**Módulo 2: Revisión evolutiva de las bases de datos**

**Introducción:**

Actualmente las organizaciones concursan por satisfacer las necesidades de los clientes con la mayor agilidad y calidad posible, ofreciendo productos y servicios personalizados y competitivos en calidad y precio. Para esto, resulta fundamental el alineamiento de las arquitecturas de los sistemas de información con las estrategias corporativas y los factores críticos de éxito, por lo que los Sistemas de Información deben constituir un modelo eficaz para tener organizaciones flexibles, contribuyendo al rediseño de los procesos de negocio.

Las bases de datos pueden almacenar información sobre personas, productos, pedidos. Muchas bases de datos comienzan como una lista en una hoja de cálculo o en un programa de procesamiento de texto. A medida que la lista aumenta su tamaño, empiezan a aparecer redundancias e inconsistencias en los datos. Cada vez es más difícil comprender los datos en forma de lista y los métodos de búsqueda o extracción de subconjuntos de datos para revisión son limitados. Una vez que estos problemas comienzan a aparecer, una buena idea es transferir los datos a una base de datos creada con un sistema de administración de bases de datos (DBMS), como Access.

En este módulo estudiaremos los conceptos de un sistema manejador de bases de datos. En estos sistemas de gestión de archivos, la definición de los datos se encuentra codificada dentro de los programas de aplicación en lugar de almacenarse de forma independiente y, además, el control de acceso y la manipulación de los datos viene impuesto por los programas de aplicación.

Esto es un gran inconveniente a la hora de tratar grandes volúmenes de información. Surge así la idea de separar los datos contenidos en los archivos de los programas que los manipulan. Es decir, que se pueda modificar la estructura de los datos de los archivos sin que por ello se tengan que modificar los programas con los que trabajan.

Además, en este módulo veremos cómo resguardar y recuperar datos. Los medios de almacenamiento han evolucionado en forma notable desde las primeras computadoras. En la actualidad existe una gran variedad de tecnologías y dispositivos nuevos, pero el disco rígido sigue siendo el "almacén" principal de la información en la computadora.

**2.1. Bases de datos: definición, características y creación**

#### Introducción:

Los conceptos alusivos a las bases de datos están muy claros y definidos, a diferencia de las bases de conocimiento. Las tecnologías de gestión de bases de datos se encuentran en una etapa muy madura, éstas han evolucionado durante los pasados 30 años, desde sistemas de archivos rudimentarios hasta sistemas gestores de complejas estructuras de datos que ofrecen un gran número de posibilidades.

Lentamente, el centro de atención de la informática, que estaba situado en el proceso, evolucionó hacia la estructuración de los datos, siendo una pieza fundamental en la solución de problemas con los que se enfrenta todo diseñador de un sistema de información. Por lo tanto, se cambia de sistemas orientados hacia el proceso a sistemas orientados hacia los datos, donde estos adquieren el protagonismo, pasando del plano oscuro y difuso en el que estaban situados a ocupar un lugar privilegiado en el interés de todo informático.

El objetivo al que se refieren los modelos de datos ha sido el modelo relacional, éste ha marcado las líneas de investigación en los últimos veinticinco años y es el que se ha asentado en la mayoría de los productos actuales. A pesar de que han surgido otros modelos, como los orientados al objeto, que presentan grandes ventajas para cierto tipo de aplicaciones, cuyos conceptos también se empiezan a incorporar a los sistemas relacionales (extendidos).

El rediseño de los procesos de negocio ayuda a la toma de decisiones, facilitando la delegación de funciones y dando lugar a organizaciones más planas, lo que conlleva a procesos más motivadores y productivos. Las bases de datos, como centro de los sistemas de información, se verán modificadas por estos cambios y deberán ofrecer un soporte adecuado a la nueva organización.

#### Objetivo:

Abordar la definición, las características y el origen de las bases de datos, con el fin de comprender sus características y la relación de la tecnología con la creación de las mismas.

Los temas a revisar para responder al objetivo de esta sesión son:

* Bases de datos: aproximación conceptual y características
* Relación de la tecnología con la creación de bases de datos

#### Integridad de las Bases de Datos

Cuando se dice que una base de datos presenta integridad se refiere a un espejo preciso de su universo de datos, asegurar la integridad es una garantía que deben ofrecer los sistemas de información modernos. El diseño para mantener la integridad es un proceso importante en el desarrollo de bases de datos; si no cambia, por ejemplo, cuando se usa para obtener únicamente información de ella, la integridad deja de ser un problema.

Por ejemplo, en nuestra base de datos académica, la integridad significa asegurarse de que nuestros sistemas de información proporcionen respuestas precisas, como cuántos alumnos están matriculados actualmente en este módulo.

#### Funciones y restricciones

Existe una interrelación entre las ideas de función de actualización y de restricción de actividad. La función de actualización tiene una serie de condiciones asociadas a ella, estas condiciones son restricciones de integridad y así mismo habrá un conjunto de acciones asociadas a la función de actualización. Estas especificarán cuál será el resultado de ser ciertas, por ejemplo:

**Función de actualización:** transferir al alumno X del módulo 1 al módulo 2

**Condiciones:** El alumno X cursa el módulo 1

El alumno X no cursa el módulo 2

Se oferta el módulo 2

**Acciones:**

Negar que el alumno X cursa el módulo 1

Confirmar que el alumno X cursa el módulo 2

#### Funciones de consulta

Otra clase de funciones usadas contra una base de datos son las funciones de consulta, las cuales no se modifican de ninguna forma, sino que principalmente se utilizan para afirmar si un hecho o un grupo de hechos se mantienen en un estado determinado de una estructura de datos, por ejemplo:

Funciones de consulta relevantes para el actual UdD podrían ser:

¿Se está ofertando el curso X?

¿El alumno X está cursando el curso Y?

#### Bases de datos

A mediados de la década de los ochenta, a partir del uso generalizado de las microcomputadoras, la elaboración de bases de datos dejó de ser un trabajo inalcanzable para muchos centros de investigación.

A su vez, en ese tiempo, las bibliotecas comenzaron a visualizar las posibilidades que brinda la computación y generaron una infinidad de bases de datos referenciales.

Los especialistas en computación empezaron a entender que no era suficiente el simple uso de las computadoras y las telecomunicaciones para organizar la información y que eran necesarios programas que permitieran estructurar la información para acceder a la misma en forma fácil, rápida y confiable. En ese momento se crearon las bases de datos y sus manejadores como: Dbase III Plus, Dbase IV y Micro-Isis.

**2.2. Evolución de los sistemas de almacenamiento**

**Introducción:**

La tecnología avanza de forma impresionante y los requerimientos de las grandes organizaciones son cada vez mayores, inclusive la exigencia de los sistemas de información; por ejemplo, los videojuegos cada vez necesitan más almacenamiento, lo que ocasiona que los [procesadores](https://www.monografias.com/trabajos5/sisope/sisope.shtml) tengan que ser más rápidos y las capacidades deben ser proporcionales. Después de las tarjetas perforadas, no pasó mucho tiempo hasta que se descubrieron las nuevas tecnologías de las cintas magnéticas y se comenzó a aplicar en el almacenamiento de datos para computadoras.

Luego de la revolución industrial llega la necesidad de realizar cálculos muy grandes que eran difíciles de realizar por el [hombre](https://www.monografias.com/trabajos15/fundamento-ontologico/fundamento-ontologico.shtml), ya que requería de años para terminar un [cálculo](https://www.monografias.com/trabajos7/caes/caes.shtml); entonces, comienzan a crear las computadoras, con ellas se realizaban las tareas matemáticas de manera más rápida y segura.

Si bien no era suficiente, se buscaba mayor rapidez, ahí es cuando se logra hacer procesadores más rápidos con [circuitos](https://www.monografias.com/trabajos10/infoba/infoba.shtml%23circuito) más pequeños, logrando tener capacidades de almacenamiento abismales en espacios físicos muy pequeños.

Todos los avances, tanto teóricos como prácticos, se respaldan en dos elementos clave de los SGBD, la arquitectura y los modelos de datos. En cuanto a la arquitectura, las propuestas de ANSI e ISO en sus modelos de referencia han predominado positivamente, no solo en las investigaciones teóricas, sino también en las aplicaciones prácticas. En otras palabras, las normas de ISO y de los productos, subyacen dos principios de los modelos de referencia, la arquitectura a tres niveles y la descripción recursiva de los datos.

El sistema Sabre es la segunda base de datos privada más grande del mundo después de la NASA, este sistema ha sido reconocido como el Sistema de Distribución Global número uno en el mundo por los “World Travel Awards”. Sabre Travel Network, una compañía de Sabre Holdings, ofrece acceso al sistema de distribución (GDS) líder en el mundo, permitiéndoles a los agentes de viajes de más de 50.000 agencias locales alrededor del mundo convertirse en expertos de viaje.

Sabre GDS, el primer sistema para conectar a compradores y vendedores de viaje incluye más de 400 aerolíneas, aproximadamente más de 76.000 hoteles, 28 compañías de alquiler de autos, 13 compañías de cruceros, 35 trenes y 220 operadores turísticos.

**Objetivo:**

Analizar la evolución de los sistemas de [almacenamiento](https://www.monografias.com/trabajos12/dispalm/dispalm.shtml) de información, con el fin de comprender la funcionalidad de las primeras unidades de almacenamiento y cómo ha sido su evolución a nivel electrónico al sistema SABRE de IBM.

Los temas a revisar para responder al objetivo de esta sesión son:

* Cintas magnéticas y las primeras unidades de almacenamiento
* Del almacenamiento electrónico al sistema SABRE de IBM

**Migración y archivos de datos**

La migración es el proceso de transferir datos desde un sistema existente a otro u otros. Esta tarea contempla los pasos de limpiar, corregir y mover los datos al nuevo sistema; cuando se cambia de base de datos, se tiene que conseguir que los datos de traslado funcionen en el nuevo entorno.

Se requiere transformar los datos a un formato conveniente para el nuevo entorno y preservar la información del viejo. Este proceso es recomendable hacerlo primero en un entorno de pruebas. Una vez que se asegura el paquete de datos puede realizarse en entorno real.

**Sistemas de almacenamiento NAS**

El sistema NAS, por sus siglas en inglés Network Attached Storage,distribuye archivos a través de la red local (LAN). Estos sistemas están compuestos de un procesador, unidades de almacenamiento y un módulo de conexión de red.

Ya conectado, el sistema NAS a la red TCP/IP detectará y ajustará los parámetros para que sea visible en todas las terminales conectadas, posteriormente estará preparado para recibir peticiones de acceso para compartir archivos.

**Acceso paralelo**

El acceso paralelo está implementado en los discos duros, este conecta todas las unidades para realizar operaciones de entrada y salida. Este conjunto de discos ofrece una tasa de transferencia muy alta, debido a que las operaciones son distribuidas entre todos de manera simultánea.

La tasa de transferencia está muy cercana a la suma de las tasas de transferencia de los discos miembros, mientras que las operaciones de entrada y salida se asemejan mucho a los índices alcanzados por un disco virtual.

**Cintas magnéticas**

Los primeros soportes de cintas magnéticas rígidas fueron de acero y de papel de aluminio. En 1932 fueron sustituidos estos componentes por los acetatos. La razón por la cual se utilizaba acetato celulósico es porque este componente protege los discos y cintas magnéticas de hongos, bacterias, polvo y también de los cambios climáticos bruscos.

En 1944 el acetato fue sustituido por cloruros de vinilo porque el acetato se reblandece a una temperatura de 601 C. El cloruro de vinilo, al igual que el acetato, es una sustancia termoplástica que soslaya a los soportes magnéticos de agentes contaminantes o nocivos. La diferencia radica en que es capaz de combatir mejor los agentes medioambientales traducidos en microorganismos, humedad y calor.

Las cintas magnéticas sólo se encuentran presentes en modelos de ordenadores antiguos porque se inventaron hace 100 años.

**2.3. Modelos y lenguajes relacionales**

**Introducción:**

Una base de datos relacional es un conjunto de tablas, semejante a las de una hoja de cálculo, estas están formadas por filas (registros) y columnas (campos). Los registros muestran cada uno de los objetos descritos en la tabla y los campos los atributos (variables y de cualquier tipo) de los objetos. En este modelo relacional de base de datos, las tablas comparten algún campo entre ellas; estos campos compartidos sirven para establecer relaciones entre las tablas que permitan consultas complejas.

Desde otro punto de vista de los datos, un sistema de información se basa en una serie de capas de registros en formato digital que representan diversas variables, a los que corresponden varias entradas en una base de datos enlazada. Estas capas corresponden en parte al mismo tipo, de manera que pueden analizarse en conjunto. De este modo puede combinarse, en un mismo sistema, información de mismas características y diferentes atributos, con orígenes y formatos muy diversos.

Las Bases de Datos actuales trabajan sobre servidores web. El número de transacciones (consultas o modificaciones) puede variar de modo considerable según su cantidad de usuarios. En aplicaciones sencillas o para sitios con poco movimiento, la actividad puede ser baja, pero en aplicaciones complejas o en sitios con muchas visitas, reducir el número de consultas a la base de datos puede reducir considerablemente el coste en hardware del sitio. Sucede también que, en sitios web con muchas consultas, la importancia de la base de datos es tal que suele situarse en un servidor aparte del servidor de la aplicación.

Si hace diez años las bases de datos ya eran importantes, actualmente casi todas las páginas web tienen detrás un gestor de contenidos que, a su vez, utiliza un Sistema Gestor de Base de Datos para almacenar la información de la web. Sea como sea, para tener un sistema web que sea fácil de actualizar y utilizar, la opción más sencilla es personalizar un gestor de contenidos y la más complicada será crearlo. En ambos casos, es importante tener conocimientos de cómo funcionan las bases de datos y cómo acceder a la información en el lenguaje del SGBD que se utilice.

**Objetivo:**

Comprender las bases de los modelos y lenguajes relacionales, entendiendo así la metodología para el diseño de una base de datos y los elementos de cambios organizacionales como parte de la gestión gerencial.

Los temas a revisar para responder al objetivo de esta sesión son:

* Modelos y lenguajes relacionales
* Elementos de cambios organizacionales

**Manejo de los datos**

Son muchos los lenguajes utilizados por los Sistemas Gestores de Bases de Datos, algunos de ellos son procedurales, quiere decir que el usuario indica exactamente qué datos quiere manipular o consultar; los no procedurales indican qué datos requiere el usuario sin establecer cómo obtenerlos.

Esto indica que el álgebra relacional es un lenguaje procedural de alto nivel, mientras que el cálculo relacional es un lenguaje no procedural.

**Álgebra relacional**

El álgebra relacional es un lenguaje formal para la serie de operadores, estos trabajan sobre una o varias relaciones para obtener otra como resultado sin cambiar las originales. Tanto los operadores como los resultados están relacionados, por tanto, una salida de operación puede ser entrada de otra.

Esto permite anidar instrucciones del álgebra, del mismo modo que se pueden anidar expresiones aritméticas. A este atributo se le conoce como clausura: las relaciones son cerradas bajo el álgebra, del mismo modo que los números son cerrados en operaciones aritméticas.

Las principales operaciones que realiza el álgebra relacional son:

* **Unarios:** operan sobre una sola relación. Abarca la selección y proyección.
* **Binarios:** se aplican a parejas de relaciones. Abarca la unión, intersección, diferencia, producto cartesiano, Join y división.

**Cálculo relacional**

El álgebra relacional y cálculo relacional son formalismos diferentes que expresan distinta representación de datos en el ambiente del modelo relacional. El cálculo relacional obtiene su nombre del cálculo de predicados, ésta es una rama de la lógica.

Existen dos tipos: el orientado a tuplas y el orientado a dominios. Un predicado es una función con argumentos que se puede evaluar a verdadero o falso; cuando los argumentos se sustituyen por valores, la función conlleva a una expresión del nombre preposición.

Por ejemplo: las frases <<Paloma Poy es una vendedora de la empresa>> y <>Paloma Poy es jefa de Natalia Guillén>> son preposiciones, ya que se puede determinar si son verdaderas o falsas. En el primer caso, la función <<es una vendedora de la empresa>> tiene un argumento (Paloma Poly) y en el segundo caso, la función >>es jefa de>> tiene dos argumentos (Paloma Poly y Natalia Guillén).