server-independent implementation, delayed-time transactions, privacy,
and others reflect the need for further study and discussion. In fact, studies about the conformance of OGC standards to the distributed systems development are still scarce.

This paper discusses the current status of service-oriented architectures as applied to interoperable GIS, or, more specifically, to the implementation of local spatial data infrastructures (LSDI). Most existing spatial data infrastructures refer to regional or country-wide data, while LSDI deals with a more complex and rich set of geographic data [Nedovic-Budic et al. 2004, Rajabifard et al. 2000]. Thus, LSDI involve a wide variety of services, while also dealing with users of a wide range of devices, such as PDAs, cell phones and personal computers.

For this discussion, we defined a real-world use case based on an urban context, and developed a services prototype following OGC’s abstract model. We used this prototype to assess the engineering guidelines for the server and client development, according to the viewpoints established by the OGC Reference Model [Percivall 2003]. To solve some of the limitations and issues that we have identified, we proposed and developed special infrastructure services which illustrate some deficiencies of OGC specifications. Nevertheless, we do not imply here that such services should become part of the standard. Using the proposed services, it is possible to enable asynchronous communication between clients and services, to access data provided by clients, and to improve on critical points of the services-oriented architecture, such as recovery from failures, dynamic service chains creation, and others, while staying within the OGC Reference Model.

This paper is organized as follows. Section 2 presents concepts about Spatial Data Infrastructures (SDI), in general, and Local Spatial Data Infrastructures (LSDI), in particular, including the ideas behind configuring SDI and LSDI as services. Section 3 introduces our discussion as to the required functionality of a local SDI, and the way to achieve that using the aforementioned infrastructure services. Finally, Section 4 presents our conclusions and indicates some research directions from the concerns presented here.

2. Related Work

2.1. Spatial Data Infrastructures

Spatial Data Infrastructures (SDI) constitute a set of policies, technologies and standards that interconnect a community of spatial information users and related support activities for production and management of geographic information [Phillips et al. 1999]. The idea behind SDI involves avoiding redundant effort and reducing production costs for new and existent datasets through the sharing of resources. In order to achieve this, it is very important that the various partners have convergent interests, agree on common rules and are allowed to make use of data or information produced by others.

Geographic information from one or several partners can be consolidated and thus form important resources for high-level decision-makers. In this case, SDI can be seen as a set of building blocks, as defined by Rajabifard et al [Rajabifard et al. 2000], in which SDI hierarchies are built through the exchange and consolidation of information from corporate and local levels, to regional and global levels. In this hierarchy, lower levels provide detailed information that helps in the consolidation of the upper, more general, levels [Rajabifard and Williamson 2001, Jacoby et al. 2002, Man 2006].

Guiding the technology standardization and, consequently, defining the key ele-
ments for spatial data infrastructures, a number of standards were proposed by the OGC, through a framework called OGC Reference Model [Percivall 2003]. This framework has been implemented successfully in state, national and regional scales [Deegree 2006, Demis 2006, Skylab 2006], but reference cases of local SDI are still scarce. Furthermore, most of them are not OGC-compliant [Kumar et al. 2005, Mansourian et al. 2006].

Local SDI has the potential of bringing a numerous group of users together, each of which with distinct needs [Davis Junior 2005]. This characteristic is in part responsible for an increased level of complexity in local SDI development and deployment [Bishop et al. 2000, Masser 2005]. Indeed, LSDI is valuable to all of the other SDI levels as a detailed information source, and the implementation of actual data sources in the local level should evolve simultaneously with, and guided by, the new demands for Geographic Information (GI) from a new range of users, such as business travelers, citizens, small companies and others [Masser 2005].

SDI can be implemented by chaining services of different sources [Alameh 2003] and integrating software components [Granell et al. 2005] that can be found in geoportals [Maguire and Longley 2005]. The emphasis on services has increased since the emergence of Web services and of the Service-Oriented Architecture (SOA) in the field of distributed systems development. This led Bernard and Craglia [Bernard and Craglia 2005] to propose a new translation for the the SDI acronym, to mean Service-Driven Infrastructures.

2.2. Web Services and OGC Services

Web services are usually seen as Web-based enterprise-wide or inter-organizational applications that use open standards (mostly based on XML) and transport protocols to exchange data with clients, thus forming a loosely-coupled information systems architecture [Ferris and Farrell 2003]. Web services, as specified by the World Wide Web Consortium (W3C) [World Wide Web Consortium 2004], use the Hypertext Transfer Protocol (HTTP) in the transport layer. This protocol only allows the client to perform synchronous calls. This characteristic represents a problem when dealing with delayed-time transactions and call resume.

Some initiatives try to emulate asynchronous calls in Web services through the use of listeners that receive responses and forward them to the requester [Ruth et al. 2005], usually over non-HTTP protocols [Brambilla et al. 2004]. Other strategies involve the use of the electronic-mail protocol, SMTP [Chung et al. 2006]. However, as far as we know, no pattern enables asynchronous communication using only Web service standards.

Regarding geographic information specifically, the OGC Web Services specification (OWS) [Whiteside 2005] describe a set of functionalities and components which support standardized information exchange. Among the OGC services, there are asynchronous communications only for sensor services [Botts et al. 2006] such as the Web Notification Service, but the communication protocol is not specified in the standard.

In this work, we are particularly interested in LSDI, i.e., a network of clients and services that deal with an urban context and typical urban applications. The urban-specific OGC proposed platform is called Open Location Service (OpenLS) [Mabrouk 2004], in which services such as directory (yellow pages service), routing services, geocoding services, and others are grouped. A proposal for urban services of interest as LSDI elements
can be found in [Davis Júnior and Alves 2005]. The variety of possibilities for using such urban services to build compact and useful geographic information-based applications is staggering. The next section presents one of them.

2.3. Use Case: a Consumer Travel Assistance Application

The OGC Reference Model [Percivall 2003] defines a scenario entitled “consumer travel assistance” [p. 83] through which a GI consumer uses a mobile client to (1) get its actual position or location; (2) get the destination address, given a telephone number; (3) get a location, given the destination address; (4) get a route between the origin and the destination; (5) determine the traffic, weather and road conditions along the way; and (6) obtain real-time travel advice. Of course, such a task potentially requires a large amount of geographic data, possibly from various sources. This is something that is well above the capabilities of a typical mobile client, thus requiring a new computational model. In this paper, we extend this scenario and use it to implement OpenLS platform clients and an OWS framework, in order to evaluate the OGC Reference Model for LSDI client development.

Our modified scenario involves two basemap services, one routing service, one geocoding service, one directory service for telephone numbers, one directory service for companies, one public transportation service, one public-services service, two emergency services, one gateway for personal location, and two kinds of clients (thin and thick).

While designing LSDI clients to deal with such a variety of services, we have identified some limitations of OGC specifications, more specifically in important engineering requisites such as service availability assurance, recovery in case of failures, privacy control, communication cost reduction and transparent support for the peculiarities of different kinds of client devices.

Even though we have not yet fully developed this scenario as to its actual functionality, we have worked on determining and implementing the ideal communications protocols for each situation, so that some of the limitations inherent to OGC’s architecture can be adequately solved beforehand. A prototype was implemented, based mostly on Web service aspects, and leaving the many related GIS issues for the near future. The goal of this prototype is to make sure the required transactions are adequately designed, and all communications issues are solved. The next section presents and discusses the alternatives we have conceived and implemented to that effect.

3. Innovative Services for LSDI

Given the anticipated needs of urban GI applications and the main limitations of GI Web services, some novel infrastructure services are useful to improve the LSDI clients development and the LSDIs themselves.

In the next three subsections, we present the infrastructure services we conceived and developed. The first, called Data Exchange Service, is used for persistence maintenance between clients and servers. The second, called Client Access Service, is used for information exchange and access control to clients by servers and other clients. The third, called Transaction Control Service, is used by clients and servers to improve engineering capabilities such as failure recovery, dynamic service chains creation, workflow
definition, and others. The fourth subsection presents details about the prototype implementation and its analysis.

3.1. Data Exchange Service

The Data Exchange Service (DXS) supports interactions between clients and servers (client to server communication), different servers (server to server communication), and different clients (client to client communication) as a workspace where data may be freely stored and retrieved. The objective of this service is to reduce the volume of interactions between clients and servers, and to minimize the connection costs in service invocations and data retrieval, even when failures occur.

Figure 1 shows an example of the DXS in action. A service A, replicated in A’ and A” for redundant availability, as it is invoked by a client. Initially, the request parameters are sent to a DXS (1). Next, the ordering of the procedures is defined by the client, in the form of a workflow (2). An invocation process (3) reaches service A, or resorts to services A’ or A” in case of failure. In any case, the parameters are recovered from the DXS (4), and the results are stored in that server for future retrieval by another service, or by the client (5).

![Figure 1. Data Exchange Service in the invocation of a replicated service](image)

The Data Exchange Service works in three different ways. The first way effectively establishes asynchronous communication between the client and the target service. The DXS mediates the communication, so that the client does not have to wait online for a response from the target service. When the processing is through, the client receives a message from the DXS, and is then allowed to retrieve the results.

The DXS also works as a temporary repository for intermediate responses in a service chain. Intermediate results are kept by the DXS for the benefit of services along the chain, but the client is only allowed access to the final results, as they become ready.

Furthermore, the DXS supports failure recovery, since it can keep information on the status of a service chain, along with intermediate results. With this, recovery and
continuation from the client side is possible by reinvoking any service that fails, using the stored parameters, thus avoiding reinitialization of the entire chain.

To understand and to specify how the Data Exchange Service works, we defined the client and service invocation patterns using an UML (Unified Modeling Language) sequence diagram. Figure 2 presents the diagram.

![Figure 2. UML Sequence Diagram for a System using DXS](image)

Messages 1 and 7, between the client and the DXS, are the same generic processes 1 and 5 shown in figure 1, which initiate the processing and collect the final results, respectively.

Messages 2, 3, 4, 5, and 6 invocate all needed services at the beginning of the process, when the client stops and the services keep running. Next, as the Locating and Telephone services can be executed in parallel, messages 2.1, 2.2, 3.1, and 3.2 can occur in any order.

Finally, messages between 4.1 and 6.2 occur sequentially. Notice that all intermediate data exchange would pass through the client if the DXS was absent. However, using this service, such intermediate response traffic is avoided. Most of the service chain can be invoked in parallel to reduce client work, and the responsibility for most invocation callbacks is transferred to the DXS.

In the highest conceptual level, the Data Exchange Service has thus the function of intercepting service responses and forwarding them to other services (including other DXS) or to the client. In this sense, the communication costs for thick and rich clients remains the same, while costs for thin clients fall drastically. The length of the lifeline remains the same, regardless of the involvement or not of the DXS. Therefore, the DXS offers advantages only when the client has energy limitations or it is slower than the DXS server.
It would be possible to implement DXS-like persistence through simple Web services, but a standardized interface that functions as an infrastructure service is important to ensure the independence between clients and OGC-services providers.

As presented, the DXS may suffer from a number of security issues, such as unauthorized access to data from third parties. However, protocol enhancements can ensure authorized access, so that only the participating services can retrieve intermediate data, and not even the client is allowed access to privileged or confidential information [Belussi et al. 2004].

3.2. Client Access Service

As mentioned in the previous section, the Data Exchange Service enables asynchronous communication between the client and the target service, notifying the client when the results are ready to use. However, asynchronous calls are not supported in W3C Web services using HTTP [Brambilla et al. 2004, Ruth et al. 2005]. Among OGC services, only the Web Notification Service implements asynchronous answers, even though it also does not use a HTTP interface [Botts et al. 2006].

Code 1 A call that return a transaction ID

```java
transactionID = service.do(params);
```

In an urban context, asynchronous services are often necessary, working as delayed-time transactions for long service chains. They are also necessary in particular applications, such as using servers as sensors to send information to clients without counting on connection-oriented communication.

Code 2 Getting the service result

```java
while (!service.isReady(transactionID))
    self.sleep(1000);
result = service.get(transactionID);
```

Even though normally a client does not have a valid IP (Internet Protocol) address, it is able to access HTTP resources. However, without an IP address it is inaccessible from other clients or services, receiving data only during its request connections. We propose an alternative to reach clients without an IP address. A client can invoke a service using a method that returns a transaction Id, but not the final result. This is illustrated in code 1.

Code 3 The client define a method for receive a notification when the result is ready

```java
transactionID = 
    service.do(responseMethod, responseURI, 
               params)
```

In a second moment, the client invokes the service again through a method that returns the final results, giving the transaction Id, as illustrated in code 2.
However, getting results is only possible when the processing has been concluded, and the client uses up network and processing resources to poll the service for a response. Figure 3 presents this communication pattern.

If the client has a valid IP address and implements the functionalities of HTTP servers, we have a different situation. The client waits for responses through a DXS-like service, precisely until it receives a response from the service, informing that the results are ready. Code 3 illustrates a call that specifies a response method (HTTP, SMTP, SMS, etc) as part of its parameters, along with an URI (a address through which the client expects the notification when the process had ends).

**Code 4 Waiting for Notification from the Gateway or Other Service**

```python
while (self.waiting(transactionID)) {
    // do nothing
}
result = service.get(transactionID);
```

Figure 4 shows the described communication protocol, and the code 4 illustrates the last call from the client to the service when the service processing becomes ready.
The last discussed pattern is dependent on a local gateway service such as a DXS, and it is adopted when the client does not have a valid IP address, or when the client is protected behind a firewall, or when the client does not want to be identified. In this case, the client can access the target service directly or indirectly, but the notification process always occurs between the chosen gateway and the client, as illustrated in figure 5. The client implements the same code presented in code 3 and in code 4. We call this approach the Client Access Service (CAS).

![Figure 5. Client without valid IP address using a Gateway as mediator](image)

The Client Access Service offers means to use clients as sensors for others, and means to enable the client to work as a Data Exchange Service provider (specially in the case of thick clients). In both cases, the client is able to provide some information to other clients and services and to support a number of concurrent processes, with adequate security and privacy constraints.

### 3.3. Transaction Control Service

Through the two services previously presented, a GI client can improve its processing capabilities and make the access to geographic information easier. However, it is hard to develop a generic client, suitable to different LSDIs. Furthermore, some capabilities, such as fault tolerance and engineering transparency, are actually the server’s responsibility, but shifting this responsibility towards the client is sometimes convenient, specially for mobile clients in an urban context.

When traveling, the client should be independent of specific service chains. In this case, when a client finds itself in a distant city, it can load and execute a chain of different services to obtain local information. To do this, services can be readily available within chains, but it would be better if each service was available individually and listed in a public catalog service. The selection and chaining of services can occur dynamically whenever the client needs to access them.

The Transaction Control Service (TCS) performs transformations in generic chaining code in order to generate a fully functional service chain, suited to a particular context. The generic code defines a basic workflow that is to be followed by the client during the execution process. That same basic workflow can serve as a template for other
service chains, counting on the Data Exchange Service for persistence and to store the processing state.

Figure 6 presents the result of using the TCS for the conversion of a generic code into client-specific service chains in the particular context of a routing application. The main objective of this service is to ensure that neither client nor service development are dependent on the local context, i.e., on peculiarities of the services that are available at each different location.

```java
// init
loadServices('auto');
DXSService dns = self.getDxSService();
// myLocation and pointOfInterest are paroled
// myLocation:
self.request(locater startPoint,inputType,finalType,dns);
// pointOfInterest:
self.request(locater destArea,inputType,finalType,dns);
sel.wait(); // waiting for the results
// query
self.request(route(dns.getItem(1),dns.getItem(2)),dns);
sel.wait(); // waiting for the final result
// do something with dns.getItem(3), where is the route
```

### Figure 6. Conversion from a Generic Code to Specific Functional Codes

With these services, LSDI clients can be implemented based on informational and computational aspects, avoiding the introduction of technological aspects into client code. Thus, thick-, rich-, and thin-clients can be implemented transparently, and their particular constraints are addressed by a generic workflow that changes easily when the client moves to another place or has its capacity improved.

### 3.4. Prototype Implementation and Analysis

We have implemented Web services for GIS as specified by the OGC, following the Abstract Model specification. The distribution patterns are (a) 1 Client to 1 Server (or known provider) to n Servers, where there is a provider which mediates resources from others, and (b) 1 Client to n Servers, where the client performs all of the tasks related to dealing access to Web services and geographic information processing.
The first pattern (a) made it easier to implement clients and servers with important engineering requisites, but introduces much dependency on known providers, in situations where no information about their resources and about external resources is available.

However, to enable client access to information about available resources, it is necessary to abandon the “known provider” role. This led us to implement the pattern (b), 1 Client to $n$ Servers. This pattern increases the costs of processing and traffic of intermediate data, because the client requests data to several services, and once it receives their responses it usually performs some operations and submits a large volume of data to other services in the service chain. However, this pattern enabled us to perceive and to define some engineering requisites on the client side.

Next, we separated design issues from technologic ones, and we grouped them into three groups, which constitute the services presented in the previous subsections.

By executing the features grouped into the Data Exchange Service, the intermediate data traffic in non-interactive processing was avoided on the client. In addition, in case of failure no previously processed data was lost. Nevertheless, we noticed no significant effect in interactive processing when there are multiple inputs during the service chain execution.

Through the Client Access Service, the delay and communication costs were reduced by avoiding the use of the network by the client when it needs to poll the service about its readiness before it really becomes ready.

Finally, we created a subset of operations and defined service chains using XML. The Transaction Control Service was then used to convert the XML data, given the client’s requisites and characteristics, to a service chain which uses the above infrastructure services, automatically.

4. Conclusion

4.1. Results

This work evaluated the engineering aspects of OWS specifications and the main OGC services in the context of urban applications. Through these applications, we identified some implementation constraints that are characteristic of GI systems, such as non-standardized fault tolerance mechanisms, clients that are strongly dependent of providers, and others.

To achieve this, we implemented a prototype based on the services abstract model for a real-world use case, and tested the usefulness of OGC standards to LSDI, specifically considering LSDI client development. As the objectives of standardization include to guide uniform application deployment and services interoperability, distribution transparencies, such as access, failure, migration, relocation, and persistence transparencies, become essential.

Since the availability of client software contributes to LSDI diffusion and implementation, we have also developed new infrastructure services, which facilitate the development of LSDI clients and interoperable GIS with urban characteristics. Thus, the new services constitute a synthesis of missing features in the standard technologies, and this prototype allows the exploration of other issues related to OGC standards and urban
GIS applications in future research efforts.

4.2. Main Contributions

Through the Data Exchange Service, clients – either thin, rich, or thick – can be transparently developed without previous distinction. There are additional advantages for thin and rich clients, which have greater limitations of energy, storage, communication, and processing capacity.

The Data Exchange Service replaces the persistence function of service providers and local storage with a third-party neutral service, through which the services chain exchanges parameters and results of its inner processing.

The establishment of the services chain is improved by two solutions. The first one is focused on the client, while the second one benefits from the communication capabilities among servers.

In the first alternative, the services chain orchestrated by the client defines the providers before and during the invocation of services. Then, the client becomes responsible for fault tolerance and it can choose alternative services according to established parameters. Additionally, when the client travels or in case of emergencies, the services and whole service set may be changed.

Nevertheless, the services chain may be orchestrated by a W3C- or OGC-like service, that calls others. Normally the client does not participate in service selection for the composition of a chain. It can only refuse a provider by selecting another. However, even this situation benefits from a shift of responsibility from the server to the client. The insertion of the Data Exchange Service into server workflow reduces the reissuing of requests in case of failures (specially if the client is thin), and increases the parallelism level of processing.

Therefore, we proposed a set of new services, which facilitate the development of GI clients for LSDI-based systems, and contribute to improve the diffusion and implementation of LSDIs.

4.3. Future Work

We identify three main directions for future work. First, we discuss the potential use of clients as information providers. Next, we propose some engineering enhancements to the development of services. Finally, we envision the possibility of developing new LSDI services, enhancing the prototype presented in this paper.

The GI client, viewed as a server of parameters for the service chain, can potentially become the server of various kinds of information. Information perceived directly by the user, such as traffic status [Tao et al. 2005], along with previously collected data, perceived quality of information, ontologies about its interests, geographic position, and others, can be passed along to other users through services. Possible applications to this are better traffic management using vehicles as real time sensors, more precise route planning by information exchange with other GI users, quality-of-service control, among others. For traffic status monitoring, information may be exchanged between these ‘sensors’ by active client-server connections (in the case of wireless sensor networks with limited energy capacity), or peer-to-peer connections (in the case of wireless sensors networks
without energy restrictions, i.e. ad-hoc networks constituted by in-transit vehicles), with new application possibilities, such as the dissemination of warnings or instructions. Peer-to-peer connections, when energy is not an issue, should improve the response time of applications [Wang et al. 2005], because previously collected data of other users is available. Geocoding, routing and locating services can then be adapted for mobile geographic objects, such as vehicles (public transportation, emergency, particular vehicles), or for personal locating (for instance, using cell phones). However, additional studies about these applications are required, in order to deal with concerns such as security and privacy. For effective privacy control, privacy parameters could be dynamically negotiated between clients and servers.

Clients, functioning as information providers, can also be the source of a measure of quality of service. Client feedback can be used to establish perceived quality, and to identify the quality parameters most valued by a user group. Ontologies on perceived quality parameters can then be developed from these observations.

As to engineering, fault tolerance parameters can be defined and negotiated according to availability contracts that should be automatically verified through a search into the service chain and evaluation of common failure points between main and substitute services. This is specially important in emergency response, critical services applications, and others.

Another engineering point regards workflow descriptions. In some cases, it is helpful that services be localized and dynamically injected into the original workflow, as in the case of the routing service, when routes cross regions not covered by the loaded data sources. Therefore, additional improvements should be made to workflow description mechanisms.

Davis Jr and Alves (2005) [Davis Júnior and Alves 2005] propose several services for LSDI, considering the specific demands for urban geographic applications. We think that further new services related to this level of spatial infrastructure can be proposed and investigated, using the prototype presented here as a basic framework.

Finally, the creation of a more complex LSDI prototype development environment is important to enable observations and experimentation on current and upcoming OGC standards. Such a framework should be based on the OGC’s implementation specifications rather than on an abstract specification model, as we did. Thus, it can be used to develop further formal studies about several concerns (privacy, performance, security, and so on) with real-world problems. It would also be important to build such a prototype development environment using free or open-source software, so that the community involved in spatial data infrastructures can benefit, while being able also to contribute to it.

Acknowledgement
The first author would like to thank CAPES for his scholarship to PUC Minas’s Master Program. The second author thanks ITC and the coordination of the Workshop on Cross-Learning Between Information Infrastructures and Spatial Data Infrastructures for the opportunity to attend the Workshop and gain important insights on its subject matter. Special thanks go to the three anonymous reviewers for their very relevant comments and suggestions.
References


