

Bases de Datos Alfanuméricas para SIT

José M. Ciampagna

Índice

1. INTRODUCCIÓN.....	2
TIPOS DE DATOS.....	4
<i>Registros, Tablas, Base de datos</i>	5
2. BASE DE DATOS	11
DEFINICIÓN DE "SISTEMAS DE BASE DE DATOS".....	13
3. ¿NECESIDAD DE MODELAR LOS DATOS?.....	15
MODELO ENTIDAD / RELACIÓN (E / R)	16
TIPOS DE RELACIONES.....	17
<i>UNO - UNO</i>	17
<i>UNO - MUCHOS</i>	18
<i>MUCHOS - MUCHOS:</i>	18
4. SISTEMAS DE ADMINISTRACIÓN DE BASES DE DATOS (DBMS)	19
<i>Ventajas</i>	21
<i>Modificación de los datos</i>	22
5. ARQUITECTURA DE BASE DE DATOS	23
6. TIPOS DE ARQUITECTURA DE BASE DE DATOS	23
BASE DE DATOS EN RED.....	24
BASE DE DATOS JERÁRQUICAS	25
BASE DE DATOS RELACIONALES	26
BASE DE DATOS ORIENTADA A OBJETOS	28

1. Introducción

Desde el comienzo de los Sistemas de Información Geográfica el paradigma de las bases de datos estuvo presente en esta tecnología. Por ejemplo: El primer software SIG comercial - el Arc-Info de ESRI - debe parte de su nombre a un Data Base Management System (DBMS) denominado Info; La vinculación de la información gráfica a la alfanumérica, proceso denominado “geocodificación”, es una de los elementos claves que definen a un SIG; En algunas de las definiciones que se han dado de un Sistema de Información Geográfica se encuentra la consideración de definirlo como una base de datos espacial; Además de considerar a la Cartografía como una herramienta para solucionar el problema de la orientación y ubicación, podemos considerarla como un inventario bajo la óptica de los Sistemas de Información, archivo que representa los rasgos de la Geografía, bajo este nuevo concepto una carta geográfica es nada más y nada menos que una base de datos. Como vemos, en los párrafos anteriores, es necesario analizar el paradigma de las bases de datos dentro del contexto que nos ocupa. Este capítulo tiene como propósito revisar los conceptos principales presentes en el tema de las bases de datos.

Desde tiempos muy remotos los humanos tenemos la necesidad de recordar de alguna forma datos, información acerca de sucesos, personas, ideas, hechos, cosas. Con el transcurrir de los años la necesidad se fue acentuando cada vez más como consecuencia de nuestra limitada capacidad de memoria y el manejo de una realidad mas compleja y promovió el desarrollo de tecnologías destinadas a tal propósito. Podríamos decir que el crecimiento ha sido considerable. Comenzamos dibujando en piedras (en la antigüedad), seguimos con la escritura para de algún modo transmitir y comunicar diversidad de cosas, luego la imprenta, hasta nuestros días, en que contamos con una computadora donde podemos almacenar la información en medios magnéticos (disquetes, discos rígidos, ópticos, cintas, etc.).

Existe una infinidad de ejemplos cotidianos que nos revelan la necesidad de almacenar las cosas que queremos recordar. Los libros, las agendas donde guardamos los números telefónicos, los ficheros, las microfilmaciones, y listas son ejemplos de medios de almacenamiento. De aquí surge un concepto de suma importancia: “la abstracción”. Para almacenar los datos de aquello que nos interesa recordar del mundo real se hace necesario “abstraer” lo que queremos recordar, es decir, necesitamos rescatar sólo lo que nos interesa y olvidarnos por un momento de todo aquello que no es relevante o no aporta nada a nuestros fines específicos. Podemos decir entonces que los datos que almacenamos son una forma de representar el mundo real, es decir una “abstracción” del mismo.

Al respecto Niclaus Wirth, creador del lenguaje PASCAL, dice:

*...."al resolver un problema, se utilice o no el computador, es necesario elegir una abstracción de la realidad y definir un conjunto de datos para representar la situación real. Esta elección debe estar guiada por el problema a resolver. Deben tenerse en cuenta las operaciones a realizar con los datos."*¹

La porción del mundo real que contiene individuos de características similares o cosas de interés para almacenar se denomina "población". Definimos como población a un conjunto de individuos. Los individuos pueden ser: entidades tangibles (Por ejemplo.: casas, autos) o bien intangibles o ideales (Por ejemplo: cuenta bancaria, cuenta corriente, etc.). Cada uno de estos individuos posee muchas propiedades o atributos, de los cuales seleccionamos aquellos que interesan.

Como ejemplo de población tenemos: "El personal de una Institución Pública". En el proceso de abstracción y para nuestro ejemplo, rescatamos:

- ☐ Nombres;
- ☐ Sexo;
- ☐ Fecha de Nacimiento
- ☐ Domicilios;
- ☐ Estudios cursados;

Se observa que, de cada individuo se han seleccionado solamente los atributos que son de interés para la Institución, sin interesar otros aspectos que son secundarios para nuestro caso, como podría ser la marca del automóvil de cada persona, su color de pelo, su número de cuenta bancaria, etc.

Ahora imaginemos una gran Institución, con miles de transacciones y trámites diarios, que deben ser "guardados" en algún lado. Evidentemente no va a ser suficiente y cómodo anotar todo en una planilla en papel o una tabla de cálculo tipo Excel.

¹ Algoritmos + Estructura de datos = Programas, Niclaus Wirth

La informática, nos provee varias soluciones para solucionar el problema. Dentro de ellas una muy particular cuando la cantidad de datos o información es mucha; crece de manera desmedida; se complica mantener un orden, o necesitamos compartirla entre varios usuarios. La solución que nos propone son las **Bases de Datos**.

Martin define una “**Base de Datos**” como:

“Una colección de datos interrelacionados en conjunto, sin redundancia perjudicial innecesaria. Su finalidad es servir a una o más aplicaciones, de la mejor manera posible. Los datos se almacenan de forma que resulten independientes de los programas que los usan y se emplean métodos bien determinados para incluir nuevos datos, extraer o modificar los datos almacenados”.

A continuación desarrollaremos el tema.

Tipos de Datos

Distinguimos tres tipos de datos elementales. Así tenemos los datos:

* Numéricos: 1, 2, 3,.....

Sobre ellos se pueden realizar todas las operaciones aritméticas y numéricas conocidas, por ejemplo: “+”, “-”, “*”, “/”.

* Caracteres: comprende todo el abecedario, incluyendo también a los números. Varios caracteres concatenados conformarán las cadenas de caracteres o “*strings*”. Con los números, si son expresados como caracteres, no podemos realizar cálculos.

* Lógicos: Son aquellos que denotan una condición de verdad o falsedad (V o F). Están relacionados con los operadores que emplean de conjunción, negación, disyunción, etc.

Específicamente una colección de ítems o datos del “mismo tipo”, reunidos bajo un único nombre colectivo, constituye una “estructura de datos”. Mediante su uso podemos asociar a un único nombre una colección entera de valores (datos).

Se pueden formar tipos de estructuras más complejas –modelos de datos- a partir de datos elementales. Algunas de ellas son:

- ❑ Arreglo Lineal (linear array)
- ❑ Listas Lineales

- ❑ Registros → Tablas → Bases de datos
- ❑ Matrices
- ❑ Otras

Observamos que únicamente hemos descrito cómo representar los datos literales y numéricos que se corresponden con el lenguaje escrito (datos alfanuméricos). No hemos considerado, por ahora, la necesidad de almacenar otro tipo de informaciones que necesitan ser representadas como gráficos, mapas, imágenes fotográficas y satelitales, etc. en formato digital. Estos últimos tipos de datos son temas a tratar en el estudio de los Sistemas de Información Geográfica.

De las estructuras mencionadas nos focalizaremos en: el registro, las tablas y las bases de datos.

Registros, Tablas, Base de datos

Para construir "el modelo" deseado puede señalarse una serie de etapas a seguir:

1) Abstracción de la realidad: seleccionar del mundo lo que interesa a nuestro propósito y agrupar los objetos de la realidad que reúnan las mismas características y sean una unidad conceptual y conformar así las entidades de nuestro modelo. Genéricamente se denomina "entidad" a una familia de objetos con iguales propiedades o características. Siguiendo con el ejemplo del "El personal de una empresa" tendremos que la entidad "Personal" contendrá como propiedades: Nombres, Sexo, Domicilios y Estudios cursados.

Las propiedades se las conoce como "atributos" y describen la entidad cualitativa y cuantitativamente.

Los atributos de acuerdo a la función que desempeñan, se pueden clasificar en:

- "Atributos de identificación" o "claves".
- "Atributos de descripción" o "valor",
- "Atributos de vinculación".

El "atributo de identificación" o "clave primaria" identifica unívocamente una instancia de todo el conjunto que abarca la entidad. Pueden coexistir varios atributos de identificación de

una misma entidad. Todo registro contiene siempre uno o varios datos clave que lo identifican unívocamente del resto, esa clave pertenece únicamente a ese registro y se lo denomina "clave primaria", no puede haber en un mismo archivo dos o más registros con la misma clave primaria. Por ejemplo, podemos identificar unívocamente a cada persona por su tipo y número de documento.

El "atributo de descripción" califica la entidad cualitativa y cuantitativamente.

El "atributo de vinculación" o "clave secundaria" permite relacionar la entidad con otras entidades distintas y sus correspondientes atributos.

2) Una vez definidas las entidades, definiremos las relaciones entre ellas y probablemente nuevas entidades que surjan de las relaciones. Ahora bien, en nuestro ejemplo además de los datos del personal, nos interesan las asistencias y horarios de llegada de cada uno de este modo tendremos otra entidad que podemos llamar "Asistencia" en la cual las propiedades de ésta serán: día, mes, año, y hora de entrada y hora de salida. En el mundo de las entidades tenemos distintos tipo de relaciones que más adelante describiremos cuando hablemos más acerca del modelo Entidad-Relación (E/R).

3) Ahora estamos en condiciones de armar las tablas que conformarán nuestra base de datos. Cada tabla dentro de la base de datos por lo general se corresponde con una entidad o bien con una relación. En el ejemplo tendremos las tablas: "Personal" y "Asistencia", que mediante atributos estarán vinculadas:

Personal	Asistencia
Id_personal	Día
Nombre	Mes
Apellido	Año
Fecha de Nacimiento	Hora de entrada
Sexo	Hora de salida
Domicilio	Id_personal
Estudios cursados	

El `id_personal` (clave secundaria) dentro de la tabla Asistencia contendrá el número de `id` (clave primaria) del personal al cual corresponde. Todos los demás atributos son de descripción.

Los atributos son la unidad de información más pequeña almacenada que recibe un nombre.

4) Una vez armado el esqueleto de la base de datos podemos guardar las ocurrencias o instancias que representan la realidad. Los datos son agrupados para formar las filas o “tuplas” o lo que comúnmente llamamos registros.

Por ejemplo, con los siguientes datos:

Id	1
Nombres	María
Apellidos	Perez
Fecha de Nacimiento	05-06-45
Sexo	femenino
Estudios cursados	clavadista

Id	2
Nombres	Juan
Apellidos	García
Fecha de Nacimiento	05-07-46
Sexo	masculino
Estudios cursados	buzo

Podemos formar:

Nro	Nombres	Apellidos	Fecha Nacimiento	Sexo	Estudios cursados
00001	Maria	Perez	05-06-45	femenino	clavadista
00002	Juan	García	05-07-46	masculino	buzo
.....					
.....					
n-1	Santiago	Sasia	04-06-30	masculino	granjero
n	Francisca	Lopez	01-02-85	femenino	Torera

Figura 3 "Modelo de tabla"

Una tabla se organiza de tal modo que cada fila representa un objeto de la realidad (instancia u ocurrencia) y cada columna representa un "atributo".

A partir de la construcción de diferentes tablas podríamos construir modelos mucho más complejos que el trivial ejemplo del personal de una empresa y sus asistencias.

A los fines de clarificar las ideas expuestas, veamos otro ejemplo: En una biblioteca, a fines de administración, se elabora una ficha por cada libro. En la misma se escribe el número de identificación del libro, el nombre del libro, el autor el tema, la editorial, y aquellos otros aspectos que pueden interesar, etc. Se ha realizado una abstracción del libro en una ficha.

Libros	Número:123456
Nombre : Dirigir Proyectos	
Autor : Andy Brucey	
Tema : Management	
Editorial : Grijalbo	
Año : 2002	
Idioma : Español	

Figura 2: "Modelo de ficha bibliográfica"

Si consideramos nuestro modelo, podemos construir una tabla donde:

- "Número de libro": atributo de identificación o clave
 - Nombre del libro
 - Autor
 - Tema
 - Editorial
 - Año
 - Idioma
- atributos de descripción

La tabla:

Nro	Nombre	Autor	Tema	Editorial	Año	Idioma
00001						
00002						
.....						
n-1						
N						

Figura 3 "Modelo de tabla"

Continuando con el mismo tema de la biblioteca, se puede además construir una nueva tabla que modele los socios de la biblioteca.

Nro.Socio	Nombre	Documento	Dirección
00001			
.....			
n-1			
N			

Figura 4 "Modelo de tabla de socios"

Por último, es factible confeccionar una nueva tabla, para relacionar "cada libro con cada socio" la cual permitirá indicar la tenencia del libro.

Nro.Libro	Nro.Socio
0001	0003
0007	0015
0123	1234
0122	1113

Figura 5 "Modelo de tabla: relación libros-socios"

En este caso podemos observar que el atributo "Nro.libro" y "Nro.Socio" son "atributos de vinculación"; el primero con la tabla de libros y el segundo con la tabla de socios.

Queda evidente que podemos formar sistemas más complejos a partir de estos conceptos básicos y elementales. Por ejemplo en nuestro caso; es posible agregar tablas de datos que informen sobre autores, editoriales, librerías, cuentas corrientes de socios, etcétera, de tal forma que el modelo cumpla con los requerimientos de una realidad más compleja.

Nota importante: A fines de aclarar diferentes formas de denominación de los conceptos empleados exponemos la siguiente tabla de equivalencias que emplearemos a lo largo de nuestro trabajo:

Términos formales de la nomenclatura relacional	Términos informales
Relación	Tabla
Tupla	Fila o registro
Cardinalidad	Número de filas
Atributo	Columna o campo
Grado	Número de columnas
Clave primaria	Identificador único
Clave secundaria	Atributo de vinculación
Dominio	Valores posibles

2. Base de datos

Para comprender el nacimiento y evolución de los sistemas de bases de datos, es necesario conocer el medio informático de los últimos años. Las aplicaciones, que necesitaban de la utilización de archivos de datos, operaban directamente con sus propios datos.

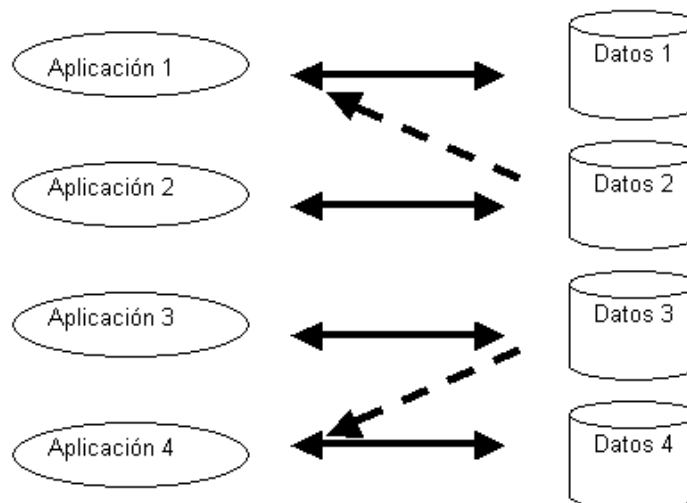
Definimos como "aplicación" al programa que realiza un proceso sobre los datos en la que se utilizan uno o más archivos, este uso puede consistir en la recuperación, actualización, mantenimiento o generación de datos, procesamiento, requiriendo la presencia de un programa que provea las instrucciones para todas las acciones que tengan lugar sobre los datos solicitados, etc.

Por otro lado, los sistemas operativos no disponían de multiproceso, es decir, el ordenador sólo podía ejecutar en un único programa por vez y la seguridad de los archivos estaba garantizada.

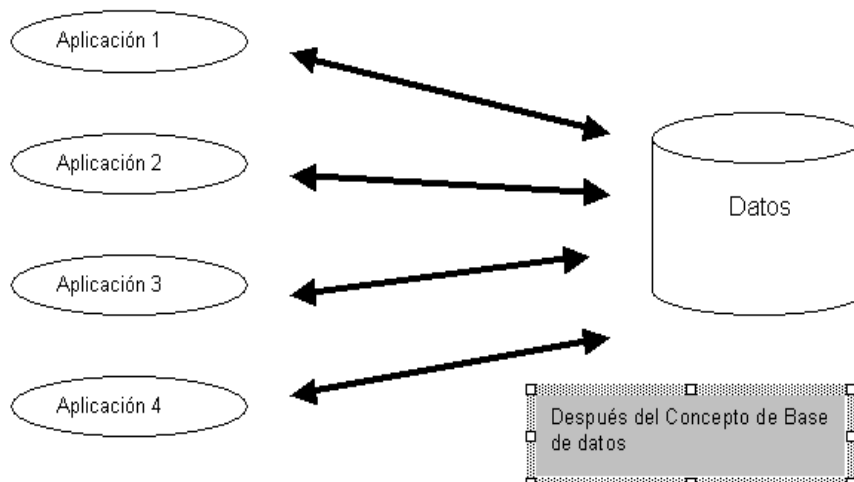
A medida que los sistemas de información crecieron, surgieron nuevos problemas y requerimientos:

- La necesidad de que varias aplicaciones convivieran en un mismo entorno.
- Inconsistencias de datos, ya que al contar cada aplicación con sus datos muchos de ellos se repetían, y las actualizaciones de los datos no siempre podían llevarse a cabo en todos lados. El peligro de contar con información errónea o desactualizada era un hecho.
- La información y datos de las empresas se encontraban dispersos por varias aplicaciones, dificultando de sobremanera la generación de informes, para conocer la situación actual de la empresa en un determinado momento.
- La custodia de los datos e información de la empresa no podía realizarse correctamente y de manera centralizada.
- Desperdicio de memoria y de medios de almacenamiento consecuente de la redundancia perjudicial de datos.
- Casi imposible se hacía hacer cumplir las políticas y lineamientos generales de la empresa, ya que cada aplicación tenía sus formatos de datos y cada una de ellas administraba los datos a su modo.
- Dificultad en el desarrollo de nuevas aplicaciones, o cambios en las existentes

- necesarias a los requerimientos cambiantes de la empresa. La situación se hizo insostenible y es así que surgieron los sistemas de Base de Datos.



Antes de las Bases de Datos



Después del Concepto de Base de datos

Definición de "Sistemas de base de datos"

Un sistema de bases de datos es básicamente un sistema para archivar en computador.; o sea, es un sistema computarizado cuyo propósito general es mantener información y hacer que esté disponible cuando se solicite. La información en cuestión puede ser cualquier cosa que se considere importante para el individuo o la organización a la cual debe servir el sistema; cualquier cosa necesaria para apoyar el proceso general de atender los asuntos de ese individuo u organización. Sus cuatro principales componentes son:

1. Información, la cual estará integrada (la base de datos puede considerarse como una unificación de varios archivos de datos, por lo demás distintos, y que elimina todo o en parte cualquier redundancia entre ellos) y también será compartida (los elementos individuales de información en la base de datos pueden compartirse entre varios usuarios distintos, en el sentido de que todos ellos pueden tener acceso al mismo elemento de información).
2. Equipo: sistemas de almacenamiento secundario (en general discos magnéticos de cabeza móvil) junto a los dispositivos de entrada/salida asociados, controladores de dispositivos, y demás; el procesador o procesadores y memoria principal.
3. Programas. El DBMS (DataBase Managment System) maneja todas las solicitudes de acceso a la base de datos formuladas por los usuarios.
4. Usuarios

Evidentemente el empleo de las bases de datos nos brindan muchas ventajas, entre las más destacadas tenemos que:

- Los datos están almacenados en diversos soportes de información de tal forma que son independientes de los programas que los manejan.
- Su utilización no está restringida a una sola aplicación, siendo posible su acceso por varias aplicaciones, incluso simultáneamente.

- Para gestionar la información en la base de datos, es decir: incluir nuevos datos, borrar ya existentes o modificarlos, se emplean procedimientos especiales.
- Es compacta
- El acceso a cualquier dato es rápido
- Menos laborioso, evita tareas tediosas como por ejemplo recorrer aplicación por aplicación para obtener un informe; actualizaciones.
- Actual: permite disponer de la información en cualquier momento y día.
- Ofrece a la empresa un control centralizado de su información: existe una persona identificable con esta responsabilidad central sobre los datos, que es el DA (Data Administrator). El conoce la información y las necesidades de la empresa en este aspecto en un nivel gerencial superior. Decide cuales datos deben almacenarse en la base de datos y establece políticas para mantener y manejar los datos una vez almacenados. También está el DBA (DataBase Administrator) que es el técnico responsable de poner en práctica las decisiones del DA, es un profesional en procesamiento de datos, su tarea es crear la base de datos en sí y poner en vigor los controles técnicos necesarios para apoyar las políticas dictadas por el DA. Además el DBA garantiza el funcionamiento adecuado de la base de datos y proporciona otros servicios.
- Es posible disminuir la redundancia, aunque a veces es necesaria.
- Es posible evitar la inconsistencia (hasta cierto punto) que ocurre cuando tengo más de una entrada de información.
- Es posible compartir datos (sharing): implica no sólo que las aplicaciones ya existentes pueden compartir la información de la base de datos, sino también que se pueden desarrollar aplicaciones nuevas para trabajar con los mismos datos almacenados.
- Es posible hacer cumplir las normas: la normalización de formatos de datos almacenados es deseable como apoyo para el intercambio de información o migración de datos entre sistemas.

- Se pueden aplicar restricciones de seguridad al tener jurisdicción completa sobre la base de datos el DBA: puede asegurar el acceso a la base de datos solo a través de los canales apropiados, y puede definir las verificaciones de seguridad por realizar cuando se intente acceder a información delicada (por ejemplo una cuenta bancaria, estudios médicos confidenciales).
- Se puede equilibrar requerimientos opuestos: el DBA puede estructurar el sistema con miras a proporcionar un servicio integral “óptimo para la empresa”.

Como puede observarse, son muchas las ventajas que encontraremos al implementar una base de datos en nuestra empresa. Pero no todo es ideal, ya que también existen desventajas que se desprenden de las ventajas anteriormente mencionadas:

- Puede peligrar la seguridad e integridad de la información manejada dentro de la empresa si no son aplicados los controles necesarios.
- Se requiere un equipo adicional y personal especializado
- El aumento de procesamiento podría ser considerable
- La operación exitosa es crucial, con ello se hace referencia a que la empresa podría ser muy vulnerable a cualquier falla del sistema.
- Con toda probabilidad el sistema será complejo, aunque deberá ser totalmente transparente para el usuario, es decir que el usuario no tendrá conocimiento alguno de mencionada complejidad.

3. ¿Necesidad de modelar los datos?

Antes que nada vamos a definir lo que es un modelo de datos: es un conjunto de estructuras lógicas que permiten al usuario describir los datos que serán almacenados en la base de datos junto con un conjunto de operaciones para manipular los datos.

A partir del concepto de base de datos, el diseño de las bases de datos es un punto muy importante y un buen diseño permite estabilizar la estructura de datos ahorrando posteriores trabajos de modificación de la estructura y sobre todo de las aplicaciones vinculadas a la base de datos. Numerosas herramientas de diseño se utilizan para modelar datos y en un proceso posterior

crear la estructura soporte de la base de datos. Estas herramientas permiten, además, independencia del modelo de datos de distintos Administradores de Bases de Datos (DBMS). Uno de los modelos más importantes es el debido a Peter Chen (1976) conocido como modelo entidad / relación (E / R), uno de sus empleos es para el diseño de las bases de datos relacionales. Analizaremos en detalle, a manera de ejemplo, el modelo (E / R).

Modelo Entidad / Relación (E / R)

Un "modelo de datos" es un proceso para crear una representación lógica de una estructura de datos. El modelo de datos entidad-relación se basa en una percepción de un mundo real que consiste en una colección de objetos básicos llamados entidades, y relaciones entre esas entidades. Estas representaciones lógicas en general se corresponden con un modelo gráfico..

Antes de avanzar, expliquemos que son las "relaciones". Una relación incluye a tres entidades: dos individuos y una relación. La relación es la forma en que los individuos se afectan entre sí.

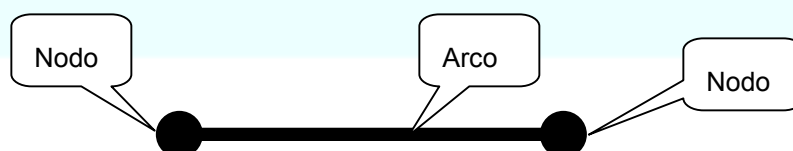
Las entidades mencionadas son:

Sujeto: es un individuo del cual parte la relación, pertenece al dominio, que es el conjunto de partida.

Relación: se caracteriza por uno de los verbos activos o pasivos que muestran cómo actúan los individuos entre sí.

Objeto: es el individuo que recibe la acción de la relación, es el conjunto de llegada.

Un modo de expresar las relaciones es por medio de un "grafo". Un grafo se compone de las figuras geométricas nodos y arcos, donde los nodos representan los individuos de la relación y los arcos la relación.



Algunos conceptos y definiciones:

Entidad: conjunto de objetos que se pueden agrupar por compartir propiedades, relaciones o aspectos de comportamiento. Por ejemplo: proveedor, estudiante, etc.

Atributo: propiedad asociada a una entidad. Por ejemplo: la entidad “estudiante” tiene de atributos DNI, nombre, apellido, edad, sexo, etc.

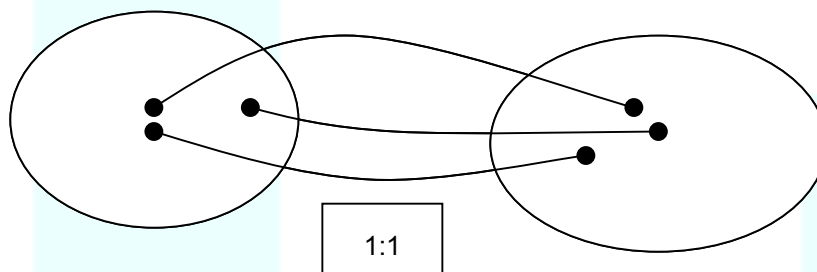
Identificador Unívoco o Clave Primaria: conjunto de atributos que identifican unívocamente a una entidad dentro de un conjunto de entidades. Por ejemplo: la entidad “estudiante” tiene de clave el DNI.

Relación (asociación): conexión semántica entre dos conjuntos de entidades. Por Ejemplo: Propietarios con Parcelas.

Tipos de relaciones

UNO - UNO

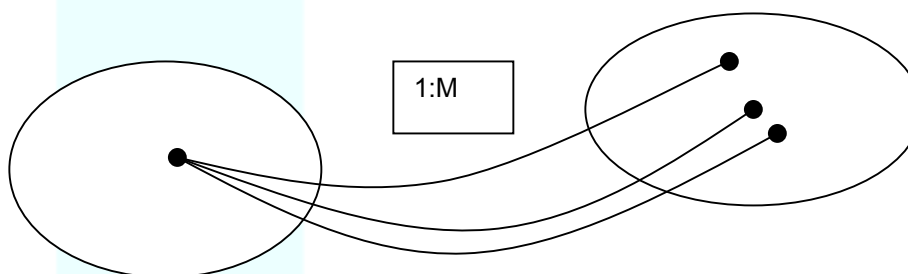
Cuando queda establecida una relación unívoca entre dos individuos. Es decir: cada entidad de la tabla A está en relación con un objeto de la tabla B.



Por ejemplo: una provincia tiene un gobernador; un gobernador gobierna una provincia.

UNO - MUCHOS

Cuando un individuo del dominio se relaciona con dos o más individuos del conjunto alcance. Cada entidad de A puede estar en relación con muchos objetos de B.

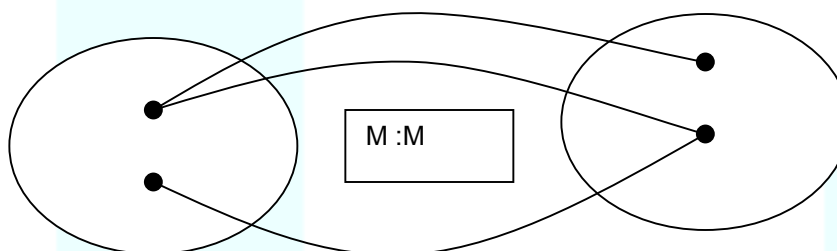


Por ejemplo: un escritor tiene muchos libros; un libro tiene un autor (suponemos que cada libro tiene un solo autor o tiene al menos uno principal).

MUCHOS - MUCHOS:

Cuando varios individuos del conjunto de dominio se relacionan con varios individuos del alcance. Cada entidad de A puede estar en relación con muchos objetos de B y viceversa.

Es aconsejable transformar esta relación de modo que queden dos relaciones uno a muchos.

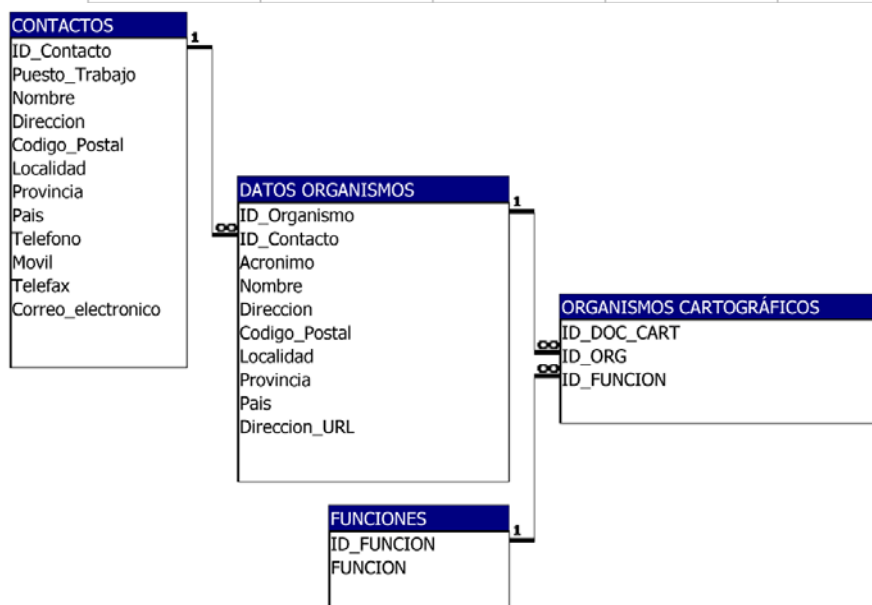


Por ejemplo: una parcela tiene varios propietarios; un propietario tiene varias parcelas.

En el siguiente ejemplo se ven las tablas y un modelo de datos realizado con las herramientas de diseño del software Microsoft ACCESS.

	ID_Organismo	ID_Contacto	Acronimo	Nombre	Direccion
▶ +	001	123456789	IGA	Instituto Geográfico	Av. La Libertad,
+	002	987654321	CMM	Cartografía del mur	C/Blanca,11
+	003	4274354867	O.S.	Ordinance Survey	13, Wellington :
*					

Codigo_Postal	Localidad	Provincia	Pais	Direccion_URL
99325	Ávila	Ávila	España	www.iga.com
28080	Madrid	Madrid	España	www.cmm.com
51654	Londres		Reino Unido	ww.os.com



4. Sistemas de Administración de Bases de datos (DBMS)

El sistema de administración de base de datos (DBMS) es el conjunto de programas que manejan todo el acceso a la base de datos. Los sistemas de administración de bases de datos se especializan en el almacenamiento y manejo de todo tipo de datos, incluyendo datos geográficos. Los DBMS (Data Base Management Systems) están optimizados para almacenar y recuperar datos y muchos SIG utilizan a estos sistemas para este propósito. Ellos no tienen las herramientas más comunes de los SIG como herramientas analíticas y de visualización gráfica.

Conceptualmente, lo que sucede es lo siguiente:

1 – Un usuario solicita acceso, empleando algún sub-lenguaje de datos determinado (por ejemplo SQL)

2 – El DBMS interpreta esa solicitud y la analiza.

3 – El DBMS inspecciona, en orden, varios esquemas (no se especificarán en esta monografía cuales son).

4 – El DBMS ejecuta las operaciones necesarias sobre la base de datos almacenada.

Las funciones básicas de un DBMS son:

Definición de datos:

El DBMS debe incluir componentes procesadores de lenguajes para cada uno de los diversos lenguajes de definición de datos (DDL). También debe entender las definiciones de DDL, en el sentido en que, por ejemplo, “entiende” que los registros externos PERSONAL contienen un campo SEXO; y debe poder utilizar estos conocimientos para interpretar y responder solicitudes de los usuarios.

DDL (Lenguaje de definición de datos): un esquema de base de datos se especifica por medio de un conjunto de definiciones que se expresan mediante el DDL. El resultado de la compilación de sentencias DDL es un conjunto de tablas las cuales se almacenan en un archivo especial llamado diccionario de datos.

Manipulación de datos:

El DBMS debe ser capaz de atender las solicitudes del usuario para extraer, y quizás poner al día, datos que ya existen en la base de datos, o para agregar en ella nuevos datos.

DML (Lenguaje de Manipulación de datos): por manipulación queremos decir: la recuperación de información almacenada en la base de datos; la inserción de información nueva en la base de datos; la supresión de información de la base de datos; la modificación de datos almacenados en la base de datos.

Seguridad e integridad de datos:

El DBMS debe supervisar las solicitudes de los usuarios y rechazar los intentos de violar las medidas de seguridad e integridad definidas por el DBA. Proporcionar mecanismos para el establecimiento de la seguridad: otorgando diferentes permisos de acceso y manipulación; proporcionando protección de accesos no autorizados mediante criptografía.

Recuperación y concurrencia de datos:

El DBMS – o en su defecto algún componente de software asociado a él, al que por lo regular se denomina *administrador de transacciones* - debe cuidar del cumplimiento de ciertos controles de recuperación y concurrencia. Gestión de la concurrencia: ayuda a resolver conflictos cuando dos o más usuarios están tratando de acceder a los mismos datos. Si no tiene mecanismos para resolverlo, se bloquea el sistema.

Diccionario de datos:

El DBMS debe cumplir con la función de *diccionario de datos*. El diccionario de datos es una base de datos por derecho propio (pero una base de datos del sistema, no del usuario) que contiene “datos de los datos”, es decir metadatos.

Desempeño: el DBMS deberá ejecutar todas las funciones descritas en la forma más eficiente posible.

Proporcionar un lenguaje de definición de datos:

(Definition Data Language). Generalmente es conocido por la sigla SQL (Structured Query Language, corresponde en castellano a: lenguaje estructurado de interrogación a base de datos).

Recuperación ante fallos:

- ⊖ Por problemas de Hardware: cambio de voltaje, corriente.
- ⊖ Por problemas de Software: debido al sistema operativo.
- ⊖ Las pérdidas accidentales se previenen mediante el mecanismo de transacción. Una transacción es una secuencia de operaciones de manipulación sobre la BD.
- ⊖ Las transacciones tienen la propiedad de que si se interrumpen antes de completarse, la BD es restaurada a un estado de auto-consistencia, usualmente el estado anterior al inicio de la transacción.
- ⊖ Las transacciones protegen los datos de fallos de la corriente eléctrica, de la red y de la concurrencia de usuarios.

Ventajas

Existen numerosas ventajas del uso de DBMS:

Respecto de los usuarios:

Usuario final: puede acceder a los datos.

Programador: elimina problemas de diseño lógico y físico. Permite la depuración de errores. Permite el mantenimiento del sistema.

Respecto del sistema:

Control centralizado, fiabilidad, consistencia, seguridad.

Criterios de homogeneización.

Generación de nuevas aplicaciones.

Equilibrio entre requerimientos.

Modificación de los datos

Debido a que las poblaciones (conjunto de datos) sufren continuas modificaciones es necesario realizar siempre en los archivos tres operaciones básicas con DML (Lenguaje de manipulación de datos):

Recuperación: es examinar uno o más registros para extraer información de algunos de sus datos, sin afectar de modo alguno al registro en el archivo.

Actualización: resulta de alterar alguno de los registros por modificaciones en alguno de los datos del individuo.

Mantenimiento: es la modificación de la población debido a la incorporación o eliminación de alguno o varios individuos. Esta operación altera el archivo para reflejar la alteración de la población.

Sin alteración de la BD:

❑ Realización de búsquedas.

❑ Realización de consultas.

Con alteración de la información contenida dentro de la BD:

❑ Modificaciones.

❑ Borrados o Bajas.

❑ Incorporaciones.

5. Arquitectura de base de datos

Se denomina así a las distintas maneras de organizar y ver los datos. Es el medio previsto para representar dentro de una BD las relaciones que existen entre y dentro de las poblaciones y que son de interés para una o más de una aplicación. El tipo de arquitectura elegido para trabajar debe ser apto para representar las relaciones más complicadas.

Hay distintos tipos de arquitectura, siendo la diferencia entre ellas el poder estar asociadas a distintos tipos de acceso, sistemas de división de los archivos, relación con archivos auxiliares, etc.

Los tipos de arquitectura de base de datos son:

- "Base de Datos en Red": los individuos de las poblaciones están relacionados entre sí por punteros, formando redes de relaciones.
- "Base de Datos Jerárquica": son estructuras que presentan los individuos en distintos niveles, relacionándose un individuo con uno o más de un nivel inferior.
- "Base de Datos Relacionales": el usuario percibe los datos en forma de tablas. Los operadores al alcance del usuario (por ejemplo recuperación de datos) generan tablas nuevas a partir de las existentes.
- "Base de datos orientada a Objetos".
- "Base de datos mixtas".

6. Tipos de Arquitectura de Base de datos

Algunos de los modelos de base de datos citados más arriba se encuentran hoy caídos en desuso o superados por modelos más modernos. Los más modernos o usuales son los modelos de base de datos relacional o el orientado a objetos; otros son modelos mixtos, como el relacional orientado a objeto

Base de datos en red

Este tipo de bases de datos están disponibles desde 1970. Nacen en la Conferencia de Lenguajes de Sistemas de Datos (CODASYL). Ellas son hábiles para manejar relaciones 1:1, 1:m y n:m sin redundancia. Esta base de datos representa una o más relaciones en forma de conjuntos, que es el prototipo para la forma de relación. El dueño se aplica al dominio para el cual se selecciona el sujeto, el miembro corresponde al alcance para el cual se selecciona uno o más objetos.

Tenemos:

- Un dueño: es el sujeto de la relación.
- Un conjunto: la relación uno-muchos.
- Varios miembros: los cuales se relacionan con el sujeto.

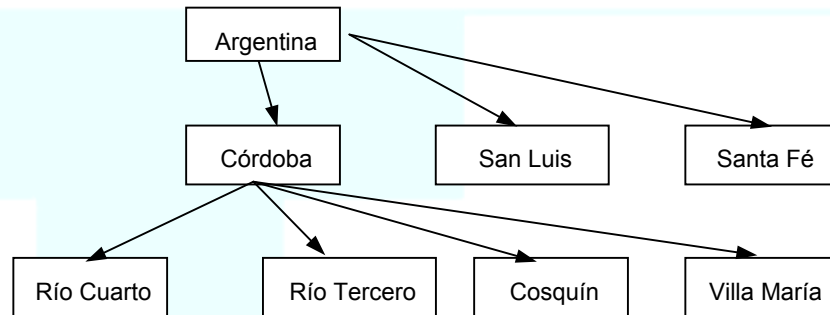
PROFESOR	dueño
ENSEÑA	relación
CLASE	miembros

Representación de ocurrencia de conjunto

La red se representa mediante segmentos y campos punteros. Los segmentos describen a los miembros y a los dueños y son los que se relacionan por medio de la estructura de conjunto. Los campos punteros sirven para relacionar o conectar un segmento con otro. Los segmentos tienen distintos tipos de punteros, los segmentos miembros tienen punteros miembros y los segmentos dueños tienen punteros dueños.

Los punteros serán colocados en extremos diferentes de los segmentos para poder diferenciar el tipo de función que cumplen.

En el ejemplo anterior podemos observar cómo se relaciona un dueño (país Argentina), con los miembros (las provincias); a su vez observamos que las provincias que participan como miembros en la primera relación figuran como dueños en la segunda, siendo los miembros de estas ciudades pertenecientes a cada una.



Características de la base de datos en red

Ventajas

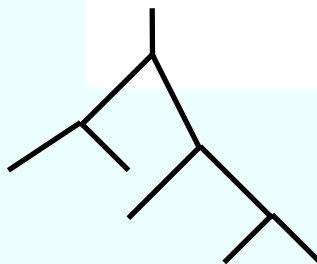
- Son rápidas.
- Todas las relaciones son posibles, no solamente las jerárquicas.
- No es necesario que exista una raíz única como en las bases jerárquicas.
- Ofrece un modelo complejo para representar el mundo real.

Inconvenientes

- Rigidez.
- Complejas de utilizar.

Base de datos Jerárquicas

Las bases de datos jerárquicas son probablemente las más antiguas, las más limitadas y han caído en desuso en este momento. Ellas representan la relación 1:1 y 1:n muy bien, pero necesitan almacenamiento redundante para las relaciones m:n. Se basan fundamentalmente en la arborescencia (árboles); cada objeto se relaciona con otros que se sitúan por encima o por debajo de él en la jerarquía, pero nunca a su propio nivel. En general son inflexibles a los cambios.



Terminología

- Raíz: vértice, no tiene arcos entrantes (a).
- Hoja: vértice que no tiene arco saliente (c, d, e).
- Padre: vértice que tiene arco saliente que lo conecta a vértices de menor nivel (b).
- Hijo: vértice al cual se conectan arcos desde un vértice padre.
- Hermanos: vértice hijo del mismo padre.

Características de la base de datos jerárquicas

Las únicas relaciones permitidas son las relaciones jerárquicas. Las relaciones son generadas por punteros.

Ventajas

- Simples.
- Rápidas.

Inconvenientes

- La realidad no es siempre jerárquica.
- No acepta las relaciones muchos a muchos.
- No disminuye la redundancia de datos, las aumenta.

Base de datos relacionales

Las bases de datos relacionales son el tipo más común y corriente de base de datos. Este tipo de base de datos se fundamenta en la confección de tablas o matrices, siendo estas últimas las relaciones existentes entre los diversos individuos.

Por lo tanto una matriz o relación, estaría compuesta por "filas y columnas".

- Filas: también llamadas registros, representan un conjunto de valores. En terminología de las bases de datos las filas representan las características o atributos de cada uno de los individuos.
- Columnas: representan, dentro de la matriz, distintos atributos registrados para cada individuo.

Para cada registro existe una celda o grupos de celdas. En cada columna, a lo largo de la fila, dicha celda contiene un valor que corresponde al atributo de esa columna para el individuo de

la fila. Un requerimiento para la BDR es que cada una de las celdas debe contener exactamente un solo valor o dato, ya sea de tipo cadena, numérico o lógico; tampoco admite valores ausentes o nulos.

El número de columnas, que es el mismo que el de células de cada fila, es un número fijo que constituye el grado de relación. El número de filas puede variar de acuerdo al alta o baja de individuos.

El conjunto de las tuplas que forman la matriz o tabla definen la relación. Una tupla es un registro convencional; un conjunto de tuplas corresponde a un conjunto de registros, por lo tanto una relación y un archivo son entidades similares, ambos contienen todos los datos acerca de individuos de una población.

Características de una relación:

- Un registro es un conjunto ordenado, los valores de atributos para todos los registros deben satisfacer el orden que se ha dado.
- Cada registro es único; todas deben diferir al menos en un atributo de las demás.
- Todas las filas deben estar completas, cuando tenemos una registro deben existir n valores para cada una.
- Sólo se admiten valores simples para cada registro.
- Varios atributos pueden tener un mismo dominio de atributo.

Características de la base de datos relacionales

Ventajas

- Son más flexibles.
- Los conceptos están mejor establecidos.
- Son más comunes.
- Existe transportabilidad de los datos entre todo tipo de ordenador.
- Gran cantidad de productos (Access, Oracle, Ingres, DB2, Info, SQL Server, Informix, Progress, Dbase, Fox, Rbase, etc.)

Inconvenientes

- Teóricamente son más lentas que los otros modelos.

Base de datos orientada a objetos

Las Bases de datos orientadas a objetos (BD OO) parten de una organización jerarquizada de la base de datos. Se crean “familias” de objetos.

Algunos conceptos y definiciones claves de este tipo de base de datos.

Objeto: colección de elementos y datos estructurados e identificados por una referencia única. Los Objetos se definen por sus variables geográficas y temáticas, así como por los métodos y operaciones que les afectan. Por ejemplo: el objeto vivienda, el objeto parcela.

Propiedad: característica o atributo de un objeto. Por ejemplo: nomenclatura catastral.

Clase o familia: grupos de objetos con idénticas propiedades. Por ejemplo: parcelas urbanas.

Método: operación asociada a una clase. Por ejemplo: cambio de propietario a una parcela.

Desencadenadores o “Demonios”: Procedimientos que se ejecutan de manera continua y que inician una acción ante una condición determinada. Por ejemplo: cambiar el código de término municipal.

Herencia: transmisión de propiedades de una clase (antecedente) hacia una subclase (descendente). Por ejemplo: las parcelas urbanas edificadas heredan las propiedades de las parcelas urbanas.

Características de la base de datos orientadas a objeto

Ventajas

- Los conceptos están mejor establecidos, pero son más complejos de definir.
- No están difundidas como las bases de datos relacionales.

Inconvenientes

- Teóricamente son más lentas que los otros modelos.

&&&&