ZIPMAP: LOCALIZING MULTIPLE POINTS OF INTEREST - MAPS AND STATISTICS

Julia FERENCZ[1], Orsolya PENZES[1], Gyula USZKAI[2], Florina DUDU[1], Evelin HORVATH[1], Florian Mircea BOIAN[3], Remus ORASAN[4]
[1] *OnkoZert GmbH, Certification Institute of the German Cancer Society, Gartenstr. 24, Neu-Ulm, 89231, Germany, j.ferencz@onkozert.de*
[2] *DotWhite SRL, Eroilor 5/6, 400129, Cluj-Napoca, Romania*
[3] *Babeş-Bolyai University, Fac. of Math. and Computer Science, M. Kogălniceanu 1, 400084, Cluj-Napoca, Romania*
[4] *Iuliu Haţieganu University of Medicine and Pharmacy, Physiology Dep., Clinicilor 1, 400006, Cluj-Napoca, Romania*

Abstract: Geocoding has gained in recent years a great importance in fields as environment, ecology, engineering and public health. There have been software applications developed, which, using geocoding, generate different maps based on postal codes or coordinates considering users' preferences, but rarely conduct further various analysis. Localizing points of interest on a map and analyzing then the coverage area can be the start of a wide variety of research and statistics. These rare applications can be usually downloaded for a fee. Our application, ZipMap, is a free web based application which can be accessed and used for generating different types of maps using postal codes as points of interest. Using postal codes to generate maps is more user friendly, because this information is more accessible and less time consuming then using latitude and longitude. Also, the maps can be divided in smaller units, such as Federal States, counties, districts and also postal codes. Beside map generation, individual statistics can be generated regarding different incidences. Excel files can be downloaded containing the number of hits for postal code and theoretical incidences regarding geographical areas and population. The field in which this application has been used until now is the medical field. Hospitals can upload their patients' postal codes. The generated maps then show the coverage area of a hospital, which are the locations from where most patients come to the specific hospital. This can be analyzed and the hospital can then take actions in order to absorb more patients to the unit. This is just an example of how useful our applications' results can be in fields that influence everyday life quality or important marketing or state decisions.

Keywords: *maps; point of interest; coverage area; zip/postal codes; geocoding.*

I. INTRODUCTION

Living in an ongoing developing world, one needs at some point in his/her life to find a location that is previously unknown for him/her. This is when the importance of a map service materializes. The use of digital maps has increased substantially in recent years. Many times we want to localize more than just one point of interest on a map [1, 2].

This is when we stumble across the problematic, that most of the available free map servicing programs allow us to localize only one or a few points of interest on a map. There are also a few software applications that can be downloaded for an additional fee, which allow the localization of more than one point of interest by uploading files which contain coordinates. But when handling thousands of locations, searching for the coordinates for each of them can be time- and nerve-consuming. In these situations a simplified search after zip codes would be very welcomed.

The problematic of localizing exact geographical points has been present since the seventeenth century. The main interest at that time was the military field [3]. The lead country in this aspect has been England, followed by France, which tried through different expeditions to the sea, to determine the latitudes and longitudes of the Earth, this being the base of modern geocoding [4]. This has not only the obvious application in localizing different points of interest, extensively studied in recent years [5], but provides also further possibilities in analyzing movement patterns. These patterns can then provide very useful information in epidemics for example, localizing ground zero and the possible high risk areas [6]. There is a wide variety of fields in which mapping can be an asset: ecology, environment, engineering, social science and public health [7].

II. DESCRIPTION OF PURPOSE

There are many software applications available online in this field, each of them offering different localization possibilities. A very detailed comparison can be found in Wikipedia [8].

Shortly, we can resume some of the well-known mapping applications regarding localization of points of interest as follows:

-Google Maps [9], OpenStreetMap [10] and Nokia-Here Maps: localization of one point by address, zip code or coordinates; route planning between two points. Supplementary, Google Maps can import lists of coordinates, but the import of zip codes doesn't work correctly for all countries, without additional information.

-Microsoft MapPoint (a discontinued [11] software program): data mapping from various sources including

Manuscript received December 20, 2015; revised February 2, 2016

Microsoft Excel and Visual Basic for Applications interface allowing automation of the MapPoint environment [3].

-MapCreator: localization of multiple points of interest by importing an Excel file with coordinates; multiple selection (step by step) of cities from a dropdown.

Figure 1. Example of zip codes map.

Our web-based application, ZipMap (www.zipmap.de) [12], is a free software which can localize many points on a map based on zip codes. Beside the conventional use of localizing one point of interest on a map, ZipMap is an application which enables the user to localize more points of interest, this way making it possible to do different analyses regarding geographical localizations, coverage area, density, incidence, theoretical probabilities. The application can also do different calculations considering not just countries, but also federal states, districts and counties. This way the search or statistic can be narrowed down to a very specific area.

The application can be used on a wide range of fields (geographical, medical and any type of market research). It can be seen as a management tool creating a link between technology and different types of market research [13]. Our point of focus and area of implementation was represented until now by the medical field.

We uploaded different lists with postal codes of different clinics and practices. This way we could generate maps for hospital types, regarding the treated organ types. Very representative maps were created showing the coverage area of hospitals in Germany. This is a very important implementation, because this way we can observe the distribution patterns of certified clinics in Germany.

This kind of analysis can be made for other fields too, showing that the implementation possibilities of our application are wide and can give important results for different analyses. Another example for usage is for the hospitals that want to see where their patients come from, by uploading their zip codes. These types of maps show the coverage area of a hospital. For example, if we would consider a hospital in city "A" and want to see the absorption area of patients, it could be represented like in Figures 1 and 2. Here we gave as example that 81 patients come from the city "A" marked with a thick black border, and the other patients came from the periphery. Regarding the federal states, we can observe that patients come from two states. This way, the hospital can analyze why his coverage area has a specific pattern. Do they collaborate with enough private practices or not? Are these sending patients to this hospital or to other hospitals for further investigations? Are they treating patients also from other counties or federal states?

Another advantage of the soft, is that one can use it simply to verify a list of postal codes from Germany if they are still accurate, as the application renders a list of the wrong postal codes that have been uploaded. The maps and the statistics can be displayed not only for one entire country, but also divided in federal states, districts, counties, or postal codes.

Internally, within our company, we use a web service of ZipMap to generate maps for another application called StudyBox (www.studybox.de) [14]. This website renders statistics about certified clinical studies.

Figure 2. Example of federal states map.

III. TECHNICAL ASPECTS

III.1 GENERALITIES, FUNCTIONALITIES AND CALCULATIONS

The application is based on the postal codes of Germany. Every introduced postal code will appear on the generated map. German population data were used from the census in 2011. Geocoding and surface areas were calculated based on the data base from "OW networks GmbH". As zip codes often change, it is very important to maintain the data base actualized [10]. The total number of cancer patients from Germany was used from the data base of GEKID

(Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V.).

Regarding user interface, the user-friendly aspect is a very important part of an application, maybe the most important part, which ensures the application's success. The user is guided from step one with intuitive buttons and messages, so he/she can follow the precise steps which must be followed in order to receive the requested results.

What do the users have to do? A list with postal codes in ".xlsx" format is necessary. This has to be uploaded on the webpage. A template for this list can be downloaded with a few zip codes as an example (Figure 3). This can then be changed and the zip codes of interest can be introduced. This file has to be saved, changed and then uploaded.

	A
1	**Zip codes**
2	80331
3	10115
4	10115
5	71032
6	88299

Figure 3. Example of upload file.

The software recognizes the valid postal codes and the invalid ones. The invalid ones can be corrected directly in the interface or in the Excel file and then re-uploaded. After the wrong zip codes are corrected, and only then, the map is generated. Then the user can choose different options for the map and customize it for federal states, districts, counties, and cities. Different analysis files in Excel can then be downloaded (Figure 4).

The analysis of postal codes shows a list with the number of appearances of each postal code and the affiliation to its county, district and federal state (Figure 5).

	A	B	C	D	E
1	**Zip code analysis**				
2	Zip codes	County	District	Federal state	Results
3	10115	Berlin, Stadt		Berlin	2
4	71032	Böblingen	Stuttgart	Baden-Württemberg	1
5	80331	München, Landeshauptstadt	Oberbayern	Bayern	1
6	88299	Ravensburg	Tübingen	Baden-Württemberg	1

Figure 5. Examples of analyses for zip codes

The analysis of counties shows a list with the number of appearances of all postal codes for each county and the affiliation to its district and federal state. It also shows the population (male and female), the surface area of that county and the theoretical incidence and probabilities. Similar analysis can be downloaded regarding the districts (Figure 6). Important to mention, is that not all federal states in Germany have districts. In these cases the calculations for districts will contain the federal states.

The analysis of federal states shows a list with the number of appearances of all postal codes for each federal state. It also shows the population (male and female), the surface area of that federal state and the theoretical incidence and probabilities (Figure 6).

In the following we can see the used formulae and the explanations with examples. For the examples we have taken a hospital which wishes to review the absorption area of its patients with breast cancer. The abbreviation used for Theoretical incidence is "Theoret. incid." and for Incindence "Incid.".

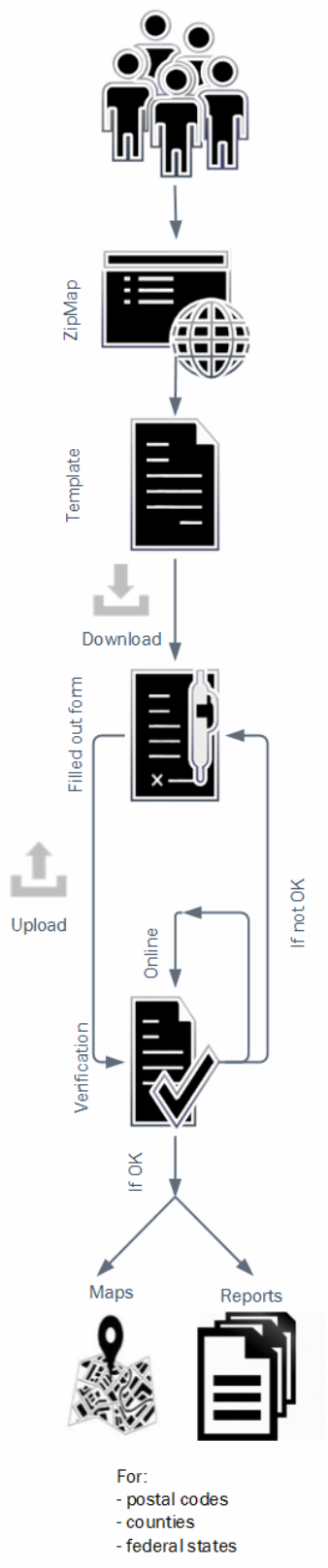


Figure 4. Application flow.

$$Theoret.\ incid. = \frac{Incid.\ in\ Germany * Population\ of\ the\ Federal\ state}{Total\ population\ of\ Germany} \quad (1)$$

Theoretically, n persons from a specific Federal State have the probability to belong to the incidence group, as in Equation (1).

For example, theoretically, n persons from a specific Federal State have the probability to have breast cancer.

$$Relative\ result = \frac{Results\ of\ the\ Federal\ state}{Theoret.\ incid.\ of\ the\ Federal\ state} * 100 \quad (2)$$

Theoretically, n percent of the persons from a specific Federal state which belong to the incidence group have the probability to be connected with the tested unit as in Equation (2).

For example, theoretically, n percent of the persons from a specific Federal State with breast cancer have the probability to be treated by the tested hospital.

$$Theoret.\ incid. = \frac{Incid.\ in\ Germany * Population\ of\ the\ County}{Total\ population\ of\ Germany} \quad (3)$$

Theoretically, n persons from a specific County have the probability to belong to the incidence group as in Equation (3).

For example, theoretically, n persons from a specific County have the probability to have breast cancer.

$$Relative\ result = \frac{Results\ of\ the\ County}{Theoret.\ incid.\ of\ the\ County} * 100 \quad (4)$$

Theoretically, n percent of the persons from a specific County which belong to the incidence group have the probability to be connected with the tested unit as in Equation (4).

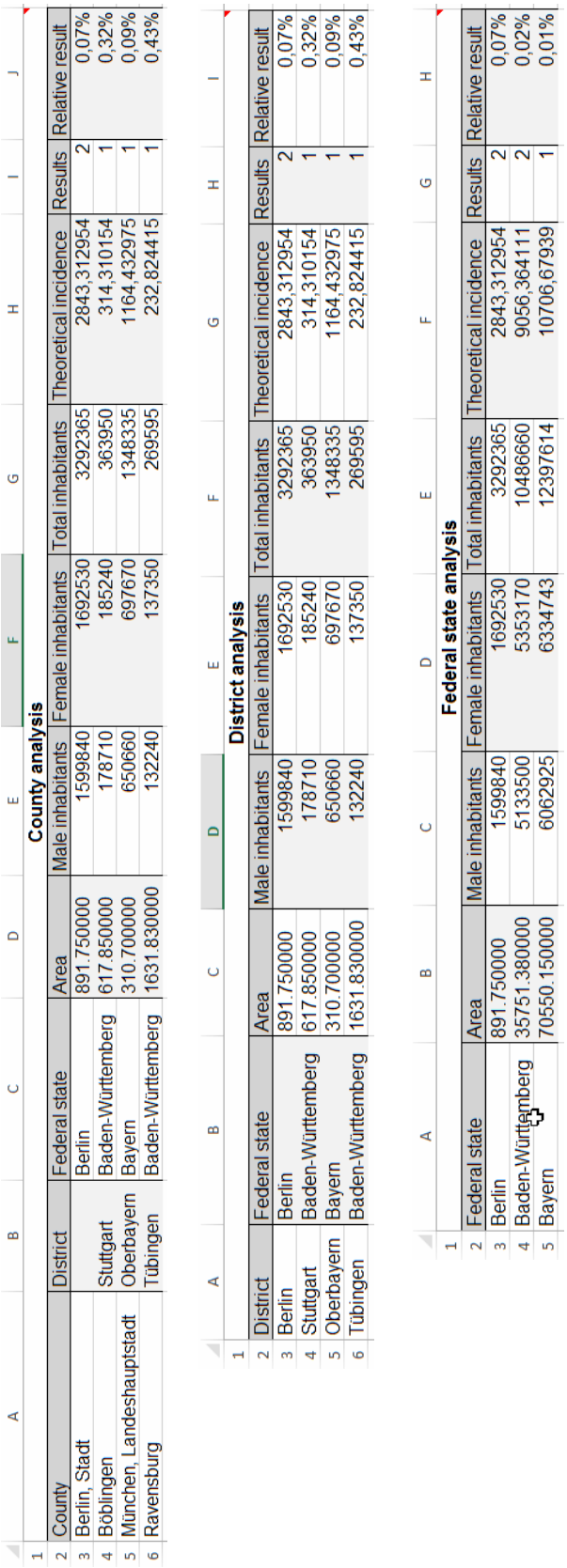
For example, theoretically, n percent of the persons from a specific County with breast cancer have the probability to be treated by the tested hospital.

The charts in this article were created using ZipMap 1.0, Microsoft PowerPoint 2013, Visio 2013 and Corel Draw x7.

III.2 ARCHITECTURE

The system is designed to be a single page web application based upon a model-view-controller (MVC) architectural pattern which with the help of front end java-script libraries and a three-tier back-end manages to present relevant data to the user in a simple but elegant fashion (Figure 7).

In this article we don't intend to detail how the MVC design pattern works [15], but the following basic concepts should be noted:



County analysis

County	District	Federal state	Area	Male inhabitants	Female inhabitants	Total inhabitants	Theoretical incidence	Results	Relative result
Berlin, Stadt		Berlin	891.750000	1599840	1692530	3292365	2843,312954	2	0,07%
Böblingen	Stuttgart	Baden-Württemberg	617.850000	178710	185240	363950	314,310154	1	0,32%
München, Landeshauptstadt	Oberbayern	Bayern	310.700000	650660	697670	1348335	1164,432975	1	0,09%
Ravensburg	Tübingen	Baden-Württemberg	1631.830000	132240	137350	269595	232,824415	1	0,43%

District analysis

District	Federal state	Area	Male inhabitants	Female inhabitants	Total inhabitants	Theoretical incidence	Results	Relative result
Berlin	Berlin	891.750000	1599840	1692530	3292365	2843,312954	2	0,07%
Stuttgart	Baden-Württemberg	617.850000	178710	185240	363950	314,310154	1	0,32%
Oberbayern	Bayern	310.700000	650660	697670	1348335	1164,432975	1	0,09%
Tübingen	Baden-Württemberg	1631.830000	132240	137350	269595	232,824415	1	0,43%

Federal state analysis

Federal state	Area	Male inhabitants	Female inhabitants	Total inhabitants	Theoretical incidence	Results	Relative result
Berlin	891.750000	1599840	1692530	3292365	2843,312954	2	0,07%
Baden-Württemberg	35751.380000	5133500	5353170	10486660	9056,364111	2	0,02%
Bayern	70550.150000	6062925	6334743	12397614	10706,67939	1	0,01%

Figure 6. Examples of analyses for County, District, Federal States

-Models: these are objects that implement the logic for the application's data domain. Many times a model object retrieves and stores the state of a model in a database. In small caliber applications usually the model is only a conceptual separation and not a physical one. If an application, for example, just reads and sends a dataset to the view, it might not have a physical model layer with classes. In such case the dataset will overtake the role of the model object.

-Views: these are the components used for displaying the user interface (UI) of the application. Normally, the UI will display the model data or part of it.

-Controllers: these components handle the interaction with users, work with the model and finally select a view for displaying the UI. The view of an MVC application only displays information, whereas the controller handles interaction with the user. We can give as an example a HTTP request which is handled by the controller and then it is passed to the model which can use this information to query the database.

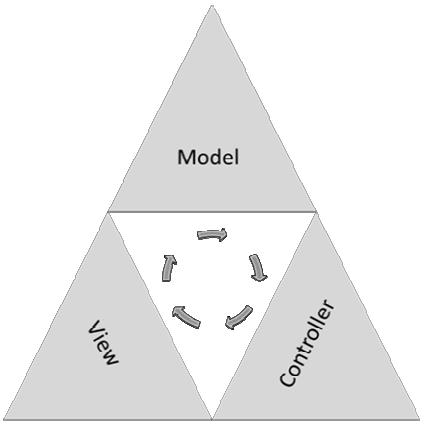


Figure 7. Model-view-controller architecture.

An MVC design enables the developer to separate an application in different layers (input, business, UI), but at the same time, provides a loose coupling between them. Through this pattern the application complexity can be easily managed while building an application. Developers can focus on one aspect of the implementation at a time, this assuring a higher precision and success.

III.3 SERVER SIDE

The back-end of the application is based on the widely used model-view-controller pattern on top of the ASP .Net MVC framework from Microsoft were a typical three-tier layer system is present. The base layer is the Data Access Layer (DAL for short) where the system manages data manipulation by connecting to an MS SQL server and over an ORM fetching and saving the business model entities. Our object relational mapper of choice was Entity Framework 5 from Microsoft (Figure 8).

Over the data access layer there is a business layer responsible for all the calculations, computation and data processing that is necessary for the transformation of the raw data into meaningful information which the user expects. This layer is responsible for deserializing the provided data from the uploaded excel file and for creating the response excel files which contain the analyzed data. For the excel file generation a third party library is used (EPPlus). The final layer also coincides with the controller part of the MVC pattern, the endpoint of the server side where inside the controller well-structured actions handle the request from the front-end side of the application. The actions besides the typical ActionResult response of this type of architecture also return JsonResult in order to efficiently communicate with HTTP web request posted by the front-end javascript services.

Even so the application is meant to be publically accessible, because it is based on the ASP .Net MVC framework that can be easily extended to contain an access control system that can differentiate between users and implement basic security with authentication and authorization if needed.

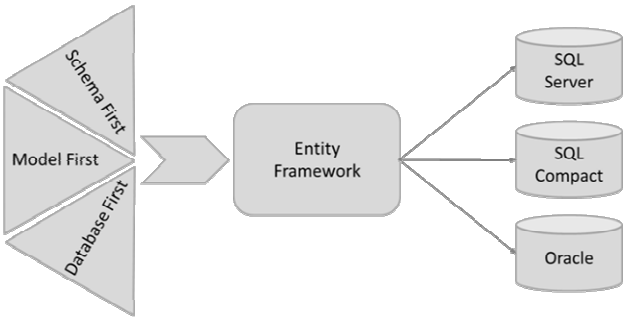


Figure 8. Microsoft Entity Framework 5 flow.

III.4 CLIENT SIDE

The front-end or client side of the application is basically a single-page application where the user will get all the requested information without leaving the original page. This is realized with the help of Document Object Model (DOM) manipulating java-script libraries such as KnockoutJS and jQuery. The basic concept of Single Page Application (SPA) is to only load a single web page with a single URL and then render HTML elements based on the user interaction. This has the advantage of reducing traffic because the common parts of the page do not need to be downloaded again and also has a positive impact on user experience by making the page more responsive and with a natural desktop-application feeling for it. Because of the nature of the application, the front end is also responsible for presenting geographical information to the user. These are information like center locations on the map or occurrence density in certain regions of certain presets of the map. For these presets GeoJson and TopoJson data is used. The basic map-related functionalities are handled with OpenLayer framework which specializes in map-related data display.

After the user downloaded and completed the template with the required zip code, during the upload of the file, it gets validated for any invalid zip codes. In this validation process an internal database with thousands of zip codes from the target area is used. In case of the invalid zip codes the user will need either to delete them or correct them before proceeding to the next step. During the data processing for each submitted zip code the geographical

coordinates are loaded into the system. Based on the selected view (Postal codes, Counties, Districts, Federal states) the system will group the input data into the required categories and create statistical information for each affected zone. The statistical information will also contain the total number of centers for each subzone of the selected view. This aggregated information is then transferred to the front-end where, with the help of the mapping javascript libraries, layers will be created both for the selected view and for the relevant information for that view. These layers will be stacked upon each other to create the graphical map representation of the data. The front-end is also responsible for creating a color scale in order to display different occurrence densities with a color marker. The more verbose form of the pre-calculated data can also be downloaded from the application using the export options where it is downloaded in excel files.

IV. CONCLUSION AND FUTURE WORK

An easy and interactive user interface can assure the success of an application. The programmers must pay attention to the clients' needs in order to find the perfect solutions for any problems encountered in the development of a map generator application [16]. An important aspect for a successful and widely used application is the possibility to generate personalized maps with users' preferences. [17]

As further aims for the program we would like to mention the idea of extending the coverage area to other countries too. This would ensure an international coverage. Another aim would be the upload of files also with coordinates instead of postal codes. This would ensure more accurate results with a wide range of possibilities for those who wish to use coordinates. Also, we would like to implement the possibility to download the maps, generate heat maps, customize the colors of the nodes, extend the supported languages, improve the user friendliness and layout, improve the web responsiveness for mobile devices, create also a desktop application, extend the analysis tools, improve zoom quality and generally improve image quality. Even though there are similar products available on the internet, these rarely provide the possibility to localize many points of interest on the map at the same time and do not have the possibility to generate different statistics regarding the density of any kind of elements on a given area.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest

REFERENCES

[1] https://en.wikipedia.org/wiki/Point_of_interest [Accessed: November 28, 2015].
[2] https://en.wikipedia.org/wiki/Map_database_management [Accessed: November 28, 2015].
[3] G. Miscione, "How come Geographic Information Systems are peripheral in Information Systems research? Problems and Prospects", *Seminar at the Department of Information Systems, University of Münster*, 2015.
[4] W.T. Pan, "Using data mining for service satisfaction performance analysis for mainland tourists in Taiwan", *Int. J. Technology Management*, vol. 64 (1), pp. 31-44, 2014.
[5] http://www.sciencedirect.com/science/article/pii/S095741741500069X [Accessed: November 28, 2015].
[6] J. I. Blanford, Z. Huang, A. Savelyev, A. M. MacEachren, "Geo-Located Tweets. Enhancing Mobility Maps and Capturing Cross-Border Movement", *PLoS One*, online, 2015.
[7] S. Wang, Y. Liu, A. Padmanabhan, "Open cyberGIS software for geospatial research and education in the big data era," *SoftwareX*, online, 2015.
[8] https://en.wikipedia.org/wiki/Comparison_of_web_map_services [Accessed: November 28, 2015].
[9] https://maps.google.com/ [Accessed: November 28, 2015].
[10] https://www.openstreetmap.org [Accessed: November 28, 2015].
[11] https://en.wikipedia.org/wiki/Microsoft_MapPoint [Accessed: November 28, 2015].
[12] www.zipmap.de [Accessed: November 28, 2015].
[13] D. Probert, N. Shehabuddeen, "Technology road mapping: the issues of managing technology change", *International Journal of Technology Management*, vol. 17 (6), pp. 646-661, 1999.
[14] www.studybox.de [Accessed: November 28, 2015].
[15] https://msdn.microsoft.com/en-us/library/dd381412%28v=vs.108%29.aspx?f=255&MSPPError=-2147217396 [Accessed: November 28, 2015].
[16] T.R. Baker, R.T. Palmer, "Coding Geo", *Techedge*, vol. 2, onlilne, 2015.
[17] D. Wilson, J. Doyle, J. Weakliam, M. Bertolotto, D. Lynch, "Personalised maps in multimodal mobile GIS", *International Journal of Web Engineering and Technology*, vol. 3 (2), pp. 196-216, 2007.