

Diplomado en Análisis de Información Geoespacial

¿Qué son bases de datos geoespaciales?

Autores:

M. en G. Cecilia Gutiérrez Nieto

M. en G. Luis A. Castellanos Fajardo

Introducción a las Bases de Datos Geospaciales.

Una Base de datos es una colección de elementos o datos interrelacionados que se pueden procesar por uno o más sistemas de aplicación. Esta funcionalidad ayuda a evitar:

- a) **Redundancia de los Datos.**- hay datos que deben aparecer una y solo una vez en el sistema.
- b) **Pobre control de los Datos.**- el mismo elemento de los datos suele tener varios nombres, dependiendo del archivo en el que esté contenido.
- c) **Capacidad inadecuada de manipulación de los datos.**- Los archivos indexados, permiten tener un control sobre las consultas, existencia de identificadores únicos.
- d) **Programación excesiva.**- para hacer consultas y manipulación más rápida y eficiente.

Las ventajas de diseñar y trabajar con una base de datos representa:

- Tener datos estructurados.
- Tener datos y procesos de división, esto significa una mayor dependencia del dato y mayor flexibilidad de procesamiento.
- Integridad de los Datos, esto es una consistencia de los datos así como seguridad y protección de los mismos.
- Datos de larga vida y duración.

Toda base de datos requiere de un manejador de base de datos y en el mercado existen muchos que acorde con las necesidades de cada usuario ofrecen ventajas y desventajas sobre el control de los datos; éstos manejadores se conforman de una

base de datos, un sistema computacional que manipula las bases de datos y el hardware.

A pesar de que cada manejador de base de datos ofrece sus propias particularidades, éstos trabajan bajo los mismos lenguajes complementarios para atender las consultas de los usuarios:

- 1) **Data Definition Language (DDL)**.- crea y estructura archivos de base de datos Ej.
“Create table Catastro”
- 2) **Data Manipulation Language (DML)**.- Modifica, Borra y Calcula datos. Ej.
Val_propiedad = Val_predio + Val_construcción”
- 3) **Structured/Standard Query Language (SQL)**.- Interroga a la base de datos. Ej.
“Select dueño from predio where dirección= “Insurgentes” y area > 100”

El trabajar con un manejador de Base de Datos ofrece ciertas ventajas que coadyuvan a un eficiente manejo de tus datos. Estas ventajas se traducen en:

- Contar con un modelo de datos.
- Crear índices.
- Utilizar lenguajes de consulta.
- Seguridad de los Datos, porque se usa solo un software.
- Mayor Independencia de datos con las aplicaciones.
- Actualización controlada.
- Una menor Redundancia.
- Acceso a grandes volúmenes de datos.
- Programación de Interfaces (Aplicaciones)

Toda base de datos requiere un diseño que incluya a los usuarios, sus necesidades, los datos, así como la estructura de la organización para la cual estará diseñada. En el

contexto de los Sistemas de Información Geográfica una base de datos es un componente primordial ya que permite transformar los datos en información, por lo que este diseño se hace aún más esencial. (Figura.1)

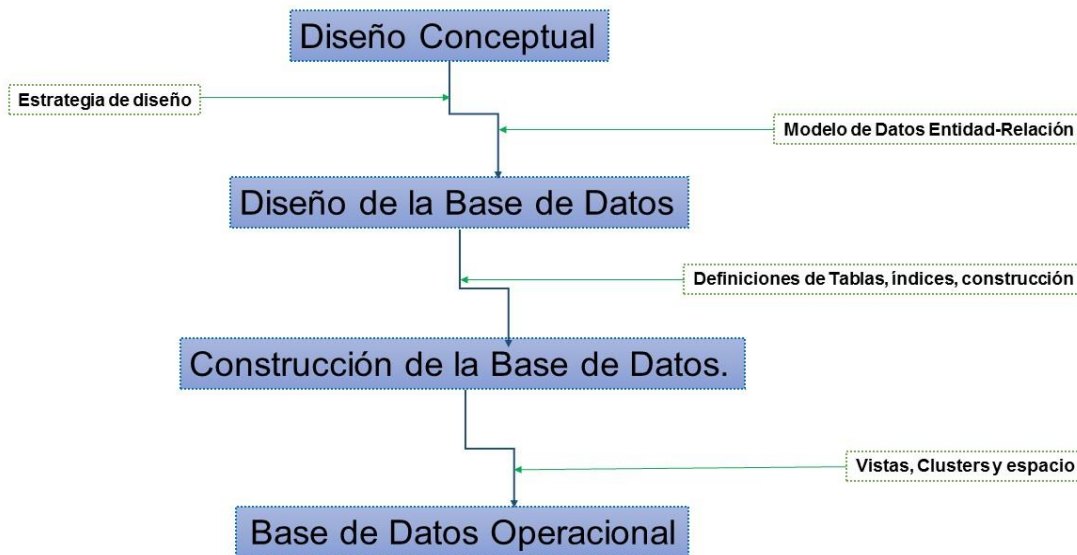


Figura 1. Gutiérrez, C. (2014). Diseño de una Base de Datos. (Ilustración)

El diseño conceptual debe incluir la estructura de la organización para saber qué datos requiere el sistema, qué documentos se generan, cómo se utilizan los datos, quien los utiliza, etc. y debe estar diseñado para que al menos resuelva los problemas más generales identificados durante la etapa de diseño.

El diseño de la base de datos debe incluir por su parte la definición de tablas, índices, vistas, clusters y espacio, en esta etapa no se debe perder de vista los objetivos, los límites y los alcances de la base de datos. Antes de la construcción de la base de datos es importante también:

- **Definición del formato de los datos:** número, texto, fecha o lógico.

- **Definición de los datos contenidos:** se le da el nombre a la columna, generalmente el nombre depende del contenido (ej. Municipio, población, superficie, etc).
- **Valores restringidos:** se aplican a restricciones sobre valores de un dato y se utilizan para validar la entrada de datos, así como también aceptar datos por omisión. (P.ej. En atributos de hora, fecha, valores negativos, intervalos, etc.)

L

os tipos de Manejadores de bases de datos se clasifican de acuerdo a la manera en cómo almacenan y manipulan los datos. Los sistemas de Información Geográfica trabajan con tres tipos:

- A. Relacional (RDBMS).** Mayormente usadas en tablas de dos dimensiones (columnas y renglones).
- B. De Objetos (ODBMS).** Emplea herramientas orientadas a objetos, muchas interfases en herramientas de SIG utilizan ya este tipo de manejadores.
- C. Relacional y Objetos (ORDBMS).** Utiliza la estructura de la Relacional con otra que soporta objetos. Ejemplos de ello son: IBM, DB2, MICROSOFT SQL Y ORACLE. Este modelo es muy empleado en SIG.

En el contexto de los SIG las bases de datos han evolucionado desde

1. Los tipos de datos espaciales se refieren a formas como puntos, líneas y polígonos.
2. La indexación espacial multidimensional se utiliza para el procesamiento eficiente de las operaciones espaciales
3. Las Funciones espaciales, planteadas en SQL, son para consulta de propiedades espaciales y relaciones.

En una primera generación de SIG todos los datos espaciales se almacenan en archivos planos con un software especial para interpretar y manipular los datos. Estos sistemas de gestión de primera generación están diseñados para satisfacer las necesidades de los usuarios en donde todos los datos requeridos en encuentran en el dominio de la organización del usuario. Son sistemas propietarios, autónomos, específicamente contruidos para el manejo datos espaciales.

En la segunda generación de sistemas se almacenan los datos no espaciales (atributos) en bases de datos relacionales, pero aún carecen de la flexibilidad que ofrece la integración directa de los datos espaciales.

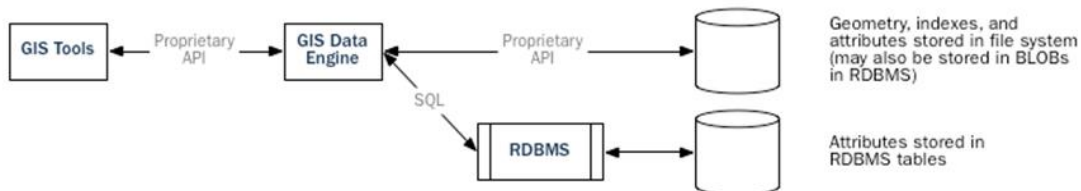
La tercera generación aparece cuando las características espaciales son tratadas como objetos de base de primera clase. Las bases de datos espaciales se integran completamente con la base de datos relacional de objetos.

Evolution of GIS Architectures

First-Generation GIS:



Second-Generation GIS:



Third-Generation GIS:



Figura 2. [Evolución de las arquitecturas para SIG]. Recuperado de <http://workshops.boundlessgeo.com/postgis-intro/introduction.html>

Bajo estas bases surge el concepto de GEODATABASE que es un modelo genérico para el tratamiento de información geoespacial, almacena objetos geográficos, sus atributos, sus relaciones (espaciales o no) y el comportamiento de cada uno de sus elementos.

En este esquema se migra de capas temáticas (colección de elementos geográficos) a entidades reales como red de carreteras, red de alcantarillado, energía eléctrica, etc. y se incorporan el concepto de entidad cuando nos referimos a transformadores, carreteras o lagos. La utilidad de este concepto refleja el tratamiento de nuestros datos para realizar análisis espacial. Pero ¿cómo trabaja el GEODATABASE?:

1. Incluye definiciones, reglas de integridad y comportamiento para una colección de datos integrados que se usan para representar una colección de capas temáticas.
2. Cada diseño incluye propiedades para mis clases de objetos, topologías, redes, catálogos, raster, relaciones y dominios, entendidos éstos elementos como importantes para construir un diseño adecuado.
3. El surgimiento del concepto **Geo-Objeto** se implementa a través del Geodatabase y cada elemento espacial es tratado como tal; cada uno tiene su propia geometría, sus propios atributos y su propio **comportamiento**.
4. Bajo este concepto ya no se habla de puntos líneas o polígonos, sino de transformadores, carreteras o lagos.

Las ventajas de trabajar con un GEODATABASE es un acceso multiusuario, ser compatible con los siguientes manejadores de bases de datos: Oracle, Informix, SQL Server, IBM DB2, Microsoft Access. Reside en un sistema gestor de bases de datos estándar, esto permite que se aprovechen todas las ventajas de los grandes sistemas gestores de BD y esto se traduce en una mayor sencillez para administrar una base de datos corporativa.

Además de estas ventajas, ofrece diversos beneficios, entre los que destacan:

1. **Gestión de datos centralizada.**- esta propiedad la da el DBMS que se esté empleando. Aun cuando cada organización tenga una política única de gestión y mantenimiento de datos geo-espaciales o no.
2. **Edición multiusuario.**- Multitud de usuarios pueden realizar tareas de edición de los datos. Se pueden realizar sesiones de edición que duren semanas o meses.
3. **Comportamiento de los objetos.**- Una vez definido el comportamiento de cada tipo de elemento se hace referencia a ríos, carreteras, parcelas, etc. Pej. los ríos tienen propiedad de flujo, caudal, etc. y solo se conectan a otros ríos o cuerpos de agua. Esto también le da propiedad de simbología determinada.

Es importante agregar que esta estructura de datos comprende dos tipos de comportamiento:

Comportamiento Sencillo:	Comportamiento Complejo:
<p>1. Emplea las reglas de validación (requisitos que debe cumplir un elemento) y son 4:</p> <ul style="list-style-type: none"> • Reglas de atributos • Reglas de Agregación o segregación • Reglas de relación • Reglas de conectividad <p>2. Implementación de topología (ayuda a mantener las relaciones espaciales)</p>	<p>1. Se definen nuevos métodos y propiedades que implementen un comportamiento determinado.</p> <p>2. Requiere de la programación y se generan elementos tipo parcela, tubería, etc.</p> <p>3. Tecnología COM* (especificación de Microsoft) que le da propiedad de integración con otros sistemas.</p>

* La tecnología COM no es otra cosa que la definición de un estándar mediante el cual un trozo de código de una aplicación proporciona servicios a otra enlazándose a ella y llamando a alguna de sus rutinas. Estos trozos de código exponen sus servicios a través de uno o más objetos tipo COM. Cada objeto COM soporta una o más interfases e incluye un determinado número de métodos. Los objetos son una idea central en COM y el objeto es una combinación de datos y métodos. Los objetos COM exponen más de una interfaz.

Sin embargo para poder elaborar y trabajar con un GEODATABASE es necesario conocer su estructura interna a través de la cual se almacena la información, se realiza topología y se puede llevar a cabo Análisis Espacial, ésta comprende:

Los Feature Classes.

- ✓ Es una Colección de Rasgos que representan los mismos elementos geográficos como: pozos, parcelas, tipos de suelo.
- ✓ Todos los elementos deben tener la misma representación espacial (llámese punto, línea o polígono) y compartir un grupo de atributos descriptivos comunes.
- ✓ Todos los elementos deben tener la misma referencia espacial.
- ✓ Elementos Individuales deben compartir relaciones espaciales con otros rasgos (p.ej. Los polígonos adyacentes comparten límites a través de reglas de integridad bien definidas).
- ✓ Una vez que has identificado el conjunto de datos o elementos y cómo los representarás, se le pueden especificar sus propiedades. Dichas propiedades incluyen la manera en que las vas a representar.

Los Feature Dataset, cuyas características son:

- ✓ Organizan una colección de Feature Classes para diversos propósitos. Uno de ellos es para manejar relaciones espaciales.
- ✓ Modela tanto estructuras de feature class como estructuras con relación compleja ya elaborada.
- ✓ Modela y representa relaciones espaciales a partir de **redes y topologías** como capacidad en un contexto de SIG.
- ✓ Las topologías definen cómo los elementos comparten geometría y controlan su integridad a partir de reglas y editando su comportamiento. Ejemplo: las manzanas no pueden traslaparse con las líneas de las calles).

- ✓ Las redes se utilizan para modelar conectividad y flujo entre elementos.
- ✓ Los grupos de datos tipo raster también utilizan una estructura relacional con un acceso masivo al igual que los vectores. También incluye un número extenso de mecanismos para agregar comportamiento.

El Geodatabase trabaja bajo este esquema o estructura:

Dataset.- son las especificaciones para un *feature class* o catálogo raster o una tabla de atributos.

Dominios.- se refiere a las características que determinan la lista o el rango de valores válidos para atributos; estas reglas controlan el comportamiento del software para mantener la integridad de los datos.

Relaciones.- son las relaciones explícitas entre atributos; propiedad de las bases de datos usadas en SIG que definen cómo las columnas se relacionan con columnas en otra tabla. Tienen dirección de cardinalidad.

Reglas Espaciales.- Modelan cómo los elementos u objetos se relacionan de manera geométrica con otros objetos. El geodatabase tiene la propiedad de establecer reglas topológicas complejas.

Capas de Mapas.- Incluyen mapas interactivos y son todas las especificaciones que determinan la representación de los objetos o elementos (símbolos para cada tipo de objeto).

En resumen, el Geodatabase es un modelo con mucha flexibilidad que permite modelar la realidad, a través de la cual la integridad de la información tanto en una base de datos alfanumérica como espacial es mucho más sencilla debido a las características de centralización, la implementación del modo comportamiento y la forma multiusuario de trabajo.

La tecnología con especificaciones COM ha roto con la brecha entre los SIG y otros sistemas de diferentes organizaciones.

Referencias:

1. Hansen James, & Hansen Gary. (1998) **Diseño y Administración de Bases de Datos**. Trad. Miguel Katrib Mora, Universidad de la Habana, Patricio Yustas Torijano, Universidad Pontificia de Madrid. Prentice Hall Editores. España 1998.
2. Longley P., Goodchild M., Maguirre D. & Rhind D. (2005) **Geographical Information Systems and Science**. 2nd edition John Wiley & Sons editors, USA.
3. Geodatabase (ESRI). Recuperado el 04 de noviembre del 2014 de: <http://www.esri.com/software/arcgis/geodatabase>
4. Introduction to PosGIS. Recuperado el 04 de noviembre del 2014 de: <http://workshops.boundlessgeo.com/postgis-intro/introduction.html>