INFRASTRUKTURA I EKOLOGIA TERENÓW WIEJSKICH
INFRASTRUCTURE AND ECOLOGY OF RURAL AREAS

Nr IV/4/2016, POLSKA AKADEMIA NAUK, Oddział w Krakowie, s. 1833–1845
Komisja Technicznej Infrastruktury Wsi

DOI: http://dx.medra.org/10.14597/infraeco.2016.4.4.137

INTEGRATED DATABASE AS THE BASIS FOR THE MODERN REAL ESTATE CADASTRE

Małgorzata Buśko
AGH University of Science and Technology

Abstract

A comprehensive IT system to run the National Cartographic Documentation Center database (PZGiK) is currently being implemented for the city of Krakow. Part of this implementation involves the conversion of the data contained in the previous system, a lossless migration of the data contained in the previously existing analog systems, as well as adjusting the data as required by the current legal provisions.

The subject of this study is to analyze the progress of this implementation in terms of building an integrated relational database and the specificity of the objects it is composed of. The analysis gave grounds to present the problems of data harmonization, associated with frequent changes in the legislation in the field of geodesy and cartography.

The paper demonstrates the problems which occur when implementing the new system combined with the conversion of the National Cartographic Documentation Center database, which require solutions in the form of strategic decision making, determining the system of work of the surveying administration in Krakow over the next years.

Keywords: database, data conversion, data migration, generating a map

INTRODUCTION

The city of Krakow is the capital of the Malopolska province, has the role of the township, and includes 4 cadastral units, which are currently districts of the city. The total number of the cadastral districts located in the cadastral units is 253. Until 2015, the main IT system to run the National Cartographic Documentation Center database was the V-SYSTEM. Additionally, individual systems, such as ENIER G and Index of plots, functioned as well. For the reasons beyond the control of the Department of Geodesy of the City of Krakow, an urgent need emerged to convert to the new system. This project has been carried out since March 2015. It aims to replace these IT and analog systems by the modern Integrated System of Resource Management. There are numerous, very interesting studies, which present the transformations of systems to run surveying databases and cadastre of real estate in different countries (Bydłosz, 2015, Bennett, Alemie, 2016, Lisec, Navratil, 2014, Zhang *et al* 2016) and their use as a database (Brown, Raymond, 2014).

A computer program which forms the basis to build the information system for the current implementation is the Geographical Information System GEO-INFO by Systherm from Poznan. However, due to the technical terms specified by the client, i.e. the City of Krakow, it is completely remodeled and adapted to the specificity of the Krakow region.

The paper presents the concept and methodology of building an integrated database in this system, with particular emphasis on the real estate cadastre data, in terms of integration and harmonization of the existing databases. It also presents the problems and dilemmas that arise during data migration, as well as the necessary decisions to be taken during the implementation of the new system. By order of the mayor of Krakow of 15 September 2015, the author of this paper acts as a consultant to the merits of the project *"Delivery and implementation of a comprehensive information system to run the District Cartographic Documentation Center database in Krakow, together with computer infrastructure, migration and adaptation of the data from the District Cartographic Documentation Center database – the Integrated System of Resource Management"* (https://przejrzystykrakow.pl), which allowed to develop and propose certain solutions.

DESCRIPTION OF THE ENTITY CARRYING OUT THE IMPLEMENTATION

The databases existing in Krakow so far were completely separate; only if put together they provided full information about the National Cartographic Documentation Center database of Krakow. The database did not include the object *Municipality/City* for Krakow. There were four *Cadastral units*, representing

districts of Krakow: Krowodrza, Podgórze, Śródmieście, Nowa Huta. Moreover, the objects of the real estate cadastre register were separate geometric and descriptive entities, recorded in separate databases: the geometric part – V MAP, and the descriptive part – VEGA. The connection between these two databases in relation to the basic objects of the real estate cadastre, or for the cadastral parcel and for the building, was based on the identifiers of the National Register of Territorial Division of the Country (TERYT). Such a situation is typical of the National Cartographic Documentation Center (PZGiK) databases in Poland, and therefore the discussed implementation can be regarded as a pioneering one. After the quantitative analysis of the PZGiK database before the conversion, in terms of the main objects of the real estate cadastre, some irregularities were identified, as illustrated in Tables 1, 2, 3.

Table 1. Non-compliance in the database before the conversion – cadastral parcels

Number of the cadastral parcels disclosed		
in the geometric part	in the descriptive part	with the corresponding geometric and descriptive parts
184 112	184 119	184 099

Source: Own study

Table 2. Non-compliance in the database before the conversion – buildings and sheds

Number of the buildings and sheds disclosed			
in the geometric part (no descriptive part and no ID)	in the geometric part and with ID (no descriptive part)	in the descriptive part and with ID	with ID and the corresponding geometric and descriptive parts
121 885	121 505	16 629	16 365

Source: Own study

Table 3. Non-compliance in the database before the conversion – record entities

Total number of record entities		
Records record entity	Entities without personal identity number PESEL	Entities appearing repeatedly with the same personal identity number PESEL
726 424	123 896	210 277

Source: Own study

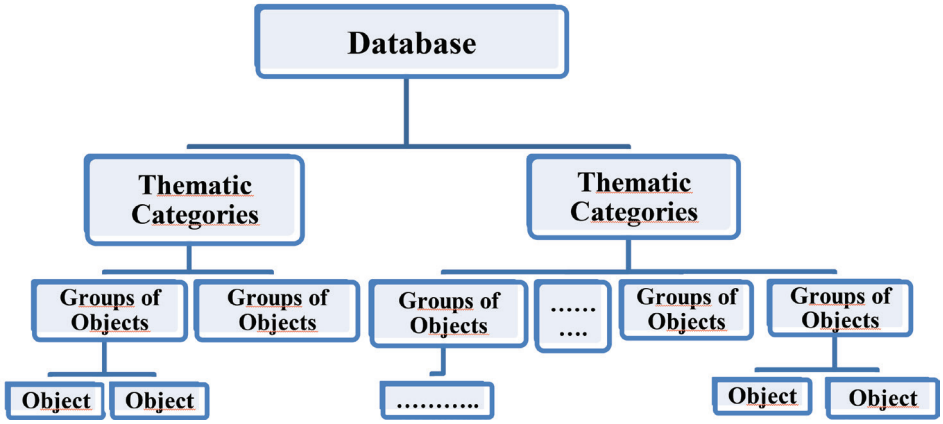
From the data contained in Tables 1 – 3, it appears that before the conversion the overall consistency did not exist in any group of the objects in the geometric and descriptive parts of the real estate cadastre. The best situation in this respect was for cadastral parcels – more than 99% of the parcels were disclosed in both parts (graphic and descriptive). The non-compliance for the cadastral parcels generally resulted from the errors of the operators of the previous system. In the case of cadastral parcels, such erroneous objects were very few, so they were analyzed individually.

The situation for the buildings and sheds appears to be much worse. Only slightly more than 13% of them were properly entered in both parts of the database before the conversion. Most of the buildings appeared only in the graphic part of the database. With regard to the buildings during the conversion, there are two solutions – one, which is more favorable for the implementation executor, the second one – for the client. They will be discussed in detail later in this study, with the description of database objects.

As far as record entities are concerned, the situation was also unfavorable. The basic criterion for distinguishing a natural person is the personal identity number PESEL. However, in the database, almost 29% of the entities occurred in different records with the same PESEL number, but with other attributes being different (for example: different addresses, different ID card numbers, etc.). In the case of record entities, a decision had to be made on merging the entities occurring many times in the database, if they had the same PESEL number. The criteria for merging and the hierarchy of their application were defined.

DATABASE AND CONVERSION

As part of the implementation, the executor provided a software license to use Oracle Database Standard Edition for dual-processor database server, which will be installed and configured on the production database server. For data safety reasons, a database replica server was also predicted, designed to exchange data within the application of the electronic service center (ECO), which is the portal to support surveying units, bailiffs, real estate appraisers, sector units and clients. The database after the final conversion must be registered with the President of Krakow. It will receive the database identifier and will be given the name. At the testing stage associated with successive conversion iterations, the database of the Department of Geodesy of the City of Krakow contained about 12 million objects. For the reasons related to the activities of the operators associated with editing of the maps and saving them to the database, the size of the database may increase up to four times. The database contains objects arranged in a certain order, related to the hierarchical division of the database (Figure 1).



Source: Own study

Figure 1. Database division in SIP GI

Table 4 summarizes the modules forming the previous IT system and the components of the new system based on GEO-INFO, indicating the areas of functionality supported by individual modules, before and after the implementation of the Integrated System of Resource Management in Krakow.

Thematic categories of databases, illustrated in Figure 1, result from the division into the records referred to in the regulations; they are also presented in Table 4 as components of the new system. A letter designation of a category is important in terms of creating object codes – determines the second letter of the GI code. Using additional division criterion, as part of a large-scale digital map, the GEO-INFO distinguished several types of objects, including: surface, linear, point, information, text. The presented division into these types has a significant effect on constructing codes of objects in GEO-INFO, as it marks the third letter of the code.

Groups of database objects are thematic layers of the numerical resource, contained within each category. The fourth letter of the GI code originates from one of the letters denoting the name of the group of objects. The objects are collected as part of the thematic groups of objects, as the most detailed elements of the database hierarchy.

In the Geographical Information System GEO-INFO, each object is assigned several types of codes, including the database-code, the code from the Regulation (Regulation, 2015), but for the operator, the most important of them is the so-called GI code, because this code is used it in everyday work with the system. The principle of creating GI codes is closely related to the hierarchical form of the database and the object belonging to a particular category, type and

group. As a result of such a way of assigning codes to the objects in GI, they are very intuitively formulated, which is of great importance to the operator working in the Polish language because they are easy to assimilate and remember. Also, attention should be paid to the GI code in the Geographical Information System for a parcel not depicted in the real estate cadastre (GESDZN). Such a code was anticipated due to the development of the Land Parcel Identification System (LPIS) which, although it does not meet the accuracy criteria, exists in some District Cartographic Documentation Centers. LPIS studies play an auxiliary role, and that is why, after the database conversion, these centers want to keep them.

Table 4. Changes in the scope of the functionality of system modules after the implementation of the Integrated System of Resource Management

Existing modules	Functionality	New components
V-Map (Graphics)	Running and supporting the digital base map and the geometric part of the geodetic network of utilities (GESUT)	**GEO-INFO Map:** **– BDOT500** **– GESUT** **– BDSOG** **– EGiB** **– RCiWN** **GEO-INFO i.Adres**
VEGA	Keeping records of the register of land and buildings (EGiB) and the Dictionary of streets (the program did not meet the requirements of the Regulation of the Minister of Administration and Digitization of 2 February 2012 on the records of the towns, streets and addresses)	
Index of plots	Reservation of parcel designations, restoration of the genealogy of parcels	
ENIER_G	Database from the previously used EGiB system (OTAGO-ENIER G) of a historic nature	
V-Biuro	Management of the Cartographic Documentation Center database	**GEO-INFO Ośrodek**
V-Dok	Management of the document circulation at the level of the Department of Geodesy	
V-WEB	Services rendered to surveyors over the Internet	**GEO-INFO i.Kerg** **GEO-INFO i.Wniosek**
RCiWN	Services rendered for sector units and customers Register of Real Estate Prices and Values	**GEO-INFO i.Rzeczoznawca** **GEO-INFO i.Komornik**
	Management of databases, operators (users), installation and updating of the metadata of objects and automatic generation of safety backup copy	**GEO-INFO Admin**

(author: Małgorzata Buśko)

In the regulation, the legislator did not provide a code for a building which exists only in the graphic part on the base map, and it is not disclosed in the descriptive part of the existing land and buildings register. In Krakow, there are more than 80% of such buildings (Table 2). In the proposed solution, the conversion executor added a specific object code to the database, which was intended for the buildings undisclosed in the descriptive part, and existing only in the graphic part (GESBNE), not provided for in the applicable regulations. As a result, buildings with that code could be moved to the integrated database, but would still be just "graphic objects". For the conversion executor it is a very convenient solution, as it allows to fully automate the conversion process. However, in consequence, at the stage of these buildings existing in the integrated database as non-record ones, no elements permanently attached to building structures, such as stairs, terraces or driveways to the underground part, could be attached to them. One should also be aware that, as a result of the absence of a code in the regulation, such non-record buildings (GESBNE) could not be transferred with the existing data exchange format GML, and therefore, could not be exchanged in this format between the executor and the National Cartographic Documentation Center, or be entered into the central repository of the Integrated Information System for Real Estate (ZSIN) at the time of its launch. According to the recommendations of the implementation executor, such buildings should be changed to register buildings (GESBZO) as soon as possible, as a result of the modernization of the cadastre or unit surveying studies. Thus, all the work and the related costs would be incurred by the client commissioning the implementation. Having analyzed this idea and the resulting threats, no consent was given to such a solution proposed by the implementation executor. Finally, the executor shall assign identifiers to all the buildings undisclosed in the real estate cadastre, create files for them, entering only those attributes which were known in the previous system based on the base map (order number, number of floors, function of the building). Thus, these buildings were entered into the database of the real estate cadastre, and during the next surveying works, their attributes will be complemented or modified. The role of a proper formulation of the specification of essential contract terms should be emphasized at this point, so that the client was able to make decisions at the appropriate stage of the implementation.

TYPES OF POINT OBJECTS IN THE GEOGRAPHIC INFORMATION SYSTEM GEO-INFO

In GI there are two types of point objects, which create the necessity to make arrangements during the conversion of the database, regarding the uniformity in their application when entering objects into the database:

- full-information point objects – these are database objects which have the code, attributes, X, Y coordinates (e.g. a boundary point of a plot),
- the so-called XY points – these are not database objects, they do not have the code, and they can be other record points (e.g. a boundary turn point) or topographic points (e.g. a slope turn point).

Entering buildings into the database is also possible in two ways – basing them on the nodes which are full-information points (representing database objects) or XY points. Both procedures have their advantages and disadvantages. The disadvantage of entering all the points as full-information points is more time-consuming for the operator, and also causes considerable expanding of the size of the database. The advantage of entering points as full-information ones is that their ID or KERG numbers, associations and relationships with other objects, as well as any attached files, are retained in the database. XY points lack such data. Due to the significantly increased time consumption associated with entering all the points as full-information ones, in Krakow it was decided to allow for the use of XY points as turn points of buildings or boundaries of record parcels.

THE MAP AS A REPORT FROM THE DATABASE

In accordance with Regulation (Regulation, 2015), the map in the Geographic Information System GI is a generated database content. In GI, there are two possibilities to configure the map, based on which the work will be carried out: the AUTO-GENERATED map – automatically by the GI system and the PRE-GENERATED map – its character is greatly determined by the system operator. When establishing the auto-generated map, as part of its configuration, the operator selects active layers to be automatically generated by the GI. Therefore, the responsibility of the operator is only to choose the database layers to be generated, whereas the selection of the objects within a given layer is automatically performed by the GI and always refers to 100% of the objects contained in the database as part of the corresponding layer. On the auto-generated map, at a great distance, the cadastral unit can only be seen as a point. As approached, the system automatically generates other smaller objects that appear on the screen. According to the hierarchical order, as the surface of the earth is being approached, the following objects appear on the map in order: cadastral unit, cadastral district, record parcels, buildings, utilities. The auto-generated map always displays the area on the screen in real time through its automatic reloading, but the reloading can also be performed on demand by using an icon. Basically, the work with the auto-generated map is faster and more convenient than with the pre-generated map. However, there are also disadvantages of this kind of map, for example, there is no possibility to filter the map content with the choice made for a single

attribute. This is one of the reasons for having to work on the pre-generated map if the operator's tasks are precisely associated with that kind of work. (Akińcza *et al*, 2015, Balawejder *et al*, 2015, Bieda *et al*, 2015).

When establishing the pre-generated map, in its configuration the operator determines which layers are to be active for generating. However, within the layers, its content can be generated with an accuracy of a single, operator-selected attribute of an object through the use of appropriate filters and queries to the database. On the pre-generated map, the operator determines the area of the map to generate from the database. Outside this area no objects will appear, despite the fact that they exist in the database. There are many other useful work applications on the pre-generated map, for example, the historical map content can be generated only from the database of this map – of any date back. Thus, different versions of a given object can be illustrated on it. The pre-generated map does not reload automatically, it is reloaded at the request of the operator, through their conscious action. A different functionality of the option should be emphasized at this point: *Generate map* and *Reload map*.

In the case of the option: *Generate map,* the GI program checks whether the objects which have been selected for generating are already on the map. If the GI program finds the object, it stops to perform additional activities and does not interfere with the object. In the case of the option: *Reload map,* the GI program also checks whether the object is on the map. However, even if the object is found, GI deletes it from the map content and generates it again from the database, in accordance with its current geometry, attributes and edition stored in the database. In conclusion – after reloading we are always dealing with the current state of the database.

INTEGRATION OF DATABASE RESOURCES DURING THE CONVERSION

The existing databases functioning in Poland before the conversion to modern IT systems generally have a high number of errors, inaccurate entries, duplicated information, and lack of the internal harmonization of the data. This is due to the coexistence of multiple non-integrated databases, where information was being recorded for many years. Before the conversion, it is necessary to specify which database will act as the reference base in a given area, and to design the procedure for automatic updating of individual databases after the conversion. Technical and legal aspects of the integration of databases and the proposed solutions will be presented on the example of the database of record entities.

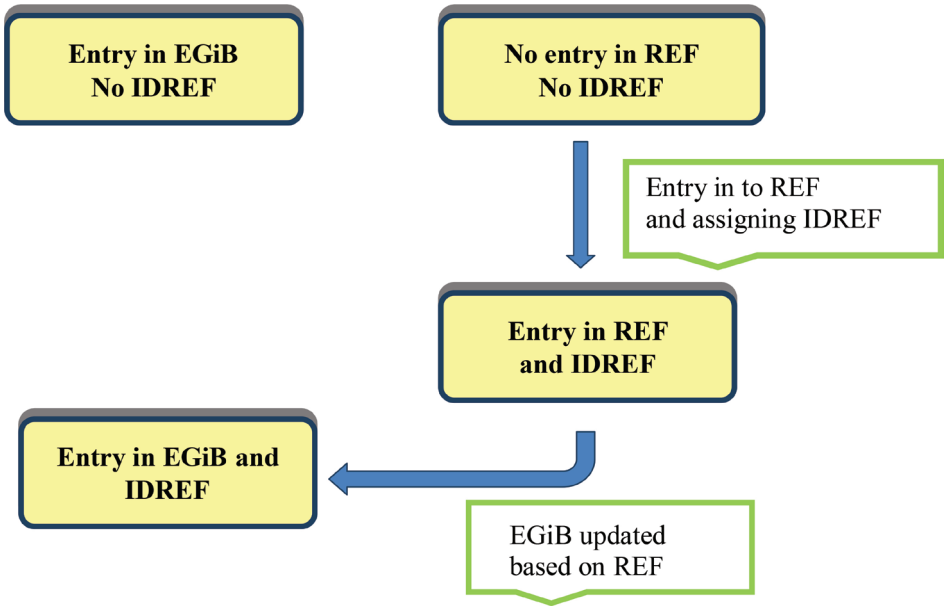


Figure 2. Entering a new person to both databases, starting with REF
(author: Małgorzata Buśko)

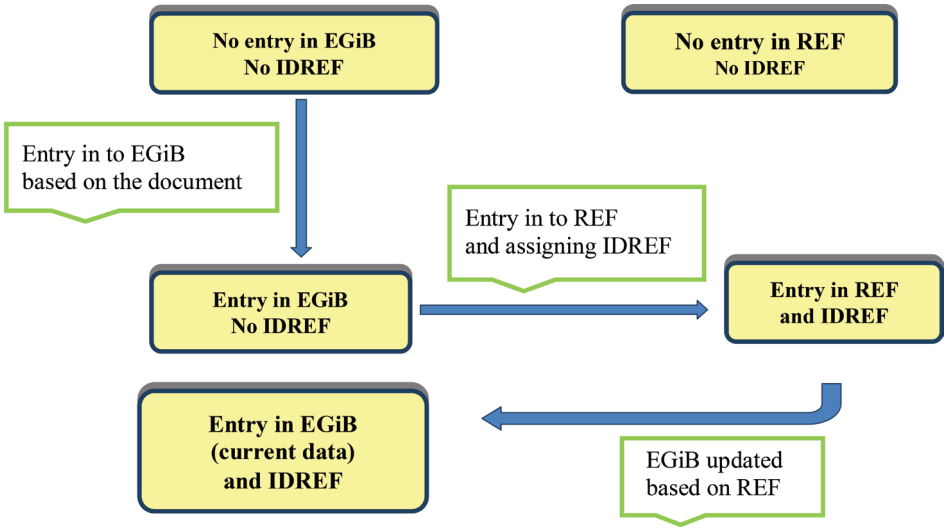


Figure 3. Entering a new person to both databases, starting with EGiB
(author: Małgorzata Buśko)

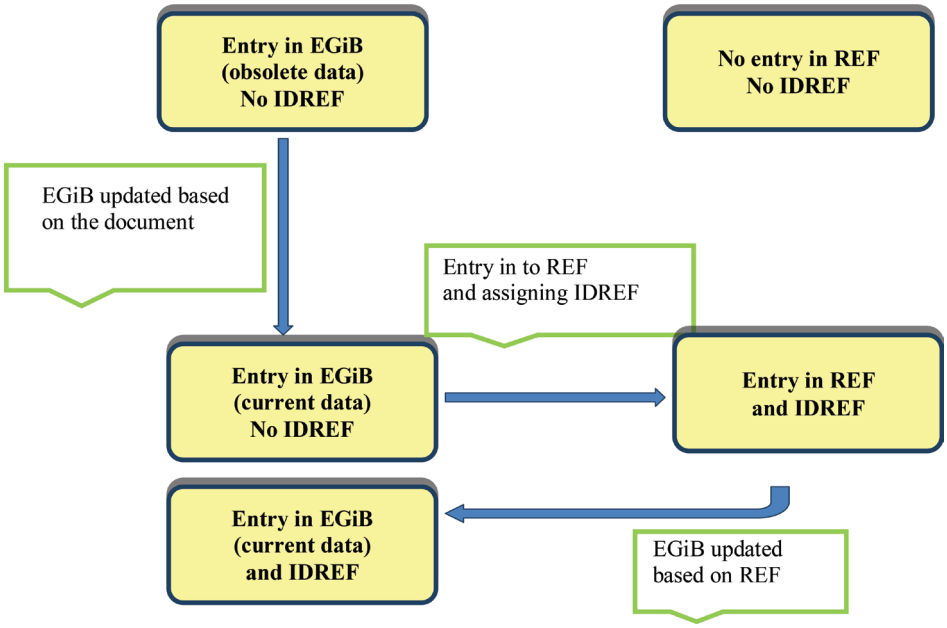


Figure 4. Updating the EGiB database, adding to REF (author: Małgorzata Buśko)

Regarding personal databases, Krakow had the so-called Personal File. However, it was not integrated with the database of the real estate cadastre (EGiB) in the scope of record entities. So far, the reference database for the (EGiB) database has been the personal identity number PESEL. However, one of the objectives of the integration, referred to in the terms of the carried out implementation, is to make Personal File the reference database for the EGiB database, which is why further in this article, it will be referred to as REF. For the implementation purposes, a mechanism of mutual updating and integration of the two databases: REF and EGiB has been proposed.

Unfortunately, the REF database is not fully updated, and also before the implementation it was not integrated at all with the database of the real estate cadastre (EGiB). There were numerous irregularities in the REF database itself, related both to the total lack of persons, their incomplete data, and the data which were obsolete. In this database, a person is assigned an identifier (IDREF), which should also be stored in the register of land and buildings (EGiB) database, to identify the person. The author has developed several models of EGiB update regarding persons, which may be used to complement, update and integrate the two databases – EGiB and REF in a variety of cases of the initial situation (Figures 4, 5, 6). The assumption of the models is that the operator in the EGiB database can change the relationship of the person found in EGiB with the reference

database (REF) by sending current data to the REF database, which may result in assigning or changing IDREF of a given person. It will be necessary to clean the EGiB database from the persons repeatedly registered due to inaccurate data, or due to lack thereof. The positive effect will also include a gradual organization of the REF database. A common mistake with regard to the persons existing in the Krakow databases is the lack of the personal identity number PESEL – before the conversion it was the case concerning nearly 50% of the total number of the people entered into the EGiB database.

SUMMARY

Conversion of the database is an extremely difficult and labor-consuming undertaking, especially in the context of such a large amount of surveying and mapping data as in the resource of the National Cartographic Documentation Center database in Krakow. An additional complication is the fact that so far the data have been collected in different registers and records, characterized by varying time of their origin, and consequently – different level of computerization. The essence of the currently conducted conversion is lossless data migration, retaining their relationship and topological coherence. It must apply not only to purely surveying databases, but also to the related databases of the city of Krakow (e.g. Personal Files). It is impossible to fully perform the quality control of the conversion in the numerical manner, and therefore, some of its components, such as the behavior of the history of objects will be verified when working with a new, integrated database. Undoubtedly, however, it is the integrated database which in the future will form the basis for running a modern cartographic documentation center, with a modern real estate cadastre as its part.

ACKNOWLEDGMENT

This research was financed by the means of Statutory Research no. 11.11.150.006

REFERENCES

Akińcza, M., Bieda, A., Buśko, M., Hannibal, H., Hanus, P., Hycner, R., Krzyżek, R., Kwartnik-Pruc, A., Łuczyński, R., Przewięźlikowska, A. (2015). Current problems of the cadastre in Poland. Monograph, Publishing House of Warsaw University of Technology.

Balawejder, M., Buśko, M., Cellmer, R., Juchniewicz-Piotrowska, K., Leń, P., Mika, M., Szczepankowska, K., Wójciak, E., Wójcik-Leń, J., □róbek, S. (2015). Current problems of the real estate management in Poland in terms of the organizational and legal changes. Monograph, WSI-E, Rzeszow.

Bennett, R. M.; Alemie, B. K. (2016). Fit-for-purpose land administration: lesson from urban and rural Ethiopia. Survey Review, Vol. 47, Issue 346, pp. 11-20.

Bieda, A., Bydłosz, J., Dawid, L., Dawidowicz, A., Glanowska, M., Góźdź, K., Przewięźlikowska, A., Stupen, M., Taratula, R., Zróbek, R. (2015). Directions in the development of the real estate cadastre. School of Engineering and Economics, Rzeszow.

Brown, G., Raymond, C. M. (2014). Methods for identifying land use conflict potential using participatory mapping. Landscape and Urban Planning, 122, 196-208.

Bydłosz, J. (2015). The application of the Land Administration Domain Model in building a country profile for the Polish cadastre. Land Use Policy, Vol. 49, pp. 598-605.

Lisec, A., Navratil, G. (2014). The Austrian Land Cadastre: From the Earliest Beginnings to the Modern Land Information System. Geodetski Vestnik, Vol. 58, Issue 3, pp. 482-516.

Regulation (2015) Regulation of the Minister of Administration and Digitization of 2 November 2015 on the database of topographic objects and the base map, Official Journal of 2015, item 2028.

Zhang, JY, Yin, PC, Li, G, Gu, HH, Zhao, H, Fu, JC. (2016). 3D Cadastral Data Model Based on Conformal Geometry Algebra. ISPRS International Journal of Geo-Information, 2016, 5(2):20.

https://przejrzystykrakow.pl/zarzadzenia/36579, 28.09.2015

Dr inż. Małgorzata Buśko
AGH University of Science and Technology
Faculty of Mining Surveying and Environmental Engineering
Department of Geomatics
Al. Mickiewicza 30
30-059 Kraków
mbusko@agh.edu.pl
+48 608 017 172

Received: 12.05.2016
Accepted: 12.12.2016