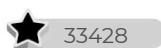


WUOLAH



TEAM_GETPPID__
www.wuolah.com/student/TEAM_GETPPID__



TEMA 6(SIF).pdf

Resúmenes temario SIF 2018



2º Sistemas de Información



Grado en Ingeniería Informática



**Escuela Politécnica Superior de Córdoba
UCO - Universidad de Córdoba**

 **escuela
de negocios**
CÁMARA DE SEVILLA

MÁSTER EN DIRECCIÓN Y GESTIÓN DE RECURSOS HUMANOS

www.mastersevilla.com

Inscríbete



BECAS

TEMA 6 Sistemas de soporte para la toma de decisiones (DSS) y Business Intelligence (BI)

SISTEMAS DE INFORMACIÓN

ISMAEL ABU-JADUR GARCÍA

CONGRESO INTERNACIONAL

DESAFÍOS A LA SEGURIDAD GLOBAL

INTELIGENCIA, TERRORISMO Y
AMENAZAS HÍBRIDAS

17 - 18 JUNIO

LUGAR

Centro Superior de Estudios de la
Defensa Nacional CESEDEN

Pº de la Castellana, 61 Madrid

MÁS INFORMACIÓN

eventos@iniseg.es

912 141 926

www.iniseg.es

Organiza:



INISEG

Instituto Internacional de
Estudios en Seguridad Global



INSTITUCIONES COLABORADORAS



OTROS COLABORADORES



Sigue las novedades del congreso

#CIS_INISEG

Contenido

TEMA 6: Sistemas de soporte para la toma de decisiones (DSS) y Business Intelligence (BI)	2
Tipos de decisiones	2
Roles gerenciales.....	2
Calidad de la información	3
Filtros gerenciales	3
Inercia y política organizacional.....	3
Fundamentos de inteligencia de negocios: Business Intelligence	3
Definición Garner	4
Business Intelligence (BI). Definición Datawarehouse Institute	4
Modelos de datos dimensionales	5
Modelo entidad-relación	5
Modelos de arquitectura	6
Big Data	7
¿De dónde provienen esos datos?.....	7
Tipos de datos de Big Data.....	7
Componentes básicos de Business Intelligence.....	8

TEMA 6: Sistemas de soporte para la toma de decisiones (DSS) y Business Intelligence (BI)

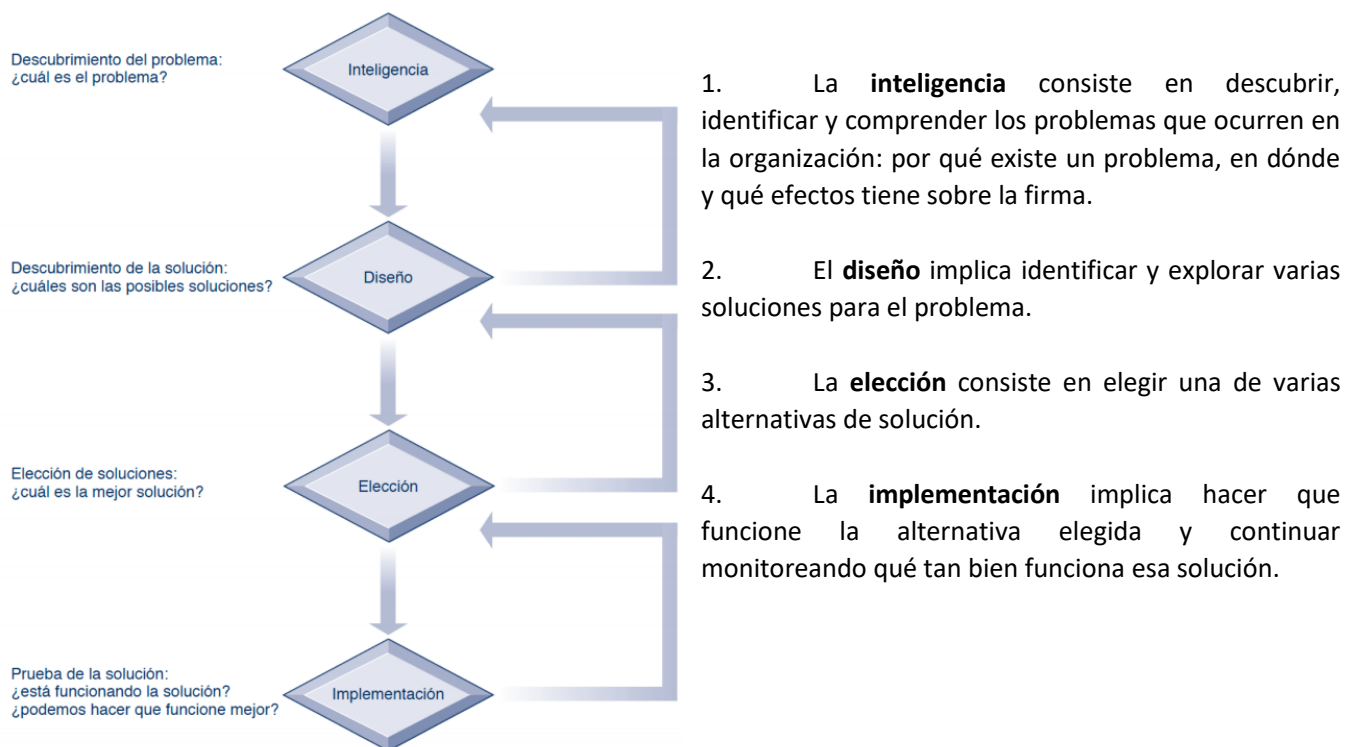
Tipos de decisiones

Teniendo en cuenta que hay distintos niveles en una organización. Cada uno tiene distintos requerimientos de información para el soporte de decisiones y responsabilidad para distintos tipos de decisiones:

- **Estructuradas:** Son repetitivas y rutinarias; además se requiere un procedimiento definido para manejarlas, y no considerarlas nuevas.
- **Semiestructuradas:** Donde sólo una parte del problema tiene una respuesta clara proporcionada por un procedimiento aceptado.
- **No estructuradas:** Aquellas en las que el encargado de tomarlas debe proveer un juicio, una evaluación y una perspectiva para resolver el problema

Según el nivel de gerencia:

- Los **ejecutivos de nivel superior** se enfrentan a muchas situaciones de decisiones no estructuradas.
- La **gerencia de nivel medio** se enfrenta a escenarios de decisiones más estructuradas, pero sus decisiones pueden incluir componentes no estructurados.
- Los **gerentes operacionales y los empleados ordinarios** tienden a tomar decisiones más estructuradas.



El proceso de toma de decisiones se divide en cuatro etapas.

En caso de que una solución dada no funcionase, siempre se podrá volver a una etapa anterior.

Roles gerenciales

Para poder entender mejor las funciones y los roles gerenciales, debemos examinar los modelos clásico y contemporáneo del comportamiento gerencial:

- **Modelo clásico:** Se describen por primera vez las cinco funciones clásicas de los gerentes como planificación, organización, coordinación, decisión y control.
- **Modelo contemporáneo:** establecen que el comportamiento actual de los gerentes parece ser menos sistemático, más informal, menos reflexivo, más reactivo y menos organizado que lo que el modelo clásico nos hubiera hecho creer.

Estos roles gerenciales se pueden clasificar en tres categorías:



- 1) Interpersonales
- 2) De información
- 3) Decisionales.

Ahora podemos ver que los sistemas de información no son útiles para todos los roles gerenciales. Y en esos roles gerenciales en donde los sistemas de información podrían mejorar las decisiones, las inversiones en tecnología de la información no siempre producen resultados positivos.

Existen tres razones principales: 1) Calidad de la información, 2) Filtros gerenciales y 3) Cultura organizacional.

Calidad de la información

Las decisiones de alta calidad requieren información de alta calidad. Si la salida de los sistemas de información no cumple con estos criterios de calidad, la toma de decisiones se verá afectada.

DIMENSIÓN DE CALIDAD	DESCRIPCIÓN
Precisión	¿Los datos representan la realidad?
Integridad	¿Son consistentes la estructura de los datos y las relaciones entre las entidades y atributos?
Consistencia	¿Están definidos los elementos de datos de una manera consistente?
Compleción	¿Están presentes todos los datos necesarios?
Validez	¿Los valores de datos se encuentran dentro de rangos definidos?
Puntualidad	¿Están disponibles los datos cuando se necesitan?
Accesibilidad	¿Es posible acceder a los datos, comprenderlos y utilizarlos?

Filtros gerenciales

Los gerentes (al igual que todos los seres humanos) absorben información a través de una serie de filtros para comprender el mundo que los rodea. Tienen atención selectiva, se enfocan en ciertos tipos de problemas y soluciones.

Inercia y política organizacional

Las decisiones que toma una firma a menudo representan un balance de los diversos grupos de interés de la firma, en vez de ser la mejor solución para el problema.

Los estudios de reestructuración de negocios han encontrado que las firmas tienden a ignorar el mal desempeño hasta que se ven amenazadas por empresas externas que tienen interés en adquirirlas.

Fundamentos de inteligencia de negocios: Business Intelligence

El objetivo básico de la Business Intelligence es apoyar de forma sostenible y continuada a las organizaciones para mejorar su competitividad, facilitando la información necesaria para la toma de decisiones. Un BI pretende convertir datos en información y a partir de la información ser capaces de descubrir conocimiento.

*“BI es un **proceso interactivo** para **explorar** y **analizar información estructurada** sobre un **área** (normalmente almacenada en un **datawarehouse**), para descubrir tendencias o patrones, a partir de los cuales derivar ideas y extraer conclusiones.*

*El proceso de Business Intelligence incluye la **comunicación** de los descubrimientos y **efectuar** los cambios.*

*Las **áreas** incluyen clientes, proveedores, productos, servicios y competidores.”*

- **Proceso Interactivo:** Análisis de información continuado en el tiempo, no sólo en un momento puntual.
- **Explorar:** Hay un momento inicial en el que por primera vez accedemos a información que nos facilita su interpretación. En esta primera fase, lo que hacemos es “explorar” para comprender qué sucede en nuestro negocio.
- **Analizar:** Descubrir relaciones entre variables, tendencias, es decir, cuál puede ser la evolución de la variable, o patrones.
- **Información estructurada y datawarehouse:** La información que utilizamos en BI está almacenada en tablas relacionadas entre ellas. Estas tablas están almacenadas en lo que conocemos como datawarehouse o almacén de datos.
- **Área de análisis:** Todo proyecto de BI debe tener un objeto de análisis concreto. Nos podemos centrar en los clientes, los productos, los resultados de una localización, etc.
- **Comunicar los resultados y efectuar los cambios:** Una vez descubierto algo, sea comunicado a aquellas personas que tengan que realizar los cambios pertinentes en la organización para mejorar nuestra competitividad.

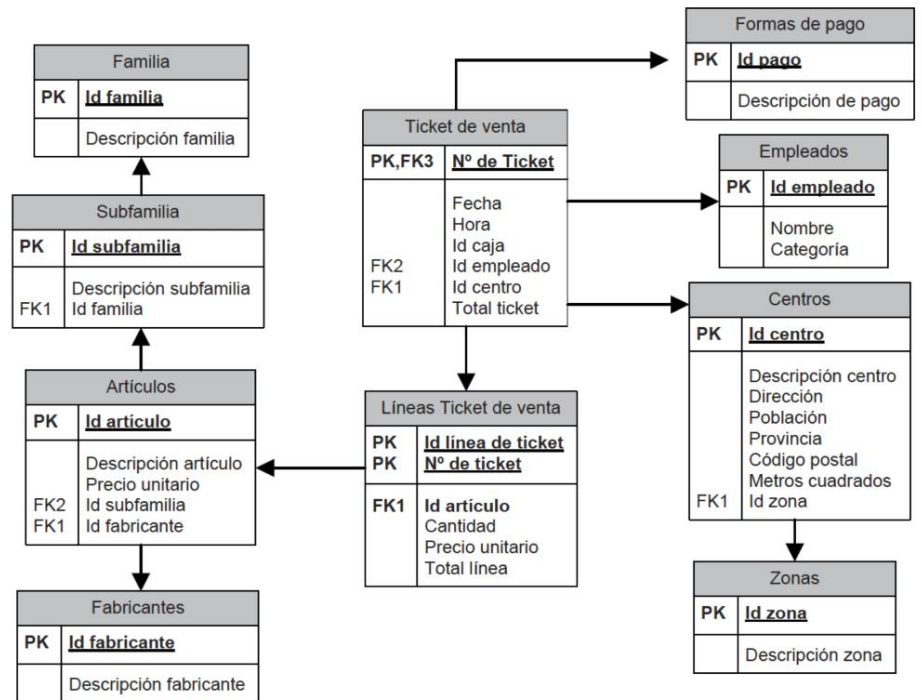
Business Intelligence (BI). Definición Datawarehouse Institute

“Business Intelligence (BI) es un término paraguas que abarca los procesos, las herramientas, y las tecnologías para convertir datos en información, información en conocimiento y planes para conducir de forma eficaz las actividades de los negocios. BI abarca las tecnologías de datawarehousing los procesos en el ‘back end’¹⁰, consultas, informes, análisis y las herramientas para mostrar información (estas son las herramientas de BI) y los procesos en el ‘front end’.”

Modelos de datos dimensionales

Modelo entidad-relación

Un ejemplo de modelo entidad-relación sería este:



Cada tabla se identifica por una clave y nos permite conocer los datos de un determinado elemento.

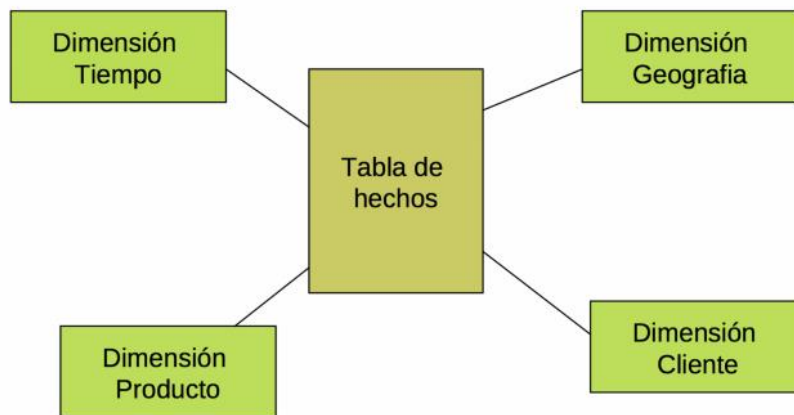
Comparación entre modelos dimensionales.

Operacionales	BI (Dimensional)
Enfocado a la actualización: Enfocado a eliminar redundancia, coordinar actualizaciones y repetir el mismo tipo de operaciones muchas veces en el día	Enfoque a la consulta
Altamente normalizadas para soportar actualizaciones consistentes y mantenimiento de la integridad referencial	Altamente desnormalizada ya que se requiere disminución de tiempos en la obtención de grandes cantidades de datos
Tiempos de respuesta en segundos o inferior	Tiempos de respuesta aceptables pueden ser segundos, minutos, horas
Almacenan pocos datos derivados	Gran cantidad de datos derivados (redundancia)
Pocos datos agregados	Agregación: Varios niveles de datos precalculados

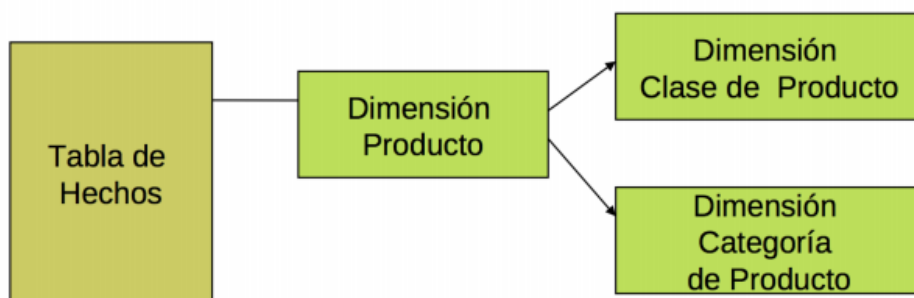
Modelos de arquitectura

Dos arquitecturas de acuerdo con la normalización de sus dimensiones:

- **Estrella:** Nos permite analizar la información de manera que podamos responder a preguntas. Consiste básicamente en utilizar estructuras de datos no normalizadas que incorporan redundancias que permiten agilizar las consultas y operaciones de análisis de datos. Para su construcción distinguimos:
 - o Las tablas de hechos (aquello que queremos medir o analizar)
 - o Las tablas de dimensiones (cómo lo queremos medir)
- Algunas características:
 - o Desnormalizado
 - o Habilidad para análisis dimensional



- **Copo de Nieve**
 - o Variación del modelo estrella
 - o Normalizado
 - o Rompe el análisis dimensional. Se usa cuando no se puede implementar un modelo estrella





Big Data

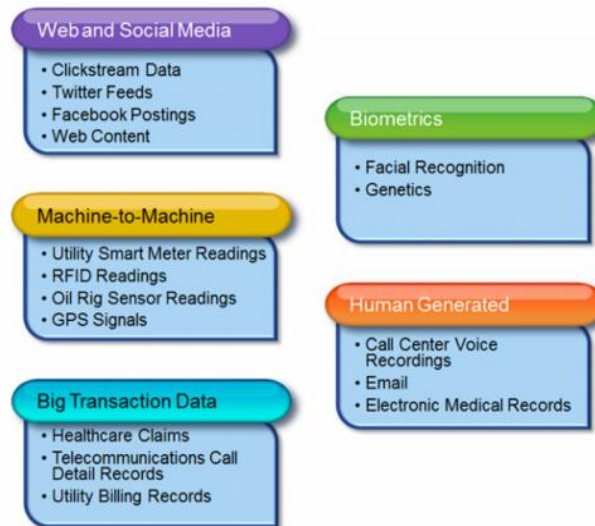
Nuevo enfoque de entendimiento y toma de decisiones, la cual es utilizada para describir enormes cantidades de datos (estructurados, no estructurados y semi-estructurados). Se aplica para toda aquella información que no puede ser procesada o analizada utilizando procesos o herramientas tradicionales.

¿De dónde provienen esos datos?

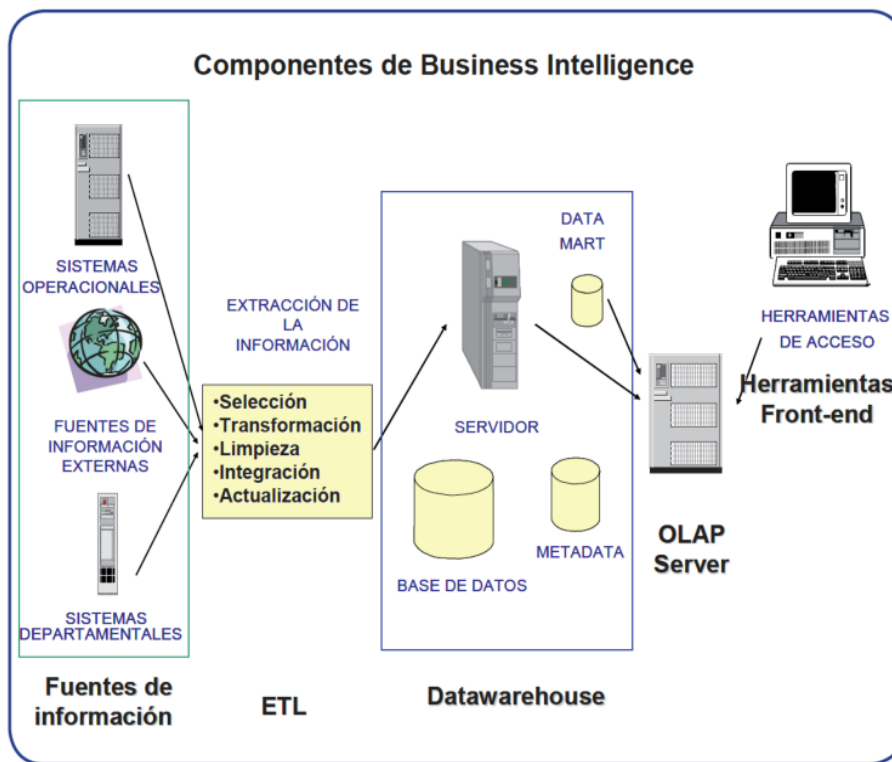
En muchos países se administran enormes bases de datos que contienen datos de: censo de población, registros médicos, impuestos, etc., y si a todo esto le añadimos transacciones financieras realizadas en línea o por dispositivos móviles, análisis de redes sociales (en Twitter son cerca de 12 Terabytes de tweets creados diariamente y Facebook almacena alrededor de 100 Petabytes de fotos y videos), ubicación geográfica mediante coordenadas GPS.

Tipos de datos de Big Data

Big Data Types



Componentes básicos de Business Intelligence



- **Fuentes de información:** De las cuales partiremos para alimentar de información el datawarehouse.

- **Proceso ETL de extracción:** Transformación y carga de los datos en el datawarehouse.

- El propio **datawarehouse** o almacén de datos.

- **El motor OLAP:** Que nos debe proveer capacidad de cálculo, consultas, funciones de planeamiento, pronóstico y análisis de escenarios en grandes volúmenes de datos.

- **Las herramientas de visualización:** Que nos permitirán el análisis y la navegación a través de estos.

Fuentes de información

Las fuentes de información a las que podemos acceder son:

- Básicamente, de los sistemas operacionales o transaccionales
- Sistemas de información departamentales
- Fuentes de información externa

Existen muchos factores que contribuyen a la complejidad de cargar la información en un datawarehouse:

- **Número de fuentes de información distintas de las que cargamos la información.** Acceder a distintas bases de datos requiere distintas habilidades y el conocimiento de distintas sintaxis de SQL. Si el número de bases de datos a las que debemos acceder es elevado, puede provocar que tanto las definiciones como las codificaciones en los distintos entornos sean diferentes.
- **Consistencia:** La definición de los distintos componentes de nuestro sistema de información no siempre es consistente a través de distintas aplicaciones, que no están integradas.
- **Información no estructurada:** cada vez más la tecnología, nos permite trabajar con información no estructurada: correos, informes, redes sociales.
- **Calidad de los datos:** La calidad de los datos en un datawarehouse es fundamental. Si en el datawarehouse hay errores, éstos se propagarán a lo largo de toda la organización y son muy difíciles de localizar.

Proceso de extracción, transformación y carga ETL

El proceso de ETL consume entre el 60% y el 80% del tiempo de un proyecto de Business Intelligence, por lo que es un proceso clave en la vida de todo proyecto.

El proceso ETL se divide en 5 subprocesos:

1. **Extracción:** Este proceso recupera los datos físicamente de las distintas fuentes de información utilizando herramientas especializadas. En este momento disponemos de los datos en bruto.
2. **Limpieza:** Este proceso recupera los datos en bruto y comprueba su calidad, elimina los duplicados y, cuando es posible, corrige los valores erróneos y completa los valores vacíos
3. **Transformación:** Este proceso recupera los datos limpios y de alta calidad y los estructura y suma en los distintos modelos de análisis.
4. **Integración:** Este proceso valida que los datos que cargamos en el datawarehouse son consistentes con las definiciones y formatos del datawarehouse ; los integra en los distintos modelos de las distintas áreas de negocio que hemos definido en el mismo.
5. **Actualización:** Este proceso es el que nos permite añadir los nuevos datos al datawarehouse.

Datawarehouse (Almacén de datos)

Normalmente la información que necesitamos proviene de distintos sistemas, pero nosotros la requerimos en un mismo entorno para facilitar su análisis.

Características

- **Orientado a un área:** Significa que cada parte del datawarehouse está construida para resolver un problema de negocio.
- **Integrado:** La información debe ser transformada en medidas comunes, códigos comunes y formatos comunes para que pueda ser útil.
- **Indexado en el tiempo:** significa que se mantiene la información histórica y se almacena referida a determinadas unidades de tiempo.
- **No volátil:** significa que los usuarios no la mantienen, como lo harían en los entornos transaccionales.

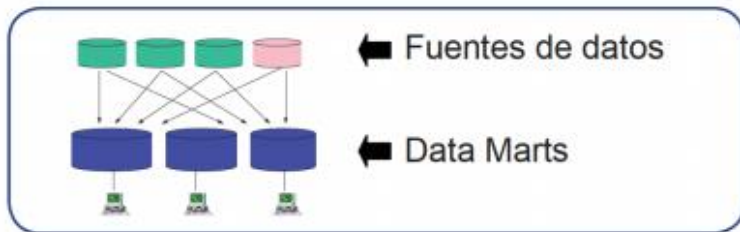
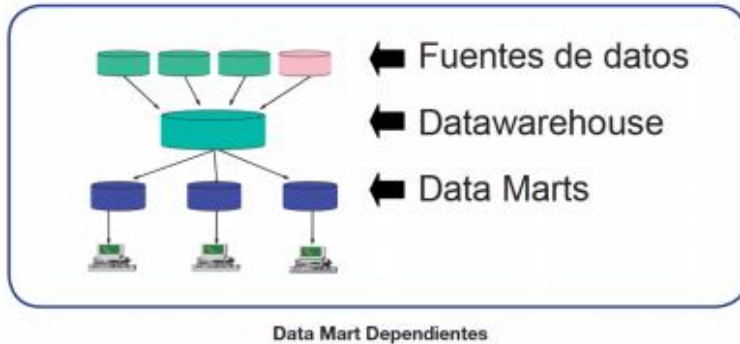
Herramientas Datawarehouse

Los datawarehouses se representan habitualmente como una gran base de datos, pero pueden estar distribuidos en distintas bases de datos.

El trabajo de construir un datawarehouse corporativo puede generar inflexibilidades, o ser costoso y requerir plazos de tiempo que las organizaciones no están dispuestos a aceptar. En parte, estas razones originaron la aparición de los Data Mart.

Data Mart

Los Data Mart pueden ser independientes o dependientes. Los primeros son alimentados directamente de los orígenes de información, mientras que los segundos se alimentan desde el *datawarehouse* corporativo. Los Data Mart independientes pueden perpetuar el problema de los “silos de información” y en su evolución pueden llegar a generar inconsistencias con otros Data Mart.



Un componente crítico de un datawarehouse es el Metadata. El Metadata es el repositorio central de información de la información. Nos da el significado de cada uno de los componentes y sus atributos que residen en el datawarehouse (o Data Mart)

OLAP

Los usuarios necesitan analizar información a distintos niveles de agregación y sobre múltiples dimensiones. Los usuarios pueden hacer este análisis al máximo nivel de agregación o al máximo nivel de detalle. OLAP provee de estas funcionalidades y algunas más, con la flexibilidad necesaria para descubrir las relaciones y las tendencias que otras herramientas menos flexibles no pueden aportar.

A estos tipos de análisis les llamamos multidimensionales, porque nos facilitan el análisis de un hecho desde distintas perspectivas o dimensiones. La representación gráfica del OLAP son los cubos.

Las herramientas OLAP nos permiten “rotar” (en inglés “slicing”) los cubos, es decir, cambiar el orden de las distintas dimensiones. También podemos seleccionar (en inglés “dicing”) sólo algunas de las celdas.

Aplicaciones de business intelligence

Muchas empresas utilizan herramientas de Business Intelligence de renombre como Microsoft, Oracle, IBM y SAS. También, aplicaciones de BI especializadas como Qlik y Tableau Software. Pero existen herramientas de código abierto más económicas y que ofrecen grandes soluciones.

Un ejemplo de BI Open Source es Pentaho.



Las principales herramientas de Business Intelligence son:

- **Generadores de informes:** Utilizadas por desarrolladores profesionales para crear informes estándar para grupos, departamentos o la organización.
- **Herramientas de usuario final de consultas e informes:** Empleadas por usuarios finales para crear informes para ellos mismos o para otros; no requieren programación.
- **Herramientas OLAP:** Permiten a los usuarios finales tratar la información de forma multidimensional para explorarla desde distintas perspectivas y periodos de tiempo.
- **Herramientas de Dashboard y Scorecard:** Permiten a los usuarios finales ver información crítica para el rendimiento con un simple vistazo utilizando iconos gráficos y con la posibilidad de ver más detalle para analizar información detallada e informes, si lo desean.
- **Herramientas de planificación, modelización y consolidación:** Permite a los analistas y a los usuarios finales crear planes de negocio y simulaciones con la información de Business Intelligence. Pueden ser para elaborar la planificación, los presupuestos, las previsiones. Estas herramientas proveen a los dashboards y los scorecards con los objetivos y los umbrales de las métricas.
- **Herramientas data mining:** Permiten a estadísticos o analistas de negocio crear modelos estadísticos de las actividades de los negocios. Data mining es el proceso para descubrir e interpretar patrones desconocidos en la información mediante los cuales resolver problemas de negocio.
- **Computación cognitiva:** es una nueva tecnología que se está utilizando para que las máquinas prácticamente puedan desarrollar procesos parecidos a los del pensamiento humano.

Minería de datos

La minería de datos o exploración de datos es un campo de la estadística y las ciencias de la computación referido al proceso que intenta descubrir patrones en grandes volúmenes de conjuntos de datos.

Utiliza los métodos de la inteligencia artificial, aprendizaje automático, estadística y sistemas de bases de datos

Proceso de minería de datos

- Selección del conjunto de datos
- Análisis de las propiedades de los datos
- Transformación del conjunto de datos de entrada
- Seleccionar y aplicar la técnica de minería de datos
- Extracción de conocimiento
- Interpretación y evaluación de datos

Técnicas de minería de datos

- **Redes neuronales:** Son un paradigma de aprendizaje y procesamiento automático inspirado en la forma en que funciona el sistema nervioso de los animales
- **Regresión lineal:** Es la más utilizada para formar relaciones entre datos. Rápida y eficaz pero insuficiente en espacios multidimensionales.
- **Árboles de decisión:** Modelo de predicción utilizado en el ámbito de la inteligencia artificial y el análisis predictivo
- **Agrupamiento o Clustering:** Procedimiento de agrupación de una serie de vectores según criterios habitualmente de distancia

- **Reglas de asociación:** - Se utilizan para descubrir hechos que ocurren en común dentro de un determinado conjunto de datos.

Aplicaciones de minería de datos

- RapidMiner
- WEKA

Computación cognitiva. Las claves: percibir, comprender y actuar

Computación cognitiva

permiten a los sistemas de información y aplicaciones percibir el mundo y recopilar datos, analizar y comprender la información recopilada, tomar decisiones.

Las herramientas de computación cognitiva son capaces de aprender de su experiencia y variar su comportamiento según lo aprendido.

Un ejemplo sería un diagnóstico médico.

IBM Watson

IBM Watson es una tecnología pionera de computación cognitiva.

Lee y entiende el lenguaje natural, con su riqueza de matices y giros lingüísticos.

Es capaz de responder a preguntas complejas en pocos segundos a partir de su avanzada capacidad para analizar grandes cantidades de datos. Ante un interrogante, formula hipótesis y escoge la respuesta en la que tiene mayor nivel de confianza, presentando su razonamiento.

Aprende de cada experiencia así que cada vez es más inteligente. También es entrenado por los profesionales que lo utilizan a medida que es "alimentado" con nueva información.