Article

Pioneering GML Deployment for NSDI — Case Study of US TIGER/GML

Lingling Guo

Public Safety Research School, Tsinghua University, 1001, Liuqing Building, Beijing 100084, China;
E-Mail: lingling.guo@yahoo.com or lept@tsinghua.edu.cn; Tel.: +86-10-6279-2859;
Fax: + 86-10-6279-2863

Abstract: The National Spatial Data Infrastructure (NSDI) is defined as the technologies, policies and people necessary to promote sharing of geospatial data throughout all levels of government, the private and non-profit sectors and the academic community. The US Census Bureau is the federal agency lead for administrative units data, one of the seven data themes identified by the NSDI framework. The administrative unit is a unit with administrative responsibilities. These units are organized as nodes/lines/areas feature data. The OpenGIS Geography Markup Language (GML) is the XML grammar to express the geographic features. This study at the US Census Bureau investigates how the general-purpose GML standard could be leveraged and extended to describe the most comprehensive geographic dataset with national coverage in the US. Challenges and problems in dealing with data volume, GML document structure, GML schema design and GML document naming are analyzed, followed by proposed solutions proven for feasibility. Our results show that one key point in making a successful GML deployment for NSDI is to reflect the characteristics of the geographic data through a carefully designed GML schema, structure and organization. The lessons learned may be useful to others transforming NSDI framework data and other large geospatial datasets into GML structures.

1. Introduction

The concept of National Spatial Data Infrastructure (NSDI) was initialized in the US and now has been widely adopted by many other countries, including Australia, Canada, Chile, China, the United Kingdom and Finland. It is defined as the technologies, policies and people necessary to promote sharing of geospatial data throughout all levels of government, the private and non-profit sectors and the academic community [1]. Geodetic control, cadastral, orthoimagery, elevation, hydrography, administrative units and transportation are seven data themes that have been identified by the NSDI framework [2,3] that forms the data backbone of the NSDI. The US Census Bureau is the federal agency working with US administrative units. An administrative unit is a geographic entity established by legal action and for the purpose of implementing administrative or governmental functions. Most administrative units have officially recognized boundaries. All areas and population of the United States are part of one or more legal units. These units include the nation, states and statistically equivalent areas, counties and statistically equivalent areas, incorporated places and consolidated cities, functioning and legal minor civil divisions, federal- and state-recognized American Indian reservations and off-reservation trust lands and Alaska Native Regional Corporations. As shown in Figure 1, the US administrative units data presents a complex internal structure. This hierarchical geographic presentation shows the geographic entities in a superior/subordinate structure. This structure is derived from the legal, administrative or areal relationships of the entities. An example of hierarchical presentation is the census geographic hierarchy consisting of a census block, within block group, within census tract, within place, within county subdivision, within county and within state. This information is presented as a series of nesting relationships.

Figure 1. Hierarchy of census geographic entities.
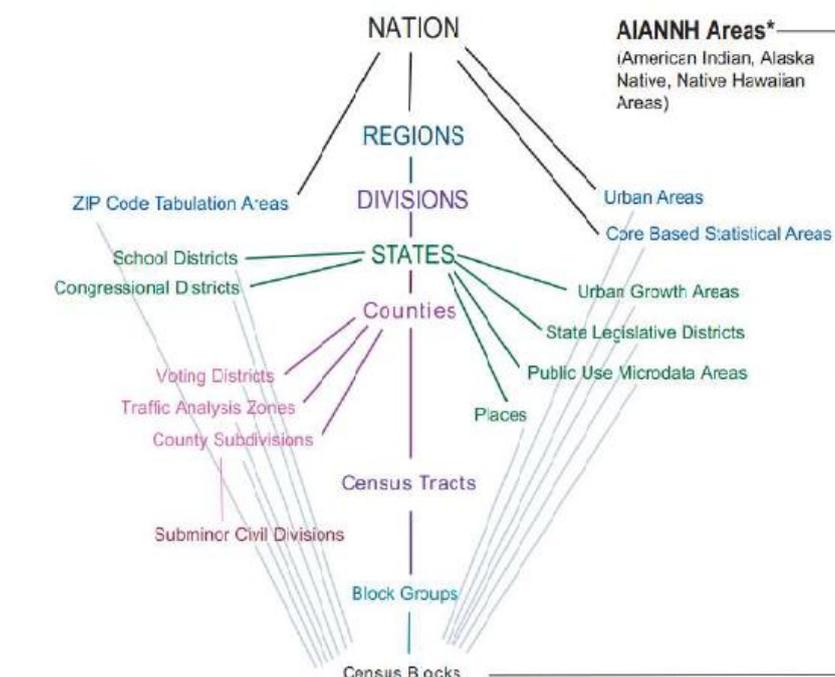
The geography division at the US Census Bureau manages the Topologically Integrated Geographic Encoding and Referencing (TIGER) system and the digital database to support the decennial census

and sample survey programs of the Census Bureau, starting with the 1990 decennial census. TIGER data is the most comprehensive geographic dataset with national coverage in the US. To make it publicly available, the Census Bureau offers several file types and an application for mapping census geographic data as TIGER/Line files, TIGER/Line Shapefiles or KML prototype files. Take TIGER/Line Shapefiles as an example: they are spatial extracts from the Census Bureau's TIGER database, containing features, such as roads, railroads, rivers, as well as legal and statistical geographic areas. To work closely with US NSDI infrastructure, the US Census Bureau started a new pilot project to utilize the general-purpose Open Geospatial Consortium (OGC) Geography Markup Language (GML) standards [4] to organize and publish TIGER spatial data.

GML is an XML grammar written in XML Schema for the modeling, transport and storage of geographic feature data [5]. Much research has been directed on how to effectively store GML documents [6–9], data compression techniques for GML documents [10], syntactic and lexicon analysis of GML documents, native support from spatial databases for GML documents [11], how GML documents can be effectively used in Web-geographic information system (GIS) environments [12–15], effective spatial query language over GML [16] and transforming GML into other open data formats, including Scalable Vector Graphics (SVG) [17]. Ahn introduced the GML extensions, a Spatial XQuery language, and its processing modules for mobile and location-based application [18]. Bardet presented a mapping from the basic geometric objects in geotechnical data to basic geometric features of GML [19]. Corcoles defined an ontology-based approach for integrating non-spatial resources with GML documents [20], and Ferri proposed a method for evaluating the semantic similarity of GML elements [21]. Huang introduced a transit network data model with GML schemas for data encoding and sharing [22]. Lake reviewed the features of GML 3.0 standards and presented its applicability to the geological sciences through several case studies [23]. Nativi defined GML-based structures for netCDF data, which is one of the primary methods of self-documenting data storage and access in the international geosciences research and education community [24]. Zhang presented a GML-based geographical information search engine over the Internet [25].

Another similar work is INSPIRE, an EU initiative to establish an infrastructure for spatial information in Europe that will help to make spatial or geographical information more accessible and interoperable for a wide range of purposes supporting sustainable development. In accordance with the INSPIRE directive, three different types or levels of metadata are distinguished: metadata 'for discovery', metadata 'for evaluation' and metadata 'for use'. Due to its extensibility and flexibility, GML is a recommended encoding for metadata 'for use' (as this kind of metadata can be quite rich and different from the metadata for discovery or evaluation, which, within INSPIRE, are less rich and more common). For other metadata encoding, the ISO/TS 19139 (and information models of ISO 19115/19119) and Dublin Core (ISO 15836) standards are used. It should be noted that, according to the INSPIRE harmonization requirements, the creation of the metadata schemas is one of the highest priorities. From the NSDI point of view, it is very desirable to thoroughly investigate the applicability of the general GML standards for complex geographic feature data with national coverage, such as the US administrative units data maintained in the Census Bureau. Little research has been directed to the deployment of GML for national scale geographic data. The purpose of this paper is to present our pioneering work in the Geography Division of the US Census Bureau to demonstrate how the general GML standards could be leveraged and extended to transform the comprehensive US national scale

Topologically Integrated Geographic Encoding and Referencing system (TIGER) geographic data to be GML-based structures.

The rest of this paper is organized as follows. The case study scenario is introduced, followed by a summary of the challenges in applying the general GML standards for the comprehensive TIGER data. The next section focuses on the proposed solutions and implementation in detail, followed by the presentation of the produced TIGER/GML products. The advantages and limitations of the proposed solution are discussed prior to the conclusions.

2. Scenario

2.1. US TIGER Data

TIGER data has been maintained at the US Census Bureau from the mid-1980s. It includes legal and statistical geographic entities, as well as transportation and hydrographic networks covering the United States, Puerto Rico and the Island Areas (American Samoa, Commonwealth of the Northern Mariana Islands, Guam and US Virgin Islands). The TIGER/Line and TIGER/Line Shapefile mapping data have been broadly used by all levels of government, the private and non-profit sectors and the academic community as one of the primary US nationwide GIS data resources.

2.2. TIGER GML

The Census Bureau performed a pilot research and implementation of TIGER/GML. The project evaluated the feasibility of generating GML structures from the massive TIGER database in the test production environment at the Census Bureau headquarters.

2.3. System Requirements

A dedicated program is needed to generate GML documents for national scale TIGER data directly from the TIGER database. It is expected to be a standalone command line program that can perform unattended GML data generation for the whole TIGER dataset in the UNIX production environment.

3. Challenges in Applying the GML Standard for TIGER Data

Analysis performed in the Geography Division of the US Census Bureau has revealed the following significant issues when designing, implementing and packaging GML documents for national scale TIGER data.

3.1. Data Volume

GML is based on XML, a text-based encoding format. GML documents tend to be much larger in size than other formats containing the same information. In the TIGER database, even county-based partitions will often be over 250 MB for counties in major metropolitan areas. Most XML utilities have been struggling to open, much less process, GML files of this size, without mentioning the file sizes for higher levels of geographic entities.

3.2. Comprehensive TIGER Organization

TIGER data has a very comprehensive organization of Census geographic areas.

One hierarchical set of Census geographic areas—Nation/Region/Division/State/County/Tract /Block Group/Block—is a completely nested structure, where the nested areas at each level below Nation are mutually exclusive and collectively exhaustive of the area above that contains them.

Another set of geographic areas—Voting Districts, Traffic Analysis Zones (TAZ), County Subdivisions and Sub-Minor Civil Division—nest within counties.

A third set of geographic areas—Congressional Districts, School Districts, Places, Alaska Native Regional Corporations (ANRCs) and State Legislative Districts (SLDs)—Upper and Lower, Urban Growth Areas (UGAs), Public Use Microdata Areas (PUMAs) and Consolidated Cities (CONC)—nest within states.

A fourth set of geographic areas—Zip Code Tabulation Areas (ZCTAs), Urban Areas and Metropolitan Statistical Areas and American Indian/Alaska Native/Native Hawaiian (AIANNH) areas—nest within the nation.

3.3. GML Document Naming

It is not a practical solution to represent TIGER data in a single GML document. Therefore, TIGER/GML has to be a suite of interrelated GML documents. How to name these individual GML documents to reflect their inner connections is another difficulty when the number of generated GML documents would be fairly large.

3.4. GML Element ID Definition

TIGER data has many built in one-dimensional (1D) and two-dimensional (2D) feature types. When generating a GML representation for TIGER data, unique values for each GML identifier must be used. The unique value, though, must allow one to directly reference the entity within the database itself, or else the identifier would become meaningless. It was therefore determined to construct the GML identifier by incorporating such information guaranteeing the success of the final GML deployment.

4. Proposed Solution

A divide-and-conquer approach was designed to deal with the aforementioned challenges: GML documents are mainly generated at the county-level; multiple GML document types are designed with each type dealing with specific TIGER features; document names consist of several parts that correspond to the GML document type and geographic unit level; and the GML element ID is developed by considering both the feature type and the corresponding level of geographic entities.

4.1. TIGER/GML Document Types

TIGER/GML data is distributed among nine different types of documents:

1. Index
2. Metadata
3. Area Features (Geographic Entities)
4. Blocks
5. Public Use Microdata Areas (PUMAs)
6. Linear Features (Roads, Railroads, *etc.*)
7. Landmarks (Point, Line and Area)
8. TIGER/Line ID (TLID) history
9. Identifier Ranges

The document structure of all documents is identical. The document types differ in which optional elements they contain at the national, state and county levels, as shown in the Table 1. Some documents may not be applicable to some level. For example, since the Tiger/GML data is extracted from county-based partitions of the TIGER database, all linear features and most landmarks are contained within a single county. They are not on any state or nation levels.

Table 1. Topologically Integrated Geographic Encoding and Referencing system (TIGER)/Geography Markup Language (GML) document/file type.

Document	County Level	State Level	National Level
Index	Link to feature collections in county Area Entities, Blocks, Linear Features, Landmarks and to TLIDhistory and IdentifierRange files	Link to feature collections in state Area Entities, PUMAs and to Index files for counties.	Link to feature collections in national Area Entities and to Index files for states.
Metadata	NA	NA	Census Metadata
AreaEntities	All geographic area entities within county, except Blocks	Multi-County Entities	Multi-State Entities
Blocks	Blocks	NA	NA
PUMAs	NA	PUMA1, PUMA5	NA
LinearFeatures	All linear features	NA	NA
Landmarks	All landmarks	NA	NA
TLIDhistory	All TLID history	NA	NA
IdentifierRange	All identifier ranges	NA	NA

4.2. TIGER/GML Schema

The Census TIGER/GML schema is an OGC GML application schema contained in five XML/Schema documents. These schemas are based on the GML version 3.1.1 specification and schemas as described in OGC document 03-105r1.

1. CensusTiger.xsd: defines census features and feature collections.
2. CensusTiger123.xsd: defines abstract and base 1, 2 and 3 dimension types
3. CensusTigerSpatialTypes.xsd: defines specialized geometry types.
4. CensusTigerBasicTypes.xsd: defines census geographic codes and metadata.
5. CensusTigerMetadata.xsd: defines census TIGER metadata types and elements

Census TIGER/GML types are extensions and restrictions of base GML types, as described in the GML specification. These schemas define the XML document structure for TIGER/GML documents, the information model for Census TIGER/GML features and the valid values for codes and other simple atomic data items.

The Census TIGER/GML schemas follow the GML <Class><property><Class> "2-step" XML encoding convention. The Class element is a GML Object with identity provided by a gml:id attribute that is an XML ID. The property element is in effect a local name for the use of the Class element it contains or references with an xlink:href. Any one property element may either contain or reference a Class element, but may not do both. An xlink:href value contains a URI prefix if it refers to an element in a different XML document, a "#" fragment identifier or an XML ID.

The names of property elements are in lowerCamelCase and the names of Class elements are in UpperCamelCase. For example, the name for property element gml:boundedBy is in lowerCamelCase format, while its Class element, gml:Envelope, is in UpperCamelCase format. A property element may not have an XML ID.

The Census TIGER/GML schemas import the XLink definitions used in GML from xlinks.xsd. This allows for that information to be referenced within and across XML documents, as well as included in-line. They import US Federal Geographic Data Committee (FGDC) metadata types via fgdc-std-001-1998.xsd. FGDC metadata use is optional for all TIGER/GML objects via the gml:metadataProperty element.

4.3. TIGER/GML Document Naming

Based on the aforementioned GML organization architecture, the following naming convention is used for these document types:

"tgr" + ssccc + docTypeName + ".xml"

where "ssccc" is the federal state and county codes (FIPS codes) for a county or the FIPS state code, appended with "000" for a state or "00000" for the nation, and "docTypeName" is "Index", "AreaEntities", "LinearFeatures", *etc*., as listed on the Document column above.

4.4. TIGER/GML Element ID Definition

GML object XML elements that represent collections will incorporate an ID (geo-id) that is a concatenation of area entity codes that uniquely identify the area. The Census Tiger Basic Types schema includes <name of area>EntityCodesType code sets for all Census area entities, where <name of area> is "State", "County", *etc*. All of the child elements, except the extracted data year and generation, are included in the ID (geo-id) for that type of area. The IDs for feature collections at the national and state levels are right justified zero filled.

The ID attributes on GML object XML elements (gml:id) in Census TIGER/GML data will be assigned to maintain global uniqueness.

5. Generated TIGER/GML Documents

In this pilot project, TIGER/GML products have been generated in the test production environment in the Geography Division of the US Census Bureau in about three weeks of continuous runtime. As stored in the ZIP archive files, TIGER/GML data for 2005 is almost 11 GB. Unzipped, it is close to 400 GB. This project of generating TIGER/GML is an internal experiment at the US Census Bureau designed to test the technology of transferring high-volume nationwide geographic data into GML format. The resulting GML will not be put online. However, the GML documents and schemas can be obtained by contacting the geography division of the US Census Bureau.

6. Discussion

6.1. Handling Cross-Boundary Features

TIGER/GML data is extracted on a county-by-county basis from the TIGER database. All linear features and most landmarks are contained within a single county. However, some geographic area features may cross county boundaries, and others may also cross state boundaries. The proportion of features that cross county and state boundaries varies with the entity/feature type and also varies across different states and regions. The multi-state features are put in nation-level GML files; the multi-county area features are in state-level GML files; the single county features are in county-level GML files. All multi-state and multi-county entities required special processing, since all of the base GML files were created county-by-county. In order to do this, it was determined to have all of the county-based TIGER/GML files read into an Oracle database. A combination of SQL and an XSLT script was then used to create a GML file containing the multi-state and multi-county entities. By evaluating the data, the scripts would combine any entity that crossed either a county or state boundary and create a record to be placed in the national GML file. This entity would be found in the counties, while not in the state file.

6.2. Coordinates in TIGER/GML

As extracted from the TIGER database system, TIGER/GML coordinate data is in the North American Datum 1983 (NAD 83), with coordinates in longitude/latitude order. TIGER is unprojected, and the coordinates are in decimal degrees. TIGER/GML was produced using Oracle Spatial and refers to its spatial reference system identifier (SRID) for NAD83, which is 8265.

TIGER/GML coordinate data may be converted to different coordinate reference systems for cartographic display. Such conversions may change the coordinate order and/or the coordinate types, e.g., to easting or northing, in many projected coordinate reference systems.

6.3. gml:boundedBy Element in TIGER/GML

The gml:boundedBy element is included on every CensusGeographyCollections feature collection element to indicate the spatial extent of all of the features contained in the collection. Coordinates in the gml:Envelope contained in a gml:boundedBy element are represented like any other TIGER/GML coordinate. The gml:boundedBy elements of contained feature collections may indicate smaller spatial

extents than the one on CensusGeographyCollections. The gml:boundedBy element is also included on every TIGER/GML feature.

6.4. Names in TIGER/GML

All features and feature collections in TIGER/GML may have one or more optional gml:name elements that support access and display of TIGER/GML data by generic GML software. For individual area features, linear features and landmark features, the gml:name will replicate name elements in specialized TIGER/GML structures. The area name element for areas that are not generally named, such as Blocks, will be the same as the gml:id, e.g., the name of the element plus the area codes that uniquely identify it.

Linear features in TIGER/GML are described by one or more CensusFeatureName elements that contain a required feature name element and optional feature prefix direction, feature type and feature suffix direction elements. The contents of the feature name element in each CensusFeatureName will be replicated in a separate gml:name element for the linear feature. Landmark features in TIGER/GML are described by a landmarkName element. There is only one landmark Name per Landmark; its contents will be replicated in a separate gml:name element for the landmark.

6.5. gml:description in TIGER/GML

General purpose GML software often relies on the optional gml:description element to describe GML data to a user. To support this use, every feature collection and feature in a TIGER/GML data document will have a gml:description element. The description for the top level CensusGeographyCollections element will indicate the contents of the document, the spatial extent, its contents cover, the data year, data generation and the date it was extracted from the TIGER database. The description for each feature collection will explain the types and extent of features in the collection. For example, for States, the description is "States and state equivalents (District of Columbia, Puerto Rico, American Samoa, Guam, Commonwealth of the Northern Mariana Islands, US Virgin Islands and the US Minor Outlying Islands) of the United States." These descriptions may replicate descriptions in the existing TIGER/Line technical documentation. The description for an individual feature will describe it and its geographic context. For example, "Census Block Group 490039601001 in Box Elder County, Utah".

6.6. CensusMetaData in Census TIGER/GML

CensusMetaData for TIGER/GML has the same basic content as the metadata for TIGER/Line 2005, amended to reflect the differences in data structure. Complete global CensusMetaData is referenced from every feature collection and feature in the current TIGER/GML data set. Partial local CensusMetaData that applies to selected features or feature collection elements in a document or data store and that differs from the global CensusMetaData only for those selected elements, may be included in-line on one of those elements in future TIGER/GML data sets. Other elements that share all of the partial local CensusMetaData may have a censusMetaDataProperty with an xlink:href that

refers to the in-line CensusMetaData. For example, elements with lower or higher than average spatial accuracy (which was 7.6 m in legacy TIGER) could have local metadata.

7. Conclusions

This paper presents research performed in the Geography Division of the US Census Bureau on how to utilize the GML standard to organize and present national scale TIGER data. We summarized the research issues, proposed solutions and introduced generated TIGER/GML experimental results. The following conclusions were reached:

1. Data volume, comprehensive data organization, GML document naming and GML element ID definition are major issues when generating a GML document for NSDI framework datasets.
2. A divide-and-conquer approach is a feasible solution to overcome the aforementioned issues.
3. Carefully designed GML schema, structure and organization that reflect the characteristics of the targeted geographic datasets are the key to making successful GML deployments for NSDI.

Acknowledgements

References and Notes

1. FGDC. Coordinating Geographic Data Acquisition and Access: The National Spatial Data Infrastructure (Executive Order 12906). Available online: http://www.fgdc.gov/policyandplanning/executive_order (accessed on 17 February 2013).
2. FGDC. Framework Introduction and Guide. Available online: http://www.fgdc.gov/framework/frameworkintroguide (accessed on 17 February 2013).
3. FGDC. The Geographic Information Framework Data Standard. Available online: http://www.fgdc.gov/standards/projects/FGDC-standardsprojects/framework-data-standard/framework-data-standard (accessed on 17 February 2013).
4. Clemens, P. OpenGIS Geography Markup Language (GML) Encoding Standards. Available online: http://www.opengeospatial.org/standards/gml (accessed on 17 February 2013).
5. Skoks, V.; Steurer, C. An overview of the use of GML in modern spatial data infrastructures. *Scientific Journal of Riga Technical University Computer Science* **2011**, *42*, 60–67.
6. Corcoles, J.E.; Gonzalez, P. Analysis of Different Approaches for Storing GML Documents. In *Proceedings of the 10th ACM International Symposium on Advances in Geographic Information Systems*, McLean, VA, USA, 4–9 November 2002; pp. 11–16.
7. Zhu, F.; Guan, J.; Zhou, J.; Zhou, S. Storing and Querying GML in Object-Relational Databases. In *Proceedings of the 10th ACM International Symposium on Advances in Geographic Information Systems*, McLean, VA, USA, 4–9 November 2002; pp. 107–114.

8. Kim, Y.K.; Jang, Y.J.; Chang, J.W. Design and Implementation of Efficient Storage Schemas and Low-level Storage Manager for GML Documents. In *Proceedings of Third International Conference on Research Challenges in Information Science*, New York, NY, USA, 22–24 April 2009.

9. Zhu, F.; Guan, J.; Zhou, S. Constraints-Preserving GML Storage in Object-Relational Databases. In *Proceedings of the 15th Annual ACM International Symposium on Advances in Geographic Information Systems*, Seattle, WA, USA, 2007; pp. 1–4.

10. Li, Y.; Imaizumi, T.; Guan, J. Spatial Data Compression Techniques for GML. In *Proceedings of Japan-China Joint Workshop on Frontier of Computer Science and Technology*, Nagasaki, Japan, 27–28 December 2008; pp. 79–84.

11. Sripada, L.; Lu, C.; Wu, W. Evaluating GML Support for Spatial Databases. In *Proceedings of the 28th Annual International Computer Software and Applications Conference*, Hong Kong, China, 28–30 September 2004; pp. 74–77.

12. Huang, C.H.; Chuang, T.R.; Deng, D.P.; Lee, H.M. Efficient GML-Native Processors for Web-Based GIS: Techniques and Tools. In *Proceedings of the 14th Annual ACM International Symposium on Advances in Geographic Information Systems*, Arlington, VA, USA, 5–11 November 2006; pp. 91–98.

13. Shekhar, S.; Vatsavai, R.R.; Sahay, N.; Burk, T.E.; Lime, S. WMS and GML Based Interoperable Web Mapping System. In *Proceedings of the 9th ACM International Symposium on Advances in Geographic Information Systems*, Atlanta, GA, USA, 9–10 November 2001; pp. 106–111.

14. Paul, M.; Ghosh, S.K. An Approach for Geospatial Data Management for Efficient Web Retrieval. In *Proceedings of the 6th IEEE International Conference on Computer and Information Technology*, Seoul, Korea, 20–22 September 2006; IEEE 28.

15. Kim, M.J.; Kim, M.; Joo I.H.; Jang, B.T. GML Encoded Traffic Information Web Services System. In *Proceedings of 2004 IEEE 60th Vehicular Technology Conference*, New York, NY, USA, 26–29 September 2004.

16. Corcoles, J.E.; Gonzalez, P. A Specification of a Spatial Query Language over GML. In *Proceedings of the 9th ACM International Symposium on Advances in Geographic Information Systems*, Atlanta, GA, USA, 9–10 November 2001; pp. 112–117.

17. Guo, Z.; Zhou, S.; Xu, Z.; Zhou, A. G2ST: A Novel Method to Transform GML to SVG. In *Proceedings of the 11th ACM International Symposium on Advances in Geographic Information Systems*, New Orleans, LA, USA, 3–8 November 2003; pp. 161–168.

18. Ahn, Y.S.; Park, S.Y.; Yoo, S.B.; Bae, H.Y. Extension of Geography Markup Language (GML) for Mobile and Location-Based Applications. In *Computational Science and Its Applications: ICCSA 2004*; Springer: Berlin/Heidelberg, Germany, 2004; Volume 3044, pp. 1079–1088.

19. Bardet, J.P.; Zand, A. Spatial modeling of geotechnical information using GML. *Trans. GIS* **2009**, *13*, 125–165.

20. Corcoles, J.E.; Gonzalez, P. Integrating GML Resources and Other Web Resources. In *Proceedings of 15th International Workshop on Database and Expert Systems Applications (DEXA'04)*, Zaragoza, Spain, 30 August–3 September 2004.

21. Ferri, F.; Formica, A.; Grifoni, P.; Rafanelli, M. Evaluating Semantic Similarity Using GML in Geographic Information Systems. In *On the Move to Meaningful Internet Systems 2005: OTM 2005 Workshops*; Springer: Berlin/Heidelberg, Germany, 2005; Volume 3762, pp. 1009–1019.

22. Huang, R. Modeling transit networks by GML for distributed transit Trip Planners. *J. Spat. Sci.* **2008**, *53*, 1–15.

23. Lake, R. The application of geography markup language (GML) to the geological sciences. *Comput. Geosci*. **2005**, *31*, 1081–1094.

24. Nativi, S.; Caron, J.; Davis, E.; Domenico, B. Design and implementation of netCDF markup language (NcML) and its GML-based extension (NcML-G(ML)). *Comput. Geosci.* **2005**, *31*, 1104–1118.

25. Zhang, J.T.; Gruenwald, L. A GML-Based Open Architecture for Building a Geographical Information Search Engine over the Internet. In *Proceedings of the Second International Conference on Web Information Systems Engineering*, Kyoto, Japan, 3–6 December 2001; Volume 2, pp. 25–32.