# OGC's GML 2.0

## A New Wave of Open Geoprocessing on the Web

**In the last three decades, dozens of data formats have been developed for digital encoding of geographic information. Most were proprietary to a particular vendor's geoprocessing software. A few were open standards created to enable inter-vendor communication. Ultimately no scheme for interoperability based on a common format was widely successful. Because no single format could be used by all geoprocessing systems, OGC was founded in 1994 to develop a new interoperability approach based not on common formats but on open, common software interfaces. OGC's OpenGIS interface specifications have become a successful foundation for interoperable geoprocessing. But through a unique marriage of the World Wide Web Consortium's XML (Extensible Markup Language) work and work done in OGC's Specification and Interoperability programs, a new common geodata format known as the Geography Markup Language (GML) has become a reality. Let's see how and why!**

by Mark Reichardt

### From HTML to XML

The Web was originally built on the HTML (Hypertext Markup Language) standard for encoding/decoding web site page descriptions that include simple images and formatted text. XML goes well beyond HTML in capability, specifying rules for encoding standards such as HTML. Thus XML enables a wide range and limitless number of data presentation schemes. Because recent versions of web browsers can interpret XML-encoded data and instructions, these browsers provide an open platform for
**1)** receiving data and
**2)** receiving and executing instructions for parsing, displaying and manipulating that data.

The data and instructions are in separate "packages," so XML enables Web developers to clearly separate content from presentation.

### Separating Content (Geodata) from Presentation (Graphics and Maps)

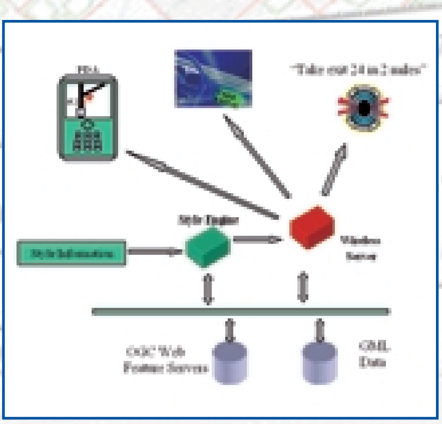
It is important to draw clear distinctions between geographic data and graphic interpretations of that data, such as maps. GML captures information about the properties and geometry of the objects that populate the world about us. How we symbolize these on a map, what colors or line weights we use, for example, is something quite different. If presentation is kept separate from content, GML data can be used for many different purposes on many different devices.

### From XML to GML

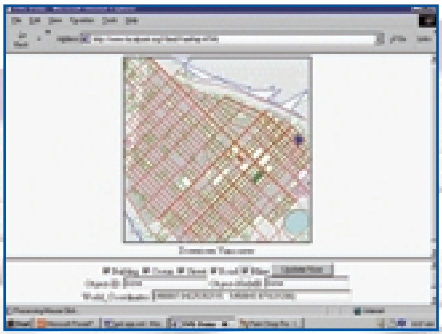
Members of OGC saw the potential: any web browser that supports XML might thus receive both geodata and the application-specific code needed to display and manipulate that data. They agreed to create the "Geographic Markup Language" (GML), an XML extension for encoding the transport and storage of geographic information, which includes geometry and properties of geographic features. GML incorporates the definitions of geometry and properties as set forth in OGC's earlier OpenGIS Simple Features Specification. For a more technical explanation of GML, see http://www.galdos-inc.com/index2.htm. OGC members from Australia, Canada, Europe, Japan and the US worked together to develop the GML specification. Development was rapid. The basic outline of GML and a Request for Comment were developed in the first (1999) phase of OGC's Web Mapping

*"Geographic Markup Language" (GML) is an XML extension for encoding the transport and storage of geographic information, which includes geometry and properties of geographic features.*

Testbed. GML 1.0, based on XML 1,0, was approved by OGC in May 2000. Following additional OGC Test Beds in the summer and fall of 2000, GML 2.0 was developed and passed as the successor to the GML 1.0 recommendation paper in February 2001. GML 2.0 was then passed in March as an OGC Adopted Specification. At the OGC meetings in Liege, Belgium in April 2001, Consortium members began work on GML Version 3.0. GML 2.0 is an "implementation specification," a working engineering specification supporting serious application development by anyone who wants to write software that encodes or decodes



*GML data available online can be interpreted by different "style engines" for various kinds of presentation - maps, written travel directions, or synthesized audio travel directions*



*GML is designed for "web-friendly geography". The map image displayed here on an ordinary web browser is not a jpeg or other image. Each feature in the map is described in an XML text file using "GML", OGC's standard way of describing features in XML. All recent versions of web browsers can interpret XML, so users need no other software to display and interact with GML-encoded geodata.*

spatial information (and processing instructions) in XML. See http://www.opengis.net/gml/01-029/GML2.html.

## Limitations of GML

GML is not intended to solve all geoprocessing interoperability problems. For one thing, GML, like all XML, is text, thus it requires greater bandwidth when transmitting spatial information over the Internet than more traditional binary formats. Many applications of course do not require high bandwidth, and the cost of obtaining higher bandwidth connections is falling with each passing day. In addition GML data, like all XML data, is highly compressible. Using common compression utilities such as zip GML can be compressed by a factor of 5 to 7. With XML specific compression techniques such as XMIL, GML data can be compressed by a factor of 15-20. Compressed XML data is then often smaller than the equivalent compressed binary data.

GML provides users the ability to create application schema to model their data. These schema are then used as the templates for the actual data itself. This requires that data providers make these schemes available to their users and that they do so on a continuous basis. GML also does not in itself solve problems of semantic interoperability. One data provider may decide to create a GML schema with road features while another provider may use street features for essentially the same entities. GML data was designed for transport over the Internet using the HTTP protocol. It is not restricted to HTTP, however, and can just was easily be transported using traditional distributed computing platforms such as

## GML 2.0 Widely Accepted

Even though GML 2.0 is relatively new as an adopted specification, it has already been accepted widely by commercial and government organizations. This is a testament to OGC's consensus process:

- Mr. Peter Woodsford of Laser-Scan notes "We have had a good initial experience with GML2.0, both for web mapping and for designing very large heterogeneous data maintenance architectures. We expect to see use of GML 2.0 take off rapidly, and to see GML become more capable in handling structured geospatial information, with extensions to cover 3D and temporal phenomena."
- The United Kingdom's Ordnance Survey has shown great interest in GML and notes on their website: "Following extensive consultation we can confirm that it is our intention to produce all DNF data in GML format."
- The US Census Bureau is considering using it to encode and distribute the Bureau's TIGER files.
- The Netherlands Society for Earth Observation and Geo-informatics (KvAG) has organized a "GML Relay" in June to explore the movement of GML data through several vendor's software implementations.

For the first time, spatial information has a truly public encoding standard positioned for wide adoption.

DCOM, or CORBA or more recent ones like Java RMI and SOAP.

## Benefits of GML

OpenGIS services that can interoperate using GML over the Internet will allow data providers to:

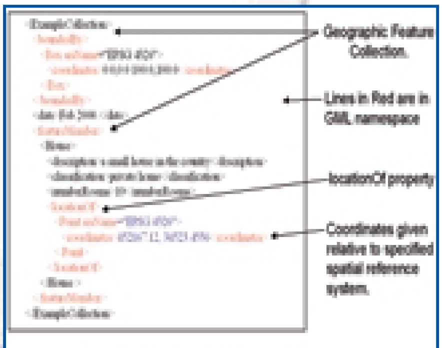
- Distribute vector-based data in GML using Web Feature Servers (WFS) instead of shipping paper or CD/ROMs in proprietary data formats, or transmitting large files of data in propriety formats over the Web.
- Support web-based transactions between WFS clients and servers. This enables, for example, live updates from the data providers' partners who maintain the data. Government organizations can improve data sharing for critical operations like disaster management and law enforcement.
- Buy Standards-Based COTS (SCOTS) products instead of building home-grown applications so that data providers will have "best of breed" components that "plug and play" over standard interfaces. This reduces product life cycle costs and makes it easier to insert new technologies.
- Plan their information infrastructure future with considerable certainty.
- Plan and budget geospatial services that will be used across the range of information environments that support their workflow, from mobile wireless to wired web browser to geospatial database server.

## Geospatial WWW

On the user side, consumers, citizens, and business and government users will benefit from the improved spatial content in information they get via the Web, as implementation of GML in commercial products allows large data set managers to buy more capable and less expensive tools, to share data, and to deliver data and services to much larger markets. GML will enable the development of a geospatial world wide web in which common data elements are shared between many data providers over the Internet, thus enabling geospatial data to be locally developed and maintained, and globally integrated. Such a geospatial world wide web is increasingly believed as essential to the emerging world of location-based services.

## OGC's Interoperability Program Provides for GML Advancement

OGC's Interoperability Program (IP) will be the catalyst for further advancements by providing a coordinated set of engineering testbed and demonstration activities for OGC members to address new GML requirements. The activities build upon the results




A simple GML fragment

*GML documents adhere to the general syntax of XML documents. GML "namespace" is a set of standard terms (shown in red) that have been approved in OGC's consensus process.*
*(Diagram courtesy of Galdos,Inc.)*

**The reasons for creating GML include:**

**1)** GML rests securely on a widely adopted public standard, XML

**2)** XML is easy to transform. Using XSLT or almost any other programming language (VB, VBScript, Java, C++, Javascript) programmers can readily perform transformations such as data visualization, coordinate transforms, spatial queries, and geospatial generalization. Other OpenGIS standard interfaces provide a GML-consistent open architecture for integrating "back-end" geo-processing with GML-based services.

**3)** GML holds promise as an important means of storing geographic information. XML's XLink and XPointer provide a means for building complex and distributed geographic data sets in which data developed for one purpose can be readily integrated with data developed for another. For example, a parcel boundary in one database may be shared as part of the boundary of a utility corridor in another, or a bus route in a database might refer to the road segments used by that route, the street segments being stored in a separate database.

**4)** GML is based on a common software model of geography (the OpenGIS Abstract Specification) which has been developed and accepted by a majority of GIS vendors in the world.

**5)** XML provides means to verify data integrity.

**6)** Because XML is Unicode, any XML document (or GML document) can be read and edited using a simple text editor.

**7)** Since XML has been applied to the encoding of many kinds of non-geographic information, it is easy to integrate GML data with most other kinds of data including text descriptions, pictures, video, and financial data.

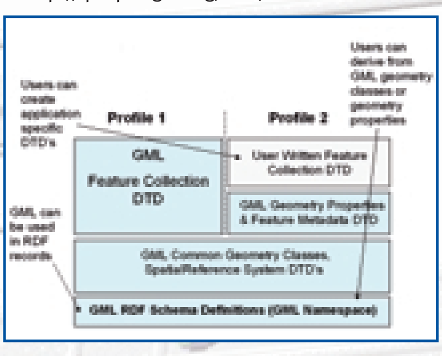
**8)** GML is a logical way to package spatial data for location-based services delivered by mobile Internet devices.

of previous OGC testbeds and pilot projects as well as on the work of the OGC Technical Committee and other standards organizations. By the end of 2001, OGC hopes to have completed GML Version 3, which will include at least some of following features, as well as others:
- GML encoding extensions for topology, temporal, coverages (including images), spatial locators, points of interest, units of measure and 3D geometry.
- GML spatial locator will support the "OpenGIS Spatial Locator", a geocoding services specification initiated in the Geospatial Fusion Services as part of the IP2000 Web Mapping Testbed Phase 2.
- A GML Point of Interest / Area of Interest standard: POI/AOI is a Spatial Locator, plus a direction, plus some digital content. Example: a point on a street corner in Boston, a compass heading, a digital photo of the shoe store on the opposite street corner, and a movie clip of crowds of people crossing the street. POI/AOI could encode the output of a Cell Phone / GPS / Laser Range Finder / Digital Compass / Digital Camera device held by a Census worker identifying housing units.

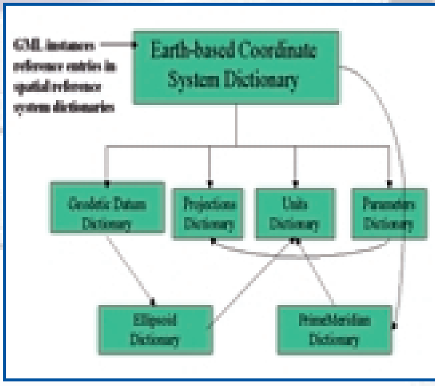
## OGC Interoperability Initiatives

In the next year, OGC Interoperability Initiatives will also address:
- Interfaces that provide open access to specific Internet-accessible services and also define a general geoprocessing service interface model.
- Interoperable service chaining (common expression and execution) and service metadata extensions for complex models (e.g., science models) that have spatial components.
- A new technical standards approach to overcoming the problem of semantic differences in geodata and associated metadata.
- Open communication among systems that offer image exploitation capabilities such as feature extraction and change detection.
- Common interfaces (and an XML-based Sensor Markup Language) for managing sensor information and metadata.

Additional information on IP sponsorship, participation, and requirements can be found at http://ip.opengis.org/ows/index.html




*A Document Type Definition (DTD) specifies the structure of an XML document in such a way that a validating parser can verify that a given document complies with this DTD. GML is specified by such a DTD. GML2 uses XML Schema, a more flexible integrity mechanism than the DTD. (Diagram courtesy of Galdos,Inc.)*




*In the old standalone model, every geoprocessing system needed to store its own earth-based coordinate system dictionaries. In the new web-based computing model enabled by OGC's OpenGIS standards, definitive versions of these dictionaries are available online to be used by any computing process that needs them. (Diagram courtesy of Galdos,Inc.)*

## The Power of GML Expands Through OGC Location Based Services

LBS providers will be major suppliers of spatial information and services, and they will benefit from GML just as other data providers will benefit. "Location Based Services" (LBS) refers to Web-based services specifically for users of location-aware, Internet-connected devices. OGC is working to establish OpenGIS standards as an integral part of the infrastructure for LBS. OGC's OpenLS Initiative involves aggressive outreach to the other standards organizations working in this space, including Location Interoperability Forum (LIF), Internet Engineering Task Force (IETF) and ISO/TC 204. By educating participants in these LBS standards efforts, OGC seeks to ensure that the standards they build will be consistent with OGC's OpenGIS Specifications. The OpenLS Initiative will include OGC-organized testbeds and demonstration projects. GML will be an important element of the expansion of spatial services to the wireless community. Mobile computing in Europe and Asia is well ahead of mobile computing in the US, so leadership by LBS providers and stakeholders in Europe will powerfully determine whether or not LBS becomes well integrated with the "deeper" Global Spatial Data Infrastructure.

*Mark Reichardt (mreichardt@opengis.org) is OGC's Director, Marketing and Public Sector Programs. He directs OGC's corporate communications and public outreach efforts. He also works with government agencies to help them understand how they might benefit from OGC membership, and once they have joined, he advises them on ways to get the greatest benefit from their membership.*