Method of Integration of Geospatial Data with Business Intelligent Systems based on Services Oriented Architecture

Dzenana Muracevic
Dept. of Computer Science & Informatics
University of Sarajevo

Fahrudin Orucevic
Dept. of Computer Science & Informatics
University of Sarajevo

Haris Kurtagic
CEO, SL-KING
Ljubljana

ABSTRACT

The need for more cost-effective, accurate and efficient methods of using geospatial information has been around a long time. However, development of GIS software solutions and information technologies enabled the distribution, access and manipulation of geospatial data over the Web.

In this paper is represented a new access to geospatial data and the possibility of integrating geospatial data with other business intelligent systems in standard and modern, Web-oriented, GIS applications. The paper reveals the motivation to show how geospatial data can be integrated into the Web environment platform and make it available to a large number of users. This paper represents the new method called GeoREST which is based on the principles of Services Oriented Architecture (SOA), or widespread Representational State Transfer (REST) architecture for geospatial data.

The proposed method called GeoREST defines new principles for manipulation with geospatial data, which will provide new ways of integration and distribution of geospatial data. This method provides an easy and open access to all types of geospatial resources through a uniform interface. GeoREST method opens new possibilities in the use of geospatial resources, both in standard desktop or Web applications, as well as business information systems. Geospatial service GeoREST API is developed and based on GeoREST method, and it defines the required application interface for simple and open access, manipulation and analysis of geospatial data. The practical implementation of services is made within the project georest.org.

Keywords

Integration, REST, GIS, GeoREST, SOA, Web service.

1. INTRODUCTION

Access to geospatial data and integration of geospatial data with other business intelligent systems [1, 2] has always represented a special problem. Geospatial data [3] are the information connected with a specific location, such as number of people in the city or residents at a particular address, parcels, roads, etc.. First programs for managing geospatial data were strictly applications, written in nonstandard formats. Data integration was not really possible, since it was based on manual or semi-automatic transfer of data from one system to the GIS system. The main goal of geographic information systems (GIS) was to establish a connection between the geospatial and attribute data, which will enable high-quality analysis and practical applying of GIS. GIS is a system that allows the input, storage, managing, searching, analysis, manipulation and presentation of geoinformations [3]. GIS differs from other information systems because it manipulates with a large amount of data, with data organized into data structures that require complex concepts in order to describe the geometry of objects and for specifying complex topological relationships among objects. The wider usage of geospatial data and analysis was not actually possible.

The disadvantages of those GIS systems are in that, that these systems are isolated for themselves. Geospatial data are stored in special files, and such data might be used only by using special programs and are usually isolated from other information systems within the company. Therefore, usually special departments are formed in accordance to manage with these data and usually 'traditional' information departmens are not dealing with such information and systems. This entails the isolation of these data, which represents a huge loss for the organization. Because of the problems in the exchange of geospatial data there has been appeared the need for standardization of records.

Open Geospatial Consortium (OGC) [4] was founded in 1994. in order to define standards that will enable the exchanging and wider usage of geospatial data. Based on these standards there was made todays support to geospatial data within relational database - relation database management system (RDBMS). Support to geospatial data has enabled the storage of geospatial data equaly with the attribute data. Using SQL language to perform GIS queries and analysis is greatly facilitated the integration of geospatial data within the organization.

Progress in the integration of geospatial data is evident from their ability to store and use within the RDBMS. The level of integration of geospatial data on the level of RDBMS is not enough if we talk about SOA, when we want to allow access to geospatial data or analysis in the form of services.

OGC has defined services such as Web Map Services (WMS) and Web Feature Services (WFS) [4] for the exchange of geospatial data. WMS is a service that allows the distribution of raster images and WFS allows the distribution of vector data. These services are based on Simple Object Access Protocol (SOAP). OGC web services have enabled a broader integration of geospatial data, but their actual usage has not spread as much as expected.

The complexity of implementation and complexity of using such services and the scalability problems are one of the reasons why Web services have not experienced widespread usage.

However with the development of Web 2.0 technology and architectures there has been observed all shortcomings of the second generation web services (WS-*), or WS-* technologies can not satisfied the requirements that are still present in the development of information systems at the global level, the internet level. REST architecture is the architecture of the Internet which provides global connectivity of the resources. The method, called GeoREST, presented in this paper defines how to facilitate access to geospatial data and analysis across the service or the internet.

2. SERVICES ORIENTED ARCHITECTURE

A geospatial SOA is supported by the efforts of GIS professionals who use GIS software, configure, and maintain GIS data, models, and applications [5]. The authored content is published on a GIS server upon which services are delivered to support mapping, query, analysis, and more. Technically, Web services are modular applications that correspond to recognizable business functions and offer a set of protocols by which they can be published, discovered, and used in a standards-based way. Organizationally, Web services are simply information technology (IT) assets that are often used as the basis for integration strategies that fuse content and capabilities in support of various business processes and initiatives. Web services provide the building blocks upon which broader IT strategies are based, such as the implementation of a service oriented architecture (SOA) [5, 6]. SOA and web services are two different things, but web services are the preferred standards-based way to realize SOA. This article provides an overview of geospatial SOA using the role of REST web services. Representational State Transfer (REST) is the next big trend in service-oriented architecture (SOA) development.

2.1 Resource Oriented Architecture and REST

This chapter has presented the concept of designing a REST architecture, under which is developed Geospatial method. REST was first introduced by Roy Fielding, in his dissertation about the architectural styles and design of network-oriented software architecture [7]. The principles and requirements of the programming system based on REST were introduced in Roy Fielding`s dissertation. Resource Oriented Architecture (ROA) is created by defining REST services based on the principles of the Web and HTTP protocols.

2.1.1 REST

REST is a hybrid style derived from several of the network-based architectural styles and combined with additional constraints that define a uniform connector interface [7]. REST system is defined as a set of conditions that the system must satisfy. REST is client-server architecture that separates the user interface from the data itself. The separation of functions allows independent development of client interface and increases server scalability. The second condition is Stateless, which means that every HTTP request happens in complete isolation. Stateless condition increases the scalability of the system, because the server does not need to keep state between requests. When client makes an HTTP request, it includes all information necessary for the server to fulfill that request. The server never relies on information from previous request.

REST defines the cache condition, in order to increase speed and efficiency of the system. Cache condition requires that the data in response on request are implicitly or explicitly marked as data with or without the possibility of caching. If the answer can be cache, this means that the same response can be used for following the same requirements. The main condition that REST differs from other network architectures is a uniform interface for interaction between client and server. Uniform interface greatly simplifies the overall system and allows separation of client and server.

2.1.2 Resource Oriented Architecture

ROA is the architecture that arises when the conditions and principles of the REST system are implemented in the Web environment [8]. REST is not necessarily linked to the Web, REST system can be implemented outside the Web environment. ROA defines Restful architecture on the Web using a URI and the HTTP protocol (Figure 1).

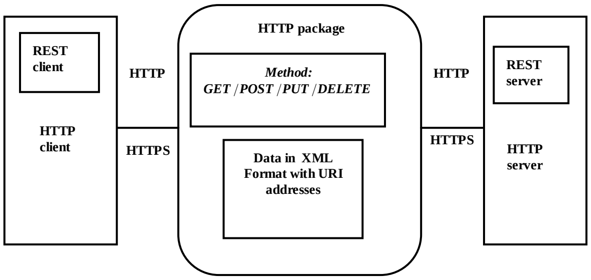

Figure 1: ROA Architecture

The picture shows the basic architecture of the REST system in which clients are using the HTTP protocol to access resources located on web servers.

3. OPEN STANDARDS AND THEIR ROLE IN GIS

This section is currently represented by the current technologies and standards for access to geospatial data in a heterogeneous and distributed environment. Since 1994, the OpenGIS Consortium (OGC) has been working on specifications for interoperability. OGC solution for exchanging geospatial information is based on the usage of web services, or more precisely the class of web services specialized for the processing of geographic data - OGC Web Services (OWS). In addition, OGC standards increasing roles in GIS has the open standards being developed with all the more rapid development of internet technologies. As Internet technologies become an indispensable part of Enterprise systems so the standards used on the Internet and the development of web applications play an increasing role in the development of geoinformation systems. Thus, the KML (Keyhole Markup Language) has become very fast the standard, supported by the OGC's. KML has due to the great support from Google's format has become indispensable in the Enterprise GIS systems. In addition to KML is often used JSON or GeoJSON, GeoRSS or RSS, Atom and AtomPub, HTML5, SVG and Canvas, etc.

4. NEW ACCESS TO GEOSPATIAL DATA AND INTEGRATION OF GEOSPATIAL DATA AND APPLICATION

A new access to geospatial data and integration of geospatial data with other business intelligent systems in standard and modern, Web-oriented, GIS applications is introduced in this chapter. A new method called Geospatial GeoREST is developed by using REST standards. Geospatial service GeoREST API is developed and based on GeoREST method, and defines the required application interface for simple and open access, managing and analysis of geospatial data.

4.1 Geospatial method GeoREST

A Geospatial method GeoREST defines new approaches in manipulation with geospatial data. GeoREST is based on REST principles with special reference to specificity of geospatial data and specific ways to use them. The basic idea of GeoREST is to represent each geospatial data and operation as an independent resource. In this way resource may be a single data such as a type of roads, telecommunication cable, spatial operations, spatial analysis (for example searching of cross sections between two objects), or resource is even analysis of the shortest path. Unique URI is assigned to each resource. Access to resources is possible through its URI. Each resource is using the same set of operations and that is the HTTP interface with its operations GET, PUT, POST and DELETE. Resources can be represented as JSON, XML, HTML or image formats such as .PNG, .JPG. The main elements of GeoREST methods are resources, HTTP interface and formats of representation resources.

4.1.1 GeoREST- resources

GeoREST defines the types of resources which are used in managing with geospatial data. Unique URL is defined for each type of resource through which is possible to access to resources. The types of resources are: spatial objects, spatial operators, geometric services and map or image services.

4.1.1.1 Spatial objects

The spatial object can be represented as a spatial component of geospatial data: point, line and polygon. Spatial object represents an actual physical object, such as parcel or road. Spatial object contains information about the objects such as attributes and geometry. Data are usually recorded in a file or a RDBMS. In the GeoREST method, each road and each parcel is one resource and that resource has its own unique URI address.

4.1.1.2 Spatial operators

Spatial operators are the types of resources used to perform spatial queries over spatial objects. Types of spatial queries match to OGC standards for spatial queries. These are operators such as: whether one geometry contains another, whether the geometries are intersecting, etc. Each operator is defined by unique URI via which access to the operator.
An example URI's can be: http://localhost/intersect.

4.1.1.3 Geometric services

Geometric services offers geometric operations on spatial objects such as distance, buffer, area, etc.. Each geometric operation is represented with its own unique URI. The parameters for the execution of operations are defined through the URI parameters. As with the spatial operator, URI parameter can be a URI pointing to a geospatial object or another geometric operation. In that way it is possible to connect multiple operations from multiple GeoREST service.

4.1.1.4 Map Services

Map services are used to define the maps which are used to display geospatial data through an image. Data Visualization is a very important part of the GIS system. Map Services is able to create and display maps that come simultaneously from multiple sources and displays them in standard format images such as TIFF, JPG or PNG format. With parameters Map service can defined which part of the map you want to get and in which kind of format. Besides the map, Map Service enables the identification of objects on the map.

4.1.2 GeoREST - uniform interface

GeoREST method uses HTTP interface for managing resources. HTTP operations GET, POST, PUT, DELETE, and OPTIONS are used to manipulate the resources. All resources may not have enabled all the commands, since it depends on the type of resources, or from the implementation of GeoREST methods. GeoREST is predicting usage of HTTP header for defining operation. Thus, the HTTP headers can be used to control the types of formats in which is possible to display resources, or, for example last modified header to control of caching HTTP requests and responses. GeoREST method predicts a complete and correct usage of HTTP protocol where in fact geospatial data and geospatial analysis are becoming a full member of the web environment.

4.1.3 GeoREST- formats of representation of the data

GeoREST method predicts the usage of standard Web-GIS formats for distribution of the data. Thus, each spatial object can be displayed as HTML, XML, JSON, PNG, etc.

4.2 Geospatial services GeoREST API

Geospatial service GeoREST API is developed on the basis of geospatial method. GeoREST API is a way of presentation of geospatial data on the web and geospatial analysis. Geospatial service supports the following features:

- Access to data and geometric attributes
- Access to image of data
- Edit of spatial and non-spatial attributes
- Enable spatial queries on data
- Enable different representation of data: PNG, HTML, KML, XML, and GeoJSON.

All abovementioned features are implemented within geospatial service GeoREST API. In GeoREST, data can be presented in several ways by using standard web presentations such as HTML, XML, or PNG, and the specific geospatial presentation of KML, GeoJSON, GML, GeoRSS, etc. Different formats of data presentations allow their usage in a wide variety of applications. One of the interesting advantages is presentation via dynamic HTMLs and generating of sitemap.xml file, which can be used to enable searching of data with a standard internet search browsers such as Google Search.

Also, by using the KML format, there is possibility to display the data on the increasingly popular Google Maps.

4.2.1 Example Application Using GeoREST API

Examples of URI requests and responses are:
HTTP GET
http://www.sl-king.si/mgrest/rest/data/parcel/29.html
This request will return HTML representation of parcel with id 29.
If we want to get XML representation of parcel we would execute.
HTTP GET
http://www.sl-king.si/mgrest/rest/data/parcel/29.xml

In similar way we can get image representing parcel by requesting .PNG as representational format to return (Figure 2.)
HTTP GET
http://www.sl-king.si/mgrest/rest/data/parcel/29.png

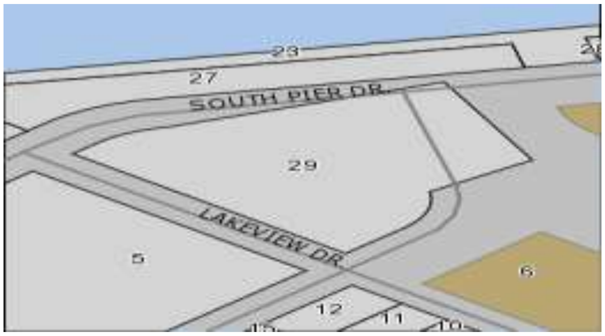

Figure 2. PNG representation of parcel

To filter data by some parameter we can use:

HTTP GET
http://www.sl-king.si/mgrest/rest/data/parcel/29.html?filter=NAME='SomeName'

To filter data by spatial parameter we can use:
http://www.sl-king.si/mgrest/rest/data/parcel/29.html?BBOX=1.1,1.1,2.2,2.2
Where BBOX represents rectangular window ((1.1,1.1) (2.2,2.2))

What is very powerful about Web Services is that we can combine more services or servers to execute one task.
For example we can have data of buildings served by one service and we can have parcels served by another server. We can create request which will return parcels which are under particular building.

URI would look like:

http://www.sl-king.si/mgrest/rest/data/parcel/.html?BBOX=http://.../data/builidng/32.xml

This request will return parcels under building with ID 32. While executing this request one service will call another service to get building with that id. We could also combine other filters and parameters while executing these requests. If we would like to update data we need to use PUT, POST or DELETE HTTP methods. These methods can be called from different clients and environments. Today, is using RESTfull Geospatial Web Services in AJAX application could be fined as the most interesting approach.
For that purpose we can use GeoJSON to query or update data.
To get data URL would look like:

HTTP GET
http://www.sl-king.si/mgrest/rest/data/parcel/29.json and
GeoJSON response is:

```
{ "type": "FeatureCollection" ,"members" :
[{"Type":"Feature","geometry": { "type":"Polygon",
"coordinates": [[[-87.708336985426598,43.747379333687],[-
87.707085159391596,43.747751654407402],[-
87.706685671813503,43.747110527794497],[-
87.707149749538601,43.746849468228802],[-
87.7071470443029,43.746837058531199],[-
87.707167092772295,43.746675799506399],[-
87.707172815197893,43.746663940457999],[-
87.707179110291307,43.746652234895599],[-
87.708637362745804,43.747259762323097],[-
87.708622790777994,43.747270160785298],[-
87.708607733831798,43.747280189565998],[-
87.708592209743003,43.747289836785399],[-
87.708576236900299,43.747299091015996],[-
87.708559834224403,43.747307941295503],[-
87.708543021145005,43.747316377140201],[-
87.708525817578106,43.747324388557502],[-
87.708508243902003,43.747331966057303],[-
87.708490320933706,43.747339100663702],[-
87.708472069904005,43.747345783925198],[-
87.7084535124322,43.747352007925102],[-
87.708434670500594,43.747357765290801],[-
87.708415566428499,43.7473630492022],[-
87.708396222845906,43.747367853400398],[-
87.708376662666396,43.747372172194297],[-
87.7083569090602,43.747376000468101],[-
87.708336985426598,43.747379333687]]] },"id": {
"Autogenerated_SDF_ID":29},"properties":
{"Autogenerated_SDF_ID":29,"URL":"","NAME":"","ID":"3220
09","RSTATE":"","RYEAR":0,"GEOEXTRA":0,"RBLDGVC":0
,"RLDESCR1":"","RLANDVC":0,"RTYPE":"","DETACHED_G
":"","RACRE":"","RLOT":"","RNAME":"","RLDESCR2":"","N
O_UNITS":0,"LAST_SALE":0,"YRBUILT":0,"RES_BED":0,"R
ES_FULL_B":0,"RPROPAD":"","RES_HALF_B":0,"KEYRLSE
Q":"","RLDESCR3":"","ATTACHED_G":"","RCITY":"","RBIL
AD":"","SQFT":0,"RSQFT":0,"RZIP":0,"RLDESCR4":"","RSEC
LN":"","GEOMAIN":0,"RES_TTL_RO":"","RPROCD":"","RW
ARD":0}}]}
```

To update data we would need to execute HTTP POST method and send GeoJSON in body of request.

HTTP PUT
http://www.sl-king.si/mgrest/rest/data/parcel/29.json
with updated GeoJSON included in body of request.
RESIDENTIAL ZONE","zone_value":"RS-1"}}]}

GeoREST API supports searching using standard Web search engines such as Google Search.

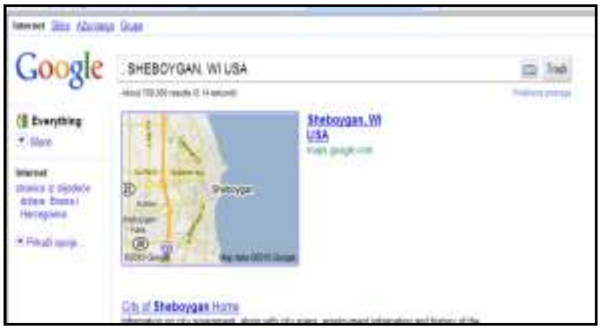

Figure 3: Example of searching geospatial data by using Google Search

5. CONCLUSION

Through years of experience in working with geospatial data, it was determined that existing architecture and interfaces have a lot of shortcomings, particularly in area of integration of the same and different environments. Research has gone towards the development of a method which is based on the principles of REST in HTTP environment, and the result was GeoREST method. GeoREST method with its principles enables widespread usage of geospatial information regardless of the type of environment in standard desktop, Web applications and in business information systems.

Unique and open geospatial services GeoREST API is based on GeoREST method, and defines the required application interface for easy and open access and managing geospatial data, thus GeoREST supports one of the main objectives, and those are Open Data and Open API.

6. FUTURE DIRECTIONS

GIS systems typically use large amounts of data. The data, except in the form of images are used in vector form. Using GeoREST APIs, vector data can be represented as XML, GeoJSON or some other text format. These formats have the advantage of the openness, readability and ease of use. However, when we want to transfer large amounts of data, records in text format are becoming too heavy. Although you can use data compression, which is automatically supported in Web services and clients, it may be insufficient for demanding GIS applications. Due to the increasing speed of data transfer and responses of those systems, it is necessary to explore possibilities for creating a new file format of geospatial data. Authentication and authorization of users is an important segment for the use and distribution of geospatial data. Besides the usual way of controlling access such as basic Client login or supports to LDAP protocol, today it is increasingly using of open web standards for access control.

It is necessary to explore the possibilities and modalities of integration of these standards in GeoREST so as to enable easy, controlled connection and using of GIS resources.

7. ACKNOWLEDGMENTS

The author owes great gratitude to the following people for their assistance in the practical verification of defined geospatial methods: Kurtagic Haris and Jason Birch (within the project georest.org).

8. REFERENCES

[1] Dejan, RISIMIC, “An integration strategy for large enterprise", Yugoslav Journal of Operations Research, Number 2, pp.209-222, 2007.

[2] Ismail Wadembere, Patrick J. Ogao, “Towards a Geometrical Spatial Integration Model for GIS”, International Journal of Computing and ICT Research, Vol. 2, No. 2, 2008.

[3] Michael J de Smith, Michael F Goodchild, and Paul A Longley, “Geospatial Analysis: A Comprehensive Guide to Principles, Techniques and Software Tools”, Troubador Publishing Ltd; 2nd edition, ISBN-13. 978-1906221522, 2007.

[4] Open Geospatial Consortium Inc., Published on the Web: http://www.opengeospatial.org, 2008.

[5] Esri, Geospatial Service-Oriented Architecture (SOA), CA 92373-8100 USA, Redlands, New York, 2007.

[6] Thomas Erl, “Service-Oriented Architecture (SOA): Concepts, Technology, and Design”, Prentice Hall/Pearson PTR, ISBN: 0-13-1858580, New York, 2005.

[7] Roy Thomas Fielding, “Architectural Styles and the Design of Network-based Software Architectures”, Doctoral Thesis, University of California, IRVINE, 2000.

[8] Leonard Richardson, Sam Ruby, “Book RESTful Web services”, O'Reilly Media, ISBN-10. 0-596-52926-0 ISBN-13. 978-0-596-52926-0, 2007.