Proceedings of the 6th AGILE
April 24th-26th, 2003 – Lyon, France

REAL-TIME INTEGRATION OF XML-ENCODED SPATIAL DATA FOR MOBILE USE

Lehto Lassi and Sarjakoski Tapani

Finnish Geodetic Institute, PO Box 15, FI-02431 Masala, Finland

1. INTRODUCTION

The European integration development obviously necessitates creation of consistent, continent-wide geospatial data services. A few European Commission-backed initiatives are already working to facilitate this process. These initiatives include projects like GINIE, GETIS, EULIS, INSPIRE, GiMoDig [1] etc. Various data harmonisation processes of the EuroGeographics also aim at same target. The main results of this work can be seen in the development of the European-wide datasets, like EuroGlobalMap, EuroRegioMap and SABE [2].

In the case of the major national datasets like the topographic map series, the starting point in the discussion about the Pan-European geospatial data provision needs to be that the data be kept and maintained in the national data model by the national authorities. Standardised access interfaces and spatial data encoding mechanisms would then be applied to achieve the desired cross-border accessibility. The INSPIRE Architecture and Standards Position Paper [3] proposes that the OGC's implementation level specification Web Feature Service (WFS), or rather the upcoming ISO-released official version of it, be used as a query interface by the data providers. According to the current WFS specification, the dataset provided by the service is to be encoded in XML format - in compliance with the Geography Markup Language (GML) specification of the OGC [4].

The challenge encountered in this approach is the question, how to integrate XML-encoded country-specific geospatial datasets in real-time. The data integration process involves at least two different aspects: the data model transformation and the coordinate system transformation. These two integration tasks are discussed in this paper, specially considering the fact that the datasets being integrated are to be encoded in XML syntax. The integration process is treated in the context of a seamless, cross-border mobile map service based on open, standards-based system architecture - thus particularly emphasising the real-time aspects of the task in hand.

2. ARCHITECTURE

The open system architecture of an on-line map service should be based on a layered service stack, in which a service would make queries to the service below it, do some processing on the data received as a response, and provide the results of this process as a service to the service layer above it. The level of detail in specifying the layers is a matter of discussion, but if the services were to be run on separate computers communicating through network, too fine-grained service definition would create a significant disadvantage in terms of overall system performance.

For the above-mentioned reason a four level system architecture is proposed (fig. 1). In the first level the data providers (e.g. NMAs) would run a Data Service providing raw spatial data in an XML –encoded form. Above the data services is the Application Service layer. The responsibilities of this layer include for instance coordinate transformations to a common reference frame and other data integration procedures, like schema transformations.

The third layer in the system architecture would be called Portal Service. The main responsibilities of this layer can be listed as: provide basic metadata service to the client, process the service requests coming from the client forwarding the request in an appropriate form to the Application Service layer below, and transform the resulting piece of geospatial data into an visual representation, according to the capabilities of the client platform in question.

On the fourth layer are finally the client applications. An advantage of the layered architecture approach is that the results can be adapted to a wide set of different client environments. For example the following three client platforms could be considered: the traditional Web browsing on a PC platform, Personal Java based clients on PDA devices, Java MIDP (Mobile Information Device Profile) clients on mobile phones.

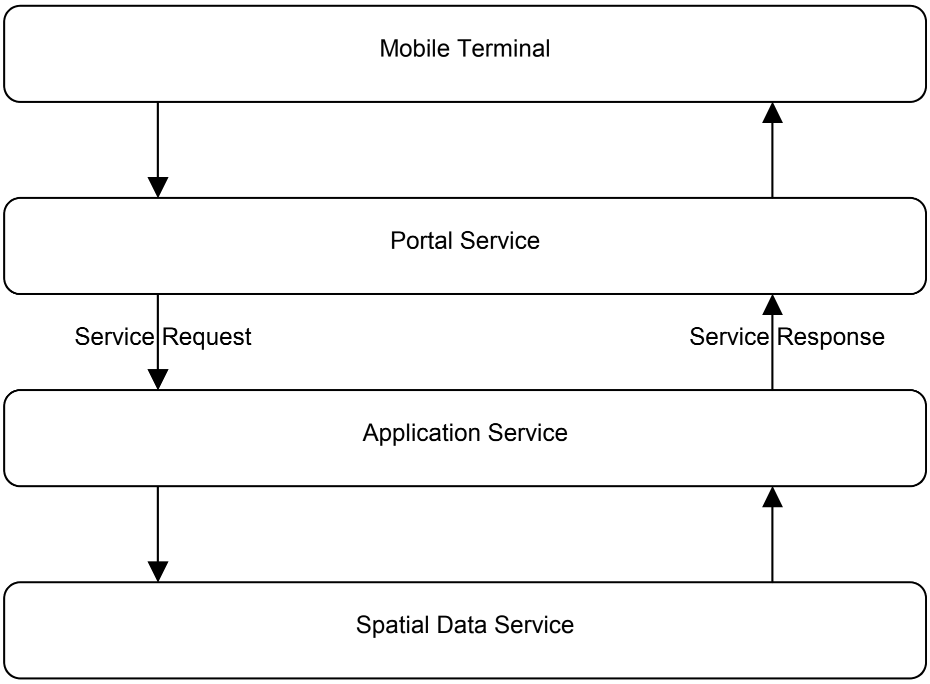

Fig. 1 Four-level service architecture for mobile map services

The communication from layer to layer must be based on internationally accepted interface standards. At the moment the main interface specifications to be considered in the architecture include Web Feature Service (WFS) specification as the access interface in the Spatial Data Service layer, Web Map Service (WMS) specification of the OGC as the query interface of the Portal Layer, and Presentation Service of the Open Location Services (OpenLS) initiative as an alternative access interface on the Portal Layer [5].

3. STANDARDISATION

Most of the existing Web and mobile map services are currently based on proprietary access interfaces and data formats. This situation largely exists due to the protective measures taken by individual service providers or software vendors in order to defend the achieved position in the market place. Extensive specifications applicable to the on-line processing of geospatial data have also been virtually non-existent. Recently the situation has changed significantly, however. Now that first standardised interface specifications have been developed, the traditionally isolated spatial software vendors and service providers start seeing mutual benefits in supporting those specifications.

The data transfer mechanisms employed in the networked environment are increasingly based on the Extensible Markup Language technologies (XML). The Open GIS Consortium (OGC) has defined an XML application also for spatial data. This specification is called Geography Markup Language (GML) [6]. Many GIS software vendors are currently developing support for this new data encoding mechanism into their products. The use of XML technology allows for structured encoding and device-independent presentation of spatial data.

OGC has also developed a map query interface specification, called Web Map Service (WMS) [7]. A WMS-compliant map service provides map data for its client applications in the form of an image. Independently of the applied internal data management solutions, a WMS-service always delivers a raster or vector map image as a response to the query message it receives. An ongoing process in the OGC, the Open Location Services (OpenLS) initiative, is currently developing a map query interface definition, called Presentation Service, specifically targeted for limited-capacity terminals.

Web Feature Service (WFS) [8] specification has been recently adopted by the OGC. A WFS-conformant service provides the client applications with real spatial data, geospatial Features in OGC’s terminology, not with pre-visualized map images like a WMS-service. The resulting spatial dataset is expressed in the form of a GML-encoded dataset.

4. DATA INTEGRATION

The data integration tasks are to be carried out on the Application Service layer of the service architecture. The resulting integrated dataset can then be easily processed on the Portal Service layer, for instance to produce visual representations of the data. In the case of the proposed European geospatial data service architecture there are two basic integration tasks. Firstly, the coordinate values presented in the individual national coordinate systems have to be transformed into a common Pan-European reference frame. Secondly, the data models applied in each country for comparable data contents, like the information normally presented on large-scale topographic maps, vary considerably from country to country. This fact necessitates schema transformations to be carried out before the datasets can be combined.

The joint efforts in the field of geodesy have produced a reference system, which is an obvious choice as the common coordinate system for a Pan-European spatial data service – this system is called the European Terrestrial Reference System 1989 (ETRS89). As a result of a thorough investigation about the existing coordinate reference systems in Europe, a Web-based service has been set up by the Bundesamt fur Kartographie und Geodäsie, Germany. Besides providing basic information about the parameters used in the coordinate systems in each country, the site also contains parameters for transforming coordinates by a seven-parameter conformal transformation from the national systems to the common ETRS89 frame [9].

Although GML is being established as a common language for encoding geospatial data in Web-related services, this does not automatically make the datasets compatible. GML provides a common geometry model and fixes in detail the XML constructs that are to be used in encoding them, but does not try to set up a common vocabulary for describing the feature level elements. For the feature structure GML only provides some very generic model concepts like Feature, FeatureCollection and Property, and states basic rules about their relationship to each other. Starting from this setting, every spatial data provider then needs to construct the local application schema. To support interoperability, GML establishes a set of rules about how those application schemas are to be developed.

As a result of the above-mentioned flexibility of the GML specification, several drastically different data models will emerge in Europe, describing the fundamental topographic datasets - even in the case of the very optimistic assumption that every country would apply GML as the data encoding standard. A flexible means for transforming XML-encoded dataset from one GML application schema to another schema will thus clearly be needed.

5. XSLT

As the number of XML-based spatial data services increases on the Web, the need to employ XML-technologies in various processes involving geospatial data becomes obvious. One of the most significant technologies developed for processing XML-encoded data is called Extensible Stylesheet Language Transformation [10], a mechanism for transforming an XML document into another XML document.

The Extensible Stylesheet Language (XSL) specification has been developed by the World Wide Web Consortium as a tool for defining presentation characteristics of an XML dataset (W3C). In connection to this work the W3C has created a specification for transforming XML documents, XSL Transformations (XSLT). XSLT is primarily designed for transforming XML documents for presentation purposes. Typical examples include dynamic creation of the table of contents, and creation of a tabular presentation of some data values in the source document. As an analogy in the geospatial data domain, XSLT could be used to transform a dataset from an application-specific data structure to the new Web vector graphics standard, Scalable Vector Graphics [11]. The other transformations being considered in geospatial applications include data model transformations, coordinate transformations, and generalisation of spatial data.

The XSLT specification is a promising tool to carry out the tasks encountered when integrating spatial datasets in real-time. Most simple integration operations are readily available. These include tasks like changing the naming system applied, grouping data from several feature classes into one class or dividing data from one feature type into several types, changing code tables etc. More sophisticated integration operations can be added via the XSLT extension mechanism. Typical examples include different coordinate manipulations, like coordinate reference system transformations, changes in geometric primitive types (e.g. area collapsed to a point) etc. The integrated datasets are written out as XML data, presented in a common GML application schema. The extension mechanism available in the XSLT process enables arbitrary, application-specific functions to be introduced into the transformation process. Several XSLT processes can also be chained together, if the task is too complicated to be expressed as one individual transformation.

An example of a simple XSLT declaration is shown in the following code sample. XSLT declarations are expressed in the form of templates. The template in the example selects all elements representing buildings ('Rakennus' in Finnish) from the source tree that match to the selection phrase (expressed in a language called XPath), then filters out all elements for which the given test phrase inside the xsl:if element do not hold. All elements inside the template not belonging to the xsl-namespace are written to the result tree. For instance, in the example the Building-element forms part of the target common vocabulary, so the effect of the transformation in this case would be a change in the naming system (from the Finnish to English terms) and a change in the collection criteria (only buildings with area ('pinta-ala' in Finnish) larger than the threshold value will be included). With the instruction xsl:apply-templates, the process continues down the XML tree. The transformation goes on until no more matching elements are found.

```
<xsl:template match="Rakennus">
   <xsl:if test="pinta-ala &gt; '200' ">
      <Building>
         <xsl:apply-templates/>
      </Building>
   </xsl:if>
</xsl:template>
```

Fig. 2 An XSLT declaration for transforming building data

6. CASE IMPLEMENTATION

The above-described four-layer system architecture is being tested in a European Union funded project "Geospatial Info-Mobility Service by Real-Time Data-Integration and Generalisation" (GiMoDig). The Finnish Geodetic Institute acts as a coordinator for the project. The other participants are the University of Hanover and the NMAs of Finland, Sweden, Denmark and Germany.

The objective of the GiMoDig project is to develop and test methods for delivering geospatial data to a mobile user by means of real-time data-integration and generalisation. The project aims at creation of a seamless data service providing access, through a common interface, to the primary topographic geo-databases maintained by the NMAs in various countries. A special emphasis will be put on providing appropriately generalised map data to the user depending on a mobile terminal with limited display capabilities.

In the GiMoDig system architecture each participating NMA provides geospatial data through the WFS interface, encoded in a country-specific XML-format (GML application schema). These datasets will be processed by a middleware service on the Application layer to integrate the pieces of data coming from individual countries into a common application schema and coordinate system. The middleware service employs the XSLT technology extensively in the process. The service is implemented as a Java servlet environment and the XSLT Processor used is a product called Xalan from the Apache community [12].

In the full paper the XSLT-based data integration transformations will be further elaborated, their performance issues discussed, and some demonstrative examples given.

7. REFERENCES

(Sites accessed Feb 28 2003)

[1] GiMoDig Project home page. http://gimodig.fgi.fi

[2] EuroGeographics home page. http://www.eurogeographics.org

[3] INSPIRE Project home page. http://inspire.jrc.it

[4] Open GIS Consortium (OGC) home page. http://www.opengis.org

[5] Open Location Services (OpenLS), home page. http://www.openls.org

[6] Geography Markup Language (GML), version 3.0. http://www.opengis.org/techno/documents/02-023r4.pdf

[7] Web Map Service (WMS), version 1.1.1. http://www.opengis.org/techno/specs/01-068r3.pdf

[8] Web Feature Service (WFS), version 1.0. http://www.opengis.org/techno/specs/02-058.pdf

[9] European Coordinate Reference Systems (CRS) home page. http://crs.ifag.de

[10] Extensible Stylesheet Language Transformations (XSLT) version 1.0. http://www.w3.org/TR/xslt

[11] Scalable Vector Graphics (SVG), version 1.0. http://www.w3.org/TR/SVG

[12] XSLT Processor Xalan-Java, version 2.4.1. http://xml.apache.org/xalan-j/index.html