SDIAnalyzer: una aplicación para el estudio de IDE.

Juan Ignacio Varela García[(1)] Fco. Alberto Varela García[(2)] Daniel Díaz Grandío[(3)] y Javier Estévez Valiñas[(1)]

[(1)] Desarrollador informático del CartoLAB. ETS de Ingenieros de Caminos, Canales y Puertos, Campus de Elviña s/n 15071 La Coruña, nachouve@gmail.com , valdaris@gmail.com
[(2)] Profesor del Área de Ingeniería Cartográfica, Geodésica y Fotogrametría. CartoLAB. ETS de Ingenieros de Caminos, Canales y Puertos, Campus de Elviña s/n 15071 La Coruña, avarela@udc.es
[(3)] CartoLAB. ETS de Ingenieros de Caminos, Canales y Puertos, Campus de Elviña s/n 15071 Universidad de Coruña, ddgrandio@yahoo.es

RESUMEN

El auge de las Infraestructuras de Datos Espaciales (IDE) en los últimos años ha provocado el despliegue masivo de este tipo de plataformas en el territorio español. A pesar de hacer uso de estándares del OGC, estar reguladas por iniciativas como la directiva INSPIRE y que existen recomendaciones y normativas para su implantación, la red de nodos IDE es muy heterogénea y se detectan grandes desigualdades.

Conscientes de esta situación surge la necesidad de un método objetivo para el estudio y análisis comparativo de IDEs. Como parte del desarrollo de esta metodología nace SDIAnalyzer, programa de software libre que permite la extracción masiva, almacenamiento y estudio de información relativa a servicios del OGC.

SDIAnalyzer (Spatial Data Infraestructure Analyzer), creado por el CartoLab de la Universidade da Coruña, ha sido aplicado en el análisis de aproximadamente 180 IDEs españolas de nivel regional elaborando mapas temáticos de diversas características y pretendemos seguir avanzando en su desarrollo. En el presente artículo, se exponen sus características principales y se lanza de forma pública a la comunidad de software libre.

Palabras clave: *SDIAnalyzer, OGC, IDEs*

INTRODUCCIÓN

El SDIAnalyzer es una aplicación para analizar, comparar y valorar con unos criterios totalmente objetivos las diferentes Infraestructuras de Datos Espaciales de cualquier marco geográfico. Creada por el laboratorio CartoLab[1] de la Universidade da Coruña, el presente artículo pretende ser la presentación del lanzamiento público de esta herramienta de software libre a la comunidad.

Entre sus características principales está la recogida masiva de información de diferentes recursos IDE. Estos datos se almacenan de manera ordenada con lo que resulta posible utilizarlos en la creación de estudios complejos, mapas temáticos, estadísticas, etc.

Para poner a prueba esta aplicación se ha realizado un estudio sobre IDEs de ámbito regional pertenecientes a la Infraestructura de Datos Espaciales de España. La metodología de análisis definida nos permite evaluar ciertos aspectos interesantes de estos servicios, establecer unos valores numéricos de ellos y poder compararlas entre sí. Tras la experiencia satisfactoria con OGCAnalyzer (versión primigenia de SDIAnalyzer) se ha hecho un cambio de plataforma de programación. En la nueva versión, aún en estado beta, las métricas utilizadas así como los desarrollos informáticos creados, dejan la puerta abierta para nuevas propuestas que amplíen o proporcionen una mejor evaluación del funcionamiento de estas plataformas.

Qué son las IDEs: importancia y marco europeo

Las Infraestructuras de Datos Espaciales (IDE) juegan un papel básico en la difusión, utilización y mejora de la Información Geográfica (IG). La disposición de una IDE eficaz partiendo de un recurso tan valioso como la IG, permite optimizar su uso a la vez que se reducen costes. En los últimos años, sobre todo a raíz del impulso en este campo que ha significado la Directiva INSPIRE (INfrastructure for SPatial InfoRmation in Europe), son muchos los organismos que han puesto en marcha plataformas de este tipo.

La Directiva INSPIRE, publicada en el Diario Oficial de la UE (DOUE) el 25 de Abril de 2007, es una iniciativa que aboga por el desarrollo de una Infraestructura de Datos Espaciales Europea. Para lograr ese objetivo, se fomenta la creación de servicios de información espacial integrados que permitan identificar y acceder a la información geográfica a nivel local y global, garantizando la interoperabilidad. Esto supone una red uniforme de IG pública, accesible a través de Internet, dentro del marco de la Unión Europea, utilizable por múltiples usuarios y aplicaciones diversas. La directiva pretende integrar tanto datos de carácter medioambiental como de otros ámbitos como el transporte, la agricultura, la economía, etc...

La directiva INSPIRE delega gran responsabilidad en las Infraestructuras Nacionales de Datos Espaciales (NSDI's) y éstas a su vez en las IDEs autonómicas, provinciales y locales. Así, se tienen en cuenta los requerimientos y necesidades de estas subentidades siendo armonizados en cierta forma con los criterios marcados globalmente. Esta estructura jerárquica basada en niveles provoca la aparición de diferencias entre nodos IDE que afectan a su organización, implementación, disponibilidad y, evidentemente, a la información ofertada. Se debe resaltar el papel del Comité Técnico ISO/TC 211 para la Información Geográfica así como del Open Geospatial Consortium (OGC)[2] en su labor de definición de estándares que garanticen la interoperabilidad. Los estándares son un punto común para todas las IDEs a la hora de afrontar su creación a nivel técnico. Sin embargo, los perfiles de los usuarios y los usos que realizan de la información suelen ser muy dispares. Además, es difícil ofertar un servicio homogéneo en una red tan amplia y compleja como la propuesta en la directiva INSPIRE. Más allá de las propias especificaciones técnicas de los servicios y estándares proporcionados por los organismos anteriormente citados, se necesitan muchas más especificaciones detalladas sobre cómo ha de desplegarse una IDE. En este sentido, la directiva INSPIRE, no solamente pone las bases para que las IDEs proliferen, sino que propone unas serie de recomendaciones sobre aspectos más concretos de cómo deben ser esas IDEs. A su vez, el Consejo Superior Geográfico también enfatiza a este respecto con unas recomendaciones

basadas en las de la propia directiva INSPIRE y desarrolladas por el "Grupo de Trabajo de la IDEE". Sin embargo, el carácter opcional de estas recomendaciones hace que su seguimiento no sea generalizado y muchas IDEs no las cumplan.

Por qué estudiar las IDEs

El auge en los últimos años de las IDEs, y la rápida aparición de nuevas plataformas en España, nos invita a realizar una reflexión acerca de la oferta actual de estos espacios de información. Una IDE es una tecnología compleja y los datos que ésta alberga exigen un tratamiento, gestión y organización no trivial. Poner en marcha y mantener una Infraestructuras de Datos Espaciales requiere muchos recursos y una notable inversión, es por ello que se debería sacar el máximo provecho posible a la misma. Dado el alto interés de la sociedad y el creciente respaldo político al acceso público de la información, parece claro que se necesitan estudios que permitan valorar estas plataformas, hacer comparativas e incluso determinar si su funcionamiento y servicios son adecuados a las expectativas que podría esperar un usuario. Además, no existen criterios estandarizados para determinar muchos de los parámetros que se tienen que ajustar al desplegar una IDE y a veces es difícil determinar si el funcionamiento es satisfactorio, si la información alojada es útil, si la jerarquía de los datos es correcta o si el tiempo de respuesta es razonable.

La Infraestructura de Datos Espaciales de España (IDEE) hace una labor centralizadora de todas estas iniciativas en el territorio nacional, siendo un lugar de referencia para la consulta y búsqueda de servicios IDE. Sin embargo, parece que estos nodos IDEs son muy heterogéneos en cuanto a su estética, forma de uso, servicios ofertados, utilidad, rendimientos, etc. Las diferencias pueden afectar al nivel de calidad y/o cantidad de datos y servicios ofrecidos en distintas regiones. Esto supone un agravio comparativo para los ciudadanos, instituciones, empresas o investigadores de esos lugares con IDEs menos eficientes. Por ello, el cuantificar estas diferencias puede suponer un gran impulso para que se tomen medidas y que situaciones de este tipo se equilibren.

Ante esta coyuntura, surge la pregunta de cómo poder elaborar una comparativa de IDEs teniendo en cuenta el gran volumen de información. Se vio la necesidad de recoger multitud de datos, así como elaborar ciertos análisis de los servicios IDE de una forma automatizada.

Medir la cantidad de datos ofertada, la disponibilidad, usabilidad y otros parámetros de una IDE es un reto difícil que con la metodología empleada se intenta superar. Esta metodología permite extraer información objetiva de carácter cuantitativo y cualitativo acerca de distintos aspectos. Todos los datos son obtenidos de fuentes públicas y accesibles libremente como, por ejemplo, las respuestas a una petición de capacidades (getCapabilities).

LA HERRAMIENTA SDIANALYZER

Como ya se ha comentado anteriormente, SDIAnalyzer es una aplicación para la recopilación de datos de infraestructuras de datos espaciales que ofrecen su información siguiendo estándares del OGC. El hecho de que estas especificaciones de estándares sean públicas permite construir implementaciones de clientes

libremente. Con SDIAnalyzer se pueden estudiar ciertas características de estos servicios siendo su objetivo principal facilitar el chequeo y extracción de datos sobre servicios de mapas.

La idea de crear una herramienta como SDIAnalyzer, está inspirada en las herramientas libresoft-tools [3] (para el estudio de las dinámicas del software libre) creadas por el grupo GSyC/LibreSoft[4] perteneciente a la Universidad Rey Juan Carlos de Madrid. Este conjunto de herramientas cuenta con utilidades para conectarse, por ejemplo, a listas de correo electrónico o repositorios de código, extraer un gran número de información de éstos y almacenarlos en una base de datos. Una vez estructurada de esta forma se pueden realizar una gran cantidad de consultas y análisis sobre esos recursos.

Primeros pasos

Siguiendo la misma línea que las libresoft-tools, OGCAnalyzer (como se llamaba en principio SDIAnalyzer) fue programada usando como lenguaje de programación Python [5] lo que permitió un desarrollo muy rápido. Tras analizar distintas alternativas de clientes/librerías para acceder a servicios OGC, se empleó la librería **owslib-0.3** [6] y que cuenta con una licencia de software libre que permite libertad para su uso, acceso al código fuente, modificación y distribución.

En un primer momento, la herramienta estaba orientada a su uso mediante línea de comandos usando archivos en texto plano para su configuración, entrada de datos y salida estándar por pantalla. Además, para aprovechar todas sus ventajas había que utilizar la herramienta en un sistema operativo GNU/linux. Al ser un instrumento de investigación todo esto no era un limitante.

En esa primera fase de desarrollo, la herramienta contaba con sólo una pequeña parte de las funcionalidades que se habían pensado en el momento de su diseño. A pesar de ello, se pudieron realizar diversos estudios muy interesantes como se verá en el apartado “Ejemplo de uso”.

Desarrollo Actual

Viendo que la aplicación podía resultar muy útil para más investigadores y otros perfiles de usuarios no técnicos, el entorno y forma de uso comentado era un poco limitante. Se decidió dotar a la herramienta de una interfaz gráfica además de ampliar sus capacidades. Como resultado de esta nueva perspectiva, se migró OGCAnalyzer de python a la plataforma Java. Este cambio se vio apoyado por la posibilidad de poder integrar en un futuro la herramienta dentro de los muchos Sistemas de Información Geográfica programados en este lenguaje que existen en la actualidad.

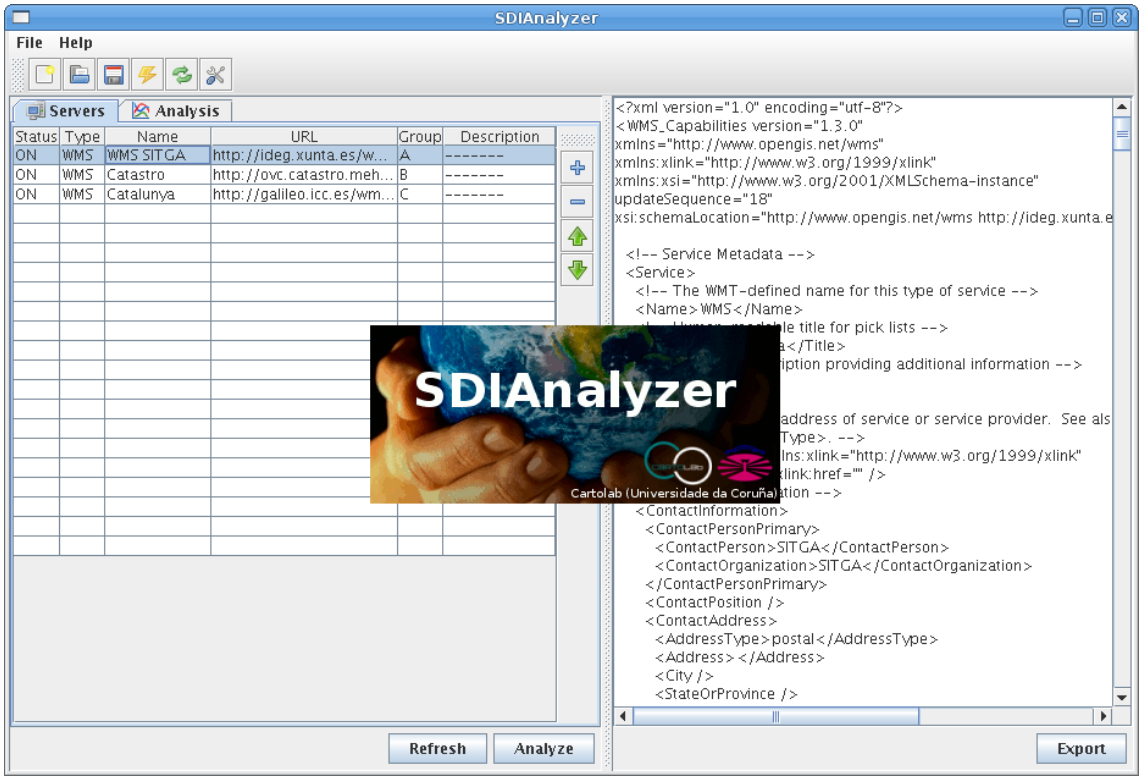

Figura 1: *Interfaz gráfica de SDIAnalyzer*

Al cambiar de tecnología, se han buscado nuevas librerías para la explotación de los servicios ofrecidos por las IDEs, esta vez en Java. Actualmente se están todavía analizando las posibilidades existentes, encontrando como las más interesantes las utilizadas en el SIG Kosmo [7] o las facilitadas por GeoTools [8]. El uso de una de estas librerías de utilidades permite a la aplicación consumir como cliente los servicios web que siguen los estándares del OGC (actualmente sólo Web Map Service, WMS). Mediante la API de estas librerías se puede acceder a metadatos y construir peticiones "GetCapabilities", "GetMap", y "GetFeature" que son la base para recopilar la información necesaria para los estudios de las Infraestructuras de Datos Espaciales.

En este proceso de cambio se introdujeron notables variaciones en la arquitectura anterior. Uno de ellos fue el uso de una base de datos SQLite[9] para el almacenamiento en tablas de los resultados obtenidos, así como los recursos a estudiar. Esto permite mayores posibilidades al poder realizar consultas más complejas, crear nuevos informes, históricos, etc.

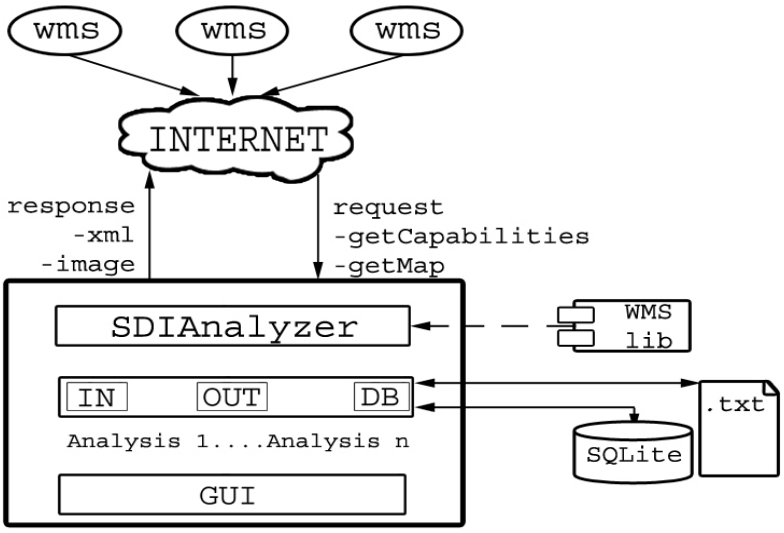

Figura 3: *Esquema básico de SDIAnalyzer*

Además, la jerarquía de clases utilizada permite ampliar las capacidades de SDIAnalyzer con nuevos análisis. Éstos pueden ser integrados fácilmente en la aplicación y aprovechar la actual interfaz gráfica de usuario.

Entre las funcionalidades que ofrece SDIAnalyzer, podemos citar:

- Comprobación de la disponibilidad del servicio
- Conteo de capas de información
- Testeo del funcionamiento de operaciones getCapabilities y GetMap
- Chequeo de porcentaje de capas que permiten getFeatureInfo
- Examinar formatos de imagen soportados
- Verificación de sistemas de referencia de coordenadas
- ...

El proyecto, que cuenta con una licencia libre GPL, será alojado en la forja OSOR, donde se centralizará su desarrollo aprovechando las herramientas disponibles: repositorio de control de versiones, wiki, sistema de bugs, listas de correo, etc. OSOR (Open Source Observatory and Repository)[10] es una plataforma de la Unión Europea cuya misión es que los desarrollos de software financiados mediante fondos públicos (creados por una administración pública, por empresas que colaboran con una administración pública o, como en el caso de SDIAnalyzer, por una una universidad pública) se realicen bajo las buenas prácticas del desarrollo colaborativo habitual en el software libre. Esto fomenta la cooperación, compartir conocimientos y la reutilización de soluciones entre las distintas administraciones[11].

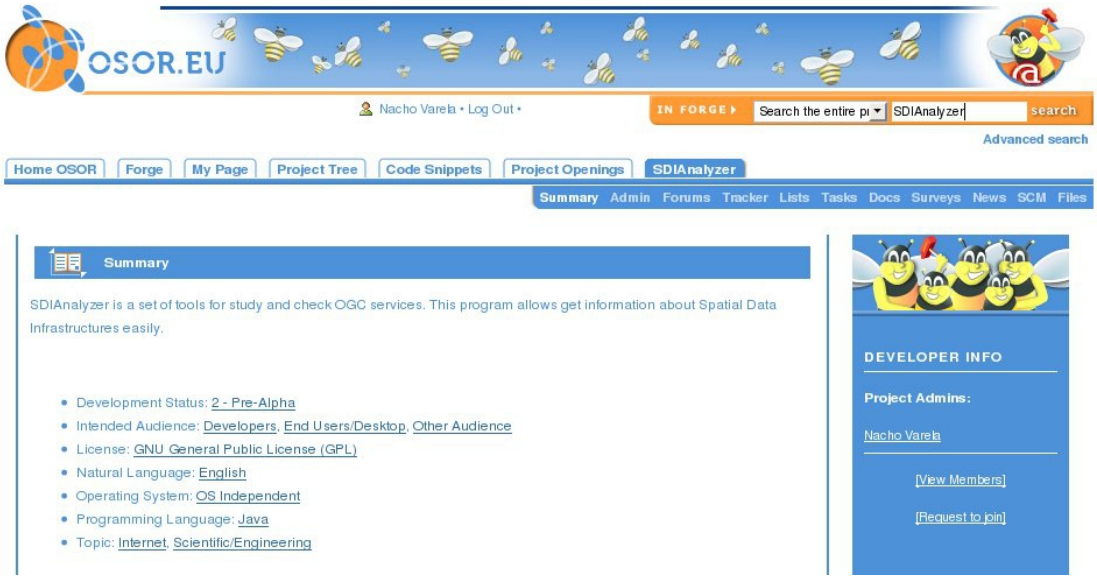

Figura 2: *Espacio de SDIAnalyzer en la forja de OSOR.*

EJEMPLO DE USO

El uso de programas de tipo SDIAnalyzer, que automatizan la recogida y análisis de datos, nos permiten por una parte reducir tiempos dedicados a este tipo de tareas que normalmente son repetitivas y, por otro, aumentar la capacidad de volumen de información tratada.

OGCAnalyzer, la versión inicial dell actual SDIAnalyzer, fue utilizada para extraer de forma masiva los datos accesibles de servicios definidos por el OGC, en este caso únicamente servicios WMS de IDEs regionales facilitadas en la página web de la IDEE. Para ello se definió un paquete de análisis a realizar sobre los servicios seleccionados, como el número de capas de información, número de ortofotografías, comprobación de formatos de salida, etc. Entre los análisis más interesantes realizados por OGCAnalyzer está el grado y calidad de metadatos ofrecidos y una comparativa número de capas de información entre Comunidades Autónomas.

Como prueba de la utilidad de esta herramienta de estudio masivo de IDEs se muestra como ejemplo el siguiente mapa temático donde queda reflejado el resultado de análisis de aproximadamente 180 URLs de servicios WMS regionales. Este estudio fue presentado en las “V Jornadas Técnicas de IDE de España” celebradas en noviembre del 2008 en Tenerife[12].

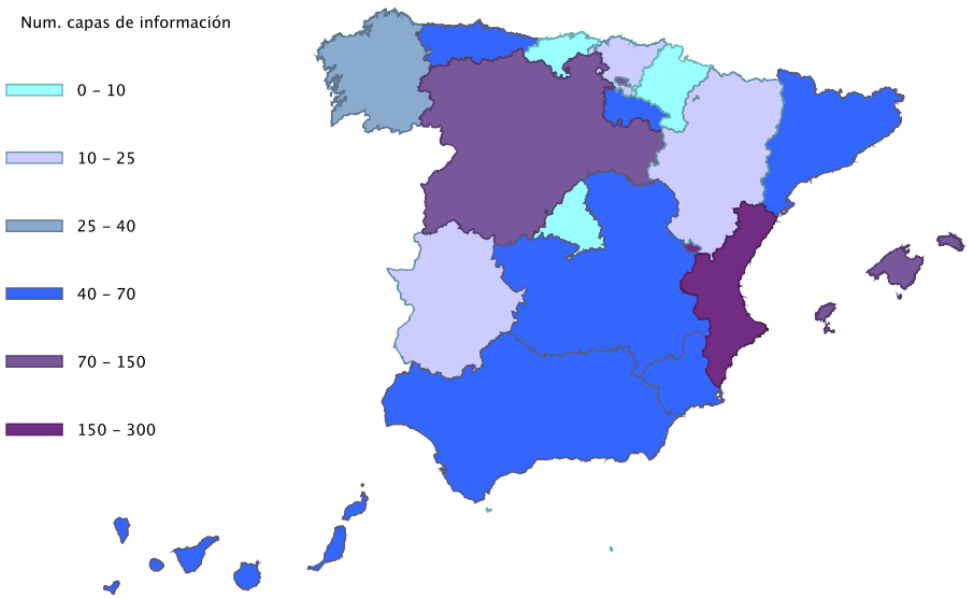

Figura 4: *Mapa de número de capas ofrecidas en WMS regionales.*

En la figura 4, podemos apreciar un mapa temático que muestra la diferencia en cuanto a número de capas de información geográfica ofertadas por Comunidad Autónoma.

DESARROLLOS FUTUROS Y CONCLUSIONES

La aplicación SDIAnalyzer se lanza a la comunidad como una interesante herramienta para estudiar Infraestructuras de Datos Espaciales. Dado que esta nueva versión es de reciente creación, aún quedan muchas cosas que hacer. Existen varias líneas para su ampliación a nivel técnico y también se tienen numerosas ideas para emplearla en casos prácticos. En este sentido, se pretenden realizar estudios periódicos a diferentes niveles administrativos y distribuir esos resultados públicamente. En cuanto a mejoras tecnológicas, se van a incorporar nuevos tipos de análisis como la medición de tiempos de respuesta, ampliar a servicios WFS y WCS, etc. Otra de las mejoras previstas será la integración de SDIAnalyzer dentro de SIGs libres como gvSIG, OpenJump y Kosmo.

El licenciamiento y distribución de SDIAnalizer como software libre, junto con el esfuerzo de documentación y soporte público, esperamos que anime a la utilización del programa así como a la colaboración de otras personas y entidades para seguir con su desarrollo y realizar mejoras. Alojar el proyecto en la plataforma de la Unión Europea OSOR contribuye a fomentar la compartición de recursos siguiendo las dinámicas propias del software libre.

REFERENCIAS

- [1] http://cartolab.udc.es/
- [2] http://www.opengeospatial.org/
- [3] http://forge.morfeo-project.org/projects/libresoft-tools/
- [4] http://libresoft.es/
- [5] http://www.python.org/
- [6] http://trac.gispython.org/lab/wiki/OwsLib
- [7] http://www.opengis.es/
- [8] http://docs.codehaus.org/display/GEOTDOC/WMS
- [9] http://www.sqlite.org/
- [10] http://www.osor.eu/
- [11] http://forge.osor.eu/docman/view.php/6/124/OSOR_dissemination-Tampere-2008.pdf
- [12] http://www.orzancongres.com/ideart/073.pdf