B04.3

GI data sharing for e-Government: Using GML to make the vision a reality

Steven Feldman, Managing Director, GDC and Eddie Curtis, Chief Technical Officer of Snowflake Software

Introduction

The UK Government has set ambitious targets for both Central and Local Government to deliver more efficient, responsive and accessible services to citizens through the implementation of Joined Up and Joint Government. Geographic Information (GI) has been recognised as a fundamental enabler in achieving these objectives.

Interoperability, the exchange of information between different computer systems, is the key to delivering Joined-up and Joint Government between Districts, Counties, Health and Police Services and a myriad of other agencies. The e-Government Interoperability Framework (e-GIF) sets out mandatory policy and standards for interoperability across the public sector and defines the architecture for joined-up and web-enabled government. These standards include XML (eXtensible Markup Language) and, for geographical information, GML (Geography Markup Language).

This paper seeks to explore the use of GML in supporting GI data sharing and exchange as part of e-Government partnerships and multi agency rapid responses.

Current Drivers for GI Data Sharing

Local e-Government Partnerships

The implementation of e-Government within the UK envisages a transformation in service delivery by central and local government. Challenging targets have been set for the introduction of e-Services by 2005 which many organisations are on target to meet either partially or fully. Two of the criteria for assessing an e-Service are that the service is delivered "Seamlessly" and "Jointly" where appropriate by local and regional partnerships.

The provision of public services within the UK is often multi-tiered with different local, regional and central government bodies providing services to citizens. A customer can face a baffling search to identify which body is responsible for the delivery of a service for example:

- Education and Highways are the responsibility of Counties in most areas but not where a Unitary authority has responsibility for all services.
- Services provided by Health and Emergency Services

Many services are not constrained by local government boundaries, when seeking a service, customers should not be concerned about which Local Government body operates that service.

An emerging element in Local e-Government service provision is the formation of partnerships between Counties, Districts, Fire Police and Health Services, Environment Agency and related bodies such as English Heritage, English Nature and the Countryside Agency. These partnerships are faced with several options for providing joint information:

1. Live connection to interoperable mapping services from each partner using either OGC Web Map Server or Web Feature Server Specifications.
2. Exchange of data between partners using either proprietary or open formats.
3. The creation and maintenance of an aggregation server holding the collective data resources of the partnership.

The role of GML and the advantages of each option are considered below.

Multi Agency Emergency Response

In a different context the ability of central and local governmental agencies to respond to major events, whether they are natural catastrophes (floods, earthquakes or storms) or man made (terrorism, accidents or industrial pollution) is dependent upon access to complete and up to date geographic information from a variety of providers.

Currently the US Government is actively promoting GI data sharing and interoperability between different agencies systems as part of its Homelands Security initiative. The Critical Infrastructure Protection Initiative (CIPI) is a programme of work funded by the US federal government and coordinated through the OpenGIS Consortium (OGC) that is working to build the necessary standards based technologies to facilitate a multi-agency response using off-the-shelf technologies. This work is being carried out by both North American and European companies, such as Snowflake Software (UK) and Ionic (Belgium).

What is GML?

Geography Markup Language (GML) is a standard published by the Open GIS Consortium (OGC) to allow the exchange of geographical information between computer systems. GML is often referred to as a data format, however, this is not entirely accurate as GML can be better thought of as a family of formats which share common elements.

To start at the beginning, GML is an extension of XML (eXtensible Markup Language). XML is a standard defined by the W3C (World Wide Web Consortium). XML defines some basic constructs for forming data formats such as the idea of a data element and basic types like strings, numbers and dates. Most importantly, XML defines a standard way of adding extensions to XML to make a more specific format; this mechanism for defining formats is called an XML Schema.

The OGC have used XML Schema to add some geographical concepts to XML. For example, GML defines the concept of a geographical feature, line, point and area geometry. These extensions are specified in the XML Schema files published by the OGC.

However, these are still abstract definitions, and so GML needs to be extended further to express the elements relevant to a particular dataset. Whilst GML defines a feature, a particular format might extend this to define a feature called “Road”. The particular format might specify that a “Road” has a geometry attribute that is a line. A particular format will be specified in one or more XML Schema files (“.xsd” files) known as the application schema.

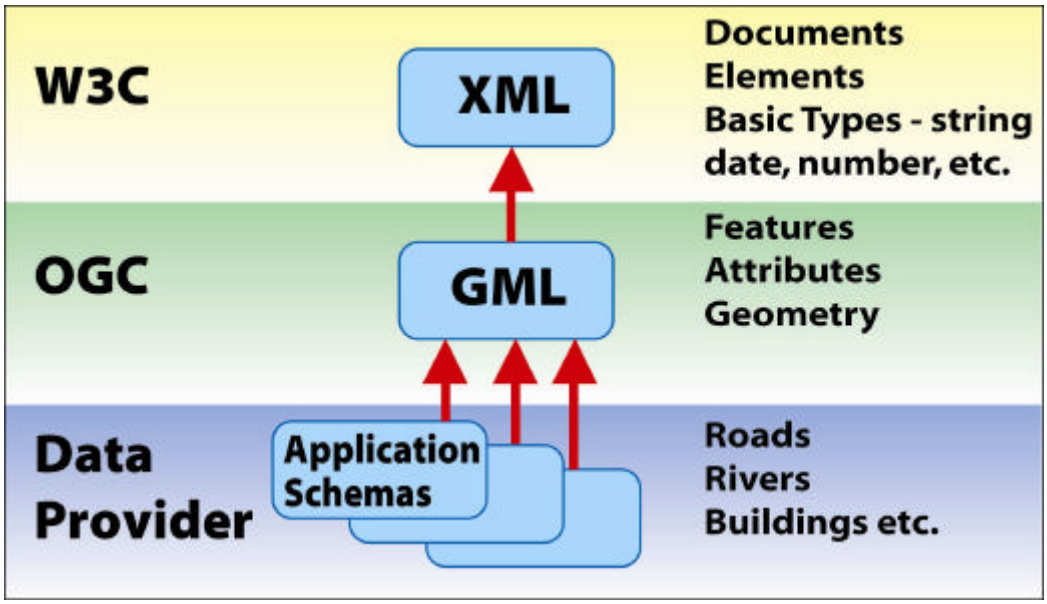

Figure 1

Why is GML different from other GIS formats?

A key element of GML is that the data specification (or schema) itself is machine readable.

Within the GIS world the conventional approach to handling data formats is to take the format specification and write software to translate data in that format to some format understood by each system being used. Each time a new format is specified a new format translator has to be written. End users must wait for their system supplier to provide them with a suitable translator before they can use the data. Sharing of geographical information between two organisations can only take place after both organisations' system suppliers have provided support for the necessary format.

XML Schema provides the opportunity for a radical shift in the way software tools are provided. By writing tools that read XML Schema it is possible for a software supplier to provide a single general purpose tool which will support the whole family of GML formats.

For example, GO Loader, from Snowflake Software, can automatically convert a new GML transfer format specification into a table model in an Oracle database. It can then load data into the database, thus making the data available to the GIS tools and other applications connected to that database without the need to write any new software each time the Schema changes.

Adopting a new GML format is quick, cheap, and low risk because schema parsing tools eliminate software development from the process. This creates the opportunity for organisations that have never published or shared GI before to do so. They are no longer dependent on software suppliers to create software specific to their data.

Options for GI Data Sharing

OGC Web Feature Server

One way to supply data in GML is through an OGC Web Feature Server. A WFS supplies collections of geographical features in GML format in response to queries in a standard structure from applications across the internet.

The main advantage of this approach (for the users) is that the users of the data do not need to have any data management system in place, all of the data management responsibility is placed on the data supplier.

Additionally a user is able to access data simultaneously from multiple data providers within the same work session. For example an emergency planner could view and analyse data provided by the Fire and Police services in real time.

Currently there are no full scale deployments of WFS within the UK public sector, however it is anticipated that WFS will become the preferred means of sharing data where the recipient does not make intensive use of the data or only wishes to access a limited geographic subset of the data.

OGC Web Map Server

A simpler approach is to use an OGC Web Map Server to generate map images which can be displayed and interrogated within compliant client software applications. Attribute information is passed to the client using HTTP, but the geographic objects are not available for incorporation into spatial analysis such as buffer queries.

The images from a WMS can be used in compliant clients as a geo-referenced raster backdrop to an organisation's own vector data. Examples of compliant client software are MapInfo Professional 7.5 and ESRI's Arc Explorer, desktop GIS capable of integrating local vector data with WMS images and Intergraph's web based WMS Viewer. If transparent images are supported by the WMS, the client can be configured to render superimposed images from multiple data providers.

WMS will be sufficient for applications where data from remote providers is incorporated into a web site for visualisation and simple queries and high level spatial analysis is not required, for example a Local Government Web Portal.

WMS server technology is starting to be deployed, with most major GIS vendors offering a WMS compliant implementation of their web mapping technology.

Data Exchange and Aggregation

There are many situations where both the data supplier and the data users already have their own data management systems in place. For example, where two organisations are sharing data at least one may have a Spatial Data Warehouse, such as GDC's GeoStore, providing management of the data holding and access to the data for desktop and web based applications. In these cases it often makes more sense for both the supplier and user to maintain synchronised copies of the data. The user's data holding can then be updated using GML files provided by the data supplier.

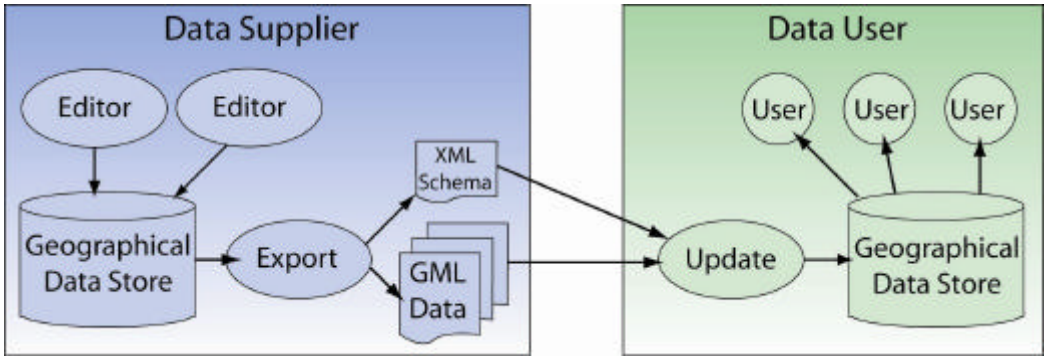

Figure 2

This approach has a number of advantages for both the data supplier and data user:

- Because the data user manages their own data holding they have control over data currency and when updates occur. Data from the supplier being used for reference will not change until it is convenient for the user, so users will not find their reference data changing half way through a task.

- The user is not forced to accept an update if they are not happy with the quality or accuracy of the new data. A data manager can apply quality checks to the incoming updates and allow the users to carry on using slightly older, but more suitable, data.
- The data is made available within the same technical environment as the users own data. If the user is managing their data using an application such as GeoStore to manage the data in Oracle, there is no need for desktop or web based applications to make a separate connection to a WMS to access partner data. All the data being used is available from a single source within the organisation.
- The user's routine business processes are not dependent on systems being run by outside organisations. If the data is supplied via a WMS then the user cannot access partner data if the data suppliers server or internet connection fails. Data exchange ensures that the user does not risk having their business processes interrupted by failures in systems which are outside of their control.
- The advantage to the data supplier of the users not directly accessing their systems is that they are not required to guarantee that their servers are always available (something which can require significant commitments of hardware and skills).

This method of data exchange can be used in two scenarios:

Firstly it can be used as means of distributing information to end users. Ordnance Survey has pioneered this method of data supply and maintenance in creating and updating its customers' holdings of MasterMap data. Most Local and Central Government users of OS data are implementing systems to receive MasterMap via GML. This model of data supply can easily scale down for smaller organisations to provide a means for them to distribute GI to other organisations in a controlled and secure way.

Secondly, this method of data exchange can be used to aggregate data from a number of organisations to allow them to provide a joint service. This allows organisations that collect and maintain data separately to pool their data resources for a shared service. The London Central Partnership LondonCABI pilot project is an example of this kind of data aggregation.

Making the vision a reality – The London Central Partnership

The LondonCABI Pilot Project

Established in January 2002, The London Central Partnership (LCP) comprises the London Boroughs of Camden and Islington, Royal Borough of Kensington & Chelsea, Westminster City Council and the Corporation of London. The partnership has received funding from the Office of the Deputy Prime Minister through the Local Government OnLine initiative and aims to demonstrate the benefits of collaborative working and joined up services to its citizens across London.

In March the partnership presented a business case to ODPM for a web-based GIS pilot project called London Council Information from A to B (LondonCABI). The project aims to identify good data management practice and develop a consistent geographical interface for the public which will deliver Council information to citizens for the whole partnership area through the Boroughs' own web-sites. The project will act as a proof of concept that the partners can work together sharing geographical data using open standards.

GDC has partnered with Oracle, Snowflake Software and MapInfo to develop a collaborative system for the partnership to collate their GI in a Spatial Data Warehouse using GML and to deliver this information to citizens and visitors via all of the Partnership web sites using a hybrid WMS implementation. The implementation is referred to as a 'hybrid' because, whilst the full OGC WMS specification is met, it also provides additional functionality above and beyond the specification.

Technical Overview

The pilot is based upon 4 main technology components:

- Snowflake GO Loader, which provides schema based upload of GML/XML files into Oracle.

- Oracle 9i Standard Edition with Locator provides a spatial data store for aggregated datasets.
- MapXtreme for Windows provides the mapping engine.
- GDC PlanAccess WMS to provide a toolset for rapid deployment of map visualisation, spatial queries and LPG searches

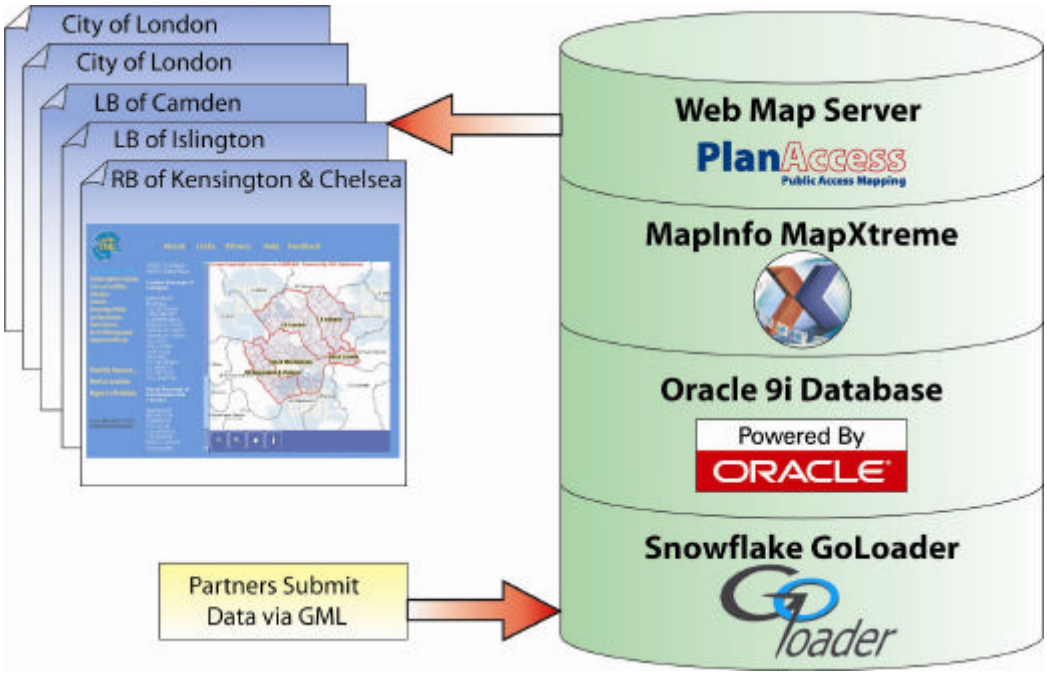

Figure 3

Each partner publishes data to their own GML application schema. GO Loader translates the different GML formats into a common database model and loads the data into the Oracle database.

The web mapping application (PlanAccess/MapXtreme) connects directly to the Oracle 9i database and provides a wide range of functionality which can be called and incorporated into the partners' web sites using either the WMS interface or by a series of URL calls directly accessing the broader range of functionality available within PlanAccess.

Partners are able to incorporate maps and spatial/LPG queries within their web sites that are based upon the aggregated spatial data within the Spatial Data Warehouse enabling participating authorities to provide a service which is truly seamless across all five boroughs.

The CABI portal can be viewed at http://cabi.graphdata.co.uk and the WMS can be accessed at http://cabi.graphdata.co.uk/webmapservice/planaccesswms.asp

Experience

At the time of writing this paper (July 2003) the LondonCABI pilot has just gone live with very positive feedback from all of the partners. The aim of producing a seamless service has been achieved. Whilst it is too early to fully assess the pilot, particularly from the end users' point of view, some guidelines for future GI partnership projects have been identified.

a. Partners experienced some difficulty in generating GML in the agreed schemas. GIS vendors need to provide integrated tools that can generate GML and validate against a Schema.

b. Partner systems will need to be upgraded to enable selection of changed only records to avoid re-supplying complete data sets each time that a record changes.

Summary

In summary, if implemented correctly, GML can join up Government by enabling the exchange of data between partners regardless of what format their geographic data is stored in. The increasing prevalence of Web Map Server technology will allow the dissemination of shared data to any partner, organisation or individual that needs it. In short GML can make the UK Governments vision of joined up e-Government a reality.

Steven Feldman: stevenf@graphdata.co.uk
www.graphdata.co.uk

Eddie Curtis: ecurtis@snowflakesoftware.co.uk
www.snowflakesoftware.co.uk