The geovisualisation window of the temporal and spatial variability for Volunteered Geographic Information activities

Beata Medyńska-Gulij, Miłosz Myszczuk

Department of Cartography and Geomatics, Adam Mickiewicz University Poznan,
Collegium Geographicum, 27 Dziegielowa St., 61-680 Poznan, Poland
e-mail: bmg@amu.edu.pl

Abstract: This study presents an attempt to design geographical visualisation tools that allow to tackle the immensity of spatial data provided by Volunteered Geographic Information (VGI), both in terms of temporal and spatial aspects. In accordance with the assumptions made at the conceptual stage, the final action was the implementation of the window entitled 'Geovisualisation of the Panoramio.com Activities in District of Poznań 2011' into the web browser. The concept has been based on a division of the geovisualisation window into three panels, of which the most important - in order to capture spatial variability - have statistical maps at the general level (dot map and choropleth map), while at the detailed level - a dot map on a topographic reference map or tourist map. For two ranges, temporal variability is presented by graphs, while a review of attributes of individual activities of the social website in question is set forward in the table panel. The element that visually interlinks all of the panels is the emphasised individual activity.

1. Introduction

Internet services that allow users to create their own content, known as Web 2.0, have pointed out a different approach to methods of collecting spatial data. Of considerable importance are social websites, the members of which use free software and simple tools to effect the implementation of geographical information into the global network. These activities are known as Volunteered Geographic Information (VGI), which according to Goodchild (2007) is the harnessing of tools to create, assemble, and disseminate geographic data provided voluntarily by individuals. The VGI phenomenon has initiated scientific research concerning connections with GIS (Elwood, 2008) and the impact on the Geographic Information Science (Coleman, 2010). Considerable interest has been aroused by the credibility of volunteered geographic information (Flanagin, and Metzger 2008), which is also of importance for identifying the changing balance of the public and private sector (Kuhn, 2007; McDougall, 2009). Spatial data infrastructures within national and European structures have come within the scope of legal standards

and norms (Bielecka, 2003). Internet communities should be treated as an important entity in the obtainment and processing of geographical information. The three basic elements of VGI include the community, spatial data, and the technologies connected by the internet. Examples of VGI service are the „WIKIMAPIA – let's describe the world!", www.flickr.com and OpenStreetMap.

This study presents an attempt to design geographical visualisation tools that allow to tackle the immensity of spatial data provided by Volunteered Geographic Information (VGI), both in terms of temporal and spatial aspects. Geographical visualisation (Geovisualisation, GeoVis or Gvis) can be defined as a private activity in which previously unknown spatial information is revealed in a higly interactive computer graphics environment (Slocum et al., 2009). In research concerning geovisualisation, emphasis is placed primarily on the representation of geographical phenomena, on visualisation, and on the interface (Dykes et al., 2005; Baranowski et al., 2009).

The forms of geovisualisation that may be used to handle multi-criterionality and multi-dimensionality of data include the following: statistical maps, diagrams, graphs with two and more variables, multi-series minimaps, and animations of data variability (Slocum et al., 2009). The basic role of the map in the process of geovisualisation is undisputed, while the statistical map and reference map continue to be the primary – because they are most easily understood – carriers of spatial information. A potentially effective method of disseminating and understanding geographical phenomena, which at the same time is a necessary component in the structure of integrated discovery of spatial knowledge, is the geovisualisation window, which allows the user to interact with individual interactive graphical panels (Medyńska-Gulij, 2011). An example of an interactive window for the web user is the atlas of incidence of cancer in Pennsylvania (MacEachren et al., 2008).

2. Methodology

2.1. Procedure for elaborating geovisualisation

The main objective of the present research was to elaborate a method for the interactive geovisualisation of temporal and spatial variability of point phenomena on the example of the Panoramio.com internet community. As it turned out, of primary importance for solving the research problem was the elaboration of procedure for creating an interactive visualisation of temporal and spatial changes that would contain cartographic and other methods of visualisation, the selection of technologies, and design of the geovisualisation window. The geovisualisation window was published in a web browser, and this fact determined the method of processing data and adapting editorial principles to the perceptive capabilities of a web user (Dickmann, 2004).

The utilitarian objective was to create of a tool that would allow each and every internet user to individually and interactively analyse the VGI phenomenon using pre-designed panels: the map panel, the table panel, and the graph panel. The process of

elaborating the geovisualisation of the temporal and spatial variability of the Panora-mio.com community comprised three primary stages (Fig. 1): conceptual; acquisition the required data and processing spatial data; and visualisation and implementation.
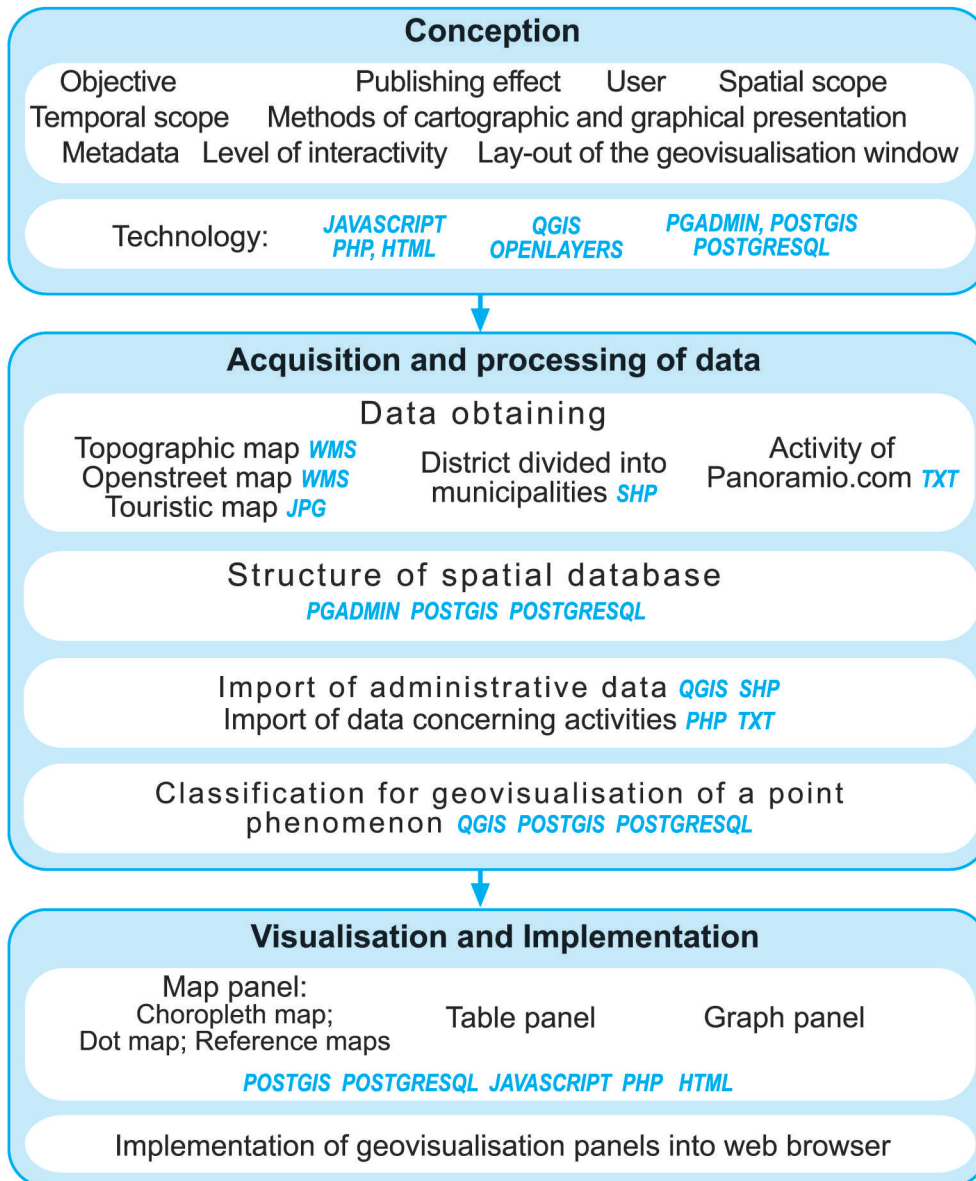
Conception

Objective   Publishing effect   User   Spatial scope
Temporal scope   Methods of cartographic and graphical presentation
Metadata   Level of interactivity   Lay-out of the geovisualisation window

Technology: *JAVASCRIPT*   *QGIS*   *PGADMIN, POSTGIS*
               PHP, HTML   *OPENLAYERS*   *POSTGRESQL*

Acquisition and processing of data

Data obtaining

Topographic map *WMS*
Openstreet map *WMS*   District divided into   Activity of
Touristic map *JPG*   municipalities *SHP*   Panoramio.com *TXT*

Structure of spatial database
PGADMIN POSTGIS POSTGRESQL

Import of administrative data *QGIS SHP*
Import of data concerning activities *PHP TXT*

Classification for geovisualisation of a point
phenomenon *QGIS POSTGIS POSTGRESQL*

Visualisation and Implementation

Map panel:
Choropleth map;   Table panel   Graph panel
Dot map; Reference maps

POSTGIS POSTGRESQL JAVASCRIPT PHP HTML

Implementation of geovisualisation panels into web browser

Fig. 1. Procedure for elaborating the geovisualisation window of the temporal and spatial variability for
VGI activity

2.2. Concept

Of key importance for elaborating the geovisualisation window were the assumptions made for the first conceptual stage, i.e.

– objective: capturing/disclosing the variability features of temporal and spatial activities of the Panoramio.com community broken down into hours and months in the year 2011 for the entire area (district of Poznań) and a randomly selected view (zoom);
– publishing effect: the interactive window in the web browser entitled 'Geovisualisation of the Panoramio.com Activities in District of Poznań 2011';
– user: a public internet user with basic internet skills;
– method of cartographic presentation: dot map, choropleth map, reference maps:
dot map – lay-out according to geographical coordinates, dot weight – individual activity = 1 dot); choropleth map broken down into municipalities, with a classification based on the Jenks's method;
– method of graphical presentation: graph presenting changes in the number of activities for individual months; graph presenting changes in the number of activities for individual hours; possibility of comparing changes in the number of activities on the graph for the district of Poznań, for the view, and for Poznań;
  – table of activities: a table of attributes for all instances of activities and for activities in the view;
  – level of interactivity: spatial navigation (moving the view rectangle on the map of the district); selection of a reference map in the view (topographic/open-street/ tourist map); selection of an individual activity (dot) on the map of the district and in the view, and of the record of an individual activity (clicking on the dot results in the dot being backlit in each available location, i.e. on the choropleth map, in the view, in the table and on activity charts);
  – technology: an open source software, which in this instance meant using desktop QuantumGIS (QGIS) software (for operations in the .shp format), PostgreSQL and PostGIS database software (for operations in .EXE, SQL and SpatialSQL), and internet technologies: OpenLayers, JavaScript, PHP, HTML i Google Charts (.js, .php); architecture: client (browser) – server with network services (database: PostgreSQL PostGIS, WMS, Google Charts);
  metadata: link to a PDF file containing descriptions of data sources, their acquisition, processing and classification according to the procedure for creating the geovisualisation window; data concerning the authors;
  – lay-out of the geovisualisation window: the title box and metadata box and the three main interactive panels: the map panel, the table panel, and the graph panel; closed, balanced horizontal composition; visual hierarchy: graphical enhancement of a selected individual activity; interactions in the window: click, scroll, move.
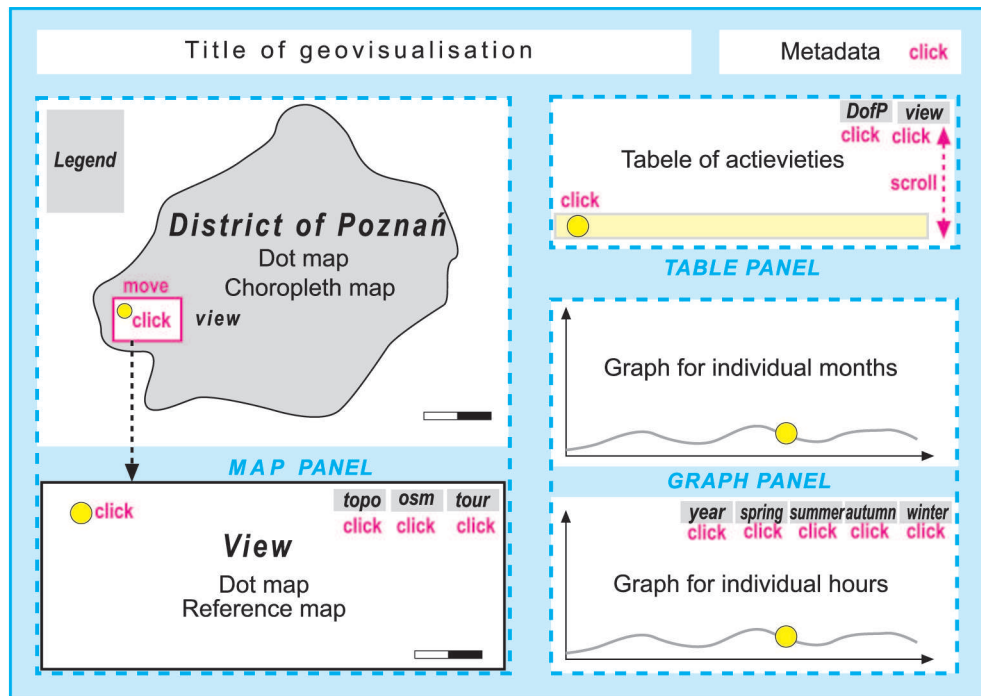
Fig. 2. Lay-out of the geovisualisation window of the temporal and spatial variability for Volunteered Geographic Information activity in the district of Poznań

2.3. Acquisition and processing of data

An administrative map of Poland in .shp format, broken down into districts and municipalities, was used for elaborating the choropleth map and the dot map. In turn, individual point objects constituted activities saved in the Java Script Object Notation (JSON) data exchange format, the strict form of which was as follows: {"count": 17046,

"photos": [{
"**upload_date**": "19 February 2007",
    "**taken_on**": "129403987321"
"**owner_name**": "Radek Szampa",
"**photo_id**": 890788,
"**longitude**": 18.222712999999999,
"**height**": 880,
"**width**": 1480,
"**photo_title**": "Kapliczka",
"**latitude**": 49.491950000000003,
"**owner_url**": "http://www.panoramio.com/user/187280",

"**owner_id**": 187280,
"**photo_file_url**": "http://mw2.google.com/mw-panoramio/photos/thumbnail/890788.jpg",
"**taken_on**": "12958762104"
    "**camera_type**": "Panasonic DMC-FZ35",
    "**eksposure**": "0.025s (1/40)",
    "**focal_length**": "20.10mm",
    "**f/Stop**": "f/3.600",
    "**ISO Speed**": "ISO200",
    "**exposure_bias**": "-0.33 EV",
    "**flash**": "no-flash",
    "**photo_url**": "http://www.panoramio.com/photo/890788"}, . . .] . . .]

The above entry is a JSON entry, which may be considered as a "light" data exchange format based on the object literal successfully replacing XML. This Java-Script subset may be used to create objects, while in the above instance the object is a photography with described features having specific attribute values. To increase interpretation opportunities for users of the VGI phenomenon geovisualisation window, a spatial reference in the following form was adopted: topographic map (WMS), OpenStreetMap (WMS) and a half-tone image of a tourist map (jpg).

Municipalities within the borders of the district of Poznań were selected by means of the functions of the QGIS programme. The administrative units thus selected were saved as an .shp file entitled "powiat-poznan.shp", which was subsequently imported into the PostGIS database using the SPIT tool available in the QGIS suite (Fig. 3).
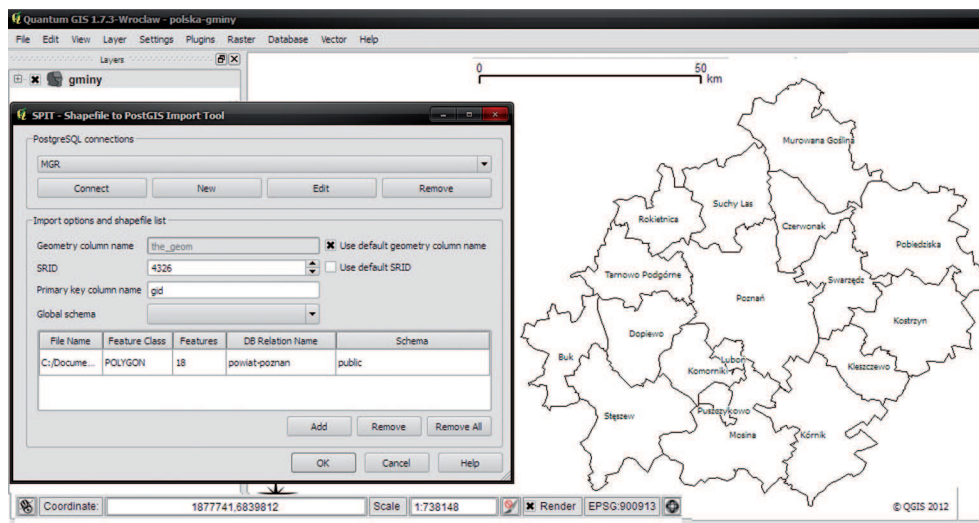

Fig. 3. Import of spatial data from the .shp file into the postGIS database

An initial classification of data concerning activities made it necessary to select attributes for further analyses, which were then saved in a .txt text file as follows:

{"count": 17046, "photos": [{
"**upload_date**": "19 February 2007", "**taken_on**": "129403987321,,"**owner_name**":
"Radek Szampa", "**photo_id**": 890788, "**longitude**": 18.222712999999999, "**height**":
880,
"**width**": 1480, "**photo_title**": "Kapliczka", "**latitude**": 49.491950000000003,
"**owner_url**": "http://www.panoramio.com/user/187280", "**owner_id**": 187280,
"**photo_file_url**": "http://mw2.google.com/mw-panoramio/photos/thumbnail/890788.jpg",
"**taken_on**": "12958762104","**camera_type"**: "Panasonic DMC-FZ35" , "**eksposure**":
"0.025s (1/40)" , "**focal_length**": "20.10mm" , "**f/Stop**": "f/3.600" , "**ISO Speed**":
"ISO200", "**exposure_bias**": "-0.33 EV', "**flash**": "no-flash", "**photo_url**":
"http://www.panoramio.com/photo/890788"}, . . .] . . .]

The next step concerned elaborating a spatial database with specifications for tables for ACTIVITIES point objects and for tables for MUNICIPALITIES surface objects. Description of individual columns for the MUNICIPALITIES tables contain the following: gid – a unique identifier of the municipality in the database, primary key; nazwa – municipality name; idgminy – municipality identifier; the_geom – geometry column to which three integrity connectors have been assigned (Table 1). In turn, the spatial reference parameters for MUNICIPALITIES have been presented in Table 2.

Table 1. Structure of the MUNICIPALITIES table

Column	Type	Not Null	Default
gid	integer	NOT NULL	nextval('gmina _gid_seq'::regclass)
nazwa	character varying(47)		
idgminy	integer		
the_geom	geometry		

Table 2. Geometry and spatial reference parameters for MUNICIPALITES

Name	Definition
Enforce_dims_the_geom	CHECK (ST_ndims(the_geom)=2)
Enforce_geotype_the_geom	CHECK (geometrytype(the_geom) = 'POLYGON'::text OR the_geom IS NULL)
Enforce_srid_the_geom	CHECK (ST_srid(the_geom) = 4326)
gmina_pkey	PRIMARY KEY (gid)

The table for point objects, i.e. activities, has the following attributes (Table 3): id – a unique identifier of an individual activity; medium – single-sign variable corresponding to: f – photography; timestamp – format of time passage, so-called epochs, i.e. the number of seconds that have passed since 1 January 1970 until the time the photograph was taken (the storage of time-related data in this format makes it easier to manipulate the date during addressing queries to the database); point – a geometry column to which three integrity connectors have been assigned; title – a photograph title; date – date stored in the DD-MM-RRRR format.

Table 3. Structure of the ACTIVITIES table

Column	Type	Not Null	Default
id	integer	NOT NULL	nextval('obiekt_gid_seq'::regclass)
medium	character varying(1)		
timestamp	integer		
point	geometry		
title	title		
date	date		

Just as in the case of MUNICIPALITES, Table 4 presents spatial reference parameters for individual activities for ACTIVITY determined by the value of the latitude and longitude coordinate pair.

Table 4. Geometry and spatial reference parameters for ACTIVITIES

Name	Definition
Enforce_dims_the_geom	CHECK (ST_ndims(point)=2)
Enforce_geotype_the_geom	CHECK (geometrytype(point) = 'POINT'::text OR point IS NULL)
Enforce_srid_the_geom	CHECK (ST_srid(point) = 4326)

Table 5 presents the result of administrative data import into the database, with exemplary records for each municipality and attribute columns, stored in the "Well-Known Binary" (WKB) format.

Table 5. Result of the import of administrative data into the database

gid	name	id
1	Kórnik	2890
2	Stęszew	2916
3	Swarzędz	2932
4	Kostrzyn	2889
5	Buk	2855
6	Mosina	2902
7	Murowana Goślina	2904
8	Pobiedziska	2905
9	Czerwonak	2869
10	Dopiewo	2870
11	Kleszczewo	2871
13	Komorniki	2872
14	Tarnowo Podgórne	2946
15	Rokietnica	2915
16	Suchy Las	2931
17	Poznań	2755
18	Luboń	3053
19	Puszczykowo	3054

In turn, the PHP script was used to import data concerning individual activities:

```
For ($i=1; $i=>17046; $i++){
$d=pg_connect('host=localhost user=postgres dbname=mgr);
pg_exec($d, "insert into obiekt (title, timestamp, medium, point, date)
values ('$array[$i][0]', $array[$i] [1],'$array[$i] [2]',
GeometryFromText('POINT($array[$i] [3] $array[$i] [4])',' $array[$i] [5])", 4326));
"); }
```

The "$array" variable is a one-dimensional table in which numbered objects, i.e. photographs, submitted in JSON format have been successively placed . The "$i" is a freely selected number from the range 1-17046, which is responsible for the execution of iteration and is at the same time a reference to a specific object located in the table. Objects placed in the table have the following values:

$array[$i][0] – title of photograph;
$array[$i][1] – timestamp 'epoch';
$array[$i][2] – medium;
$array[$i][3] – longitude;
$array[$i][4] – latitude;
$array[$i][5] – date: YYYY-MM-DD

The resulting table of activity contained 17046 records of all types, 418 records had the same error, consisting in the entry of the date attribute 01/01/70, which means that photographs added to panoramio.com did not have a correctly entered date in the exif data, or that it was not defined in any way (Tab. 6). This repeatable error was excluded by the removal of specific records from the table following the application of the SQL instruction: DELETE FROM object WHERE timestamp = '0' OR timestamp = 'NULL'

Classification queries were formulated at the data processing stage, and in the course of subsequent operations they allowed to create the geovisualisation panels.

2.4. Visualisation and implementation

The visualisation stage concerned the distribution of panels in the geovisualisation window in accordance with the designed lay-out (Fig. 3) and adapting graphics to visually strengthen the selected activity. The objective of graphical parameters management was to show the user a single selected activity in relations with other activities with reference to spatial entities on topical maps and graphs.

The technology of operation of the geovisualisation window application is based on JavaScript entries using asynchronous queries to the WWW server and the database server (Fig. 4). Such possibilities are provided by AJAX (Asynchronous JavaScript and XML), which enables refreshing individual panels without the need of reloading the entire page.

Table 6. Resulting table in the database following the import of data concerning activities

Id	medium	Title	timestamp	Date
19226	z	Swarzędz - ogródek nad Cybiną	1241191587	2009-05-01
19227	z	Kościół w Sobocie	1220764038	2008-09-07
19228	z	Nature	1187176755	2007-08-15
19229	z	Zachód słońca	1313432801	2011-08-15
19230	z		0	1970-01-01
19231	z	kościół św. Floriana w Wirach	1314113344	2011-08-23
19232	z	GŁĘBOCZEK	0	1970-01-01
19233	z	Wielkopolska, Kopanina	1185701772	2007-07-29
19234	z	Mur	1289738405	2010-11-14
19235	z	Und los geht's....	1149654190	2006-06-07
19236	z	Starynka	1298113468	2011-02-19
19237	z	Nasze polskie bociany	1214554353	2008-06-27
19238	z	Sapowice	0	1970-01-01
19239	z	jezioro(lake) Góreckie	1167480396	2006-12-30
19240	z	Błażejewko - Jezioro Bnińskie	1248617728	2009-07-26
. . .	. . .	. . .	. . .	. . .

The usage of interactive geovisualisation elements is connected with the initiation of queries to PHP scripts, and it is this stage that is responsible for the creation of queries, e.g. indicating the season of the year for which data is to be collected from the database, or indicating the location of the view (Fig. 4). The PHP script executes queries to the database with previously specified parameters, in the SQL language and taking into consideration spatialSQL. The response from the database is returned in text form. The transformation of TXT responses consists in the appropriate processing of the returned query from the database through the PHP script. Processing enables straightforward implementation during the next stage. Finally, JavaScript supports implementation into the geovisualisation window, which consists in introducing values processed during the preceding stage into specific geovisualisation panels. This may be considered as "refreshing" of the geovisualisation panel for the view adopted by the user on the dot map – choropleth map.

The key action for initiating the query on the part of the user consists in shifting the rectangle of the view on the general map and clicking the selected temporal scope. Then the appropriate view data are collected from the database on the side of the server. The query comprises the WHERE condition, in the which the temporal range is defined, e.g. spring:

WHERE (

(EXTRACT(MONTH FROM timestamp::abstime::timestamp) BETWEEN 4 AND 5) OR

((EXTRACT(DAY FROM timestamp::abstime::timestamp) BETWEEN 21 AND 31) AND (EXTRACT(MONTH FROM timestamp::abstime::timestamp)=3)) OR
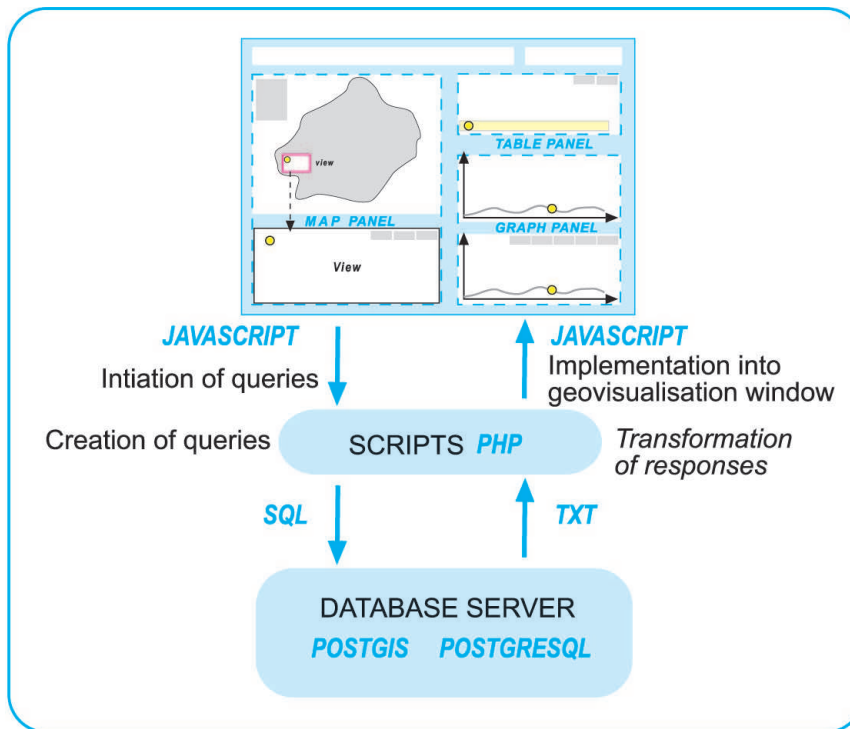
Fig. 4. Technological operation chart of the initiation of queries

((EXTRACT(DAY FROM timestamp::abstime::timestamp) BETWEEN 1 AND 22) AND (EXTRACT(MONTH FROM timestamp::abstime::timestamp)=6))
)

The WHERE condition also serves to collect only those activities from the database that are contained in polygons defined by the geometry column for MUNICIPALITY tables:

WHERE ST_Contains(a.the_geom, b.point)

During the transformation of responses there are created the appropriate view lines (in two graphs) and a table of attributes, which corresponds to the contents of the view panel.

The second key action that may be performed by the user is clicking a randomly selected dot (activity) on the view, which then changes its colour from red to yellow – both on the general map and on the enlarged view. In addition, the record corresponding to a given activity in the table of attributes is highlighted. Furthermore, the marked activity is placed on municipality activity graphs, broken down into both years and hours of activity. This effect was achieved by using JavaScript and manipulating on objects defined in the html language.

The graphs were elaborated dynamically using the Google Charts Api tool. This tool allows it possible to generate and personalise graphs by means of the appropriate

url link, which connects to Google sites where the graph is prepared as a png image with parameters corresponding to those placed in the url link. The "refreshing" of new graph values for the interactive view window consists in effecting the transformation of responses from the database into a url link with specific parameters. The table of attributes has been executed using the html language and the css styles sheet. The "refreshing" of the table consists in effecting the transformation of responses from the database into a html code, which is subsequently implemented into the panel.

3. Results

In accordance with the assumptions made at the conceptual stage, the final action was the implementation of the window entitled 'Geovisualisation of the Panoramio.com Activities in District of Poznan 2011' into the web browser (Fig. 5). The map presents the spatial lay-out of activities on the social website: uploading photographs with locational information and other attributes connected with data of this type.
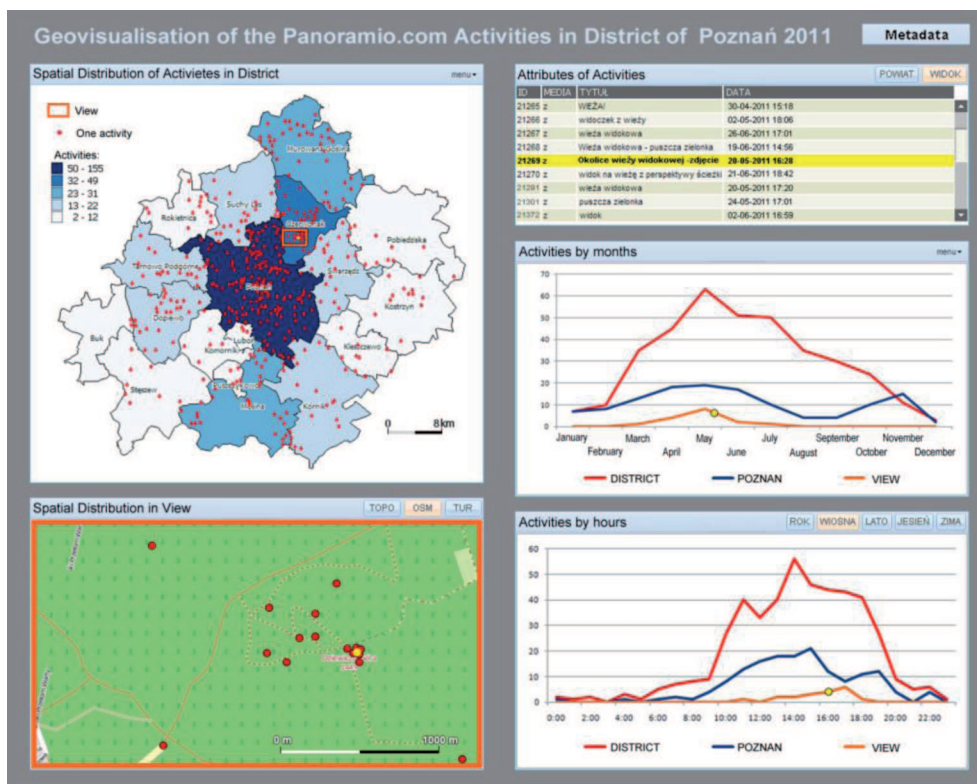

Fig. 5. Geovisualisation window in the web browser

In accordance with the principle of learning about space progressing from the general to the specific, the choropleth map identifies municipalities with maximal and

minimal activities, and assigning sets of points within these entities. Using the view rectangle, the user selects the area in which he intends to investigate the detailed location of dots on the map on a larger scale and, by clicking the appropriate reference map, may interpret the distribution of activities. The example presented in Figure 5 shows a set of activities at a touristically attractive object – Dziewicza Gora mountain. Each dot on the general and detailed map may be clicked, which also results in the backlighting of the attributes of this activity in the table, thus making it possible among others to read the comment to the photograph, e.g. "Vicinity of the viewing tower".

The presentation of activities on two graphs for the district, for Poznań, and in the view window facilitates a comparison and interpretation of the number of activities in these three spatial ranges, broken down into both months and hours for the year 2011. The distribution of activities over the course of months indicates that the internet community of the researched area most readily takes pictures in spring. The hourly distribution attains a maximum in the afternoon hours. An interesting dependence is that the decrease in activity between 16:00 and 18:00 in Poznań is accompanied by an increase in the number of activities in the view area, in which in this particular case is a recreational zone.

4. Conclusions

Until now, research related to the geographical visualisation for web users has focused on the search for new data management methods by users with at least a general knowledge of GIS (Andrienko and Andrienko 2006). In contrast, our study is also addressed to members of social media, as spatial data also relates to VGI. The authors of this study believe that it is imperative to guide the direction of analysis by appropriately placing panels in the geovisualisation window and by adapting interactive functions.

The fundamental advantage of the elaborated method of interactively geovisualising the temporal and spatial variability of point phenomena retrieved from a Web 2.0 application using VGI is its versatility, for it may be used to analyse each and every internet community. The procedure presented in Figure 2 contains successive stages of elaborating an interactive visualisation, which is viewed by the user of the web browser. The concept has been based on a division of the geovisualisation window into three panels, of which the most important – in order to capture spatial variability – contain statistical maps at the general level (dot map and choropleth map), while at the detailed level – a dot map on a topographic reference map or tourist map. For two ranges, temporal variability is presented by graphs, while a review of attributes of individual activities of the social website in question is set forward in the table panel. The element that visually links all of the panels is the selected individual activity. Such graphical highlighting may be considered as a deliberate focusing of the observer's attention used in topical maps (Dent, 1999), as well as the utilisation of the principles of creating point signs on maps (Medyńska-Gulij, 2008).

The processing of data took place in an open source software thanks to the utilisation of free desktop software, database software, and internet technologies. The

technology of operation of the geovisualisation window application is based on AJAX, that is the utilisation of asynchronous queries to the WWW server and the database server, which ensures the efficient "refreshing" of individual panels. The user of the web browser has received a tool that may be easily used to analyse the VGI phenomenon for Panoramio.com. Research may be continued with the objective of testing the functionality of the geovisualisation window with respect to data features or descriptions of members of each social media, which are becoming an ever more important source of spatial data.

Acknowledgments

References

Andrienko, N. & Andrienko, G. (2006). *Exploratory Analalysis of Spatial and Temporal Data: A Systematic*
    Approach. Berlin: Springer-Verlag.

Baranowski, M. Bielecka, E. & Dukaczewski, D. (2009). *Methods of portrayal spatial data used in official geoinformation services in Poland – comparative study*. Lecture Notes in Geoinformation and Cartography. (Eds.), G. Gartner, F. Ormeling, Cartography for Central and Eastern Europe. Heldeiberg Dordrecht London New York: Springer, 41-62.

Bielecka, E. (2003). *Spatial data usability in the Polish Spatial Information System*. CODATA Data Science Journal, 2, 128-135. Retrieved April 2, 2012, from J-STAGE on the World Wide Web: https://www.jstage.jst.go.jp/browse/dsj/2/0/_contents.

Coleman, D.J. (2010). The Potential and Early Limitations of Volunteered Geographic Information. *Geomatica*, 64(2), 27-39. Special Issue on 50th Anniversary of the Department of Geodesy and Geomatics Engineering at the University of New Brunswick.

Dent, B. (1999). *Cartography: Thematic Map Design* (5th ed.). Boston: McGraw-Hill.

Dickmann, F. (2004). Mehr Schein als Sein? – Die Wahrnehmung kartengestützter Rauminformationen aus dem Internet, *Kartographische Nachrichten*, 2, 61-67.

Dykes, J. MacEachren, A.M. & Kraak, M-J. (2005). *Exploring geovisualization*, Amsterdam: Elseviers. Imprint: PERGAMON.

Elwood, S. (2008). Volunteered Geographic Information: Future Research Directions Motivated by Critical, Participatory, and Feminist GIS. *GeoJournal* 72(3 & 4), 173-183.

Flanagin, A. Miriam, J. & Metzger, J. (2008). The credibility of volunteered geographic information. *GeoJournal* 72(3 & 4), 137-148.

Goodchild, M.F. (2007). Citizens as sensors: the world of volunteered geography. *GeoJournal* 69, 211-221.

Kuhn, W. (2007). *Volunteered Geographic Information and GIScience*. NCGIA, UC Santa Barbara, 13-14 December, 2007.

MacEachren, A.M. Crawford, S. Mamata Akella & Lengerich, G. (2008). Design and Implementation of a Model Web-based, GIS-Enabled Cancer Atlas. *The Cartographic Journal* 45, 246-260.

McDougall, K. (2009). *Volunteered Geographic Information for Building SDI*. In: Ostendorf B., Baldock P., Bruce D., Burdett M. and P. Corcoran (Eds.), Proceedings of the Surveying & Spatial Scien-

ces Institute Biennial International Conference, Adelaide 2009, Surveying & Spatial Sciences Institute, 645-653.

Medyńska-Gulij, B. (2008). Point Symbols: Investigating Principles and Originality in Cartographic Design. *The Cartographic Journal* 45, 62-67.

Medyńska-Gulij, B. (2011). *Cartography and geovisualization*. Waszawa: PWN.

Slocum, T.A. McMaster, R.B. Kessler, F.C. & Howard, H.H. (2009). *Thematic Cartography and Geographic Visualization* (3rd ed.). Upper Saddle River: Pearson Prentice Hall.

Okno geowizualizacji czasowo-przestrzennej zmienności dla działalności Volunteered Geography Information

Beata Medyńska-Gulij, Miłosz Myszczuk

Zakład Kartografii i Geomatyki,
Uniwersytet im Adama Mickiewicza w Poznaniu,
Collegium Geographicum, ul. Dziegielowa 27, 61-680 Poznań
e-mail: bmg@amu.edu.pl

Streszczenie

Problemem podjętym w tych badaniach stało się wykorzystanie metod z nurtu geograficznej wizualizacji do wskazania cech fenomenu VGI w zakresie zmienności czasowo-przestrzennej.

Zgodnie z założeniami poczynionymi w etapie koncepcyjnym finalnym działaniem stało się zaimplementowanie do przeglądarki internetowej okna pod tytułem: "Geowizualizacja aktywności społeczności Panoramio.com w powiecie poznańskim w 2011 roku". Koncepcja została oparta na podziale okna geowizualizacji na trzy panele, z których najważniejsze znaczenie dla uchwycenia zmienności przestrzennej na poziomie ogólnym ma kartogram, natomiast na poziomie szczegółowym mapa kropkowa wyświetlana na podkładzie mapy topograficznej lub turystycznej. Zmienność czasową w dwóch zakresach prezentują wykresy, a przegląd atrybutów poszczególnych aktywności prezentowanego portalu społecznościowego zapewnia tabela. Elementem spajającym wizualnie wszystkie panele jest wyeksponowana graficznie pojedyncza aktywność.

Przetwarzanie danych odbyło się w środowisku open source, a technologia funkcjonowania aplikacji okna geowizualizacji bazowała na zastosowaniu asynchronicznych zapytań do serwera WWW oraz serwera baz danych co zapewnia sprawne „odświeżanie" poszczególnych paneli.

www.czasopisma.pan.pl

PAN
POLSKA AKADEMIA NAUK

www.journals.pan.pl

The statistical geoportal and the "cartographic added value" – creation of the spatial knowledge infrastructure

Anna Fiedukowicz[1], Jędrzej Gąsiorowski[2], Paweł Kowalski[1], Robert Olszewski[1], Agata Pillich-Kolipińska[1]

[1]Warsaw University of Technology, Faculty of Geodesy
and Cartography, Department of Cartography
Plac Politechniki 1, 00-661 Warsaw, Poland

[2]Institute of Geodesy and Cartography,
Department of Spatial Information Systems and Cadastre
27 Modzelewskiego St., 02-679 Warsaw, Poland
e-mail: r.olszewski@gik.pw.edu.pl; p.kowalski@gik.pw.edu.pl;
a.pillich@gik.pw.edu.pl; jedrzej.gasiorowski@igik.edu.pl

Abstract: The wide access to source data, published by numerous websites, results in situation, when *information* acquisition is not a problem any more. The real problem is how to transform information in the useful *knowledge*. Cartographic method of research, dealing with spatial data, has been serving this purpose for many years. Nowadays, it allows conducting analyses at the high complexity level, thanks to the intense development in IT technologies, The vast majority of analytic methods utilizing the so-called *data mining* and *data enrichment* techniques, however, concerns non-spatial data. According to the Authors, utilizing those techniques in spatial data analysis (including analysis based on statistical data with spatial reference), would allow the evolution of the **Spatial Information Infrastructure** (SII) into the **Spatial Knowledge Infrastructure** (SKI). The SKI development would benefit from the existence of statistical geoportal. Its proposed functionality, consisting of data analysis as well as visualization, is outlined in the article. The examples of geostatistical analyses (ANOVA and the regression model considering the spatial neighborhood), possible to implement in such portal and allowing to produce the "cartographic added value", are also presented here.

Keywords: geostatistical analysis, geostatistical atlas, geoportal, spatial knowledge infrastructure

1. Introduction

In order to meet the demands of the modern information society, analyzing the space is equally important as its cartographically correct modeling. At present, the term "modeling" is understood, first of all, as creation of reference spatial data bases, which describe particular components of the natural environment and relations between them. Its second meaning is development of sectoral GIS systems, basing on the topographic

canvas, defined that way. Thematic data, stored in such systems, contains information of economic, demographic, social conditions etc. The wide access to source data, published by numerous websites, results in situation, when ***information*** acquisition is not a problem any more. The real problem is how to transform information in the useful ***knowledge***.

The so-called, cartographic method of research (CMR) has been known for more than a half of a century (Saliszczew, 1955, 1968; Berlant, 1978, 1986). This method consists of utilizing maps (and nowadays also geographical data bases) for description, analysis and scientific recognition of phenomena, discovery new regularities of their locations and mutual relations as well as forecasting changes. The essence of the CMR relies on inclusion of an intermediate tool, a map – as a model of researched phenomena – into the process of investigating the reality. Nowadays it is usually a digital map. The contemporary analytical CMR systems, using IT technologies, allow conducting analyses at the high complexity level (Poole et al. 1998). One of the ways of such analysis is to utilize the, so-called, *data mining* and *data enrichment*, i.e. to detect regularities, rules and structures "hidden" in the data base. The concept which is wider than data mining, is *knowledge discovery from databases* – KDD. It is based on the assumption that information is "hidden" in the data base in the form of the data *structure*. Data mining is often equalized with KDD; however, it is only one of the stages of knowledge discovery (Miller and Han, 2001). Data mining is the process of „data analysis which leads to getting information", while KDD is the „higher order process, which comprises analysis of information acquired by data mining and transforming it into *knowledge"*.

The vast majority of contemporary research works, which utilize data mining techniques, concern analysis of non-spatial data, i.e. simple tables which contain, for example, commercial information. Analysts detect structures and relations hidden in big files of data and they state, for example, that "customers, who buy the product A often also decide to buy products B or C". Research performed by the authors enlarges the scope of applicability of that method to analysis of spatial data, stored in geographical data bases. Utilization of data mining techniques allows researching large data resources (such as digital maps related to economic, demographic issues) in order to find patterns and interrelations (Wachowicz, 2001). Detection of rules "hidden" in the geographical data base, allows construction of a decision system, which, for example, enables to forecast trends of development of particular agglomerations or regions.

Therefore, data processing, including spatial data processing, serves for ***explicite*** presentation of relations, which ***implicite*** occur in the source data base (Olszewski, 2006; 2008). *Spatial data mining* techniques, often explained as the explorative analysis of spatial data, are the indispensable element of the "*data enrichment*" process. Presentation of spatial structures proves the power of the cartographic method of research. Analysis of spatial information, understood in this way, is therefore the source of creation of the "cartographic added value" – the knowledge which explains source data, its spatial arrangements and which allows forecasting trends of development. That

concept is fully compliant with the idea of the geoinformation structure development, proposed by Iwaniak (2011). He states, that after the period of development of the web **Spatial Data Infrastructure** (SDI) in the nineties of the 20th century, allowing for locating, publishing and getting the access to spatial data, the beginning of the new century was connected with intensive development of the **Spatial Information Infrastructure** (SII), which allowed combining services, harmonizing and complex data processing. This also allowed for widening data publication over the web, with developed cartographic methods. However, following Iwaniak (2011), the further step towards the development of the geoinformation infrastructure is of key importance. Orchestration of spatial services, which has been recently observed, successive implementation of ontology, thesauruses and context processing of data published in the web, allowed evolution of the SII into the **Spatial Knowledge Infrastructure** (SKI). As a result of utilization of sophisticated algorithms of analyses and access to metadata, the infrastructure, understood in this way, is capable to "understand" data and services supplied by institutions, which distribute spatial data.

Following the authors' opinions, conditions of development of the SKI, defined by Iwaniak (2011) should be considered as necessary, but not sufficient. It is necessary not only to simultaneously utilize many "atomized" services, offered by various institutions and their mutual orchestration, but also to develop specialized dedicated services, which would support the process of spatial data analysis and transformations. In the authors' opinion, services of this type include implementation of statistical algorithms in digital mapping, which allows creating the knowledge base with the use of the "cartographic added value".

2. The role of a map in a website

Usually a user who reaches for geographical, statistical information or who intends to spatially locate events does not expect a form of information transfer other than a map or an atlas. Thus, the cartographic form of transfer meets the user's expectations. This is one of the important prerequisites of the product usefulness. Another positive result of utilizing maps in websites is related to intuitive maintenance and effectiveness of distribution of spatial information.

Due to diversity of data published in geoinformation websites, a map usually plays the role of integrating particular information components. The map is also the basic element of the website composition. Other elements, such as: text fields, queries, functional buttons, illustrations etc., are placed around it. Similarly to other elements of the user interface, a map may be used for operating the website. Its particular components: symbols, labels, specified legend fields, may be connected with specified operations. They can involve starting predefined functions, calling links to other parts of presentation etc. Lastly, a map remains an independent information transfer. Therefore, the competent combination of advantages of a conventional map and a useful application interface, is the basic challenge at the stage of designing the discussed systems.

2.1. Issues concerning website's functionality

The basic assumption concerning functionality (designing the usefulness) of websites is to achieve the required product quality, which is expressed by such properties as: functionality, quality and content timeliness, accessibility and practicality of use. Thus, the widely understood usefulness of a website is the resultant value of above features. It does not result only from the number and development level of functions or options, which characterize a given tool or a thing. As the validating feature, the usefulness is the total value, which specifies the number of benefits the user can get from a product, assuming low inputs of time and labor. In order to achieve the expected usefulness level, clearly specified rules of user centered design (UCD) should be applied considering the user expectations, as well as possibilities and limitations of particular customers (Kowalski, 2006).

Being an element, which constitutes the website interface, a map is the subject of both web and cartographic functionality. The cartographic usefulness covers such features as: the scope of content, the generalization level, cartographic presentation methods, graphical variables, cartographic projection, the legend scope and outlay and auxiliary elements. For the group of interactive maps all functional elements should be additionally considered, which – by analogy to a cartographic network – may be called the map's functional network. Usually, if the basic rules of map editing are not considered, the map readability decreases, its interpretation is more difficult and, as a result, the transfer efficiency is also lower (Kowalski and Mostowska, 2010). The incorrectly prepared functional network may be noticed at every stage of work with a map, leading to unexpected results.

2.2. The functional network of a map in the website

The functional network of an interactive map is wrapped on the entire cartographic image and on marginal elements (according to the conventional terminology). Every active element of presentation within the map area (signature, surface symbol, name, description), in the legend field, in the map frame or in additional toolbars, may cause a specified, internal or external action. Internal actions include, for example, the change of view parameters (zooming, range modifications), the content scope modification (operations on information layers), re-symbolization etc. External operations include hypermedia navigation – evoking hyperlinks to another (text, sound or image) document in the system. Such a high functional potential of a map strengthens its role in information systems, particularly on the web.

The following components may be considered as the **basic elements** of the cartographic user interface (Figure 1):
– a panel of information layers – a map control panel,
– a legend, which is sometimes included in the list of layers in the control panel,
– toolbars, including the bar of basic navigation tools,

Fig. 1. The example of an advanced cartographic interface, applied in the web portal of the Małopolska Spatial Information Infrastructure (geomalopolska.pl)

– a toolbar (or a panel) of selection of cartographic profiles in websites which offer a wide set of data,
– simple text labels.

The following components may be considered as **auxiliary elements** of the cartographic user interface, which directly depend on the scope of website's functionality (Figure 1):

– panels which control functions of the website; usually graphically coherent with the panels of information layers,
– legends of thematic overlays,
– additional toolbars and dialog boxes used in order to configure maps and user's objects (such as navigation map or comments),
– advanced information (text-and-image) labels.

The legend plays a particular role among marginal elements. In the case of interactive visualization it is closer connected with an cartographic image than in the case of conventional maps. It does not only play the role of a pattern of symbols, but it also comments the current application status or presented events (Kowalski, 2003). Additional modules are also necessary, which may be, for example, related to measurements of time and to controlling of animated maps. Special functions of interactive changes of views may be played by a common linear scale: zooming in and out or

changes in information resolution (e.g. size of reference units). Not all elements of a website, which occur together with a map (such as search forms) will be considered as the functional unity. Fields of spatial-and-time navigation and other buttons, which are instrumentally connected with a map, are also map components – particularly when their functionality without a map would be unreasonable.

At present, application map components are those elements, which mostly contribute to popularization of geographical information over the web (Kowalski, 2008). This results from many reasons. Starting from technical opportunities such as increase of web application possibilities, pragmatic features such as independence from specialized applications towards a universal web browser, to such unserious reasons as fashion. Regardless of the reasons, the marriage of cartography and teleinformation benefits both sides.

3. The geostatistical portal development

3.1. Conceptual assumptions and research challenges

Social-and-economic development, which happens at the era of the information society, requires not only the wide access to data and information, but, first of all, development and distribution of tools, which allow transformation of those source data into the useful and applicable knowledge. A good example of such approach is the governmental strategy of national development "Poland 2030", published in the web (http://zds.kprm.gov.pl/node/19), and its more detailed amendment, the Concept of Spatial Management of Poland (http://www.mrr.gov.pl/rozwoj_regionalny/polityka_ przestrzenna/ kpzk/aktualnosci/strony/default.aspx). The authors of those documents have formulated the thesis that development of the information infrastructure is not only a technological issue, but, rather, it is the civilization challenge.

In the strategic document the thesis has been formulated that the condition required for fast convergence of the Polish economy is the assumption of the "polarization diffusion" model, which should substitute the "sustainable development" model. This practically means to support "leading and dynamically developing regions", which will become local development "engines". However, the question – **which** agglomerations, municipalities or districts should be considered as those leading centers – remains very interesting. Similarly, it is worth considering which **particular** regions of the country require the diffused inflow of knowledge, technology or investments.

For example, one of development challenges considered in the document "The Concept of Spatial Management of Poland" is the analysis of "the high professional activity and adaptiveness of labor resources". Such analysis requires that data related to the employment and unemployment levels, as well as average salary levels, are considered. Diversification of the Polish labor market, including so-called development clusters (development centers, suburbs, cities, ex-state farms, areas of prevailing low agricultural productivity or rural-and-industrial eras), should be also considered

together with the analysis of participation of particular levels of education and forms of temporal employment. It is obvious that more similar issues may be formulated. However, in order to answer unit questions, advanced spatial data analysis as well as cartographic visualization and **interpretation** of obtained results should be performed – with respect to, for example, the uniformity of spatial distribution of results.

Questions of such type may be answered with an appropriately developed geoinformation structure of spatial knowledge. Following the authors' ideas, the step towards creation of the SKI is the development of a thematic geoportal, which – through the required orchestration of spatial services – will allow analyzing data distributed by various institutions. Utilization of *spatial data mining* techniques for analysis of GUS (The Main Statistical Office) or MRR (The Ministry of Regional Development) data will allow creating the "cartographic added value". This will contribute to development of the knowledge data base to support decision making processes.

At the beginning of this project's implementation it should be remembered that utilization of exploring data analysis techniques in geographic sciences is highly difficult and complex. Many issues which are analyzed using data mining techniques contain multidimensional data, however particular features are usually independent. Spatial data, which describes a four-dimensional space-time, is highly correlated. Analysis of such data requires that the existing relations are considered.

Statistical data is important for development of the information society. At present, such data is distributed by various institutions in various forms and ranges. One of the most important institutions, which distributes statistical data in the web in Poland, is the Main Statistical Office. However, such data is also distributed by other institutions, such as the State Election Commission or the Ministry of Regional Development. Although heterogeneity of data sources and methods of data distribution has some benefits, it may also result in serious difficulties in data access by an average user. The need to create a unified interface, allowing data selection using substantial criteria (such as themes of data, spatial range of data, level of details or temporal coverage) without the necessity to know all data sources, is the first prerequisite to develop an interactive statistical atlas. Its main task would be to become a point of the common access to distributed statistical data, like geoportal.gov.pl plays the similar role with respect to geospatial data.

Apart from data integration through geoinformation services, one of the main tasks performed by the discussed atlas would be cartographic visualization of collected data. Therefore, it should ensure simple visualization tools, allowing the creation of choropleth or diagrammatic maps, as well as selection of colors for those presentations, transparency or the type of reference background. Such presentation, which would follow the rules of cartographic art, would allow better perception of retrieved data by not highly qualified users, thanks to creation of the so-called "cartographic added value". The operation of data visualization itself contributes to better data perception and creates the possibility to draw basic conclusions or to notice regularities, which govern the data, as a result of the correctly selected graphical form of transfer. In order to achieve above results it is necessary to integrate (using for example TERYT codes)

statistical data from various institutions, with geometric data of administrative units' borders, for which they are mostly acquired and aggregated.

The manner of presenting statistical data only in tabular form dominates in existing statistical portals. Not so many provide spatial reference for the data, and even fewer offer partial interactivity in creating thematic visualizations. A short research in this field, conducted by (Fiedukowicz and Gąsiorowski, 2012), confirms the need to develop the concept of geostatistical portal.

At the stage of designing the geostatistical portal, the basic assumption was to utilize – to the maximum possible level – web services organized in the form of an interactive statistical atlas, basing on diversified sources of statistical data and methods of their processing. The idea of the public access to data acquired by various institutions was to be also implemented with the use of an interactive, user friendly cartographic interface of the website (Figure 2). The characteristic feature of such interface would be the richer selection of tools available to the user, comparing to the most popular mapping location websites. The tool set should contain not only elements controlling the cartographic image, but also a wide set of buttons used for controlling tasks of geoprocessing source data (the statistical analyses subsystem).
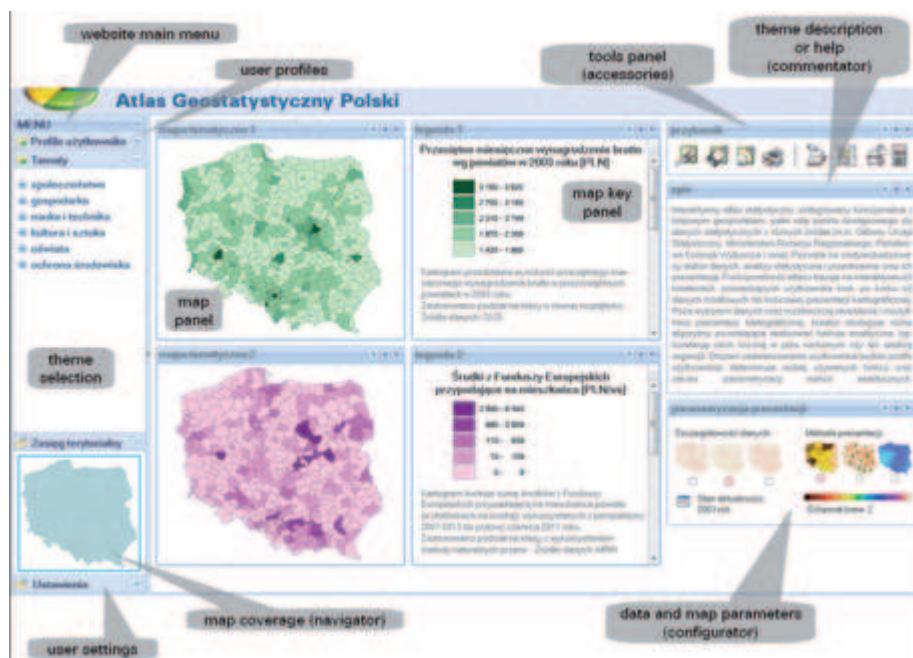

Fig. 2. The design of the User Interface of the Geostatistical Atlas of Poland

An important application component of the designed website would be the creator of thematic maps. It should consider the possibility of independent user selection of such parameters as: data selection, method of geostatistical analyses (e.g. measures of

correlation and regression, movable field etc.) and specification of the cartographic visualization (choropleth map, diagrammatic map, isoline map etc.).

An important aspect of design is consideration of composing schemes, following the rules of cartographic symbolization and generalization. This apparent limitation of the user freedom would guarantee, however, the correct readability and logic of the final cartographic image. At the same time, it would accelerate work for the majority of users who are interested in the final result and not in experimenting with visualization parameters. It will be also possible to utilize pre-made presentations, placed in the basic version of the geostatistical atlas, which do not require the use of the map creator.

3.2. Application components of the developed website

As it was mentioned above, destination of the designed website, different from typical geoinformation portals, will result in modification of some control bars' roles and in development of the user toolbar. In the interface's design there are various fields of lists and fields of selection, which are used for the user interaction with the analyses subsystem. The subsystem of presentation will also answer to the user requests in a modified scope (Figure 3). For example: in a typical map portal the basic operations of the control panel include the change of scale and the level of details of the view (zooming in and out). In the geostatistical website the map scaling mainly influences the change of the information resolution (e.g. the size of data reference units – municipality, district, voivodeship) and not the change of the content.

The information layer panel (the map management panel) – in the case of reference data – is used mainly to switch the visibility of thematic layers (the content change) (Kowalski, 2008). In the designed atlas, the majority of presentations will be limited to several thematic layers. Therefore, tools for modification the ways of presentation and symbolization should be dominating.

The important difference will be noticeable in the scope of the legend field. In many geo-location websites the legend is reduced, displayed optionally or it does not exist at all (Kowalski and Mostowska, 2010). In the geostatistical atlas the presence of the legend results not only from the necessity to explain symbols, but also from the possibility to adapt it to control the presentation. That could involve changing the limits of intervals for steps of measuring scales (choropleth maps, diagrammatic maps). A very important role will be played by the toolbars for selecting cartographic profiles: as a standard – to select a satellite image background or shaded terrain relief. In the case of the discussed project they will be used for the initial organization of the user interface. It will include the basic version of the atlas, the advanced interactive atlas (with the possibility to activate a two- or four-sheet composition for the need of comparing), and the user atlas (with the option of thematic layers creation).

Other tools are listed below:
– tools of names and map descriptions (labeling): simple text labels, advanced information (text-and-image) labels,
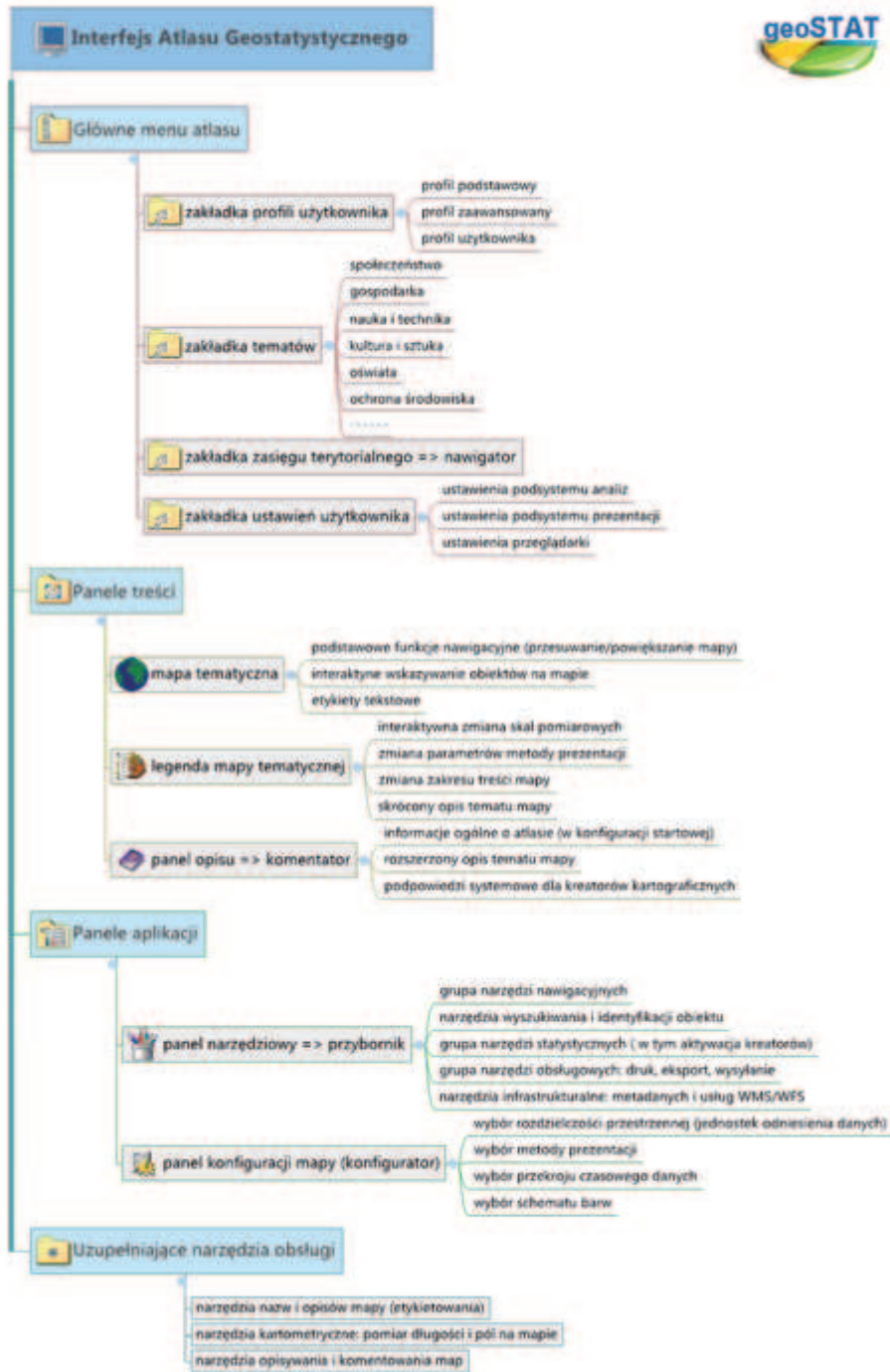– cartometric tools: measurements of distances and areas on a map,

Fig. 3. The functional diagram of the website accessible in the form of the Geostatistical Atlas of Poland

– searching tools according to descriptive or topological criteria, tools of information about the object (object identification tools),
– metadata tools for information files and thematic layers,
– infrastructural tools: adding WMS/WFS services to the user atlas,
– additional toolbars for comments, saving presentation, printing and exporting a map.

3.3. Examples of website's functions and ways of their utilization

Besides outlining the idea of the website and development of its graphical interface, the authors also attempted to practically utilize *data mining* techniques to analyze spatially located statistical data.

The example described below discusses (at the level of a municipality) the percentage of individuals of more than 65 years of age in the population of the Masovian Voivodeship. Three various approaches of presentation of this phenomenon using the statistical geoportal have been proposed. Those approaches are the results of certain statistical analyses, performed for source data, starting from the simplest analysis and ending with the more complex analysis which requires calculations of a series of values and statistical characteristics.
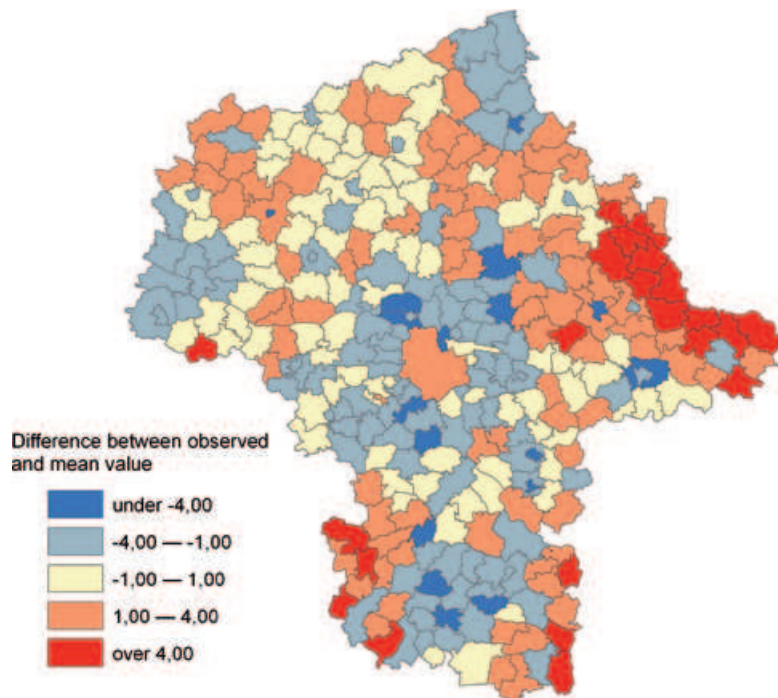

Fig. 4. Percentage of individuals of more than 65 years of age expressed as the difference between values observed in particular municipalities and the mean value

The fist variant assumes presentation of the percentage of people above 65 years of age in the entire population for each municipality, in the form of differences between the observed values and the mean arithmetic value, calculated basing on all observations (for municipalities within the analyzed area). The results of such analysis are presented in Figure 4. This simplest analysis requires that all values of the required attribute are collected by the web service, that the mathematical mean value is calculated basing on values of those attributes, as well as that differences between the mean value and values observed for each municipality are calculated; results are than displayed by the WMS service, with the use of defined symbols.

The second variant is based on the analysis of the influence of a certain factor on the observed percentage of individuals of more than 65 years in the population for each municipality. The discussed factor is the characteristics of a municipality, which classifies it with respect to the urbanization level. Three classes of the municipal status: urban, urban-and-rural, rural, are distinguished. Division of municipalities with respect to their status in the Masovian Voivodeship is presented in Figure 5.
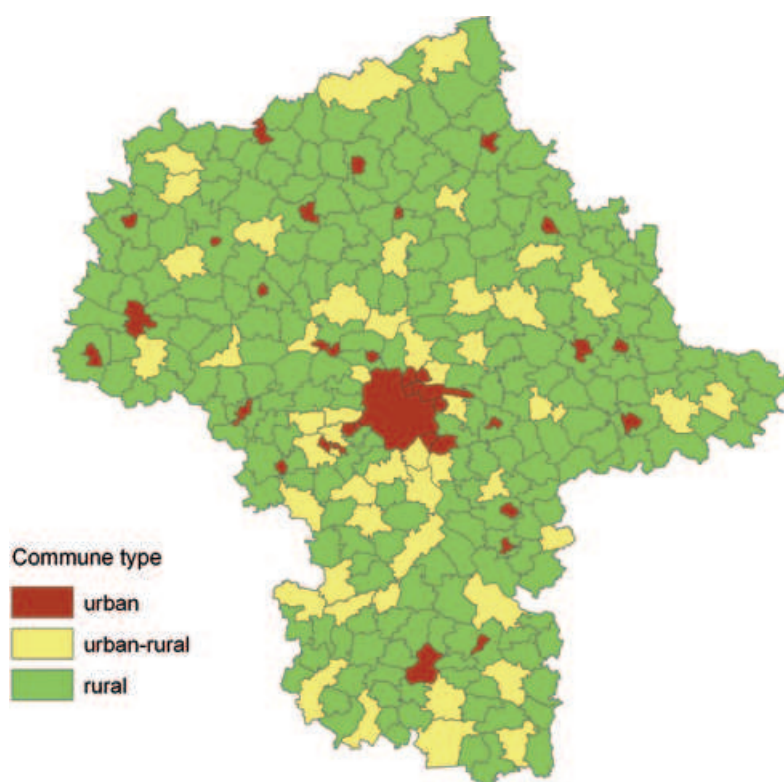

Fig. 5. Municipalities divided by status

This variant requires that all values of corresponding attributes are received by the web service. In this case, it's the attribute representing the percentage of individuals of

more than 65 years of age in the population and the attribute which represents the status of a municipality. Then a series of calculations, required by the single-factor **an**alysis **o**f **va**riance (ANOVA) model are performed (the integer total of squares of deviations, the number of freedom levels, the mean square of deviations, the statistics F). The essence of this analysis is to investigate the importance of differences between mean values calculated on the basis of several groups (three groups in the discussed case), into which observations (municipalities) were divided. The next step should comprise verification of the validity of the hypothesis saying about the influence of the factor (the independent variable – status of a municipality) on the investigated phenomenon (the dependent variable – participation of individuals >65 years). For this purpose the F value should be compared with the limit value, read from the F-Snedecor distribution matrix (existing in the repository) for theassumed level of importance, basing on the levels of freedom of the numerator and the denominator. When the F value is greater than the critical value, calculations of mean values in groups, as well as differences between the mean and observed values will be performed and then the result will be displayed by the web service in the geoportal. In the opposite case, a message should be displayed, informing about the lack of justification for such visualization due to the lack of relations between the variables.
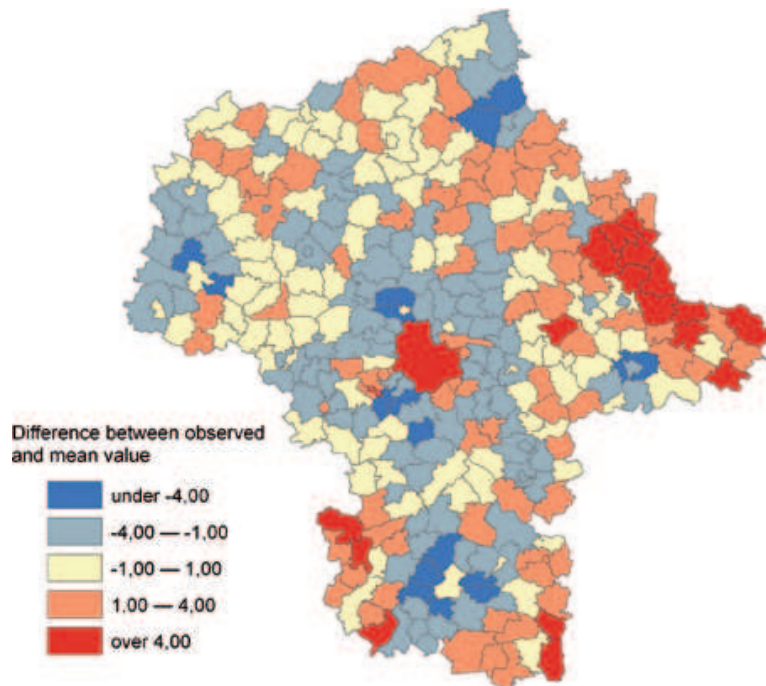

Fig. 6. The diagram of interactions between the percentage of individuals above 65 years of age expressed as differences between the values observed in particular municipalities and mean values for groups determined by the status of a municipality

Visualization of the performed analysis will illustrate the difference between values observed in municipalities and forecasted values (in ANOVA model these will be the mean values in particular groups). Such visualization is presented in Figure 6.

The graphical interpretation of identified interactions between the dependent variable (participation of individuals >65 years), and the independent variable (status of a municipality) is presented in Figure 7.
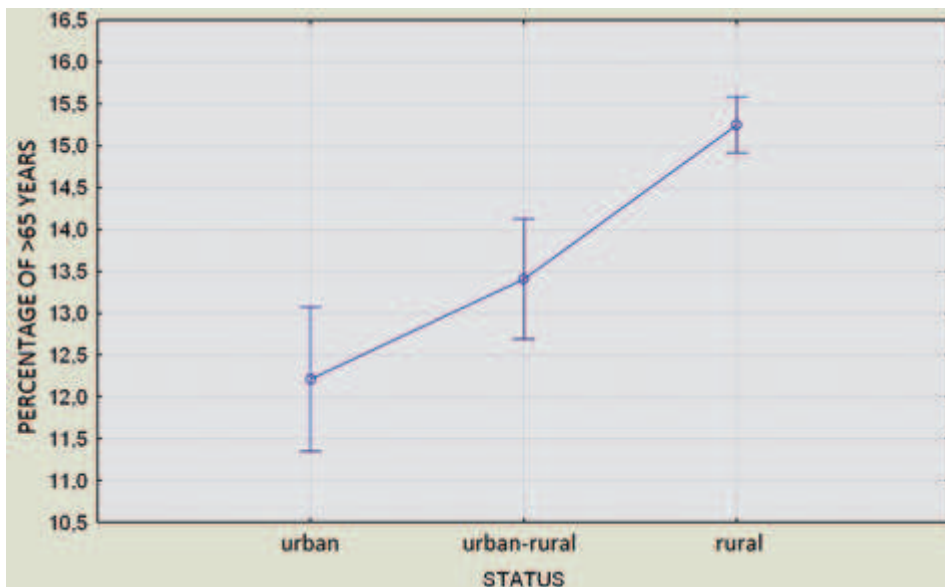

Fig. 7. The diagram of interactions between the percentage of individuals above 65 years of age and the status of a municipality

The third and the most advanced variant of visualization assumes the simultaneous influence of two factors on the dependent variable, which is the percentage of individuals of more than 65 years of age in the population. Similarly to the previous variant, the ANOVA model will be utilized for verification of the hypothesis concerning the simultaneous influence of those two factors; in this case this will be the multi-factor model for permanent results. Besides the status of a municipality, the effect of distances from the administrative center (the Capital City of Warsaw) will be also analyzed. One of three values, which characterize these distances, has been assigned to each municipality in the voivodeship: close, medium, far. Figure 8 illustrates the division of municipalities of the Masovian Voivodeship with respect to distances from the administration center.

The result of this analysis will be visualization presenting differences between values observed in particular municipalities and forecasted values (Figure 9). Similarly to the previous variant, they will be mean values in particular groups. In this case 9 groups will be formed, what results from assumption of all possible combinations (3 groups distinguished for both independent variables).