

Capítulo 3. Trabajos Relacionados

Este capítulo presenta trabajos que de alguna manera influyeron o determinaron las bases para el desarrollo de este proyecto.

3.1 Usando XML para la distribución de datos geo-espaciales

Este artículo fue presentado por Zhang et. al [2000] en el ACM Symposium on Advances in Gis en el 2000. El principal interés en este artículo consiste en la arquitectura que presentan y el uso de XML con información geo-referenciada como protocolo de comunicación entre sus componentes.

El artículo presenta un prototipo para encapsular y visualizar datos geo-referenciados en un ambiente distribuido usando tecnología XML. Uno de los objetivos del proyecto se basa en la idea de que el uso de datos geo-referenciados permitiría en un momento dado incorporar información particular a un mapa geo-referenciado de dominio público. De esta manera los clientes pueden aportar datos que se incorporan a un mapa base, común para todos los usuarios.

Se ha demostrado que XML es un buen mecanismo para la integración de información in Internet. XML generalmente se utiliza para estructurar información y para presentarla en HTML se utilizan plantillas de traducción. Sin embargo, hay muchos elementos (particularmente información visual) que no pueden desplegarse en HTML. Por esta razón el sistema utiliza un applet para visualizar la información espacial.

Este artículo propone un prototipo para integrar y visualizar información geo-referenciada en un ambiente distribuido (WWW). Este sistema está formado de tres componentes (ver figura 3.1) que incluye: un encapsulador de datos de información geo-referenciada, un mediador y un cliente como una interfaz de visualización. En este prototipo XML se utiliza como protocolo de comunicación el encapsulador se encarga de recuperar los datos entre sitios Internet distribuidos y entonces se provee la información a los mediadores como documentos XML. También XML es utilizado entre el mediador y los clientes.

Como detalle de implementación se utilizaron Java Servlets para traducir la información recopilada a documentos XML. La información originalmente se encuentra almacenada en archivos Shapefile para ArcView. Se tiene pensado realizar esta conexión con una base de datos también. El componente encapsulador consiste de un Servlet que recupera la información de los diferentes sitios distribuidos y la convierte a un formato en XML a petición de un cliente. Este mediador también puede analizar, integrar o modificar estos documentos en XML para construir nuevos documentos de acuerdo a las necesidades del cliente. Al momento de que el cliente recibe los datos del servidor (Servlet), la aplicación interpreta el documento XML y lo dibuja en un applet contenido en un browser.

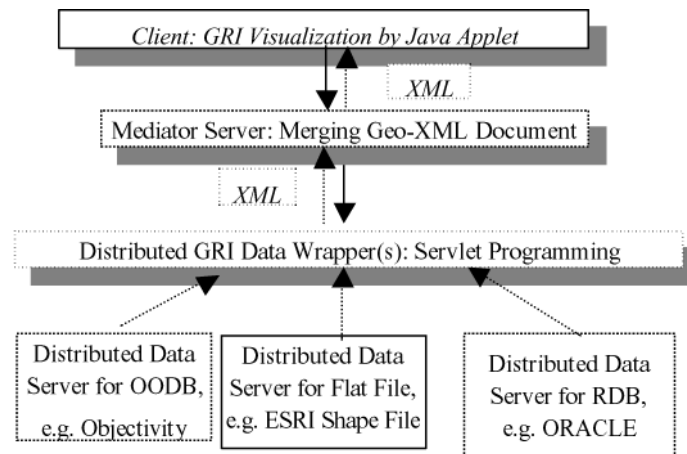


Figura 3.1 Arquitectura

Para el uso de información geo-referenciada se han hecho esfuerzos continuos para integrar múltiples fuentes y múltiples formatos. Uno de estos estándares es propuesto por el OpenGIS Consortium. En su artículo mencionan que sería muy interesante si toda la información tanto descriptiva como geo-referenciada fuera expresada en un formato XML; las tareas de integración serían más fáciles. Hasta donde les fue posible revisar no se encontró ningún trabajo que integrara y visualizara información geo-espacial usando XML.

Los metadatos juegan un rol importante en aplicaciones de información geo-referenciada. En este artículo se menciona también que existen propuestas tanto de OpenGis como de FGDC (Federal Geography Data Committee) y de ISO/TC 211. Estas propuestas separan los metadatos de los datos pero es posible con XML integrar ambos en un solo documento.

Como mencione anteriormente este proyecto utiliza XML como protocolo de comunicación entre el encapsulador y el mediador encargado de estructurar la información. También se utiliza XML entre el applet en Java y el mediador.

Basándose en sus experiencias el modelo basado en árbol de DOM (Document Object Model) es más apropiado para la integración de información geo-referenciada con XML. La razón es que un objeto geo-referenciado generalmente contiene niveles múltiples de atributos que muy difícilmente se pueden ordenar en un solo elemento de XML. En el método basado en SAX (Simple Api for XML) cada elemento es procesado secuencialmente FIFO y por lo tanto es muy difícil mantener la jerarquía de los objetos geo-referenciados. Por el contrario el modelo DOM permite tener la representación del documento en forma de árbol en memoria. Esta es una buena representación para los múltiples niveles de atributos de un objeto geo-referenciado.

El prototipo final contiene las siguientes características:

- Se implementaron tres entidades: punto, polilínea y polígono y se adoptó la definición de un DTD(Data Type Definition) subconjunto de Oracle GeoXML.
- El formato de la información se utilizó fue Shapefile.
- Se utilizaron clases desarrolladas en la universidad de Leeds, UK para leer en memoria los archivos y transformarlo a documentos XML
- El encapsulador es un servlet que recupera diferente número de documentos XML.
- El mediador es otros servlet que integra los documentos enviados por el encapsulador de acuerdo a búsquedas en metadatos sobre estos.
- El applet se comunica con el mediador, analiza el documento XML recuperado y lo visualiza en 2D. Estas funciones incluyen zoom in y zoom out. (ver figura 3.2)

El cliente puede ser cualquier programa que pueda interpretar los documentos XML o puede ser un objeto incrustado en un browser como un applet.

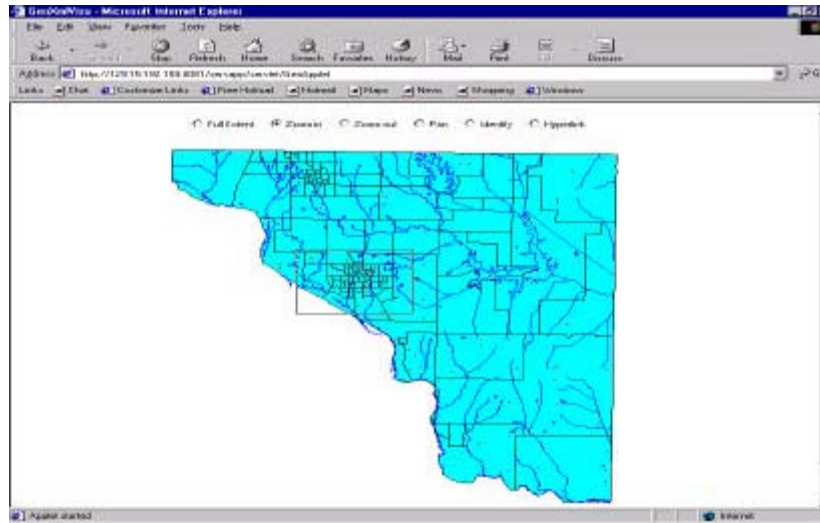


Figura 3.2 Applet para visualizar mapa.

Esta arquitectura permite la implementación de un sistema de información geográfico a gran escala basado en Internet que ofrezca consultas complejas y análisis espaciales el cual es su objetivo a largo plazo.

Las razones que mencionan para utilizar XML como protocolo de intercambio común para información geo-referenciada son las siguientes:

- Un protocolo de intercambio de información basado en texto permite una independencia de plataforma y una fácil implementación
- El proveedor de datos puede ser cualquier sistema, manejando cualquier formato siempre y cuando pueda generar documentos XML
- El documento generado en XML es independiente de aplicaciones ya que cualquier aplicación puede usarlo mientras pueda reconocer el documento
- Es más eficiente almacenar y transferir datos geo-referenciados integrados en un formato XML porque generalmente hay muchos atributos en una base de datos que no están asociados a su objeto geográfico correspondiente
- Una de las ventajas de utilizar documentos de XML, es posibilidad de validarlo al momento de su construcción gracias a su DTD (Data Type Definition)
- El cliente puede ser cualquier programa que pueda interpretar los documentos XML o puede ser un objeto incrustado en un browser como un applet.

3.2 GML, un caso práctico de visualización

En esta sección se describirá un trabajo que muestra un esfuerzo por traducir documentos en XML a documentos SVG. SVG son las siglas de Scalable Vector Graphics [W3C 2001], esta especificación permite la representación de imágenes vectoriales en 2D utilizando XML. Entre sus características más importantes se encuentran las siguientes:

- Figuras básicas: puntos, líneas, círculos, paths
- Múltiples niveles de acercamiento: zoom
- Texto no se altera: búsquedas, copiar y pegar
- Independencia de dispositivo de salida: monitor o impresora
- Control de color, estilos y patrones
- Interactividad y animación por medio de lenguaje script

Este trabajo es el resultado de sesiones de trabajo de verano 2001 en el Institute for Geoinformation de Vienna University of Technology en Austria. No presentan todavía ninguna publicación pero la presentación y la página del desarrollo del proyecto se encuentra en: <http://www.geoinfo.tuwien.ac.at>. El interés en este proyecto consiste en el trabajo que realizaron con plantillas de traducción con XSLT para generar una representación gráfica de un documento en XML utilizando el GML (Geographic Markup Language) propuesto por OpenGis Consortium [www.opengis.org].

Este trabajo consiste en la definición de plantillas en XSLT [W3C 2001], que es un lenguaje de traducción estándar para XML, para transformar documentos en GML a un documento SVG. GML es la especificación en XML propuesta por el OpenGis Consortium y su objetivo es integrar información descriptiva y espacial en un mismo documento. GML contempla entidades espaciales simples, puntos, líneas y polígonos y el uso de sistemas de referencia espacial. GML se encuentra en dos versiones 1.0 y 2.0. La diferencia principal es que el primero está basado en DTD y el segundo en Schemas [W3C 2001].

XSLT es una especificación del World Wide Web Consortium [W3C 2001] y se encuentra en su versión 1.0. El objetivo de XSLT es transformar un documento en XML.

XSLT es una especificación en XML e incluye componentes que nos permiten reconocer tags y aplicarles una instrucción, ya sea que esta consista en obtener su información o en manipularla. El manejo de plantillas asemeja a un lenguaje script. Scalable Vector Graphic SVG es otra especificación del W3C en su versión 1.0 permite definir en un documento figuras simples y complejas con atributos enfocados a su presentación vectorial en 2D. En SVG es muy importante el uso del formato vectorial para la creación de figuras. También incluye el manejo de texto y animación. Las imágenes que crea SVG no son raster: se puede hacer zoom a múltiples detalles, el texto es extraíble y permite copiarse, tiene gran variedad de atributos: color, textura, tamaño, transparencia, y su principal ventaja es su estandarización.

OpenGIS propone en su especificación de GML que se utilice SVG o X3D para traducir documentos en GML a su representación visual, esta traducción se puede obtener de manera directa con la tecnología propuesta por W3C que son las plantillas de traducción en XSLT.

XSLT tiene las siguientes características:

- es un documento XML (documento de texto),
- no necesita ser compilado (aunque se puede hacer) es interpretado,
- ofrece una alternativa a la implementación de un programa traductor,
- manipulación directa de XML,
- funciones adicionales de manipulación de cadenas,
- formato de números, etc.,
- es estándar,

En el sitio de Internet del laboratorio se presenta el trabajo de Clemens Nothegger donde muestran cómo es posible a través de una plantilla en XSLT transformar un documento en GML v 1.0 a SVG. Las entidades que manejan son las siguientes: Puntos, polígonos y texto. El ejemplo que manejan es un documento GML que contiene información catastral de una manzana. Entonces podemos encontrar en el documento, polígonos que representan

las propiedades, puntos que representan los límites de las propiedades, y el nombre del dueño y la dirección.

La plantilla de traducción define primeramente el tamaño de la imagen final, después obtiene del total de coordenadas a utilizar (ya sean de líneas o puntos) el “bounding box”. Estos valores se almacenan en variables globales ya que sirven posteriormente para escalar las coordenadas.

```
<featureMember typeName="Parcel">
  <Feature typeName="Parcel">
    <name>Bauplatz 5</name>
    <property typeName="EZ" type="string">2139/14</property>
    <property typeName="Owner" type="string">Stadt Wien</property>
    <geometricProperty typeName="Boundary">
      <Polygon srsName="EPSG:XXXX">
        <outerBoundaryIs>
          <LinearRing>
            <coordinates>9557.743,5334329.953
9566.831,5334326.288 9635.264,5334302.503 9631.897,5334297.010
9588.745,5334226.630 9535.357,5334245.186 9526.168,5334248.574
9557.743,5334329.953</coordinates>
          </LinearRing>
        </outerBoundaryIs>
      </Polygon>
    </geometricProperty>
  </Feature>
</featureMember>
```

Para extraer la información descriptiva, que en este caso es la dirección y el nombre del dueño, se busca el elemento Feature para obtener su atributo name y owner. Para determinar la ubicación del texto obtiene el valor medio de las coordenadas “x” y “y” de las coordenadas en el elemento Feature que en este caso pertenecen a un polígono. Crea al momento de obtener esta información un elemento text de SVG y define el estilo y la alineación (ver figura 3.3-1).

Para los polígonos se identifica el elemento LinearRing, como podemos ver en el fragmento de código, este elemento está contenido en el “outerBoundaryIs” del elemento polígono. Este polígono es la propiedad geométrica del elemento Feature de tipo parcela. Al momento se recuperan las coordenadas y se procesan de manera recursiva para obtener un path al estilo SVG. Un path en SVG consiste en un punto inicial y los siguientes puntos definidos en términos del punto anterior. Al momento de tener el path completo se crea el

elemento SVG path con atributos de trazo como color y ancho, además del relleno (ver figura 3.3-2).

Para obtener los límites de la parcelas se recuperan nuevamente las coordenadas de los polígonos, pero ahora por cada pareja de x-y se crea un elemento círculo en la posición x-y con atributos de trazo y relleno (ver figura 3.3-3).

Para generar la imagen en el documento SVG es necesario transformar todas las coordenadas del documento GML a las coordenadas de pantalla, esto se realiza también en la plantilla.

El resultado final lo podemos ver en la Figura 3.3.

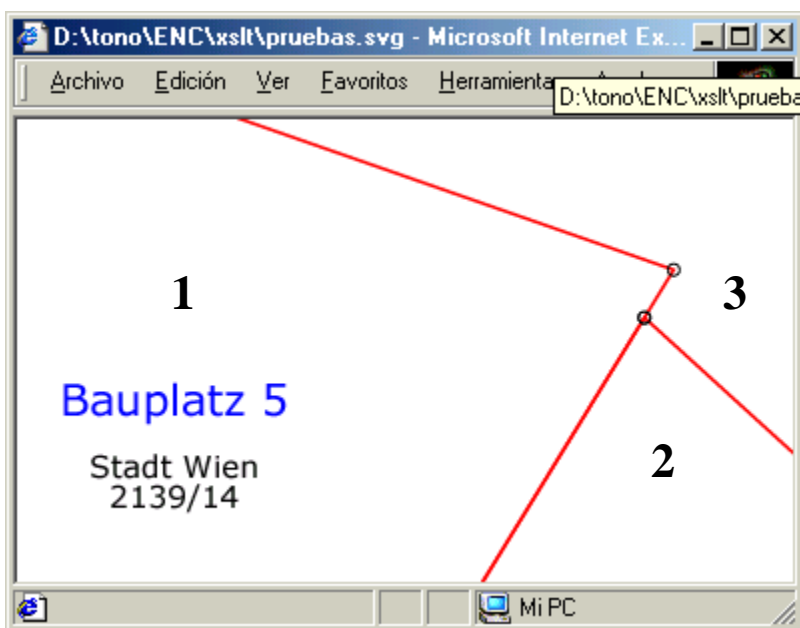


Figura 3.3 SVG generado con XSLT.

De los resultados de este trabajo obtuve un listado de aspectos a considerar en el caso de utilizar GML y en general XML para el manejo de información geo-referenciada.

GML

- Esta definido por OGC, universidades y compañías reconocidas,
- es un lenguaje para transmisión y almacenamiento, no para presentación,
- maneja entidades simples (puntos, líneas, polígonos) ,

- útil para servidores de mapas en Internet,
- estándar internacional y específico del área,
- integra información espacial con no espacial,
- puede ser transformado a través de XSTL.

XML

- Información Estructurada jerárquicamente en un documento de texto,
- XML es texto, aunque legible, su ventaja esta encaminada a su lectura por parte de aplicaciones,
- puede generarse documentos muy grandes, pero estos se comprimen hasta en un 90% con gzip,
- se basa en SGML definido en 1986,
- no necesita licencia para utilizarse,
- permite una independencia de plataforma,
- existen muchos desarrollos, librerías, aplicaciones y servidores que lo manejan.

3.3 Un SIG en 3D basado en Internet

Este trabajo presentado en ACM GIS 98 por VolkerCoors y Sascha Flick presenta un prototipo basado en Java y VRML para interactuar con un sistema de información geográfica en tres dimensiones por Internet. Este trabajo llamó nuestra atención por su interfaz vía Internet, el uso de VRML y los fundamentos teóricos planteados para la implementación de un sistema de información geográfica tridimensional.

Usando un modelo de datos que soporta tanto geometría como topología tridimensional, se desarrolló un sistema de información geográfica con una interfaz vía Internet. Se utilizó para esta interfaz, java y VRML. Debido a la gran cantidad de información que pueden llegar a tener los modelos tridimensionales se definieron múltiples niveles de detalle para cada modelo y un protocolo de transmisión para los escenarios en VRML.

El trabajo presenta la necesidad de un SIG en 3D y sus posibles aplicaciones. En particular presenta un sistema de apoyo a la planeación urbana. Internet simplifica el acceso a la información por una gran variedad de usuarios, presenta además una oportunidad para presentar y discutir proyectos. Critican abiertamente la confusión, intencional o no, entre sistemas de visualización tridimensionales y sistemas de información tridimensionales, ya que las herramientas actualmente desarrolladas no permiten interactuar de manera intuitiva con los modelos y realizar consultas sobre estos. Un sistema de consultas espaciales en tercera dimensión no sólo contempla aspectos geométricos sino todo tipo de relaciones entre los objetos. Estas relaciones espaciales pueden ser:

Métricas basadas en la geometría euclidiana (eg. cercanía) o

Topológicas como adyacencia, conectividad o pertenencia.

Un verdadero SIG-3D debe contemplar lo siguiente :

- Manejo completo de información tri-dimensional vectorial
- Integración de información raster como fotografías e imágenes de satélite
- Visualización y modelado de diferentes representaciones de la información
- Relaciones topológicas (tri-dimensionales)
- Consultas información descriptiva y tri-dimensional

En este proyecto se desarrollo un sistema enfocado al modelado de información tri-dimensional y un sistema de consulta para Internet que provee acceso e interacción con una base de datos tri-dimensional. Para este último desarrollo se utilizo Java para la interfaz y CORBA para la comunicación entre el cliente y el servidor de datos. Se implementó también una extensión para integrar diferentes niveles de detalle al modelo de datos.

La clase principal se denomina geo-object e incluye toda la información descriptiva, geométrica, topológica y de niveles de detalle. Cada geo-objeto puede manejar diferentes vistas aunque las consultas geométricas se realizan sobre su representación geométrica original.

El objetivo de la interfaz por Internet fue poner a disposición de una gran variedad de usuarios el SIG-3D. Por este motivo el software del cliente fue realizado utilizando Java y

VRML. La ventaja que ofrece Java, es que es posible cargarlo desde cualquier máquina y ejecutarse de manera local, evitando de esta manera tener que instalar cualquier otro programa. Los geo-objetos se representan en un nodo para VRML definido para este caso. Una vez que se genera el mundo VRML el usuario puede navegar por este (ver Figura 3.4). Sin embargo el objetivo principal está en interactuar con los geo-objetos, una primera consulta que se puede realizar sobre estos objetos es solicitar información adicional sobre un objeto seleccionado. Esto se puede realizar con una referencia a una página HTML o con una consulta a la base de datos. Otra manera de consultar la información consiste en definir características de objetos, de manera que la respuesta nos muestra los objetos que cumplan este criterio.

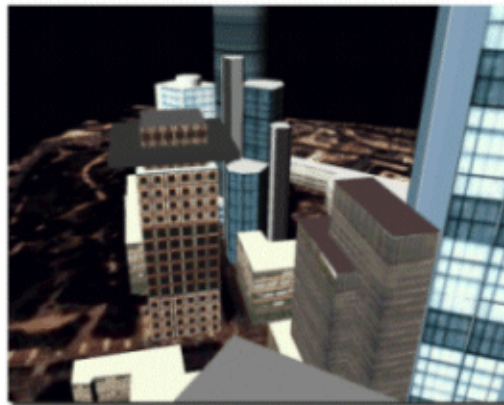


Figura 3.4 Interfaz de 3D-GIS

3.4 Construyendo Repositorios Espacio-temporales

Este artículo fue presentado en el 8th ACM Symposium on GIS por Brodeur et al. [2000] y presenta una propuesta de modelado de información geo-espacial en UML y tomando como base estándares internacionales de OpenGis e ISO. Este artículo resalta la importancia de modelar y documentar adecuadamente el contenido de un repositorio de datos espaciales, definiendo esquemas, diccionarios y metadatos.

La disponibilidad de una gran diversidad de información geo-espacial hace difícil a los usuarios encontrar la que se ajusta a sus necesidades. Es necesario conocer a detalle y un mejor entendimiento de los datos para definir la funcionalidad de una colección. Los metadatos, definidos como datos sobre los datos permiten estructurar los datos de acuerdo a

contenido, fuente de datos, proceso de adquisición, calidad, resolución y formato. De la misma manera si se agrega a esta información la descripción de la semántica, geometría, temporalidad e integridad de los objetos se puede obtener un buen diseño de base de datos. Este artículo presenta un modelo para representar el contenido de una base de datos que contemple los elementos mencionados y que se ajuste a los estándares internacionales en geomática. Esta herramienta se denomina Perceptory y utiliza el diagrama de clases de UML extendido con estereotipos espaciales y temporales.

Un repositorio de datos geo-espaciales se define como colecciones de metadatos estructurados de manera que permitan obtener la semántica y la estructura de los datos contenidos en una base de datos, junto con sus propiedades geométricas y temporales. Entre estas características se puede mencionar, los nombres y definiciones de las clases de objetos, los nombres y definiciones de sus atributos, descripción de los valores y dominios de estos atributos, tipos de datos, operaciones, geometría (figuras y especificaciones, proyecciones), tiempo (dimensiones y especificaciones, unidades, resolución), relaciones, limitantes. Estas características constituyen un esquema y el diccionario. El esquema es una representación abstracta de un objeto. El diccionario presenta la colección de metadatos de la información almacenada en la base de datos.

Los repositorios deben permitir recuperar la descripción del contenido de los objetos para determinar su utilidad, o la incorporación total o en parte a nuevas colecciones. Para evitar problemas de interoperabilidad y portabilidad los repositorios de datos geo-espaciales deben establecerse de acuerdo a estándares relacionados con geomática y las tecnologías de información (por ejemplo OpenGIS, ISO/TC 211).

El trabajo de Perceptory se compone de los siguientes tres componentes:

- una descripción del proyecto,
- metadatos del catálogo y esquema,
- un esquema y diccionario.

El artículo presenta el esquema conceptual basado en UML que incorpora estereotipos para el manejo de propiedades espaciales y temporales. Con Perceptory es posible definir el contenido y la estructura de un repositorio de datos geo-espaciales con un Metamodelo.

Este trabajo se enfoca en la importancia de documentar el contenido de una base de datos geo-espacial. Una base de datos sin una descripción completa de su contenido no es útil para otras personas y organizaciones. Los esquemas conceptuales y diccionarios de datos son componentes claves de los repositorios geo-espaciales. La documentación de los datos geo-espaciales debe ajustarse a los estándares internacionales en geomática para garantizar su interoperabilidad. El acceso a los repositorios vía Internet es un aspecto que se debe estudiar, así como el uso de XML para la definición de diccionarios.

El objetivo de este capítulo fue recopilar experiencias de proyectos relacionados de alguna manera con el trabajo de esta tesis. Por ser una propuesta que abarca muchos aspectos los trabajos relacionados son diversos. Destaca el uso de nuevas tecnologías como son XML, XLST, VRML y UML lo que determinaron muchos aspectos del diseño final del sistema. En el siguiente capítulo se presentará el análisis y diseño del proyecto GISELA-X3.