



Organizaron de Datos

Organización de Datos

Técnicas de representación y almacenamiento de datos en archivos para su recuperación, resguardo y transmisión en formas eficientes y seguras.



Conceptos Fundamentales

Dato:

Representación de una **cosa real o ideal** o de un **evento** ocurrido o programado para que ocurra, en una unidad lógica de manipulación llamada **registro**, en términos de características descriptivas (también llamadas **atributos**) que se representan en unidades llamadas **campos**.



Conceptos Fundamentales

Archivo de Datos

Unidad lógica de almacenamiento permanente de **registros**, administrada por un **Sistema Operativo**. Dentro de un archivo los registros pueden organizarse en otras unidades lógicas llamadas **bloques o páginas**.



Conceptos Fundamentales

Clasificación de Archivos por Función:

Se Pueden clasificar como de:

Datos Maestros

Datos Transaccionales

Reporte

Trabajo

Control de Datos

Intercambio de Datos

Recursos de Programa

Productos de Programas

Empaquetado de Archivos



Conceptos Fundamentales

Archivos de Datos Maestros:

Datos de un sistema de información que representan entidades de existencia real o ideal, por ejemplo productos o servicios, o valores de Referencia para determinar características o atributos de otros datos (dominios de atributos definidos por extensión).

Archivos de Datos Transaccionales:

Registros de hechos o eventos relacionados con datos maestros, por ejemplo de ventas de productos o de prestaciones de servicios.



Conceptos Fundamentales

Archivos de Reporte:

Información editada para su presentación al usuario (en general en formatos pdf, html o de texto).

Archivos de Trabajo:

Resultados parciales o intermedios de procesamiento, o datos de intercambio entre programas.



Conceptos Fundamentales

Archivos de Control de Datos:

Para almacenar metadatos (definiciones de datos), administrar espacios libres, registrar identificadores de registro vacantes o acceder al contenido de otro Archivo (índices y tablas de acceso).



Conceptos Fundamentales

Archivos de Intercambio de Datos:

Representar datos en formatos estándar de manera que puedan ser procesados libremente conociendo el estándar.

Generalmente son archivos de texto, con alguna convención para rotular o delimitar datos, que pueden incluir o no definiciones sobre la estructura de la información contenida.

(un estándar actual es el XML: eXtensible Markup Language).



Conceptos Fundamentales

Archivos de Recursos de Programa:

También conocidos como de Unidades Grandes de Información. (imágenes, audio, vídeo).

Archivos de Productos de Programas:

Son archivos con tipo asociado a un programa o aplicación (.doc, .xls, etc.).

Archivos de Empaquetado de Archivos:

Se utilizan para agrupar, normalmente en forma Comprimida, archivos y directorios, con propósitos de transmisión o respaldo (.zip, .rar, etc.).



Conceptos Fundamentales

Clasificación de Archivos por Tipos de Datos:

Se Pueden clasificar como de:

Texto (ASCII).

Binarios.

Tipificados y Sin Tipo

Clasificación de Archivos por tipo de Registros:

Registro de Longitud Fija.

Registro de Longitud Variable.



Conceptos Fundamentales

Clasificación de Archivos por Tipo de Acceso:

Secuencial

Directo

Indexado

Relativos

Stream (flujo de datos).



Conceptos Fundamentales

Sistema Operativo:

Es el responsable del manejo de recursos de una computadora.

Los recursos pueden ser físicos (el hardware) o lógicos (carpetas y archivos).

Provee una interfaz controlada entre usuarios y aplicaciones, a nivel lógico, y dispositivos de Almacenamiento secundario, a nivel físico.



Conceptos Fundamentales

Sistema Operativo: Sistema de Archivo

Los usuarios y programas tienen una visión lógica de la Información almacenada, y los dispositivos de Almacenamiento tienen una visión física.

El sistema operativo es el mediador.

El conjunto de programas del sistema operativo encargados de proveer la visión lógica de la información almacenada a usuarios y programas conforman:
el Sistema de Archivo.



Conceptos Fundamentales

Funciones de un S.O.

- ▮ **Identificación y Ubicación de archivos:**
(Directorio Estructura Jerarquica o catalogo)
- ▮ **Brindar Seguridad:**
(Usuarios y Grupos; permiso de acceso a los datos)
- ▮ **Asignación de espacio en dispositivos de almacenamiento:**
(Administra lo que ocupa cada archivo, y las Unidades de asignación libre)
- ▮ **Coordinación de transferencia:**
(Atender y resolver las solicitudes simultaneas;
Priorizar lectura/escritura de registros)
- ▮ **Administrar la comunicación e/ la CPU y los dispositivos de almacenamiento:**
(Administrar el acceso concurrente de procesos a un mismo registro físico)



Conceptos Fundamentales

Visión Lógica de Archivos

Atributos

Nombre, Tamaño, Fecha/Hora, Dueño, Grupo, Indicadores, etc.

Operaciones

*Crear, eliminar, abrir, cerrar, leer, escribir,
Escribir al final, posicionarse,
Obtener atributos, definir atributos, renombrar.*



Organización de Archivos

La Organización de Archivos consiste en:

***Abordar técnicas de organización de datos en archivos
con el objetivo de optimizar
la eficiencia de almacenamiento, recuperación y
resguardo de los datos,
comenzando con las bases conceptuales y
procedimentales sobre las que se apoyarán dichas técnicas.***



Organización de Archivos

Diseño de Datos

Se realiza en dos fases o etapas:

Diseño conceptual o de alto nivel.

La definición conceptual de datos implica la definición de atributos, sin especificar tipos o dominios

Diseño lógico (de programación) o de bajo nivel

Definición de bloques o unidades de organización dentro de un archivo.

Ambas son independiente del lenguaje de Programación



Diseño de Datos

Diseño Conceptual de Datos

Atributos:

*Especificación que define una **propiedad de un objeto, elemento**
O archivo.*

También puede referirse a el valor específico para una instancia determinada de los mismos.

Tipo de Atributos:

Simple. una propiedad ejem.(Nombre, Apellido, Peso,Altura)

Compuestos. Una propiedad compuesta ejem (Fecha->dia,mes,año)

Opcional. Pude o no estar al momento de registrarse.

Polivalentes. Depende de una lista de valores del mismo tipo.

Externos. Relacionan un dato con otro u otros del mismo conjunto o de otro conjunto.



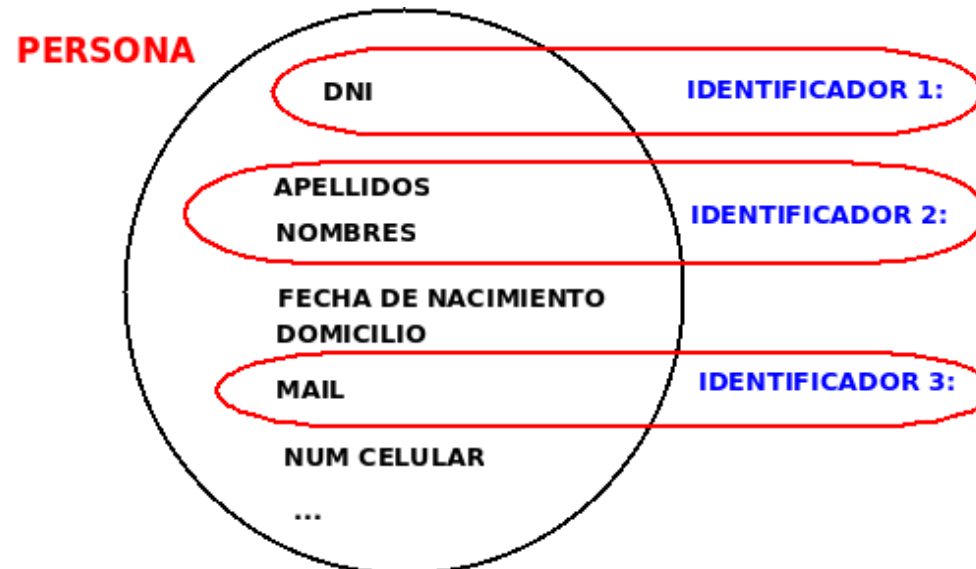
Diseño de Datos

Diseño Conceptual de Datos

Identificador:

Cada cosa o evento debe poder **distinguirse** de los demás. Debe haber uno o más **atributos** que identifiquen unívocamente a los datos.

Un **Conjunto de Datos** puede **tener mas** de **Un Identificador**.





Diseño de Datos

Definición Conceptual de Datos

Para cada atributo se indica la identidad (nombre).
La estructura (en caso de que sean compuestos).
La cardinalidad.
También se puede definir la extensión

LOS CALIFICADORES

DE IDENTIFICACIÓN

- i** Identificador
- ie** Identificador externo.
- d** Definido por extensión en otro archivo

DE CARDINALIDAD

- ?** Opcional
- *** Ninguno o varios (0 a N)
- +** Uno o varios (1 a N)

Persona((dni)i, ((apellido)+, (nombres)+)i, etc..)



Diseño de Datos

Definición Conceptual de Datos

Persona((dni)i,
((apellido)+,
(nombres)+)i,
(fecha de nacimiento(mes,
 dia,
 año))),
(domicilio(tipo('Real' | 'Legal'),
 calle,
 numero,|
 piso,
 localidad,
 (provincia)d,
 (c postal)?),
(mail)*,
(num celular)?))



Diseño de Datos

Definición Lógica de Datos

Para cada atributo (dentro de un archivo):
Implica la definición de bloques o unidades de organización
Especificando tipos o dominios para atributos, y la determina
la convención de almacenamiento y recuperación.

INDEPENDIENTE DEL LENGUAJE DE PROGRAMACIÓN

Tipos de valores convencionales		
En	<i>Enteros</i>	Complemento a dos en n bytes
Fn	<i>Fraccionarios</i>	Punto flotante en n bytes
Cn	<i>Caracteres</i>	Con longitud exacta n
CV	<i>Caracteres Variables</i>	Hasta 255, con prefijo de longitud
T	<i>Texto</i>	Cantidad ilimitada de caracteres, incluyendo caracteres de control como salto de línea, retorno de carro, tabulación, fin de texto.
L	<i>Lógicos</i>	0: Falso o No, 1: Verdadero o Sí
B	<i>Binario</i>	imagen, audio, vídeo, etc.



Diseño de Datos

Definición Lógica de Datos

Ejemplo con Registro de Longitud Fija:

```
Persona( dni:E4,  
         apellido:C64,  
         nombres:C64,  
         fecha de nacimiento(mes:E1,  
                               dia:E1,  
                               año:E2),  
         (domicilio(tipo('Real ' | 'Legal'):(  
             calle:C64,  
             numero:E1,  
             piso:E1,  
             localidad:C64,  
             provincia:C64,  
             c postal:C64 )3,  
         (mail:C32)2,  
         (num celular:C16) )
```

Ejemplo con Registro de Longitud Variable:

```
Persona( cantreg:E1,  
         ( dni:E4,  
         apellido:CV,  
         nombres:CV,  
         fecha de nacimiento(mes:E1,  
                               dia:E1,  
                               año:E2),  
         cantDoms:E1,(domicilio  
                       (tipo('Real ' | 'Legal')):CV,  
                       calle:CV,  
                       numero:CV,  
                       piso:CV,  
                       localidad:CV,  
                       provincia:CV,  
                       c postal:CV),  
         cantMail:E1, (mail:CV)*,  
         cantCel:E1, (num celular:CV)* )  
)
```



Organización de Registros

Longitud Variables Vs Longitud Fija

Los archivos de registros de longitud fija:

Son elementos del mismo tamaño y almacenan la información en los archivos mediante un encabezado.

Se introducen uno a uno los registros ubicados en posiciones Consecutivas.

En el tamaño del campo produce un desperdicio de espacio. Facilita la dirección y extracción de la información del campo



Organización de Registros

Longitud Variables Vs Longitud Fija

Los registros de longitud variable:

Almacenan elementos de varios tipos en un archivo y permite uno o más campos de longitudes variables y los campos pueden ser repetidos.

La longitud de los registros debe estar definida correctamente para poder leer y escribir de forma efectiva.

Aquí hay un aprovechamiento del espacio pero se complica la localización y extracción de la información.



Organización de Registros

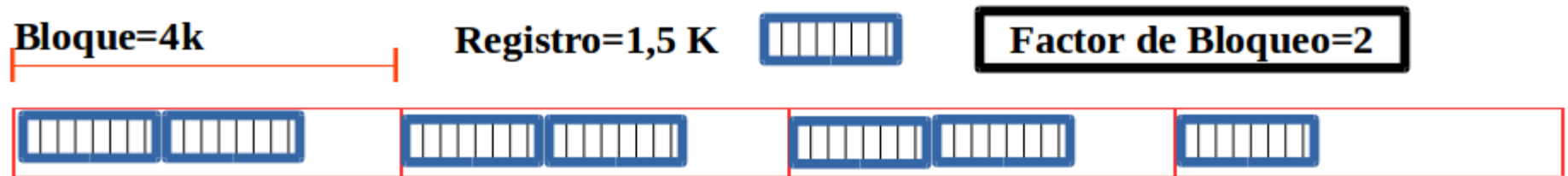
La Problemática...

Los registros de longitud variable:

Representación de valores nulos: mapas de nulidad en cada registro.

Factor de Bloqueo:

*Determinación de tamaños de bloque: $512 * 2^n$ bytes
(coordinación con buffering del disco)*



Administración de espacio libre:

Encadenamiento de registros vs mapas para registros de longitud fija

Desperdicio de espacio libre vs archivo auxiliar para registros de longitud variable



Organización de Registros

Independencia Lógica de Datos

Datos lógicamente Dependientes:

Dependen de una aplicación propietaria (sólo la aplicación conoce la estructura y organización de los registros).

Datos lógicamente independientes

*Hay una definición de los datos (**metadatos**) pública y en una convención estándar (por ejemplo XML).*

Opciones: encabezados en archivos vs diccionario.



Organización de Archivos

Conceptos de Organización de Archivos

Dónde almacenar registros nuevos y cómo encontrar registros dentro del archivo para eliminarlos, modificarlos o recuperarlos para consulta.

“Son los modos de disponer los registros del fichero en el soporte.”

Existen tres modos principales:

Secuencial: *Un registro a continuación de otro.*

Directo: *Los registros binarios no se disponen en el soporte atendiendo a un algoritmo de cálculo. (-seek)*

Indexado: *Los registros generalmente se almacenan secuencialmente y van con un índice.*



Organización de Archivos

Primitivas de Organización de Archivos

De Creación: creación y carga inicial sin validación de unicidad ni búsqueda de espacio libre.

De Actualización de Registros: inserción con validación de unicidad y búsqueda de espacio libre, modificación y supresión.

De Recuperación de Registros: consulta o recuperación unitaria de registros, y reporte o recuperación comprensiva.

De Mantenimiento: reestructuración (reconstrucción), depuración (archivos transaccionales) y respaldo.