"Medidas para impulsar la utilización del Núcleo Español de Metadatos (NEM)"

Daniela Ballari†, Alejandra Sánchez Maganto‡, Javier Nogueras-Iso‡‡, Antonio Rodríguez Pascual‡, Miguel Ángel Bernabé†.

† Laboratorio de Tecnologías de la Información Geográfica (LatinGEO)
Escuela Técnica Superior de Ingenieros en Topografía, Geodesia y Cartografía
Universidad Politécnica de Madrid
Autovía de Valencia Km. 7,5, 28031 Madrid
Tlf: 913311968 Fax: 913311968.
e-mail: daniela@topografia.upm.es; ma.bernabe@upm.es

‡ Instituto Geográfico Nacional C/General Ibáñez Ibero, 3; 28003; Madrid Tlf: 91 5979664 Fax: 91 5979764 e-mail: asmaganto@fomento.es; afrodriguez@fomento.es

‡‡Departamento de Informática e Ingeniería de Sistemas Universidad de Zaragoza C/María de Luna 1, 50018 - Zaragoza Tlf: 976762358 Fax: 976761914. e-mail: jnog@unizar.es

Resumen

El presente artículo se centra en la descripción de las distintas iniciativas y acciones que está realizando la Infraestructura de Datos Espaciales de España (IDEE) para promover la adecuada y generalizada creación de metadatos en los organismos pertenecientes a la Administración General del Estado (AGE), en la Administración Regional y Local. Se describe el trabajo desarrollado para impulsar, crear metodologías y facilitar la creación de metadatos con el objetivo de nutrir el catálogo de metadatos de la IDEE. Se resume el bagaje de experiencia acumulada en el proceso de creación de metadatos aplicando el Núcleo Español de Metadatos (NEM).

Palabras clave: Infraestructuras de Datos Espaciales, IDE, IDEE, Metadatos, Perfiles de aplicación de metadatos, Catálogos, Núcleo Español de Metadatos, NEM.

1 Introducción

Uno de los requisitos principales para el establecimiento de una Infraestructura de Datos Espaciales (IDE) es contar con metadatos que describan la Información Geográfica, siendo este un requisito esencial para localizar, describir y evaluar los datos disponibles [1].

A pesar del reconocimiento general de la importancia de los metadatos, su generación dista de ser una tarea sencilla y atractiva. Diferentes obstáculos se interponen a los creadores de metadatos, algunos de ellos son [2]:

- Los estándares de metadatos son demasiado difíciles y extensos de implementar.
- La producción de metadatos requiere tiempo y otros recursos.
- Hay pocos beneficios tangibles e incentivos para producir metadatos.
- Falta de personal capacitado.
- Dificultad de utilizar las herramientas de generación de metadatos.

Teniendo en cuenta estos obstáculos, el Subgrupo de Metadatos del Grupo de Trabajo para el establecimiento de la IDEE, perteneciente a la Comisión de Geomática del Consejo Superior Geográfico, ha puesto en marcha una serie de actividades tendientes a superar las dificultades antes nombradas.

De esta forma se han promovido diferentes iniciativas para facilitar e impulsar la labor de creación de metadatos. Una de ellas está siendo llevada a cabo por el Instituto Geográfico Nacional en colaboración con la Universidad Politécnica de Madrid y la Universidad de Zaragoza, por medio del proyecto de investigación titulado: "Convenio para el desarrollo complementario de la tecnología y metodología de captura de metadatos de Información Geográfica y del Catálogo de metadatos de la IDEE y del Nodo del Instituto Geográfico Nacional de distribución e intermediación de datos y servicios geográficos"

Algunas de las acciones del citado proyecto de investigación, que serán detalladas en el presente artículo, son:

- Creación de una guía de usuario del perfil "Núcleo Español de Metadatos (NEM)" de la norma ISO 19115.
- Creación de un equipo de expertos para la ayuda y soporte a los distintos organismos implicados en la generación de metadatos.
- Establecimiento de un Plan de Formación del personal de la Administración General del Estado, de la administración Regional y local.

Apoyo al desarrollo y mantenimiento de la herramienta "Open Source"
 CatMDEdit para la generación de metadatos de acuerdo a las normas de metadatos geográficos más utilizadas en el contexto de la información geográfica.

A continuación se detallan las características principales del Núcleo Español de Metadatos, como origen de las iniciativas de implementación, seguidamente se describe cada una de las acciones que se están llevando a cabo para impulsar su utilización. Se finaliza el artículo con una serie de conclusiones obtenidas a partir del trabajo de investigación realizado.

2. El Núcleo Español de Metadatos.

NEM, acrónimo de Núcleo Español de Metadatos, es una recomendación de metadatos aprobada por el Consejo Superior Geográfico, a través de su Comisión de Geomática [3,11].

El NEM es un perfil de la norma internacional ISO19115:2003 [10] que constituye el conjunto de metadatos "mínimo" aconsejable por su utilidad y relevancia que permitirá realizar búsquedas, comparaciones, etc., a partir de metadatos que proceden de diferentes fuentes, sobre distintos conjuntos de datos, de una manera rápida, práctica, fácil y fiable. Se define, para ser utilizado por todos los catálogos generados en las diferentes organizaciones relacionadas con la información geográfica de manera que se consiga la interoperabilidad de metadatos en toda España.

El perfil NEM está formado por elementos de las dos normas de metadatos más importantes que existen hoy en día en esta materia, la Norma Internacional ISO 19115:2003 "Metadata" y la Norma ISO 15836:2003 "The Dublin Core Metadata Element Set", teniendo en cuenta además los proyectos más relevantes en curso en materia de metadatos, especialmente la Directiva Marco del Agua y las iniciativas IDE regionales. Este conjunto de elementos han sido considerados según la justificación de su inclusión (ver Figura 1):

- Elementos pertenecientes al núcleo de la norma ISO19115 (Core Metadata for Geographic Datasets) para la descripción de conjuntos de datos geográficos. Se trata de un conjunto de 22 elementos, 7 de ellos de carácter obligatorio y otros 15 elementos de carácter opcional y condicional.
- Elementos adicionales incluidos por compatibilidad con la norma Dublin Core [10,12], que no tenían equivalente en el subconjunto ISO19115 Core Metadata for Geographic Datasets. Por tanto, se incluyeron elementos adicionales como son: información de agregación, créditos y constricciones sobre el recurso.

- Elementos adicionales para descripción detallada de la calidad de los recursos de información geográfica [4]. Se consideró que el elemento "genealogía" (lineage) incluido dentro del subconjunto ISO19115 Core Metadata for Geographic Datasets era insuficiente para detallar con profundidad la calidad de un recurso y se propuso la inclusión de elementos adicionales del modelo general de ISO19115, relativos a la calidad.
- Propuestas de inclusión de elementos adicionales del modelo general de ISO19115 realizadas por expertos Subgrupo de Trabajo del NEM. Entre estas propuestas se incluyen los siguientes elementos: "palabras clave", la "forma de presentación geoespacial", el "nivel jerárquico" para identificar el tipo del recurso descrito, el "propósito del recurso", y la "descripción de los usos específicos del recurso". Varias de estas propuestas coinciden con las realizadas con otras recomendaciones y guías como la Guía GIS de la Directiva Marco del Agua (una de las primeras directivas en exigir de forma explícita la recopilación de información geográfica en la forma de mapas) [15], los borradores de las reglas de implementación para metadatos de la propuesta de directiva INSPIRE [14], o los perfiles de metadatos propuestos en el contexto del proyecto SDIGER [13,16].

Dublin Core

Calidad

Elementos
Adicionales

Figura 1: Definición del NEM

A continuación se describen las diferentes iniciativas para impulsar la utilización del NEM, la creación de metadatos y, como consecuencia de ello, el progreso y ampliación del catálogo de la IDEE.

3. Medidas para impulsar la utilización del Núcleo Español de Metadatos (NEM)

En la introducción del presente artículo se ha mencionado una serie de obstáculos a los que se enfrentan los generadores de metadatos. En base a ellos se plantearon distintas iniciativas para lograr su superación, tal como se muestra en Tabla 1.

Tabla 1. Tabla de obstáculos e iniciativas lanzadas

Obstáculo	Iniciativa
Los estándares de metadatos son extensos	Núcleo español de Metadatos. (Apartado 2)
Los estándares de metadatos son complicados.	Guía de Usuario del Núcleo Español de Metadatos. (Apartado 3.1)
Falta de personal capacitado	Creación de un grupo de expertos catalogadores de la Información Geográfica y Plan de Formación para el personal de la Administración General del Estado, regional y local (Apartado 3.2)
La producción de metadatos requiere tiempo y otros recursos	Elaboración de una metodología de creación de metadatos basada en Cuestionarios (Apartado 3.3)
Dificultad de utilizar las herramientas de creación de metadatos	Desarrollo y mantenimiento de la aplicación Open Source CatMDEdit (Apartado 3.4)

3.1 Guía de Usuario del Núcleo Español de Metadatos

La implementación de metadatos es una tarea difícil y complicada que requiere cierta especialización y considerable dedicación, pues, además de conocer bien las características técnicas y básicas del recurso a metadatar, hay que saber qué información hay que recoger en cada ítem de metadatos, cómo y con qué criterios. Ha surgido así, dentro del Subgrupo de Metadatos la necesidad de elaborar un documento "Guía de Usuario NEM" [5] que describe para cada uno de los elementos que constituye el NEM, los criterios a seguir para rellenarlo, facilitando así, el trabajo de los responsables de la creación de metadatos en cada organización y haciendo posible la interpretación común del resultado final.

Esta guía se estructura, esencialmente, en los siguientes puntos principalmente:

- Breve descripción de los diferentes niveles de información para los cuales se pueden crear metadatos conforme a la recomendación NEM,
- Tabla con los elementos que componen el NEM, describiendo los criterios a seguir para rellenar cada uno de ellos así como una serie de recomendaciones para los mismos.
- Ejemplos de aplicación de dichos criterios para diferentes productos asociados a la información geográfica. [4]

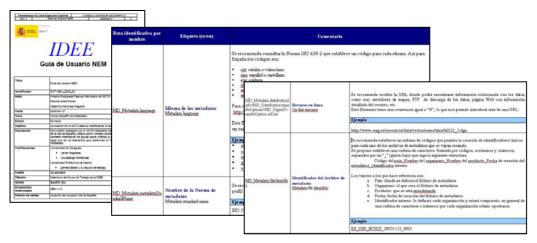


Figura 2. Campos de la Guía de Usuario del NEM

Este documento así como el documento del Núcleo Español de Metadatos se pueden descargar en el Geoportal de la IDEE (www.idee.es).

3.2 Creación de un grupo de expertos catalogadores de la Información Geográfica.

Los productores de datos se encuentran en la actualidad con la necesidad de crear metadatos para que la información geográfica sea localizable a través de catálogos en la Web, pero al mismo tiempo se enfrentan con la escasez de tiempo para generarlos [1], la dificultad propia de creación de los metadatos, y la falta de personal capacitado que sirva de referencia.

Teniendo esto presente, se creó un grupo especialistas para la generación de Metadatos de la Información Geográfica. Este grupo es formador y consultor en materia de Metadatos de la Geoinformación, para dar apoyo y asistencia a los órganos y organismos tanto de la Administración General del Estado como de la Administración Regional y local.

El objetivo de este grupo es ayudar a los productores de datos a dar los primeros pasos en la generación de metadatos, a través de:

- Realización de Cursos de iniciación en Metadatos
- Realización de Cursos de alto nivel de Metadatos
- Asesoramiento a las organizaciones para la implementación del Núcleo Español de Metadatos (NEM) y la Norma ISO 19115.

 Consultorías individualizadas para la definición de perfiles y plantillas de productos y organismos concretos.

El Grupo de Catalogadores creado pertenece al Laboratorio de Tecnologías de la Información Geográfica (LATINGEO) – Universidad Politécnica de Madrid [6] y cuenta con el apoyo de personal del Instituto Geográfico Nacional para la realización de las tareas descritas anteriormente.

3.3 Generación de un Cuestionario de Metadatos

Para superar la barrera de la escasez de tiempo y falta de personal capacitado en materia de metadatos, se ha diseñado una metodología de creación de metadatos basada en un Cuestionario [7]. De esta forma los productores de datos proveen información documental no estructurada, sin necesidad de pasar por los rigores de unos metadatos formales completamente estructurados [1], dejando esta tarea a los Catalogadores de la Información Geográfica.

En el cuestionario se recoge la información necesaria para cumplimentar los Metadatos a partir de preguntas simples, familiares y facilitando ejemplos, que guiarán al productor de datos en el momento de su completado.

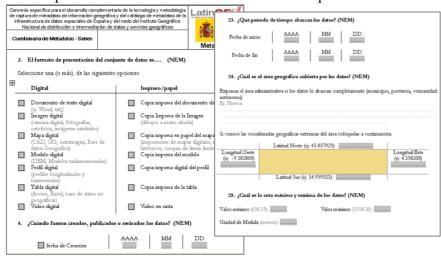


Figura 3. Cuestionario de Metadatos

El cuestionario (ver Figura 3) se encuentra a disposición de todos aquellos que deseen recoger información para generar metadatos de la información geográfica basados en la Norma ISO 19115 y el Núcleo Español de Metadatos, con el objeto de nutrir el catálogo de metadatos de la IDEE.

La finalidad de estos cuestionarios es que el productor de datos, o persona que mejor conozcan las características de los productos, sea capaz de aportar toda la información necesaria para crear los metadatos de sus productos sin tener la necesidad de conocer el perfil NEM, colaborando así en la creación de la documentación descriptiva de sus productos, y en consecuencia en el enriquecimiento de los catálogos existentes en las IDE.

3.4 CatMDEdit: herramienta para la creación de metadatos.

CatMDEdit [8] es una herramienta de edición de metadatos que facilita la documentación de recursos, haciendo especial énfasis en la descripción de los recursos de información geográfica. Es una herramienta Open Source (código abierto) que ha sido desarrollada por el consorcio TeIDE¹ y bajo el apoyo de varias instituciones y proyectos (ver http://catmdedit.sourceforge.net/), destacando entre ellos el apoyo otorgado por el Instituto Geográfico Nacional en su labor de coordinador para la creación de la Infraestructura de Datos Espaciales Española (http://www.idee.es/).

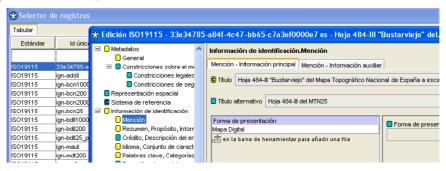


Figura 4. Ventana de visualización y edición de Metadatos

La herramienta, cuyo aspecto se muestra en la Figura 4, ha sido desarrollada íntegramente en Java (su arquitectura se describe en [8]) y presenta las siguientes funcionalidades:

¹ TeIDE es un consorcio constituido por grupos de investigación de la Universidad de Zaragoza (Depto. de Informática e Ingeniería de Sistemas), la Universitat Jaume I (Depto. de Lenguajes y Sistemas Informáticos), y la Universidad Politécnica de Madrid (Depto. de Ingeniería Topográfica y Cartografía), cuyo objetivo es fomentar el la investigación y el desarrollo tecnológico de las Infraestructuras de Datos Espaciales.

- Edición de metadatos de acuerdo a la norma internacional "ISO19115. Geographic Information Metadata". Se proporcionan cuatro interfaces de edición especializadas para esta norma:
 - o Una interfaz detallada de acuerdo al modelo general de ISO19115.
 - Una interfaz reducida de acuerdo al "Núcleo Español de Metadatos" (NEM).
 - Una interfaz de acuerdo al perfil de aplicación de metadatos SDIGER INSPIRE, para la descripción de los recursos espaciales que serán incluidos en la propuesta de la directiva INSPIRE. Este perfil se ha producido dentro del marco del proyecto SDIGER [13,16].
 - Una interfaz de acuerdo al perfil de aplicación de metadatos SDIGER WFD, desarrollado igualmente bajo el marco del proyecto SDIGER. Es un perfil de aplicación de ISO19115 para la descripción de los recursos exigidos por la Directiva Marco del Agua [15].
- Edición de metadatos conformes al perfil de aplicación de metadatos SDIGER Dublin Core para minería de datos geográficos. Tomando como punto de partida el "Dublin Core Spatial Application Profile" definido por el CEN/ISSS Workshop on Metadata for Multimedia Information Dublin Core (WS/MMI-DC), este perfil desarrollado bajo el marco del proyecto SDIGER está orientado a la minería de datos geográficos.
- Intercambio de registros de metadatos de acuerdo a distintos estándares y formatos: ISO19115, Dublin Core y CSDGM (Content Standard for Digital Geospatial Metadata).
- Diferentes estilos de presentación de registros de metadatos en HTML y Excel.
- Herramientas adicionales para facilitar la edición de metadatos: Agenda de contactos, gestión de tesauros y conversión de sistemas de coordenadas.
- Generación automática de metadatos para algunos formatos de transferencia de datos como Shapefile, DGN, ECW, FICC, GeoTiff, GIF/GFW, JPG/JGW, o PNG/PGW.
- Soporte multiplataforma (Windows, Unix), gracias al uso de Java como lenguaje de programación.
- Soporte multilingüe. Actualmente, se puede acceder a la aplicación en 6 idiomas: castellano, inglés, francés, portugués, polaco y checo.

Cabe destacar que tanto esta herramienta, como otras con similares funcionalidades, son herramientas en constante evolución y desarrollo. La importancia de los metadatos para unir los distintos recursos de una IDE, la complejidad de las normas de metadatos y la escasa madurez de estas normas (la norma ISO19115 se aprobó hace tan solo tres años) obligan a una mejora constante

de las herramientas con el objeto de hacerlas más eficaces y transparentes al usuario. El objetivo ideal sería conseguir integrar perfectamente los procesos de creación de datos y metadatos, consiguiendo automatizar lo máximo posible la creación de metadatos.

Así pues, dentro del proyecto de investigación donde se circunscribe este trabajo se ha decido impulsar el desarrollo de esta herramienta en los siguientes aspectos:

- Maximizar la ergonomía de la aplicación con el objeto de facilitar al máximo posible la interacción con el usuario. Los diferentes actores de una IDE comparten una percepción generalizada de que la creación de metadatos es una tarea tediosa y pesada. Por ello, resulta muy importante investigar sobre la usabilidad de las herramientas de edición, para que los usuarios las consideren realmente como mecanismos de ayuda para realizar sus tareas. En base a estudios de usabilidad [9] y normas de accesibilidad se intentará minimizar el tiempo requerido por un usuario para la creación de un registro de metadatos y mejorar la usabilidad de la herramienta.
- Impulsar la implementación de la especificación técnica ISO19139 que describe la serialización sobre XML del modelo de metadatos propuesto por ISO19115. Se espera, que tras varios años de elaboración de borradores, se disponga de una versión definitiva en Septiembre de 2006 que permita la interoperabilidad entre distintas herramientas de catalogación.
- Impulsar la creación de metadatos para la descripción de servicios. Distintas normas e iniciativas han surgido recientemente para la descripción de servicios. Se pretende estudiar y analizar cuales son las más idóneas e integrarlas dentro de la herramienta CatMDEdit.
- Facilitar la descripción de series cartográficas y los componentes de esas series.
- Impulsar avances tecnológicos en aspectos de calidad, multilinguismo y clasificación de información en el ámbito de los metadatos de datos geoespaciales. Se avanzará en aspectos como la estimación de la calidad de metadatos, la recuperación multilingüe y la elaboración y puesta en marcha de Sistemas de Organización de Conocimiento (tesauros, taxonomías y ontologías) para la clasificación de la información.

Conclusiones

Este trabajo ha presentado un conjunto de actividades orientadas a fomentar la creación de metadatos, uno de los pilares fundamentales de una Infraestructura de Datos Espacial, ya sea a nivel local, institucional, regional, nacional o global.

Este conjunto de medidas adoptadas ha tendido a la creación de metadatos que cumpliesen al menos con el Núcleo Español de Metadatos, un perfil cuya primera

versión fue publicada en 2005 y que tras un año de vida dentro de la comunidad de usuarios española se puede afirmar que es un perfil consolidado, concensuado, estable y no restrictivo.

Entre las medidas adoptadas se ha descrito como se han creado una serie de guías que facilitan la descripción de este Núcleo Español de Metadatos. Asimismo, se ha contado con un grupo de Expertos catalogadores capaces de dar formación y soporte en materia de metadatos y que se encuentra a disposición de todos los interesados. Una de las labores de este grupo de Expertos catalogadores ha sido la elaboración de un cuestionario de metadatos para la recogida de información, que facilitará y agilizará la creación de metadatos.

Por último, se ha descrito como se ha apoyado el desarrollo y evolución de la tecnología de captura de metadatos de información geográfica desde el proyecto de investigación descrito en este trabajo, destacando asimismo aquellos aspectos más innovadores donde hay que seguir invirtiendo esfuerzos.

Los catálogos de metadatos, pueden parecer en un principio grandes colosos cuya construcción resulta difícil de abordar, sin embargo creemos que el conjunto de medidas implementadas facilita su creación y supone un avance significativo en apoyo de los creadores de metadatos. Aprovechando la experiencia acumulada y la retroalimentación de los usuarios, esperamos estar en disposición de continuar trabajando en esta línea y ser capaces de proporcionar más medios y herramientas de apoyo que faciliten la labor de los metadatadores.

Referencias

- [1] Global Spatial Data Infraestructure. "El Recetario IDE". Capítulo Tres: Metadatos Describiendo Datos Geoespaciales "Versión 2.0. Enero 2004
- [2] Lynda Wayne. "Institutionalize metadata. Before it institutionalizes you." Federal Geographic Data Committee. Noviembre 2005.
- [3] Consejo Superior Geográfico (Ministerio de Fomento). "SGTNEM_2005_01: Núcleo Español de Metadatos". Infraestructura de Datos Española. 2005. http://www.idee.es/resources/recomendacionesCSG/NEM.pdf
- [4] Curso de IDEs AECI-IGN-UPM 2006. Núcleo Español de Metadatos
- [5] Consejo Superior Geográfico (Ministerio de Fomento). Guía de Usuario del Núcleo Español de Metadatos.Borrador.

http://www.idee.es/resources/recomendacionesCSG/BorradorGuiaNEM.pdf

- [6] Laboratorio de Tecnologías de la Información Geográfica LatinGeo http://www.latingeo.com
- [7] Cuestionario para la Generación de Metadatos. http://mapas.topografía.upm.es/proyectometadatos/cuestionario
- [8] Zarazaga-Soria FJ, Lacasta J, Nogueras-Iso J, Torres MP, Muro-Medrano PR (2003). A Java Tool for Creating ISO/FGDC Geographic Metadata. In Geodaten- und Geodienste-Infrastrukturen von der Forschung zur praktischen Anwendung. Beiträge zu den Münsteraner GI-Tagen 26./27. Juni 2003, pp 17-30. IFGIprints 18, Münster, Germany.
- [9] M. Manrique, M.A. Bernabé, D.Ballari. (2006) Guía de procedimiento de evaluación heurística basado en perspectivas concretas. Informe de Usabilidad para la Herramienta CatMDEdit. (Sin Publicar)
- [10] International Organization for Standardization (ISO). Geographic information Metadata. ISO 19115:2003.
- [11] A. Sanchez Maganto, A. Rodriguez Pascual, P. Abad Power, E. López Romero. El Núcleo Español de Metadatos, perfil mínimo recomendado de metadatos para España.
- [12] International Organization for Standardization (ISO). Information and documentation The Dublin Core metadata element set. ISO 15836:2003.
- [13] M.A.Latre, F.J.Zarazaga-Soria, J.Nogueras-Iso, R. Béjar, P.R.Muro-Medrano. SDIGER: A cross-border inter-administration SDI to support WFD information access for Adour-Garonne and Ebro River Basins, Ver documento. Proceedings of the 11th EC GI & GIS Workshop, ESDI Setting the Framework. Alghero, Sardinia (Italy), 29 June -1 July 2005.
- [14] Commission of the European Communities (CEC), 2004. Proposal for a Directive of the European Parliament and of the Council establishing an infrastructure for spatial information in the Community (INSPIRE). COM(2004) 516 final, 2004/0175 (COD).
- [15] Vogt, J. (ed.). "Guidance Document on Implementing the GIS Elements of the Water Framework Directive", European Communities, 2002.
- [16] F.J. Zarazaga-Soria, J. Nogueras-Iso, M.Á. Latre, A. Rodríguez, E. López, P. Vivas, P.R. Muro-Medrano. Providing SDI Services in a Cross-Border Scenario: the SDIGER Project Use Case. Aceptado para su publicación en *Research and Theory in Advancing Spatial Data Infrastructure Concepts*, ESRI Press (editor: H. Onsrud), 2006.