GML-Based Interoperable Geographical Databases

C Zhang, W Li,
M J Day, Z-R Peng

Many geographical databases have been developed for different programs and applications/ but data acquisition and data sharing are still a big problem because no interoperability exists among these different databases. This study presents a GML (Geography Markup Language) approach to build a geographical database in order to enable interoperability. As an open/ non-proprietary industry standard/ GML overcomes the problems of current CIS software proprietary data models and database structures. Compared with other standards/ such as the Geographic Oata File (GOF) and Spatial Oata Transfer Standard (SOTS)/ the GML approach has the advantage of enabling on-fine data exchange. GML holds promise in providing a standard way to share and use existing spatial data over the World Wide Web. A GML-based interoperable geographical database for the conservation of the Stone Forest Landscape is implemented as a case study. It shows how the public can access and use the GML-based spatial database through a user-friendly interface and that GML can deliver high quality vector data on the Web.

C Zhang, W Li, M J Day
Department of Geography
University of Wisconsin-Milwaukee
czhang@uwnl.edu
weidong6616@yahoo.com
mickday@uwm.edu

Z-R Peng
Department of Urban Planning
University of VVisconsin-Milwaukee
zpeng@uwm.edu

INTRODUCTION

With rapid developments in GIS (Geographic Information Systenl) and its applications, more and more geographical databases have been developed for different programs and applications, but data sharing and acquisition are still significant problems for the development of GIS applications. Not that data are not available, there is a huge amount of geographical data stored in different places and in different formats, but data reuse for new applications and data sharing are daunting tasks because of the heterogeneity of existing systems in terms of data modeling concepts, data encoding techniques and storage structures, etc. (Devogele *et al/ 1998).*

Currently, several commercial desktop GIS software systems dominate the geographical information (GI) industry, such as ESRI ArcInfo and ArcView, Smallworld GIS, Intergraph GeoMedia, MapInfo Professional and Clark Labs Idrisi. It is unlikely that all GIS applications will use the same software (Tarnoff, 1998). Different vendors have their own proprietary software designs, data models and database storage structures. Thus, geographical databases based on these designs cannot communicate without data conversion. In order to exchange information and share computational geo-database resources among heterogeneous systems, conversion tools have to be developed to transfer data from one format into another.

Furthermore, these diverse desktop GIS database structures make remote data exchange and sharing more difficult because of limited accessibility and requirement for data conversion.

The development of the World Wide Web creates a unique environment for sharing geospatial data. Users can use the World Wide Web to download data for viewing, analysis or manipulation. Many commercial Internet-based GIS programs, such as ESRI's MapObject IMS and ArcIMS, AutoDesk's MapGuide, Intergraph's Geomedia WebMap, MapInfo's MapXtreme, GE SmallWorld's Internet Application Server, and ER Mapper's Image Web Server, are developed to offer better tools for data sharing over the Web. But like the desktop GIS software, these Internet GIS programs also have the problems of proprietary software designs, data models and database storage structures. Sharing of data, facilitated by the advances in network technologies, is hampered by the incompatibility of the variety of data models and formats used at different sites (Choicki, 1999).

In addition, two other problems result directly from non-interoperability of databases. One is about data precision. This includes coordinate precision, errors of omission, missing or wrong attribute names, and incorrect topology, after data are converted from one format to another (Noronha, 2000). The other problem is that a lot of money and time have been wasted on data conversion or developing data conversion tools. Most investment by today's GIS users lie in three areas: data conversion, development of application specific extensions to general purpose GIS products, and the learning of applications of software and data to enhance productivity. Among these three areas data exchange and conversion account for a very high percentage (Siki, 1999).

Interoperability is *the ability* of a *system/ or components* of a *system, to provide information portability and inter-application cooperative process control* (Bordie, 1992). Two kinds of interoperability can be distinguished. For a program, data interoperability means the ability to utilize a range of data formats. For a data set, program interoperability means that it can be used by different types of programs (Laurini, 1998). An interoperable database refers to the data level interoperability. It can be used by different types of programs and applications. With interoperable databases users can request and integrate data easily no matter whether the databases are stored locally or remotely. The interoperability of data from heterogeneous sources is extremely important in the context of geographical applications, because there exist large amounts of spatial data of different geographical formats and there are increased demands for re-use of these existing spatial data.

How to realize the goal of data interoperability? There are two approaches to data interoperability-database integration and standardization (Devogele *et all 1998).* Database integration is the most sophisticated approach. A very basic approach is to provide users with a global catalogue of accessible information sources, where each source is described by associated metadata, including representation mode, scale, last update date, and data quality level, etc. (Stephan *et all* 1993; Uitermark, 1996). Current database integration has evident drawbacks related to lack of scalability, consistency and duplication (Devogele *et all* 1998). The second approach to interoperability is through standardization. The definition of standard data modeling and manipulation features provide a reference point which facilitates data exchange among heterogeneous systems (Devogele *et all* 1998).

In the past, several useful standards have been developed to facilitate data exchange. Among them, the Geographic Data File (GDF) and the Spatial Data Transfer Standard (SOTS) are widely used and accepted. GDF is specifically designed for spatial data exchange for Intelligent Transportation Systems (ITS). It defines a set of spatial features, attributes and relationships that are particularly relevant to ITS applications, and specifies a set of useful data structures and data formats. This makes it readily usable for off-line data exchange. SOTS is a general purpose standard that is flexible and adaptive (NIST, 1994). With anticipated extensions and refinements, SOTS was expected to become an important data format for ITS spatial data transfer or a neutral format for data archiving (Arctur *et all* 1998), however several barriers blocked the popularity of SDTS. These barriers include the complexity of SOTS, slowness in the development of practical SDTS profiles, restriction of each SOTS dataset to a single profile, lack of a clear definition of geospatial features in SOTS, and ambiguity in the means of specifying cardinality of relationships in a data model (Arctur *et all* 1998). Currently, both GDF and SOTS are not as widely used as originally anticipated. The creation of a new standard data exchange format-Geography Markup Language (GML)-represents another important step taken by the geospatial community towards data interoperability. The GML is *an XML grammar written in XML Schema for the modeling, transport, and storage of geographic information including both the spatial and non-spatial properties of geographic features* (OGC, 2003). It is developed as a Data Exchange Standards Interface by the Open GIS Consortium (OGC) to achieve data interoperability and reduce costly geographic data conversions between different systenls. In the spirit of aGC, interoperability is achieved by means of common specifications that programs and data must follow (Buehler and McKee, 1996). *aGC* takes a new spatial interoperability approach, which is not based on a common format, but based on open and common software interfaces. The interface specification largely eliminates the need for data format standards and costly batch data conversion. With the development of XML (Extensible Markup Language) by the World Wide Web Consortium (W3C), the creation of the Geography Markup Language '"mplementation Specification by OGe represents a significant step in the development of interoperable architectures for the use of spatial information between different applications. GML holds promise to support mapping from a wide variety of sources and enable sharing of geospatial data for on-line information exchange.

Unlike current proprietary comnlercial Internet GIS programs, the OpenGIS GML specification is a public open standard for coding and sharing spatial data. GML is a good alternative to expensive, proprietary web-based mapping solutions because:

- GML is an open source standard. Users can use it for free, but for other commercial Internet GIS programs there is a purchase cost to users. For example, ESRI ArcIMS Internet software is so expensive that many users cannot afford it, but they need to provide online spatial data services. GML is a good alternative for these users. With GML they can provide the online spatial data services without buying these proprietary software
- GML data are stored in text format, which is a universal format. Thus it is easy to integrate GML data into other data across a variety of platforms and devices
- As a standard data exchange format GML reduces the costly conversion processes among different format databases. Figure 1 illustrates that one data format conversion is necessary between any two different-format conventional databases for the exchange of spatial data, while with the help of GML only one conversion is needed among all these different databases
- Although GML specifications take the standardization approach to data interoperability as do GOF and SDTS, it goes further and supports interoperable solutions that geo-enable the Web. While GOF and SOTS are useful for off-line data exchange, GML is capable of facilitating real-time data sharing and exchange on the Web because it uses XML grammar which is widely supported on the Web. GML can enable an accessible Geo-Web (Lake, 2002; Peng and Tsou, 2003; Shekhar *et al,* 2001).

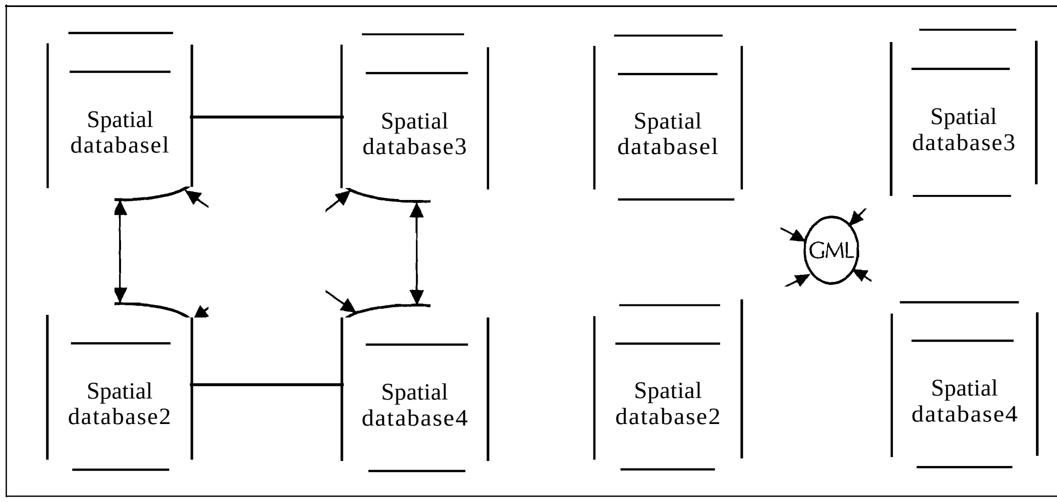

Figure 1. *CML reduces the costly conversion processes among different format databases.*

In addition, GML can deliver vector data over the Internet by styling the data into Scalable Vector Graphics (SVG) format. Most current Internet GIS programs deliver spatial data through transmission of raster images such as GIF and jPEG formats over the World Wide Web. There are several advantages for delivering SVG vector GIS data over the Web compared with raster GIS data:

- Compatibility. SVG also uses text-based XML format, which is compatible with other formats. It can be seamlessly integrated with current Web technologies, such as HTML, javaScript, jSP, ASP, jPEG and GIF
- Graphic quality. SVG format graphics are scalable and resolution-independent.- This kind of data can be scaled without loss of quality across different platforms and devices. But coarse raster images are low quality because of a low image resolution. Especially when users zoom in too many times, images become blurred and pixelated. Additionally, a raster image with high resolution usually has a larger file size since it needs to store information as finer pixels. The speed of delivering such large files over the Web becomes slow, so it is not practical to use high-resolution images for Internet GIS. The need for delivering high quality vector graphic maps over the Internet is becoming pressing as data availability and global sharing increases (Bertolotto and Egenhofer, 2001)
- SVG vector data can be accessed in a more interactive and dynamic way_ Some dynamic functions can be integrated into SVG documents so that the SVG graphics are animated on the Web. For example, a SVG graphic can interact with users by mouse over if a mouseover function is added in the SVG document. By combining SVG with other web technologies like HTML, JavaScript, jSP or ASP, a GML-based database can provide users with an extremely rich interactive graphic interface.

In general, GML-based databases have many advantages compared with other alternatives. Firstly GML-based databases can be easily shared and reused. They have no proprietary data models and database structures. Because of the proprietary software design, databases created by current commercial GIS software are difficult to share. To share data among such databases, many data conversion processes are necessary. Since GML-based databases are text format, they can be easily integrated with other format data across a variety of platforms. Secondly GML-based databases can be shared and exchanged online in real time, however databases based on other standards, such as GDF and SDTS, can only be shared and exchanged off-line. Although current Internet GIS programs can allow users to share spatial data online, they have the aforenlentioned proprietary data model and database structure problems. Thirdly, GML-based databases can let users exchange data at feature level, while current commercial Internet GIS programs cannot. For example, from a large GML-based database, users can query and download just one feature such as a specific road, while from other alternatives users must download the whole data set. Sharing and exchanging data at feature level in real time are especially important for emergency services, they can greatly reduce the time spent on data-acquisition processes. Fourthly, by styling GML data into SVG, GML-based databases can provide users with a more sophisticated interactive graphic interface and deliver higher quality graphic maps over the Web than most other online alternatives. Fifthly, GML is more flexible than other alternatives. It only defines a basic geographical feature schema and geometry schema, which are convenient for users to utilise. Based on these schemas users can define their own specific schemas for their spatial data documents.

It has been widely recognized that GML will play an inlportant role as a future Web data exchange standard (Clemens, 2002; Lake, 1999; Meneghello, 2001; Murray and Chow, 2002; Smith *et all* 2002). This paper will talk about what mechanisms of GML lead to data interoperability and provide a real application of a GML-based database by building a GML-based interoperable geographical database for the conservation of the Stone Forest Landscape in Lunan, China.

MECHANISMS OF GML FOR DATA INTEROPERABILITY

As mentioned previously, the GML specification is an important step taken by the geospatial community towards the vision of widespread spatial data interoperability. GML-based geographical databases can communicate with each other. The mechanisms of GML for data interoperability are given as the following:

- GML provides a common schema framework for encoding geo-spatial features. GML uses the W3C XML Schema Definition Language to define and constrain the contents of its XML documents. The GML v2.0 Specification provides two basic XML Schemas: the GML Feature Schema (feature.xsd) and GML Geometry Schema (geometry.xsd). Users can develop their own application schemas according to GML v2.0 Specification conformance requirements. GML schemas reconcile the need for standardization with the need for diversity by providing a standard means of extending the GML format. The direct consequence of applying schemas with GML is that it becomes possible for organizations to define formats to suit their needs and exchange geographic information without the need to involve software developers to create translators for that specific format. This has impacts on both the cost and risk of exchanging data (Curtis, 2003).
- While GML builds on XML Schema, it provides a more constrained model. GML is based on a common abstract model of geography (GGe Abstract Specification), which describes the world in terms of features. A geographic feature is an abstraction of a real world phenonlenon; it is a geographic feature if it is associated with a location relative to the Earth (OGC Abstract Specification, 2001). A feature has both simple properties and geometric properties. Simple properties refer to the usual name, type and value description. Geometric properties are composed of points, curve (/inestring) and surface (polygon). By looking at feature schemas and properties, one can readily compare features and integrate data.
- GML is based on an XML standard. XML is a universal format for structured documents and data on the Web. XML is easy to transform. Using XSLT or almost any other programming language (VB, VBScript, Java, C++, Javascript), users can transform XML from one form to another. By adhering to an open, non-proprietary standard, GML documents can be manipulated, transformed and presented in the same flexible way as XML contents.

- GML provides XLink and XPointer mechanisms as does XML. The linking mechanism of HTML (one web page linking to another), is one of the key foundations of the Web. GML goes further by providing a mechanism for linking multiple distributed resources into a complex association. As HTML is important to the Internet as a linked collection of web pages, GML can enable the development of a Geo-Web as a linked collection of geo-spatial features. Through XLink and Xpointer, different features and feature collections, which may be located remotely, can be associated together at the feature level (Peng, 2003). XLink and XPointer hold great promise for building complex and distributed geographic data sets (Lake, 1999). They make it possible to access and seamlessly integrate data from different departments, cities, states and countries.
- GML provides a means to transport geospatial data over the Web. With the help of Web Feature Server (WFS databases with different for transparently communicate with by being converted to GML-form the fly. Figure 2 illustrates that enables the Web by transporting a geospatial data with different fo GML. As XML is an important In transport technology, GML makes for real-time data access and trar Internet environment at feature the GML-marked geospatial transported, all the markup ele describe every spatial and n feature, geometry and spatial systems of the data are also tra the recipient (Peng, 2003).
- GML data are stored in plain t vendor-neutral, so informatian GML is not locked into a propriE format. Since GML is text-ba readily integrate geospatial d wide variety of non-spatial (

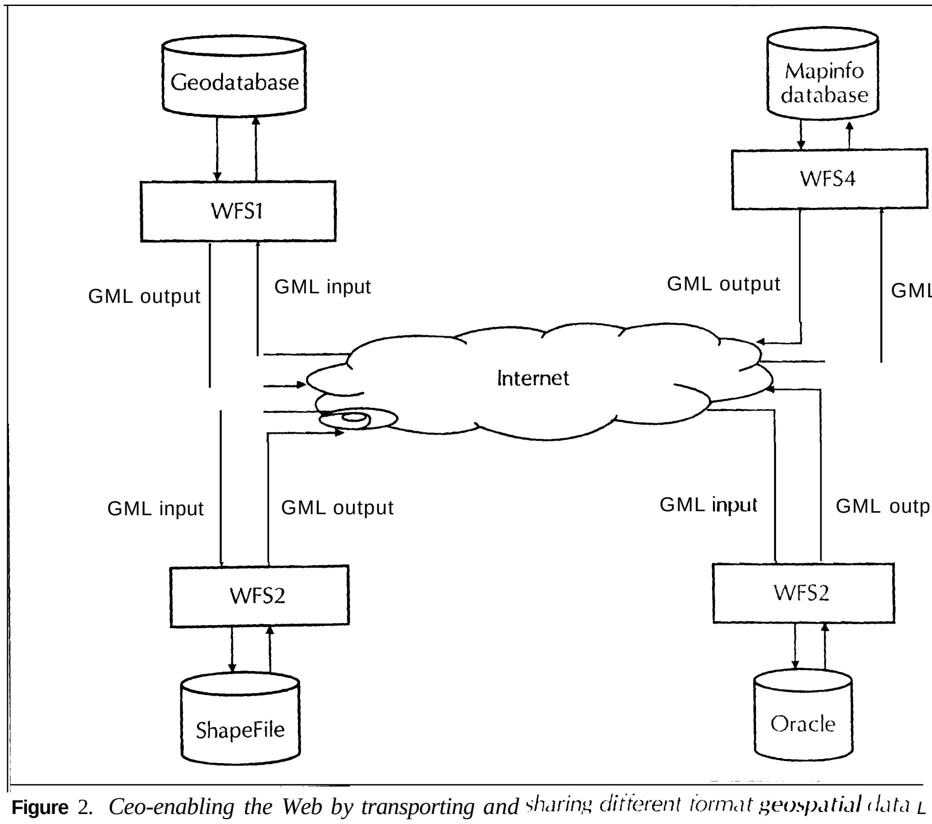

Figure 2. *Ceo-enabling the Web by transporting and sharing different format geospatial data L* including text, business transactions, graphics, audio, voice and more. This capability would greatly enhance the value and accessibility of geospatial information. For example, users can easily insert a map in a financial report, or vice versa (Peng and Tsou, 2003). In addition, as a text format, GML can be easily transnlitted across a variety of platforms over the Internet. Thus GML enables disparate systems to share information easily.

GML mechanisms allows users to build a large, global map that is stored and processed in a scalable and redundant distributed architecture. GML makes it possible that all spatial data across the world could be integrated into one map (Misund and Johnsen, 2003). The inherent transformability and accessibility of GML opens a new domain for the geo-community (Lake, 1999). The following section provides a case study of building such an interoperable geographical database using GML. Because of the above mechanisms of GML, this database can be easily shared and used by different programs.

A CASE STUDY-A GML-BASED INTEROPERABLE GEOGRAPI-IICAL DATABASE FOR CONSERVATION OF THE LUNAN STONE FOREST LANDSCAPE

Objectives

The first objective of this case study is to implenlent an interoperable database with GML for the conservation of the Lunan Stone Forest Landscape that allows for the database to be easily shared and re-used in the future. The second objective is to illustrate that GML plays an important role as a vector feature distribution format and that GML-based interoperable databases can serve better quality maps through delivering SVG vector data over the Internet. The third objective is to show that the GML-based interoperable database offers a user-friendly interface, thus the public can easily access and use the existing spatial data to perform GIS analyses. The fourth objective is to illustrate that the interoperable database can be accessed and queried at the feature level from the Web Feature Service (WFS) server over the Internet.

There are several reasons for building the GML-based interoperable geographical database:

- the database can be easily re-used in the future
- data developed on a local scale can be readily integrated into those on a regional or global scale in the future
- data developed for one application can be readily integrated with data developed for another application. Any other GML-based database can conlmunicate with this database
- the interoperable database can allow other organisations and application programs to easily share and integrate the data. It can also supply the public and the decision makers with data resources in real-time such as those required for the conservation of the Lunan Stone Forest Landscape.

Background

Lunan Stone Forest is a forest of intensively corroded limestone pinnacles up to some ten metres high. The Lunan Stone Forest is one of three unique landscapes in China, and is a very special type of karst landform in the world. It has a total area of 350 square km and is by far the largest area of pinnacle karst in the world (Figure 3). It is currently under consideration for UNESCO World Heritage designation and is presently attracting 1500 000 visitors each year (Song, 1997).

As a karst landscape, the Lunan Stone Forest is inherently fragile (Huntoon, 1992, 1993). The conservation of the Stone Forest Landscape is a very complex problem. The conservation of the landscape involves many factors, such as geomorphology, geology, soil, vegetation, hydrology, population, tourism and the economy. To effectively protect the landscape various data are required, which come from different government departments. An interoperable database can enable convenient sharing and re-use of these spatial data.

Figure 3. *A picture of the Stone Forest Landscape in Lunan, China (Joo, 2002).*

Spatial heterogeneity is the most fundamental characteristic of all landscapes, and scale multiplicity is inherent in spatial heterogeneity. Multi-scale analysis is imperative for understanding the structure, function and dynamics of landscapes. A GML-based database can easily integrate different scale data (Lake, 1999). Thus, a GML-based interoperable database will play an important role in a comprehensive study of landscape conservation problems.

The conservation program of the Stone Forest Landscape has created large quantities of data and involves a range of different departments, many that may need or already have their own different databases. Thus, data sharing between these departments is very necessary. Since GML-based databases are interoperable, such a database not only supplies a model for the departnlents involved to develop their own databases so as to better share data, but also provides a common database to share among them. Any other GML-based databases can also communicate with this database.

Public participation will play an important role in the conservation of the Stone Forest Landscape. A web-enabled geo-database can supply the public and the decision makers with the data resources in real-time. This will create for the public the conditions of attending data analyses and expressing their opinions for the conservation program through the Internet.

Database Construction

Figure 4 illustrates the procedure for building a GML-based interoperable geographical database. The first step is to collect and prepare the spatial data for the conservation of the Stone Forest Landscape. The data collected for this case study include a Stone Forest distribution map, a river map, a lake map, a village distribution map, a geology map and a geomorphology map. The format of these data is ArcView Shapefile.

The second step is to install and setup the WFS server to support the database. Currently there are several commercial WFS software programs available. For this research the

an important
Stone Forest
eo-database
he decision
in real-time.
he conditions
d expressing
tion program

e for building
geographical
o collect and
e conservation
pe. The data
nelude a Stone
r map, a lake
ap, a geology
p. The format
file.

l and setup the
base. Currently
WFS software
s research the

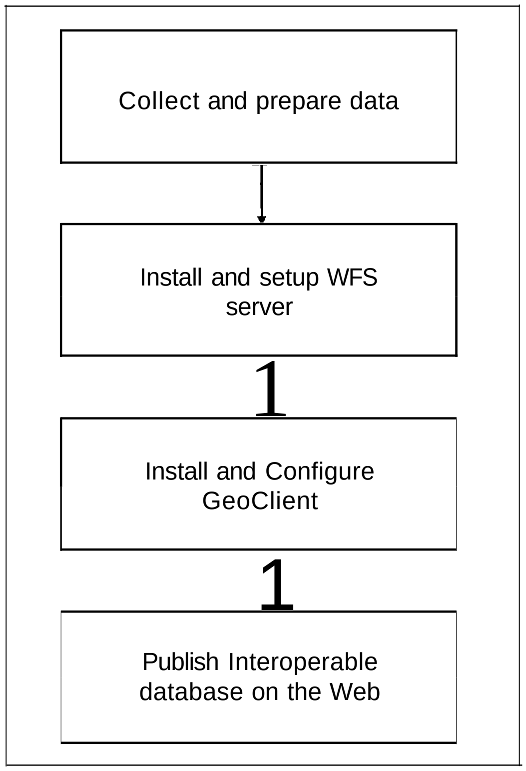

Figure 4. *Procedure of building a GML-based interoperable geographical database.*

GeoServerLite (http://www.mycgiserver.com/~amri/) was selected, which is a simple WFS server based on the Open GIS Consortium standard. The GeoServerLite is open-source software with a graphic client. All supporting software can be downloaded for free. GeoServerLite is written in the broadly available PHP scripting language and is based on the MySQL database. To setup GeoServerLite, it was necessary to first select and setup a HTTP Web Server. In this case study the Apache HTTP Web Server (http://httpd.apache.org/), currently a popular web server on the Internet, is used. To completely set up GeoServerLite, it was also necessary to install and setup the PHP scripting language environment and MySQL database.

The third step is to install and configure a client interface. The GeoClient (http://www.mycgiserver.com/-amri/) was selected, which is a graphic interface for accessing and querying GML data over the Internet. GeoClient is written in SVG and ECMAscript/JavaScript, and can run natively in a web browser if SVG support is available. Currently it needs support of the Adobe SVG Viewer Plug-in.

The final step is to build the interoperable database. Two methods are used in this study (Figure 5). The first is to use ToWKT, an extension of ArcView developed by the Geoclient project. ToWKT can export Shapefile data into a MySQL database in text format. It can also transform Shapefile data into the GML data format or SVG data format. The PHP-based GeoServerLite can be connected to the MySQL database. The feature level data required by users can be extracted from the MySQL database and then transferred into GML formaton the fly by the GeoServerLite. The GML format data are delivered out through the Internet to clients. On the client side, the data are further dynamically transferred to SVG maps by client browsers. This method can serve feature level data in real-time, and is easier to implement without changing existing databases. The second method is using the FME software plus some customized Visual Basic code to first convert Shapefile data into GML data. The GML data are then converted into SVG files with the help of the XSLT processor and style sheet. The GeoServerLite directly serves the SVG data to users. This method is more flexible but can not serve data at the feature level.

To browse SVG maps on the Internet, users need to download and install the Adobe SVG Viewer Plug-in. This is a free browser Plug-in (http://www.adobe.com/svg/). There are also a number of stand-alone SVG viewers available.

Results

Based on the above procedure, a GML-based interoperable database for the conservation of the Lunan Stone Forest Landscape was constructed. Figure 6 shows a user-friendly interface provided by the GeoClient for the GML-based interoperable database. It can be accessed and queried at the feature level in real time from the WFS server over the Internet. Such a GML-based database is interoperable and different programs on various platforms can use it remotely through the Internet. Other geo-databases that can accept GML formatted data can communicate with the GML-based database. Of course, the GML-based database can also be re-used easily in the future for other purposes.

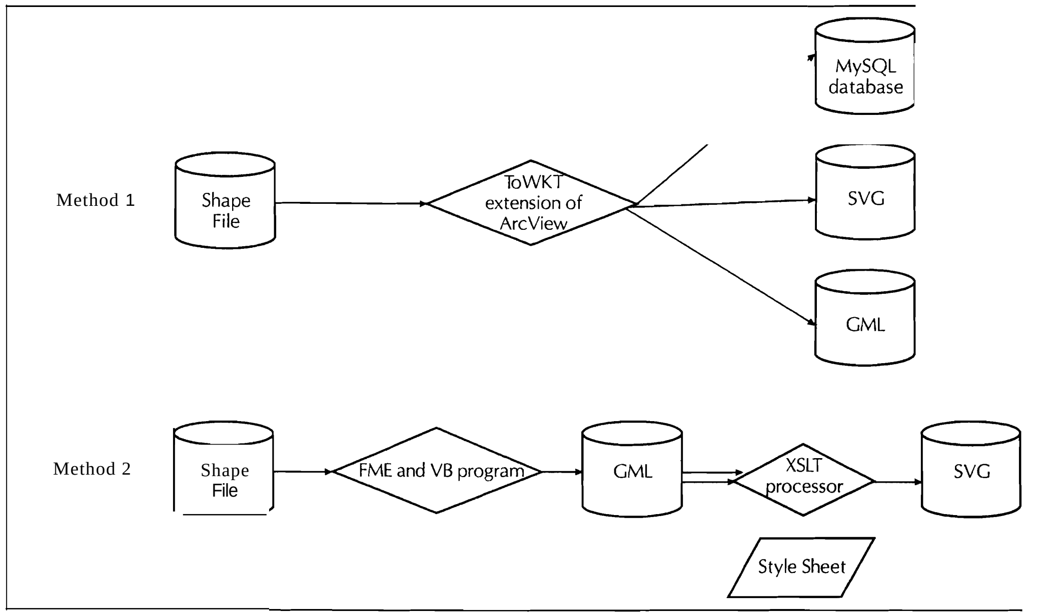

Figure 5. *The two methods for building an interoperable database.*

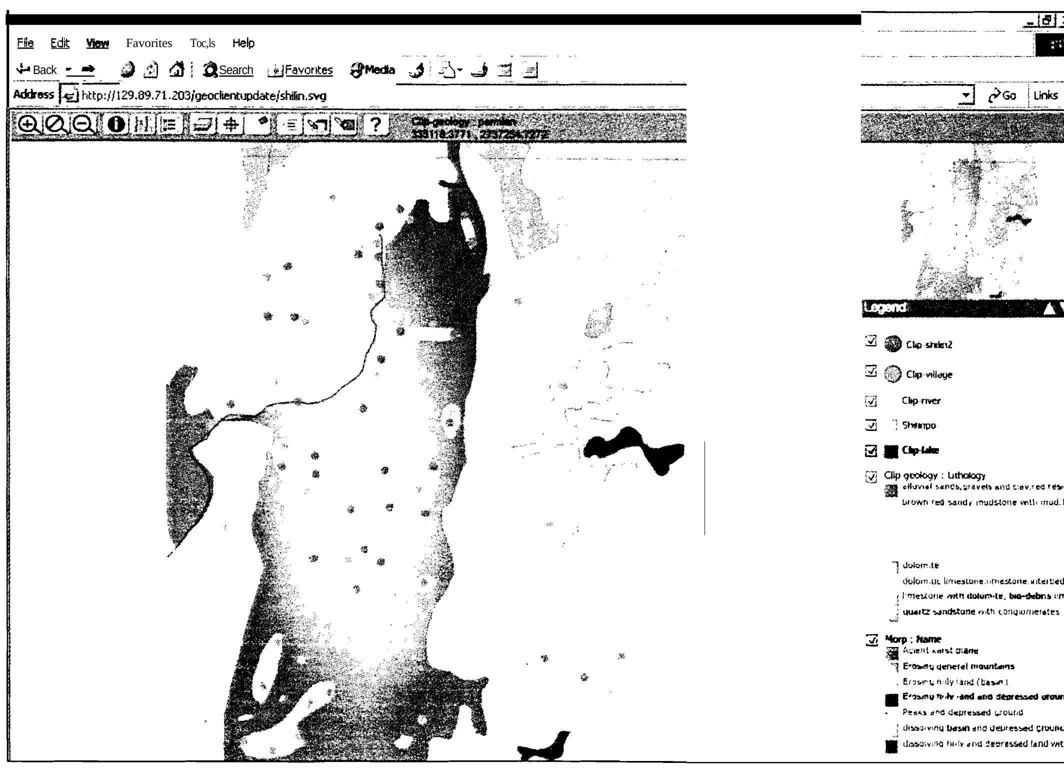

Figure 6. *Interface of the GML-based interoperable database.*

Currently the GeoClient software can only provide some basic GIS functions for the GML-based database on the Web. X and Y coordinates can be viewed by moving the mouse over a map. The map can be labeled automatically through the graphic interface (Figure 7). Users can zoom in, zoom out, pan and query the map. When users click on a feature the interface will display an attribute table for the feature (Figure 8). Because the development of the Geoclient software is only in its initial stages, other GIS analysis functions like buffer analysis, are not supported and need to be added in the near future.

The case study also shows that the GML-based interoperable database can serve scalable and high quality maps through delivering SVG vector data over the Internet. To make maps from GML data, the eXtensible Stylesheet Language Transformations (XSLT) processor was used to style the GML data into a SVG graphical format. As a vector format, SVG can let users zoom in on any portion of the GML data without any degradation in the quality of maps. The SVG vector maps can be printed with very high quality at any scales. No matter how many times they are zoomed in, a resolution-independent high quality map can always be provided (Figure 9). Naturally, the quality of information is still restricted by the original data, but the vector maps from the database never have the staircase (pixilated) effects users see when printing enlarged pixel-based GIF and JPEG image. So this has solved the blur problem of raster image maps when the scale is changed (especially amplified). Further, like XML, GML separates content from presentation. GML is only concerned with the content of the geographic data. How to present the map data is decided by users, which is an advantage of GML. This case study indicates that users can control how the data are displayed in a Web browser by changing the symbols of features using the Geoclient graphic interface. By delivering vector data

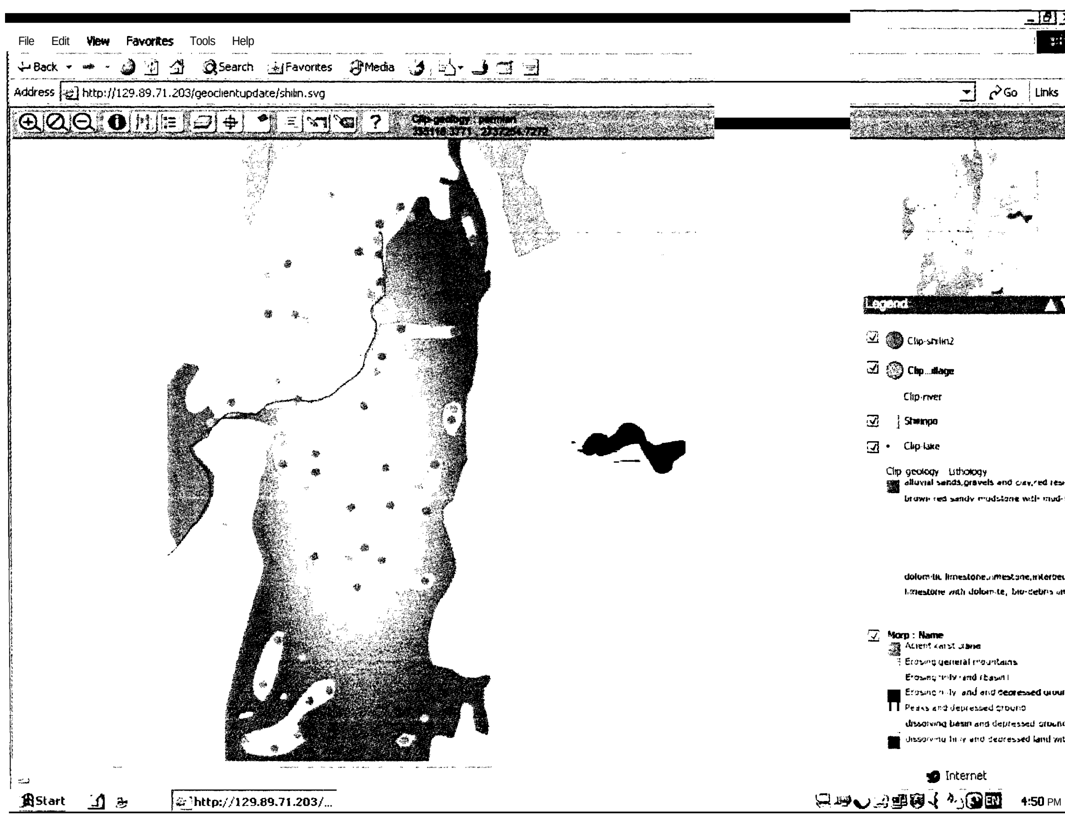

Figure 7. *Labeling the map.*

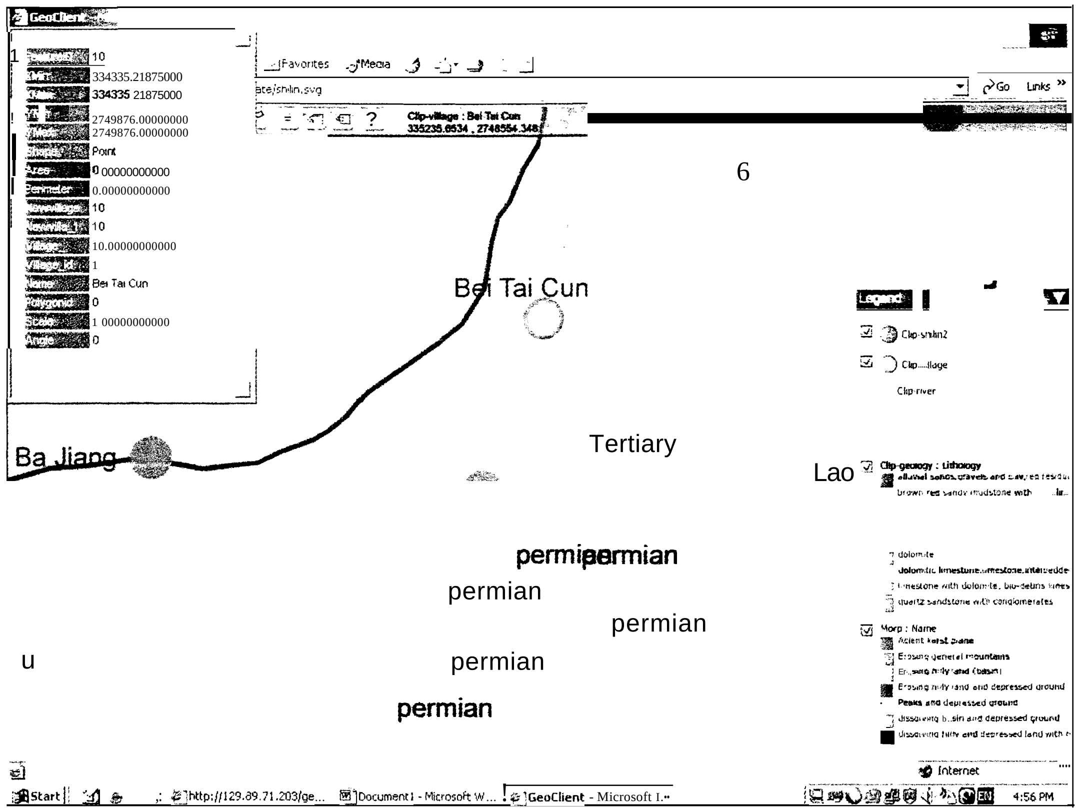

Figure 8. *Querying the GML -based interoperable da tabase.*

over the Internet, GML gives users the capability to publish higher quality spatial maps quickly, dynamically and economically. Any changes to the GML data can be instantly reflected in the SVG maps, but since the GeoClient software is still in its infancy, in this case study the advantages of the GML-based database can not fully demonstrated.

Effective conservation of the Stone Forest Landscape requires knowledge of local ecosystems on different temporal and spatial scales. Further research is needed to use GML to effectively store temporal data so as to build a 4-dimensional GML-based database for the conservation of the Stone Forest Landscape.

CONCLUSION

This paper introduces the issues of data interoperability, advantages of GML, and its mechanism for data interoperability. A simple GML-based database is designed and constructed as a case study to demonstrate the interoperability of GML-based databases. The GeoClient software, which is the only one we can find currently on the Internet for free, is used to serve as an graphic interface for the GML-based database. The case study shows that the GML-based interoperable database can be displayed as a SVG map, which is very high quality, scalable and resolution-independent, on a user-friendly interface provided by the GeoClient. The database can be accessed and queried at the feature level in real time from the WFS server over the Internet. The information can be accessed by a range of programs on different piatforms via the Internet. Basic GIS functions provided by the GeoClient, such as zooming, panning, labeling and querying, can be performed using the database on the Web. The GML-based database can serve high quality maps through delivering SVG vector data over the Internet. It can be shared and re-used easily in the future.

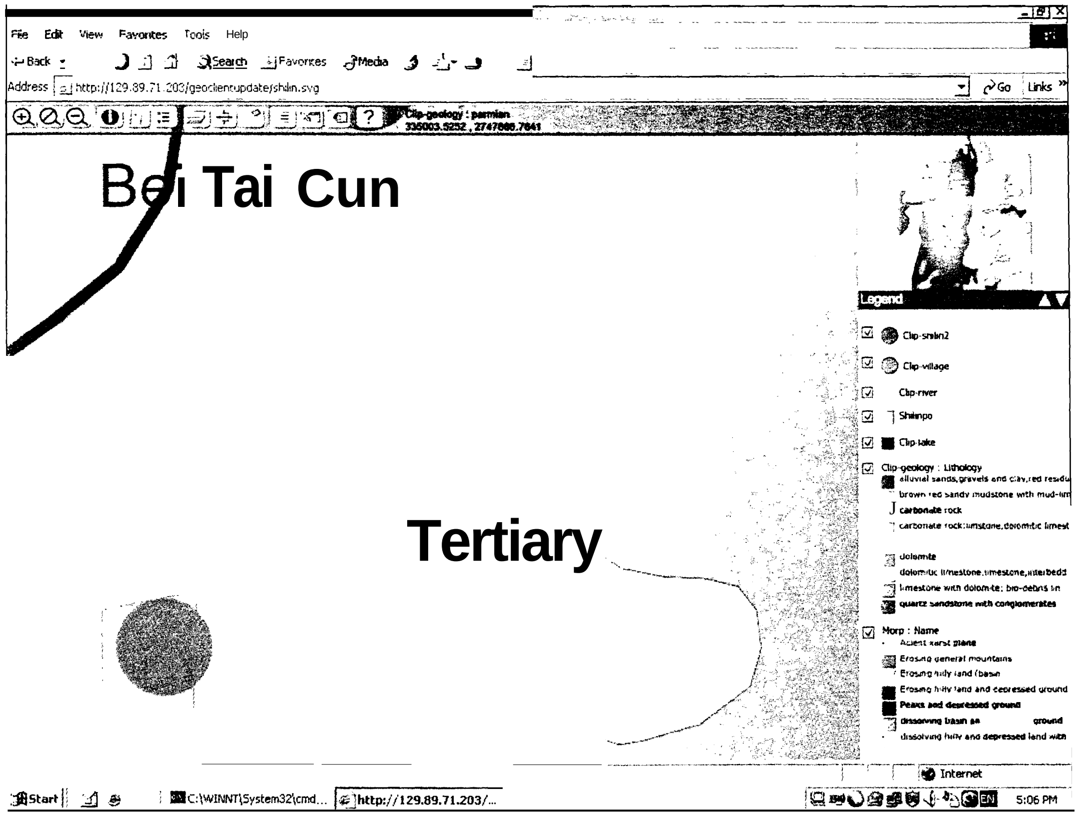

Figure 9. *Zooming in the map without any degradation in map quality.*

As an interoperability standard, GML allows the gaps among different data sources, vendors, databases and formats to be bridged. The database built in this case study can communicate with other databases through converting ArcView Shapefiles into GML data. GML can give users the capability to easily and dynamically publish and exchange data in an open, non-proprietary industry-standard format on the Web, thus maximizing the re-use of geospatial data, eliminating time-consuming data conversion and reducing associated costs. The high quality and colourful SVG nlap transformed from the GML-based database shows a very nice interface to users, which can irrlprove the public accessibility to existing data. GML holds promise to lead an exciting interoperable future via online interactive Web maps and spatial Web services. But because the development of support software systenls for GML-based databases are still at an early stage, the advantages of GML-based databases cannot be fully displayed in this case study.

As a new interoperability approach, GML still has some limitations. GML is not intended to solve all geo-processing interoperability problems. It still can not fully solve the problem of semantic interoperability. For example, GMt provides users the ability to create application schemas to model their data, but different users (i.e., data providers) may use different names to represent the same feature, e.g., one user may decide to create a GML schema with a *building* feature while another user may use a *house* feature for essentially the same entities. Thus the second user must know the schema created by the first user in order to integrate the data from the first user into his. Without knowledge of these schemas users cannot fully understand what the GML represents.

The real data inter.operability is to provide seamless communication between remote GIS databases without having prior knowledge of their underlying semantics. A real interoperable GIS database should provide transparent comnlunications at data model and application semantics level (Bishr, 1998). Further research needs to be done for a GML-based database to become a real interoperable GIS database.

REFERENCES

Arctur, D., Hair, D., Timson, G., Martin, E.P., and Fegeas, R. (1998) Issues and prospects for the next generation of the spatial data transfer standard (SOTS). *Int.). Geographical Information Science; vol.* 12; *no.* 4; *pp. 403-425.*

Bertolotto, M. and Egenhofer, M. j. (2001) Progressive Transmission of Vector Map Data over the World Wide Web. *GeoInformatica.* 2001. *vol.* 5; *no.* 4. *pp* 345-373.

Bishr, Y. (1998) Overcoming the semantic and other barriers to GIS interoperability. *Int.). Geographical Information Science; vol.* 12; *no.* 4; *pp.* 299-314.

Brodie, M.L., (1992) The promise of distributed computing and the challenge of legacy Information systems. *Proceedings of the IFIP WG2-6 Database Semantics Conference on Interoperable Database Systems (05-5); Lorne; Victoria; Australia (Amsterdam: North Holland); pp.* 1-31.

Buelher, K. and McKee, L. (1996) Introduction to Interoperable Geoprecessing. *The OpenGIS Guide/* Open GIS Consortium.

Choicki, j. (1999) Constaint-based Interoperability of Spatiotemporal Databases. *Geoinformatica/ vol.* 3; *no. 3/ pp.211-243.*

Clemens, I. (2002) The Elegance of XML-based Mapping *Proceedings of URISA 2002 Annual Conference and Exposition Chicago/ Illinois. Oct. 2002 pp.* 320-324

Curtis, E. (2003) XML Schema: Reconciling Diversity with Standardization. *gml dev days 2003.* http://www.gmldev.org/GMLDev2002/presentations/

Devogele, T., Parent, C., and Spaccapietra, S. (1998) On spatial database integration. *Int.). Geographical.fnformation Science; 1998; vol.* 12/ *no.4/ pp335-352.*

Huntoon, P. W. (1992) Hydrogeologic Characteristics and Deforestation of the Stone Forest Karst Aquifers of South China. *Ground vol.* 3D; *pp. 167-176.*

Huntoon, P. W. (1993) The hydrogeological impacts of deforestation on the sensitive stone forest karst aquifers in south China. In: Song, L. and Ting, H. (eds), *Karst landscape and cave tourism.* China Environmental Science Press, pp. 112-118.

joo, I (2002) Shi Lin(Natural Stone Forest). http://www.virtualtourist.com/vt/f7238/?0=2&i=1#T

Lake, A. (2002) Will GML enable an accessible Geo-Web. GEO World, july 2002 http://www.geoplace.com/gw/2002/0207/0207gml.asp.

Lake, R. (1999) Introduction to GML Geography Markup Language, http://www.webmapping.org/GMLIntroduction.html.

Laurini, R. (1998) Spatial multi-database topological continuity and indexing: a step towards seamless GIS data interoperability. *Int.). Geographical Information Science; vol.* 12; *no.4/ pp. 373-402.*

Meneghello, M. (2001) XML (eXtensible Markup Language): The New Language of Data Exchange, *Cartography, Vol. 3D; No.1.*

Misund, G., and johnsen, K.-E. (2003) The One Map Project. *gml dev days 2003.* http://www.gmldev.org/GMLDev2002/presentations/

Murray, D. and Chow, j. C. (2002) GML User Perspectives. *Proceedings of URISA 2002 Annual Conference and Exposition Chicago; Illinois. Oct. 2002,pp. 324-332*

NIST (National Institute of Standard and Technology) (1994) Federal Information Processing Standard Publication Spatial Data Transfer Standard. US Department of Commence, Washington D.C.

accapietra, S.
itegration. *Int.*
science, 1998,

ldrogeologic
tation of the
South China.
7-176.
rogeological
the sensitive
south China.
(eds), *Karst*
rism. China
p. 112-118.
tone Forest).
/vt/f7238/?

enable an
World, July
place.com/

n to GML
Language,
/GMLIntrodu

ti-database
exing: a step
teroperability.
tion Science,

nsible Markup
age of Data
,No.1.
003) The One
days 2003.
ev2002/pre

2) GML User
URISA 2002
Exposition
.324-332
tandard and
Information
tion Spatial
partment of

Noronha, V. (2000) Towards ITS Map Database Interoperability-Database Error and Rectification. *GeoInformatica, vol. 5, no.* 4, *pp.* 345-373.

aGC Abstract Specification (2001) http://www.opengis.net/gml/01-029/GML2.html

aGC (2003) http://www.opengis.org/techno/documents/02-023r4.doc

Peng, Z.-R. (2003) A Framework of Feature-Level Transportation Geospatial Data Sharing Systems. (Paper presented at the Transportation Research Board Annual Meeting January 2003).

Peng, Z.-R., and Tsou, M.-S. (2003) *Internet GIS: Distributed Geographic Information Services for the Internet and Wireless Networks,* New York: John Wiley & Sons.

Shekhar, S., Vatsavai, R. R., Sahay, N., Burk, T. E., and Lime, S. (2001) WMS and GML based interoperable web mapping system. *Proceedings of the ninth ACM international symposium on Advances in geographic information systems. Atlanta, Georgia, USA. 2001. pp 106-111*

Siki, Z. (1999) GIS data exchange problems, solutions. *Periodica Polytechnika.* http://bme-ge9d.agt.bme.hu/staff_hisiki/gisexch.htm.

Smith, C., Aberline, P. and Bellman, C. (2002) Investigation and Prototyping of Emerging Web Standards for Web GIS, *Cartography, Vol.* 31, *No.1.*

Song, L. (1997) Stone Forest A Treasure of Natural Heritage. In: Song, L. et ai, *Stone Forest Atreasure of Natural Heritage.* China Environmental Science Press, pp. 1.

Stephan, E., Vckovski, A., and Bucher, F. (1993) Virtual data set: an approach for the integration of incompatible data. *Proceedings of Auto Carto 11, Bethesda: American Congress on Surveying and Mapping,* pp. 93-102.

Tarnoff, P. J. (1998) Evolution or *Traffic Technology International.* February/March, 1998, pp. 32-35.

Uitermark, H. (1996) The integration of geographic databases. *Proceedings of Second Joint European Conference.* Amsterdam, 105 Press, pp. 92-95.