

# Aplicación del estándar ISO 19139 a un modelo relacional de capa, tablas y campos

Alaitz Zabala<sup>1</sup> y Joan Masó<sup>2</sup>

<sup>1</sup> Dept. Geografia, Edifici B, Univ. Autònoma de Barcelona, 08193 Bellaterra  
a.zabala@miramon.uab.es

<sup>2</sup> Centre de Recerca Ecològica i Aplicacions Forestals, Edifici C,  
Univ. Autònoma de Barcelona, 08193 Bellaterra  
joan.maso@uab.es

**Abstract.** El estándar de metadatos ISO 19115:2003 aplicado a una capa vectorial no cubre la documentación de los atributos temáticos de una base de datos relacional (dado que no describe tablas, campos y relaciones entre tablas). Se revisan las opciones existentes y se propone un esquema de aplicación basado en el reciente pre-estándar ISO 19139 que sugiere el uso del *General Feature Model* (GFM, descrito en el pre-estándar ISO 19109) para la descripción completa de este tipo de datos.

## 1 Introducción: Aproximaciones

Los metadatos son la base para el intercambio, la catalogación y la búsqueda de información geospacial. El estándar ISO 19115:2003 [5] define qué informaciones forman parte de los metadatos de la información geográfica.

El objetivo del trabajo es describir, usando los estándares existentes (aprobados o pre-estándares) los metadatos necesarios para describir completamente una capa vectorial (o una geodatabase) asociada a una bases de datos relacional que contiene los atributos temáticos de las entidades vectoriales (véase Fig. 1).

### 1.1 ISO 19115:2003

El estándar ISO 19115:2003 [5] es un modelo conceptual que permite dos aproximaciones para describir los metadatos de las entidades y los atributos. En primer lugar es posible que los metadatos pueden describir información geográfica a diferente nivel jerárquico, particularmente serie, capa (*dataset*), y también tipos de entidades (*feature type*), tipos de atributos (*attribute type*), entidades (*feature instances*) y atributos (*attribute instances*). Esta vía nos ‘obligaría’ por ejemplo a generar un conjunto completo de metadatos para cada tipo de atributo o atributo a describir.

En segundo lugar se puede referenciar la descripción de las entidades y atributos realizada en una catálogo de entidades (*feature catalog*). Esta aproximación proporciona la cita a un catálogo que describe las entidades y atributos y algunos otros me-

tadatos como el idioma del catálogo de entidades o la lista de entidades del catálogo representadas en la capa (opcional). También se puede indicar si este catálogo de entidades es conforme al pre-estándar ISO 19110 *Methodology for feature cataloguing* [3, p. 69] (recomendable).

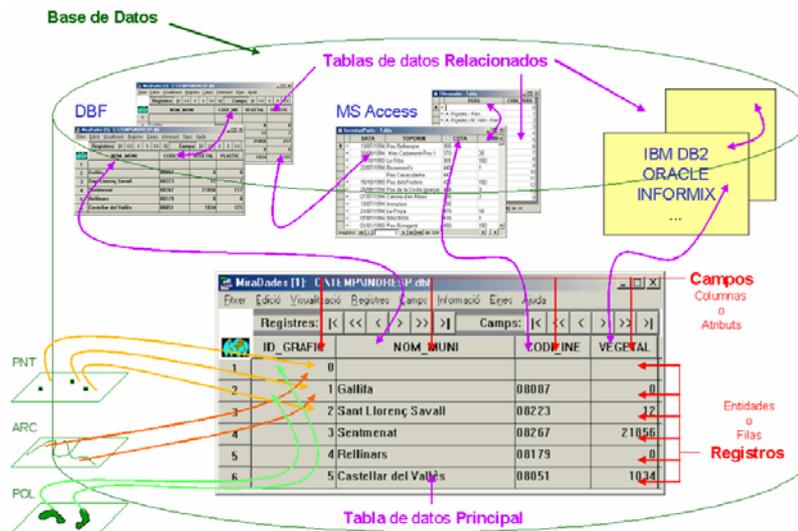


Fig. 1. Representación gráfica del modelo de datos de una capa vectorial asociada a una base de datos relacional (Fuente: elaboración propia)

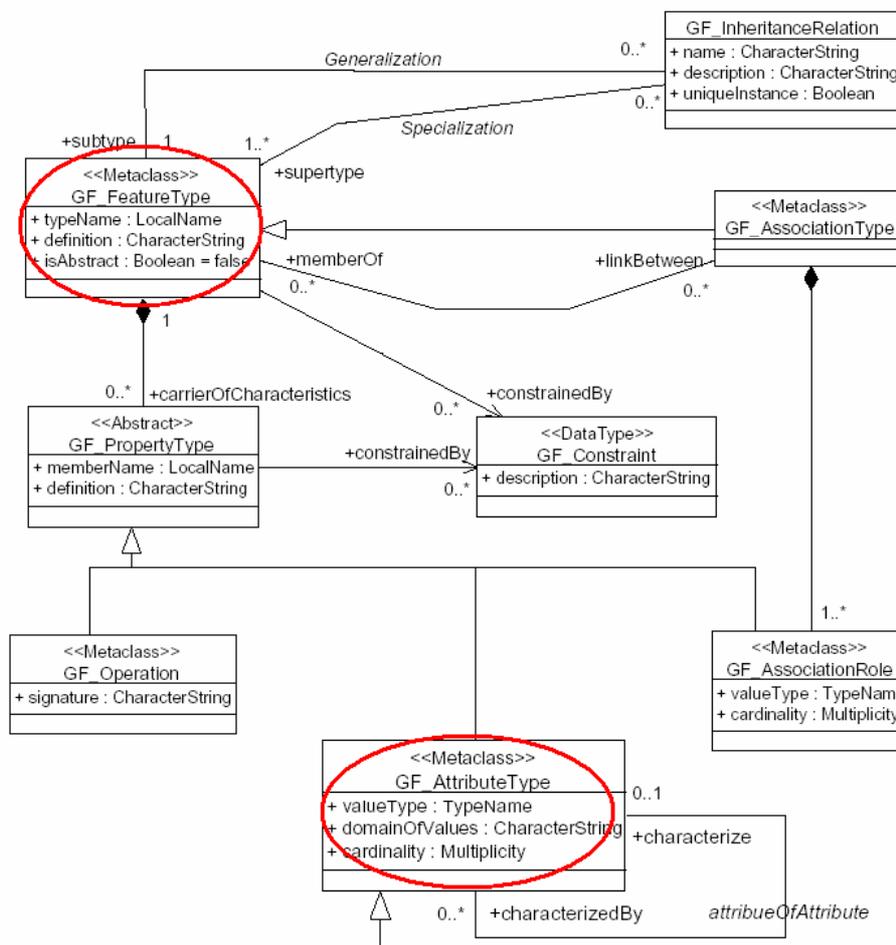
## 1.2 ISO 19109 y ISO 19110

El estándar ISO 19115:2003 [5] no define por sí mismo cómo se debe llevar a la práctica la descripción de las entidades ni del conjunto de atributos (y sus relaciones) vinculadas con una capa (*dataset*).

La descripción general de las entidades y sus atributos se realiza en el pre-estándar ISO 19109 *Rules for application schema* [4], que incluye el *General Feature Model* (GFM). Este pre-estándar define las reglas para crear y describir un 'esquema de aplicación' y incluye los principios para la definición de las entidades y los atributos (véase Fig. 2). Los esquemas de aplicación proporcionan la descripción formal de la estructura de los datos y del contenido requerido por una o más aplicaciones.

El pre-estándar ISO 19110 *Methodology for feature cataloguing* [3] define la metodología para la catalogación de tipos de entidades y utiliza conceptos que son realizaciones del GFM (es una implementación sobre este modelo general, véase Fig. 3).

Existen diferencias entre el modelo general descrito por ISO 19109 y el modelo implementado en ISO 19110. Particularmente la definición de atributos de atributos, que aparece en el modelo general, no aparece en el modelo del catálogo de entidades. Esta característica es importante para describir totalmente los atributos temáticos de una capa vectorial, contenidos en una base de datos relacional.



**Fig. 2.** Fragmento del *General Feature Model* (ISO 19109)  
 (Fuente: elaboración propia a partir de diagramas UML del pre-estándar ISO 19109)

### 1.3 CSDGM – FGDC

El estándar norteamericano de metadatos *Content Standard for Digital Geospatial Metadata* (CSDGM, desarrollado por FGDC [1]), anterior a la serie de estándares ISO 19100, considera en la sección 5. *Entity and Attribute Information* como describir entidades y sus atributos.

El elemento ‘*attribute of attribute value*’ del estándar permite definir información adicional para una entidad con un valor determinado en un atributo. Por ejemplo una entidad ‘pozo’ puede tener un atributo ‘producto’. Si el valor del atributo ‘producto’

es 'agua', entonces este atributo puede tener otros atributos que determinen las características del agua (pH, salinidad, etc.).

Esta aproximación es más cercana al problema considerado aquí que la aproximación de ISO 19115:2003 [5] pero de todas maneras continua sin permitir describir completamente las tablas, campos y relaciones entre tablas del modelo relacional. En particular no dispone del concepto equivalente al de 'tabla'.

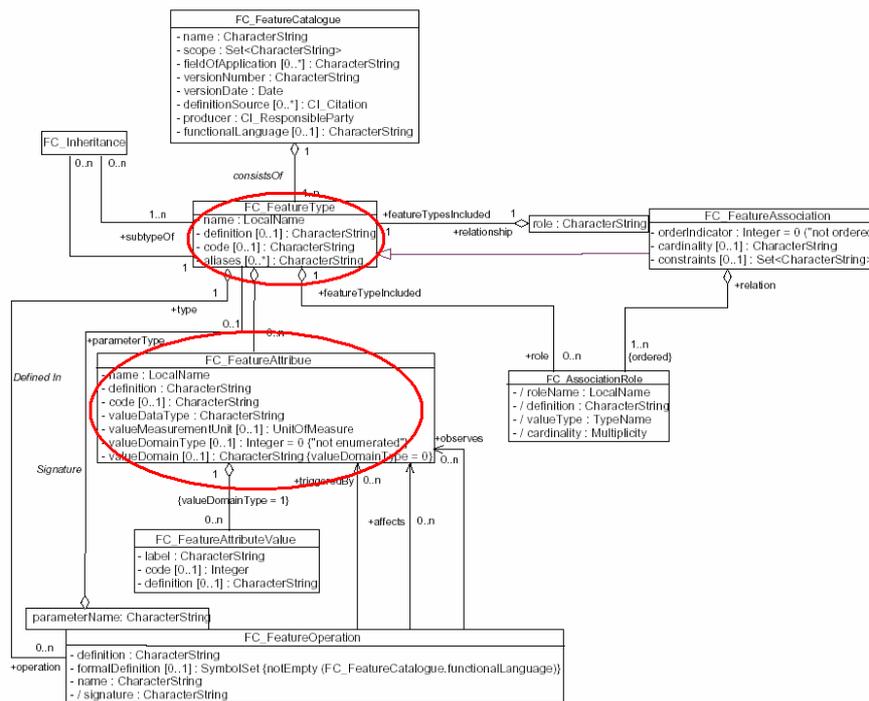


Fig. 3. Modelo conceptual del catalogo de entidades (ISO 19110)  
(Fuente: pre-estándar ISO 19110)

#### 1.4 Propuesta de FGDC-to-ISO Metadata Crosswalk

El documento *FGDC-to-ISO Metadata Crosswalk* [2] propone equivalencias entre los elementos del estándar CSDGM de FGDC [1] y los de ISO 19115:2003 [5] (y relacionados), equiparando los elementos de la sección 5. *Entity and Attribute Information* (de CSDGM) con los elementos de la descripción del catálogo de entidades ISO 19110 [3].

Este documento no indica a qué elemento de la serie de estándares ISO se debe equiparar el elemento 'attribute of attribute value' del estándar norteamericano. Esta omisión puede ser debida al hecho de que el 'crosswalk' se basa en 19110, que no considera el elemento 'attributeOfAttribute' del modelo GFM (ISO 19109 [4]).

## 1.5 Propuesta de ISO 19139 (v.1.0)

El pre-estándar ISO 19139 [6] es una especificación técnica que todavía no está aprobada aunque cuenta con el respaldo de la mayoría de los países miembros de ISO según una reciente votación. Pretende desarrollar una implementación en XML del modelo de metadatos descrito por ISO 19115:2003 [5] y, como tal, define un conjunto de esquemas XML que declaran elementos XML para describir los metadatos de una capa (ficheros xsd).

Sin embargo, ISO 19139 incorpora elementos que no estaban presentes en el modelo abstracto de metadatos ISO 19115:2003. En relación al tema que nos ocupa es importante destacar que ISO 19139 incorpora directamente las descripciones de entidades (*featureType*) y atributos (*attributeType*) basados en el esquema GFM (ISO 19109) y que se sitúan como elementos del elemento compuesto MD\_Metadata.

Siguiendo esta propuesta, para cada conjunto de metadatos que describe una capa es posible describir las entidades y los atributos que forman esta capa. Además incorpora de manera natural el concepto, basado en ISO 19109, que permite describir atributos de atributos, aspecto esencial que nos permite en el presente artículo realizar una propuesta para describir completamente las tablas, campos y relaciones de la base de datos relacional.

## 2 Modelo de bases de datos relacional (DBRM)

Este apartado propone la descripción de la base de datos relacional que contiene los atributos temáticos de una capa vectorial, a partir de las herramientas proporcionadas por el pre-estándar ISO 19139 [6].

### 2.1 Entidad vectorial, campos de la primera tabla asociada

A partir de un identificador de entidad vectorial es posible obtener los campos que aportan atributos temáticos a esta entidad. Los valores de estos atributos temáticos se pueden describir más extensamente en otras tablas que se vinculan o relacionan a esta tabla aportando otros atributos. Cuando existe una relación con otra tabla, uno o diversos campos se enlazan con esta otra tabla según el valor del campo o campos en los que se inicia la relación, y se aportan otros atributos a la capa (véase Fig. 1).

Se puede asimilar la entidad vectorial al tipo de entidad GF\_FeatureType definida en el modelo GFM y concretamente al elemento XML *<featureType>* definido en el esquema XML *MetadataEntity.xsd*.

Tal y como el modelo general GFM contempla, para cada entidad es posible definir uno o más atributos. En el esquema XML esto se concreta con uno o más elementos *<carrierOfCharacteristics>* que contienen elementos de tipo *<attributeType>* (en este caso) que describen cada uno de los campos de la primera tabla de la base de datos relacional.

## 2.2 Campos de las tablas relacionadas y ‘attributeOfAttribute’

La diferencia entre el modelo de relacional de tablas y campos y el modelo de ‘atributo de atributo’ definido por CSDGM (FGDC) radica en que, en el modelo relacional, los campos asociados a una tabla se definen para cualquier valor del campo (atributo) original, aunque, evidentemente para algunos valores no existirá ningún registro en la tabla asociada y por lo tanto el valor de estos ‘atributos de atributos’ quedará indefinido.

El concepto de ‘atributo de atributo’ de CSDGM aplicado al modelo relacional de bases de datos es el mismo concepto de ‘atributo de atributo’ que describe el GFM (ISO 19109), y por tanto, el que usa ISO 19139.

En el GFM un atributo de un atributo (véase Fig. 2) es un vínculo entre dos atributos (no entre un valor de atributo y un atributo, como para CSDGM) uno de los cuales define alguna característica del otro. Por ejemplo si un atributo contiene la posición de una entidad, un ‘atributo de este atributo’ podría contener la calidad de esta posición o las unidades del parámetro de calidad cuantitativo que define la calidad del posicionamiento de la entidad.

Por lo tanto se pueden considerar ‘atributos de atributos’ tanto los campos (atributos) de tablas relacionadas como los atributos que definen otras características del campo (nombre del campo, descripción, tipo de valores, unidades, calidad, etc.).

La implementación de este concepto se concreta en el elemento XML `<characterizedBy>` que forma parte del elemento `<attributeType>` y que a su vez (y en este caso) contiene elementos de tipo `<attributeType>` que permiten describir cada uno de los campos de las tablas asociadas de la base de datos relacional. Este elemento presenta cardinalidad múltiple.

Aunque se ha comentado que estos atributos de atributos se pueden utilizar para describir también características de campo, se debe considerar que los elementos `<attributeType>` están formados por otros elementos (además de `<characterizedBy>`) que nos permiten describir de manera directa algunas de las características interesantes de los campos, por ejemplo el nombre, descripción, tipo de valores, dominio o cardinalidad. Aquellas características sin un lugar específico dentro de los elementos que forman `<attributeType>` se pueden siempre añadir como ‘atributos de atributos’.

El elemento `<valueType>` se utiliza para definir el tipo de atributo en cada caso. En nuestra implementación, todos los elementos `<attributeType>` que describan campos tienen el elemento `<valueType>` igual a “DBFIELD”. De la misma manera, usamos un atributo de atributo para definir las unidades de los campos, usando siempre un `<valueType>` igual a “DBUNITS”.

## 2.3 Descripción de las tablas relacionadas

Hasta ahora se ha relacionado el modelo de datos relacional y el modelo del GFM (ISO 19109) y se ha presentado la manera de describir los campos asociados del modelo relacional de bases de datos siguiendo las directrices del pre-estándar ISO 19109 y ISO 19139.

También es interesante poder describir las “tablas” relacionadas de manera directa (y no sólo a través de sus campos) puesto que puede ser necesario conocer de qué

tabla proviene cada campo, así como algunas características de la tabla: tipo de tabla, origen de los datos, código de caracteres, etc.

Es posible considerar los campos asociados de una misma tabla de atributos temáticos de forma conjunta, de manera que se describe el 'registro' o 'conjunto de registros' como un todo, y posteriormente se describen los campos de este registro. Este concepto es similar al concepto de ROWSET usado en SQL.

Así, se utiliza un atributo para describir el registro (es decir, toda la tabla) y luego se caracteriza este atributo con otros atributos que son, por un lado, los campos que la forman y por otro, algunos otros elementos interesantes (tipo de datos, etc.). En este caso también es necesario utilizar un *<valueType>* único y conocido para cada atributo particular, en nuestra propuesta se usa 'DBROWSET' para definir que el atributo describe registros de una tabla.

## 2.4 Ejemplo

A continuación se expone como se describe una capa vectorial de puntos con una base de datos de atributos temáticos asociada usando la propuesta presentada (Fig. 4).

The image shows a GIS software interface. On the left, a tree view displays the structure of a vector layer named 'Capa vectorial de puntos: CATIEFC'. The tree includes a main layer 'CATIEFC.dbf' with an internal graphic identifier, and several attribute tables: 'Altitud', 'Orientación', 'Pendiente', 'Código municipal', 'municipi.dbf' (containing 'Nombre del municipio', 'Código INE del municipio', 'Código comarcal', and 'Código provincial'), 'Censo91.dbf' (containing 'Código municipal' and 'Núm. de habitantes (censo 1991)'). A red arrow points from the text 'Objeto gráfico, entidad vectorial' to the 'Capa vectorial de puntos: CATIEFC' layer name.

On the right, a metadata window titled 'Inventario Ecológico y Forestal de Catalunya' provides details for a specific record. It shows the file path 'C:\Mapes\IEFC\CATIEFC.REL.pnt', the element ID 'Elemento gráfico 1 de 10583 (id=0)', and notes 'Encontrados 1 registros. Éste es el núm 1.'. Below this, it lists 'Datos la estación de muestro' (Altitud (m): 1160, Orientación (?): 50, Pendiente (?): 30) and 'Datos del municipio en el que está situada' (Nombre del municipio: Albanyà, Código INE del municipio: 17003, Núm. de habitantes (censo 1991): 109).

Fig. 4: Capa vectorial de puntos con base de datos de atributos temáticos

En primer lugar, existe un elemento XML *<featureType>* que es hijo del elemento general *<MD\_Metadata>* (véase Fig. 5). Este elemento se usa para representar la entidad o entidades vectoriales que forman la capa que se describe.

```

<?xml version="1.0" encoding="ISO-8859-1"?>
<smXML:MD_Metadata xmlns:smXML="http://www.isotc211.org/smXML" xmlns:asXML="http://www.isotc211.org/asXML"
xmlns:scXML="http://www.isotc211.org/scXML" xmlns:xlink="http://www.w3.org/1999/xlink" xmlns:xsi="
http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.isotc211.org/smXML
C:\XML_MD\ISO19139\smXML\metadataEntity.xsd">
  <fileIdentifier>
  <language>
  <contact>
  <contact>
  <contact>
  <dateStamp>
  <metadataStandardName>
  <metadataStandardVersion>
  <dataSet>
  <featureType>
  <identificationInfo>
  <dataQualityInfo>
  <referenceSystemInfo>
  <has>
</smXML:MD_Metadata>

```

Objeto gráfico, entidad vectorial

Fig. 5: Descripción en el documento XML de las entidades vectoriales que forman la capa

El primer y único atributo que complementa a esta entidad (*<featureType>*) es un atributo (*<asXML:GF\_AttributeType>*) de tipo 'DBROWSET' que define la primera tabla de la base de datos con atributos temáticos (Fig. 6).

Fig. 6: Descripción del primer conjunto de registros 'DBROWSET' (o tabla)

Este atributo de tipo 'DBROWSET' contiene como mínimo tantos *atributos de atributos* (elementos *<characterizedBy>* hijos de *<asXML:GF\_AttributeType>*) como campos forman el 'DBROWSET' (cinco en el ejemplo presentado, aunque en el documento XML se muestra parcialmente).

Cada uno de estos elementos *<characterizedBy>* describe uno de los campos del DBROWSET (o tabla, véase Fig. 7). En caso de que alguno de estos campos tenga una relación con algún otro DBROWSET, se añade un *atributo de este atributo* (el campo) que contiene la descripción del DBROWSET (o tabla) relacionado. En el ejemplo, el último campo del primer DBROWSET (campo 'Código municipal') tiene

relaciones con dos tablas (o DBROWSETs) y por lo tanto presenta dos atributos de atributos (*<characterizedBy>*) que los describen (véase Fig. 7, anotación naranja). A su vez cada uno de estos DBROWSET contiene la descripción de los campos que lo forman.

The image shows a screenshot of a GIS software interface. On the left, a tree view displays a vector layer named 'Capa vectorial de puntos: CATIEFC'. Underneath, there are several fields listed, including 'Identificador gráfico interno', 'Altitud', 'Orientación', 'Pendiente', 'Código municipal', 'municipi.dbf', 'Nombre del municipio', 'Código INE del municipio', 'Código comarcal', 'Código provincial', 'Censo91.dbf', and 'Código municipal'. A green arrow points to the 'Código municipal' field, with a green annotation 'Un campo del DBROWSET'. To the right, an XML snippet is shown, describing a 'DBFIELD' with 'memberName' 'CODI\_MUN' and 'valueType' 'DBFIELD'. A red box highlights the 'characterizes' elements, with an orange annotation 'DBROWSETS relacionados con este DBFIELD'.

Fig. 7: Descripción del último campo (DBFIELD, Código Municipal) del primer DBROWSET

## 2.5 Consideraciones finales

Se pueden usar los ‘atributos de atributos’ para otras finalidades. En primer lugar, de momento no ha sido explicado cómo indicar el campo de la tabla relacionada que se vincula con el campo origen de la relación. Para realizar esta función se utiliza otro atributo de tipo definido, concretamente “DBFIELD\_LINK\_TARGET”. Este atributo se utiliza para describir, a nivel de ‘DBROWSET’, el nombre del campo de la tabla destino que forma el enlace.

En el caso de relaciones entre tablas descritas por que utilizan un campo clave combinado no tiene mucho sentido atribuir el DBROWSET destino de la relación a uno de los campos clave (¿a cuál de ellos?). Por este motivo se propone usar un atributo de tipo DBLINK para definir la relación entre ambos DBROWSET’s. De esta manera un DBROWSET tiene un atributo DBLINK. Éste, a su vez, tiene un atributo DBROWSET (tabla destino de la relación) además de dos o más atributos del tipo DBFIELD\_LINK\_SOURCE para describir cuales de los campos del DBROWSET original utilizados para enlazar con el DBROWSET destino. La definición del campo o campos de enlace en el DBROWSET destino se realiza como atributo del DBROWSET destino, como en el caso relación más simple. El caso de relación simple (campo clave único) se podría definir también siguiendo este esquema aunque se recomienda la anteriormente descrita porque resulta más intuitiva.

En caso que una tabla sea referenciada desde más de un campo, la tabla se describe una sola vez atribuyéndole un identificador que es usado posteriormente para indicar los otros vínculos a la misma tabla evitando redundancias y circularidades.

De la misma manera que ISO 19139 propone crear registros de unidades y de valores de CodeList, las diferentes comunidades de usuarios podrían también describir un registro de valores de *valueType* con el listado de valores ‘especiales’ a usar en este elemento XML para la descripción de campos, tablas y relaciones. Los valores usados en esta implementación son:

- **DBROWSET**: tipo de atributo que define el conjunto de registros de una tabla. Puede definir (ser ‘hijo’ de) a una entidad (*featureType*) o a un atributo de tipo ‘DBFIELD’ o ‘DBLINK’ (relación entre dos DBROWSETs, campo clave combinado)
- **DBFIELD**: tipo de atributo que define las características de un campo de la base de datos. Se usa para definir los campos de un atributo de tipo ‘DBROWSET’
- **DBLINK**: tipo de atributo que define las características de un enlace entre dos DBROWSETs a partir de un campo clave combinado. Puede definir a un DBROWSET.
- **DBFIELD\_LINK\_SOURCE**: tipo de atributo que define los campos del DBROWSET original que actúan de campos enlace en una relación a partir de un campo clave combinado. Puede definir a un atributo de tipo ‘DBLINK’
- **DBFIELD\_LINK\_TARGET**: tipo de atributo que define los campos del DBROWSET destino que actúan de campos enlace en una relación a partir de un campo clave combinado. Puede definir a un atributo de tipo ‘DBROWSET’

Un lector atento habrá notado la discrepancia entre el elemento XML denominado *<characterizedBy>* en el texto y escrito como *<characterizes>* en los XML mostrados en las figuras. Los esquemas XML proporcionados el elemento con cardinalidad múltiple es *<characterizes>* en vez de *<characterizedBy>* aunque nuestra interpretación del pre-estándar ISO 19139 nos lleva la conclusión contraria.

Este trabajo es fruto de un convenio de colaboración entre el *Centre de Recerca Ecològica i Aplicacions Forestals* (CREAF), la *Infraestructura de Dades Espacials de Catalunya* (IDEC) y miembros del FGDC.

## Bibliografía

1. Federal Geographic Data Committee: Content Standard for Digital Geospatial Metadata. CSDGM Version 2: FGDC-STD-001-1998. Washington (1998)
2. Federal Geographic Data Committee: ISO-FGDC-METADATA-CROSSWALK-V4 (2003)
3. International Organization for Standardization: Draft International Standard: Geographic information – Methodology for feature cataloguing. ISO/DIS 19110. Technical Committee 211 (2001)
4. International Organization for Standardization: Draft International Standard: Geographic information – Rules for application schema. ISO/DIS 19109. Technical Committee 211 (2002)
5. International Organization for Standardization: International Standard: Geographic information – Metadata. ISO 19115:2003. Technical Committee 211 (2003)
6. International Organization for Standardization: Draft Technical Specification: Geographic information – Metadata – XML schema implementation. ISO/PTDS 19139. Technical Committee 211 (2004)