

Generación de metadatos según las Reglas de Implementación de Metadatos de la directiva INSPIRE en el marco del Departament de Medi Ambient i Habitatge (DMAH) de la Generalitat de Catalunya. Algunas consideraciones

Laura Ortiz¹, Alaitz Zabala¹, Pilar Casanovas²

¹Dept. de Geografia, Universitat Autònoma de Barcelona

²Dept. de Medi Ambient i Habitatge, Generalitat de Catalunya

INTRODUCCIÓN

Si en los últimos años el foco de atención principal en las tecnologías de la información recaía directamente sobre los Sistemas de Información Geográfica (SIG), actualmente se asiste a un cambio de paradigma en el que se imponen las Infraestructuras de Datos Espaciales (IDE). Una IDE puede definirse como el conjunto de datos, servicios, metodologías, normas y estándares, que permiten acceder, consultar y analizar la información geográfica publicada en Internet, según estándares bien definidos.

Los metadatos, o "datos sobre los datos", consisten en informaciones que acompañan a cada capa de datos espaciales informando sobre sus características. Los datos espaciales bien documentados permiten un ahorro de tiempo y esfuerzos a la hora de acceder a los mismos, y posibilitan un mejor aprovechamiento al quedar definidos parámetros importantes al emplear la información geográfica, como son el ámbito geográfico, la escala, los procesos realizados sobre la capa, la extensión temporal de los datos, etc.

El aumento del volumen de datos geográficos que se produjo a raíz del auge de los SIG hace unas décadas hace ahora necesario conocer y definir las características de cada base geográfica de manera estandarizada para posibilitar un mejor acceso y un aprovechamiento más eficaz, además de su compilación en IDEs bien organizadas y estructuradas. El acceso a los metadatos por parte de los usuarios se produce mediante el uso de catálogos desde las IDE.

INSPIRE

Dentro de este marco surge INSPIRE, la Directiva 2007/2/CE, de 14 de marzo de 2007, por la que se establece una infraestructura de información espacial en la Comunidad Europea. Esta Directiva es de aplicación para datos espaciales referentes al territorio de la UE, en poder de las Administraciones Públicas de los estados miembros, comprendidos en los temas especificados en sus Anexos I, II y III. Uno de sus elementos clave es la generación de metadatos conformes a unos requerimientos fijados por las reglas de implementación (IR) de la misma Directiva.

La Directiva entró en vigor el 15 de mayo de 2007. Después de diversas revisiones, comentarios por parte de estados miembros, etc. la Regla de Implementación de metadatos (IR) de INSPIRE se aprobó en mayo de 2008 por la Comisión de la UE. La aprobación

definitiva (por parte del Parlamento Europeo) y su publicación se esperan para agosto de 2008.

El claro objetivo de la Directiva es suministrar medidas de aplicación que faciliten la utilización de datos espaciales de diversas fuentes en todos los Estados miembros, apostando por la interoperabilidad de los mismos mediante la aplicación de estándares internacionales, como las normas de metadatos ISO-19115, ISO-19119 y ISO-15836 (Dublín Core).

La norma ISO-19115:2003 Geographic Information - Metadata, y la ISO 19119:2005 Geographic information - Services (y sus adendums) son dos de los estándares en los que se asienta la implementación de la norma INSPIRE. Por un lado, ISO-19115 aporta la terminología necesaria para describir los datos geográficos a nivel de capa, serie o servicio; mientras que ISO-19119 amplía la descripción de los servicios de información geográfica. Además es necesario considerar el estándar ISO-19139:2007 Geographic information -- Metadata -- XML schema implementation que describe la implementación XML de ISO-19115 a nivel de capa. Estos documentos son, por tanto, las "guías técnicas" a seguir cuando se definen los aspectos básicos de los metadatos. Todos los requerimientos exigidos por INSPIRE son conformes a estos estándares, si bien son las IR las que fijan los elementos mínimos que debe contener un archivo de metadatos para ser incluido en la IDE europea.

DMAH

Con anterioridad a la aparición de INSPIRE, Cataluña ya contaba con su propia infraestructura de datos espaciales, la IDEC (Infraestructura de Datos Espaciales de Cataluña). Anteriormente, el Departament de Medi Ambient i Habitatge de la Generalitat de Catalunya (DMAH) ya ha trabajado conjuntamente con IDEC en la generación de metadatos en formato XML conformes con ISO-19139, quedando definido el perfil de metadatos para la infraestructura catalana.

El DMAH a través de su página web viene poniendo a disposición del público, desde 1999, unas 650 bases geográficas acompañadas de sus correspondientes metadatos siguiendo los pre-estándares europeos (CEN) en primer lugar y el estándar ISO 19115 desde antes de su aprobación.

OBJETIVOS

El DMAH ha querido ahora anticiparse a la fecha de aprobación definitiva del reglamento de metadatos de la Directiva INSPIRE, documentando un volumen de 60 capas de información espacial, de entre las ya disponibles en su portal de descarga de cartografía, siguiendo las directrices de la versión borrador de mayo de 2008. Además, algunas de las capas corresponden a series que se documentan atendiendo al modelo de cuatro niveles propuesto por el Dept. de Geografia de la Universitat Autònoma de Barcelona (UAB) y por el CREA: multiserie, serie, hoja y capa (y presentado en las III Jornadas Técnicas de la Infraestructura de Datos Espaciales de España, JIDEE 2006). Con este modelo, se evitan

redundancias en la documentación de las capas que pertenecen a series al definirse herencias de elementos de metadatos comunes para los niveles inferiores.

El trabajo se ha llevado a cabo durante el verano de 2008, y ha dado lugar a esta comunicación, cuyo objetivo es explicar el proceso de implementación del reglamento de metadatos en el marco del DMAH. Se presentan además las implicaciones de esta implementación y se realizan consideraciones sobre la misma: problemas, dudas, "errores" o puntos a mejorar.

CONSIDERACIONES

El proceso de implementación ha consistido, en primer lugar, en la generación de la plantilla de exportación a formato XML (eXtended Mark-up Language) de los metadatos documentados para las capas en formato del software de SIG y Teledetección MiraMon. Para este fin se han seguido las especificaciones de las IR de INSPIRE y las normas ISO citadas anteriormente.

La plantilla de exportación consiste en un documento de base, en XML, que contiene todas las clases y subclases que INSPIRE requiere para la definición de MD y que provienen de ISO. Una vez documentados los datos, en nuestro caso con la aplicación de gestión de metadatos de MiraMon, el GeMM, se exportan a XML resultando un archivo de metadatos que contiene todas las entradas de información exigidas por la Directiva. Cuando se documentan los metadatos de una capa de manera incompleta, el GeMM no permite la generación del documento XML por no ser válido según la plantilla.

La ventaja de documentar los metadatos con esta aplicación de metadatos, es que los metadatos están directamente relacionados con los datos (y no en un catálogo desvinculado de éstos). Esta conexión directa entre datos y metadatos permite que muchos de los metadatos se puedan deducir automáticamente, reduciendo la posibilidad de errores y el esfuerzo de generación de metadatos, sin perder la posibilidad de generar, de manera automática, una exportación XML de metadatos siguiendo los estándares ISO-19115/19193 y INSPIRE IR para alimentar los catálogos de las IDEs (IDEC, IDEE o INSPIRE).

Asimismo y como referencia, además del documento de las IR de INSPIRE y del de implementación XML de ISO (ISO-19139), se ha empleado el editor online de MD que recientemente (junio de 2008) ha puesto a disposición del público general el portal de INSPIRE, accesible desde <http://www.inspire-geoportal.eu/inspireEditor.htm>, que permite exportar a XML los metadatos introducidos manualmente, generándose así un documento XML contra el que validar nuestra plantilla.

Algunos de los problemas encontrados a la hora de crear y validar nuestra plantilla incluyen leves inconsistencias a nivel de formulación de clases y subclases en formato XML en los documentos generados por el editor online (e.g. referencias a code lists introducidos en algunos casos y no en otros), divergencias respecto a las "buenas prácticas" recomendadas en el estándar ISO-19139 (e.g. codificación de los idiomas de los datos y metadatos como texto libre y no como un código de idioma referido a un code list) o la

falta de especificidad en la definición de algunos elementos en las IR, como la conformidad con especificaciones (la Conformity en el texto de las IR).

Los estándares ISO son muy generales y abiertos porque están pensados para que puedan ser útiles en cualquier disciplina relacionada con el territorio en cualquier país, siendo las IR de INSPIRE las que deben concretar qué entradas de metadatos serán las requeridas para los metadatos que formen parte de la infraestructura. Por todo esto, la falta de especificidad en la definición de elementos a incluir por parte del texto de las IR constituye un problema que podría afectar a la homogeneidad deseada de los metadatos, y a su interoperabilidad.

Otra dificultad ha sido adaptar los requerimientos de INSPIRE a la compleja estructura organizativa del DMAH, concretamente a la especificación del organismo relacionado con los datos. Debido a los cambios constantes que se producen (nomenclatura de las unidades administrativas, movilidad de personal, etc.) se ha optado por establecer unos criterios genéricos y válidos para cualquier momento en el tiempo.