

Unidad 1: Enfoques de Bases de Datos.

1.1 Enfoque tradicional de procesamiento de datos

Las organizaciones al incorporar sistemas de información administrativa, lo hacen con el fin de resolver problemas puntuales que apoyen la toma de decisiones. La planificación de un SIA utiliza dos enfoques tradicionales denominados enfoque por agregación y enfoque de base de datos.

Para iniciar el tema es necesario que demos una mirada introductoria a algunos conceptos elementales de análisis de sistemas tradicionales que son la base para una adecuada comprensión del enfoque por agregación y del enfoque de base de datos.

Las Empresas e Instituciones, organizan su estructura interna en pos de la eficiencia tecnológica, económica y administrativa para alcanzar los objetivos y metas que justifican su existencia como tal. Esto determina la búsqueda de herramientas técnicas y metodológicas que faciliten el proceso de toma de decisiones. Los sistemas de información administrativos, SIA, son las herramientas que apoyan la toma de decisiones.

¿Qué es un Sistema de Información Administrativo?

Se entiende por SIA, las personas, estructura organizacional, máquinas manuales y computarizadas, procedimientos administrativos, recursos logísticos, que en su conjunto tienen como finalidad la recolección, clasificación, preparación, almacenamiento, modificación, actualización, recuperación y transmisión de datos que apoyan la toma de decisiones en la organización.

Recursos logísticos:

Los recursos logísticos son los que permiten cumplir la transformación de los datos. En general estos recursos son de tipo humano, físicos, equipos computacionales, software, datos históricos, algoritmos y procedimientos, que posibilitan guiar la secuencia y la forma de diferentes acciones que determinan la forma en que se transforman los datos

Procesos de transformación:

Se nutren de datos de entrada y recursos logísticos que logren la transformación de los datos.

Datos de entrada:

Los datos de entrada se obtienen de tres vertientes, datos primarios o los provenientes de procesos internos, de datos obtenidos de cómo se tomaron las decisiones pasadas y desde los resultados de las acciones llevadas a buen termino por la organización.

Objetivos y Metas:

Toda empresa debe cumplir sus metas y objetivos, de lo contrario no se justifica, por tanto debe cautelar que los resultados de la gestión de cada uno de los niveles administrativos de la organización sea lo más eficiente y efectiva posible. Las metas y objetivos son el resultado de acciones generadas por

decisiones tomadas por los diferentes niveles de la estructura de toma de decisiones.

Toma de decisiones:

Las decisiones que se adoptan a través del proceso de toma de decisiones son apoyadas y los problemas que surgen son enfrentados con información (relevante y oportuna).

Procedimientos administrativos:

Para llevar a buen fin las actividades administrativas mencionadas la organización implementa un conjunto de procedimientos administrativos.

Toma de decisiones y Sias

La toma de decisiones en un SIA se establece en tres niveles, las estructuradas o programables, las semi-estructuradas, las no estructuradas o no programables.

En al marco de la toma de *decisiones estructuradas* se desarrollan modelos reglados que establecen la forma de tomar las decisiones.

Respecto de la toma de *decisiones no estructuradas*, son aquellas que se toman por expertos y no es posible desarrollar un algoritmo que automatice tal proceso heurístico.

Niveles de decisión:

Los Sias proporcionan información para la toma de decisiones a tres niveles de decisión, nivel de planeamiento estratégico, niveles de control de gestión y operativo.

Nivel de planeamiento estratégico:

Los informes que proporcionan los Sias a este nivel contienen en general: información actualizada de la base de datos, estimaciones a futuro, basada en la información actualizada de la base de datos e información que pone el énfasis en situaciones excepcionales.

Nivel de control de gestión:

Se relaciona con la información necesaria para el uso eficiente y efectivo de los recursos que permitan cumplir los objetivos y metas de la organización. Los informes emanados por los Sias para este nivel son aquellos que contengan información para: analizar y efectuar acciones correctivas sobre la operación y estado de las funciones correspondientes además de información que refleje estados de transacciones pasadas.

Nivel de control operacional:

Esta relacionado con la implementación de las diversas actividades que componen la operación de la organización para lograr los objetivos aplicando los recursos de acuerdo a las políticas establecidas. Los informes que apoyan la toma de decisiones a nivel de control operacional se caracterizan por incluir: transacciones rutinarias, datos utilizados en tareas simples y repetitivas cuyo origen esta establecido claramente.

Tipos de SIAS en el enfoque tradicional de datos

SIA puntual, se caracteriza por apoyar la toma de decisiones en una función específica dentro de la organización. Por ejemplo: Sias para la gestión de existencias.

SIA integral, se caracteriza por cubrir todas las actividades de la organización, pudiendo incluir los denominados SIAS puntuales.

En general y a grosso modo un SIA debe contemplar los siguientes elementos:

Explicitación de metas y objetivos de la organización o función administrativa.

1. Determinación de medidas de eficiencia y efectividad para evaluar el logro de objetivos y metas.
2. Estructuración del proceso de toma de decisiones que será utilizado.
3. Identificación y caracterización de la información relevante provista por el SIA.
4. Determinación de datos de salida, entrada, procesos de transformación, tipos y cantidad de recursos a emplear, talque satisfagan los requerimientos de información relevante.

5. El SIA provee todos los procedimientos administrativos y documentación necesaria que hacen posible operar las diferentes actividades del SIA.

Enfoque por agregación

Cuando la organización implementa los SIAS por primera vez lo hacen para resolver problemas puntuales que apoyan la toma de decisiones y controlar el logro de sus metas y objetivos.

Ahora la planificación para el desarrollo de un SIA que aplica el *enfoque por agregación*, se caracteriza por implementar SIAS puntuales independientes uno de otros con una interacción mínima entre ellos y que apenas comparten recursos. Estos SIAS puntuales se desarrollan uno sobre el otro a medida que se van necesitando, originando problemas como la conocida duplicidad de información.

La expansión de las organizaciones produce naturalmente la evolución progresiva de los SIAS, implicando problemas puntuales de procesamiento de información (cuellos de botella), al desarrollar estas soluciones bajo el enfoque por agregación se han producido los siguientes inconvenientes:

1. Los SIAS se desarrollan en forma independiente entre sí, sin compartir recursos ni interacción.
2. Se produce consecuentemente duplicidad de información, un dato se encuentra en varios archivos.
3. Se produce como corolario de lo anterior problemas con la consistencia de la información ya que los datos duplicados no serán actualizados al mismo tiempo.
4. Además la responsabilidad de la actualización de estos datos recae en muchas personas.

Otras consecuencias relativas al contexto de los datos:

1. Los datos satisfacen SIAS que responden a necesidades específicas del área, departamento o función de la organización.
2. Pueden existir datos con la misma denominación pero con valores distintos por provenir de fuentes distintas, ser interpretados en forma distinta, poseer procesos de actualización que obedezcan a acontecimientos distintos.
3. Los mismos datos pueden derivar en resultados diferentes dependiendo del SIA y sus procesos.

Respecto del diseño de los SIAS aplicando el enfoque por agregación, surgen las siguientes consecuencias:

Sobre el diseño lógico:

1. Al diseñar un SIA bajo este enfoque resulta compleja la delimitación del mismo, dado que las funciones administrativas están interrelacionadas entre sí.
2. Los datos se encuentran distribuidos en diversos SIAS, lo que implica dificultades al momento de establecer las fuentes de información u origen de datos para el sistema.
3. Aumenta la necesidad de relacionar los datos para satisfacer nuevos requerimientos.
4. La identificación y caracterización de datos se vuelve inorgánico.
5. Se complejiza el diseño de procedimientos administrativos.

Respecto del diseño físico:

1. Implica la creación de nuevos archivos con datos ya existentes en otros, así como nuevos datos.
2. El uso de diferentes lenguajes de programación produce incompatibilidad en los formatos de almacenaje.
3. Al modificar programas de aplicación generalmente es necesario modificar también sus archivos de datos influyendo a otros programas que usan los mismos archivos.

Sistema de procesamiento de archivos

En la década del 60 el tratamiento de la información se caracterizó por la aplicación de programas denominados BALANCE LINE, que se caracterizan por operar con dos tipos de archivos clásicos de la época, denominados archivos maestros y de transacciones. La lógica de operación de este tipo de programa conocidos hoy como *pareo de archivos* se basa en la actualización de uno o más archivos maestros a partir de uno o más archivo de transacciones. Es el caso de una cuenta corriente y sus movimientos respectivamente.

Otro programa de esta era de los sistemas de procesamiento de archivos es el conocido *corte de control*, aplicado para producir informes de acuerdo a un criterio de agrupamiento de datos. Es el caso de la cartola mensual de una cuenta corriente, allí las transacciones u operaciones son ordenadas por fecha.

Desventajas del enfoque tradicional de datos

1. Redundancia no controlada
2. Inconsistencia de datos
3. Inflexibilidad
4. Escasa posibilidad de compartir datos
5. Pobre estandarización
6. Baja productividad del programador
7. Excesiva Mantención

1.2 Enfoque de bases de datos

Elementos del enfoque de Banco de Datos:

La administración, control y uso de los datos en la organización basado al enfoque de base de datos se rige de acuerdo a los siguientes consideraciones:

Los datos de la organización son contemplados como un recurso fundamental de esta, del mismo modo que el capital, los recursos humanos y otros. Por lo tanto se le da un manejo, control y uso eficiente y efectivo.

En consecuencia se requiere un nivel de decisiones dentro de la organización cuya responsabilidad sea administrar el recurso información.

Todos los datos de la información se encuentran almacenados en archivos centralizados, que permiten el acceso de las aplicaciones que las necesitan.

Los archivos centralizados son accesibles por las aplicaciones y los usuarios según sus necesidades.

Contempla un sistema de identificación, descripción y definición de los datos de la organización.

Incluye dispositivos de acceso directo y pantallas que facilitan la interrogación por parte del usuario.

Permite establecer distintos tipos de usuarios con distintos tipos de accesos centralizados.

Incluye software que facilita la interrogación de la base de datos para los distintos niveles de usuarios.

Implementa condiciones de seguridad e integridad de los datos y procedimientos de recuperación de datos en caso de error.

Comprende un almacén centralizado que incluye toda la información necesaria de los datos de la base de datos con el fin de evitar problemas en su administración a programadores, analistas de sistemas y otros especialistas.

Implementación del enfoque de Banco de Datos:

Antes de contemplar los elementos del enfoque de base de datos es necesario examinar las funciones que deben incluirse en la implementación de este enfoque. Para la implementación del enfoque de base de datos se debe distinguir las siguientes funciones:

1. Administración de la información:

Encargada de caracterizar, identificar y estandarizar los datos contemplados para la base de datos

2. Almacenamiento de datos:

Centraliza los datos de la base de datos en archivos integrados, que genéricamente se denominan base de datos.

3. Supervisión del almacenamiento y recuperación de los datos:

Proporciona las facilidades necesarias para definir, acceder, manejar, recuperar y controlar los datos que se encuentran en la base de datos. Esta función es apoyada por el denominado SABD sigla de sistema administrador de base de datos, este software interactúa fuertemente con el sistema operativo.

4. Administración de la implementación computacional de la base de datos:

Identifica, caracteriza, estructura y estandariza computacionalmente aquellos datos que nutren la base de datos y que estarán bajo el control del SABD, por lo cual se llama ASABD, es decir administrador el SABD. Esta función además se encarga de administrar el hardware y software asociado que permite operar al SABD, así como aquellos archivos que este origina.

5. Demanda:

Debe agrupar todos los usuarios de la base de datos, que aprovechan las facilidades provistas por el SABD. Se entiende por usuarios a los que toman decisiones, los sistemas de información administrativos, los programadores, analistas de sistemas y otros.

Elementos del enfoque:

A. Administrador de la Información (AI):

El administrador de la información debe identificar, caracterizar y controlar los datos incluidos en la base de datos, tal que los usuarios finales encuentren en ella los datos necesarios para la toma de decisiones y los SIAS encuentren los datos para opera. El AI centraliza los datos evitando la duplicidad descontrolada, ambigüedad e inconsistencias de la información.

Las actividades del AI:

1. Determinar y estructurar los requerimientos de información de los usuarios.
2. Especificar los requerimientos de Información.
3. Diseñar los procedimientos administrativos para que los usuarios puedan utilizar los datos de la base de datos y del diccionario de datos. (que contiene identificación, caracterización y estructura de los datos de la base de datos)

En la determinación de requerimientos de información el AI tiene en cuenta que la creación y Mantenión de la base de datos debe ser segura, confiable y fiable. Entre las actividades del AI para identificar la información a incluir en la base de datos están:

1. Determinación de las necesidades de información de cada usuario.
2. Establecimiento de estándares o medidas de los datos de la base de datos.
3. Determinación, análisis y filtrado de los datos a incluir en la base de datos.
4. Producir un inventario de los datos incluidos en la base de datos.

En la especificación de requerimientos el AI, en conjunto con usuarios y el ASABD, identifica y caracteriza los datos que irán en la BD, documentándolos de manera unívoca mediante el diccionario de datos, el cual se transforma en la fuente de información que tiene la organización, en cuanto a la disponibilidad de datos. La especificación de requerimientos de información se realiza a través del diccionario de datos, cuya Mantenión y uso es responsabilidad del AI.

El diseño de procedimientos administrativos realizado por el AI esta dirigido a:

1. Definición y control de estándares para la identificación y caracterización de los datos a incluir en la base de datos
2. Modificación de la estructura de la base de datos.
3. Procedimientos para el manejo y acceso de los datos.
4. Determinación de responsabilidades sobre ciertos datos, de manera de asegurar la confiabilidad de los valores asignados a cada dato.
5. Determinación de procedimientos que regulen el acceso, lectura, inserción, modificación y eliminación de datos de la BD.
6. Determinación de procedimientos que permitan al AI conocer el uso dado a los datos de la base de datos.
7. Analizar las alternativas costo / beneficio para la organización acerca de tener una base de datos que satisfaga los requerimientos de todos los usuarios.
8. Analizar los errores encontrados por los usuarios, con el fin de colaborar en futuras reestructuraciones de la base de datos.

Elementos de una Base de Datos:

Los elementos de una base de datos son los archivos integrados y el catálogo.

- Se entiende por archivos integrados aquellos archivos que han sido modelados y estructurados de tal forma que se encuentran relacionados entre sí permitiendo su interrogación. Los SABD proporcionan las herramientas necesarias para la producción de este tipo de archivos, denominadas lenguajes de definición de datos (LDD), además de lenguajes de manipulación de datos (LMD) para interrogar la base de datos.
- El catálogo es un archivo creado y mantenido por el SABD, en el que se mantienen las características físicas de los datos de la base de datos. Estas características son usadas por el SABD en la traducción y ejecución de aplicaciones computacionales.

El catálogo es producido por un conjunto de rutinas del SABD mediante las descripciones proporcionadas por el ASABD. Para ser accedido por otro conjunto de rutinas para efectos de su mantención y lectura. En general la descripción de los datos almacenados en el catálogo incluye: nombre del dato, tipo, largo, nivel de seguridad, fecha de origen, archivo de residencia, modo de acceso y almacenamiento.

B. Administrador de SABD

La persona encargada de esta función tiene la responsabilidad de la implementación y operación del SABD. El ASABD administra el producto de software denominado SABD, realiza la creación física y Mantención de la base de datos.

1. Las principales responsabilidades del ASABD son las siguientes:
2. Desarrollo, estructuración y crecimiento de la base de datos de acuerdo a las facilidades del SABD y la situación de la organización.
3. Habilitación de facilidades que originen una óptima implementación del SABD, como interfaz de usuarios, mecanismos de seguridad, integridad, privacidad, validación, verificación entre otros.
4. Supervisión del uso dado por el usuario de las facilidades otorgadas por el SABD.
5. Preparación y difusión de procedimientos para la operación del SABD.
6. Asistencia técnica a los usuarios del SABD
7. Medición periódica del desempeño del SABD.

El ASABD, en conjunto con el AI deben determinar como traducir y satisfacer los requerimientos de información de los usuarios. Para ello, previo a la implementación del SABD, tanto el ASABD como el AI tienen las siguientes responsabilidades:

1. Producir el inventario de datos de la organización.
2. Coordinar el manejo y seguridad de los datos.

3. Crear y mantener un diccionario de datos
4. Coordinar los procesos de codificación y estandarización de datos.
5. Preparar normas y procedimientos para verter archivos tradicionales a la base de datos.
6. Experimentar y difundir en forma piloto las funciones del AI y el ASABD.

El diseño físico de la base de datos es labor del ASABD, realizando las siguientes actividades:

1. Coordinación y apoyo en el desarrollo de SIAS para que estos aprovechen las facilidades del SABD y la BD.
2. Mantener el contacto con los proveedores del SABD y de otros SABD
3. Mantener información para los usuarios respecto de la organización de la BD.
4. Mantener un control total sobre el DDL
5. Definir las características e identificación de los datos
6. Analizar el contenido de la BD.
7. Mantener el software de apoyo al diccionario de datos.
8. Preparación y Mantenimiento del catálogo de la base de datos.
9. Determinar la estructura física de la base de datos
10. Especificación de la estructura de la BD.
11. Controlar el acceso a los datos de la BD.
12. Coordinación entre las el SABD y el sistema operativo empleado
13. Definición de procedimientos de protección contra destrucción o accesos no autorizados.
14. Definición de niveles de seguridad para el acceso a la BD.
15. Establecer procedimientos para la seguridad y protección física de los datos.
16. Participación en las pruebas de programas de aplicación.
17. Establecer procedimientos de respaldo para la BD.
18. Analizar y controlar el seguimiento de trazas y errores.
19. Mantenimiento actualizado de los procedimientos de recuperación de la BD.
20. Determinar procedimientos que permitan detectar violaciones a las reglas de seguridad e integridad, buscando la identificación del causante e informando a los niveles que corresponda.

B. Diccionario de datos (DD)

Este elemento del enfoque de base de datos es el conjunto centralizado de atributos lógicos que especifican la identificación y caracterización de los datos que se manejan en la BD. La BD contiene el valor de los datos, el DD contiene meta datos, es decir los atributos lógicos de dichos datos.

Entre las ventajas del DD se tiene:

1. Es un medio centralizado de tener información sobre los atributos lógicos de los datos de la BD.
2. Es un medio de estandarización en el manejo y uso de los datos
3. Es un medio expedito de almacenamiento y recuperación de proposiciones de atributos lógicos originados por analistas de sistemas en el diseño de un SIA.
4. Representa una ayuda para analistas y programadores en el momento de desarrollo de un SIA.
5. Permite introducir procedimientos estandarizados en el manejo de datos, informes y documentación de procesos y aplicaciones.

Los usuarios del DD son: el AI, el SABD, usuarios finales, Analistas de Sistemas y programadores entre otros.

El diccionario de datos contiene para cada dato los siguientes atributos lógicos o meta datos:

Información respecto de: identificación, control administrativo, seguridad, validación y sobre relaciones lógicas y físicas.

Atributos de identificación comprende el nombre completo del dato, nombre abreviado, sinónimos, identificador o clave, fecha de última actualización.

Atributos de información para control administrativo incluye: unidad de origen del dato, nombre del programa o transformación que lo origina, nombre del documento que lo contiene por primera vez en la organización, las unidades organizacionales y programas de aplicación que lo usan, Cardinalidad del dato.

Atributos de relación física: largo, tipo, nombre para programación, reglas de edición, unidad de medida del dato, precisión.

Atributos de seguridad identificación de las personas autorizadas para cambiar las características del dato, accesarlo o actualizarlo, fecha de última actualización e identificación del usuario que efectuó esta actualización.

Atributos de validación contienen lista o rango de valores permitidos, nombre de los programas validadores que actúan sobre él.

Atributos de relaciones lógicas algoritmos de derivación, identifica la forma de generación del dato, estructuras lógicas, grupos y jerarquías donde el dato es miembro.

Beneficios y riesgos de usar Banco de Datos.

Un banco de datos esta constituido por todos los datos formales, relevantes para la toma de decisiones. Los datos del banco de datos se encuentran dispersos en la organización soportados en diversos medios, como, archivadores, formularios, documentos, dispositivos de almacenamiento digital y otros.

La base de datos se constituye por todos los datos del banco de datos, almacenados en archivos centralizados altamente disciplinados, de tal forma que puedan ser requeridos de diversas maneras lógicas, con el fin de satisfacer las consultas de los distintos usuarios de la base de datos.

Los beneficios del banco de datos son amplios y casi innumerables, el banco de datos como se señalo en l párrafos anteriores representa toda la información relevante y formalizada de la organización, entiéndase por datos de la constitución de la empresa hasta los relativos al pago de patentes pasando por datos de acreedores y deudores.

El riesgo del banco de datos es que el volumen de información va aumentando paulatinamente y se hace inmanejable si no es vertida a un sistema de base de datos.

1.3 Concepto Data – Warehouse

La traducción literal es almacenamiento de datos. Como es sabido existen muchos lugares donde podemos buscar datos, pero ha surgido la idea que exista un mayorista al interior de la empresa que acumule toda la información, bodega de datos inteligente.

Existen muchas definiciones de DW, pero la más completa es la de Bill Inmón, la cual dice: “Data Warehouse es una tecnología orientada a temas específicos, integrada, variante con el tiempo y es una colección no volátil que soporta la administración del proceso de toma de decisiones dentro de las organizaciones

¿Cuándo y porqué nace?

Día a día van surgiendo nuevos problemas en una organización y junto a ello nuevas formas de solucionarlos, la idea de DW data de hace mucho tiempo pero la razón de que hoy día sea un tema de actualidad es que hoy existen tecnologías de HW y SW suficientemente poderosas para depurar esta información.

Él porque se asocia a la necesidad de mejorar la información analítica a través de un medio computacional, la mayoría de la información útil en una empresa

está encerrada en viejas aplicaciones y los usuarios creían que bastaba con crear nuevas formas de acceso pero no es así porque además tienen las siguientes características, complejidad en la estructura de los sistemas, diseño de sistemas orientados al rendimiento óptimo, información dependiente, información a menudo dispersa en múltiples o diversos sistemas, definición inconsistente y la solución fue crear un almacén de datos, en el cual los datos fueran transformados, integrados y cargados a un dispositivo en donde tuvieran sentido para aquellas personas que lo necesiten como soporte a la toma de decisiones.

Su creación se ha estimulado gracias a la necesidad de sistemas de información que apoyen la toma de decisiones de una organización.