Land Use Policy 49 (2015) 565–576

Contents lists available at ScienceDirect

Land Use Policy

journal homepage: www.elsevier.com/locate/landusepol

A model for the creation and progressive improvement of a digital cadastral data base

Rodney James Thompson[a,b,*]

[a] Delft University of Technology, Jaffalaan 9, 2628 BX Delft, The Netherlands
[b] Department of Natural Resources and Mines, Cadastral and Geodetic Services, GPO Box 2454, Brisbane, QLD 4001, Australia

ARTICLE INFO

Keywords:
Cadastral schema
Progressive development
Topology
Land administration domain model
Data cleansing
Open source

ABSTRACT

A digital cadastral data base (DCDB) is a big investment for a jurisdiction tasked with the administration of land boundaries. In the past, the development of such a database produced no real pay-back on investment before many years of time, and millions of dollars in cash had been committed.

The Land Administration Domain Model (LADM) (ISO-TC211, 2012) provides a schema in which the progressive creation and improvement of a DCDB is possible; to allowing benefits to be obtained even in the early stages of effort. It also incorporates the necessary structure to ensure that a useful historical record of the cadastre can be kept. This paper explores issues to be faced in the development of software based on the LADM, which retains the history of the cadastre, and allows for progressive improvement of the data. From experience gained in the development of cadastral databases of the Queensland (Australia) Department of Natural Resources and Mines, and the Netherlands Kadaster, a suggested logical schema is presented and discussed with respect to the requirements of a progressively developed and refined cadastral database.

Rather than each cadastral jurisdiction developing its own database structure from basic geometric primitives, this paper proposes the establishment of a cadastral schema, based on the LADM, which can support all levels of encoding, variable accuracy and topological purity, while maintaining a comprehensive history. This would allow data quality to vary by geographic and temporal location and would be configurable to allow for country profiles under ISO 19152; thus permitting local terminology and language to be retained. Many jurisdictions are having extreme difficulty in successfully creating a cadastral database, so an open source type of software development may be indicated and desirable.

This paper presents findings based on theoretical consideration and the construction of a proof of concept database, which indicate that such a schema is a practical proposition for the development of a digital cadastral data base.

Introduction

Typically, a DCDB is repository which is developed as an adjunct to the administration of interests in land “It usually includes a geometric description of land parcels linked to other records describing the nature of the interests, the ownership or control of those interests, and often the value of the parcel and its improvements. It may be established for fiscal purposes (e.g. valuation and equitable taxation), legal purposes (conveyancing), to assist in the management of land and land use (e.g. for planning and other administrative purposes), and enables sustainable development and environmental protection” (Österberg et al., 1995). Such a rich set of data is also frequently used for many other purposes, providing background mapping for assets management, network administration, and other high value activities.

A problem has been that database structures chosen to support a DCDB have been such that data cannot be stored until it has passed stringent tests of validity, therefore much manual cleansing and correcting is necessary. This is exacerbated by the fact that a DCDB is of limited usefulness until it is complete. The classical approach to data capture in a spatial database has been for the incoming data to be validated against a set of rules, usually set by the database vendor, and often not well defined. Any failure of these rules results in the data being rejected.

Unfortunately, this puts a giant hurdle in front of any organisation. If the data cannot be entered without being correct, it

* Correspondence to: Department of Natural Resources and Mines, Cadastral and Geodetic Services GPO Box, 2454 Brisbane, QLD 4001, Australia.
Tel.: +61 07 38963286; fax: +61 07 38963697.
E-mail address: Rod.Thompson@qld.gov.au

http://dx.doi.org/10.1016/j.landusepol.2014.12.016
0264-8377/

cannot be made visible to a wider audience, and no pay-back can be obtained. Correcting data on input is a difficult (and therefore expensive) process, and the only eyes on the data are those of the data capture operators. As an indication, the Queensland DCDB took about 12 years to capture to an acceptable standard of quality, at the estimated cost of AU$50 million in the currency of the 1980s (Diggles, 2008, p. 209 in Part 3). (This equates to approximately $25 per parcel.)

By contrast, many of the uses to which cadastral data is put do not necessarily do not need highly validated data, and can accept small imperfections such as "knots", "overshoots", "undershoots", etc. It is important to remember that spatial data invariably has an intrinsic limit to its accuracy. For example, various mapping functions, including Web Map Services (WMS), Web Feature Services (WFS) and cadastral maps, searches etc., may be adequately supported by data with small imperfections. It is also important to be aware that validity and correctness are distinct concepts.

Traditionally, where information is publicly owned and maintained, and particularly when that data provides a legal framework for decision making, it has been the aim to prevent the release of data that might not be completely correct. An alternative viewpoint is that the possibility of errors in the database could be a reason to allow public viewing, so users might detect and report these errors (especially the ones that cannot be detected automatically).

This line of thinking supports the OpenCadastre concept (Keenja et al., 2012). As occurs in the OpenStreetMap, volunteers can enter data. Similarly, as in OpenStreetMap, users may correct each other's entries. In cases of 'conflicts', cadastral experts could be consulted to resolve these issues. This may be counter-intuitive as cadastre is about authoritative registration and the guarantee of land ownership and title, but provided a distinction can be made in the metadata between volunteered and authoritative information, and this distinction can be held in the public view, it may be an effective way to achieve clean and complete data. At the very least, making data visible to the public and providing an error reporting mechanism will lead eventually to higher quality data.

There are several possible ways to encode the geographic information in a cadastral database. The LADM (ISO-TC211, 2012) defines 5 levels of encoding:

1. "Text-Based" Spatial Unit
2. "Point-Based" Spatial Unit
3. "Line-Based" Spatial Unit
4. "Polygon-Based" Spatial Unit
5. "Topology-Based" Spatial Unit

(with "Sketch-Based" as a sub category of Text-Based).

These are discussed and described in Thompson (2013).

As a DCDB matures, it can be expected that its quality will be improved both in terms of its accuracy (Tarbit and Thompson, 2006), and in terms of its topological correctness (Thompson, 2013). This may also involve changes in the level of encoding. For example, a jurisdiction with polygon-based encoding might convert to a topology based form. In the past, an improvement in encoding would require a reworking of the database, with re-programming, data conversion and very probably a loss of history.

It is an important part of any cadastral database (though sometimes overlooked) to maintain the historical record of land use in digital format. In providing this functionality the Queensland Government, like the Netherlands Kadaster, adopted forms of what is now known as the "Versioned Object" pattern (van Oosterom, 1997). The LADM itself uses this pattern, permitting a permanent and efficient storage of cadastral history within the database. In a progressively developed database with history, it must be recognised that older historic data will usually be of a lower state of accuracy and topological purity than has been achieved later and may contain errors that have subsequently been detected and corrected. It is however commonly accepted that history of "the cadastre as we knew it" is a valuable resource. One important issue with history is that it must not be necessary to jettison many years of that history if the level of encoding is changed, or to partition the database into incompatible layers to allow progressive improvement.

With regard to the LADM, Lemmen raises a critical question "Is the design implementable and applicable in a real life situation?" (Lemmen, 2012, p. 14). The "FLOSS Cadastre Project" (Steudler et al., 2010) and "OSCAR" (Hay and Hall, 2009) argue that this is the case, and that a practical database can be built on the principles of the Social Tenure Domain Model (STDM) which is a profile of LADM. This paper explores the question further in terms of the LADM support of a fully mature cadastral database.

Original research presented in this paper includes: (1) The building of a database closely based on the LADM structure, and the loading of that database with realistic data quantities. (2) The use of that database to explore complexity issues. (3) The finding that levels of encoding can co-exist within the same cadastral database and that 2D and 3D parcels can be mixed.

The paper is structured as follows: the next section discusses the issue of data quality within a cadastral database; third section considers the requirement to record history; fourth section proposes a data model, and explores its capabilities; fifth section presents findings of an experimental database used to investigate the model; sixth section summarises the conclusions and the last section suggests further work.

Quality of a developing DCDB

Positioning accuracy

Measurement accuracy has improved over the years, but there is and will always remain a limit to the accuracy that can be achieved by any measurement. Typically, as a DCDB is being developed, the accuracy of the earliest capture will be lower than that of data added later (Effenberg and Williamson, 1996). The other major issue in this regard is that a DCDB may be the most useful and complete (or only) base mapping layer available, so it is often used as a background for assets management and for the positioning of street furniture by local authorities, electricity supply organisations, telecom, etc. (Priebbenow, 1993). It is usually true that the local relative accuracy of a DCDB is significantly higher than its absolute positional accuracy. This is certainly true in Queensland, where individual surveys are carried out to high accuracy, but the positioning of the property in absolute terms may have been done using a significantly lower accuracy technique (Diggles, 2008).

As an example, in Fig. 1A, an underground cable junction may be positioned 1 m from a property boundary. If a later survey is done which improves the positioning of the land parcels, it is not acceptable to lose this relativity between the cable junction and the parcels (Fig. 1B). The approach used in Queensland was that each time a vertex in the database was moved, a "point movement" record was generated, giving the old and new location of the vertex. These could be processed by the infrastructure authority to keep its asset locations up to date. The approach was not totally satisfactory, as it relied on the update operators maintaining point integrity, and not simply deleting linework and entering new points and lines. The result of this is that the major (paying) customers for the DCDB data are loath to see a large number of small adjustments to the positions of cadastral boundaries, and prefer that point positions be held. Thus the update process in use in Queensland is that incoming survey information is adjusted to fit the existing (probably lower accuracy) points. Only when a certain number of new surveys are available is a general adjustment of a region carried out.