| PROGRAMA DE FORMACIÓN | Gestión de analítica de datos |
| --- | --- |

| COMPETENCIA | 220501113 – Administrar base de datos de acuerdo con los estándares y requisitos técnicos | RESULTADOS DE APRENDIZAJE | 220501113-01 Modelar base de datos de acuerdo con especificaciones técnicas  220501113-02 Establecer el entorno de la base de datos de acuerdo especificaciones técnicas |
| --- | --- | --- | --- |

| NÚMERO DEL COMPONENTE FORMATIVO | CF09 |
| --- | --- |
| NOMBRE DEL COMPONENTE FORMATIVO | Fundamentos de modelamiento de datos |
| BREVE DESCRIPCIÓN | El modelamiento de datos, fundamental para el éxito del sistema de información organizacional, requiere definir los datos necesarios, qué decisiones se esperan tomar con ellos y cómo deben ser almacenados y procesados. En este módulo se presentarán los conceptos básicos y herramientas que permitirán realizar estas actividades, familiarizando al aprendiz con las metodologías y técnicas para su aplicación. |
| PALABRAS CLAVE | Bases de datos, Datos, gestión de la información, SQL. |

| ÁREA OCUPACIONAL | Ciencias naturales, aplicadas y relacionadas |
| --- | --- |
| IDIOMA | Español |

**Tabla de contenidos**

**Introducción**

1. **Introducción a la arquitectura de datos**
2. **Modelos de datos**
3. **Sistemas de gestión de datos SGBD (DBMS)**
4. **Modelo relacional de datos**
5. **Bases de datos NO relacionales**

# Introducción

| Cuadro de texto |
| --- |
| Estimado aprendiz, reciba una cordial bienvenida a este recurso de aprendizaje orientado a los fundamentos en el modelamiento de datos. Durante su desarrollo se abordarán los ejes temáticos de arquitectura y modelamiento de datos, sistemas de gestión de bases de datos, bases de datos relacionales y no relacionales. Al finalizar, estará en capacidad de modelar una base de datos conforme a los requerimientos. Antes de empezar, se sugiere revisar el siguiente video para identificar el contexto de aprendizaje:  ¡Le deseamos una experiencia de aprendizaje significativa y memorable! |

**Guion video introductorio**

**Ver Anexo 1-CF9-Guion Introductorio**

**Desarrollo de contenido**

# Introducción a la arquitectura de datos

| Tipo de recurso | Cajón de texto de color |
| --- | --- |
| Avances tecnológicos como el internet de las cosas (IoT), el comercio electrónico y las redes sociales hacen que cada minuto se produzcan enormes cantidades de información de vital importancia para las organizaciones y para la vida diaria. Una parte de esta información es estructurada y tiene un modelo definido que se puede organizar en forma de tablas, hojas de cálculo y bases de datos relacionales. La otra parte de la información, la no estructurada, no tiene una estructura específica como lo son las fotografías, documentos de texto, videos y comentarios en redes sociales, pero al igual que la información estructurada, brinda elementos importantes para el análisis y toma de decisiones. La incorrecta administración de cantidad ingente de datos, hace que su manejo resulte abrumador. | |
| Un minuto en Internet en 2020  Imagen recomendada, tomada de: [**https://www.statista.com/chart/17518/data-created-in-an-internet-minute/**](https://www.statista.com/chart/17518/data-created-in-an-internet-minute/)  Imagen: 228131\_ i1 | |
| Las empresas pueden (y deben) sacar provecho de los datos que generan día a día y esto es posible por medio de los sistemas de información, los cuales permiten recolectar, almacenar, procesar y explotar los datos disponibles para que sirvan de insumo para la planificación y la toma de decisiones. Así las decisiones estarán sustentadas en información y no en la intuición. | |

## Conceptos básicos: datos e información

| **Tipo de recurso** | Carrusel de tarjetas | |
| --- | --- | --- |
| **Introducción** | Los siguientes son algunos conceptos básicos necesarios para comprender cómo trabajar con los diversos tipos de datos que existen en cada organización y cómo se pueden gestionar a través de los sistemas de información para aprovecharlos en la gestión estratégica y en la toma de decisiones: | |
| Imagen: 228131\_ i2  **Datos - Información - Conocimiento**  **Imagen de muestra para reproducir algo similar que represente los 3 conceptos**  **Fuente:** <https://cursos.clavijero.edu.mx/cursos/area-basica/005_cia/modulo1/contenidos/12_diferencia_entre_dato_informacin_y_conocimiento.html> | | |
| **Datos**  Un dato es una representación simbólica de un atributo o característica de un elemento o evento. Por sí solo carece de significado. Los datos pueden ser de diferentes tipos: números, letras, objetos, figuras, entre otros. | | Imagen: 228131\_ i3  Datos  Imagen de muestra, tomada de CANVA |
| **Información**  La información es un dato o conjunto de datos con un contexto que les da significado. Son datos que han sido organizados, etiquetados y puestos en contexto. | | Imagen: 228131\_i930  Información  Imagen de muestra, link: <https://www.freepik.es/vector-> gratis/ilustracion-concepto-almacenamiento-memoria\_9712734.htm#query= informaci%C3%B3n%20almacenada&position=4&from\_view=search |
| **Conocimiento**  El conocimiento es la interpretación de la información. En el conocimiento se involucran la información, la reflexión, la experiencia para proporcionar valor. | | Imagen: 228131\_i931  Conocimiento  Imagen de muestra, link: <https://www.freepik.es/foto-gratis/maestro-sento-> mesa-ensenar-estudiantes-aula\_10069729.htm#page= 2&query=conocimiento%20saber&position= 10&from\_view=search |
| **Ejemplo**: dato versus conocimiento   | Dato | 21 | Sin un contexto e información adicional, carece de sentido. | | --- | --- | --- | | Información | Utilidad año 2021: 21% | Al darle contexto como el nombre de la variable y tipo, adquiere significado. | | Conocimiento | La empresa superó el año anterior el porcentaje de utilidad esperado | Se demuestra la comprensión y utilización de la información. | | | Imagen: 228131\_i932  Dato Vs Información Vs Conocimiento  Imagen de muestra que trate de ilustrar el concepto  Link: [https://www.freepik.es/foto-gratis/persona-dibujando-simbolos-que-salen-bombilla-encendida-encima-libro\_985251.htm#query= dato%20informaci%C3%B3n%20conocimiento&position](https://www.freepik.es/foto-gratis/persona-dibujando-simbolos-que-salen-bombilla-encendida-encima-libro_985251.htm#query=%20dato%20informaci%C3%B3n%20conocimiento&position)= 2&from\_view=search |
| **Datos estructurados**  Se consideran datos estructurados cuando tienen una estructura bien definida, siguen un orden y son de fácil acceso y uso para los sistemas de información y las personas.  Se suelen mostrar en filas y columnas con títulos. Por ejemplo: transacciones de compra / venta, base de datos de contactos, inventario de implementos. | | Imagen: 228131\_i933  Datos estructurados  Imagen de muestra que represente datos en filas y columnas, link: https://www.youtube.com/watch?v=7jVdLEWRY94 |
| **Datos no estructurados**  Este tipo de datos no tienen un modelo o estructura muy definida y pueden tener diferentes formatos como texto, imágenes, sonido, video lo que hace más difícil su procesamiento y análisis.  La mayor parte de la información relevante para la organización se puede clasificar como “no estructurada”. Por ejemplo: correos electrónicos, documentos de texto, imágenes de catálogo, audios, videos, comentarios en redes sociales. | | Imagen: 228131\_i934  Datos no estructurados  Imagen de muestra que presente información en diferentes formatos, link: <https://www.freepik.es/vector-premium/carpeta-almacenamiento-nube-archivos-diferentes-formatos_28953841.htm#query>= informaci%C3%B3n%20diferentes%20formatos&position= 14&from\_view=search |

## ¿Qué es la arquitectura de datos?

| Tipo de recurso | Cajón de texto de color |
| --- | --- |
| La arquitectura de datos es un conjunto de métodos que permiten definir:   1. Cómo va la organización a trabajar con los datos disponibles. 2. Cómo se van a almacenar, procesar y presentar estos datos, teniendo en cuenta los recursos y necesidades de los interesados. 3. Cómo se van a integrar modelos, políticas y reglas disponibles. 4. Su principal objetivo es el de modelar el sistema de datos de la organización, identificando:    1. Tipos de datos.    2. Fuentes de datos.    3. Las relaciones entre los datos y cómo estos interactúan. | |
| Realizar correctamente la definición de la arquitectura de datos de la organización permite:   1. Que se pueda sacar provecho a la cantidad de datos existentes. 2. Que los analistas realicen las actividades de procesamiento y análisis de una forma eficiente. 3. Una ventaja competitiva en comparación con las empresas u organizaciones que no cuentan con arquitecturas de datos claras. | |

| **Tipo de recurso** | Slider pasos | |
| --- | --- | --- |
| **Introducción** | La arquitectura de datos definida debe cumplir con las siguientes características:  **Arquitectura de datos**  **Imagen opcional – Elaboración propia**  Imagen: 228131\_i935 | |
| **Slide 1** | **Integración de datos**  Se deben identificar las diversas fuentes de datos existentes en la organización y las relaciones entre los datos que ofrecen. Por medio del uso de interfaces de programación de aplicaciones (API: Application Programming Interface por sus siglas en inglés), se pueden brindar mecanismos que permitan la comunicación entre diferentes componentes del sistema de información. | Imagen: 228131\_i936  **Integración de datos**  Imagen de muestra que dé cuenta del concepto  Link: https://www.freepik.es/vector-gratis/ilustracion-concepto-integracion-software\_18352184.htm#query=integraci%C3%B3n%20de%20datos&position=28&from\_view=search |
| **Slide 2** | **Seguridad**  Se deben identificar las amenazas existentes que puedan afectar la calidad, disponibilidad o confidencialidad de la información y desarrollar herramientas que disminuyan los riesgos de estas amenazas. | Imagen: 228131\_i937  **Seguridad de los datos**Imagen de muestra que ejemplifique el concepto de seguridad  Link: https://www.freepik.es/vector-gratis/seguridad-datos-global-seguridad-datos-personales-ilustracion-concepto-linea-seguridad-datos-ciberneticos-seguridad-internet-o-privacidad-proteccion-informacion\_12953630.htm#query=seguridad%20en%20datos&position=0&from\_view=search |
| **Slide 3** | **Escalabilidad**  Los sistemas de información están en constante crecimiento. La arquitectura de datos definida debe estar preparada para el aumento en la cantidad de datos y en la capacidad de procesamiento. Las soluciones en la nube facilitan esta escalabilidad brindándole al usuario acceder a los recursos bajo demanda (en caso de necesitar más espacio o mayor capacidad de almacenamiento). | Imagen: 228131\_i938  **Escalabilidad**  Imagen de muestra del concepto y de almacenamiento en la nube  Link: https://www.freepik.es/vector-gratis/gran-fuente-datos-centro-datos-computacion-nube-concepto-almacenamiento-nube-rack-sala-servidores\_4102872.htm#query=datos%20en%20la%20nube&position=32&from\_view=search |
| **Slide 4** | **Adaptabilidad**  La arquitectura de datos debe ser flexible para ajustarse a las diferentes necesidades de la organización y para los diversos roles o tipos de usuario. | Imagen: 228131\_i939  **Adaptabilidad**  Imagen de muestra que ejemplifique diversos roles en el manejo de los datos  Link: https://www.freepik.es/vector-gratis/ilustracion-concepto-abstracto-colaboracion-nube\_20770245.htm#query=roles%20manejo%20de%20datos&position=37&from\_view=search&track=ais |

## Modelos, políticas y reglas

| **Tipo de recurso** | Pestañas o tabs horizontales | |
| --- | --- | --- |
| **Introducción** | La arquitectura de datos es una rama de la arquitectura empresarial que abarca los modelos, políticas, reglas y estándares que orientan la recolección, el almacenamiento, la disposición, integración y el uso de los datos al interior de las organizaciones y su objetivo es traducir las necesidades de la organización en requerimientos de datos y herramientas para aprovechar la información y así atender dichas necesidades.  Existen diferentes marcos de referencia que definen las reglas y características de los sistemas de información a desarrollar, así como las mejores prácticas para la realización de este proceso de diseño e implementación de una arquitectura de datos acorde a los requerimientos de la organización. Estos marcos de referencia sirven como guías para un entendimiento detallado de las alternativas de arquitecturas disponibles según las características de cada organización y cómo realizar su implementación paso a paso para lograr un uso adecuado de las TIC y la información, así como el cumplimiento de las necesidades de la organización.    Imagen: 228131\_i940  **Modelos, políticas y reglas - Imagen de muestra**  El análisis y conocimiento de los diferentes marcos de referencia disponibles permitirá seleccionar la mejor opción según las necesidades de la organización, adaptando los modelos a las características particulares del negocio. | |
| **DMBOOK** | **DMBOK (Data Management Book of Knowledge)**  Creado por la Asociación de profesionales en gerencia de datos (DAMA Internacional), presenta una visión estándar sobre las funciones de la gestión de datos incluyendo: terminología y definiciones, funciones, procesos y prácticas, roles y responsabilidades, métricas*.*  Este marco de referencia propone las siguientes áreas de conocimiento:   * Gobierno de datos. * Arquitectura de datos. * Almacenamiento y operación de datos. * Seguridad de datos. * Integración e interoperabilidad de datos. * Documentación y contenido. * Referencias y datos maestros. * Bodega de datos e inteligencia de negocios. * Metadatos y calidad.   Además, define elementos de contexto soportados sobre tres pilares que son: personas, procesos y tecnologías. | **DMBOK**  Imagen: 228131\_i941  Utilizarla o recrear una imagen del concepto  Nota: adaptada de Data Management Book of Knowledge. (2022). <https://www.dama.org/events/enterprise-data-world-2022> |
| **DGF** | **DGF (Data Governance Framework)**  Provee una estructura lógica para clasificar y ejecutar actividades para la toma de decisiones involucrando estructuras, personas y procesos de la organización.  El marco de referencia DGF está formado por 10 componentes universales. Independientemente del enfoque de la organización, se incluirán los siguientes 10 componentes en el modelo, aunque el énfasis de cada uno varía según los requerimientos de la organización. Estos componentes son:   * Misión y visión. * Objetivos y métricas. * Reglas (definiciones y políticas). * Derechos de decisión. * Responsabilidades. * Controles. * *Stakeholders.* * Oficina de gobierno de datos. * Administradores de datos. * Procesos de gobernanza. | Nota*:* adaptada del Marco de gobierno de datos de la DGI (s.f.).  **Data Governance**  Imagen: 228131\_i942  Utilizarla o recrear una imagen del concepto |
| **TOGAF** | **TOGAF (The Open Group Architecture Framework)**  Define el proceso de creación de la arquitectura de datos como parte integral de la arquitectura de la organización.  Está sustentado bajo el modelo TAFIM (Technical Architecture Framework for Information Management), desarrollado por el Ministerio de Defensa de EE.UU. y a la fecha es administrado por The Open Group en su versión 9.2.  Este modelo de referencia utiliza un modelo iterativo de buenas prácticas conocido como “Método de desarrollo de arquitectura” (Architecture Development Method – ADM en inglés), con las siguientes fases:   * Fase preliminar. * Gestión de requerimientos:   1. Visión de arquitectura.   2. Arquitectura de negocio.   3. Arquitectura de sistemas de información.   4. Arquitectura tecnológica.   5. Oportunidades y soluciones.   6. Plan de migración.   7. Gestión de cambio de la arquitectura. | Nota: adaptada de El estándar TOGAF (2018).  **Gestión de Requerimientos (ADM)**  Imagen: 228131\_i943  Utilizarla o recrear una imagen del concepto |
| **COBIT** | **COBIT (Control Objectives for Information and related Technology)**  Es empleado como modelo para auditar la gestión y control de sistemas de información. Este marco de gestión de tecnologías de información, desarrollado por ISACA en 1996, permite a las organizaciones desarrollar, organizar e implementar sus estrategias en torno a la gestión de la información y la gobernanza.  Sus cinco componentes principales son:   1. Marco. 2. Descripción de los procesos. 3. Objetivos de control. 4. Directrices de gestión. 5. Modelos de madurez.   Tiene como principios claves:   * Cumplir con las necesidades clave de los interesados. * Cubrir la empresa de extremo a extremo. * Integrar múltiples marcos bajo un paraguas. * Fomentar un enfoque holístico para los negocios. * Alejar al gobierno de la administración. | **Principios de COBIT 5**  Imagen: 228131\_i944  Utilizarla o recrear una imagen del concepto  Nota: adaptada de Jiménez (2018). Geniusitt. <https://geniusitt.com/blog/que-es-cobit-5/> |

## Ciclo de desarrollo de la arquitectura de datos

| Tipo de recurso | Cajón de texto de color |
| --- | --- |
| El proceso de arquitectura de datos se debe desarrollar durante las fases iniciales de diseño y planificación del sistema de información, con el fin de establecer de manera oportuna y clara cómo se procesarán, utilizarán y almacenarán los datos. De la correcta realización de esta actividad dependerá en gran parte el éxito o el fracaso del sistema a implementar. Un sistema de información sin claridad en el origen y tipo de datos que utiliza o sin conocimiento de las políticas de uso de la información no será de utilidad para la organización. | |
| Proceso de arquitectura de datos  Imagen sugerida o de muestra  Tomada de: <https://www.freepik.es/vector-premium/colaboracion-trabajo-especialistas-devops-programadores-empresarios-huge-laptop-ilustracion-plana-dibujos-animados_13482569.htm#query=ciclo%20de%20desarrollo%20software%20pasos&position=8&from_view=search>  Imagen: 228131\_i945 | |

| **Tipo de recurso** | Slider Hitos/ Línea de tiempo horizontal | |
| --- | --- | --- |
| **Introducción** | Este proceso está dividido en cuatro etapas, similares a las etapas de desarrollo de arquitectura de *software*, pero enfocadas en los datos:    Imagen propia  Imagen: 228131\_i946 | |
| **Requerimientos** | Se enfoca en la captura, documentación y priorización de requisitos que influyan en la arquitectura de datos.  En esta etapa es importante:   1. Conocer las expectativas de los *stakeholders* (personas interesadas y/o impactadas por el proceso). 2. Los objetivos del sistema a desarrollar (qué información o *insights* se podrán obtener con la información procesada y analizada). 3. Los flujos de trabajo que describen cómo se mueve la información a través de los diferentes procesos de la organización. 4. Las condiciones o restricciones existentes en la organización para el uso y difusión de la información. 5. Aquí también se analizarán los orígenes de los datos (fuentes de información), su tipo (estructurados o no estructurados), dimensiones y periodicidad.   Esta etapa debe ser documentada de forma clara y detallada para no omitir detalles que puedan impactar negativamente el diseño y desarrollo del sistema de información. | **Fase de Requerimientos**  Imagen de muestra de la fase de requerimientos de la arquitectura de datos  Link: <https://www.freepik.es/vector-gratis/ilustracion-composicion-concepto-isometrico-desarrollo-web_13805027.htm#query=desarrollo%20de%20software&position=4&from_view=search>  Imagen: 228131\_i947 |
| **Diseño** | En esta etapa se toman los requerimientos documentados y se definen las estructuras necesarias para el funcionamiento del sistema de información.  Esta definición se hace utilizando patrones de diseños según las necesidades de la organización. En esta fase también se decide cuáles serán las tecnologías a utilizar para la gestión y almacenamiento de los datos y se identifican las relaciones entre los datos y cuál será el flujo de la información desde su origen hasta el final del proceso. | Imagen: 228131\_i948  Fase de Diseño  Imagen de muestra de la fase de diseño  Link: https://www.freepik.es/vector-gratis/desarrollo-web-ingenieria-programadores-sitio-web-codificacion-pantallas-interfaz-realidad-aumentada-desarrollador-ingeniero-proyectos-software-programacion-o-diseno-aplicaciones-ilustracion-dibujos-animados\_10798281.htm#query=desarrollo%20de%20software&position=1&from\_view=search |
| **Documentación** | Se documenta el diseño realizado, utilizando representaciones ya establecidas y estandarizadas para facilitar la comunicación del diseño a los demás actores del proceso.  Algunas herramientas que permiten plasmar el diseño realizado son:   * Mapas de procesos. * Diagramas de flujo. * Diagramas de entidad – relación. | Imagen: 228131\_i949  Fase de documentación  Imagen de muestra de la fase de documentación (mapas)  Link: <https://www.freepik.es/vector-premium/diagramas-flujo-conjunto-6-esquemas-diagramas-flujo-diagramas-elementos-infograficos-editables-color-simple_30249004.htm#query=mapas%20de%20procesos%20diagrama%20de%20flujo&position> =17&from\_view=search |
| **Evaluación** | Esta etapa permite detectar de manera temprana fallas y oportunidades de mejora en el diseño realizado. Esto permitirá identificar y corregir posibles problemas antes de pasar a la fase de desarrollo. | Imagen: 228131\_i950  Fase de documentación  Imagen de muestra de la fase de evaluación y corrección de errores  Link: https://www.freepik.es/vector-premium/desarrollo-software-sistema-solucion-software-todo-modernizacion-sistema-central\_30836747.htm?query=correcci%C3%B3n%20de%20errores |

# Modelos de datos

| **Tipo de recurso** | | Pestañas o tabs Verticales |
| --- | --- | --- |
| **Introducción** | | El modelado de datos es un proceso que permite analizar y definir los datos generados por la organización y la relación entre estos, con el fin de identificar sus características y dependencias. Un modelo de datos bien definido permite crear sistemas de información eficientes con estructuras de datos simplificadas mejorando así el rendimiento del sistema y por consiguiente los resultados que arroja. |
| **Modelado de datos**  **Imagen que represente modelado de datos**  **Link:** [**https://www.freepik.es/vector-gratis/plantilla-portada-facebook-privacidad-datos-diseno-plano\_28206831.htm#page=2&query=modelos%20de%20datos&position=8&from\_view=search**](https://www.freepik.es/vector-gratis/plantilla-portada-facebook-privacidad-datos-diseno-plano_28206831.htm#page=2&query=modelos%20de%20datos&position=8&from_view=search)  Imagen: 228131\_i951 | | |
| **Niveles de abstracción** | Existen tres niveles de abstracción de modelado de datos que son:   * Modelado conceptual. * Modelado lógico. * Modelado físico.   Estos modelos permiten determinar y visualizar cómo se gestionan los datos de la organización. | |
| **Modelo conceptual** | Este tipo de modelo corresponde al nivel de abstracción más alto, en el cual se representan la estructura general y el contenido, pero sin mucho detalle. En este, se identifican los conjuntos de datos y el flujo de estos en la organización. La herramienta utilizada para esta representación es el diagrama entidad – relación.  En este modelo se pueden encontrar:   * Datos del sistema de información. * Atributos de los datos (condiciones y restricciones). * Reglas empresariales (relaciones entre los datos). * Cómo se organizan los datos. * Requerimientos de seguridad e integridad de datos. | |
| **Modelo lógico** | Este modelo describe los datos con más detalles que el modelo conceptual, pero sin incluir especificaciones propias de la base de datos. Esto con el fin de poder aplicar el modelo a diferentes tecnologías. En este modelo se describen los atributos de los elementos modelados en el nivel de modelo conceptual.  En este modelo se pueden encontrar:   * Tipos de datos de los atributos (número, cadena, fecha). * Relaciones de datos. * Llaves primarias de los datos. | |
| **Modelo físico** | En este, se describe de forma detallada, cómo se implementará el modelo lógico. Incluye las estructuras de almacenamiento y método de acceso a los datos y depende del sistema de gestión de bases de datos (SGBD) específico, que se haya seleccionado. De un solo modelo lógico, se pueden derivar varios modelos físicos, según el SGBD a utilizar.  En este modelo se pueden encontrar:   * Tipos de datos de los atributos (según el SGBD seleccionado). * Relaciones de datos (según el SGBD seleccionado). * Información adicional del modelo. | |

## ¿Qué son modelos de datos?

| Tipo de recurso | Cajón de texto de color |
| --- | --- |
| Un modelo de datos es una representación visual del sistema de información para identificar las estructuras de datos que lo componen, así como las conexiones entre estas estructuras.  Estos modelos pueden tener diferentes niveles de abstracción según la etapa del proceso de diseño en la que se encuentren. Los modelos de datos conceptuales (de más alto nivel) permiten visualizar las relaciones entre los datos mientras que los modelos lógicos representan mejor la forma en la que serán almacenados los datos en el servidor. | |
| **Modelo de Datos**  **Imagen de ejemplo**  **Fuente:** [**https://www.lucidchart.com/pages/database-diagram/database-models**](https://www.lucidchart.com/pages/database-diagram/database-models)  Imagen: 228131\_i952 | |

## Tipos de modelos de datos y marcos de referencia

| **Tipo de recurso** | Slider Presentación | |
| --- | --- | --- |
| **Introducción** | Existen varios tipos de modelos de datos que se diferencian por la forma en la que se vinculan los registros entre sí:   * Modelo jerárquico. * Modelo de red. * Modelo relacional. * Modelo orientado a objetos, entre otros. | |
| **Modelo relacional**  Clasifica los datos en tablas que constan de columnas y filas. Las columnas son los atributos del elemento en cuestión (conocido también como entidad). Uno de los atributos (o una combinación de ellos) será la llave primaria. Cada fila de la tabla (también conocida como tupla o registro) recoge los datos sobre una instancia específica de la entidad.  En el ejemplo de la imagen, se observa que, en el caso de la primera tabla, la llave primaria es el “ID Vendedor” el cual es un atributo único (no se repite en las otras filas). Esa llave primera se usa en la segunda tabla para relacionar la información del vendedor con las ventas realizadas. El vendedor con el ID 577856 tiene dos registros en la segunda tabla (ventas). Para el vendedor con el ID 1203565 no se registran ventas en la tabla. | | **Modelo Relacional**  **Imagen propia**  Imagen: 228131\_i953 |
| **Modelo jerárquico**  El modelo jerárquico organiza los datos en una estructura similar a un árbol, donde cada registro tiene un solo padre o raíz. Cada padre puede tener varios hijos, pero cada hijo viene de un solo padre (relación 1: N también llamada 1 a muchos).  En el ejemplo de la imagen, se observa que el DEPTO1 es el padre de los elementos PROFESOR 1, PROFESOR 2 y PROFESOR 3. Cada elemento “Profesor” pertenece solo a 1 departamento (un solo padre).  **Modelo en red (*Network*)**  Es una expansión del modelo jerárquico basado en la teoría de conjuntos. Este modelo, a diferencia del modelo jerárquico si permite relaciones muchos a muchos (N: N). | | Imagen: 228131\_i954  **Modelo Jerárquico**  **Imagen de ejemplo, cambiar el estilo, pero reproducir con los mismos textos explicados en el ejemplo**  **Fuente: https://es.m.wikipedia.org/wiki/Archivo:DB\_jerarquica.png** |
| **Modelo entidad - relación (ER)**  Este modelo utiliza diagramas para representar las relaciones entre las entidades (personas, objetos, conceptos) de una base de datos. Existen diferentes herramientas para la realización de este tipo de modelos entre ellas: GitMind, Lucidchart y Visual Paradigm. | | Imagen: 228131\_i955  **Modelo Entidad relación**  **Imagen de ejemplo**  **Link: https://upload.wikimedia.org/wikipedia/commons/thumb/f/f6/Ejemplo\_Diagrama\_E-R\_extendido.PNG/400px-Ejemplo\_Diagrama\_E-R\_extendido.PNG** |

**Elementos del modelo entidad - relación (ER)**

| **Tipo de recurso** | | Pestañas o tabs Verticales |
| --- | --- | --- |
| **Introducción** | | El modelo entidad relación, tiene los siguientes elementos: |
| **Elementos del modelo E-R**  Imagen: 228131\_i956 | | |
| **Entidades** | Se refiere a un elemento u objeto del mundo real (o un concepto abstracto) que se puede diferenciar de otros objetos.  Las entidades tienen propiedades o atributos que las caracterizan. Por ejemplo: Programas académicos, estudiantes, profesores. Las entidades se representan por medio de rectángulos. | |
| **Atributos** | Son las características o propiedades que tiene una entidad. Cada elemento de la entidad tiene los mismos atributos, pero con valores propios para cada uno. Por ejemplo, para la entidad estudiante se tienen como atributos: identificación, nombre, apellido, fecha de nacimiento, género. Existe un atributo principal (o llave primaria para las bases de datos) cuyo valor es único para cada elemento de la entidad y no pueden existir dos elementos con el mismo valor. Los atributos se representan por medio de óvalos. | |
| **Relación** | Es el vínculo que existe en dos entidades. Describe cómo es la dependencia entre ellas y, por lo general, se presentan como un verbo. Las relaciones tienen una cardinalidad, es decir, la cantidad de elementos de una entidad A, que se relacionan con elementos de la entidad B. Esta cardinalidad puede ser: uno a uno (1:1), uno a muchos (1:M) o muchos a muchos (N:N). Las relaciones se representan por medio de rombos. | |

## Selección del modelo de datos adecuado para el negocio

| Tipo de recurso | Cajón de texto de color |
| --- | --- |
| Es necesario analizar las fortalezas de cada modelo en particular y los requerimientos del negocio para seleccionar aquellos que mejor describan a la organización. No existe un modelo más idóneo que otro, pero sí algunos que por sus características permiten describir mejor los datos que tiene la organización y las relaciones entre estos. | |
| De la correcta selección y elaboración del modelo de datos de la organización, dependerá el éxito del sistema a desarrollar.  Aquí, es importante tener en cuenta el sistema de gestión de base de datos que se esté utilizando (si ya existe uno en uso en la organización), para validar que sea compatible con el modelo de datos a utilizar. | |
| La siguiente tabla compara algunos de los modelos de datos disponibles para facilitar su análisis:  **Tabla 1** *Comparación de modelos de datos*    Imagen: 228131\_i957 | |

# Sistemas de gestión de bases de datos SGBD (DBMS)

| **Tipo de recurso** | Tarjetas Animadas | | |
| --- | --- | --- | --- |
| **Introducción** | Los sistemas de gestión de bases de datos - SGBD (o DBMS por sus siglas en inglés: DataBase Management System)son fundamentales para la administración de la información de una organización. Son herramientas de *software* que permiten configurar, administrar y utilizar bases de datos y así sacar provecho de la información que allí se almacena. | | |
| Imagen: 228131\_i958  Sistemas Gestores de Bases de Datos (SGBD)  Imagen de muestra  Link: https://iutablog.wordpress.com/2016/04/26/sistema-gestor-de-base-de-datos/ | | | “Al almacenar los datos en un SGBD, en vez de en un conjunto de archivos del sistema operativo, se pueden utilizar las características del SGBD para gestionar los datos de un modo robusto y eficiente. A medida que crecen el volumen de los datos y el número de usuarios el apoyo de los SGBD se vuelve indispensable”. (Ramakrishnan, 2011). |
| Imagen: 228131\_i959  Propiedades de los SGBD  Imagen de muestra  Link: http://dbadixit.com/modelo-acid-transacciones/ | | | Los SGBD deben mantener cuatro propiedades en las transacciones para garantizar que los datos se procesen de manera fiable.  Estas propiedades son conocidas como ACID, por sus siglas en inglés, y son:   * Atomicidad. * Consistencia. * Aislamiento. * Durabilidad. |
| Imagen: 228131\_i960  Atomicidad  Imagen de muestra  Link: http://contenidos.sucerman.com/nivel2/web1/unidad1/leccion1.html | | **Atomicidad**  Un cambio se debe realizar en su totalidad o no modificar nada en absoluto. | |
| Imagen: 228131\_i961  Consistencia  Imagen de muestra  Link: http://contenidos.sucerman.com/nivel2/web1/unidad1/leccion1.html | | **Consistencia**  Todo cambio debe conducir o terminar en un estado válido de la base de datos. | |
| Imagen: 228131\_i962  Aislamiento  Imagen de muestra  Link: http://contenidos.sucerman.com/nivel2/web1/unidad1/leccion1.html | | ***Isolation* (Aislamiento)**  Un cambio no debe afectar a otros cambios que se estén realizando sobre la base de datos de manera concurrente. | |
| Imagen: 228131\_i963  Durabilidad  Imagen de muestra  Link: http://contenidos.sucerman.com/nivel2/web1/unidad1/leccion1.html | | **Durabilidad**  Una vez realizado el cambio, sus efectos deben persistir incluso si se producen fallos en la base de datos. | |

## 

## Evolución de los sistemas de gestión de bases de datos

| Tipo de recurso | Cajón de texto de color |
| --- | --- |
| Los sistemas de gestión de bases de datos (SGBD) vienen en evolución desde sus inicios como sistemas manuales hasta los sistemas actuales en línea y distribuidos, que permiten la gestión de enormes cantidades de información. | |
| **Figura 1**  *Evolución de los sistemas manuales*    Imagen: 228131\_i964  Línea de tiempo de los SGBD  Imagen de muestra para reproducir una propia  Link: <https://www.semanticscholar.org/paper/Evoluci%C3%B3n-de-las-Bases-de-Datos%3A-de-Fijas-a-M%C3%B3viles-Fern%C3%A1ndez-Moreno/4e26288271493a84bc43973b5471ca8462116853> | |
| El siguiente es un listado de los SGBD más utilizados en el mercado:  **Tabla 2**  *Listado de los sistemas de gestión de bases de datos de mayor uso*   | **SGBD** | **Nombre** | **Tipo** | | --- | --- | --- | | Imagen: 228131\_i965  Imagen del logo  Link https://aws.amazon.com/es/rds/mysql/ | MySQL | Relacional | | Imagen: 228131\_i966  Imagen del logo  Link https://es.wikipedia.org/wiki/PostgreSQL | PostgreSQL | Relacional y orientado a objetos | | Imagen: 228131\_i967  Imagen del logo  Link https://www.infobip.com/es/partnership/partners/oracle | *ORACLE* | Relacional | | Imagen: 228131\_i968  Imagen del logo  Link https://es.wikipedia.org/wiki/Microsoft\_Access | Microsoft Access | Relacional | | Imagen: 228131\_i969  Imagen del logo  Link https://mvpcluster.com/sql-server-2016-enlaces-de-descarga-de-sql-server-management-studio-y-sql-server-data-tools/ | Microsoft SQL Server | Relacional | | Imagen: 228131\_i970  Imagen del logo  Link https://aws.amazon.com/es/rds/mariadb/ | MariaBD | Relacional | | Imagen: 228131\_i971  Imagen del logo  Link https://www.genbeta.com/desarrollo/mongodb-que-es-como-funciona-y-cuando-podemos-usarlo-o-no | MongoDB | Orientado a documentos | | Imagen: 228131\_i972  Imagen del logo  Link https://ubunlog.com/rdm-una-herramienta-de-administracion-de-escritorio-redis/ | Redis | Almacenamiento clave - valor | | |

## Componentes de los SGBD

| **Tipo de recurso** | Acordeón tipo 1 |
| --- | --- |
| **Introducción** | Los siguientes son algunos de los componentes que hacen parte de los sistemas de gestión de bases de datos: |
| **Componentes de un SGBD**  **Imagen Propia elaborada con SmartArt**  Imagen: 228131\_i973 | |
| **Diccionario de datos**  Es un conjunto de esquemas que reflejan las características de los datos incluidos en la base de datos (nombre, tipo, tamaño y relaciones entre datos). En este repositorio se incluye, además, la información de los permisos de uso que tiene cada usuario que utiliza el sistema. | |
| **Lenguajes de datos**  Son los lenguajes disponibles para los diversos tipos de manipulación de datos:   | Marca de verificación | Definición (DDL) que permite crear, eliminar y modificar estructuras. | | --- | --- | | Marca de verificación | Control (DCL) que permite especificar permisos sobre los objetos y crear o eliminar usuarios. | | Marca de verificación | Manipulación (DML) que permite la inserción, actualización, eliminación y consulta de datos. |   Estos tres lenguajes se engloban en el lenguaje SQL (Structured Query Language o Lenguaje de consulta estructurada). | |
| **Herramientas de administración**  Son las herramientas que permiten gestionar tareas como: seguridad, control de recuperación, gestión de diccionarios de datos, importación y exportación de datos, programación de aplicaciones. | |
| **Objetos**  Son los elementos que permiten la definición y manipulación de datos: tablas, consultas, funciones, disparadores. | |
| **Planificadores**  Herramientas que permiten programar y automatizar la ejecución de tareas como puede ser cálculo de resultados o copias de seguridad. | |

## Funciones del SGBD

| **Tipo de recurso** | Acordeón tipo 2 |
| --- | --- |
| **Introducción** | Las características de un SGDB posibilitan el cumplimiento de unas funciones que pueden agruparse de la siguiente manera: |
| **Figura 2**  *Funciones del SGBD*    Imagen: 228131\_i974  Características SGBD  Imagen de muestra que ilustra el tema  Link: <https://homeworkdatabase.wordpress.com/2015/06/27/sgbd-o-smbd-sistema-manejador-de-base-de-datos/> | |
| **Almacenamiento de los datos**  La base de datos debe almacenar textos, documentos, contraseñas y demás datos digitales de forma segura para que puedan consultarse. | |
| **Manipulación de los datos**  La base de datos debe permitir la edición de los datos almacenados según los permisos de los usuarios. Se podrán realizar operaciones de adición, supresión, actualización, extracción, consulta, entre otras. | |
| **Seguridad e integridad de los datos**  La base de datos debe garantizar la seguridad de la información evitando el acceso a personas no autorizadas. Además de cifrar la información, deberá tener un esquema de permisos de diferentes niveles para acceder a ésta. También se deben garantizar la coherencia e integridad de la información almacenada por medio del cumplimiento de las cuatro propiedades de las transacciones (ACID). | |
| **Recuperación y restauración de los datos**  La base de datos debe brindar herramientas para el respaldo y recuperación de la información en caso de posibles fallos. | |

# Modelo relacional de datos

| **Tipo de recurso** | Slider Presentación | |
| --- | --- | --- |
| **Introducción** | Un modelo de datos relacional es una recopilación de elementos de datos con relaciones predefinidas. En este tipo de modelo los elementos se organizan en tablas con columnas y filas. Cada columna de la tabla almacena un determinado tipo de dato (atributo o característica) y cada campo almacena el valor de dicho atributo para cada registro (fila o tupla). | |
| Las bases de datos relacionales utilizan las restricciones para garantizar la integridad de los datos y las transacciones que allí se realizan, se manejan como secuencias de operaciones que forman una sola unidad lógica de trabajo. Esto quiere decir que, la transacción se debe ejecutar completamente y no de forma parcial para evitar así inconsistencias en la base de datos. | | Imagen: 228131\_i975  Ejemplo de BdeD Relacional  Imagen de muestra tomada de Wikipedia  Fuente: https://es.wikipedia.org/wiki/Base\_de\_datos\_relacional |
| El lenguaje utilizado para comunicarse con las bases de datos relacionales es el lenguaje SQL o lenguaje de consulta estructurada. Este lenguaje es utilizado para agregar, actualizar o eliminar los datos de la base de datos, así como para recuperar subconjuntos de datos como consultas. | | Imagen: 228131\_i976  SQL Logo  Imagen de muestra  Fuente: https://blog.desafiolatam.com/12-lenguajes-de-programacion-mas-populares/sql-logo/ |
| Si se puede incrustar aquí tarjetas animadas, se ponen los términos en las tarjetas:  Los siguientes son algunos términos y definiciones relacionados con este tipo de base de datos: | | Imagen: 228131\_i977  Bases de datos en SQL  Imagen de muestra  Fuente: https://www.freepik.es/vector-gratis/dibujado-mano-ilustracion-sql-diseno-plano\_21901976.htm#query=sql&position=5&from\_view=keyword |
| **Tabla**  Arreglo bidimensional de datos. Está conformada por un encabezado con el nombre de las columnas y el cuerpo que contiene las filas de datos. Cada fila de la tabla representa un elemento independiente con los valores de sus atributos. | | Imagen: 228131\_i978  Tabla  Imagen de ejemplo  Link: [**http://codejavu.blogspot.com/2019/08/uso-del-data-table-en-jsf.html**](http://codejavu.blogspot.com/2019/08/uso-del-data-table-en-jsf.html) |
| **Tipo de dato**  Define el conjunto de valores que pueden tomar los atributos (columnas) de una tabla. Cada columna de la tabla se asocia con un tipo de dato.  Los tipos de datos más comunes en SQL son:   | Información | char (texto de longitud fija). | | --- | --- | | Lista | varchar (texto de longitud variable). | | Dólar | float (datos numéricos con precisión flotante). | | Calendario diarioReloj | date/time (fecha y hora). | | Matemáticas | integer (enteros). | | Cara sonriente con relleno sólidoCara sonriente sin relleno | boolean (booleanos). | | | Imagen: 228131\_i979  Tipos de datos SQL  Imagen de muestra – reconstruida a partir de: https://www.plsql.biz/2006/10/tipos-de-datos-en-plsql.html |
| **Relación**  Conexión entre las filas de dos tablas. Se muestran mediante los valores de una columna en una tabla que coinciden con los valores de una columna en otra tabla. | | Imagen: 228131\_i980  Relación  Imagen de muestra – reconstruida a partir de: |

## Operadores de álgebra relacional

| **Tipo de recurso** | Slider Presentación | | |
| --- | --- | --- | --- |
| **Introducción** | El álgebra relacional es un lenguaje que define una serie de operaciones que se realizan por medio de “operadores” aplicados sobre uno o varios conjuntos de datos (tablas), produciendo como resultado un nuevo conjunto de datos (una nueva tabla). Las operaciones de álgebra relacional son similares a las del álgebra numérica, pero en este caso se aplican a tablas y no a números. | | |
| El álgebra relacional agrupa los operadores en tres categorías: | | | Imagen: 228131\_i981 |
| **Operadores de restricción y proyección**  Estos operadores generan subconjuntos de una tabla y son ampliamente usados, pues muchas veces los usuarios desean observar subconjuntos de una tabla que cumplan con ciertos criterios o condiciones.  La restricción genera un subconjunto de renglones o filas del conjunto de datos, mientras que la proyección genera un subconjunto de columnas. La restricción usa una condición o expresión lógica para seleccionar las filas a mostrar. La proyección usa una lista de nombres de columnas para indicar las columnas a mostrar. | | | Imagen: 228131\_i982  **Operadores de restricción y proyección**  **Imagen de muestra: https://docplayer.es/13940008-Bases-de-datos-tema-4-lenguajes-relacionales-francisco-ruiz-dic-2000-uclm-esi-f-ruiz.html** |
| **Operador de producto cruz extendido**  Este operador puede combinar dos tablas. La tabla resultante muestra todas las posibles combinaciones de los datos de las dos tablas generando, en muchos casos, datos excesivos. | | | Imagen: 228131\_i983  **Operador de producto cruz extendido**  **Imagen de muestra: https://docplayer.es/13940008-Bases-de-datos-tema-4-lenguajes-relacionales-francisco-ruiz-dic-2000-uclm-esi-f-ruiz.html** |
| **Operador de unión o enlace (join)**  Es el operador más usado para combinar tablas, pues, a diferencia del operador de producto cruz extendido el operador, join requiere de una condición de coincidencia sobre las filas de dos tablas. | | | Imagen: 228131\_i984  **Operadores de enlace**  **Imagen de muestra: https://docplayer.es/13940008-Bases-de-datos-tema-4-lenguajes-relacionales-francisco-ruiz-dic-2000-uclm-esi-f-ruiz.html** |
| En la siguiente tabla, se puede apreciar la equivalencia entre nomenclatura de álgebra relacional y SQL: | | Imagen: 228131\_i985 | |

## Sintaxis SQL

| **Tipo de recurso** | Slider Presentación | |
| --- | --- | --- |
| **Introducción** | SQL es el lenguaje estándar de consultas a las bases de datos. Estas consultas permiten almacenar, manipular y recuperar datos de BD relacionales. Al ser SQL un lenguaje estandarizado permite desplegar implementaciones de este mismo lenguaje, en diferentes sistemas con cambios mínimos. (Orbegozo, 2019)  Las sentencias SQL se pueden agrupar por funcionalidades: | |
| **Lenguaje de manipulación de datos (DML):**   * SELECT. * INSERT. * UPDATE. * DELETE. | | Imagen: 228131\_i986  **DML**  **Imagen de muestra**  Fuente: https://es.dreamstime.com/stock-de-ilustraci%C3%B3n-dml-image92510885 |
| **Lenguaje de definición de datos (DDL):**  CREATE, ALTER, DROP que se aplican a los componentes principales de una base de datos:   * Base de datos (DATABASE). * Tablas (TABLE). * Vistas (VIEW). * Índices (INDEX). * Procedimientos almacenados (PROCEDURE). * Disparadores (TRIGGER). | | Imagen: 228131\_i987  **DDL**  **Imagen de muestra**  **Fuente: https://www.shutterstock.com/es/image-vector/ddl-mean-data-definition-language-computer-1842168220** |
| **Lenguaje de control de datos (DCL):**   * COMMIT. * ROLLBACK. * GRANT. * REVOKE. | | Imagen: 228131\_i988  **DCL**  **Imagen de muestra**  **Reconstruida a partir de: https://www.shutterstock.com/es/image-vector/ddl-mean-data-definition-language-computer-1842168220** |
| Las sentencias en SQL no son sensibles a mayúsculas y minúsculas, pero se recomienda usar estos comandos en mayúsculas sostenidas para diferenciarlos de los nombres de tablas y columnas.  Ejemplo: CREATE DATATABASE nombrebasedatos; | | Imagen: 228131\_i989  **Sentencias SQL**  **Imagen de muestra y escribir curso, nombre, nota en MINÚSCULAS**  **Fuente: https://desarrolloweb.com/articulos/tipos-de-sentencias-sql.html** |
| **Consultas y operaciones básicas con SQL**   * Traer TODAS las columnas de una tabla:   SELECT \* FROM tabla;   * Traer dos o más columnas de una tabla (los nombres de las atributos o columnas se separa con comas):   SELECT atributo1, atributo2 FROM tabla;   * Traer dos columnas de la base de datos y ordenar por una de ellas de forma descendente (Por defecto el comando *GROUP BY* ordena de forma ascendente):   SELECT atributo1, atributo2 FROM tabla ORDER BY atributo2 DESC;   * ALIAS (AS): Permiten renombrar una columna o una tabla:   SELECT atributo1 AS nuevonombre FROM tabla;   * Filtros: Se pueden usar filtros en las consultas para traer solo los registros que cumplen con estos:   SELECT nombre FROM estudiante WHERE edad > 10;    SELECT nombre FROM estudiante WHERE edad > 10 AND ciudad = “Cali”; | | **Consultas con SQL**  Imagen: 228131\_i990  Imagen de muestra  Fuente: https://www.campusmvp.es/recursos/post/Fundamentos-de-SQL-Como-realizar-consultas-simples-con-SELECT.aspx |
| **Funciones de agregación**   * AVG (promedio de los valores)   SELECT AVG (salario) FROM nomina   * COUNT (contar elementos)   SELECT COUNT (salario) FROM nomina   * MAX (valor máximo)   SELECT MAX (salario) FROM nomina   * MIN (valor mínimo)   SELECT MIN (salario) FROM nomina   * SUM (sumar valores)   SELECT SUM (salario) FROM nomina | | **Funciones de agregación**  Imagen: 228131\_i991  Imagen de muestra  Fuente: https://www.campusmvp.es/recursos/post/Fundamentos-de-SQL-Agrupaciones-y-funciones-de-agregacion.aspx |

| **Tipo de recurso** | | Infografía estática |
| --- | --- | --- |
| **Texto introductorio** | | La siguiente infografía resume los principales comandos del lenguaje de consulta SQL con ejemplos de su forma de uso. |
| **Imagen**    Imagen: 228131\_i992  Imagen de muestra de la Infografía.  Se puede utilizar cualquier imagen de Bases de Datos SQL | | |
| Fundamentos de modelamiento de datos  Comandos Básicos SQL  SELECT \* FROM cliente WHERE nombre = "Jake" | Imagen: 228131\_i993  **Comandos SQL**  **Se puede el texto sencillo como está en la infografía de muestra o se puede una imagen como ésta, con comandos SQL**  **Link: https://www.facebook.com/metricaandina/photos/a.111166203601712/485921316126197/?type=3** | |
| **CREATE**  Permite crear una nueva tabla o una nueva base de datos.  CREATE DATABASE [IF NOT EXISTS] database\_name | Imagen: 228131\_i994  **Imagen de muestra – Base de datos - Imagen de muestra** | |
| **SELECT – FROM**  Permite hacer consultas de una tabla específica recuperando las columnas requeridas.  SELECT \* FROM students  SELECT id, name FROM students | Imagen: 228131\_i995  **Base de datos relacional - Imagen de muestra** | |
| **WHERE**  Permite aplicar filtros a las consultas para recuperar solo los registros que cumplen la condición.  SELECT name FROM student WHERE id = 001 | Imagen: 228131\_i996  **Filtros - Imagen de muestra** | |
| **Funciones de agregación**   * avg(expr): promedio de los valores * count(expr): cuenta los valores * max(expr): valor máximo * min(expr): valor mínimo * sum(expr): suma de los valores | Imagen: 228131\_i997  **Funciones**  **Imagen de muestra** | |
| **ORDER BY**  Permite ordenar los resultados de una consulta de forma ascendente (por defecto) o descendente.  SELECT name, id FROM student ORDER BY age | Imagen: 228131\_i998  **Ordenar**  **Imagen de muestra** | |
| **Orden lógico de las operaciones**   1. FROM, JOIN 2. WHERE 3. GROUP BY 4. Funciones de agregación 5. HAVING 6. SELECT 7. DISTINT 8. UNION / INTERSECT / EXCEPT 9. ORDER BY 10. OFFSET 11. LIMIT / FETCH / TOP | Imagen: 228131\_i999  **Link:** fuente imágenes, elementos CANVAS | |

## 

## Procedimientos almacenados y disparadores

| Tipo de recurso | Cajón de texto de color |
| --- | --- |
| Los procedimientos almacenados son programas recolectados en la base de datos y su implementación depende del sistema de gestión de bases de datos que se emplee.  Este tipo de procedimientos se ejecuta directamente en el SGBD y por tanto tiene acceso directo a los datos. Esto mejora el tiempo de respuesta.  Ejemplo:  CREATE PROCEDURE nombre([parámetro1, parámetro2,....])  [Atributos]  BEGIN Instrucciones  END | |
| Los disparadores son un tipo de procedimiento almacenado que se invoca o ejecuta automáticamente en respuesta a una modificación concreta en la base de datos. Este tipo de procedimientos están compuestos de tres partes:   | **Evento:** | modificación en la BD que activa el procedimiento. | | --- | --- | | **Condición:** | consulta o verificación que se ejecuta al activar el disparador. | | **Acción:** | comandos que se ejecutan al activar el procedimiento y verificar que la condición se cumple. |   Ejemplo:  CREATE TRIGGER contar AFTER INSERT ON estudiantes  WHEN (edad > 10)  FOR EACH ROW  BEGIN  cuenta = cuenta + 1;  END | |

| Cuadro de texto |
| --- |
| En el siguiente video, podrá aprender más sobre la creación de Bases de datos con SQL y la realización de consultas:  **Video: 228131\_Comandos SQL** |

# Bases de datos NO relacionales

| Tipo de recurso | Cajón de texto de color |
| --- | --- |
| Existen tipos de datos no estructurados. Como se indica, no tienen una estructura o arquitectura identificable y por esto no se ajustan a un modelo de datos predefinido. Por esta falta de organización, este tipo de datos no es apto para una base de datos relacional como las gestionadas con SQL. Un gran porcentaje de los datos que generan las organizaciones son datos no estructurados, lo que hace más compleja su gestión y explotación. | |
| Algunas fuentes de datos no estructuradas son:   | Paleta | Imágenes | | --- | --- | | Cámara de vídeo | Videos | | Presentación con gráfico de barras RTL | Informes | | Periódico | Documentos de texto (Word, PDF) | | Marketing | Audios | | Micrófono de radiocomunicación | Notas de voz | | Red | Comentarios en redes sociales | | Correo electrónico | Correos electrónicos | | |
| Para poder aprovechar todos estos datos surgen las bases de datos no relacionales (NoSQL) ofreciendo nuevas formas de almacenamiento y de gestión de esta información. | |

## ¿Qué son las bases de datos NO relacionales?

| Tipo de recurso | Cajón de texto de color |
| --- | --- |
| Joyanes (2015), define las bases de datos NoSQL (Not only SQL) como “…una categoría de sistemas de gestión de bases de datos que no utilizan SQL como lenguaje de consulta principal. Estas bases de datos no requieren esquemas de tablas fijas y no soportan operaciones Join. Están optimizadas para operaciones de lectura/escritura escalables en lugar de pura consistencia”. | |
| Las bases de datos NoSQL son más flexibles que las bases de datos relacionales ya que permiten gestionar información en otros formatos como grafos, documentos, clave-valor y columnas. Este tipo de bases de datos son idóneas para aplicaciones que requieren manipulación de grandes cantidades de datos y por esta razón es usada por empresas como Facebook, LinkedIn o Google.  Imagen: 228131\_i1000  Imagen de muestra  Fuente: https://www.grapheverywhere.com/bases-de-datos-nosql-marcas-tipos-ventajas/ | |
| Las principales características de estas bases de datos son:   | Internet | Capacidad de almacenamiento de grandes cantidades de datos. | | --- | --- | | Internet | Escalabilidad sin afectación del rendimiento. | | Internet | Acceso rápido. | | Internet | Distribución y manipulación de datos no estructurados. | | |

## 

## Tipos de bases de datos NO relacionales

| **Tipo de recurso** | | Pestañas o tabs Verticales |
| --- | --- | --- |
| **Introducción** | | Existen cuatro categorías de bases de datos NoSQL: almacenes clave-valor, almacenes de documentos, almacenes de grafos y familia de bases de datos columnares. |
| Imagen: 228131\_i1001  **Tipos de BdeD No Relacionales**  **Imagen de muestra** [**https://docs.microsoft.com/es-es/dotnet/architecture/cloud-native/relational-vs-nosql-data**](https://docs.microsoft.com/es-es/dotnet/architecture/cloud-native/relational-vs-nosql-data)  **Link: https://aprendeinformaticas.com/modelo-entidad-relacion/** | | |
| **Almacén clave - valor** | Es la forma más sencilla de las bases de datos NoSQL. Los datos se representan como colecciones de pares clave-valor. Este tipo de base de datos es más rápido y óptimo para los accesos. En este caso la clave (*key*) sirve como identificador único.  Las bases de datos NoSQL más populares de este tipo son*:* Amazon DynamoDB y Cassandra. | |
| **Almacenes de documentos** | Este tipo de base de datos guarda la información como un listado de documentos desestructurados. Aquí se combinan formatos de marcación estándar como XML (Lenguaje de marcado extensible o Extensible Markup Language) y JSON (Notación de objetos en JavaScript o JavaScript Object Notation) con formatos propietarios como PDF y Word de Microsoft.  La base de datos de este tipo más utilizada es MongoDB. | |
| **Almacenes de grafos** | Este tipo de base de datos organiza la información en estructuras de grafos dirigidos. En estas estructuras los objetos se conocen como nodos y aristas (*edges*). El principal beneficio de estas bases de datos es la velocidad y se usan en aplicaciones de redes sociales como Twitter.  Algunas bases de datos de grafos son: Neo4J y AllegroGraph. | |
| **Bases de datos columnares** | Consisten en un sistema de almacenamiento distribuido para manipulación de datos estructurados, diseñado para grandes volúmenes de datos. Son ampliamente utilizadas para aplicaciones de *Big data*.  Algunas bases de datos de este tipo son: Apache HBase, Hypertable y Cassandra. | |

| **Tipo de recurso** | Infografía estática |
| --- | --- |
| **Texto introductorio** | La siguiente infografía permite ver los motores de bases de datos NoSQL y su clasificación: |
| Búsqueda | EDteam  Imagen: 228131\_i1002  Imagen de muestra – Reproducir una imagen similar que agrupe todos los motores de BdeD.  Link: https://edteam-media.s3.amazonaws.com/community/original/2d14aa20-8e75-4844-822c-b5cd8b715ada.jpg | |
| **Código de la imagen** | Construya aquí el código de la imagen |
| **MOTORES DE BASES DE DATOS NoSQL:** Están optimizados para operaciones de lectura/escritura rápidas y escalables en lugar de consistencia. | Imagen: 228131\_i1003  **Bases de datos NoSQL**  **Imagen de muestra**  **Tomada de: https://www.diegocalvo.es/caracteristicas-y-comparativa-de-las-bases-de-datos-nosql/** |
| **MongoDB**:  Motor orientado a documentos.  Los datos son almacenados en BSON, una representación binaria de JSON.  (Tipo almacén de documentos) | Imagen: 228131\_i1004 |
| **Amazon DynamoDB**:  Servicio “*server-less*”.  Provee escalabilidad automática y permite transacciones seguras.  (Tipo clave-valor) | Imagen: 228131\_i1005 |
| **Cassandra**:  Motor multiplataforma de código abierto desarrollado por APACHE.  Permite grandes volúmenes de datos en forma distribuida.  (Tipo clave-valor - Columnar) | Imagen: 228131\_i1006 |
| **Neo4J**:  Carga los datos en grafos permitiendo ejecutar automáticamente consultas para marcar patrones específicos de eventos en la red.  (Tipo grafos) | Imagen: 228131\_i1007 |
| **AllegroGraph**:  Está diseñada para proteger los datos más sensibles dentro del entorno flexible de una base de datos de grafos / documentos  (Tipo grafos – Almacén de documentos) | Imagen: 228131\_i1008 |
| **Apache HBase**:  Forma parte del proyecto Hadoop de APACHE.  Es un almacén de *Big data* distribuido y escalable de forma masiva.  (Columnar) | Imagen: 228131\_i1009 |
| ***Redis***:  Motor de base de datos en memoria.  Incorpora un conjunto de estructuras de datos en memoria que le permiten crear con facilidad diversas aplicaciones personalizadas. | Imagen: 228131\_i1010 |

Recuerde explorar los demás recursos que se encuentran disponibles en este componente formativo; para ello, diríjase al menú principal, donde encontrará la síntesis, una actividad didáctica para reforzar los conceptos estudiados, material complementario, entre otros.

**Síntesis**

| **Tipo de recurso** | Síntesis |
| --- | --- |
|  | |
| **Introducción** | El siguiente mapa integra los criterios y especificidades de los conocimientos expuestos en el presente componente formativo. |
| Imagen: 228131\_i1011  *Nota:* Adaptada de *Arquitectura de Datos.* S/f (2014-2015). <https://cmapscloud.ihmc.us:443/rid=1YDGC7H8R-14H70Q6-J6GP25> | |

**Actividad didáctica**

**Anexo 2: Actividad Didáctica**

**Material complementario**

| Tipo de recurso | Material complementario | | |
| --- | --- | --- | --- |
| Tema | Referencia APA del material | Tipo | Enlace |
| Introducción a la arquitectura de datos | GitMind. (s.f.). *Mejor programa para hacer mapa mental y conceptual. GitMind.* | Sitio Web | <https://gitmind.com/es/> |
| Introducción a la arquitectura de datos | ISACA. (2022). *COBIT Control objectives for Information Technologies.* ISACA Resources. | Sitio Web | <https://www.isaca.org/resources/cobit> |
| Introducción a la arquitectura de datos | The Data Governance Institute. (2022). D*ata Governance Framework & Components.* | Sitio Web | <https://datagovernance.com/data-governance-framework-components/> |
| Introducción a la arquitectura de datos | The Open Group. (2022). *TOGAF*. The Open Group. | Sitio Web | <https://www.opengroup.org/togaf> |
| Introducción a la arquitectura de datos | Visual Paradigm. (s.f.). *Design database with professional ERD Software*. | Sitio Web | <https://www.visual-paradigm.com/features/database-design-with-erd-tools/> |

**Glosario**

| **Tipo de recurso** | Glosario |
| --- | --- |
| Datos: | un dato es una representación simbólica de un atributo o característica de un elemento o evento. |
| Datos estructurados: | se consideran datos estructurados cuando tienen una estructura bien definida, siguen un orden y son de fácil acceso y uso para los sistemas de información y las personas. |
| Datos no estructurados: | este tipo de datos no tiene un modelo o estructura muy definida y puede tener diferentes formatos como texto, imágenes, sonido, video lo que hace más difícil su procesamiento y análisis. |
| Información: | la información es un dato o conjunto de datos con un contexto que les da significado. |
| JSON: | notación de objetos en JavaScript o JavaScript Object Notation en inglés. |
| Modelado de datos: | es un proceso que permite analizar y definir los datos generados por la organización y la relación entre estos con el fin de identificar sus características y dependencias. |
| NoSQL: | not only SQL son una categoría de sistemas de gestión de bases de datos que no utilizan SQL como lenguaje de consulta principal. |
| Sistema de información: | es un conjunto de componentes interrelacionados que trabajan juntos para recopilar, procesar, almacenar y difundir información para apoyar la toma de decisiones. |
| SQL: | es el lenguaje estándar de consultas a las bases de datos. |
| XML: | lenguaje de marcado extensible o Extensible Markup Language en inglés. |

**Referencias bibliográficas**

| **Tipo de recurso** | Bibliografía |
| --- | --- |
| Bisson, A. (2021). *SQL - Fundamentos del lenguaje con ejercicios corregidos* (3ra edición). <https://www-eni-training-com.bdigital.sena.edu.co/portal/client/mediabook/home> | |
| DAMA Internacional. (2022). *Imágenes de rueda DMBOK v2*. <https://www.dama.org/cpages/dmbok-2-wheel-images> | |
| Data Governance Institute. (s.f.). *Marco de Gobierno de Datos de la DGI*. <https://datagovernance.com/the-dgi-data-governance-framework/> | |
| Fernández, A. (2020). *Ciencia de datos para la ciberseguridad*. 1RAMA. <https://www.alphaeditorialcloud.com/reader/ciencia-de-datos-para-la-ciberseguridad-1628020600?location=eyJjaGFwdGVySHJlZiI6IngwMiIsImNmaSI6Ii80W3gwMl0vMi8yW19pZENvbnRhaW5lcjAwMl0vMiJ9> | |
| García, J. (2018). *Ciencia de datos Técnicas analíticas y aprendizaje estadístico en un enfoque práctico*. Editorial Altaria. <https://www-alphaeditorialcloud-com.bdigital.sena.edu.co/reader/ciencia-de-datos?location=1> | |
| Jiménez, F. (2018). *Principios de COBIT 5*. https://geniusitt.com/blog/que-es-cobit-5/ | |
| Jones, H. (2019). *Ciencia de los datos: la guía definitiva sobre análisis de datos, minería de datos, almacenamiento de datos, visualización de datos.* Publicación independiente. <https://biblioteca.sena.edu.co/F/QK1UI7RDS4Q5XCEJHTEG681M74XTARHJQABMVRV1F1YV8U8ELM-13705?func=full-set-set&set_number=003407&set_entry=000001&format=999> | |
| Joyanes, L. (2015). *Sistemas de información en la empresa.* Alfaomega. | |
| Mannino, M. V. (2007). *Administración de bases de datos*. McGraw-Hill Irwin. | |
| Orbegozo Arana, B. (2019). *Gestión de bases de datos con SQL, MySQL y Access: curso práctico*. Alfaomega. | |
| Piattini, M. y Ruiz, F. (2020). *Gobierno y gestión de las tecnologías y los sistemas de* *información*. Ra-Ma. <https://www-alphaeditorialcloud-com.bdigital.sena.edu.co/reader/gobierno-y-gestion-de-las-tecnologias-y-los-sistemas-de-informacion-1591724977> | |
| Ramakrishnan, R. (2011). *Sistemas de gestión de bases de datos.* McGraw-Hill España. | |