

Prototipo de GeoPortal de Mapas Antiguos.

M.A. Manso, M.A. Bernabé

Departamento de Ingeniería Topográfica y Cartografía.
Universidad Politécnica de Madrid
Campus Sur, Carretera de Valencia, km 7. 28031- Madrid
m.manso@euitto.upm.es, mab@euitto.upm.es

Abstract: En este documento se tratan la problemática y la metodología utilizada para crear dos Prototipos de GeoPortales con capacidad de búsqueda, descripción y visualización de la cartografía antigua digitalizada y catalogada procedente de varias fuentes. Se describen y analizan los tiempos y dificultades de las siguientes tareas: exploración óptica, rectificación geométrica y georreferenciación, la compresión de imagen, la creación de los metadatos, la configuración del servidor de mapas Web (WMS) con los metadatos ISO19119 del servicio, el registro del servicio WMS en un catálogo de servicios OGC, la carga de los metadatos ISO19115 asociados a los datos, en un catálogo y el diseño o personalización de los portales. También se describen los componentes software WMS de la Universidad de Minnesota [8], Catálogo de servicios de Ionic [6], GeoPortal Open Source GeoNetwork [4], editor de metadatos CatMDEdit [1] y otras herramientas de desarrollo propio utilizadas para la automatización de tareas.

1. Introducción.

Actualmente existen muchas iniciativas a distintos ámbitos geográficos: nacionales, regionales, locales y corporativos para crear de Infraestructuras de Datos Espaciales (IDE) que satisfagan Estándares Internacionales. La mayoría de estas IDEs están destinadas a la explotación de la Información Geográfica (IG) de naturaleza vectorial: mapas topográficos, coberturas temáticas, inventarios de datos georreferenciados, etc... También pueden incluir información de naturaleza matricial procedentes del proceso de orto rectificación de fotografías aéreas o de imágenes de satélite. Por lo general, en todos estos casos, se está haciendo un considerable esfuerzo en la cartografía más moderna o actual, dejando en los armarios, archivadores o en los cajones de las organizaciones, la IG antigua con mucha riqueza histórica, absolutamente necesaria para realizar diversos tipos de análisis.

En este documento se sintetizan los esfuerzos realizados para construir y alimentar con datos y metadatos, un GeoPortal Web que utiliza componentes conformes con las especificaciones definidas por el Open Geospatial Consortium (OGC). Las funcionalidades previstas para el Geoportal son:

2 M.A. Manso, M.A. Bernabé

- a.- Capacidad de búsqueda de I.G. en base a los metadatos que describen los datos o en base a los metadatos que describen los servicios,
- b.- Acceso a la I.G. a través de clientes de servicios WMS que permitan visualizar de forma combinada datos de distintas fuentes y
- c.- El acceso a los propios metadatos.

El conjunto de datos objetivo de este estudio son mapas impresos y que se encuentran depositados en la Cartoteca José Martín López de la Escuela Universitaria de Ingeniería. Técnica Topográfica de la UPM.

El documento se estructura de la siguiente forma:

En primer lugar se describe el proceso de exploración óptica, indicando el modelo de escáner, la resolución de exploración y el formato gráfico que se ha utilizado en esta fase. Como consecuencia de este trabajo se obtiene un indicador de costes relacionado con consumo temporal del proceso.

En segundo lugar se describe el proceso de rectificación geométrica, identificando la aplicación utilizada, el tipo de transformación polinómica utilizada en la rectificación, el número de puntos de control utilizados y el formato de almacenamiento de la imagen ya rectificadas. Como consecuencia de este trabajo se obtiene el segundo indicador relativo al coste temporal de esta operación.

En tercer lugar se describe el proceso de creación de los metadatos, indicando con qué herramientas se han creado los mismos. Para este trabajo se han utilizado dos metodologías diferentes con el objeto de identificar posibles discrepancias en los costes temporales y económicos que de esta fase se puedan derivar. El tercer indicador que se obtiene es el tiempo y la cualificación técnica del personal necesario para crear los registros de metadatos.

En cuarto lugar se describe el proceso de puesta en marcha del servicio WMS que ofrecerá los mapas digitalizados, rectificadas y descritos en forma de metadatos en los apartados anteriores. Se describirá la metodología utilizada para crear de forma automática el fichero de configuración para el Servidor de Mapas de la Universidad de Minnesota, incluyendo los metadatos del servicio para cada capa (mapa).

En quinto lugar se describen los trabajos realizados para adaptar y preparar el GeoPortal basado en el proyecto OpenSource GeoNetwork. Estos trabajos han consistido fundamentalmente en la traducción a Español de los mensajes de las interfaces, de la ayuda, la inclusión de las numeraciones de los campos conforme a ISO19115, el marcado de los campos que pertenecen al Núcleo Elemental de Metadatos (NEM), la inclusión de algunas informaciones de ayuda relativas a la cumplimentación de ciertos campos. Otros trabajos realizados sobre el proyecto han permitido la carga de los metadatos y las previsualizaciones de los mapas en modo encadenado secuencial, o el borrado de intervalos de registros.

En sexto lugar se describe la preparación del GeoPortal basado en los productos de Registro de catálogo y cliente de Catálogo de Ionic. Se describe como se personaliza la interfaz del cliente, como se registran los servicios WMS, WFS, WCS o Contextos en el catálogo de servicios y como este proceso, de forma automática, recupera los metadatos del servicio de cada una de las capas. Se describen los trabajos que se han tenido que realizar en relación con el acceso a los

metadatos que describen los contenidos de las capas en los metadatos del servicio, para el proyecto OpenSource Mapserver de la universidad de Minnesota.

Para finalizar se presentan las conclusiones del trabajo, se presentan las líneas de actuación futuras y se presentan las referencias bibliográficas.

2. Exploración óptica de los mapas.

El proceso de digitalización de los mapas en formato papel que se ha realizado en este prototipo, se ha utilizado un escáner Océ CS4035 formato A0+, que es el disponible en las instalaciones de la Cartoteca y el que dispone de más calidad. Las principales características técnicas del mismo son: 800 ppp de resolución geométrica y 24 bits de resolución radiométrica.

Los conjuntos de mapas que se han explorado son:

- La serie 8C (1:800.000) del Servicio Geográfico del Ejército (SGE), compuesta por 9 hojas.
- La serie 4C (1:400.000) del SGE, compuesta por 30 hojas.
- La serie 2C (1:200.000) del SGE, compuesta por 91 hojas.
- Un conjunto de 414 cartas náuticas del Instituto Hidrográfico de la Marina (IHM) a distintas escalas que proporciona cobertura global a pequeñas escalas y cobertura de puertos marítimos a grandes escalas.

La resolución de exploración adoptada es 300 puntos por pulgada, el formato gráfico adoptado para almacenar temporalmente el mapa hasta ser rectificado geoméricamente es Tiff siendo el tamaño medio de los archivos generados con este formato es 200MBytes.

El tiempo que consume el proceso de exploración, depende de la velocidad del dispositivo, del interfaz de comunicación con el ordenador, del tamaño del soporte, de la resolución de exploración y del formato de archivo de almacenamiento. Teniendo en cuenta que el equipo es capaz de explorar a una velocidad de 0.8m/minuto a 400 ppp, que el interfaz de comunicación es USB 1.1, que el formato de almacenamiento es TIFF, solo queda por definir en esta ecuación el tamaño del documento.

Así para los mapas de las series cartográficas del SGE cuyo tamaño está normalizado, el tiempo medio de exploración y almacenamiento es de 5 minutos por mapa. Para los mapas procedentes del IHM, los tamaños son muy variables, motivo por el cual no el valor medio del tiempo de exploración es de 6 minutos con una alta varianza (2).

3. Rectificación geométrica de los mapas digitalizados.

La rectificación geométrica de las imágenes exploradas se ha realizado con la aplicación GlobalMapper [10]. Esta aplicación permite visualizar una gran variedad de formatos de ficheros (más de 23) matriciales y vectoriales (más de 20) que

contengan la información necesaria para definir el Sistema de Referencia Espacial por si mismos, mediante una consulta al propio usuario o en último lugar remitiéndole al procedimiento de rectificación de la imagen con una herramienta sencilla e intuitiva.

La herramienta dispone de los siguientes algoritmos de rectificación: Lineal, Afín, Polinómico y Triangulación. Los algoritmos de re-muestreo disponibles son: vecino más próximo y bilineal. Es capaz de manejar un elevado número de datums, proyecciones y zonificaciones.

El algoritmo de rectificación seleccionado es Afín con 4 puntos de control, de manera que existe un mínimo de redundancia para realizar un ajuste de mínimos cuadrados. El algoritmo de re-muestreo seleccionado es bilineal para suavizar la imagen al rectificarla. Los puntos de control se toman en los cruces de las rejillas de los mapas y las coordenadas se leen sobre los mismos. El formato de salida de las imágenes una vez rectificadas es GeoTiff. Se ha adoptado este formato no comprimido de imagen por almacenar la información necesaria de georreferenciación y la del Sistema de Referencia Espacial dentro del archivo imagen. Posteriormente este fichero se transforma a un formato comprimido como ECW o JPEG2000, utilizando herramientas de libre distribución entre las que se encuentran ECW Compressor [2] y la librería gdal[3].

Para realizar estos procesos prestando un mínimo de atención a los residuos que introduce cada punto de control en el tipo de ajuste es necesario destinar por cada mapa un tiempo medio de 6 minutos. Los procesos de compresión de las imágenes GeoTiff a ECW o JPEG2000 se pueden automatizar y ejecutar por lotes en modo no asistido, razón por la cual este tiempo puede despreciarse.

4. Creación de los metadatos relativos a los mapas.

Para crear los metadatos asociados a las imágenes rectificadas procedentes de la exploración óptica de los mapas se han utilizado dos técnicas distintas.

- La primera ha consistido en crear uno a uno los registros de metadatos asociados con cada imagen utilizando la herramienta CatMDEdit y la utilidad de replicación de registros de metadatos con el objetivo de optimizar tiempos.
- La segunda ha consistido en diseñar una tabla (Microsoft Access o Excel) que contenga las características que puedan recogerse de forma individual desde los mapas como son: el identificador, el título, la fecha, la escala, el depósito legal o isbn, etc...), la herramienta GuardaMeta (todavía en desarrollo) y un conjunto de utilidades de automatización todas ellas de desarrollo propio.

Las metodologías utilizadas son las siguientes:

a. La utilidad de replicación de la herramienta CatMDEdit permite que los usuarios puedan crear a partir de un registro de metadatos un duplicado sobre el que se pueden realizar las modificaciones necesarias para ajustar las descripciones a las del objeto a describir. De esta manera el trabajo que consume más tiempo se realiza una sola vez

(el primer registro de una colección) y para los demás registros solo se modifican los campos del registro que varían como son: el título, el identificador de elemento, la escala, las coordenadas del rectángulo mínimo, el conteo de píxeles en cada dimensión, las resoluciones, etc.. Esta metodología ha sido aplicada sobre las series de mapas del SGE 8C, 4C y 2C.

Un estudio realizado sobre el tiempo y la completitud de los metadatos de estas series indican que la creación del primer registro en término medio consume 40 minutos y los siguientes 10 minutos/registro.

b. La segunda alternativa utilizada consistente en cumplimentar los campos de una tabla, previamente diseñada a tal efecto, que contiene los campos significativos para la colección que se desea describir utilizando como fuente la lectura del propio mapa de un modo sistemático. En segundo lugar se extraen las propiedades de la información geográfica auto contenidas en el los ficheros imagen de forma automática (coordenadas del rectángulo mínimo, conteo de filas y columnas, resolución del pixel, el sistema de referencia espacial) de modo que se añadan a los datos recuperados de forma manual. Tras realizar estas operaciones se dispone de toda la información propia de cada elemento (mapa). Como tercer paso se utiliza la herramienta GuardaMeta, diseñada para gestionar colecciones de metadatos, creando una nueva serie a la que se asocia toda la información común: información de identificación, de distribución, responsable de los metadatos, informes de calidad, fuentes de datos y procedimientos, palabras claves comunes, etc... Como cuarto paso se hace uso de la utilidad para recuperar los datos que previamente se habían almacenado en la tabla, con la información de cada elemento de la serie. La figura 1 muestra de forma sintética las fases de los trabajos.

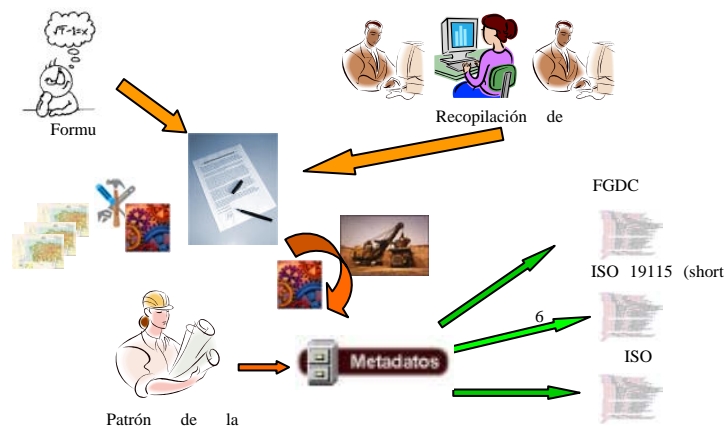


Fig 1. Diagrama de flujo de las fases de la metodología.

Esta metodología ha sido aplicada a un conjunto de 416 cartas náuticas, invirtiendo los siguientes tiempos: diseño de tabla inicial 15 minutos; relleno de los 416 registros 30 horas; extracción automática de información 5 minutos; relleno de los datos de una

serie 40 minutos, e importación de datos (los datos particulares de cada elemento) 2 minutos. En términos absolutos $3 \text{ horas} * 60 \text{ minutos} / 416 \text{ registros} \sim 5 \text{ minutos/registro}$.

5. Puesta en marcha de los servicios WMS con los mapas.

Para satisfacer el requisito de visualización de los mapas georreferenciados en el entorno de Internet de una forma interoperable se ha de instalar una aplicación que proporcione dicho servicio (WMS). Existen múltiples aplicaciones que proporcionen servicios WMS conformes con las especificaciones del OGC como puede observarse en el registro de implementaciones de servicios del OGC [7]. Cada una de estas aplicaciones están diseñadas y/o están siendo desarrolladas para trabajar con distintos tipos de fuentes/almacenes de datos; Así, por ejemplo, GeoServer [5] está especializado en proporcionar servicio WMS de datos vectoriales almacenados en Bases de Datos (Oracle, PostGis, MySQL, ArcSDE, Shapefile, etc...). Si entre los objetivos de la institución que proporciona el servicio está el uso de OpenSource, existen varias aplicaciones disponibles como: Mapserver, originalmente de la Universidad de Minnesota, y Deegree, originalmente de la Universidad de Bonn. Se trata de dos tecnologías muy distintas: CGI frente a Servlet y como tales debe realizarse la comparación: Mapserver es el más antiguo y puede manejar más de 30 formatos matriciales, 9 formatos vectoriales, acceso a BBDD PostGis, Oracle y el resto de formatos que se puedan gestionar a través de FME [9], Deegree puede acceder a un conjunto de formatos gráficos más reducidos entre los que se encuentran: Geotiff, Ecw y los formatos gráficos que reconoce la librería JAI (java advanced imaging)

Para instalar y poner en marcha Mapserver como WMS, primero debe de descargarse la última versión estable del producto (4.2) para la plataforma de servidor de que disponga (Linux, Windows) y compilarla o instalarla siguiendo las recomendaciones. Una vez que se ha instalado el CGI mapserv se ha de crear el fichero en el que se define la configuración del mismo (fichero .map). Este fichero incluye una cabecera en la que se definen las coordenadas del rectángulo mínimo que encierra los datos, las unidades de las coordenadas, el formato por defecto de las imágenes, el sistema de referencia espacial por defecto y el tamaño de las imágenes que se ofrecerá. Otra de las secciones importantes es aquella en la que se definen los metadatos del servicio, para posteriormente proceder a describir todas y cada una de las capas con los tipos de datos, las fuentes, los estilos, el sistema de referencia espacial y los metadatos de la capa.

Para ayudar a construir el archivo de configuración, se ha creado una herramienta que toma los datos de la tabla con la información relevante de cada mapa y la escribe con el formato apropiado sobre dicho archivo (.map). De este modo se crea de una forma segura dicho fichero ya que se evitan los errores tipográficos de las transcripciones mecanográficas.

Las principales dificultades encontradas en esta fase están relacionadas con las limitaciones del número de capas de la aplicación Mapserver (100) y en segundo lugar el hecho de no estar implementado el metadato asociado a la capa que contiene el enlace a los metadatos de los datos (Sección 7.1.4.5.10 de la especificación WMS 1.1.1). La solución de estos problemas ha consistido en compilar una versión modificada del proyecto Mapserver que amplíe el número de capas a 1000 y que acepte el enlace al fichero de metadatos de la imagen. Esta solución ha tenido una complejidad añadida nada despreciable ya que dicho proyecto depende de otros muchos proyectos OpenSource y librerías de uso libre que comportan dificultades de personalización.

6. Preparación de GeoNetwork como GeoPortal.

Hasta el momento se dispone de: las imágenes rectificadas, los metadatos y el servidor de mapas web que proporciona un acceso estandarizado e interoperable. El siguiente paso consiste en crear un Portal Web que permita localizar la I.G. disponible, evaluar la adecuación con los objetivos requeridos para poder localizar el servicio WMS que proporciona acceso a la misma y usar los datos localizados junto a otros sobre un cliente Web de servicios WMS.

Tras realizar algunas búsquedas en la web se ha localizado el proyecto OpenSource GeoNetwork utilizado por la FAO-UN. Se trata de una herramienta multilingüe que dispone de un editor de metadatos web (ISO19115), que importa o registra metadatos almacenados en ficheros XML y que posee una interfaz de búsqueda y otra de visualización de los datos. Se puede destacar que la interfaz de búsqueda está desarrollada de forma que se pueden realizar búsquedas distribuidas sobre otros catálogos federados y fusionar los resultados utilizando el estándar ISO23950 (Z39.50).

Los trabajos realizados para preparar el GeoPortal con GeoNetwork han consistido en: realizar las traducciones de los mensajes, imágenes y ayudas de la herramienta al español, incluir la numeración de los campos conforme a la norma ISO19115 de metadatos, indicar cuáles de los mismos corresponden con el NEM, añadir descripciones que ayuden a la redactar el contenido de ciertos campos. Finalmente se ha configurado la herramienta InterMap (cliente de servicio WMS) para que gestione por defecto un conjunto de datos vectoriales global (todo el mundo) así como modificar la apariencia del portal para personalizarla (ver figura 2).

La carga de los metadatos sobre el catálogo se ha realizado con una utilidad de GeoNetwork que posibilita la importación de todos los metadatos contenidos en ficheros XML dentro de un directorio local. Esta utilidad no prevee la carga simultánea de las imágenes de pre-visualización de los conjuntos de datos. Como aportación al proyecto se ha incluido esta opción facilitando esta tarea. Dado que el borrado de metadatos del catálogo solo se podía realizar de forma unitaria, y que parecía necesaria la posibilidad de borrar un conjunto de elementos dentro de un

rango de identificadores, se ha añadido esta facilidad a la herramienta. Ambas modificaciones junto con los trabajos de traducción de los mensajes y ficheros de ayuda han sido enviadas al administrador del proyecto, de modo que puedan ser incluidas en la próxima distribución que se publique de la herramienta.



Fig 2. Geo Portal basado en Geonetwork personalizado.

7. Preparación de Ionic como GeoPortal.

En este apartado se trata de exponer como se pueden aprovechar los trabajos de los apartados 2, 3, 4 y 5 para construir con los mismos datos, metadatos y servicios WMS un GeoPortal con un producto comercial como es el cliente Web de Registro de servicios de Ionic (WRSCClient).

Este producto comercial incluye los siguientes componentes conformes con el OGC: WMS con mapas índices para definir los dominios espaciales de búsqueda, WFS-G (Gazetteer) para realizar búsquedas por nombres de lugares, WMS Client para visualizar de una forma interoperable los resultados de las búsquedas y un Registro de servicios WMS, WFS, WCS y Contextos. Toda la información se almacena en una Base de Datos OracleSpatial.

Los trabajos realizados han consistido en personalizar la interfaz y realizar la traducción de los mensajes y ayudas de la herramienta (ver figura 3).

El proceso de registro de los servicios es muy simple: indicar el URL del servicio que se desea registrar. La herramienta realiza la consulta GetCapabilities sobre el servicio y almacena los metadatos del servicio, de las capas y los enlaces a los metadatos de los datos. Aunque la idea es simple, por distintas razones aún no está operativa al 100% en el servidor piloto utilizado:

- la complejidad de la instalación de los servicios WRS y WRSClient,
- los problemas derivados de la codificación de los ficheros que contienen los metadatos,
- el esquema XML utilizado para la codificación de los metadatos (ISO19139 versión 0.6).

Tras realizar un amplio intercambio de mensajes, con el fin de solucionar problemas de instalación y primeros fallos, se espera que la próxima versión del cliente de catálogo permita realizar de esta forma tan sencilla el registro de un servicio que proporciona acceso a un amplio conjunto de capas y que de forma automática se almacenen y se acceda a los metadatos del conjunto de datos de cada capa.

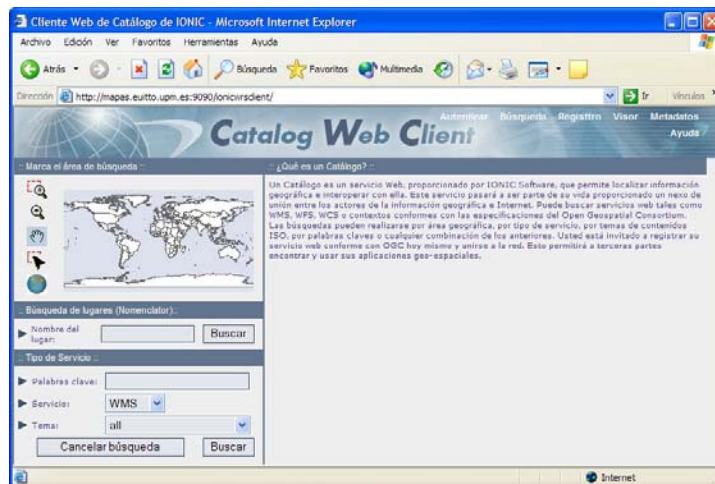


Fig 3. Traducción de la interfaz del Catálogo de Ionic al Español.

8. Conclusiones.

Con la realización de este prototipo de GeoPortal para mapas antiguos podemos concluir que:

- Se han obtenido un conjunto de estimadores de coste temporal relativos a los trabajos de digitalización de las cartotecas, los trabajos de rectificación geométrica y la creación de los Metadatos.

- Se han analizado dos metodologías paralelas que posibilitan la creación de metadatos de una forma sistemática para series cartográficas con un número de elementos mediano y grande que permiten reducir de forma drástica el coste temporal.
- Se demuestra que es posible construir un Portal Web para la Información Geográfica (Geo-Portal) con herramientas OpenSource. Se ha constatado que aunque el software puede usarse de forma gratuita, en algunos casos se deben de invertir recursos humanos para adaptar los proyectos a las necesidades específicas como ha ocurrido en los casos de Mapserver y GeoNetwork.
- Se hace patente la falta de madurez de la norma de metadatos y la incertidumbre existente entorno a las codificaciones de los mismos, limitándose en unos casos a utilizar las etiquetas abreviadas de ISO19115 (GeoNetwork), o borradores de la norma de implementación ISO19139 (Ionic, CatMDEdit). Esta incertidumbre se disipará cuando vea la luz la primera versión oficial de la especificación (finales de 2004, según las previsiones de ISO). Mientras tanto, es recomendable trabajar en la creación de los metadatos utilizando herramientas que nos garanticen que en el futuro inmediato pueden exportar/importar registros conforme con esta norma.
- Desde el punto de vista práctico de la creación de metadatos para series cartográficas, es muy interesante disponer de herramientas de edición y almacenamiento que permitan modificar un campo sobre todo un conjunto de registros de metadatos.
- Desde el punto de vista de la interoperabilidad de los catálogos en las búsquedas distribuidas, de las dos plataformas estudiadas solo GeoNetwork es capaz de realizarlo utilizando el protocolo Z39.50. Actualmente se están realizando esfuerzos para sustituir este protocolo de bajo nivel por otro tipo de tecnologías de nivel superior como *Electronic Business XML* (ebXML) + WebServices de modo que las comunicaciones entre repositorios o catálogos puedan estar garantizadas a nivel de protocolo HTTP, evitando de esta forma los problemas de los *Network Address Translation* (NAT) y *Firewalls*.

9. Referencias.

1. Consorcio TeIDE (Tecnologías para las Infraestructuras de Datos Espaciales). CatMDEdit. Editor de Metadatos para información Geográfica. <http://teide.unizar.es>
2. Earth Resource Mapping. Free ECW Compressor v2.6. <http://www.ermapper.com/ecw>
3. GDAL: Geospatial Data Abstraction Library. <http://gdal.maptools.org>
4. Geonetwork: Geographic Metadata Catalog OpenSource project. <http://sourceforge.net/projects/geonetwork>
5. GeoServer: OpenGIS-compliant web services layer. <http://sourceforge.net/projects/geoserver/>
6. Ionic Corp. RedSpider Catalog 2.0. http://www.ionicssoft.com/products/main_catalog.jsp
7. Open Geospatial Consortium (OGC). <http://www.opengeospatial.org>
8. Universidad de Minnesota. Mapserver OpenSource project. <http://mapserver.gis.umn.edu>
9. Safe software Corp. Feature Manipulation Engine FME. <http://www.safe.com>
10. GlobalMapper. <http://www.globalmapper.com>