Unidad 2. Análisis

Clave:

Telemática

21141209 / 22141209

Desarrollo de Software

15141211 / 16141211

Universidad Abierta y a Distancia de México

Índice

Presentación de la unidad 2
Propósitos de la unidad 3
Competencia específica 3
Logros 3
2.1. Lineamientos metodológicos de recopilación de requerimientos 4
2.1.1. Análisis. Concepto 5
2.1.2. Estructura de las entradas, procesos y salidas 6
2.1.3. Herramientas de estudios de factibilidad (operacional, financiero y económico) 8
2.1.4. Reporte del estudio de factibilidad 9
2.1.5. Requerimientos de hardware y software 10
2.2. Técnicas de modelado de datos 11
2.2.1. Modelado de objetos 12
2.2.2. Modelado entidad relación (E-R) 13
2.2.3. Normalización 16
2.2.4. Modelado relacional 19
Cierre de la unidad 22
Fuentes de consulta 23

Presentación de la unidad

Se ha llegado a la segunda unidad del curso, la cual permitirá internar a los estudiantes en la información sobre los conceptos adquiridos en la unidad anterior con mayor detalle y practicidad, ya que se abordará lo relacionado con el análisis para la creación del prototipo, haciendo uso de la terminología básica de la asignatura.

Antes de entrar de lleno a la elaboración del análisis como parte del proceso para la creación del prototipo, es necesario conocer algunas cuestiones que preceden a todas las demás actividades del análisis, como lo son los lineamientos metodológicos de recopilación de información, el estudio de sistemas y el análisis, lo que permite tener una idea más clara sobre lo que se pretende lograr en esta unidad.

Asimismo, se mencionan los conceptos acerca de qué es la estructura de las entradas, los procesos y las salidas, las herramientas de recopilación de información, el estudio de factibilidad operacional, financiero y económico, así como la forma de realizar un reporte de este estudio. También se hace un análisis sobre los requerimientos de hardware y software, para finalmente abarcar las técnicas de modelado de datos, donde se hablará de: modelado de objetos, modelo entidad relación, normalización y modelo relacional, siendo este último uno de los más utilizados en la actualidad empresarial.

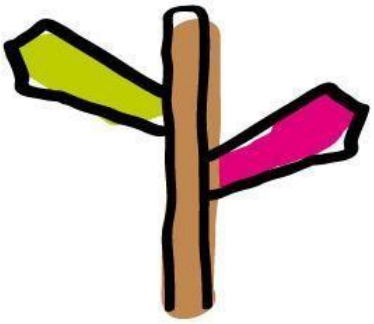
Los temas antes mencionados son abordados a lo largo de la unidad y aplicados en un caso de estudio, con el fin de seleccionar el Sistema Gestor de Base de Datos más adecuado de acuerdo a sus características y contribuir así en la solución del caso.

¡Aviso importante!

Para el uso y actualización de actividades, es necesario considerar versiones de Sistemas operativos y gestores, en concordancia con las actualizaciones de software y hardware, estables que permitan llevar a cabo las actividades aquí propuestas.

Bases de datos
Unidad 2. Análisis

Propósitos de la unidad



Continuar con el proceso para la elaboración del prototipo de Base de datos documental al elegir un caso de estudio, llevar a cabo un análisis y un estudio de factibilidad, identificar los requerimientos del usuario, administrar tiempos, así como revisar las técnicas de modelado, que son: modelado de objetos, modelado entidad relación, normalización y modelado relacional; lo que permitirá documentar el prototipo.

Competencia específica

Analiza problemas de gestión de información para elegir la alternativa que responda a los requerimientos de acuerdo al resultado de los datos obtenidos por medio del estudio de factibilidad (financiero, operacional y económico).

Logros

- Elegir el caso por los elementos que lo integran.
- Organizar la información del caso que se le presente, con base en las herramientas del estudio de factibilidad establecidas.
- Identificar mediante el estudio de factibilidad, los requerimientos del usuario y procesamiento de la información significativa para determinar tiempos en la realización de las actividades de la fase inicial del prototipo.
- Aplicar el modelado para abstracción de información del prototipo.

2.1. Lineamientos metodológicos de recopilación de requerimientos

En este apartado se revisa de manera general el procedimiento para realizar un adecuado análisis, partiendo desde la forma en la cual se logra recopilar la información necesaria, con el fin de profundizar en los siguientes subtemas.

Ahora bien, antes de realizar un protocolo de base de datos se debe anticipar todo un flujo de información y motivos, que hacen que se requiera un cambio de operación de cualquier empresa, esta acumulación de información se denomina “estudio de sistemas”, y es la que precede a todas las demás actividades del análisis.

Las bases de datos hacen mucho más que resolver problemas, con frecuencia se solicita ayuda para planificar la expresión de la organización, es entonces cuando se valoran de manera cuidadosa las necesidades futuras de la empresa y los cambios que deben considerarse para satisfacer esas necesidades. El tiempo, costo y beneficio son factores determinantes para desarrollar una opinión. Al final, la administración es quien decide cuál opinión aceptar. Una vez tomada la decisión se hace un plan para implantar la recomendación, el plan incluye saber con claridad cuáles son los requerimientos del usuario y, en este caso, se puede hacer por medio de un análisis que permita identificar las entradas, los procesos y las salidas, llevando al especialista a elegir la herramienta de recopilación de información que mejor se adapte al caso en cuestión, algunos ejemplos son las encuestas, entrevistas y cuestionarios.

Una vez aportados los criterios a utilizar, se continúa con el estudio de factibilidad, recomendado para dar un reporte informativo y detallado sobre los requerimientos del usuario. Cabe mencionar que es de suma importancia el reconocimiento de los requerimientos de hardware y software, los cuales permiten saber qué base de datos hacer y qué tipo de gestor utilizar, ya que cada uno de éstos tiene sus propias características. De esta manera, la documentación que se va elaborando

lleva al desarrollo de las técnicas de modelado, las cuales se abordan en el último apartado de la unidad.

2.1.1. Análisis. Concepto

¿Qué es el análisis?

Análisis es una fase del diseño de sistemas, que podemos definir como un conjunto de procedimientos que examinan hechos, principios y reglas clasificadas de manera ordenada y lógica; muestra resultados de solución a los requerimientos de información. Este proceso de análisis contiene varias **herramientas** a utilizar, según sea el caso a resolver, entre las que se encuentra el estudio de factibilidad (operacional, financiera y económica).

El estudio de factibilidad es una de las primeras etapas del desarrollo del prototipo, deberá incluir los alcances, objetivos y restricciones del requerimiento del sistema.

Ahora bien, **¿qué se debe de tomar en cuenta en el análisis?**

De primera instancia se deben detectar los hechos relevantes relacionados con la actividad de la empresa, la función es reunir información y determinar los requerimientos, para así pasar al análisis y el diseño, con la responsabilidad de que sean personas que saben programar, ya que sus conocimientos les permiten formular especificaciones mejores y más completas para las nuevas aplicaciones.

También es importante determinar a los usuarios finales; estos se agrupan en cuatro categorías:

1. " Usuarios primarios. Son los que interactúan con el sistema (base de datos). Ellos alimentan la base con datos de entrada, y reciben salidas por medio de una terminal.
2. Usuarios indirectos. Son aquéllos que se benefician de los resultados o reportes generados por estos sistemas, pero que no interactúan de manera directa con el hardware o software. Para este tipo de usuarios se deben incorporar consideraciones adicionales, tanto para la interacción, como para proteger de cualquier riesgo a la organización que proporciona el servicio.
3. Usuarios gerentes. Son los que tienen la responsabilidad administrativa en los sistemas de aplicación, este tipo de usuario es el que debe participar en los esfuerzos de desarrollo de la base de datos.
4. Usuarios directivos. Son aquellos que tienen mayor responsabilidad en los sistemas de información.

Los cuatro tipos de usuarios son importantes, ya que cada uno posee información esencial sobre las funciones de la organización y hacia dónde se dirige esta" (Hernández, s/f: 4)

Teniendo en cuenta los factores antes mencionados, a continuación se definirán los siguientes elementos que conforman los lineamientos metodológicos.

2.1.2. Estructura de las entradas, procesos y salidas

Para realizar la estructura de las entradas, procesos y salidas, se debe tener toda la información significativa de la empresa, la cual se procesará para obtener los diferentes resultados establecidos en el requerimiento de usuario.

También se deben tener en cuenta aquellas operaciones matemáticas y contables requeridas para el proceso de la información y generación de salidas. En esta parte es necesario realizar una tabla, en la cual se lleve a cabo un listado de aquellas salidas de información, y verificar si existe algún cálculo a realizar, es importante agregar la(s) fórmula(s) a aplicar.

Dentro de la estructura, las **entradas** se definen como variables constantes, que son la base que permitirá ejecutar las operaciones básicas de las bases de datos, las cuales son: altas, bajas y cambios; en casos más específicos, actualizaciones y modificaciones, convirtiéndose en la estructura de procesos.
Al hablar de las entradas de información en las especificaciones de la preparación del ingreso de información para su procesamiento, es elemental resaltar la importancia que tiene su correcta realización, porque de ella depende que los resultados de la información requeridos sean eficientes para los diferentes usuarios, y así tomar decisiones correctas que beneficien a la empresa. El buen diseño de entradas de información permitirá que el(los) usuario(s) encargado(s) de alimentar la base de datos aporte(n) de manera consistente e íntegra la información.

Los **procesos** son el juego que hacen las entradas para poder obtener las salidas correspondientes, dependiendo de los requerimientos.

Las **salidas** de información se refieren al diseño de la presentación de información ya procesada, requerida por el usuario, o los diferentes usuarios, del sistema; estas pueden ser reportes, mensajes, gráficos y estadísticos. Las salidas deben ser diseñadas lo más claramente posible y con gran precisión, para la toma de decisiones. Las salidas de información no solo son en pantalla, también se deben considerar en impresiones, audio y video.

En la siguiente tabla se muestra un ejemplo de la estructura general de las entradas, procesos y salidas.

ENTRADAS	PROCESOS	SALIDAS
DATOS DE ALUMNOS CALIFICACIONES	PROMEDIO = CÁLCULOS DE CALIFICACIONES	PROMEDIOS

Entradas, procesos y salidas.

Elección de caso

Ya teniendo los conceptos básicos y específicos de las bases de datos, así como sus características, ventajas y desventajas, se puede llevar a cabo la elección de un caso, por lo que a continuación se destacan algunas puntualidades que se deben tener en cuenta.

Para iniciar con esta elección es preciso saber qué es un **caso** y como estructurarlo**.** Pues bien, el caso es una técnica que nos permite visualizar problemas del pasado y presente, permitiendo estudiar la problemática y generar posibles soluciones, es decir, la realización de un prototipo de base de datos.

Para la construcción del caso es necesario considerar antecedentes, contexto, consecuencias y solución. Con el fin de realizar una buena estructura para la edificación del caso, el diseñador debe realizar las siguientes preguntas:

¿Qué es lo que sucede?
¿Cuál es el asunto o problema?
¿Quiénes están involucrados?
¿A qué situaciones se enfrentan?
¿Cuáles son los conflictos, intereses y posibilidades en juego?

Es importante que al realizar las especificaciones del caso esto se haga de manera coherente, detallando las cosas más sobresalientes, pero sin salirse de contexto. Al tomar en cuenta estos elementos se podrán generar preguntas clave o cruciales para su análisis o estudio.

Con base en las consideraciones anteriores, en la siguiente actividad se deberá elegir un caso de estudio.

2.1.3. Herramientas de estudios de factibilidad (operacional, financiero y económico)

Este apartado tratará sobre la importancia de contar con herramientas que permitan facilitar la recolección de la información, ya que no solamente es necesario conocer el caso para saber las especificaciones de los requerimientos de los usuarios, sino buscar técnicas y herramientas con las que se acceda a la información, y nos permitan recolectarla, y, con ella, determinar qué papel se jugará en este proceso, es decir, identificar los requerimientos operacionales, financieros y económicos, así como de hardware y software; por tal razón, en este tema vamos tomar como base el estudio de factibilidad.

El estudio de factibilidad está integrado por tres aspectos: operacional, financiero y económico. Tiene por objetivo verificar si un sistema de automatización de información es operacional, ya que para la implementación de un nuevo sistema de información es importante tener en cuenta que se requerirá de una inversión, y en este sentido el estudio de factibilidad permitirá saber, desde el punto de vista técnico y operacional, si es factible o no realizar dicha inversión, dependiendo de la rentabilidad que tenga para la empresa. El estudio deberá contemplar el costo de investigación o análisis, costo de software, costo de hardware, pago a personal calificado, etc., si no hay factibilidad económica el proyecto no existirá.

Para recopilar la información requerida para este estudio, se cuenta con el apoyo de los cuestionarios, entrevistas y encuestas. Los **cuestionarios** son utilizados con mayor frecuencia, dado que son menos costosos, estos fueron diseñados para realizar cuantificación de información y ahorrar tiempo, ya que permiten a las personas llenar o contestar las preguntas de manera escrita, sin la ayuda ni la intervención directa del investigador. Por otro lado, la **entrevista** es la realización de preguntas de forma directa, de persona a persona (entrevistador-entrevistado), en la cual se obtiene la información a analizar para el requerimiento del caso de manera más específica.

Finalmente, la **encuesta** es una herramienta que recopila información más enfocada a estudios de mercado y de opinión pública, la
cual es cuantificada.

Para obtener los requerimientos operacionales, financieros y económicos, observemos las siguientes preguntas, como ejemplo de algunas que pueden generarse:

Háblame de la empresa.

1. ¿Cómo describiría a la empresa?
2. ¿Qué hace diferente a esta empresa de otras del mismo giro?
3. Describa el logro de la empresa del cual esté más orgulloso.
4. ¿La empresa cuenta con bases de datos?
5. ¿Cuentan con el poder adquisitivo para el desarrollo de otra base de datos?
6. ¿Cuentan con hardware y software para la operacionalidad?

7. ¿Cuentan con hardware y software para el desarrollo?
8. ¿Considera que tener una nueva base de datos incrementará los ingresos de la compañía?

Este tipo de preguntas pueden llevarse a cabo en la entrevista, en el cuestionario y en la encuesta, la elección depende de los diseñadores de la base de datos, por tal razón, se sugiere investigar y realizar algunos ejemplos de estas herramientas para el caso en cuestión.

2.1.4. Reporte del estudio de factibilidad

En este subtema se muestran los pasos a realizar una vez que se ha elaborado el instrumento de recopilación de información.

Ahora bien, ya que se tiene el instrumento es necesaria su aplicación, porque la información que de él se obtenga, junto con su análisis, servirá para obtener los resultados de los requerimientos operacionales, financieros y económicos, que en el reporte de estudio de factibilidad se muestran punto por punto.

Estos resultados son los más valioso de la investigación, porque determinan si el sistema de información requerido es factible en lo operacional, financiero y económico. Para aclarar lo anterior observemos el siguiente ejemplo:

HARDWAR	SOFTWARE	OPERATIVIDAD	FINANCIERO	ECONÓMICO
Impresoras Servidores PC´s	- Sistema Operativo Paquetería Antivirus	- La Base de datos facilitará la administración de tiempos de la organización	- La organización incrementará sus activos económicos utilizando esta base de datos	- Dadas las características y Sistema Gestor de Base de Datos a utilizar, fue sustentable el costo de la misma

Ejemplo de reporte del estudio de factibilidad.

2.1.5. Requerimientos de hardware y software

Para recabar la información con respecto a los requerimientos de hardware y software, es necesario tener dos panoramas: el del usuario de la base de datos y el del programador.

En cuanto a los requerimientos de usuario, es importante cuestionar acerca de qué tipo de computadora tiene, así como las características del hardware y software, de manera que el programador tenga la visión de qué tipo de base de datos desarrollará, y qué manejador, dependiendo de sus características.

Para toda implementación de sistemas de información es importante definir los requerimientos de software y hardware, por lo que estos se abordarán a continuación.

Se entiende como requerimiento de hardware a los equipos de cómputo, servidores, impresoras, etc., considerando las diferentes características establecidas para la implementación del sistema. Mientras que el requerimiento de software se refiere a los programas necesarios para el funcionamiento del sistema de información, como lo son el Sistema Manejador de Base de Datos (SMBD), anti-virus, sistema operativo, etcétera.

Ahora bien, para definir los requerimientos de software y hardware es preciso realizar un estudio, el cual se lleva a cabo a través de tablas comparativas de las diferentes alternativas de solución, las cuales contienen los costos, características, ventajas y desventajas del software y hardware, así como los servicios de soporte ofrecidos por parte de los proveedores al realizar la compra de estos productos.

2.2. Técnicas de modelado de datos

Las técnicas de modelado de datos son las diferentes herramientas que se tienen para la abstracción conceptual de la información de manera gráfica, ayudan a la valoración y visualización de las especificaciones necesarias de información para la solución del caso; tienen un procedimiento lógico a seguir, que es utilizado en la ciencia del modelado.

La parte fundamental del análisis es integrar un modelado de sistemas que represente la información necesaria y requerida por el usuario, abstrayendo las cuestiones más importantes y significativas, que se utilizará para la generación de los diferentes modelos, y que permite visualizar el flujo de información.

Entre los diferentes modelados podemos encontrar:

- Modelado de objetos
- Modelado entidad relación (E-R)
- Normalización
- Modelado relacional
- Diagrama de flujo de datos

- Modelado entidad relación extendida (ERE)
- Modelado UML (*UnifiedModelingLanguage*), entre otros.

A continuación se profundizará en los primeros cuatro modelados, sin embrago es preciso que el estudiante investigue sobre el resto de ellos.

El modelado de datos corresponde a una serie de preguntas específicas importantes para cualquier aplicación de procesamiento de datos:

¿Cuáles son los objetos de datos primarios que van a procesar el sistema?
¿Cuál es la comprensión de cada objeto de datos?
¿Qué atributo describe el objeto?
¿Dónde residen actualmente los objetos?
¿Cuál es la relación entre los objetos y los procesos que lo transforman?

Para responder estas preguntas, los métodos de modelado de datos hacen uso del diagrama de entidad relación (DER). El DER, descrito con detalle posteriormente, permite que un ingeniero del software identifique objetos de datos y sus relaciones mediante una notación gráfica. En el contexto del análisis estructurado, el DER define todos los datos que se introducen, se almacenan, se transforman y se producen dentro de una aplicación.

2.2.1. Modelado de objetos

La técnica de modelado de objetos propone una forma abstracta de pensamiento acerca de problemas a resolver, empleando conceptos del mundo real, y no conceptos técnicos informáticos. Estos modelados son utilizados en los niveles conceptuales y de visión, se caracterizan por proporcionar restricciones de datos explícitamente, se fundan en pensar acerca de problemas a resolver. La esencia del desarrollo orientado a objetos es la identificación de conceptos de objetos del dominio de
aplicación; esta aplicación se ha centrado, por lo general, en lenguajes de programación, es una premisa básica para la detección de errores.
Dentro del modelado de objetos tenemos varios conceptos básicos que es necesario conocer para entender el tema, por lo que en esta parte se comenzará por definir el término **objeto,** hasta llegar al concepto de **instancia**.

Objeto (instancia de clase) es algo real, abstracto, acerca del cual se almacenan datos y métodos; por otra parte, cuando se dice **orientado a objetos,** significa que el software se organiza como una colección de objetos que contiene estructuras de datos y comportamiento.

Otro término relevante es el de **identidad**, el cual se refiere a que los datos están cuantificados en entidades discretas y distinguibles denominadas objetos, y que cada uno de estos posee su propia identidad inherente, es decir, los objetos son diferentes, aunque los valores de todos sus atributos, tales como nombre y tamaño, sean idénticos. En este contexto también se tiene el concepto **herencia,** que se trata de un mecanismo que permite definir nuevas clases a partir de otras ya definidas, y que la clase *padre* tiene atributos que

heredará a las clases *hijas*. Este término nos lleva al de **clase,** que es la unidad básica que encapsula toda la información de un objeto, a través de esta se puede modelar el entorno en un estudio (casas, cuentas, muebles, etc.).

A**tributos** son los valores asociados a los objetos de una clase, del cual se describen**.** Mientras que **clasificación** es la categorización de objetos con la misma estructura de datos (atributos) y comportamiento (operaciones). Finalmente, **instancia** es un objeto en el cual están contenidos todos los elementos que conforman un objeto.

Para concluir con este tipo de modelado se puede agregar que este provee un uniforme para modelar el sistema desde la captura de requerimientos en la etapa inicial del análisis, hasta su implementación, a través de todo el ciclo de desarrollo de sistemas. Para el modelado de objetos es necesario tomar en cuenta el análisis del negocio, que es el reconocimiento de los elementos claves de este, y la generación de la abstracción de entidades apropiadas u objetos- entidad.

Para entender la aplicación de estos términos en la conformación del propósito de este curso, es importante recordar que el **análisis** es una fase de la creación de sistemas informáticos en el que se identifican los objetos, clases, atributos y entidades a utilizar, esto permitirá realizar el **diseño lógico**, que es la integración de todas las clases requeridas para la aplicación.

2.2.2. Modelado entidad relación (E-R)

El modelado entidad relación por lo general es reconocido sólo por sus siglas (E-R), es un modelado conceptual de datos, considerado de alto nivel, que es utilizado como base en el diseño de bases de datos relacionales; simboliza información real a través de una representación gráfica, que es el dibujo que se hace empleando la terminología de **entidades,** que podemos definir como cualquier tipo de objeto sobre el cual se desea obtener información, también se le define como sustantivo. Por ejemplo: coches, casas, excursiones, empleados, clientes, empresas, oficios, diseños de productos, conciertos, etcétera.

Las entidades se representan gráficamente mediante rectángulos y su nombre aparece en el interior. Un nombre de entidad puede aparecer solo una vez en el esquema conceptual. Existen dos tipos de entidades: fuertes y débiles.
Una **entidad débil** es una entidad cuya existencia depende de la existencia de otra, y que por lo general carece de atributos. Mientras que una **entidad fuerte** es aquélla de la cual surgen las entidades débiles.

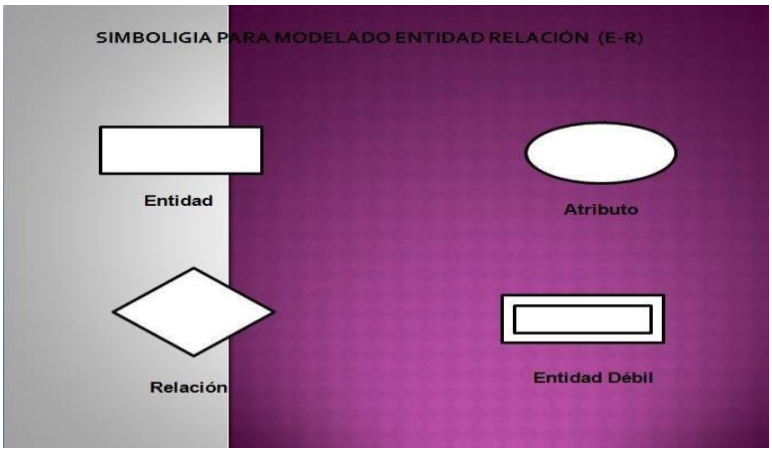
Las entidades se clasifican en:

- Tangibles (Ejemplo: persona, vehículo, material)
- Intangibles (Ejemplo: proyecto, materia)
- Eventuales (Ejemplo: compra, venta)

- Rol (Ejemplo: cliente, vendedor)
- Localización (Ejemplo: sucursal, estado, bodega)

Toda entidad está conformada de **atributos**, los cuales son las características que definen e identifican a las entidades (Ejemplo: tenemos a la entidad *alumno* que tendrá atributos como: nombre, sexo, fecha de nacimiento, CURP, RFC, etc.). Los atributos también son conocidos como características.

A continuación, se muestra la simbología a utilizar para el desarrollo de los modelados entidad relación (E-R).



Simbología del modelado entidad relación (E-R).

Estas entidades se encontrarán ligadas entre sí, y se utilizará el término **relación**, también conocido como **verbo,** para designar a la conexión que exista entre dos o más entidades. Las relaciones son representadas gráficamente con un rombo, dentro de él se describe la relación que existe entre las entidades. Existen básicamente tres tipos de relaciones:

Relaciones 1-1 Relaciones 1-n Relaciones n-n

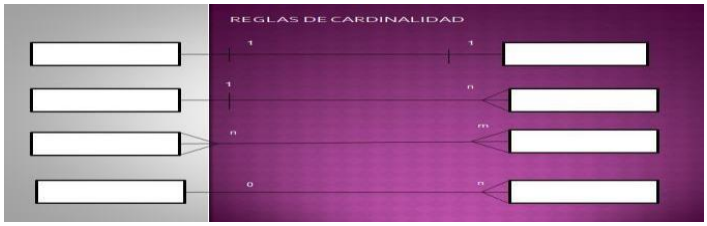
Relaciones 1-1. Las entidades que intervienen en la relación se asocian una a una (Ejemplo: la entidad ALUMNO y la entidad CARRERA. La relación que existe es que el alumno pertenece a una carrera).

Relaciones 1-n. Una ocurrencia de una entidad está asociada con muchas (n) de otra (Ejemplo: la entidad ALUMNO y la entidad MATERIAS, entre ellos la relación es que los alumnos cursan muchas materias).

Relaciones n-n. Cada ocurrencia, en cualquiera de las dos entidades de la relación, puede estar asociada con muchas (n) de la otra y viceversa (Ejemplo: la entidad ALUMNO y la entidad PROFESOR, entre ellos la relación es que un alumno tiene muchos profesores y viceversa); a la existencia de relaciones se llama **cardinalidad**, que describe la dimensión cuantitativa. Se visualiza esta cardinalidad con una entidad de manera jerárquica, como el esquema padre e hijo visto en la base de datos jerárquica.

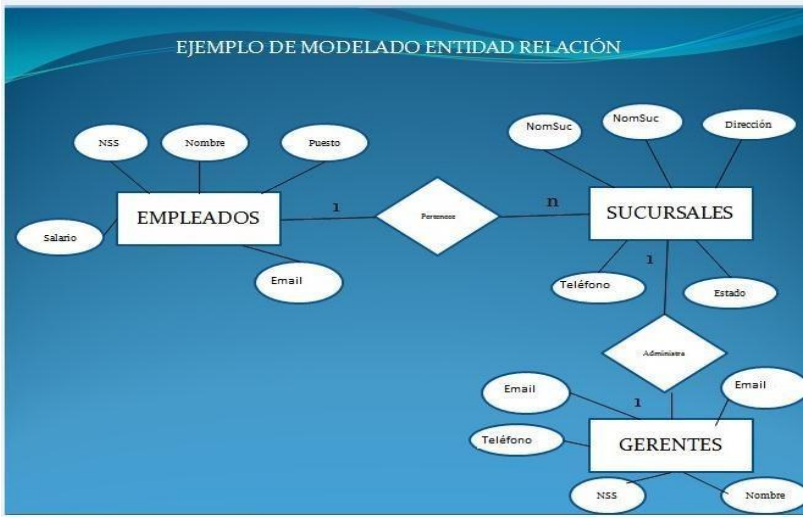
Reglas de cardinalidad.

Reglas de cardinalidad.



Reglas de cardinalidad

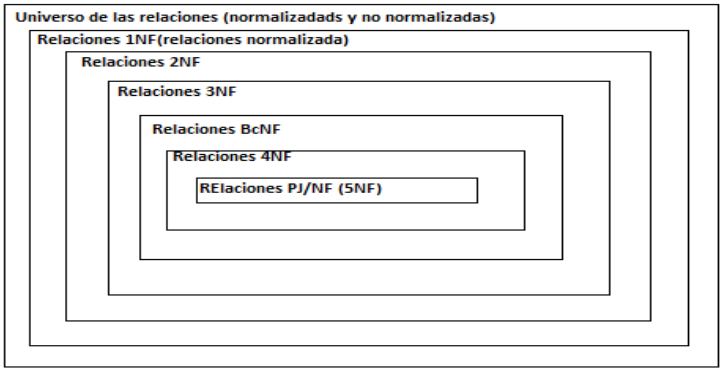
Al realizar el modelado E-R es importante considerar que se debe generar un **campo identificador,** que tendrá como funciónevitar la duplicidad de información (Ejemplo: en la entidad ALUMNO, el campo identificador será su matrícula). Con el campo identificador se podrán identificar entidades con el mismo nombre.
Por lo expuesto anteriormente, se puede concluir que una entidad es un sustantivo, los atributos son adjetivos y las relaciones son verbos.
En la siguiente figura se muestra un ejemplo sencillo de un modelado entidad relación (E-R).



Ejemplo de modelado entidad relación.

2.2.3. Normalización

La teoría de la normalización tiene como base el concepto de formas normales, mismo que dice que una relación está en una determinada forma normal si satisface un cierto conjunto de restricciones, de este modo, una relación está en primera forma normal (abreviada 1NF) solo si satisface la restricción de que sus dominios simples subyacentes contengan solo valores atómicos.



Universo de la normalizaciónen forma gráfica.

Se ha definido un gran número de formas normales, sin embargo, Codd definió la primera, segunda y tercera (1NF, 2NF, 3NF); la idea que perseguía Codd con su propuesta era la de mostrar las ventajas de las relaciones en 3FN respecto a las relaciones en forma normal inferior: mínima redundancia y mínimas anomalías al actualizar la base de datos. Debido a que aún persistían los problemas en las relaciones en 3FN, Codd introdujo en 1974 una definición más restrictiva de la tercera formal normal, que se denominó forma normal de Boyce-Codd (FNBC).

Fagin(1977, referido en Kroenke, 2003) introduce la cuarta forma normal (4FN), y posteriormente, en 1979, la quinta (5FN), basadas en otro tipo de dependencias distintas de las funcionales: las dependencias multivaluadas y las dependencias de proyección-combinación, respectivamente.

Cuando un esquema de relación está en una forma normal, implícitamente también está en las formas normales inferiores a esta, es decir, un esquema de relación en FNBC, está en 3FN, 2FN y 1FN; lo contrario no es cierto, un esquema de relación en 2FN no puede estar en 3FN.

Ya que se presentaron algunas generalidades, a continuación se formalizará la definición de las tres primeras formas normales y la de Boyce y Codd, que son las que apoyan en las dependencias funcionales (Kroenke, 2003).

La primera forma normal (1FN) es una restricción inherente al modelo relacional, por lo que su cumplimiento es obligatorio y afecta al número de valores que pueden tomar los atributos de una relación. Recordemos que para que una tabla pueda ser considerada una relación no debe admitir grupos repetitivos, esto es, que debe estar en primera forma normal; por ejemplo, si un estudiante solicita más de una beca, se tienen grupos repetitivos, y para pasar a 1FN habrá que repetir el resto de atributos de la tupla para cada uno de los valores del grupo repetitivo.

Definición:

Se dice que una relación está en 1FN cuando cada atributo solo toma un valor del dominio simple subyacente.

La segunda forma normal (2FN) está basada en el concepto de dependencia plena y en las interrelaciones existentes entre los atributos principales (que se encuentran en alguna de las claves) y no principales (que no se encuentran en ninguna clave) de una relación.

Definición:

Se dice que una relación está en 2FN si:

- Está en 1FN
- Cada atributo no principal tiene dependencia funcional completa respecto de cada una de las claves

De esta forma, cualquier relación binaria siempre se encuentra en 2FN; de igual forma la relación en la que todas las claves son simples, es decir, que contienen un sólo atributo; también, cualquier relación en la que todos sus atributos son PRINCIPALES, o dicho de otra manera, que forman parte de alguna clave.

La segunda forma normal no se cumple si algún atributo no principal depende funcionalmente de algún subconjunto de la clave. Sin embargo, se puede transformar un esquema de relación que no se encuentre en 2FN, en esquemas de relación en 2FN, sin que cause pérdida de información ni de dependencias.
Así pues:

“Sea el esquema de relación ESTUDIANTE_BECA (AT, DEP) donde:

AT= (cod_estudiante, cod_beca, fecha_sol, titulo)

DEP= (cod_estudiante, cod_beca, fecha_sol, cod_estudiante, titulo)

Que refleja las becas que solicitan los estudiantes, la fecha en que lo han hecho y la titulación del estudiante.

La clave de la relación ESTUDIANTE_BECA es cod_estudiante, cod_beca. Se puede observar que el atributo título no es un hecho (una información) acerca de la totalidad de la clave, si no acerca de parte de ella (en este caso del atributo cod_estudiante). Esta relación no está en 2FN” (Ibánez, 2007: 43).

La tercera forma normal (3FN) está basada en el concepto de dependencia transitiva.

Definición:

Un esquema de relación R está en la tercera forma normal sólo si:

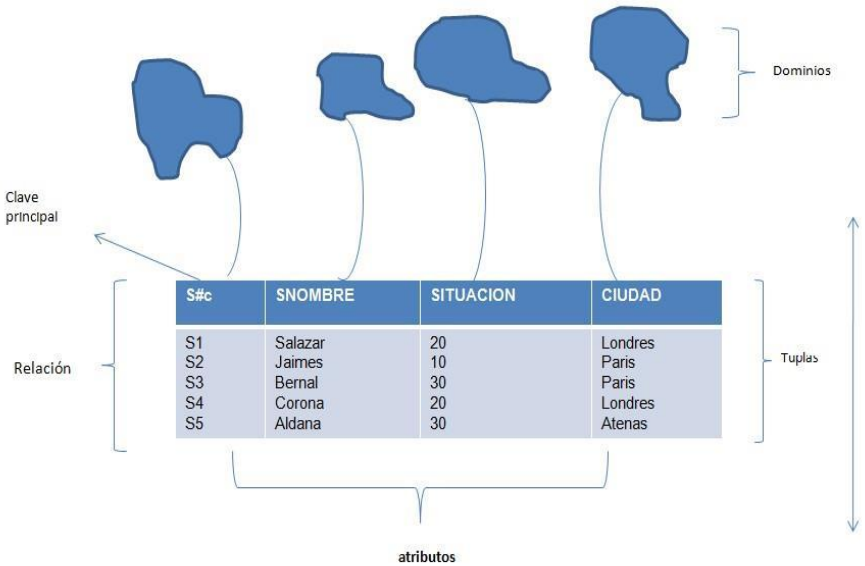
- Está en 2FN
- No existe ningún atributo no principal que dependa transitivamente de alguna clave R

Considerando la definición tenemos que, toda relación binaria se encuentra en 3FN, del mismo modo, toda relación cuyos atributos son todos principales, o bien cuando hay un único atributo no principal.

La tercera forma normal no se cumple cuando hay atributos no principales que dependen funcionalmente de otros atributos no principales. No obstante, se puede transformar un esquema de relación que no está en 3FN, en esquemas de relación en 3FN, sin que se originen pérdidas de información ni de dependencias funcionales.

2.2.4. Modelado relacional

El modelado relacional se divide en tres partes, que se ocupan de la estructura, la integridad y la manipulación de los datos, respectivamente. Cada una de las partes tiene sus propios términos especiales, estos conceptos son: relación, tupla, cardinalidad, atributo, grado, clave primaria y dominio.



S#c	SNOMBRE	SITUACION	CIUDAD
S1	Salazar	20	Londres
S2	Jaimes	10	Paris
S3	Bernal	30	Paris
S4	Corona	20	Londres
S5	Aldana	30	Atenas

Relación de proveedores S. A continuación se explica de manera general cada uno de los términos:

- Una relación corresponde a lo que hasta ahora hemos llamado tabla.
- Una tupla corresponde a una fila de esta tabla y un atributo a una columna. El número de tuplas se denomina cardinalidad, mientras que el número de atributo se llama grado.
- La clave primaria es un identificador único para la tabla, es decir, una columna o combinación de columnas con la siguiente propiedad: nunca existen dos filas de la tabla con el mismo valor en esa columna o combinación de columnas.
- Por último, un dominio es una colección de valores de los cuales uno o más atributos (columnas) obtienen sus valores reales. Por ejemplo, el dominio marcado con S# en la figura, es el conjunto de todos los números de proveedores legales; y el conjunto de valores que aparecen en el atributo S#, de la relación S, en cualquier momento es algún subconjunto de ese conjunto. De manera similar, el conjunto de valores que aparecen en el atributo S#, de la relación SP, en cualquier momento dado es también un subconjunto de ese conjunto.

Esta terminología se resume en la siguiente figura, sobre esta se harán dos aclaraciones:

Término relacional formal	Equivalentes informales
Relación	Tabla
Tupla	Fila o registro
Cardinalidad	Número de filas
Atributo	Columna o campo
Grado	Número de columnas
Clave primaria	Identificador único
Dominio	Fondo de valores

Terminología de la estructura de datos (Diseño de bases de datos, 2010: 6)

1. Debe entenderse que las “equivalencias” mostradas son solo aproximadas, porque los términos formales del modelo relacional, situados a la izquierda, tienen definiciones precisas, pero los “equivalentes” informales de la derecha solo poseen definiciones aproximadas, aunque prácticas. Así, por ejemplo, una relación y una tabla no son en realidad la misma cosa, aunque en la práctica muchas veces es conveniente hacer como si lo fueran.

2. La noción “dominio”, nos sirve para ilustrar una cosa muy importante: no todos los sistemas relacionales se ajustan a todos los aspectos del modelo relacional; DB2, por ejemplo, no maneja en absoluto los dominios, de hecho tampoco lo hace INGRES ni la mayor parte de los sistemas actuales.

Definición formal:
Dominio. El punto de partida para nuestro tratamiento formal de la estructura de datos tradicional es la menor unidad semántica de información, la cual suponemos es el valor de un acto individual (como el número de un proveedor individual o el peso de una parte individual o el nombre de una ciudad individual o la cantidad de un envío individual). Llamaremos a estos valores escalares (aunque este término no se utilice mucho en la literatura relacional).

Estructura de las bases de datos relacionales

Una base de datos relacional radica en un conjunto de tablas, donde a cada una de ellas se les asigna un nombre exclusivo; cada fila de la tabla representa una relación entre un conjunto de valores. Puesto que cada tabla es un conjunto de dichas relaciones, hay una correspondencia directa entre el concepto de tabla y el concepto matemático de relación, del que adopta su nombre el modelo de datos relacional.
Las bases de datos relacionales se usan varias relaciones diferentes para mostrar los conceptos propios del modelo de datos relacional, dichas relaciones constituyen parte de una entidad bancaria.

Cierre de la unidad

Aquí concluye la segunda unidad de la asignatura de Bases de datos. A manera de síntesis, se revisaron los lineamientos metodológicos de recopilación de requerimientos, se definió lo que es un análisis y la estructura de las entradas, procesos y salidas; asimismo se eligió un caso de estudio, se seleccionó y elaboró el instrumento para la recopilación de información y se abarcó lo que es un estudio de factibilidad. Finalmente, se abordaron cuatro técnicas de modelado de datos, con lo cual, en la siguiente unidad, se le dará continuidad al caso de estudio elegido.

Si los temas que se acaban de señalar son familiares, ya se está listo(a) para seguir con la unidad posterior, en donde se continuará con la generación del prototipo de base de datosdocumental. En caso de que no se tenga suficientemente claro alguno de los temas que se abordaron a lo largo de la unidad, se recomienda realizar un repaso con el fin de reforzar el aprendizaje.

Fuentes de consulta

- Celma, M.; Casamayor, J.C.; Mota, L. (2003) *Bases de datos relacionales.* Madrid: Pearson-Prentice Hall.
- Camps Paré, R. Et al. (2005). Bases de datos. Barcelona: Universitat Oberta de Catalunya https://www.uoc.edu/pdf/masters/oficiales/img/913.pdf
- Hernández, M. (s/f) *Introducción al desarrollo de sistemas de información.* Recuperado el 29 de abril de 2011.
- Ibáñez, A. (2007) *Tesis de grado "Modelo de optimización para el diseño de bases de datos relacionales".* La Paz, Bolivia: Universidad Mayor De San Andrés, Facultad De Ciencias Puras Y Naturales Carrera De Informática.
- Kroenke, D. (2003) *Procesamiento de Bases de Datos. Fundamentos, diseño e implementación.* México: Pearson Educación.
- Pérez L, César. (2008) *Oracle 10g: administración y análisis de bases de datos*. Segunda edición. México: Alfaomega.
- Quiroz, Javier. (2003) "El modelo relacional de bases de datos" en *Boletín de Política Informática* Núm. 6 (Versión electrónica). Recuperado el 26 de enero de 2011, de http://www.doanalytics.net/Documents/Modelo_Relacional.pdf
- Real Academia Española (2001) *Diccionario de la lengua española*. Vigésima segunda edición (Versión digital). Recuperado el 19 de enero de 2011, de http://buscon.rae.es/draeI/SrvltConsulta?TIPO_BUS=3&LEMA=inform%E1tica
- Silberschatz, Abraham. (2006). *Fundamentos de Bases de Datos*. España: McGraw-Hill.
- "Definición de requerimiento de hardware" en *Diccionario de informática*. Recuperado el 26 de junio de 2016, de http://www.alegsa.com.ar/Dic/requerimiento%20de%20hardware.php