**TEMA 3 – ALMACENAMIENTO DE LA INFORMACIÓN**

1. **ANÁLISIS Y DISEÑO DE BASE DE DATOS**

El **Nivel lógico** o **conceptual** describe la estructura completa de la DBA a través de lo que llamamos **Esquema Conceptual**, que se encarga de representar la información de una manera totalmente independiente del SGBS.

2 fases de trabajo al desarrollar una DBA:

* **Fase de Análisis**
  + Análisis de entidades, trata de localizar y definir las entidades y atributos.
  + Análisis de relaciones, se definen las relaciones existentes entre entidades.
  + Obtención del Esquema Conceptual a través del modelo E-R.
  + Fusión de vistas, se reúnen en un único esquema todos los esquemas existentes en función de las diferentes vistas de cada perfil de usuario.
  + Aplicación del enfoque de datos relacional.
* **Fase de Diseño**
  + Diseño de tablas.
  + Normalización.
  + Aplicación de retrodiseño, si fuese necesario.
  + Diseño de transacciones, localización del conjunto de operaciones o transacciones que operará sobre el esquema conceptual.
  + Diseño de sendas de acceso, se formalizan los métodos de acceso dentro de la estructura de datos.

1. **MODELO E/R**

Es una herramienta de referencia para la representación conceptual de problemas del mundo real, su objetivo principal es facilitar el diseño de DBA permitieno la representación de un esquema que representa la estructura lógica completa de una DBA.

Este esquema permite las descripciones textuales de la realidad, que establecen los requerimientos del sistema, buscando ser lo más fiel posible al comportamiento del world real para modelarlo.

1. **ENTIDADES**

Objeto real o abastracto acerca del cual se desea guardar información con características diferenciadoras capaces de hacerse distinguir de otros objetos.

**Un conjunto de entidades** serán un grupo de entidades que poseen las mismas características o propiedades, por lo general se suele emplear el termino entidad para identificar conjuntos de entidades, cada elemento del conjunto de entidades será una ocurrencia de entidad.

Represenación de una entidad en el modelo Entidad/Relación:

Texto

Descripción generada automáticamente con confianza media

* 1. **Tipos: Fuertes y Débiles**
* **Entidades Fuertes o Regulares**:

Aquellas que tienen exsistencia por sí mismas, su existencia no depende de la existencia de otras entidades.

Texto

Descripción generada automáticamente con confianza media

* **Entidades Débiles**:

Tipo de entidad cuyas propiedades o atribuots no la identifican completament, sino que sólo la identifican de forma parcial, esta entidad debe participar el una relación que ayude a identificarla.

Un dibujo de una persona

Descripción generada automáticamente con confianza media

**Tipos de dependencia de las Entidades Débilies:**

* **Dependencia en Existencia**, entre entidades, si desaparece una instancia de entidad fuerte desapareceren las de entidad débilies que dependen de la primera.

Su representación incluirá una E en el interiro de la relación débil.

Diagrama

Descripción generada automáticamente

* **Dependencia en Identificación**, debe darse una dependencia en existencia y además, una ocurrencia de la entidad débil no puede identificarse por sí misma, debiendo hacerse mediante la clave de la entidad fuerte asociada.

Su representación incluirá una ID en el interior de la relación débil.

Una señal de tránsito

Descripción generada automáticamente con confianza baja

1. **ATRIBUTOS**

Un atributo cada una de las propiedades o características que tiene un tipo de entidad o un tipo de relación, toman valores de uno o varios dominos.

Se utilizan para guardar información sobre alguna característica o propiedad de una entidad o relación.

En el modelo E/R son representandos mediante el nombre del atributo rodeado por una elipse que se conecta con la entidad mediante una recta.

Diagrama

Descripción generada automáticamente

**Dominio**, el el conjunto de valores permitidos para un atributo. Los posibles valores que puede tomar el atributo deben estar dentro del dominio.

Es recomendable establecer unos límites adecuados para que el sistema gestor de da Dba lleve a cabo las verificaciones oportunas en los datos que se almacenen, garantizando así la integridad de éstos.

* 1. **Tipos de Atributos**
* **Obligatorios**, ha de estar siempre definido para una entidad o relación.
* **Opcionales**, podra ser definido o no para la entidad; puede o no haber ocurrencias de entidad para las que el atributo no esté definido o no tenga valor.
* **Simple o Atómico**, no puede dividirse en otras partes o atributos, presenta un único elemento, no es posible extraer partes más pequeñas que puedan tener significado.
* **Compuesto**, pueden ser divididos en subpartes, que constituyen otros atributos con significado própio.
* **Monovaluado**, puede tomar un único tipo de valor para cada ocurrencia de entridad.
* **Multivaluado**, puede tomar diferentes valores para cada ocurrencia de entidad.

Conceptos a tener en cuenta en este tipo de atributo:

* + **Cardinalidad de un atributo**, indica el nº mínimo y máximo de valores que puede tomar cada ejemplar de la entidad o relación a la que pertenece.
  + **Cardinalidad mínima**, la cantidad de valores del atributo que debe existir para que la entidad sea válida.
  + **Cardinalidad máxima**, la cantidad máxima de valores del atributo que puede tener la entidad.
* **Derivados, Calculados o Almacenados**, el valor de estos puede ser obtenido del valor/es de otros atrubutos relacionados.
  1. **Claves**

Atributo especial obligatorio, cuyos valores deben ser tales que permitan identificar unívocamente a la entidad.

Conceptos:

* **Superclave**, cualquier conjunto de atributos que permite identificar de forma única a una ocurrencia de una entidad.

Puede tener atributos no obligatorios, que no identifican por si solos una ocurrencia de entidad.

* **Clave Candidata**, una superclave de la que no es posible obtener ningún subconjunto que sea a su vez una superclave.

Para la eleccion de claves candidats se tendrá en cuenta:

* + Sus valores deben ser conocidos -> distintos de nulos.
  + La memoria que ocupen debe ser la menor posible.
  + Tienen codificación sencilla.
  + El contenido de sus valores no debe variar.
* **Primary key**, de todas las keys candidatas, es un atributo o conjunto de ellos, que toman valores únicos y distintos para cada ocurrencia de entidad, identificandola unívocamente.

No puede toma valores Null.

Entre varias candidatas de las mismas cualidades se deben tener en cuanta los siguientes criterios para elegir la clave primaria:

* + Elegir la de menor longitud.
  + Elegir las simples sobres las compuestas.
  + Numéricas sobre no númericas.
  + Codificadas sobre no codificadas.
  + Las de ámbito local sobre las de ámbito más genera.

Una ves elegida Primary Key, las restantes candidatas son denominadas **Alternativas** o **Secundarias.**

Se puede representar de 2 formas:

* Se usan elipses para representar atributos, se subrayan aquellos que formen la clave primaria.
* Se usan circulos para representar atributos, se usa uno negro en los que formen la clave primaria

Diagrama

Descripción generada automáticamenteDiagrama

Descripción generada automáticamente

* 1. **Atributos de una Relación**,

Una relación puede tener atributos que la describan:

Diagrama

Descripción generada automáticamente

1. **RELACIONES**

Una asociación entre diferentes entidades, no pueden aparecer dos veces relacionadas las mismas ocurrencias de entidad.

La representación grafica es un rombo en cuyo interior se encuentra inscrito el nombre de la relación expresado con un verbo, conectado con las entidades a las que relaciona mediante líneas rectas que podrán o no acabar en punta de flecha según el tipo de relación.

Se interpreta de izquierda a derecha o de arriba abajo.

Un dibujo de una persona

Descripción generada automáticamente con confianza baja

Para describir y definir adecuadamente las relaciones existentes entre entidades es imprescindible conocer los siguientes conceptos:

* 1. **Grado de una Relación**

El número de entidades que participan en una relación

Tipos de relaciones en función del grado:

* **Relación Unaria o de Grado 1.**
* **Relación Binaria o de Grado 2.**
* **Relación Ternaria o de Grado 3.**
* **Relación N-aria o de Grado N.**
* **Relación Doble**, 2 entidades se relacionan a traves de 2 relaciones.
  1. **Cardinalidad de Relaciones**

El número máximo de ocurrencias de cada entidad que pueden intervenir en una ocurrencia de relación.

Viene expresada siempre para relaciones entre dos entidades.

Dependiendo del número de ocurrencias de cada una de las entidades pueden existir relaciones:

**Uno a uno** (**1:1**), sean entidades A y B, una instancia de A se relaciona únicamente con otra de B y viceversa.

**Uno a muchos** (**1:N**), Sean entidades A y B, una ocurrencia de A se relaciona con muchas de B y una ocurrencia de B sólo se relaciona con una de A.

**Muchos a uno** (**N:1**), Sean entidades A y B, una ocurrencia de A está asociada a una única de B y una de B está relacionada con muchas de A.

**Muchos a muchos** (**N:N**), Sean entidades A y B, una ocurrencia de A está relacionada con muchas de B y viceversa.

* 1. **Cardinalidad de Entidades**

Especifica el número mínimo y máximo de correspondencias en las que puede tomar parte cada ejemplar de dicha entidad. Indica el número de relaciones en las que una entidad puede aparecer.

Sean entidades A y B, la participacion de A en una relacion es **obligatoria (total)** si la existencia de cada una de sus ocurrencias necesita mínimo una ocurrencia de B, en el caso contrario la participacion es **opcional (parcial)**.

* **Cardinalidad Mínima**, el número mínimo de asociaciones en las que aparecerá cada ocurrencia de la entidad, el 0 solo se pone cuando la participación es opcional.
* **Cardinalidad Máxima**, el número máximo de relaciones en las que puede aparecer cada ocurrencia de la entidad.

Imagen de la pantalla de un celular de un mensaje en letras negras

Descripción generada automáticamente con confianza baja

1. **RESUMEN SIMBOLOGÍA DEL MODELO E/R**

Diagrama

Descripción generada automáticamente

Diagrama

Descripción generada automáticamente

Interfaz de usuario gráfica

Descripción generada automáticamente

1. **EL MODELO E/R EXTENDIDO**
   1. **Restricciones en las Relaciones**

* **Restricción de Exclusividad**

Cuando existe una entidad que participa en dos o más relaciones y cada ocurrencia sólo puede pertenercer a una de las relaciones únicamente.

Si la ocurrencia de entidad pertenece a una de las relaciones, no podrá formar parte de la otra.

**Se produce una relacion o se produce otra pero nunca ambas a la vez.**

* **Restricción de Exclusión**

Se produce cuando las ocurrencias de las entidades sólo pueden asociarse utilizando una única relación.

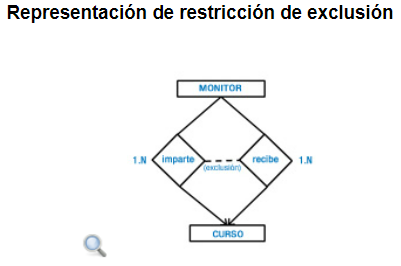
* **Restricción de Inclusividad**

Se aplican cuando es necesario modelar situaciones en las que para que dos ocurrencias entidad se asocien a través de una relación, tengan que haberlo estado antes a través de otra relación.

* **Restricción de Inclusión**

Restricción mas fuerte que la de inclusividad

Diagrama

Descripción generada automáticamente

Diagrama

Descripción generada automáticamente

* 1. **Generalización y Especialización**

**Superclase** o **Supertipo**, conjunto de entidades de nivel superior que engloban a conjuntos de entidades de nivel inferior con caracteristicas comunes.

**Subclase** o **Subtipo**, conjuntos de entidades de nivel inferior.

Existe la posibilidad de realizar una especialización de una superclase en subclases y análogamente, establecer una generación de las subclases en superclases.

La **Generalización** es la reunión en una superclase o supertivo de entidad de una serie de subclases o subtipos de entidades, que poseen características comunes, las subclases tendrán otras claracteristicas que las diferenciarán entre ellas.

**Herencia** relación a traves de la cual los atributos de una superclase de entidad son heredados por las subclases, si una superclase interviene en una relación, las subclases también lo harán.

Una generalización/especialización podrá tener las siguientes **restricciones semánticas:**

* **Totalidad**, si todo ejemplar de la superclase pertenece a alguna de las sublases.
* **Parcialidad**, si no todos los ejemplares de la superclase pertenencen a alguna de las subclases.
* **Solapamiento**, si un mismo ejemplar de la superclase puede pertenecer a más de una subclase.
* **Exclusividad**, si un mismo ejemplar de la subclase pertenece sólo a una subclase.
  1. **Agregación**

Es una abstracción a través de la cual las relaciones se tratan como entidades de nivel más alto, siendo utilizada para expresar relaciones entre relaciones o entidades y relaciones.

2 clases de agregaciones:

* **Compuesto/Componente**, unión de diversas partes, pueden ser objetos distintos y que desempeñan papeles distintos en la agregación.
* **Miembro/Colección**, unión de diversas partes del mismo tipo y que desempeñan el mismo papel a la agregación.

1. **ELABORACIÓN DE DIAGRAMAS E/R**
   1. **Identificación de Entidades y Relaciones**

**Entidades**

Analizar la especificacion de requerimientos en busca de nombres y sustantivos, si estos se refieren a objetos importantes serán entidades. Tambien podremos identificar entidades localizando objetos o elementos que existen por sí mismos, acerca de los cuales interesa guardar info.

Se tendran en cuenta además que los nombres referidos a características, cualidades o propiedades no se convertiran en entidades.

3 condiciones para considerar los sustantivos como entidades:

* Existencia própia
* Cada ejemplar debe poder ser diferenciado del resto.
* Todos los ejemplares de un tipo de entidad deben tener las mismas propiedades.

**Relaciones**

Establecer las relaciones entre las entidades, buscando **verbos** o **expresiones verbales que conecten unas con otras.**

Cada una de las relaciones establecidas deberá tener asignado un **nombre**, preferiblemente un **verbo en minúsculas**, representativo del significado o acción de la relación.

* 1. **Identificación de Atributos, Claves y Jerarquías**

**Atributos**

Buscar nombres relativos a características, propiedades, identificacores o cualidades de entidades y relaciones.

Un atributo deberá tener asignado un nombre en minúsculas, representativo de su contenido o función, siempre es recomendable recopilar la siguiente información de cada atributo:

* Nombre y descripción.
* Atributos simples que lo componen, si es compuesto.
* Método de cálculo, si es derivado o calculado.

**Claves**

Establecer una o varias claves candidatas, escogiéndo una como clave primaria de la entidad.

La clave primaria estará formada por uno o varios atributos que identificarán de manera unívoca cada ocurrencia de entidad.

**Jerarquías**

Expresar en el esquema las particularidades de diferentes ejemplares de un tipo de entidad creando subclases o subtipos de una superclase o supertipo.

* 1. **Metodologías**

Las metodologías o estrategias parara la elaboracion del esquema conceptual:

* **Top-Down**, partir de un esquema general e ir descomponiendo éste en niveles, cada uno de ellos con mayor número de detalles.

Se parte de objetos abstractos que se refinan hasta llegar al esquema final.

* **Bottom-Up**, se parte del nivel más bajo, los atributos, se van agrupando en entidades, despues crear relaciones entre éstas y las posibles jerarquías hasta obtener un diagrama completo.

Se parte de objetos atómicos y se obtienen abstracciones u objetos de mayor nivel de abstracción que forman el esquema.

* **Insede-Out**, se comienza a desarrollar el esquema en una parte de papel y a medida que se analiza la especificación de requerimientos, se va completando con entidades y relaciones hasta ocupart dodo el documento
* **Mixta**, empleada en problemas complejos, se dividen los requerimientos en subconjuntos para analizarlos independientemente.

Se usan técnica descendente para dividir los requerimientos y en cada subconjunto de ellos, se aplicará la técnica ascendente.

* 1. **Redundancia de Diagramas E/R**

**Reduncancia**, reproducción, repetición, reiteración, insistencia, reincidencia, reanudación: en base de datos hace referencia al almacenamiento de los mismos datos varias veces en diferentes lugares.

Problemas que puede provocar:

* **Aumento de la carga de trabajo.**
* **Gasto extra de espacio de almacenamiento.**
* **Inconsistencia.**

Elementos que pueden presentar redundancia:

* **Atributos redundantes**, cuyo contenido se calcula en función de otros, un atributo derivado puede ser origen de redundancia.
* **Un** ciclo, varias entidades unidas circularmentes o cíclica a traves de varias relaciones. No siempre que exista uno estaremos ante una redundancia, es necesario analizarlo para determinarlo:
  + El significado de las relaciones no sea el mismo
  + Si eliminamos la relacion, el significado restante sigue siendo el mismo.
  + Si la relación eliminada tenia atribuos, sean asignados a alguna entidad participante sin que pierda significado.

No todas las redundancias son perjudiciales -> **Redundancias controladas.**

* 1. **Propiedades Deseables de un Diagrama E/R**

Desglose de las Características o Propiedades deseables:

* **Completitud**, si es posible verificar que cada uno de los requerimientos está represntado en el diagrama y viceversa.
* **Correccción**, si emplea adecuadamente todos lo elementos del modelo E/R.
  + Corrección Sintáctica, cuando no se produzcan representaciones erróneas en el diagrama.
  + Corrección Semántica, cuando las representaciones signifiquen exactamente lo que está estipuado en los requerimientos.
* **Minimalidad**, si se puede verificar que al eliminar algún concepto presente en el diagrama, se pierde información.
* **Sencillez**, si representa los requerimientos de manera fácil de comprender, sin artificios complejos.
* **Legibilidad**, si puede interpretarse fácilmente.
* **Escalabilidad**, si es capaz de incorporar posibles cambios derivados de nuevos requerimientos.

1. **PRIMEROS PASOS DEL DIAGRAMA E/R AL MODELO RELACIONAL**
   1. **Simplificación Previa de Diagramas**

Conjunto de procedimientos y normas necesarios de aplicar a diagramas E/R para que su transformación al model lógico basado en el modelo relacional, sea correcta y casi automática.

* **Transformación de atributos compuestos**, descomposicion de los atributos compuestos en los simples por los que estan formados.
* **Transformación de atributos multivaluados**, se ha de convertir los atributos multivaluados en entidades relacionadas con la entidad de la que procede.
* **Transformación da relaciones jerárquicas**, transformar las relaciones con cardinalidad **N-M** en relaciones de cardinalidad **1-N**.
* **Transformación de relaciones cíclicas**, se crea una nueva entidad cuya clave estara formada por dos atributos que contendrán las claves de las ocurrencias relaciondas para eliminar la relación cíclica.
* **Transformación de relaciones ternarias**, pueden considearase como una relación binaria a la que se le asocia una entidad.
* **Transformación de entidades débiles en fuertes**, añadir a la entidad débil los atribuots clave de la entidad que hace posible la identificación de las ocurrencias.

1. **PASO DEL DIAGRAMA E/R AL MODELO RELACIONAL**

Hay que tener en cuenta las siguientes cuestiones:

* Toda entidad se transforma en una tabla
* Todo atributo se transforma en clumna.
* El atributo clave se convierte en clave privada y se representará subrayado.
* Cada entidad débil generá una tabla que incluirá todos sus atributos, añadiendo a ésta los atribuots que son clave primaria de la entidad fuerte con la que esté relacionada.

Estos atribuos se constituyen como clave foránes que referencia a la entidad fuerte, seguidamente se escogerá una clave primaria para la tabla creada.

* Las relaciones 1-1 podrán generar una nueva tabla o prolongar la clave en función de la cardinalidad de las entidades.
* Relaciones 1-M podrán generar una new tabla o prolongar la clave.
* Las relaciones reflexivas o cíclicas podrán generar una o varias tablas en función de la cardinalidad de la relación.
* Las jerarquías generarán la reunión, eliminación o creacion de relaciones 1-1.
* Las relaciones M-N se transforman en una tabla que tendrá como clave primaria las claves primarias de las entidades asociadas.
* Las relaciones N-arias que agrupan 3 o más entidades, cada entidad se convierte en una tabla y tambien la relación que contendrá sus atributos propios más las claves de todas las entidades.

La clave principal será la concatenación de la claves de las entidades, pueden darse 2 casos dependiendo de las cardinalidades:

* + Si la relación es N:M:N, si todas las entidades participan con cardinalidad máx M, la clave de la tabla es la unión de las claves de las entidades relacionadas.
  + Si la relacion es 1:M:N, si una de las entidades participa con cardinalidad máxima 1, la clave de esta no pasa a formar parte de la clave de la tabla resultante, pero forma parte de la relación como un atributo más.

1. **NORMALIZACIÓN DE MODELOS RELACIONESLES**

La **Normalizacion** es el proceso consistente en imponer a las tablas del modelo Relacional una serie de restricciones a través de un conjunto de transformaciones consecutivas.

Garantiza que las tablas contienen los atributos necesarios y suficientes para describir la realidad de la entidad que representan, permitiendo separar aquellos atributos que por su contenido podrían genrar la creación de otra tabla.

El proceso de normalización se basa en el análisis de de las dependencias entre atributos.

* 1. **Tipos de Dependencias**
* **Funcional**, dados atributos A y B, B depende funcionalmente de A, si, y solo si, para cada valor de A sólo puede existir un valor de B.

Siempre se establece entre atributos de la misma tabla.

El atributo A se denomina determinante.

* **Funcional Completa**, dados los atributos Ak y B, se dice que B depende funcionalmente de forma completa de Ak, si y solo si B depende funcionalmente del conjunto de atributos Ak, pero no de ninguno de sus posibles subconjuntos.
* **Transitiva**, dados tres atributos A, B y C, se dice que existe una dependencia transitiva entre A y B, si B depende funcionalmente de A y C depende funcionalmente de B.

A,B y C podrían ser un solo atributo o un conjunto de ellos.

* 1. **Formas Normales**
* **1º Forma Normal**
* **2º Forma Normal**
* **3º Forma Normal**
* **Forma Normal de Boyce Codd**
* **Otras Formas Normales**