

El servicio de descarga Inspire basado en ATOM y OpenSearch del Centro Nacional de Información Geográfica

Experiencias y conocimientos adquiridos a través de la implementación de un servicio de descarga de conjuntos de datos predefinidos del Centro Nacional de Información Geográfica, utilizando la aproximación Atom más OpenSearch

LÓPEZ ROMERO, Emilio; GONZÁLEZ GARCÍA, Julián; ABAD POWER, Paloma; PAVO LÓPEZ, Marcos Fco.; GARCÍA DÍAZ, David

Uno de los productos más utilizados, reconocidos y valorados del Centro Nacional de Información Geográfica es el Centro de Descargas de productos cartográficos (<http://centrodedescargas.cnig.es>). A través de este sitio web, el usuario puede buscar, localizar y descargar miles de ficheros de datos de manera gratuita. Hoy en día se trata de una aplicación creada mediante un desarrollo propio y que no está sujeta a normalización. Sin embargo, es necesario que este servicio adopte las recomendaciones dictadas por la guía técnica de implementación de servicios de descarga INSPIRE y, en concreto, las definidas para la utilización del formato ATOM junto con la tecnología OpenSearch para la creación de un servicio de descargas de conjuntos de datos predefinidos. En este artículo, se mostrarán las acciones llevadas a cabo, incluyendo los aspectos técnicos exigidos, la creación y configuración de los ficheros ATOM, y se mostrarán ejemplos de los resultados obtenidos.

PALABRAS CLAVE

Servicio, descarga, INSPIRE, ATOM, OpenSearch, predefinidos.

INTRODUCCIÓN

La Directiva 2007/2/EC del Parlamento Europeo y del Consejo [1], adoptada el 14 de marzo de 2007 tiene como objeto el establecimiento de una Infraestructura para Información Espacial en la Comunidad Europea (INSPIRE) para políticas medioambientales u otras políticas y actividades que tengan impacto en el medio ambiente.

La Directiva INSPIRE [1] establece en su artículo 11 que los Estados miembros establecerán y gestionarán una red con los siguientes servicios, orientados a los conjuntos de datos espaciales y servicios relacionados con ellos para los que se hubieran creado metadatos, de acuerdo con lo dispuesto en la presente Directiva:

- a) servicios de localización que posibiliten la búsqueda de conjuntos de datos espaciales y servicios relacionados con ellos partiendo del contenido de los metadatos correspondientes, y que muestren el contenido de los metadatos;
- b) servicios de visualización que permitan, como mínimo, mostrar, navegar, acercarse o alejarse mediante zoom, moverse o la superposición visual de los conjuntos de datos espaciales, así como mostrar los signos convencionales o cualquier contenido pertinente de metadatos;
- c) **servicios de descarga que permitan descargar copias de conjuntos de datos espaciales, o partes de ellos y, cuando sea posible, acceder directamente a ellos;**
- d) servicios de transformación, que permitan transformar los datos espaciales con vistas a lograr su interoperabilidad;
- e) servicios que permitan el acceso a servicios de datos espaciales.

Para asegurar que las infraestructuras de datos espaciales de los Estados miembros son compatibles y

útiles en un contexto transfronterizo y comunitario, la Directiva precisa la adopción de unas reglas de implementación comunes en diversas áreas, entre las que se incluye el Reglamento por el que se ejecuta la Directiva INSPIRE en lo que se refiere a los servicios de red [2].

En ese reglamento, se determinan los requisitos para el establecimiento de los servicios de red previstos en la Directiva INSPIRE, así como las obligaciones relacionadas con la disponibilidad de esos servicios por parte de las autoridades públicas de los Estados miembros y terceros. En concreto, se determinan para cada uno de los servicios de red, los requisitos de calidad, los plazos para la provisión de los servicios y las operaciones que debe soportar.

OPERACIONES DE LOS SERVICIOS DE DESCARGA

Los servicios de descarga proporcionarán como mínimo las siguientes operaciones:

Obtener metadatos del servicio de descarga (*Get Download Service Metadata*)

Proporciona toda la información necesaria sobre el servicio y los conjuntos de datos espaciales disponibles, y describe las capacidades del servicio.

Obtener conjunto de datos espaciales (*Get Spatial Data Set*)

Permite recuperar un conjunto de datos espaciales.

Describir un conjunto de datos espaciales (*Describe Spatial Data Set*)

Devuelve la descripción de todos los tipos de objeto espacial contenidos en el conjunto de datos espaciales.

Conectar con servicio de descarga (*Link Download Service*)

Permite a una autoridad pública o a un tercero dar a conocer la disponibilidad de un servicio de descarga para la descarga de conjuntos de datos espaciales o, cuando sea practicable, de objetos espaciales, a través del servicio de descarga del Estado miembro, manteniendo la capacidad de descarga en la ubicación de la autoridad pública o del tercero.

PLAZOS DE LOS SERVICIOS DE DESCARGA

A más tardar el 28 de junio de 2012, los Estados miembros proporcionarán los servicios de descarga con una capacidad operativa inicial, esto es, ofrecerán toda la funcionalidad sin garantizar la calidad del servicio o sin garantizar el acceso al servicio a todos los usuarios a través del geoportal INSPIRE.

A más tardar el 28 de diciembre de 2012, los Estados miembros proporcionarán los servicios de descarga tal como dispone el Reglamento por el que se ejecuta la Directiva INSPIRE en lo que se refiere a los servicios de red.

REQUISITOS DE CALIDAD DE LOS SERVICIOS DE DESCARGA

Se aplican los siguientes criterios de calidad de servicio en relación con el rendimiento, la capacidad y la disponibilidad para los servicios de descarga de conjuntos de datos predefinidos:

Rendimiento

El tiempo para enviar la respuesta inicial para la operación <<obtener metadatos del servicio de descarga>> (*Get Download Service Metadata*) será de 10 segundos como máximo en una situación normal.

El tiempo para enviar la respuesta inicial para la operación <<obtener conjunto de datos espaciales>> (*Get Spatial Data Set*) será de 30 segundos como máximo en una situación normal y, a continuación, aún en una situación normal, el servicio de descarga mantendrá una respuesta sostenida superior a 0,5 Megabytes por segundo o superior a 500 objetos espaciales por segundo.

El tiempo para enviar la respuesta inicial para la operación <<describir conjunto de datos

espaciales>> (*Describe Spatial Data Set*) será de 10 segundos como máximo en una situación normal y, a continuación, aún en una situación normal, el servicio de descarga mantendrá una respuesta sostenida superior a 0,5 Megabytes por segundo o superior a 500 descripciones de objetos espaciales por segundo.

Capacidad

El número máximo de peticiones simultáneas a un servicio de descarga que deben atenderse en conformidad con los criterios de calidad del servicio relativos al rendimiento será de 10 por segundo. El número de peticiones procesadas en paralelo podrá limitarse a 50.

Disponibilidad

La probabilidad de que un servicio de red esté disponible será el 99% del tiempo total.

LA GUÍA TÉCNICA PARA LA IMPLEMENTACIÓN DE UN SERVICIO DE DESCARGA INSPIRE

Además de las reglas de implementación, la Comisión Europea llevo a cabo la creación del *IOC Task Force* compuesto por representantes de los Estados miembros, responsables del diseño del diseño de la arquitectura y de la implementación de los servicios de las infraestructuras de datos espaciales nacionales. El propósito de este equipo es ayudar y proporcionar la implementación de INSPIRE en los Estados miembros.

Fruto del trabajo del *IOC Task Force* se han desarrollado una serie de guías técnicas para la implementación de servicios de red INSPIRE. Estas guías técnicas, no vinculantes, describen los detallan los aspectos de la implementación y las relaciones con las prácticas, tecnologías y estándares existentes.

Una de las guías técnicas publicadas trata de la implementación de servicios de descarga INSPIRE. Atendiendo a las operaciones que se implementan, se definen dos tipos de servicios de descarga; los que proporcionan los requisitos mínimos y los que ofrecen todos los requisitos funcionales. Estos últimos deberán implementarse cuando sea factible. Los dos tipos se conocen como:

- a) Servicios de descarga de conjuntos de datos predefinidos. Proporcionan la descarga de conjuntos de datos predefinidos (o partes de un conjunto de datos predefinido), sin permitir consultas o la selección de subconjuntos definidos por el usuario. Un conjunto de datos predefinido o un subconjunto de datos predefinido puede ser, por ejemplo, un fichero almacenado en un repositorio, que puede descargarse como una unidad completa sin la posibilidad de cambiar el contenido, la codificación, el sistema de referencia, etc.
- b) Servicio de descarga de acceso directo. Este servicio extiende la funcionalidad del servicio de descarga de conjuntos de datos predefinido permitiendo consultar y descargar subconjuntos de los conjuntos de datos.

Este artículo hace referencia únicamente al primer tipo de servicios de descarga, esto es, aquellos que simplemente permiten la descarga de conjuntos de datos predefinidos. Además de lo expuesto anteriormente, los conjuntos de datos predefinidos se caracterizan por dos condiciones:

- a) Tienen un metadato y pueden ser localizados usando un servicio de descubrimiento INSPIRE.
- b) El metadato contiene un enlace (URL) al conjunto de datos o parte de él que puede ser descargado inmediatamente mediante una petición HTTP GET.

La guía técnica propone dos posibles alternativas para implementar servicios de descarga de conjuntos de datos predefinidos:

- a) Utilizando el formato de sindicación ATOM [4]
- b) Utilizando las normas ISO 19142 *Web Feature Service* [5] y ISO 19143 *Filter Encoding Specification* [6].

Este artículo sólo hará referencia al primero de los métodos, dado que ha sido el elegido en la implementación del servicio de descarga de conjuntos de datos predefinidos del CNIG.

IMPLEMENTACIÓN CON ATOM DE UN SERVICIO DE DESCARGA DE CONJUNTOS DE DATOS PREDEFINIDOS

El formato de sindicación Atom [4] proporciona un mecanismo fácil y muy conocido para la publicación de información en la web en la forma de fuentes (*feed*), de tal manera que es compatible con las arquitecturas web existentes y con muchas de las herramientas de desarrollo disponibles.

Atom es un documento basado en el formato XML que describe listas de información conocidas como fuentes. Esas fuentes están compuestas de un conjunto de elementos llamados entradas (*entry*) y cada una de esas entradas tiene un conjunto de elementos que contiene información sobre ella. Por ejemplo, cada entrada tiene un título. Además, las entradas pueden contener fuentes adicionales.

Por otro lado, en la guía técnica se propone la utilización del estándar *Open Search* [7] para proporcionar las operaciones de manera convencional y documentarlas siguiendo los requisitos del Reglamento sobre servicios de red INSPIRE. *Open Search* permite especificar, mediante un documento XML, las operaciones y sus parámetros de forma estructurada e interoperable. El mecanismo OpenSearch es reconocido por los navegadores más importantes como Mozilla Firefox, Internet Explorer, Safari y Chrome.

El documento de descripción *Open Search* se enlaza desde la “fuente del servicio de descarga”. Además del documento de descripción, es necesario implementar un servicio simple para satisfacer las operaciones <<obtener conjunto de datos espaciales>> (*Get Spatial Data Set*) y <<describir conjunto de datos espaciales>> (*Describe Spatial Data Set*).

La guía técnica recomienda la utilización de las fuentes Atom para poner a disposición conjuntos de datos predefinidos siguiendo las siguientes premisas:

- Se publica una única fuente Atom al nivel más alto como “fuente del servicio de descarga”.
- Esta fuente contiene un enlace a un documento de descripción de OpenSearch [7] que proporciona los metadatos de las operaciones para el servicio de descarga. El documento de descripción de OpenSearch proporciona información sobre las operaciones implementadas por el servicio de descarga.
- Esta fuente contiene una o más entradas Atom: una por cada conjunto de datos predefinido.
- Cada una de esas entradas Atom debe contener un enlace a otra fuente Atom (una “fuente de un conjunto de datos”) que describe un determinado conjunto de datos predefinido.
- Cada una de las “fuentes de conjuntos de datos” debe contener entradas Atom con enlaces para la descarga de conjuntos predefinidos en diferentes formatos (por ejemplo, en GML, ShapeFile, etc.) y en diferentes sistemas de referencia de coordenadas (CRS). Se debe proporcionar un enlace para cada combinación formato /CRS.
- Se pueden proporcionar fuentes en múltiples idiomas.

En la siguiente figura se muestra la estructura de las fuentes Atom para un servicio de descarga de conjuntos de datos predefinidos.

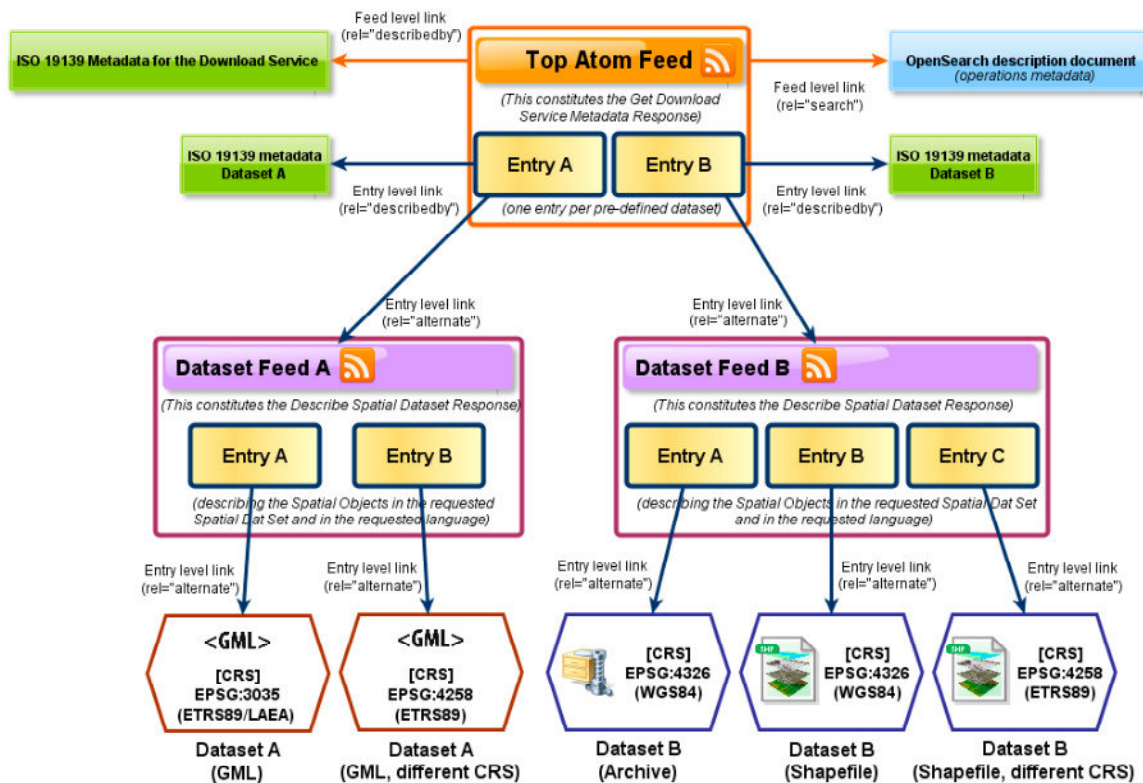


Figura 1: Esquema jerárquico de la estructura de las fuentes Atom.

PASOS PREVIOS A LA GENERACIÓN DE LAS FUENTES ATOM EN EL CNIG

En la actualidad el CNIG dispone de un centro de descargas desde el cual se da acceso a los productos del IGN. Habitualmente estos productos no pueden ser descargados en un solo fichero sino que se distribuyen por varios ficheros, debido a limitaciones en el tamaño máximo permitido por el formato del fichero, o por ser más conveniente para su distribución y posterior uso. La información se distribuye por los ficheros espacialmente atendiendo a dos criterios principalmente:

- División por provincias.
- División por hojas, de las escalas 1/25.000 o 1/50.000.

Después de analizar la guía técnica para servicios de descarga INSPIRE, se decidió implementar el servicio de descarga utilizando ficheros ATOM. Por varias razones:

- En el caso de que ya existan conjuntos de datos predefinidos para descargar, tal y como sucede en el centro de descargas del CNIG, parece la alternativa más fácil de implementar.
- No necesita de software específico, con un simple editor de textos pueden crearse todas las fuentes. En el caso de grandes volúmenes de ficheros será necesario programar la generación de la parte de las fuentes de los conjuntos de datos donde se especifican los vínculos de las descargas a partir de la información guardada en una base de datos o similar.
- La gran cantidad de ejemplos que se dan en la guía.
- La existencia de un servicio de ejemplo en la siguiente dirección: <http://inspire-geoportal.ec.europa.eu/demos/ccm/>.

Como primera tarea se identificaron como conjunto de datos los productos o familias de productos

disponibles en el centro de descargas del CNIG. De todos los productos disponibles se seleccionó solo una parte atendiendo a los siguientes criterios:

- Que contengan datos vectoriales. En el caso de que exista la misma serie o familia en formato vectorial y ráster sólo se dará acceso al vectorial.
- Que contengan datos actuales. En el caso de que exista la misma serie o familia actualizada y como cartografía histórica sólo se dará acceso a la serie actualizada.

La lista definitiva de conjuntos de datos disponibles a los que se dará acceso desde el servicio de descarga será:

- PNOA MÁXIMA ACTUALIDAD
- PNOA MÁXIMA RESOLUCIÓN
- CORINE LAND COVER
- BCN25/BTN25
- MDT05/MDT05-LIDAR
- MDT25
- MDT200
- CARTOCIUDAD
- BCN200
- SIOSE
- BCN500
- BTN100

Además se consideró que sería necesario incluir el Equipamiento Geográfico de Referencia Nacional (EGRN) en el servicio de descarga. Al igual que en el caso de los productos del centro de descargas, se seleccionó una parte del EGRN atendiendo a los siguientes criterios:

- Que contengan información en formato que permita la representación espacial directa en sistemas de información geográfica (ShapeFile, Geodatabase, etc.)
- Que contengan datos vectoriales. En el caso de que exista la misma serie o familia en formato vectorial y ráster sólo se dará acceso al vectorial.
- Que no sean provisionales.

La lista definitiva de ficheros del EGRN a los que se dará acceso desde el servicio de descarga serán:

- CUADRICULAS CARTOGRÁFICAS MTN25 Y MTN50
- LÍNEAS LIMITE MUNICIPALES
- NOMENCLÁTOR GEOGRÁFICO CONCISO DE ESPAÑA

Antes de comenzar con la generación de las fuentes fue necesario crear los metadatos del servicio conforme a la norma ISO 19139 ya que hay una referencia a los mismos dentro de la fuente del servicio, el documento de descripción OpenSearch del motor de búsqueda y el motor de búsqueda.

ESTRUCTURA DE LA FUENTE DEL SERVICIO DE DESCARGA

A continuación se explica de forma detallada la estructura de la fuente del servicio y como se dotó de contenido a cada elemento.

La fuente del servicio contiene los elementos, “*title*”, y “*subtitle*”, que dan una breve descripción del servicio. Para generar su contenido se utilizó el contenido de los siguientes elementos del metadato del servicio:

- `//gmd:identificationInfo/srv:SV_Serviceldentification/gmd:citation/gmd:CI_Citation/gmd:"title"/gco:CharacterString`
- `//gmd:identificationInfo/srv:SV_Serviceldentification/gmd:abstract/gco:CharacterString`

Los atributos “*href*” de los vínculos (“*link*”) hacia los metadatos, hacia la propia fuente del servicio y hacia la descripción OpenSearch del motor de búsqueda, se fijaron teniendo en cuenta la petición al catálogo de metadatos del IGN donde está publicado el metadato del servicio, y las respectivas URLs de la fuente del servicio y del documento OpenSearch. El resto de los valores de los atributos se indican en la guía en función del recurso al que apunta cada vínculo.

Como identificador (contenido del elemento “*id*”) se utilizó la URL de la fuente, ya que se cumple que el identificador de la fuente debe ser único al ser la URL única, tal y como indica la guía técnica.

Los derechos de propiedad, restricciones de uso, etc. del servicio se indicaron como contenido del elemento “*rights*” y coinciden con los que aparecen en el resto de servicios publicados por el IGN/CNIG.

La marca de tiempo de la fuente presente en el contenido del elemento “*updated*” se generó en el momento de creación del fichero.

La información referente al autor de la fuente se codifica en el contenido del elemento “*author*” y suele ser suficiente con informar sobre el nombre de la organización y una dirección de correo electrónico.

ENTRADAS EN LA FUENTE DEL SERVICIO DE DESCARGA

En la fuente del servicio aparecen tantas entradas como conjuntos de datos estén disponibles para la descarga. Cada entrada contendrá a su vez otros elementos hijos:

- “*title*” (título del conjunto de datos)
- “*inspire_dls:spatial_dataset_identifier_code*” (código que identifica el conjunto de datos dentro de un cierto espacio de nombres)
- “*inspire_dls:spatial_dataset_identifier_namespace*” (espacio de nombres que identifica un conjunto de elementos de metadatos)

Se utilizaron los contenidos de los siguientes elementos del metadato del servicio:

- `//gmd:MD_Metadata/gmd:identificationInfo/gmd:MD_DataIdentification/gmd:citation/gmd:CI_Citation/gmd:"title"/gco:CharacterString`
- `//gmd:MD_Metadata/gmd:identificationInfo/gmd:MD_DataIdentification/gmd:citation/gmd:CI_Citation/gmd:identifier/gmd:RS_Identifier/gmd:codeSpace/gco:CharacterString`
- `//gmd:MD_Metadata/gmd:identificationInfo/gmd:MD_DataIdentification/gmd:citation/gmd:CI_Citation/gmd:identifier/gmd:RS_Identifier/gmd:codeSpace/gco:CharacterString`

Cada entrada debe tener un identificador único “*id*”. Para facilitar la gestión de estos identificadores se decidió utilizar como identificador la URL de la fuente del conjunto de datos al



Figura 3: Visualización de la fuente Atom en el navegador Mozilla Firefox.

Internet Explorer presenta la fuente aplicando una hoja de estilo y generando un documento en el que aparecen los contenidos de los elementos *“title”* de la fuente y, por cada entrada, el contenido de los elementos *“title”*, *“updated”* y *“summary”* de manera secuencial. Sin embargo, no se insertan hipervínculos en los contenidos de los elementos *“title”* de cada entrada hacia las fuentes de los conjuntos de datos, imposibilitando la navegación. Además, ofrece la posibilidad de filtrar por categorías, palabras clave, etc. y ordenar por título o por fecha.



Figura 3: Visualización de la fuente Atom en el navegador Internet Explorer.

ESTRUCTURA DE LA FUENTE DE LOS CONJUNTOS DE DATOS

La información sobre la ubicación de los ficheros y los formatos en los que están disponibles cada conjunto de datos se expresa en las fuentes de los conjuntos de datos.

La generación de los contenidos de los elementos *“title”*, *“id”*, *“rights”*, *“updated”*, y *“author”* y del vínculo (*“link”*) a la propia fuente se realizan de una manera similar a como se crearon en la fuente del servicio.

Para informar sobre los tipos de objetos espaciales contenidos en el conjunto de datos se utilizan vínculos *“link”* que hacen referencia a temas o tipos de objetos existentes en el diccionario de fenómenos disponible en el registro de INSPIRE (<https://inspire-registry.jrc.ec.europa.eu/registers/FCD/items>), aunque se pueden sustituir por un vínculo a un catálogo de fenómenos local.

ENTRADAS EN LAS FUENTES DE LOS CONJUNTOS DE DATOS

En el caso de las fuentes de los conjuntos de datos se generó una entrada por cada una de las combinaciones de formatos de fichero y sistemas de referencia disponibles.

Las entradas en este fichero, así como el contenido y los elementos hijos se generaron utilizando la información de la base de datos con la que se gestiona el centro de descargas del CNIG, explotándola mediante una aplicación Java y consultas SQL. En dicha base de datos, para cada conjunto de datos se especifican los ficheros en los que está repartido y para estos se almacena el sistema de referencia y el formato del fichero.

Cada entrada se genera utilizando consultas resumen por cada conjunto de datos, sistema de referencia y formato de ficheros. Y para cada una de ellas se especifica:

- Un identificador “*id*” que se generó concatenando a la localización de los ficheros en el centro de descargas del CNIG, el código EPSG y el formato de los ficheros a descargar.
- El sistema de referencia mediante el elemento “*category*”, al igual que se hace en la fuente del servicio.
- La manera en la que se distribuye el conjunto de datos en el caso de que haya varios ficheros por conjunto de datos en el elemento “*content*”.
- Los vínculos “*link*” de descarga de los ficheros.
- Un resumen en formato HTML con información relevante del conjunto de datos, así como vínculos HTML a los ficheros de manera que los ficheros del conjunto de datos estén disponibles.

REFERENCIAS

- [1] INSPIRE, Directiva 2007/2/EC del Parlamento Europeo y del Consejo, de 14 de marzo de 2007, por la que se establece una Infraestructura de Información Espacial en la Comunidad Europea
- [2] Reglamento sobre servicios de red INSPIRE, número 976/2009, de 23 de noviembre de 2010, modificada por el Reglamento, número 1088/2010 relativa a los servicios de descarga y los servicios de transformación.
- [3] Guía Técnica para la implementación de servicios de descarga INSPIRE del IOC Task Force
- [4] IETF RFC 4287 The Atom Syndication Format
- [5] ISO 19142:2010 Información geográfica - Servicio web de fenómenos
- [6] ISO 19143:2010 Información geográfica - Codificación de filtros
- [7] formato de documento de descripción OpenSearch, <http://www.opensearch.org/Specifications/OpenSearch/1.1>

AUTORES

Emilio LÓPEZ ROMERO
elromero@fomento.es
CNIG

Julián GONZÁLEZ GARCÍA
jgonzalezg@fomento.es
CNIG

Paloma ABAD POWER
paabd@fomento.es
CNIG

Marcos Fco. PAVO LÓPEZ
marcosf.pavo@cnig.es

David GARCÍA DÍAZ
tahbit.dg.externo@cnig.es

CNIG

TAHBIT