Development of Spatial Database for Regional Development in Romania

Anda BELCIU (VELICANU), Vlad DIACONITA, Ion LUNGU, Aura-Mihaela VIRGOLICI (MOCANU)
Bucharest University of Economic Studies, Bucharest, Romania
anda.velicanu@ie.ase.ro, diaconita.vlad@ie.ase.ro, ion.lungu@ie.ase.ro,
mocanuaura@yahoo.com

Geographical Information Systems are used in solving many regional development related problems, around the world. Starting from some national programs to famous international ones, such as INSPIRE program, each such initiative uses geospatial data as well in the process of building regional development strategies. This paper presents the main technical components of a geographical information system, meaning the spatial database, the web mapping server and the APIs used to embed the maps into web applications. The development steps for a pre-alpha version of a web GIS application dedicated to the regional development in Romania are also shown. The software tools which were integrated in order to develop the online application were Oracle Spatial, where geospatial data was stored, GeoServer, an open source web mapping server used to generate the map out of the data from Oracle Spatial’s tables and ASP.NET as a web framework for building the website.

Keywords: Regional Development, GIS, Spatial Database, Web Mapping Server

1 Introduction
In the new knowledge based economy, the information systems for data handling, and for knowledge discovery became a “must have” for making decisions and elaborate development strategies in both public and private sectors.

Among the classical data types, the geospatial data became very interesting for the governmental decision factors, especially after the launch of the INSPIRE (Infrastructure for Spatial Information in the European Community) directive of the European Parliament in 2007. This kind of data shows “where” certain objects are placed on the surface of the Earth, from the pipes of a city to the best possible location to open a new business.

The Geographical Information Systems (GIS) have the technical capacity of storing, manipulating and visualizing the geospatial data, together with the descriptive/economic data. GIS is used for example in urban and regional development by creating different planning maps [1], in utilities management (electricity, gas, water) [2], when choosing the best locations for a certain investment ([3], [4]), in the environment studies [5], in assuring national health programs ([6], [7], [8]). The essential role of geospatial information in pressing issues such as: land cover, access to water, food production and distribution, biodiversity, is the main topic in many conferences in this field such as AGIT symposium “Geoinformation for Development (gi4dev)” held in 2006 in Austria.

This paper discusses the use of GIS tools in regional development and details the GIS architecture by analyzing its main software components: spatial databases, web mapping servers and the application programming interfaces (APIs) used to embed maps in web applications. In the last chapter we propose system architecture for building an online GIS for regional development in Romania. The application will offer the possibility of maintaining both spatial data and economical/statistical data regarding different variables of interest in different counties and to visualize statistical thematic maps which could be the basis of the decision making process.

2 Regional Development and GIS
Regional development’s aim is to offer assistance and to modernize regions which are less economically developed than others,
at international or domestic level. In order to comply with EU regional policy, in Romania there were identified eight regions which are administrated by the nongovernmental, nonprofit organizations called *Regional Development Agencies (ADR)* which are functioning in compliance to the Law no. 315/2004 on the regional development of Romania. The purposes of the Regional Development Agencies are detailed in paper [9] but all the tasks could be included into the main one and that is the creation of regional development plans.

A regional development plan starts with an analysis of the existing situation in a certain region by investigating the natural resources, infrastructure, physical capital, human capital, knowledge capital and the productive capital of the region. All this information (descriptive, statistical, geospatial) could be successfully integrated into a geographical information system. The GIS system could not only show the data on the maps, but it can also make complex statistical reports based both on economical and spatial / topological information.

The advantages and disadvantages of using the GIS technology in regional development are synthesized in [10] as shown in Table 1.

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Through GIS, one can make the connection between a regional element and its location;</td>
<td>• GIS software license costs;</td>
</tr>
<tr>
<td>• The maps offer objectivity when showing the geographical information;</td>
<td>• Costs for personnel training in GIS;</td>
</tr>
<tr>
<td>• Maps could be generated for every geographical problem (where data is available);</td>
<td>• The lack of certain technical equipment in the public institutions.</td>
</tr>
<tr>
<td>• The visualization of the data through a geographical context;</td>
<td></td>
</tr>
<tr>
<td>• The creation of the spatial data and inclusion of the non-spatial data as well;</td>
<td></td>
</tr>
<tr>
<td>• The possibility of the geo-statistical analysis of the information;</td>
<td></td>
</tr>
<tr>
<td>• Finding simple solutions for complex issues;</td>
<td></td>
</tr>
<tr>
<td>• One can build forecasts, predictions, trends of a phenomenon in time and space;</td>
<td></td>
</tr>
<tr>
<td>• Increase the level of detailed data;</td>
<td></td>
</tr>
<tr>
<td>• Making decision process is improved;</td>
<td></td>
</tr>
<tr>
<td>• Facilitates the communication between the stakeholders in regional development;</td>
<td></td>
</tr>
<tr>
<td>• Automate work.</td>
<td></td>
</tr>
</tbody>
</table>

At the beginning of 2010, the Romanian ADR Center and the Ministry of Public Finance have signed the contract for the project entitled “Development of modern planning tools for Region Center using GIS technology”. The project aims to create a connection between the six district councils, in terms of data sharing and spatial planning and also to facilitate the access of local government and of citizens to statistical information about Region Center (www.adrcentru.ro).

The regional policy of the EU underlines how important it is to promote growth and jobs by making regions more attractive for investments (http://europa.eu/pol/reg/index_en.htm). The most important factors when choosing a location for a new investment were found in a research by Weber and Chapman [3]: physical capital (infrastructure and facilities), human capital (labor market, demographic

DOI: 10.12948/issn14531305/18.3.2014.05
situation), knowledge capital (educational demographic situation, the skills of the inhabitants), and productive capital (competitors’ location).

GIS is successfully used in finding the best locations for new businesses, for example it was used for making strategic planning of wind farm sites in Great Britain [11], for urban aquaculture development in Bangladesh [4], for identification of potential wave energy locations in Malaysia [12] or for selecting camps for ecotourism [13].

3 GIS Software Tools
Open Geospatial Consortium (OGC) is a nonprofit organization who develops publicly available interface standards in order to enable geo-processing technologies to interoperate (www.opengeospatial.org). OGC provides specifications for both geospatial data formats and for GIS software tools.

According to these standards, there are the following types of encoding geospatial data:

- XML-based encodings: Geography Markup Language (GML) for transport and storage, Keyhole Markup Language (KML) which encodes both what to show and how to show in an Earth browser, Scalable Vector Graphics (SVG) for specifying how to visualize vector geospatial data and others;
- encodings for raster geospatial data: Joint Photographic Experts Group (JPEG) - a .jgw (world file) has to be added in order to geo-reference an image, Geo Tagged Image File Format (GeoTIFF);
- Well Known Binary (WKB) and Well Known Text (WKT) representations – the Well Known Binary Representation for Geometry (WKBGeometry) provides a portable representation of a geometric object as a continuous stream of bytes. It permits geometric object to be exchanged between an SQL/CLI client and an SQL-implementation in binary form.

3.1 Spatial Data Representation
One way to represent spatial data is exposed in [20], where spatial data is modeled with constraints.

Bi-dimensional spatial data are primarily represented by points \((x, y)\) or set of points \((x_i, y_i)\). These set of points can be infinite, which imposes the problem of storing this kind of data in a database. For solving this situation, we must define the finite representation of infinite relations.

In a relational database table, the rows must be seen as simple logical formulas. For example, for a finite set of four points \((x, y)\), we have the expression:

\[
\Phi_{\text{spatial_table}} = (x=1 \; \land \; y=1) \lor (x=7 \; \land \; y=3) \lor (x=4 \; \land \; y=4) \lor (x=7 \; \land \; y=1)
\]

In the database table the points look like in table 2.

**Table 2** (x, y) points stored in a table

<table>
<thead>
<tr>
<th>x</th>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
</tr>
</tbody>
</table>

Using predicates like <= and arithmetic operators (+, *), complex spatial data (with infinite points, for example, \(x \in \{1, 1.001, 1.002 \ldots\}, y \in \{3, 3.001, 3.002 \ldots\}\)) can be represented. For a polygon, the constraints can be:

\[
(y \leq x) \\
\land (x \leq 7) \\
\land (y \geq 1) \\
\land (x + 3y - 16 \leq 0)
\]

By using these four constraints, a polygon (considered as infinite spatial data) was represented in a finite way. Basically, all these constraints have a standard definition, which is \(ax + by + c \leq 0\), and must be logically intersected through the “and” operator (\(\land\)). This works only for spatial convex objects. A non-convex object can be vectorial represented like in the following example:

\[
N = \langle [4,3], [6,5], [8,2], [10,5] \rangle
\]

which will be represented using constraints and logical disjunction (\(\lor\)):

\[
((4 \leq x \leq 6) \land (3 \leq y \leq 5) \land (x-y-1=0))
\]
\[\lor ((6 \leq x \leq 8) \land (2 \leq y \leq 5) \land (3x+2y-28=0)) \lor ((8 \leq x \leq 10) \land (2 \leq y \leq 5) \land (3x-2y-20=0))\]

After considering these examples, we can rank the representational level of spatial data, thus:

- abstract level, which can be finite or infinite (when having a multitude of points that compose an object);
- symbolic level, which allows a finite representation of infinite relations;
- physic level, which allows storing relations in a database and querying the data.

Therefore, no matter the nature of the spatial data (finite or infinite), we showed that, at symbolic level, data can be processed and brought to a finite representation, in order to be properly stored and analyzed through a Database Management System.

### 3.2. Spatial databases

The spatial databases are a recent development in the databases software. They are extensions of the classical databases which contain special objects, procedures, mechanisms in order to store and manipulate geospatial data. We will further make an inter-comparison between the two most known spatial databases: Oracle Spatial which is a commercial tool and PostgreSQL/PostGIS on open source solution.

PostGIS extension for the open source PostgreSQL database stores the spatial object in the standard Well Known Binary (WKB) format. The Well Known Binary representation for Geometry is obtained by serializing a geometric object as a sequence of numeric types drawn from the set \{Unsigned Integer, Double\} and then serializing each numeric type as a sequence of bytes using one of two well defined, standard, binary representations for numeric types (NDR, XDR) [14]. When filling a column with geometry data, one will use the following syntax:

\[
\text{ADDGEOMETRYCOLUMN}(\text{table_name}, \text{column_name}, \text{srid}, \text{data_type}, \text{dimension_number});
\]

The spatial reference system id, SRID, associated with a geospatial column must exist in the PostGIS’s SPATIALREFSYS technical table (equivalent of Oracle’s MDSYS.CS_SRS table). One can find out the Well-Known Text (WKT) representation of a certain SRS by using the following select statement:

- in Oracle Spatial database:
  \[
  \text{select WKTEXT from MDSYS.CS_SRS where SRID = 4326};
  \]
- in PostgreSQL/PostGIS database:
  \[
  \text{select SRTEXT from SPATIALREFSYS where SRID = 4326};
  \]

and the result is:

```sql
"GEOGCS["WGS84",DATUM["WGS_1984",SPHEROID["WGS84",6378137.295,298.257223563,AUTHORITY["EPSG","7030"]],TOWGS84[0,0,0,0,0,0,0],AUTHORITY["EPSG","6326"]],PRIMEM["Greenwich",0,AUTHORITY["EPSG","8901"]],UNIT["degree",0.01745329251994328,AUTHORITY["EPSG","9122"]],AUTHORITY["EPSG","4326"]]
```

The data types used in PostGIS database are: POINT, LINESTRING, POLYGON, MULTIPOLYGON, MULTILINESTRING, and MULTIPOLYGON.

Both databases Oracle Spatial and PostGIS offer the possibility for mass upload of the data from shapefiles (world’s most famous commercial GIS company, ESRI’s proprietary file format storing geospatial data), by providing special tools such as: `shp2sdo` (Oracle – Figure 1) or `shp2pgsql` (PostGIS).

**Fig. 1.** Loading a shapefile into the Oracle Spatial database
Among the usual database constraints (primary keys, foreign keys, semantic constraints, and user defined constraints), the spatial databases are also characterized by the topological consistency relations, which are listed in paper [15] as follows: everything must be bounded by two nodes (start node and end node); for every arc, there exist two polygons (left polygon and right polygon); every polygon has a closed boundary consisting of an alternating sequence of nodes and arcs; around every node, there exists an alternating closed sequence of arcs and polygons; arcs do not intersect except at nodes.

RTree indexes are used in both databases for the indexing of geospatial data. This index approximates each geometry with a *Minimum Bounding Rectangle (MBR)*.

Create an index on a "geometry" column in Oracle Spatial:

```sql
CREATE INDEX roads_ro_idx ON roads_ro(geom)
INDEXTYPE IS MDSYS.SPATIAL_INDEX;
```

and in PostGIS:

```sql
CREATE UNIQUE INDEX roads_ro_idx ON roads_ro(geom)
USING GIST(geom);
```

GiST (*Generalized Search Trees*) indexes break up data into "things to one side", "things which overlap", "things which are inside" and can be used on a wide range of data-types, including GIS data. PostGIS uses an R-Tree index implemented on top of GiST to index GIS data (http://postgis.refractions.net/docs/ch04.html #id2638756).

Spatial analysis functions are very important feature of these databases and the one that makes a difference between the databases. For example, the most advanced spatial database from this point of view is Oracle Spatial, which offers a special package SDO_SAM with many subprograms which can be used in spatial analysis and data mining. PostGIS does not have a special package with functions to be used in the spatial analysis but it can be done by writing some smart SQL queries using the spatial functions `ST_DISTANCE`, `ST_INTERSECTS`, `ST_OVERLAPS` [16].

### 3.3 Web Mapping Servers

OGC International Standard for web mapping servers, WMS, [17] defines the output generated by a web mapping service as maps of spatially referenced data dynamically from geographic information. In reference work [18], it is specified that the OGC’s services are based on a set of standards that are popularly called RESTful web services. This means that all the queries are going to be simple HTTP GET requests; the service will be called using an URL prefix to which additional parameters are appended in order to construct a valid operation request.

The structure of WMS request using HTTP GET is described in OGC standard as:

http://host[:port]/path[?{name=[value]&}].

The reference standard defines three operations for a WMS: mandatory `GetCapabilities` and `GetMap`, and optional `GetFeatureInfo`. The definition of `GetCapabilities` operation according with [17] is operation to obtain service metadata, which is a machine readable (and human-readable) description of the server’s information content and acceptable request parameter values. A `GetCapability` operation is a XML document showing what the server offers.

In simple words, the WMS is a bridge between a geospatial data source and a client application showing the map out of the stored data, as it was shown in [19]. It is the piece of software that “reads” the data from the spatial data source (database or file), and it “draws” the map of the data in different formats (png, gif, jpeg and others). Table 3 shows a comparative study between two web mapping servers: Oracle MapViewer and GeoServer from different point of views.
<table>
<thead>
<tr>
<th>WMS</th>
<th>Oracle MapViewer</th>
<th>GeoServer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input formats</td>
<td>Oracle Spatial</td>
<td>Vector sources: ESRI Shapefiles, GML, PostgreSQL, MySQL, SQL Server 2008, Oracle Spatial, ArcSDE and others</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Raster sources: GeoTIFF, ImageMosaic, WorldImage, GTOPO30 and others</td>
</tr>
<tr>
<td>HTTP Methods</td>
<td>GET, POST</td>
<td>GET, POST</td>
</tr>
<tr>
<td>Operations</td>
<td>GetCapabilities,</td>
<td>GetCapabilities, GetMap, GetFeatureInfo, DescribeLayer, GetLegendGraphic</td>
</tr>
<tr>
<td></td>
<td>GetMap,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GetFeatureInfo</td>
<td></td>
</tr>
<tr>
<td>Output formats</td>
<td>Png, gif, jpeg,</td>
<td>Png, atom xml, openlayers, pdf, rss, xml, google-earth.kml, google-earth.kmz, geotiff, gif, jpeg, svg, tiff and others</td>
</tr>
<tr>
<td></td>
<td>svg+xml</td>
<td></td>
</tr>
<tr>
<td>SLD</td>
<td>SLD support</td>
<td>SLD support</td>
</tr>
</tbody>
</table>

### Table 3. Comparative study between: Oracle MapViewer and GeoServer

#### 4 Development of a spatial database for a Romanian Regional Development Agency

The standard architecture of GIS is the three layer architecture containing:

- the spatial database, core of every GIS, where the geospatial data is stored and maintained;
- the web mapping server which generates the map out of the spatial database;
- APIs, on client side, to embed the GIS maps into web applications.

In order to embed a map generated by a WMS out of a spatial database, into a web application, one must use APIs, which are programs library for map displaying in web browser. APIs are available for specific geospatial servers such as ArcGIS JavaScript API, MSN Virtual Earth APIs, or Google Maps API, and for general geospatial servers, such as OpenLayers JavaScript library which implements industry-standard methods for geographic data access, such as the OpenGIS Consortium’s Web Mapping Service (WMS) and Web Feature Service (WFS) protocols.

Our goal was to develop a first version of a web GIS application to be used in the optimization of the regional development strategy making process. The main functional requirements for the system were:

- maintain the data regarding the statistics macroeconomic indicators (population, labor market, education) of each county;
- maintain geospatial information regarding the counties of Romania, the main points of interest, cities and main roads;
- visualize thematic statistics maps;
- visualize reports with economical and geospatial data needed in the process taking decision.

The technical architecture of the new system is integrating a spatial database, Oracle Spatial, where the geospatial data will be stored, a web mapping server, GeoServer, which will generate the map out of the data stored in the Oracle Spatial and a web framework which is used to develop the web application which will embed also the maps generated by GeoServer.

First, we have created the conceptual model of the new system by drawing the Entity-Relationship diagram with Oracle Designer software product (Figure 2).
This diagram has to be read as follows: the new system will store the geospatial data regarding the counties of Romania (COUNTIES_RO); each county has roads (ROADS_RO) and different points of interest (POI_RO) for which descriptive and geospatial data will be stored as well; for each county there are available a set of macroeconomic indicators describing the economic development of the county, regarding the population (STAT_POP_RO), labor market (STAT_LABOUR_RO), or education level (STAT_EDU_RO).

The normalization technique was then used in order to refine the conceptual model and obtaining minimal redundancy database tables.

After creating the structure of the database we have loaded the tables with geospatial data regarding the counties, roads and points of interest and also with statistic indicators values; the data sources which were used to load the initial data are listed in Table 4.

<table>
<thead>
<tr>
<th>No.</th>
<th>Document name</th>
<th>Source</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>judete_ro.shp, judete_ro.shx</td>
<td>geo-spatial.org</td>
<td>annually</td>
</tr>
<tr>
<td></td>
<td>judete_ro.dbf</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>romania_highway.shp</td>
<td><a href="http://downloads.cloudmade.com/">http://downloads.cloudmade.com/</a></td>
<td>annually</td>
</tr>
<tr>
<td></td>
<td>romania_highway.shx</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>romania_highway.dbf</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>romania.poi.shp, romania.poi.shx</td>
<td><a href="http://downloads.cloudmade.com/">http://downloads.cloudmade.com/</a></td>
<td>annually</td>
</tr>
<tr>
<td></td>
<td>romania.poi.dbf</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>Statistical Yearbook of Romania</td>
<td>National Institute of Statistics</td>
<td>annually</td>
</tr>
</tbody>
</table>

The next step was to configure GeoServer in order to connect to the Oracle Spatial database which was previously built. Also, there were defined the styles in which the maps will be rendered by writing style layer descriptor (SLD) files.

After connecting GeoServer with Oracle Spatial, the web application was developed in ASP.NET who was also connected with both Oracle Spatial using ADO.NET technology and with GeoServer, using the web services. The users have the possibility to upload the macroeconomic data in the database by using the web application (Figure 3).
Macroeconomic Data

<table>
<thead>
<tr>
<th>County</th>
<th>Population</th>
<th>Employees</th>
<th>Unemployed</th>
<th>Students</th>
<th>Faculties</th>
<th>Life expectancy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Select ALBA</td>
<td>379189</td>
<td>87000</td>
<td>13011</td>
<td>6362</td>
<td>12</td>
<td>71,89</td>
</tr>
<tr>
<td>Select ARAD</td>
<td>459286</td>
<td>120000</td>
<td>7036</td>
<td>21447</td>
<td>26</td>
<td>70,93</td>
</tr>
<tr>
<td>Select ARGES</td>
<td>646320</td>
<td>137000</td>
<td>16055</td>
<td>16714</td>
<td>18</td>
<td>72,25</td>
</tr>
<tr>
<td>Select BACAU</td>
<td>723518</td>
<td>121000</td>
<td>13759</td>
<td>8235</td>
<td>9</td>
<td>71,01</td>
</tr>
<tr>
<td>Select BIHOR</td>
<td>595685</td>
<td>152000</td>
<td>7647</td>
<td>23230</td>
<td>30</td>
<td>70,38</td>
</tr>
</tbody>
</table>

Fig. 3. “Macroeconomic Data” maintenance web page

In order to visualize a certain statistic thematic map (Figure 4), it was called the GeoServer “GetMap” operation, from the web application page, as follows:


Fig. 4. The thematic map “Employees per county”

Other web pages show information about the counties area (calculated from the database using `sdo.geom.sdo_area` Oracle Spatial’s function) or the points of interest from a
certain county (using a join between COUNTY_RO and POL_RO table and sdo_geom.sdo_relate function).

5 Conclusions
The following software technologies: Oracle Spatial (to store the geospatial data), ASP.NET (to develop web applications) and the open source web mapping server, GeoServer (to generate the maps), could successfully be integrated in order to build a web GIS application, as support for regional development strategies making process.

The first version of the developed web application described in chapter 4 was only intended to explore the capacity of using the technologies together in order to visualize in a web page the map generated from the data stored into the “geometry” columns of a spatial database. We further intend to develop the web application, by embedding an interactive map using OpenLayers library, being then able to release an alpha version of the proposed web site.

References


**Anda BELCIU** is co-author of 4 books, has 11 articles published in prestigious journals included in international recognized databases (SCOPUS, Elsevier, EBSCO, ProQuest, or DOAJ) and also 14 papers in the volumes of national and international scientific manifestations, of which 1 is indexed Thomson ISI Web of Science. Her scientific fields of interest and expertise include database systems, e-business, e-learning, spatial databases. She has experience in 6 research projects, participating as a team member.

**Vlad DIACONITA** is a member of the IEEE and INFOREC organizations and Member of the technical team of the Database Systems Journal. As member of the research team has taken part in 4 UEFISCDI funded grants and published more than 30 papers in peer reviewed journals and conference proceedings, many indexed in ISI or SCOPUS. Since July 2014 he is pursuing post-doctoral research financed by EU through the Excelis program with the project entitled: “Distributed analysis of large volumes of data for decision support”. He is the author of a book and the co-author of three books.
Ion LUNGU, senior researcher, 62 years old is since 1993 full professor and since 1999 PhD Conductor at AES Bucharest, field of Informatics and Cybernetics. Since 1999 he completed over 20 research topics in doctoral work. He is the author of 41 books or book chapters in prestigious publishing houses (InTech - Open Access Publishing Company), has published over 80 articles in professional journals (of which 22 in ISI journals or included in international databases) and over 80 studies presented at international conferences, some cited in journals included in the ISI database or prestige. He participated in the 28 contracts of which 20 were won by competition.

Aura-Mihaela VIRGOLICI has since 2013 a PhD in the domain of Economic Informatics. Her areas of interest are: Geographical Information Systems, Databases, Information Systems integration, Programming languages.