

Programa Ejecutivo de Business Intelligence & Big Data

Introducción al Business Intelligence

Ing. Arturo Rojas Medrano







"Algo peor que no tener información disponible es tener mucha información y no saber que hacer con ella "



- En los 70's se da un gran salto al aparecer los dispositivos de Acceso Directo (DASD)
- En los 90s las grandes compañías contaban con Centros de Información.





- ➤ Entre 1992 y 1993 surgió el concepto de Data Warehouse, pieza esencial del BI.
- ➤ En 1996 el concepto de Business Intelligence empezó a difundirse como una evolución de los EIS (Executive Information Systems).



- El término Business Intelligence comenzó a abarcar más herramientas como:
 - ✓ EIS (sistema de informaciones ejecutivas), soluciones
 - ✓ DSS (sistema de soporte a las decisiones),
 - ✓ Planillas Electrónicas,
 - ✓ Generadores de Consultas y de Informes,
 - ✓ Data Marts, Data Mining, Herramientas OLAP



➤ La historia del BI también está profundamente ligada a los ERP (Enterprise Resource Planning -Planeación de Recursos Empresariales), sigla que representa los sistemas integrados de gestión empresarial cuya función es facilitar el aspecto operacional de las empresas.



- Un ERP utiliza la información a través de la organización de forma más proactiva en áreas claves.
- ➤ El reto es unir elementos de las áreas y proporcionarle a los usuarios una manera universal de utilizar la información almacenada en diferentes sistemas.

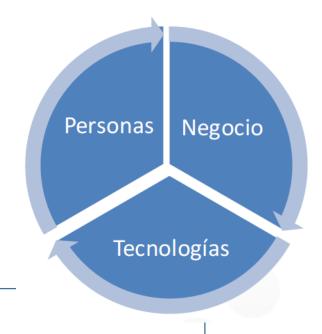


- Un ERP concentra todas las actividades de su negocio en una base de datos **común**, con lo cuál todas los usuarios de la organización trabajan sobre la misma información.
- Esto se le conoce como "One single number concept".



¿ QUE QUIERE DECIR INTELIGENCIA DE NEGOCIOS ?

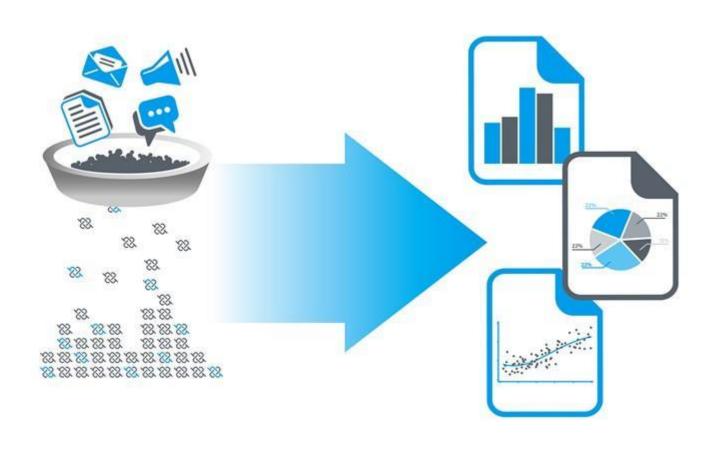
"Usar datos de ayer y de hoy para tomar decisiones para mañana "



Transforma los datos de la organización en información y distribuirla a lo largo de la cadena de valor.



¿ QUE ES BUSINESS INTELLIGENCE ?





¿ QUE ES BUSINESS INTELLIGENCE ?

"Business Intelligence es una herramienta crítica para el éxito y sobrevivencia de su organización hoy en día, no se trata solo de herramientas y tecnología, sino también de organización"

Gartner Group



TECNOLOGIAS DEL BUSINESS INTELLIGENCE

									Apps	Form factor	PSO	
Delivery	Desktop gadgets		Office	Office suites Mo		ile	Di	sconnected			>	
	Portals I		Interac	Interactive voice response,			e, ATM, point-of-sale				Strategy	
Reporting Dashboard		ds	Alerts			ata visualization		icati	Appliance	Stre		
	Search Geospatial Reporting — ad hoc, analytical, production						appl	hppli		logy		
Performance management	Metrics/KPIs		Planning		ig S		Scorecards		ical	4	S	Methodology
on seed to the territory of the territory	Strategy/objectives management							ver			Meth	
Supporting applications	Collaboration Life-cycle mgt. Localization QA Version control						Industry vertical applications		urcing			
	Metadata — integration, repositories											
Analytics	Data/text r	mining	nining Guided decisions NLP Guided				uided search		MSP/application outsourcing			
	Time series	Operatio	tional DSS Predictive analytics				BI SaaS					
	Usage analytics Statistical analysis Web analytics						_	BIS	icati			
Discovery		Accelerators/query optimization Adapters/tool kits						, ER		appl		
and integration	BAM/CEP BPM/BRE int							SCM		WSP/		
•									RM,		97.	
	DQ — cleansing, profiling EAI/SOA EII ETL/CDC								Enterprise applications: ERP, CRM, SCM, ERM	SP)	Center of excellence	Governance
	Integration — third-party applications											
	Operational data stores (ODS), data warehouses (DW), data marts (DM)											
Data	Report mining			Services registry and repository				ise applic	Hosted BI (ASP)	Center of		
	Columnar DBMS Hie			erarchical/XML In-memory DBMS								
	Multidimensional OLAP Multivalue DBMS RDBMS						Enterpr	Hos				
	Streaming DBMS Search DBMS								ВРО			
Infrastructure	Network			Servers Storage			torage			8		



BENEFICIOS

- ✓ Crear un circulo virtuoso de la información.
- ✓ Permite una visión única conformada, histórica, persistente y de calidad de toda la información.
- ✓ Crear, manejar y mantener métricas, indicadores claves de rendimiento (KPI key performance indicador) e indicadores claves de metas (KGI key goal indicador) fundamentales de para la empresa
- ✓ Aporta información actualizada tanto a nivel agregado como en detalle.
- ✓ Mejora la competitividad de la organización: diferencia lo relevante de lo superfluo, acceder más rápido a la información, mejorar la toma de decisiones.



¿ QUIENES NECESITAN BI?

- √ ¿Pasa más tiempo recolectando y preparando información que analizándola?
- √ ¿En ocasiones le frustra el no poder encontrar información que usted está seguro que existe dentro de la empresa?
- √ ¿Quisiera tener una guía sobre las cosas que han sucedido cuando los administradores anteriores implementaban determinada estrategia?
- √ ¿No sabe qué hacer con tanta información que tiene disponible en la empresa?



¿ QUIENES NECESITAN BI?

- √ ¿Quiere saber qué productos fueron los más rentables durante un periodo determinado?
- √ ¿No sabe cuáles son los patrones de compra de sus clientes dependiendo de las zonas?
- √ ¿Ha perdido oportunidades de negocio por recibir información retrasada?
- √ ¿Trabaja horas extras el fin de mes para procesar documentos o reportes?
- √ ¿No sabe con certeza si su gente está alcanzando los objetivos planeados?



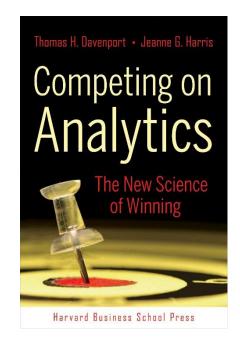
BUSINESS INTELLIGENCE, BUSCA RESPONDER

- ¿Qué pasó?
- ¿qué pasará ahora?
- ¿Por qué pasó?
- ¿Qué pasará?
- ¿Qué quiero que pase?

Pasado

Presente

Futuro



"Una nueva forma de estrategia está emergiendo basada en el uso de la estadística descriptiva, modelos productivos y complejas técnicas de optimización, datos de alta calidad y toma de decisiones basadas en hechos.



ESTRATEGIA DE BUSINESS INTELLIGENCE

Desplegar un proyecto BI en una organización no es sencillo, se requiere una estrategia.

Se requiere coordinar de forma efectiva las tecnologías, el uso, los procesos de madurez.





¿ COMO DETECTAR QUE NO EXISTE UNA ESTRATEGIA ?

- Los usuarios identifican a TI como el origen de los problemas.
- ➤ La dirección considera que la Inteligencia de negocios es otro centro de coste.
- ➤ TI sigue preguntando a los usuarios finales sobre necesidad de los informes.
- No es posible medir el retorno de la inversión (ROI) de un proyecto BI.
- No hay un plan para desarrollar, contratar, retener y aumentar el equipo de BI



¿ COMO DETECTAR QUE NO EXISTE UNA ESTRATEGIA ?

- No existe un responsable funcional (o el asignado no es el adecuado).
- Existen múltiples soluciones en la organización distribuidas en diferentes departamentos que repiten funcionalidad.
- Los usuarios creen que la información del Data Warehouse no es correcta.



PRIMEROS PASOS

- Crear un centro de competencias de BI (aunar conocimientos en tecnologías, metodologías, etc).
- > Establecer estándares de BI en la organización.
- ➤ Identificar que procesos de negocios necesitan diferentes aplicaciones analíticas.
- Desarrollar un Framework de métricas a nivel empresarial como pilar de una gestión del rendimiento a nivel corporativo.
- Incluir los resultados de aplicaciones analíticas en los procesos de negocios para añadir valor a todo tipo de decisiones



PRIMEROS PASOS

- Alinear el departamento TI y negocio en caso de no poder organizar un centro de competencias.
- Evangelizar a la organización.
- Poner atención a las necesidades que requieren BI.

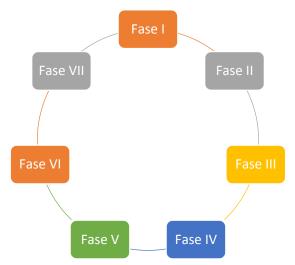


TENER RESPUESTAS A ESTAS PREGUNTAS

- Que problemas o necesidades de negocio se buscan resolver mediante la estrategia BI.
- Cual es el impacto por la falta de dicha información.
- Qué fuentes de datos son necesarias para responder las preguntas (marketing, finanzas, operaciones, etc).
- ➤ En que medida las diferentes entidades de información (cliente, producto, etc) están alineadas.
- Cual es el grado de calidad de los datos.
- Que cantidad de datos actual e histórica debe ser guardada y con qué frecuencia hay cambios



➢ BIMM permite clasificar nuestra organización desde el punto de vista del grado de madurez de la implementación de sistemas Business Intelligence.





- > FASE 1 : NO existe BI
 - Los datos se encuentran en los sistemas OLTP.
 - Las decisiones se basan en la intuición, experiencia pero no en datos consistentes, carecen de herramientas.



OLTP

Características



- Orientado al registro de datos.
- Modelos de datos estructurados.
- Normalización.
- RBDMS.
- Consultas rápidos.
- Alta concurrencia.

Terminología habitual	Modelo relacional					
Tabla	Relación					
Fila	Tupla					
Columna	Atributo					
Número de filas	Cardinalidad					
Valores posibles	Dominio					



- > FASE 2 : NO existe BI, pero los datos son accesibles.
 - No existe proceso formal para acceder a los datos, pero algunos usuarios tienen accesos a información de calidad y justifican decisiones con dicha información.
 - Se utiliza Excel o algún tipo de reporting para este proceso.
 - Se intuye que deben existir soluciones para mejorar este proceso.



- ➤ **FASE 3** : Aparición de procesos formales de toma de decisiones basadas en datos .
 - Se establece un equipo que controla los datos y que permite hacer informes contra los mismos que permiten tomar decisiones fundamentales.
 - Los datos son extraídos directamente desde los OLTP sin data cleansing.
 - No existe un data warehouse.



- > FASE 4 : Data Warehouse.
 - El impacto negativo de consultar directamente los OLPTP lleva a la conclusión de la necesidad de un repositorio de datos DW.



- FASE 5 : Data Warehouse crece y el reporting se formaliza.
 - El DWH funciona y se desea que todos se beneficien del mismo.
 - Se fomaliza el reporting corportativo.
 - Se habla de OLAP, pero solo algunos identifican su beneficio.



- > FASE 6 : Despliegue de OLAP.
 - Ni el acceso al DW ni el reporting son satisfactorias para responder a preguntas sofisticadas.
 - OLAP se despliega para dichos perfiles.
 - Las decisiones empiezan a impactar de forma significativa en los procesos de negocio de toda la organización.



- > FASE 7 : Business Intelligence se formaliza.
 - Aparecen la necesidad de implantar otros procesos de BI como Data Mining, Balanced ScoreCard, etc y procesos de calidad de datos que impacten en procesos CRM, SCM.



Elementos de una solución de Business Intelligence

ETL Data Warehouse **OLAP** Reporting Data Mart Calidad y Gobierno de datos





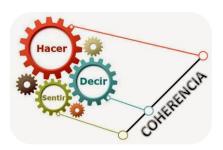
Es un repositorio de datos que proporcionan una visión global, común e integrada de los datos de la organización independientemente de cómo se vayan utilizar posteriormente por los usuarios



Propiedades:



Estable



Coherente



Información histórica



Fiable





Alto volumen de datos



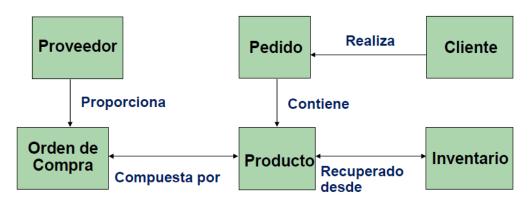




Orientado a un tema



Los Usuarios piensan en términos de "cosas" y sus "relaciones", no en términos de procesos, funciones o aplicaciones





Integrado



Incluyen datos de múltiples orígenes y presenta consistencia de datos.

- Convención de nombres.
- Descripciones.
- Atributos físicos de los datos.
- Valores de los datos consistentes.



Variable en el tiempo



Se realizan fotos de los datos basadas en fechas o hechos



Variable en el tiempo

Entorno Operacional

- Datos con valores actuales
- Horizonte de 30 90 días
- Exactitud en los accesos



Id de cliente nombre dirección teléfono ratio de crédito

Data Warehouse

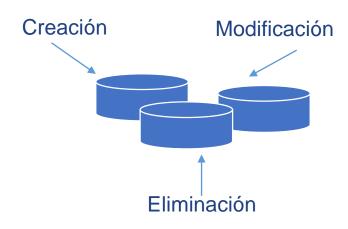
- Datos en "fotos"
- Horizonte de 5 10 años
- Refleja la perspectiva desde un momento en el tiempo

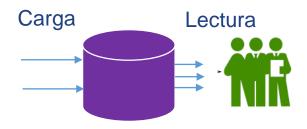


Id de cliente
fecha desde
fecha hasta
nombre
dirección
teléfono
ratio de crédito



No volátil





Sistemas OLTP (Dinámico) Sistemas OLAP (Más estático)



DATA WAREHOSING, BUSINESS INTELLIGENCE Y MODELADO DIMENSIONAL

- Un Data Warehouse debe considerar las necesidades de negocio.
- Luego trabajamos hacia atrás a través de los díseños lógicos y físicos, junto con las decisiones de tecnología y hardware.





DATA WAREHOSING, BUSINESS INTELLIGENCE Y MODELADO DIMENSIONAL



- La información principal activo de una organización.
- Se utiliza para mantenimiento de registro operativo y toma de decisiones.
- Los sistemas operacionales están optimizados para procesar transacciones rápidamente.



DATA WAREHOSING , BUSINESS INTELLIGENCE Y MODELADO DIMENSIONAL



- Los usuarios DW/BI miran las ruedas de la organización para evaluar rendimiento, comparan, cuestionan.
- Los sistemas DW/BI están optimizados para consultas de alto rendimiento.
- Preservan contexto histórico para evaluaciones.
- Los DW/BI no deben ser copias de los OLTP.



OBJETIVOS DE UN DATA WAREHOSING / BUSINESS INTELLIGENCE

- Hacer accesible la información.
- Presentar información creíble.
- Un DWH/BI debe adaptarse al cambio.



- Los DW/BI no deben ser copias de los OLTP.
- Presentar información de manera oportuna.
- Debe proteger los activos de información.
- Controlar accesos a la información confidencial.



- Comprender a los usuarios comerciales:
 - Sus objetivos, responsabiliades.
 - Que decisiones toman.
 - Identifique a los mejores usuarios.
 - Encuentre nuevos usuarios potenciales.
 - Concéntrese en las capacidades del sistema DWH/BI



- Entregar información y análisis de alta calidad accesibles a los usuarios:
 - Elija datos mas robustos de todas las fuentes.
 - Hacer interfaces de usuario simples, basadas en plantillas.
 - Asegurar que los datos sean precisos y monitoréelos.
 - Adaptarse a los usuarios, requisitos y prioridades comerciales cambiantes.



- Sostenga el entorno DWH/BI:
 - Tome parte del crédito de las decisiones comerciales por usar el DWH para justificar gastos, etc.
 - Actualice el almacen DWH / BI.
 - Mantenga la confianza de los usuarios comerciales.
 - Mantenga feliz a los usuarios, sponsor y a la administración TI.



• ¿Qué pasaría si omitimos una responsabilidad de la lista?.

 ¿Contraste esta vista con su entorno de trabajo?.



- Un Data Warehouse tiene un patrón de diseño diferente a un OLTP.
- La información debe estar desnormalizada para optimizar las consultas.



Aceptado por que cumple dos requisitos:

- Entrega datos que sean comprensibles para los usuarios comerciales.
- Ofrece un rendimiento de consulta rápido

Consideraciones:

- Técnica para simplificar las bases de datos.
- La simplicidad es fundamental para que los usuarios entiendan los datos



"Vendemos productos en varios mercados y medimos nuestro desempeño a lo largo del tiempo":

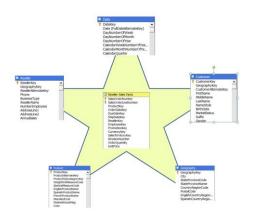
- Producto mercado tiempo.
- Un modelo de datos que empiece simple tiene la posibilidad de permanecer simple hasta el final del diseño.
- La simplicidad es fundamental para que los usuarios entiendan los datos.

 "Haz que todo sea lo mas simple posible, pero no mas simple" Albert Einstein.

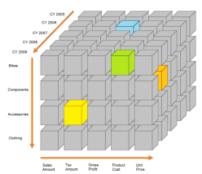


- Los modelos Dimensionales son bastantes diferentes de los modelos de tercera forma normal (3FN) que buscan eliminar redundancia de datos.
- 3FN divide la información en muchas entidades.
- La diferencia con un Modelo dimensiona es el grado de normalización.
- Los modelos normalizados son muy complicados para las consultas BI.
- Un modelo dimensional contiene la misma información que un modelo normalizado, pero empaqueta los datos en un formato que ofrece comprensión a los usuarios.





 Esquema estrella: Modelos dimensionales implementados en SGBD relacionales



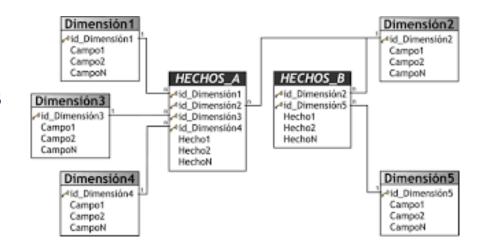
 Cubos de procesamiento analíticos OLAP: Modelos dimensionales implementados en entornos de base de datos multidimensionales.

Igual diseño lógico con dimensiones reconocibles, pero diferente implementación física



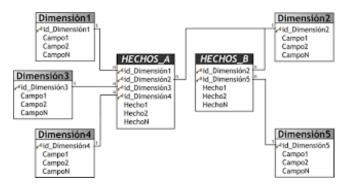
Tablas de Hecho:

 Almacenan las métricas resultantes de los procesos de negocios de una organización, por ejemplo una venta, inventario, planilla.



- Cada final de la tabla de la tabla de Hecho representa una medición.
- Estos datos se encuentran a un nivel de detalle denominado granularidad.





Dimensiones:

- Representan una perspectiva o una vista para cierto proceso de negocio, por ejemplo el cliente que ha comprado un producto, en que fecha realizó la compra, que producto, el vendedor.
- Describen "el quién, qué, donde, cuando, como y por que" asociado a un evento.
- El DWH es tan bueno como los atributos de las dimensiones

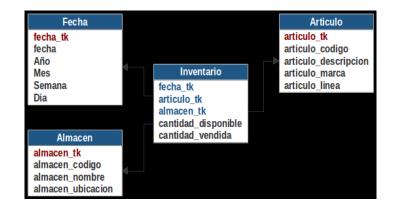


- El reducido número de tablas y el uso de descriptivas comerciales hacen mas facil navegar y menos probable que se encuentren errores.
- La simplicidad de un modelo dimensional tiene muchos beneficios de rendimiento.
- El DWH es tan bueno como los atributos de las dimensiones



Métrica:

- Son los indicadores de un proceso de negocio.
- Conceptos cuantificables que permiten medir nuestro proceso de negocio, ejemplo cantidad vendida





Esquema Estrella:

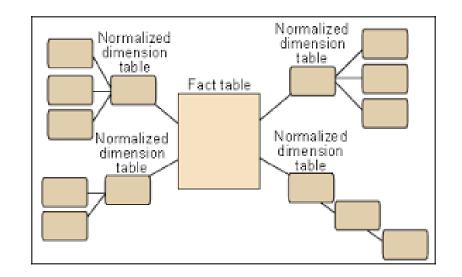
- Estructurar la información en procesos, vistas como una estrella.
- Se tiene una tabla de hecho en el centro y varias dimensiones por cada punto de vista de análisis que participa en la descripción del hecho.





Esquema en copo de Nieve:

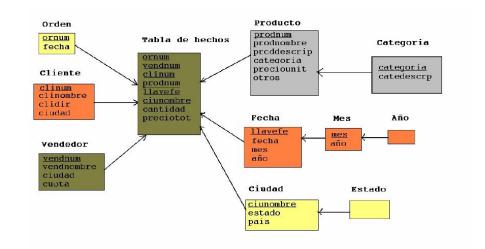
- Derivado del esquema estrella.
- Las tablas de dimensión se normalizan en múltiples tablas.
- Aparecen nuevas uniones





Esquema en copo de Nieve:

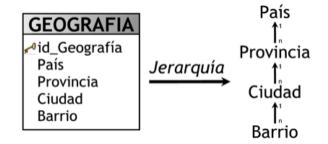
- Derivado del esquema estrella.
- Las tablas de dimensión se normalizan en múltiples tablas.
- Aparecen nuevas uniones





Esquema en copo de Nieve:

- Derivado del esquema estrella.
- Las tablas de dimensión se normalizan en múltiples tablas.
- Aparecen nuevas uniones







TIPOS DE METRICAS

Métrica:

- Valores que recogen el proceso de una actividad o resultado de la misma.
- Estas medidas proceden del resultado de la actividad de negocio.





TIPOS DE METRICAS

Métrica de realización de actividad (Leading):

 Ejemplo, la participación de una persona en un evento.





Métrica de resultado de actividad (Lagging):

 Ejemplo, la cantidad de puntos de un jugador en un partido.



INDICADORES CLAVES

- Valores correspondientes que se tiene que alcanzar (objetivos).
- Proporcionan información sobre el rendimiento de una actividad o sobra la consecuencia de una meta.



KPI, Indicadores de rendimiento, nos definen unos valores que nos explican en qué rango optimo de rendimiento nos deberíamos situar al alcanzar los objetivos, métricas de proceso.



KGI, Indicadores de metas, definen mediciones para informar a la dirección general si un proceso TIC ha alcanzado sus requisitos de negocio (criterios de información).



MITOS DEL MODELO DIMENSIONAL

- Los modelos dimensionales son solo para valores resumidos, agregados.
- Son departamentales no empresariales.
- No son escalables.
- No son inttegrados



PROCESO DE DISEÑO DIMENSIONAL

- Seleccione el proceso de Negocio (comprender las necesidades de negocio, así como las realidad de los datos fuentes), son actividades operacionales que realiza su organización, este proceso define un objetivo de diseño especifico y permite declarar la granularidad, las dimensiones y las fact.
- Declare la granularidad de la información, el grano establece lo que representa una sola fila de la tabla de hechos, se debe hacer antes de declarar las dimensiones, porque cada dimensión debe ser consistente con la granularidad.



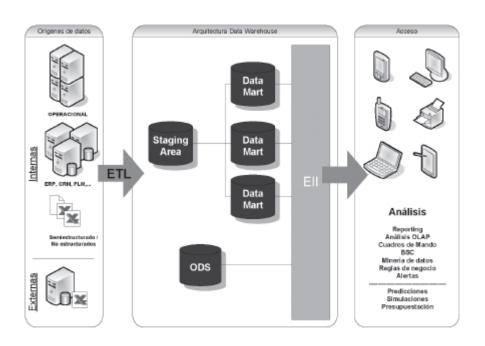
PROCESO DE DISEÑO DIMENSIONAL

- Identifique las dimensiones, proporciona el contexto de "Quién, qué, dónde, cuándo, por qué y cómo".
- Contienen los atributos descriptivos utilizados por los reportes para filtrar y agrupar hechos.
- Identifique las FACT, los hechos son las medidas que resultan de un proceso de negocio y casi siempre son numéricas.
- Una única fila tiene relación de uno a uno con un evento de medida como se describe en la granularidad de la tabla de hechos.



ARQUITECTURA DATA WAREHOUSE

 Entrerprise Bus Arquitecture (Data Warehouse Virtual / Federado), conocido como MD (Multidimensional Arquitecture).



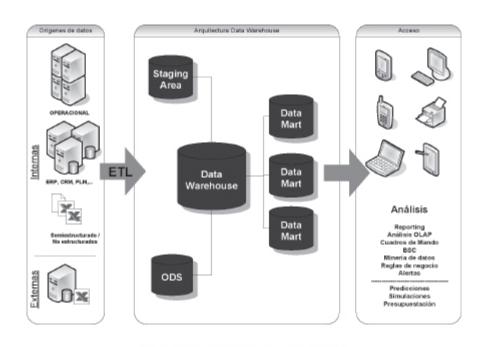
Federado: Hace uso de herramienta de Integración para realizar las consultas como si se tratará de un único Data Warehouse.

Fuente: Josep Curto, Enterprise Bus Architecture



ARQUITECTURA DATA WAREHOUSE

Corporate Information Factory (Entrerprise Data Warehouse).



Existe un Data Warehouse corporativos y unos Data Marts (Incluso cubos OLAP).

Fuente: Josep Curto. Enterprise Data Warehouse



FASES DE UN PROYECTO BI





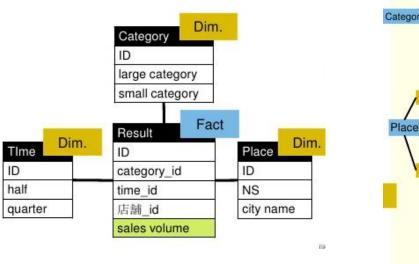
DISEÑO ANALISIS OLAP

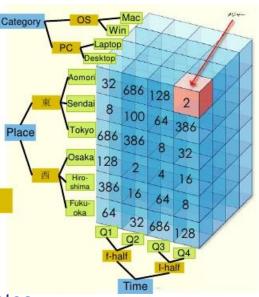


- Tecnología que permite organizar datos sobre objetos u organización multidimensional (tablas matriciales), cuyo objetivo es recuperar y manipular datos y combinaciones de los mismos a través de consultas.
- Responde preguntas como: ¿Por qué pasó?.
- Formada por un Motor y un visor para consultar, manipular, filtrar.



DISEÑO ANALISIS OLAP





- Orientado a la consulta masiva de datos.
- Las estructuras OLAP permiten realizar preguntas que serían sumamente complejas mediante SQL
- Permite comparar datos entre dimensiones.
- Los visores proporcionan libertad a los usuarios finales independientes al departamento de TI.



TIPOS DE OLAP

MOLAP (Multidimensional OLAP)

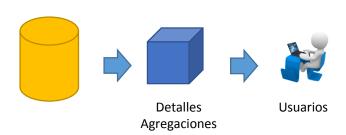
Utiliza cubos con todas las consultas precalculadas, los atributos de los cubos son las dimensiones.

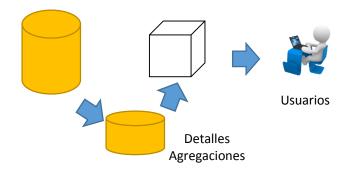
ROLAP (Relational OLAP)

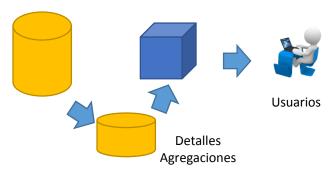
Trabaja directamente con las base de datos relacionales y crean nuevas tablas para guardar la información agregada.

HOLAP (Hybrid OLAP)

Combina las dos primeras de la mejor manera.









LA INFORMACION QUE NECESITAN

Alta Dirección

- Información agregada y destructurada.
- Decisiones estratégicas de mercado y productos.

Mandos Intermedios

- Información operativa y semi estructurada.
- Indicadores de seguimiento de la gestión.

Equipos y colaboradores

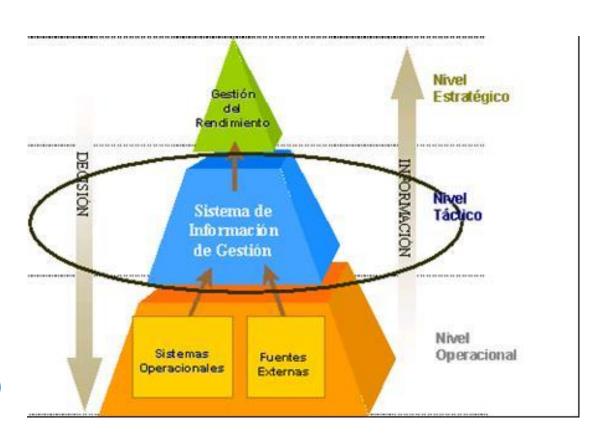
- Información estructurada
- Decisiones dentro de protocolos y reglas



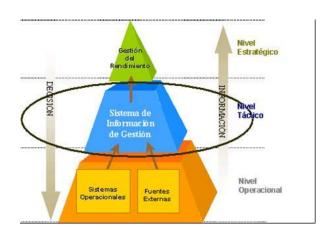
5 % - 10 % Usuarios (Ejecutivos)

15 % - 25 % Usuarios (Analistas)

65 % - 80 % Usuarios (Colaboradores)



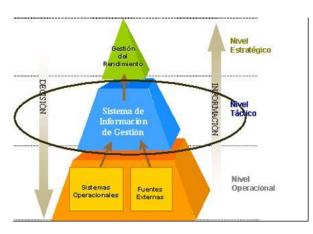




Información Estratégica

- Toma de decisiones en las áreas directivas para alcanzar la misión empresarial.
- Ejecutivos de primer nivel

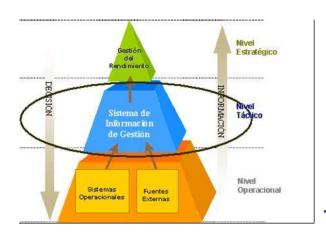




Información Táctica

- Coordinación de actividades y el plano operativo de la estrategia.
- Gerentes o subdirectores





Información Técnico Operacional

- Operaciones tradicionales que son efectuadas de modo rutinario en las empresas.
- Jefaturas o coordinaciones operativas o de tercer nivel.



Programa Ejecutivo de Business Intelligence & Big Data

Introducción al Business Intelligence

Ing. Arturo Rojas Medrano