Data Service: GML

Peter van Oosterom
Department of Geodesy, TU Delft, The Netherlands.

NCG/GIM,Ravi studiedag 20 juni 2002, Wageningen, The Netherlands

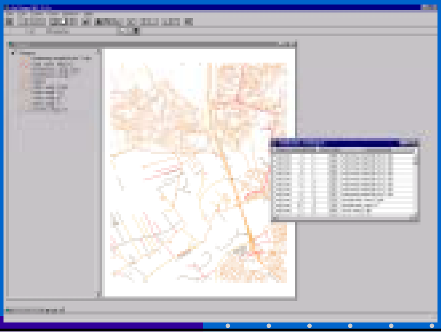
Current situation geo-information exchange

- Problems with current exchange formats
  - official standards (NEN/CEN): not well supported by products
  - vendor-specific standards: not fair for competition (other vendors) and may change over time (and may be not be the same on different platforms)
- From Internet GIS (web-mapping) the solution was offered: Geography Markup Language (GML)

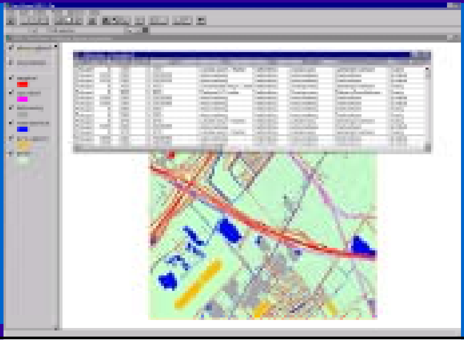
NEN 1878 = present exchange format

```
01278.        N200001200001  100000101top10vector            01
020110100000         0         0         1         0        101
03M26113    G030000   D20000101                              00
04I1        X161310549Y475005000                             01
03M25013    G030000   D20000101                              00
04I1        X183335225Y475005000                             01
03M25263    G030000   D20000101                              00
04I1        X160005000Y494859812                             01
03M24044    G110000   D19930101                              00
04I1        X178992927Y484947328I2        X178816060Y48527769001
03M24044    G110000   D19930101                              00
04I1        X178882670Y484888126I2        X178705879Y48521931801
03M24044    G110000   D19930101                              00
04I1        X177288961Y484137212I2        X177135174Y48447923501
03M24044    G110000   D19930101                              00
04I1        X177175005Y484085750I2        X177021186Y48442777401
03M24044    G110000   D19930101                              00
04I1        X175141766Y483726716I2        X175123406Y48410108701
03M24044    G110000   D19930101                              00
04I1        X175274285Y483736331I2        X175255829Y48411070501
03M24044    G110000   D19930101                              00
04I1        X173463187Y483656494I2        X173449143Y48403110801
```

Old situation: only geometry + linking-code



New situation: geometry + feature-info



TDN Project: restructuring Topographic data

- User requirements:
  - more object-oriented
  - also non-spatial attributes (names, road type, …)
  - metadata
  - history / temporal queries
  - easy aggregation (thematic, geometry)
  - linking with other data
  - unique id
  - conformance to standards: OpenGIS, CEN

OpenGIS and ISO

- Standards in this presentation are OpenGIS
- OpenGIS has two levels of standards:
  - Abstract specifications (domain knowledge)
  - Implementation specifications (for certain DCPs)
- ISO (especially TC211) has/develops also a set of geographic standards, comparable to OpenGIS abstract specs
- ISO and OpenGIS harmonize: abstract specs and ISO standards are (or will become) equal

Project ObjectTop10: *background case study new topo map*

- Geo-data transfer needs occurred immediately after creation of the digital map
- Many different GI systems, standards needed (de-facto and official), problems remain
- Later on explicit separation: DLM and DCM
- World-wide acceptance of OpenGIS
- TOP10 project:
  - accommodate new user requirements (CGI),
  - develop new logical model/DLM (ITC),
  - develop GML prototype for new DLM (TUD)

why XML?

- GML is based two important standards: OpenGIS geography model and XML
- XML enables well-formed check and valid (according to schema) check
- XML documents are text (simple editors)
- Integration with non-spatial data (XML-based)
- GML/XML is easy to transform (form DLM to DCM) using XSLT

Simple GML fragment, first example

Two examples of primitive geometry elements (valid according to standard geometry.xsd)

```
<Point srsName=“EPSG:4326”>
  <Coord><x>5</x><y>10</y></Coord>
</Point>

<LineString srsName=“EPSG:4326”>
  <coordinates>0,2 3,4 4,-1</coordinates>
</LineString>
```

GML, historic overview

- May 2000: GML 1.0 (Recommendation)
- December 2000: GML 2.0 (Candidate Recommendation)
- December 2000: Ordnance Survey Initiative “DNF” (Digital National Framework) also based on GML 2.0
- March 2001: GML 2.0 (Implementation spec.)
- April 2002: GML 2.1.1 (current version)

GML: schema definition, *GML limitations*

- Only linear elements (no arcs)
- No topology (e.g. NPR3611: chain + wheel)
  - Different types for different purposes:
    - represent a planar partition (polygon reconstruction of a face possible without looking at value of the coordinates)
    - represent a linear network of cables (or roads,…)
  - Main advantages:
    - avoid redundancy and maintain consistency
    - facilitate complex operators (e.g. shortest path)
- No temporal aspect (full delivery/updates)

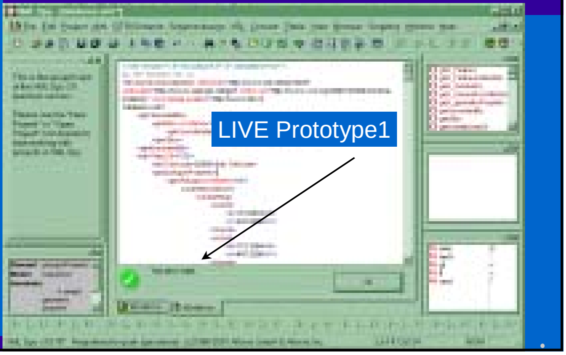
GML: schema definition, *TOP10 schema*

- Define application features (classes) with attributes (properties)
- Can solve standard GML limitations by application schema solutions for topology and/or temporal aspect (do not expect every GIS to be able to handle this)
- Application schema classes can inherit from abstract (OpenGIS) classes
- XML schema definition in a ‘xsd’ file
- XML data in a ‘xml’ file

GML: document, *validation*

- Check if document is well-formed according to xml standard. (match tags)
- Validate GML (.xml) file against Schema Definition (.xsd)
- Schema Definition is a collection of schema’s:
  - GML Feature
  - GML Geometry
  - Application (top10)
- XML Spy used for this purpose (commercial software)

GML: document, *validation*

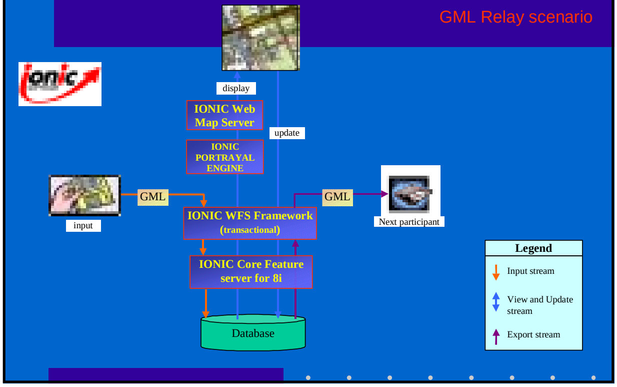


LIVE Prototype1

GML relay = proof of concept

- Goal: test interoperability of the GML data
- Method: a relay
- How: GML file to be imported, edited and exported by application 1, after that by GIS-product 2, after that by …
- Public event at KvAG meeting (June 12, 2001)
- Invited: number of large and small vendors

GML Relay scenario



Before the relay: first GML prototype

- srsName: full URL (<= examples in GML spec)
- <Coord> - pairs (<= examples in GML spec)
- (forgot the namespaces)

```
<tdn:extentOf>
  <gml:Polygon
srsName="http://www.opengis.net/gmp/srs/epsg.xml#28992">
    <outerBoundaryIs>
       <LinearRing>
         <coord><X>8.7004214E7</X><Y>4.52334491E8</Y></coord>
         <coord><X>8.7000476E7</X><Y>4.52335882E8</Y></coord>
```

GML prototype: data (.xml)

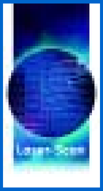
```
<tdn:top10Member>
  <tdn:terrein fid="TOP10.6000409">
     <tdn:top10_id>6000409</tdn:top10_id>
     <tdn:landgebruik>Overig</tdn:landgebruik>
        ...
     <tdn:hoogteniveau>0</tdn:hoogteniveau>
     <tdn:extentOf>
        <gml:Polygon srsName="EPSG:28992">
           <outerBoundaryIs>
              <LinearRing>
                 <coordinates>
                    85092.549,454965.491 85074.054,454988.376 85029.762,454953.805
                    85007.86,454980.098 85023.435,454992.271 85042.904,454970.847
```

GML prototype: schema (.xsd)

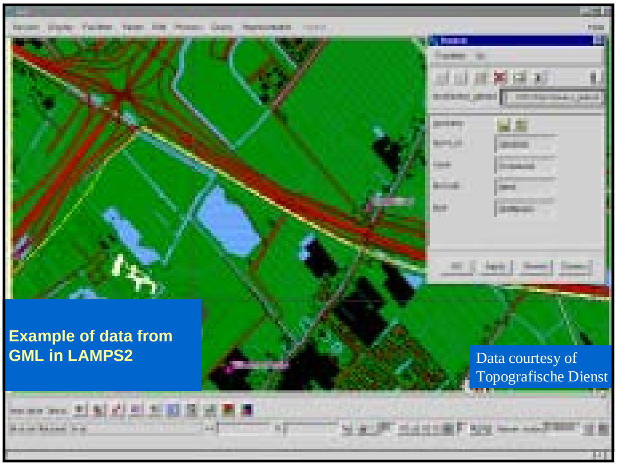
```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<!-- File: top10.xsd  -->
<schema targetNamespace="http://www.tdn.nl/top10test"
        xmlns="http://www.w3.org/2000/10/XMLSchema"
        xmlns:xlink="http://www.w3.org/1999/xlink"
        xmlns:gml="http://www.opengis.net/gml"
        xmlns:tdn="http://www.tdn.nl/top10test"
        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="top10Member" type="tdn:Top10MemberType"
      substitutionGroup="gml:featureMember"/>

   <element name="_Top10Feature" type="gml:AbstractFeatureType"
      abstract="true" substitutionGroup="gml:_Feature"/>

   <element name="terrein" type="tdn:terreinType"
      substitutionGroup="tdn:_Top10Feature"/>

   <complexType name="terreinType">
      <complexContent>
         <extension base="gml:AbstractFeatureType">
```

GML relay 12 June 2001 … may be too early
Next relay 13 December 2002

IONIC Software



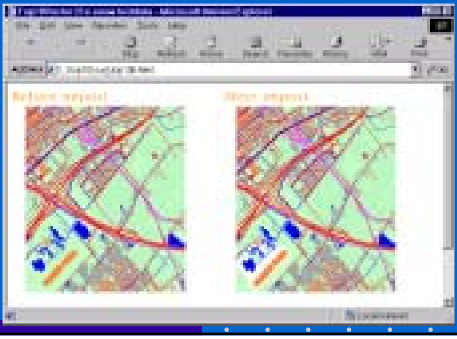
Laser-Scan



Example of data from GML in LAMPS2

Data courtesy of Topografische Dienst

Features with >100 coordinates



Issues / discussions / difficulties

- <Coord> vs <Coordinates>
- full or short srs-reference
- nesting / hierarchy:
  - ‘featureMember’ necessary as ‘association property’
  - substitutionGroup mechanism
- type inheritance: extension vs restriction
- namespaces

Attribute domain constraints

TDN_STRICT.XSD (definition typeBebouwing):

```
<simpleType name="typeBebouwing">
   <restriction base="string">
     <enumeration value="Gebouw"/>
     <enumeration value="Huizenblok"/>
     <enumeration value="Toren"/>
     <enumeration value="Installatie"/>
   </restriction>
</simpleType>
```

Status Top10vector project

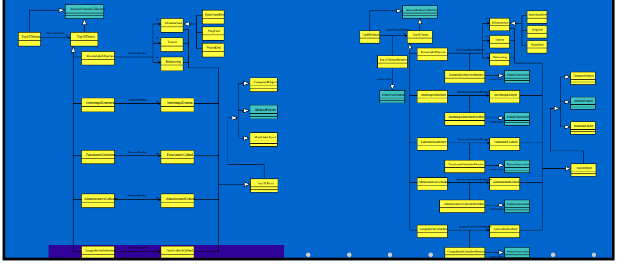
- evaluation of data model and GML
  - which user requirements are met, which not
  - no topology, no 3D
- feedback of user groups
- new projects:
  - cartography / visualization
  - other scale levels (1:50.000 - 1:500.000)

8 design issues for the Top10 Model

1. Use featureMember restrictions or not?
2. How do we map multiple inheritance to GML?
3. Do we flatten the object structure?
4. Do we use multiple GML files for one area?
5. How do we solve stacked features?
6. How do we solve Z-values with slopes?
7. What sort of temporal model do we use?
8. What do we do with multiple representations?

Design Issues (1/8): Use featureMember restrictions?

- Complex rich model vs simple model
- Do we use featureMember restrictions?

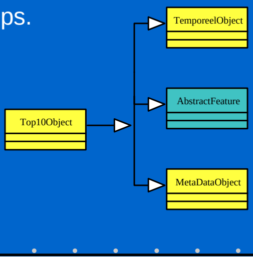


Design Issues (2/8): Multiple Inheritance

In UML one can model Multiple Inheritance, XSD does not allow this:

- We solved it with groups.

```
<group name ="TemporeelObject">
    <sequence>
      <element name="begindatum" type="string"/>
      <element name="einddatum" type="string"/>
    </sequence>
</group>

<group name ="MetadataObject">
    <sequence>
      <element name="brontype" type="string"/>
      <element name="bronbeschrijving" type="string"/>
      <element name="nauwkeurigheid" type="string"/>
      <element name="actualiteit" type="string"/>
      <element name="tdncode" type="integer"/>
    </sequence>
</group>
```



Design Issues (3/8): 'Flattening' GML Files

'Flattened' vs nested .gml file:

```
<Feature>
  <startTime>24 JUN 2001</startTime>
  <endTime/>
  <provider>TDN</provider>
  <accuracy>3m</accuracy>
  …
</Feature>
```

```
<Feature>
   <temporalObject>
      <startTime>24 JUN 2001</startTime>
      <endTime/>
   </temporalObject>
   <metaData>
      <provider>TDN</provider>
      <accuracy>3m</accuracy>
   </metaData>
   …
</Feature>
```

Design Issues (4/8): Multiple GML Files

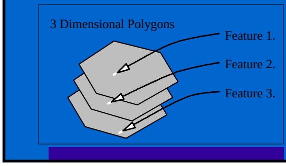
One or more .gml files per area:

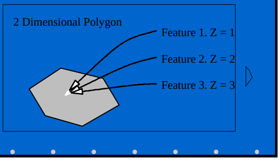
- Many users (and applications?) expect one map layer per GML file.
- GML allows a Collection of FeatureCollections, but will users (and GIS vendors) appreciate it?

Design Issues (5/8): Stacked Features

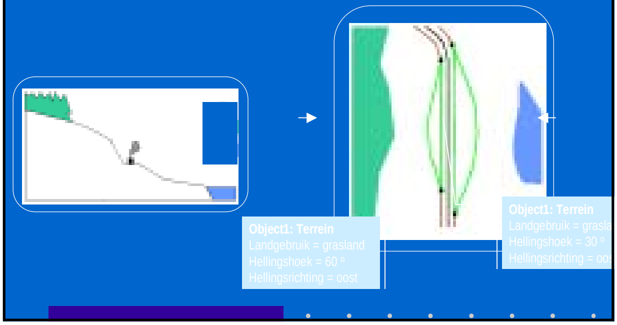
How do we model Features on top of each other? (viaducts, railroad crossings, etc.):

- Each feature gets a 3-dimensional geometry.
- All features reference to the same geometry and have a separate z-coordinate.





Design Issues (6/8): Z-Values in case of slope



Object1: Terrein
Landgebruik = grasland
Hellingshoek = 60 º
Hellingsrichting = oost

Object1: Terrein
Landgebruik = grasla
Hellingshoek = 30 º
Hellingsrichting = oos

Design Issues (7/8): Temporal Behaviour

How do we track objects through time:

- Does the ID of the object change or does it get a new version number.
- Which time instance do we record
  - The time the object is changed.
  - The time the object is surveyed.
  - The computer operator time.

Probably solved in GML 3.0

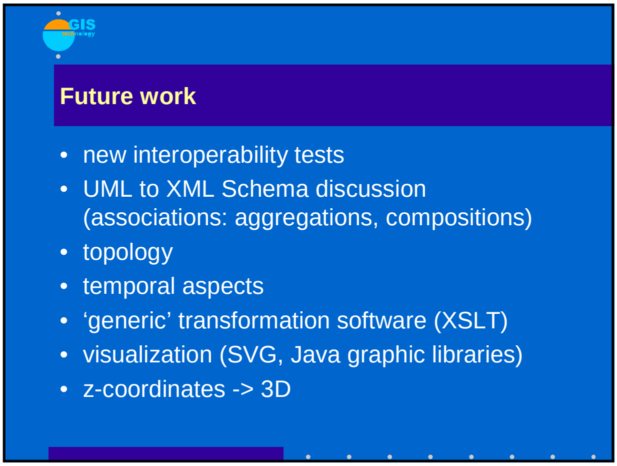
Design Issues (8/8): Multiple Geometries

How do we handle features that have more than one geometry?

- We can reference to multiple geometries, but will the GIS systems be able to handle this?

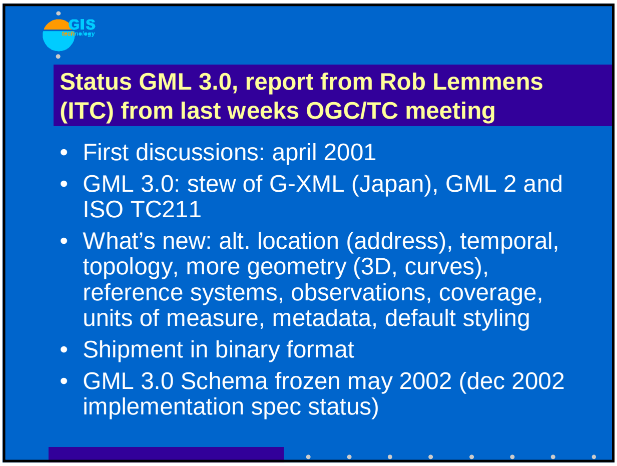
Conclusions

- GML has the future
- Beside the standard itself we need many examples / ‘best practice’
- GML 3 is very welcome:
  - topology
  - temporal aspects
- Cooperation between the NMA’s could help (same problems, same answers ?)



Future work

- new interoperability tests
- UML to XML Schema discussion (associations: aggregations, compositions)
- topology
- temporal aspects
- ‘generic’ transformation software (XSLT)
- visualization (SVG, Java graphic libraries)
- z-coordinates -> 3D



Status GML 3.0, report from Rob Lemmens (ITC) from last weeks OGC/TC meeting

- First discussions: april 2001
- GML 3.0: stew of G-XML (Japan), GML 2 and ISO TC211
- What’s new: alt. location (address), temporal, topology, more geometry (3D, curves), reference systems, observations, coverage, units of measure, metadata, default styling
- Shipment in binary format
- GML 3.0 Schema frozen may 2002 (dec 2002 implementation spec status)