

REPOSITORIO ACADÉMICO DIGITAL INSTITUCIONAL

Administración de bases de datos en sistemas de información

Autor: Fernando Víctor Medina Ortiz

**Tesina presentada para obtener el título de:
Lic. En Sistemas computarizados [sic]**

**Nombre del asesor:
Sergio Fco. Barraza I.**

Este documento está disponible para su consulta en el Repositorio Académico Digital Institucional de la Universidad Vasco de Quiroga, cuyo objetivo es integrar, organizar, almacenar, preservar y difundir en formato digital la producción intelectual resultante de la actividad académica, científica e investigadora de los diferentes campus de la universidad, para beneficio de la comunidad universitaria.

Esta iniciativa está a cargo del Centro de Información y Documentación "Dr. Silvio Zavala" que lleva adelante las tareas de gestión y coordinación para la concreción de los objetivos planteados.

Esta Tesis se publica bajo licencia Creative Commons de tipo "Reconocimiento-NoComercial-SinObraDerivada", se permite su consulta siempre y cuando se mantenga el reconocimiento de sus autores, no se haga uso comercial de las obras derivadas.





UNIVERSIDAD VASCO DE QUIROGA

LICENCIATURA EN SISTEMAS COMPUTARIZADOS

**“ADMINISTRACIÓN DE BASES DE DATOS
EN SISTEMAS DE INFORMACIÓN”**



**QUE PARA OBTENER EL TÍTULO DE:
LICENCIADO EN SISTEMAS COMPUTARIZADOS**

**PRESENTA
FERNANDO VICTOR MEDINA ORTIZ**

**ASESOR
ING. Y M.A. SERGIO FCO. BARRAZA I.**

INDICE

A Dios,

por iluminar siempre mi camino y por permitirme lograr mis metas

A mis padres,

por sus consejos y apoyo, que han sido la base para el desarrollo de mi vida profesional

A mi asesor de tesis,

por la asesoría brindada para el desarrollo del presente trabajo y por su orientación durante mis estudios universitarios

VII. DISEÑO DE BASE DE DATOS	pág. 26
vii.i) Principio de diseño de base de datos	pág. 26
vii.ii) Arquitectura de base de datos.	pág. 27
vii.iii) Lenguajes	pág. 31
vii.iv) Interacción con un programa de aplicación	pág. 36
VIII. INFLUENCIA DE LOS MODELOS DE DATOS	pág. 38
EN EL DISEÑO DE BASE DE DATOS	
viii.i) Estructura de datos	pág. 38
viii.ii) Diseño mediante un modelo de datos	pág. 40
relacional	
viii.iii) Diseño mediante un modelo jerárquico	pág. 43
de datos	
viii.iv) Diseño mediante un modelo de datos	pág. 46
reticular (o de red)	
IX. ADMINISTRACIÓN DE BASE DE DATOS	Pág. 49
ix.i) Designación de un administrador de	pág. 51
base de datos (DBA).	
ix.ii) Establecimiento de un diccionario de	pág. 56
datos	
ix.iii) Validación de las entradas	pág. 63
ix.iv) Medios de respaldo.	pág. 65
ix.v) Seguridad y confidencialidad.	pág. 66
ix.vi) Control de operaciones concurrentes.	pág. 70

CONCLUSIONES pág. 73

BIBLIOGRAFIA pág. 75

I. INTRODUCCION

La era actual se caracteriza por generar, recopilar y procesar una gran gama de información. La cantidad de estas actividades que hoy en día se realiza no tiene precedentes en la historia. Más aún, el volumen aumenta a un ritmo vertiginoso y continuará así en el futuro.

Debido a la necesidad de contar con medios que proporcionen información rápida y confiable, el hombre ha ideado métodos que le ayuden a optimizar el manejo, almacenamiento y procesamiento de la misma. Una de las formas que más han auxiliado al ser humano a mejorar dichos procesos ha sido el uso de la computadora.

De hecho, la informática nace de la idea de ayudar al hombre en aquellos trabajos rutinarios y repetitivos, generalmente de cálculo y gestión, donde es frecuente tener tareas reincidentes.

Las computadoras son una herramienta muy útil en muchos aspectos de la vida del hombre moderno. Éstas se utilizan para una simple impresión de una carta, para controlar y manejar el lanzamiento de cohetes espaciales, o ayudar a resolver problemas específicos a través de sistemas expertos.

A pesar de que éstas son un instrumento de vital importancia para cualquier empresa, es necesario contar con sistemas especializados que apoyen en la toma de decisiones de cualquier administrador, como pueden serlo las bases de datos.

En los últimos años ha surgido una tecnología esparcida y diversa para estructurar bancos de datos integrados y sistemas con esa orientación, que comprende conocimientos, herramientas, prácticas y sistemas comerciales de programación (software).

El manejo acertado de datos es la clave para el apoyo exitoso de la información administrativa. Sin un manejo adecuado, datos vitales y esenciales no estarán disponibles para respaldar las operaciones de una empresa u organización.

De igual modo, la administración deficiente de datos a menudo produce pérdida de la integridad y de la confiabilidad de los mismos y de la información que están basados en dicha administración.



La administración de los datos es la base de la integración eficaz del almacenamiento, en los entornos con automatización. Las estructuras de archivos tradicionales son muy eficientes para el procesamiento de transacciones pero limitan el uso de los datos requeridos en un entorno integrado. Las bases de datos, por lo tanto, están diseñadas para destacar las relaciones naturales entre los datos y para facilitar la formulación de preguntas de diversos sucesos que ayuden a los administradores a tomar decisiones.

Los sistemas de administración de base de datos no son sistemas de información, pero a menudo los dos se utilizan conjuntamente. Mientras que un sistema de base de datos no da ningún significado a los datos, esa es la misión de un sistema de información, ofrece los medio para almacenar y recabar datos y torna más sencilla la definición, por parte del usuario, del contenido de reportes.

Recientemente, se ha iniciado la interrogación de diversos sistemas de aplicaciones computarizados y sus archivos asociados, dentro de sistemas de información más capaces y provechosos centrados en bancos de datos, con vías a satisfacer a las comunidades de usuarios que crecen cada vez más y son más sofisticados.

El presente trabajo está enfocado a conocer la importancia del manejo de las bases de datos dentro de los sistemas de información dentro de las actividades administrativas.

Los capítulos por los que está compuesto, constan de la siguiente temática:

En el capítulo cinco, se definirán le concepto de sistema y de sistema de información.

En el capítulo seis se manejará la definición de base de datos, también su propósito y sus objetivos.

En el capítulo siete, se analizará la evolución de los sistemas de bases de datos, las características de los sistemas de bases de datos y los tipos de sistemas de bases de datos.

En el capítulo ocho, se estudiará la arquitectura de los bancos de datos y los diferentes lenguajes relacionados con el manejo de los mismos.

En el capítulo nueve, se examinarán las estructuras de datos, y los modelos de diseño de bases de datos.

II. ANTECEDENTES

El capítulo diez, tratará del administrador de bases de datos, validación de las entradas, diccionario de datos, medios de respaldo, seguridad y confidencialidad de la información y control de las operaciones concurrentes.

Por último, se manejará una sección con las conclusiones de la información analizada.

II. ANTECEDENTES

El presente trabajo no es pionero en el tema a tratar, se han hecho múltiples investigaciones al respecto. A continuación se citan algunos acontecimientos importantes que sirven de base para la elaboración y desarrollo de éste:

- Antes de que surgieran las computadoras de la tercera generación, la primera base de datos fue instalada en 1965. Los archivos estaban organizados en forma secuencial y el software ejecutaba las operaciones de entrada/salida de los dispositivos de almacenamiento y pocas cosas más. La codificación incluida en los programas era la responsable de la organización de los datos y no había independencia entre ellos.
- La mayor parte de los archivos servían sólo para una aplicación. Para poder ser utilizados en otras aplicaciones debían estar organizados de otra forma y con campos diferentes, lo que llevaba consigo un alto grado de redundancia entre los archivos.
- La llegada de los dispositivos de direccionamiento directo permitió al usuario el acceso a un determinado registro sin tener que pasar por los demás, pero era tarea del programador el especificar el tipo de direccionamiento utilizado.
- Al reconocerse la naturaleza cambiante de los archivos y del dispositivo de almacenamiento, el software permitió modificar la distribución física de los datos sin tener que alterar la estructura lógica. Los archivos seguían estando diseñados para una aplicación determinada o para un grupo de aplicaciones muy similares.
- La evolución del procedimiento de los datos comerciales llevó a independizar los programas de aplicaciones no sólo de los cambios del hardware y del aumento del volumen de los datos, sino también de la adición de nuevos campos a la estructura lógica del archivo.
- A partir de los últimos años de la década de los sesenta muchas organizaciones gubernamentales, industriales y educativas han desarrollado grandes sistemas de información.

III. OBJETIVO GENERAL

Analizar la importancia de las bases de datos, como apoyo a las empresas, para optimizar el manejo y procesamiento de información.

iii.i) Objetivos específicos

- Analizar los principales objetivos de las bases de datos.
- Explicar las funciones del administrador de base de datos y su relación con la alta gerencia.
- Determinar cómo se planean, organizan y controlan los bancos de datos.
- Establecer formas de lograr la seguridad y confidencialidad de la información.
- Conocer la relevancia que tienen los sistemas de información dentro de las empresas.

IV. SISTEMAS DE INFORMACIÓN

Un sistema es un conjunto de componentes que interaccionan entre sí para lograr un objetivo común. Todo sistema organizacional depende, en mayor o menor medida, de una entidad abstracta denominada sistema de información.

Este sistema es el medio por el cual los datos fluyen de una persona o departamento hacia otros y puede ser cualquier cosa, desde la comunicación interna entre los diferentes componentes de la organización y líneas telefónicas hasta sistemas de cómputo que generan reportes periódicos para varios usuarios. Los sistemas de información proporcionan servicio a todos los demás sistemas de una organización y enlazan todos los componentes en forma tal que éstos trabajen con eficiencia para alcanzar el mismo objetivo.

En cualquier organización se deben tomar decisiones sobre muchos asuntos que se presentan con regularidad y para hacerlo se requiere de cierta información. Dado que los procesos de decisión están claramente definidos, entonces se puede identificar la información necesaria para formular las decisiones.

Los sistemas de información están formados por subsistemas que incluyen hardware, software, medios de almacenamiento de datos para archivos y bases de datos.

La mente de un administrador recibe datos de una gran variedad de fuentes; procesa y guarda esta información y después la comunica a otros, estas características de la mente humana las comparten los sistemas de cómputo. Éstos y los administradores procesan, guardan y comunican conjuntamente la información de la organización, y por esta razón aquí se ve a los administradores como partes integrales de los sistemas de información de la organización.

Los sistemas de información de una organización son completos y diversos y el desarrollo de éstos es un reto difícil. Los sistemas de información de una organización deben proporcionar información acerca de sus múltiples medios a sus administradores, y mientras más dinámicos sean estos ambientes, más crítica se convierte la información.

.....

iv.i). Tipos de Sistemas de Información

Sistemas de Procesamiento de transacciones

Una transacción es un proceso que afecta o implica a una organización o empresa.

Los sistemas de procesamiento de transacciones procesan los datos referentes a las actividades de la empresa. El procesamiento de transacciones se centra en la operación ordenada de cualquier organización.

Sistemas de Información gerencial

El procesamiento de información proporciona información a los gerentes o administradores para decidir qué acción realizar en una situación particular.

El procesamiento de información analiza los datos almacenados de las transacciones para evaluar las posibilidades y seleccionar el curso de acción que se va a seguir.

Los sistemas de información gerencial o sistemas de reportes de gerencia se enfocan al apoyo de toma de decisiones.

Un sistema para poder dar la información requerida se puede diseñar y desarrollar para satisfacer la necesidad de fundamentar las decisiones por parte de la gerencia. Las decisiones se presentan reiteradamente, los reportes se producirán con base a un periodo agregando cada vez nuevos detalles de lo ocurrido desde la última vez que se generó información.

Sistemas de Apoyo para la toma de decisiones

Al formular decisiones estratégicas, parte del problema consiste en determinar qué información se debe reunir.

Los sistemas para el apoyo de toma de decisiones son sistemas de información que auxilian a los administradores con las decisiones únicas que no se repiten y que carecen de estructura, y son de vital importancia para el control y planeación de procesos.

Estos sistemas existen con el fin de responder a situaciones inesperadas y propias de la información.

Los sistemas expertos utilizan programas de computación que almacenan datos con el objeto de reproducir el proceso de toma de decisiones de los humanos. También tratan situaciones caracterizadas por un alto grado de incertidumbre y se debe aplicar una combinación de experiencia y criterio para tomar una decisión. Se enfocan a un área limitada, estudiando reglas, suposiciones y hechos que permitan deducir datos que lleven a una decisión.

Sistemas de Información para oficinas

Evolucionaron a partir del procesamiento de palabras para elaborar correspondencia, reportes, relaciones y documentos especiales. Están integrados por el procesamiento de palabras y la teletransmisión de datos.

Sistemas de Información computarizados

El objetivo de la computadora es que opera sobre datos para reestructurarlos y presentarlos en forma especial.

La combinación de los circuitos en una computadora representan los datos como pulsaciones electrónicas. Los datos pueden ser números, letras o símbolos. Todos los datos son convertidos en números los que a su vez quedan representados como impulsos en el sistema de cómputo.

Las computadoras almacenan interiormente datos e instrucciones y ejecutan las instrucciones de manera automática.

Las instrucciones especifican o definen una operación básica que va a ser efectuada e identifica los datos sobre los cuales se actuará y donde éstos se encuentran almacenados en la máquina. Al conjunto de instrucciones que indica a la computadora que debe de afectar una serie de operaciones o procesos se le denomina procedimiento.

.....

V. DEFINICION DE BASE DE DATOS

La unidad más pequeña de datos que se considera por lo general es el *dato elemental*, llamado también campo o atributo. Un conjunto de datos elementales constituye un *registro lógico o entidad*. Un tipo de registro lógico es un registro con una constitución particular de datos elementales. Un archivo es una colección de ocurrencias de un mismo tipo de registros. Un *banco de datos* es un conjunto de ocurrencias de varios tipos de registros en el cual tanto los tipos de registros como sus ocurrencias están inter-relacionados mediante *relaciones* específicas. El término *base de datos* es sinónimo de banco de datos.

Una base de datos es una colección integrada de datos almacenados en diferentes tipos de registro. Los registros se interrelacionan por medio de relaciones propias de datos, y no mediante su ubicación física en el almacenamiento.

Las bases de datos se distinguen de los archivos maestros comunes y de transacciones en cuatro formas significativas.

Primera, un archivo está destinado al almacenamiento. Los registros que contienen un archivo se almacenan juntos y son recabados por medio de un método de acceso (secuencial o aleatorio).

Segunda, agregar registros a un archivo para agrandarlo no lo convierte en una base de datos. La existencia de una base no está determinada por el número de registros almacenados; no es cuestión de tamaño.

Tercera, los registros referentes a distintas entidades de interés pueden almacenarse dentro de una base de datos.

Cuarta, tener bases de datos no elimina la necesidad de archivos en un sistema de información.

v.i) Propósito de una base de datos

El objetivo de almacenar y organizar datos en una base es representar las relaciones entre las entidades de interés para la empresa o institución, un resultado que no se puede alcanzar con archivos maestros individuales.

En el nivel de transacciones, la creación de archivos maestros que reflejen las relaciones físicas en los datos ofrece una manera eficaz de almacenar y recabar datos basándose en los objetivos de uso.

Por el contrario, las aplicaciones administrativas no se centran en el almacenamiento y en la eficiencia del procesado, sino en la recuperación de la información necesaria para procesar preguntas.

Debido al énfasis sobre un área específica en una base de datos integrada, el término base de datos de área se aplica algunas veces con miras a subrayar que las bases de datos no se deben enfocar a aplicaciones individuales, como lo hacen los archivos maestros, sino a áreas integradas.

El efecto general de las bases de datos es la integración del sistema de información que a su vez sirve tanto para integrar como para alterar los procesos administrativos, aumentando el acceso a los datos y transformando la administración en un proceso real.

v.ii) Objetivos de la administración de bases de datos

La administración de bases de datos consiste en la planeación, organización y control de las bases de datos de una empresa u organismo.

El efecto general de las bases de datos es la integración del sistema de información que a su vez sirve tanto para integrar como para alterar los procesos administrativos, aumentando el acceso a los datos y transformando la administración en un proceso real.

Como se mencionó, las bases de datos mejoran la capacidad de los administradores para responder a problemas administrativos especiales, por lo que proporcionan una mejor capacidad de realización de informes especiales.

.....

La información necesaria para resolver problemas especiales no puede anticiparse en forma razonable como la información necesaria para las actividades de control administrativo y de planeación. Los datos que deben usarse en problemas especiales no pueden colocarse en una base de datos con la seguridad de que serán necesarios en una forma o un formato particulares, o que serán necesarios alguna vez. Las bases de datos pueden servir a los administradores ayudándoles a organizar el sistema de información alrededor de tareas administrativas clave.

Las bases de datos deben estar estructuradas en torno a las tareas clave en cada nivel, en general con una base de datos para cada tarea clave.

En el mundo actual, la cantidad de datos almacenados en una empresa crece casi en progresión geométrica. El proceso administrativo de toma de decisiones depende de la calidad y cantidad de información. La información que se puede extraer de la base de datos es uno de los recursos más valiosos de la empresa.

La base de datos se debe diseñar, procesar y mantener adecuadamente para proporcionar información correcta en el momento oportuno a las personas autorizadas.

La función de administración de la base de datos está compuesta de personas y procedimientos; el objetivo de una base de datos es suministrar información a diferentes usuarios en las distintas áreas funcionales de la organización.

El alcance de la actividad de la administración de datos es la totalidad de la organización y el alcance de la actividad de la administración de la base queda restringida a una base de datos en particular, y a los sistemas que la procesan. La administración de la base de datos opera dentro del marco proporcionado por la administración de datos y así facilita el desarrollo y el uso de una base de datos en particular, junto con sus aplicaciones.

Muchas organizaciones tienen sistemas de información muy exitosos y no han tenido necesidad de invertir en administración de bases de datos. Con todo, para la mayoría de las empresas la única forma eficaz y eficiente de lograr el nivel deseado de apoyo a los sistemas de información es mediante la administración de bases de datos. Los siete objetivos de administrar bases de datos muestran las ventajas que pueden obtenerse: evitar la redundancia innecesaria, proporcionar flexibilidad de acceso, aportar relacionabilidad, mantener la independencia de los datos y garantizar la seguridad de los mismos.

-
- 1) Evitar la redundancia innecesaria.- A medida que en las organizaciones aumenta la dependencia con respecto a los sistemas de información se implementan más aplicaciones.

Con frecuencia cada aplicación se diseña para utilizar sus propios archivos maestros y de transacciones – tal vez debido a que los analistas no saben que los mismos datos existen en otros archivos, o porque los usuarios desean tener más control sobre sus propios archivos, o bien porque los archivos son un poco distintos. El resultado, sin embargo es la abundancia de datos que no son de interés ni para los usuarios ni para la organización.

Pero más importante aún, cuando los datos se duplican a través de los archivos no es probable que se hagan cambios en forma simultánea en todos los archivos.

Es costoso almacenar datos redundantes, y su mantenimiento resulta caro y consume mucho tiempo. Y si tienen que actualizarse diferentes archivos cuando ocurren los eventos habrán múltiples etapas de preparación e introducción de datos, y se deberá dedicar más tiempo de cómputo para procesar las actualizaciones.

En vez de permitir la redundancia no controlada, el diseño del sistema debe posibilitar compartir datos a través de las aplicaciones.

Se pueden satisfacer las necesidades comunes de datos si los mismos son accesibles a todas las aplicaciones que los requieran.

- 2) Proporcionar flexibilidad de acceso.- Frecuentemente los administradores y otros usuarios que solicitan información necesaria para la toma de decisiones suponen que pueden obtenerla porque saben que los datos requeridos se almacenan en una modalidad entendible para la computadora. Por desgracia, a menudo descubren que tales demandas no pueden ser satisfechas con rapidez debido a que:

- ❖ Los datos se hallan en diferentes archivos.
- ❖ Se tiene que escribir un programa para extraer los datos de los diversos archivos (y después puede que tenga que escribirse un programa adicional para elaborar reportes.)
- ❖ Tomará tiempo desarrollar el programa.
- ❖ Los programadores deben terminar primero otros proyectos.

Resulta irónico que los datos que sirven para las decisiones que tienen el mayor beneficio potencial, frecuentemente no sean accesibles de modo inmediato.

Estas necesidades exigen el acceso a cualquier dato sin considerar dónde se encuentra almacenado, o cómo se debe tener acceso a él.

- 3) Aportar relacionabilidad.- La relacionabilidad es la capacidad de definir las relaciones entre entidades o tipos de registros y de recabar datos basándose en esas relaciones. Es obvia la importancia de describir los atributos de una entidad cuando se definen registros. Definir la relación entre las entidades constituye la siguiente etapa.
- 4) Mantener la independencia de los datos.- Los usuarios no tienen por qué limitar su consideración de los datos por la modalidad en que el sistema de cómputo almacena los detalles. La independencia de los datos se refiere a la capacidad de tener separados los aspectos lógico y físico de los datos. Estos términos se utilizan de la siguiente manera:
 - ❖ Aspecto lógico de los datos: es el aspecto conceptual que el usuario tiene de los datos, incluyendo los nombre y el contenido de los elementos de los datos así como la manera en que están organizados; se elige cierto aspecto debido a la forma en que los datos se utilizarán por parte de una persona o en una aplicación.
 - ❖ Aspecto físico de los datos: es la manera en que los datos se organizan y almacenan en la computadora y en el almacenamiento secundario, incluyendo las estructuras de archivos que almacenan los datos y los métodos de acceso que recaban los detalles.
- 5) Asegurar el desarrollo futuro.- El cambio es un concepto clave para la comprensión de la necesidad de alcanzar el objetivo de asegurar el desarrollo futuro de las bases de datos. Es posible prever qué cambio ocurrirá en muchas áreas:
 - ❖ Claves de registros.
 - ❖ Estructuras de almacenamiento físico.
 - ❖ Dispositivos de almacenamiento.
 - ❖ Lenguajes de programación.
 - ❖ Programas de aplicación.

Rara vez se puede prever la necesidad de cambios específicos. Aún así el mantenimiento – y efectuar ajustes y mejoras en los sistemas de información – ocupa gran porción del tiempo del programador en muchas empresas.

La necesidad de cambios nunca se eliminará. Pero el objetivo es minimizar la dificultad para efectuar modificaciones cuando se vuelven necesarias.

- 6) Prever la integridad de los datos.- La seguridad con respecto a las bases de datos es el resultado de controlar el derecho de acceso a la base y la capacidad de las personas para recabar, cambiar, agregar o eliminar registros.
- 7) La seguridad es especialmente trascendente, y difícil de alcanzar en los entornos de grandes bases de datos donde los usuarios interactúan con el sistema a través de líneas de comunicación.

El objetivo de garantizar la seguridad de los datos es evitar todo acceso o empleo no autorizado de la base de datos, ya sea accidental o intencional.

Objetivos primarios de la organización de la base de datos:

- Los datos podrán utilizarse de múltiples maneras.- Diferentes usuarios, que perciben diferentemente los mismos datos, pueden emplearlos de distinta maneras.
- Se protegerá la inversión intelectual.- No será necesario rehacer los programas y las estructuras lógicas existentes (que representan muchos hombres-años de trabajo) cuando se modifique la base de datos.
- Bajo costo.- Bajo costo del almacenamiento y el uso de los datos y minimización del costo de los cambios.

-
- Menor proliferación de datos.- Las necesidades de las nuevas aplicaciones se satisfarán con los datos existentes más bien que creando nuevos archivos, evitándose así la excesiva proliferación de datos que se advierte en las bibliotecas de cintas actuales.
 - Desempeño.- Los pedidos de datos se atenderán con la rapidez adecuada según el uso de que ellos habrá de hacerse.
 - Claridad.- Los usuarios sabrán que datos se encuentran a su disposición y los comprenderán sin dificultad.
 - Facilidad de uso.- Los consumidores tendrán fácil acceso a los datos. Las complejidades internas son ajenas al usuario, gracias al sistema de administración de la base.
 - Flexibilidad.- Los datos podrán ser utilizados o explorados de manera flexible, con diferentes caminos de acceso.
 - Rápida atención de interrogantes no previstos.- Los pedidos espontáneos de información se atenderán sin necesidad de escribir un programa de aplicación, sino utilizando un lenguaje de alto nivel para averiguación o generación de reportes.
 - Facilidad para el cambio.- La base de datos puede crecer y variar sin interferir con las maneras establecidas de usar los datos.
 - Precisión y coherencia.- Se utilizarán controles de precisión. El sistema evitará las versiones múltiples de los mismos ítems de datos con diferentes estados de actualización.
 - Reserva.- Se evitará el acceso no autorizado a los datos.
 - Protección con pérdida o daño.- Los datos estarán protegidos contra fallos y catástrofes, contra delincuentes, vándalos, incompetentes y personas que intenten falsearlos.
 - Disponibilidad.- Los datos se hallarán inmediatamente disponibles para los usuarios casi todas las veces que los necesiten.

Objetivos secundarios:

- Independencia física de los datos.- El hardware de almacenamiento y las técnicas físicas de almacenamiento podrán ser modificados sin obligar a la modificación de los programas de aplicación.
- Independencia lógica de los datos.- Podrán agregarse nuevos ítems de datos, o expandirse la estructura lógica general.
- Redundancia controlada.- Los ítems de datos serán almacenados una sola vez, excepto cuando existan razones técnicas o económicas que aconsejen el almacenamiento redundante.
- Adecuada rapidez de acceso.- Los mecanismos de acceso y los métodos de direccionamiento serán lo suficientemente rápido, habida cuenta de los usos previos.
- Adecuada rapidez de exploración.- La conveniencia y necesidad de exploración espontánea se incrementarán en la medida que se difunda el uso interactivo de los sistemas.
- Normalización de los datos dentro de un organismo.- Se necesita un acuerdo interdepartamental sobre los formatos y las definiciones de datos.
- Diccionario de datos.- Se necesita un diccionario de datos que defina todos los ítems de datos.
- Interface de alto nivel con los programadores.- Los programadores de aplicaciones deben disponer de medios sencillos para pedir datos y estar aislados de las complejidades internas de organización y direccionamiento de los archivos.
- Lenguaje del usuario final.- Un lenguaje de averiguación de alto nivel o en lenguaje para la generación de reportes permitirá que los usuarios finales se vean libres de tener que escribir un programa de aplicación convencional.
- Controles de integridad.- Siempre que sea posible, se recurrirá a chequeos de límites y otros controles para asegurar la exactitud de los datos.

.....

-
- Fácil recuperación en caso de fallo.- Recuperación automática sin pérdida de transacciones
 - Afinación.- La base de datos debe ser afinable, para mejorar su desempeño sin exigir la reescritura de los programas de aplicación.
 - Ayudas para el diseño y la supervisión.- Ayudas que permitan al diseñador o al administrador de datos predecir y opinar el desempeño
 - Migración o reorganización automática.- Migración de datos u otra reorganización física previstas para la mejora del desempeño.

VI. SISTEMAS DE ADMINISTRACION DE BASE DE DATOS

Un sistema de administración de base de datos (SABD) es un sistema programático que posibilita la creación, empleo y mantenimiento de bases de datos. Sin embargo, en vista de que es independiente de la aplicación puede utilizarse en una gran variedad de entornos y áreas de aplicación. No depende de ningún programa de aplicación o archivo específico, pero se puede usar para hacer que los datos estén disponibles para varios programas de aplicación.

Es un programa que no está ligado con ningún conjunto de archivos en particular. Elimina la necesidad de vincular estrechamente la estructura de los datos en el programa, y los datos almacenados en dispositivos de almacenamiento secundario. Las características integradas en sistemas de administración de base de datos hacen posible separar las consideraciones programa/usuario de los datos, de la modalidad en que los datos en cuestión están almacenados.

vi.i) Evolución de los sistemas de administración de base de datos

Etapa 1.- Cuando se introdujeron las computadoras, se tenía que confiar exclusivamente en los programadores para desarrollar las aplicaciones de la computadora en la organización. Los programas de lenguaje de máquina especificaban la dirección de todas las posiciones de almacenamiento en números binarios. La programación no estaba generalizada para manejar una variedad de datos. Por lo tanto, toda operación de lectura/escritura, por ejemplo, tenía que volverse a codificar en detalle junto con las revisiones para la detección de errores. Los registros, las direcciones de almacenamiento, los códigos de procesamiento numérico, etc., tenían que volver a formularse cada vez que se efectuaba una operación –nada era implícito o transparente para los usuarios.

Etapa 2. – Muchas actividades de procesamiento se repiten dentro de los programas y entre éstos.

Ya que una de las clases más comunes de actividades es el ingreso de datos y egreso de resultados, se deduce que la generalización de los procedimientos para la entrada/salida, con instrucciones sencillas para controlar estas operaciones, ahorrarán tiempo en la programación.

Tales prioridades basadas en el ahorro de tiempo presentadas por los usuarios de los sistemas de cómputo, las que llevaron a la creación de una programación generalizada de nombre sistemas de control de entrada/salida (SCES).

Estos sistemas se sirven de macroinstrucciones (una macroinstrucción reemplaza a varias instrucciones detalladas de nivel inferior) para reducir la complejidad de la entrada/salida y para ejecutarlas con mayor eficiencia cuando un programador las escribe una por una. El SCES fue creado para reemplazar el complejo conjunto de instrucciones en lenguaje de máquina en la descripción de archivos, dispositivos, formas y direcciones de almacenamiento direccionadas.

Debido a la generalidad del SCES, los mismos procedimientos se pueden emplear una y otra vez en los mismos o en diferentes programas sin tener que reprogramar las instrucciones.

Etapa 3. – Cuando el uso de las computadoras creció, las capacidades de procesamiento también aumentaron. En cierto sentido, los sistemas de cómputo se tornaron más complejos debido al creciente refinamiento electrónico del CPU y de los dispositivos periféricos.

El desarrollo de organizaciones de archivos y de métodos de acceso, agregaron complicación al sistema de control de entrada/salida. Los métodos de acceso hicieron posible trabajar sobre registros de datos sin considerar el contenido real del registro.

Etapa 4. – La programación de uso general llegó a ser cada vez más importante. Adicionalmente a los métodos de acceso se desarrollaron sistemas de manejo de archivos con el objeto de minimizar los problemas que se presentaban al trabajar con un único archivo y con los registros individuales que contiene. Los sistemas de manejo de archivos exigen que el programador o el usuario proporcione el formato de los datos de un archivo y el que se desea para los resultados. La programación generalizada del sistema realiza la lectura, la extracción y las operaciones de cálculo o comparación necesarias para reducir la salida solicitada.

Los sistemas de manejo de archivos operan únicamente sobre un archivo a la vez. Esta restricción se tornó frustrante a medida que los sistemas evolucionaban hacia el empleo de datos almacenados a través de múltiples archivos.

Etapa 5. – Los sistemas de administración de base de datos surgieron de la necesidad de una generalización aún mayor de la programación que fuera capaz de procesar en forma simultánea los datos de varios archivos. Además, tuvieron cada vez mayor importancia los medios que permitían que varios programas introdujeran los datos de diversas maneras y que posibilitaban a los usuarios disponer de datos en diversos formatos. Junto con estas necesidades se presentó el deseo de mayor seguridad y control de la integridad de los datos y de mayor flexibilidad en los formatos de salida. Los sistemas para la administración de bases de datos se transformaron para satisfacer las demandas, éstos son significativos, debido a que difieren de los sistemas de manejo de archivos, en que pueden ocuparse de más de un conjunto de datos a la vez.

vi.ii) Características de los sistemas de administración de base de datos

Un sistema de manejo de base de datos se define por sus peculiaridades, las cuales incluyen las aptitudes para desarrollar estructuras de datos, definir datos, e interrogar y actualizar la base de datos después que ha sido creada, así como de manejar las definiciones de la estructura del almacenamiento.

- Creación de bases de datos. Una base de datos debe ser conocida por el Sistema de administración de bases de datos, de suerte que aquélla pueda quedar bajo el control del SABD y emplearse para el procesamiento. Con el fin de elaborar una base de datos, es menester definir sus especificaciones valiéndose de un lenguaje de definición de datos que sea parte del SABD.

La creación también puede incluir la recreación de una base de datos.

Algunos sistemas de administración de base de datos dejan que los conjuntos de datos existentes, elaborados por otros programas se redefinan para mostrar otra estructura y queden bajo el control del SABD. La creación de una base de datos, se funda en una estructura de datos subyacente.

- Estructuras de datos. La peculiaridad de estructuras de datos de un SABD permite a los usuarios utilizar datos sin tener que preocuparse de cómo está almacenada la información. Los administradores de las bases de datos pueden organizar los datos de acuerdo a las necesidades de los usuarios valiéndose de subesquemas, y el SABD agrupa los elementos y los registros necesarios para el usuario.

No interesa cómo están almacenados los datos, qué archivo contiene, o si el registro contiene más elementos que los necesarios para una determinada aplicación. Simplemente se requiere especificar los datos que se desean; el SABD los ubicará y reunirá para su uso.

- Definición de los datos. Mediante la independencia de los datos, éstos se mantienen separados de los programas que los procesan. Así que los datos pueden variar sin que necesariamente se tengan que efectuar cambios en los programas que los emplean. Del mismo modo, los usuarios pueden transformar los requerimientos de los datos en programas sin tener que reorganizar la manera en que la información está almacenada, pudiendo estar disponible para atender a varias personas con diferentes necesidades sin que se tenga que alterar la estructura original.

Para conseguir la independencia de los datos, los SABD se encargan de la definición de datos estandarizada. Es decir, todos los elementos de los datos se deben amoldar a las especificaciones estándares de tipo y longitud según se establece en la definición.

Los datos se definen según uno de los diversos tipos de definición, los cuales son de forma libre, forma narrativa, de palabra clave, de con separadores y de posición fija.

La **forma narrativa** es una posición o enunciado en que sigue la estructuración del idioma inglés aplicado a la información acerca de la estructura de los datos, con algunas restricciones respecto a la sintaxis utilizada en particular.

.....

La **forma de palabra clave** consiste en una sucesión de pares de atributo-valor (descripción) pero eliminando la mayoría de las "palabras de relleno" (términos no esenciales), que se encuentran en la forma narrativa.

Si se fija la sucesión de los términos, o sea, si se deben señalar especificaciones valiéndose de los comandos dispuestos de cierta manera, se utiliza la **forma de separadores**. Un símbolo especial indica el fin de cada comando, se requiere que se inserten los pares atributo-valor, separados por un símbolo especial. Finalmente, en algunos sistemas cada elemento de la definición aparece en una posición fija para su ingreso. Con frecuencia se proporciona una forma impresa o cuestionario con objeto de guiar al usuario al enunciar las especificaciones.

- Definición de la estructura del almacenamiento.- Los datos pueden almacenarse con diferentes estructuras de almacenamiento (por ejemplo, secuencial, aleatorio o indexada), o modos de preservación física.

Los administradores de bases de datos seleccionan estructuras de almacenamiento de acuerdo con la forma en que los datos se emplearán en las aplicaciones.

- Interrogación.- En la interrogación, los datos se escogen de una base de datos y se les extrae o copia para su procesamiento; después los resultados se presentan con formato en pantallas de visualización, en reportes impresos o en una forma utilizable por máquina para procesamiento posterior. A fin de que pueda ocurrir la interrogación, la parte de la base de datos sobre la cual se va a efectuar dicha interrogación se debe identificar y proporcionar a la computadora. Después de esto, se exponen a la computadora los criterios y las instrucciones para el procesamiento de los datos extraídos.
- Actualización.- Actualizar una base de datos significa cambiar los valores de todos o algunos de los elementos de los datos de dicha base. Esto no significa cambiar la estructura lógica de una base, los criterios de selección o los procedimientos de seguridad. La función de actualización requiere el uso de cinco clases de información:
 - a) Descripción de la parte de los datos de la base que va a actualizarse.
 - b) Datos que se hallan actualmente almacenados en la parte de la base de datos que se actualizará.

-
- c) Descripción de los datos actualizados que se van a aplicar a la base de datos.
 - d) Datos actualizados.
 - e) Reglas de procesamiento a seguir al aplicar los datos actualizados a la base de datos.

Esta característica de actualización hace posible agregar nuevos datos a la base, cambiar elementos ya existentes de los datos o bien eliminar datos obsoletos. Las actualizaciones se efectúan procesando los datos de transacciones contra la base de datos, sirviéndose de las definiciones que se han especificado para el tipo, la longitud, etc.

vi.iii) Tipos de sistemas de administración de bases de datos.

- a) **Sistemas de lenguaje anfitrión.**- Éstos son extensiones o mejoras de los lenguajes orientados a procedimientos que se utilizan comúnmente. El COBOL ha sido el lenguaje anfitrión predominante, aunque algunos sistemas de administración de bases de datos hacen uso de otros lenguajes.

El lenguaje orientado a procedimientos es un "anfitrión" en el sentido de que proporciona el "alojamiento" o marco básico para las actividades de administración de bases de datos. Adiciones o extensiones a la estructura básica de comandos del lenguaje permiten desarrollar la administración de bases de datos. Para crear la interfaz conveniente, se producen con frecuencia nuevos enunciados y se combinan con las instrucciones existentes para efectuar las actividades de almacenamiento y acceso a los datos.

Las actividades para la administración de bases de datos pueden realizarse añadiendo al lenguaje de procedimientos, que sirve como anfitrión, los comandos para la descripción y el manejo de los datos.

El usuario de un sistema de lenguaje anfitrión se considera como un programador de aplicaciones; escribe conjuntos de enunciados o instrucciones que se ejecutan secuencialmente como en un típico lenguaje orientado a procedimientos. Aunque los usuarios están aislados de la estructura de almacenamiento de los datos, ejercen casi el mismo grado de control de procedimientos sobre los programas que si trabajaran en COBOL.

-
- b) **Sistemas autónomos (o autosuficientes).**- Éstos no dependen de un lenguaje anfitrión de procedimientos para interfaz, sino que se sirven de un nuevo lenguaje propio. Los sistemas autónomos están destinados a manejar cierto conjunto de funciones para la administración de bases de datos por lo que la programación por procedimientos (o lenguajes) ordinarios no será necesaria.

Estos sistemas están organizados alrededor de un conjunto de comandos generales como los que permiten la consulta de la base de datos o la actualización de registros. Estas actividades, a su vez, están controladas mediante comandos de alto nivel, por lo que no es posible especificar operaciones con el detalle que caracteriza a un sistema de lenguaje anfitrión. Por el mismo motivo, el usuario no tiene que escribir instrucciones muy detalladas sobre cómo introducir y procesar los datos.

- c) **Interfaces alternas.** Cada tipo de sistema de administración de base de datos debe incluir modalidades para que programadores profesionales o usuarios finales, interactúen con el sistema a fin de lograr sus objetivos. A los programadores les interesa desarrollar programaje que funcione eficientemente. Los usuarios finales, por lo contrario, desean recuperar con facilidad información sin tener que especificar gran número de detalles técnicos.

La interfaz del programador es una función dentro de los sistemas autónomos o de alto nivel que hace posible programaciones de bajo nivel cuando así se requiere para llevar a cabo ciertas funciones de procedimiento que no se encuentran en los lenguajes autónomos. Exige el desarrollo de programas orientados a procedimientos en el nivel de detalle del programador de computadora, y no en el usuario.

Los sistemas de lenguaje huésped ofrecen mayor flexibilidad puesto que no se limitan a un conjunto específico de funciones u operaciones que pueden ser efectuadas en la base de datos.

La interfaz del usuario permite a un programador no profesional o a un usuario no programador interactuar con el sistema de base de datos.

Una modalidad alterna de interfaz del usuario está dirigida al sistema: el programaje de aplicación formula interrogantes, brinda alternativas al usuario, el cual responde seleccionando la acción adecuada. Los generadores de aplicaciones con frecuencia se desempeñan de esta manera.

Una interfaz de lenguaje natural posibilita la comunicación con el sistema mediante un lenguaje similar al utilizado en la vida ordinaria.

Por medio de una interfaz de lenguaje natural, los usuarios estarán en posibilidad de hacer preguntas y solicitar reportes con un lenguaje similar al utilizado en la conversación. El tema de cómputo a su vez traduce la narrativa, verbal o tecleada, en comandos comprensibles por él.

VII. DISEÑO DE BASE DE DATOS

vii.i) Principio de diseño de base de datos

Principios de desarrollo de sistemas especialmente relevantes para el desarrollo de las bases de datos que sirven a propósitos administrativos:

1. *Tomar una perspectiva global.* Cuando se prevé una nueva aplicación, debe verse como parte de todo el esquema del sistema de información. Con frecuencia se desarrollan nuevas aplicaciones en forma independiente cuando la necesidad aparente, y por lo tanto, el sistema total resultante es un mosaico no planeado sin un patrón coherente. Sin una perspectiva general de las operaciones de la organización y de las necesidades de bases de datos, y sin un plan general que esto implica, es muy alto el riesgo de que las aplicaciones estén unidas en una combinación errónea para formar una base de datos.
2. Usar un diseño de "arriba hacia abajo". *Este se enfoca primero a las necesidades de información de los administradores altos e intermedios y después a las necesidades de operaciones.*
3. *Pensar en la capacidad de obtener informes selectivos.* La base de datos debe diseñarse, en lo posible, para que pueda proporcionarse un informe principal para cada actividad administrativa. Este Informe especial debe contener toda la información relevante necesaria para desarrollar la tarea administrativa para la cual se generó el informe y no debe contener información de otro tipo.
4. *Proporcionar muchas bases de datos diferentes para diferentes niveles de administración.* La diferencia entre los tipos de información necesaria para las actividades administrativas en diversos niveles, y entre las actividades de control y planeación, llevo a concluir que las bases de datos deben diseñarse por tipo actividad administrativa, en un nivel particular de administración. En muchas situaciones al menos dos tipos de bases de datos en cada nivel serían útiles; una o más bases de datos pueden necesitarse para realizar actividades de planeación.

vii.ii) Arquitectura de base de datos.

La arquitectura establece la estructuración del entorno de una base de datos. Fundamenta la separación de las relaciones lógicas y físicas, y define las bases de datos mediante el esquema y los subesquemas.

Un sistema de gestión de base de datos (DBMS) es un conjunto de procedimientos, ayudas de documentación, lenguajes y programas de software que administran los archivos de base de datos.

Uno de los objetivos de un **DBMS** es proporcionar a los usuarios una visión abstracta de la información, ocultar ciertos detalles referentes a la forma en que los datos se almacenan y mantienen, pero siempre permitiendo una recuperación eficaz de la información.

La arquitectura de una base de datos debe poseer tres niveles: **interno, conceptual y externo**. Cada uno de ellos pertenece a un tipo de vista diferente: almacenamiento físico, del usuario y el del programador.

- Nivel interno.- Este es el nivel más bajo de abstracción de la información.

El nivel interno es la representación interior de una base de datos, por ello es el más cercano al almacenamiento físico. Permite describir los datos tal como están almacenados en la computadora.

El nivel interno es descrito por el **DBMS** por medio de un **esquema interno o vista interna**. Hay que recordar que para un programador un registro interno almacenado difiere de un registro lógico.

La operación de transformar registros lógicos en físicos y viceversa se llama **transformación de datos o mapeo**.

- Nivel conceptual.- El administrador de la base de datos define el nivel conceptual por medio de un esquema o vista conceptual, al decidir qué información se guarda en la base de datos.

Este nivel corresponde a la estructura organizacional de los datos de la base obtenida al reunir los requerimientos de todos los usuarios de una empresa, sin preocuparse de su organización física ni las vías de acceso.

El esquema conceptual podría contener:

- a) Los datos elementales que definen los campos, **atributos**, de los objetos de la empresa.
- b) Los datos compuestos que permiten reagrupar los campos para describir los registros, las **entidades** del mundo real.
- c) Relaciones entre los datos para conectar o relacionar registros en el procesamiento de archivos múltiples.

Nota: los archivos conceptuales no existen físicamente.

- Nivel externo.- El nivel externo representa la percepción individual de cada usuario o programador de la base de datos.

Describe únicamente la parte de los datos de interés para un usuario o grupo de usuarios. Los usuarios pueden imaginar que los archivos externos utilizados en sus programas existen tal como ellos los perciben. Pero los archivos externos tampoco existen físicamente.

El DBMS es el encargado de extraer los datos requeridos por los registros lógicos externos de uno o más registros físicos, de la base de datos, cada vez que se ejecuta una operación de E/S en un programa específico.

Por cada programa es necesario especificar un **esquema externo o subesquema o vista externa** para poder acceder de forma selectiva a un subconjunto de datos de la base integrada.

Para una base específica, hay un único esquema interno y un único esquema conceptual, pero puede haber varios esquemas externos. Cada uno de los esquemas externos puede estar definido para un grupo de usuarios.

- **Separación de relaciones lógicas físicas.-** Se sabe que los usuarios desearían considerar los datos en diferentes formas; esto es, querrán múltiples enfoques.
- Las necesidades de cualquier usuario o los requisitos de un programa de aplicación específico típicamente harán uso de solo una porción de la BD. Rara vez, si acaso, se necesitarán todos los tipos de registros o cada una de las relaciones por parte de un sistema de información.

Asimismo, se sabe que de vez en cuando se efectuarán cambios en el sistema. Todo cambia, el equipo, el software, los usuarios, y desde luego, los datos.

Pero el cambio de una área no debe afectar a las otras, ni el punto de vista de un usuario acerca de un conjunto de datos debe restringir el acceso de otros usuarios.

Es posible satisfacer cada interés precedente diseñando la arquitectura de la base de datos de manera que las relaciones lógicas estén separadas de las relaciones físicas.

Especificar las características físicas del almacenamiento y utilizar los datos significa que los usuarios deben describir los elementos de datos contenidos en un registro, incluyendo las especificaciones de longitud y tipo, que son muy importantes desde el punto de vista del diseñador de sistemas, no deben limitar el modo en que los usuarios tratan a los datos y la información.

La consideración lógica en los datos necesarios para una aplicación particular, en vez de en los detalles de almacenamiento o de la forma de acceso.

1. Consideración lógica de los datos por el usuario: forma en que cierto usuario visualiza un conjunto de datos requeridos para un programa de aplicación de una parte de los datos de la base de datos.
2. Consideración lógica de una base de datos por el administrador: forma en que es apreciada toda la base de datos cuando la administra la persona responsable del diseño de la base y del logro de los objetivos de la administración de los datos.
3. Consideración física de la base de datos: el modo en que los datos están almacenados y organizados en los dispositivos físicos.

Esquema y subesquema. La arquitectura de una base de datos se expresa o se define por medio de un esquema y un subesquema.

El **esquema** es la descripción de la consideración lógica de toda la base de datos; consiste en una lista de los nombres y atributos de cada entidad y de las relaciones que existen entre entidades en torno a las cuales se almacenan los datos.

Los casos de registros se crean cuando los valores se asocian con las estructuras de los registros. Esto hace al registro útil para un usuario final. Los casos de un esquema se producen cuando los valores de los datos se asocian con el marco de trabajo, es decir, cuando se asignan valores específicos a cada uno de los valores señalados en el esquema.

Un **subesquema** representa el aspecto lógico de los elementos de datos y los registros conservados por un usuario o por un grupo de usuarios en especial. Es un subconjunto de la base de datos y por consiguiente un subconjunto del esquema. Sólo algunos de los elementos del esquema se incluyen típicamente en un subesquema individual.

Se puede considerar que el esquema equivalente a un mapa o un plano de una ciudad, en el cual se muestran todas las avenidas, calles, puentes y viaductos, así como las relaciones existentes entre ellos.

Trazar un mapa con objeto de mostrar el camino desde la universidad hasta la casa únicamente representar algunas de las calles que aparecen en el plano de la ciudad, esto es, un subesquema.

Mediante la combinación de un esquema y subesquemas, puede posibilitarse que otros usuarios y programas de aplicación elaboren diferentes perspectivas o estructuras de datos descritos en el esquema. Por lo tanto se puede elaborar un segundo subesquema referente a la relación entre los artículos pedidos y a quien fueron ordenados.

El esquema y el subesquema son ambos utilizados por el sistema de administración de base de datos, cuya función primaria es la de servir a los programas de aplicación ejecutando sus operaciones de datos.

Vistas del sistema

El sistema de administración de una base de datos tendrá que atender ordinariamente múltiples pedidos simultáneos de datos. Debe por tanto organizar sus almacenes intermedios de modo que en ellos puedan procesarse varias operaciones al mismo tiempo.

Es posible que múltiples programas de aplicación se hallen en ejecución al mismo tiempo y que ellos utilicen diferentes subesquemas.

El sistema debería estar organizado de manera que la vista del programador de aplicaciones se restrinja al programa de aplicación, su área de trabajo y su subesquema.

Por otra parte, el programador de sistemas y otros especialistas tienen como campo de actividad la organización física de los datos.

vii.iii) Lenguajes

Para cumplir sus objetivos la base de datos necesita varios lenguajes:

Lenguaje de descripción de datos.

La descripción lógica global de los datos debe ser transformada conforme al almacenamiento físico. Tal lenguaje especificaría la asignación de los datos a los dispositivos y los espacios disponibles en los medios de registro; especificará también la manera de controlar las transferencias a los almacenes intermedios, la paginación y el desborde; finalmente, describiría las técnicas adoptadas para el direccionamiento y la exploración.

Los detalles de un esquema se comunican a través de un *lenguaje de descripción de datos*; en ocasiones también llamado *lenguaje de definición de datos*. Todo sistema de administración de bases de datos tiene su propio lenguaje de descripción de datos. Para definir un esquema, primero se le debe dar un nombre. Puesto que a veces un esquema puede resultar demasiado extenso, a menudo es útil dividirlo en subconjuntos conocidos como áreas o dominios, que son subdivisiones lógicas del esquema total.

Después que se han determinado el nombre y el dominio mediante el LDD o lenguaje de descripción de datos, se requiere identificar cada uno de los registros del esquema y los elementos de datos que constituyen cada registro.

Cada subesquema también debe ser definido, valiéndose de un LDD que contiene los comandos que hacen posible la especificación del nombre del subesquema, el nombre del esquema que se utiliza y las características de los elementos de datos que están siendo empleados en cada registro.

El DBMS está diseñado como un paquete generalizado de software que permite a los usuarios procesar de hecho una base integrada de acuerdo con sus requerimientos específicos. El administrador de la base de datos tiene la tarea de describir una estructura de la base de datos en el nivel externo apropiado para cada usuario, usando un lenguaje de descripción de datos proporcionado por el DBMS.

Idealmente, un esquema únicamente describe la estructura lógica de la base de datos. El mapeo de almacenamiento de base se debería excluir de reorganizaciones de la base.

En el sistema de base de datos, los datos requeridos por los programas individuales se describen por separado con un subesquema, cuando un programa emite un comando de E/S, el DBCS del sistema administrador de la base buscará en los directorios esquema y subesquema de la definición de los datos así como la autorización de acceso antes de acceder la base..

Lenguaje de control de medios para dispositivos

En virtud de que la conservación en dispositivos físicos está separada de las definiciones lógicas establecidas por el lenguaje de descripción de datos. Lenguaje de control de medios para dispositivos, sirve para almacenar los datos, es un conjunto de comandos aplicado por los programadores de sistemas para almacenar datos. Una presentación detallada de los comandos supera el alcance de esta descripción, se debe tener presente que la asignación de datos a los dispositivos físicos se realiza con este lenguaje. Por medio del LCMD se distribuye el espacio y se establecen organizaciones de almacenamiento, y se direccionan para indexado y excedentes de información. Los usuarios pueden suponer que todo el almacenamiento de los datos esta automáticamente bajo el control de los programadores de sistemas que se sirven del lenguaje del control de medios para dispositivos. Esta es otra manera en la cual se puede lograr independencia entre los requerimientos para el procesado de datos y las aplicaciones.

Lenguaje para el manejo de datos.

Los programas de aplicaciones comunes, que ejecutan cálculos y procesan datos, emiten comandos de SABD con miras a recabar los datos necesarios de los dispositivos de almacenamiento secundario.

.....

Esto es posible mediante la inclusión en el programa de aplicación de comandos de un lenguaje de manejo de datos, los cuales especifican la recuperación, modificación, almacenamiento o eliminación de los datos almacenados en una base de datos. También establecen y suprimen relaciones entre los datos.

El lenguaje de manejo o manipulación de los datos es el lenguaje que usa el programador de aplicaciones para acceder a los datos almacenados en la BD de suerte que pueda tener lugar el proceso en el programa de aplicaciones. Igualmente, el LMD es útil para transferir datos desde el programa de aplicaciones hacia la base de datos.

Por si solo no lleva a cabo ningún proceso o cálculo, sino que señala con precisión lo que el sistema de administración de bases de datos debe realizar con la finalidad de auxiliar el programa de aplicaciones.

El DBMS suministra un lenguaje de manejo de datos LMD, que permite a los usuarios manejar o tener acceso, a la base de datos. La manipulación consiste en insertar nueva información, o acceder, modificar o borrar la información ya existente.

Existen dos tipos de LMD que no son transferibles de un sistema a otro:

- **De procedimientos:** El usuario especifica la forma de obtener los datos deseados.
- **Sin procedimientos:** El usuario especifica los datos que desea obtener, pero no la forma de obtenerlos.

El DBMS proporciona un lenguaje de manejo de datos, para que emita proposiciones de E/S para acceder la base de datos, puede ser una extensión del lenguaje de la computadora central o un lenguaje totalmente independiente. Las proposiciones del lenguaje LMD de algunos sistemas relacionales se pueden teclear desde una terminal como comandos interactivos o se pueden emitir desde un programa en lenguaje de alto nivel como si fuera parte del mismo lenguaje.

Los programas de aplicación generalmente no son transferibles de un sistema a otro. En otras palabras, la independencia de los datos de un sistema a otro es nula. Por lo tanto, antes de adquirir un sistema DBMS para una empresa es importante hacer una evaluación cuidadosa.

.....

Lenguajes de consulta.

Así como el LMD, el DBMS proporciona el lenguaje de consulta para efectuar operaciones de E/S. Es un lenguaje semejante al inglés, que permite a los no programadores recuperar información de la base de datos sin necesidad de escribir programas; características generales:

1. Son lenguajes no algorítmicos con los cuales el usuario puede teclear posiciones de consulta conversacionales para efectuar recuperaciones o actualizaciones de base.
2. Son convenientes para recuperaciones orientadas a transacciones aisladas en vez de procesamientos por lotes voluminosos.
3. Sus proposiciones se pueden emitir directamente desde una terminal para peticiones de E/S interactivas. Sin embargo, las proposiciones de consulta de algunos sistemas relacionales pueden estar inmersas en un programa en lenguaje de alto nivel.

Lenguajes de programación

Hallándose en primer término los lenguajes que deben usar los programadores de aplicaciones. Estos escribirán sus programas utilizando un lenguaje de programación convencional, o tal vez un lenguaje de ensamble.

La mayoría de los sistemas administradores de bases de datos para macrocomputadoras proporcionan interfaces con lenguajes de alto nivel. Para operaciones de E/S, las proposiciones LMD, pueden estar en programas escritos en estos lenguajes, sin embargo, la mayoría de los DBMS para microcomputadoras no tienen interfase con lenguajes de alto nivel.

Lenguajes de Bases de Datos

El sistema DBMS proporciona al usuario un lenguaje especial de descripción de datos, para describir esquemas conceptuales y externos, así como distintos lenguajes de manejo de datos, para manejar los datos de la base. Los distintos lenguajes varían extensamente en su sintaxis, formato y capacidades, todas tienen gran influencia tanto en la rapidez y la flexibilidad de desarrollo del programa como en el desempeño del sistema después de su implantación.

El tipo de estructura lógica de la base que usa un DBMS no sólo dicta el diseño de su esquema y subesquema LMD sino también el modo de sus operaciones de Entrada/Salida.

Lenguaje para la manipulación de datos

El programa de aplicación debe dar instrucciones al sistema de administración de la base y disponer de medios para interpretar los mensajes de estado con que éste replica, diciéndole si se ha cumplido o no satisfactoriamente con el pedido. Para estos fines, el sistema de administración proveerá un conjunto de macroinstrucciones o enunciados de invocación para el programador de aplicaciones.

Estos pueden ser considerados como una ampliación del lenguaje de programación de aplicaciones, como un sublenguaje adicional, o simplemente como un conjunto de enunciados de invocación o llamada provistos por un sistema de administración en particular.

Lenguaje para la descripción de subesquemas

El programador de aplicaciones debe estar capacitado para describir los datos que usa, esto es, declarar su subesquema. Tipos de recursos disponibles:

1. Recursos para la declaración de datos en el lenguaje de programación de aplicaciones
2. Recursos provistos por el sistema de administración de base.

3. Un lenguaje independiente.

Probablemente el método de uso más común y el más conveniente es el que se basa en la inclusión de recursos para la declaración de datos en el lenguaje de programación. Suele resultar conveniente ampliara estos recursos de modo que permitan una mejor descripción de las relaciones entre datos.

Lenguaje para la descripción de esquemas

El administrador de datos debe estar en condiciones de hacer la descripción lógica global de los datos: el esquema. Algunas veces tendrá que describir tipos de relaciones o características de datos que el programador de aplicaciones no describe. Para esta finalidad es necesario tener un lenguaje para la descripción de esquemas.

Este lenguaje podría ser una ampliación de un lenguaje de programación o un recurso del sistema de administración o un lenguaje independiente.

El lenguaje para la manipulación de datos

La interface entre el programa de aplicación y el sistema de administración de la base de datos, que hemos llamado lenguaje para la manipulación de datos, está embebida en un lenguaje huésped. Es deseable que aquel lenguaje tenga una sintaxis compatible con la del huésped, porque en los programas de aplicación el lenguaje huésped y el lenguaje para la manipulación de datos aparecerán íntimamente entremezclados. En realidad, esta mezcla debería presentarse al programador como si se tratase de un lenguaje único. No debieran existir requerimientos de entrar y salir para pasar de un lenguaje a otro.

vii.iv) Interacción con un programa de aplicación

El área de trabajo del usuario en el programa de aplicación reside en la memoria de la computadora. Cada programa tiene su propia área de trabajo del usuario, en la cual los datos son movidos en todas direcciones por el sistema de administración de BD.

En el procesamiento de los datos según se especifica en los programas de aplicación, el sistema de administración de base de datos surge como una interfaz entre el programa y los datos almacenados. El proceso ocurre de la siguiente manera.

1. - Los programas de aplicación solicitan ayuda al sistema de administración de bases de datos para recabar los datos requeridos para el procesamiento.
2. - El sistema de administración de base de datos acepta y examina la solicitud. Se efectúa una comparación con el fin de asegurarse de que los datos que se solicitan hayan sido definidos con acierto en el esquema y en el subesquema.
3. - El sistema de administración de base de datos solicita operaciones de entrada/salida del sistema operativo de la computadora.
4. - El sistema operativo a su vez accede al dispositivo de almacenamiento secundario correspondiente para transferir datos a los buffers de entrada/salida del sistema de cómputo.
5. - El sistema de administración de BD lleva los datos al buffer el área del trabajo del usuario según lo solicitado en los comandos iniciales para el manejo de datos.
6. - Los datos en el área de trabajo del usuario se procesan de acuerdo con las instrucciones contenidas en el programa de aplicación.

Esto sucede exactamente como habría sido ejecutado en cualquier programa de aplicación ordinario.

VIII. INFLUENCIA DE LOS MODELOS DE DATOS EN EL DISEÑO DE BASES DE DATOS

El esquema y subesquema definen la base de datos (BD) describiendo entidades y relaciones. Sin embargo, el diseño real de la BD, el cual es el esquema expresado, es influido por los modelos de datos que pueden representar un SABD específico.

viii.i) Estructura de datos

Se ha recalculado la separación de los aspectos lógico y físico de los datos. Como organizar los datos en forma lógica, esto es, las estructuras de datos que se utilizan para implementar los modelos de datos. La estructura de listas es el elemento básico para la elaboración.

Organización de las listas. Los señaladores o indicadores, que son direcciones identificativas de registros, para reducir el tiempo de búsqueda para el procesamiento de los registros en organizaciones indexadas. El uso de señaladores en las organizaciones de registros sirve de fundamento a la organización de listas. Una lista es simplemente un conjunto ordenado de datos.

Existen cinco tipos básicos de listas: lista simple, lista de dos vías, lista circular, multilista y lista invertida. Una **lista simple** es un conjunto de registros, cada uno de los cuales contiene un indicador que señala el siguiente registro en sucesión lógica. Los diferentes registros pueden no estar situados uno a continuación de otro en una organización secuencial lógica, y aquí los señaladores sirven para resolver esta separación física de datos, permitiendo que sean procesados según una secuencia lógica. El primer registro señala el segundo registro (determinado lógicamente); el segundo, el tercero; el tercero, el cuarto, y así sucesivamente.

La inserción y supresión de registros son relativamente fáciles en la estructura de lista. Los registros pueden almacenarse en cualquier lugar del medio de almacenamiento. Se enlazan con la lista cambiando el señalador del registro al que debe seguir el nuevo registro.

Con frecuencia es deseable tener unidos elementos de la lista, de suerte que el último registro en la lista señale el primero. La estructura de **lista circular** hace que esto sea posible. Contenido en el último registro se encuentra un indicador que se enlaza con el primer registro.

Se conservan símbolos especiales para designar al primero y último registros de la lista, pero esta estructura torna imposible entrar en la lista y buscar desde cualquier punto.

Una variante en la estructura de anillo es la **lista de dos vías** (o de dos sentidos), la cual contiene no únicamente indicadores de avance sino también indicadores de retroceso, esta estructura posibilita mover hacia adelante o hacia atrás desde cualquier punto de la lista. La inserción y la eliminación son más complejas, debido a que se debe cambiar el doble de indicadores. Aun así, la capacidad para buscar desde cualquier punto de vista y en cualquier dirección pueda dar como resultado un tiempo de búsqueda más eficaz.

Multilista. Una multilista se sirve de la estructura básica de las listas para unir a todos los registros que tienen un atributo en común.

Cada lista es como un cordón. Tírese de un extremo del mismo y todos los registros referentes al mismo atributo se identificarán o reportarán. Cada eslabón de la cadena es un registro que cumple los requisitos de tener un atributo común. Los eslabones se originan en forma lógica por medio de indicadores. Físicamente los registros pueden guardarse en la base de datos con cualquier estructura de almacenamiento.

Las multilistas enlazan puntos comunes en un archivo. Una multilista es como una cadena, en donde cada eslabón es un registro que cumple con los requerimientos especificados por el usuario mediante el programa de aplicación. La única diferencia es que los eslabones están físicamente adyacentes entre sí, mientras que en la lista de la base de datos están ligados lógicamente.

El nombre de la multilista se refiere a la capacidad de seguir muchas multilistas en una base de datos.

Archivo invertido. La estructura de datos de archivo invertido (llamada también lista invertida) se vale de un índice para almacenar información referente a la ubicación física de registros que tienen atributos particulares. Existen dos tipos de archivo invertido. Un archivo completamente invertido tiene un índice para cada atributo en los registros y una entrada en el índice para cada registro que tenga un valor de datos específico para tal atributo. Cada registro del índice contiene las direcciones de almacenamiento de cada uno de los registros del archivo que satisface tal atributo.

.....

En un archivo parcialmente invertido, sólo algunos de los atributos están en el índice. Esta variación se selecciona frecuentemente porque en muchos casos es improbable que ciertos elementos incluidos en la base de datos vayan a utilizarse alguna vez para recuperar datos. En consecuencia, los diseñadores no elaboran un índice sobre tales elementos.

Este enfoque utiliza un índice para almacenar la información acerca de la ubicación de registros con atributos particulares. En un archivo completamente invertido existe un índice por cada tipo de datos en el conjunto de datos.

viii.ii) Diseño mediante un modelo de datos relacional

El modelo de datos relacional representa la base de datos como un conjunto de tablas.

Una base de datos relacional consiste en un conjunto de tablas, que tienen asignado un nombre único. El esquema de una relación es una lista de atributos y sus correspondientes dominios.

Cuando se diseña una base de datos y se establece el esquema, es importante conocer el modelo de datos subyacente del sistema de administración de bases de datos que será utilizado. El modelo de datos define cómo se muestran las relaciones entre las entidades de la base de datos.

Este modelo puede evolucionar, ya que las relaciones entre datos no necesitan estar predefinidas, además utilizar valores de los datos para implicar las relaciones. El modelo relacional de datos, se basa en una relación: una tabla bidimensional. Los renglones de la tabla representan los registros y las columnas muestran los atributos de la entidad. Las bases de datos relacionales utilizan un modelo para mostrar cómo se relacionan lógicamente los datos de un registro.

El modelo de datos relacional, creado en 1970 por E.F. Codd, es el modelo más importante utilizado entre los sistemas de base datos que en la actualidad adquieren las empresas u organizaciones. El modelo de datos en sí está basado en una relación, la que se presenta como una tabla de dos dimensiones. Los renglones de la tabla citada constituyen los registros, y las columnas, los atributos de la entidad.

.....

En una base de datos relacional, el esquema se comunica al SABD estableciendo las relaciones.

Para elaborar la tabla para una relación, las anotaciones descriptivas de los datos deben dar nombre a la relación a fin de que ésta se pueda agregar al directorio de la BD.

Una afinidad es una tabla de dos dimensiones. Cada hilera en la tabla tiene datos que pertenecen a alguna cosa o a una parte de alguna cosa. Cada columna de la tabla contiene datos referentes a un atributo. Algunas hileras se denominan tuples, y las columnas se llaman atributos.

Para que una tabla se una afinidad debe cumplir con ciertas restricciones. Primero, las celdas de las tablas deben ser de valor único, no se permite repetir grupos ni tener arreglos como valores. Todos los ingresos en cualquier columna (atributo) deben ser del mismo tipo. Cada Columna posee un nombre único y no es importante el orden de las columnas en la tabla. En una tabla no pueden ser idénticas dos hileras (tuples) y no es importante el orden de las hileras.

Para entender el modelo relacional se deben comprender dos términos: dependencia funcional y clave. Estos conceptos aluden a las relaciones entre atributos en una afinidad; éste es un buen ejemplo de por qué la terminología en la tecnología de las bases de datos puede ser confusa. Una afinidad es una tabla. No confundirla con el término relación que es una asociación entre cosas.

- Dependencias Funcionales.- Una dependencia funcional es una relación entre uno o más atributos.
- Claves.- Una clave es un grupo de uno o más atributos que identifican de modo único a una hilera.

El orden de los datos en la tabla no es significativo y tampoco implica un orden cuando los registros están incluidos en la relación. Al hacer una solicitud de información, el sistema produce una tabla que contenga la información. Además, la relación es una estructura lógica. Cuando se utilizan los datos incluidos en la relación, el usuario no necesita preocuparse acerca de los detalles del almacenamiento.

Cuando se emplean dos elementos de datos para distinguir un registro de otro, se denomina *clave concatenada* al mecanismo que incorpora este par de elementos de datos.

El proceso de cristalización de las entidades y sus relaciones en formatos de tabla usando los conceptos relacionados se llama proceso de normalización. La teoría de **normalización** está basada en la observación de que cierto conjunto de relaciones presenta mejores propiedades en un medio de actualización que las presentan otros conjuntos de relaciones que contienen los mismos datos.

La normalización es el proceso de simplificar la relación entre los campos de un registro. Por medio de la normalización, un conjunto de datos en un registro se reemplaza por varios registros que son más simples y predecibles y, por lo tanto, más manejables.

Al planear la organización de los datos que van a almacenarse, el analista debe prever la necesidad de acceder los datos para cumplir con requerimientos inesperados, objetivo que se puede alcanzar mediante la normalización de los datos.

Manejo de datos. La característica distintiva del lenguaje de manejo de datos en las BD relacionales es el empleo de tres operadores de relación: SELECT, JOIN, y PROJECT.

- *Relaciones de SELECT.* La selección conceptualmente produce una nueva tabla, una nueva relación, que se crea por medio de los renglones de la tabla inicial que satisfacen una condición. Una proyección origina conceptualmente una nueva tabla que está constituida por las columnas de la tabla original que cumplen los requisitos del procesamiento. La operación de JOIN conceptualmente produce una nueva tabla a partir de los renglones de las dos tablas que satisfacen una condición para el procesamiento.
- *Relaciones / PROJECT.* El comando Project forma una nueva tabla a partir de los datos extraídos utilizando atributos especificados en la pregunta. En otras palabras, PROJECT toma columnas de una relación.
- *Relacionales JOIN.* Las operaciones de JOIN crea una nueva relación combinando dos relaciones existentes, eligiendo los registros que cumplen con los criterios establecidos en la pregunta y eliminando los registros repetidos o duplicados. (Recuérdese que un modelo de datos relacionales no existen registros duplicados; esta característica es parte de la teoría de las bases de datos relacionales.

Ventajas del modelo relacional.

El modelo relacional es importante por dos razones. Primero debido a que las estructuras del modelo relacional son amplias y generales y pueden usarse para expresar diseños independientes de DBMS. Segundo, el modelo relacional es la base de una importante categoría de productos DBMS.

Las características más importante el modelo relacional son:

- Las bases de datos relacionales son conceptualmente muy simples y fáciles de entender.
- Las relaciones no requieren ser predefinidas, la base de datos puede transformarse para satisfacer condiciones cambiantes.
- Las relaciones quedan implícitas en los valores de los datos.

Simplicidad. El usuario final se representa con un modelo de datos. Un modelo relacional es lo que el usuario ve, y no necesariamente lo que implantará físicamente.

Consultas no planéadas. Debido a que no hay una dependencia de posición entre las relaciones, las consultas no tienen que reflejar ninguna estructura preferida y por lo tanto pueden ser de tipo procedimientos no estándar de consulta.

Interdependencia de datos. El modelo relacional elimina los detalles relativos a la estructura del almacenamiento y la estrategia del acceso de la interface con el usuario. Este modelo proporciona un grado de independencia de los datos relativamente más alta. Sin embargo, para poder usar esta propiedad del modelo relacional, el diseño de las relaciones debe ser completo y exacto.

viii.iii) Diseño mediante un modelo jerárquico de datos

El modelo jerárquico de datos sigue una estructura genealógica que relaciona entidades por medio de una relación superior/subordinado o padre/hijo.

El modelo jerárquico, los datos y las relaciones entre ellos se representan por medio de registros y ligas, respectivamente, los registros se organizan para formar conjuntos de árboles.

Un registro es un conjunto de campos (atributos), cada uno de los cuales contiene un solo valor. Una liga es una asociación entre dos registros, exclusivamente.

La estructura jerárquica de árbol se construye con nodos y ramas. Un nodo es una colección de atributos de datos que describen a la entidad en este nodo. El nodo más alto de una estructura jerárquica de árbol se reconoce como RAIZ. Los nodos dependientes se encuentran en niveles más bajos en el árbol. El nivel de estos nodos depende de su distancia del nodo raíz.

Un diagrama de estructura de árbol es el esquema de una base de datos jerárquica. Este tipo de diagrama está formado por dos componentes básicos:

- **Cuadros**, que corresponden a tipos de registro.
- **Líneas**, que corresponden a ligas.

Un diagrama de estructura de árbol especifica la estructura lógica general de la base de datos. Un diagrama de estructura de árbol es similar a uno de estructura de datos. La diferencia principal radica que en este último los tipos de registro se organizan en forma de una gráfica arbitraria, mientras que en el primero se organizan en forma de un árbol con raíz. La gráfica no puede contener ciclos. Para cada diagrama existe únicamente una instancia del árbol de base de datos.

Gráficamente, el modelo de datos jerárquico aparece como un árbol invertido, en el que el nivel superior del croquis corresponde paradójicamente a la raíz. Los nodos del árbol representan entidades. Se pueden dar dos tipos de relaciones entre los nodos del arborigrama:

- Uno a uno (\longleftrightarrow)

Una entidad en un nivel se encuentra relacionada con una entidad en otro nivel.

- Uno a muchos ($\longleftarrow\leftarrow$)

.....

Una entidad en un nivel está relacionada con ninguna, con una o con más entidades en el siguiente nivel.

Una estructura jerárquica de árbol tiene que satisfacer las condiciones siguientes:

- Un modelo jerárquico siempre comienza con un nodo raíz.
- Cada nodo consiste de uno o más atributos que describen a las entidades en ese nodo.
- Los nodos dependientes pueden aparecer en dos niveles consecutivos. El nodo en el nivel precedente se convierte en el nodo padre de los nuevos nodos dependientes. Los nodos dependientes se pueden añadir tanto horizontal como verticalmente sin ninguna limitación, el nivel sólo puede tener un nodo, al que llamamos nodo raíz.
- Cada nodo que se presenta en el nivel 2 tiene que conectarse uno y solo un nodo que se presente en el nivel 1. Cada nodo que se presenta en el nivel 3 tiene que conectarse con uno y sólo un nodo del nivel 2, y así sucesivamente. Debido a que como máximo puede haber una sola conexión entre nodos cualesquiera, los arcos no necesitan tener flechas de orientación.
- Un nodo padre puede tener uno o varios nodos hijos bajo su dependencia. Si no tiene ningún nodo bajo su dependencia no es un nodo padre.
- Cada nodo, excepto la raíz, tiene que accesarse a través de su nodo padre.

En el diseño de bases de datos los analistas determinan las entidades que se van a incluir en la base de datos y después establecen cada relación entre las entidades. Los nodos representan casos de registros que contienen los elementos de datos apropiados según lo determinado por el analista de sistemas.

Al diseñar bases de datos jerárquicas, los analistas están obligados a elegir entre diferentes alternativas que influyen en la accesibilidad de los datos.

Las bases de datos jerárquicas incluyen anomalías con respecto a:

- Inserción de registros: un registro dependiente no puede ser agregado a la base de datos sin un nodo progenitor.

-
- Eliminación de registros: La eliminación de un nodo progenitor (padre) de la base de datos también elimina a todos sus descendientes.

La repetición de registros tiene dos desventajas principales:

- Puede producirse una inconsistencia de los datos al llevar a cabo la actualización.
- Será inevitable el desperdicio de espacio.

Ventajas de un modelo jerárquico.

La ventaja mayor de un modelo jerárquico es la existencia de sistemas de manejo de base probados que usan el modelo jerárquico como estructura básica.

La relativa simplicidad y facilidad de uso del modelo jerárquico y la familiaridad de los usuarios del procesamiento de datos con la jerarquía, son ventajas importantes.

Existe una reducción de la dependencia de los datos.

viii.iv) Diseño mediante un modelo de datos reticular (o de red)

Es un método en que los datos se expresan por medio de una serie de registros, y las relaciones entre los datos, mediante ligas.

Una base de datos en red consiste en una serie de registros que están conectados entre sí por medio de ligas (links). Todo registro es un conjunto de campos (atributos), cada uno de éstos contiene únicamente el valor de un dato. Una liga es una asociación entre dos registros exclusivamente.

El modelo de datos reticular es semejante al modelo jerárquico, excepto que una entidad puede tener más de un progenitor. Por ello, los miembros pueden pertenecer a más de una relación (es decir, tener más de un propietario).

En las bases de datos de red, como en las bases de datos jerárquicas, las relaciones entre las entidades deben establecerse en el momento en que se define el modelo de los datos y se crea la base de datos.

Un diagrama de estructura es un esquema que representa el diseño de una base de datos de red. Estos diagramas están formados por dos componentes básicos:

- **Cuadros**, que corresponden a tipos de registro.
- **Líneas**, que corresponden a ligas.

Una modalidad de base datos del tipo de red que ha sido extensamente aplicada es la designada por CODASYL (acrónimo del inglés *Conference On Data System Languages*), que es una organización independiente de proveedores, investigadores y agencias gubernamentales de Estados Unidos que se estableció en 1959 en Washington, D.C., en un encuentro de los representantes de 40 importantes usuarios de computadoras, fabricantes y departamentos de gobierno que han trabajado para establecer directrices y recomendaciones para el software de computadoras. Es una organización voluntaria y esta patrocinada por las empresas a las que representa

Al trabajo de CODASYL y de sus comités, podemos decir, se deben los conceptos de bases de datos. La intención del Comité es desarrollar y recomendar técnicas y lenguajes para el análisis, implantación y operación de sistemas de procesamiento de datos, así como proporcionar estas especificaciones a los Grupos Normativos en forma de Informe Técnico de Desarrollo.

CODASYL estableció un modelo para los sistemas de administración de base de datos que destaca los conjuntos propietario-miembro. Los "propietarios" son análogos a los progenitores en el modelo jerárquico y los "miembros" dependen de los propietarios; a la relación propietario-miembro se le conoce como *conjunto*. Tanto los propietarios como los miembros son entidades. Para recabar datos de un conjunto se debe especificar al propietario y al miembro.

.....

Ventajas del modelo reticular.

La principal ventaja del modelo reticular consiste en que, como para el modelo jerárquico, existen sistemas de manejo de base de datos que utilizan, con éxito, el modelo reticular como estructura básica.

La relación varios a varios, que ocurre con frecuencia en la vida real, se puede implantar fácilmente.

Desventajas del modelo reticular.

La principal desventaja del modelo reticular es su complejidad. El programador tiene que conocer su posición en las ocurrencias de un conjunto determinado, al moverse a través de la base de datos.

Otra complicación consiste en que, cuando la base de datos se reorganiza, es posible, a menos que se tenga mucho cuidado, que se pierda la independencia de datos. Esto ocurre cuando los conjuntos se eliminan.

.....

IX. ADMINISTRACIÓN DE BASE DE DATOS

La administración de base de datos se refiere a la tecnología de computadoras requerida para organizar, almacenar, recopilar y manipular datos.

La base de datos debe administrarse como un recurso para la organización - un recurso que tiene valor y utilidad. El administrador de los datos debe desarrollar directrices y tomar las acciones convenientes, que van desde la separación de los aspectos lógico y físico de los datos para proporcionar integridad y seguridad a los datos.

La administración de las bases datos empieza con la designación de un administrador para la base de datos. Establecer un diccionario de datos, encargarse de la validación de las entradas, suministrar medios de respaldo, brindar seguridad y confidencialidad a los datos y controlar las operaciones concurrentes.

El éxito de la tecnología de las bases de datos en una organización depende en gran parte del administrador de bases de datos. Cuando se integran los datos para realizar múltiples aplicaciones en los archivos de las bases, ningún grupo debe controlar segmentos de los datos, y la distribución eficiente de éstos a todos los usuarios que pueden beneficiarse con ellos no deben inhibirse por intereses propios. También debe establecerse procedimientos explícitos para cambiar una base de datos, como puede ser para modificar los formatos de los registros o agregar y borrar estos, porque los cambios afectan a muchos usuarios. El administrador de bases de datos (**DBA**) asegura que los datos estén disponibles para que toda la organización los comparta; también está en la mejor posición de asegurar que la base de datos se desarrolle con propósitos administrativos.

La administración de la estructura de la base de datos incluye en el diseño inicial de la misma y su puesta en práctica, así como controlar y administrar sus modificaciones. El DBA queda muy pronto involucrado en el desarrollo de la base de datos y sus aplicaciones, participando en los estudios de requerimiento, ayudando a evaluar alternativas, incluyendo los DBMS a utilizar, y ayudando en el diseño de la estructura de la base de datos.

Una vez diseñada la base de datos, es puesta en práctica utilizando productos de DBMS, y los datos de la base de datos son creados.

El DBA participa en el desarrollo de procedimientos y controles para asegurar la alta integridad y calidad de la base de datos.

- Documentación.- La responsabilidad final de un administrador de bases de datos en la administración de la estructura de las mismas es la documentación. Es muy importante saber qué modificaciones han sido efectuadas, cómo fueron realizadas, y cuando fueron establecidas. Para identificar un cambio resulta de importancia mantener un registro de los procedimientos de pruebas y de las ejecuciones de pruebas efectuadas.

Otra razón para documentar con atención las modificaciones en la estructura de la base de datos es el uso adecuado de los datos históricos.

Administración de la actividad de datos.

Aunque el DBA protege datos, no los procesa. El DBA no es usuario del sistema. En consecuencia no administra valores de datos. El DBA administra actividad de datos. La base de datos es un recurso compartido, y el DBA proporciona estándares, guías de acción procedimientos de control; documentación para asegurar que los usuarios trabajan de una forma cooperativa y complementaria al procesar datos de la base de datos.

Otro aspecto de la administración de la actividad de datos es establecer derechos de proponentes de datos, derechos de acceso y de modificaciones. Dado que en los datos son recursos compartidos, pudieran aparecer problemas en relación con el proceso de derechos.

Otra preocupación de importancia para el DBA en la administración de la actividad de datos es el diseñar técnicas y procedimientos de recuperación. A pesar que DBMS lleva a cabo una parte del proceso de recuperación, los usuarios son parte crítica. El DBA, deberá anticipar fallas y desarrollar procedimientos estándar para su manejo, y los usuarios deberán saber qué hacer cuando el sistema se encuentre caído y qué es lo primero por hacer cuando el sistema vuelva a ponerse en marcha.

El DBA es responsable de la publicación y del mantenimiento de la documentación en relación con la actividad de datos, incluyendo los estándares de la base de datos, los derechos de recuperación, y el cumplimiento de políticas.

.....

ix.i) Designación de un administrador de base de datos (DBA).

El administrador de una base de datos puede ser una persona o un grupo de individuos o ser un trabajo de tiempo parcial. Su tarea consiste en administrar y proteger la base de datos a la vez que buscar el máximo beneficio para todos los usuarios. La elección depende del número de funciones que se asignen al DBA.

Características de un DBA.

El papel señalado aquí para el DBA exige principalmente comprender las necesidades del usuario y la capacidad de manejar estas necesidades. Un DBA que no puede manejar en forma efectiva la mayoría de las solicitudes de los usuarios, no cubre la responsabilidad del puesto, ya que son demasiados los detalles que necesitarán atención a la gerencia. El DBA debe poseer suficiente autoridad para manejar la mayoría de las necesidades del usuario, y esto implica la capacidad para obtener o afinar los recursos de recolección de datos, programación, computación y comunicación.

Aquellas personas que asumen las funciones de un DBA tendrán que estudiar continuamente por su cuenta, de manera que sigan siendo competentes desde el punto de vista organizativo y técnico.

Responsabilidad del Administrador de la base de datos

La primera tarea importante del administrador de la base de datos es resolver las diferencias entre varias funciones de la organización, con el fin de desarrollar una estructura conceptual y, más tarde, lógica del modelo de base de datos para la empresa. Tiene que ser el negociador entre los diferentes departamentos, no sólo para obtener un acuerdo entre ellos, sino también para llegar al acuerdo "correcto" sobre las entidades de la empresa; es función del DBA conocer exactamente cómo van los negocios y cuáles son las políticas de la administración a alto nivel y la futura dirección de la empresa. En la fase inicial del diseño de la base de datos, el DBA se deberá concentrarse en:

.....

-
- La definición de los campos de datos y las entidades de la empresa
 - La determinación de los distintos nombres que se usarán para referirse a los mismos elementos.
 - La definición de las relaciones entre los campos de datos.
 - El establecimiento de la descripción textual de los campos de datos.
 - El conocimiento de los departamentos o los usuarios que serán responsables de mantener la exactitud de los datos.
 - La determinación del uso de los campos de datos con propósitos de control y planeación.

El DBA tiene que saber cuándo y cómo usar el poder y la influencia, e implica que no debe ser un individuo completamente ajeno a los matices de la política que se practica dentro de la organización. La selección de un DBA deberá ser tomada muy en serio por el cuerpo directivo. Los criterios de selección para la posición del DBA deberán ser los mismos o semejantes a los utilizados para ocupar otras posiciones directivas dentro de la empresa. Las cualidades administrativas son necesarias debido a que el DBA tendrá tanta (o más) importancia en la satisfacción de las necesidades a largo plazo de la empresa como en la consecución de los objetivos a corto plazo.

Funciones del administrador de base de datos

El DBA actúa formalmente como mecanismo de control de un sistema que maneja ciclos de información de los usuarios. Las entradas a la función de control son resultado de la actividad de vigilancia y de las solicitudes de los usuarios insatisfechos. Para realizar la función de control, el DBA puede reestructurar la base de datos y obtener facilidades adicionales para análisis, procesamientos o comunicación. La inercia de las diferentes actividades de control puede diferir considerablemente.

El esquema es la herramienta principal de que dispone el DBA para la asignación de recursos. El esquema interno proporciona control sobre la eficiencia y la responsabilidad de los procesos, y la matriz asociada de protección controlada el acceso. La selección del esquema externo determina la visión operativa que un usuario tiene del modelo de la base de datos.

El administrador de la base de datos es responsable de equilibrar los puntos de vista de todos los usuarios:

- ❖ *Programador de sistemas*: Se encarga del almacenamiento de datos en la BD, trabajando con el enfoque físico en vez del aspecto lógico de los datos. Esta persona organiza los datos aplicando la estructura de almacenamiento acordada de manera que sea la que mejor satisfaga las necesidades de todos los demás usuarios, selecciona los dispositivos de almacenamiento y especifica los detalles del almacenado.
- ❖ *Programador de aplicaciones*: desarrolla el software, por lo común con lenguajes de tercera y cuarta generación, para producir informes, actualizar registros y ejecutar otras funciones en las que intervienen datos almacenados en la BD
- ❖ *El usuario especificador (un usuario final)*: Un agente o empleado que interroga (inquire) a la base de datos con objeto de obtener información para la toma de decisiones o que tiene acceso a ciertos registros para actualizarlos. Este usuario no requiere entender los lenguajes de programación comunes pero debe ser capaz de introducir ciertas palabras clave o códigos para obtener o modificar los registros.
- ❖ *Usuario paramétrico (un usuario final)*: funciona con base en preguntas y estructuras definidas previamente (por ejemplo, "¿De qué año son las ventas que desea utilizar?") Para interactuar con la base de datos; puede introducir o extraer datos, pero sólo a través de preguntas formuladas por el sistema. Usuarios paramétricos típicos son agentes de reservaciones de las compañías de aviación, cajeros de bancos y corredores de las casas de bolsa.

El administrador de una base de datos controla la estructura de la BD para satisfacer los requisitos y necesidades de todos los usuarios. Se deben desarrollar y hacer cumplir ciertas normas. Entre las áreas de interés más importantes se encuentran:

- a) Creación de los nombres de datos, longitudes y criterios de validación.
- b) Determinación de la propiedad de los datos.
- c) Designación de la responsabilidad para el acceso y la modificación.

El administrador de una base de datos no establece el contenido de la base de datos o los valores de datos. Esto es responsabilidad de los encargados del desarrollo de los sistemas y de los usuarios.

Las responsabilidades más importantes del DBA son las de un agente de cambio: introducir una nueva tecnología, demostrar sus beneficios y persuadir a los miembros de la organización para que la acepten. También tiene otras responsabilidades.

1. Guiar el diseño inicial de la base de datos y desarrollarla y extenderla cuando esto sea necesario.
2. Establecer estándares y sus políticas en las bases de datos.
3. Decidir el contenido de la base de datos, asegurándose de que los datos que deban incluirse al inicio se reúnan en la forma necesaria y se conviertan en una parte de la base de datos, y asegurarse de que excluyan los datos que no son convenientes por su costo.
4. Establecer y monitorear las políticas y los procedimientos de control y seguridad de las bases de datos diseñadas para asegurar que a) la actualización de la base de datos esté controlada y sea exacta; b) los datos estén disponibles para los usuarios que tengan necesidad de ellos y no para otros, y c) los datos destruidos o perdidos puedan recobrase.

Administración del DBMS. (Sistema de Manejo de Bases de Datos)

Además de administrar la actividad de datos y la estructura de la base de datos, el DBA debe administrar el DBMS mismo. Deberá compilar y analizar estadísticas relativas al rendimiento del sistema e identificar áreas potenciales de problema.

El DBA deberá analizar las estadísticas de tiempo de ejecución sobre la actividad de la base de datos y su rendimiento. Cuando el fabricante del DBMS en uso anuncie nuevas características de producto, el DBA deberá considerarlas a la luz de las necesidades generales de la comunidad de usuarios. Si decide incorporar estas nuevas características del DBMS, los usuarios deberán ser notificados y capacitados en su uso.

El DBA y la dirección.

La dirección se define como el conjunto de individuos de los cuales depende al DBA, ya sea directa o indirectamente. El DBA necesita imperiosamente apoyo y autorización de la dirección para administrar la base de datos, y a cambio debe obtener su confianza. Para diseñar una base de datos que pueda amoldarse a cambios futuros en la manera de operación de la empresa, la dirección debería idealmente proveer al DBA de la información necesaria referente a planes futuros, como:

- Prioridades principales de la empresa.
- Compromisos de la nueva empresa con respecto al tiempo del desarrollo de la nueva base de datos o de su expansión
- Limitaciones presupuestales para el proyecto (incluyendo gente, software, hardware).
- Compromisos con otras partes.
- Planes futuros.
- Posibilidades de reestructurar la organización.

El DBA se deberá comunicar con la dirección para asuntos relacionados con los datos, el diseño de la base de datos, su implantación y operación.

El DBA y el grupo de usuarios.

La base de datos se diseñara para los usuarios. La tarea del DBA consiste en diseñarla y mantenerla de manera adecuada. Los usuarios deberán tener gran confianza en la comunicación con el DBA, sentir que están tratando con una autoridad imparcial que va a ayudarlos, a ser más productivos. El DBA debe decirle a los usuarios la forma en que la base de datos satisfará sus necesidades, así como las limitaciones de ésta.

El DBA y el personal del desarrollo de las aplicaciones.

El DBA se comunica con todos los departamentos de desarrollo de aplicaciones, en relación con el medio de la base de datos. Tiene una visión general del desarrollo de aplicaciones y, por lo tanto, se puede comunicar con cada uno de los departamentos de desarrollo de aplicaciones de manera efectiva.

El DBA y el personal de sistemas.

La base de datos será establecida, preservada y accesada físicamente en un medio compuesto de operaciones de computadoras, junto con los especialistas de sistemas.

ix.ii) Establecimiento de un diccionario de datos

En un medio de base de datos, uno de los principales objetivos es el que muchos usuarios compartan datos comunes. Otro objetivo importante es proporcionar datos correctos a estos usuarios. Para realizar los objetivos de tener datos correctos, redundancia mínima y control del uso de los datos, es indispensable un mecanismo central de control.

El primer paso en el diseño de una base de datos es recabar información sobre la empresa, esto es, acerca del uso, relaciones y significado de los datos. Al avanzar el proceso de diseño es necesario almacenar información sobre modelos conceptual, lógico, interno y externo en un lugar central. La herramienta que da posibilidad de controlar y manejar la información sobre los datos en la fase de diseño, implementación, operación y expansión de una base de datos, se llama diccionario de datos.

Un diccionario de datos es una colección de las definiciones de los datos almacenados en la BD. Las definiciones incluyen la especificación de las características de los elementos de datos, tales como longitud y tipo (de caracteres o numéricos), descripciones y nombres alternos para los elementos de los datos. Constituye lo que se llama una meta-base; una base que describe a las otras bases.

El diccionario de datos almacena información sobre los datos relativos al origen de éstos, descripción, relación con otros datos, uso, responsabilidad y formato.

Un diccionario de datos también almacena información de referencia cruzada, la cual es importante para la administración de las bases de datos, ya que permite que el administrador determina con rapidez dónde se usa cierto conjunto de datos.

Un diccionario de datos DBMS es una herramienta de importancia para el administrador de la base de datos, es un catálogo accesible para el usuario de datos relacionados con la base de datos.

Algunos diccionarios de datos son activos, y algunos pasivos. Un diccionario de datos activo es un diccionario cuyas entradas son modificadas en forma automática por el software, siempre que ocurran modificaciones en la estructura de la base de datos. Los diccionarios de datos pasivos necesitan ser actualizados en forma separada, para hacer modificaciones en la base de datos, de lo contrario no reflejarán con exactitud el estado de la base de datos. Los diccionarios de datos activos cuestan más, pero aseguran su actualidad; no están disponibles con todos los productos DBMS. Los diccionarios de datos pasivos son menos costosos que los activos, pero se requiere de mayor esfuerzo para mantenerles actualizados.

Objetivos básicos de un diccionario de datos.

Uno de los objetivos básicos es permitir el manejo y la documentación de los datos. Puesto que la base de datos sirve a varios usuarios, es vital que cada uno de ellos entienda precisamente qué son los datos y que significan. Es aconsejable comenzar a recabar la información sobre los datos en un diccionario de datos en el mismo día en que el proyecto se echa a andar.

El diseñador y el usuario deben estar convencidos de que usan un término se refieren precisamente a lo mismo; de otra manera el diseñador puede construir un sistema que no le es necesario al usuario. Esto destaca otro objetivo básico de un diccionario de datos: ayudar a establecer una comunicación efectiva entre el diseñador y los usuarios, y entre usuarios. Los dos objetivos básicos de un diccionario de datos son la administración y el control de los datos como recurso, en un lugar central, a través de las fases del diseño, realización y operación, así como el establecimiento de una comunicación efectiva entre todos los que estén interesados en la base de datos.

Un diccionario de datos instalado en cada lugar puede proveerse con la información sobre los lugares físicos donde los datos están almacenados, así como el control de acceso y las restricciones de seguridad y privacidad. Con ayuda de esta información, el diccionario de datos puede decidir inteligentemente si satisfacer una consulta a la base de datos local o si el usuario está autorizado para acceder los datos remotos, en cuyo caso dicha solicitud se manda al lugar remoto.

Con el fin de lograr que el diccionario de datos sea aceptado y utilizado en un medio de procesamiento de datos, necesita tener una entrada central. La información puesta en esta única fuente se puede usar a lo largo de todas las fases de diseño y los programadores la pueden copiar en sus programas para utilizarla en estructuras de datos.

El sistema de manejo de la base de datos y el diccionario de datos.

El diccionario de datos también se puede usar en un medio ajeno a una base de datos. El diccionario de datos es un lugar central de información sobre descripciones de los datos, tales como significado, relación con otros datos y responsabilidad de tener los datos actualizados, así como tener registrados los orígenes. En un medio de base de datos, la información almacenada en un diccionario de datos es sobre los datos almacenados en la base.

Es necesario instalar software para crear y manejar el diccionario de datos de una base de datos. El software también se conoce como diccionario de datos.

El diccionario de datos debe contener todo lo que el DBA quiera saber sobre la base de datos:

- Las descripciones externa, conceptual e interna de la base de datos
- Las descripciones de los campos, registros y referencias cruzadas entre los registros de varios archivos.
- Códigos de autorización y seguridad de los datos, así como la definición de apodos, para que un determinado elemento se pueda referir con nombres distintos en programas diferentes.

-
- Los esquemas externos que son empleados por cada programa, quiénes son sus usuarios y qué autorizaciones tienen.

El diccionario de datos ideal; sus requisitos y organización.

- Modo Conceptual.- La información sobre los datos que se necesitan en el proceso de diseño del modelo conceptual incluye entidades, campos de datos o atributos que representan a las entidades y a las relaciones entre los campos.
- Modelo Lógico.- Las agrupaciones de campos de datos con los elementos clave (estas agrupaciones pueden ser los subconjuntos de las agrupaciones específicas en el modelo conceptual), el modelo de datos, los modelos externos contenidos en el modelo lógico (esto es, las diferentes trayectorias lógicas para el procesamiento de información), las transacciones lógicas, los programas y los módulos. Deben almacenarse también las interdependencias de las transacciones, los programas y los módulos.

Un diccionario de datos ideal debe ser parte integral del medio de la base de datos completa. El administrador de la base de datos debe ser responsable de la implantación del diccionario de datos. Siendo el diccionario de datos la parte central del medio de la base de datos, debe ponerse a salvo mediante copias de respaldo para evitar efectos desastrosos debido a un mal funcionamiento de sistema o la destrucción accidental de la versión activa del diccionario.

Sistema ideal de diccionario de datos.

- El diccionario de datos debe apoyar a los modelos conceptual, lógico, interno y externo.
- El diccionario de datos debe estar integrado en el sistema de manejo de la base de datos.
- El diccionario de datos debe contener varias versiones de la documentación (por ejemplo, las versiones de prueba, la versión de producción).

.....

-
- El diccionario de datos debe prestar apoyo a la transferencia eficiente de información al sistema de manejo de la base de datos.

Un sistema de diccionario de datos debe iniciar la reorganización de la versión de producción de la base de datos como resultado de los cambios a la descripción de la base de datos.

El diseño de la de un diccionario de datos está sujeto a los mismos criterios del de una base de datos. El primer estudio que debe llevarse a cabo es el de las consideraciones costo/beneficio. El segundo problema que queda por resolver es cómo y cuando un diccionario de datos puede acoplarse dentro de las operaciones de la empresa.

El software del diccionario contiene un **administrador del diccionario** que interactúa mediante:

- El **procesador del lenguaje de definición** encargado de procesar las proposiciones de entrada al diccionario.
- Un **procesador de consulta** que se encarga de procesar las peticiones de los usuarios.
- Un **generador de reportes para la realización de reportes**.
- Un **traductor del lenguaje de manejo** que permite a los usuarios modificar el contenido del diccionario de datos sin necesidad de cargarlo nuevamente.

Formas de entrada de datos al diccionario.

Existen las siguientes formas de introducir datos en el diccionario:

- Por **lote**, donde se especifican los nuevos datos que se van a añadir al diccionario.
- Por **comandos en línea**, de forma interactiva desde un terminal, indicando las propiedades del dato.
- Mediante **pantallas diseñadas** para tal fin, donde se incluyen todas las características de los nuevos datos.

-
- Por **extracción** de los bloques del control de las bases de datos mediante programas escritos expresamente para ello.

Formas de salida de datos del diccionario

- Un **listado** con la descripción de los campos, las referencias cruzadas entre archivos, los nombre de los programadores que tienen acceso a un archivo determinado, etc.
- **Consultas en línea**, bien mediante pantallas diseñadas o bien mediante búsquedas por teclado.
- Generación de los **bloques de control** para comprobar el perfecto funcionamiento de la base de datos.
- Mediante **programas** que pueden leer las anotaciones del diccionario.

Partes de un diccionario de datos

- *Campo de datos.* Atributo que describe a una entidad. Cada Campo tiene un nombre único o etiqueta.
- *Campo de grupo.* Agrupación de datos lógicamente relacionados.
- *Campo derivado.* Un campo de datos que se origina como resultado de una formula o de un calculo.
- *Sinónimos o alias.* Un campo de datos con designación distinta con el mismo significado que otro campo.
- *Homónimos.* Dos campos de datos con la misma designación pero con distinto significado.
- *Descripción del modelo conceptual.* El modelo describe las entidades de la empresa, las relaciones entre ellas y el flujo de información.
- *Descripción del modelo lógico.* El modelo es un posible subconjunto del modelo conceptual, y para utilizarlo se transforma al paquete del sistema de manejo de base de datos.

-
- *Descripción del (los) modelo (s) externo (s)*. Describen los puntos de vista de los programas de aplicación, las formas en las distintas aplicaciones ven las relaciones entre las entidades.
 - *Descripción del modelo interno*. Es la conversión física de la base de datos. Describe las relaciones entre, los métodos de acceso que se han de usar, los elementos de grupos y otros detalles de conversión física

Uso del diccionario

- En el caso de que se produzcan fugas de datos confidenciales, el diccionario debe dar la relación del personal de la empresa o de los usuarios que tienen acceso a dichos datos.
- Como fuente de información para los analistas de sistemas y los diseñadores de programas para conocer los campos, archivos, etc., que ya existen en la base de datos como punto de partida para el desarrollo de las nuevas aplicaciones.

Beneficios

Mejora las comunicaciones entre los diseñadores y los usuarios y entre los usuarios mismos. Esto trae como consecuencia menos repeticiones en la fase de diseño.

Hace más completa y sistemática la retención de la información sobre los datos.

Es más fácil implantar estándares para regir los nombres de los campos de datos, sus descripciones, usos, etc., en un lugar central que en lugares dispersos.

Capacita para implantar cambios en la base de datos más rápido, debido a la capacidad de las referencias cruzadas.

Mejora la habilidad de crear registros de la información accesada y de las personas que los hicieron.

El mejor momento de empezar a utilizar el diccionario de datos es en la fase de definición de los requisitos.

El uso de un diccionario de datos en definición de requerimientos proporcionará una sólida línea de fondo para el control del proyecto y normas, así como un denominador común para comunicaciones.

Características útiles adicionales:

- Disminuye la redundancia y la incongruencia de los datos facilitando la comunicación entre usuarios y diseñadores respecto al almacenamiento y al empleo de los datos.
- Ofrece un depósito central de información de diseño.
- Permite que los diseñadores puedan determinar cuál será el efecto de un cambio en los requerimientos de los datos de una aplicación sobre otras aplicaciones y sobre la propia
- Proporciona reportes y listados impresos del contenido de la base de datos en diferentes formatos.
- Suministra especificaciones de datos interpretadas por computadora que en algunos sistemas pueden estar integrados en la generación automática de sistemas de aplicación.

ix.iii) Validación de las entradas

La validación de las entradas, es el conjunto de acciones tomadas durante el ingreso de los datos con la finalidad de detectar errores en los datos antes de que entren en la BD. La detección de los errores en este nivel evita costosas actualizaciones después que los datos han sido aceptados en el almacenamiento y facilita corregir errores rápidamente, sin tener que esperar un largo periodo en el sistema y a una degradación de la base de datos tan grave que la BD ya no sea un recurso preciso y confiable para la organización.

.....

Validación de entradas: validación de transacciones, revisión secuencial, totales por lotes, revisiones de formato, verificaciones de razonabilidad, lista de cambios para respaldo, procesamiento duplicado y dígitos de comprobación.

Entrada de datos

Las dos características indispensables que debe tener cualquier base de datos son ofrecer un método fácil para captar datos y otro para sacarlo nuevamente.

Decisión acerca de lo que se almacenará en el archivo.

Lo primero que se tiene que hacer es definir la estructura del archivo. Esto se lleva a cabo mediante una cuidadosa consideración de los datos que se desea almacenar en el archivo o archivos y como se pretende identificarlos.

Diseño de formas para la captación de datos.

Este tipo de registro es bastante estándar y no requiere un esquema complicado. Sin embargo, se requieren algunos tipos de registros para enumerar los datos que siempre se habían recopilado en el papel de manera estandarizada. El recurso proporcionado para realizar esta función se conoce como generador de formas. Este permite al usuario crear una forma que aparezca en la pantalla como la imagen de una forma de papel.

Establecimiento de los atributos del campo.

Es posible establecer ciertos atributos del campo si se transfiere parte de la necesidad del usuario de asegurar la precisión a la computadora. Un campo podría tener un atributo que hiciera obligatoria una entrada.

.....

Otra manera de afectar los valores que se toman como datos es indicarles al programa que sólo acepte los valores que verificó y encontró en primera instancia en una lista de entradas aceptables mecanografiada en una etapa previa por el usuario o el diseñador.

- Valores por omisión.- El establecimiento de atributos y la validación de las entradas de los campos es un tipo de medida que ayuda a la rápida y precisa captación de los datos. Otra es la de los valores de inicio. Un valor de inicio es un valor adoptado por el sistema a falta de uno proporcionado por el usuario. De esta manera, el sistema de la base de datos inserta automáticamente un valor en un campo, que solo puede cambiarse si se escribe otro en su lugar.

Cálculo del tamaño de un archivo específico en términos del espacio de almacenamiento.

Al calcular la cantidad de espacio que ocuparía un determinado archivo, es importante recordar que los encabezados y los espaciamentos de la pantalla, son muy prodigiosos sin tomar más que una cantidad fraccional del almacenamiento en disco, ya que se almacenan como un conjunto individual de encabezados o una sola forma, que se borra conforme se inserta cada registro.

ix.iv) Medios de respaldo.

Siempre son necesarias precauciones para cerciorarse de que los datos almacenados en una BD estén disponibles cuando se les requiera.

Una manera de lograr la protección es conservar duplicados de los archivos y de las bases de datos denominadas copias de respaldo, las cuales pueden ser introducidas en el sistema en caso de pérdida o destrucción del material original.

- Registración dual.- Como su nombre lo indica, los mismos datos se registran en forma simultánea en dos dispositivos de almacenamiento. A medida que las transacciones se procesan contra la base de datos y ocurren actualizaciones o cambios, las modificaciones se hacen en ambas copias de los archivos.

Desafortunadamente, en caso de una falla de energía eléctrica, un problema de programática que evita que los datos se inscriban o en el caso de un desastre natural, tal vez no sería suficiente este recurso para la protección.

- Vaciado de información.- Un medio alterno de protección es copiar con cierta periodicidad la base de datos, por ejemplo, diaria o semanalmente, y almacenar la copia. Se requiere mantener un anotador hasta que se haga la siguiente copia de la base de datos, como una constancia de todas las transacciones provisionales, procesadas contra la BD. Si se pierden los datos, la BD se puede restaurar con facilidad montando la copia de respaldo y procesando todas las transacciones inscritas recientemente en el anotador contra la copia, actualizándola así. El vaciado continuo de la base de datos es, sin embargo, costoso y prolongado; adicionalmente pueden ocurrir errores que producirán datos inexactos mientras están copiando los archivos.
- Lista de cambios (tira de auditoria). - El contenido de todas las transacciones de entrada se registra a medida que los datos ingresan en el sistema, se conservan en un anotador como el que se uso en el vaciado de la información. Se lleva una lista de todas las transacciones de entrada, y se hace una copia de cada registro antes y/o después de que se anote cada actualización a estas copias de registro se les conoce como anotaciones previas y anotaciones posteriores, respectivamente. Si ocurre alguna falla de cualquier tipo en el sistema, este tipo de respaldo posibilita volver a elaborar la porción dañada de la base de datos valiéndose de la computadora para intercalar las cintas de las transacciones y las copias de los registros actualizados.

ix.v) Seguridad y confidencialidad.

Las medidas de seguridad están orientadas a preservar la información en los datos impidiendo cualquier intromisión que pudiera conducir a la destrucción de archivos y de base de datos. La protección de confidencialidad consiste en las medidas que se toman para prevenir o evitar la difusión o distribución no autorizadas de los datos. La seguridad se refiere al acceso ilegal a los archivos de la computadora, ya sea físicamente o de infiltración en un sistema en línea, con el propósito de destruir, modificar o tener acceso a los datos sin permiso. La confidencialidad o privacidad, por otra parte entraña el derecho de controlar la distribución o divulgación de los datos.

Se incurre en violación de la confidencialidad si datos accedidos para un fin se distribuyen para otras intenciones.

La administración de la base de datos implica la protección de la privacidad y de seguridad. El confinamiento es el acto de alojar los datos en una ubicación física a la que no tengan fácil acceso personas no autorizadas. Mantener la base de datos en un recinto de concreto armado o una bóveda metálica de seguridad es una manera de llevar a cabo el confinamiento.

Las tres faces de la reglamentación son: identificación, autorización y vigilancia. La identificación de las personas que desean tener acceso a los datos se puede efectuar por medio de contraseñas y claves o bien, por credenciales, tarjetas o llaves.

La identificación va unida por lo general a la autorización, que es la aprobación para tener acceso a archivos en particular y hacer determinados usos de los datos. Alguien debe decidir si se concede a un individuo en particular permiso para consultar todo un archivo, ciertos registros del mismo o nada más parte del mismo registro.

Una tercera fase de la reglamentación es la vigilancia. La conservación de registros de todas las personas que utilizan los datos y el examen de los registros de modo periódico es una forma relativamente sencilla de vigilar el acceso y tal vez detectar actividades indebidas. A través de la vigilancia también se pueden detectar otros problemas, como la identificación de personas que realizan intentos para utilizar la base de datos sin mostrar medios aceptables de identificación personal.

Mediante cifrado o encriptación, personas no autorizadas pueden tener dificultades para servirse de los datos que hayan extraído sin permiso o incluso para determinar cuales son los registros a los que tuvieron acceso.

La seguridad de los datos es costosa puesto que requiere equipo y tiempo especiales. Por ello el personal de sistemas y los gerentes o directores tienen que decidir hasta que punto el costo de la seguridad (y de la privacidad) excede el valor de los datos que se protegen.

Seguridad de la base de datos.

El procesamiento de la base de datos puede proporcionar una productividad mucho más alta que otros tipos de procesamiento, pero este procesamiento aumenta la vulnerabilidad de la empresa. En una base de datos, los datos de la empresa quedan centralizados y quedan accesibles. De hecho, los productos DBMS están diseñados para maximizar tal accesibilidad.

Esta situación es excelente para el usuario autorizado. Pero los productos DBMS también resultan sencillos para los usuarios no autorizados, y para el uso de delincuentes.

No todos los usuarios (y programadores) que tengan autorización de acceso a un sistema de base de datos tendrán acceso completo a todos los datos.

- **Sujetos.** - Un sujeto es cualquier usuario identificable o grupo de usuarios identificables que pueden procesar la base de datos. Es difícil determinar si alguien que afirma ser un sujeto es dicho sujeto. Las personas pueden identificarse mediante huellas dactilares, huellas bucales y contraseñas.

En el entorno de la base de datos, el sujeto que se desea tener acceso a un objeto proporciona su identificador (nombre o contraseña). Como verificación, el DBMS revisa nombre del objeto y de otros datos en la tabla, y si lo encuentra, permite el acceso del sujeto a dicho objeto.

- **Objetos.** - Ejemplos de objetos: las bases de datos, las afinidades, las hileras, las columnas, las vistas, el DBMS, los programas, las transacciones y el conocimiento del estado.

El término granularidad, a veces se utiliza para referirse al tamaño de los objetos de seguridad. Un sistema de seguridad que sólo permite (o impide) acceso a la totalidad de la base de datos como una unidad tiene una granularidad grande. En comparación, un sistema de seguridad que permite el acceso o lo impide a un atributo en particular tiene pequeña granularidad.

La selección de la granularidad de un sistema de seguridad involucra un intercambio entre una seguridad muy estrecha y la eficiencia del procesamiento. Conforme aumenta la seguridad la eficiencia se reduce.

-
- Acciones. - Identifica lo que un puede hacer con el objeto. Las posibilidades son: lectura, inserción, eliminación modificación, destrucción y concesión.

La acción de conceder se refiere a otorgar un permiso: como el permiso de modificar datos a otro sujeto.

Limitantes de autorización.

Las limitantes de autorización especifican limitaciones en los permisos en relación con el sujeto, el objeto y la acción.

- a) Seguridad orientada a sujetos.- Con la seguridad orientada a sujetos, los sujetos quedan definidos respecto al DBMS, y cada uno de ellos recibe permisos.
- b) Seguridad orientada a objetos. - Cada objeto tiene una matriz de autorización que muestra las acciones que pueden realizar varios sujetos sobre dicho objeto.

Protección, privacidad y secreto.

La protección de los datos requiere ejercer control sobre la lectura, escritura y empleo de esa información.

El secreto se logra cuando no existe acceso a todos los datos sin autorización. La privacidad adecuada puede lograrse cuando los datos que puedan obtenerse no pueden enlazarse a individuos específicos o no pueden utilizarse para imputar hechos acerca de ellos.

La protección de los datos siempre es limitada. Mientras más protección se implante para reducir las violaciones accidentales y deliberadas de acceso, mayor será el costo del sistema.

El valor de la protección de un elemento dato puede determinarse de acuerdo al beneficio obtenido por un intruso deliberado o de acuerdo a la pérdida sufrida por el propietario en caso de una exposición accidental. En el caso de que se presente una incapacidad para lograr el acceso a los datos, este valor depende de nuevo del beneficio obtenido por quien los conserva, o por la pérdida sufrida por el accesor debido a la omisión deliberada o errónea de los datos

Elementos que deben protegerse

El espacio de datos, al que se dirige un accesor solicitando un tipo específico de acceso contiene los objetos o elementos que se proporcionarán o se restringirán. Es posible distinguir varios tipos de tales objetos:

- ❖ **Datos.** En primer lugar, se encuentran los hechos mismos registrados. Un hecho resulta interesante para el accesor sólo cuando está asignado a un conjunto de atributos que lo identificarán, de manera que este hecho pueda relacionarse con una persona u objeto del mundo real.
- ❖ **Ruta de acceso.** Además de los elementos dato que representan valores dato, existen elementos en la base de datos que indican vínculos y que aparecen como apuntadores en los archivos. Existen otros apuntadores en índices de los datos. Estos apuntadores presentan una clase diferente de objetos que se debe asegurar.
- ❖ **Programas.** Ya que los programas pueden ser elementos activos en un sistema de base de datos, su protección es de igual preocupación. Cuando la protección para los datos esté razonablemente completa, si los programas se colocan bajo las mismas reglas, también estarán protegidos.
- ❖ **Esquemas.** En este caso, hay que preocuparse acerca de la asignación adecuada de objeto en un medio de base de datos de esquema, cuando la protección se proporcione a través del esquema mismo. La asignación de privilegios de acceso al esquema es sumamente importante. Muchos accesores tendrán privilegios de lectura a los descriptores, pero muy pocos tendrán privilegios de modificación o escritura. Un error al modificar un esquema puede hacer que toda la base de datos quede inaccesible.

ix.vi) Control de operaciones concurrentes.

Es significativo particularmente cuando múltiples usuarios solicitan los mismos datos al mismo tiempo. Cuando dos usuarios están efectuando cambios en una misma base de datos, es posible que se pierda una de las actualizaciones.

Si se otorga a un usuario el uso exclusivo de un conjunto de datos en particular mientras otra persona esta esperando el mismo conjunto de datos y otro usuario controla un conjunto de datos que necesita el primer usuario, puede ocurrir un estancamiento. En un estancamiento cada usuario esta esperando que otro libere los datos que esta necesitando aquel y, salvo otra acción ambos usuarios podrían esperar indefinidamente.

Control de concurrencia o simultaneidad

Uno de los avances en el tratamiento de la información viene dado por la idea de la **multiprogramación**, que consiste en la ejecución concurrente de dos o más programas diferentes e independientes en la misma computadora, con el fin de incrementar las prestaciones del procesador y realizar un número mayor de transacciones en un intervalo del tiempo dado.

Es tarea del DBMS controlar la interacción entre las transacciones concurrentes para conservar la coherencia de la base de datos.

Problema de la concurrencia

El problema de la concurrencia es la pérdida de operaciones y por tanto de actualizaciones, con lo cual la base de datos va a perder su coherencia. Surge no solamente cuando se trata de recuperar información, sino cuando dos o más usuarios intentan actualizar un dato al unísono.

Solución al problema de la concurrencia

El DBMS está diseñado para que, una vez iniciada una actualización con un determinado elemento de datos, cualquier intento de iniciar otra actualización sobre el mismo dato no pueda efectuarse por encontrarse el acceso bloqueado o cerrado hasta que la primera haya concluido; es decir, se trata de secuenciar las actualizaciones.

Las transacciones que desean acceder a los mismos datos y se encuentran con el seguro cerrado estarán esperando en una cola para que sus peticiones puedan ser atendidas. Cuando la primera actualización termina, quita su cerradura sobre los datos dejando el acceso libre para la siguiente actualización bloqueada en la cola de espera.

Esta vuelve a bloquear el acceso al registro y comienza la actualización de datos, sobre los datos modificados en la anterior transacción.

El nivel o granularidad de la cerradura puede variar en función de las necesidades.

Por granularidad se entiende al tamaño de las unidades controladas individualmente por un controlador o manejador de la seguridad, que es un módulo del sistema encargado de controlar los accesos concurrentes a los datos.

Bloqueo mutuo

La granularidad ocasiona otros problemas que no existían antes, como el bloqueo mutuo o paro.

El bloqueo mutuo tiene lugar cuando se origina una espera circular entre dos o más transacciones y cada una solicita una actualización sobre el mismo archivo.

Solución al problema del bloqueo mutuo

Para resolver el problema del bloqueo mutuo se pueden seguir dos técnicas:

- Impedir que suceda. Pero impedir el bloqueo no siempre es fácil, que unido a otras funciones de desempeño, no hace aconsejable impedir que se produzca el bloqueo mutuo.
- Detectarlo cuando aparece y tomar las decisiones adecuadas para conseguirlo. El bloqueo puede detectarse por dos caminos.
 1. A través de un reloj que mide el intervalo de tiempo que dura una transacción de tal forma que cuando supera unos límites se considera que está en paro.
 2. Por medio de un procedimiento para determinar si la transacción empieza y termina en la misma consulta

.....

CONCLUSIONES

Los sistemas proporcionan información tanto de problemas como de oportunidades. Más que desarrollar un nuevo producto, es posible simularlo con el consiguiente ahorro de tiempo, dinero y errores.

Con el surgimiento de las bases de datos dentro de las organizaciones, ha sido posible mejorar muchos procesos que anteriormente requerían de mucho tiempo y esfuerzo, y con esto los administradores tienen más herramientas para tomar mejores decisiones.

Las bases de datos tienen un gran potencial para las actividades de apoyo a las operaciones y para las actividades de planeación y control de los administradores, y se tiene la necesidad de estructurar las bases de datos para respaldar las tareas administrativas más importantes.

Los sistemas de información automatizados, a través de su papel central en la economía de la información están llevando a cabo cambios en los siguientes aspectos fundamentales:

- a) Las personas trabajan de manera más inteligente, ya que tienen a la mano más y mejores alternativas en la toma de decisiones.
- b) Existe un cambio global en el concepto de industria.
- c) Tanto las ideas como la información están tomando más importancia que el dinero.
- d) Las personas que trabajan con la información dominan la fuerza de trabajo.

Debido a la gran importancia que tiene la información dentro de las empresas, es importante que sea administrada por alguna persona o departamento.

Al responsable de manejar y optimizar la información, se le designa administrador de base de datos, el cual actúa formalmente como mecanismo de control de un sistema que maneja ciclos de información de los usuarios.

.....

El administrador de base de datos es responsable de resolver las diferencias entre varias funciones de la organización, proporcionando a cada departamento alternativas de solución a diversos problemas que puedan presentarse, tomando en cuenta los distintos puntos de vista de cada área involucrada.

Para poder realizar una óptima administración de la información el administrador de base de datos, debe tener acceso a todo tipo de información, incluyendo la que es confidencial, para lo cual necesita imperiosamente apoyo y autorización de la dirección. Ésta debe indicarle cuáles son las principales metas de la organización, limitaciones de presupuesto, planes a corto y largo plazo, entre otros.

El manejo de información es una gran responsabilidad y el administrador de base de datos debe decidir a quién se tiene que proporcionar y en qué medida. Éste tiene que idear formas para preservar la información en los datos impidiendo cualquier intromisión que pudiera conducir a la destrucción de archivos y de base de datos, previniendo la difusión o distribución no autorizadas de los mismos.

Con el presente trabajo se logró establecer la importancia que tienen los sistemas automatizados dentro de una empresa, para el manejo de las distintas funciones administrativas.

Cabe mencionar que la automatización de procesos debe ser incorporada a todas las empresas, sino se corre el riesgo de quedar rezagado ante la competencia.

BIBLIOGRAFIA

1. - Principios de Sistemas de Información

Scott, George M.

1° Edición

Edit. MacGraw-Hill

México, 1988

2. - Informática básica

Alcalde Eduardo, García Miguel

2° Edición

Edit. MacGraw-Hill

España, 1994

3. - Diccionario de computación

Freedman, Alan

5° Edición

Edit. MacGraw-Hill

España, 1993

4. - Introducción a las computadoras y al procesamiento de información

Long, Larry

4° Edición

Edit. Prentice-Hall Hispanoamericana

México, 1995

5. - The facts on file dictionary of microcomputers

Chandor, Anthony

5° Edición

Edit. Library of Congress Catalogin in Publication Data

U.S.A., 1993

6. - Sistemas de Bases de Datos: Administración y uso

Tsai, Alice Y. H.

1° Edición

Edit. Prentice-Hall

México, 1990

-
7. - Procesamiento de Bases de Datos, fundamentos, diseño e instrumentación
Kroenke, David M.
5° Edición
Edit. Prentice-Hall
México, 1996
8. - Organización de las Bases de Datos
Martin, James
1° Edición
Edit. Prentice-Hall
México, 1997
9. - Análisis y diseño de Sistemas de Información
Senn, James A
2° Edición
Edit. MacGraw-Hill
México, 1992
10. - Bases de Datos: Uso y administración en centros de cómputo
Deakin, Rose
1° Edición
Edit. Trillas
México, 1988
11. - Fundamentos de Bases de Datos
Korth, Henry F.; Silberschatz, Abraham
1° Edición
Edit. MacGraw-Hill
México, 1988
12. - Técnicas de Bases de Datos, estructura en diseño y administración
Atre, Shakuntala
1° Edición
Edit. Trillas
México, 1988

13. - Introducción a la computación

Levine G., Guillermo

3° Edición

Edit. MacGraw-Hil

México, 1991

14. - Informática, Presente y Futuro

Sanders, Donald H.

3° Edición

Edit. MacGraw-Hil

México, 1992

15. - Introducción a los Computadores Electrónicos

Davis, Gordon B.

4° Edición

Edit. Continental

España, 1992

16. - Estructura de la Información

García de Sola, Juan F.

2° Edición

México, 1990