# The Battle of Masterdata



GIS Professionals within large corporate organisations sometimes find themselves at odds with 'general IT' management. Some CIOs, IT Managers or Enterprise Architects lack a full understanding of GIS and, as a result, they try to apply rules to GIS that are better suited to other classes of systems, writes Nathan Heazlewood.

Over recent years a new area of conflict has opened up, one that threatens to undermine the effectiveness of GIS: the 'IT Generals' have issued an order that ALL DATA MUST be 'mastered' in one system (and that system is generally not GIS). Some organisations are increasingly taking this command to extreme lengths: removing data from GIS systems and restricting the access to that data once it has been transferred somewhere else. It is now time to find ways to stop retreating and counter this attack through improving the understanding and appreciation within 'general IT management' about the capabilities of GIS, about why data has to be available in certain formats and about the serious collateral damage if access to this data is forbidden.

This battle is particularly prevalent within larger organisations that are implementing 'Enterprise Resource Planning' (ERP) systems. ERP systems were originally designed to provide one system to carry out functions that are common to almost all organisations, such as purchasing, inventory, sales, marketing, finance and human resources. Often, an ERP will depose multiple legacy systems that only carried out one or two of these functions.

Intelligence reports from the frontline are suggesting that some IT Solution Architects are making statements such as "in a GIS the only thing that you need to hold is the X, Y". In extreme cases, Solution Architects are saying "well, our ERP system has a map so why do we need this GIS stuff?" IT architects can confuse basic digital maps and rudimentary GIS functions available in some ERP systems with a fully-fledged GIS. The result of such misinformed strategy is that GIS teams are often forced to surrender attribute data that they have traditionally managed over to management within ERP systems, and in some cases, worryingly enough, they are not even permitted to hold a copy of that attribute data within GIS. Instead they get some form of access to that attribute data via an API or other connection, often resulting in limiting the GIS functions that can be used. The result of the above effectively means that GIS teams are operating with 'one hand tied behind their backs'. Not advisable in any battle.

Some data is of legitimate interest to both a GIS and an ERP - the street addresses of clients serves as an obvious example. If the UN was in charge then what should be happening is for one or other system to be designated as the system where data can be created, updated or deleted and for both systems to be able to 'read' the data. Another option is for both systems to be permitted to edit, with a reconciliation process.

Part of the misunderstanding is that ERP systems are often 'transaction' based systems of record that typically deal with easily definable, repeatable and uniform processes. This may include frequent tasks such as taking an order from a new client, implementing an authorisation transfer process when a staff member resigns. Some ERP systems are also being used for tasks related to land, such as valuation of land parcels. Unfortunately, very often the experts tasked with designing those systems don't understand the strengths or functionality of a GIS that are both less easily defined and not always uniform. This includes tasks such as ad-hoc analysis, scenario based design and analysis, spatial analysis etc.

The importance of this is difficult to justify to people that have a mindset that 'all processes must be designed in advance'. Somehow we need to implant the understanding that one of the main strengths of GIS is the ability to answer ad-hoc questions and to perform previously unimagined analysis.

As an example, one very common use of GIS in 'real' military planning is 'scenario based analysis' i.e. future scenarios that may or may not happen. Like Churchill in his war-room, many military strategists use maps to display the location of ships or units, and then role play multiple different actions that the enemy might make, and what counter-attacks they can make with their own forces and then take into account of other geographic factors. How stretched will our supply lines get? What would happen if it rains tomorrow? Would these roads become unsuitable for tanks?

A civilian example of scenario based analysis is currently occurring where I live in Auckland. A new shipping port needs to be built and 28 possible locations have been compared using GIS to analyse various scenarios. As any GIS professional will understand, that analysis creates a lot of data, much of which will be discarded once a decision has been made. However, in one of those scenarios the 'draft' data will be the starting point for the 'real' data. If all of the records for these 28 scenarios need to be created in another system and then transferred to GIS, just so that coordinates can be added before analysis is then carried out in another system, then this leads to a lot of unnecessary extra work. A critical element of this is where and how 'object IDs' or unique identifiers are created. This process alone can cause havoc for the GIS professional since it can impact versioning and object ID history managed geographically (for example, the relationships between land parcel records when a single land parcel is 'split' to form two new land parcels).

Many spatial analysis queries and processes rely on a combination of spatial functions and 'standard' SQL queries. If the data to perform these functions is restricted or only available using certain limited API connections then a lot of analysis becomes impossible. Even if some form of connection is permitted for GIS to access the attribute data in another system then the performance of spatial queries can still often be painfully slow. For example, if a spatial query is generated in a GIS, and utilises many buffers, and intersects against geospatial data,

but also needs to run multiple nested SQL queries against many tables in an external system, then the time taken to execute the query can be many magnitudes greater than the time it would take if the data was structured and held within the GIS. One recommendation is that measurements of tasks like this is some of the ammunition that we need.

The GIS industry needs to get organised to coordinate and respond to the threat posed by 'misinformed' IT managers. The solution is to find ways to educate general IT management so that they understand the issues outlined above. GIS Professionals need to do things like encouraging more general IT people to attend GIS conferences. Throughout history many wars have been started through miscommunication. Therefore GIS professionals need to do more to learn the language used by our IT brethren. Our industry is heavily out-numbered by the general IT industry meaning that we must, therefore, push for new peaceful alliances with our IT colleagues, who are, after all, on the same side.

*Views expressed do not necessarily reflect the opinion of the author's employers or any 3rd party.*

*This article was published in GIS Professional June 2018*