Integration of GIS: a showcase study on GML based WebGIS

Conference Paper *in* Proceedings of SPIE - The International Society for Optical Engineering · November 2008

DOI: 10.1117/12.812533

4 authors, including:


Tohid Ahmed Rana
Yasar University


Tuncay Ercan
Yasar University


Yumei Sun
Huazhong Agricultural University

Integration of GIS: A showcase study on GML based WebGIS

Tohid Ahmed Rana[a], Murat Komesli[a], Tuncay Ercan[a], Yumei Sun[b]
[a]Yasar University, DYO Kampus : Kazim Dirik Mah. 364 Sok. No:5 35500 Bornova,Izmir, Turkey;
[b]Institute of RS and GIS, Peking University, Beijing, China, 100871.

ABSTRACT

The growth of Geographical Information Systems (GIS) emerges as analytical decision-making tools utilizing highly specialized geographic data. However, a simple system still does not exist for the integration/sharing information among GIS. Different organizations working within various domains have built their own data models for collecting and analyzing data. Many data storage, analysis, and delivery issues have prevented these organizations from effectively sharing their data. This presentation aims to show how to share the geo referenced information by utilizing GML (Geography Markup Language) technology among heterogeneous GIS. At present, the updates are retrieved manually; often failing entirely results in information loss and major inconsistencies or referenced data set can be purchased from geographic information producers. The problems and expense associated with the integration of updates for geographic databases are documented. Moreover, a showcase study on a GML based Web-GIS application system has also been developed in order to show how to share geographic information by using GML. The application includes various coverage layers online maps, important textual information and selected case studies

Keywords: GIS, XML, GML, heritage, integration.

1. INTRODUCTION

The development of the Internet has changed the environment for Geographical Information Systems (GIS), with the emphasis shifting from analysis to the sharing of data and information over the Internet thus making GIS more mobile and powerful. The Geography Mark-Up Language (GML) was developed as the standard language and is emerging as the foundation for Internet GIS.

Nowadays, Geographical Information Systems (GIS) are considered to be true analysis and decision-making tools. For that reason, a growing number of organizations invest in such systems and add specific information necessary to the tasks for which they have responsibility. However, propagation of geographic data in a GIS project takes %85 of overall project life cycle. Mostly, geo data is propagated by utilizing proprietary GIS software in its own format. This limits the usage and exchange of geo data among the heterogeneous GIS's. Furthermore, unless you own the commercial software, you can not even open the geo data in your GIS application. Actually, there are some means (like AutoCad DXF-Data Exchange Format) in order to exchange or import/export geo data among the different systems, but this way of exchange mostly requires manual corrections [7].

After announcement of GML (Geographic Markup Language) by Open Geospatial Consortium (OGC), which is simply a dialect of XML (eXtended Markup Language), the geo data will not only be labeled with tags, but can also be exchanged freely through explorer software on the Internet. Geo data exchange among GISs by utilizing GML will enable the developers to share their geo data without any proprietary GIS software license cost. By this way of integration, once the geo data have been propagated, there will be no need for proprietary GIS software in order to share geo data/information among the heterogeneous GIS's.

Geoinformatics 2008 and Joint Conference on GIS and Built Environment: Geo-Simulation and Virtual GIS Environments, Lin Liu, Xia Li, Kai Liu, Xinchang Zhang, Aijun Chen, Eds., Proc. of SPIE Vol. 7143
71430C · © 2008 SPIE · CCC code: 0277-786X/08/$18 · doi: 10.1117/12.812533

2. GEOGRAPHIC MARKUP LANGUAGE (GML).

GML is a specification of the Open Geospatial Consortium (OGC) [9], and becoming a world standard for the encoding, transport and storage of all forms of geographic information [3]. This is true increasingly for vector based data.

GML has three main roles with respect to geographic information. First, as an encoding for the transport of geographic information from one system to another; second as a modeling language for describing geographic information types; and third as storage format for geographic information. These different roles are essential to the understanding of the purpose and possible application of GML in any geographic domain. For data modeling, GML can be thought of as a framework that provides some fundamental components and patterns that are used to construct a domain specific vocabulary of geographic objects. Instances of the concrete geographic objects are then the substance of what is exchanged (GML data transport) or stored in a GIS.

2.1 Why GML?

Due to the lack of interoperability between GIS software, GML solution implies to define a specific information structuring for each format. Moreover, it often leads to the delivery of exchanges diffucult to integrate because GIS data formats have not been originally designed for the transfer of updating information. For that reason, the International Hydrographic Organization (1996) has designed a specific data exchange format for hydrographic data, which includes the delivery of evolutions. It allows real time updating of geographic databases, but is specifically devoted to the management of hydrographic data sets, which prevents its application for other geographic data [10].

Together with the development of these solutions, the computer science community has seen the emergence of XML (World Wide Concortium-W3C) [8], a new data format for the exchange of structured documents on the web. XML stands for eXtended Markup Language. While HTML is powerful for formatting a document, XML describes the document's structure and meaning [4].

GML is a dialect of XML (eXtended Markup Language). By utilizing GML, geo data will not only be labeled with tags, but can also be exchanged freely through explorer software on the Internet. Geo data exchange among GIS by utilizing GML will enable the developers to share their geo data without any proprietary GIS software license cost. By this way of integration, once the geo data have been propagated, there will be no need for proprietary GIS software in order to share geo data/information among the heterogeneous GIS's.

GML is designed for the Internet and directly embraces the ideas of interconnected and distributed information elements [1,2,5]. GML was not developed simply as a means of shipping data between separate, otherwise non-interoperable geographic information systems, although it can and is used in this role. GML is intended to form the fabric of a geographic world wide web, interconnecting geographic data sets and services, much as HTML has done for the wider world of the Internet and textual information access.

2.2 GML Encoding

GML is build on the abstract model of geography which is produced by the OGC. This describes the world in terms of geographic entities called features. Essentially a feature is just like a list of properties and geometries. Properties have the usual name, type, value description. Geometries are composed of basic geometry building blocks such as points, lines, curves, surfaces and polygons.

There are two basic XML schema [6] in order to model a geographic feature in GML. These are:

- GML Feature Schema (feature.xsd.)
- GML Geometri Scheme (geometry.xsd)

Although, a considerable amount of gegraphic information can be coded by means of these two schema, features can coded without geometry. Let us take as an example, how we can code unspatial features in GML. This is known a classical “Dean” example. The feature, named “Dean” has properties “familyName” in string type and “age” in integer type. Additionally, it has one more property in string type, “nickName”. So, an instance of Dean feature in XML code can be as follows:

```
<Dean>
   <familyName>Gerekee</familyName>
   <age>48</age>
   <nickName>Gereky</nickName>
   <nickName>Whitehead</nickName>
</Dean>
```

Related XML Schema can be:

```
<element name="Dean" type="ex:DeanType" />
<complexType name="DeanType">
    <sequence>
        <element name="familyName" type="string"/>
        <element name="age" type="integer"/>
        <element name="nickName" type="string" minOccurs="0" maxOccurs="unbounded"/>
    </sequence>
</complexType>
```

However, it can be described in detail how feature and property can be identified by using “feature” scheme of GML. In Dean Example, “dean” is a feature type, where as “age” is a property. This can be shown in GML as follows:

```
 <element name="Dean" type="ex:DeanType"
     substitutionGroup="gml:_Feature" />
<complexType name="DeanType">
   <complexContent>
     <extension base="gml:AbstractFeatureType">
        <sequence>
           <element name="familyName" type="string"/>
           <element name="age" type="integer"/>
           <element name="nickName" type="string" minOccurs="0" maxOccurs="unbounded"/>
       </sequence>
```

```
    </extension>
  </complexContent>
</complexType>
```

Only “instance” document is validated by this XML schema. Dean type content model in both XML schema definitions are the same. The use of feature schemas out of GML makes possible to use the predefined descripted attributes. For example, features can be identified with an attribute “fid”, and pre-described property “gml:description”. These capabilities are produced by the inheritance from “gml:AbstractFeature Type”. To make the situation more clear, “Deanextends gml:AbstractFeature” can be used.

```
<Dean fid="D1123">
   <gml:description>A nice old chap</gml:description>
   <familyName>Gerekee</familyName>
   <age>48</age>
   <nickName>Gereky</nickName>
   <nickName>Whitehead</nickName>
</Dean>
```

2.3 Geometry Encoding

Basic geometric elements such as Point, LineString, LinearRing, Polygon can be defined in GML bu utilizing OGC Simple Feature Model specifications. In addition to these elements, there is one more element called as Box, in order to code the coordinates with the elements <coordinates> and <coord>.

The coordinates of instance of any geometry class is encoded as an only string either in <coordinates> elements or as a sequence of <coord> including tuple sections. The advantage of using <coord> element is to enable the control as a validating XML parser. Related schema parts are shown as follows:

```
<element name="coord" type="gml:CoordType" />
<complexType name="CoordType">
  <sequence>
    <element name="X" type="decimal"/>
    <element name="Y" type="decimal" minOccurs="0"/>
    <element name="Z" type="decimal" minOccurs="0"/>
  </sequence>
</complexType>
```

A point element is actually a whole coordinate tuple:

```
<Point srsName="http://www.opengis.net/gml/srs/epsg.xml#4326">
     <coord><X>5.0</X><Y>40.0</Y></coord>
</Point>
```

Moreover, coordinates can be identified as a single string. Normally, coordinates in a tuple are separated by commas from each other. Coordinates is shown on the following XML schema:

```
<element name="coordinates" type="gml:CoordinatesType"/>
  <complexType name="CoordinatesType">
    <simpleContent>
      <extension base="string">
        <attribute name="decimal" type="string" use="default" value="."/>
        <attribute name="cs" type="string" use="default" value=","/>
        <attribute name="ts" type="string" use="default" value="&#x20;"/>
    </extension>
  </simpleContent>
</complexType>
```

Lastly, a point will be as follows:

```
<Point srsName="http://www.opengis.net/gml/srs/epsg.xml#4326">
   <coordinates>5.0,40.0</coordinates>
</Point>
```

A single feature like an airport might thus be composed of other features such as taxi ways, runways, hangers and air terminals. The geometry of a geographic feature can also be composed of many geometry elements. A geometrically complex feature can thus consist of a mix of geometry types including points, line strings and polygons.

To encode the geometry of a feature like a building we simply write:

```
<Feature   fid="142" featureType="building"  Description="A university">
    <Polygon name="extent" srsName="epsg:2137">
      <LineString name="extent" srsName="epsg:2137">
        <CData>
          491888.999999459,5458045.99963358 491904.999999458,5458044.99963358
          491908.999999462,5458064.99963358 491924.999999461,5458064.99963358
          491925.999999462,5458079.99963359 491977.999999466,5458120.9996336
          491953.999999466,5458017.99963357 </CData>
      </LineString>
    </Polygon>
</Feature>
```

3. APPLICATION: WEB-GIS

Another scope of this work is to develop an on-line spatial database with multidisciplinary data (e.g. Administrative, topographical, physical, geological, hydrological, climatological, geomorphological, agricultural, etc.) according to the above mentioned prospectuses. The online spatial database will act as information rendering and data exchanging medium and will provide its' users with adapted, rich multimedia experiences. In the light of the availability of enormous amount of data deriving from a variety of sources (maps, published books and articles, individual databases, electronic journals, disperse information on the web) but being all pertinent to environmental issues, a unified conceptual and technological system is imperative for providing equal opportunities to education, research institutes involved in the study, preservation and management of natural resources, universities (faculties of geology and geoenvironment, geophysics, engineering and architecture, agriculture, telematics and innovative learning strategies) along with private companies (like consultants, experts in various environmental disciplines, computing applications and geoinformation) cooperate on establishing an innovative way of e-learning, specifically orientated to the needs of earth sciences . We will also try to provide a complete and practically feasible solutions with regard to innovative training programs based on the use of Web-GIS technologies. A series of e-courses based on GIS tools on the topics of geology, geomorphology, agriculture, etc. are developed and available online. The courses are designed either to act as a supplementary tool for traditional classes or to act as autonomous courses for open learning.

The end users of the system are teachers and students at all educational levels (under/post graduate), individual researchers and educational institutes. Fortified case studies would be about teaching, learning, studying earth sciences via a web-based GIS platform and that will include various modules on the topics of natural and cultural resource management, applied geomorphology and other earth sciences are taught with online educational methods. The modules are designed in such a way so as to constitute both distance learning material and supplementary material for traditional courses.

The first module, the training on Web-GIS, is a pre-requisite for all the participants of the rest modules. The users can download (or directly view) the theory about each topic. Extended text accompanied by images, maps and other necessary material (bibliography, etc.) is available. Having read the theory, the user/student will follow the practical, self-paced exercises on the taught topic while using the Web-GIS. The user undertakes the practical on line, then saves the files (GIS Information layers) to their personal space and informs the tutor that he is finished with the specific practical. The tutor afterwards checks the files and sends feedback (corrections on GIS files).

All the modules involve both a theoretical and a practical part. The practical are utterly associated to the web-based GIS platform. The Web-GIS facilities provide a user-friendly environment even for a novel user. After taken the introductory course, each student is capable of handling any type of geographical data with Web-GIS tools.

The contact details of the tutor will be online, one per course, so as the users/students will be able to contact him/ her in case of queries, etc. The contact details of the course's participants will also be available on the web, so as students undertaking the same course could contact each other.

The proliferation of web-GIS in e-learning will provide a set of opportunities for the protection and understanding of natural/cultural heritage places and be used by various people from different disciplines.

Within the scope of this application, Bodrum region in Turkey is taken as an example and developed some GIS layers including Forestry, Contour lines, Slope, Coastline, Soil type, and Populated areas as shown Figure-1. MapInfo program is used while developing the application.

Integration of heterogeneous GIS will be much easier by utilizing GML technology. In Figure-2, a typical integration model used in the project is shown. There are four heterogeneous GIS databases in the model system. The updated data (this can be geographic coordinates or non-graphical or tabular data in GIS databases) can easiliy be shared by sending it in GML tagged documents to each other.

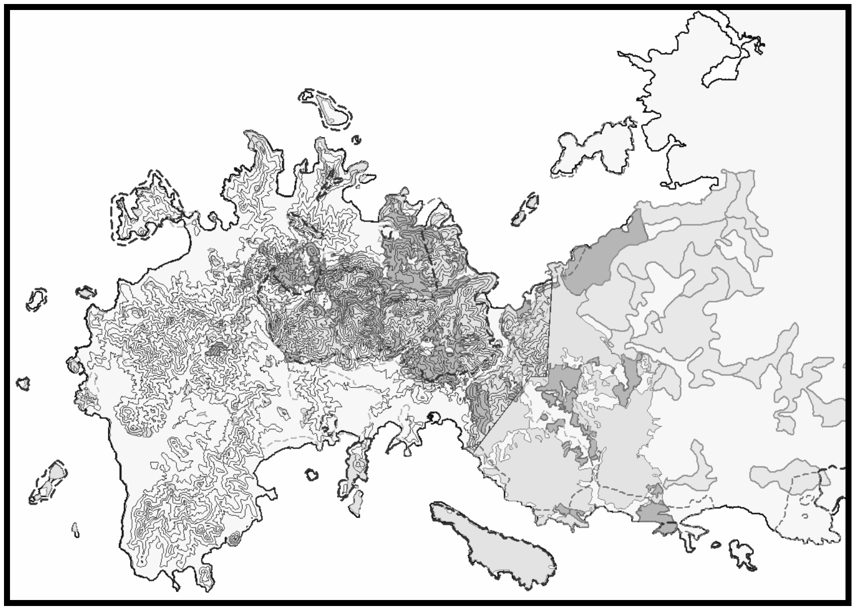

Fig. 1. Geo data including coastline, forestry, vegetation and contour lines layers produced in the project.

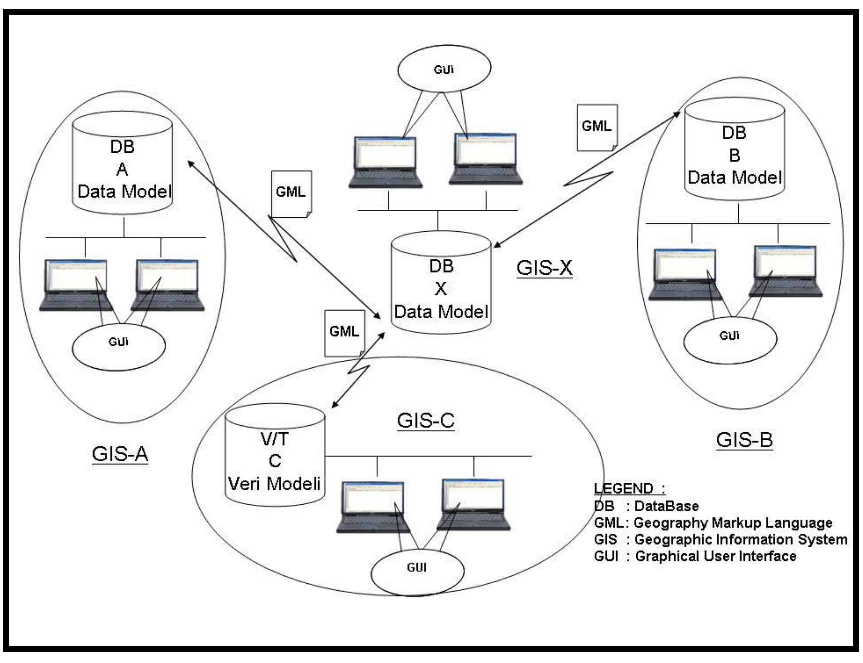

Fig. 2. Integration of heterogeneous GIS by means of exchanging geo referenced data in GML.

4. CONCLUSION AND FUTURE WORKS

GML is a powerful way to look at spatial information using XML encoding. It promises, however, much more than a mere encoding standard. The inherent transformability and accessibility of GML will open a whole new domain in geo-spatial information management.

The further development on this subject can be Semantic Web based integration of GIS's. Semantic web is considered as the future of the classical web technology. Nowadays, there are some considerable works ongoing under the W3C umberalla on this issue. But, this requires an additional workload such as developing geographic ontologies, mapping mechanism and moreover inference mechanisms. So that, machines (in this case, program software on the computers) will understand the data and exchange them between each other. The expected result from semantic web based geographic information systems is to make queries or infer the geo referenced data from heteregoneous databases. Such a simple query which is " Provide me all harbour coordinates on the Mediterrenean coast." will return not only the coordinates of the harbours, but also information about their capacities and connected ships.

REFERENCES

[1] Lake, R., 1996. Information communities support information sharing. GIS World 9 (2), 72–73.

[2] Lake, R., 2001. Enabling the geo-spatial web. GeoSpatial Solutions July2001 http://www.geospatial-online.com/geospatialsolutions /article/articleDetail.jspid8156.

[3] Lake R., 2004, "The application of geography markup language (GML) to the geological sciences", Computers and Geosciences 31 (2205) 1081-1094.

[4] Badard T., Richard D., 2001, "Using XML for the exchange of updating information between geographical information systems", Computers, Environment and Urban Systems 25 (2001) 17-31.

[5] Raper, J., 1999. GIS Without Borders. Geoeurope 3 (8), 22–23.

[6] Thompson, H.S., Beech, D., Maloney M., Mendelsohn N., 2004. XML schema part 1: structures, second ed. W3C Recommendation Document, October 2004 http://www.w3. Org/tr/xmlschema-1/.

[7] Maquire D., Goodchild M., Rhind D.W.; An Overview and Definition of GIS, Longman, 1991.

[8] World Wide Web Consortium (W3C), Extensible Markup Language, 1998, http://www.w3c.org/TR/1998/REL-xml-11980210.html

[9] OGC, OpenGIS Consortium, Geography Markup Language (GML V.2.1.1), 2002, http://www.opengis.net/gml/02-009/GML2-11.html.

[10] Galdos Systems, http://www.galdosinc.com/reasons_for_gml.htm, 2001.