Workflows for Ensuring Consistency of Cadastral Data

Saša VRANIĆ, Hrvoje MATIJEVIĆ, Croatia

Keywords: Cadastre, Business Process, Workflow, Transaction, Spatial Data

SUMMARY

Cadastre is the essential land register and the basis for many services in the area of land administration and land management. Although European parcel based cadastral system is used in many countries, differences in business processes between countries exist. Through a long period of time these differences have been adapting to development of technology, legislation and other conditions within a certain country. With the development of information technology, business processes have been partially improved. This means that certain operations were improved but the business processes were still adapted to the analogue way of thinking.

Database management systems (DBMS) enable the storage of data into a digital, structured format. Databases operate using transactions which are meant to be short. With the development of technology the business processes became more and more complex, resulting in a strong requirement to enable long duration of transactions. Various advanced transaction models were developed in order to fulfil this requirement, but with limitations. In order to mitigate these limitations workflow management systems (WFMS) were developed.

WFMS provide flexibility and a modular design which enables covering of wide range of possible applications. WFMS's are mainly investigated in the area of non-spatial data. Although there are some possible solutions for basic maintenance of spatial data, it would not be possible to implement these solutions to cadastral data. Cadastral data have several layers of complexity due to the fact that they have legal and spatial component, various external stakeholders may impose some restrictions on cadastral data, etc.

In this paper we present an introductory research on the subject of modelling of business processes in cadastre and establishment of cadastral workflow management system.

Workflows for Ensuring Consistency of Cadastral Data

Saša VRANIĆ, Hrvoje MATIJEVIĆ, Croatia

1 INTRODUCTION

Cadastre is the basis for any land administration system (Enemark et al 2005). Cadastral systems have been set up in different time periods (Williamson and Enemark 1996) and throughout history have had different purposes and therefore different business rules. Most of the cadastral systems are based on similar foundations but also have details which are specific to each system. In order to support this variety of business rules and slight differences in data models, a cadastral information system should be flexible. Cadastre is the basic infrastructure supporting different segments in the area of land administration and land management by improving its underlying information system all related systems will get improved (Enemark 2001). International standard ISO 19152 – Land Administration Domain Model (LADM) defines generic conceptual model for static component of land administration system. Although it defines attributes for modelling of dynamic and temporal component of data, modelling of process for maintenance of data is outside of the scope of LADM (ISO/TC 211 2012).

Database Management Systems (DBMS's) are transaction based, i.e. each operation over the data is executed by strictly defined transactions. The basic transaction model as described in Gray (1981) implies short duration of a transaction. Due to increased need for support of long transactions, various transaction models have been developed. These transaction models are implemented on the database level, i.e. each transaction is a set of SQL commands (in relational DBMS's).

Figure 1 shows the trends in development of information systems. Modern information systems use DMBS's which reside in the second layer. DBMS's offer functionalities which used to be in tailor-made applications so the programmers have less components to develop. In that sense, the emphasis shifted from programming to assembling of complex software systems. Another trend represents the shift from the data to processes (van der Aalst et al 2003). The seventies and eighties were dominated by data-driven approaches and business processes were often neglected. Modern information systems tend to be process driven. Business process reengineering as a procedure of continuous improvement of business processes illustrates the increased emphasis on processes. By using the business process reengineering methodology for development of information system a fewer systems need to be built from scratch. Cadastral systems are influenced by legislative changes so the aforementioned development methodology fits their needs. Also, though European concept of cadastre is implemented in many countries, there are differences in data models and business processes. These differences require flexible systems with generic base components and additional components if needed to support specific requirements.

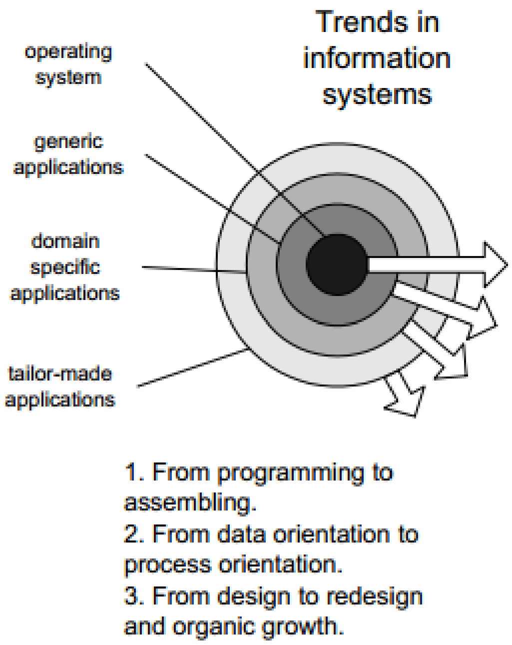

Figure 1: Trends in Business Process Management (van der Aalst et al. 2003)

By using the paradigm of business process management, where each part of a process (activity) can be assigned to a different user, one can also include external users into the process flow. Furthermore, it is possible to implement web services for certain activities to make some actions publicly available, leading in the direction of establishment of Semantic Web. The aim of the Semantic Web is to allow more advanced knowledge management systems (Antoniou and van Harmelet 2004).

The remainder of the paper is organized as follows. In the second section, the existing transaction models as well as their deficiencies for maintenance of cadastral data are analysed. In the third section the concept of workflow management system is explained. The fourth section gives an overview of requirements for cadastral data maintenance and explains the applicability of workflow management system for maintenance of cadastral data. In the last section results of this research summarized.

2 EXISTING TRANSACTION MODELS

The basic transaction model as defined by Gray (1981) has to comply with the following properties:

- Atomicity – all or nothing is executed
- Consistency – a correct transformation of data
- Isolation – effects of transaction hidden until successful completion
- Durability – effects survive failure

If several transactions are executed concurrently, i.e. at the same time, the results must be the same as if they were executed serially. Soon after its establishment, it became clear that the basic transaction model has deficiencies in cases when transactions have very long duration (days, months). This is especially important for cadastral transactions since they tend to be of a long duration. For that reason, many extended and relaxed transaction models have been developed. Extended transaction models allow grouping of operations into hierarchical structure, while relaxed transaction models “relax” (some) ACID properties (Eder and Liebhart 1994).

2.1 Advanced transaction models

Nested Transactions (Elmagarmid 1991) allow structuring of a transaction in a tree of subtransactions which may recursively contain other subtransactions. A child transaction may start after its parent has started and a parent transaction can terminate only after all its children terminate. Open Nested Transactions (Moss 2006) make the results of committed subtransactions visible to other concurrently nested transactions which enhances concurrency between long running transactions. As a consequence, the changes of subtransactions become globally visible and the root transaction must be undone by manually scheduling compensation transactions.

A long lived transaction, or a saga (Alonso et al 1996), is seen as a sequence of subtransactions that can be interleaved in any way with other transactions. Each subtransaction is an ACID transaction that preserves database consistency. Partial executions of a saga are undesirable so if the saga aborts then subtransactions that have committed must be compensated. Thus, each subtransaction has a compensation subtransaction associated with it, which undoes any changes introduced by the subtransaction. If any of the subtransactions aborts, the previous subtransactions are undone by automatically scheduling compensation transaction.

Split-Transactions and Join-Transactions (Pu et al 1988) is a combined concept which was designed for open-ended activities characterized by uncertain but usually very long duration, unpredictable development and interaction with other activities. A transaction may split into two separate transactions and later join another transaction.

Flexible transactions work in the multidatabase environment where each local database acts independently from the others. A flexible transaction provides alternative execution paths, if a subtransaction is aborted, then a different subtransaction can be submitted in the hope that it will be successful. A flexible transaction commits if either the main subtransactions or their alternatives commit.

Polytransactions (Elmagarmid 1991) have been proposed as a mechanism to support maintenance of interdependent data in a multidatabase environment. It is assumed that interdatabase consistency requirements are specified as a collection of Data Dependency Descriptors (D^3). Each D^3 contains description of the relationships among data objects, together with consistency requirements and consistency restoration procedures. The main advantage of polytransactions is that they transfer the responsibility for preserving interdatabase consistency from an application programmer to the system.

The ConTract model (Reuter and Schwenkreis 1995) is not an advanced transaction model as those previously described. Rather, it is a mechanism for grouping transactions into a multitransaction activity. Contracts consist of sets of predefined actions (with ACID properties) called steps, and an explicitly defined execution plan called a script.

2.2 Analysis of transaction models

Advanced transaction models are geared mainly towards processing the entities stored within a DBMS (Rusinkiewicz and Sheth 1995). All of the analyzed advanced transaction models have been developed with the purpose of mitigation of deficiencies of a basic transaction model from a certain perspective, having advantages or disadvantages depending on its application. Alonso et al (1996) point out that the main deficiency of all advanced transaction models is that they are too heavily dependent to a DBMS. The desire to overcome these limitations was a motivation for development of workflow models.

3 BUSINESS PROCESSES AND WORKFLOW MANAGEMENT

Workflow is computerised facilitation or automation of business process in whole or a part. A Workflow management system (WFMS) is a system that completely defines, manages and executes “workflows” through the execution of software whose order of execution is driven by a computer representation of the workflow logic (Hollingsworth 1995).

WFM systems have the similar purpose as advanced transaction models, but provide a wider set of possibilities. For instance, an activity in a workflow context can be any operation that can be monitored (user action, computer program, etc.) while a transaction in the database sense can be only a set of SQL commands.

Van der Aalst et al (2003) define Business Process Management (BPM) as follows: Supporting business processes using methods, techniques, and software to design, enact, control, and analyse operational processes involving humans, organizations, applications, documents and other sources of information”. It is important to note that BPM is broader than WFM. WFM is only a segment of BPM dealing with automation of business processes (Figure 2). BPM represents a continuous process of monitoring, analysing, improving, modelling and automation of business processes.

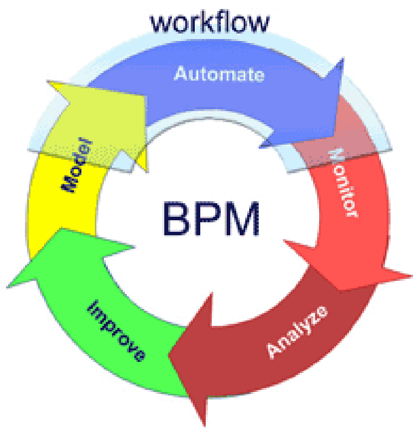

Figure 2: Relation between BPM and WFM

Reference model of a WFMS as defined by Workflow Management Coalition (WFMC) is shown below (Figure 3).

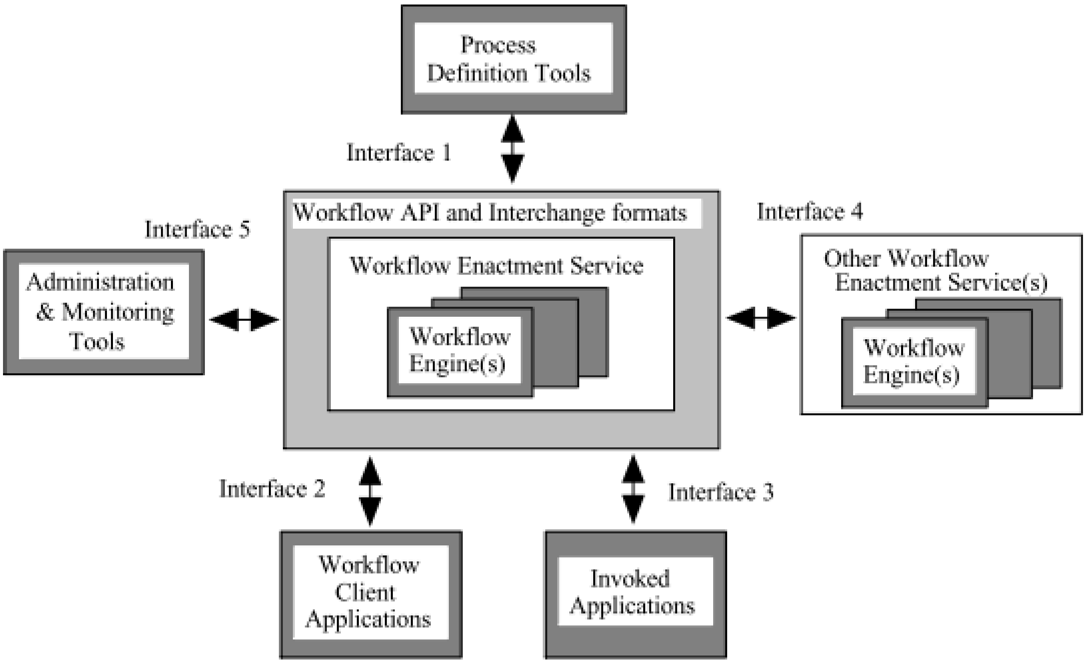

Figure 3: WFMS reference model (Hollingsworth 1995)

Workflow Enactment Service consists of one or more workflow engines with the purpose of creating, managing and executing workflow instances. Applications may interface to this service via the workflow application programming interface (WAPI) (Hollingsworth 1995). Furthermore, other Workflow Enactment Services can communicate to this service. Interaction with external resources is accessible to the particular enactment service via Client Application Interface through which a workflow engine interacts with a worklist handler, responsible for organising work on behalf of a user resource or an Invoked Application Interface, which enables the workflow engine to directly activate a specific tool to undertake a particular activity.

Processes are dominant factor in workflow management, hence it is important to use an established framework for modelling and analysing processes. There are several frameworks which can be used for modelling and analysing processes. Petri Nets was introduced by Carl Adam Petri in the nineteen sixties. Since then, Petri nets have been used to model all kinds of processes and have been extended several times to facilitate the modelling of complex processes (van der Aalst 1998). Van der Aalst (1998) notes several advantages of Petri nets in definition of workflow processes: formal semantics, graphical nature, expressiveness, vendor independent, etc.

Another process oriented notation is Business Process Management Notation (BPMN) which is based on a flowcharting technique tailored for creating graphical models of business process operations (White 2004). A very similar notation is used in Unified Modeling Language Activity Diagram (UML AD). Major difference between BPMN and UML AD is that BPMN is process oriented while UML AD is object oriented.

For the subsequent researches we intend to use UML because it supports a wide range of diagram types and other compatible modelling languages such as Object Constraint Language for definition of constraints. However, should a more formal definition of processes be needed Petri Nets seem to be more suitable for the purpose.

4 WFMS ON CADASTRAL DATA

Cadastral data are an important set of spatial data because they represent spatial extent of legal facts, they can serve as a basis for creating the topographical databases (van Oosterom and Lemmen 2001), and generally represent one of the bases of spatial data infrastructures. Considering the variety of statuses the cadastral systems in the world are in, the Land Administration Domain Model - LADM (ISO/TC 211 2012) provides a sound platform for modelling of static component of the land administration systems. Modelling of dynamic component cadastre is outside of the scope of the LADM so there is a strong requirement for a framework for modelling of business processes in land administration systems.

A business process is a set of logical tasks which have a common goal. Regarding the definition of business process, Phoung (2011) considers a land transaction to be a business process with several activities related to land. A business process can be considered a long term transaction (e.g. split a parcel and transfer ownership) while an activity is considered to have much shorter duration and ACID properties could be applied to it (e.g. assign a parcel to a transaction).

The legal domain of real estate transactions (Zevenbergen et al 2007) is well researched. Phoung (2011) published a research on enhancing transparency of land transactions by introducing WFMS into its execution. Dynamic aspect of spatial cadastral data is investigated through various aspects but there is no systematic research on maintenance of spatial cadastral data. For instance, in order to define the activities in a geometry editing workflow, or even to define the geometry editing workflow types, the types of changes on geometry would need to be known. Types of changes on geometry of cadastral data were investigated by several authors (Renolen 2000, Spery et al 2001, Zhou et al 2008, Fan et al 2010). Using the results of these researches could present a good starting point for definition of process types within Process Definition Tools of a WFMS.

Various aspects of WFMS have to be reconsidered and adapted to the cadastral data because cadastral data include a spatial component which imposes additional conditions on the maintenance of data. The first aspect which needs to be reconsidered is the correctness of data. Since the cadastral data have a spatial component it is important to control, besides the usual textual data correctness constraints, also the correctness of geometry and topology. Furthermore, if multiple and interrelated object classes are being edited, which is often a case, it is important to control the referential integrity as well as spatial relations between objects. Another aspect which needs to be reconsidered is the control of concurrent processes on spatial objects. When dealing with concurrent processes on spatial data it may be required to consider object being edited as well as its adjacent objects in order to preserve spatial correctness of data. This can either be done by explicit locking (van Oosterom and Lemmen 2001) or by version management (Peerbocus et al 2004).

Besides the spatial correctness requirements, the execution of transactions in cadastre must follow strict, legally defined procedures and comply with a variety of formal rules (Navratil and Frank 2004). Furthermore, different types of changes on the spatial cadastral data usually need to comply with different external requirements, e.g. urban planning. Hence in order to create a seamless procedure, the system which controls the consistency of spatial cadastral data should be able to invoke external services for checking of these special requirements.

The results of processes on cadastral data should not be visible to external users until the process has successfully finished. The isolation of a process is essential especially when producing official excerpts from the data. The excerpts can contain only the officially valid cadastral data and are usually the subject of payments. Because of that, producing of excerpts also needs to be modelled as a process. In that way, a process can be easily monitored, logged and payment can be recorded within the system.

WFMS provides a flexibility to cover the needs of cadastral information system which has many stakeholders involved in maintenance, checking and distribution of data. WFMS can be upgraded with the support for a certain transaction model. By combining WFMS with a DBMS transaction model, more rigid rules for activities can be introduced while maintaining the flexibility of a WFMS. WFMS also provides the ability to include external stakeholders into cadastral processes. Examples of such integration would be an urban planning department checking the specific spatial requirements on cadastral parcels prepared to be split, or editing of geometry of cadastral parcels by the system's external actors (e.g. private surveyors).

5 CONCLUSION

A cadastral information system must provide a seamless maintenance of cadastral data which may involve various stakeholders, internal and external to the institution itself. Those include cadastral officers, employees of other interested public institutions but potentially also private surveyors who prepare the changes to cadastral data.

Such complex set of activities grouped into business processes cannot be implemented by using database transaction models. WFMS provide possibility to extend the standard database transaction model and include other activities (e.g. web services, human interventions, etc.) as well as multiple users of different types in the process of data maintenance. The main deficiency of WFMS is inability to ensure consistency, isolation and forward recovery if they are not supported by some of the DBMS transaction models. By combining the workflow management system with advanced transaction models these deficiencies can be mitigated.

This work has been done as an introduction to a PhD research with a goal to define concepts of a Workflow Management System on cadastral data, in the first phase specifically its geometric part. The research will go in the direction of specification of WFMS for maintenance of spatial cadastral data in sense of process definition, correctness criteria, etc. Further to that we plan to investigate possible solutions for dealing with concurrent processes, distribution of parts of processes to external users (e.g. private surveyors, lawyers, etc.).

REFERENCES

Alonso, G., Agrawal, D., Abbadi, A. E., Kamath, M., Gunthor, R., and Mohan, C. (1996). Advanced transaction models in workflow contexts. In Proceedings of the 12th International Conference on Data Engineering, pages 574–581.
Antoniou, G., van Harmelet, F. (2004). A semantic web premier. England: The MIT Press Cambridge.
Eder, J., Liebhart, W. (1994). A transaction-oriented workflow activity model: Advanced transaction concepts for workflow activity management. Information Systems. pages 9-16.
Elmagarmid, A. K. (1991). Transaction models for advanced database applications.
Enemark, S. (2001): Land Administration Infrastructures for Sustainable Development. Property Management. 19(5). pages 366-383.
Enemark, S., Williamson, I., Wallace, J. (2005). Building modern land administration systems in developed economies. Journal of Spatial Science. 50(2), 51-68, Taylor & Francis.
Fan, Y.T., Yang, J.Y., Zhang, C. and Zhu, D.H., 2010. A event-based change detection method of cadastral database incremental updating. Mathematical and Computer Modelling 51 (2010), 1343-1350.
Gray, J. (1981). The transaction concept: Virtues and limitations. In VLDB, volume 81, pages 144–154.
Hollingsworth, D. (1995). Workflow management coalition - the workflow reference model. Technical report, Workflow Management Coalition.

ISO/TC 211, 2012. Geographic information – Land Administration Domain Model (LADM) – international standard ISO 19152:2012. ISO.
Moss, J. E. B. (2006). Open nested transactions: Semantics and support. In Workshop on Memory Performance Issues. Vol. 28.
Navratil, G. and Frank, A. U., 2004. Processes in a cadastre. Computers, environment and urban systems, 28 (5), 471-486.
Peerbocus, M. A., Medeiros, C. B., Jomier, G. and Voisard, A., 2004. A system for change documentation based on a spatiotemporal database. GeoInformatica, 8(2), 173-204.
Phoung, T. H. (2011). Enhancing transparency in land transaction process by reference architecture for workflow management. PhD thesis, University of Twente.
Pu, C., Kaiser, G. E., Hutchinson, N. C. (1988). Split-Transactions for Open-Ended Activities. In VLDB. 88. pages 26-37.
Renolen, A., 2000. Modelling the real world: Conceptual modelling in spatiotemporal information system design. Transactions in GIS, 4(1), 23-42.
Reuter, A. and Schwenkreis, F. (1995). Contracts – a low-level mechanism for building general-purpose workflow management systems. IEEE Data Engineering Bulletin, 18(1). 17
Rusinkiewicz, M., Sheth, A. P. (1995). Specification and Execution of Transactional Workflows. Modern database systems, 1995. pages 592-620.
Spery, L., Claramunt, C. and Libourel, T., 2001. A Spatio-Temporal Model for the Manipulation of Lineage Metadata, GeoInformatica, 5(1), 51-70.
van der Aalst,W. M. P. (1998). The application of petri nets to workflow management. Journal of circuits, systems, and computers, 8(01):21–66.
van der Aalst,W. M. P., ter Hofstede, A. H. M., Weske, M. (2003). Business process management: A survey. Business Process Management Journal, 2687. 1–12.
van Oosterom, P. J. and Lemmen, C. J. (2001). Spatial data management on a very large cadastral database. Computers, Environment and Urban Systems, 25(4):509–528.
White, S. A. (2004). Introduction to BPMN. Technical report, IBM Corporation.
Williamson, I. and Enemark, S. (1996). Understanding cadastral maps. Australian surveyor, 41(1):38–52.
Zevenbergen, J., Frank, A. and Stubkjaer, E., 2007. Real property transactions: procedures, transaction costs and models. IOS Press.
Zhou, X., Chen J. and Madden, M., 2008. Classification and Identification of Cadastral Structure Change. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. XXXVII (B4), 1437-1446.

BIOGRAPHICAL NOTES

Saša Vranić graduated at University of Zagreb, Faculty of Geodesy in 2009th with diploma thesis Interface of cadastral database. After graduation he worked in Croatian geoinformatic company Geofoto L.L.C. for several years as GIS Consultant on projects of implementing spatial information systems. Since 2012, he is employed at the University of Zagreb, Faculty of Geodesy as University Assistant on Chair for land surveying. He is also PhD student and his main research interests are Spatial DMBS and web applications in service of cadastral data management.

Hrvoje Matijević is currently working as Project Manager at the Croatian geoinformatics company IGEA L.L.C. and is a guest lecturer at the Faculty of Geodesy in Zagreb. Prior to this Hrvoje has worked in another Croatian geoinformatics company, Geofoto L.L.C. as Senior Consultant, and at the Faculty of Geodesy in Zagreb as University Assistant. He received masters degree in 2004 with thesis "Cadastral Data Modeling" and defended in 2006 his Ph.D. thesis "Modeling Changes in Cadastre" at same university. His main research interests were and still are GIS and Spatial DBMS technology in service of cadastral data management.

CONTACTS

Saša Vranić, MSc
University of Zagreb Faculty of Geodesy
Kačićeva 26
Zagreb
CROATIA
Tel. +385 1 4639 210
Fax + 385 1 4639
Email: svranic@geof.hr
Web site: http://www.geof.unizg.hr

Assistant professor Hrvoje Matijević, PhD
IGEA LLC
Frana Supila 7/b
Varaždin
CROATIA
Tel. +385 1 4639 210
Fax + 385 1 4639
Email: hmatijev@geof.hr
Web site: http://www.igea.hr