**Universidad Nacional Mayor de San Marcos**

**Facultad de Ingeniería de Sistemas e Informática**

**E.A.P. de Ingeniería de Sistemas**



**Docente:** Dr. Vega Huerta, Hugo

**Alumna:** Ruitón Carrera, Joel Hernán

**Curso:** Proyecto deTesis II

**Proyecto:** Implementación de Inteligencia de Negocios para optimizar la toma de decisiones en el área de operaciones de Ipsos

# **RESUMEN**

**IMPLEMENTACIÓN DE INTELIGENCIA DE NEGOCIOS PARA OPTIMIZAR LA TOMA DE DECISIONES EN EL AREA DE OPERACIONES DE IPSOS**

En esta Tesis se implementa una solución de Inteligencia de Negocios que nos permitirá analizar y entender los datos obtenidos de sistemas transaccionales para tomar mejores decisiones y así cumplir con los objetivos del área de Operaciones de Ipsos. Se ha considerado como caso de estudio una compañía de servicios BPO en la cual se ofrecen servicios evaluación de calidad a la empresa Ipsos y una metodología de trabajo en la que se llegará al modelo dimensional y a la presentación de una solución que responderá las preguntas del negocio. El resultado será una toma de decisiones eficaz para el cliente Ipsos y mejora en la rentabilidad de la empresa GSS.

Como solución de Business Intelligence se diseña un datamart de gestión de encuestas, luego se realizarán los procesos de extracción, transformación y carga de datos, logrando un estado donde la explotación de los datos mediante reportes permitan hacer el análisis de la información.

**Palabras claves:** Business Intelligence, Toma de decisiones, Metodología de Ralph Kimball, Metodología Rapid Warehousing, Metodología Bill Inmon, Transaccional, Información.

**ÍNDICE**

[**RESUMEN** ii](#_Toc517448291)

[**ÍNDICE DE FIGURAS** v](#_Toc517448292)

[**ÍNDICE DE CUADROS** vi](#_Toc517448293)

[**GLOSARIO** vii](#_Toc517448294)

[INTRODUCCIÓN 1](#_Toc517448295)

[CAPÍTULO I: VISIÓN DEL PROYECTO 2](#_Toc517448296)

[1.1. Planteamiento del problema 2](#_Toc517448297)

[1.1.1. El Negocio 2](#_Toc517448298)

[1.1.2. Los Procesos de Negocio 3](#_Toc517448299)

[1.2. Formulación del Problema 3](#_Toc517448300)

[1.2.1. Realidad Problemática 3](#_Toc517448301)

[1.2.2. Descripción del Problema 4](#_Toc517448302)

[1.3. Objetivos del Proyecto 5](#_Toc517448303)

[1.3.1. Marco Lógico 5](#_Toc517448304)

[1.3.2. Objetivo general 6](#_Toc517448305)

[1.3.3. Objetivos específicos 6](#_Toc517448306)

[1.4. Importancia del proyecto 6](#_Toc517448307)

[1.4.1. Justificación académica 6](#_Toc517448308)

[1.4.2. Beneficios tangibles 6](#_Toc517448309)

[1.5. Alcance del proyecto 7](#_Toc517448310)

[CAPÍTULO II: MARCO TEÓRICO 8](#_Toc517448311)

[2.1. Business Intelligence 8](#_Toc517448312)

[2.1.1. Según (Arisa Shollo 2013) (ISBN: 978-87-92977-32-8) 8](#_Toc517448313)

[2.1.2. Según (Carlo Vercellis 2009) (978-0-470-51138-1) 11](#_Toc517448314)

[2.2. Metodologías de Business Intelligence 15](#_Toc517448315)

[2.2.1. Según (Ralph Kimball y Margy Ross) (ISBN: 978-1-118-53080-1) 15](#_Toc517448316)

[2.2.2. Según (W. H. Inmon) (ISBN-13: 978-0-7645-9944-6) 18](#_Toc517448317)

[CAPÍTULO III: ESTADO DEL ARTE 22](#_Toc517448318)

[3.1. Artículos 22](#_Toc517448320)

[3.1.1. Modelo de Inteligencia de Negocios de Gestión de Consultoría para una Empresa Analítica (Leonardo Mora, Oswaldo Díaz, Carlos Montenegro, 2013) (ISSN: 1390-5236) 22](#_Toc517448321)

[3.1.2. Modelamiento dimensional en competencias TIC (Martha Mendoza, Jaime Mendoza, Deyro Zuñiga, Jorge Moreno, Carlos Cobos, 2013) (ISSN: 1794-1237) 27](#_Toc517448322)

[3.1.3. Implementación de un Datamart como una solución de Inteligencia de Negocios para el área de logística de T-Impulso (Julio Yalan Castillo, Luis Palomino Paniora, 2012) (ISSN: 1815-0268) 30](#_Toc517448323)

[REFERENCIAS BIBLIOGRÁFICAS 35](#_Toc517448324)

# **ÍNDICE DE FIGURAS**

[Ilustración 1: Beneficios de usar un sistema de inteligencia de negocios 11](#_Toc515784962)

[Ilustración 2: Arquitectura de Business Intelligence común 12](#_Toc515784963)

[Ilustración 3: Componentes principales de un sistema BI 12](#_Toc515784964)

[Ilustración 4: Ilustración simplificada de una arquitectura de datamarts independientes 14](#_Toc515784965)

[Ilustración 5: Estructura de un Data Warehouse 17](#_Toc515784966)

[Ilustración 6: Fenómeno día 1- día n 19](#_Toc515784967)

[Ilustración 7: Aqruitectura del Sistema de Gestión de Consultoría 21](#_Toc515784968)

[Ilustración 8: Diagrama estrella para ventas 27](#_Toc515784969)

**FIGURA 30.** Ventana de Explorador de Soluciones………………………………...……94

**FIGURA 31.** Editor de Dimensiones para DIM ESPECIFICACION………………..……94

**FIGURA 32.** Editor de Dimensiones para DIM ZONA ………………………………..….95

**FIGURA 33.** Editor de Dimensiones para DIM TIEMPO……………………………..…..95

# **ÍNDICE DE CUADROS**

[Tabla 1: Resultados de Búsqueda 7](#_Toc517447929)

[Tabla 2: Tipos de definiciones de BI 9](#_Toc517447930)

[Tabla 3: Software empleado en el Sistema de Gestión de Consultoría 22](#_Toc517447931)

[Tabla 4: Matriz de comparación de Metodologías para el Módulo de Registro de Horas 24](#_Toc517447932)

[Tabla 5: Matriz de comparación de Kimball vs Inmon 24](#_Toc517447933)

[Tabla 6: Definición y ponderación de variables de Comparación de herramientas BI 32](#_Toc517447934)

[Tabla 7: Cuadro de puntuación de las herramientas BI 32](#_Toc517447935)

[Tabla 8: Cuadro comparativo de herramientas BI 32](#_Toc517447936)

# **GLOSARIO**

B

BI: Business Intelligence, viii, xvi, xvii, xviii, xxi, xxx, xxxi, xxxii, xxxiii

BPO: La subcontratación de procesos de negocios o externalización de procesos de negocio, ii, xi

C

CRM: es un software / programa / herramienta / aplicación en el que cualquier conversación que un compañero de tu empresa tenga con un cliente (o potencial cliente) se guarda en una zona común y accesible para todo el mundo en tu empresa., x

D

DSS: Sistema de apoyo de toma de decisiones, xxvii, xxviii, xxxvi

G

GSS: Empresa Global Sales Solutions, ii, iii, xi, xii

# **INTRODUCCIÓN**

El presente trabajo de investigación tiene como objetivo principal el desarrollo de Business Intelligence, aplicando la metodología de Ralph Kimball para mejorar la toma de decisiones en el área de Operaciones de Ipsos .

En todas las organizaciones se toma decisiones a diario, trascendentes o intranscendentes, pero todas ellas sin estar exentas de riesgo. Los encargados de tomar las decisiones requieren minimizar este riesgo, teniendo para esto la mayor cantidad de información, la cual debe ser oportuna, eficiente y además que agregue valor. A pesar de que vivimos en la era de la información, donde crecientes volúmenes de información están a nuestra disposición, frecuentemente luchamos para comprender el significado de los datos.

La finalidad de usar Business Intelligence consiste en convertir los datos contenidos en las bases de datos corporativas de las organizaciones, en información y ésta a su vez en conocimiento útil en el proceso de toma de decisiones estratégicas.

# CAPÍTULO I: VISIÓN DEL PROYECTO

## Planteamiento del problema

### El Negocio

El Grupo GSS (Global Sales Solutions) es una multinacional de origen español dedicada a la gestión CRM. Como tal está presente en la gran mayoría de los países de Latinoamérica. Sus principales centros de trabajo están situados en Madrid y Lima aunque cuenta con centros en varias ciudades de España​ y del mundo. Alcanzó en el año 2010 un volumen de 6.300 empleados​ la mayor parte de ellos en España, aunque pretende potenciar sus centros de Chile, México y Colombia. La empresa se ha destacado desde el principio por una agresiva política de captación, motivación y retención de recursos humanos.

#### Misión

Ayudar a nuestros clientes a optimizar la gestión de clientes, las ventas, la atención y la experiencia de sus clientes con sus marcas.

#### Visión

La visión de GSS es liderar en el ámbito internacional la oferta de soluciones de gestión de clientes en tiempo real, de manera integral y multicanal; mediante la innovación continua en nuevas soluciones, tecnologías y procesos.

#### Valores

* Grupo independiente de empresas gestionado por socios profesionales  líderes desde 1.986, con experiencia en la implantación y gestión de centros de relación con clientes.
* Nacimos con el objetivo de aportar a nuestros clientes un valor diferencial, hacemos  soluciones personalizadas, únicas. No somos meros proveedores de servicio, nos convertimos en partners de nuestros clientes.
* Nuestra experiencia de clientes es el resultado de una larga serie de éxitos. La experiencia del Grupo GSS y de sus profesionales le garantiza el éxito de todos sus proyectos, dado que ayudamos a nuestros clientes a optimizar la gestión de clientes, las ventas, la atención y la experiencia de sus clientes con sus marcas.

### Los Procesos de Negocio

Las principales áreas usuarias beneficiarias de GSS involucradas en el planteamiento del problema y su solución son el área de ventas, el área de operaciones, el área contable, el área de calidad y el área de tecnología.

El área de ventas se encarga de vender los productos y/o servicios de calidad de la empresa con el fin de conseguir rentabilidad y fidelizar a los clientes. Una vez concretada la venta es el área de operaciones la que asume la responsabilidad de brindar el servicio contratado, gestionar nuevos requerimientos y trasladar la información requerida por el cliente, toda nueva solicitud es gestionada de la mano con el área contable de modo que se puedan establecer los costos de cada una de estas.

El área de calidad colabora constantemente con la elaboración de reportes, están dirigidos por la gerente de calidad. Entre sus responsabilidades están la de elaborar reportes ricos en valor para cliente (paretos, dashboards, resultados mensuales, etc), ésta tarea es bastante dinámica debido a que mensualmente se definen nuevos estándares de medición de calidad en el servicio brindado.

Por último tenemos al área de tecnología, quienes son responsables del desarrollo de los nuevos requerimientos, mantenimiento de la página web mediante la cual el área de operaciones realiza sus labores y modificación mensual de los reportes transaccionales de los cuales debe disponer el área de calidad para realizar su gestión.

* 1. **Formulación del Problema**

### Realidad Problemática

Actualmente en nuestro país existe un entorno competitivo en el área de servicios BPO, debido a la existencia de un gran número de multinacionales dedicadas al customer service. Por este motivo, las compañías deben enfocar su trabajo a satisfacer las necesidades de sus clientes de tal manera que permanezca la fidelidad de éstos con la empresa.

Una de las necesidades es el oportuno y eficiente servicio de atención al cliente, en los cuales se registran los requerimientos (reportes, aplicaciones, entre otros) mediante comunicación personal, escrita, telefónica o vía web.

Un cliente, al hacer un contrato con GSS, da un primer paso tras el cual se abre una relación que hay que saber ampliar y mejorar. Para ello se debe, por una parte, asegurar la calidad del servicio mediante la supervisión de la red, la resolución de los incidentes, y, por otra parte, cualquier requerimiento debe ser atendido con agilidad y utilizado como una fuente de información sobre el cliente para permitir mejorar el servicio que se le está prestando.

Las demoras en atención y mala calidad de ésta son unas de las principales fuentes de pérdida de productividad en las organizaciones, mientras más rápido un cliente sea atendido y mayor sea la calidad del servicio brindado, la aceptación de este aumentará. GSS cuenta procedimientos para la solicitud de servicios e intermediarios para el reporte de problemas que deben ser atendidos, en el mejor de los casos de manera inmediata. Todas las auditorías de calidad realizadas por parte de GSS hacia las llamadas del cliente se encuentran almacenadas en una base de datos, diariamente se maneja un volumen muy grande de evaluaciones, información que a su vez debe estar debidamente organizada por diversos criterios que aumentan a petición del cliente.

La jefatura del centro calidad lleva un control de una cantidad numerosa de reportes de diferentes campañas, para tener conformidad de todo lo que se ha realizado; sin embargo el proceso que abarca su construcción y presentación del mismo hacia la alta jefatura del centro de operaciones del cliente no es óptimo, debido a que el valor de la información presentado no es de mucha ayuda, lo que hace dificultoso la toma de acciones.

### Descripción del Problema

Actualmente GSS brinda servicios de auditoría de calidad a la empresa Ipsos de Ecuador, nuestra área de calidad envía mensualmente varios reportes, estos se elaboran de manera manual sobre Excel, para ello se solicita al área de tecnología la actualización de los datos transaccionales ya que los cambios se dan con mucha continuidad, esto tiene como resultado un tiempo prolongado de entrega de la información al cliente lo que conlleva a una deficiente toma de decisiones en el área de operaciones de Ipsos (variable 1: tiempo de entrega de reportes, valor 1: 3 días; variable 2: tiempo de espera para visualizar reportes, valor 3: 40 segundos), lo cual genera malestar en la relación de la empresa con el cliente.

Asimismo se evidencia una información histórica insuficiente, (variable 3: antigüedad de la información histórica, valor 3: 2 meses) este comportamiento afecta directamente al proceso de toma de decisiones del cliente debido a que no cuenta con información rápida y completa.

## Objetivos del Proyecto

### Marco Lógico

#### Árbol de problemas

Malestar en la relación con el cliente

Lentitud para poder visualizar resultados

Dependencia de GSS para obtener reportes

**Ineficiencia en el proceso de toma de decisiones de Ipsos** (variable 1: tiempo de entrega de reportes, valor 1: 3 días) (variable 2: antigüedad de la información histórica, valor 2: 2 meses) (variable 3: tiempo de espera para visualizar reportes, valor 3: 40 segundos)

Información disponible insuficiente

Demora para brindar información al cliente

Carga lenta de datos

#### Árbol de objetivos

Rapidez para visualizar los reportes

Mayor satisfacción del cliente

Independencia de GSS para obtener reportes

**Mejora en el proceso de toma de decisiones de Ipsos** (variable 1: tiempo de entrega de reportes, valor 1: 15 minutos) (variable 2: tiempo de espera para visualizar reportes, valor 2: 8 segundos) (variable 3: antigüedad de la información histórica, valor 3: 15 años)

Reportes disponibles online (actualizados de manera automática diariamente)

Carga rápida de datos

### Objetivo general

Para solucionar esta situación se desarrollarán y se publicarán en línea los reportes requeridos por nuestro cliente dentro de la aplicación web a la cual se accede para realizar la gestión diaria, con ello se optimizará la toma de decisiones del cliente (variable 1: tiempo de entrega de reportes, valor 1: 15 minutos; variable 2: tiempo de espera para visualizar reportes, valor 2: 8 segundos), lo cual reforzará la relación y el bienestar del cliente.

Para cumplir de manera óptima con los objetivos aplicaremos inteligencia de negocios, con lo que nos aseguraremos de contar con información histórica suficiente (variable 3: antigüedad de la información histórica, valor 3: 15 años) para poder tomar decisiones de una manera más adecuada.

### Objetivos específicos

* Analizar la base de datos transaccional para determinar el modelado dimensional del datamart.
* Creación de procesos de carga para poblar el datamart (ETL).
* Implementar un Cubo que explote la información almacenada en el datamart.
* Diseñar y desarrollar un Dashboard con información de fácil entendimiento y manejo para el cliente.
* Evaluación y validación de resultados obtenidos.

## Importancia del proyecto

### Justificación académica

Dar a conocer la importancia y las ventajas que nos facilita actualmente la inteligencia de negocios y cómo es que nos ayuda a tener mejores resultados que los reportes transaccionales convencionales, de modo que se facilite enormemente el proceso de toma de decisiones.

### Beneficios tangibles

1. Incrementar la velocidad de carga de los reportes.
2. Incrementar la velocidad de entrega de informes al cliente.
3. Mayor cantidad de información disponible en los reportes.
4. Beneficios intangibles
5. Mejorar la satisfacción del cliente
6. Mejorar el índice de aceptación de la empresa GSS

## Alcance del proyecto

La presente tesis tiene como alcance la creación de los reportes requeridos por el cliente, así como un dashboard a partir de un datamart, el cuál reflejará información concreta y de fácil entendimiento para el cliente de modo que se simplifique su proceso de toma de decisiones.

# CAPÍTULO II: MARCO TEÓRICO

## Business Intelligence

### Según (Arisa Shollo 2013) (ISBN: 978-87-92977-32-8)

El autor, intentando llegar a una definición literal de Inteligencia de Negocios, realizó una revisión del concepto y su evolución desde sus etapas iniciales hasta su uso contemporáneo, logrando con esto identificar dos flujos los cuales difieren a cómo es que se ve y se entiende la inteligencia de negocios, un punto de vista tecnológico y otro que se centra en el proceso que se lleva a cabo cuando se aplica inteligencia de negocios. El punto de vista tecnológico se centra en las tecnologías de inteligencia de negocios que permiten la recopilación, almacenamiento, recuperación y análisis de datos, mientras que la de proceso se enfoca en el proceso de recuperación y análisis de datos para generar valor al proceso de toma de decisiones. A continuación se presentará la visión general que el autor de la a la inteligencia de negocios y después se ahondará en los puntos de vista anteriormente mencionados.

#### Resumen de la literatura de BI

Desde un primer vistazo a la literatura (Davenport y Prusak 1998) uno entiende que BI está relacionado con la toma de decisiones, la gestión estratégica y la gestión del rendimiento. En su investigación, el autor intenta aclarar el estado actual de la inteligencia de negocios en relación a la toma de decisiones, la gestión estratégica y de rendimiento; sin embargo, debido a que el campo de inteligencia de negocios es aún bastante novedoso, el autor optó por realizar una búsqueda exhaustiva en bases de datos bien establecidas, así como en las 10 principales revistas de los sistemas de información. Como se estableció una tabla la cual incluye un periodo de 20 años entre los años 1991 y 2010, en la primera fila se tienen la cantidad de resultados de la búsqueda, en la segunda los resultados de filtrar las frases “Inteligencia de negocios” y “decisiones” en el título o resumen, en la última fila se realizó un escaneo para determinar cuáles artículos son de importancia para la investigación, como resultado se tuvo un total de 103 artículos entre bases de datos y revistas:

|  |  |  |  |
| --- | --- | --- | --- |
|  | Resultados de búsqueda | Filtrado en títulos y resúmenes | Escaneo de resúmenes |
| Bases de datos | 3542 | 144 | 56 |
| Top revistas | 152 | 73 | 47 |
| Total | 3694 | 217 | 103 |

Tabla 1: Resultados de Búsqueda

#### El concepto de Inteligencia de Negocios

Desde un punto de vista histórico, el concepto de inteligencia no es nuevo ya que desde tiempos remotos el hombre ha desarrollado procesos, técnicas y herramientas de recolección y análisis de información para tomar decisiones más eficientes, sobre todo en tiempos de guerra (Kinsinger 2007).

Uno de los principales problemas en la definición de Inteligencia de negocios es que existe una confusión entre esta y la inteligencia competitiva (CI). La CI es un proceso y a la vez un producto. Como proceso es el conjunto de métodos legales y éticos que una organización utiliza para aprovechar la información que le ayuda a alcanzar el éxito en un entorno global; como producto es la información de fuentes públicas y privadas sobre las actividades de competidores, su alcance es el comportamiento presente y futuro de competidores, proveedores, clientes, tecnologías, adquisiciones, mercados, productos y servicios, y el negocio general del entorno. (p 109)

En contraste con CI y "espionaje industrial" que están orientados a la reunión de información y explotación de fuentes externas, BI comenzó como una utilización de todas los datos transaccionales acumulados y producidos dentro de una organización (Yermish et al. 2010). Por lo tanto, BI era en ese punto bastante diferente de CI en el alcance. Sin embargo, hoy las cosas han cambiado. Los sistemas transaccionales incluyen mucha información externa y, con el crecimiento de Internet, el potencial de BI se ha expandido significativamente (Negash 2004). Por lo tanto se observa CI como un subconjunto de BI o una "rama especializada de BI" (Negash 2004) y concuerda con Choo (2002) que BI tiene el alcance más amplio entre los conceptos de inteligencia.

BI se trata tanto de conocer las fortalezas y debilidades de una organización como el estado de sus competidores u otros factores externos (tales como el medio económico y político) (Choo 2002; Negash 2004).

Es por ello que el autor considera que no existe un acuerdo sobre la definición de inteligencia de negocio, ya que, a pesar de no ser un concepto nuevo, ha experimentado cambios y adiciones a través del tiempo. En la investigación realizada por Shollo se define a BI como un proceso o conjunto de tecnologías; pero sin centrarse en el producto obtenido, ya que se le considera integrado a estos. De acuerdo a lo expuesto por el autor existen dos perspectivas que principales que intentan definir BI, la que definen BI como un proceso, esta reconoce el producto de este proceso de BI; y la que define BI como conjunto de tecnologías, esta se refiere solo implícitamente al producto de esas tecnologías:

|  |  |
| --- | --- |
| Tipos de definiciones | Definiciones de BI |
| BI como proceso | "Business Intelligence (BI) se puede definir como el proceso de convertir datos en información y luego en conocimiento. "(Golfarelli et al. 2004, p.1)  "... como la actividad continua de recopilación, procesamiento y análisis de datos - respaldado por un sistema BI. "(Dekkers et al., 2007, p.626)  "El término BI se puede usar para referirse a:  1. Información relevante y conocimiento que describe el negocio del entorno, la propia organización y su situación en relación con sus mercados, clientes, competidores y problemas económicos  2. Un proceso organizado y sistemático por el cual las organizaciones adquieren, analizan y difunden información tanto interna como externa de fuentes de información significativas para sus actividades comerciales |
| BI como conjunto de tecnologías | "Business Intelligence (BI) es una amplia categoría de tecnologías, aplicaciones, y procesos para recopilar, almacenar, acceder y analizar datos para ayudar a su los usuarios toman mejores decisiones". (Wixom y Watson 2010, p.14)  "La inteligencia de negocios abarca todas las aplicaciones de software y tecnologías que una empresa utiliza para recopilar, proporcionar acceso y analizar datos e información sobre sus operaciones. "(Pemmaraju 2007, p.14)  "La inteligencia de negocios abarca un conjunto de herramientas, técnicas y procesos para ayudar a aprovechar esta amplia gama de datos y permitir a los responsables de la toma de decisiones convertir a información y conocimiento útiles. "(Clark et al., 2007, p.589) |

Tabla 2: Tipos de definiciones de BI

### Según (Carlo Vercellis 2009) (978-0-470-51138-1)

#### Business Intelligence

El autor expone que actualmente, la llegada de tecnologías de almacenamiento de datos y el uso de internet ha facilitado el acceso, tanto a personas como a organizaciones, a una considerable cantidad de datos de diferente índole y que sería de mucha utilidad poder convertirlos en información y conocimiento que apoye al proceso de toma de decisiones con el fin de optimizar la administración pública y de gobernanza de las empresas.

En este sentido el autor define a la inteligencia de negocios como un conjunto de modelos matemáticos y metodologías de análisis que explotan los datos disponibles que explotan los datos disponibles para generar información y conocimiento útil para procesos complejos de toma de decisiones.

#### Decisiones efectivas y oportunas

En todo tipo de organizaciones se toman decisiones diariamente, estas pueden ser de diferente nivel de criticidad y tomadas por personas de diferente nivel jerárquico, quienes, en su gran mayoría, llegan a dichas decisiones principalmente a través del uso de metodologías fáciles e intuitivas las cuales a su vez se basan en la experiencia, dominio de aplicaciones e información disponible. No obstante, los procesos de toma de decisiones actuales son muy complejos y dinámicos para ser tratados por este enfoque intuitivo, por lo que se requiere usar en su lugar metodologías analíticas y modelos matemáticos.

* + - * 1. **Decisiones efectivas**

La aplicación de metodologías analíticas brinda información de confianza a las personas a cargo de la toma de decisiones, lo que conlleva a que se alcancen los objetivos trazados de una manera más efectiva. Al aplicar este tipo de metodologías es necesario describir los criterios para evaluar opciones alternas y mecanismos de regulación del caso bajo investigación, este nivel de profundidad de pensamiento conduce a una compresión más amplia de la lógica del proceso de toma de decisiones.

* + - * 1. **Decisiones oportunas**

Actualmente las organizaciones operan en un entorno muy dinámico y competitivo, por lo que se debe tener una capacidad de respuesta rápida por parte de estas ante el accionar de los competidores y los cambios en el entorno. Para tomar decisiones, los responsables de ellas se cuestionan una serie de interrogantes y realizan un análisis de ello, lo que conlleva a plantear diversas alternativas y seleccionar la mejor de ellas. La adopción de la inteligencia de negocios facilitaría esta labor, ya que además de proporcionar información confiable, brinda una gama más amplia de opciones que permitirían conclusiones más precisas y decisiones oportunas.

A continuación el autor nos presenta un gráfico con los beneficios de usar inteligencia de negocios para la toma de decisiones:

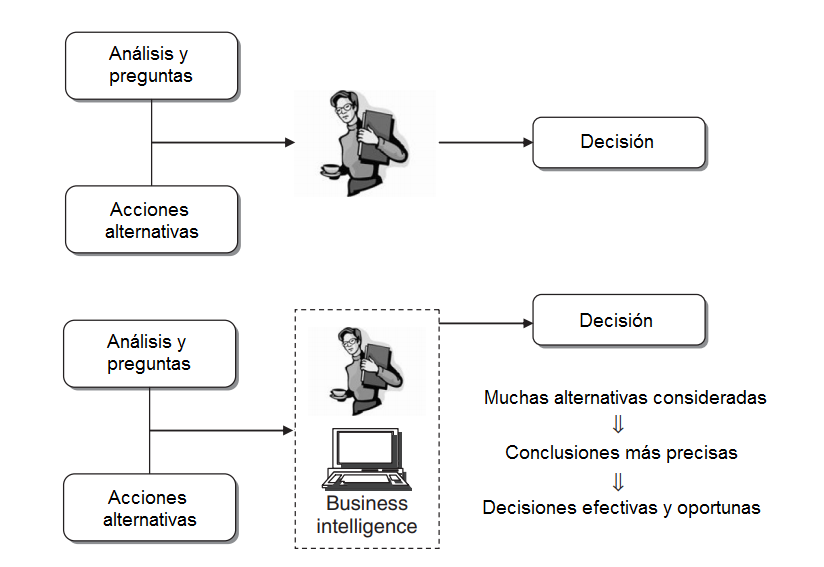


Ilustración 1: Beneficios de usar un sistema de inteligencia de negocios

#### Arquitecturas de Business Intelligence

La arquitectura de la inteligencia de negocios tiene 3 componentes principales:

* + - * 1. **Origen de datos**

Para empezar se necesita reunir e integrar los datos disponibles, los cuales mayormente son heterogéneos y provienen de diferentes tipos de fuentes.

* + - * 1. **Data warehouses y Data marts**

Usando herramientas de extracción y transformación de datos (ETL), se almacenan los datos de los orígenes de datos a bases de datos destinadas a apoyar el análisis de inteligencia de negocios, a estas bases se les llama comúnmente data warehouses y data marts.

* + - * 1. **Metodologías de BI**

Los datos finalmente son recuperados para el análisis destinado a brindar soporte a los responsables de la toma de decisiones. En inteligencia de negocios existen diferentes aplicaciones de apoyo que sirven de apoyo, las principales son:

• Análisis de cubo multidimensional

• Análisis exploratorio de datos

• Análisis de series temporales

• Modelos de aprendizaje inductivo para la minería de datos

• Modelos de optimización

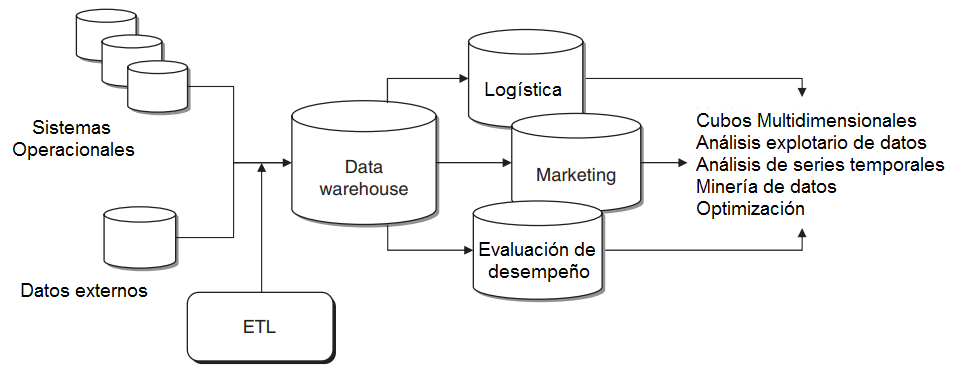


Ilustración 2: Arquitectura de Business Intelligence común

La pirámide en la Ilustración 3 muestra los componentes básicos de una inteligencia comercial sistema. Hasta ahora, hemos visto los componentes de los primeros dos niveles al discutir la Ilustración 2. Pasamos ahora a la descripción de los niveles superiores.

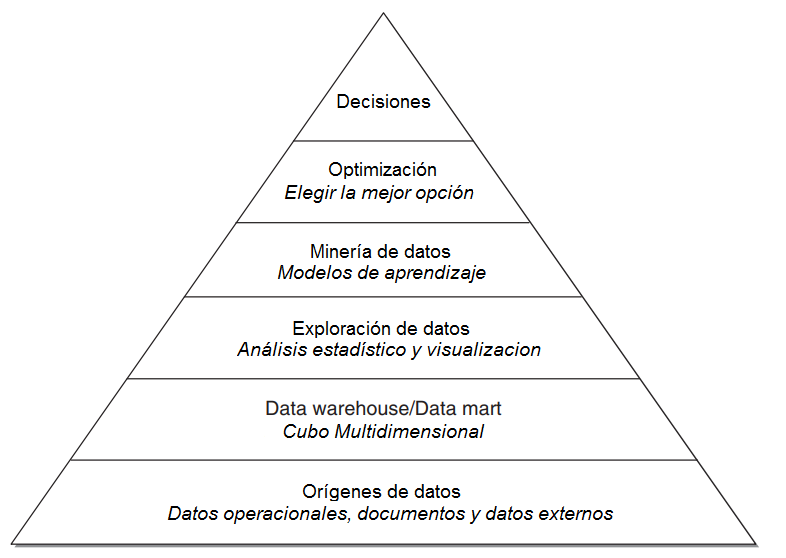


Ilustración 3: Componentes principales de un sistema BI

* + - * 1. **Exploración de datos**

Son las herramientas para realizar un análisis pasivo de inteligencia de negocios, son llamados así porque las personas a cargo de la toma de decisiones deben hacer hipótesis o definir criterios de extracción de datos para luego utilizar estas herramientas para buscar respuestas y confirmar su visión original.

* + - * 1. **Minería de datos**

Este nivel incluye metodologías activas de inteligencia de negocios, su propósito es la extracción la información y el conocimiento de los datos, son llamados activos debido a que no se requiere que las personas a cargo formulen alguna hipótesis previa para ser verificado en el futuro. Su función es expandir la toma de decisiones de los tomadores de decisiones.

* + - * 1. **Optimización**

Son los modelos que nos permiten encontrar la mejor solución entre todas las alternativas disponibles.

* + - * 1. **Decisión**

Por último, tenemos la elección y adopción de una decisión específica, la cual es el resultado final del proceso de toma de decisiones. A pesar de haber aplicado exitosamente las metodologías disponibles de inteligencia de negocios, la decisión final pertenece a los tomadores de decisiones, por lo que no sería extraño que estos adapten y modifiquen las conclusiones logradas.

## Metodologías de Business Intelligence

### Según (Ralph Kimball y Margy Ross) (ISBN: 978-1-118-53080-1)

Según los autores la implementación de un datamart debe ser independiente por departamentos dentro de la organización, es decir sin la preocupación de tener que integrar y compartir información con otros departamentos. Para ello es necesario que dicha área despeje las dudas y requerimientos necesarios para obtener información que satisfaga sus necesidades con el área tecnológica. Es común que más de un área esté interesada en tener su propio datamart para administrar de forma más efectiva su información, por lo que se realiza el mismo proceso en las mismas con ligeras cambios estructurales; sin embargo esto no significa que se tenga el mismo desempeño organizacional en todas las áreas debido a las evidentes diferencias en las reglas del negocio.

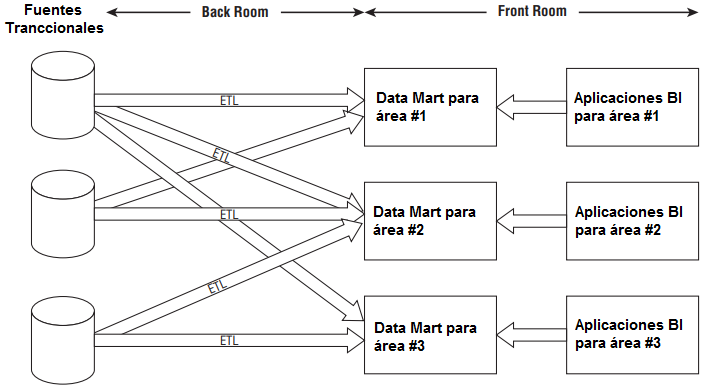


Ilustración 4: Ilustración simplificada de una arquitectura de datamarts independientes

#### Planificación del proyecto

La planificación busca identificar la definición y el alcance del proyecto de data warehouse, incluyendo justificaciones del negocio y evaluaciones de factibilidad. La planificación del proyecto se focaliza sobre recursos, perfiles, tareas, duraciones y secuencialidad. El plan de proyecto resultante identifica todas las tareas asociadas e identifica las partes involucradas.

#### Definición de los requerimientos del negocio

Los usuarios finales y sus requerimientos impactan siempre en las implementaciones realizadas de un data-warehouse. Según la perspectiva de Kimball, los requerimientos del negocio se posicionan en el centro del “universo del data warehouse”. Como destaca siempre el autor, los requerimientos del negocio deben determinar el alcance del data warehouse (qué datos debe contener, cómo debe estar organizado, cada cuánto debe actualizarse, quiénes y desde dónde accederán, etc.).

Kimball da consejos y técnicas para descubrir eficazmente los requerimientos del negocio. Estas tácticas y estrategias se focalizan sobre las entrevistas de relevamiento (diferentes tipos, preparación de la entrevista, roles a cubrir.

#### Modelado dimensional

La definición de los requerimientos del negocio determina los datos necesarios para cumplir los requerimientos analíticos de los usuarios. Diseñar los modelos de datos para soportar estos análisis requiere un enfoque diferente al usado en los sistemas operacionales.

#### Diseño físico

El diseño físico de las base de datos se focaliza sobre la selección de las estructuras necesarias para soportar el diseño lógico. Algunos de los elementos principales de este proceso son la definición de convenciones estándares de nombres y seteos específicos del ambiente de la base de datos.

#### Diseño y desarrollo de presentación de datos

Se definen como procesos de transformación los procesos para convertir o recodificar los datos fuente a fin poder efectuar la carga efectiva del modelo físico. Por otra parte, los procesos de carga de datos son los procesos requeridos para poblar el data warehouse.

#### Diseño de la arquitectura técnica

Los ambientes de data-warehousing requieren la integración de numerosas tecnologías. Se debe tener en cuenta tres factores: los requerimientos del negocio, los actuales ambientes técnicos y las directrices técnicas estratégicas futuras planificadas.

#### Selección de productos e instalación

Utilizando el diseño de arquitectura técnica como marco, es necesario evaluar y seleccionar componentes específicos de la arquitectura como ser la plataforma de hardware, el motor de base de datos, la herramienta de etl o el desarrollo pertinente, herramientas de acceso, etc.

#### Especificación de aplicaciones para usuarios finales

No todos los usuarios del warehouse necesitan el mismo nivel de análisis. Es por ello que en esta etapa se identifican los diferentes roles o perfiles de usuarios para determinar los diferentes tipos de aplicaciones necesarias en base al alcance de los diferentes perfiles (gerencial, analista del negocio, vendedor, etc.).

#### Implementación

La implementación representa la convergencia de la tecnología, los datos y las aplicaciones de usuarios finales accesible desde el escritorio del usuario del negocio. Hay varios factores extras que aseguran el correcto funcionamiento de todas estas piezas, entre ellos se encuentran la capacitación, el soporte técnico, la comunicación, las estrategias de feedback.

#### Mantenimiento y crecimiento

Data Warehousing es un proceso (de etapas bien definidas, con comienzo y fin, pero de naturaleza espiral) pues acompaña a la evolución de la organización durante toda su historia. Se necesita continuar con los relevamientos de forma constante para poder seguir la evolución de las metas por conseguir. Según afirma Kimball, “si se ha utilizado el ciclo de vida de business dimensional el data warehouse está preparado para evolucionar y crecer”. Al contrario de los sistemas tradicionales, los cambios en el desarrollo deben ser vistos como signos de éxito y no de falla.

#### Gerenciamiento del proyecto

El gerenciamiento del proyecto asegura que las actividades se lleven en forma y sincronizadas. Como lo indica el diagrama, el gerenciamiento acompaña todo el ciclo de vida. Entre sus actividades principales se encuentra el monitoreo del estado del proyecto y la comunicación entre los requerimientos del negocio y las restricciones de información para poder manejar correctamente las expectativas en ambos sentidos.

### Según (W. H. Inmon) (ISBN-13: 978-0-7645-9944-6)

El autor de esta investigación menciona que, a lo largo del tiempo, el hombre ha creado y aprovechado sistemas de apoyo para la toma de decisiones y su arquitectura ha ido evolucionando hasta el día de hoy. Inmon identifica cuatro niveles de datos y procesamiento en esta arquitectura: nivel operativo, nivel de data warehouse, nivel departamental (datamart) y el nivel individual.

#### Estructura de un Data Warehouse

La Ilustración 5 muestra que hay diferentes niveles de detalle en el almacén de datos. Hay un nivel de detalle más antiguo (por lo general en almacenamiento alterno, a granel), una corriente nivel de detalle, un nivel de datos ligeramente resumidos (el nivel del centro de datos) y un nivel de datos altamente resumidos. Los datos fluyen al almacén de datos desde el entorno operativo. Por lo general, se produce una transformación significativa de los datos en el paso del nivel operacional al nivel del almacén de datos. Una vez que los datos envejecen, pasa de los detalles actuales a los más antiguos. Como los datos son resumidos, pasa del detalle actual a datos ligeramente resumidos, luego de datos ligeramente resumidos a datos muy resumidos.

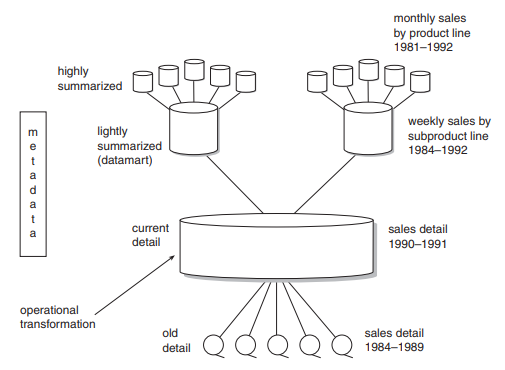


Ilustración 5: Estructura de un Data Warehouse

#### Fenómeno día 1 – día n

Un data warehouse no se construye todos a la vez. En cambio, está diseñado y poblado un paso a la vez, son evolutivos, no revolucionarios. Los costos de construir un data warehouse de una vez, los recursos requeridos y la interrupción para el medio ambiente todos dictan que el almacén de datos se construya en una modo iterativo ordenado, paso a paso.

La Ilustración 6 muestra el proceso típico de construcción de un data warehouse. En el día 1 hay un políglota de sistemas heredados que hace esencialmente operativos, transaccionales tratamiento.

El día 2, las primeras tablas del primer tema de los datos el almacén está poblado En este punto, se levanta una cierta cantidad de curiosidad, y los usuarios comienzan a descubrir almacenes de datos y procesamiento analítico.

El día 3, se puebla más el data warehouse y con la población de más datos provienen más usuarios. Una vez que los usuarios encuentran que hay una fuente integrada de datos de fácil acceso y tienen una base histórica diseñada para mirar datos con el tiempo, hay más que curiosidad. Aproximadamente en este momento, el analista serio de DSS se siente atraído por el data warehouse.

El día 4, a medida que se llena más del almacén, algunos de los datos en el entorno operativo se colocan adecuadamente en el almacén de datos. Y el almacén de datos ahora se descubre como una fuente de haciendo procesamiento analítico. Todo tipo de aplicaciones DSS surgen. De hecho, muchos usuarios y muchas solicitudes de procesamiento, junto con un gran volumen de datos que ahora reside en el almacén, parece que algunos usuarios están aplazados por el esfuerzo requerido para llegar al almacén de datos. La competencia para llegar al almacén se convierte en un obstáculo para su uso.

El día 5, comienzan a florecer las bases de datos departamentales (data mart u OLAP). Los departamentos encuentran que es más barato y más fácil hacer que su procesamiento sea realizado por llevar los datos del almacén de datos a su propio departamento de procesamiento ambiente. A medida que los datos van al nivel departamental, algunos analistas de DSS son atraídos.

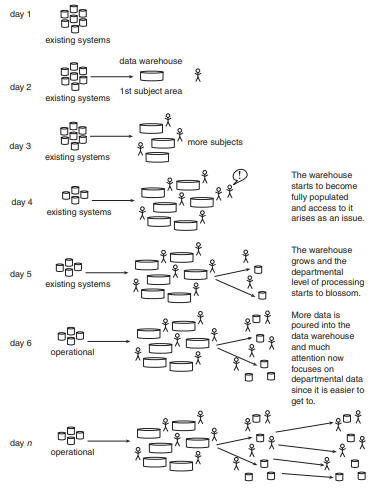


Ilustración 6: Fenómeno día 1- día n

#### Granularidad

El aspecto más importante del diseño de un almacén de datos es la cuestión granularidad. De hecho, el problema de la granularidad impregna toda la arquitectura que rodea el entorno del almacén de datos. Este concepto se refiere al nivel de detalle o resumen de las unidades de datos en el almacén de datos. Cuanto más hay detalles, menor es el nivel de granularidad. Cuanto menos detalles hay, el mayor el nivel de granularidad.

Muchas organizaciones se sorprenden al descubrir que el almacenamiento de datos proporciona una base invaluable para muchos tipos diferentes de procesamiento de DSS. Las organizaciones pueden construir un data warehouse para un propósito, pero descubrir que puede ser utilizado para muchos otros tipos de procesamiento DSS. Aunque la infraestructura para el almacén de datos es costosa y difícil de construir, solo tiene que construirse una vez. Después de que el almacén de datos se ha construido correctamente, proporciona a la organización una base extremadamente flexible y reutilizable.

Los datos granulares encontrados en el almacén de datos son la clave de la reutilización, porque puede ser utilizado por muchas personas de diferentes maneras. Por ejemplo, dentro de una corporación, los mismos datos podrían usarse para satisfacer las necesidades de mercadeo, ventas y contabilidad. Los tres departamentos miran los mismos datos básicos. Marketing puede desear ver las ventas mensualmente por distrito geográfico, Ventas pueden querer ver ventas por vendedor por distrito de ventas semanalmente, y Finanzas puede desear ver ingresos reconocibles trimestralmente por línea de productos. Todos estos tipos de información están estrechamente relacionados, aunque ligeramente diferentes.

Un beneficio relacionado es la capacidad de conciliar datos, si es necesario. Una vez que exista una base única en la que todos confían, si es necesario explicar una discrepancia en los análisis entre dos o más departamentos, luego la reconciliación es relativamente simple.

Otro beneficio relacionado es la flexibilidad. Supongamos que el marketing desea alterar cómo se ve en los datos. Tener una base en el lugar permite que esto se logre fácilmente.

Otro beneficio de los datos granulares es que contiene un historial de actividades y eventos en toda la corporación. Y el nivel de granularidad es lo suficientemente detallado que los datos pueden ser reformados en toda la corporación para muchas necesidades diferentes.

Pero tal vez el mayor beneficio de una base de almacenamiento de datos es que el futuro de los requisitos desconocidos puede ser acomodado. Supongamos que hay un nuevo requisito para ver los datos, o la legislatura estatal aprueba una nueva ley, o los cambios de la OPEP sus reglas para la asignación de petróleo, o las caídas del mercado de valores. Hay un constante flujo de nuevos requisitos de información porque el cambio es inevitable. Con el data warehouse en su lugar, la corporación puede responder fácilmente al cambio. Cuando surge un nuevo requerimiento y existe la necesidad de información, los datos en el almacén ya están disponibles para el análisis, y la organización está preparada para manejar los nuevos requisitos.

# CAPÍTULO III: ESTADO DEL ARTE



## Artículos

### Modelo de Inteligencia de Negocios de Gestión de Consultoría para una Empresa Analítica (Leonardo Mora, Oswaldo Díaz, Carlos Montenegro, 2013) (ISSN: 1390-5236)

En el trabajos presentado por los autores se muestra el diseño e implementación de un modelo BI para automatizar el proceso de Gestión de Consultoría de una empresa analítica. Para conseguir esto se siguieron 2 metodologías, Programación Extrema para el Módulo de Registro de Horas; y la metodología de Ralph Kimball, para el modelamiento dimensional de Módulo BI de Gestión de Consultoría, ya que se enfoca en la creación de un data warehouse especializado que se adapta a las necesidades de su proyecto.

#### Arquitectura de la solución

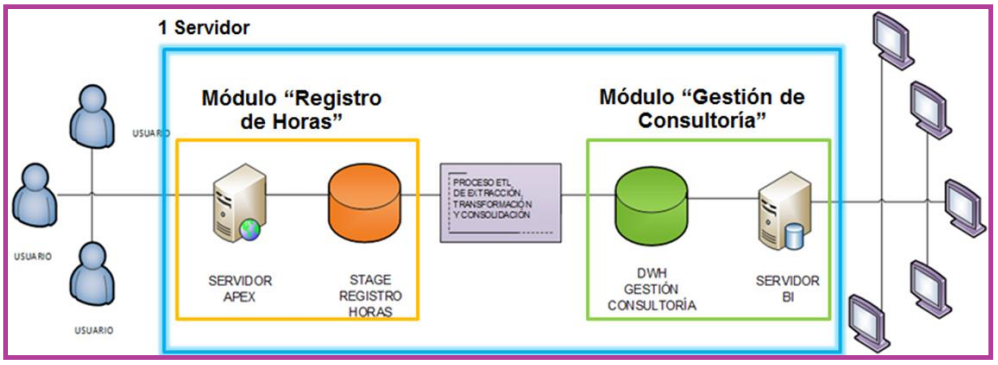


Ilustración 7: Aqruitectura del Sistema de Gestión de Consultoría

Como se puede observar, el proyecto consta de dos módulos prioritarios Registro de Horas y Gestión de Consultoría. Las herramientas necesarias para el desarrollo del aplicativo Web y para el modelo de BI se instalarán en un servidor proporcionado por la empresa el cual cuenta con una base de datos Oracle.

* **Módulo de Registro de Horas**: Es una aplicación Web desarrollada sobre Oracle Application Express, que permite el registro de las actividades diarias por hora del personal de la empresa, la gestión general de proyectos, esto involucra planificaciones, evaluación de tiempoes y costos de ejecución.
* **Módulo de Gestión de Consultoría**: Es el resultado del modelamiento dimensional a partir del modelo entidad – relación del módulo anterior. Este módulo abarca el modelamiento físico y lógico del modelo de BI, considerando el modelamiento dimensional, ETLs, flujos de trabajo, calendarizaciones, creación y publicación de reportes y cuadros de mando.

#### Materiales y Métodos

##### Evaluación de herramientas de BI

A continuación, se analizan y comparan Oracle vs Pentaho:

Las dos herramientas comparten un crecimiento tecnológico en los últimos años. Así por ejemplo, Oracle BI es una herramienta muy útil para el desarrollo de prototipos rápidos y puede extenderse a informes, análisis, data mining o bigdata. Por su parte, Pentaho, ofrece una replicación completa de la Suite BI de Oracle, pero basado en Open Source, siguiendo planteamientos tradicionales como la metodología de Kimball e Inmon. Por otro lado, Pentaho, posee una ventaja en cuanto al precio de licenciamiento, pero hay que tener en cuenta que la diferencia no es tan amplia, si se habla de la totalidad del desarrollo del proyecto, ya que puede resultar en un costo más alto, evaluando el personal que conozca la herramienta, capacitación y cantidad de desarrollo. En relación a la interfaz de Oracle BI es más elegante y sencilla de usar. En relación a los reportes, Pentaho, no consigue informes a gran nivel de detalle, para grandes listados de información. Así mismo, Pentaho, no dispone de un buen sistema de versionado y control del código fuente. Tanto Oracle BI como Pentaho, son accesibles y disponen de visualización desde iPad y Android. Tanto Oracle BI como Pentaho pueden ser implementadas y usadas en semanas, dependiendo de la complejidad y volúmenes de datos [18]. Para el desarrollo del presente proyecto se ha determinado el uso de herramientas de la línea de Oracle por las razones descritas y por poseer un licenciamiento gratuito otorgado por una empresa analítica, como se muestra en la Tabla 3. A continuación se describe las herramientas de software empleado. En el siguiente apartado se detallan las metodologías empleadas y el sustento de su selección:

Tabla 3: Software empleado en el Sistema de Gestión de Consultoría

|  |  |
| --- | --- |
| SOFTWARE | DESCRIPCIÓN |
| Oracle Database SE One 11g R2 | La base de datos líder del mundo incluye características propias para un data warehouse [19]. |
| Oracle Warehouse Builder 11g | Una de las herramientas líderes en procesos de ETL (“extraer, transformar y cargar”) permitiendo construir y mantener un data Warehouse eficiente y de alta calidad [21]. |
| Oracle BI Server SE One 11g | El servidor de BI es un poderoso servidor de análisis y consulta capaz de integrar múltiples fuentes de datos heterogéneos en una visión única y simplificada de información. Soporta el acceso directo a fuentes de datos Oracle y no Oracle [21]. |
| Oracle Dashboards 11g | Oracle BI Interactive Dashboards brinda una interfaz Web personalizada, basada en roles, 100% para los usuarios finales, con inclusión de las visualizaciones en forma de estimaciones, cuadros, informes de resumen e incluso análisis basados en condiciones [21]. |
| Oracle Answers 11g | Oracle BI Answers es una solución de análisis e informes estáticos, totalmente integrada con Interactive Dashboards. Los usuarios finales pueden crear rápidamente sus propios informes, luego realizar desgloses, analizar, visualizar e incorporar los resultados en sus propios tableros de control personalizados [21]. |
| Oracle BI Server Administrator 11g | Es un sistema de repositorio digital, de código abierto, que permite recoger material digital, distribuirlos sobre la Web a través de un sistema de búsqueda y recuperación. [21]. |
| Oracle Application Express 4.2 | Es un entorno de desarrollo orientado a Web de 4ta generación, que permite la creación de aplicaciones bajo un marco de desarrollo rápido [22]. |

##### Metodología XP para el Módulo de Registro de Horas

Para el desarrollo de este módulo, se seleccionó la metodología XP, basándose en los criterios de análisis que se muestran en la Tabla 4. En esta evaluación XP, obtiene una mejor puntuación sobre el resto de metodologías ya que presenta mejores características de adaptabilidad según las exigencias del desarrollador y los requerimientos del cliente; y sobre todo porque propone un manejo flexible de la documentación, motivo por el cual se la seleccionó como un marco de referencia de buenas prácticas adaptables al proyecto.

Tabla 4: Matriz de comparación de Metodologías para el Módulo de Registro de Horas

|  |  |  |  |
| --- | --- | --- | --- |
| Criterios: Alto=3, Medio=2, Bajo=1, Nulo=0 | Metodologías | | |
| XP | RUP | MFS |
| Pocos roles y flexibilidad | 3 | 1 | 2 |
| Iterativa | 3 | 1 | 1 |
| Entorno amplio de proyectos de Software | 2 | 2 | 2 |
| Adaptable sobre cualquier tecnología | 2 | 2 | 2 |
| Nivel de trabajo en equipo | 2 | 1 | 1 |
| Afinidad con el sistema | 3 | 2 | 2 |
| Adaptable a cambios | 3 | 2 | 2 |
| Comunicación con el cliente | 3 | 2 | 2 |
| TOTAL: | 21 | 13 | 14 |

##### Metodología Ralph Kimball para el Módulo de Gestión de Consultoría

Para la selección de la metodología a seguir en este módulo se tomaron en cuenta dos características. La primera, es que la empresa analítica cuenta con un licenciamiento gratuito de las herramientas, y porque dentro de la suite de Oracle Business Intelligence (OBI), se emplea de forma embebida la metodología de Ralph Kimball. Además, se realizó un análisis comparativo entre la metodología de Bill Inmon y Ralph Kimball, obteniendo como resultado, que la metodología de Kimball, se adapta mejor a las necesidades del modelamiento de BI; enfatizando en la creación de un datawarehouse especializado para el proceso de Gestión de Consultoría, y no un data warehouse global donde se involucran todos los procesos de la empresa [17], como propone Inmon. La Tabla 5, muestra la comparación entre Kimball Vs Inmon.

Tabla 5: Matriz de comparación de Kimball vs Inmon

|  |  |  |
| --- | --- | --- |
| Especificación | Inmon | Kimball |
| Generalización | General a detalle | Detalle a general |
| Arquitectura orientado a | Datawarehouse | Datamart |
| Complejidad de implementación | Compleja | Simple |
| Usabilidad para el usuario | Baja | Alta |
| Orientado a | Orientado a temas | Orientado a procesos |
| Modelamiento | Tradicional | Dimensional |
| Esquemas de modelamiento | Normalizados | Desnormalizados |
| Manejo de cambios en dimensiones | Continuo y discreto | Dimensiones cambiantes |
| Dirigido a | IT | Usuarios finales |
| Tiempo de desarrollo | Largo plazo | Corto y mediano plazo |
| Ayuda a la toma de decisiones | Estratégicas | Tácticas |
| Flexibilidad | Baja | Alta |
| Costo de implementación | Alto | Bajo |
| Equipo de desarrollo | Especialistas | Generalistas |

**Utilidad para el proyecto de tesis**

Del artículo se tomará como referencia el uso de la metodología de Ralph Kimball, implementando el análisis y modelado dimensional estipulado en esta, para de esta manera tener un repositorio de datos especializado para nuestra investigación.

### Modelamiento dimensional en competencias TIC (Martha Mendoza, Jaime Mendoza, Deyro Zuñiga, Jorge Moreno, Carlos Cobos, 2013) (ISSN: 1794-1237)

Los almacenes de datos son sistemas que permiten homogeneizar e integrar datos de organizaciones desde diferentes fuentes de datos ya sean internas o externas, y están basados en una representación simple y detallada de la empresa, la cual permite extraer información útil para la toma de decisiones.

A pesar de que no existe un estándar para los la terminología del modelo multidimensional, a continuación los autores presentaran nociones clave muy populares. Los principales conceptos del modelado son los hechos y las dimensiones, los cuales deben estar claramente diferenciados para facilita la compresión y análisis del modelo multidimensional. Al tomar un hecho con sus dimensiones relacionadas se obtiene lo que se denomina como **esquema multidimensional o estrella**. Una dimensión es un concepto que se usa como una perspectiva de análisis de los hechos, los cuales normalmente se descomponen de jerarquías, niveles y descriptores:

* Las jerarquías de dimensiones son gradaciones de niveles que representan diferentes niveles de detalle de datos.
* Los descriptores son atributos de cada uno de los niveles de la dimensión.

En la Ilustración 6 se aprecia un esquema estrella para el hecho de ventas, cuyas dimensiones son: fecha, cliente, producto y almacén. Además, la dimensión fecha presenta una jerarquía compuesta por los niveles: año, semestre, mes y día; y un atributo descriptor que es la fecha completa. Los hechos también contienen medidas de análisis, (Romero and Abelló 2009; Kimball, et al., 1998). En la tabla de hechos de Ventas, se muestran las claves foráneas de las dimensiones y medidas relacionadas con la venta: Cantidad\_Ordenada y Valor\_Venta.

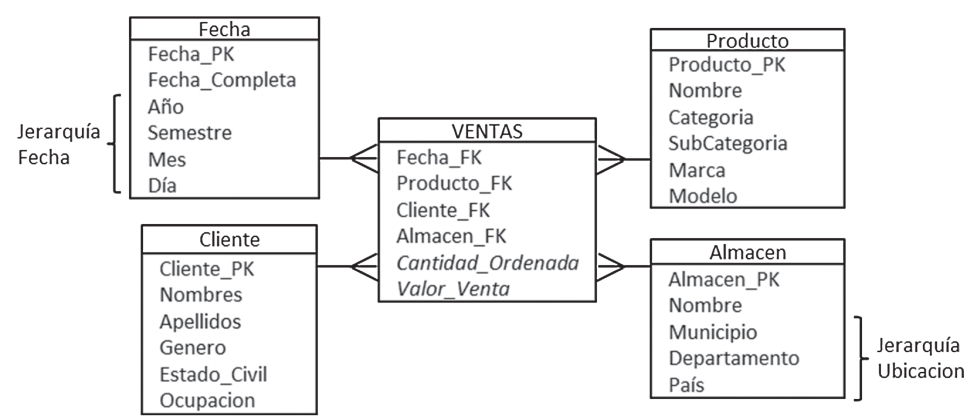


Ilustración 8: Diagrama estrella para ventas

Un aspecto muy importante en el desarrollo de una bodega de datos es el proceso de extracción, transformación y carga (ETL, por sus siglas en inglés, Extraction, Transform and Load) (Kimball, et al., 1998); que busca extraer los datos relevantes desde las fuentes de datos (archivos planos, bases de datos relacionales, etc.), aplicarles un proceso de transformación para homogeneizarlos (también se pueden generar nuevos datos calculados) y finalmente cargarlos o almacenarlos en la bodega de datos. Este proceso se presenta en dos tiempos: una carga inicial que permite que los datos sean cargados desde las fuentes de datos, y posteriormente, cargues incrementales que toman solo los datos nuevos desde el último cargue.

Las herramientas de procesamiento analítico en línea (OLAP, por sus siglas en inglés, Online Analytical Processing) se han utilizado por años para consultar las bodegas y han mostrado ser apropiadas para que los usuarios exploren en forma adecuada los datos almacenados en ellas. Las herramientas OLAP se basan en una vista multidimensional de los datos, definiendo un modelo «sencillo» de diseño comprensible y fácil de usar por parte de los usuarios que toman las decisiones en las organizaciones (Romero and Abelló 2009; Kimball, et al., 1998). Entre las operaciones básicas de OLAP se encuentran: (i) slice, selección de las dimensiones para generar una vista del cubo; (ii) dice, selección de valores sobre una dimensión; (iii); drill-down, que disminuye el nivel de agregación a través de una o más jerarquías de dimensión; (iv) roll-up, que permite incrementar el nivel de agregación; y (v) pivot, rota las dimensiones para proveer una presentación alternativa de los datos. Las bodegas de datos, como parte de un sistema de soporte a la toma de decisiones (DSS, por sus siglas en inglés Decision Support Systems), han sido utilizadas principalmente como apoyo al análisis de grandes cantidades de datos históricos en entidades comerciales y financieras (Kimball, et al., 1998). Por ejemplo, para definir estrategias de mercadeo, identificar patrones del comportamiento de los clientes y definir el impacto de algún producto novedoso. Actualmente, las bodegas se están empleando con más frecuencia en contextos diferentes a los negocios, que también necesitan el análisis de gran cantidad de datos, como, en la gestión de información geográfica (Kyung, et al., 2012; Glorio, et al., 2012), la investigación bioinformática (Bustos, et al., 2011), las misiones espaciales (Liu, et al., 2012), problemas de diseño en agricultura (Nilakanta, et al., 2008), información clínica de cuidado intensivo (De Mul, et al., 2012), análisis de la interacción de los estudiantes en un sistema gestor de aprendizaje (Mendoza, et al., 2006) y en la gestión de las competencias requeridas por profesionales en el mercado laboral (Georgescu and Sbughea, 2012). Debido al auge de las tecnologías de la información y comunicación (TIC), cada vez es más importante su incorporación al sistema educativo, permitiendo el acceso a los recursos digitales (UNESCO, 2004). Por esto, los gobiernos, entre ellos los países latinoamericanos, están desarrollando programas que permitan que los niños y jóvenes que asisten a escuelas y colegios tengan acceso a las TIC. Para verificar la eficacia de estos programas es necesario medir y analizar las competencias en TIC adquiridas por los integrantes de estas comunidades educativas. En este artículo se presenta un modelo dimensional de bodega de datos para competencias en TIC de docentes y estudiantes de instituciones educativas que permite apoyar el proceso de toma de decisiones de los encargados de estos programas del gobierno. Para realizar el modelamiento de la bodega de datos, se tomó como referente el conjunto de casos de diseño propuestos por Ralph Kimball, (Kimball, et al., 1998). Sin embargo, cuando se realizó el modelamiento dimensional en el contexto específico (competencias en educación), se encontraron retos en la adaptación de los casos de diseño y apareció un nuevo caso no contemplado previamente en la literatura. De otra parte, la investigación que modela las competencias requeridas por profesionales en el mercado laboral en una bodega (Georgescu and Sbughea, 2012), presenta casos de diseño muy sencillos y, por lo tanto, no aporta significativamente al desarrollo del proyecto en este aspecto. Además, en este artículo se presenta un escenario de ejemplo, mostrando los componentes principales de un sistema de soporte a la toma de decisiones desarrollado para el Programa Computadores para Educar (CPE) (MINTIC, 2012) en alianza con la Universidad del Cauca, el cual puede ser replicado en otros programas con objetivos similares. Este sistema busca medir el desarrollo de habilidades y desempeños, expresados en competencias adquiridas por estudiantes y docentes en las TIC, soportado por una bodega de datos. A continuación, en la sección 2 se presentan los casos de diseño identificados en el modelado dimensional de competencias en TIC y el nuevo caso de diseño encontrado. En la sección 3 se presenta el sistema de soporte a la toma de decisiones para el programa CPE-UniCauca y, finalmente, en la sección 4 se presentan las conclusiones de la investigación y el trabajo futuro.

**Utilidad para el proyecto de tesis**

Del artículo se tomará como referencia la aplicación del esquema estrella, implementando el modelado de nuestro datamart de acuerdo a este esquema, para de esta manera agilizar el acceso a los datos.

### Implementación de un Datamart como una solución de Inteligencia de Negocios para el área de logística de T-Impulso (Julio Yalan Castillo, Luis Palomino Paniora, 2012) (ISSN: 1815-0268)

#### Introducción

El trabajo presentado por los autores se implementa un DataMart como una solución para aplicar Inteligencia de Negocios en el área de Logística de T-IMPULSO, empresa encargada de brindar servicios de Recursos Humanos principalmente. Esta propuesta nace a raíz de la necesidad de explotar la información de forma eficiente para tener una rápida, acertada y oportuna toma de decisiones bajo el manejo de datos confiables.

#### Planteamiento metodológico

El problema principal es el actual proceso de extracción e integración de los datos para la realización de los reportes, el cual se realiza manualmente, resultando muy complicado, provocando retraso en la atención de requerimientos haciendo que en muchos casos la entrega de la información no sea oportuna; por otro lado la elaboración de algunos reportes requiere mayor conocimiento en el manejo de las herramientas informáticas, por lo que son derivados al área de Tecnología de Información, ocasionando más pérdida de tiempo.

#### Marco referencial

##### Inteligencia de negocios

Este concepto no es un resultado de desarrollos en el mundo de las Ciencias Administrativas, sino que es un producto del progreso de la informático de la denominada "infotecnología".[23]

Tampoco es un concepto nuevo originado ya que su origen data de la publicación en el IBM Journal de 1958, del artículo de Hans Peter Luhn titulado "A Business Intelligence System" dond ese define con detalle el concepto con una perspectiva, que solo en nuestros días, ha sido posible su plena utilización.[23]

##### Data Warehouse

Es una base de datos corporativa que se caracteriza por integrar y depurar información de una o más fuentes distintas, para luego procesarla permitiendo su análisis desde infinidad de perspectivas y con grandes velocidades de respuesta. La ventaja principal de este tipo de bases de datos radica en las estructuras en las que se almacena la información (modelos de tablas en estrella, en copo de nieve, cubos relacionales... etc.). Este tipo de persistencia de la información es homogénea y fiable, y permite la consulta y el tratamiento jerarquizado de la misma (siempre en un entorno diferente a los sistemas operacionales).

##### Data mart

Es una base de datos departamental, especializada en el almacenamiento de los datos de un área de negocio específica. Se caracteriza por disponer la estructura óptima de datos para analizar la información al detalle desde todas las perspectivas que afecten a los procesos de dicho departamento. Un datamart puede ser alimentado desde los datos de un data warehouse, o integrar por sí mismo un compendio de distintas fuentes de información. [24]

#### Estado del arte metodológico

##### Posible s herramientas a utilizar

###### SSIS

El software SQL Server Integration Services (SSIS) [SQL 2007], permite la integración de los datos de cualquier fuente. SSIS provee una plataforma escalable y extendible que capacita al equipo desarrolladora construir, mantener, y desplegar soluciones de integración para alcanzar soluciones de integración únicas de acuerdo a las necesidades.

###### Data Stage

DataStage [DAT2007] es una herramienta que permite soportar la información que necesita la compañía, y construir un Data Warehouse en "tiempo real". El DataStage es una herramienta ETL que utiliza notación gráfica para construir integración de datos para dar soluciones, y está disponible en varias versiones, como Server Edition y Enterprise Edition.

###### SUNOPSIS

Sunopsis [SUN 2007] ofrece un alto desempeño y una integración efectiva, cubriendo las necesidades de integración.

###### Microstrategy

MicroStrategy Business Intelligence Solutions permite mejorar y predecir el comportamiento del negocio, poniendo información en las manos de toda persona de negocios en la empresa.

###### Cognos

Cognos 8 Business Intelligence [COG 2007] es una plataforma del grupo IBM que permite la generación y visualización de reportes, cubos, dashboards y Balance scorecards, además de la gestión de permisos y usuarios necesaria para la implementación de la plataforma

###### Bussiness Objects

Bussiness Objectses se caracteriza por ofrecer distintas funcionalidades según el tamaño de la empresa que la adquiere y la licencia.

###### Pentaho

Pentaho incluye todos los principales componentes requeridos para implementar soluciones basadas en procesos. Estas se componen fundamentalmente de una infraestructurade herramientasde análisis e informes integrado con un motor de workflow de procesos de negocio.[25]

###### Octopus

Octopus es, al igual que Pentaho,una herramienta libre pero sólo se centra en los procesos ETL. Está basada en Java y por lo tanto se puede conectar a cualquier fuente JDBC.

##### Definición de variables de comparación

En la Tabla 6 se muestra la definición y ponderación de las variables de comparación para las herramientas antes mencionadas.

Tabla 6: Definición y ponderación de variables de Comparación de herramientas BI

|  |  |  |
| --- | --- | --- |
| Variables | Ponderación | Definición |
| ETL | 1 | Permite la Extracción, Transformación y Carga. |
| Elaboración de reportes | 1 | Permite la elaboración de reportes |
| Open source | 1 | El código fuente es de código abierto |
| Conocimiento de herramienta | 2 | Tiene una ponderación de 2 porque consideramos una ventaja el manejo de la herramienta para realizar el proyecto |

##### Criterio de puntuación

En la Tabla 7 se muestra una puntuación que se asigna al valor de las variables.

Tabla 7: Cuadro de puntuación de las herramientas BI

|  |  |
| --- | --- |
| Criterios de puntuación | Puntaje |
| SI | 1 |
| NO | 0 |

##### Puntuación de las herramientas

En la Tabla 8 se aprecia que SQL Server Integration Services obtiene el mayor puntaje. Tomando como base las características de estas herramientas y el puntaje obtenido se optó por utilizar la herramienta SSIS (SQL Server Integration Services)

Tabla 8: Cuadro comparativo de herramientas BI

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Herramienta | ETL | Elaboración de reportes | Open source | Conocimiento de herramienta | Puntuación |
| SSIS | 1 | 1 | 0 | 2 | 4 |
| Data Stage | 1 | 0 | 0 | 0 | 1 |
| Sunopsis | 1 | 0 | 0 | 0 | 1 |
| Micro Strategy | 0 | 1 | 0 | 0 | 1 |
| Bussiness Objects | 0 | 1 | 0 | 0 | 1 |
| Cognos | 0 | 1 | 0 | 0 | 1 |
| Octopus | 1 | 0 | 0 | 0 | 1 |
| Pentaho | 1 | 1 | 1 | 0 | 3 |

**Utilidad para el proyecto de tesis**

De la tesis se tomará como referencia la utilización de SSIS, aprovechando las facilidades de manejo de datos de esta herramienta para la construcción de los procesos de extracción, transformación y carga de datos.

# REFERENCIAS BIBLIOGRÁFICAS

[1] Rouhani,S (2012). Review Study: Business Intelligence Concepts and Approaches

[2] Kimball, R. (2004) .The Data Warehouse ETL Toolkit

[3] Kimball, R. (2013) .The Data Warehouse Toolkit (3er edition)

[4] Ranjan, J. (2009) Business Intelligence: Concepts, Components, Techniques and Benefits, Vol 9. No 1. (pp 060 - 070)

[5] Madhuri, V. (2013). Data Mining and Business Intelligence: Applications in Telecommunication Industry

[6] Gamarra, L. (2013) .Diseño e implementación de una aplicación móvil para la presentación de estadísticas del módulo de incidentes de un Sistema de Gestión de Servicios.

[7] Ishaya, T & Folarin, M. (2012) .Business Intelligence in Telecoms Industry: A Service Oriented Approach.

[8]Bustamante, R. (2009) .Análisis, Diseño e Implementación de un Sistema de Administración de Incidente en atención al cliente para una empresa de telecomunicaciones.

[9]Ramírez, P. (2011).Inteligencia de negocios y toma de decisiones de los gerentes en la banca universal en Venezuela

[10]Gamarra, A. (2011). Solución integral para explotar eficientemente la información de los contactos con los clientes utilizando Datamart en Telefónica del Perú.

[11] La inteligencia de negocios aplicada a las organizaciones en Latinoamérica.

[12] El impacto de las herramientas de inteligencia de negocios en la toma de decisiones de los ejecutivos.

[13] Análisis, Diseño e Implementación de un sistema gerencial basado en una suite integrada de DataMarts para las áreas de finanzas, contabilidad, recursos humanos y comerciales.

[14] Análisis, Diseño e Implementación de un DataMart para el área de mantenimiento y logística de una empresa de transporte público de pasajeros.

[15] The Role of Business Intelligence in Organizational Decision-making (Arisa Shollo)

[16] Inteligencia de Negocios y Toma de Decisiones (Dr. Ramiro Rollano PhD).

[17]. Espinosa, R. (2010). Kimball vs Inmon. Aplicación de conceptos del modelo dimensional. Recuperado el 16 de noviembre, de http://churriwifi.wordpress.com/2010/04/19/15-2- ampliacion-conceptos-del-modelado-dimensional/

[18]. Pentaho. (2013). Pentaho Business Analytics. Recuperado el 06 de marzo de 2013, de <http://www.pentaho.com/>

[19]. Oracle. (2013). Oracle Database. Recuperado el 23 de febrero de 2013, de http://www.Oracle.com/technetwork/database/enterprise-edition/downloads/index.html

[20]. Oracle. (2013). Oracle Warehouse Builder. Recuperado el 23 de febrero de 2013, de <http://www.Oracle.com/technetwork/developertools/warehouse/downloads/software/index.html>

[21]. Oracle. (2013). Oracle Business Intelligence. Recuperado el 19 de febrero de 2013, de <http://www.Oracle.com/technetwork/middleware/bi-enterprise-edition/overview/index.html>

[22]. Oracle. (2013). Oracle Application Express. Recuperado el 23 de febrero de 2013, de http://apex.Oracle.com/i/index.html

[23] Conceptos y tecnología de inteligencia de negocio, Agosto 2005, de <http://www.bi-spain.com>

[24] Productos de Oracle para Inteligencia de Negocios, Julio 2008, de http://www.oracle.com/technology/ products/bilindex.html

[25] Página oficial de Pentaho http://www.pentaho. com