

UML FOR BETTER COMMUNICATION AMONG STAKEHOLDERS

Data Modelling for Water Management

In 2004 the government agency responsible for water management in The Netherlands began a project to create an integrated geo-database for holding all geo-information on water-related objects. The set-up of the geo-database had to be sufficiently flexible for use at different management levels and geo-data had to be accessible using Web services. UML has played an important role in the project: it helps discussions, forcing designers and end-users to be specific.

Because of its geographic location in a low-lying delta, water management is of prime importance for The Netherlands. Rijkswaterstaat (RWS), the Directorate-General for Public Works and Water Management, is responsible for maintenance of main infrastructure networks, including the water network of rivers, canals, lakes, and North Sea shore flood defences such as dikes and dams, bridges and locks. The maintenance of these water-related assets requires adequate geo-information and an appropriate information infrastructure.

Beheerkaart-Nat

Organisationally, RWS is subdivided into local districts that carry out the actual maintenance work. In the past these districts were allowed rather autonomously to develop their geo-information and Information and Communication Technology (ICT) infrastructure. The disadvantages of the resulting colourful palette of approaches have since become clear: they obstruct both information exchange and aggregation for use at higher levels of management. In response, the Geo-information and ICT department of Rijkswaterstaat (RWS-AGI) is setting up a Geo-Information Infrastructure (GII) to harmonise use, content and format of the databases and digital maps. One of the geo-information datasets in this GII will be the Beheerkaart-Nat, which will contain information on all RWS water-related objects. The other datasets created within the GII contain information on topography (digital topographic maps and aerial photographs), permits, cables and pipelines, observations, samples and inspections. It is expected that the new initiative will considerably reduce ICT costs. The strategy is built on principles of uniform models, open standards, server-based computing and central data hosting and maintenance. A GII based on the Open Geospatial Consortium (OGC) Services Architecture has been established using both Open Source Software (OSS) and vendor components. The Beheerkaart has also to fit into this strategy.

Requirements

An important requirement for the data model was visualisation and analysis of the data from four different perspectives: organisational, functional, topological and system. Three organisational levels are distinguished: national (RWS-area), regional (regional departments) and local (districts). The functional perspective refers to the subdivision of object categories into subcategories. For example: 'bridge' and 'pump station' are subcategories of the category 'manmade objects'; 'dam' and 'quay' are subcategories of the category 'shore'. In total, six categories and thirty subcategories are distinguished. Thirdly, in order to maintain a correct topological structure the real world is represented within five vertically differentiated layers. The top layer consists of bridges, the next is 'water', the next the 'surface' layer, below this the 'waterbed' layer and at the lowest level there is the 'below-surface' layer. No overlap is allowed between objects on any level. Finally, there is the system perspective. From a hydrological point of view, the water network consists of water systems. Three levels are distinguished: 'main water system', 'water system' and 'water system part'. In addition, the data model has to fulfil several other requirements. The content should be in line with the RWS-system for managing and budgeting for maintenance of RWS water-related objects. Further, the model should be flexible and extendable; it should, for example, be possible to add management boundaries such as water quality and nautical management data at later stages. It should also be possible to generate aggregated information, especially for the use of higher management, at national network level. Aggregation should be enabled at various levels to serve each of the three management levels. A last requirement is that several types of queries should be possible, for example on the total number of bridges, or the total length of navigable rivers within a region.

Data Model Design

Based on the above requirements, first a conceptual model was designed of which the implementation details were not considered important. Next, a physical storage data model was developed to implement the conceptual model in a geo-database for which Oracle Spatial 9i was chosen. A main challenge in the design phase was the selection of the basic criterion for grouping and categorising RWS maintenance objects. Should this be based on their function and physical appearance, such as 'water', 'terrain', 'bridge', and 'lock' or on the 'topological layer' they belonged to? Another major concern was the modelling of objects to enable presentation to end-users at different levels of aggregation, each with the most optimal visualisation style: as polygons, curves/lines or symbols (points). Other concerns included how to link with meta-data on the Beheerkaart objects, such as date of capture, date of modification, accuracy, and responsible organisation unit, and how to model the 'history' or 'life cycle' of Beheerkaart

objects.

The choice of modelling and storage of the "history" of objects depends on the use-case requirements, and on the environment within which the system has to function. A first option is to monitor only the changes made during editing, so that action can be taken on these changes. Another option is to keep all past versions of objects in the database, even when they have been deleted. The last option enables visualisation of changes over time, for example display of all situations between 1st January 2000 and 1st January 2001. Furthermore, this option enables retrieval of geo-information for one specific moment in time, for example the situation on 1st July 2001. Both alternatives have been investigated during the design process and a definitive choice will be made during the implementation phase.

UML Model

Figure 3 shows part of the conceptual model in a Unified Modelling Language (UML) class diagram. The most important class is "Beheerobject". One of the subclasses is Lmaaiveldobject, a "topological layer" with two constraints: the objects in this class may not overlap, and they should together form a complete coverage of that layer, so that no holes are present. The diagram also shows that each Beheerobject can have three geometries - geom, vis1_geom and vis2_geom - for use at different levels of visualisation. To guarantee consistency among these three geometries during editing, one can only be edited by the end-user. The other two are ideally derived automatically using database triggers. Triggers consist of software code activated when an edit is made. The derivation of points from polygons is unproblematic, except for MultiPolygons, but these are avoided during editing. The derivation of lines from polygons is more complex because the quality of automated centreline extraction depends on the form of the polygon.

OGC Compliance

The Beheerkaart will be centrally stored; however, local districts will be able to edit the database from their own office via Web services, particularly OGC-set specifications like WFS-T (Web Feature Service-Transactional). Publication through an OGC Web Feature Service (WFS) requires the Beheerkaart-Nat data model to be in accordance with the GML model, since GML is the WFS data exchange format. Figure 4 shows the four steps involved in transferring a concept to a real WFS application. The GML (ISO 19136) model implements a subset of the General Feature Model (GFM) defined in ISO 19109. That norm uses a series of UML class diagrams to define the notions of "feature" (identifiable object in the world) and "feature type" (feature class). The UML for the GFM itself describes the "meta model", which defines the creation of feature types for particular domains.

Feature-centric Model

Traditional GIS systems use a geometry-centric meta-model in which the items of interest are primarily geometric (points, lines, polygons, etc), with additional attributes often stored in database tables. Features, on the other hand, are primarily typed by their conceptual significance within the application domain, such as "measurement", "borehole", "geological boundary" and "mine". The emphasis is thus on the semantics of geographic objects rather than on their geometry and the representation of that geometry as polygon, line, point or other shape. The geometry is actually just a property of the feature type. Further, a feature type can have more than one geometric property, either representing different spatial properties of the feature or implementing different representations of a property, for example at different scales or using different geometry models. This feature-centric model with objects having multiple geometric properties has allowed the Beheerkaart-Nat data model to fulfil the important requirement of data visualisation from different perspectives needed for a powerful and useful GIS.

Concluding Remarks

In the implementation phase the UML models will be used to automatically generate, for example, the XML schemas for GML (in the case of an OGC Web Feature Service). As from mid-2006 the data from the Beheerkaart and Web services for viewing and editing will be available for the entire RWS organisation.

Further Reading

- P. van Asperen, M. Grothe, S. Zlatanova, M. de Vries, T. Tijssen, P. van Oosterom and A. Kabamba, Specificatie datamodel Beheerkaart Nat (main part in English), RWS Report AGI-2005-GPMP-017, Delft, 2005, 130 p. Online: www.gdmc.nl/publications/reports/GIS31.pdf.