

Nomenclátor Geográfico Conciso de España conforme a INSPIRE

Antonio Villena¹, Antonio F. Rodríguez¹, Emilio López¹, Paloma Abad¹, Alejandra Sánchez¹, Marta Juanatey¹, Carolina Soteres¹, Cristina Ruiz¹, Lorena Hernández¹, Inmaculada Serra¹

¹Centro Nacional de Información Geográfica (Instituto Geográfico Nacional)

antonio.villena@cnig.es, afrodriguez@fomento.es, elromero@fomento.es, pabad@fomento.es, asmaganto@fomento.es, mjuanatey@fomento.es, csoteres@fomento.es, cruiz@fomento.es, lhquiros@fomento.es, inmaculada.serra@cnig.es

Resumen

Los nombres geográficos son un elemento clave en el campo de la geografía, ya que constituyen la base para los sistemas de referencia basados en identificadores geográficos, muy importantes para la búsqueda e identificación de información geográfica.

Entre las bases de datos geográficos de que dispone el IGN, se encuentra el Nomenclátor Geográfico Conciso de España, que contiene 3.667 topónimos correspondientes a una escala 1/1.000.000.

La Directiva INSPIRE, a través de la Norma de Ejecución *Reglamento (UE) nº 1089/2010 en lo que se refiere a la interoperabilidad de los conjuntos y los servicios de datos espaciales*, establece los criterios generales que debe cumplir todo conjunto de datos perteneciente a cada uno de los temas de los anexos.

A fin de garantizar la interoperabilidad y la armonización entre los distintos temas de datos espaciales, se han desarrollado documentos técnicos (Especificaciones de datos), para cada uno de los temas de los anexos de INSPIRE, que desarrollan la parte legal definida en la Norma de Ejecución.

En esta comunicación se describen las principales características de la Especificación de datos sobre Nombres Geográficos y se describe el proceso que está llevando a cabo el IGN para la adaptación de la base de datos del Nomenclátor Geográfico Conciso de España v1.0 (NGCE v1.0) al modelo INSPIRE en el tema de Nombres Geográficos (*Geographical names*) del Anexo I de la Directiva.

Palabras clave: nomenclátor geográfico conciso, *geographical names*, INSPIRE, armonización, interoperabilidad, modelo de datos.

1 Introducción

El reglamento que se refiere a la interoperabilidad de los conjuntos y los servicios de datos espaciales, es el Reglamento (UE) nº 1089/2010 y establece los requisitos sobre las disposiciones técnicas relativas a la interoperabilidad y, cuando sea practicable, la armonización de los conjuntos de datos espaciales y los servicios de datos espaciales de los temas de los anexos I, II y III de la Directiva INSPIRE.

Los documentos técnicos (*Guidance Documents*) describen los requisitos comunes: identificación, modelo de datos, metadatos, representación, calidad, etc. relativos a la interoperabilidad y otros conceptos que son aplicables a cada uno de los temas de datos espaciales. En concreto, para el tema de Nombres Geográficos se ha elaborado el documento *INSPIRE Data Specification on Geographical Names - Guidelines v 3.0.1*¹, aprobado en 2010 y que establece los requisitos a cumplir para nombres geográficos.

El NGCE v1.0 contiene 3.667 topónimos almacenados en una base de datos Access creada en 2006. Está basado en la cartografía oficial de España a escala 1/1.000.000 del IGN; este nomenclátor cumple el Modelo de Nomenclátor de España. El NGCE v1.0 no contempla los nombres de vías de comunicación, ya que tradicionalmente nunca se han incluido en los nomenclátors. Los atributos de los topónimos, acordes con el MNE,

1

http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_GN_v3.0.1.pdf

han sido confrontados con los atributos del esquema de aplicación de INSPIRE según la especificación de datos de Nombres geográficos (*Geographical names*) pertenecientes al Anexo I.

2 Atributos del NGCE

Los atributos de los que consta la base de datos del Nomenclátor Geográfico Conciso de España son:

- *IdEntidad*, identificador de la entidad geográfica.
- *Nombre Preferente*, nombre utilizado como identificador de cada entidad geográfica según el organismo.
- *Idioma Nombre Preferente*, lengua en que está expresado el nombre de cada elemento geográfico.
- *Variante e idioma variante*, nombre menos conocido o de uso restringido y la lengua en a que está.
- *Estatus*, indica si es Oficial o Normalizado.
- *Tipo Entidad*, nombre del tipo del topónimo, según la clasificación jerárquica del Catálogo de Entidades.[7]
- *Sistema de Referencia* así como sus coordenadas expresadas en Latitud y Longitud.
- *Municipio, provincia y comunidad autónoma*, nombre oficial de la unidad administrativa donde se encuentra el punto representativo elegido de la entidad geográfica.
- *Geometría*, se indica si la entidad es un punto, una línea o una superficie.
- *Mapa*, número de la hoja del Mapa Topográfico Nacional de escala 1/50.000 del Instituto Geográfico Nacional en la que está el punto representativo.
- *Fuente Nombre y de la Variante*, organismo o documento del cual se ha tomado la denominación introducida en el campo «Nombre».
- *Antes e idioma antes*, nombre oficial anterior al vigente en la fecha de finalización de la versión 1.0 del NGCE (agosto 2006) y su lengua.
- *Catálogo de Entidades*, indica la URL de la tabla de clasificación de las entidades geográficas.[7]

El «Nombre Preferente» es el nombre o topónimo de cada entidad geográfica. «Variante» es el nombre menos conocido o de uso restringido, llegando a tener varios campos en función del idioma, etc. «Antes» es el nombre oficial anterior al vigente, es decir, el nombre histórico de la

entidad geográfica. Estos atributos se deberán incluir en el mismo tipo de entidad. Por tanto, una misma entidad geográfica alberga tantos textos como posea, diferenciándose el estatus (oficial o histórico), el idioma, fuente, etc.

3 Atributos y entidades INSPIRE

En las especificaciones de datos «[Nombres Geográficos](#)» del modelo INSPIRE tenemos como clase principal «*NamedPlace*» junto con otras clases llamadas: «*GeographicalName*», «*SpellingOfName*» y «*PronunciationOfName*».

El elemento central del modelo de datos es la clase «***NamedPlace***», que representa cualquier objeto geográfico del mundo real al que se refieren uno o más nombres. Tiene tres atributos obligatorios, *inspireId*, *name* y *geometry* (más adelante se explica cada atributo).

Cada fenómeno de nombre geográfico (*NamedPlace*) puede estar asociado con uno o más nombres, modelados con el tipo de datos de nombre geográfico (*GeographicalName*), si tiene diferentes lenguas, diferentes formas de ortografía, diferentes alfabetos, etc. Además cada tipo de datos de nombre geográfico (*GeographicalName*) debe tener un texto (*SpellingOfName*) obligatoriamente y si lo tiene, debe introducirse la pronunciación (*PronunciationOfName*), al ser «*voidable*».

En la Figura 1 se muestra el diagrama de clases UML de esta especificación.

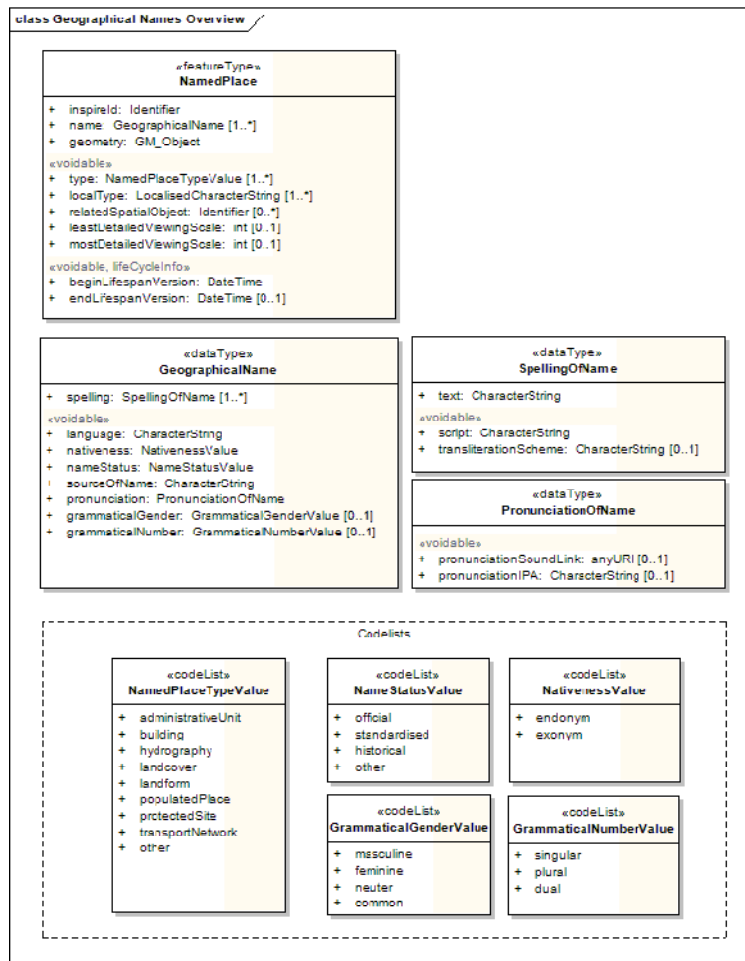


Figura 1. Diagrama de clases UML: vista general del esquema de aplicación de Nombres geográficos.

3.1 NamedPlace, representación del mundo real al que se refieren uno o varios nombres propios, descrito por los siguientes atributos:

- *inspireId*, identificador del objeto espacial.
- *name*, nombre o nombres de la entidad que se está describiendo. El *name* es del tipo *GeographicalName*.
- *geometry*, descripción geométrica.
- *type*, tipo de entidad representada a partir de una lista controlada.

- *localType*, tipo proporcionado por cada organización.
- *relatedSpatialObject*, identificadores de los objetos espaciales que representan la misma entidad que el *NamedPlace* pero que pertenecen a otros temas de INSPIRE.
- *LeastDetailedViewingScale* y *MostDetailedViewingScale*, representan el nivel de escala al cuál están representados .
- *beginLifespanVersion* y *endLifespanVersion*, es el ciclo de vida.

3.2 GeographicalName, nombre propio del *NamedPlace*, descrito por los siguientes atributos:

- *spelling*, representa las distintas formas en las que se puede escribir el *GeographicalName*.
- *language*, idioma del *GeographicalName*.
- *Nativeness*, posibles valores: endónimo o exónimo.
- *nameStatus*, posibles valores: oficial, estandarizado, histórico u otro.
- *sourceOfName*, representa el origen del que el nombre geográfico esta tomado.
- *pronunciation*, forma correcta o estándar de pronunciación del *GeographicalName*.
- *grammaticalGender*, posibles valores: masculino, femenino, neutro o común.
- *grammaticalNumber* posibles valores: singular, plural o dual.

3.3 SpellingOfName, representa la forma correcta de escribir un *GeographicalName*, descrito por los siguientes atributos:

- *text*, ortografía textual propiamente dicha.
- *script*, alfabeto en la que está procesada la ortografía.
- *transliterationScheme*, define el método usado para la conversión de la ortografía de una escritura a otra.

3.4 PronunciationOfName, pronunciación en el Alfabeto Fonético Internacional (IPA), descrito por los siguientes atributos:

- *pronunciationIPA*, expresa la pronunciación en el Alfabeto Fonético Internacional.
- *pronunciationSoundLink*, expresa la pronunciación mediante un link a un fichero de sonido.

3.5 Lista controlada **NamedPlaceTypeValue**, lista de valores que expresan a qué tipo de elemento geográfico se refiere el *NamePlace*:

- *administrativeUnit*, (unidades administrativas).
- *building* (edificios),
- *hydrography* (hidrografía),
- *landcover* (cubierta),
- *landform* (relieve), características geomorfológicas del terreno.
- *populatedPlace* (lugares habitados),
- *protectedSite* (sitio protegido),
- *transportNetwork* (red de transporte),
- *other* (otros),

3.6 Lista controlada **NameStatusValue**, estado de un nombre geográfico. Es la información disponible para discernir qué crédito debe darse al nombre con respecto a su normalización o a su temática. Valores que puede tomar:

- *historical* (histórico),
- *official* (oficial),
- *other* (otro),
- *standardized* (normalizado),

3.7 Lista controlada **NativenessValue**, origen de un nombre geográfico. Valores que puede tomar:

- *endonym* (endónimo),
- *exonym* (exónimo),

3.8 Identifier (identificador), identificador único de objeto publicado por el organismo competente, el cual debe ser usado por aplicaciones externas que referencien el objeto espacial. (Ver DS-D2.5 [8]).

4 Proceso de correspondencia

La adaptación de la base de datos del Nomenclátor Geográfico Conciso de España v1.0 (NGCE v1.0) al modelo de datos establecido por INSPIRE se ha realizado utilizando los programas: MS Office Access versión 2007 y el ETL (*Extract Transform Load*) de [FME](#) versión 2011. El proceso seguido ha consistido en relacionar los elementos del modelo de datos del NGCE con el del modelo INSPIRE para Nombres Geográficos. Circunstancias encontradas durante el cotejo de ambos modelos debido a la discrepancia entre sus elementos:

- No se han podido relacionar con el modelo INSPIRE algunos atributos de la Base de datos del NGCE como Municipio, Provincia, Comunidad Autónoma, Mapa, Geometría, Serie, o la URL del Catálogo de Entidades y Observaciones.
- Se han introducido nuevos atributos, manual o automáticamente, al completar el modelo con los atributos ***beginLifespanVersion***, ***endLifespanVersion*** y ***nativeness***.
- Se ha podido realizar la correspondencia entre los valores del NGCE y los valores de las listas controladas (*codeList*), tales como ***NamedPlaceTypeValue***, ***NameStatusValue*** y ***NativenessValue***.
- Se ha realizado la transformación de las coordenadas del NGCE, que se encontraba en coordenadas geodésicas en ED50 (EPSG:4230) a coordenadas geodésicas en ETRS89 (EPSG:4258). En España el Real Decreto 1071/2007² que regula el sistema geodésico de referencia oficial en España establece que a partir del 2015, toda la información geográfica producida o actualizada por las Administraciones Públicas contenga la referencia a ETRS89, además, los conjuntos de datos espaciales en el ámbito de aplicación de INSPIRE se deben representar mediante coordenadas geodésicas en ETRS89 en el continente europeo y WGS84 (EPSG:4326), equivalente a ETRS89, en territorios de ultramar.

² <http://www.boe.es/boe/dias/2007/08/29/pdfs/A35986-35989.pdf>

La relación de atributos del NGCE con los atributos INSPIRE se muestra en la Figura 2.

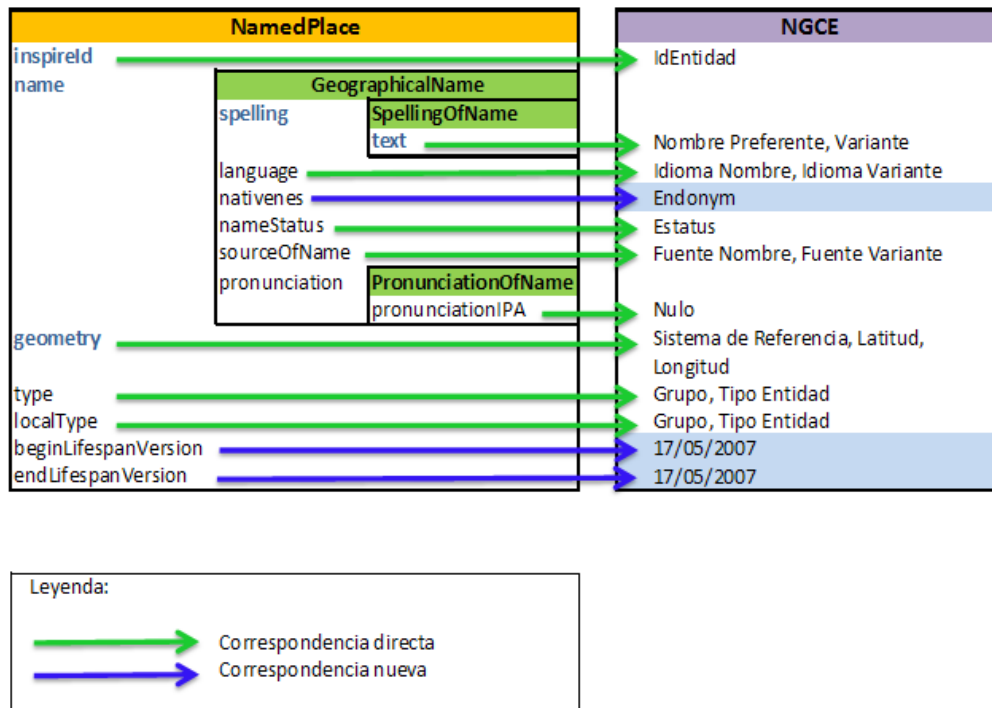


Figura 2. Cuadro de la correspondencia de atributos INSPIRE – NGCE.

Una vez realizado el proceso correctamente se ha generado la codificación GML para cada entrada del NGCE.

5 Problemas encontrados

Al hacer la correspondencia, se ha detectado que algunos campos que componen la BD de partida, no eran datos «puros», como por ejemplo el atributo «Idioma Nombre» del NGCE, para el que había celdas con dos nombres en dos lenguas oficiales separados por una barra inclinada, como «spa/cat», ya que en la celda correspondiente a «Nombre Preferente» figuraba «Chilches/Xilxes». En este caso se han tenido que desdoblar el topónimo y el idioma en dos, ya que existen dos topónimos, Chilches y

Xilxes en el ejemplo. Se han considerado como dos *GeographicalName* específicos dentro del *NamedPlace*.

El problema de los nombres geográficos compuestos en dos idiomas, como «Donosti-San Sebastián», se ha resuelto almacenando «Donosti-San Sebastián» como un único *GeographicalName* y almacenado en el campo que describe el idioma del nombre el código ISO 639-2 valor «mul» (múltiple).

Otro problema ha sido el de depurar la base de datos en Access. Por ejemplo, en los campos de Latitud y Longitud, algunos registros no contenían los caracteres «^o» o «'», y se han tenido que introducir para que la conversión con FME no diese problemas.

En el atributo «*geometry*» de INSPIRE, se ha tenido que realizar un compendio de atributos correspondientes en el NGCE, como «Latitud», «Longitud» y «Sistema de Referencia». Con estos atributos se han generado objetos *GM_Object*, concretamente *GM_Point*.

Por otro lado, teniendo en cuenta el formato, no se estaba teniendo en cuenta el RD 4/2010 que define el Esquema Nacional de Interoperabilidad (ENI) y que establece que las Administraciones Públicas no deben publicar datos utilizando como única opción un formato que no sea un estándar abierto, como ocurre con los ficheros Access. En ese sentido el ENI obliga a implementar al menos una alternativa basada en estándares abiertos. Se ha optado por publicar también un fichero en formato GML, siguiendo lo establecido por las Especificaciones INSPIRE de Nombres Geográficos en cuanto a formato por defecto.

En cuanto a la correspondencia de los tipos de objetos geográficos considerados en el NGCE y los valores definidos para *NamedPlaceType* (5.2.2.3.3), sería:

NGCE	<i>INSPIRE NamedPlaceTypeValue</i>
División administrativa	administrativeUnit
Poblaciones	populatedPlace
Orografía	landform
Hidrografía	hydrography
Costa (cabo, golfo, playa, isla)	landform
Costa (Mar)	hydrography

Parque natural	protectedSite
Aeropuerto	transportNetwork
Puerto (marítimo)	transportNetwork

6 Otros aspectos

Este proyecto se ha centrado en primer lugar en la correspondencia entre los dos modelos de datos, el del IGN y el de INSPIRE, pero hay otros aspectos que se han de tener en cuenta para que el conjunto de datos sea conforme a INSPIRE:

- La calidad de los objetos espaciales ha de reflejar las evaluaciones de la compleción y de la exactitud posicional. Se ha revisado exhaustivamente el contenido del NGCE en relación al Mapa General de España 1:1.000.000, que constituye el Universo de Discurso en este caso, y se puede asegurar que no hay errores de compleción ni por omisión ni por comisión. En cuanto a la exactitud posicional absoluta, se puede estimar en la propia del mapa original, es decir 200 metros.

- Los metadatos deben cumplir el Reglamento (CE) 1205/2008 de la Comisión de 3 de diciembre de 2008 por el que se ejecuta la Directiva 2007/2/CE del Parlamento Europeo y del Consejo en lo que se refiere a los metadatos y los ítems adicionales citados en la especificación de datos. No se han generado todavía los metadatos INSPIRE de la nueva versión del NGCE, pero dado que se dispone de los metadatos de la versión anterior es una tarea que se completará rápidamente.

- En caso de implementar un servicio de visualización (WMS) habría que tener en cuenta los estilos por defecto y el nombre INSPIRE de la capa, pero el marco legal sólo obliga a publicar los datos de nombres geográficos con servicios en red, en general, y se ha decidido publicar un servicio de nomenclátor, basado en un WFS, y un servicio de descarga de ficheros.

- Cada una de las entidades geográficas (*NamedPlace*) del NGCE deben tener un identificador único INSPIRE que se formará mediante la concatenación de los identificadores del país, del productor, del producto y un identificador interno

- Los datos conformes a la especificación de datos de INSPIRE deben estar disponibles a través de un servicio de red INSPIRE, el Nomenclátor Conciso se ofrecerá como un servicio de descarga, que cumpla la especificación WFS de OGC. Este servicio sustituirá al WFS que actualmente está disponible.

- De acuerdo a la política de datos del IGN definida en la Orden Ministerial FOM/956/2008 y por formar parte del Equipamiento Geográfico de Referencia Nacional (EGRN), los datos del NGCE pueden ser utilizados de manera gratuita y para todo tipo de aplicaciones, incluidas las comerciales, con la única condición de mencionar la autoría del IGN.

- En cuanto a la actualización del NGCE, dada la baja tasa de variaciones que se prevé, se contempla que se realice cada vez que cambie el nombre oficial de alguna entidad (Unidad administrativa, capital o población) y cuando se detecte que ha habido un número suficiente de modificaciones.

Ultimando estos detalles, documentándolos en los metadatos, y habiendo solucionado ya los aspectos relativos al CRS y a la descarga en formato GML, se dispondrá de una versión del NGCE conforme a INSPIRE en este mes de octubre y a continuación se pondrá a disposición pública en el Centro de Descargas del CNIG.

7 Conclusiones

Con este trabajo, el IGN ha puesto en práctica lo desarrollado teóricamente en un documento técnico de INSPIRE, contribuyendo a la implementación del resto de especificaciones de datos que son necesarias llevar a cabo en esta organización.

El trabajo de reingeniería de datos que se ha tenido que invertir para hacer que el NGCE sea conforme al marco INSPIRE no ha sido excesivo, debido a su pequeño volumen, y ha sido un buen banco de pruebas de las especificaciones y métodos de trabajo ensayados.

La consistencia del tema Nombres geográficos con los nombres de las instancias en el resto de temas INSPIRE, es una cuestión que se resuelve a través de la clase «*GeographicalName*», que enlazará con otros temas, como Parcelas Catastrales(*CadastralZoning*), Hidrografía (*HydroObject*),

Redes de Transporte (*Transport Network*), Elevaciones (*Elevation*) o Regiones Marinas (*Sea Area Name*).

8 Referencias bibliográficas

[1] <http://inspire.jrc.ec.europa.eu/>

[2] http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_GN_v3.0.1.pdf

[3] <http://www.ign.es/ign/layoutIn/herramientas.do#NOMENCLATOR>

[4] www.ideo.es

[5] www.ign.es

[6] <http://www.langsci.ucl.ac.uk/ipa/>

[7]

<http://www.ign.es/ign/resources/herramientas/Nomenclator/CatalogodeEntidades.xls>

[8]

http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/D2.5_v3.4rc2.pdf