

# **GESTION DE DATOS**

## **MODELOS DE DATOS**

### **AUTORES**

Ing. Vilma Martín

Esp. Lic. Fabiana María Riva

## MODELOS DE DATOS

### EJE CONCEPTUAL

#### Modelos conceptuales y de representación de datos

### OBJETIVOS

Que el alumno logre:

- Describir metodologías para el modelado de datos
- Conocer los símbolos y el vocabulario de la notación de Chen para el modelo Entidad-Relación
- Obtener habilidad para traducir el modelo de entidad-relación en el esquema de tablas relacionales
- Aplicar técnicas de normalización y obtener el modelo relacional

### INDICE

<b>Modelos de Datos .....</b>	<b>1</b>
Definición .....	1
Categorías .....	1
<b>Modelo Entidad-Relación .....</b>	<b>2</b>
Componentes y Representación.....	2
Tipos de Atributos. Valores nulos.....	3
Restricciones Estructurales .....	4
Atributos de las Relaciones o Vínculos .....	7
Entidades Fuertes y Entidades Débiles.....	7
Otros Tipos de Vínculo .....	8
Resumen de la notación básica de un modelo e-r .....	11
<b>Modelo E-R Extendido.....</b>	<b>12</b>
Superclases y Subclases.....	12
Generalización y Especialización .....	12
Jerarquías y Retículas .....	16
Categorías (unión).....	17
Agregación .....	18
<b>Modelo Relacional .....</b>	<b>20</b>
Dependencias Funcionales .....	21
Normalización .....	24
<b>Transformación de un Modelo E-R en Relacional .....</b>	<b>29</b>
Reglas de transformación .....	29
Análisis comparativo de los enfoques de conversión de jerarquías de generalización ...	34

## MODELOS DE DATOS

### DEFINICIÓN

Un modelo de datos es un sistema formal y abstracto que permite describir los datos de acuerdo con reglas y convenios predefinidos.

Es formal pues los objetos del sistema se manipulan siguiendo reglas perfectamente definidas y utilizando exclusivamente los operadores definidos en el sistema, independientemente de lo que estos objetos y operadores puedan significar.

Un modelo se obtiene a través de un proceso de abstracción, mediante el cual en el modelo sólo se representan aquellos aspectos del mundo real que son relevantes para el objetivo del sistema o dominio en estudio.

Un modelo de datos es una combinación de tres componentes:

1. una colección de estructuras de datos;
2. una colección de operadores o reglas de inferencia, los cuales pueden ser aplicados para consultar o derivar datos de cualquier parte de estas estructuras en cualquier combinación deseada;
3. una colección de reglas generales de integridad, las cuales explícita o implícitamente definen un conjunto de estados consistentes. Estas reglas algunas veces son expresadas como reglas de insertar-actualizar-borrar.

### CATEGORÍAS

- **Modelos de datos de alto nivel o conceptuales**, muy cercanos al modo como la mayoría de los usuarios percibe los datos. Entre estos existen:
  - a. **Basados en objetos:**
    - i. **Modelo de Dominio**
    - ii. **Modelo Entidad-Relación**
  - b. **Basados en registros:**
    - i. **Jerárquico:** datos en registros, relacionados con punteros y organizados como colecciones de árboles
    - ii. **Redes:** datos en registros relacionados por punteros y organizados como grafos
    - iii. **Relacional:** datos en tablas relacionados por el contenido de ciertas columnas
- **Modelos de datos de bajo nivel o físicos:** proporcionan conceptos que describen los detalles sobre cómo se almacenan los datos en la computadora

## MODELO ENTIDAD-RELACIÓN<sup>1</sup>

En 1976 Peter Chen publicó por primera vez el Modelo Entidad Relación (E-R). El modelo es comprehensivo, elude las complicaciones de almacenamiento y consideraciones de eficiencia, las cuales son reservadas para el diseño físico de la base de datos. Posteriormente algunos autores extendieron el modelo incrementando sus capacidades.

El enfoque del Modelo E-R hace hincapié simplifica el proceso de modelado desarrollando un diagrama entidad-relación (DER) que una representación abstracta del dominio analizado y además es independiente del almacenamiento y las consideraciones de eficiencia.

El DER puede más tarde ser trasladado hacia modelos más apropiados para su implementación en una Sistema de Gestión de Base de Datos (SGBD o DBMS por sus siglas en inglés).

### COMPONENTES Y REPRESENTACIÓN

Presentaremos las características del modelo aumentando la complejidad a través de ejemplos.

**ENTIDAD:** se define como un tipo de información que tiene interés para la empresa. "Algo" del mundo real con existencia independiente, tener existencia física (ej.: cliente, alumno, coche, empleado) o existencia conceptual (un puesto de trabajo, una carrera universitaria).

Puede ser denominada, según los autores, como Tipo de Entidad, Conjunto de entidades (aludiendo a que cada instancia que conforma el objeto es una entidad) o simplemente Entidad (entendiendo que se conforma con todas las instancias del objeto).

Una de sus representaciones es la de un rectángulo con el nombre de la Entidad (más adelante veremos otro tipo de representación según el tipo de entidad).



**RELACIÓN:** Se define como una correspondencia, **conexión o vínculo** entre dos o más entidades. Se las representa con un rombo con el nombre de la relación y se las relaciona con las entidades que conecta por medio de líneas. Por ejemplo, al modelar la regla

✓ *Los clientes realizan solicitudes de artículos*  
obtendríamos:

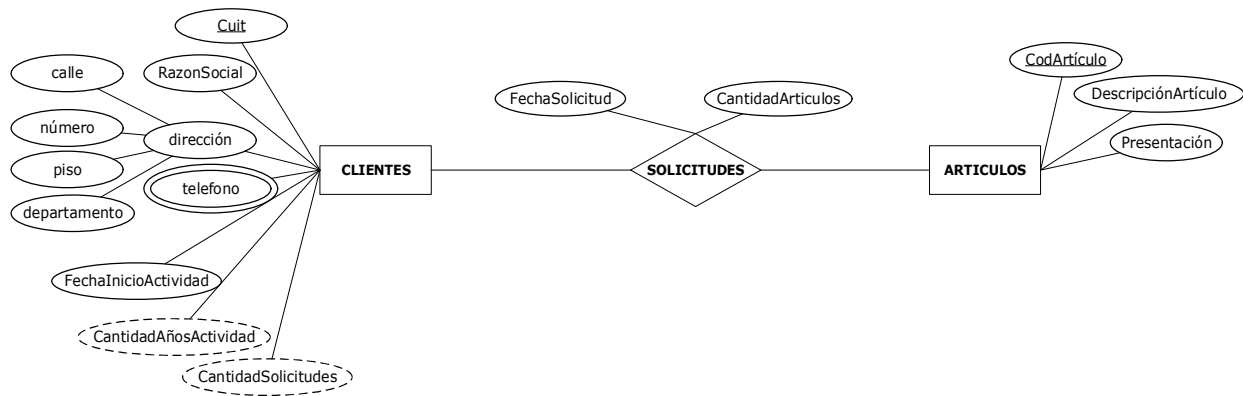


**Figura 1**

Optaremos por poner un nombre representativo al vínculo, en este caso se eligió SOLICITUDES, pero podría haber sido CLIENTES\_ARTICULOS por las entidades que vincula.

**ATRIBUTOS:** Caracterizan a la entidad o relación.  
Se los representan con elipses unidas a cada entidad o relación mediante líneas.

<sup>1</sup> La lectura de este apunte debe complementarse con los capítulos 3 y 4 de "Fundamentos de Sistemas de Bases de Datos. 3ra. Edición. Ramez Elmasri-Shamkant Navathe. Ed. Addison Wesley (2002) ISBN: 84-7829-051-6 Disponible en la Biblioteca bajo el Nro. 23619"



**Figura 2**

Las reglas asociadas al ejemplo serían:

- ✓ Los **clientes** se identifican por su número de Cuit y se conoce de los mismos su razón social, dirección, teléfono, fecha de inicio de actividad y cantidad de años de actividad.
- ✓ Los **clientes** realizan **solicitudes** de **artículos** en una fecha y por una cantidad.
- ✓ Los **artículos** están identificados por un código y además poseen una descripción y presentación.
- ✓ Se requiere conocer, además, la cantidad de solicitudes de artículos realizadas por el cliente.

## TIPOS DE ATRIBUTOS. VALORES NULOS.

### Compuestos vs. Simples:

Un atributo puede ser **simple o atómico** cuando admite valores que no pueden subdividirse, por ejemplo: *CantidadArticulos*.

Un atributo es **compuesto** cuando puede dividirse en atributos más pequeños. Por ejemplo *Dirección* que puede subdividirse en *calle*, *número*, *piso* y *departamento*.

### Monovaluados vs. Multivaluados:

Un atributo es **monovaluado** si sólo admite un valor para una instancia de una entidad en particular. Ej.: *CantidadArticulos* es simple y monovaluado. *Dirección* es compuesto y monovaluado.

**Atributo multivaluado:** pueden existir atributos que asuman varios valores para una misma entidad. Por ejemplo: *Teléfono* puede asumir un único valor para una instancia de una entidad o varios si por ejemplo el cliente tiene varias líneas telefónicas. *Teléfono* es entonces un atributo simple y multivaluado.

Un atributo multivaluado puede tener un límite inferior y superior en el número de valores para una entidad individual. Podríamos suponer para el ejemplo que un cliente puede tener como mínimo un *Teléfono* y como máximo cuatro.

### Almacenados vs. Derivados.

Los **atributos almacenados** son aquellos que deben guardarse con la entidad pues no pueden calcularse ni de otros atributos de la misma entidad u otras, ni de su relación con otras entidades. Por ejemplo: *FechaInicioActividad* es un atributo simple, monovaluado y almacenado.

Los **atributos derivados** se calculan en función de otro atributo como por ejemplo: *CantidadAñosActividad* es un atributo derivado de *FechaInicioActividad*. *CantidadSolicitudes* puede calcularse de la cantidad de instancias en la relación "*Clientes realizan solicitudes de Artículos (SOLICITUDES)*".

**Valores nulos:** En algunos casos un atributo podría no tener valor para una instancia de una entidad. Por ejemplo en el atributo *Domicilio*, *piso* podría admitir valores nulos (NULL) para alguna instancia de la entidad. Por otra parte existen atributos que no deberían admitir valores nulos (NOT NULL o NN), por ejemplo *RazonSocial*.

**Identificadores o claves:** Cada entidad **debe** tener un único identificador que permita individualizar cada una de sus instancias. En el ejemplo son *Cuit* y *CodArtículo*. No necesariamente deben ser un único atributo, pueden ser un conjunto de atributos a lo que se denomina clave compuesta. Todos los atributos que conforman una clave deben ser monovaluados y siempre deben tener valor (son NOT NULL). Completaremos el concepto de identificadores o claves al referirnos al Modelo Relacional.

## RESTRICCIONES ESTRUCTURALES

- **Razón de Cardinalidad**
- **Restricción de Participación**

Estas restricciones se explicitan para las relaciones que vinculan las entidades.

Analizando el ejemplo de la Figura 2:

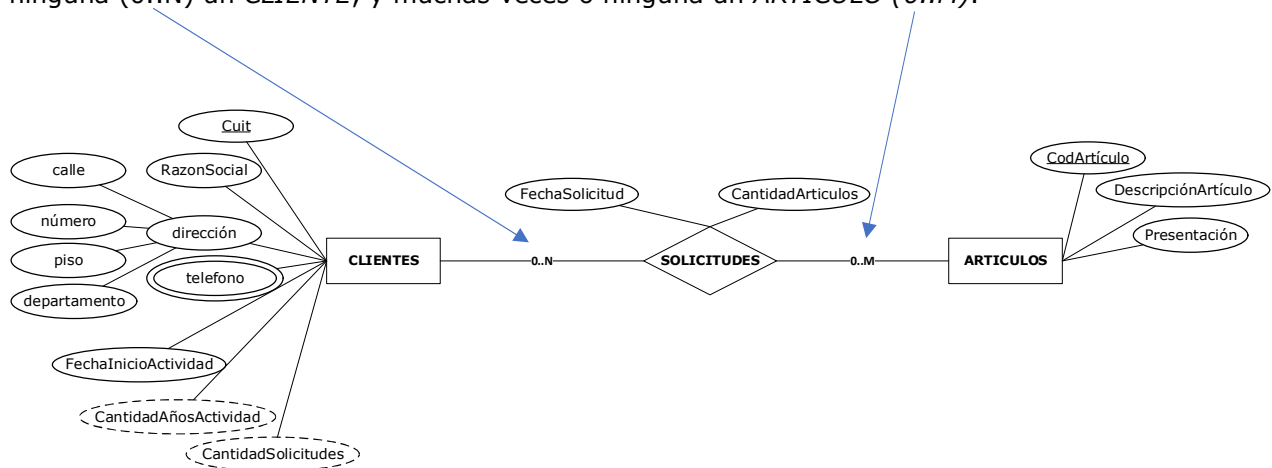
*SOLICITUDES* es un **vínculo binario** ya que relaciona dos entidades: *CLIENTES* y *ARTICULOS*. Más adelante veremos que pueden existir relaciones que vinculen otro número de entidades. Al número de entidades que participan en una relación se lo denomina **grado del vínculo**.

Podríamos agregar al ejemplo las siguientes reglas:

*"Un Artículo puede ser solicitado por varios clientes o por ninguno.*

*Un cliente puede realizar la solicitud de varios artículos o no haber realizado ninguna solicitud".*

Estamos entonces ante un vínculo en cuyas instancias pueden participar muchas veces o ninguna (0..N) un *CLIENTE*, y muchas veces o ninguna un *ARTÍCULO* (0..M).



**Figura 3**

Al número de instancias del vínculo en los que puede participar cada entidad de una relación o vínculo, se lo denomina **Razón de Cardinalidad**.

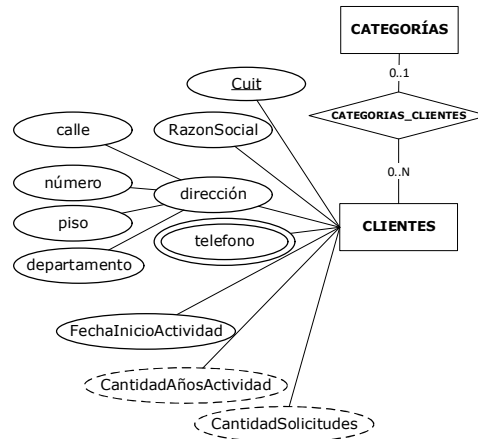
Debido a que no existe una amplia aceptación de un estándar para representar las cardinalidades en una relación optamos por la marcada en el dibujo: **0..M** significa cardinalidad mínima=0 y cardinalidad máxima=M. Para mencionar la cardinalidad del vínculo lo haremos con sus valores máximos.

En el ejemplo la cardinalidad del vínculo será: **N a M**.

Agregamos al ejemplo anterior las siguientes reglas:

"Periódicamente la empresa categoriza a sus clientes. Las categorías están codificadas y poseen una descripción. Inicialmente un cliente puede no tener una categoría asignada. Pueden existir categorías que no tengan clientes asignados".

Esta regla genera un vínculo **1 a N**. Gráficamente:

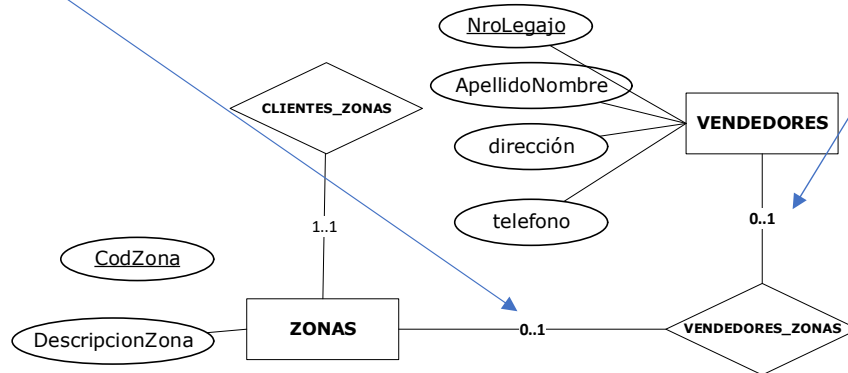


**Figura 4**

Otro tipo de vínculo binario que se puede establecer es el **1 a 1**.

Por ejemplo con las reglas:

- ✓ La empresa tiene definidas zonas para realizar las ventas y vendedores que las atienden.
- ✓ Las zonas están codificadas y tienen una descripción.
- ✓ Los vendedores se identifican con número de legajo y se tienen además nombre y apellido, dirección y teléfono.
- ✓ Un vendedor puede atender una única zona o ninguna. Una zona puede o no tener asignado un vendedor.

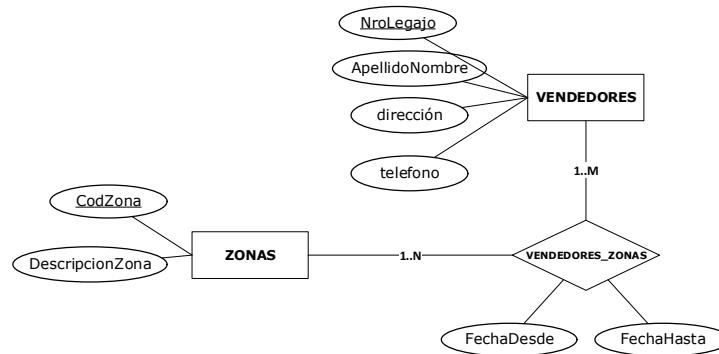


**Figura 5**

Véase qué importante son las reglas y cómo puede cambiar el modelo.

Por ejemplo, si se cambia la última regla mencionada por esta:

- ✓ Cada zona tiene asignado un único vendedor y un vendedor atiende una única zona desde una fecha y por un período determinado, es decir hasta una fecha. Puede ser que un vendedor vuelva a atender una misma zona en diferentes períodos.



**Figura 6**

Otra restricción estructural es la **Restricción de Participación** que especifica si la existencia de una entidad depende de que esté relacionada con otra entidad.

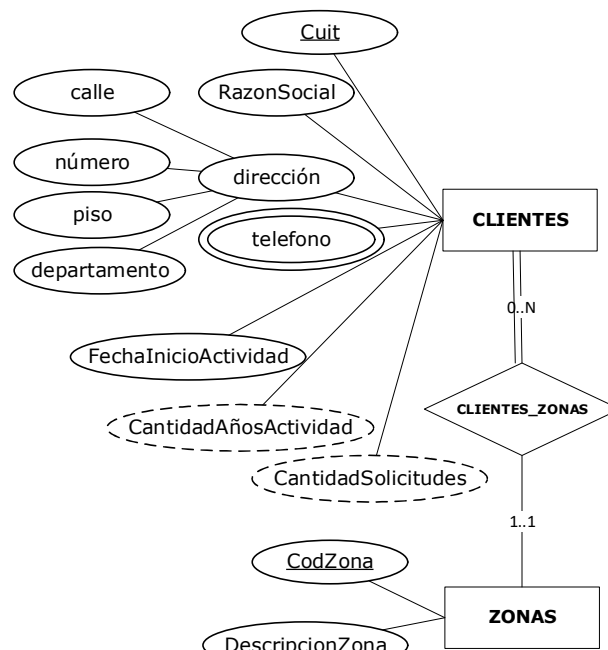
Hay dos clases de restricciones de participación **total** o **parcial**.

La restricción de participación **total** recibe el nombre de **dependencia de existencia**.

Si se establece la regla:

- ✓ Todo cliente **debe** pertenecer a una zona, pero una zona puede no tener clientes

gráficamente el modelo quedaría:



**Figura 7**

Es decir, una instancia de la entidad **CLIENTES** solo puede existir si participa en una instancia del vínculo **CLIENTES\_ZONAS**. Diremos que existe una **restricción de participación total o dependencia de existencia**. La dependencia de existencia se representa con línea doble.

No pasa lo mismo con **ZONAS** en la relación ya que puede existir una instancia de **ZONA** que no participe del vínculo **CLIENTES\_ZONAS**, es decir, no tiene ningún cliente. Diremos que la participación de **ZONA** en el vínculo **CLIENTES\_ZONAS** es **parcial**.



## ATRIBUTOS DE LAS RELACIONES O VÍNCULOS

Recordemos la regla:

- ✓ Los **clientes** realizan **solicitudes** de **artículos** en una fecha y por una cantidad.

Queda claro en este caso que los atributos *Cantidad* y *FechaSolicitud* pertenecen al vínculo *SOLICITUDES* (cardinalidad N a M) y no pueden trasladarse a ninguna de las entidades que participan del mismo.

En el ejemplo de la Figura 5 (Vínculo 1 a 1) si agregamos la regla "*un vendedor puede atender una única zona y lo hace determinados días de la semana*" podríamos agregar el atributo *DíasAtención* en la relación o trasladarlo a *VENEDORES* o a *ZONAS* indistintamente.

En los casos de vínculos 1 a 1, un atributo del mismo podrá ser trasladado a cualquiera de las entidades que participan del vínculo.

Tomando el ejemplo de la Figura 7 (Vínculo 1 a N), si agregáramos la regla "*debe conocerse la fecha en la que se asignó la zona al cliente*" podríamos agregar *FechaIngreso* como atributo en la relación *CLIENTES\_ZONAS*, pero preferiremos trasladarlo a la entidad *CLIENTES*.

En los casos de vínculos 1 a N un atributo del mismo sólo se podrá trasladar a la entidad del lado N del vínculo.

## ENTIDADES FUERTES Y ENTIDADES DÉBILES

Las entidades vistas hasta ahora como ejemplo son entidades denominadas **Entidades Fuertes**. Las entidades fuertes tienen un atributo o conjunto de atributos que identifican a sus instancias unívocamente, es decir un identificador o clave.

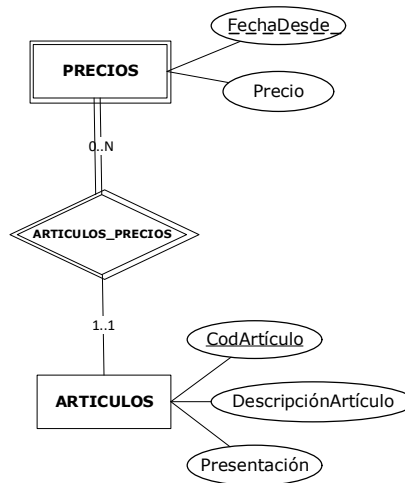
Es posible que una entidad no posea un identificador o clave asociado y su identificación dependa del vínculo con otra entidad. A estas entidades se las llama **Entidades Débiles** y se dice que hay **dependencia de identificación**. Al vínculo con la entidad de la que depende su identificación se lo denomina **vínculo identificador**.

Una entidad débil siempre tiene una **restricción de participación total** o **dependencia de existencia** con otra entidad a través de su vínculo identificador. Sin embargo, no toda dependencia de existencia da lugar a un tipo de entidad débil (como el ejemplo de la Figura 7).

Agregamos la regla:

- ✓ *Un artículo tiene para una fecha un precio y se desea poder tener un histórico para conocer los cambios de precios.*

Gráficamente para las entidades débiles, el rectángulo con que se representa la entidad y el vínculo identificador se marcan con doble línea y se refleja la dependencia de identificación como dependencia de existencia, también con doble línea:



**Figura 8**

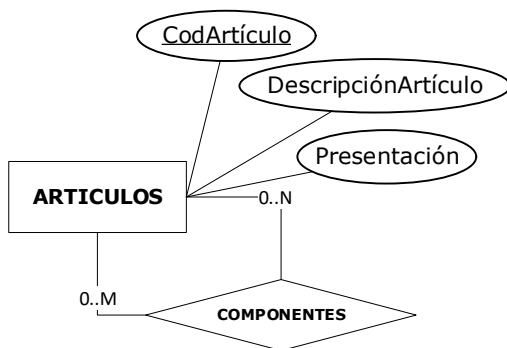
## OTROS TIPOS DE VÍNCULO

### RELACIONES RECURSIVAS O REFLEXIVAS

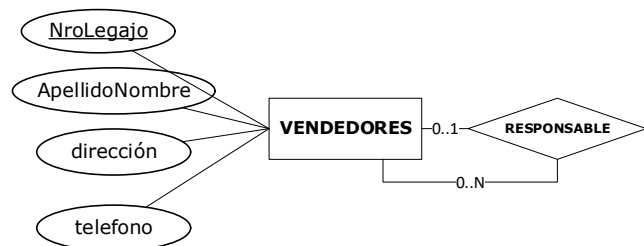
Supongamos que en el ejemplo que estamos trabajando los *ARTÍCULOS* que los clientes solicitan pueden ser parte de artículos compuestos. La regla podría indicar que: *"Los artículos pueden ser individuales o pueden ser compuestos por otros artículos"*. Esta relación sigue siendo binaria pero en este caso la relación vincula a la misma entidad (Figura 9).

Otro ejemplo de vínculo recursivo puede ser:

- ✓ *Los vendedores tienen un vendedor responsable superior identificado. Solo los vendedores encargados no tiene ningún responsable superior* (Figura 10).



**Figura 9**



**Figura 10**

## TIPOS DE RELACIÓN DE GRADO SUPERIOR A DOS

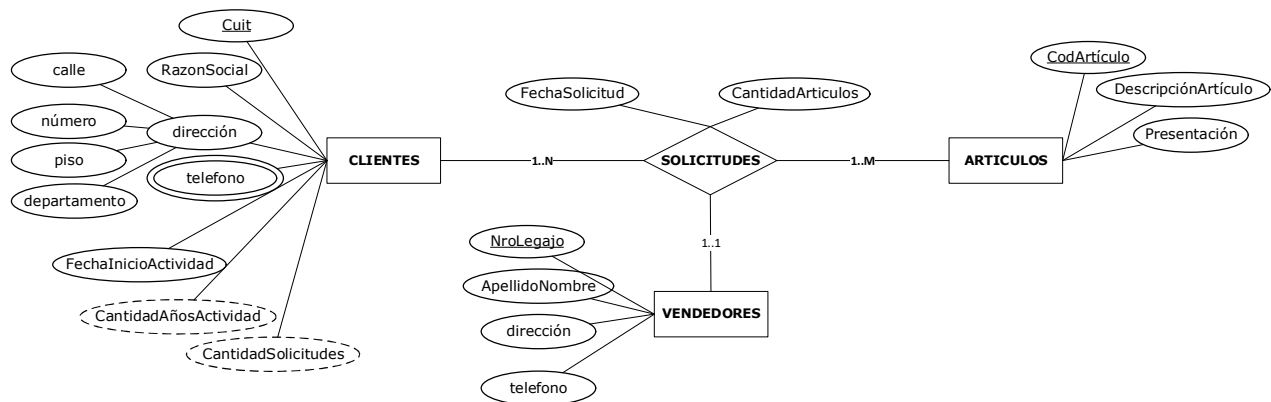
En varias ocasiones es complicado decidir si una relación deberá ser representada como un tipo de relación de grado N o si debería descomponerse en varios tipos de relación de grados inferiores. Se debe basar la decisión en la semántica o en el significado de la situación particular que se representa. En una relación ternaria las tres entidades deben estar relacionadas.

### RELACIONES TERNARIAS

Supongamos agregar la siguiente regla:

- ✓ *Cada solicitud de artículos, en el momento que se realiza, es registrada por un vendedor.*

Esto significa que deben vincularse en el mismo momento las entidades: clientes, artículos y vendedores.



**Figura 11**

Para minimizar la complejidad del análisis, las cardinalidades mínimas de una relación ternaria las consideraremos siempre 1.

Las cardinalidades de un vínculo ternario podrán ser todas 1..N, solo podrá existir una con cardinalidad 1..1, como en el ejemplo la cardinalidad a VENDEDORES.

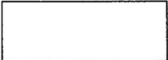
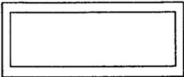
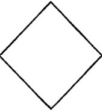




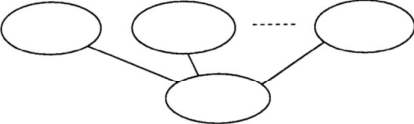

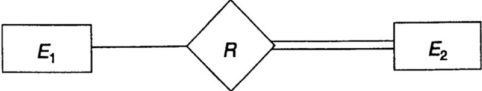
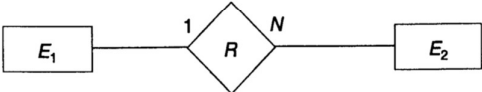
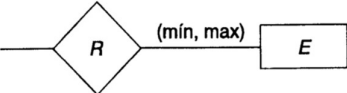
No necesariamente las entidades que participan de este tipo de relaciones deben ser entidades fuertes.

El modelo completo trabajado hasta aquí puede verse en la Figura 12.

Analice cómo quedaría el modelo si cambiáramos algunas reglas del ejemplo por estas:

- ✓ *La empresa compra los artículos a proveedores. Los proveedores se identifican con un CUIT y se conoce de los mismos Razón Social, Dirección y Teléfono.*
- ✓ *Los artículos están codificados y su código depende del proveedor al que se le compra, tienen además una descripción y una presentación*

## RESUMEN DE LA NOTACIÓN BÁSICA DE UN MODELO E-R

<u>Símbolo</u>	<u>Significado</u>
	TIPO DE ENTIDADES
	TIPO DE ENTIDADES DÉBIL
	TIPO DE VÍNCULOS
	TIPO DE VÍNCULOS IDENTIFICADOR
	ATRIBUTO
	ATRIBUTO CLAVE
	ATRIBUTO MULTIVALUADO
	ATRIBUTO COMPUESTO
	ATRIBUTO DERIVADO
	PARTICIPACIÓN TOTAL DE $E_2$ EN $R$
	RAZÓN DE CARDINALIDAD 1:N PARA $E_1;E_2$ EN $R$
	RESTRICCIÓN ESTRUCTURAL (mín, máx) DE LA PARTICIPACIÓN DE $E$ EN $R$

## AGREGACIÓN

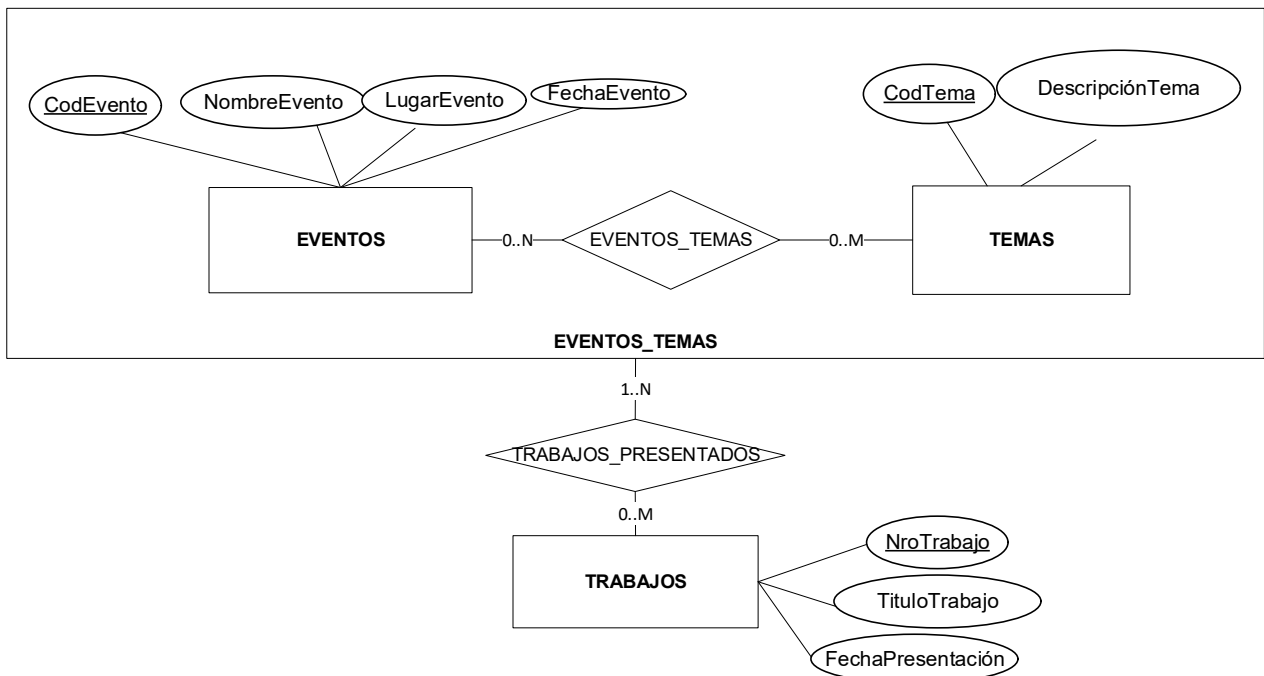
La **agregación** es una abstracción a través de la cual las relaciones se tratan como entidades de nivel más alto. Es útil cuando el objeto agregado de más alto nivel se debe relacionar con otro objeto posteriormente (no en el mismo momento como en el caso de los vínculos ternarios).

### IMPORTANTE:

- Una **agregación** hace referencia al **vínculo N a M** de las entidades que agrega (aunque su nombre puede diferir del nombre del vínculo son lo mismo).
- La agregación puede estar haciendo referencia a un vínculo **de cualquier grado** (por ejemplo a un vínculo ternario) y las **entidades** participantes del mismo pueden ser **de cualquier tipo** (fuertes, débiles, o también otra agregación).
- La agregación puede luego vincularse con otras entidades (fuertes, débiles o también otra agregación) y participar de vínculos de cualquier grado. También puede ser parte de una jerarquía.

Trabajemos sobre las siguientes reglas:

- ✓ La Asociación de Profesionales del Arte de Curar (APAC) tiene definidos temas de interés para la asociación. Los temas están codificados y tienen una Descripción.
- ✓ La APAC organiza eventos que están codificados, tienen un nombre, un lugar donde se realizará y la fecha de realización.
- ✓ Para completar la organización del evento la APAC selecciona los temas de los que tratará cada evento en particular.
- ✓ Una vez realizada la apertura a presentación de trabajos, los Profesionales podrán presentar trabajos al evento para los temas definidos hasta un mes antes de la fecha del evento. Los trabajos estarán asociados a uno o más temas definidos para el evento."

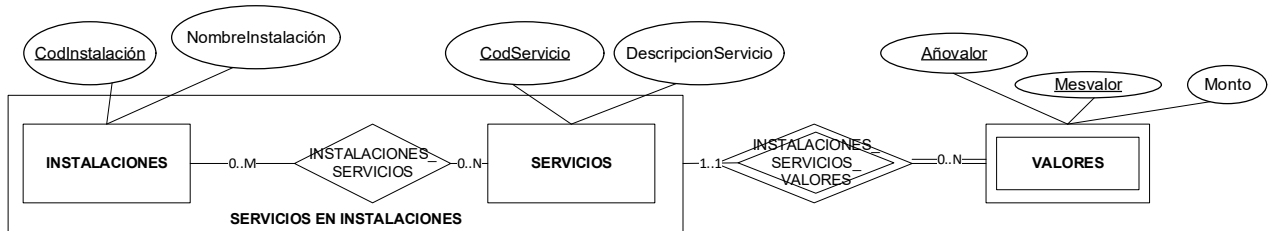


**Figura 17**

En este caso *EVENTOS* y *TEMAS* son entidades fuertes que se "agregan" para modelizar una entidad de nivel superior *EVENTOS\_TEMAS* para los cuales los profesionales pueden presentar *TRABAJO*. *TRABAJO*, en este caso es una entidad fuerte.

Otro ejemplo:

- ✓ Un club tiene diferentes instalaciones donde se realizan actividades. Las instalaciones están codificadas y tienen un Nombre
- ✓ Además cuenta con distintos servicios que brinda en las instalaciones mencionadas.
- ✓ Los servicios están codificados y tienen una Descripción.
- ✓ Los precios de los servicios dependen de la instalación en se brinden y son modificados mensualmente.



**Figura 18**

En este caso *INSTALACIONES* y *SERVICIOS* son entidades que se “agregan” para modelizar una entidad de nivel superior: *SERVICIOS EN INSTALACIONES* (o *INSTALACIONES\_SERVICIOS*) de la cual dependen la entidad *VALORES* que es una entidad débil. Esto quiere decir que *VALORES* tiene una restricción de participación total (dependencia de existencia) con respecto a su vínculo identificador: *INSTALACIONES\_SERVICIOS\_VALORES*.