

Geospatial Supply Chain to the Complex Animal of Data Quality

They sort of knew that the subject of cleaning legacy data and the role of automation in the supply chain would be a tough one. Thus, UK's Laser-Scan took its customers and partners out on a punting trip, booked a world champion snooker player-turned-comedian and did some golfing just before everybody went home. The outcome of three days of lectures and debate: the golden egg of your organisation is in the existing data you have already invested in.

By Remco Takken



Aerial Data of the Forties

With 100 billion euros invested in geospatial data throughout Europe, stated Laser-Scan's Steven Ramage, it should be worthy to maintain that investment, and maybe even get more out of it by making it do a whole lot of other things than we would have expected earlier on.

Chris Corbin of European GI Policy came up with a neat example. "In taking the RAF and Luftwaffe aerial data of the forties, and those of the USSR after the collapse of communism, these old data suddenly became valuable again when it became available online." He also pointed out that there is economic value to be gained from exploiting public sector information. "The current ageing demographic in the EU calls for a smaller public sec-

tor. We need to release people for other activities. E-Government, interoperability of systems and greater conformance to standards should be helpful in that regard," he said. Corbin was one of the first speakers of the Laser-Scan User and Partner Conference 2006, held between June 27th and June 29th at Wyboston Lakes, just outside of Cambridge, UK.

A Complex Animal

Paul Watson, chief scientist with Laser-Scan, called the spatial data supply chain "fundamental. It is a business process that manufactures a product out of sources of geospatial data new and old." For the difficulties that rise with legacy systems placed within new systems, and the need to combine the existing data with new information in order to produce

the required product. Watson named this supply chain "quite a complex animal". Nevertheless he did succeed in describing all steps from data entry up to final extraction and encoding. Data Integration Warehousing and making an inventory out of it came just after the core of what Laser-Scan is doing: the part of quality assurance and data reconciliation.

Radius Studio

During the conference, Laser-Scan's Chris Tagg and eSpatial's Bryan Hughes demonstrated what it is exactly that a product like the recently launched Radius Studio does with faulty data. For their presentation Tagg and Hughes chose cable information which had been drawn on vectorized raster maps. When

Fore at Laser-Scan Conference

displaying these data onto a modern digital map, all kinds of distortions and unwanted placement occurred, rendering the data unusable. By applying a set of rules, which can add up to an amazing 6,000 in some cases, Radius Studio was able to sort out the 'spaghetti data'. Tagg and Hughes repeatedly checked, reviewed and repaired some 8,000 mistakes each session, which ran just under a minute.

This is not 'a dinosaur with wings that is starting to fly' as one lecturer aptly stated, but before you can run quick little checks like these, a lot of work has been done by figuring out all the business rules that apply to the data. The collection and implementation of the rules is a tedious job, because most of the knowledge about spatial data is spread out over point applications, documents and the heads of individual workers all through an organisation. But once the business rules are saved in Radius Studio's Knowledge Management Tool, they can be useful for years to come in order to ensure quality checks of future data. Radius Vision is an additional editing environment for manual correction for data that isn't suited to automated processing. It also exists in a mobile version for field use, ensuring data consistency in newly captured data.

Value Chain Management Prayer

Tom McGuffog is an economist who has been working with Nestlé and Tesco, and is a Visiting Professor at the University of Glasgow in Scotland. He looked at geospatial data from an outsider's perspective, and hit the nail on the head when he commented that the "investments in data have made people frightened to share. If you look at joint use, we might get more out of it when sharing this expensive data for free, for instance to be able to provide better customer service. All parties would be better off once they agreed on this." Not only did McGuffog come up with great one liners like 'Opportunity and risk are the mother of inventory', he even provided the baffled audience with a 'value chain management prayer', quoted here in full: "Give us the will and insight to remove uncertainty caused by lack of cooperation, the character to accept irreducible uncertainty caused by dynamic users and markets, and the wisdom to know the difference."

From Core to Daily Life

Perhaps due to their well-known partnership with Oracle, Laser-Scan is more known in the GIS than in the CAD world. Having Autodesk's VP of Engineering Gary Lang talking about their fairly recent move into Geospatial Open Source was therefore informative to most people attending this conference.

According to Lang, all things that seem free these days have moved outside of the marketplace because it has moved from what's core into normal daily life. "MapServer is not unique anymore, it is a commodity. What we ascribe value to changes over time," Lang said. Most notable is the fact that Lang understatedly admitted that geospatial is not seen as core business to Autodesk by 'giving away' the code to MapGuide in Open Source. Of course, only parts of the solution are 'free', that is the 'commodity' components.

Geographic Swiss Army Knife

Another recent development important for platform neutral companies like Laser-Scan is the gradual move of mapping agencies towards the platform neutral Geographic Markup Language (GML). While Galdos CEO Ron Lake characterised GML as being all-round like a Swiss Army Knife, it was Snowflake's Ian Painter who made a likely portrayal of GML in the context of major implementations by mapping agencies that require reading and writing of GML data.

In the Netherlands, NEN 3610 was recently adapted to GML as a base for GI exchange. Painter explained how TOP10NL is a master map with five subdivisions. Each of those comes with a full collection of pre-built objects for different users, like the agricultural and the archeological worlds.

While Painter heralded NEN3610s compliance to ISO and OGC standards, he also found that "some of the problems are that they are doing too much. There are too many very vague definitions and options in order to keep everyone happy."

As for the German format NAS-AA, Painter explained that this particular example was "more than a move into a new application schema, because it has new data content, makes use of generalisation and is scale independent." The 190 feature types within it are, as opposed to the Dutch model, very specific, and they cover all possibilities. There are still



Tom McGuffog provided the audience with a value chain management prayer.

issues of adoption though, and Painter suggested that with the considerable re-engineering task Laser-Scan's Radius Clarity could be of help in Germany. Also, about 90 percent of the possible rich attribution in NAS-AAA does not yet exist in the data.

Golden Eggs

Laser-Scan CEO Mike Sanderson only took the stage for a couple of minutes, right at the end of the conference. He gathered that there were so many great speakers available, that he could suffice to just sit there and listen. In that brief moment, he summarized the three past days as follows: "We have heard that we are indeed living in a brave new world, with SOA, driven by the web and so forth, and that we can do more with less. But in those and other ugly words like GML, SOA, Ontology, and Semantics we still must find our recipes in moving forward."

Throughout those sunny days in June it became apparent that the fake eggs (they were chocolate, not stone) lying around on all tables, symbolized golden eggs. All attendees were unknowingly sitting on them in their organisations. It is the existing geospatial data which has to be treasured and maintained.

Remco Takken (rtakken@geoinformatics.com) is editor of GeoInformatics. Have a look at www.laser-scan.com for more information on the topics discussed in this article.