FORMATTING DISPLAY OF SPATIAL METADATA BASED ON XML

Xiaodong ZHOU [a,*], Chuncheng YANG[a], Nina MENG[b]

[a]Xi'an Institute of Surveying and Mapping, Xi'an 710054,China
[b]College of Geology Engineering and Geomatics, Chang'an University, Xi'an 710054,China

KEY WORDS: XML; XSLT; GIS; Spatial Metadata; Data Transformation; Formatting Transformation

ABSTRACT:

XML is the abbreviation of Extensible Markup Language, and is technology creating structural data. When XML appeared, it was widely used in some fields, for example: e-commerce, data exchange, multimedia transmission, etc. Spatial Metadata is description about geographic spatial data and related information resources, and it helps and promotes us to effectively locate, evaluate, compare, obtain and use geographic related data when it describes and explains content, quality, condition, position and other features about geographic spatial data. XML expressing metadata, not only is exact and can provide effective methods for metadata communication between heterogeneous systems, but also can strongly support for metadata transformations and formatting display. On the basis of mutual information definition and understanding structure, with XSLT we could define relevant XSL files to let processor to do relevant transforming operation and output HTML (other required format), and browse transforming results by IE browser. In the paper, we expressed spatial metadata with XML and used XSLT technology, then realized spatial metadata transformations by defining relevant XSL files and visualization by IE browser.

1. INTRODUCTION

XML stands for Extensible Markup Language, which is a technology of creating the structured data. The term "structure" means in a clear and no ambiguity way to identify the various separate part of information. XML has been widely used in many fields such as e-commerce, data exchanged and multimedia transmission. Metadata describes the content, quality, condition and other characteristics of data. Spatial metadata is descriptive information on geospatial data and related resources. It help and promote us effective positioning, evaluation, comparison, access and use of geographic data through description of content, quality, condition, location, and other features of geospatial data. Since the development of the Internet, metadata has from methods of data description and indexing expanded to one of the indispensable tools and methods throughout the process of the network information include data discovery, data conversion, data management and use of data. Using XML to express metadata not only can accurately express metadata for providing an effective method for metadata exchange among heterogeneous system, but also support metadata conversion and formatting display, and provide a rapid navigation role for the sharing of geospatial database in distributed network. XSLT is a very important XML language in many XML workflows, and it has been an W3C recommendation.

If there is no appropriate style sheet, that is no the formative conversion, a simple XML document displays a tree structure in the Web browser, and there is not much significance. Users may not want to face this vapid data, and they need to have a personalized form of expression and color, they need a critical information can be found. On the basis of understanding of the structure and definition information of XML document of spatial metadata, we can define corresponding XSL documents to do the conversion work, output HTML or other format for direct using Internet Explorer 6.0 or other browser to explore. The main content of this paper is based on XML and XSLT technologies, research how to present complex spatial metadata content quickly, effectively, vividly to the users, so that users can easily position, evaluation, comparison, access and use of spatial data.

2. THE SPATIAL METADATA MODEL BASED ON XML

On the content systems of metadata, there are major differences between international organizations currently. Metadata systems are divided into 10 parts by CEN/TC287, 7 parts by FGDC and 8 parts by ISO/TC211 in the process of research. The spatial metadata model in the paper is developed on the basis of "National Infrastructure Geographic Information System (NFGIS) Metadata Standard". This model stores spatial metadata using XML document, and text editor will be able to complete the editing of the model, which is the model of cross-platform editing. Microsoft XML Core Services (MSXML), System.xml in the Microsoft.net framework, as well as the JAXP in the Java environment all support application development of based on XML.

2.1 Model structure

Spatial Metadata includes two parts, namely, the basic metadata and complete metadata. Basic metadata provides the minimum element sets of basic geographic data source required, which includes the thematic information, regional information, temporal information, and other information of spatial data, mainly for the network query of spatial data. Complete metadata, including eight main subsets of non-repetition using and three secondary subsets can reusing, mainly describes data sets, data sets series, elements and attributes in detail. Model uses tree structure as organizational forms of data, that each XML document is a directed tree, and only has a root node that is spatial metadata information, as shown in Figure 1.
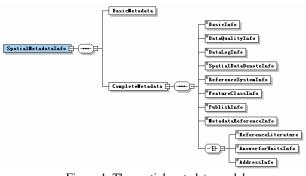
---

* Corresponding author. zhouxdm@163.com.

Figure 1. The spatial metadata model

2.2 Data modeling and verification

Data model provides the expression of logic structure of a XML document, and it designates elements that XML document can contain and the relationship between these elements. XSD (XML Schema Definition Language) is a recommended standard by W3C, and it is a very powerful and flexible norm to the definition of the data structure of XML documents, it is the important technical on designation of data model, and its support rapid expanding. XSD is from two aspects of the structure and data types to describe and bound a group of XML documents, and it is also an XML document. XSD is a powerful tool to define data structure and verification mechanism of XML, and supports namespace, a variety of data types and secretive data communication.

Using XSD can set the syntax of XML document of spatial metadata that must follow to and ensure that the format of documents is good and effective. Whether the creation of XML documents of spatial metadata or read process, we must verify it according to the definition of XSD document. Under normal circumstances, the process that in accordance with the restriction of mode to verify the XML document includes three steps: loading XML files, loading model documents and in accordance with the mode to validate XML documents.

3. FORMATTED DISPLAY TECHNOLOGY

The purpose of formatted display is to present the content of spatial metadata to users with visual effects, allowing users to see the contents of their interest, and shielding complex data structures in the interior of system. Formatted display technology mainly through using XSLT technology, an XML document is converted into different HTML data stream by the styles processor, and achieve different displaying styles in the browser, because of the styles processor has been loaded different style sheets.

3.1 About XSLT

XSLT (Extensible Style sheet Language Transformations) is the style and conversion language of the serial XML technologies. At present, XSL Transformations(XSLT) version2.0 is the latest version which was promulgated by The W3C in Nov, 2003. However, the most popular one used among most people is the XSLT1.0 standard processor released in the late 1999. XSLT is based on declaration-type programming model. In this kind of model, there is no designation about how to manipulate or what sequence to take. In XSLT, the conditions proposed will be manipulated once they are met, otherwise, there will not be any manipulations as if the relevant codes do not exist. In other words, sequence is of no importance in XSLT. We need only to add the matched templates, and the sequence will be determined by the processor. The templates matching the nodes of the higher-level in the hierarchical structure will be implicitly manipulated first. Therefore, the use of XSLT is of great flexibility in that it can provide particular formatting display for the users according to the specific templates.

3.2 Formatting conversion

The purpose of using XSLT is to provide powerful and easy-to-use style sheet syntax for XML document. Because XSLT can be re-ordered, reproduced, compressed, and increase the element of classification, conversion can be very powerful and complex. On the one hand, XSLT is independent of any other output format. In fact, even the use of XSLT output is not in XML format, when receiving XML documents, if receiving application needs comma-separated variable (CSV) format, it will be converted into the format. On the other hand, we can use the rules or mode of XSLT (Pattern) to inquire the structure and data in XML source document, when found in a specific mode (or find matching), it may do a specific operation. Having so much powerful and complex functions in XSLT, in addition to formatting, such functions should be useful in many applications. In this paper, we will make XML documents into HTML format in order to use the Internet Explorer 6.0 display.

For the XML document as the expressive object, it is considered as a tree having many nodes by XSLT, known as the source tree. The root elements and child elements composed XML document all are the nodes of the source tree. When we design XSL file to display XML document, get out of the data we need from the corresponding nodes in the source tree, and these data also form a tree, known as the result tree. Once the result tree is structured, the source tree and the result tree are independent of each other, and the operation of the result tree will not influence the corresponding data of the source tree. In this way, XSLT realizes the completely isolation between contents and manifestations of the XML document.

3.2.1 Construction of result tree

At this step, the XML tags and data of the XML document including spatial metadata information will be loaded or imported through DOM technology. DOM(Document Object Model) regards the XML document as an object, the data of which will be manipulated through the interface provided. Once the tree structure and relevant data is read by DOM objects, the system will rearrange the structure of XML tree (for example/such as, selection, operation, sorting etc.) according to the signification of the control elements in XSLT style sheet and then produce a temporary tree structure that is the so-called result tree. XSLT empowers the newly added tags in result. So, if there are other new tags and data, they will be added to the result tree after the construction of which is finished. Once the ultimate result tree is formed, the work of this step is finished and will turn to next.

3.2.2 Formatting and conversion of the result tree

Result tree and source tree both are arboreal structure. The data obtained from the source tree is presented as performance manner of node in the results tree, and retains a space of recording style, and this space of recording style is template. System processes the result tree same using DOM technology and fills specified styles of XSLT style sheet into the template.

Finally, the system converts the result tree to a specific format, such as HTML file.

3.2.3 Result showing

After the metadata of digital published map is converted, the result of the formatting conversion of HTML files can be directly showed with Internet Explorer 6.0.

3.3 XSLT Style Sheet

Through transformation of XSLT style sheet, one XML document of metadata of digital published map can show all the contents, or show only the contents of a user's interest after folding, or show with the format of print. XSLT style sheet also is the XSL file that is a good format of XML document, so it is identical in the syntax with XML document. The purpose of using XSL file is to provide a powerful and easy-to-use style sheets for XML document. Some attributes of the structure of XSL file as follows:

1. The root element "style sheet" in XSL namespace indicates that it is XSL document;
2. A series of template elements are matching with the model in objective document of XML;
3. Some function templates are separated to complete some necessary outputs;
4. Other XSL elements included in the templates accomplish other functions.

XSLT has a large number of elements, and common elements include: pattern, matching and template, these are the important components of style design. XSLT patterns using XPath syntax specify a node set, and returned result must be the node set, that is to say, not allow the type of boolean or string be returned. If the designated mode return the non-empty node set, called matching. Each model states a template, and defines a rule. The rule contains the instructions to identify what operations can be used when find a specific mode (or find matching) in the XML source documents.

XSLT transformations accept a document tree as input and produce a tree as output. From the XSLT point of view, documents are trees built of nodes. XSLT recognizes seven types of nodes, here are those nodes, along with how XSLT processors treat them:

a) The document root—This is the very start of the document.
b) Attribute—This node holds the value of an attribute after entity references have been expanded and surrounding whitespace has been trimmed.
c) Comment— This node holds the text of a comment, not including <!-- and -->.
d) Element—This node consists of all character data in the element, including character data in any of the children of the element.
e) Namespace— This node holds the namespace's URI.
f) Processing instruction—This node holds the text of the processing instruction, which does not include <?and ?>.
g) Text— This node holds the text of the node.

3.4 XSLT Processor

XSLT processor is the core to achieve formatting conversion of XML document, the construction of a good XSLT processor need to have a strong API support. Currently, the API used to achieve XSLT are: MSXML (Microsoft XML Parser), Xalan,

libxml2. MSXML is Microsoft software components to provide core XML services, and is a COM (Component Object Model) object library. It encapsulates all the necessary objects for XML analysis, and be used in the form of MSXML.DLL.

Below are some codes to achieve XSLT processor using MSXML in the VC6.0 :

CoCreateInstance(__uuidof(XSLTemplate), NULL, CLSCTX_SERVER, __uuidof(IXSLTemplate), (void**)(&pIXSLTemplate)); note: create XSLT template file

CoCreateInstance(__uuidof(FreeThreadedDOMDocument), NULL, CLSCTX_SERVER, __uuidof(IXMLDOMDocument2), (void**)(&pStyleSheet)); note: create DOM object of XSL style document

pStyleSheet->load(_variant_t(szFileXSL), &sResult); note: read XSL style document

pIXSLTemplate->putref_stylesheet(pStyleSheet); note: set the quote of XSLT template to DOM document

pIXSLTemplate->createProcessor(&pXSLPro); note: through XSLT template to create object pXSLPro of XSLT processor

4. THE EXAMPLE

The method described in this paper has been applied in developing independent software "spatial metadata management system". The system introduces MSXML4.0 parser in the VC6.0 to realize the analysis, conversion and formatting show of XML document of spatial metadata. System full use of XML technology and standards, and the workflow of system is presented in Figure 2. XML1.0 norms is mainly used to ensure the standardization of generated XML document in format; XSD is the guarantee of its effectiveness; XSLT is to achieve conversion from XML format to HTML format, so that can directly display in the browser; XQuery/XPath is used to realize rapid enquiries of spatial metadata.
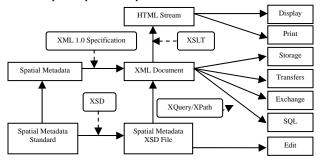

Figure 2. The workflow of system.

The system mainly realizes the functions to spatial metadata such as automatic creation, reading and writing, editing, store, formatting conversion, display and query. The structure of System is presented in Figure 3.
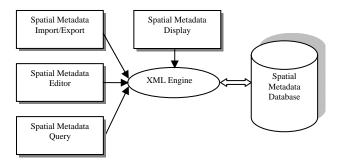

Figure 3. The structure of system.

"Spatial Metadata Display" is the module of formatting display, and its class structure is presented in Figure 4.
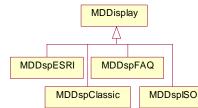

Figure 4. The class structure of the module of formatting display.

The system has been realized four kinds of showing styles, namely: ERSI, FAQ, Classic Style and ISO. The effects of four displays are presented in Figure 5.

- ESRI

- FAQ

- Classic

- ISO

Figure 5. Four kinds of showing styles.

5. CONCLUSION

This paper discusses and implements a new method of spatial metadata formatting display based on XML and XSLT technology. XML has many characteristics and applications in various fields have been more and more. Using XML file to organize spatial metadata has many advantages. XSLT technology not only solves the problem of spatial metadata formatting display, but also can achieve the formatting conversion of spatial metadata for the exchange between heterogeneous spatial data providing a new channel, and that require further study.

REFERENCES AND/OR SELECTED BIBLIOGRAPHY

Chen Huirong, You Xiong., 2003. Researches on Geo2spatial Metadata and Its Related Techniques[J]. Journal of Institute of Surveying and Mapping, 20(4), pp. 290-292.

Chinese Managing Center of 21st Century Agendum., 1999. A Study on the Spatial Metadata Standards of China. Beijing: Science Press.

FGDC. Content Standard for Digital Geospatial Metadata V2. 1998,6. http://www.fgdc.gov/constan.html

H.M.Deitel P.J.Deitel., 2002. XML Programming Technology. Beijing: Tsinghua University Press.

http://msdn.microsoft.com/xml.

http://www.oreilly.com/catalog/xslt

http://xml.apache.org.

http://xmlsoft.org.

National Fundamental Geography Information System Standard. http://nfgis.nsdi.gov.cn

O'Reilly&Associates. XSLT.

Peter G.Aitken., 2002. XML—the Microsoft Way(ISBN 0-201-74852-5). Addison Wesley,Inc.

Steven Holzner., 2003. Real World XML. New Riders Publishing.

W3C.XSLTransformations(XSLT)version2.0[EB/OL]. http://www.w3.org/TR/2003/WD-xslt20-20031112

XSL Transformation. http://www.w3.org/TR/xslt

Xiaodong ZHOU （1977- ）, male, Assistant Researcher, majors in cartography and GIS, orientation is spatial data mining.
Tel: +86-29-84749138, E-mail:zhouxdm@163.com