

Bloque 1. Introducción: Información Geográfica, Sistemas de Información Geográfica y Web Semántica

Capítulo 1C. Las Infraestructuras de Datos Espaciales

Unidad 1.C.1: Introducción a las IDEs

Dr. Miguel A. Bernabé Poveda (Grupo de Investigación MERCATOR. UPM)

5.1. Introducción

“Existe un reconocimiento cada vez mayor de que los desafíos más importantes de la sociedad moderna, tales como la protección del entorno, el incremento de la seguridad, la mejora del transporte, el desarrollo socialmente justo y la ampliación de servicios para los ciudadanos, requieren que quienes toman decisiones identifiquen dónde es más acuciante la necesidad y los medios para intervenir allí del modo más efectivo, monitorizar resultados y evaluar impactos. Para todas estas tareas, la información geográfica es crucial. Dicha información no sólo debe existir, sino que además debe ser fácil el identificar quién la tiene, si es apropiada o no para el propósito que se persigue, cómo se puede acceder a ella, y si puede o no ser integrada con otra información”. (GINIE. Informe Directivo “Infraestructuras de Datos Espaciales: De lo local a lo global. Recomendaciones para entrar en acción. 2004)

La Información Geográfica (IG) y su intercambio parecen ser uno de los temas de más interés en los últimos años. Frases como la que afirma que *“más del 80% de las bases de datos almacenadas en las administraciones son susceptibles de ser georreferenciadas”* conducen a pensar que de pronto la IG toma una importancia que hasta hace poco no parecía tener.

La llegada de los SIG supuso disponer de una herramienta (local) con la que poder realizar análisis en base a la información almacenada en las bases de datos del ordenador y mostrar el resultado de manera gráfica. Dicho de manera grosera:

- El usuario de un SIG dispone de datos almacenados en bases de datos.
- El usuario puede realizar preguntas a las bases de datos. El sistema le responde de manera numérica (tables, resultados de operaciones) o de manera gráfica (gráficos sobre mapas)

Los SIG se convirtieron enseguida en sistemas capaces de proporcionar ayuda en los procesos de toma de decisión (gestión de territorios, gestión de emergencias, negocios, etc), pues es más fácil y en momentos críticos es más útil, ver el resultado global sobre un mapa de manera gráfica que disponer de respuestas en forma de interminables listas de variables.

La capacidad de los SIG para ayudar a decidir con mayor o menor acierto en situaciones en las que el territorio está afectado, proporciona una seguridad que se transforma en ganancia de tiempo y dinero. Esto es posible gracias a la posibilidad de visualizar múltiples circunstancias que afectan al territorio y analizar las consecuencias de cada decisión teniendo en cuenta todas esas circunstancias.

Sin embargo, la multiplicidad de formatos de los datos, la diversidad de marcas y sistemas operativos de las plataformas sobre las que corren los programas, los diferentes sistemas de captura, almacenamiento y distribución de los datos, hacen que compartirlos sea un problema, que en el mejor de los casos, genera una pérdida de tiempo. Es necesario afrontar el problema de la interoperabilidad en el campo de la Información Geográfica y eso es lo que organizaciones tradicionales como ISO y otras específicas como el OGC persiguen con sus trabajos.

Por último debe ser posible la localización y acceso de manera sencilla y eficaz (con las restricciones que imponga el dueño) a la información geográfica existente sobre un territorio.

5.1.1 Actualización y accesibilidad de la IG

5.1.1.1 Información actualizada. La información geográfica es una información altamente cambiante. Ya sea por la acción del hombre o por causas naturales, las características de la Tierra son poco estables y para tomar decisiones fundamentadas es necesario disponer de datos actualizados. La IG suele ser cara

pues son costosos los medios para conseguirla. Teniendo en cuenta esa afirmación, las tomas de decisión y en general, las preguntas que se realicen a un SIG deben hacerse sobre conjuntos de datos actualizados. La actualización implica siempre un gasto considerable

5.1.1.2 Información instantánea. También en los momentos críticos de tomas de decisión instantánea, se requiere que la información esté disponible de manera instantánea. Esto implica que los centros de distribución de información deben estar cercanos al lugar de la crisis o tener agilidad de entrega.

5.1.1.3 El acceso ubicuo como solución. La solución a esos dos inconvenientes sería la de acceder instantáneamente a la información más actual. Esta información está en manos de quien la produce o la distribuye (instituciones, organismos, empresas, universidades) y el acceso más rápido, generalizado y ubicuo es el que se realiza por medio de las redes de Internet. Así pues, la actualización y accesibilidad de la IG se puede mejorar mediante el acceso a la información proporcionada por el distribuidor vía Internet.

Sin embargo, desaparecidos esos problemas nos encontramos de nuevo con otros nuevos relacionados con el concepto de Interoperabilidad.

5.2. El entendimiento entre sistemas: la Interoperabilidad

Acceder a los datos que definen una característica de un territorio, almacenados en las bases de datos de una institución a través de internet y disponer de esos datos para compararlos o unirlos a otros datos de otra característica del mismo territorio almacenados en las bases de datos de otra institución, sería una buena solución. Sin embargo, se deben solucionar ciertas características asociadas a la IG para que ésta pueda ser fácilmente compartida.

5.2.1 Captura de la IG.

La información geográfica o georreferenciable que llega a un SIG es capturada de muy diferentes maneras y dispone de muy diferentes formatos. Para que la información sea utilizable en un SIG y ofrezca resultados a las preguntas del usuario, la información debe estar homogeneizada. La información espacial puede haber sido recogida de las siguientes formas:

5.2.1.1 Mediante levantamientos topográficos. Las colecciones de puntos que conforman las características de un territorio, de los que se conocerán sus coordenadas, se almacenan en las memorias de las estaciones totales. Los formatos de almacenamiento de esos datos, dependen de la marca de la Estación y se procesan por medio de programas informáticos. La homogeneización de los datos captados con una u otra estación suele hacerse por medio de un formato estandarizado tipo .dxf o .dgn compatible con los programas de CAD.

5.2.2.2 Por medio de levantamientos fotogramétricos. A partir de las fotografías aéreas los restituidores recogen las características del territorio y las almacenan en formatos diferentes y generalmente propietarios de cada marca. Los programas de SIG son capaces de reconocer los formatos más usuales. En caso de no ser reconocidos se deben exportar a un formato de paso. En este proceso puede perderse información y casi siempre tiempo.

5.2.2.3 Utilizando imágenes de satélite. Multitud de tipos sensores, cada uno con sus características, recogen información de la superficie terrestre. La información recogida por cada uno de ellos varía en cuanto a resolución, banda del espectro electromagnético recogida, formato en el que se guarda la información, etc. Para trabajar con diferentes imágenes y poder realizar consultas al sistema los archivos deben disponer de ciertas características que permitan comparar y sacar conclusiones cuantificables.

5.2.2.4 Mediante medidas realizadas GPS o similares. Capturas con receptores tipo GPS, Glonass o GALILEO. Cada casa constructora almacena la información recibida en sus formatos favoritos. La información proveniente de dos receptores diferentes debe ser homogénea u homogeneizable en algún tipo estándar.

5.2.2.5 Por medio de sensores que miden características asociadas espacialmente al territorio (temperatura, humedad, presión, conductividad...)

5.2.2.6 Mediante informes escritos, tablas, listas. Cuanquier información sobre el territorio que pueda asociarse a una posición e informar de sus características (nomenclátors, descriptores, tesauros, etc.,)

5.2.2 La dificultad debida a los formatos de los datos.

Casi cada casa comercial diseña el formato que más le interesa. En algunos casos este formato es traducible directamente a un formato estandarizado y la pérdida de información es inapreciable. En otros casos no es sencilla la conversión a un formato estándar y si ésta se puede llevar a cabo, se realiza con pérdidas de diferente consideración (pérdida de las capas de información, pérdida de estilos, pérdida de descriptores). En otros casos simplemente no es posible transformar un formato a otro estándar.

5.2.3 Limitación debida a las diferentes plataformas.

Los software que permiten poner en marcha un SIG pueden funcionar sobre diferentes plataformas informáticas o pueden exigir estar instalados en plataformas concretas (Wintel, Unix, etc.). Debido a eso, la exportación de ficheros de un programa que funciona en una plataforma a otro que funciona en otra, puede tener también un grado de dificultad de diferente consideración.

5.2.4 Los diferentes sistemas de referencia.

Una vez que la información ha sido recogida, ésta debe ser referenciada respecto al sistema geodésico que mejor se adapte a las características geográficas del país. Los datos pasan por un proceso de transformación. Si existen varios datum en las colecciones de datos, dispondrán de diferentes meridianos origen y diferentes ecuadores dependiendo de los elipsoides. Esto conducirá a que un punto real de la superficie de la Tierra tenga una latitud y longitud distinta dependiendo del datum al que se haga referencia

Los modernos sistemas de referencia, como ITRS2000, con resoluciones centimétricas son sensibles a los cambios en la posición de un punto debido a fenómenos geológicos como: movimiento de placas (deriva continental), elevación post glacial etc. que nos conduciría a preguntarnos sobre si deberíamos cambiar las coordenadas de cada vértice geodésico cada año, en función de esos movimientos.

El sistema de referencia ETRS89 ha sido concebido para evitar este inconveniente práctico publicándose con cada realización ITRF de cada año los parámetros de transformación a ETRS89 aprovechando así mismo el continuo incremento de precisión de los sucesivos ITRF (Marco de Referencia Terrestre Internacional)

5.2.4 Interoperabilidad: La solución a la diversidad

Hacer que la gestión de la información geográfica no tenga limitaciones debido a las restricciones impuestas por los sistemas de captura, proceso, almacenamiento, distribución o visualización, es un paso muy importante para que las Instituciones puedan compartir la información.

La interoperabilidad es la condición mediante la cual, sistemas heterogéneos pueden intercambiar procesos o datos. Tal ocurre en la actualidad con las páginas web, en las que los usuarios de Windows y los de UNIX pueden acceder y leer el mismo tipo de información a pesar de usar plataformas diferentes.

En los procesos de gestión de la IG debe existir interoperabilidad por medio de:

- Normas que definan los procesos y acoten sus características de calidad
- Especificaciones de los procesos que describan las posibilidades y los resultados esperados.
- Formalización de los descriptores de los datos mediante Modelos de Información
- Estándares de bases de datos que definan los esquemas de los datos
- Protocolos que permitan la utilización de diferentes plataformas o hardware.
- Redes definidas por protocolos universales (HTML, XML, FTP/IP)
- Interfaces que garanticen que cualquiera que sea el sistema utilizado para el acceso a la información ésta se visualizará de la misma manera

Pero también debe existir interoperabilidad Institucional a nivel local, regional, estatal o global. Esta interoperabilidad, posiblemente más difícil de conseguir que la interoperabilidad técnica, requiere del

acuerdo de los productores, integradores y usuarios de datos espaciales del ámbito territorial del que se trate.

5.2.4.1 El Open Geospatial Consortium <http://www.opengeospatial.org/>

El Open Geospatial Consortium (OGC) es una organización compuesta por más de 300 instituciones comerciales, gubernamentales, académicas y otras, dedicadas a proporcionar un consenso de colaboración mediante el desarrollo de especificaciones que garanticen la interoperabilidad de contenidos y servicios de la Información Geográfica.

5.3 Las Infraestructuras de Datos Espaciales (IDEs)

Partiendo de la base de que se deben unificar los procesos relacionados con la IG, que ésta debe ser accesible (con las limitaciones que imponga el dueño de la información) a través de internet y que debe existir un consenso entre instituciones para compartir información, el término Infraestructura de Datos Espaciales (IDE) es utilizado para nombrar la colección de tecnologías, políticas y estructuras institucionales que facilitan la disponibilidad y acceso a la información espacial. Las IDEs deben proporcionar una base para la localización de datos espaciales, su evaluación y su utilización por los usuarios, a todo tipo de niveles: gubernamental, sector comercial, sector no lucrativo, área académica y ciudadanos en general. Dicho de forma sencilla: Una Infraestructura de Datos Espaciales (IDE) es una iniciativa que reúne acuerdos políticos, tecnologías, datos y servicios estandarizados que permiten el acceso e intercambio a diferentes niveles de uso de información geográfica.

Tecnológicamente una IDE es un sistema integrado por un conjunto de recursos técnicos (catálogos, servidores, programas, datos, aplicaciones, páginas Web,...) dedicados a gestionar la Información Geográfica (mapas, ortofotos, imágenes de satélite, topónimos,...), disponibles en Internet, que cumpla una serie de condiciones de interoperabilidad (normas, especificaciones, protocolos, interfaces,...). Este sistema debe permitir que un usuario, utilizando un simple navegador, pueda utilizarlos y combinarlos según sus necesidades y licencias de uso. (<http://www.idee.es>)

5.3.1 Dominio de las IDEs.

Conforme ese sistema llamado IDE se contruya para servir la información geográfica de un municipio o de una provincia o de una nación, el sistema será local, provincial o nacional. También puede ocurrir que una IDE se construya sobre una característica especial transadministrativa, como por ejemplo, la cuenca de un río. Tal es el caso del Proyecto SDIGER (<http://sdiger.unizar.es/index.vm?page=home&lang=es>) que une las cuencas del Ebro y del Garona ambas con territorios en Francia y España. En todo caso, cada IDE debe poder integrarse (gracias a los estándares de los que dispone) dentro de otra IDE superior hasta llegar a una IDE regional (INSPIRE para Europa) o a una mundial (GSDI a nivel Global).

El establecimiento de una IDE, a nivel local, regional, estatal o global, requiere del acuerdo de los productores, integradores y usuarios de datos espaciales del ámbito territorial en el que se establece. Este acuerdo debe considerar también las IDE definidas, o en definición, en otros ámbitos territoriales superiores, hacia las cuales deberá converger.

5.3.2 Justificación del establecimiento de las IDEs

Con anterioridad se ha hablado de la necesidad de disponer de la IG existente de manera ubicua. La justificación del establecimiento de una IDE, está ligada fundamentalmente a dos ideas:

- Se debe acceder fácil, cómoda y eficazmente a los datos geográficos existentes. La Información Geográfica, que es un recurso de costosa producción y difícil acceso (por motivos de formatos, modelos, políticas de distribución, falta de información) debe dejar de serlo
- La Información Geográfica debe poder reutilizarse una vez que ha servido para el proyecto para el que ha sido adquirida. Se debe reutilizar la Información Geográfica generada en un proyecto para otras finalidades diferentes, dado el alto coste de su producción.

5.3.3 Principios de las IDEs

(Desde aquí hasta el final del capítulo corresponde casi exactamente a los contenidos de la página web de la IDEE http://www.idee.es/show.do?to=pideep_Info_IDEs.ES)

Todas las iniciativas para el establecimiento de una IDE incluyen unos principios comunes:

Marco Institucional: el establecimiento de acuerdos entre los productores de información geográfica, especialmente entre los productores oficiales, para generar y mantener los datos espaciales fundamentales para la mayoría de las aplicaciones basadas en sistemas de información geográfica.

Estándares: el establecimiento de normas a las que deberá ajustarse la información geográfica, los intercambios de ésta y la interoperación de los sistemas que la manejan.

Tecnologías: el establecimiento de la red y mecanismos informáticos que permitan: buscar, consultar, encontrar, acceder, suministrar y usar los datos espaciales o geográficos. Como por ejemplo permitir incorporar los metadatos organizados en catálogos y ofrecerlos a través de servidores en red.

Política de datos: El establecimiento de las políticas, alianzas y acuerdos de colaboración necesarios para aumentar la disponibilidad de datos espaciales y compartir los desarrollos tecnológicos.

5.3.4 Componentes de las IDEs

En esencia una IDE está compuesta de:

- Datos
- Metadatos
- Servicios

5.3.4.1 Datos

En la actualidad existe un consenso internacional que clasifica los datos espaciales que pueden manejar las IDEs en:

5.3.4.1.1 Datos de referencia

Son los datos sobre los que construir o referenciar cualquier otro dato fundamental o temático. Constituyen el marco de referencia que proporciona el contexto geográfico a cualquier aplicación.

La iniciativa europea INSPIRE ha definido los temas que deben ser considerados como Datos de Referencia, en los Anexos I y II en la Propuesta de Directiva por la que se establece una Infraestructura de Información Espacial de la Comunidad (INSPIRE):

- Sistema de Coordenadas.
- Cuadrículas Geográficas.
- Nombres geográficos.
- Unidades Administrativas.
- Redes de Transporte.
- Hidrografía.
- Lugares Protegidos.
- Elevación.
- Identificadores de Propiedad.
- Parcelas Catastrales.
- Cubierta Terrestre.
- Ortoimágenes.

5.3.4.1.2 Datos temáticos

Son los datos específicos que explotan la Información Geográfica con una finalidad concreta. Proporcionan información de un fenómeno concreto (clima, educación, industria, etc.) de una región o de todo el país. Incluyen valores cualitativos y cuantitativos que se referencian espacialmente con los datos de referencia.

5.3.4.2 Metadatos

Los metadatos de la Información Geográfica informan a los usuarios sobre las características de los datos geográficos existentes. Con esta información, los usuarios pueden entender “qué es lo que representan” y “cómo lo representan” y puedan buscar y seleccionar los datos que más les interesan. Deben así mismo sean capaces de explotarlos de la manera más eficaz posible. Para ello la información incluida en los metadatos debe describir:

- la fecha de los datos (captura, edición, actualización)
- el contenido,
- la extensión geográfica que cubren,
- el sistema de referencia espacial,
- el modelo de representación espacial de los datos,
- su distribución,
- restricciones de seguridad y legales,
- frecuencia de actualización,
- calidad, etc.

5.3.4.2.1 Los objetivos de los metadatos son:

- **La búsqueda de conjuntos de datos:** saber qué datos existen, qué datos hay disponibles de una cierta zona, de un tema determinado, a una escala, de una fecha o en general de unas características específicas que el usuario demanda. Para ello los metadatos almacenan información sobre el conjunto de datos: el qué es dicho conjunto, el por qué se ha elaborado, el cuándo, el quién lo ha producido y el cómo, etcétera.
- **La comparación entre distintos conjuntos de datos entre sí,** de modo que se pueda seleccionar cuáles cumplen los requisitos del usuario de manera más adecuada para el propósito perseguido.
- **El acceso a la descripción de todas características técnicas de los datos,** de la manera más objetiva, más amplia y completa, con la finalidad de permitir su explotación eficaz. Sirve de ayuda a los usuarios de los datos tanto en la obtención de resultados como en su mantenimiento y actualización

5.3.4.2.2 Normas y perfiles de metadatos de la IG

La estructura y el contenido de los metadatos deben estar basados en una norma internacionalmente aceptada.. Uno de los beneficios de las normas es que son fruto de la experiencia y del consenso, ya que han sido desarrolladas y revisadas por grupos internacionales de expertos. En particular, las normas ISO19100 relativas a Información Geográfica proporcionan una base desde la que pueden desarrollarse perfiles, o particularizaciones de la norma, nacionales y sectoriales.

En la actualidad existen diferentes normas y perfiles dentro del campo de los metadatos que es interesante mencionar:

- ISO 19115
- Núcleo Español de Metadatos
- Iniciativa Dublin Core de Metadatos

5.3.4.3 Servicios

Una IDE es en realidad un conjunto de servicios, que ofrecen una serie de funcionalidades que resultan útiles e interesantes a una comunidad de usuarios de datos geográficos. En tanto que en los primeros años se ponía el énfasis en los datos a los que el usuario podía acceder, ahora se pone el énfasis en los servicios, en la utilidad. Al usuario no le interesa ya tanto descargarse los datos en su sistema (puesto que éstos estarán siempre disponibles y actualizados), sino que le interesa obtener directamente las respuestas que un servicio le ofrece a sus preguntas. Estos servicios IDE ofrecen funcionalidades accesibles vía Internet mediante un simple navegador o browser, sin necesidad de disponer de otro software específico.

Se describen a continuación someramente los Servicios más importantes que son todos especificados por el Open Geospatial Consortium (OGC), posiblemente la organización actual más importante en el establecimiento de estándares relacionados con la IG.

5.3.4.3.1 Servicio de Mapas en Web (WMS)

Su objetivo es poder visualizar Información Geográfica. Proporciona una representación que puede provenir de un fichero de datos de un SIG, un mapa digital, una ortofoto, una imagen de satélite,... Está organizada en una o más capas, que pueden visualizarse, hacerse transparente y ocultarse una a una. Se puede consultar cierta información disponible y las características de la imagen del mapa.. Permite superponer visualmente datos vectoriales o ráster, en diferente formato, con distinto Sistema de Referencia y Coordenadas y situados en distintos servidores.

5.3.4.3.2 Servicio de Fenómenos en Web (WFS)

Permite acceder y consultar todos los atributos de un fenómeno geográfico como un río, una ciudad o un lago, representado en modo vectorial, con una geometría descrita por un conjunto de coordenadas. Habitualmente los datos proporcionados están en formato GML, pero cualquier otro formato vectorial puede ser válido.

5.3.4.3.3 Servicio de Coberturas en Web (WCS)

Es el servicio análogo a un WFS para datos ráster. Permite no solo visualizar información ráster, como ofrece un WMS, sino además consultar el valor de atributos o atributos almacenados en cada píxel.

5.3.4.3.4 Servicio de Nomenclátor (Gazetteer)

Permite localizar un fenómeno geográfico mediante su nombre. Devuelve la localización, mediante unas coordenadas, del fenómeno en cuestión. Adicionalmente, la consulta por nombre permite fijar otros criterios como la extensión espacial en que se desea buscar o el tipo de fenómeno dentro de una lista disponible (río, montaña, población,...).

5.3.4.3.5 Servicio de Catálogo (CSW)

Permite publicar y buscar información de datos, servicios, aplicaciones y en general de todo tipo de recursos. Los servicios de catálogo son necesarios para búsquedas y llamar a los recursos registrados dentro de una IDE.

5.3.4.3.6 Descriptor de Estilo de Capas (SLD)

Esta especificación describe un conjunto de reglas que permite al usuario definir estilos personalizados de simbolización de las entidades geográficas.

Los servicios OGC pueden encadenarse y combinarse en un Geoportal, ofreciendo por ejemplo la posibilidad de: buscar un fenómeno por nombre (Nomenclátor) y visualizar el resultado sobre unos datos de referencia (WMS); localizar un producto seleccionando algunas características (Catálogo) y visualizarlo en pantalla (WMS o WCS). También es posible basarse en un servicio OGC para implementar servicios que ofrezcan funcionalidad adicional, por ejemplo desarrollar un servicio de camino mínimo por carretera basado en un WFS que acceda a todos los atributos de un conjunto de datos de poblaciones y carreteras.

5.4. Bibliografía y enlaces

Enlaces:

- El Recetario IDEs (Traducción de The SDI CookBook. Nevert, D.D. Editor). 2001. <http://redgeomatica.rediris.es/metadatos/publica/recetario/html/>
- El Recetario IDEs (2004). Versión nueva y actualizada de la referencia anterior. <http://redgeomatica.rediris.es/metadatos/publica/recetario/html/>
- La IDE de España <http://www.idee.es/>