

### 3 Información geográfica

*Bengt Rystedt, Suecia*

#### 3.1 Introducción

Con la expresión «información geográfica» nos referimos a información que tiene una ubicación geográfica. La ubicación debe suministrarse en forma matemática, de tal manera que pueda ser utilizada en un ordenador. Lo más útil es emplear la latitud y la longitud. El concepto de «ubicación» se describe con más detalle en el próximo capítulo. Una manera fácil de describir cómo se trabaja en los ordenadores con la información geográfica es pensar en «capas» (véase la figura 3.1), es decir, viendo el paisaje estructurado en diferentes capas.

A continuación, podemos comenzar a trabajar con las capas topográficas, es decir con aquellas capas para las áreas administrativas, carreteras, lagos, ríos, etc. respectivamente. Los diferentes datos temáticos que describen la geología, el uso del suelo y la vegetación pueden constituir, por ejemplo, otras capas adicionales. En la figura 3.1, se ve el principio de un modelo digital del paisaje basado en diferentes capas. Esta idea y forma de organizar los datos geográficos se introdujo por primera vez en Canadá en la década de 1960, cuando se elaboró el «Inventario de tierras de Canadá» (*Canada Land Inventory*) como herramienta fundamental para todo tipo de planificación territorial y gestión de los recursos nacionales.

Las capas, además de proporcionar datos de las dimensiones geográficas, deben también contar con los datos de los atributos, que se almacenan en tablas relacionales. Un área en una capa y los datos de sus atributos se conectan a través de un identificador único que es conocido generalmente como «número de identificación». Se produjo un gran avance en el manejo de la información geográfica y en el análisis geográfico por ordenador cuando Jack Dangermond llegó a la conclusión que se puede gestionar la geometría en una

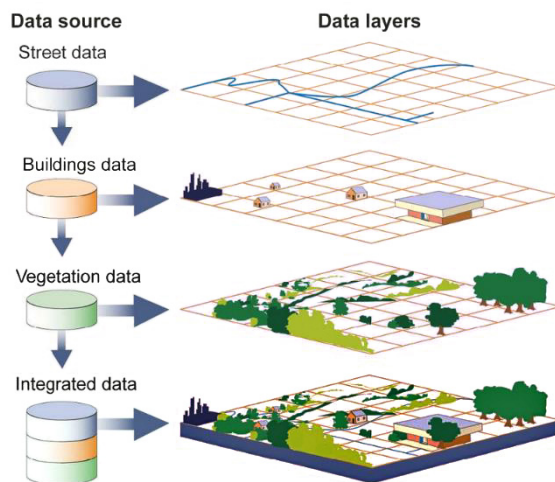


Figura 3.1. La figura muestra el principio de un modelo digital del paisaje. Cada capa contiene tanto los datos de localización como un conjunto de datos de atributos. Fuente: <http://education.nationalgeographic.com/education/photo/new-gis/>.

base de datos y los datos de los atributos en otra base de datos de forma simultánea. Llamó al sistema ARC/INFO: ARC para la geometría e INFO para los datos de atributos en una base de datos relacional. Más tarde, se emplearon muchos otros sistemas.

#### 3.2 Modelado de datos

Antes de que la información geográfica se pueda utilizar tanto para el análisis como para la cartografía, se debe diseñar un modelo de datos geográficos. En la figura 3.1 se presenta uno, en el que se muestra un principio de modelo de datos estructurado en diferentes capas. Lo siguiente es definir todos los objetos que van a ser incluidos. Los objetos se construyen a partir de elementos puntuales, lineales y superficiales.

La parte más importante de un modelo de datos geográficos es su *topología*, que describe cómo los diferentes elementos encajan entre sí para formar redes y estructuras. En una red, tal como un sistema de carreteras, los puntos de las intersecciones se denominan «nodos» y la topología informa de qué vías (carreteras, caminos, etc.) están conectadas a cada nodo, como se muestra en la figura 3.2.

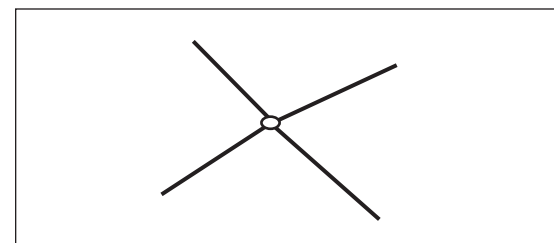


Figura 3.2. Principio de una estructura de carreteras con un nodo en el centro y cuatro carreteras conectadas. Los nodos y las carreteras deben tener una identidad (por ejemplo, un número de identidad que sea fácil de encontrar en una base de datos) y también pueden tener atributos.

En una estructura de áreas, cada área se encuentra rodeada por varias áreas vecinas. Si se sigue el contorno de una de ellas en una determinada dirección, siempre se encontrará un área a la izquierda y un área a la derecha. Cuando se calcula la topología de una estructura de áreas, cada línea se considera dos veces, una para cada dirección, teniendo así un área a la izquierda y un área a la derecha. Eso puede parecer innecesario, pero es realmente necesario para conseguir un sistema que se pueda utilizar para el análisis geográfico en un «Sistema de Información Geográfica» (SIG). La figura 3.3 muestra una estructura de áreas para un municipio con dos distritos.

Una clasificación completa de las áreas administrativas puede ser: *nación, provincia, municipio, distrito y parcela de terreno*. Eso también quiere decir que cada una de esas áreas puede tener a otras como límites y deben ser

introducidas en la base de datos (por ejemplo, en una estructura de datos jerárquica, como se muestra en la figura 3.4). La figura también muestra que las líneas y puntos deberán ser registrados varias veces en la base de datos y que el tamaño de la base de datos crecerá más rápidamente que si se configura de manera lineal.

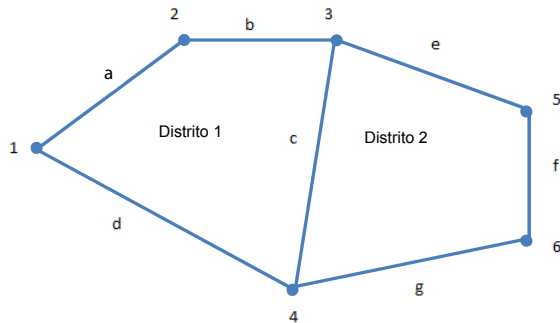


Figura 3.3. Un municipio con dos distritos. Siguiendo los límites de cada distrito en sentido horario, podemos ver que el límite «c» tiene dos direcciones, mientras que los externos solo tienen una dirección.

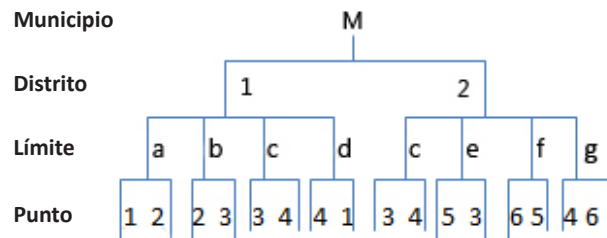


Figura 3.4. Estructura de datos jerárquica para los dos distritos que se muestran en la figura 3.3. También se muestra que el límite «c» está representado dos veces y que mientras que todos los puntos están representados dos veces, los puntos «3» y «4» se representan 4 veces.

Tabla 3.1. Tablas de una base de datos relacional. Las coordenadas X e Y son solo una suposición.

Municipio	Distrito 1	Distrito 2
Nombre del municipio	Nombre del distrito	Nombre del distrito

Superficie	Línea	Línea	Línea	Línea
Distrito 1	a	b	c	d
Distrito 2	c	e	f	g

Límite	Punto	Punto	Distrito 1	Distrito 2
a	1	2	1	
b	2	3	1	
c	3	4	1	2
d	4	1	1	
e	3	5	2	
f	5	6	2	
g	6	4	2	

Punto	Coord. X	Coord. Y	Línea	Línea	Línea
1	80	229	a	d	
2	221	121	a	b	
3	375	119	b	c	e
4	372	295	c	d	g
5	517	127	e	f	
6	544	228	f	g	

Hasta ahora hemos mostrado lo que es una estructura de datos y dado a entender que los datos se deben almacenar en una base de datos. Uno de los tipos más comunes de organización para una base de datos es la base de datos relacional. Eso significa que los datos se manejan en tablas y que las relaciones muestran la conexión entre tablas. Para el ejemplo anterior, se proporciona en la tabla 3.1 una base de datos relacional. El número de columnas vendrá definido según el número de áreas, de líneas, etc. que vayan a existir. Las coordenadas son solo un ejemplo.

### 3.3 Encontrar las coordenadas en una base de datos

Las tablas descritas anteriormente se organizan por orden de identidad de cada ítem. Cada tabla se almacena como un fichero en la base de datos, por lo que será más fácil encontrar un objeto. La cosa se complica con las coordenadas. La coordenada «x» se define a lo largo de la distancia desde el ecuador hacia el polo (Norte o Sur). La coordenada «y» da la distancia Este-Oeste desde el meridiano seleccionado y la proyección que haya sido adoptada (más detalles en el capítulo 9). Es obvio que las coordenadas no se pueden organizar en una tabla. El problema se ha resuelto mediante la organización en *quad-trees* (árboles cuaternarios). En primer lugar, se divide el área en cuatro cuadrados y luego, esos cuatro cuadrados a su vez se dividen en otros cuatro cuadrados, con lo que tenemos 16 cuadrados y así sucesivamente, hasta que solo tenemos un par de coordenadas en cada cuadrado. Utilizamos el sistema binario para dar un identificador a los cuadrados. Después de la primera división, obtenemos los números 00, 01, 10 y 11. Mediante el uso de *quad-trees* es fácil encontrar las coordenadas en una base de datos con solo hacer un clic en la pantalla. No vamos a dar aquí ejemplos de *quad-trees*. Si se desea saber más sobre ellos, se recomienda consultar los trabajos de Worboys y Duckham (2004).

### 3.4 Modelos de información

Una base de datos geográficas debe estar basada en el mundo real y debe diseñarse según sea el tipo de análisis que se requiera. A modo de ejemplo, podemos tomar un sistema para la gestión de un cable de fibra óptica en un determinado barrio. Eso incluirá objetos

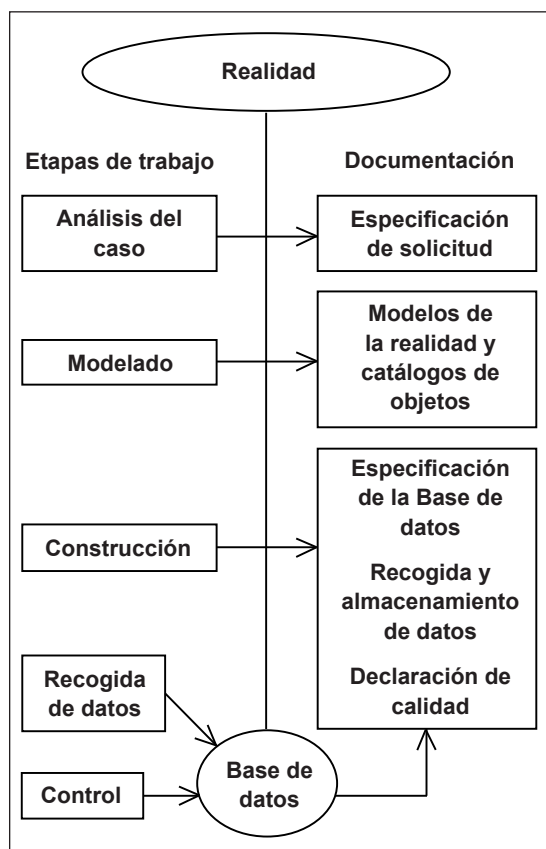


Figura 3.5. Etapas de trabajo en el modelado de información, a la izquierda, y documentos que deben presentarse, a la derecha.

tales como bienes inmuebles, sus propietarios (o usuarios), la ubicación del cable, los acuerdos de gestión y los costes. El análisis debe discutirse con el futuro usuario del sistema. También es importante documentar todo el proceso. Tanto los diferentes pasos que se han de llevar a cabo en el trabajo como la documentación requerida, se muestran en la figura 3.5.

El modelado de información ha sido estudiado intensamente durante los últimos años. Un Comité Técnico de la ISO (el ISO/TC211) ha elaborado y publicado numerosas normas, que han sido utilizadas por todos los productores de datos geográficos. Para nuestro ejemplo con el cable de fibra, se han mencionado algunos tipos de objetos, que deben ser documentados en el catálogo de objetos.

### 3.5 Metadatos y calidad

No vamos a recorrer todas las normas de información geográfica, sino que solo vamos a dar una breve descripción de los metadatos y de la calidad de la información. Los metadatos proporcionan un resumen de qué tipo de datos se incluyen en la base de datos y, además, proporcionan una visión general de cuáles son los datos que pueden ser útiles para una determinada aplicación.

Los metadatos (datos sobre los datos) proporcionan una descripción de la base de datos y pueden incluir:

- El nombre de la base de datos.
- La organización que los gestiona.
- El área geográfica cubierta.
- Una lista del catálogo de objetos.
- El sistema de coordenadas.

- Reglas para la descarga y aplicaciones.
- Costes.

La calidad de los datos es también un tipo de metadatos y pueden incluir:

- *Origen*, que cita las fuentes básicas para los datos, cómo se han recogido y qué organización es la responsable de ello.
- *Exactitud de la ubicación*, que proporciona la exactitud especificada de las coordenadas (tanto sobre el plano, como la elevación).
- *Actualidad*, que da la actualidad de los datos y la planificación de su actualización.
- *Integridad*, que especifica si se incluyen todos los objetos o no, la corrección de la clasificación y si la topología está completa (si, por ejemplo, las vías de comunicación constituyen una red completa).

El mismo tipo de medición de la calidad se puede dar para la información de los atributos incluidos en los objetos. La calidad también puede incluir información sobre el sistema de control de calidad empleado. En conjunto, la información sobre la calidad le dirá al usuario si la base de datos se puede utilizar o es apta para la aplicación deseada.

### 3.6 Recolección de datos

Una base de datos geográficas contiene tanto la geometría como los atributos de los datos. Además, la geometría puede venir dada ya sea en formato vectorial o en formato ráster. El formato vectorial es más natural y más cercano a la geometría que conocemos desde la escuela. Los datos ráster vienen dados sobre una malla

de pequeños cuadrados llamados píxeles y dan una representación no tan detallada de la geometría.

Los datos geográficos pueden recogerse de muchas maneras. La calidad mejor vendría dada por las mediciones de datos en el campo. La digitalización es, no obstante, más común y se digitalizan tanto fotos aéreas como mapas. Los topógrafos han creado sistemas de administración de tierras en los que pueden medir parcelas y utilizar aplicaciones de planificación como, por ejemplo, de ubicación de casas, carreteras y puentes. Las líneas que discurren bajo el suelo (soterradas) como las redes eléctricas, de teléfonos y de alcantarillado deben medirse también. El municipio almacena la ubicación de esas líneas y un empresario que quiera excavar en el terreno tendrá que solicitar un permiso para ello, y recibirá un mapa donde se encuentran dichas líneas junto con el permiso para excavar. Sin embargo, un mapa que incluya todos estos cables y conducciones no es de acceso público, ya que podría ser utilizado por delincuentes para buscar zonas en donde se pudieran dañar los intereses vitales del municipio. En las grandes ciudades, hay túneles llenos de cables y conducciones de diferentes tipos.

Otras fuentes de datos geográficos son las fotos aéreas y de satélite. Se utilizan en agricultura y selvicultura para determinar los usos del suelo y la vegetación. *Google Earth* nos ofrece una buena idea de sus posibilidades. Sin embargo, el uso de imágenes de alta resolución puede estar restringido por razones militares o por razones de privacidad. Con el progresivo aumento de la alta resolución podemos llegar a ver demasiado. Observar las imágenes es gratuito, pero no se permite recoger información relativa a la defensa nacional y, más aún, transformar los datos sin una autorización por escrito dada por la autoridad competente que tenga que ser consultada, conforme establezca la legislación de cada país.

Para el procesamiento de datos geográficos debemos disponer de un Sistema de Información Geográfica (SIG)

que pueda gestionar la información geográfica de una manera eficiente (véase el capítulo 15). El resultado del tratamiento de la información geográfica se puede mostrar como mapas complementados con datos tabulares, como sucede en un sistema de tipo Atlas (véase el capítulo 7). Cuando calculamos la ruta más corta entre dos puntos, obtenemos un mapa que nos muestra el camino más corto y datos tabulares que nos muestran las distancias entre todas las desviaciones.

Los datos geográficos también pueden recogerse mediante el uso de GPS junto con un dispositivo manual para el registro de datos. Ya de vuelta en casa, podemos descargar los datos al ordenador y posteriormente, cuando los datos se consideren satisfactorios, transferirlos al sistema o a un mapa público ([www.openstreetmap.org](http://www.openstreetmap.org)), haciendo públicos los datos y pudiendo ser así utilizados por todo el mundo. Se dan más detalles acerca de los mapas de *OpenStreetMap* en el capítulo 16.

## Referencias

Worboys, M.F. and Duckham, M., 2004: *GIS: A Computing Perspective*, Second edition. London: CRC Press. ISBN 0-415-28375-2.