Data Integration within Public Property Management Systems

Elzbieta BIELECKA, Poland

Key words: Geoinformation, standards, data integration, property

SUMMARY

Public property management requires systemic approaches, which enables capture the relevant data about the properties and their surroundings as well as develop specialized applications to data updating, checking, and analysis. Currently, data about properties is stored in digital form and managed by the various administrative bodies. It usage in property management applications requires solving a number of problems relating to data integration. The article describes those public registers that store information about the property, conducted from the perspective of data collecting for property management systems, shows the stages of data integration, and indicates problems to solve.

Data Integration within Public Property Management Systems

Elzbieta BIELECKA, Poland

1. INTRODUCTION

An information system, including also a property management system, is usually dedicated to specific users, as there are no universal systems that would meet the requirements of many different user groups. The implementation of such a system requires collecting pertinent data and purchasing specialised software and hardware. The data should be kept up to date and it should comprehensively describe the objects and scenarios that are relevant to the user. The software should allow for spatial and temporal analysis of data as well as for visualising data in the form of reports, summaries, charts and thematic maps. The hardware should quickly process the collected data and ensure operational reliability.
System for public property management should contain a number of data about property originally stored in many source public registers. It usage in property management applications requires solving a number of problems relating to data harmonisation and integration.

The issue of data integration is not new – it came into being as early as in the initial period of IT systems development. However, it was only in the recent years that, following the promotion of information society, e-administration and digital transformation, methods of solving different problems related to the data, systems or services integration have being debated as part of international initiatives. Since the beginning of the 2000s, a range of activities have been undertaken with the aim to develop standards that would facilitate data exchange between heterogeneous database systems. For instance, the OGC (Open Geospatial Consortium), W3C (World Wide Web Consortium), International Organisation for Standardisation (ISO) and OMG (Object Management Group) carry out such activities, as well as the government of many countries in order to develop spatial data infrastructure. Systems developed according to such standards and specifications are interoperable, which means that they allow for exchanging data and information between heretogenius database systems. Such activities are also carried out in Poland, as part of spatial information infrastructure that is being developed. The results of such actions will, however, be only visible in a few years.

2. PUBLIC REGISTERS CONTAINING DATA ABOUT PROPERTIES IN POLAND

2.1. The scope of data

In Poland, data about property is stored and managed by surveying and mapping administration, the Ministry of Justice, local authorities, and property managers and administrators. Its scope is regulating by acts and governmental regulations, e.g. Geodetic and Cartographic Law (Journal of Laws of 2010 No 76, entry 1287), Land and Mortgage Registers and Mortgage Act of 6 July 1982 (Journal of Laws of 2001 No. 124, entry 1361).

A property (parcel, building and premises) is a spatial object (ISO/FDIS 19152: 2008) so it is necessary to record its location. Location of a parcel and a building is determined by coordinates of the boundaries of a parcel or a building outline in a given coordinate reference system. Location of a premises is determined indirectly, i.e. by reference to another spatial object, i.e. a building. Data on property location as well as some atributte characteristics of a property (area, owners, technical standard) is stored in lands and buildings register maintained by district administration (starost). Rights to a property and information about owners – in land and mortage register maintained by district court; taxation – in register of real estate prices and values; intended use – local development plan. The list of public sourse resourses storing data about property is listed in Table 1.

Table 1. Sources of data about property

Name of a public register	Scope of data	Responsible entity
Land and mortage register		Court
Lands and buildings register	Location and general characteristic of a property	Starost (city mayor)
Database of Topographic Objects	Data concerning a building's location and address	Voivodeship marshal
Register of property prices and values	Data about real property value	Starost (city mayor)
State Register of Borders and Areas of Territorial Division Units	Administrative units' borders and names	Surveyor General of Poland
Aerial or satellite orthoimages	Image of a real property item	Surveyor General of Poland, voivodeship marshal, starost

Apart from the data sources listed in Tables 1, data concerning costs of, outlays on, technical characteristics of a property also play a very important role in property management systems. of high importance are also various documents (e.g. scanned copies of agreements, results of inspections and audits, technical documentation) collected by a person who holds a property or an entity that manages a property on the person's behalf.

2.2. Land and mortage register

A land and mortgage register is the most important document containing information about the legal status of a property. The registers hold data such as property designation, rights attached to a property, designation of its owner or perpetual lessee as well as entries concerning limited property rights or restrictions on disposal of a property. Article 354.4 of the Land and Mortgage Registers and Mortgage Act of 6 July 1982 (Journal of Laws of 2001 No. 124, entry 1361) provides for the possibility of direct access to the Central Land and Mortgage Register Database by all bodies that keep land and property registers. This provision guarantees the compatibility of property identyfication number in two main property registers: land and mortgage register and lands and buildings register.

2.3. Lands and buildings register

The lands and buildongs register (named also as cadastre) is kept in a decentralised way, by 454 administration bodies, including: 378 starosts and city county presidents as well as 76 commune heads (mayors or city mayors). A detailed scope of information covered by the register, the structure of database as well as rules of exchanging data are specified in the Regulation of the Minister of Regional Development and Construction on land and property register dated 29 March 2001 (Journal of Laws of 2001 No. 30, entry 454). The quality of recorded data is varied. Data about lands (parcels) is recorded in digital form for the whole country, whereas data on buildings and premises is incomplete and stored both in digital and analogue form. Only in 22% of districts (Pakuła-Kwiecińska, 2007) the geometrical and descriptive databases are integrating, as recommended by Cadastre 2014 (Kaufmann and Steudler, 1998). Data exchange between other system is possible due to SWDE format. For many information systems it has the primary and reference nature (Pietrzak et al, 2012).

2.4. Register of property prices and values

The register of property prices that are specified in notarial deeds and values that are established by property appraisers in property valuation reports, according to the regulation on the property and land register (Journal of Laws No. 30, item 454), is kept by the district government (starost). The register comprises the prices and values as well as the addresses of a property, register numbers of land plots that are part of the property, the type of property, area of land property, date of signing a notarial deed or drawing up a property valuation report, and other available data about property. Usually, this register is integrated with lands and buildings register. Data is made available in the form of text files.

2.5. Database of Topographic Objects

The Database of Topographic Objects (DBTO) has been created in Poland since 2004. The scope of data corresponds to a topographic map at a scale of 1:10 000 and comprises the following thematic data: administrative units, road and railway networks, buildings and non-building structures, land cover, land use, watercourse networks, protected areas, geodetic control network, other objects, utility networks and address points.
As far as the management system for public real property is concerned, three classes of objects play a key role: buildings and non-building structures, address points, as well as road network (Bielecka 2012). One of the most important problems with data use is a big number of different application schemas. Currently, data are made available in sexteen TBD GML application schemas. The differences between application schemas are considerably large and they are related to, e.g. discrepancies between classes, attributes and related tables, also the definition of some attributes may be different. For property management, different scopes of definition of the *detailed function* attribute for the class: buildings pose problems. As a consequence, the DBTO data entered at the same time by different ordering parties are collected in different technical conditions and recorded by using different application schemas.

Data in the DBTO is made available in the form of the GML file, but the large number of TBD GML versions makes it difficult to upload to applications. An alternative format for making the DBTO data available is SHP, which is, however, a non-topological format. Therefore, relations between data recorded in this format are lost and, consequently, not all important relation can be transferred. Such a situation is especially inconvenient with regard to data about buildings and address points. This data is recorded in seven tables (buildings, address points, road sections, carriageway sections, areas of inhabited places, inhabited places and streets), so rebuilding the relations between tables requires the knowledge of the topological database structure and is very time consuming. Another difficulty in using the DBTO data for property management is that there is a lack of integration with the National Official Register of the Territorial Division of the Country as far as street names and inhabited places' identifiers are concerned, which results in the outdatedness of such data or incorrect spelling of such names and identifiers.

2.6. Orthoimages

In public property management, orthoimages serve a function of referential data which make it possible to verify the location of buildings and to preliminarily verify a property boundaries (fig 1). There are many websites that provide image data free of charge (e.g. geoportal.gov.pl, Google Earth, Zumi, wikimapia, igeomap). However, of special importance are orthophotomaps provided by the Polish official geoportal: geoportal.gov.pl because it guarantees data quality.

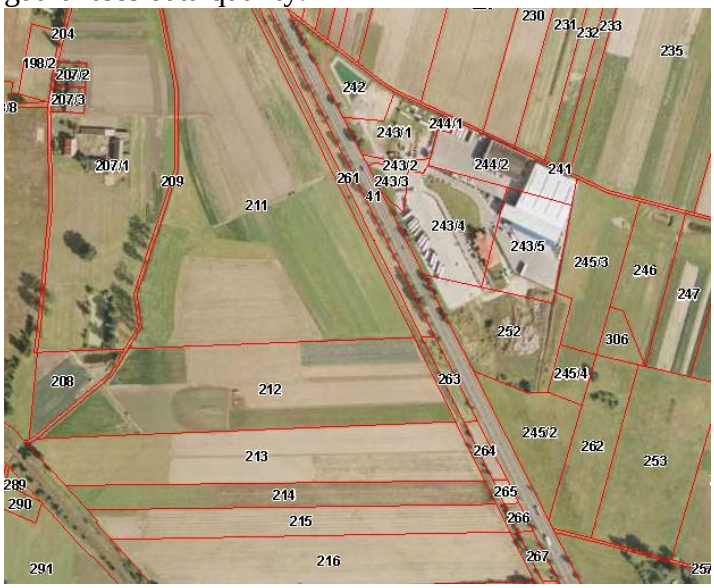

Figure 1. ortoimage and parcel boundarysie (source: http://mapy.geoportal.gov.pl/imap/)

The data comes from the central geodetic and cartographic documentation resources and constitute images with a ground pixel of 1.00 m, 0.50 m and 0.25 m. They are recorded in a coordinate system "2000" or "1992" and provided in digital form in the geoTIFF format (http://www.codgik.gov.pl).

2.7. State register of borders and areas of territorial division units

Data concerning the administrative division of a country is very useful in systems for public property management or administration. Its allow for visualizing properties in selected units, presenting data on property in the form of thematic maps, and for conducting various comparative analyses. The data is provided by the Main Geodetic and Cartographic Documentation Centre and made available in the formats: SHP, DXF, TAB and MIF (http://www.codgik.gov.pl/).

2.8. Data concerning property characteristics

Data about the technical and economic characteristics of a property play a very important role in public property management. Such data comprises, for instance, maintenance costs, repairs, renovation, equipment/fittings, technical or development conditions, and scheduled renovations. This data is collected in various forms – very often in the form of spreadsheets, and for various types of reporting periods – monthly, quarterly, half-yearly or yearly.

3. DATA INTEGRATION

3.1. Problem formulation

Data about property is dispersed: it is collected in the public registers (Table 1) as well as in registers kept by entities that manage property. Those who create property management system are faced with the problem of how to effectively use data collected in public registers and connect them with the data kept by a given administrative body. In other words, the problem is how to integrate data coming from heterogeneous sources in one system and how to ensure its up-to-dateness.
Etymologically, the word "integration" is derived from the Latin word *integratio* – "renewal"; *integer* means "untouched" or "complete", and *integrare* means "renew" or "recreate" and can be defined as "merging; the process of creating a whole from parts or incorporating an element into a whole" (Kopalinski, 1989). In information technology, the term refers to systems, applications, services and databases that function in a heterogeneous software and hardware environment. Data integration entails connecting data from different sources in one coherent data system. New data resources created in this way add value to the source data.

3.2. Data integration process

The process of data integration has many aspects, in particular it comprises three basic dimensions: conceptual, logical and physical. On the conceptual level, data integration entails establishing a common language to translate the terms of a conceptual model of a property management system into objects described by source data. As part of conceptual integration, it should be verified whether the basic concepts used in both systems are understood in the same way, e.g. if the definitions of a property, building, owner and costs are the same in all systems. Conceptual integration of data also comprises eliminating data which is irrelevant from the point of view of the adopted conceptual model, removing redundancies in the source

data, as well as establishing quality control system for data entered into the database (e.g. the required completeness of data). On the logical level, data integration entails removing the conflicting names of objects and attributes (Lenzerini, 2002). It refers both to homonyms (one name referring to different concepts) and synonyms (different names referring to one concept). Logical integration should cover all data that are uploaded to the property management system. Establishing relations between corresponding objects and their attributes in source and target databases is also referred to as mapping or synchronisation.
On the physical level, integration entails determining the physical connections between IT systems and reconciling the technical methods of data transfer. It is usually preceded by the cleansing and unifying of source data. Integration requires appropriate tools and qualified staff. On the technical level, the cleansing and unification of source data is carried out by ETL programs (*Extraction, Transformation, Load*). Their task is to download and process data from varied sources (relational databases, data from standing alome systems existing in various units, text files, spreadsheets, recording devices, binary files, etc.).
Currently, there are only few tools for harmonising spatial data: FME (Feature Manipulation Engine) produced by Safe Software, the Data Interoperability module offered by ESRI and GeoIntegrator produced by Intergraph. The tools offered as free software do not, so far, provide full functionality which is necessary for integrating data and they enable only limited integration in the form of changing a data encoding method and coordinate system.

3.3. Data integration within public property management system problems

Physical data integration entails in particular:
- transforming or converting data so that they are consistent with the agreed coordinate reference system;
- eliminating geometrical and topological inconsistency;
- changing the values of attributes which resulted from the use of different units or errors in data;
- subdividing data from one, common text field (e.g. street name, building and unit number or a person's name and surname);
- completing data on the basis of external data sources (e.g. completing a missing name of an inhabited place based on the "TERYT" code);
- coding Polish characters according to the adopted international (Unicode) encoding standards – utf-8;
- encoding data according to the adopted standard.

4 CONCLUSIONS

The purpose of computerization of public property management is to improve the efficiency and effectiveness of property administration and management.
Despite the lack of legal barriers to access to the data stored in a many public resources, significant technological barriers still hinder the use of this data in the property management database system. Different coordinate reference systems, data exchange formats, incompatible standards of encoding of Polish letters, timeliness, topological inconsistency, the lack of documentation, make the data useless for property managers.

The problem of obtaining data from multiple sources will decrease significantly after the implementation of Spatial Information Infrastructure. Spatial data services that provide harmonized data from national geodetic and cartographic resources will bring data integration in property management systems only to physical integration, performed once when loading data into the database.

REFERENCES

1. Bielecka E. (2012). The possibility of use the spatial data stored in state geodetic and cartographic resource for state property management. Studia i materiały Towarzystwa Naukowego Nieruchomości, vol.20, No 4, 19-30.
2. Geodetic and Cartographic Law (1989). Journal of Laws of 2010 No 76, entry 1287.
3. Land and Mortgage Registers and Mortgage Act (1982) Journal of Laws of 2001 No. 124, entry 1361.
4. ISO/FDIS 19152: 2008 Geographic Information – Land Administration Domai Model.
5. Kaufmann J., Steudler D. (1998). Cadastre 2014. A vision to a future cadastral system. Commision 7 FIG.
6. Kopaliński W. (1989) Słownik wyrazów obcych i zwrotów obcojęzycznych, Państwowe Wydawnictwo Wiedza Powszechna, Warsaw 1989, p. 232. *[Dictionary of Foreign Words and Phrases of Foreign Origin]*.
7. Lenzerini M. (2002). Data Integration: A Theoretical Perspective. PODS 2002. pp. 233–246.
8. Pakuła-Kwiecińska K. (2007). ¼ systemów do wymiany od zaraz. Systemy EGiB. GEODETA, nr 10(149), 4-6.
9. Pietrzak L., Hopfer A., Cegielski S. (2012). Reforms on a Real estate cadastre in Poland. Geodesy an Cartography vol. 61 No 2., 117-126. DOI: 10.277-012-0022-9.

BIOGRAPHICAL NOTES

Profesor at Military University of Technology, Faculty of Civil Engineering and Geodesy. Professional interests and skills: GIS and land administration system, spatial analysis, Spatial Data Infrastructure, data modelling according to ISO19 100 standards, land cover/use. Project leader of many national and international project in the field of GIS. Member of INSPIRE Thematic Working Group on land cover, member of national committee of ICA/ICC, CODATA, national standardisation committee TC 211 and others.

CONTACTS

Prof. Elzbieta Bielecka
Military University of Technology
Faculty of Civil Engineering and Geodesy
St. Gen. S. Kaliskiego 2
00-908 WARSAW,
POLAND
Email: ebielecka@wat.edu.pl