Spatio-Temporal Data Exchange Standards

Albrecht Schmidt   Christian S. Jensen

Department of Computer Science, Aalborg University, Denmark
{*al,csj*}*@cs.auc.dk*

Abstract

We believe that research that concerns aspects of spatio-temporal data management may benefit from taking into account the various standards for spatio-temporal data formats. For example, this may contribute to rendering prototype software "open" and more readily useful. This paper thus identifies and briefly surveys standardization in relation to primarily the exchange and integration of spatio-temporal data. An overview of several data exchange languages is offered, along with reviews their potential for facilitating the collection of test data and the leveraging of prototypes. The standards, most of which are XML-based, lend themselves to the integration of prototypes into middleware architectures, e.g., as Web services.

1 Introduction

It is often important to be able to test new spatial and temporal query processing techniques with real-world data. This contributes to understanding how the techniques will perform in specific production settings. Over the years, a number of data sets have become publicly available that, along with synthetic data, can serve as a basis for experimenting with software prototypes. However, since many of these data sets are only available in custom formats, a fair amount of domain knowledge is required to make use of them. Recent standardization efforts, which typically employ XML-based technologies and digital library data exchange formats, aim to facilitate the communication between applications by specifying data exchange standards. Especially, XML-based geo-enabled applications can be deployed and enriched with the plethora of query languages, constraint definition languages, and extensibility opportunities that are part of the standards infrastructure.

This paper surveys spatio-temporal data exchange standardization efforts. They all have different histories and stem from different sub-communities. They thus come with both advantages and disadvantages in the eyes of a potential user. The standards not only deliver exact specifications of data structures, but also address meta-data issues. For example, it is often useful to know the validity, accuracy, and reliability of a measurement or its origin. In this sense, data exchange standards do not only serve as an exchange mechanism for data structures, but are also helpful in more advance applications like, e.g., decision support systems.

In particular, we survey the Geography Markup Language [12], which is the foremost candidate for addressing the manifold needs of the geographical community. We also survey the Scalable Vector Graphics (SVG) standard [8]. SVG is an XML-based language for the description and presentation of *two-dimensional* vector

Bulletin of the IEEE Computer Society Technical Committee on Data Engineering

and raster data. Next, we cover the ISO Technical Committee 211, which is central to the standardization of geographic data [1]. The US Federal Geographic Data Committee [3], another active player, is also covered briefly. At least a handful of other (XML-based) standards exist that do not seem to be driven by large communities, but are still of interest. For reference and completeness, these are covered briefly.

Generally, we observe a trend towards XML-based standards. A recent book [11] provides a very readable introduction to data exchange in geographical applications. Many of the XML languages covered in this paper adhere to the fundamental principle of separating content from presentation: data are encoded according to semantic criteria without constraining how the data may be presented to a human or consumed by an application.

Sections 2 and 3 cover GML and SVG, respectively. Then the ISO Technical Committee 211 and the Federal Geographic Data Committee are covered in Sections 4 and 5, respectively. Section 6 presents the "minor" standards, and Section 7 summarizes the paper.

2 Geography Markup Language (GML)

The Geography Markup Language (GML) [12] is an XML language created under the auspices of the Open GIS Consortium, whose mission statement is to facilitate the "full integration of geospatial data and geoprocessing resources into mainstream computing and to foster the widespread use of inter operable geoprocessing software and geodata products throughout the information infrastructure" [4]. GML, which could be considered the flagship effort in geoprocesssing, is a multi-stage effort that has reached version 3.0. By designing the language in multiple stages, the standardization body wants to make sure that the language evolves naturally and incorporates more and more features over time. For example, the so-called *simple* features have been defined and been integrated in the releases leading up to version 2.0.

GML's definition is based on XML Schema and tries to take advantage of its full feature set. For example, it makes it possible for users to tailor the language to their needs by means of XML schema's extensibility. So it is possible to incorporate well-defined GML schema fragments into documents that follow a user's application schema.

Therefore GML is able to describe a wide variety of geographical objects by combining its built-in data types, which implement the OGC Simple Features model [4], with the extensibility features of XML. Being an application of XML, GML is designed to take advantage of the XML standards framework; documents can be readily transformed into a variety of presentation formats, most notably vector formats and raster graphics such as maps.

Future versions of GML are planned to include additional aspects. Version 3.0 has added features for temporal GIS, including time stamps, events, and histories, as well as units of measure and the possibility of grouping features into layers, to name but a few. But since GML is used in practice and competing standards converge with it, many users consider it mature at its current stage and take advantage of its features.

For researchers, the advantages of using GML include the availability of test data sets and the incentive to use the modeling knowledge of the geographic data management community in their prototypes, which can in turn be more flexibly leveraged by other researchers and industrial partners alike. Due to the strict semantics of GML, there is also the potential to benefit from the domain modeling that is part of the standard.

3 Scalable Vector Graphics (SVG)

SVG (Scalable Vector Graphics) is a standard developed under the auspices of the World Wide Web Consortium (W3C) [5]. It defines features and syntax for describing *two-dimensional* vector and raster graphics in XML. It is thus primarily oriented towards presentation.

SVG knows three kinds of graphical objects: composite vector elements, e.g., paths that consist of lines and curves, as well as images and text. The graphical objects can be grouped, styled, transformed, and composited with various operations like nested transformations, clipping paths, alpha masks, filter effects, templates as well as previously rendered objects. Furthermore, SVG objects can be interactive and dynamic. This is achieved through event-handlers and a well-defined interface: Care has been taken to ensure interoperability with scripting languages. To this end, SVG features event handlers for user interaction, e.g., using a mouse-related events like "onmouseover", and the SVG Document Object Model for well-defined transformations and manipulations by means of scripting languages. Additionally, animations can be defined and triggered either declaratively, i.e., by embedding SVG animation elements into SVG content or by using the aforementioned scripting interface.

SVG comprises a comprehensive list of built-in graphical elements: path, text, rect, circle, ellipse, line, polyline, polygon, image, and use – a mechanism to reference SVG elements. A particularly important aspect of SVG drawings is *reusability*. Through symbol mechanisms, users may build libraries of graphical elements and re-use them without having to register at a centralized directory. To blend with the other elements of applications at client-side rendering of presentations, SVG also allows declarative definitions of filters for rasterization. Additionally, SVG provides sophisticated handling of fonts so that the original text is preserved with respect to indexing, search, and the graphical appearance intended by the authors. SVG content may be included via a stand-alone document, inline embedding, references like links, or style sheets.

The benefits researchers get from using SVG in their projects range from software engineering considerations to data re-use and integration opportunities. For example, the time it takes to prototype visual applications can be significantly reduced by using an SVG viewer instead of a custom-built user interface. Viewers are available for many different types of devices, including mobile phones, personal digital assistants, and navigators embeddable in conventional graphical user interfaces. Additionally, the tight integration and interoperability of the XML standards facilitates the development of robust applications by providing a language glue that is expressive enough to capture the semantics of the different domain models. For instance, SVG applications can directly benefit from the availability of meta-information encoded in GML documents. Furthermore, the ability of SVG to combine different source data, such as raster graphics, vector graphics, and fonts, in a single data models provides application programmers with a uniform interface and greatly reduces the need to learn and understand more than one API.

4 ISO/TC 211

ISO is the primary international standards organization for information technology. The ISO Technical Committee 211, on Geographic information/Geomatics, has as its objective to create standards for "[...] information concerning objects or phenomena that are directly or indirectly associated with a location relative to the Earth" [1]. Because the ISO/TC 211 adopts a service-oriented view of geoprocessing, its standards concern a wide variety of uses of geographic data, including methods, tools, and services that relate to data management, acquiring, processing, analyzing, accessing, presenting, and exchanging data. As a result, the ISO 19100 series standards produced by the ISO/TC 211 go well beyond plain data modeling. Further, GML, which is known as ISO 19136 in this context, is only one component of this larger series of standards developed under the auspices of the ISO/TC 211.

The interoperability of standards is of particular interest to the ISO/TC 211. For example, the committee has to ensure that the standards for position services (ISO 19116), location-based services (ISO 19132/3), and multimodal services (ISO 19134) integrate smoothly with GML.

The overall aim is to increase the availability, access, integration, and sharing of geographic information, and to enable interoperability of geospatially enabled computer systems [1] as well as to contribute to the creation of a community that supports the dissemination of geographic understanding and enables sustained development.

The focus of the ISO/TC 211 on providing an overarching standardization framework has several notable advantages. The adoption of such a framework is expected to result in lower training costs, better comparability of results, increased planning safety by relying on a mature foundation as well as the existence of a knowledge-able and experienced community for consulting and collaboration. So planning and implementation will be less error-prone, and interoperability is achieved more easily. In this sense, the standardized framework is the answer to many practical questions.

Also, the ISO/TC 211 can be seen as an umbrella intended to cover all areas native to geoprocessing. Consequently, basic familiarity with this effort provides an interesting overview of a multi-faceted research area.

5 Federal Geographic Data Committee (FGDC)

The Federal Geographic Data Committee [3] is a US governmental, inter-agency committee. The FGDC is central to the effort that develops the US National Spatial Data Infrastructure. New standards are developed by the FGDC only when there are no existing standards that are suitable for governmental use. A number of geographic data and metadata standards have been developed and are available online (see [3]).

As part of its objectives, the FGDC aims to enable the simultaneously reduction of data production costs and improvement of data quality by providing a foundation for sharing geographic data among government agencies. A complementary goal was to increase data availability, not only to the public sector, but also to academia and the private sector.

6 Standards for Specific Application Domains

We proceed to cover some efforts that are not part of the "big" standards. All of these standards are XML based. The three first are early examples of XML being deployed in the transportation industry. In addition to offering data models, these standards show how domain specialists model different application domains.

POIX (Point Of Interest eXchange Language) POIX [9] is an XML representation for places and routes. It was proposed by a Japanese consortium interested in motor vehicles, traffic, and road networks; car navigation systems is the application area of primary interest. The purpose of POIX is to provide markup for both the locations of an object and location-related information. Examples of location-related information include, e.g., opening hours and other descriptions for shopping locations. POIX also provides facilities for handling localized information such as geodetic or angle units, which may vary from country to country.

NVML (NaVigation Markup Language) NVML (NaVigation Markup Language) [13] is a markup language for describing navigation information in vehicle information systems. The main objective of NVML is the description of different types of routes. Examples include routes from a current position to a destination, the way to a shop from the nearest train station, transportation courses, sightseeing courses, and tour schedules. Special emphasis is on any-time/anywhere services for a variety of end-user devices.

RWML (Road Web Markup Language) RWML [6] has been being developed at the Civil Engineering Research Institute of Hokkaido. This language was developed as part of an effort to support smooth and safe driving under a variety of circumstances. Consequently, RWML includes standard representation syntax for weather and disaster information as well as for messages and announcements from public services.

G-XML The G-XML effort [2] is also Japanese; it has influenced developments in the context of ISO TC/211 and tries to maintain harmonization with GML as it itself evolves. The goals of the G-XML effort are similar in spirit to those of many XML applications: creating an encoding of spatial data, thus providing "a method for freely accessing and using geographic information over the Internet," according to [2]. GML 3.0 is a G-XML converged version, which implies that the underlying data models have been aligned.

Sensor Modeling Language (SensorML) SensorML [7] aims to implement an XML data exchange language for geolocating data that stems from dynamic sensors. The objective is to cover most or all types of existing sensors, including those installed on satellites, aircrafts, ships, land vehicles, and stationary platforms. It covers means for describing sensor-specific information such as location, rotation, timing, target, sensor geometry, dynamics, and radiometry. It also addresses dependencies among variables. SensorML offers database researchers an overview of a field that so far has been outside the community, and it encourages proper application data models. Again, the standard enables data sharing and data exchange as well as increased system modularity and interoperability.

7 Summary

A rich variety of standards that enable the exchange of geographic data exists. Although standardization has been initiated by committees representing different organizations and communities, in order to serve different purposes, the broad acceptance of XML appears to have caused many standards committees to join forces, which has resulted in some consolidation of standards. However, because fundamentally different application domains exist, such as the description of two- or three-dimensional visualizations and the representation of geographic or sensor data, a diversity of standards remains.

We briefly surveyed some of the geoprocessing standards that are of special interest to researchers in spatio-temporal query processing. As part of this, we pointed out in which application domains the standards can be deployed and what advantages, be it technical, educational, or social, they bring about.

References

[1] ISO TC 211. ISO TC 211 Web Page. http://www.isotc211.org/, 2003.

[2] Database Promotion Center. G-XML – Geospatial-eXtensible Markup Language Version 2.0. available at http://gisclh01.dpc.or.jp/gxml/contents-e/, 2003.

[3] Federal Geographic Data Committee. Metadata. http://www.fgdc.gov, 2003.

[4] Open GIS Consortium. Homepage. http://www.opengis.org/, 2003.

[5] The World Wide Web Consortium. Homepage. http://www.w3c.org, 2003.

[6] RWML Working Group. Road Web Markup Language (RWML). available at http://rwml.its-win.gr.jp/eng/index.htm, October 2001.

[7] Open GIS Consortium Inc. Sensor Modeling Language (SensorML). available at http://stromboli.nsstc.uah.edu/SensorML/, 2002.

[8] J. Ferraiolo (Editor). Scalable Vector Graphics (SVG). available at http://www.w3.org/TR/SVG/, 2001.

[9] H. Kanemitsu and T. Kamada. POIX: Point Of Interest eXchange Language Specification. available at http://www.w3.org/TR/poix/, June 1999.

[10] Ron Lake. Geography Markup Language (GML) 2.0 – Enabling the Geo-spatial Web. tutorial available at http://gislounge.com/ucon/ucgmlintro.shtml, 2003.

[11] Zhong-Ren Peng and Ming-Hsiang Tsou. *Internet GIS: Distributed Geographic Information Services for the Internet and Wireless Networks*. Wiley, 2003.

[12] S. Cox, A. Cuthbert, R. Lake, and R. Martell. Geography Markup Language (GML) 3.0. available at http://www.opengis.net/gml/, 2001.

[13] M. Sekiguchi, K. Takayama, H. Naito, Y. Maeda, H. Horai, and M. Toriumi. NaVigation Markup Language (NVML). available at http://www.w3.org/TR/NVML, August 1999.

References including URLs are current as of May 21, 2003.