

# Topographic data, object orientation and GML

**Marian de Vries, Theo Tijssen, Jantien Stoter, Wilko Quak, Peter van Oosterom**

Delft University of Technology, Faculty of Civil Engineering and Geosciences

Department of Geodesy, Section GIS Technology

Thijssseweg 11, 2629 JA Delft, The Netherlands

Email: M.E.deVries@geo.tudelft.nl

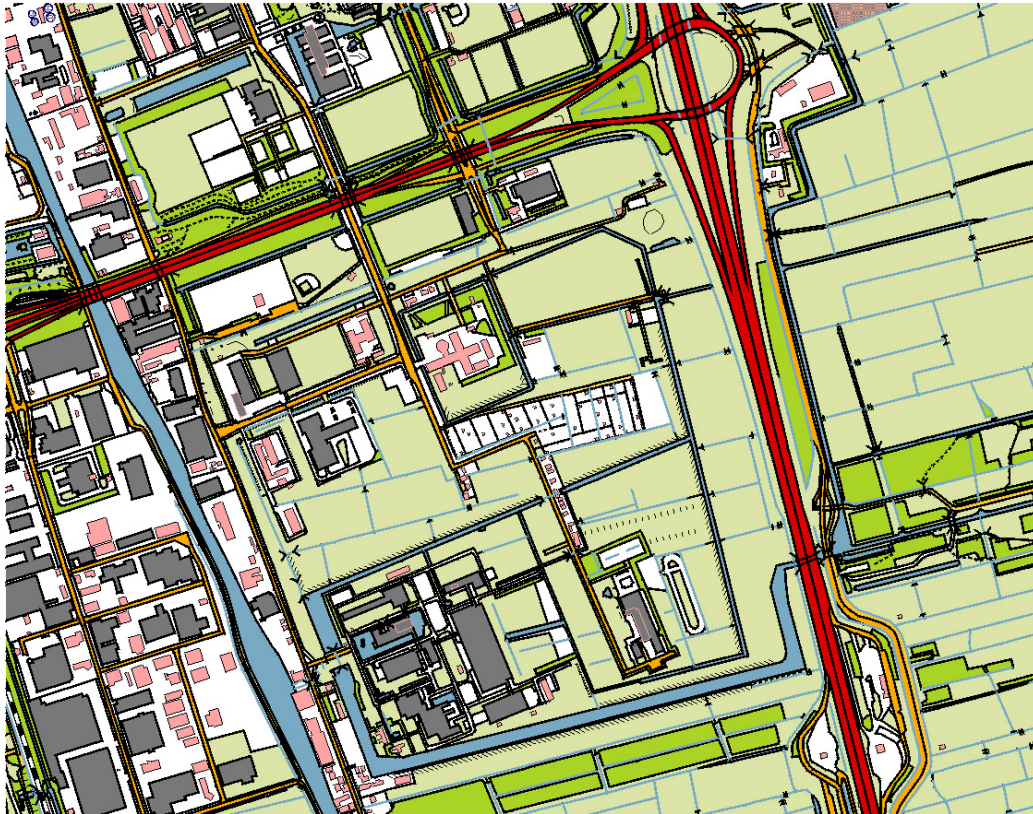
Tel: +31 15 2784268; Fax: +31 15 2782745

## Abstract

Since the creation of the first digital topographic data set there has been a need for exchange formats. This has always been a difficult issue and point of discussion. At the same time users have expressed additional requirements. Most current topographic data products are a mixture of geographic data and cartographic presentation. In the design for a renewed product in the Netherlands a clear distinction is made between a Digital Landscape Model (DLM) and a Digital Cartographic Model (DCM). The focus is on the design of the DLM meeting current and future user requirements. In parallel the exchange problem is tackled by using a new OpenGIS standard, Geography Markup Language (GML), which will be supported in future GISs.

## 1. Background

The Dutch Topographic Service (TDN) is the supplier of the TOP10vector, a digital vector file with topographical information of the Netherlands territory at a scale of 1:10,000.



*Figure 1. Map Fragment TOP10vector*

At the moment the data in TOP10vector is stored as lines, points and text-features. The current TOP10vector is not maintained in a relational or object-oriented database. Object (=geometric) types and attributes are stored by means of element coding. Polygons, as well as geo-objects, can only be formed as a result of the interconnection of linear elements. Users have to perform this operation themselves after the data has been imported in their own application environment.

However, other limitations exist. The only attribute information available is a feature code (road, building, pasture, river, etc.) [TDN 1998]. Most importantly, lines or (implied) polygons do not have (unique) object-IDs. This makes it hard to join or link the topographic data with other kinds of external geo-data such as addresses,

parcel-numbers or taxation records. User organizations in the Netherlands have asked the TDN to re-engineer their topographic data into a more object-oriented data model. For this purpose a new data model is defined for the TOP10vector that meets these requirements. Since there is also a growing demand to distribute the data in a more open transfer format, a prototype of this data model is implemented using GML 2.0. GML 2.0 is accepted as recommendation by the OpenGIS Consortium in February 2001 [OGC 2001]. The rationale behind the choice of GML is the fact that it is based on the world-wide accepted XML standard and that a rapidly increasing number of tools is available to generate, check and interpret XML/GML. ISO [ISO 2001] and OpenGIS [OGC 1998] have harmonized their models related to spatial data.

Section 2 of this paper describes the different roles of the participants in the TOP10 renewal project. In Section 3 the characteristics of the new DLM are explained at a conceptual level. The realization of the GML prototype will be discussed in Section 4. The 'proof' of the concept, a GML relay, is introduced in Section 5. Finally some conclusions are drawn and future developments are sketched in Section 6.

## 2. The project

The project 'Object Orientation TOP10vector' is carried out by the Section GIS Technology of Delft University of Technology (TU Delft) in co-operation with the Centre for Geo Information of Wageningen University (CGI), the International Institute for Aerospace Survey and Earth Sciences Enschede (ITC) and the Dutch Topographic Service (TDN). The project has a number of deliverables: a new conceptual model for TOP10 data, an XML Schema definition as implementation of the new model and some GML example files with real-world data.

## 3. New DLM

Important characteristics of the new conceptual model are: unique object-IDs, a partitioning of the surface as the basis for geometry (exceptions occur in case of overlap, e.g. road segments in tunnels or road segments on bridges), 2.5D objects with 3D coordinates, possibility of complex features (an aggregation of road segments into one – or more – 'named' roads) and the incorporation of metadata and temporal data for each object instance (versioning).

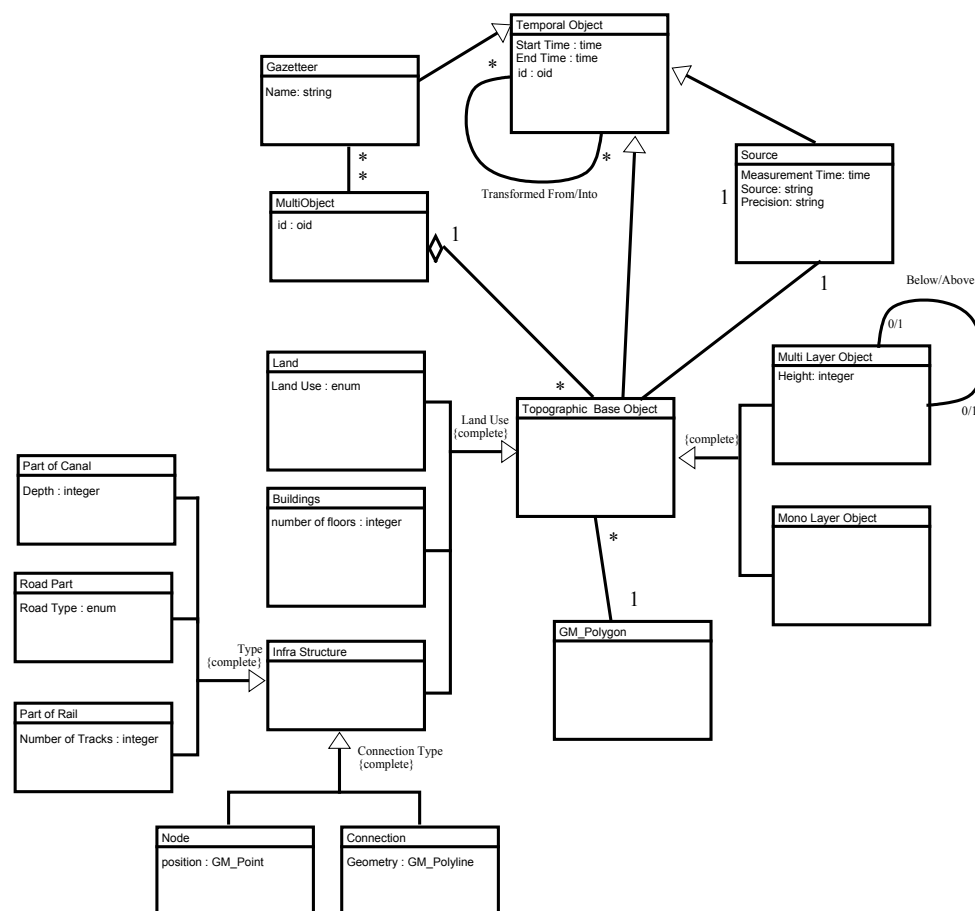


Figure 2. The new topographic model based on the geometry object model of OGC

The last characteristic opens the way for ‘change only’ updates distributed to user organizations [see also Badard 2001]. Many of the improvements suggested for the TOP10vector bear a strong resemblance to the developments of the model of the Dutch Cadastre [Oosterom 1997].

A new conceptual model was constructed using UML class diagrams and the OpenGIS Simple Feature specification [OGC 1999]. This specification contains standards to represent geospatial features and geospatial feature collections. Figure 2 shows the TU Delft proposal for the new conceptual model of the topographic data (DLM) based on the geometry object model specified in the OpenGIS Simple Feature specification.

#### 4. GML prototype

The process to create a GML prototype starts with the creation of a sample data set by the Dutch Topographic Service in accordance with the new conceptual model. In the next step this data is loaded into an Oracle 8i database at TU Delft. Then database views are created to model the data after the UML schemas. Finally the data is retrieved from the database and converted into GML. This is done by means of a Java program that first makes a JDBC connection to the database. Then it reads the data via SQL queries and writes the result to a GML file.

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<!-- File: tdndemo.xml -->

<tdn:top10vectorobjecten
  xmlns:tdn="http://www.tdn.nl/top10test"
  xmlns:gml="http://www.opengis.net/gml"
  xmlns:xsi="http://www.w3.org/2000/10/XMLSchema-instance"
  xsi:schemaLocation="http://www.tdn.nl/tdndemo.xsd">

  <gml:boundedBy>
    <gml:Box srsName="rd">
      <gml:coordinates>0,300 300,600</gml:coordinates>
    </gml:Box>
  </gml:boundedBy>

  <tdn:Weg OID="33">
    <tdn:TdnCode>02900</tdn:TdnCode>
    <gml:polygonProperty>
      <gml:Polygon srsName="rd">
        <outerBoundaryIs>
          <LinearRing>
            <coord> <X>173.0739</X> <Y>447.5921</Y> </coord>
            <coord> <X>173.0771</X> <Y>447.5889</Y> </coord>
            <coord> <X>173.0786</X> <Y>447.5833</Y> </coord>
            ...
            <coord> <X>173.0649</X> <Y>447.5932</Y> </coord>
            <coord> <X>173.0649</X> <Y>447.5932</Y> </coord>
            <coord> <X>173.0689</X> <Y>447.5935</Y> </coord>
            <coord> <X>173.0739</X> <Y>447.5921</Y> </coord>
          </LinearRing>
        </outerBoundaryIs>
      </gml:Polygon>
    </gml:polygonProperty>
  </tdn:Weg>
  ...
</tdn:top10vectorobjecten>
```

Figure 3. Fragment of GML code generated by Java program

For references to datums and coordinate systems the ‘well known text representation’ and IDs (e.g. SRSids) as defined by EPSG (European Petroleum Survey Group) are used. Apart from generating the GML file, the Java program also generates an XML Schema definition file. This schema definition file is partly based on database metadata (views that describe the tables, attributes and data types in the database). The file that defines the format of the GML file extends the OpenGIS GML schema definitions (feature.xsd and geometry.xsd). These documents themselves are based on schema definitions (xlink.xsd) delivered by the W3C organization [W3C 2000].

```

<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<!-- File: tdndemo.xsd -->
<schema targetNamespace="http://www.tdn.nl/top10test"
  xmlns="http://www.w3.org/2000/10/XMLSchema"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:gml="http://www.opengis.net/gml"
  xmlns:tdn="http://www.tdn.nl/top10test"
  elementFormDefault="qualified"
  version="0.4">

  <import namespace="http://www.opengis.net/gml" schemaLocation="feature.xsd"/>

  <element name="top10vectorobjecten"
    type="tdn:top10vectorobjectenType"
    substitutionGroup="gml:_FeatureCollection"/>
  <element name="Weg"
    type="tdn:WegType"
    substitutionGroup="gml:_Feature"/>

  <complexType name="top10vectorobjectenType">
    <complexContent>
      <extension base="gml:AbstractFeatureCollectionType">
        <sequence>
          <element ref="tdn:Weg" minOccurs="0" maxOccurs="unbounded"/>
        </sequence>
      </extension>
    </complexContent>
  </complexType>

  <complexType name="WegType">
    <complexContent>
      <extension base="gml:AbstractFeatureType">
        <sequence>
          <element name="TdnCode" type="integer"/>
          <element ref="gml:polygonProperty"/>
        </sequence>
        <attribute name="OID" type="integer" use="required"/>
      </extension>
    </complexContent>
  </complexType>
</schema>

```

Figure 4. Part of XML Schema definition file

The final step is to validate the generated GML file against the schema definition. Various standard XML tools can be used for this purpose. In our case XML Spy software was used to check whether the files generated are 'well formed' and 'valid'.

## 5. GML relay

In order to prove the concept of interoperability, the Netherlands Society for Earth Observation and Geo-informatics (KvAG) will organize a GML-relay. The GML-relay will take place on the 12<sup>th</sup> of June 2001 at the University of Wageningen, the Netherlands. The results will be presented at the EOGEOWorkshop during Digital Earth 2001, 24<sup>th</sup> to 28<sup>th</sup> June at Fredericton, New Brunswick, Canada (provided the abstract is accepted). At least one month before the seminar, each participant (3 to 5 geo-ICT vendors) receives data in GML 2.0 format (12<sup>th</sup> of May). The seminar itself starts with a general introduction of GML. Next, each participant can explain his/her views and present a demo, which consists of at least the following parts:

- read GML (2.0) data from a diskette into the system;
- edit this data (add or modify a few features);
- save the edited data in GML 2.0 and write it to a diskette as input for the next participant.

If the next participant can read this, the OpenGIS GML concept did prove to work (with these two independent products). If not, then the original data (of 12<sup>th</sup> May) can always be used as input. If a data set cannot be read, there will be no discussions about the reasons why it did not function. After the meeting the parties involved

will investigate this and publish the results on a website within one month (and if possible present the findings in Fredericton).

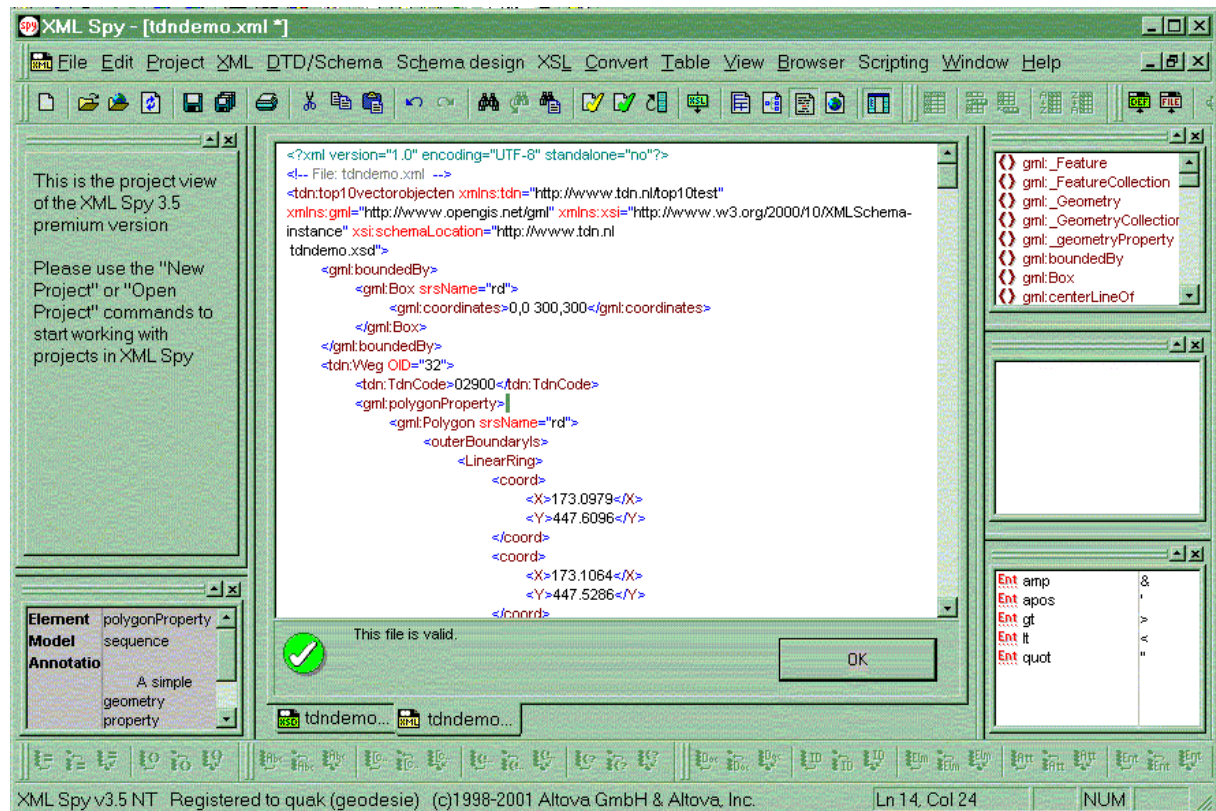


Figure 5 Example of check by means of XML Spy

## 6. Conclusion and future developments

With the project described the first steps have been taken to improve the (model for) topographic data in the Netherlands. After careful evaluation of user requirements by CGI, the ITC has developed a new conceptual schema in UML. The Topographic Service in the Netherlands has enriched a sample data set, which was converted into a GML (2.0) prototype by TU Delft. This prototype cannot cover certain aspects in a standard manner, because these are not yet part of GML 2.0. So, they are currently part of the application part of the XML schema. However, GML 3.0 developments have started with work items such as Topology, Temporal, Geometry Extensions, Units of Measure, Spatial Locator, Metadata Mechanisms, Default Styling and Points of Interest/Areas of Interest. It is obvious that standardization in these areas will further improve true interoperability. The work item 'Default Styling' may form a bridge between the DLM and DCM.

## References

- Badard (2001) *Using XML for the exchange of updating information between geographical information systems*, T. Badard, D. Richard; in Computers, Environment and Urban Systems, vol. 25 (2001), pages 17-31.
- ISO (2001) ISO/DIS 19107: *Geographic information - Spatial schema*, Final text of CD 19107, International Organization for Standardization, Januari 2001.
- OGC (1998) *The OpenGIS Guide, third edition, An introduction to Interoperable Geo-processing*, edited by K. Buhler, L. McKee, The OGC Project Technical Committee of the OpenGIS Consortium.
- OGC (1999) *OpenGIS Simple Features Specification for SQL*, revision 1.1, OpenGIS Project Document 99-049.
- OGC (2001) *Geographic Markup Language (GML) 2.0*, OGC Recommendation Paper, OGC Document Number 01-029, 20 February 2001.
- Oosterom (1997) *Maintaining Consistent Topology including Historical Data in a Large Spatial Database*; in Proceedings Auto-Carto 13, 8-10 April 1997, pages 327-336.
- TDN (1998) *Productbeschrijving en Objecten catalogus TOP10vector*, Versie 1.0, Paul van Asperen, Nicolien Haasbroek, Topografische Dienst Nederland, September 1998.

