W05.1

Charting a new frontier: the management of data in the MasterMap environment – a service provider's perspective

Philip Williams, Business Development Consultant, TerraQuest Group

Introduction

The vast majority of datasets relating to land and property in the UK are referenced to Ordnance Survey basemapping. The investment by an organisation in this type of data frequently runs to many hundreds of thousands if not millions of pounds. Despite this significant investment in what for many organisations is a core business asset there is a general lack of proactive spatial data management to maintain the continuing accuracy and fitness for purpose of the data.

Modern business applications supporting land and property processes increasingly use spatial data processing to create records, maintain linkages between records and to generate reports. The removal of geometric inconsistencies and the recording of real world topographic change in Ordnance Survey data without equivalent adjustment in pre-existing derived user data can result in misleading information being reported and erroneous decisions being made.

To take an example: a land charges application is likely to use the co-ordinate geometry of a property search polygon, derived from current basemapping, to identify relevant records with intersecting geometry. The details of the intersected records are then provided to the enquirer in response to the search.

Changes arising from geometric rectification and/or the larger scale movements resulting from the Positional Accuracy Improvement Programme are likely to result in intersections with records which do not actually relate to the property about which the enquiry is being made. This will result in the enquirer being provided with information which is potentially misleading or which might fail to reveal relevant information which may affect the value or use of the property to which the enquiry relates.

With the rapidly increasing requirements for the exchange of spatial information between organisations and with the General Public in support of the N-initiatives the fitness for purpose of data and the quality assurance of the processes that act upon it become critical issues. Failure by an organisation to manage these issues will at best result in loss of confidence and at worst in potential litigation.

MasterMap - the business opportunity

MasterMap provides a framework with the potential to facilitate significant improvement in the ongoing management of spatial data. For the first time, change only updates are available which can be globally applied to an organisation's Ordnance Survey map holding and selectively applied to derived user data. While the framework for significant process improvement now exists, much remains to be done in order to achieve the potential business benefit and diminution of risk.

MasterMap the challenge - basemapping

The efficient maintenance of basemapping via the application of change only update will require significant modification to the way that basemapping is managed within an organisation.

In a LandLine environment, the dissemination of basemapping updates is a relatively simple process, often merely requiring the replacement of files by more recent versions. In a MasterMap environment, the required processes are very similar to those prevalent within large-scale database management environments. GML, while an excellent vehicle for the initial supply of MasterMap and for the dissemination of change only update, is not a suitable medium for the storage and ongoing management of the basemapping. The application of change only update dictates that the MasterMap data be stored within a spatial database capable of transaction based processing. The database must also be capable of maintaining historical as well as currently active data.

Today, departmental or workgroup, based map servers are the norm. In South Gloucestershire Council, a relatively small organisation, 11 local map servers are currently deployed. This scenario is sustainable in the LandLine update environment due to its infrequent and file-based character. It is however illogical to maintain this scenario within a MasterMap change only update environment due to the greatly increased data management overheads and the potential for much more frequent update.

In the twenty-first century e-business based organisation, where internal and external exchange of spatial data is becoming increasingly commonplace, there is strong argument for the centralised management of basemapping. By definition, basemapping will then be consistent across the organisation and both it and the data that is derived from it can be qualified in terms of update history and the degree of currency.

The logical data management model is for, the potentially automated, maintenance of a central quality assured map store (spatial data warehouse) where the change only update is acquired and applied and from which application or task specific, for example, 'browser friendly', basemapping is served to satellite systems. The central service should also provide satellite systems with the change data, perhaps filtered, to facilitate the maintenance of locally managed derived data. In such a manner, the established techniques of data warehousing can be applied to produce a robust solution for the maintenance and dissemination of basemapping within an organisation.

MasterMap the challenge – derived data.

While MasterMap provides the potential to update spatial data not all business data should be modified to reflect all of the changes which will be present within an OS change only update.

While it is likely that positional accuracy improvement changes should be applied to all derived data, it is likely that other changes will not be applied in every case. The business rules that define the change application process will need to be defined for each class of user data within the organisation and processes implemented to ensure that these rules are applied.

The successful management of derived data within a MasterMap environment is dependent on the data being first synchronised with a known MasterMap release. Recent trials by TerraQuest using data from a number of Local Authorities have identified significant levels of inconsistency (up to 100%) between derived data and the referenced Ordnance Survey dataset. Studies have however shown that 80% - 90% of discrepancies within client specified tolerances can be resolved by the use of automated processes.

Larger discrepancies between the derived data and the MasterMap topography identified via the automated process must be resolved by the application of clearly defined and documented dataset specific business rules. While it may be possible to apply some of these rules via automated process, current experience suggests that it is more likely that the majority of decisions will require at least a degree of user intervention. It should, however, be emphasised that those discrepancies requiring manual intervention typically represent only 10% - 20% of the total present. Of these the majority only require a rapid visual inspection in order to accept positional accuracy movements that exceed the user imposed tolerances or the adherence to a simple resolution rule base. To put this in perspective, if the management of these automated processes were to be outsourced, then it is likely that on average only 2% - 5% of discrepancies would require any intervention from the client to assist in their resolution.

In order to facilitate derived data maintenance via change only update, the derived data's geometry must be cross-referenced to MasterMap TOIDs. This cross-referencing can easily be achieved as part of the automated and user assisted synchronisation processes. TOID referencing will, however, require significant change to the existing data structures and processes within application software, specifically the addition of TOID lookup tables and the tools to manage them. While it may initially appear desirable to attempt to derive a standard for TOID lookup tables discussions to date suggest that this would be impractical given the variety of user requirements, data capture companies and software vendors. It is however believed that it would be useful to discriminate between full and Partial TOIDs within the lookup tables. It is, also suggested that an XML based standard should defined to facilitate the exchange of derived data together with TOID references and versions between systems.

Application providers will also need to modify data schemas to facilitate a flag to denote dataset specific applicable levels of change. While it is likely that positional accuracy changes would be universally applied the application of real world topographic change is likely to be much more restricted. In reality it may well be that the business rules relating to topographic change may be so complex as to require manual intervention in most cases. Application providers will need to create or revise processes in order to facilitate this.

TOID based data capture and maintenance tools – some proposals.

Once TOID referencing is implemented, there is then significant potential for the development of editing tools to aid spatial data maintenance processes.

Discussions led by Peter Roberts between Powys County Council, South Gloucestershire Council, TerraQuest and MVM resulted in the following proposals for enhanced data capture and maintenance tools. While discussions focused on the provision of such tools within a MapInfo platform, the proposals are equally applicable to other software environments. It is envisaged that the proposed tools would be implemented via a toolbar or menu providing the appropriate tool selection, task completion, termination etc.

The processes outlined below assume the following basic concepts:

Dynamic Data: data to which real world topographic changes will be applied. A Change Flag controls the types of changes that will be applied.

Frozen Data: data to which change is restricted, for example, only positional accuracy changes might be applied. An example of this class might be historical planning application polygons. Frozen data is indicated via setting a Change Flag.

Capture Table(s): the table(s) into which the geometry and attributes will be stored.

TOID Lookup Table: A lookup table comprising the derived data object identifier, the TOID identifiers and version that in full or in part make up the geometry stored in the capture table(s).

Change Flag: Attribute stored in the capture table that indicates the change types that will be applied to derived geometry.

Polygon Capture

Single Polygon.

Derived polygon exactly matches a single MasterMap TOID.

Click within the required polygon, geometry is transferred to the capture table(s). The TOID reference, version and derived polygon identifier are transferred to the TOID lookup table. Associated data entry is facilitated. Change flag set dependant on data class.

Multiple Polygon

Derived polygon exactly matches multiple MasterMap TOID's.

Click within each polygon required creating the boundary extent. Indicate end of polygon capture. Polygons are merged and the resultant geometry is transferred to the capture table(s). The TOID references versions and derived polygon identifier is transferred to the TOID lookup table. Associated data entry is facilitated. Change flag set dependant on data class.

Complex polygon.

Derived polygon is composed of both full and at least one partial TOID.

Option 1.

Define the polygon by clicking on MasterMap line features where they define the correct boundary and by digitising new line work where no MasterMap line exists. Indicate end of first phase of capture. Topology is built and 'dangling' lines removed. User selects polygons to be included. Indicate end of second phase of capture. Included polygons are merged and the resultant geometry is transferred to the capture table(s). Both the full and partial TOID references versions and derived polygon identifier are transferred to the TOID lookup table. Associated data entry is facilitated. Change flag set dependant on data class.

Option 2.

Click within polygons that completely lie within the required boundary. Indicate end of first phase of capture. Accurately digitise line work to create fill-in polygons. Indicate end of second phase of capture. Polygons are merged and the resultant geometry is transferred to the capture table(s). Both the full and partial TOID references versions and derived polygon identifier are transferred to the TOID lookup table. Associated data entry is facilitated. Change flag set dependant on data class.

Line Capture.

A considerable uncertainty remains over the life cycle of the TOID associated with linear features, but given that this is resolved.

Single Line.

Derived line is composed of a single full or partial TOID

Click within a tolerance to select the line. If necessary unwanted, lengths of line work are edited out. Any additional line work, which does not relate to any MasterMap feature, is added. Indicate end of line editing. Geometry is transferred to the capture table(s). The TOID reference, version and derived line identifier are transferred to the TOID lookup table. Associated data entry is facilitated. Change flag set dependant on data class.

Multiple Line.

Derived line is composed of a multiple full or partial TOIDs.

Click within a tolerance to select each line. Indicate end of line capture. If necessary unwanted, lengths of line work are edited out. Any additional line work, which does not relate to any MasterMap feature, is added. Indicate end of line editing. Geometry is transferred to the capture table(s). The TOID references, version and derived line identifier are transferred to the TOID lookup table. Associated data entry is facilitated. Change flag set dependant on data class.

Point Capture.

Click on location for point. Geometry created and is transferred to the capture table(s). The TOID reference for the polygon within which the point resides, version and derived point identifier are transferred to the TOID lookup table.

Polygon Conversion.

Overlay existing polygon object onto MasterMap data and show all polygons that completely lie within boundary. User accepts or rejects each polygon in turn. Show all polygons that cross the existing polygon. User accepts or rejects each polygon in turn. Accept resultant polygon set. Polygons are merged and the resultant geometry transferred together with the full and partial TOID references. Associated data is transferred. Change flag set dependent on data class.

Data Management.

Dynamic data.

For positional change: locate new versions of all complete TOIDs and merge together, locate all geometry representing a partial TOID, apply OS shift file to relocate it, erase any part of the polygon that now falls outside the TOID within which the object fell, merge with the remainder of the object.

For real world change: update lookup with new TOID version numbers and remerge all TOIDs that comprise the object. If the TOID change impinges on an incomplete TOID object then modify this accordingly.

Frozen data.

For positional change: Check version number and change history of TOIDS that make up the object. If no real world change has occurred, combine the new objects together. Else, apply shift files to incomplete polygons, merge with full TOIDs and treat as per a data conversion exercise.

For real world change: Restructure and/or revise the lookup table to account for new TOID references.

Conclusions.

MasterMap provides a framework with the potential to enhance spatial data management across the organisation. At South Gloucestershire, it has been demonstrated that derived data can be synchronised with a given MasterMap release utilising a high degree of automation. However, without application providers enhancing their products to support TOID referencing and implementing tool sets to exploit TOIDs in supporting change only update, MasterMap will remain just cartographically enhanced wallpaper.

If organisations, both public and private sector, fail to recognise the urgent requirements for quality assured data and spatial data management and the opportunities provided by MasterMap to facilitate a solution then there is a real possibility for significant failure of crucial components of the N-initiatives and joined up Government.

Ordnance Survey has not yet announced a termination date for the supply of LandLine data, however, they have publicly stated that the policy decision has been made. If an organisation wishes to continue to use current basemapping to support its business processes then the transition to MasterMap will have to be made.

As has been discussed above and in the complementary paper by Barbara Jones the issues facing organisations are complex but the potential for significant business benefit and risk reduction is high. Today few application providers have progressed very far in addressing support for the management of MasterMap and data derived from it. Organisations therefore still have the opportunity to influence the processes that will be implemented. It seems therefore logical for organisations to plan ahead, assess their existing data holdings and the business process that act upon them, define potential business benefit and implement a well managed transition to the MasterMap environment.

Philip W. Williams.

Business Development Consultant.

TerraQuest Group plc.