

Desarrollo de aplicación de control de calidad semántico con Software libre para la explotación SIG de los productos cartográficos del Instituto Geográfico Nacional.

Rafael Fernández Mejías⁽¹⁾ y Jimena Martínez Ramos⁽¹⁾

Agradecimientos a Jose Antonio Merino Martín⁽²⁾

⁽¹⁾Sinfogeo S.L. Paseo de la Habana, 9-11, edificio unicentro, 28036 Madrid

⁽²⁾Instituto Geográfico Nacional. Subdirección de cartografía

RESUMEN

El control de calidad en los proyectos de producción cartográfica es una fase que, en muchas ocasiones, no está bien resuelta en los organismos productores. Estos organismos cuentan con diferentes aplicaciones y procedimientos para controlar la calidad de los productos que las instituciones, departamentos o empresas van ejecutando. Y estas aplicaciones suelen estar basadas en software propietario y no permiten la interoperabilidad entre Modelos de Datos, ni son fácilmente actualizables a las nuevas necesidades.

El presente trabajo desarrolla una aplicación de control de calidad semántico para los productos Base Cartográfica Numérica del IGN (escalas 1/100.000 y 1/200.000 en principio), basada en software libre. De este modo se pretende fomentar el uso de herramientas de código abierto en los procesos de producción cartográfica, favoreciendo la reutilización de código para implementaciones futuras. Esto será beneficioso tanto para los organismos productores, que tendrán más control sobre sus productos, como para las empresas que se encarguen de producir y realizar controles de calidad, que podrán trabajar sobre aplicaciones ya desarrolladas y mejorar los procesos de control.

La aplicación está desarrollada en C++ y python, utilizando QT y Quantum GIS y consta de:

- 1.- Un importador ligero de los ficheros utilizados por el IGN.*
- 2.- Algoritmos de lectura y validación de datos, que es el motor del control semántico, que validan la estructura de la BBDDs y cada uno de los campos de las tablas que componen el modelo de datos de BCN.*
- 3.- Un generador sencillo de los metadatos asociados a la información geométrica y a los procesos de calidad seguidos, cuyos ficheros xml resultantes se generan conforme a las instrucciones INSPIRE.*

Palabras clave: Control de calidad, QT, Quantum GIS, IGN, Metadatos, INSPIRE

1. INTRODUCCIÓN

La proliferación de fuentes de información como resultado de los avances en la web 2.0, ha llevado a cambios significativos en la cantidad, disponibilidad, y naturaleza de la información geográfica. Entre los cambios más significativos está la cantidad cada vez mayor de información geográfica y, aunque el volumen de información ha mejorado los datos geográficos, también ha suscitado preocupaciones con respecto a su calidad, confiabilidad y usabilidad (Flanagin y Metzger 2008).

En el presente artículo se realiza una breve exposición del proyecto denominado Base Cartográfica Numérica a escala 1:200.000 (BCN200) del Instituto Geográfico Nacional (IGN), como ejemplo de un moderno modelo de datos aplicable a otros muchos productos.

Se hace especial referencia a la evolución del mismo desde el año 2006 hasta finales del año 2010, donde las mejoras que se han introducido por las partes implicadas han logrado que sea un producto viable económicamente para la Propiedad.

Una vez expuestos los antecedentes, se justifica el empleo de una herramienta *Open Source* de conversión de ficheros Geomedia a un formato público, para después emplear sobre esos ficheros, la herramienta de control de calidad semántico para el producto BCN200, ambas basadas en software libre y realizadas en lenguaje python y Quantum GIS (QGIS) (PyQGIS, 2009) la primera y en C++ y QT la segunda.

Por último, se presenta también una aplicación de ayuda para generar los metadatos necesarios bajo las normativa del Núcleo Español de Metadatos (NEM) (Serrano y Valcarce 2000).

2. EL PROYECTO BCN (*)

“La Base Cartográfica Numérica BCN200 es un sistema de información geográfica multipropósito que alberga datos topográficos y temáticos, y es capaz de servir de soporte tanto a consultas geográficas como a la producción de diversos productos cartográficos, además de constituir la base de referencia para otros sistemas de información geográfica, tanto dentro del IGN como en otras Administraciones y empresas privadas.

La base de datos que sustenta BCN200 está basada en el modelo entidad-relación. La entidad se define como un fragmento de fenómeno con atributos invariables, por lo que cada tabla contiene una clase de entidad. Todas ellas contienen geometrías simples y los atributos asociados a estas geometrías pueden ser comunes a todas las clases de entidad (es el caso del ID, ID_BD, FECHA DE ALTA, FECHA DE BAJA, etc.) o específicos de cada una de ellas (CURSO, en el caso de los ríos o ITINERARIO_EUROPEO en el caso de las carreteras, por poner un par de ejemplos).



Fuente: documentación de proyecto BCN200. IGN

Para la definición de las clases de entidad se ha procurado respetar las ya establecidas previamente para la BTN25 del IGN, con el fin de favorecer la interoperabilidad entre bases de datos. De hecho, muchas de las clases de entidad así como sus atributos y el propio nombre de éstas se realizaron teniendo siempre presente la hermana mayor de las bases cartográficas del IGN.

La actualización de BCN200 se realiza a través de contratos a empresas privadas, cuya adjudicación se resuelve en función de factores como la descripción de los trabajos y metodología empleada, el control de calidad desarrollado y el sistema on-line de seguimiento de trabajos implementado. Se da importancia al desarrollo de herramientas con código abierto en entornos no propietarios, tanto dentro de la metodología como para el control de calidad.” (Merino Martín, 2010)

(*) Extraído del artículo 'BCN200 del IGN-E: base cartografica multipropósito', de las I Jornadas Ibéricas de Infra-estructuras de Datos Espaciales, redactado por los responsables del proyecto del IGN.

3. HISTORIA DEL PROYECTO

El primer contrato de BCN200 se produjo en el año 2006, aquel modelo de datos era más complejo en cuanto a geometrías y menos interoperable con otros productos. Desde entonces hasta ahora la situación ha cambiado considerablemente.

Características BCN200	2006	2010	
Nº de entidades catálogo	533	46	Catálogo
Temas a tratar	4	2	
Dependencia de BTN25	<i>Mucha</i>	<i>Poca</i>	Madurez
Tipo de modelo	<i>Basado en Elemento</i>	<i>Basado en Entidad-relación</i>	
Tipos de elementos	<i>P, L, Perimetral, S</i>	<i>P, L, S</i>	
Tipo de cartografía	<i>Derivada</i>	<i>Derivada, con matices</i>	
Ejemplos de captación	<i>No</i>	<i>Si</i>	Claridad
Riqueza del Pliego	<i>Escasa</i>	<i>Suficiente</i>	
Flujo de procesos definido	<i>No</i>	<i>Si</i>	
Responsables implicados en obtener fuentes fiables	<i>No</i>	<i>Si</i>	Iniciativas
Armonización con Nomenclátor geográfico	<i>No</i>	<i>Si</i>	
Aplicaciones desarrolladas	<i>No</i>	<i>Si</i>	
Predisposición a la convergencia de producto	<i>Sólo algunas de las partes implicadas</i>	<i>Todas las partes implicadas</i>	
Control de calidad	<i>Visual. Analítico muy pobre</i>	<i>Definido.</i>	Gestión
Seguimiento del proyecto	<i>Escasa trascendencia</i>	<i>Mucha trascendencia</i>	

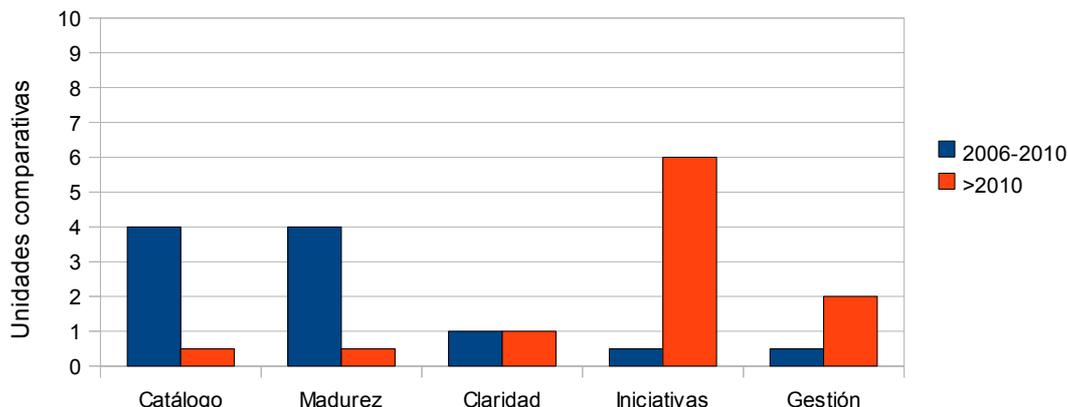
Fuente: propia

Resulta alentador que las mejoras introducidas hayan permitido paulatinamente reducir el tiempo de producción, a la vez que ha aumentado la calidad.

Desde el año 2006 hasta el año 2010, toda España ha sido actualizada por provincias al menos una vez. Pero más importante que la actualización temporal en sí, que se ha producido mediante imágenes SPOT5, ha sido la adecuación a la escala y al Modelo (mejoras descritas en catálogo, madurez y claridad)¹, por lo que para contratos posteriores a 2010 existirá un considerable ahorro de tiempo con respecto a los contratos anteriores, por el simple hecho de que la información ahora cuenta con más calidad y son necesarios menos pasos para su ejecución, es decir, menos tiempo dedicado. El incremento de calidad también, al igual que las mejoras en iniciativas y gestión, capacita a las empresas para producir más eficientemente.

¹Ver última columna de la tabla

Influencia de las mejoras en la reducción del tiempo de ejecución BCN200



Fuente: Propia en base a experiencias del proyecto desde el año 2006

A veces un producto puede no ser viable. Aplicado a la cartografía, el IGN podría disponer de presupuesto para producir BCN200 a un precio por unidad alto, pero seguro que repercutiría negativamente en el resto de sus productos, lo cual resulta del todo inaceptable. Por lo que la disminución del tiempo de producción era objetivo tanto de Sinfgeo como del IGN en aquella primera adjudicación, buscando por ambas partes la convergencia de producto, algo necesario para que BCN200 fuera económicamente viable.

Las mejoras introducidas de forma paulatina, fruto del esfuerzo entre IGN y Sinfgeo, permitirán que la Propiedad pueda licitar más unidades de producción con mayor cadencia, objetivo perseguido para alimentar la IDEE, el SIGNA, el PEIT, EUROREGIONALMAP, INTA, INECO, IGME, además de universidades, ingenierías e incluso ONGs.

4. INICIATIVAS DE ÉXITO

Bajo esta denominación se encuentran todas las características citadas en la tabla anterior, ya que todas han permitido reducir el tiempo de producción y aumentar la calidad hasta tener un producto maduro, como es ahora BCN200.

Pero una de las iniciativas aplicadas por Sinfogeo y permitida por el IGN, ha sido el desarrollo de aplicaciones geográficas orientadas a la producción, a la resolución de problemas innatos de la Serie o al control de calidad.

La Subdirección de Cartografía del IGN trabaja, entre otros pero principalmente, con el software Geomedia (Intergraph).

Existen muchos ejemplos de proyectos implementados por terceros o por la Propiedad, con el software Geomedia en los procesos productivos o de control de calidad dentro del IGN.

Así, se puede citar:

Mapa de España a escala 1:500.000 con Geomedia Professional 6.0

❖ **(ME500) PRESENTACIÓN**

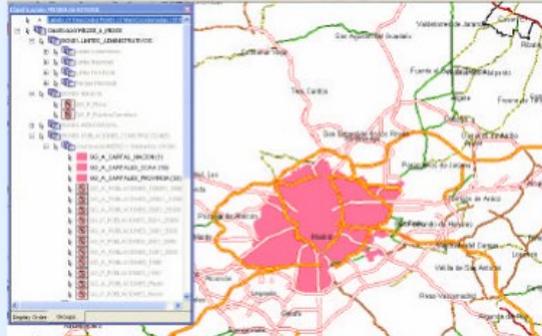

- Es un producto creado para el Atlas Nacional
- Constituye un proyecto piloto de cartografía completamente realizada en entorno SIG
- Nueva era en cartografía: pasamos de mapas impresos y divididos en hojas a una cartografía continua

1994 Mapas Impresos x Hojas



➔

Cartografía SIG 2007



Instituto Geográfico Nacional



La presentación referenciada de (Núñez Maderal, 2007), que se presentó en la reunión de usuarios de Intergraph, es el flujo de trabajo basado en Geomedia Professional, para conseguir el producto ME500 como cartografía derivada de BCN200 y consta de una serie de comandos desarrollados que se integran en el menú personalizado de la interfaz de Geomedia.

Fuente: Ver referencia a (Núñez Maderal, 2007)

Comando de Geomedia para Control de Calidad SIOSE

**Comando de GeoMedia para
Control de Calidad SIOSE**



Utilización de GeoMedia para el control de calidad

Mayo de 2008



Security, Government & Infrastructure



Como indica la propia presentación (INTERGRAPH, 2008), el desarrollo de este comando (un comando no es mas que una dll escrita en visual basic que Geomedia es capaz de integrar en su interfaz para ser utilizada como si de un plugin o extensión se tratara), el cual en este caso forma parte de la utilización de Geomedia para el control de calidad del producto SIOSE.

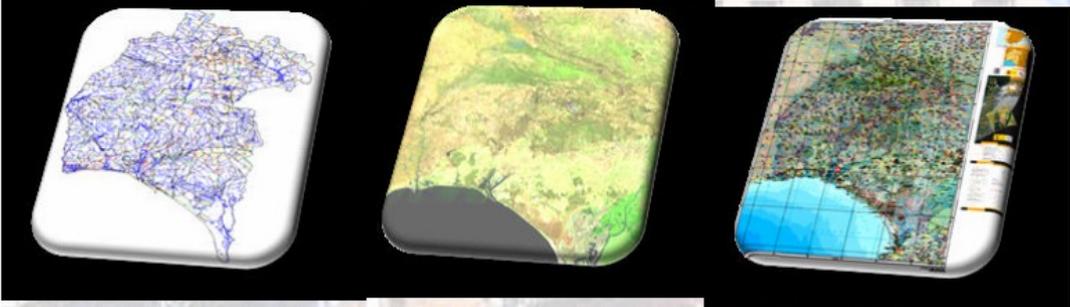
Fuente: Ver referencia a (INTERGRAPH, 2008)

Producción de Cartoimágenes en entornos GIS: Provincia de Huelva



Instituto Geográfico Nacional

Producción de Cartoimágenes en entornos GIS: Provincia de Huelva.



Jesús Celada Pérez¹ - Pilar de Lózar² - Isaac Barredo³ - Julio Vieco⁴ - Francisco J. González⁵
¹jcelada@ign.es - ²mpdlozar@ign.es - ³ibarredo@ign.es - ⁴jvieco@ign.es - ⁵fjgmatesanz@ign.es
Instituto Geográfico Nacional – Subdirección General de Producción Cartográfica

Reunión Usuarios Intergraph 2007 – Madrid, 5 noviembre 2007 1

❖ 1. Introducción



■ Cartoimagen...¿qué es?:

- Combinación de información ráster + información vectorial.
- Ráster: Imagen satélite SPOT5 (2.5m/pixel - 500Mb)
- Vector: ficheros DGN BCN200
- Hasta 2006: Producidas con ficheros DGN en Microstation

■ ¿Por qué Geomedia?:

- Geomedia 6.0:
 - Modifica los procesos y el flujo de trabajo.
 - Reduce el tiempo de producción e incrementa la producción.
- Requerimientos previos:
 - Manejo de Bases de Datos.
 - Definición de un Modelo de Datos BCN200
 - Cargar ficheros DGN en la BDG (ORACLE)
 - Manejo de Geomedia

Instituto Geográfico Nacional



Reunión Usuarios Intergraph 2007 – Madrid, 5 noviembre 2007

3

Donde se puede observar en la propia presentación (Celada Pérez, 2007) , la utilización del software Geomedia para los procesos productivos descritos.

Fuente: Ver referencia a (Celada Pérez, 2007)

También Sinfogeo en sus primeros desarrollos para el IGN, realizó aplicaciones con la librería de objetos de Geomedia (*Geomedia Object Model*). Se desarrollaron numerosas herramientas, muchas de ellas para dar solución a problemas particulares que tenían los almacenes BCN200 en origen:

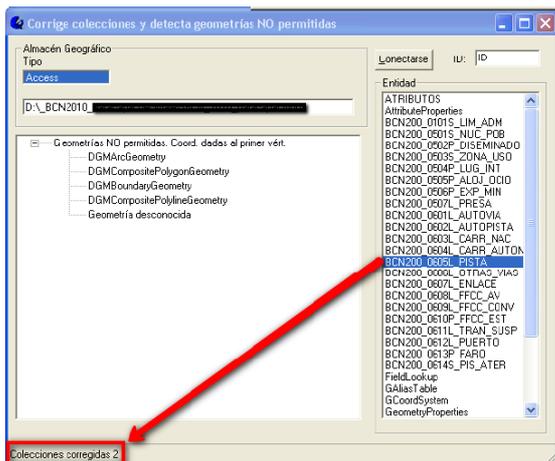
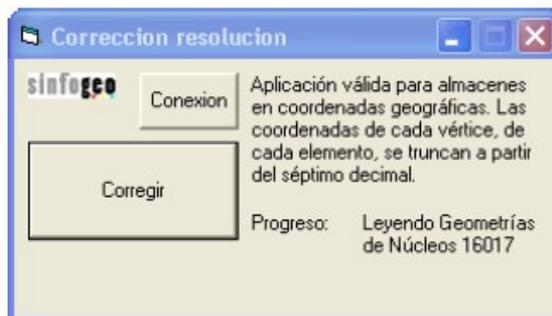


Integridad de almacenes

Aplicación que lee las tablas de metadatos de un almacén access (warehouse) y determina si cumple con los criterios relacionales de la estructura Geomedia.

Corrección de la resolución

El problema de la injerencia. Cambio de resolución en un almacén. La aplicación va leyendo cada una de las tablas de geometría y para cada elemento encontrado, determina su tipo gráfico. En función del tipo gráfico, va entrando en los vértices que lo componen y trunca la coordenada, volviendo a escribir ese elemento en el almacén.

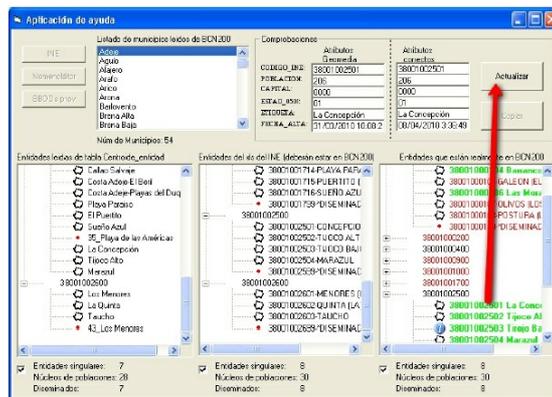


Corrección de elementos de tipo Collection

Detecta geometrías no permitidas por el catálogo de entidades y corrige colecciones de elementos, convirtiendo estos a elementos simples.

sinfoINE

Herramienta que facilita la incorporación de los códigos INE, población, nombre y titularidad de la capitalidad del municipio. Esta aplicación ha sido de capital importancia para fijar unos criterios objetivos comunes de captación de entidades singulares.





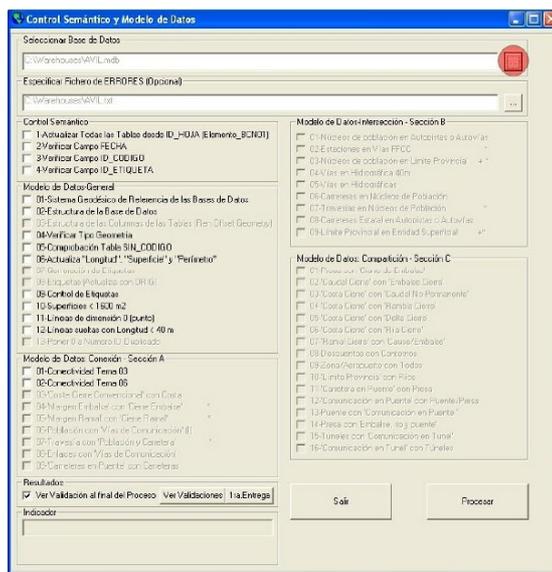
ActualizaFecha

Comando para dentro de gráficos en Geomedia. Siempre que se actualice alguna entidad, ya sea en su geometría o en sus atributos, se modificará la FECHA_ALTA al día de la modificación, con el formato dd/mm/aaa.

Qcv1.0

Esta aplicación permite detectar errores geométricos, topológicos y semánticos de acuerdo con las especificaciones técnicas del proyecto.

Está escrita utilizando el motor de objetos Geomedia y programada en entorno propietario (VB).



En el año 2010 es cuando Sinfogeo comienza a cargar procesos sobre el software Feature Manipulation Engine (FME Safe) para tareas específicas de este proyecto, como por ejemplo:

- Recortar zonas de cartografía a partir de límites administrativos
- Detectar y corregir típicos errores de injerencia
- Detectar y corregir errores geométricos (undershoot, overshoot, overlaps, etc.)
- Áreas mínimas
- Geometrías desconocidas o no válidas
- Líneas de longitud nula
- Áreas de superficie nula
- Puntos duplicados
- Kickbacks
- Entidades lineales de longitud mínima
- Kinks, loops
- Intersección de geometrías no permitidas
- Faltas de coincidencia
- Caras compartidas no permitidas
- Slivers, Gaps

A la vez, y contando con el apoyo del IGN, es cuando se empieza a pensar en buscar soluciones libres en el ámbito del control de calidad, empezando por algo útil y relativamente sencillo como es el Control de Calidad semántico, que es lo que da sentido a este artículo.

5. CAMBIO DE FORMATO PRIVADO A PÚBLICO

Las empresas que participen en la producción de BCN200, podrán ejecutar el contrato con el software más conveniente para ellas. Pero cualquier acción que se quiera realizar sobre los almacenes originales pasa por 'leer' su geometría y atributos y pasar de un formato privado a otro público; y si bien es verdad que podría ser el propio IGN el que realizara la conversión, se ha considerado adecuado realizar una herramienta gráfica de conversión de Access Geomedia (.mdb) a Esri Shapefiles (.shp), liberando de ese tiempo a los técnicos del IGN y contribuyendo de paso, a la promoción del software libre, tal como se aconseja en el documento oficial de la Unión Europea "Guideline on public procurement of Open Source Software" (Edler et al. 2005).

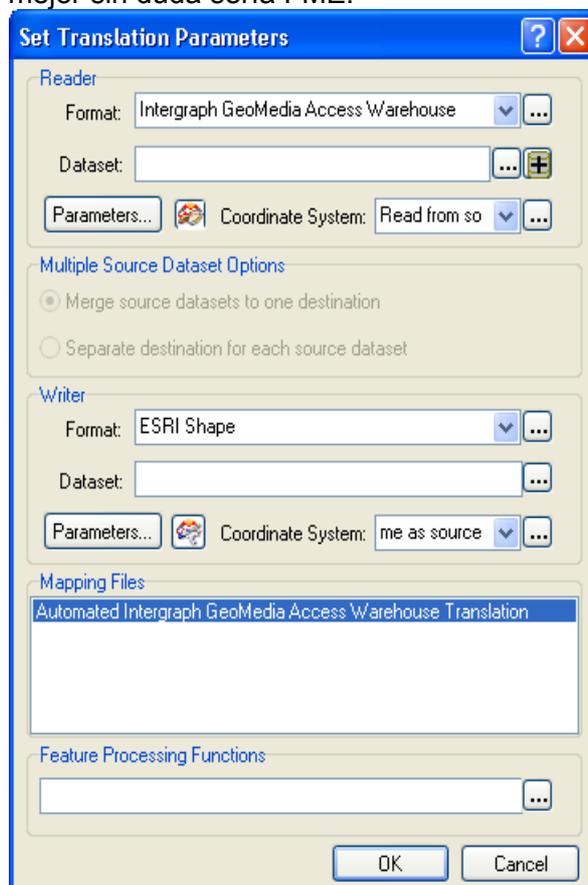


Guideline on public procurement of Open Source Software

March 2010
(revised June 2010)

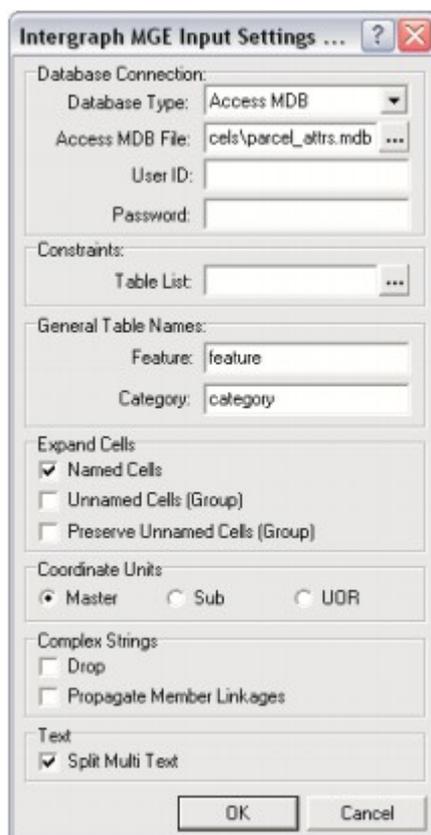
5.1 ¿Existe algo similar en el mercado?

Si se busca en software privativo, se pueden encontrar algunas alternativas al propio Geomedia, el mejor sin duda sería FME.



Fuente: Software FME Safe. Universal Translator

Pero ni uno ni otro software podrían estar al alcance de muchas empresas, las cuales en sus procesos de producción cartográfica pueden emplear Microstation, DIGI, ArcGis (por citar algunos de los más usuales) y ninguno de ellos cuenta con un conversor de Geomedia a SHP, únicamente ArcGis en su versión 10 sólo si se dispone de la licencia *Data interoperability Connection* (ESRI, 2006). La cual está basada a su vez en la integración con el motor de conversión de FME (SAFE Software, 2011).



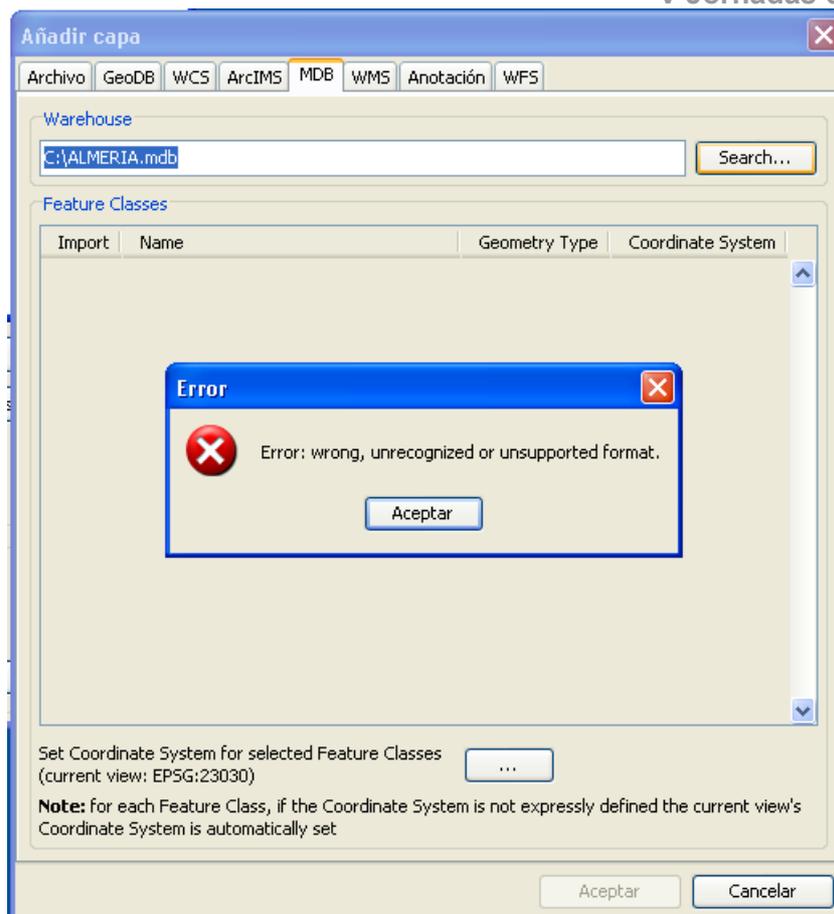
Fuente: Ver referencia (SAFE Software, 2011)

En el panorama del software libre sí que han existido iniciativas para lograr esta conversión, y así podemos encontrar la extensión **extMDB** para gvSIG (gvSIG.org, 2009):



Fuente: Ver referencia (gvSIG.org, 2009)

La cual, probablemente por motivos de versionado (gvSIG 1.9) o por la estructura de tablas de las listas codificadas del formato Intergraph, no llega a funcionar con las BBDDs probadas:



Fuente: Programa gvSIG (licencia GPL)

Algo más se podría esperar de GeoKettle (en lo que respecta a la conversión), ya que al tratarse de una herramienta ETL al estilo de FME, tendría más formatos soportados que los oficiales de OGR.



GeoKettle
ETL for Geospatial Data

© Copyright 2008-2009 GeoSOA research group, Laval University
Provided under the terms of the
GNU Lesser General Public License (LGPL; see below)

This software tool is based on:



Pentaho Data Integration
Previously Kettle

Version 3.2

© Copyright 2008 Pentaho Corporation. All rights reserved.
Originally developed by Pentaho Corporation provided under the terms of the
GNU Lesser General Public License (LGPL), Version 2.1 or any later version.
You may not use this file except in compliance with the license. If you need a
copy of the license, please go to <http://www.gnu.org/licenses/lgpl.txt>. The
Initial Developer is Pentaho Corporation.

Software distributed under the GNU lesser General Public License is distributed
on an "AS IS" basis, WITHOUT WARRANTY OF ANY KIND, either express
or implied. Please refer to the license for the specific language governing your
rights and limitations.

No es así, y aunque el esfuerzo realizado por sus desarrolladores es ímprobo, GeoKettle cuenta con un extenso listado de conexiones a BBDDs, pero actualmente no dispone de alta capacidad para convertir formatos. Lo cual es común en Kosmo, UDIG y QGIS, ya que estos también utilizan las librerías GDAL/OGR (<http://www.gdal.org/ogr/>) y dentro de los formatos soportados por esta librería, se puede encontrar el formato Geomedia MDB:

OGR Vector Formats

Format Name	Code	Creation	Georeferencing	Compiled by default
Aeronav FAA files	AeronavFAA	No	Yes	Yes
ESRI ArcObjects	ArcObjects	No	Yes	No, needs ESRI ArcObjects
Arc/Info Binary Coverage	AVCBin	No	Yes	Yes
Arc/Info E00 (ASCIT) Coverage	AVCE00	No	Yes	Yes
Atlas BNA	BNA	Yes	No	Yes
AutoCAD DXF	DXF	Yes	No	Yes
Comma Separated Value (.csv)	CSV	Yes	No	Yes
DODS/OPeNDAP	DODS	No	Yes	No, needs libdap
EDIGEO	EDIGEO	No	Yes	Yes
ESRI Personal GeoDatabase	PGeo	No	Yes	No, needs ODBC library
ESRI ArcSDE	SDE	No	Yes	No, needs ESRI SDE
ESRI Shapefile	ESRI Shapefile	Yes	Yes	Yes
FMEObjects Gateway	FMEObjects Gateway	No	Yes	No, needs FME
GeoJSON	GeoJSON	Yes	Yes	Yes
Geoconcept Export	Geoconcept	Yes	Yes	Yes
Geomedia .mdb	Geomedia	No	No	No, needs ODBC library
GeoRSS	GeoRSS	Yes	Yes	Yes (read support needs libexpat)

Fuente: www.gdal.org

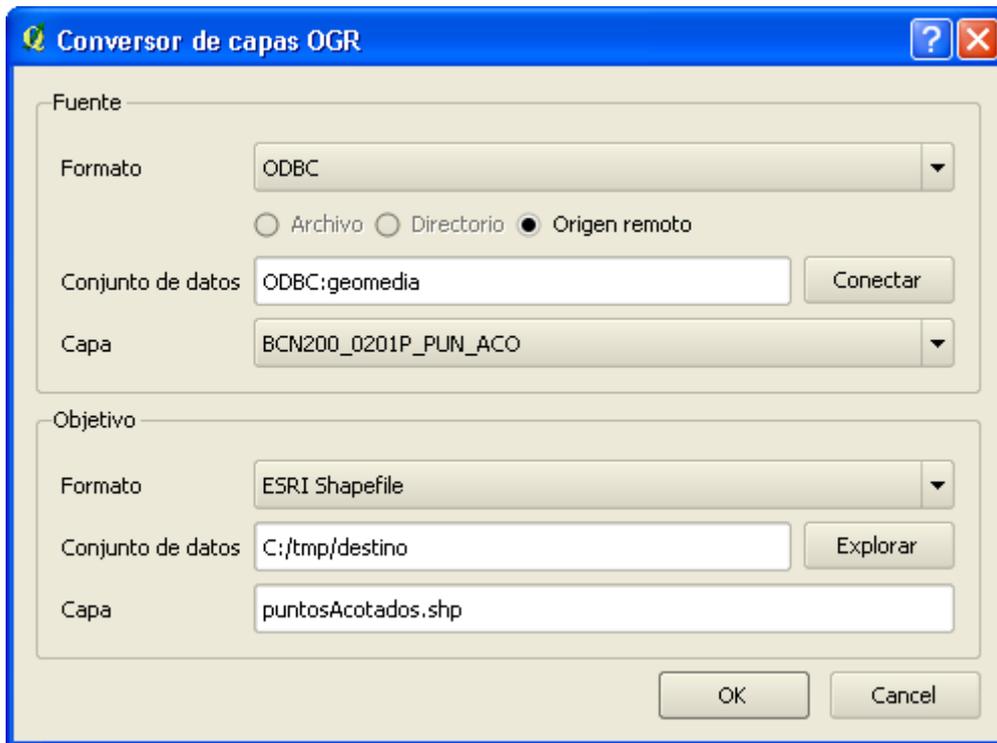
Que está soportado a partir de la versión 1.9, aunque sólo se soporta utilizando ODBC. Ninguno de los software libres mencionados está actualmente utilizando la librería 1.9 de GDAL.

Por ejemplo, Quantum GIS en su última versión (v1.6) incorpora una eficiente herramienta de conversión, donde se puede observar que el conjunto de formatos soportados es numeroso:



Fuente: Programa QGIS (licencia GPL)

Entre ellos se encuentra ODBC, necesario para obtener la conexión con el almacén Geomedia:



Fuente: Programa QGIS (licencia GPL)

Y si se pide a la aplicación “Conversor de capas OGR” de QGIS, que realice la conversión solicitada según la imagen, aun cuando se ha reconocido el DNS de nombre “geomedia” que apunta a la BBDDs access correcta, la cual nos proporciona las definiciones de cada capa o tabla. Obtendremos el siguiente error, ya que la librería GDAL/OGR que utiliza la última versión de QGIS es la 1.6, no la 1.9 donde supuestamente se soportaría el BLOB de la geometría de Geomedia.



Fuente: Programa QGIS (licencia GPL)

Por lo tanto, se justifica el desarrollo de una herramienta de conversión de almacenes Access de Geomedia a un formato público, aunque no libre, como es ESRI Shapefiles (ESRI, 1998).

La argumentación del formato destino elegido hay que buscarla en que SHP es popular, utilizado tanto en organismos públicos como privados, interpretado por todos los clientes ligeros o pesados y fácilmente exportable a su vez a Bases de Datos geográficas.

6. LA HERRAMIENTA DE CONVERSIÓN GEOMEDIA->SHP

El objetivo principal de esta aplicación es convertir un almacén Geomedia (warehouse) en ficheros de formato Shapefiles y se realiza en principio para disponer de la principal fuente vector de actualización de la que dispone el IGN, para el proyecto BCN200, en un formato público utilizable por empresas implicadas en la producción cartográfica.

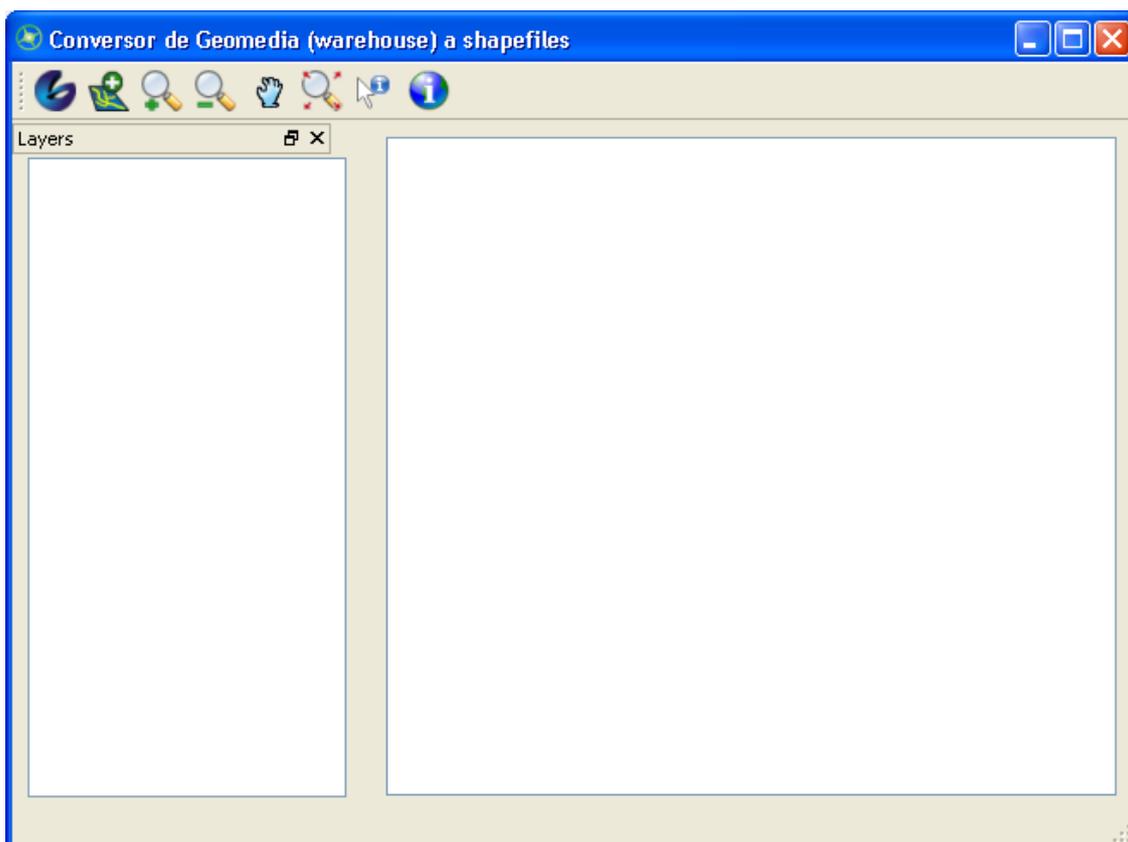
El objetivo secundario de esta herramienta general, al igual que para la siguiente que es más específica del producto BCN, es apoyar desde Sinfogeo la geomática basada en aplicaciones libres de código abierto.

Por último, esta aplicación de conversión es de uso general y válida por tanto, no sólo para el Modelo del IGN en su producto BCN200, sino que por homogeneización de productos, se podrá utilizar para todas las series cartográficas BCN.

Explicación técnica de la herramienta de conversión

Está desarrollada en lenguaje python, utilizando las librerías QT y el motor de Quantum GIS (PyQGIS) del paquete OsGeo.

(<http://download.osgeo.org/osgeo4w/osgeo4w-setup.exe>). Se incluye QGIS 1.6 y todas las dependencias necesarias, junto con Python, GRASS, GDAL, etc.



Fuente: Elaboración propia

La interfaz, en cuanto a su esqueleto, está basada en lo descrito en el “blog” de Geotux, titulado “Construcción de un visor de Shapefiles con herramientas libres: QGIS, Python y Qt” (Germán Carrillo, 2009) al cual los autores de este artículo le reconocen su trabajo.

Las versiones de las dependencias son las que existen en el paquete Osgeo en el momento de realizar el desarrollo (finales de 2010):

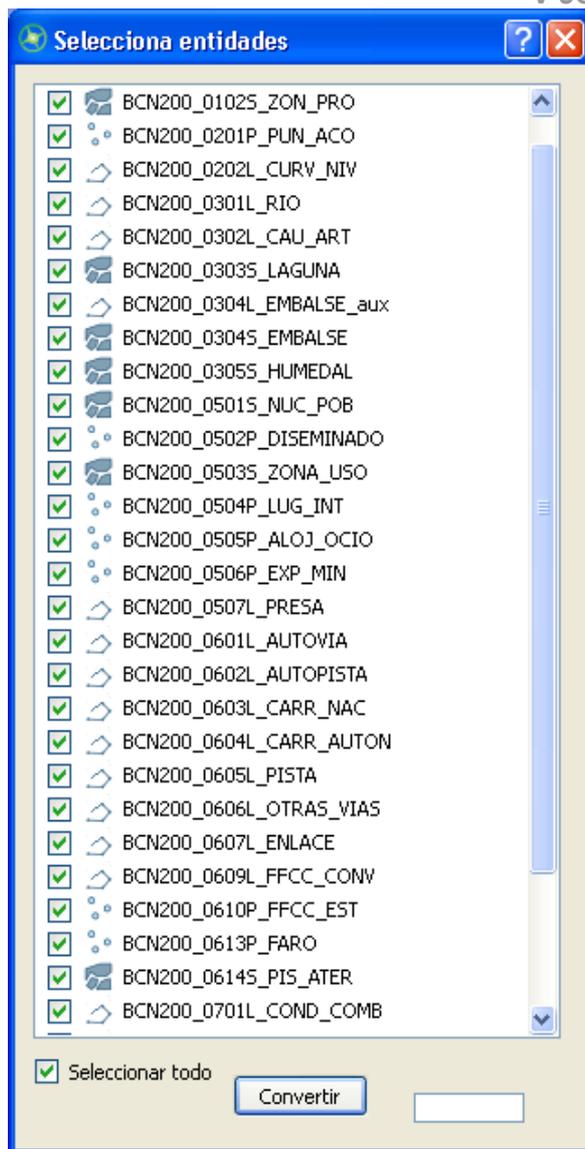
Dependencia	Versión
QGIS	Versión 1.6
Python	Versión 2.5
GDAL	Versión 1.7

Fuente: <http://download.osgeo.org/osgeo4w/>

Algoritmo: La lectura del almacén Geomedia se realiza iterando sobre cada una de las tablas del fichero MDB, conservando en memoria sólo aquellas que incluyen un campo binario largo, que es donde se almacenan las geometrías de las entidades. Una vez en memoria dichas tablas, se recorre cada una de ellas y por medio del binario se determina el tipo de geometría que almacena. La estructura del binario, por ingeniería inversa, es tal que los primeros 16 bytes se corresponden con el GUID de la entidad, que nos daría el tipo de elemento geomedia. Con el tipo de elemento y utilizando la librería OGR, se puede determinar el tipo de elemento Shapefile que le corresponde, construyéndose así la geometría destino.

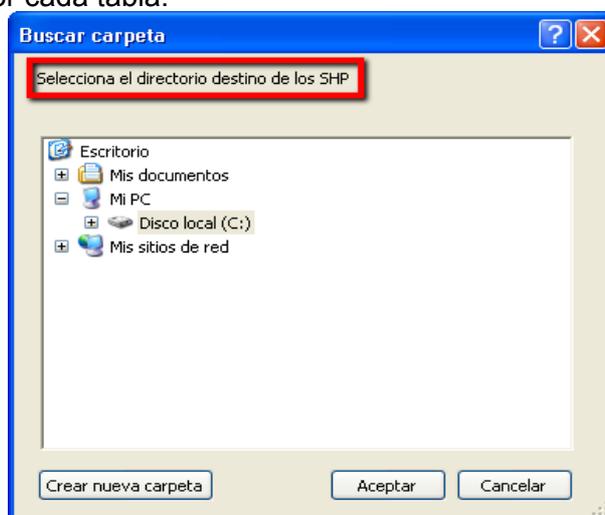
El copiado de atributos se realiza primero determinando las propiedades de los campos que configuran cada tabla, para lo cual se itera inicialmente sobre las columnas para lograr la definición del DBF y posteriormente se recorren para 'copiar' los mismos en cada entidad leída. Todo queda estructurado en diferentes clases y funciones Python.

El resultado para el usuario es la presentación de una ventana con las tablas que tienen geometría, con un icono al margen que indica el tipo de entidad y con la posibilidad de convertir sólo las que seleccione.



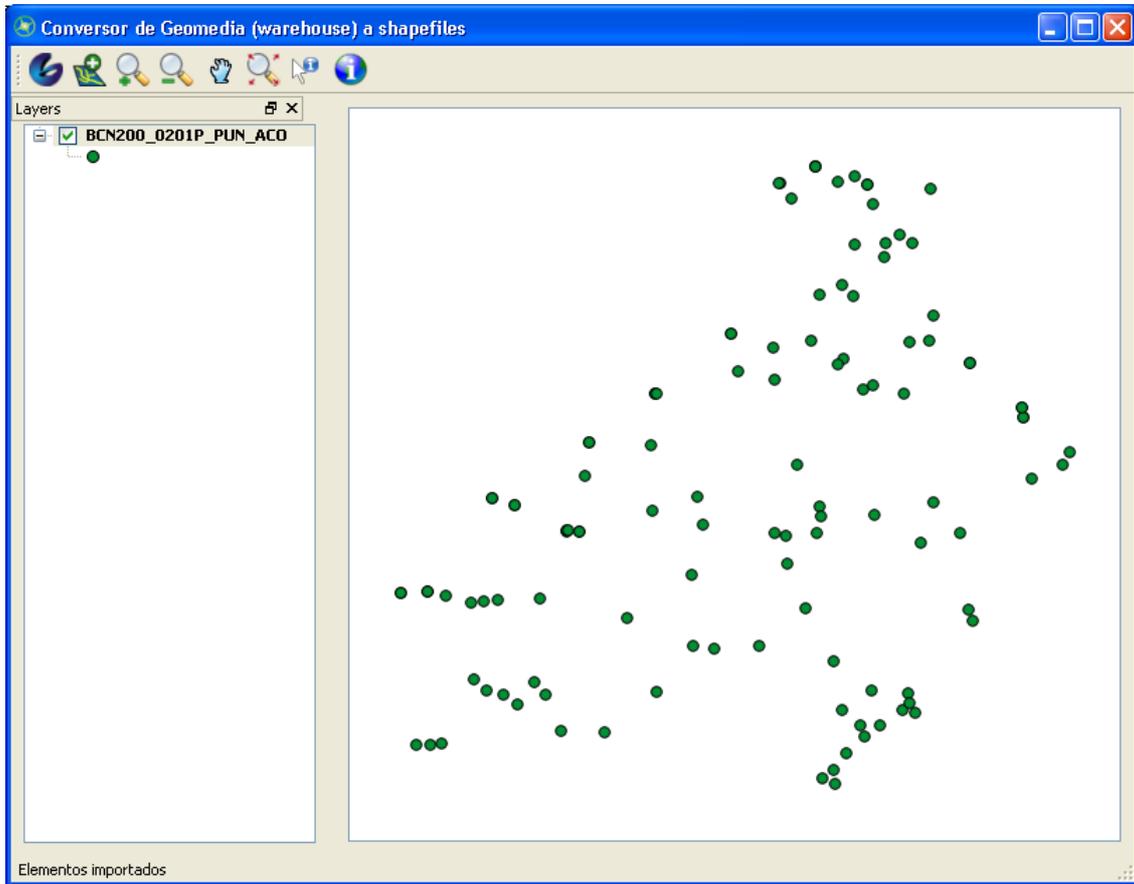
Fuente: Elaboración propia

Se deberá elegir el directorio destino donde la herramienta generará los diferentes ficheros SHP, uno por cada tabla:



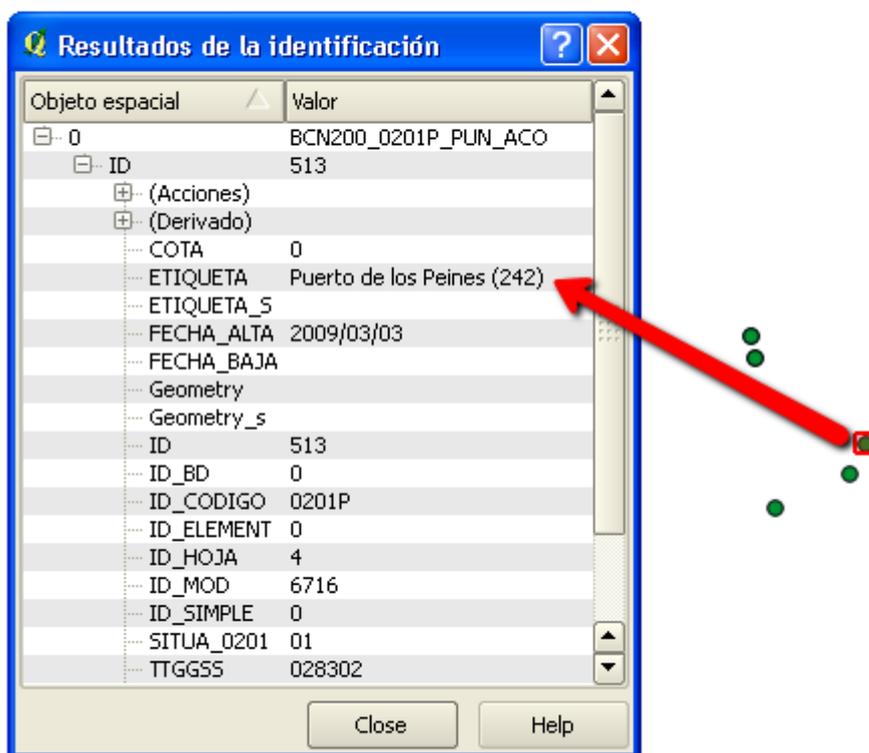
Fuente: Elaboración propia

El resultado de la conversión se puede observar aquí:



Fuente: Elaboración propia

El conversor no realiza cambio de Sistema Geodésico de Referencia (SGR), así que si en el almacén origen era sistema coordenado geográfico y elipsoide ETRS89, tal cual aparecerá en el canvas de la herramienta. También es posible comprobar el resultado del copiado de atributos de cada entidad:



Fuente: Elaboración propia

Con los ficheros (geometrías y atributos) convertidos a formato Shapefiles, se posibilitan los procesos productivos cartográficos basados en software libre, lo cual por los casos de éxito que se han producido en otros proyectos, se espera que contribuya a la mejora general del proyecto.

Esta aplicación quedará disponible para su uso público con licencia GPL (Wikipedia), en un repositorio que todavía está por definir, y que contará con un instalador para sistemas operativos windows que incluirá todas las dependencias necesarias para que el usuario no tenga que instalarse en su ordenador ni Quantum GIS, ni ninguna de las librerías utilizadas en el desarrollo.

7. HERRAMIENTA DE CONTROL DE CALIDAD SEMÁNTICO

La calidad semántica se refiere a la exactitud en la terminología de las clases de entidad con relación al tema que representan, así como la precisa utilización de los acentos, mayúsculas, minúsculas y partículas de enlace entre los nombres. El término exactitud semántica es un tema de actualidad que se encuentra en estudios realizados para determinar la calidad en la información geográfica generada voluntariamente y lo definen como “Medida que vincula la forma en que se captura el objeto y se representa en la base de datos conforme a su significado y a la forma en que debe ser interpretado.” (Haklay, M. 2008).

Esta aplicación resuelve los condicionantes de control semántico definido en el producto BCN200 a través de su Pliego de Prescripciones Técnicas (PPT). Hasta ahora, este proceso se ha realizado por medio de otros recursos, como por ejemplo a través de consultas por atributos concatenadas, donde se presuponían ciertos condicionantes de entrada.

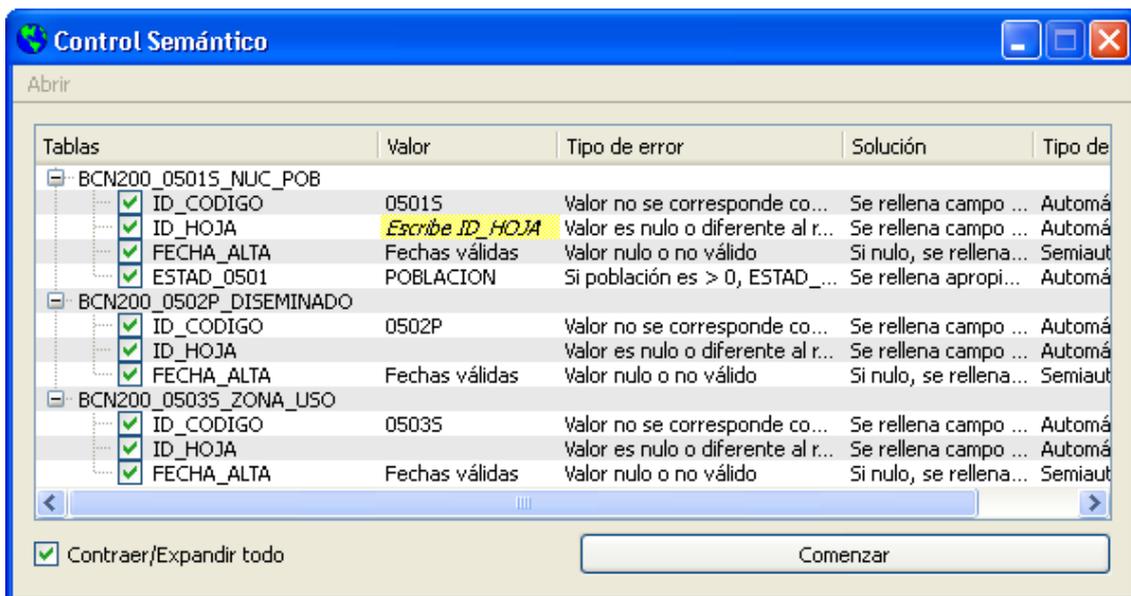
Resulta entonces conveniente disponer de una herramienta que pueda utilizar tanto la propiedad como los proveedores y que sin ser responsabilidad final del IGN, éste participe en su definición, modificación y mejora, como hasta ahora lo ha hecho.

El objetivo principal perseguido es descentralizar la dedicación de recursos y por tanto buscar esa viabilidad económica comentada al principio del artículo.

El objetivo secundario para esta aplicación no es otro que servir de punto de partida para disponer de un control de calidad lo más completo posible, donde estarían implicados agentes del IGN en la definición, y agentes de las empresas productoras en calidad de desarrollo informático.

Explicación técnica de la herramienta de control de calidad semántico

Está desarrollada en lenguaje C++, utilizando las librerías QT y el entorno de desarrollo (SDK) de Nokia, licenciado bajo LPGL (<http://qt.nokia.com/downloads>).



Fuente: Elaboración propia

Las versiones de los entornos utilizados son, en el momento de realizar el desarrollo (2010):

Producto	Versión
QT Creator	Versión 2.0.0
QT	Versión 4.7.0

Fuente: Elaboración propia

Algoritmo: La aplicación comienza leyendo cada una de las capas de un directorio y comparando el nombre de la tabla con el nombre que tendría según lo definido en el modelo de datos de BCN200. Si el nombre es correcto, se puede pasar al análisis de cada uno de los controles. El tipo de error encontrado se corresponderá siempre con un valor esperado en alguno de los atributos. En función del tipo de valor la solución al error podrá ser automática o semiautomática. Por ejemplo, un error en el identificador de hoja es corregible de forma automática porque se puede asignar de forma inicial por medio de la caja de texto habilitada en la columna 'valor'. Al final del proceso la aplicación habrá generado un documento PDF con el informe de los errores encontrados.

La explicación del control que se realiza a las entidades del modelo BCN200 es la siguiente (sobre un ejemplo de control sobre los elementos del tema 5 “Edificaciones y núcleos urbanos”):

Entidad	Control
0501S_NUC_POB	Este valor junto con el de la consulta sobre 0502S debe ser igual al número de registros del listado MAP.xls de capitales de municipio de la provincia. Se trata de que figuren en mapa todas las entidades oficialmente existentes.
0501S_NUC_POB	Comprobar que el único ID_CÓDIGO que hay en la tabla es 0501S.
0501S_NUC_POB	Comprobar que ID_HOJA es único e igual al COD_PROV.
0501S_NUC_POB	Detectar que la fecha de alta no está vacía y no tiene valores fuera del rango permitido en la actualización.
0501S_NUC_POB	Comprobar que si hay atributo DESHABITADO la POBLACION es 0 y que si es HABITADO la POBLACION es diferente de 0.
0501S_NUC_POB	Comprobar que todos los COD_INE acaban en 1, 2, 3...pero no en 0 ni en 99.
0501S_NUC_POB	Revisar que solo hay una capital de provincia.
0501S_NUC_POB	Revisar que no hay etiquetas vacías.
0501S_NUC_POB	Cada municipio tiene su capital, una y solo una. Pueden ser capitales también los diseminados en algún caso.
0501S_NUC_POB	Comprobar que no quedan COD INE vacíos.
0502S_DISEMINADO	Este valor junto con el de la consulta sobre 0501S debe ser igual al número de registros del listado MAP.xls de capitales de municipio de la provincia. Se trata de que figuren en mapa todas las entidades oficialmente existentes.

0502S_DISEMINADO	Comprobar que el único ID_CÓDIGO que hay en la tabla es 0502S.
0502S_DISEMINADO	Comprobar que ID_HOJA es único e igual al COD_PROV.
0502S_DISEMINADO	Detectar que la fecha de alta no está vacía y no tiene valores fuera del rango permitido en la actualización.
0502S_DISEMINADO	Comprobar que si hay atributo DESHABITADO la POBLACION es 0 y que si es HABITADO la POBLACION es diferente de 0.
0502S_DISEMINADO	Comprobar que todos los COD_INE acaban en 99.
0502S_DISEMINADO	A igual COD INE debe haber el mismo atributo POBLACION también. No debe haber etiquetas repetidas en varios diseminados
0502S_DISEMINADO	Comprobar que no quedan COD INE vacíos.
0503S_ZONA_USO	Comprobar que el único ID_CÓDIGO que hay en la tabla es 0503S.
0503S_ZONA_USO	Comprobar que ID_HOJA es único e igual al COD_PROV.
0503S_ZONA_USO	Detectar que la fecha de alta no está vacía y no tiene valores fuera del rango permitido en la actualización.
0503S_ZONA_USO	Detectar ETIQUETA o TIPO_0503S vacías - Seleccionar registros cuya ETIQUETA no comienza por uno de los genéricos - Detectar incongruencias de ETIQUETA y TIPO_0503S.
0504P_LUG_INT	Comprobar que el único ID_CÓDIGO que hay en la tabla es 0504P.
0504P_LUG_INT	Comprobar que ID_HOJA es único e igual al COD_PROV.
0504P_LUG_INT	Detectar que la fecha de alta no está vacía y no tiene valores fuera del rango permitido en la actualización.
0504P_LUG_INT	Detectar ETIQUETA o TIPO_0504P vacías - Seleccionar registros cuya ETIQUETA no comienza por uno de los genéricos - Detectar incongruencias de ETIQUETA y TIPO_0504P.
0505P_ALOJ_OCIO	Comprobar que el único ID_CÓDIGO que hay en la tabla es 0505P
0505P_ALOJ_OCIO	Comprobar que ID_HOJA es único e igual al COD_PROV.
0505P_ALOJ_OCIO	Detectar que la fecha de alta no está vacía y no tiene valores fuera del rango permitido en la actualización.
0505P_ALOJ_OCIO	Detectar ETIQUETA o TIPO_0505P vacías - Seleccionar registros cuya ETIQUETA no comienza por uno de los genéricos - Detectar incongruencias de ETIQUETA y TIPO_0505P.

0506P_EXP_MIN	Comprobar que el único ID_CÓDIGO que hay en la tabla es 0505P.
0506P_EXP_MIN	Comprobar que ID_HOJA es único e igual al COD_PROV
0506P_EXP_MIN	Detectar que la fecha de alta no está vacía y no tiene valores fuera del rango permitido en la actualización.
0506P_EXP_MIN	Detectar ETIQUETA o TIPO_0506P vacías - Seleccionar registros cuya ETIQUETA no comienza por uno de los genéricos - Detectar incongruencias de ETIQUETA y TIPO_0506P.
0507L_PRESA	Comprobar que el único ID_CÓDIGO que hay en la tabla es 0505P
0507L_PRESA	Comprobar que ID_HOJA es único e igual al COD_PROV
0507L_PRESA	Detectar que la fecha de alta no está vacía y no tiene valores fuera del rango permitido en la actualización.
0507L_PRESA	Detectar ETIQUETA o TIPO_0507L vacías - Seleccionar registros cuya ETIQUETA no comienza por uno de los genéricos - Detectar incongruencias de ETIQUETA y TIPO_0507L. La etiqueta debe comenzar por PRESA... y debe NO ESTAR VACIA

Fuente: Documentación de proyecto BCN200

En general, se trata de detectar que:

1. La suma de las entidades codificadas como capitales de municipio tiene que ser igual en nombre y número que las enumeradas oficialmente (<http://www.dgal.map.es/cgi-bin/webapb/webdriver?Mlval=munprov>).
2. Comprobación de la coherencia del atributo COD_INE con respecto a los datos oficiales de nomenclátor.
3. El ID_CODIGO se corresponde con el código de cada entidad.
4. El ID_HOJA es igual para todas las entidades y se corresponde con el código de provincia o de la unidad cartográfica correspondiente.
5. La fecha de alta del elemento está dentro del rango de la actualización.
6. Existe etiqueta para las entidades cuyo requisito es obligatorio.
7. La etiqueta tiene que ser congruente con la realidad representada.
8. La etiqueta deberá estar correctamente formateada con relación a los artículos y determinantes.

La aplicación, una vez que ha finalizado, genera un informe en formato PDF con el resultado del análisis:



mié 2. mar 18:04:51 2011

Informe sobre el Control semántico de Extremadura_ETRS89.mdb

BCN200_0501S_NUC_POB

ID_CÓDIGO: Se han leído 484 elementos
y se han corregido 0 elementos

ID_HOJA: Se han leído 484 elementos
y se han corregido 484 elementos

Fecha fuera de rango en ID:1836

Fecha fuera de rango en ID:1839

Fecha fuera de rango en ID:1841

Fecha fuera de rango en ID:1850

Fecha fuera de rango en ID:1852

Fecha fuera de rango en ID:1857

Fecha fuera de rango en ID:1860

Fecha fuera de rango en ID:1872

FECHA_ALTA: Se han leído 484 elementos
y se han encontrado 8 fechas incorrectas

Fuente: Elaboración propia

7.1 ¿Por qué es importante el nomenclátor geográfico en este proyecto?

El nomenclátor es una fuente oficial de documentación para el proyecto BCN200. El territorio nacional se encuentra dividido administrativamente en comunidades autónomas, provincias, municipios y otras entidades locales de ámbito territorial inferior al municipal, cuya delimitación, denominación, organización y competencias se describen y regulan con detalle en la legislación vigente en materia de régimen local.

Desde un punto de vista estadístico, esta división es insuficiente para conocer de qué forma se asienta la población en los municipios, debiendo descender a una subdivisión de los mismos, que no posee carácter oficial, pero sí gran tradición: las entidades colectivas y singulares de población, así como los núcleos y diseminados de estas últimas.

Como se comentaba al principio del artículo, en la historia del proyecto, la armonización con el nomenclátor ha sido una de las iniciativas que ha permitido tener objetividad en la captación de entidades de población, especialmente en las aisladas.

Para consultar las entidades de nomenclátor, se accederá a la siguiente dirección:

<http://www.ine.es/nomen2/index.do>

donde desplegando al año que nos interese, el sistema nos presentará los datos solicitados:

Municipio	Unidad Poblacional	AÑO 2008	Población total
083 Mérida	000000 MERIDA		55568
083 Mérida	000100 CAMPOMANES		14
083 Mérida	000199 *DISEMINADO*		14
083 Mérida	000800 ESTACION DE ALJUCEN		8
083 Mérida	000801 ESTACION DE ALJUCEN		8
083 Mérida	000900 ESTACION DE EL CARRASCALEJO		4
083 Mérida	000999 *DISEMINADO*		4
083 Mérida	001200 MERIDA		54979
083 Mérida	001201 MERIDA		54710
083 Mérida	001299 *DISEMINADO*		269
083 Mérida	001500 CARRION		21
083 Mérida	001501 CARRION		21
083 Mérida	001600 MRALRIO		40
083 Mérida	001601 MRALRIO		40
083 Mérida	001700 PRADO (EL)		4
083 Mérida	001701 PRADO (EL)		4
083 Mérida	001799 *DISEMINADO*		0
083 Mérida	001800 PROSERPINA		151
083 Mérida	001801 PROSERPINA		151
083 Mérida	001900 PSIGUATRICO		224
083 Mérida	001901 HOSPITAL PSIGUATRICO		224
083 Mérida	002000 VIRGEN DE LA LUZ		20
083 Mérida	002001 VIRGEN DE LA LUZ		20
083 Mérida	002100 VIVERO (EL)		103
083 Mérida	002101 VIVERO (EL)		103

Población total Igual a

Criterio Denominación

Unidad poblacional cuyo nombre Contenga

Seleccione años 2006 2007 2008

Fuente: <http://www.ine.es/nomen2/index.do>

Los diseminados serán entidades de tipo puntual y se corresponden con las edificaciones o viviendas de una entidad singular de población que no pueden ser incluidas en el concepto de núcleo de población.

Los diseminados recogen todas las poblaciones contempladas en el INE como diseminadas y en MP200 como cortijos, casas y demás entidades que no sean los núcleos de población. Se incluyen como un único punto y, de forma general, deben captarse al menos todos los que vienen en los listados del INE.

Al igual que el resto de iniciativas reseñadas, el hecho de que el IGN disponga de una BBDDs de toda España armonizada con nomenclátor, implica que ya hay un fuerte trabajo desarrollado y que en próximos

La armonización con el nomenclátor, debido a la dificultad para encontrar la correspondencia geográfica de los listados INE, ha sido hasta el año 2010, motivo de un gran consumo de recursos en producción.

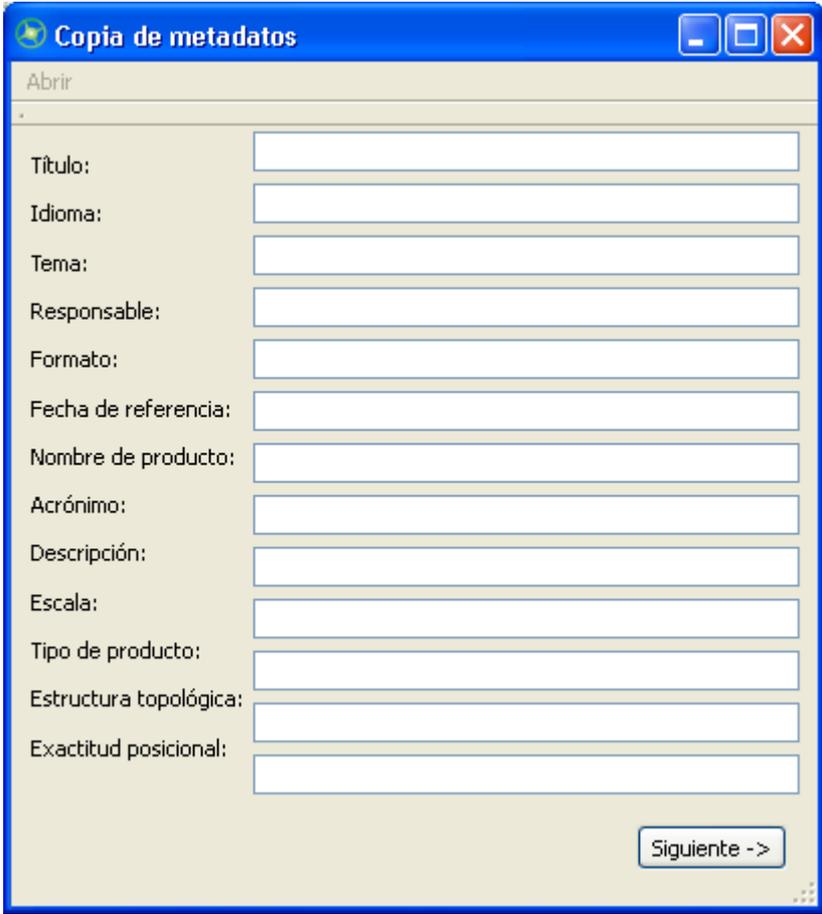
contratos no será necesario dedicar más que unos pocos recursos para actualizar numéricamente las variaciones con respecto a los nuevos censos del INE.

8. GENERACIÓN DE METADATOS

Si se describen correctamente los datos y se obtiene un conjunto de metadatos acordes con las normas establecidas, nos encontramos con un sistema que puede ser tratado de manera más sencilla, más eficiente y con un menor uso de los recursos existentes.

En el caso de que se desee realizar algo más específico, como mantener esos metadatos en bases de datos, es posible el desarrollo de aplicaciones que trabajen directamente en XML (si se trabaja con ISO 19115) y que creen, actualicen y eliminen esos datos automáticamente. Diversos gestores de bases de datos permiten el uso de XML en sus llamadas o si se desea implementar por software, diversos lenguajes de programación tienen la capacidad de trabajar con objetos XML que facilitan su uso. Este es el caso de la herramienta desarrollada por Sinfogeo para la generación de los metadatos asociados a BCN200, conforme a INSPIRE, a la Infraestructura de Datos Espaciales de España (IDEE) y al Núcleo Español de Metadatos (NEM).

La aplicación, desarrollada en C++, utiliza las librerías QT, el SDK de Nokia y el módulo QtXml.



Metadato	Valor
Título:	<input type="text"/>
Idioma:	<input type="text"/>
Tema:	<input type="text"/>
Responsable:	<input type="text"/>
Formato:	<input type="text"/>
Fecha de referencia:	<input type="text"/>
Nombre de producto:	<input type="text"/>
Acrónimo:	<input type="text"/>
Descripción:	<input type="text"/>
Escala:	<input type="text"/>
Tipo de producto:	<input type="text"/>
Estructura topológica:	<input type="text"/>
Exactitud posicional:	<input type="text"/>

Siguiete ->

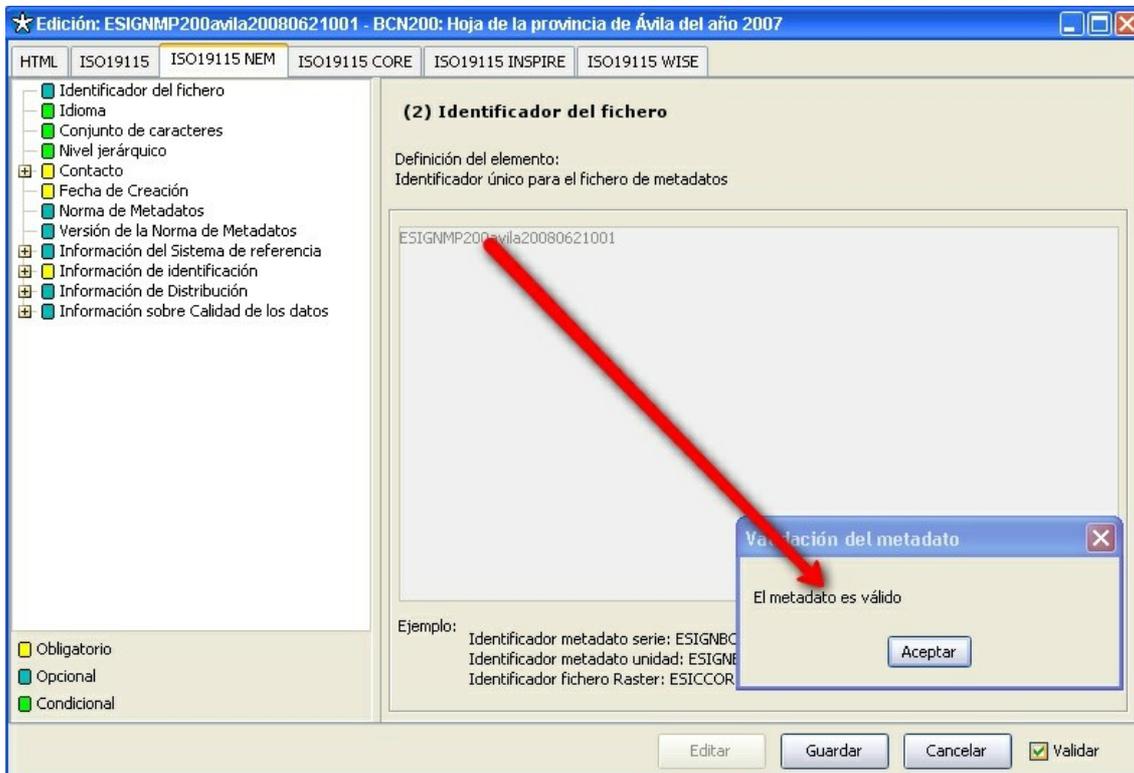
Fuente: Elaboración propia

Algoritmo: La aplicación 'lee', interpreta y presenta la información proveniente de un fichero XML previamente formateado y validado con CatMDEdit, que será el metadato plantilla. Evidentemente tendremos información que:

- Variará con cada unidad cartográfica y que habrá que rellenar adecuadamente de forma manual, como el título y las fechas.

- Variará con cada unidad cartográfica y se podrá rellenar de forma automática, como el boundary o el número y tipo de entidades.
- Será fija para todos los elementos de la Serie y no habrá que modificar nada, como la descripción de la exactitud posicional.

Una vez guardado el nuevo metadato en formato XML, es posible abrirlo con la herramienta CatMDEdit y validarlo para comprobar que es correcto:



9. HACIA BCN100 (ACTUALMENTE BASE100)

BCN100 es un modelo de BCN200 ampliado, que da cabida a aquellas clases de entidad que o bien han adquirido sentido al ampliarse la escala a 100.000 o bien tienen una necesidad fundamental para uso militar.

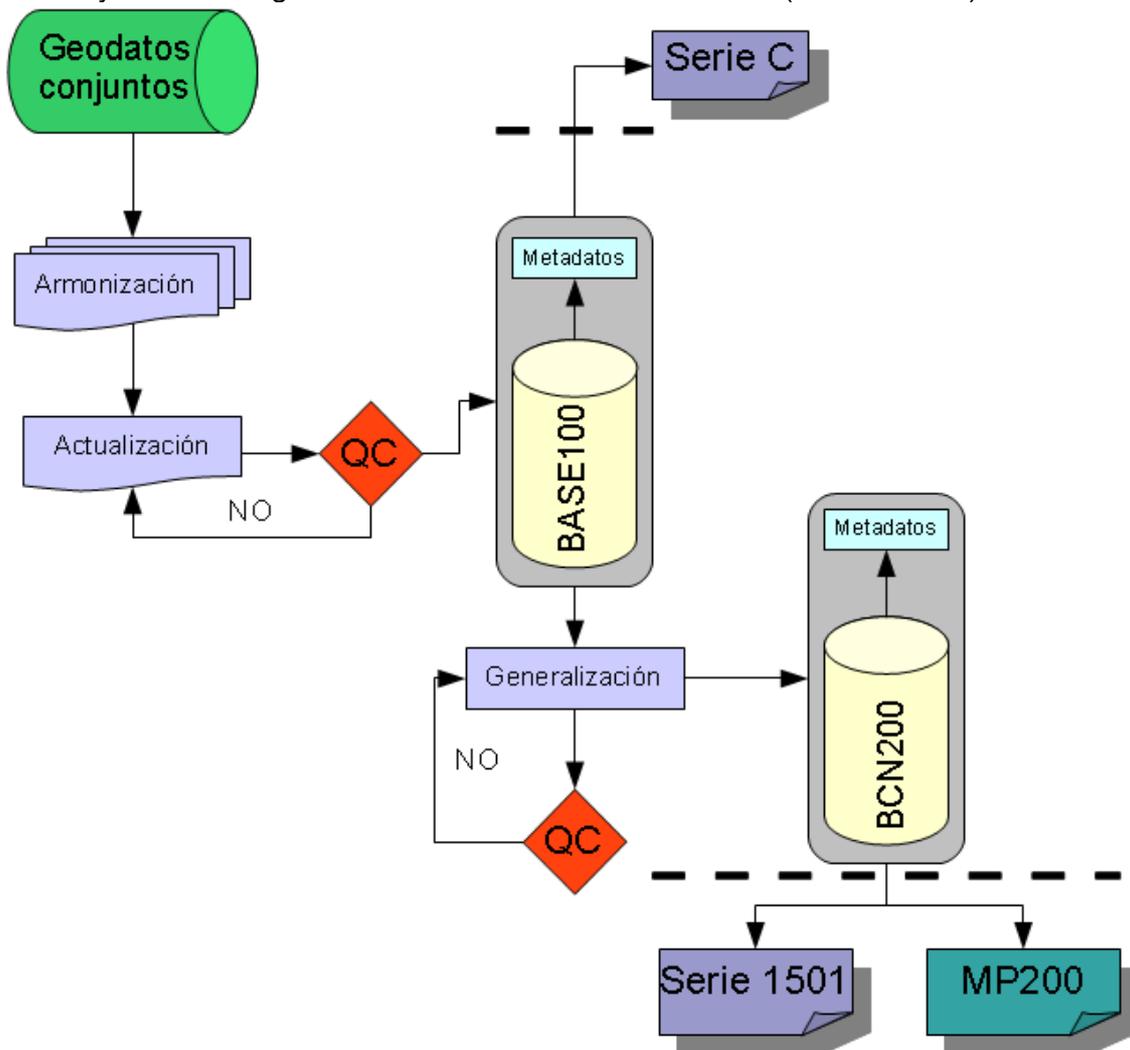
Precisamente este producto denominado inicialmente BASE100, es el resultado del convenio de colaboración entre IGN y Centro Geográfico del Ejército (CEGET), donde cada organismo aportará una serie de entidades cartográficas para disponer finalmente de dos productos:

1. Mapa de mando a escala 1:100.000, con distribución por hojas
2. BCN200 a escala 1:200.000, con distribución por provincias

Ambos, tendrán el modelo de datos de BCN200, por lo que para las entidades aportadas por el CEGET se realizará un cambio de esquema así como una adecuación geométrica.

Una de las grandes diferencias de este proyecto (BASE100) con respecto a los anteriores (BCN200) es el volumen de entidades a tratar. Si hasta hace muy poco sólo se trataban los Temas 5 (Edificaciones y Construcciones) y Tema 6 (Comunicaciones), ahora se deben tratar también temas como la Hidrografía (Tema 3), Usos del suelo (Tema 4), Altimetría (Tema 2) y Conducciones (Tema 7). El problema aquí no será tanto el tiempo dedicado a captar dichas entidades, sino los

recursos que se emplearán para llevar a coincidencia elementos de los diferentes Temas y a cómo las geometrías interactúan unas con otras (armonización).



Fuente: Documentación de proyecto BASE100

El nacimiento de este producto bajo la BCN200, demuestra la continuidad del modelo de datos por parte del IGN y justifica los recursos dedicados en las mejoras introducidas en estos 4 últimos años.

10. CONCLUSIÓN

El trabajo realizado por los organismos implicados en la mejora de la BCN200 ha sido colosal. El método utilizado siempre ha sido de consenso técnico entre las partes responsables, pensando no sólo en la ejecución de los contratos, sino en la viabilidad futura del proyecto, prevaleciendo el criterio más adecuado para el producto.

Con este objetivo en mente, se han dedicado muchos recursos a analizar, informar, modificar, rehacer, desarrollar y un sin fin de actuaciones que han ido mucho más allá de la paciencia de otros proyectos.

Este esfuerzo tiene su resultado en el actual modelo de datos y flujo productivo de BCN200, que a buen seguro tendrá repercusión en otros productos del IGN, como ahora lo empieza a tener en BASE100.

11. AGRADECIMIENTOS

En el ámbito de este artículo, los autores agradecen especialmente a Jose Antonio Merino Martín su actitud por permitir demostrarle las ventajas del software libre aplicado a BCN200 y por tener predisposición a la mejora continua.

Así mismo, en el ámbito del proyecto BCN200, los autores agradecen a todas las partes implicadas del IGN, por la profesionalidad demostrada y por compartir su ilusión en el proyecto con nosotros.

12. REFERENCIAS

Celada Pérez, Jesús. 2007. Producción de Cartoimágenes en entornos GIS: Provincia de Huelva.

<http://www.intergraph.com/global/es/assets/local/ProduccionDeCartoimagenesIGN.pdf>.

Edler, J., L. Tsipouri, L. Hommen, y J. Rigby. 2005. Innovation and Public Procurement – Review of Issues at Stake. *ISI Fraunhofer Institute Systems and Innovation Research, Karlsruhe*.

ESRI. 1998. ESRI Shapefile Technical Description.

<http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf>.

———. 2006. *ArcGIS Data Interoperability Tutorial*.

http://webhelp.esri.com/arcgisSDEsktop/9.2/pdf/DataInterop_Tutorial.pdf.

Flanagin, Andrew, y Miriam Metzger. 2008. The credibility of volunteered geographic information. *GeoJournal* 72, no. 3 (Agosto 1): 137-148.

<http://dx.doi.org/10.1007/s10708-008-9188-y>.

Germán Carrillo. 2009. GeoTux | Construcción de un visor de Shapefiles con herramientas libres: QGIS, Python y Qt. http://geotux.tuxfamily.org/index.php?option=com_myblog&show=construcci%F3n-de-un-visor-de-shapefiles-con-herramientas-libres-qgis-python-y-qt.html&Itemid=59.

gvSIG.org. 2009. extMDB. <http://www.gvsig.org/web/plugins/downloads/extmdb-1.0>.

Haklay, M. 2008. How good is Volunteered Geographical Information? A comparative study of OpenStreetMap and Ordnance Survey datasets.

INTERGRAPH. 2008. Comando de Geomedia para Control de Calidad SIOSE.

http://www.ign.es/siose/Herramientas_y_aplicaciones/Documentacion_Herramientas/Manual_de_la_aplicacion_SIOSE-Modulo_GEOMEDIA/Ayuda_CC_SIOSE_v4.5.pdf.

Merino Martín, Jose Antonio; Gullón Muñoz-Repiso, Tania; Ruiz Ramírez, Ángela del Carmen; González Matesanz, Francisco Javier 2010. BCN200 del IGN-E: base cartografica multipropósito.

http://www.01.idee.es/resources/presentaciones/JIIDE10/ID451_BCN200_del_IGN_E_base_cartografica_multiproposito.pdf

Núñez Maderal, Eduardo. 2007. Mapa Espana Escala 1500.000 con GeoMedia Profesional 6.0 presented at the Reunión de usuarios de Intergraph.

<http://www.intergraph.com/global/es/assets/local/MapaEspanaEscala1500.000ConGeoMediaProfesional60IGN.pdf>.

PyQGIS. 2009. PyQGIS Developer Cookbook — PyQGIS v documentation.
<http://www.qgis.org/pyqgis-cookbook/>.

SAFE Software. 2011. Tight Integration with ArcGIS. <http://www.safe.com/solutions/for-products/esri-arcgis/>.

Serrano, R. M, y E. V Valcarce. 2000. *Técnicas econométricas para el tratamiento de datos espaciales: La econometría espacial*. Edicions Universitat Barcelona.

Wikipedia. GNU General Public License.
http://es.wikipedia.org/wiki/GNU_General_Public_License.