

INFRAESTRUCTURAS DE DATOS ESPACIALES: LA INICIATIVA EUROPEA INSPIRE Y EJEMPLO DE SU APLICACIÓN¹

Rubén Béjar, Pablo Gallardo, Javier Nogueras-Iso, P.R. Muro-Medrano, F.J. Zarazaga-Soria

Dpto. de Informática e Ingeniería de Sistemas
Universidad de Zaragoza
María de Luna 1, 50018 Zaragoza

Introducción

Los sistemas de información geográfica (SIG) han demostrado, durante casi cuatro décadas, su utilidad en el inventario y gestión del territorio. Sin embargo, su puesta en marcha sigue causando problemas y siendo cara, en su mayoría, debido al coste de la adquisición y mantenimiento de los datos. No es que los datos no existiesen, sino más bien que no han sido fácilmente accesibles y esto ha conllevado la práctica habitual de duplicación de grandes bases de datos, con su correspondiente duplicación de esfuerzos y presupuestos.

La Comisión Europea, en reconocimiento al problema de la falta de reutilización, y acceso a la gran cantidad de información archivada por el sector público, publicó en 1.999 un informe verde² dedicado a la necesidad de crear formatos y servicios de metadatos para facilitar la indexación, búsqueda, recuperación e intercambio de estos datos. De acuerdo con las conclusiones del informe verde, y con el nacimiento de nuevas tecnologías basadas en la WWW, varias organizaciones de las más representativas global, europea y nacionalmente, en el campo de la Información Geográfica, se han puesto de acuerdo para impulsar la creación de *Infraestructuras de Datos Espaciales* (IDE, SDI en sus siglas anglosajonas) con el fin de maximizar la explotación y el intercambio de estos datos, que en adelante llamaremos *geodatos*. Más que la última *moda* del mundo de las tecnologías de la información, estas infraestructuras definen un marco de normas y estándares a seguir para guiar a los productores de geodatos (en su mayoría del sector público) en la recolección, mantenimiento e intercambio de los datos necesarios, de forma que se cumpla con sus mandatos en una manera coherente, **sostenible** e **interoperable** (NSDI, 1994).

Los planes para la adopción de la legislación marco de INSPIRE en 2.006-07 están forzando a los gobiernos nacionales y regionales de la Unión Europea a dar pasos

¹ La tecnología de base de este proyecto ha estado parcialmente financiada por el proyecto TIC2003-09365-C02-01 del Plan Nacional de Investigación Científica y Desarrollo Tecnológico del Ministerio de Ciencia y Tecnología de España.

² [http://europa.eu.int/ISPO/docs/policy/docs/COM\(98\)585/index.html](http://europa.eu.int/ISPO/docs/policy/docs/COM(98)585/index.html)

efectivos hacia la creación de las infraestructuras de datos espaciales y de servicios que esta legislación establece. Actualmente esto ya es viable debido a que los estándares y las arquitecturas que serán adoptados ya han sido desarrollados, propuestos y comprobados, y a que existen implementaciones para los distintos componentes necesarios.

La iniciativa INSPIRE

INSPIRE: Concepto y principios

La iniciativa INSPIRE (“INfrastructure for SPatial InfoRmation in Europe”: <http://www.ec-gis.org/inspire/>) surge ante la situación en la que se encuentra, con carácter general, la información espacial en Europa. Una situación que nace, entre otros, de la fragmentación de los datos, falta de armonización entre datos de diferentes escalas geográficas, problemas con la disponibilidad de información y duplicación de la misma. Esta iniciativa, promovida por la Comisión Europea y desarrollada en colaboración con los Estados Miembros, intenta incentivar la creación de una infraestructura europea de información espacial que proporcione a los usuarios una serie de servicios integrados que permitan identificar y acceder a la información geográfica. Los usuarios de esta infraestructura serían los políticos, los planificadores locales, nacionales y europeos, los ciudadanos y sus organizaciones.

Los principios de INSPIRE son los siguientes:

- Los **datos (básicos) deben ser recogidos una vez** y mantenidos al nivel donde esto pueda ser realizado de forma más eficaz.
- Debería ser posible **combinar, de forma simple, información espacial** de distintas fuentes en Europa y compartirla entre muchos usuarios y aplicaciones.
- Debería ser posible **que la información recogida a un nivel sea compartida entre distintos niveles**, detallada para investigaciones y general para propósitos estratégicos.
- La **información geográfica** necesaria para el gobierno a todos los niveles debería ser **abundante** y estar disponible **bajo condiciones que no frenen su uso masivo**.
- Debería ser sencillo **descubrir qué información geográfica está disponible**, cumple con las necesidades de un uso particular y bajo qué condiciones puede ser adquirida y usada.
- Los datos geográficos deberían ser **sencillos de entender e interpretar** y poder ser visualizados en el contexto apropiado a distintos usuarios.

Arquitectura

Una de las primeras decisiones del grupo de trabajo de INSPIRE ha sido la elección de un modelo conceptual hacia el cuál enfocar la iniciativa. El modelo elegido fue el Modelo de Referencia del proyecto *Digital Earth* de la NASA (ver *Figura 1*). Fue elegido por su alto nivel de simplicidad y síntesis, representando en unos pocos gráficos y tablas el abanico de posibles normas y estándares.

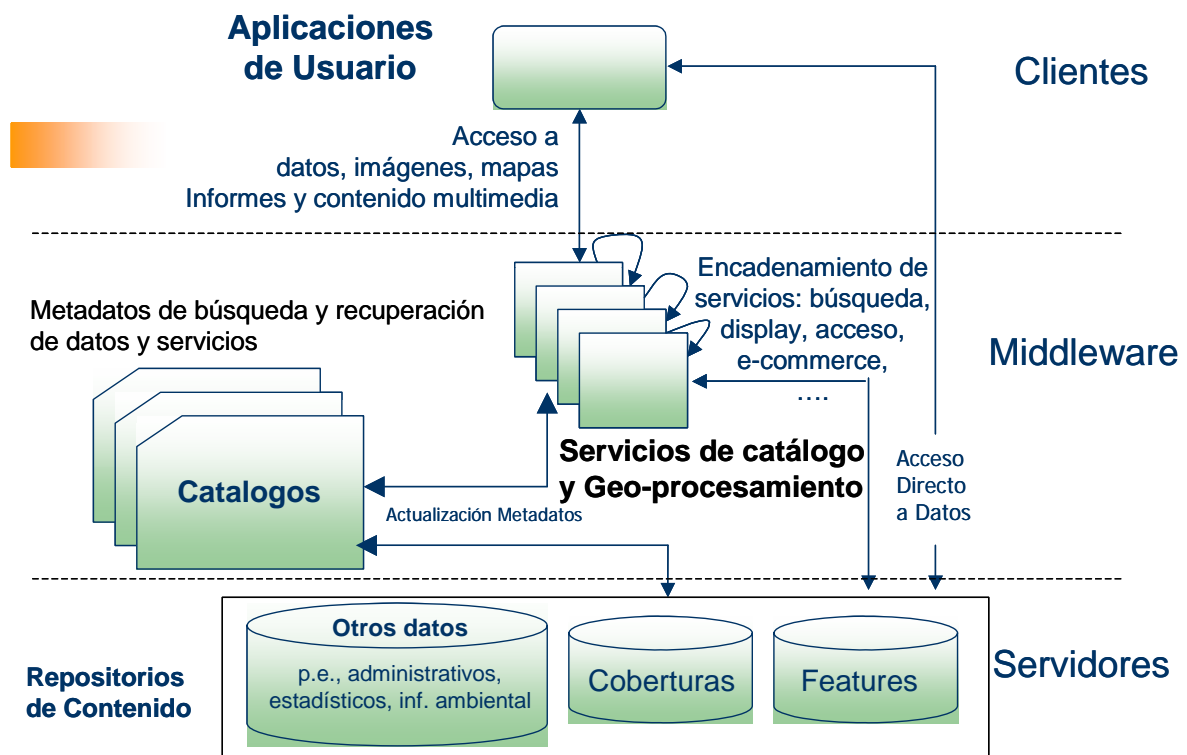


Figura 1 Esquema arquitectónico de una IDE, según la recomendación de la iniciativa europea INSPIRE y basado en el Modelo de Referencia del Digital Earth (NASA, 2001).

El término arquitectura en este contexto se entiende como los modelos, los estándares, la tecnología, las especificaciones y los procedimientos usados para representar, transformar, mantener y hacer uso de la información en formato digital.

Aplicaciones de usuario

Por aplicación de usuario se entiende el software que normalmente percibirá el usuario, es decir, las interfaces para las búsquedas y consultas, las herramientas para administrar las bases de datos y algunas aplicaciones de análisis. Las más importantes serán:

- **Publicación de datos y metadatos:** Ésta es una de las funcionalidades más importantes. Al permitir publicar los datos y metadatos se está estableciendo la conexión entre los que producen los datos y los que los necesitan. Para este propósito, es necesario que los productores de datos sean los responsables de elaborar sus correspondientes metadatos. La experiencia ha dejado claro que una buena documentación de los datos facilita en gran medida la gestión de los mismos.
- **Búsquedas:** Aquellas funcionalidades que nos permiten encontrar la información que necesitamos (y que por supuesto haya sido previamente publicada) Las búsquedas, generalmente, se realizan siguiendo unos criterios; espaciales (mapas de Andalucía, fotos aéreas de Portugal...), temáticos (mapas forestales, geológicos, de carreteras...), tipo (mapas vectoriales, ortofotografías, imágenes de satélite...).

- **Visualizaciones:** Las visualizaciones de mapas en Internet se realizan a través de servicios conocidos como Web Map Servers (WMS, 2002). Esta funcionalidad permite normalmente seleccionar los temas que queremos ver, superponerlos, hacer zooms y moverlos por el mapa.
- **Adquisición de datos:** En muchos casos a un usuario le basta con ver el mapa mediante un Web Map Server para obtener la información que necesita. Pero algunos usuarios necesitarán tener esa información, y no sólo visualizarla mediante una interfaz, sobre todo en casos en los que haga falta un análisis más complejo de los datos. Para ello se ha pensado en servicios de encargo de datos, en los que el usuario, tras encontrar los datos que necesita, remite su solicitud al responsable de dichos datos y este se los envía. La transferencia se puede realizar mediante e-mail o en caso de grandes volúmenes de datos (como las imágenes de satélite) mediante CDs. Los servicios Web de acceso a datos se basan en los estándares denominados Web Feature Server (WFS, 2002) para la recuperación de entidades de carácter vectorial y Web Coverage Server (WCS, 2003) para el caso de coberturas de tipo ráster.
- **Análisis:** Relacionado con el concepto de Web Map Server, pero yendo algo más lejos en términos de funcionalidad, está el uso de herramientas de análisis en Internet. Por ahora, los análisis que se pueden realizar son simples, pero poco a poco, se irán incrementando estas funcionalidades.
- **Geoportal de la Unión Europea:** Todas las funcionalidades mencionadas anteriormente, y otras, deberían estar localizadas en un portal de Internet en el que confluyeran, tanto a nivel de usuarios como de ofertantes, todos los países de la Unión Europea. Sería el llamado Geoportal de la Unión Europea, donde confluirían todas las funcionalidades conformando una única aplicación de usuario.

Servicios de geoprocésamiento

Estos servicios podrían definirse como el motor que mueve todas las funcionalidades de la aplicación de usuario descrita en el apartado anterior. Mientras la aplicación de usuario es en cierta medida la interfaz que los usuarios pueden ver e interactuar, el servicio de geoprocésamiento está por debajo, realizando las funciones requeridas por la aplicación, y no es visible para los usuarios.

Catálogos y servicios de catálogo

Los catálogos pueden ser considerados como el corazón de una infraestructura de datos espaciales, puesto que en ellos se almacenan los metadatos que permiten las búsquedas y el manejo de la información espacial.

El consorcio OpenGeospatial (OGC, anteriormente denominado OpenGIS Consortium, (OGC, 1996)), define un catálogo como el *conjunto de servicios que soporta la organización, la búsqueda y el acceso a la información espacial*. Se puede pensar en un catálogo como una base de datos especializada de información referente a recursos geoespaciales (metadatos) disponibles para una comunidad de usuarios (Nogueras, 2005).

Como venimos explicando, los metadatos son la materia prima de un catálogo y su principal finalidad es hacer que los datos existentes se puedan encontrar y compartir.

Almacenes de datos

Como el nombre sugiere, el almacén de datos es el lugar donde se guardan todos los datos geoespaciales. Por datos geoespaciales no sólo se entiende los mapas en formato vectorial, o los mapas ráster, o las ortofotos, sino todo aquel dato que tenga una referencia espacial implícita; datos tabulares con un identificador geográfico, fotografías o videos de algún lugar en concreto...

Estándares

ISO (International Standardisation Organisation) define estándares como *acuerdos documentados que contienen especificaciones técnicas u otros criterios precisos para ser usados como reglas, directrices o definiciones de características que aseguran que los materiales, procesos, productos y servicios encajan con sus respectivos propósitos* (ISO/TC 211, <http://www.isotc211.org>). Así, los estándares nos proporcionan una serie de métodos y herramientas para manejar, adquirir, procesar, analizar, acceder, presentar y transferir datos entre diferentes usuarios, diferentes sistemas y diferentes lugares.

Para que los distintos organismos se acomoden a los nuevos estándares y especificaciones se debe crear y mantener la denominada “INSPIRE Guidelines and Best Practice for Architecture and Standards”. Esta directriz se elaborará y corregirá en estrecha colaboración con los expertos de cada uno de los países miembros.

Aspectos importantes para la implementación

Existe toda una serie de aspectos que, por su relevancia a la hora de construir los sistemas que den soporte a las infraestructuras de datos espaciales deben ser tenidos en cuenta. Los principales son:

- **Interoperabilidad semántica:** habitualmente la semántica de los datos espaciales, los nombres y topónimos que manejan, vienen muy definidos por la comunidad a la que pertenezcan (cada país tendrá su idioma, cada sector temático tendrá sus propias palabras). Para que sea posible compartir los datos, buscar y acceder a ellos, es necesario que sus metadatos sean comprensibles para cualquier usuario de la unión europea, sin importar de que país sea, o en que sector temático trabaje. Para conseguir la interoperabilidad semántica debemos considerar los aspectos multilingües y el uso de tesauros o palabras controladas (Nogueras, 2004^a) (Nogueras, 2004b).
- **Identificadores únicos:** Cada información geográfica individual (un determinado mapa de carreteras, una determinada imagen de satélite) deberá tener su identificador único. La finalidad de éstos es:
 - Identificar cada información geográfica para poder acceder a ella y manipularla dentro de una base de datos.
 - Poder agregar datos; hacer de un conjunto de datos individuales, uno único que los abarque a todos.
 - Soportar las actualizaciones de los datos temáticos.
- **Actualización y versiones de los datos:** Las características temáticas y la geometría de los objetos de la superficie terrestre (tanto los usos de suelo, como la vegetación, las infraestructuras...) están en constante

cambio. Por lo tanto, la labor de actualización de los datos es un trabajo interminable. En la mayoría de las comunidades se sigue teniendo la costumbre de volver a realizar toda la cartografía de un área cuando pasa el tiempo y las características cambian. Esto es una gran pérdida de tiempo y esfuerzo si se considera que a lo largo del tiempo hay objetos que cambian, y que deben actualizarse, pero hay otros que permanecen iguales. Para solventar este problema se propone crear un sistema mediante el cual, en un conjunto de datos, sólo se actualicen aquellos objetos que cambian. Además, deberían quedar almacenadas las versiones de épocas anteriores, para posibilitar el acceso a datos de diferentes momentos y poder analizar los cambios en el tiempo o combinarlos con otros datos de la misma época.

La IDE de la Consellería de Medio Ambiente de la Xunta de Galicia (IDE Rimax)

En esta sección se presenta el caso de la IDE de la Consellería de Medio Ambiente de la Xunta de Galicia, denominada IDE Rimax, que es un ejemplo de cómo un SIG corporativo puede desarrollarse siguiendo arquitecturas estándar siendo de esta forma una IDE de pleno derecho, lista para integrarse en una futura IDE de Galicia, de España y de Europa.

La Consellería de Medio Ambiente de la Xunta de Galicia se ha encontrado con la misma clase de problemas que INSPIRE apunta: existencia de datos en formatos incompatibles, dificultades en la distribución de la información entre sus usuarios (con el agravante de que se trata de un departamento muy descentralizado), dificultades en la búsqueda de información relevante, etc. La solución adoptada para resolver estos problemas ha consistido en el desarrollo para esta Consellería de un sistema de información basado en los principios y recomendaciones de INSPIRE en lo que respecta a arquitectura y estándares utilizados, o lo que es lo mismo una IDE.

La Consellería de Medio Ambiente de la Xunta de Galicia

Las principales responsabilidades de la Consellería de Medio Ambiente, en adelante CMA, de la Xunta son las concernientes a la conservación de la naturaleza, tratamiento de residuos, uso de las aguas y gestión de montes e industrias forestales, siendo quizás esta última la más importante, debido al hecho de que aproximadamente el 69% de la superficie gallega está cubierta de bosques.

Con el objeto de atender las necesidades (reforestación y limpieza de bosques, lucha contra incendios forestales) de una superficie arbolada tan grande, la CMA está dividida en diecinueve distritos forestales. Se trata, por tanto, de un departamento muy descentralizado, con tan sólo un 10% de su personal trabajando en la sede central de Santiago de Compostela, y con el resto repartido en las delegaciones provinciales y en los distritos forestales. Actualmente, este Departamento está terminando de cablear las delegaciones provinciales y los distritos forestales, de modo que todas las delegaciones y distritos puedan conectarse a la intranet de la Consellería.

La descentralización de la CMA agrava seriamente los problemas habituales con la información geográfica existentes en grandes organizaciones y administraciones públicas: para los usuarios es difícil encontrar la información geográfica que precisan, e incluso saber si ésta existe. En algunos casos, como en el de los empleados de los

distritos forestales que trabajan fuera de la sede central, los usuarios no tienen ningún tipo de acceso a información geográfica que podría hacer su trabajo mucho más fácil.

Otro problema añadido es que el retraso en la construcción de una solución integrada al problema ha propiciado que algunos de los distritos hayan adoptado distintas soluciones utilizando productos SIG diferentes, e incluso, en alguno de ellos, a no adoptar ningún tipo de solución en absoluto. Esta situación ha terminado añadiendo problemas de formato de datos e interoperabilidad a los problemas existentes.

Requisitos

El primer paso para la construcción de la IDE de la Consellería de la Xunta de Galicia consistió en el análisis de requisitos de usuario. A continuación se presenta un resumen de las más relevantes:

- Toda la información geográfica, tanto vectorial como *ráster*, debe estar almacenada en una base de datos con capacidad espacial.
- Debe ser posible realizar búsquedas y encontrar la información geográfica relevante disponible.
- Los usuarios deben disponer de herramientas que les permitan visualizar la información geográfica que necesiten desde la intranet de la CMA. Deben ser igualmente capaces de realizar consultas espaciales y no espaciales sobre esa información.
- Los usuarios deben tener la posibilidad de descargar la información geográfica (tanto vectorial como *ráster*) en el formato en el que suelen trabajar.
- Los usuarios avanzados deben tener acceso a servicios de mapas simples, diseñados para ser compatibles con la información que posean localmente y tendrán la posibilidad de realizar consultas de la información visualizada en esos mapas.
- El software del que la CMA posea licencias deberá ser utilizado si es posible. Este software incluye: *ArcSDE* 8.x y *ArcIMS* 4.x de *ESRI*, *Oracle* 8i/9i espacial y el traductor universal *FME* y su versión de web *Spatial Direct* de *Safe Software*.

Asimismo, tras estudiar las necesidades de la CMA y las directrices ofrecidas por INSPIRE y GSDI (Global Spatial Data Infrastructure, <http://www.gsdi.org>) se añadieron otros requisitos a los establecidos por la CMA:

- En el caso de que estén disponibles, y si es posible, se deben utilizar estándares; especialmente aquellos recomendados por INSPIRE y GSDI.
- La arquitectura debe seguir las recomendaciones de INSPIRE: datos, catálogo de metadatos y servicios encadenables que ofrezcan, al menos, visualización, acceso y búsqueda de información.

Diseño de servicios Web

La Figura 2 muestra una visión orientada a servicios de la IDE de la CMA. Esta visión pone de relieve algunos aspectos básicos que forman el núcleo de una IDE:

- Los principales componentes de una IDE son los servicios web encadenables. A través del establecimiento de un conjunto de servicios web conforme a los estándares del OGC, se consigue fácilmente la interoperabilidad sintáctica. Esto facilita la integración de distintos servicios, ya sean procedentes de la propia infraestructura o de otra externa, de forma que se puedan desarrollar de forma rápida aplicaciones temáticas o adaptadas a la evolución de las necesidades concretas de los usuarios.

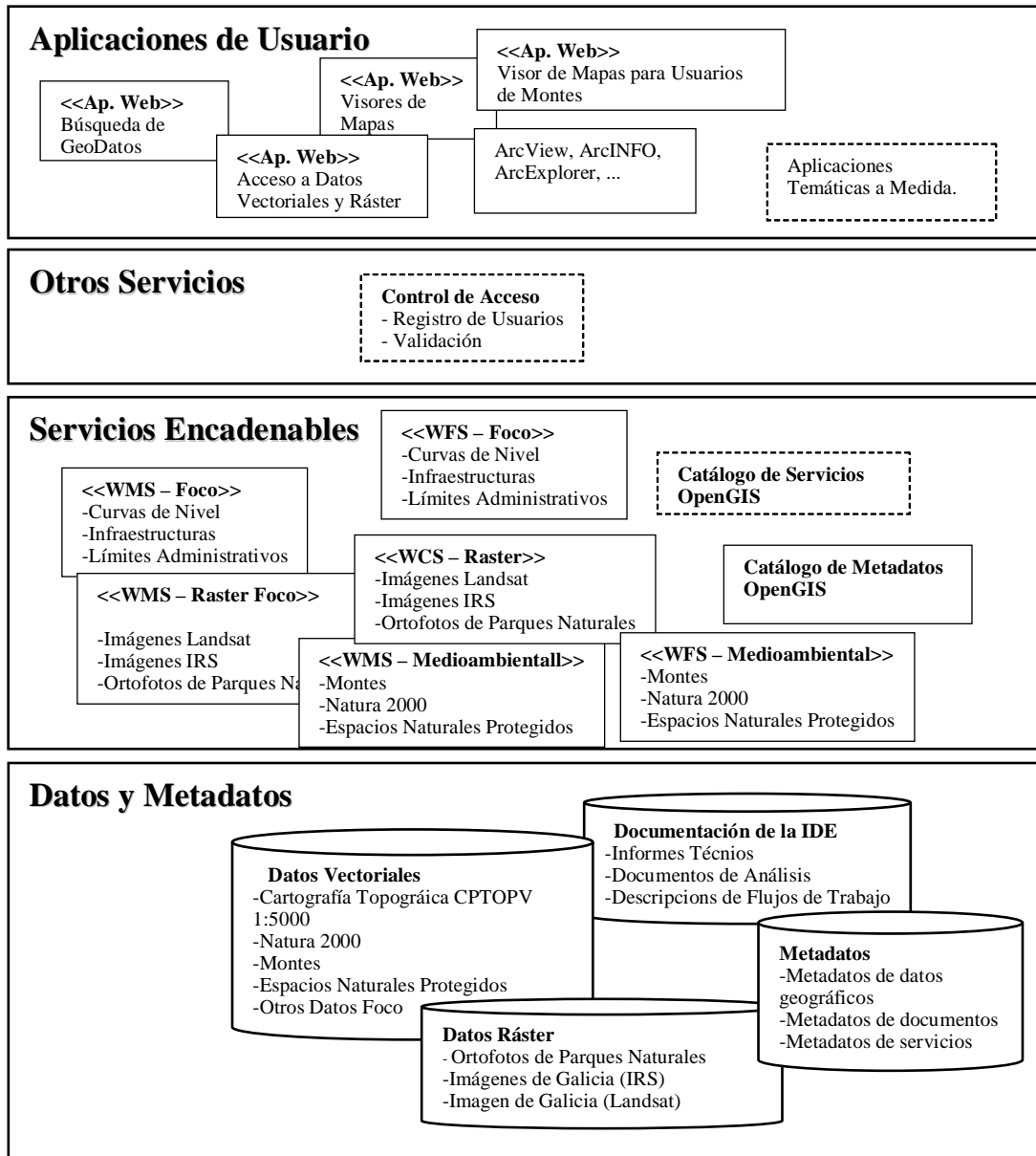


Figura 2: Vista de la arquitectura orientada a servicios

- También destaca que estos servicios encadenables se construyen sobre información geográfica y sobre metadatos, posibilitando una mayor interoperabilidad semántica, ya que la información geográfica es descrita por los metadatos que tiene asociados. La inclusión de metadatos junto con información geográfica temática y de referencia es también una recomendación de INSPIRE.

- Las aplicaciones de usuario se construyen sobre los servicios distribuidos, tanto sobre los servicios encadenables estándar como sobre los servicios de integración. Los servicios de integración son los que se proporcionan funcionalidades adicionales, es decir, no cubiertas por los servicios estándar pero que no están vinculadas a una aplicación final de usuario.

Después del estudio de las necesidades de usuarios y de los datos geográficos disponibles, cuya discusión se retomará más adelante, se definió el conjunto de servicios web de visualización y de acceso a datos (siguiendo los estándares de OGC sobre servidores web de mapas (WMS), de entidades (WFS) y de coberturas (WCS), teniendo como objetivos:

- permitir la visualización y el acceso a todo tipo de coberturas vectoriales y *ráster*,
- facilitar la integración y adaptación de los diferentes datos,
- establecer un número adecuado (ni excesivamente alto, ni demasiado bajo) de servicios web, con el objeto de facilitar su mantenimiento y
- agrupar elementos relacionados con el fin de hacerlo más intuitivo.

Se establecieron tres categorías conceptuales simples para los datos (una con información vectorial de referencia y topografía básica, otra con información vectorial medioambiental relativa a la CMA y una tercera con información *ráster*), diseñándose consecuentemente los servicios web. Los servicios de visualización y acceso planeados se muestran en la Tabla 1.

	WMS	WFS	WCS
Datos de referencia vectoriales	X	X	
Datos medioambientales vectoriales	X	X	
Datos <i>ráster</i>	X		X

Tabla 1: Relación entre servicios planeados y tipos de datos

Además del acceso y visualización de los datos, la otra funcionalidad principal requerida es la búsqueda de datos y la gestión de sus metadatos. Las especificaciones sobre catálogos del OGC cubren estas áreas, por lo que se incluyó un catálogo de metadatos (del mismo modo se incluirá en el futuro un catálogo de servicios).

Datos y metadatos

Antes de construir una infraestructura de datos espaciales, es importante obtener una relación de los datos existentes y de aquellos a los que podría accederse, ya sea por compra o solicitud. En este último caso, por ejemplo, la Consellería de Política Territorial, Obras Públicas e Vivenda (CPTOPV) disponía de cartografía a escala 1:5000 que incluye información topográfica, límites administrativos, infraestructuras de comunicación, entre otras. La información estaba almacenada en un total de 4000 archivos en formato *dgn* que cubren todo el territorio de Galicia, aunque más orientada a ser utilizada por un sistema CAD que por un SIG (por ejemplo, las líneas de contorno están cortadas para ubicar etiquetas de texto), lo que dificulta la extracción de entidades de ellos.

Además la propia CMA disponía de una ingente información, principalmente en formato *shapefile*, sobre los montes donde realiza la gestión, bien sea por ser de propiedad pública, por existir un convenio o consorcio en ellos o por ser espacios incluidos en la Red Natura 2000.

Para complementar la información vectorial disponible, la CMA, en conversaciones con otros entes públicos de Galicia y compañías privadas, obtuvo las imágenes satélite de Landsat e IRS que cubrían todo el territorio gallego y que se complementaban en algunas zonas con ortofotografías más precisas, dotando así a los usuarios de completa información *ráster* a diferentes resoluciones o escalas.

De acuerdo con el espíritu de la iniciativa INSPIRE, la compra de datos no debería haber sido necesaria, ya que la labor de proporcionar acceso a esta información debería haber sido responsabilidad de otros departamentos de la Xunta, pero, dado que la de la CMA es la primera iniciativa de construcción de una IDE en Galicia, y que ésta debe proporcionar a los usuarios un servicio lo más útil y completo posible, hubo que tomar decisiones como ésta.

Uno de los requisitos de usuario es el de que toda la información debe ser almacenada en una base de datos espacial. Las dos secciones siguientes describen el proceso seguido para el almacenamiento de la información disponible, tanto vectorial como *ráster*, en la base de datos. Finalmente hay una sección que describe el proceso de creación de los metadatos, necesarios para poder realizar búsquedas de datos en el catálogo.

Almacenamiento de datos vectoriales

El objetivo era el de facilitar el proceso mediante la utilización de un software traductor universal, *FME*, de forma que pudieran almacenarse todas las coberturas vectoriales en el formato de *Oracle Spatial*. Su utilización fue necesaria ya que en la recopilación de la información disponible, existían datos en varios formatos diferentes. Además la herramienta de traducción se apoyó en un software mediador (*ArcSDE*) con el objeto de facilitar el acceso a la información tanto por software directamente compatible con *Oracle Spatial*, como por software compatible con *ArcSDE*.

Este proceso, brevemente descrito, consistió en el diseño de filtros *FME* (traductores semánticos y sintácticos que describen el proceso de transformación entre un formato de origen y otro de destino) y en aplicarlos a todos los ficheros de datos que debían ser transformados.

La realización de estas tareas tomó mucho más tiempo del que inicialmente había sido planeado: aproximadamente 12 meses-hombre fueron necesarios para el almacenamiento de la información vectorial, cuando inicialmente se estimó que, como mucho, serían necesarios entre 4 y 6 meses-hombre. Los principales problemas, y soluciones aplicadas, fueron los siguientes:

- Especificación de los datos pobre: La cartografía 1:5000 del CPTOPV, principal conjunto de datos que debía ser almacenado en *Oracle*, estaba muy pobremente especificada. La intención inicial era la de dividir las coberturas temáticas originales en entidades, con el objeto de permitir una mejor manipulación de las mismas. Para conseguirlo, es necesario un gran conocimiento de la cartografía. Fue necesario mantener entrevistas con los técnicos gestores de la información, y gracias a éstas se conoció que en determinados casos la creación de una misma capa fue contratada

a varias compañías distintas, en años distintos, teniendo como resultando datos cuya calidad variaban espacialmente según el territorio asignado a cada compañía creadora. Fruto de ello fue una falta de información fidedigna, que forzó a la utilización de un lento sistema de prueba y error para extraer las entidades de esas coberturas. Las entidades extraídas fueron comparadas visualmente con las de los ficheros originales con el objeto de asegurar un mínimo grado de calidad.

- Errores de geometría: “Falsos” segmentos lineales (segmentos de longitud inferior a un valor umbral muy pequeño), intersecciones de límites de polígonos consigo mismos, intersecciones en las curvas de nivel y otros problemas similares fueron encontrados a lo largo del proceso de almacenamiento. Esta clase de problemas impedía a *FME* almacenar los datos correctamente en *Oracle* o impedían a *Oracle* crear índices espaciales (de vital importancia para garantizar un acceso eficiente) sobre la misma, impidiendo la creación de procesos automatizados de almacenamiento para efectuar “cargas masivas”. La solución adoptada fue la de hacer más robustos los filtros *FME*, incluyendo comprobaciones geométricas previas a las traducciones. Otros errores fueron remitidos a los creadores o mantenedores de las coberturas para que los corrigieran.

Almacenamiento de datos ráster

El almacenamiento de los datos *ráster* (fotografías aéreas e imágenes de satélite de distintas resoluciones) necesitó una aproximación diferente debido a que *FME* no permite transformar este tipo de datos. Aunque esto en principio no supuso un problema ya que su inserción en la base de datos no presentaba problemas semánticos, surgieron otros problemas con los datos *ráster*:

- Creación de mosaicos: Las imágenes de satélite y fotografías aéreas que cubren extensas áreas suelen presentarse divididas en teselas o ficheros de tamaño manejable, que pueden ser regulares o no y que pueden presentar solapes entre ellas, marcos blancos alrededor de las mismas y otro tipo de problemas similares. Las imágenes disponibles, archivos en formato *GeoTIFF*, que iban a ser incorporadas a la IDE de la CMA, presentaban algunos de estos problemas:
 - La imagen del satélite Landsat que cubre la totalidad del territorio de Galicia consistía en un único fichero.
 - Las imágenes del IRS disponibles estaban compuestas por teselas regulares, perfectamente georreferenciadas y ocorrectificadas.
 - Las ortofotografías aéreas de parques naturales con una resolución de 1 metro, estaban formadas por teselas irregulares, con solapes y diferencias de contraste, brillo y sombra, pero no presentaban ni bordes blancos ni otro tipo de problemas más serios.

Al no disponer de software adecuado para la creación de un mosaico a partir de las fotografías aéreas, se utilizó una solución más creativa para tratar, principalmente, las ortofotografías de los parques naturales. Se optó por una aproximación basada en *ArcObjects*, los componentes

software proporcionados por los productos *Arc* de *ESRI*. Se escribió un pequeño programa en Visual Basic con el objeto de aprovechar las capacidades ofrecidas por *ArcObjects* para crear mosaicos a partir de las distintas imágenes georreferenciadas, incluidas aquellas que presentaban solapes, y para introducirlas en la base de datos a través de *ArcSDE*. Con mayor disponibilidad de tiempo y de presupuesto se habría podido hacer una aproximación más sofisticada que tuviera en cuenta el balanceado de color y brillo de las fotografías aéreas, etc., aunque la solución finalmente adoptada fue suficiente para cubrir las necesidades de los usuarios.

- Acceso rápido: Una de las principales preocupaciones con las imágenes, debido a sus grandes tamaños, es posibilitar un acceso y visualización rápidos a las mismas a cualquier resolución. La solución habitual a este problema consiste en la creación de pirámides de imágenes compuestas por la original y versiones derivadas de la misma, todas ellas adecuadamente divididas y a diferentes resoluciones; de esta forma sólo hay que proporcionar los trozos de imágenes a la resolución requerida por el usuario, incrementando la eficiencia en el acceso y posterior visualización de la información *ráster*. *ArcSDE* es capaz de crear automáticamente estas pirámides, por lo que se adoptó esta solución en la construcción de la IDE del CMA.

Proceso de creación de metadatos

La creación de metadatos o catalogación de la información espacial tiene dos objetivos principales. El primero es el de mantener y organizar la inversión económica derivada de la creación y adquisición de los datos, evitando la pérdida de valor de la información ocasionada por la paulatina pérdida del conocimiento existente sobre la misma, así como el de potenciar un uso apropiado de la misma. El segundo objetivo es el de proporcionar a los catálogos de datos la información necesaria, con el fin de permitir búsquedas y, en un horizonte más lejano, el de facilitar la integración de esta IDE en otras, siguiendo las recomendaciones INSPIRE sobre servicios distribuidos interoperables, donde se incluyen servicios de catálogo distribuidos.

Existen varios estándares de metadatos geoespaciales (algunos de ellos todavía no definitivos) como son el CSDGM del Federal Geographic Data Committee, norteamericano, o el definido por el comité ISO/TC 211, borrador del estándar internacional DIS 19115, cuyas aplicaciones eran de especial interés. Antes de comenzar el proceso de creación de metadatos se deberían tomar éstas y otras decisiones y directrices. El objetivo, por lo tanto, era crear unos metadatos tan significativos y completos como fuera posible, con un esfuerzo razonable y facilitando al máximo las búsquedas de datos que se realizaran gracias a ellos. Otro de los puntos a tener en cuenta fue el que los técnicos forestales no estaban acostumbrados a crear y gestionar metadatos, por lo que necesitaban formación, directrices claras y una aproximación sencilla a los metadatos (considerando que a medida que maduraran como usuarios de metadatos, sería más sencillo para ellos comprender, completar y mejorar los metadatos ya creados). Por lo tanto, las principales decisiones tomadas a este respecto fueron:

- Crear como mínimo los campos obligatorios del estándar ISO de catalogación (unos 40 de los aproximadamente 500 campos del estándar completo), siempre que fuese posible.
- Utilizar *Dublin Core* (un estándar genérico de metadatos que proporciona una descripción mínima de recursos para ser utilizada en procesos de búsqueda y descubrimiento) como guía para una catalogación mínima.
- Impartir seminarios a los usuarios de la CMA, así como facilitarles las herramientas adecuadas, con el fin de promover el uso de metadatos. La herramienta escogida para la creación de metadatos fue un software llamado *CatMDEdit*, desarrollado íntegramente en nuestro laboratorio, que permite la creación de metadatos y la importación y exportación desde y hacia los distintos estándares más relevantes (ISO, FGDC, *Dublin Core*) y trabajar con archivos XML y con bases de datos (*Access*, *Oracle*).

Implementación

Como el catálogo de metadatos es, según las recomendaciones técnicas de INSPIRE, el núcleo de la IDE, éste servirá como punto de partida para describir la implementación. Con el objeto de proporcionar búsquedas de datos en el catálogo se desarrolló un motor de búsquedas temático (ver Figura 3). Este motor de búsquedas proporciona una interfaz que combina temas (infraestructuras, cartografía forestal, de ordenación del territorio, ...), áreas (toda la comunidad, provincias o distritos), escalas y fechas de actualización para permitir búsquedas restringidas completamente personalizadas. Un ejemplo con los resultados de una búsqueda puede verse en el lado derecho de la Figura 3. Este listado muestra algunos metadatos para cada una de las coberturas encontradas (título, resumen, escala de trabajo, formato, fecha de actualización y creador). Haciendo clic sobre el título de cada uno de los elementos se abre una nueva ventana con los metadatos completos como se muestra en la Figura 4 (lado izquierdo).

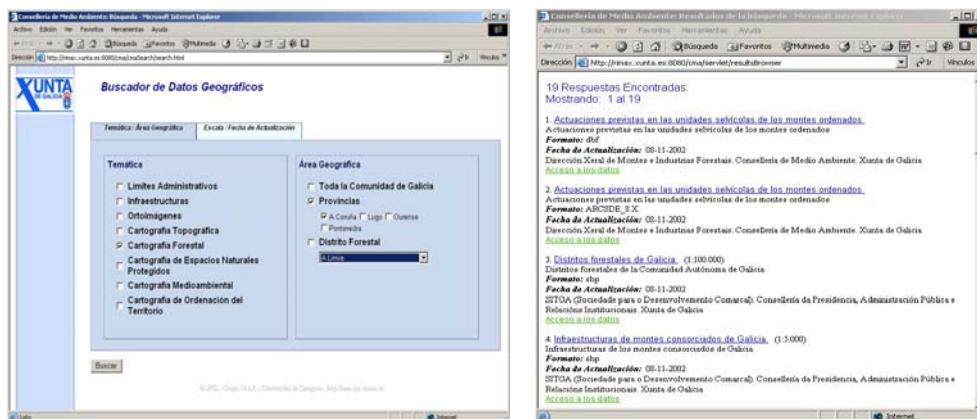


Figura 3: Búsqueda temática de datos y resultados encontrados

La principal ventaja de una arquitectura IDE es el uso de servicios “encadenables”. Esto resulta todavía más evidente cuando los usuarios solicitaron mecanismos que permitiesen conexiones entre las búsquedas de datos y los servicios de mapas, permitiendo a los usuarios que, tras la búsqueda en el catálogo y una vez encontrada la información deseada, pudieran acceder de un modo sencillo a un servicio de mapas para visualizarla. La forma de implementar esta posibilidad fue

proporcionando un enlace entre la ventana que muestra los resultados de la búsqueda (Figura 3) y una segunda ventana que muestra los servicios de mapas disponibles que presentan tal información (ver Figura 4 derecha). Para ello se desarrolló una extensión del catálogo que incluyese esta funcionalidad, con la intención de implementar un catálogo de servicios en un futuro próximo que posibilite esta y otras funcionalidades.

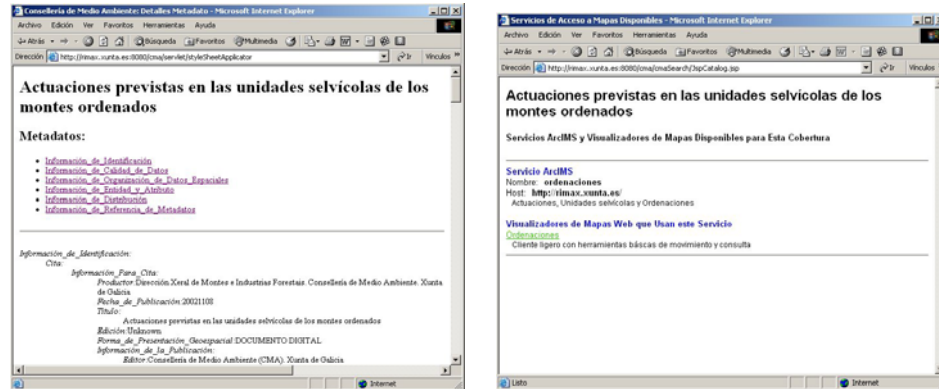


Figura 4: Metadatos completos y servicios de mapas disponibles para un conjunto de datos

Tras encontrar un servicio de mapas adecuado a sus necesidades, el usuario puede, por supuesto acceder al propio servicio siguiendo un enlace. Un ejemplo puede verse en la parte izquierda de la Figura 5, donde aparece un servidor de mapas mostrando un mapa de las provincias gallegas y sus autopistas. Se trata del visualizador de mapas en HTML del *ArcIMS*, ligeramente adaptado para satisfacer las necesidades de los usuarios de la CMA. El usuario puede detenerse aquí, si sólo necesita visualizar la información, o puede descargarse la información que le está siendo mostrada. Un enlace de hipertexto facilita el acceso al formulario de descarga de *Spatial Direct* (Figura 5 a la derecha), adaptado para proporcionar el área y la información que se encuentran seleccionados en el cliente del servidor de mapas. De este modo, todos los servicios (búsqueda, visualización y acceso) se encuentran conectados (“encadenados”), proporcionando al usuario una visión de homogeneidad e integración de los elementos que componen la IDE.

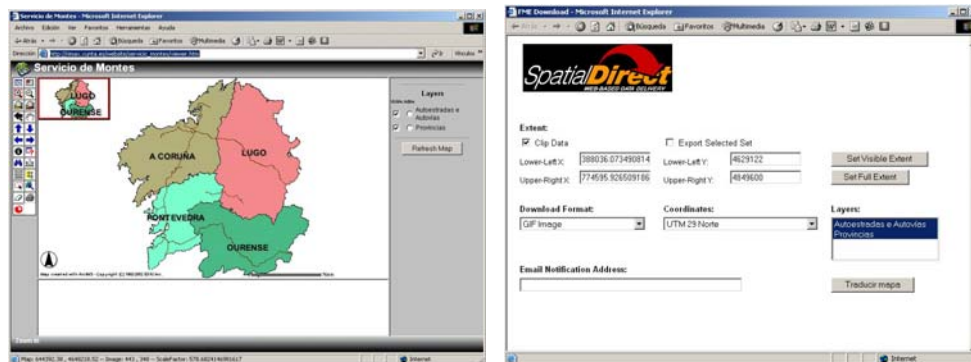


Figura 5: Visualizador web de mapas y formulario de descarga de datos

Conclusiones CMA

La Infraestructura de Datos Espaciales desarrollada para la Consellería de Medio Ambiente de Galicia ha sido diseñada e implementada para resolver los problemas habituales de gestión y uso de información geográfica encontrados en organizaciones y administraciones públicas grandes y descentralizadas, siguiendo las recomendaciones de INSPIRE. Los requisitos han sido satisfechos y el sistema está siendo probado en estos

momentos por un grupo reducido de usuarios que sugerirán posibles mejoras y detectarán potenciales errores.

Seguir las recomendaciones de INSPIRE sobre arquitectura y estándares ha demostrado ser una estrategia adecuada, tanto desde una perspectiva técnica como estratégica:

- Por una parte, la arquitectura de servicios web encadenables ha facilitado la integración de todos los elementos en la infraestructura, a la par que ha preparado el sistema para su futura integración en iniciativas mayores. Por otra, el énfasis hecho en la creación de un catálogo de metadatos como núcleo de la infraestructura se ha mostrado útil tanto al permitir una descripción semántica de los datos mucho más rica (permitiendo así una utilización apropiada) como al proporcionar un componente central alrededor del cual organizar el resto.
- La de la CMA es la primera IDE de Galicia. Haber elegido las recomendaciones de INSPIRE como directrices para su desarrollo la han colocado en una posición estratégica imbatible para convertirse en el núcleo de la IDE de esta Comunidad Autónoma, una vez que estas recomendaciones se conviertan en legislación Comunitaria.

Pese a que actualmente su utilización sea exclusivamente interna, esta IDE ha sido diseñada para permitir su apertura en el momento en que sea necesario. Todos los servicios de mapas y de acceso implementados pueden ser accedidos a través de las interfaces estándar de OpenGIS de servicios web de mapas y de entidades. El catálogo también es conforme a los estándares de OGC, de modo que será capaz de interoperar con otros cuando sea necesario. Se han dado, pues, los pasos técnicos apropiados para hacer de esta IDE el futuro núcleo de la de Galicia, compatible con la legislación de INSPIRE cuando se establezca. Factores políticos y económicos, pero no dificultades técnicas, determinarán ahora el que esto se convierta en una realidad.

Conclusiones

Es bien sabido que la gestión de los datos espaciales es un tema complejo de abordar. Si a esto añadimos las propias dificultades que conlleva cualquier estudio medioambiental, la complejidad aumenta en mayor medida. La situación actual de las distintas políticas medioambientales, en lo que a datos geográficos se refiere, tiene inconsistencias graves que deben ser solucionadas. De estas, las principales son, en primer lugar la propia escasez de datos espaciales susceptibles de tratamiento informático y en segundo la falta de conexión entre los distintos organismos.

La iniciativa de INSPIRE, para el desarrollo de una IDE europea, ha sido pensada para solucionar, justamente, esta clase de problemas, en tanto en cuanto proporcionará las infraestructuras necesarias para una relación más estrecha, no sólo entre los diferentes sectores u organismos que usen datos espaciales, sino también entre diferentes países.

En este documento se ha presentado además un caso real de aplicación de tecnología de infraestructuras de datos espaciales (IDE): la Consellería de Medio Ambiente de la Xunta de Galicia. Lo más relevante de este proyecto es que se ha diseñado un Sistema de Información Geográfica adecuado a las necesidades de esta entidad, pero sobre los estándares y la arquitectura, basada en servicios Web, propuesta por la iniciativa europea INSPIRE. De esta forma se ha cumplido un doble propósito:

por un lado se han proporcionando soluciones SIG a los usuarios de este organismo y por otro se ha diseñado este SIG de forma que sea abierto y fácilmente accesible, pudiendo desde el principio formar parte de las futuras IDE española y europea.

Bibliografía

- (NASA, 2001) The new Digital Earth Reference Model. Ed: John D. Evans, NASA Digital Earth Office. Version 0.4, May 2001. <http://www.digitalearth.gov/derm/v04/>
- (Nogueras, 2004a) "Exploiting disambiguated thesauri for information retrieval in metadata catalogs". J. Nogueras-Iso, J. Lacasta, J. A. Bañares, P. R. Muro-Medrano, F. J. Zarazaga-Soria. Lecture Notes on Artificial Intelligence (LNAI), volumen 3040, páginas 322-333. 2004.
- (Nogueras, 2004b) "Metadata Standard Interoperability: Application in the Geographic Information Domain". J. Nogueras-Iso, F. J. Zarazaga-Soria, J. Lacasta, R. Béjar, P. R. Muro-Medrano. Computers, Environment and Urban Systems, 2004 (In Press)
- (Nogueras, 2005) "OGC Catalog Services: a key element for the development of Spatial Data Infrastructures". J. Nogueras-Iso, F. J. Zarazaga-Soria, R. Béjar, P. J. Álvarez, P. R. Muro-Medrano. Computers and Geosciences. Aceptado para su publicación en marzo de 2005
- (NSDI, 1994) "Executive Order 12906. Coordinating Geographic Data Acquisition and Access the National Spatial Data Infrastructure (U.S.)". U.S. Federal Register. The April 13, 1994, Edition of the Federal Register, Volume 59, Number 71", 1994, pages 17671-17674
- (OGC, 1996) "The OpenGIS Guide. Introduction to Interoperable Geoprocessing. Part I of the Open Geodata Interoperability Specification (OGIS)." K. Buehler and L. McKee(eds.). OGIS Project 6 Technical Committee of the OpenGIS Consortium Inc. OGIS TC Document 96-001. 1996
- (WCS, 2003) "Web Coverage Service, Version 1.0.0", John D. Evans (Eds). Open GIS Consortium Inc. OpenGIS© Project Document: OGC 03-065r6. Agosto 2003
- (WFS, 2002) "Web Feature Server Implementation Specification. Version 1.0.0". P.A. Vretanos(Eds). Open GIS Consortium Inc. OpenGIS project document OGC 02-058. September 2002.
- (WMS, 2002) "Web Map Server Implementation Specification. Version 1.1.1", J. Beaujardiére(Eds). Open GIS Consortium Inc. OpenGIS project document 01-068r3. January 2002

José A. Taboada González • José Manuel Cotos Yáñez
(Editores)

Sistemas de Información Medioambiental



netbiblo