INTEGRATING GEO-DATA FROM MULTIPLE SOURCES

Web-based GIS

The next phase in GIS development on the internet should restore user freedom and give a fully comprehensive overview of geo-objects originating anywhere and containing all relevant information. It should be fully integrated and freely applicable. The author envisages data from multiple resources integrated for process support by combining (geo)graphical and other (administrative) data using Web-technology.

Six years ago our firm began, like others of its kind, with application developments based on the first available map servers. User-interfaces, specific applications and administrator modules were built on an ad hoc basis, with demand-driven Internet Map Servers (IMS). Great effort had to be put into correction of the â€˜badâ€™ behaviour of these IMS, which are actually client-like software specialised in producing web-enabled pictures and tabular information (Figure 1). Examples of IMS requiring â€˜behaviourâ€™ correction were those which showed severe performance problems in generating small-scale maps (1:100,000). We solved this by defining another resource with raster data for geographical orientation and questioning the â€˜real resourceâ€™ at a scale of 1:5,000. Another problem was unequal response times, so we split short and long transactions, i.e. server-side plotting. One of our clients, an example described below, needed a response time of two seconds for 1,000 users mostly based on Smallworld systems. Fortunately, IMS have improved substantially over the past two years and less attention has to be paid to their â€˜behaviourâ€™.

Internet Software
Our company has two GIS-departments and a Multimedia-Internet department (MM&I). Initial discussions conducted on the subject of this paper ran something as follows:

- GIS: "GIS is a very complex matter; it will be hard to get this on internet."
- MM&I: "All right; what will we derive from your complex GIS?"
- GIS: "Well, some pictures and some tabular information."
- MM&I: "OK, whatâ€™s new or difficult here?"

So we designed and built a system with an empty viewport filled by external resources presenting pictures and tabular information. If we obtain a picture with topographical information from one source, and a picture simply showing rivers from another, we can present them together in one View (through transparency). If both sources offer a filled-area picture we have to open another viewport. Because we expect that this software will become dominant, we decided to focus on internet software rather than on IMS-functions.

Market Strategies
In order to know which way we are heading, we have to focus on a place on the horizon to fix our direction. This implies having to make assumptions about future developments. In the field of hardware and network technology we expect few problems, although mobile communications should improve significantly (UMTS). The Open Geospatial Consortium defines standards for the use of GIS on the internet. WebServices (WMS, WFS) are a very important step forward; the XML-based GIS-specification (GML) may even have more impact in the future. If these standards are broadly adopted and implemented, real integration of different resources becomes possible. Client-based rendering and use of GIS-functionality in the browser, independent of software-vendor, can then occur. This brings us to the software vendors. The GIS vendors will have a double feeling: Is there a world to win or is there one to lose? Vendors of database software are rapidly implementing geo-objects and geo-functionality (see GIM International January 2005, pp. 40-41). Internet software is being produced to handle XML information. The open-source and public domain movements eagerly anticipate developments. So software vendor strategy and, of course, consumer requirements and demand will be decisive.

Internet Concept
Based on this discussion, and anticipating market developments, the next concept arises; the following is an explanation of Figure 2, from bottom to top.

- *Different sources.* Every possible digital resource can be integrated (streaming, 3D, GIS, databases etc.) The only requirements are availability on inter or extranet and the ability to respond correctly to a request. The quality of the resource is not relevant to this concept.
- *Geo-Connector* handles source-identification (IP-address and IMS), gets available information (GetCapabilities) and translates Open GIS-questions into a form this resource can handle, and visa versa. It knows when this resource can answer (geographical routing) and handles layout questions.
- *WebGIS admin* is an interface for configuration, available resources, user groups, functionality per user-group, etc.
- *Content Management System* hand-les content per group, including mailing lists, forum and other common CMS functions.

- *WebGIS.* The GIS and tabular interface present the integrated information in a form of tabs, including basic functionalities such as pan, zoom and identify. Specific WebServices can be added (as network-trace and overlay).
- *Programming Environment.* A reusable library (Fusebox 4) for ColdFusion code (Java/J2EE) meant to add quick specific functionality or process support.
- *Client Appearance.* The way the system is presented to a user group.

Working Examples

A certain utility company has five different GIS-systems and three Oracle databases carrying important additional information. Over ten years much money had been spent in getting assets digitised. It was time for a return on investment. The system needed to appear as one, integrated system. To prevent damage (websites KLIC) a contractor in the Netherlands is obliged to check for the presence of underground assets everywhere he plans to dig. Such a process could be 90% automated. Public authorities are responsible for issuing permission on all kinds of environmental issues and of course there is a need for surveillance. One problem is that much information and surveillance is spread over different organisations, creating a need for communication and sharing of geo-information. Surveillance is mainly done on a project basis (petrol station, transport of dangerous goods etc). A system has been built based on the concept that every project has its own site, geo-information and communication tools, of which the administrator has configured the content. The same concept, extended with 3D animations, interactive design tools, streaming video and 3D moving through aerial pictures, will soon be used for Participative Spatial Planning. A final example is a Regional Command Centre for Disaster Management. This organisation carries little data of its own but can access almost every data source available (utilities, buildings, environmental information, road maps, inhabitants etc) if the need arises. It needs communication tools to give orders, record response, communicate with personnel on location, track vehicles (police cars) and so on.

Middleware Level

GIS-software suppliers are working hard on improvements. Nevertheless, it is unlikely that they will make realisation of the above concept a priority amongst established internet software. The possibility of data integration and built functionality at middleware level (for example, at a data warehouse) also creates a different role for underlying (GIS-) systems. Further, the OGC reasons explicitly, in the case of WebServices, on from integration at middleware level, and not from the costly, â€˜classicâ€™ method of integration in which only data was involved. The final consequence of this is that in the case of standards being broadly implemented the IMS could be partly replaced by the GML in/out mechanism, in addition to making available Webservices.

The Crucial Step

The internet business is open to public domain movements and international standards. Database manufacturers such as Oracle and Microsoft - for market technical reasons â€“ also want to conform to OpenGIS standards. A major advantage of using non-GIS-specific database storage of spatial data is that the X, Y, Z-values can again be classified under simple and clear data structures. There is a growing feeling of regaining a grip on geometry and matter. In the future, many more purchased GIS-software components will work with a generic database. This is a major advantage, especially for Web-based environments, and research institutes, universities and businesses can again begin freely convert their own ideas into (public domain) software. An implementation structure such as GML is a crucial step for anyone who supports progress in this area of trade. For example, in The Netherlands, GML-like publishing is a very popular concept, despite a lack of co-ordination and exchange of know-how.

Citizens Pay Twice

Another crucial aspect is availability of data. In The Netherlands, and elsewhere, the existence of public authorities has been made dependent upon selling data. With GML data is transported direct to the client - and this is seen as a threat. Of course, Dutch taxpayers have financed such data and the price is so high that only other public authorities can afford it; so that the citizen pays again (indirectly), twice, thrice and so on, yet still without access to these desirable datasets. Let us hope the focus will change from selling data to improving and keeping it up to date, and that money will be used for this. Looking at the whole picture, open standards will become a reality in the near future. Database and internet specialists will rapidly offer alternatives to current GIS-internet solutions. Geo-information being the connecting key, this may prove valuable to the entire information supply sector, and new markets and applications will then open up as a result - to the further benefit of all.