

INSTITUTO TECNOLÓGICO Y DE ESTUDIOS

SUPERIORES DE MONTERREY

CAMPUS MONTERREY

PROGRAMA DE GRADUADOS EN MECATRÓNICA Y TECNOLOGÍAS DE  
INFORMACIÓN



**TECNOLOGICO  
DE MONTERREY®**

*Factores Críticos de Éxito para la Implementación del Data Warehouse en las  
Empresas del Sector Financiero de Monterrey*

**TESIS**

PRESENTADA COMO REQUISITO PARCIAL PARA OBTENER EL GRADO

ACADÉMICO DE:

MAESTRO EN ADMINISTRACIÓN DE TECNOLOGÍAS DE INFORMACIÓN

POR:

*Isaías Velázquez Reyes*

MONTERREY, N. L.

Noviembre de 2010

FACTORES CRÍTICOS PARA LA IMPLEMENTACIÓN DEL DATA  
WAREHOUSE EN LAS EMPRESAS DEL SECTOR FINANCIERO DE  
MONTERREY



**TECNOLOGICO  
DE MONTERREY®**

TESIS POR:

*Isaías Velázquez Reyes*

PRESENTADA AL PROGRAMA DE GRADUADOS EN MECATRÓNICA Y  
TECNOLOGÍAS DE INFORMACIÓN

ESTE TRABAJO ES REQUISITO PARCIAL PARA OBTENER EL GRADO DE  
MAESTRÍA EN ADMINISTRACIÓN DE TECNOLOGÍAS DE INFORMACIÓN

INSTITUTO TECNOLÓGICO DE ESTUDIOS SUPERIORES DE MONTERREY

MONTERREY, N. L.

Noviembre de 2010

## **Dedicatoria**

### **A mi padre...**

Por su apoyo incondicional durante todo mi proyecto de vida, incluida esta investigación. Sin ti no hubiera sido posible la culminación de mis estudios.

### **A mi madre...**

Por ese apoyo que cualquier hijo desearía. Siempre presente en los buenos y malos tiempos, criando a todos tus hijos por el buen camino.

### **A mí prometida Erika...**

Que con sus palabras y dedicación impulsan mi voluntad para exigirme superar mis propios límites.

### **A mis hermanos Uriel y Lysette...**

Cada uno de su manera peculiar hizo entretenido este proceso de vida.

### **A mi hermano Isaac...**

Donde quiera que se encuentre, tengo la seguridad de la satisfacción que le causa este logro.

## **Agradecimientos**

### **A mi asesor...**

Dr. David Alanís, por todo el tiempo dedicado al proyecto y su experiencia convirtió al estudio en producto de valor. Muchas gracias por su apoyo.

### **A mis sinodales...**

Dr. Macedonio y Dr. Miguel Ángel por el esfuerzo dedicado a este estudio. Gracias.

### **Al Lic. Rico Salomón...**

Que aunque tuve el gusto de conocerle una ocasión, su asesoría fue muy importante para encontrar a los encuestados.

### **A los encuestados...**

Por hacer posible obtener estos resultados de alta calidad.

## Resumen

La fuerte competencia empresarial exige tener información confiable y acertada. De igual manera en el Sector Financiero, según Hwang, Ku, Yen y Cheng (2004), la tecnología tiene un rol crucial para mejorar la calidad de sus servicios y proveer de una ventaja competitiva. Dada esta necesidad, el *data warehouse* es una tecnología con la capacidad de generar una sola versión de la verdad (Turban, Sharda, Aronson & King, 2008), para así mejorar su capacidad productiva, reducir costos y mejorar el apoyo a la toma de decisiones (Feinberg, 2009a).

Una vez definido el panorama se optó por realizar la presente investigación, con el objetivo de identificar y determinar los factores críticos al momento de la implementación de la tecnología del *data warehouse* en las empresas del Sector Financiero; siendo los objetivos específicos reportar evidencia empírica del nivel de avance de la implementación, y realizar una serie de recomendaciones.

Se seleccionaron 16 empresas del Sector Financiero, enfocando el estudio a sólo aquellas que tuvieran implementado un *data warehouse*. Al final de la investigación se identificaron veinte factores críticos, siendo la inversión en capacitación, cambio en la forma de trabajo, participación de personal interno y externo, consistencia e integración de los datos, inversión en infraestructura tecnológica, y capacidad de integración con el resto de los sistemas, y los aspectos de mayor relevancia para lograr una implementación exitosa. Este estudio es importante para que las organizaciones consideren las variables con un efecto positivo o negativo en esta clase de proyectos, además de servir como referencia a empresas de otros sectores con planes de incorporar al *data warehouse* a su estrategia de negocios.

## Tabla de contenidos

Contenido de la primera página.....	i
Contenido de la segunda página.....	ii
Contenido de la tercera página.....	iii
Dedicatoria.....	iv
Agradecimientos.....	v
Resumen.....	vi
Tabla de contenidos.....	vii
Lista de figuras .....	x
<b>1. Introducción.....</b>	<b>1</b>
1.1. Definición del problema .....	1
1.2. Objetivo de la investigación .....	3
1.3. Justificación de la investigación .....	3
1.4. Limitaciones.....	4
1.5. Producto final de la investigación .....	5
1.6. Estructura del documento.....	5
<b>2. Conociendo el <i>data warehouse</i>.....</b>	<b>7</b>
2.1. Historia.....	7
2.2. Definición del <i>Data Warehouse</i> .....	11
2.3. Beneficios del <i>Data Warehouse</i> .....	12
2.4. Problemas del <i>Data Warehouse</i> .....	15
2.5. Tendencias del <i>Data Warehouse</i> .....	16

2.6.	Características del <i>Data Warehouse</i> .....	23
2.7.	La Inteligencia de Negocios y el <i>Data Warehouse</i> .....	24
2.8.	Arquitectura del <i>Data Warehouse</i> .....	28
2.9.	Elementos Importantes de la Arquitectura.....	32
2.9.1.	Procesos básicos realizados por un <i>data warehouse</i> .....	33
2.9.2.	Procesamiento analítico en línea.....	36
2.9.3.	<i>Data marts</i> .....	39
2.9.4.	<i>Data mining</i> .....	42
3.1.	Metodología de Desarrollo en Espiral y el Enfoque de los Siete Flujos .....	50
3.2.	Metodología del Ciclo de Vida Dimensional del Negocio.....	53
3.3.	Metodología de Cascada Bajo el Enfoque de Rainardi.....	55
<b>4.</b>	<b>Factores críticos a considerar en la implementación de un <i>data werehouse</i>.....</b>	<b>58</b>
4.1.	Enfoque Hwang, Ku, Yen, y Cheng.....	59
4.2.	Enfoque Ponniah .....	64
4.3.	Enfoque Wixom y Watson.....	65
<b>5.</b>	<b>Proveedores de herramientas de <i>data warehouse</i>.....</b>	<b>68</b>
<b>6.</b>	<b>Casos de éxito.....</b>	<b>75</b>
6.1.	Caso Banco de Baroda.....	75
6.2.	Caso Compañía de Servicios Financieros.....	77
<b>7.</b>	<b>Investigación de campo .....</b>	<b>81</b>
7.1.	Metodología .....	81
7.1.1.	Tipo de investigación .....	81
7.1.2.	Población .....	81
7.1.3.	Selección de muestra.....	82

7.2. Herramienta de Investigación.....	83
7.3. Modelo Particular.....	83
7.4. Variables de la Investigación.....	84
7.5. Estrategia para la Recolección de Datos .....	86
<b>8. Análisis de resultados .....</b>	<b>87</b>
8.1. Nivel de Implementación .....	87
8.2. Aspectos económicos .....	91
8.3. Aspectos sociales .....	97
8.4. Aspectos tecnológicos.....	109
8.5. Factores críticos para la implementación del <i>data warehouse</i> en el Sector Financiero.....	118
8.6. Recomendaciones para la implementación exitosa del <i>data warehouse</i> en el Sector Financiero .....	123
<b>9. Conclusiones y trabajos futuros.....</b>	<b>131</b>
9.1. Conclusiones Generales.....	131
9.2. Trabajos Futuros.....	134
<b>Apéndices.....</b>	<b>137</b>
Apéndice 1. Formato de la encuesta .....	137
<b>Referencias .....</b>	<b>148</b>
<b>Vita .....</b>	<b>153</b>

## Lista de figuras

Figura 2.1 Prioridades de tecnologías de información en 2009.....	17
Figura 2.2 ¿Cuál de los siguientes productos de DBMS tu organización utiliza actualmente o planea utilizar en los próximos 12 meses? .....	19
Figura 2.3 Organizaciones de TI que ya están utilizando <i>software</i> de código abierto....	22
Figura 2.4. Porciones estratégicas y tácticas de ambiente de inteligencia de negocios ..	27
Figura 2.5 Componentes del <i>Data Warehouse</i> .....	29
Figura 2.6 Marco comparativo de técnicas del <i>data mining</i> .....	48
Figura 3.1 Mapa de una proyecto de <i>data warehouse</i> basado en el enfoque de los siete flujos .....	51
Figura 3.2 Diagrama del ciclo de vida dimensional del negocio .....	53
Figura 3.3 Metodología de cascada con el enfoque de proyectos de <i>data warehouse</i> ....	55
Figura 4.1 Modelo de investigación.....	60
Figura 5.1 <i>Magic Quadrants</i> .....	69
Figura 5.2 <i>Magic Quadrant</i> para los DBMS de <i>Data Warehouse</i> del 2008.....	71
Figura 7.1. Fórmula para obtener la muestra probabilística .....	82
Figura 7.2. Modelo de Investigación.....	84
Figura 8.1 Sistemas de información existentes en las organizaciones .....	88
Figura 8.2 Nombre de la herramienta de <i>data warehouse</i> que utilizan.....	89

Figura 8.3 Etapa en que se encuentra la implementación del <i>data warehouse</i> .....	90
Figura 8.4 Importancia de la implementación de un <i>data warehouse</i> .....	90
Figura 8.5 Grado de importancia que tiene la información obtenida del <i>data warehouse</i> en las actividades de enfoque al negocio .....	91
Figura 8.6 Necesidad de inversión en capacitación al momento de la implementación.	93
Figura 8.7 Suficiencia de recursos al momento de la implementación .....	94
Figura 8.8 Grado de conformidad del costo de la implementación .....	94
Figura 8.9 Duración de la implementación conforme a lo planeado .....	95
Figura 8.10 Tiempo que duró la implementación del <i>data warehouse</i> .....	96
Figura 8.11 Análisis de costo/beneficio para definir la suficiencia de los recursos para soportar la implementación .....	97
Figura 8.12 Interés y compromiso de los niveles directivos .....	98
Figura 8.13 Apoyo recibido de la administración en las actividades del proyecto.....	99
Figura 8.14 Apoyo recibido de la administración en las etapas del proyecto .....	100
Figura 8.15 Grado de importancia en las actividades para elaborar expectativas realistas .....	102
Figura 8.16 Etapa del proyecto de implementación en la que se involucró a los usuarios .....	103
Figura 8.17 Cantidad de personas involucradas en el proyecto del <i>data warehouse</i> ....	104

Figura 8.18 Tipo de personal involucrado en la implementación del <i>data warehouse</i> .	105
Figura 8.19 Habilidades y la experiencia del personal involucrado en la implementación del <i>data warehouse</i> .....	106
Figura 8.20 Beneficios proporcionados por el <i>data warehouse</i> .....	107
Figura 8.21 Situaciones relacionadas con la resistencia al cambio durante la implementación .....	108
Figura 8.22 Cambios e inversiones en infraestructura tecnológica para el buen desempeño del <i>data warehouse</i> .....	110
Figura 8.23 Magnitud de la inversión en <i>software</i> y/o <i>hardware</i> .....	111
Figura 8.24 Alineación del <i>data warehouse</i> con la visión, misión, objetivos y estrategias .....	112
Figura 8.25 Compatibilidad del <i>data warehouse</i> con el resto de los sistemas de la organización .....	113
Figura 8.26 Estrategia de implementación del <i>data warehouse</i> .....	114
Figura 8.27 Presencia de una estrategia para la selección del proveedor del <i>data warehouse</i> .....	115
Figura 8.28 Presencia de un tercero o intermediario para apoyar al proceso de selección del proveedor del <i>data warehouse</i> .....	116
Figura 8.29 Criterios para la selección del proveedor de <i>data warehouse</i> .....	117
Figura 8.30 Atención/servicio del proveedor de <i>data warehouse</i> .....	118

## 1. Introducción

### 1.1. Definición del problema

Cada organización se basa en la información para realizar estrategias empresariales y de esta manera poder funcionar, crecer y transformar sus negocios; la obtienen de manera manual o computarizada, con diferente nivel de detalle y distintas fuentes. Las organizaciones necesitan identificar, comprender y administrar estas fuentes de información; además de diseñar, desarrollar, y realizar pruebas de despliegue de sus procesos y servicios (Blechar & Thompson, 2009).

Debido a la fuerte competencia empresarial, es necesario tener información confiable y acertada para la toma de decisiones; sin embargo, las bases de datos tradicionales han sido incapaces de responder a las demandas de extracción, acceso, actualización y mantenimiento de la información debido a su crecimiento exponencial. Esta incapacidad ha impactado en los negocios, al no permitir a la alta gerencia utilizar la información de forma efectiva, eficiente y oportuna; es por eso que la administración masiva de la información se ha convertido en un factor crítico para los negocios (Hwang, Ku, Yen & Cheng, 2004).

Dada la necesidad de contar con un sistema de base de datos con alta flexibilidad y adaptabilidad, la industria y la academia han sugerido como solución la tecnología del *data warehouse* (Hwang et al., 2004). Actualmente la mayoría de las herramientas de *data warehouse* son utilizadas para reportar o integrar datos, para la inteligencia de negocios y mejorar la relación con el cliente (Rainardi, 2008).

En particular el Sector Financiero se ha caracterizado por las altas demandas de procesamiento de información, lo cual ha tomado un rol clave para mejorar la calidad de los servicios y obtener una ventaja competitiva (Hwang, Ku, Yen & Cheng, 2004).

Estudios e informes de la industria han señalado frecuentemente fallas en las Tecnologías de Información (TI), así como altos costos. Se estima que los proyectos de informática registran índices de fracaso del 30%, y un rango de 46-53% en proyectos con problemas (Goldfinch, 2007; Standish Group, 2001).

La implementación de sistemas de TI requiere que las empresas sean dinámicas para cambiar y capaces de adoptar nuevas tecnologías a los procesos del negocio (Rivard, 2006; Tarafdar & Gordon, 2007). Por consiguiente, cuando se publica el fracaso de la implementación, se envía una señal al mercado acerca de la incapacidad de la organización para utilizar un recurso tecnológico (Bharadwaj, Keil & Mähring, 2009).

Las inversiones en sistemas de TI representan opciones de crecimiento para la empresa (Benaroch, 2002) y por tanto, un fracaso en la implementación tiene a menudo implicaciones a corto, como a largo plazo. En el corto plazo, una empresa puede incurrir en costos adicionales durante la ejecución del proyecto; mientras en el largo plazo, se afectan las partes críticas de la infraestructura de la empresa. Hablando del caso de los sistemas de información para la toma de decisiones, cualquier implicación podría reducir la posibilidad de la empresa a entrar a nuevos mercados u ofrecer nuevos productos o servicios (Litan & Rivlin, 2001).

Considerando el panorama antes mencionado, la inversión en un *data warehouse* involucra oportunidades y riesgos para el negocio; por lo tanto, este estudio se describe como: ¿Cuáles factores económicos, sociales y tecnológicos deben considerarse para la

implementación exitosa de un *data warehouse*, que apoye una estrategia de Inteligencia de Negocios en el Sector Financiero de Monterrey?

## **1.2. Objetivo de la investigación**

El objetivo principal de la investigación consiste en la identificación y determinación de los factores críticos que deben ser tomados en consideración por las empresas del Sector Financiero de Monterrey, al momento de la implementación de la tecnología del *data warehouse*; siendo los objetivos específicos reportar evidencia empírica del nivel de avance de la implementación, y realizar una serie de recomendaciones.

## **1.3. Justificación de la investigación**

La cantidad de información que se maneja en las grandes empresas requiere ser administrada por una tecnología con la capacidad de manipularla, integrarla y explotarla; Hwang, Ku, Yen y Cheng (2004) mencionan que los sistemas de administración de datos no tienen la capacidad de cubrir dichas necesidades.

En respuesta a la problemática la tecnología de *data warehouse* resulta conveniente, ya que las organizaciones están empezando a reconocer que su manera de administrar los datos es fundamental para mejorar su capacidad productiva, reducir costos y mejorar el apoyo a la toma de decisiones (Feinberg, 2009a).

Conforme a lo mencionado por Feinberg (2009a), hoy en día el *data warehouse* y la inteligencia de negocios, se consideran áreas de misión crítica en las organizaciones. Estos se definen como sistemas que apoyan al incremento de los ingresos y/o a los procesos del negocio; su ausencia por un periodo determinado puede causar pérdidas económicas. La promoción de estos sistemas es el resultado de varios factores:

1. El incremento en el uso de la tecnología de *data warehouse*, como parte de la infraestructura de los sistemas de procesamiento de transacciones en línea (en inglés, *Online Transaction Processing*, abreviado *OLTP*)
2. Las nuevas aplicaciones para la administración de los procesos y la inteligencia de negocios
3. El uso del *data warehouse* para reducir los costos e incrementar la competitividad

Cada año, el mercado del *data warehouse* tienen nuevos participantes, que van de nuevos productos de sistemas de administración de bases de datos (en inglés, *Database Management Systems*, abreviado *DBMS*) a equipos completos de *hardware* y *software*. El mercado ha evolucionado para ofrecer soluciones a organizaciones con poca o ninguna experiencia en la tecnología. Aunque se trata de un mercado maduro en el cual los grandes fabricantes ofrecen soluciones escalables y robustas, los pequeños ofrecen soluciones con un enfoque más preciso e innovador.

Por tratarse de una tecnología con un mercado que está creciendo alrededor del 14% por año (Feinberg, 2009a), el costo de una implementación fallida resulta elevado; siendo imprescindible identificar oportunamente los factores críticos de éxito con el fin de disminuir la probabilidad de aparición de riesgos (Goldfinch, 2007).

#### **1.4. Limitaciones**

La presente investigación sólo se enfocará en los factores económicos, sociales y tecnológicos que deben ser tomados en consideración, para la implementación exitosa de un proyecto de *data warehouse*.

Durante la recolección de datos algunos encuestados cancelaron su participación en el estudio por falta de voluntad o disponibilidad, siendo necesario localizar nuevos prospectos para completar el tamaño de la muestra especificada. En ningún momento la investigación buscó elaborar una guía técnica para el proceso de implementación.

### **1.5. Producto final de la investigación**

Es un documento que presente los factores críticos identificados, una serie de recomendaciones para mejorar la implementación, y evidencia empírica del nivel de avance de dicho proceso en las empresas del Sector Financiero en Monterrey.

### **1.6. Estructura del documento**

El documento se estructura en nueve capítulos que se distribuyen de la siguiente manera:

- I. **Introducción**, se explica la definición del problema, los objetivos, la justificación, las limitaciones y el producto final de la investigación.
- II. **Conociendo el *Data Warehouse***, aborda la historia, definición, ventajas y desventajas, tendencias, características, elementos importantes, el procesamiento analítico en línea, *data marts*, *data mining*, su arquitectura, y los procesos del *data warehouse*.
- III. **Metodologías de Implementación de un *Data Warehouse***, se resalta la importancia de seguir un método para la implementación de esta tecnología, bajo la perspectiva de diferentes autores.

- IV. **Factores Críticos para la Implementación de un *Data Warehouse***, se explica la importancia de los factores críticos, origen, definición, y aquellos que han sido identificados por diversos autores.
- V. **Herramientas para la Explotación del *Data Warehouse***, se enlista el *software* especializado en *data warehouse*, así como los principales proveedores de esta tecnología disponibles en el mercado.
- VI. **Casos de Éxito**, se documenta una serie de casos de éxito de la implementación de esta tecnología.
- VII. **Investigación de Campo**, se explican los detalles del estudio de la encuesta aplicada a las empresas del Sector Financiero de Monterrey.
- VIII. **Análisis de Resultados**, se presentan los resultados de las encuestas, los factores críticos identificados y una serie de recomendaciones.
- IX. **Conclusiones y Trabajos Futuros**, se dan a conocer las conclusiones generales y los trabajos futuros, derivados de los resultados obtenidos en la investigación.

Como se puede observar, el marco teórico lo componen los primeros seis capítulos, seguido de los capítulos de la investigación de campo y el análisis de los resultados, para finalizar con el capítulo de conclusiones y trabajos futuros.

## **2. Conociendo el *data warehouse***

### **2.1. Historia**

El concepto de *data warehouse* no se centra en la generación de nuevos datos, sino utilizar los existentes de forma masiva para transformarlos en una forma conveniente para proveer información estratégica, es así como nace (Ponniah, 2001).

Al respecto, Han y Kamber (2006) opinan que el *data warehouse* fue el resultado de la exigencia de los ejecutivos para proveer herramientas que pudieran organizar, entender y utilizar los datos, para ser competitivos. En los últimos años varias empresas han gastado millones de dólares en la construcción de su propio *data warehouse*, el cual les sirve para conocer las necesidades de los clientes, dado el incremento en la competencia de los sectores.

Conforme a Rainardi (2008), el *data warehouse* surgió por la necesidad de poder almacenar la historia de una organización, ya que la mayoría de los sistemas transaccionales en un principio solo podían almacenar algo de información.

Ponniah (2001) asegura que el *data warehouse* existe para proveer a los usuarios acceso a los indicadores de la organización desde una sola versión unificada, almacenar su historia adecuadamente y ver los datos de diferentes perspectivas.

La evolución del *data warehouse* se dio a través de los años, de acuerdo a Inmon, Strauss y Neushloss (2008), consta de seis etapas, a continuación se explica cada una de ellas.

#### **1. Los Primeros Medios de Almacenamiento**

En un inicio había mecanismos de recolección de datos muy simples. Las tarjetas perforadas se utilizaban para el almacenaje de la información, sin embargo resultaban ser muy limitadas en su capacidad y con un costo muy alto.

Las cosas cambiaron con la introducción de las cintas magnéticas, las cuales permitían almacenar grandes volúmenes de datos a un bajo costo, caracterizándose por eliminar las restricciones importantes sobre el formato en qué se registraban los datos, además la información podía ser escrita y modificada. Las cintas magnéticas representaron un gran salto hacia el futuro hablando de métodos de almacenamiento; sin embargo, sólo se podían acceder a los datos secuencialmente, aparte de no ser el medio de almacenamiento más estable, ya que la presencia de óxido podía inutilizar la cinta.

Otro paso importante fue el almacenamiento en disco, debido a que se podía escribir, modificar, y acceder en masa.

## **2. Los Sistemas de Almacenamiento de Bases de Datos**

En poco tiempo el almacenamiento en disco fue acompañado por un programa de *software* llamado “Sistema de Administración de Bases de Datos”, creado con el propósito de administrarlo; constaba de actividades como:

- Identificar la ubicación adecuada de los datos
- Resolver los conflictos cuando dos o más unidades de datos se asignan en la misma ubicación física
- Permitir que los datos logran borrarse
- Expandir una ubicación física, cuando un registro de datos no se ajustaba a un espacio físico limitado

De los puntos anteriores, el beneficio más importante del almacenamiento en disco fue la capacidad de localizar rápidamente los datos.

### **3. Las Aplicaciones en Línea**

Una vez que los datos se podían acceder directamente, pronto se evolucionó a las aplicaciones en línea; las cuales dependían de la computadora para acceder a los datos rápida y consistentemente.

Con la llegada de las aplicaciones en línea, las instituciones avanzaron al siglo veinte, prestando servicios, volviéndose poderosas y populares; y en poco tiempo dieron origen a muchas aplicaciones entrelazadas. Los usuarios finales comenzaron a tener problemas al localizar los datos, porque estos podían estar en cualquier lugar y no se tenía la certeza que fueran correctos.

### **4. Las Computadoras Personales y la Tecnología de Cuarta Generación**

Con la finalidad de solucionar los problemas de los usuarios finales con respecto al acceso de los datos, surgieron dos tecnologías: las computadoras personales y la tecnología de lenguajes de programación de cuarta generación (en inglés, *Fourth-Generation Programming Language*, abreviado *4GL*).

Las computadoras personales facilitaron las necesidades personales de procesamiento y almacenamiento, eliminando con esto el requisito de tener departamentos especializados para estas tareas.

4GL hizo posible que cualquier persona pudiera desarrollar sistemas y programas de una manera sencilla; siendo el usuario final libre de depender del departamento de TI.

La idea original radicaba en que los usuarios finales conseguirían crear las herramientas que necesitaban, ya que contaban con las computadoras personales y la tecnología 4GL, pero se encontraron con los siguientes problemas:

- Los datos incorrectos podían ser muy engañosos
- Los datos incompletos e inoportunos no eran muy útiles
- Cuando hay múltiples versiones de los mismos datos puede dar lugar a malas decisiones
- Los datos sin documentación, eran cuestionables

Quedaba claro que la participación del departamento de TI continuaba siendo imprescindible.

## **5. El Ambiente de “Telaraña”**

El resultado de las etapas anteriores fue un gran lío, este desorden fue llamado "el ambiente de telaraña"; se llamó así por las múltiples conexiones a todos los lugares, lo cual hace semejanza a una tela de araña.

El ambiente de telaraña pasó a ser increíblemente complejo en muchos entornos corporativos. La frustración del usuario final, el profesional de TI y los administradores; dio lugar a un movimiento para desarrollar una arquitectura de información diferente; iniciando así el *data warehouse*.

## **6. El Comienzo del Ambiente del *Data Warehouse***

El *data warehouse* representó un cambio importante en el pensamiento de los profesionales de TI. Antes de su llegada, se pensó que una base de datos debía ser algo que sirviera para todos los datos; pero el *data warehouse* puso en claro que hay

diferentes tipos de bases de datos, por lo que inició un movimiento para unificarlas; lo cual hizo posible la transición de un ambiente de “telaraña”, a uno que formaba los cimientos de los sistemas que apoyan actualmente a la toma de decisiones.

Resumiendo, el proceso de evolución del *data warehouse* comenzó con el progreso de diferentes herramientas tecnológicas, para cubrir las necesidades de información de las empresas.

## **2.2. Definición del *Data Warehouse***

Bill Inmon (2005), considerado el padre del *data warehouse*, lo define originalmente como “una colección de datos que apoya a la toma de decisiones caracterizándose por ser orientada al tema, integrada, no volátil y variable con respecto al tiempo”, además menciona que los datos pueden ser utilizados para diferentes propósitos en la misma organización. Posteriormente, Inmon junto con Strauss y Neushloss (2008) añaden la característica de poder soportar las decisiones administrativas al concepto de *data warehouse*.

Rainardi (2008) define a un *data warehouse* como “un sistema que recupera y consolida datos periódicamente de sistemas fuente, a un repositorio de datos dimensional o normalizado.” Rainardi explica que los sistemas fuentes suelen ser aquellos que tienen años de historia de la organización, que sirven para la inteligencia de negocios y las actividades analíticas.

El *data warehouse*, según explica Silvers (2008), es “un activo de la empresa que existe para beneficiarla en su totalidad, y no sólo ha algunas entidades de la misma.” Su definición se enfoca entorno a que los datos deben ser promovidos en toda la

organización, con el fin de poder ser utilizados por todos los miembros durante el ciclo de vida del repositorio.

Para Han y Kamber (2006), el *data warehouse* es “un repositorio de información que es construido mediante un proceso de limpieza, integración, transformación, carga y actualización de datos”; enfatizan que se trata de un repositorio originado de diversas fuentes, almacenado bajo un esquema unificado y que reside en un sitio.

Mundy, Thornthwaite y Kimball (2006) indican que el *data warehouse* es “la plataforma de la inteligencia de negocios que incluye todo lo necesario de la fuente original de información, que es extraída para ser presentada a los usuarios mediante aplicaciones.”

Ponniah (2001) señala que el *data warehouse* es “un ambiente informativo que provee una vista total e integrada de la empresa, facilita el acceso a la información actual e histórica, hace posible las transacciones para la toma de decisiones, transforma la información de la empresa a un estado consistente, y presenta una fuente de información estratégica, flexible e interactiva.”

En conformidad a lo anterior, un *data warehouse* se puede definir como el repositorio de información para la inteligencia de negocios, se compone de diversas fuentes; construido mediante un proceso de limpieza, integración, transformación y carga; caracterizado por ser orientado al tema, integrado, no volátil y variable con respecto al tiempo

### **2.3. Beneficios del *Data Warehouse***

El *data warehousing* se ha vuelto popular en la industria debido a los beneficios que representa para el procesamiento analítico de la información y el apoyo a la toma de

decisiones (Inmon, Strauss & Neushloss, 2008), debido a que provee una alternativa interesante para el enfoque tradicional de integración de base de datos heterogéneas (Han & Kamber, 2006).

En lugar de utilizar un enfoque orientado a la consulta, en el *data warehouse* la información proviene de múltiples fuentes heterogéneas que está integradas por adelantado y almacenadas en un repositorio directo para consulta y análisis (Han & Kamber, 2006). Inmon, Strauss y Neushloss (2008) sostienen que los principales beneficios de un *data warehouse* consisten en su capacidad de realizar búsquedas, la flexibilidad de los datos y su naturaleza para almacenar datos históricos.

Si se trata de realizar búsquedas a través de toda la organización, el *data warehouse* representa una alternativa adecuada debido a que integra los datos bajo un mismo esquema, lo que permite tener una verdadera vista de éstos, dando la impresión de que llegaron de una única fuente bien definida.

La flexibilidad de los datos se consigue gracias al nivel de detalle con el que se pueden hacer las búsquedas. Al estar los datos en una forma granular, estos pueden ser examinados por varios grupos de personas. Granular significa que hay un sólo conjunto de datos, una sola versión de la verdad.

La capacidad de poder almacenar datos de varios años convierte al *data warehouse* en la base para las herramientas analíticas. Para Han y Kamber (2006), las organizaciones utilizan los datos procedentes de un *data warehouse* para apoyar a las actividades relacionadas con la toma de decisiones, los beneficios que se obtienen son:

- La información obtenida puede ser utilizada para analizar los patrones de compra del cliente (como el tiempo, periodos y preferencias de compra)

- Es posible reposicionar y administrar los portafolios de los productos mediante la comparación de los resultados de las ventas por trimestre, por año, y por regiones geográficas, a fin de afinar las estrategias de producción
- Facilita las operaciones de análisis y búsqueda de fuentes de financiamiento
- Permite la administración de las relaciones con los clientes, con el fin de hacer correcciones conforme a la industria, la administración y el costo de los activos de la empresa
- Provee de un alto rendimiento para integrar un DBMS; debido a que los datos se copian, son preprocesados, integrados, anotados, resumidos, y reestructurados semánticamente en un repositorio de datos
- El procesamiento de las consultas realizadas por un *data warehouse* no interfiere con las fuentes de datos locales
- Un *data warehouse* puede almacenar e integrar la información histórica y dar soporte a consultas complejas
- El *data warehouse* se puede utilizar para el procesamiento de la información (consulta y presentación de informes), el procesamiento analítico (permite a los usuarios navegar a través de datos detallados y resumidos), y el *data mining* (apoya el descubrimiento de conocimiento)

Analizando los beneficios se puede observar que en su mayoría sirven para proveer de ventajas competitivas a la empresa, mediante el uso adecuado y eficiente de la información.

## **2.4. Problemas del *Data Warehouse***

Los principales problemas al momento de querer construir un *data warehouse* son (Inmon, Strauss & Neushloss, 2008):

### **1. La integración de los datos**

El problema que se enfrentan las organizaciones al momento de querer construir un *data warehouse* es la necesidad de integrar los datos; ya que la mayoría cuentan con sistemas que han sido utilizados por muchos años, con la particularidad de no poder ser modificables.

Aunque en algunos casos resulta imposible realizar esta transición, es posible superar este obstáculo, lo que se recomienda es un involucramiento profundo con el viejo ambiente de estos sistemas y la realización de una investigación a detalle de la información con la que cuenta, para poder convertir los datos de esas aplicaciones en datos útiles al *data warehouse*.

### **2. El volumen de los datos**

El volumen de datos es el resultado de integrar la información de los diferentes departamentos de una organización, la mayoría de los profesionales de TI nunca han hecho frente a este problema.

### **3. Un enfoque de desarrollo diferente**

El tercer problema consiste en la dificultad de definir la forma en cómo se construirá el *data warehouse*. La metodología de construcción del sistema más utilizado alrededor

del mundo es la recolección de requerimientos, sin embargo este proceso difiere con la de un sistema tradicional.

Si bien es cierto que se debe tener cuidado en los aspectos de integración, volumen de datos y enfoque al negocio; al momento de la implementación estas cuestiones se presentan en mayor o menor proporción dependiendo de la empresa en cuestión.

## **2.5. Tendencias del *Data Warehouse***

### ***Escenario actual del data warehouse***

Hoy en día, el *data warehouse* y la inteligencia de negocios (BI) se consideran áreas de misión crítica en las organizaciones. Se definen sistemas de misión crítica a los sistemas de apoyo a la generación de ingresos y los procesos del negocio, por lo que sin ellos, en un período de tiempo determinado hay aumento en las pérdidas de ingresos o aumentos inaceptables de los costos empresariales (Feinberg, 2009c).

El mercado de DBMS, al cual pertenece el *data warehouse*, seguirá creciendo a una tasa de casi un 9% durante los próximos cinco años, con ingresos totales de casi \$30 billones para 2011. Actualmente se visualiza sólo una pequeña disminución en los ingresos debido a la situación económica mundial en 2008 y 2009. Los ingresos de DBMS seguirán siendo fuertes, ya que las organizaciones entienden el valor de las tecnologías de información (TI), sus datos de BI y los sistemas de *data warehouse* (Feinberg, 2009b).

De acuerdo a los datos proporcionados por Feinberg (2009b), las principales prioridades para los Directores de Informática en el año del 2009 son las que se muestran en la figura 2.1.

## BI and Applications Top the List — Again!

To what extent is each of the following technologies a priority for you in 2009?	LA Rank 2009	Change	LA Rank 2008	WW Rank 2009
BI applications, analytics and data mining	1	←	1	1
Enterprise Applications (ERP, SCM, CRM)	2	←	2	2
Server and storage technologies (virtualization)	3	▲	8	3
Networking, voice and data communications (VoIP)	4		*	6
Legacy application modernization, upgrade and replacement	5	▼	3	4
Security technologies (access control, authentication)	6	▲	7	8
Collaboration technologies	7	▼	5	5
Service-oriented applications and architecture (SOA, SOBA)	8	▼	4	9
Mobile workforce applications and devices	9		*	**
Workflow management	10	←	10	**

▲ ▼ Selected change in 2009 ranking compared with 2008. \*Not in LA top 10 for 2008

Source: Gartner Executive Programs CIO Survey, 2009.

\*\*Not in WW top 10 for 2009

**Gartner.**

Figura 2.1 Prioridades de tecnologías de información en 2009

Fuente: Latin America Scenario (Feinberg, 2009a)

Del estudio se pueden observar dos puntos relevantes. Primero, es que cada año América Latina y las primeras diez prioridades de TI a nivel mundial están alineadas, aunque se encuentran en distinto orden. Segundo, por quinto año consecutivo el BI está en las dos primeras categorías. Curiosamente, muchas de las otras prioridades, como la administración del flujo de trabajo, la colaboración, la modernización de aplicaciones heredadas, las aplicaciones empresariales y la arquitectura orientada al servicio (SOA) también son afectadas por la inteligencia de negocios; o tienen un efecto en ésta.

Debido a que el BI tiene una influencia tan amplia dentro de la empresa, seguirá siendo una de las dos principales prioridades en varios años, algo importante de señalar es que hay una falta de conocimiento sobre estos temas en América Latina.

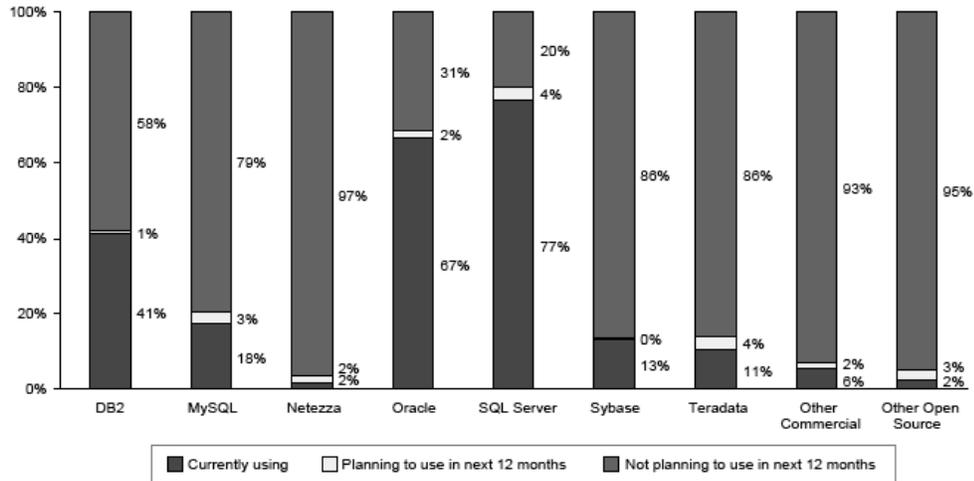
Feinberg (2009b) prevé que a finales del año 2012, dos terceras partes de los proveedores de *software* de administración de datos dependerá fundamentalmente de los ingresos de servicios y soporte; debido a que los DBMS de código abierto han alcanzado su madurez, escalabilidad y fiabilidad, por lo que la presión sobre los distribuidores comerciales DBMS crecerá para reducir los precios de las licencias y las actualizaciones de sus productos.

### ***Panorama actual de los productos de data warehouse***

En la figura 2.2 Feinberg (2009b), mediante la respuesta a la pregunta ¿cuál de los siguientes productos de DBMS tú organización utiliza actualmente o planea utilizar en los próximos 12 meses?, presenta el panorama actual de los productos de *data warehouse*.

## DBMS Deployment Plans Show Growth in Smaller Vendors

Which of the following DBMS products does your organization currently use or plan to use in the next 12 months?



314 Respondents

Source: 2008 Business Intelligence Multiclient Study

Gartner

*Figura 2.2 ¿Cuál de los siguientes productos de DBMS tu organización utiliza actualmente o planea utilizar en los próximos 12 meses?*

*Fuente: The Future of DBMS for Data Warehousing (Feinberg, 2009b)*

El mercado de Microsoft SQL Server crece a un promedio excedido del 28%, seguido por la oferta de Oracle con un 15%, siendo ésta última la más cercana al ritmo de crecimiento del líder del mercado. Cabe mencionar que IBM, mediante su solución llamada DB2, tiene potencial de abarcar mayor mercado a pesar de sólo estar creciendo a un 8.8%.

### *El futuro del data warehouse*

Por su parte, Rainardi (2008), Thoo, Friedman, Beyer, Feinberg, Bitterer, Bell, Gassman, Newman, Blechar, Radcliffe y White (2008), y Feinberg (2009c) afirman que las tendencias que le depara el futuro al *data warehouse* se enfocan en:

### **1. Datos no estructurados**

Los datos están organizados en estructuras de bases de datos. Los datos no estructurados, no se encuentran ordenados (filas y columnas, estructuras de árbol, clases o tipos) lo que dificulta su extracción. Algunos ejemplos de datos sin estructura son documentos, imágenes (fotos y diagramas), audio (canciones, discursos y sonidos), películas de vídeo (animaciones), flujos de datos, texto, correos electrónicos y sitios Web de Internet.

La gente ha comenzado recientemente a darse cuenta que la mayoría de sus datos se almacenan en la clasificación de no estructurados y sólo un poco de ellos son estructurados. Por lo tanto, los datos no estructurados se han convertido en una de las tendencias actuales del *data warehouse*.

### **2. Arquitectura orientada a servicio (SOA)**

SOA es un método de construcción de aplicaciones que utiliza una serie de pequeños componentes independientes que hablan unos con otros mediante la oferta y la demanda de sus servicios; estos componentes pueden ser distribuidos, de hecho, pueden estar ubicados en diferentes partes del mundo.

Un sistema de *data warehouse* se forma de muchos componentes: sistemas origen; sistemas extracción, transformación y carga (en inglés, *Extract, Transform & Load*, abreviado *ETL*); mecanismo de calidad de datos; sistema de metadatos; sistemas de auditoría y control; portales de BI; solicitudes de informes; aplicaciones analíticas;

aplicaciones de *data mining*; y sistemas de base de datos. Para la construcción de un *data warehouse* se puede optar por construirlo como una aplicación gigante con todos los componentes estrechamente relacionados, sin embargo, al momento de querer sustituir uno de los componentes se afecta a otros.

### **3. *Data warehouse* de tiempo real**

Un *data warehouse* en un principio fue creado para actualizarse una vez por día o por semana; en los últimos dos a tres años, ha habido más demandas para aumentar la frecuencia de las actualizaciones. Los usuarios desean ver los datos en el *data warehouse* actualizado cada dos minutos e incluso en tiempo real, este último ofrece actualizaciones al momento que una transacción ocurre, en el sistema origen.

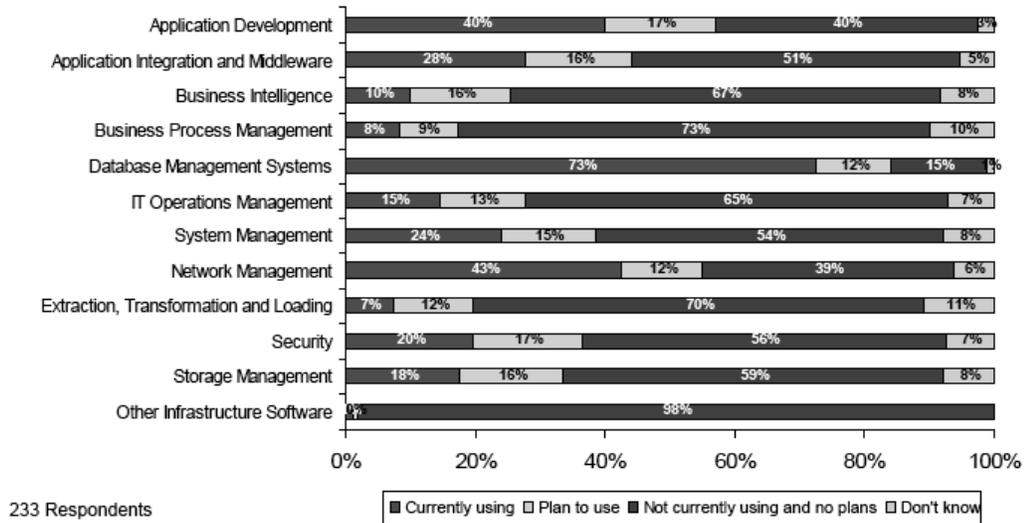
### **4. DBMS de código abierto**

Se define como el *software* de código abierto que se utiliza en aplicaciones, incluyendo aquellas de misión crítica, tales como el procesamiento de transacciones *online* y *data warehouse*.

El mercado los DBMS de código abierto ha madurado en varios niveles de aplicación. En la figura 2.3 se muestran los resultados, presentados por Feinberg (2009c), con respecto a una encuesta realizada a las organizaciones de TI sobre el uso del *software* de código abierto.

# Open-Source Use Survey 2008

## Open-Source Use in IT Organizations Already Using Open-Source Software



**Note: All survey participants currently use open source or plan to use open source within the next year**

Source: Gartner Open-Source UNW Study, 2008

**Gartner.**

Figura 2.3 Organizaciones de TI que ya están utilizando software de código abierto

Fuente: *The Truth About Open-Source BI/DI/DW* (Feinberg, 2009c)

Un número creciente de organizaciones de TI están encontrando que el código abierto es una opción válida y rentable en muchos aspectos de los mercados de infraestructura de *software*.

### 5. Artefactos de *data warehouse*

Las organizaciones de TI al momento de pretender reemplazar un antiguo *data warehouse* por uno nuevo, se enfrentan a problemas de compatibilidad por la enorme variedad de proveedores en el mercado, sin embargo, la solución se encuentra en los artefactos de *data warehouse*; conocidos en inglés como *data warehouse appliances*; éstos se definen como un conjunto equilibrado de *hardware* preempacado o

preconfigurado (servidores, memoria, almacenamiento y canales de entrada/salida), *software* (sistema operativo, DBMS y *software* de administración), servicio y soporte, que se vende como una unidad con redundancia incorporada para alta disponibilidad como una plataforma para el *data warehouse* (Thoo et al., 2008).

## 6. Base de Datos como Servicio

La idea que está detrás de la arquitectura orientada al servicio (SOA) hizo posible la formulación de una nueva tecnología: una base de datos como servicio. Tecnología conocida en inglés como *Database as a Service*, abreviada *DaaS*, Feinberg (2009b) la define como una instancia de base de datos alojada fuera del sitio del cliente y administrada por el vendedor en lugar del cliente; que es accedida por los usuarios y las aplicaciones a través de la Web.

Contextualizando las tendencias descritas anteriormente, se puede observar que el *data warehouse* como apoyo a las aplicaciones de inteligencia de negocios sigue siendo prometedor, y con prioridad de inversión para los Gerentes de Informática.

### 2.6. Características del *Data Warehouse*

De acuerdo a William H. Inmon (2005), "Un *data warehouse* es una colección de datos orientada hacia un tema, integrada, no volátil, y variable con respecto al tiempo en apoyo a la toma de decisiones empresariales". En esta definición están presentes las principales características de un *data warehouse*; a continuación se detallan.

**1. Orientado hacia un tema:** En lugar de concentrarse en las operaciones del día a día, un *data warehouse* se centra en organizar los datos en base a temas de negocios de la empresa, lo que proporciona una visión simple y concisa de las principales

áreas de la empresa, que es útil en el proceso de apoyo a la toma de decisiones (Silvers, 2008).

2. **Integrado:** Un *data warehouse* esta generalmente construido por múltiples fuentes heterogéneas integradas, como bases de datos relacionales, archivos planos, y registros de transacciones en línea. Las técnicas de limpieza y de integración de datos se aplican para garantizar la consistencia en las convenciones de nomenclatura, las estructuras, los atributos, etc. (Inmon, 2005).
3. **Variable con respecto al tiempo:** Un *data warehouse* refleja el estado histórico y actual de la empresa mediante la inserción de nueva información; cada estructura clave contiene implícita o explícitamente, un elemento de tiempo (Silvers, 2008).
4. **No volátil:** Los datos, una vez escritos en un *data warehouse*, no se suprimen o actualizan; debido a este factor, se conservan los datos de la historia de la organización. (Silvers, 2008).

Las características mencionadas, son las que distinguen al *data warehouse* de otros sistemas de almacenamiento, tales como los sistemas de bases de datos relacionales, de procesamiento de transacciones y de archivos.

## 2.7. La Inteligencia de Negocios y el *Data Warehouse*

La inteligencia de negocios de acuerdo a Rainardi (2008), es un conjunto de actividades para comprender situaciones de negocios mediante la realización de diversos tipos de análisis de datos de la empresa, con el fin de apoyar al proceso de la toma de decisiones. Esto incluye la recopilación, análisis, comprensión y administración de datos sobre el rendimiento de las operaciones, las actividades de los clientes y los proveedores, los resultados financieros, los movimientos del mercado, la competencia, el cumplimiento de regulaciones y los controles de calidad.

Rainardi (2008) clasifica a las distintas aplicaciones de inteligencia de negocios en seis categorías: las aplicaciones de informes, las aplicaciones analíticas, las aplicaciones de minería de datos, los *dashboards*, las alertas, y el portal.

- **Aplicaciones de informes:** Consultan los datos en el *data warehouse* y los presentan en formato tabular o de pivote.
- **Aplicaciones analíticas:** Consultan el *data warehouse* repetidamente y de forma interactiva, presentando los datos en formatos flexibles.
- **Aplicaciones de *data mining*:** Exploran el *data warehouse* en busca de patrones y relaciones que describan a los datos.
- ***Dashboards*:** Son una categoría de aplicaciones de inteligencia de negocios que dan un rápido resumen de alto nivel sobre el rendimiento empresarial; mediante indicadores, gráficas, y mapas codificados por color. Al seleccionar un indicador, se puede profundizar a un mayor nivel de detalle.
- **Alertas:** Son notificaciones a los usuarios cuando ocurren ciertos eventos o condiciones.
- **Portal:** Es una aplicación que funciona como una puerta para acceder y administrar los informes de inteligencia de negocios, análisis, *data mining*, y aplicaciones de escritorio, así como las suscripciones a alertas.

Ya establecidas las distintas clases de aplicaciones de inteligencia de negocios, es importante mencionar su historia. Ponniah (2001) afirma que se trató de un proceso evolutivo por el que la mayoría de las empresas pasaron, en su intento por obtener información estratégica.

- **Reportes *Ad Hoc*.** Cuando los usuarios realizaban informes especiales, enviaban solicitudes al departamento de TI; quien escribía programas (para cada solicitud), y producía los reportes *ad hoc*.
- **Programas Especiales de Extracción.** El departamento de TI escribía un conjunto de programas, y los ejecutaba periódicamente para extraer datos de distintas aplicaciones, esto con la finalidad de anticiparse a las solicitudes de informes especiales.
- **Pequeñas aplicaciones.** El departamento de TI creaba aplicaciones sencillas basadas en archivos extraídos, las cuales permitían a los usuarios manipular los parámetros para cada informe especial que ellos requerían. Algunas aplicaciones avanzadas también permitían a los usuarios ver la información en línea.
- **Centros de Información.** El centro de información, creado en la década de los sesenta, era un lugar donde los usuarios podían ir a solicitar informes *ad hoc* o ver información especial en pantallas. El personal de TI estaba presente en estos centros de información para ayudar a los usuarios a obtener la información deseada.
- **Los Sistemas para el Apoyo a la Toma de Decisiones.** En esta etapa, las compañías comenzaron a construir sistemas más sofisticados destinados a proporcionar información estratégica. Los sistemas tenían interfaces de menú, que facilitaban información en línea y tenían la capacidad de imprimir informes especiales.
- **Sistemas de Información para Ejecutivos.** Construidos con la finalidad de llevar la información estratégica al escritorio de los ejecutivos, este tipo de sistemas mostraban la información clave de cada día y proporciona los informes de una forma sencilla. Sin embargo, sólo pantallas e informes preprogramados estaban disponibles, lo cual ocasionaba frustración entre sus usuarios.

Dada la importancia de las diferentes tecnologías y las necesidades particulares de cada negocio, se vio obvio que un repositorio de datos (de calidad, confiable, flexible y reutilizable), debía ser el punto de partida para apoyar y mantener un ambiente de inteligencia de negocios. Es entonces cuando el *data warehouse* se incorporó a la arquitectura de inteligencia de negocios como lo muestra la figura 2.4 (Imhoff, Galemme & Geiger, 2003).

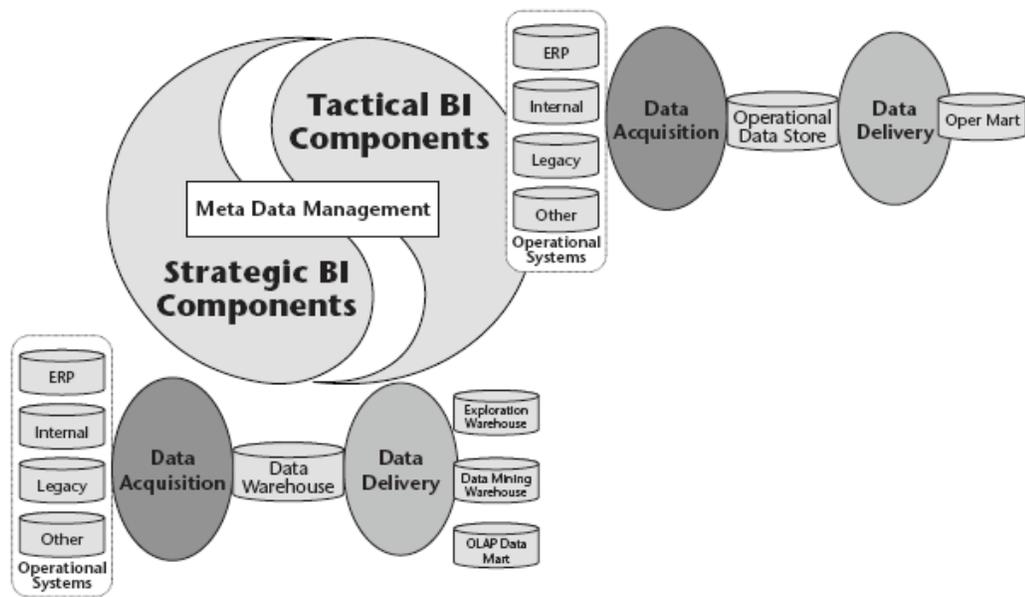


Figura 2.4. Porciones estratégicas y tácticas de ambiente de inteligencia de negocios

Fuente: *Mastering Data Warehouse Design* (Imhoff, Galemme & Geiger, 2003)

Siendo la parte fundamental de la arquitectura, el *data warehouse* actualmente es utilizado principalmente para ayudar a los ejecutivos a entender mejor su negocio, ayudarlos a tomar mejores decisiones en el área operativa, táctica y estratégica, y para ayudarles a mejorar su rendimiento empresarial (Rainardi, 2008).

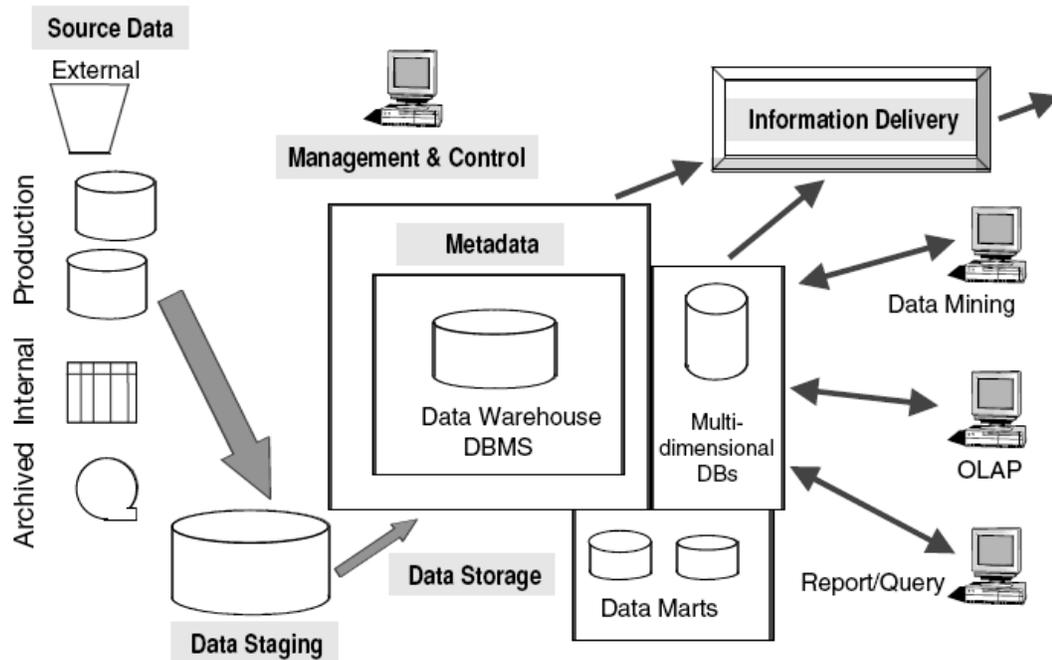
## 2.8. Arquitectura del *Data Warehouse*

El *data warehouse*, conforme a Inmon, Strauss y Neushloss (2008), se ha convertido en una parte esencial de la infraestructura de las tecnologías de información, originalmente creció en respuesta a la necesidad de información de las empresas, además de tratarse de una construcción que suministra datos integrados, granulares e históricos de la empresa.

Sin embargo, el problema con el *data warehouse* es que se tienen muchas versiones de éste en la actualidad. Este problema ocasiona que existan versiones diferentes de cómo debe ser la estructura, ocasionando que los mismos conceptos del *data warehouse* tengan distinto significado dependiendo de la versión que se maneje. Existen tantas versiones del *data warehouse* como proveedores de *software* y *hardware*.

Ponniah (2001) menciona que una arquitectura es la forma adecuada de como se deben arreglar los componentes, además esta estructura debe ser de cierta forma para obtener el máximo beneficio posible, y adaptarse a los requisitos de la organización. En la figura 2.5 se ilustran los seis componentes básicos de un entorno de *data warehouse*.

*Architecture is the proper arrangement of the components.*



*Figura 2.5 Componentes del Data Warehouse*

*Fuente: Data Warehouse Fundamentals Ponniah (2001)*

## 1. Componente de origen de datos

El origen de los datos que entran en el *data warehouse* pueden agruparse en cuatro grandes categorías.

- a) **Datos de Producción.** Esta categoría de datos proviene de los diferentes sistemas operacionales de la empresa; pueden estar en distintos tipos de formatos, residir en diferentes plataformas de *hardware*, sistemas de bases de datos y sistemas operativos.
- b) **Datos Internos.** Estos datos se encuentran en toda la organización y pueden ser de utilidad para el *data warehouse*, además de que pueden ser desconocidos por la misma organización.

- c) **Datos Archivados.** Son datos antiguos y archivos comprimidos que provienen de los sistemas operaciones, los cuales son fundamentales para la organización y periódicamente son almacenados. El período en el que los datos se quedan en el DBMS varía de uno a cinco años.
- d) **Datos Externos.** Estos datos provienen de fuentes externas, y representan un alto porcentaje de información que la mayoría de los ejecutivos utilizan. Se utilizan para crear estadísticas relacionadas a su industria; son producidas por organismos externos, para conocer el mercado de sus competidores, y para obtener indicadores financieros con el fin de comprobar el rendimiento de su negocio.

Sin importar su categoría, la principal prioridad de este componente es la captura de los datos de los sistemas origen, los cuales capturan las transacciones diarias de la empresa.

## 2. Componente de concentración de datos

Después de extraer datos de diferentes sistemas operacionales y de fuentes externas, se tienen que preparar para almacenarlos en el *data warehouse*. Los datos extraídos de diferentes fuentes necesitan ser cambiados, convertidos y dispuestos en un formato adecuado, para poder ser almacenados; con el fin de realizar consultas y análisis.

Tres funciones principales deben realizarse en el área de concentración, que sirve para limpiar, cambiar, combinar, convertir, duplicar y preparar a los datos.

- a) **Extracción de datos.** En esta función se emplean una serie de técnicas de extracción para cada fuente de datos. Dado que los datos origen pueden estar en diversos formatos, residir en sistemas de bases de datos relacionales, estar en sistemas heredados de otros modelos y jerarquías de datos, estar en archivos planos,

o encontrarse en hojas de cálculo, la extracción de datos puede llegar a ser bastante compleja.

- b) Transformación de Datos.** Durante esta función se lleva a cabo la limpieza, estandarización, y combinación de los datos como parte de la transformación de datos.
- c) Carga de Datos.** Dos grupos distintos de tareas forman la función de carga de datos. Al completar el diseño y la construcción del *data warehouse*, se realiza la carga inicial de datos en el repositorio. La carga inicial mueve grandes volúmenes de datos utilizando pequeños periodos de tiempo.

Una vez realizadas las funciones del componente de concentración, los datos están preparados para su almacenamiento y uso.

### **3. Componente de carga de datos**

Un repositorio de datos tradicional para sistemas operacionales normalmente contiene datos actuales, además de estructurados en formatos altamente normalizados para un procesamiento rápido y eficiente. En cambio, en un repositorio de datos para un *data warehouse*, es necesario mantener grandes volúmenes de datos históricos y guardarlos en estructuras adecuadas para su análisis; y no para una rápida recuperación de piezas individuales de información. Por lo tanto, el almacenamiento de datos para el *data warehouse* se mantiene por separado de los sistemas operacionales.

### **4. Componente de entrega de información**

Con el fin de proporcionar información a la amplia comunidad de usuarios del *data warehouse*, el componente de entrega de información incluye diferentes métodos. Se pueden emplear informes orientados principalmente para el principiante y los usuarios

ocasionales; consultas complejas, análisis multidimensionales y estadísticos para atender las necesidades de los analistas de negocios y usuarios avanzados; o para proporcionar datos a las aplicaciones de *data mining*.

## **5. Componente de metadatos**

Los metadatos en un *data warehouse* son similares al diccionario o al catálogo de datos, en un sistema de administración de base de datos. En el diccionario de datos, se mantiene la información acerca de las estructuras lógicas, los archivos, las direcciones y los índices, además de contener metadatos en la base.

## **6. Componente de administración y control**

Este componente de la arquitectura del *data warehouse* se encuentra en la parte superior. El componente de administración y control, coordina los servicios y las actividades dentro del *data warehouse*. Las funciones consisten en controlar la transformación y la transferencia de los datos durante el almacenamiento, controlar el movimiento de datos, e interactuar con el componente de los metadatos.

Se puede observar que los componentes de la arquitectura definen un ciclo de transformación de los datos, que abarca desde el proceso de recolección de las fuentes, a la entrega; en la dimensión de trabajo entendible para los usuarios finales.

### **2.9. Elementos Importantes de la Arquitectura**

Adicional a los componentes básicos, la arquitectura del *data warehouse* tiene otros elementos relevantes. Este conjunto lo forman los procesos básicos de transformación de datos, el procesamiento analítico en línea, los *data marts*, y el *data mining*. Dada su natural complejidad, a continuación se explican cada uno de ellos.

### 2.9.1. Procesos básicos realizados por un *data warehouse*

La información contenida en un *data warehouse* proviene de los mismos sistemas operacionales, sin embargo de acuerdo a Ponniah (2001), no puede ser utilizada directamente para proporcionar información estratégica.

Rainardi (2008) menciona que los procesos de extracción, transformación y carga (en inglés, *Extract, Transform and Load*, abreviado *ETL*) son un sistema que tiene la capacidad de leer los datos del *data warehouse*, transformarlos, y cargarlos en otro repositorio. Conforme a Ponniah (2001), sin un sistema ETL no sería posible remodelar los datos relevantes en información útil, no habría ninguna información estratégica, los datos origen no podrían ser extraídos correctamente, limpiados, e integrados en los formatos adecuados; y el procesamiento del *data warehouse* no sería posible.

El sistema ETL se encuentra en el proceso de adquisición y el almacenamiento de datos, dentro del ambiente del *data warehouse*. Después de la transformación de los datos, es posible mover físicamente los datos al *data warehouse*. Cada de uno de los procesos de un sistema ETL se explica a continuación (Ponniah ,2001).

#### 1. Extracción de datos

Los datos son extraídos de diferentes fuentes, lo cual se va haciendo en cargas de trabajo incrementales. La extracción efectiva de los datos es la clave para el éxito del *data warehouse*, por lo tanto hay que prestar atención a las siguientes cuestiones:

- **Identificación de la fuente:** Identificar aplicaciones y estructuras de origen.
- **Método de extracción:** Para cada fuente de datos, definir si el proceso de extracción es manual o se basa en una herramienta.

- **Frecuencia de extracción:** Para cada fuente de datos, establecer con que frecuencia se debe hacer la extracción de los datos (diaria, semanal, trimestral, y así sucesivamente).
- **Ventana de tiempo:** Para cada fuente de datos, denotar el periodo de tiempo para el proceso de extracción.
- **Secuencia de tareas:** Determinar si el comienzo de una tarea en un flujo de tareas de extracción tiene que esperar hasta que la tarea anterior haya finalizado exitosamente.
- **Manejo de excepciones:** Determinar cómo manejar los registros de entrada que no se pueden extraer.

Las cuestiones mencionadas conforman la estrategia de extracción de datos.

## 2. Transformación de los datos

Realizar varias transformaciones de datos de acuerdo a los estándares es inevitable, ya que provienen de diferentes sistemas, y es indispensable verificar que al reunirse no violen las normas del negocio.

Para la transformación es necesario realizar una serie de tareas básicas, que son independientes de la variedad y complejidad de los sistemas operacionales, y la magnitud del *data warehouse*. El conjunto de tareas básicas se compone de:

- **Selección.** Esto tiene lugar al comienzo de todo proceso de transformación de datos. Se seleccionan los registros enteros o las partes de varios registros de los sistemas origen.
- **División/Unión.** Esta tarea incluye los tipos de manipulación de datos que se necesitarán en las partes seleccionadas de los registros origen.

- **Conversión.** Incluye una gran variedad de conversiones rudimentarias de datos por dos razones principales; una es la estandarización a través de las extracciones de los diferentes sistemas origen; y la otra para hacer los campos utilizables y comprensibles a los usuarios.
- **Resumen.** A veces no es posible mantener los datos en el nivel más bajo de detalle en el *data warehouse*. Puede ser que ninguno de sus usuarios necesite datos con la menor granularidad para su análisis o consulta.
- **Enriquecimiento.** Esta tarea es la reorganización y la simplificación de los registros individuales para hacerlos más útiles. Se pueden utilizar uno o más campos del registro de la misma entrada para crear una mejor vista.

El proceso de transformación concluye al finalizar de realizar estas tareas básicas, alistando los datos para el proceso de carga.

### 3. Carga de datos

El último paso de los procesos del sistema ETL consiste en la preparación y aplicación de los datos. Se crean varias imágenes que corresponden a los archivos objetivo que se cargan en el *data warehouse*.

Todo el proceso de mover los datos al repositorio del *data warehouse* es llamado de varias maneras. La carga inicial consiste en poblar todas las tablas del *data warehouse* la primera vez; la carga incremental en la que se realizan los cambios de manera periódica; y actualización completa que borra por completo el contenido de una o más tablas y cargarlas con nuevos datos.

Concluyendo, Kimball y Caserta (2004) sostienen que el sistema ETL es la base del *data warehouse*. Un sistema ETL adecuado impone la calidad y coherencia con los

estándares, conforma los datos de modo que las fuentes independientes se puedan utilizar en conjunto y, finalmente los proporciona en un formato de presentación listo; para que los desarrolladores puedan crear aplicaciones y los usuarios finales puedan tomar decisiones.

### **2.9.2. Procesamiento analítico en línea**

De acuerdo a los datos de Ponniah (2001), los usuarios necesitan tener la capacidad de realizar análisis multidimensionales con cálculos complejos, sin embargo las herramientas tradicionales como los reportes, los productos de consulta, hojas de cálculo, las interfaces de lenguaje, y el entorno de un *data warehouse* tradicional resultan insuficientes. Se necesita un conjunto diferente de instrumentos y productos que están específicamente destinados a un análisis serio, y dados los beneficios que proporciona el *data warehouse*, la respuesta a esta necesidad es la integración del procesamiento analítico en línea (en inglés, *On-Line Analytical Processing*, abreviado *OLAP*).

Al integrar OLAP en el entorno del *data warehouse* se obtienen los siguientes beneficios:

- Permite a los analistas, ejecutivos y gerentes obtener información útil a partir de la presentación de los datos.
- Puede reorganizar métricas a lo largo de varias dimensiones y permitir que los datos sean vistos de diferentes perspectivas.
- Apoya el análisis multidimensional.
- Es capaz de profundizar dentro de cada dimensión.
- Es capaz de aplicar fórmulas matemáticas y los cálculos de las medidas.

- Proporciona respuestas rápidas, facilitando la velocidad del pensamiento analítico.
- Complementa el uso de otras técnicas de entrega de información como el *data mining*.
- Mejora la comprensión de los conjuntos de resultados a través de presentaciones visuales de gráficos y tablas.
- Puede ser implementado en la Web.
- Diseñado para análisis altamente interactivos.

Con la relación de OLAP y el *data warehouse* bien establecida, más sus beneficios resulta conveniente describir los aspectos clave de OLAP. Enseguida se hace mención de la definición de diferentes autores, y de los tipos de tecnologías que se derivan de ésta.

### ***Definición de procesamiento analítico en línea***

Imhoff, Galemno y Geiger (2003) sostienen que OLAP es un término introducido por E. F. Codd que se refiere, a una categoría de tecnologías de *software* que permite a los analistas, gerentes y ejecutivos conocer mejor los datos a través de un acceso rápido, consistente y interactivo en una amplia variedad de puntos de vista de la información; con la finalidad de reflejar la dimensión real de la empresa tal como la entiende el usuario.

Conforme a Rainardi (2008), OLAP es una actividad de análisis, agregación, navegación, y reordenamiento de los datos del negocio, almacenados en el *data warehouse* para la toma de decisiones tácticas y estratégicas.

Por su parte, Silvers (2008) afirma que OLAP es un conjunto de aplicaciones que precalculan, y almacenan respuestas de las permutaciones de las dimensiones de un *data warehouse*.

OLAP, de acuerdo a Mundy, Thornthwaite y Kimball (2006), es una tecnología para almacenar, administrar y consultar datos; diseñada específicamente para apoyar la inteligencia negocios.

Se puede entender en resumen por procesamiento analítico en línea (OLAP) como un conjunto de aplicaciones, que permite analizar, navegar y reordenar los datos para apoyar al proceso de toma de decisiones empresariales.

### ***Tipos de tecnologías OLAP***

Los datos en las aplicaciones OLAP son precalculados y almacenados en estructuras multidimensional, también conocidas como cubos OLAP. Al momento de construir un cubo OLAP, el cliente puede elegir entre tres diferentes tecnologías OLAP (Silvers, 2008).

- 1. Procesamiento analítico multidimensional en línea:** En inglés, *Multidimensional Online Analytical Processing*, abreviado *MOLAP*, almacena todos los conjuntos de resultados de todas las permutaciones de la dimensión en un cubo OLAP. MOLAP requiere una amplia capacidad de almacenamiento, de ciclos de CPU, I/Os, y de memoria.
- 2. Procesamiento analítico relacional en línea:** En inglés, *Relational Online Analytical Processing*, abreviado *ROLAP*, no almacena conjuntos de resultados, mas bien identifica los datos dentro de un *data warehouse* para verificar si hay una intersección entre dimensiones, traduciendo esa información en una sentencia SQL.

ROLAP requiere la menor capacidad de almacenamiento en el servidor OLAP, sin embargo, ROLAP transfiere el consumo de los ciclos de CPU y de I/O al *data warehouse* al cliente, proporcionando un rendimiento más lento.

- 3. Procesamiento analítico híbrido en línea:** En inglés, *Hybrid Online Analytical Processing*, abreviado *HOLAP*, es una combinación de MOLAP y ROLAP. Al precalcular y almacenar la mayoría de los resultados, HOLAP logra un equilibrio entre capacidad, rendimiento, y número de permutaciones de las dimensiones disponibles para el cliente.

Cada una de las distintas tecnologías ofrece diferentes características de almacenamiento y rendimiento.

### **2.9.3. Data marts**

La relación de un *data mart* con un *data warehouse*, se origina por la incapacidad de este último de poder proporcionar los datos de la manera requerida por los clientes y el negocio. La justificación común para la existencia de *data mart* consiste en la entrada de datos de un área del negocio, la cual realiza análisis y pronósticos; y su capacidad de segregación, que permite limitar los datos sensibles de acuerdo a las particularidades de cada área (Silvers, 2008).

Observada su importancia en el ambiente de *data warehouse*, en esta sección se describe a detalle la definición, las ventajas y desventajas, los tipos, y el proceso de construcción de un *data mart*.

#### ***Definición de data mart***

Conforme a Inmon, Strauss y Neushloss (2008), el *data mart* es el entorno donde el usuario final tiene acceso directo y controla sus datos analíticos, y configura un

conjunto de expectativas generales de los usuarios sobre la forma que deberían tener los datos.

Un *data mart*, según explica Silvers (2008), es un subconjunto físico o lógicamente independiente de un *data warehouse* que recibe al menos algunos de sus datos de éste último, siendo la fuente de datos los clientes u organizaciones externas a la empresa que lo utilizan.

Han y Kamber (2006) definen a un *data mart* como un subconjunto de datos a nivel empresarial con tendencia a estar resumidos, que son de valor para un grupo específico de usuarios.

Kimball y Ross (2002) mencionan que originalmente el *data mart* se definió como un subconjunto de un alto nivel de agregación de datos, sin embargo esta definición resultó inviable debido a que dio lugar a *data marts* inflexibles e incompatibles entre sí. Ellos redefinen al *data mart* como un conjunto flexible de datos que pueden unirse a través de técnicas, llegando a representar a los datos de los procesos del negocio.

Imhoff, Galemno y Geiger (2003) señalan que el *data mart* es donde se llevan a cabo la mayoría de las actividades de análisis de la inteligencia de negocios, siendo los datos de este adaptados a una capacidad o función en concreto.

En resumen, un *data mart* es un subconjunto de datos donde realizan la mayoría de las actividades de análisis para la inteligencia de negocios, siendo estos de utilidad para un grupo específico de usuarios por su nivel de agregación y flexibilidad.

### ***Ventajas y desventajas del data mart***

Las ventajas de un *data mart*, según Inmon, Strauss y Neushloss (2008), es el bajo costo de procesamiento al momento de extraer los datos, en comparación a extraerlos directamente del *data warehouse*; la capacidad delegar el procesamiento del *data mart* a otras maquinas para dejar el libre el entorno del *data warehouse*; y el sentido de propiedad que proporciona a los diferentes departamentos, ya que cada uno puede manipular y realizar su propio procesamiento.

Ellos también indican que la desventaja de los *data marts* es que los datos del *data warehouse* suelen ser más detallados. El *data warehouse* permite a las personas tener una forma en común de visualizar los datos, lo cual resulta ser eficiente y conveniente, porque de esta manera siempre están estructurados y en el formato de la manera en que uno quiere verlos. El usuario no tiene que preocuparse acerca de la labor de la obtención detallada y de estructuración.

### ***Tipos de data marts***

Dependiendo de la fuente de los datos, los *data marts* pueden clasificarse como independientes o dependientes. Los *data marts* independientes se obtienen de los datos capturados de uno o más sistemas operativos, de proveedores de información externos, o de los datos generados a nivel local dentro de un departamento o área geográfica. Los *data marts* dependientes se obtienen directamente del *data warehouse* de la organización (Han & Kamber, 2006).

### ***Construcción del data mart***

Silvers (2008) afirma que un *data mart* puede elaborarse por dos métodos básicos. El primer método consiste en la creación de un conjunto físico de bases de datos y tablas, las cuales se encuentran en una plataforma separada del *data warehouse*. Este método

proporciona el máximo aislamiento posible de los datos, sin embargo resulta ser muy caro ya que se requiere una plataforma independiente, aplicaciones para la transportación y mantener la plataforma.

El segundo método consiste en definir un conjunto de interfases que basan sus datos en el *data warehouse*. Si bien este método no incrementa los gastos generales y no requiere de una plataforma independiente, una interfase de un *data mart* no es igual que un *data mart* independiente, debido a que la interfase necesita compartir los recursos con el *data warehouse*.

Analizando los dos métodos de construcción, los recursos y el nivel de intrusión en el *data warehouse* es lo que los diferencia entre sí.

#### **2.9.4. Data mining**

Ponniah (2001) sostiene que el *data mining* se ajusta bien y juega un papel importante en el entorno del *data warehouse*. Un *data warehouse* limpio y completo forma la base para el *data mining*, ya que el primero permite que las operaciones del segundo se efectúen exitosamente. Los siguientes son algunos de los principales factores de esta relación.

- Los algoritmos de *data mining* necesitan grandes cantidades de datos, más que en el nivel de detalle. La mayoría de los *data warehouses* contienen los datos en el nivel más bajo de granularidad.
- El *data mining* integra y limpia los datos. Si las funciones de ETL del *data warehouse* se llevaron a cabo correctamente, es muy conveniente para el *data mining*.

- La infraestructura del *data warehouse* ya es robusta, con tecnología para procesamiento paralelo y potentes sistemas de bases de datos relacionales. Debido a que el *hardware* de este es escalable, no son necesarias nuevas inversiones para apoyar al *data mining*.

Las dos tecnologías se apoyan mutuamente, que adicionalmente de los factores mencionados, de acuerdo a Han y Kamber (2006), normalmente cuanto más tiempo un *data warehouse* se ha utilizado, más se han desarrollado, y es más benéfico para las herramientas de *data mining* empleadas para el descubrimiento de conocimiento y la toma de decisiones estratégicas.

En la presente sección se describen los aspectos clave del *data mining*, en donde se incluyen los beneficios, la definición, la clasificación de los sistemas, los patrones y las técnicas que pueden ser encontrados utilizando esta tecnología.

### ***Beneficios del data mining***

Aún tratándose de una tecnología de uso moderado en la comunidad científica, ésta tiene amplias aplicaciones en el ámbito comercial (Ponniah, 2001). Listando sus beneficios, Rainardi (2008) menciona los siguientes:

- Conocer que productos pueden ser comprados en conjunto, ya sea mediante el análisis de los datos, teniendo en cuenta la probabilidad, o mediante el análisis de los datos del pedido.
- Conocer el patrón entre la delincuencia, el lugar, la tasa de delincuencia y de diversos factores, en un esfuerzo por reducirla.
- La puntuación del cliente en términos de lealtad y poder de compra, en base a sus pedidos, atributos geográficos y demográficos.

- La puntuación de crédito en la industria de las tarjetas de crédito para clientes de acuerdo con las actitudes acerca de la exposición al riesgo, los comportamientos préstamos y la capacidad para pagar sus deudas.
- Investigar la relación entre los tipos de clientes y los productos y/o servicios; en un esfuerzo para crear nuevos y para diseñar una estrategia de marketing.
- Creación de un patrón de llamadas en el sector de las telecomunicaciones, en términos de segmentos de tiempo y área geográfica (diaria, semanal, mensual y estacional de los patrones) con el fin de administrar los recursos de la red (ancho de banda, el mantenimiento, y atención al cliente) adecuadamente.

Se puede observar que la mayoría de las aplicaciones tienen como meta el sector comercial en las áreas de segmentación de clientes, análisis del mercado, detección del fraude, monitoreo de actividades delictivas, y predicción de la demanda.

### ***Definición del data mining***

Según Rainardi (2008), el *data mining* es un proceso de exploración de datos para encontrar los patrones y las relaciones que los describan y predecir lo desconocido o el valor futuro de los datos. El punto clave en el *data mining* es su capacidad de entender por qué algunas cosas sucedieron en el pasado, y predecir lo que sucederá en el futuro.

Para Han y Kamber (2006), el *data mining* es el proceso de descubrimiento de conocimientos interesantes de grandes cantidades de datos almacenados en las bases de datos, los *data warehouses* o los repositorios de información.

Al respecto Kimball y Ross (2002) definen que el *data mining* es una clase de consultas no dirigidas, a menudo en contra de los datos más atómicos, que trata de encontrar patrones inesperados en ellos.

Por su parte, Silvers (2008) define al *data mining* como una búsqueda de patrones y asociaciones de datos que no son inmediatamente evidentes, o pueden estar ocultos por completo.

Ponniah (2001) afirma que el *data mining* es un proceso de descubrimiento de conocimiento, que ayuda a comprender el contenido de los datos de una forma especialmente insospechada, ya que descubre patrones y tendencias en bruto que no se sabía que existían.

Finalmente, se puede definir al *data mining* como un proceso orientado a la búsqueda de conocimiento, que se apoya en la búsqueda de patrones, asociaciones y correlaciones, encuentra datos que a menudo no se visualizan.

### ***Clasificación de los sistemas de data mining***

Los sistemas de *data mining*, según explican Han y Kamber (2006), pueden clasificarse según el tipo de bases de datos, conocimiento, técnicas utilizadas, y aplicaciones adaptadas.

- 1. Clasificación según el tipo de bases de datos:** Un sistema de *data mining* se pueden clasificar según el tipo de bases que analiza, tales como modelos de datos, tipos de datos o aplicaciones implicadas, pudiendo requerir cada uno su propia técnica de *data mining*.
- 2. Clasificación según el tipo de conocimiento:** Los sistemas de *data mining* se pueden clasificar según el tipo de conocimiento que es extraído, tales como la caracterización, la discriminación, el análisis por asociación y correlación, la clasificación, el *clustering*, el análisis de valores atípicos, y el análisis evolutivo.

3. **Clasificación de acuerdo a los tipos de técnicas utilizadas:** Los sistemas de *data mining* se pueden clasificar de acuerdo a las técnicas de *data mining* utilizadas, las cuales se pueden describirse de acuerdo al grado de interacción con el usuario del cual se trate, o los métodos empleados para el análisis de datos.
4. **Clasificación según las aplicaciones adaptadas:** Los sistemas de *data mining* también se pueden clasificar de acuerdo a las aplicaciones a las que se adaptan, ya que a menudo se requiere la integración de métodos específicos por aplicación.

Han y Kamber (2006) sostienen que una clara clasificación de los sistemas de *data mining*, ayuda a los usuarios potenciales a distinguir el que mejor se adecua a sus necesidades.

### ***Patrones del data mining***

Los tipos de patrones que pueden encontrarse por *data mining* se categorizan en (Han & Kamber, 2006):

1. **Patrones por frecuencia:** Los patrones por frecuencia, como su nombre lo sugiere, se producen por la frecuencia que existe en los datos, por lo que conducen al descubrimiento de interesantes asociaciones y correlaciones. Hay muchos tipos de patrones de frecuencia, incluyendo conjuntos de elementos, subsecuencias, y subestructuras.
2. **Análisis por asociación:** Se descubren correlaciones entre pares de atributos de valores asociados. Por lo general, las reglas de asociación son descartadas si no proveen un grado mínimo de confianza.
3. **Clasificación y predicción:** La clasificación es el proceso de encontrar un modelo o una función que describa y distinga las clases de datos o conceptos, con el fin de

poder utilizar el modelo para predecir la clase desconocida. El modelo derivado puede ser representado de diversas formas, tales como reglas de clasificación, árboles de decisión, fórmulas matemáticas o las redes neuronales.

4. **Análisis de cluster:** A diferencia de la clasificación y de la predicción, que analizan la clase de objetos de datos, el análisis de cluster analiza los objetos de los datos sin consultar a una clase conocida.
5. **Análisis de valores atípicos:** Una base de datos puede contener objetos de datos que no cumplan con el comportamiento general, o con el modelo de los datos. Estos objetos de datos son valores atípicos. La mayoría de los métodos de minería descartan los valores atípicos por considerarlos como ruido o excepciones.
6. **Análisis evolutivo:** El análisis evolutivo de datos describe y modela regularidades o tendencias para los objetos cuyo comportamiento cambia con respecto al tiempo.

Han y Kamber (2006) afirman que en ocasiones, los usuarios no tienen idea de los tipos de patrones en sus datos que pueden ser interesantes, por ello los múltiples patrones de *data mining* permiten cubrir las expectativas de los diferentes usuarios y aplicaciones.

### ***Técnicas del data mining***

Las técnicas utilizadas por el *data mining*, acorde a Ponniah (2001), se clasifican en seis: *clustering*, árboles de decisión, razonamiento basado en memoria, análisis por enlace, redes neuronales, y algoritmos genéticos. En la figura 2.6 se detallan las técnicas.

<u>Data Mining Technique</u>	<u>Underlying Structure</u>	<u>Basic Process</u>	<u>Validation Method</u>
<b>Cluster Detection</b>	Distance calculations in n-vector space	Grouping of values in the same neighborhood	Cross validation to verify accuracy
<b>Decision Trees</b>	Binary Tree	Splits at decision points based on entropy	Cross validation
<b>Memory-based Reasoning</b>	Predictive structure based on distance and combination functions	Association of unknown instances with known instances	Cross validation
<b>Link Analysis</b>	Based on linking of variables	Discover links among variables by their values	Not applicable
<b>Neural Networks</b>	Forward propagation network	Weighted inputs of predictors at each node	Not applicable
<b>Genetic Algorithms</b>	Not applicable	Survival of the fittest on mutation of derived values	Mostly cross validation

*Figura 2.6 Marco comparativo de técnicas del data mining*

*Fuente: Data Warehousing Fundamentals (Ponniah, 2001)*

En todas las técnicas se busca elaborar patrones, asociaciones y correlaciones, en donde la diferencia radica en hacer o no distinción entre variables independientes y dependientes, clasificar los datos en grupos de acuerdo a un conjunto de reglas, y la imitación de estructuras de pensamiento.

Resumiendo tenemos que la herramienta del *data warehouse* tuvo que pasar por todo un proceso evolutivo, dando origen a una solución enfocada en mostrar una sola versión de los datos. Indudablemente esta tecnología es de gran impacto para la toma de decisiones, y las tendencias son del todo favorables. Es probable que en el futuro se presente bajo diferentes esquemas, tales como un servicio, código libre, o SOA, sin embargo su capacidad de entender, organizar, y utilizar los datos son ventajas de gran

impacto difíciles de obtener con otra herramienta. La implementación continuará siendo compleja por la interacción con los procesos que involucran a los *data marts*, ETL, OLAP, y el *data mining*, por lo que obtener el máximo beneficio posible dependerá de la definición de la arquitectura tecnológica.

### **3. Metodologías para la implementación de un *data warehouse***

Un *data warehouse* tiene elementos especiales que incrementan la complejidad y los riesgos del proceso de implementación, es por eso que la aplicación de una metodología resulta esencial.

En ocasiones el equipo del proyecto cree que no es necesario definir un plan para el proceso de implementación, pero son los usuarios finales los que determinan su éxito o fracaso. En la presente sección se exponen las metodologías de distintos autores para implementar un *data warehouse*.

#### **3.1. Metodología de Desarrollo en Espiral y el Enfoque de los Siete Flujos**

El enfoque de siete flujos, de acuerdo a Inmon, Strauss y Neushloss (2008), es una estructura para diseñar un proyecto de *data warehouse* que utiliza aspectos de la metodología de desarrollo en espiral.

La metodología de desarrollo en espiral, conforme a Pressman (2005), es un proceso de *software* donde se construyen prototipos de forma iterativa, controlada y sistemática. Cuando se aplica esta metodología el *software* se desarrolla en una serie de entregables, produciendo versiones cada vez más completas en las últimas iteraciones.

Un “flujo” es un conjunto de actividades relacionadas, que al combinarse con la metodología de desarrollo en espiral, producen un mapa completo de todas las actividades del proyecto como se muestra en la Figura 3.1.

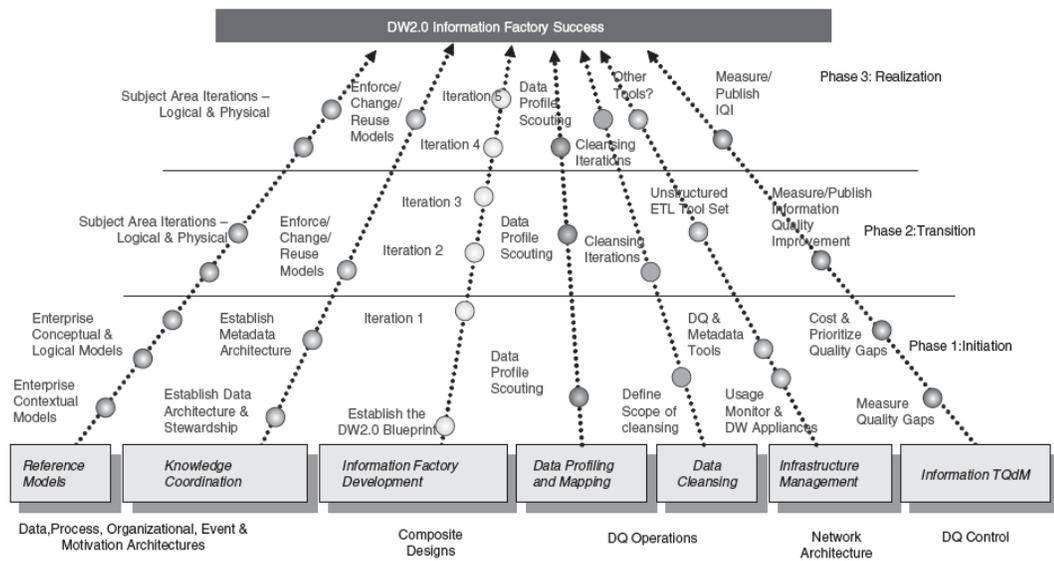


Figura 3.1 Mapa de una proyecto de data warehouse basado en el enfoque de los siete flujos

Fuente: DW 2.0: The Architecture for the Next Generation of Data Warehousing  
(Inmon, Strauss & Neushloss, 2008)

Cada “flujo” produce información clave para definir los entregables en base a las prioridades de la organización. En seguida se explican cada uno de los “flujos”.

1. **Flujo de los modelos de referencia:** El primer flujo se refiere a la creación y el mantenimiento del modelo de datos de la organización. Este modelo es construido de manera iterativa.
2. **Flujo de coordinación del conocimiento:** Es el encargado de la administración de los “artefactos”, con el fin de tener datos confiables de la organización. Un artefacto es el resultado de la identificación de datos originados de los flujos de los modelos de referencia, de la fábrica de desarrollo de la información, y del perfil y mapeo de datos.

3. **Flujo de la fábrica de desarrollo de información:** El siguiente flujo, como su nombre lo indica, es donde se da lugar a la construcción de la información.
4. **Flujo de perfil y mapeo de datos:** En este flujo es donde los sistemas son examinados para entender a los datos en términos de calidad e integridad. La salida del flujo permite mapear los datos al flujo de la fábrica de desarrollo de la información.
5. **Flujo de corrección de datos:** Este flujo involucra a la fuente de datos para determinar que datos necesitan ser corregidos, completados o modificados de acuerdo a un conjunto de reglas.
6. **Flujo de infraestructura:** En este flujo se definen las actividades de soporte de la infraestructura tecnológica, tomando en consideración a la gente, recursos, plataformas, herramientas, políticas, estándares y procedimientos.
7. **Flujo de la administración total de la calidad de la información:** Finalmente, en este flujo se monitorea la calidad de los datos y los procesos.

La metodología del desarrollo en espiral forma parte del enfoque de los siete flujos, dando como resultado los siguientes beneficios:

- La aceleración del proceso de evaluación del modelo de datos corporativos
- La coordinación de los conocimientos de la empresa
- El desarrollo de perfiles y el mapeo de los datos
- El proceso de limpieza de datos
- La administración de la infraestructura de forma preventiva
- La administración total de calidad de los datos

Al obtenerse este conjunto de beneficios, disminuye la cantidad de obstáculos en los proyectos de *data warehouse*, ya que las anomalías generadas por los modelos de datos, normas y definiciones pueden conocerse por adelantado.

### 3.2. Metodología del Ciclo de Vida Dimensional del Negocio

Kimball y Ross (2002) sostienen que el uso de un mapa es extremadamente útil al dirigirse a un rumbo desconocido, como en un proyecto de *data warehouse*. Ellos desarrollaron esta metodología en base a tres principios clave para una implementación. Primero, los proyectos deben enfocarse en las necesidades del negocio. Segundo, los datos presentados a los usuarios deben estar en la dimensión del negocio. Finalmente, el proyecto debe realizarse en un tiempo determinado.

La figura 3.2 muestra las doce actividades que se realizan bajo la metodología del ciclo de vida dimensional del negocio.

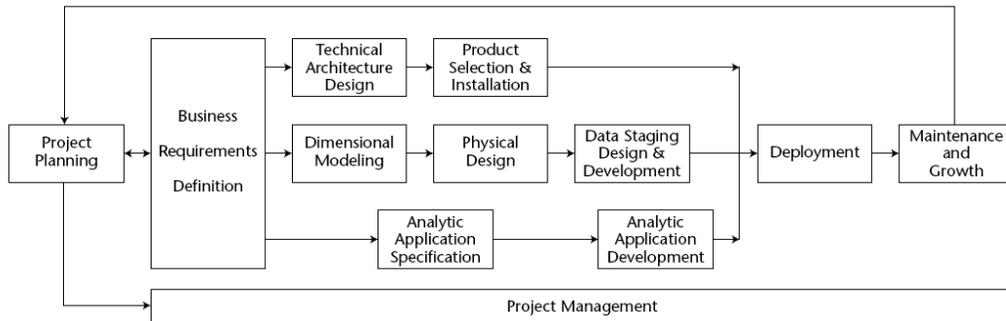


Figura 3.2 Diagrama del ciclo de vida dimensional del negocio

Fuente: *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling*  
(Kimball & Ross, 2002)

Enseguida se explica a detalle cada una de las actividades.

1. **Planeación del proyecto:** Se evalúa la voluntad de la organización con respecto a una iniciativa de *data warehouse*, estableciendo un alcance y justificación, obteniendo recursos y presentando un plan preliminar del proyecto.
2. **Administración de proyectos:** Se sigue una disciplina de administración de proyectos durante todas las actividades de la implementación.
3. **Definición de los requerimientos del negocio:** Se alinea el proyecto de *data warehouse* con los requerimientos del negocio, en donde es crucial entender las necesidades de sus usuarios y trasladarlas al diseño.
4. **Diseño técnico de la arquitectura:** Establece la estructura completa que soportará al *data warehouse*.
5. **Selección e instalación del producto:** Utilizando el diseño técnico de la arquitectura se evalúan y selecciona el producto.
6. **Modelación dimensional:** Se trasladan los requerimientos del negocio a un modelo dimensional.
7. **Diseño físico:** Se utiliza el modelo dimensional para la elaboración de una estructura física que incluya estrategias para mejorar rendimiento.
8. **Diseño y despliegue del área de concentración de datos:** En base al diseño físico, se diseñan y desarrollan los procesos de ETL.
9. **Especificación de las aplicaciones analíticas:** Tomando los requerimientos del negocio, se diseñan las aplicaciones que responderán a las necesidades analíticas de los usuarios.
10. **Despliegue de aplicaciones analíticas:** Se desarrollan las aplicaciones analíticas definidas en el punto anterior.
11. **Despliegue:** Se procede a unir la tecnología, los datos y las aplicaciones analíticas para arrancar el servicio de *data warehouse*.

**12. Mantenimiento y crecimiento:** Se proporciona mantenimiento al *data warehouse* para satisfacer las necesidades de la comunidad de usuarios. Adicionalmente, se evalúa continuamente su capacidad de crecimiento.

La metodología ha resultado ser un éxito porque concentra los procesos clave y las mejores prácticas de un proceso de implementación.

### 3.3. Metodología de Cascada Bajo el Enfoque de Rainardi

Rainardi (2008) propone una adaptación de la metodología de cascada para un proyecto de *data warehouse*, la cual incluye los procesos de manejo del repositorio de datos, el sistema ETL y las aplicaciones de usuario final. En la figura 3.3 se ilustran las fases que componen a la metodología.

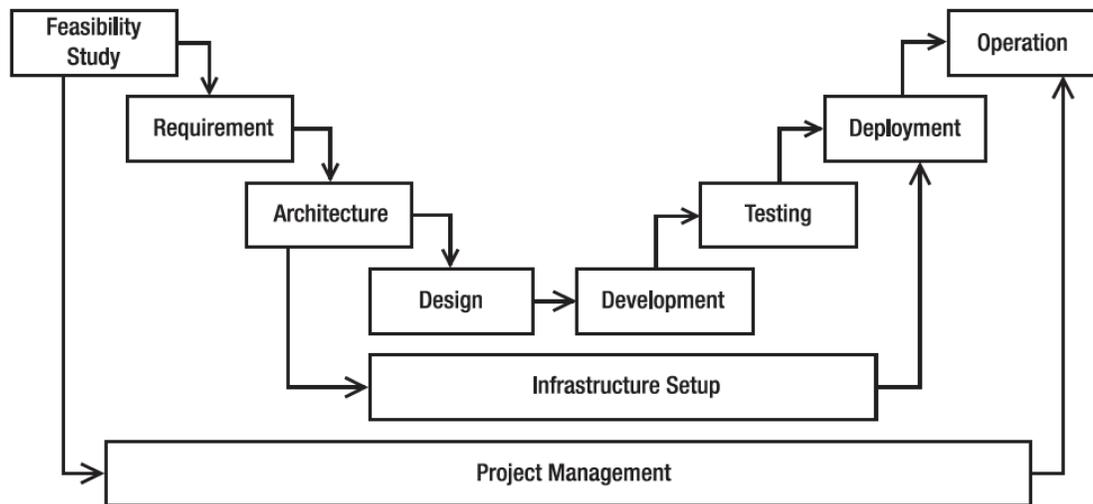


Figura 3.3 Metodología de cascada con el enfoque de proyectos de data warehouse

Fuente: *Building a Data Warehouse: With Examples in SQL Server* (Rainardi, 2008)

Se puede observar que el modelo lo componen diez fases, del cual sobresalen la fase de configuración de la arquitectura y la de administración de proyectos. Los detalles de cada fase se listan en seguida.

- 1. Estudio de viabilidad:** En esta fase se analizan los requerimientos y el volumen de datos desde una perspectiva de alto nivel. Además se determina la arquitectura y las herramientas adecuadas, y se obtiene una proposición consolidada.
- 2. Requerimientos:** Se habla con los usuarios para entender los detalles de los procesos del negocio, datos y otros factores importantes. Es aquí donde se listan los requerimientos funcionales y no funcionales definidos durante las entrevistas.
- 3. Arquitectura:** Se determina que arquitectura se utilizará, incluyendo la especificación de los servidores de bases de datos, el tipo de red, la solución de almacenamiento, entre otros.
- 4. Diseño:** En esta fase se diseñan las tres partes principales del *data warehouse*; el repositorio de datos, el sistema ETL y las aplicaciones de usuario final.
- 5. Desarrollo:** En base al diseño de la fase anterior se procede a la construcción del repositorio de datos, el sistema ETL y las aplicaciones de usuario final.
- 6. Pruebas:** Dado que es la primera vez que los usuarios finales interactúan con el *data warehouse* y utilizan el ambiente de producción y pruebas, esta fase se encarga de arreglar los problemas de los componentes desarrollados hasta el momento.
- 7. Despliegue:** Una vez que el sistema está listo, se procede a integrar todos los componentes al ambiente de producción. Como resultado de la fase se crea una guía de usuario, de operación, y de solución a problemas.
- 8. Operación:** En esta fase se maneja la interacción de los usuarios del *data warehouse* y sus aplicaciones.

**9. Configuración de la infraestructura:** Esta fase consiste en la elaboración de la arquitectura del sistema, conectar las redes, y probar y documentar los detalles de la infraestructura. Es común que muchas personas subestimen la complejidad de la fase, ocasionando retrasos en el proceso de implementación.

**10. Administración del proyecto:** Durante esta fase se mantiene un plan del proyecto, una buena comunicación con los accionistas y se atacan agresivamente los riesgos.

A pesar de tratarse de una variante de la metodología más antigua de desarrollo de *software* su aceptado uso facilita su rápida adopción.

Concluyendo, cada una de las metodologías antes mencionadas definen un conjunto de actividades para implementar exitosamente un *data warehouse*. Los autores, bajo su propio enfoque, sugieren una forma particular de realizar el proyecto, pero todos incluyen actividades de planeación, modelación, construcción y despliegue. Es importante destacar que el éxito de un proyecto no depende exclusivamente del capital humano o de los productos, sino del procedimiento para la elaboración de las actividades.

#### **4. Factores críticos a considerar en la implementación de un *data warehouse***

De acuerdo a R. Crandall y W. Crandall (2008), a medida que la competencia se hace más difícil, es vital para las empresas se enfoquen en sus actividades principales. Las empresas han descubierto que no pueden ser competitivos en todo, por lo tanto, necesitan decidir que tareas deben hacer excelentemente.

Los factores críticos de éxito son una alternativa que proporcionará a la empresa una ventaja competitiva por un período limitado. Los autores definen a los factores críticos de éxito como un concepto formal del proceso de establecer y mantener las prioridades organizacionales. A mayor detalle, afirman que son eventos internos o externos que pueden afectar a la empresa positiva o negativamente, y por lo tanto requieren especial atención.

Durante la implementación de cualquier proyecto de TI, conforme a Wixom y Watson (2001), hay diversos factores que desempeñan un papel crítico para su éxito, y más aún en el caso del *data warehouse*, que cuenta con características únicas.

Aunque el concepto de *data warehouse* está bien establecido, según Bitterer (2005), muchas organizaciones experimentan los mismos problemas en el proceso de implementación. Aunque la escalabilidad de volumen de datos se ha dado en pocos años, las organizaciones necesitan investigar si sus infraestructuras son lo suficientemente capaces de abordar los retos de administración de volumen, velocidad y variedad de datos.

Beyer, Feinberg y Berg (2008) afirman que un *data warehouse* es una colección de proyectos que operan bajo una misma administración de datos y una iniciativa de integración. Algunos de estos proyectos serán simplemente nuevas fuentes de datos y

*data marts* dependientes, y otros se centrarán en su actualización, pero no cabe duda que el *data warehouse* se está convirtiendo en un elemento central en la administración de la información empresarial.

Dada la importancia de los factores críticos para la implementación de un *data warehouse* para este estudio, se expone el punto de vista de diferentes autores con respecto al tema.

#### **4.1. Enfoque Hwang, Ku, Yen, y Cheng**

Hwang, Ku, Yen y Cheng (2003) realizaron un estudio para identificar los factores críticos en la adopción de la tecnología del *data warehouse* en la industria bancaria de Taiwán. Para la investigación ellos propusieron el modelo de la figura 4.1, el cual muestra una serie de hipótesis relacionadas con los factores que consideraron relevantes para este caso.

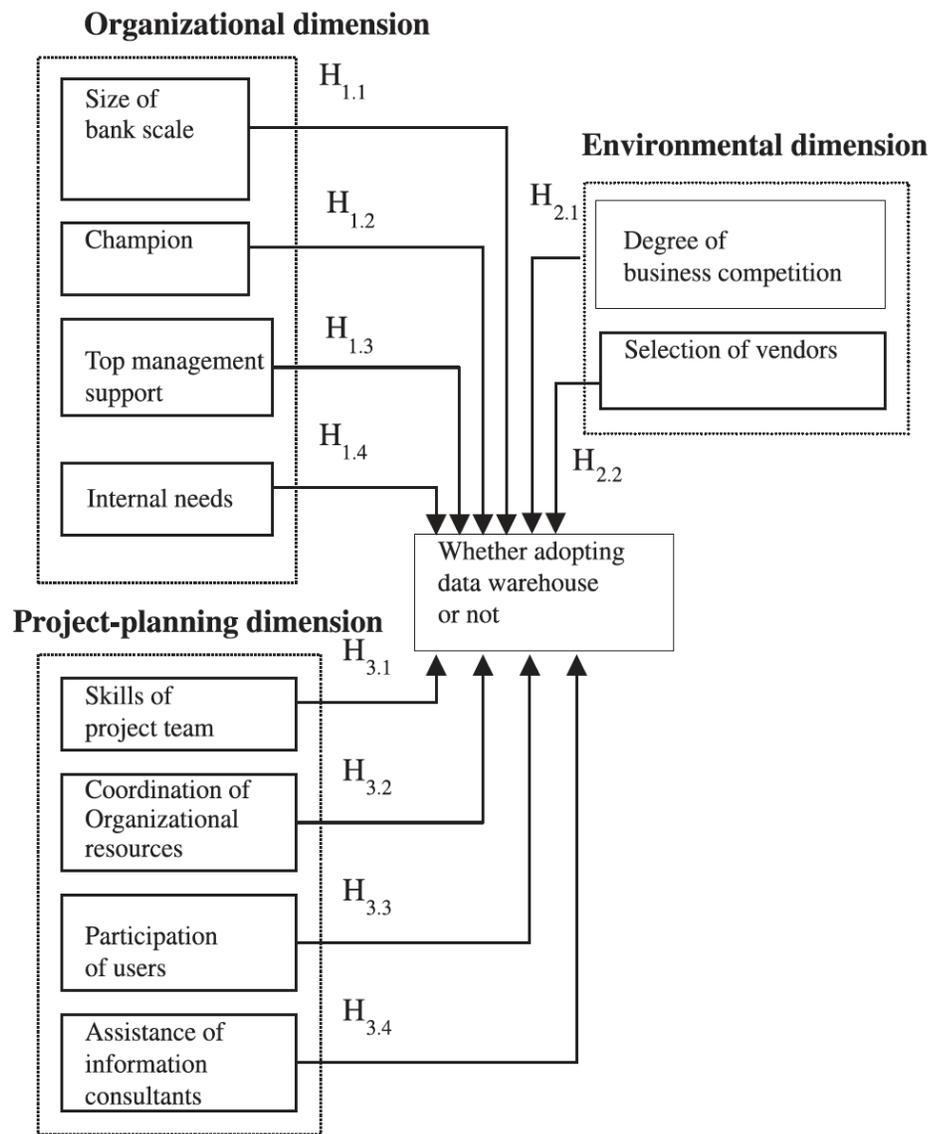


Figura 4.1 Modelo de investigación

Fuente: Critical factors influencing the adoption of data warehouse technology

(Hwang, Ku, Yen & Cheng, 2003)

En la figura 4.1 se puede observar como los factores se agrupan en tres dimensiones. La dimensión organizacional trata con los factores internos que directamente pueden afectar la adopción del *data warehouse*; la dimensión de planeación de proyectos evalúa

factores relacionados al proyecto, debido a la gran inversión involucrada; y finalmente, la dimensión ambiental trata con los factores relacionados al contexto. A continuación se mencionan los detalles de las dimensiones, los factores analizados, y los resultados encontrados después del estudio.

## **1. Dimensión organizacional**

Para cualquier organización, la adopción de la tecnología de *data warehouse* es un trabajo muy costoso y complejo. Además, la dificultad en el aprendizaje de una nueva tecnología dará lugar a la resistencia de los usuarios durante el proceso. Como resultado de ello, es esencial asegurar el soporte y el compromiso de la administración, y fomentar la participación entusiasta de los usuarios. Al considerar todos los factores organizacionales durante la ejecución, los obstáculos y las barreras se lograrán reducir. Los factores de la dimensión organizacional analizados fueron:

### **a. El tamaño del banco afecta a la adopción de la tecnología de *data warehouse*.**

Entre más grande sea el tamaño de la organización, mayor será la cantidad de recursos que pueden ser asignados para la adopción de una nueva tecnología.

### **b. La existencia de campeones afecta a la adopción de la tecnología de *data***

***warehouse*:** La presencia de personas que aprecian las tecnologías emergentes puede estimular e influenciar directamente la adopción.

### **c. El soporte de la administración afecta a la adopción de la tecnología de *data***

***warehouse*.** En numerosos estudios se comprobó que el soporte de la administración puede disminuir la dificultad y complejidad de los problemas durante la adopción.

### **d. Las necesidades internas del banco afectan a la adopción de la tecnología de**

***data warehouse*.** Para adoptar una nueva tecnología, los ejecutivos deben

comprender las necesidades internas de su organización, ya que pueden influenciar positiva o negativamente.

Todos los factores de la dimensión organizacional resultaron importantes desde una perspectiva práctica, además resultaron ser los de mayor nivel de relevancia con respecto a las otras dimensiones.

## **2. Dimensión ambiental**

En el entorno dinámico actual, las organizaciones deben hacer todo lo posible para medir y reducir la incertidumbre, y por lo tanto crear ventajas competitivas. Los factores de la dimensión ambiental analizados fueron:

- a. El grado de competencia entre las empresas afecta a la adopción de la tecnología del *data warehouse*:** La adopción de una nueva tecnología puede proveer a las empresas de ventajas competitivas con respecto a sus competidores.
- b. La selección de los proveedores afecta a la adopción de la tecnología del *data warehouse*:** Las empresas deben de evaluar la posibilidad de desarrollar el *data warehouse* en formato *outsourcing* o al interior de la organización. Si optan por *outsourcing*, resulta importante seleccionar correctamente a los proveedores.

El estudio descubrió que el grado de competencia para implementar nuevas tecnologías es un factor influyente, contrario a la selección de proveedores de *data warehouses*. El grado de competencia resultó importante debido a la fuerte disposición de los bancos para adoptar nuevas tecnologías de información; contrario a la selección de proveedores, la cual no resultó relevante debido a la cantidad limitada de proveedores de esta tecnología en Taiwán.

### 3. Dimensión de planeación de proyectos

La dimensión de planeación de proyectos está asociada a resolver todos los problemas o cuestiones relacionadas sobre el análisis, desarrollo, administración y control de proyectos en el proceso de adopción. Los factores analizados de esta dimensión fueron:

- a. **Las habilidades del equipo del proyecto afectan a la adopción de la tecnología del *data warehouse*:** El equipo del proyecto juega un papel muy importante durante la planeación, ya que su nivel de conocimientos puede influenciar positivamente en la adopción de una nueva tecnología.
- b. **La coordinación de los recursos de la organización afecta a la adopción de la tecnología del *data warehouse*:** La adopción de una nueva tecnología puede incurrir en grandes gastos y un alto consumo de tiempo, siendo de importancia tener una adecuada coordinación de los recursos.
- c. **La participación de los usuarios finales afecta a la adopción de la tecnología del *data warehouse*:** La participación de los usuarios puede ayudar a cumplir con sus demandas y expectativas con respecto a la tecnología.
- d. **La asistencia de consultores de la información afecta a la adopción de la tecnología del *data warehouse*:** Los consultores pueden proveer de ideas o asistencia a las organizaciones, lo cual disminuye los riesgos involucrados en la adopción de la tecnología.

Los resultados revelaron que aunque los factores de esta dimensión son muy importantes para desarrollar cualquier tecnología, estos no influyen considerablemente durante la adopción.

Bajo el enfoque de Hwang, Ku, Yen y Cheng (2003), se pueden observar los factores de mayor influencia en la industria bancaria asiática, sin embargo los hallazgos encontrados sirven de referencia para otras industrias que están considerando adoptar el *data warehouse*.

#### **4.2. Enfoque Ponniah**

Ponniah (2001) propone una serie de factores críticos de éxito, que están conformados por:

- No iniciar el *data warehouse* a menos que la compañía esté preparada para ello.
- Encontrar al mejor patrocinador ejecutivo que garantice la continuidad y el apoyo a largo plazo.
- Hacer hincapié en los aspectos del negocio, no en los tecnológicos.
- Elegir un administrador de proyectos orientado al negocio.
- Tomar una visión amplia de los requerimientos de la empresa.
- Tener un enfoque pragmático y organizado durante la ejecución.
- Comunicar expectativas realistas a los usuarios.
- No ir demasiado lejos al justificar y predecir el retorno de la inversión.
- Institucionalizar los métodos apropiados y eficaces de comunicación.
- A lo largo del ciclo de vida del proyecto, mantenerlo como un esfuerzo en conjunto entre la tecnología y los usuarios.
- Adoptar las tecnologías probadas; evitar las tecnologías muy recientes.

- Reconocer la importancia fundamental de la calidad de los datos.
- No ignorar el potencial de las fuentes de datos externas.
- No subestimar el tiempo y el esfuerzo para las funciones de extracción, transformación y carga de datos (ETL).
- Seleccionar la arquitectura correcta para el entorno de la organización.
- La arquitectura va primero, después la tecnología, y hasta entonces las herramientas.
- Determinar claramente una estrategia de capacitación.
- Tener cuidado con la "parálisis por análisis".
- Comenzar el despliegue con un piloto adecuado que visualice los beneficios rápidamente.
- No descuide la importancia de la escalabilidad.
- Centrarse en las consultas, no en las transacciones.
- Definir y administrar claramente las consideraciones de propiedad de datos.

Se puede apreciar que el enfoque se centra en tratar aspectos relacionados con la organización, los usuarios, y con la buena administración de un proyecto.

### **4.3. Enfoque Wixom y Watson**

Wixom y Watson (2001) incluyeron siete factores en su modelo de investigación para el éxito de la implementación de un *data warehouse*. Enseguida se listan los factores.

1. Apoyo de la administración
2. Campeón
3. Recursos
4. Participación de los usuarios
5. Las habilidades de equipo
6. Sistemas fuente
7. Desarrollo de tecnología

El enfoque considera en mayor parte los factores que se relacionan con aspectos organizacionales y de planeación del proyecto, sin entrar a detalles al lado técnico que se requiere para la implementación del *data warehouse*.

Contextualizando, los autores coinciden que los factores críticos para la implementación del *data warehouse* son los relacionados con aspectos de la organización, el apoyo de la administración, los usuarios, y los tecnológicos. Definitivamente cada enfoque se ve influenciado por la cultura y experiencia de los autores, pero sirven como marco de referencia para conseguir una implementación exitosa.

Recordando los capítulos anteriores, la tecnología del *data warehouse* ha evolucionado de tarjetas perforadas a un ambiente sofisticado que permite el almacenaje de información confiable y oportuna. Hay una tendencia generalizada por adoptar herramientas de inteligencia de negocios, que tienen su fundamento en esta tecnología. En ocasiones es subestimada la relevancia de los procesos que involucran la extracción, transformación y carga de datos, sin embargo la calidad de las consultas dependen de

los resultados de estos. Un metodología de implementación adecuada a los proyectos del *data warehouse* disminuye la resistencia al cambio, dando lugar a una mejor adopción por parte de los usuarios finales.

## 5. Proveedores de herramientas de *data warehouse*

El mercado del *data warehouse*, según afirma Feinberg (2009b), ha evolucionado de diferentes maneras. Aunque la atención está centrada en los grandes proveedores, los participantes más pequeños con poca o ninguna experiencia en *data warehouse*, ofrecen un enfoque más preciso con soluciones innovadoras.

Feinberg (2009b) propone clasificar a los proveedores mediante la herramienta de Cuadrantes Mágicos (en inglés, *Magic Quadrants*), por lo que resulta importante definirlos.

De acuerdo a Hawkins (2008), evaluar al mercado y a sus participantes resulta una tarea complicada, debido a la diferencia entre los proveedores con respecto a su tamaño, nivel de complejidad y estrategia. Los *Magic Quadrants* solucionan estos problemas al ofrecer una “fotografía” del mercado y de sus participantes, permitiendo al cliente poder clasificar las fortalezas de los proveedores conforme a sus necesidades.

Definiéndolos puntualmente, los *Magic Quadrants* representan al mercado mediante una matriz bidimensional que evalúa a los proveedores en base a la integridad de su visión, capacidad de respuesta, desarrollo del producto, canales de venta, clientes base, habilidad de ejecución y el grado de influencia en el mercado. En la figura 5.1 se muestran estos criterios.

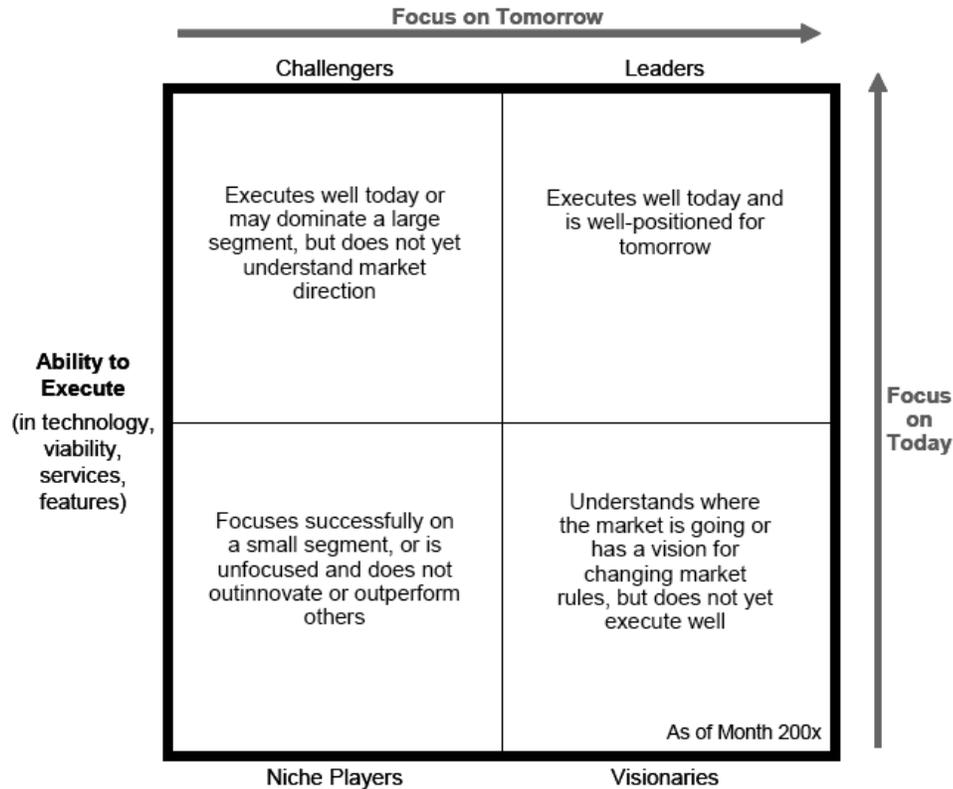


Figura 5.1 Magic Quadrants

Fuente: *Magic Quadrants and MarketScopes: How Gartner Evaluates Vendors Within a Market* (Hawkins, 2008)

Una vez comparados los criterios con las características de cada proveedor, éstos se pueden agrupar en alguno de los cuatro cuadrantes. En detalle, los cuadrantes se componen de:

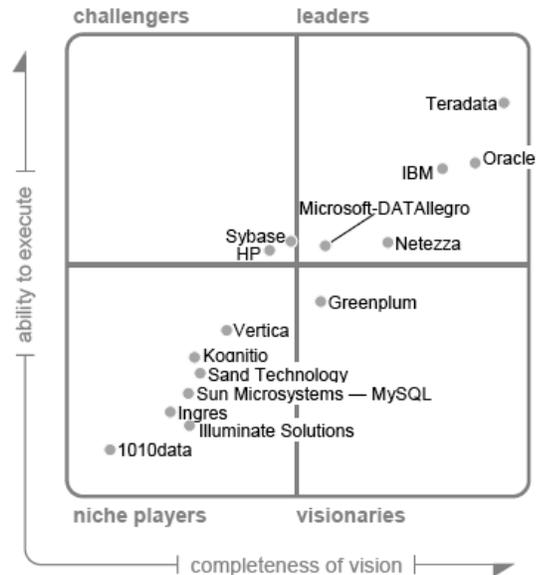
1. **Líderes:** Los líderes proponen ofertas maduras que se adecuan a la demanda del mercado, así como demuestran tener una visión para mantener su posición. Al observar a los líderes se puede entender como evolucionarán los productos de *data warehouse*.

2. **Retadores:** Los retadores tienen una fuerte capacidad de ejecución, pero no tienen un plan para mantener su posición y atraer a nuevos clientes. Los retadores se pueden convertir en líderes si desarrollan su visión.
3. **Visionarios:** Los visionarios se alinean conforme a la evolución del mercado, pero no han probado tener las capacidades necesarias para alcanzar esa visión. Para que lleguen a convertirse en retadores o líderes, dependerá si tienen la capacidad de desarrollar nuevas tecnologías atractivas para los clientes o si logran alianzas estratégicas para complementar sus fortalezas.
4. **Jugadores del nicho:** Son proveedores que se desempeñan bien en un segmento del mercado, pero tienen una habilidad limitada para innovar o superar a otros competidores. Si fueran capaces de establecer una visión entorno a la dirección del mercado, pudieran crecer o convertirse en visionarios.

Localizada la posición del proveedor en el *Magic Quadrant*, el cliente puede encontrar una oportunidad de inversión que se adecue a la estrategia de su negocio. Hay que considerar que enfocarse en el cuadrante de los líderes no siempre es lo mejor, ya que hay buenas oportunidades para los retadores, además de que los jugadores del nicho pueden operar mejor que un líder del mercado.

Una vez comprendido el funcionamiento de la herramienta, en la figura 5.2 se puede observar la posición que ocupa cada proveedor en el mercado de *data warehouse* (Feinberg, 2009b).

# Data Warehouse: Today's Decision



As of December 2008

From "Magic Quadrant for Data Warehouse Database Management Systems," 23 December 2008

Gartner.

Figura 5.2 Magic Quadrant para los DBMS de Data Warehouse del 2008

Fuente: *The Future of DBMS for Data Warehousing* (Feinberg, 2009b)

Para tener una mayor comprensión del *Magic Quadrant*, a continuación Feinberg y Beyer (2008) hacen mención de los aspectos más relevantes de cada proveedor.

1. **1010data:** Proveedor de servicios de *data warehouse* con siete años de experiencia, con sede en la ciudad de Nueva York. Cuenta con una solución integrada por tecnologías de DBMS y BI dirigida a la parte comercial de las organizaciones, principalmente al Sector Financiero. Como tal, tiene un modelo operacional diferente de otros proveedores.

2. **Greenplum:** Greenplum propone un DBMS de *data warehouse* que realiza procesamiento paralelo masivo de datos, se basa en el DBMS de código abierto PostgreSQL, y se puede ejecutar en Linux y Unix.
3. **HP:** HP introdujo Neoview al mercado a principios del 2007. HP Neoview debutó en 2008 en el cuadrante de los “Retadores” y debe demostrar su aceptación en el mercado.
4. **IBM:** IBM sigue siendo un líder en los DBMS de *data warehouse*. En este mercado altamente competitivo y maduro, IBM ha posicionado su oferta como una solución completa de administración de la información, en donde el repositorio de datos es parte clave de su estrategia.
5. **illuminate Solutions:** Es un pequeño proveedor de *software* con sede en España, illuminate Solutions tiene un DBMS de *data warehouse* integrado y herramientas de Inteligencia de Negocios.
6. **Ingres:** Es proveedor de una solución de DBMS de código abierto con 30 años de historia, conocido también por sus sistemas de administración de base de datos relacionales (en inglés, *Relational Data Base Management System*, abreviado *RDBMS*). La empresa tiene muchos clientes que ejecutan aplicaciones críticas, incluyendo las relacionadas al *data warehouse*.
7. **Kognitio:** Hoy en día, tiene una mezcla de clientes que utilizan su DBMS por separado, clientes de DBMS de *data warehouse*, así como el servicio de esta tecnología bajo el esquema de *outsourcing*.

8. **Microsoft-DATAllegro:** Microsoft está surgiendo en este mercado debido a su bajo precio de entrada y los cambios en la demanda. La adquisición de DATAllegro promete un DBMS de *data warehouse* con capacidad de escalabilidad de *hardware*.
9. **Netezza:** Sigue aprovechando su estrategia de aceleración de *hardware* con procesamiento múltiple. Concretamente, trabaja con proveedores de servicios de Internet.
10. **Oracle:** Oracle sigue siendo un líder en *data warehouse*, ocupando más de 48% del mercado de RDBMS. En 2008, Oracle añadió una aplicación de *data warehouse* a sus productos clave.
11. **Sand Technology:** Es un pequeño proveedor de DBMS que ha estado en el mercado aproximadamente ocho años.
12. **Sun Microsystems-MySQL:** MySQL fue adquirida por Sun Microsystems en enero del 2008, y es el DBMS más utilizado de código abierto.
13. **Sybase:** Sybase fue el primero en ofrecer sistemas de DBMS de almacenamiento en columnas. Con un mayor número de participantes en esta tecnología, el mercado parece favorecer su solución.
14. **Teradata:** Ofrece varias aplicaciones de *data warehouse* que combinan *hardware*, sistema operativo y un DBMS.
15. **Vertica:** Es el participante más nuevo en el mercado de DBMS de *data warehouse*. Su oferta incluye una serie de capacidades funcionales que agregan alto rendimiento y disponibilidad.

De lo anterior, conforme a Feinberg (2009b), se observa que Oracle, IBM y Microsoft son los líderes del mercado, continuarán creciendo en el futuro y el número de nuevos participantes se incrementará.

Las tendencias relevantes que involucran a los líderes del mercado son: en primer lugar, Teradata continuará creciendo en gran parte por el soporte a clientes de Unix y Linux; en segundo lugar, los tres proveedores de DBMS en código abierto (MySQL y Ingres) continúan creciendo sustancialmente; y finalmente, aunque Microsoft ocupe el tercer lugar del mercado, su sistema operativo Windows está presente en casi el 50% del mercado.

Cada año el mercado continuará incrementándose por nuevos participantes, que van desde proveedores de nuevo *software* de DBMS a productos integrados de *hardware/software*. DATAlegro fue adquirida por Microsoft en 2008. Ahora, una empresa pública, Netezza ha ganado una buena parte del mercado con más de 200 clientes. MySQL, adquirida por Sun Microsystems en 2008, tiene muchas referencias sólidas de *data warehouse*, la mayoría de sus sistemas se implementan en bases de datos de pequeñas y medianas empresas.

Concluyendo, en el 2009 el mercado de DBMS del *data warehouse* ha recuperado su dinamismo y una renovada competencia entre los proveedores. Tras varios años de liderazgo de los grandes proveedores, los nuevos participantes han empezado a "sacudir las cosas" con algunos productos interesantes, sobre todo la ola iniciada por Netezza. La competencia entre los principales proveedores de DBMS sigue siendo intensa. Además, la tendencia hacia productos de código abierto está presionando los precios de los líderes del mercado.

## 6. Casos de éxito

Una nueva era en la implementación del *data warehouse* está comenzando. Conforme a un estudio realizado por Gartner a empresas financieras, el 25% se encuentra implementándolo por primera o segunda ocasión; mientras que otro 30% está realizando un proceso de optimización de datos o de expansión de aplicaciones (Beyer, 2009).

Complementando lo anterior, resulta conveniente investigar a detalle casos de éxito de empresas del Sector Financiero. En la siguiente sección se muestran los casos del Banco de Baroda y de la Compañía de Servicios Financieros, que mediante un *data warehouse* obtuvieron beneficios significativos.

### 6.1. Caso Banco de Baroda

Raval (2008) afirma que el Banco de Baroda (BoB) es una compañía proveedora de servicios financieros a personas, negocios y corporaciones. El banco opera principalmente en Estados Unidos, Inglaterra y las regiones del Pacífico de Asia. Sus instalaciones principales están ubicadas en Baroda, India, y emplea alrededor de 36,838 personas.

#### ***Problemática***

BoB no contaba con una herramienta con la capacidad de realizar modelos de pronósticos, por lo que implementó un sistema para el apoyo a la toma de decisiones. Los principales retos a superar consistieron en:

- Complementar los datos de la organización, ya que era necesario almacenar datos de al menos cinco años de antigüedad para desarrollar los modelos
- Elaborar un perfil completo del cliente debido a que los datos no se encontraban actualizados, en especial en áreas como la información demográfica de los clientes
- Mejorar la calidad de los datos, ya que los errores provocaban fallas al momento de querer obtener información concisa de los clientes
- Mejorar la visión de la corporación, para pensar en escenarios de tres a cinco años de distancia

Estaba claro que de superar estos obstáculos la organización podría realizar pronósticos o análisis acertados, además de ayudar al banco a visualizar los planes de segmentación del mercado.

### ***Solución***

El banco seleccionó la solución de *data warehouse* de Hewlett-Packard llamada HP Neoview. Inició el proceso de despliegue en Junio del 2008, planeando tener el primer reporte del sistema en Enero del 2009, lo cual aceleró el proceso de implementación. Se optó por dividir el proceso en cinco fases:

- 1. Involucramiento de la organización:** Se involucró a los principales funcionarios, entre los que se encontraban personas proactivas, analistas de tendencias y gente influyente de la organización.
- 2. Aseguramiento de la calidad de los datos:** El banco contrató a una agencia externa que se encargaría de la tarea. Adicionalmente, se implementó un programa de incentivos para el personal encargado de recolectar los datos.

3. **Definición del negocio:** Durante esta etapa el equipo discutió con las unidades de negocio de crédito, depósitos, tarjetas, administración del riesgo, tesorería y unidades internacionales sobre los requerimientos para elaborar los modelos.
4. **Definición de la información:** En este punto se iniciaron los procesos de extracción, registro, mapeo y almacenamiento de la información.
5. **Entrega de producto final:** Finalmente, las fases anteriores generaron los sistemas de administración de la relación con el cliente (en inglés, *Customer Relationship Management*, abreviado *CRM*), de planeación estratégica y de análisis predictivos.

Básicamente, la solución obtenida dio como resultado un modelo de pronósticos basado en diversas fuentes. La compañía basó su información en el sistema principal de la Banca, de administración de recursos humanos, entre otras aplicaciones.

### ***Beneficios***

Después de la implementación del *data warehouse* se mejoró el proceso de toma de decisiones; se mejoró el contenido de los sistemas en lo que respecta a patrones de comportamiento del cliente y tendencias; se obtuvo la capacidad de elaborar un análisis comparativo entre la historia de la organización y las tendencias actuales; se mejoró la administración del banco; y se mejoraron los productos.

### **6.2. Caso Compañía de Servicios Financieros**

Mejor conocida por su nombre en inglés, *Financial Services Company (FSC)*, Watson, Goodhue y Wixom (2002) mencionan que se trata de una empresa con sede en el sureste de Estados Unidos. Opera en seis estados del sur y sureste del país con

aproximadamente 7,500 empleados. Además de ser la principal comercializadora de productos de inversión y de procesamiento de pagos de salud.

### ***Problemática***

En 1999, FSC perdió 60 millones de dólares y funcionaba bajo acuerdos regulatorios. Se comenzó a rumorar que la compañía se declararía pronto en banca rota. En respuesta a esta situación, el Consejo Administrativo reestructuró a su equipo directivo para proporcionar una nueva dirección al banco. El primer paso de este equipo consistió en reducir los gastos.

### ***Solución***

El Consejo Administrativo reconoció que necesitaba una estrategia a largo plazo si quería garantizar la supervivencia de la compañía. FSC tuvo que identificar un nicho de mercado en el cual pudiera sobresalir.

La estrategia seleccionada fue enfocarse entorno a sus clientes de forma excepcional. El nombre dado a esta estrategia fue “Soluciones a la Medida del Cliente” (en inglés, *Tailored Client Solutions*, abreviado *TCS*). Se basó en la premisa de que los productos se tienen que adaptar a las necesidades de los clientes, creando así relaciones más sólidas y obteniendo beneficios sostenibles a largo plazo. Los productos financieros, canales de distribución, y la venta y prestación de servicios se integraron adecuadamente para satisfacer las necesidades y preferencias de los clientes, alcanzando los niveles de rentabilidad planteados. El sistema que proporcionó los datos necesarios para la ejecución del TCS fue un *data warehouse*.

El TCS se introdujo por primera vez en la planeación de la estrategia del banco para 1995. El equipo directivo del FSC aprovechó la reunión para discutir y perfeccionar los planes de crecimiento, llegando a la conclusión de que el banco tenía que ser gestionado de una nueva manera.

El Consejo Administrativo reconoció la falta de habilidades internas de su recurso humano. La estrategia TCS requirió a un nuevo director ejecutivo de mercadotecnia, con experiencia en el desarrollo y la implementación de tecnologías de información. La búsqueda fue larga, pero una vez reclutado formó un equipo de nuevos gerentes y profesionistas. El equipo comenzó con el desarrollo del *data warehouse* VISION, compuesto de:

- **Un módulo de datos transaccionales:** Para cada cliente el banco tenía almacenadas todas sus transacciones bancarias, sin considerar que el proceso fuera mediante retiros de un cajero automático o usando una tarjeta crédito.
- **Un módulo de datos demográficos:** El banco creó un perfil demográfico de sus clientes, combinando datos de objetivos de inversión, género, edad, código postal y datos externos.
- **Un módulo de datos de comportamiento.** A través del análisis de más de 3,000 clientes, el banco pudo tener una mejor comprensión de sus preferencias.
- **Un módulo de datos de ingresos y costos:** Los datos fueron capturados para apoyar al cálculo de rentabilidad de sus productos y clientes.

VISION fue desarrollado en fases con objetivos y resultados tangibles. Esto le dio una imagen aún más precisa de quiénes eran sus mejores clientes, y cuáles eran los productos más importantes.

### ***Resultados obtenidos***

Mediante VISION, usuarios tales como analistas financieros y de mercadotecnia, gerentes y personal operativo obtuvieron los siguientes beneficios:

- Se mejoró la capacidad de hacer análisis financieros, con la finalidad de evaluar la rentabilidad de los productos actuales y futuros
- Se pudieron descubrir oportunamente las cuentas bancarias de mayor rentabilidad, revelando que el tamaño promedio de saldo de las cuentas era un factor diferenciador
- Utilizando los datos de VISION, los analistas pudieron definir el límite de tasa de interés anual por cliente y poner esta información a disposición del personal operativo
- Se mejoraron los análisis de mercado, al grado de poder determinar acciones bajo diferentes escenarios.
- Con los datos de VISION se pudo identificar los productos que se podían vender a los clientes
- Usando VISION, el sistema pudo calcular la manera óptima de distribuir los productos en base a las preferencias de los clientes

En los casos presentados se puede observar una fuerte necesidad por obtener información oportuna para la toma de decisiones. En ambos casos los ejecutivos obtuvieron la capacidad de hacer análisis y modelos a gran escala, mostrando una perspectiva más amplia de la industria, y proporcionándoles una ventaja competitiva en el Sector Financiero.

## **7. Investigación de campo**

El presente capítulo tiene como propósito describir la metodología que se utilizó para realizar la actual investigación. Se detallan los elementos de la metodología utilizada tales como el tipo de investigación, selección de la muestra, entre otros.

### **7.1. Metodología**

Se trata de una investigación basada en el método científico, definido por McKean (2005) como una metodología compuesta por la observación sistemática, la medición y la experimentación. La finalidad es formular, probar o modificar una hipótesis.

#### **7.1.1. Tipo de investigación**

La investigación será de tipo *probabilística*, la cual de acuerdo a Hernández, Fernández y Baptista (2006), todos los elementos de la población tienen la misma probabilidad de ser seleccionados. Por medio de los resultados se podrá describir el nivel de implementación; los factores críticos de éxito de índole económica, social y tecnológica; y una serie de recomendaciones para la implementación de esta tecnología.

#### **7.1.2. Población**

El Sector Financiero Mexicano, también conocido como Sistema Financiero Mexicano, de acuerdo al Banco de México (BANXICO) se compone de:

- Organismos de Regulación, Supervisión y Control
- Fondos y Fideicomisos Públicos
- Sector Bancario
- Sector de Ahorro y Crédito Popular

- Sector de Intermediarios Financieros No Bancarios
- Sector Bursátil
- Sector de Derivados
- Sector de Seguros y Fianzas

Para la investigación se delimitó la población a empresas que conforman esta estructura y con sede en Monterrey, Nuevo León. Se buscó que las empresas encuestadas tuvieran implementado un *data warehouse*, y una participación significativa en el Sector Financiero.

### 7.1.3. Selección de muestra

Una vez identificado el tamaño de la población, se procedió a la obtención de una muestra aleatoria simple, conforme al procedimiento de selección de muestras probabilísticas de Marcos (1998). La muestra obtenida fue el resultado de aplicar la ecuación de la figura 7.1.

$$n = \frac{Z^2 * P(1-P) * N}{e^2(N-1) + Z^2 * P(1-P)}$$

*Figura 7.1. Fórmula para obtener la muestra probabilística*

*Fuente: Manual para la elaboración de tesis (Marcos, 1998)*

Donde:

- n= tamaño de la muestra
- Z= nivel de confiabilidad; utilizando 95%

- $P$ = probabilidad de éxito de que las empresas seleccionadas tengan las características requeridas
- $e$ = error esperado
- $N$ = tamaño de la población

Sustituyendo las variables:

- $Z$ = .95
- $P$ = .95
- $e$ = .05
- $N$ = 298

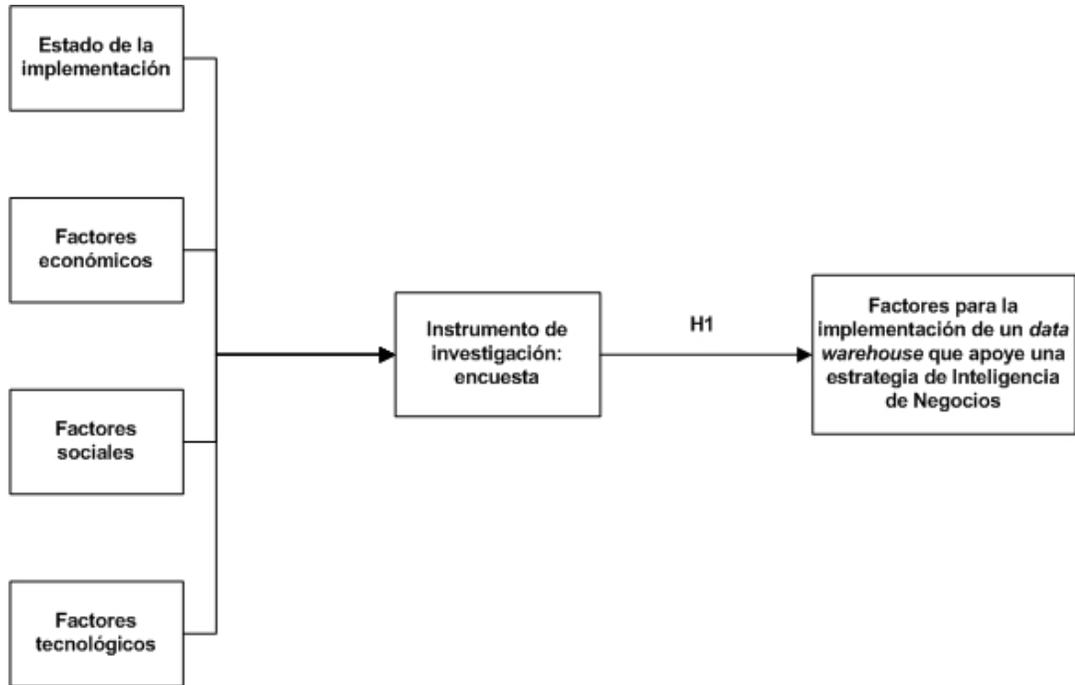
Se obtiene una muestra de 16 empresas.

## **7.2. Herramienta de Investigación**

La herramienta de investigación seleccionada para la recolección de datos fue la encuesta. El formato de la encuesta se muestra en el Anexo 1, y fue dirigida a las personas involucradas en la administración del *data warehouse*.

## **7.3. Modelo Particular**

Con la finalidad de alcanzar los objetivos establecidos, se definió un modelo de investigación que permita analizar al Sector Financiero de Monterrey, para así obtener resultados que puedan validar la hipótesis. A continuación la figura 7.2 muestra el modelo particular para realizar esta investigación:



*Figura 7.2. Modelo de Investigación*

La hipótesis a comprobar es:

H1. Los factores económicos, sociales y tecnológicos no son tomados en consideración al momento de la adopción de la tecnología del *data warehouse*.

#### **7.4. Variables de la Investigación**

##### ***A. Variables dependientes***

- Nivel de avance en la implementación de un sistema de *data warehouse* en las empresas del Sector Financiero de Nuevo León
- Factores críticos para la implementación exitosa de un *data warehouse* que apoye una estrategia de Inteligencia de Negocios

##### ***B. Variables independientes***

### Nivel de la Implementación

- Tipos de sistemas con los que cuenta la empresa
- Existencia de una herramienta de *data warehouse*
- Etapa en la que se encuentra la implementación

### Factores

- Factores Económicos
  - Coordinación de recursos
  - Enfoque al negocio
  - Costo
  - Tiempo
  - Estudio de factibilidad
- Factores Sociales
  - Apoyo al proyecto por parte de la administración
  - Elaboración de expectativas reales
  - Involucramiento de los usuarios
  - Beneficios recibidos
  - Resistencia al cambio
- Factores Tecnológicos
  - Involucramiento de personas con experiencia
  - *Software y hardware* necesario
  - Alineación de la tecnología con el negocio
  - Compatibilidad con el resto de los sistemas

- Estrategia de implementación
- Proveedores

### **7.5. Estrategia para la Recolección de Datos**

Elaborar una estrategia de recolección de datos implica definir un plan detallado de los procedimientos, dirigido a reunir los datos con un propósito específico (Hernández, Fernández & Baptista, 2006). El plan definido para esta investigación consiste en:

1. Identificar las empresas financieras para la aplicación de la encuesta
2. Buscar contactos que estén laborando en alguna de las empresas identificadas, recurriendo a redes de personas y contactos directos
3. Una vez establecidos los contactos, justificar los motivos de la investigación y presentar documentos que avalen el estudio
4. Conseguir la aprobación de las empresas
5. Hacer llegar las encuestas electrónicamente o en formato impreso
6. Recolectar las encuestas contestadas

Finalmente una vez terminado el proceso de recolección, se iniciará el análisis de los datos, para cerrar con la elaboración de las conclusiones y trabajos futuros.

## **8. Análisis de resultados**

Una vez definida la metodología de la investigación de campo, se inició con la aplicación de la encuesta, la cual fue definida como el instrumento de recolección de datos.

Se seleccionaron 16 empresas del Sector Financiero de Monterrey, dirigiendo el estudio a aquellas que tuvieran implementado un *data warehouse*. Las personas que respondieron la encuesta fueron las involucradas en la planeación, implementación, y mantenimiento del *data warehouse* en dichas organizaciones.

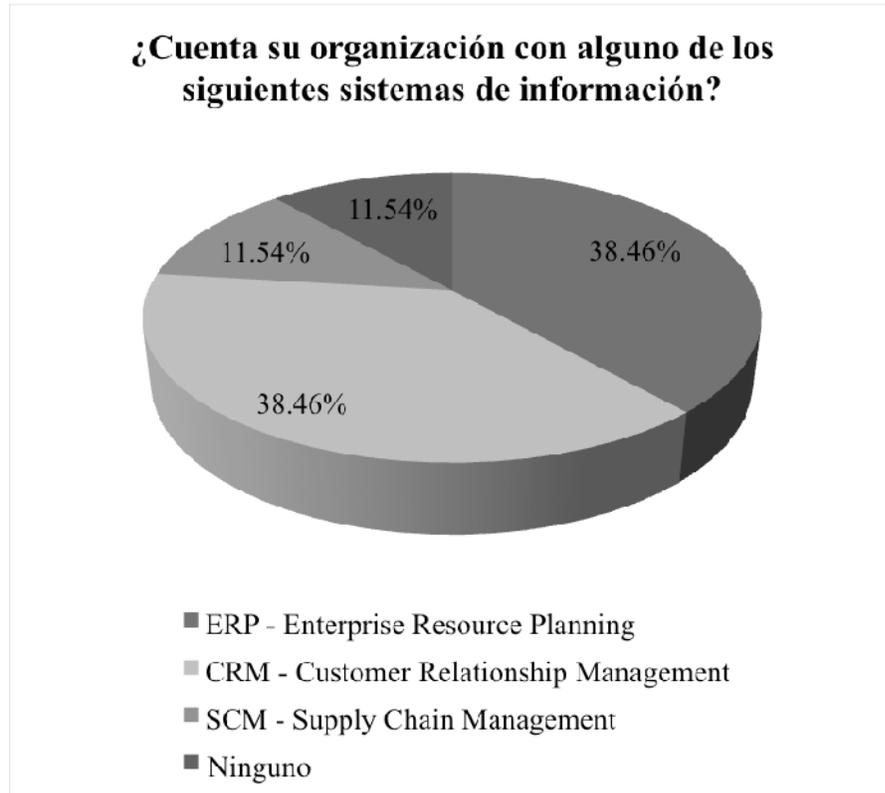
En el presente capítulo se darán a conocer las respuestas obtenidas de la aplicación de la encuesta, además de los factores críticos identificados y una serie de recomendaciones. Los resultados se agruparon por nivel de implementación, aspectos económicos, aspectos sociales, y aspectos tecnológicos.

### **8.1. Nivel de Implementación**

En este apartado se evaluaron los tipos de sistemas con los que cuentan las organizaciones, el nombre de la herramienta de *data warehouse* y la etapa en la que se encuentra la implementación. Los resultados se presentan a continuación.

#### ***Tipos de sistemas con los que cuentan las organizaciones***

Los sistemas de información son herramientas que proveen de nuevas capacidades a una organización. En base a lo anterior, se les preguntó a los encuestados sobre la existencia de sistemas ERP (en inglés, *Enterprise Resource Planning*), CRM (en inglés, *Customer Relationship Management*), y SCM (en inglés, *Supply Chain Management*) en sus organizaciones, obteniendo los resultados ilustrados en la figura 8.1.

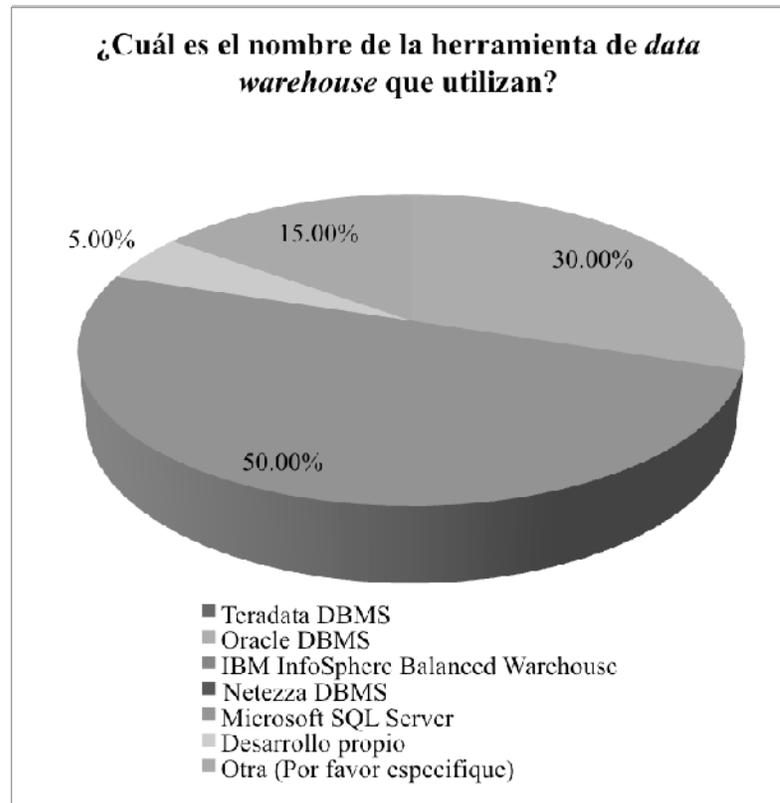


*Figura 8.1 Sistemas de información existentes en las organizaciones*

En las respuestas obtenidas se aprecia el interés de las empresas financieras en mejorar la relación con el cliente, e integrar las operaciones de la organización mediante soluciones tecnológicas. Además resalta la baja presencia de sistemas SCM, ya que en sus operaciones no hay una dependencia crítica en las actividades de producción o manejo de inventarios.

***Nombre de la herramienta de data warehouse***

La variedad de herramientas de *data warehouse* existentes en el mercado pone en perspectiva las características que son de utilidad para la organización. Al preguntar entre los encuestados sobre el nombre de su herramienta, se obtuvieron los resultados que se muestran en la figura 8.2.



*Figura 8.2 Nombre de la herramienta de data warehouse que utilizan*

En forma general se observa que dos grandes proveedores de soluciones tecnológicas dominan el mercado financiero, y menos de una tercera parte tiene implementada una herramienta no tan conocida. Microsoft es el líder debido a la popularidad de sus aplicaciones, seguido de cerca por Oracle gracias a sus sistemas de bases de datos. En los casos donde se optó por desarrollos propios u otras alternativas, las herramientas de Artus y SAP BW son lo suficientemente robustas para satisfacer las necesidades de la minoría.

***Etapas en la que se encuentra la implementación***

No todos los proyectos de informática tienen el mismo grado de importancia, de ahí la relevancia de conocer la etapa en la que se encuentra la implementación del *data*

*warehouse*. Al preguntar al respecto, se obtuvieron las respuestas que se señalan en las figuras 8.3 y 8.4.

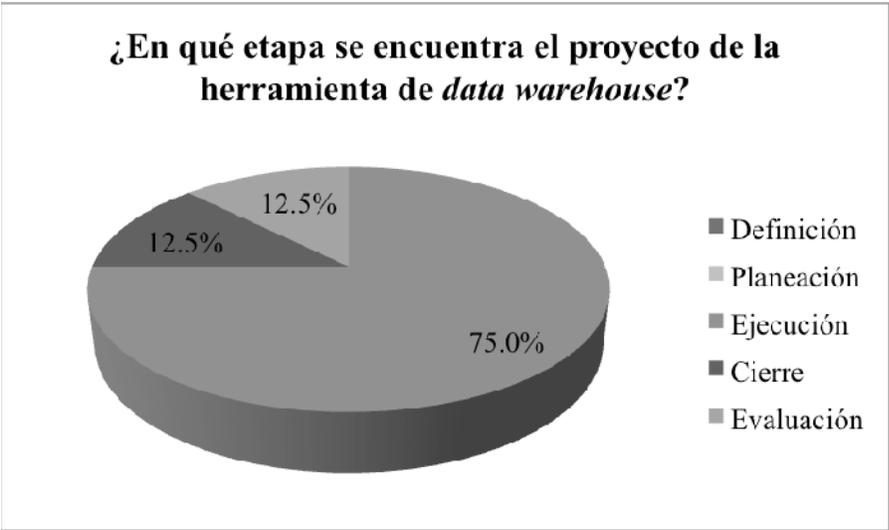


Figura 8.3 Etapa en que se encuentra la implementación del *data warehouse*

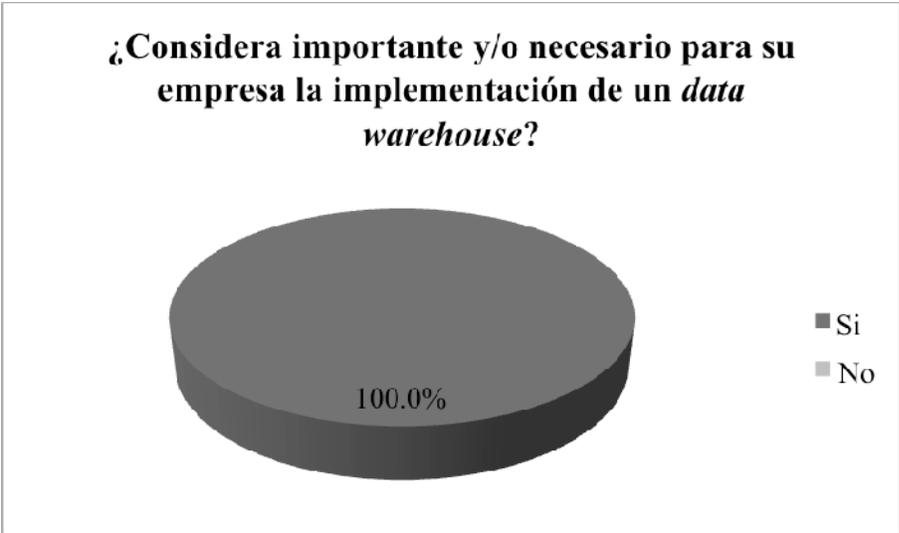


Figura 8.4 Importancia de la implementación de un *data warehouse*

Analizando la información presentada en ambas preguntas, se visualiza un alto grado de madurez de la tecnología de *data warehouse* en las empresas financieras, además de la importancia de la herramienta. En base a lo anterior se puede inferir que ésta es

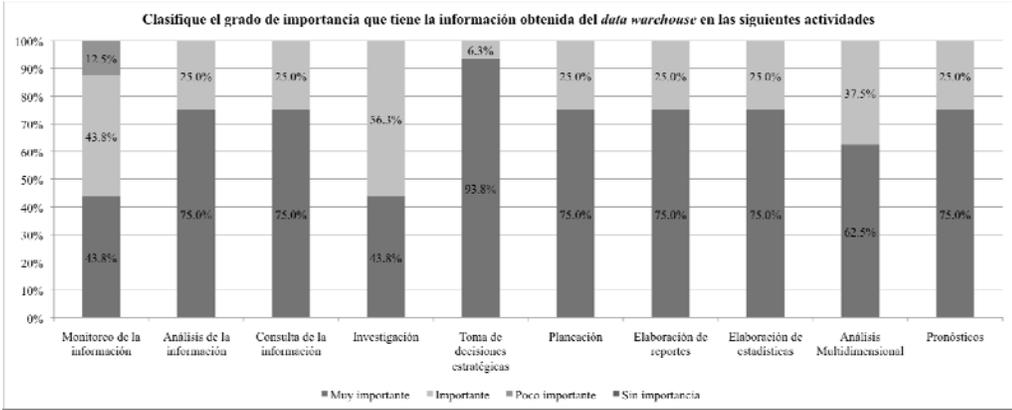
utilizada para la elaboración de estrategias, y se ha integrado adecuadamente a las operaciones de las organizaciones.

**8.2. Aspectos económicos**

En esta sección se evaluaron los aspectos de enfoque al negocio, la coordinación de los recursos, el costo y tiempo en la implementación, y los estudios de factibilidad. A continuación se analiza cada uno de estos aspectos.

*Enfoque al negocio*

El aspecto fundamental de toda herramienta tecnológica, es el valor de la información que ésta genera. A los encuestados se les solicitó clasificar el grado de importancia que tiene la información obtenida del *data warehouse* en las actividades de monitoreo, análisis, consulta, investigación, toma de decisiones estratégicas, planeación, elaboración de reportes, elaboración de estadísticas, análisis multidimensional, y pronósticos. Las respuestas obtenidas se expresan en la figura 8.5.



*Figura 8.5 Grado de importancia que tiene la información obtenida del data warehouse en las actividades de enfoque al negocio*

Con respecto a la pregunta, todas las actividades fueron críticas para los encuestados, siendo mínimas las diferencias entre los resultados. Son tres los puntos a destacar.

En primer lugar, no existió duda que la información obtenida del *data warehouse* es la más importante para el proceso de toma de decisiones estratégicas. Los ejecutivos financieros prestan especial atención en este aspecto, ya que dicha información les permite tener un mayor control de los procesos de la organización, además de hacer posible la implementación de tecnologías más complejas, como las relacionadas con la inteligencia de negocios.

Segundo, los resultados sobre el rubro de análisis multidimensional fueron interesantes. Esta clase de análisis permite a sus usuarios manipular la información de diferentes formas, para así obtener perspectivas que se adecuen a sus exigencias. Al ser éste uno de los principales beneficios de un *data warehouse*, se esperaba que fuera apreciado en igual grado que en el caso para la toma de decisiones, sin embargo existe una barrera tecnológica para los usuarios.

Finalmente, los resultados de monitoreo e investigación ocuparon el menor grado de importancia. Esto indica que la tecnología del *data warehouse* no se explota para esta actividad, en su lugar utilizan otros medios (boletines informativos, asesorías, *dashboards*, etc).

### ***Recursos necesarios***

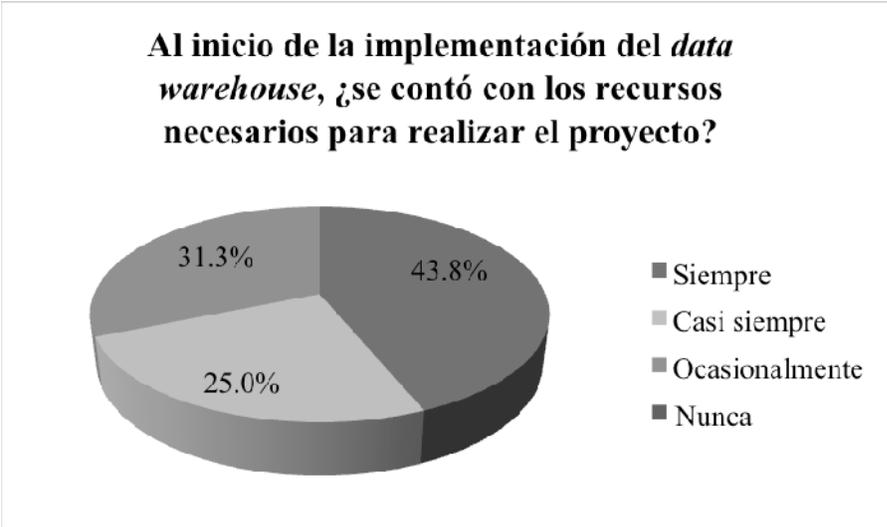
No siempre es posible contar con todos los recursos al momento de arrancar un proyecto de informática, y una inversión en capacitación es una exigencia. En cuanto al caso particular del *data warehouse*, se les cuestionó esta necesidad al momento de la implementación. La figura 8.6 muestra los resultados obtenidos.



*Figura 8.6 Necesidad de inversión en capacitación al momento de la implementación*

Como lo señala la gráfica anterior, se tuvo que invertir en capacitación al momento de la implementación. Los financieros no deben de escatimar el presupuesto destinado a esta finalidad.

Por segundo aspecto se tiene la suficiencia de recursos al inicio de la implementación. Los resultados obtenidos de los encuestados se muestran en la figura 8.7.

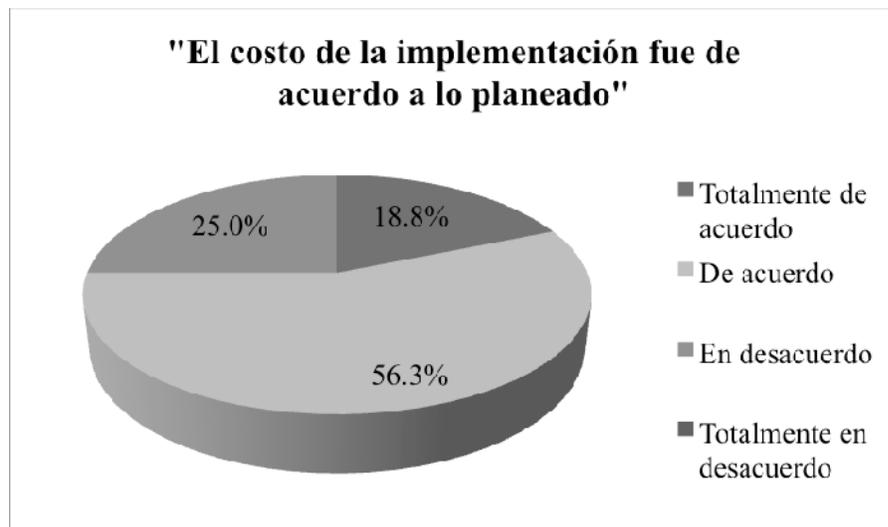


*Figura 8.7 Suficiencia de recursos al momento de la implementación*

En lo referente a la segunda pregunta, la mayoría de los ejecutivos financieros conocen los riesgos derivados de una implementación deficiente, además se observa la buena voluntad de la administración al momento de la asignación de presupuesto para proyectos de índole estratégica.

### **Costo**

El costo total es el factor más observado por la alta dirección, antes de tocar este rubro, se opta por modificar el alcance o el tiempo del proyecto. Al cuestionar sobre el grado de conformidad del costo de la implementación del *data warehouse*, se obtuvieron los resultados mostrados en la figura 8.8.

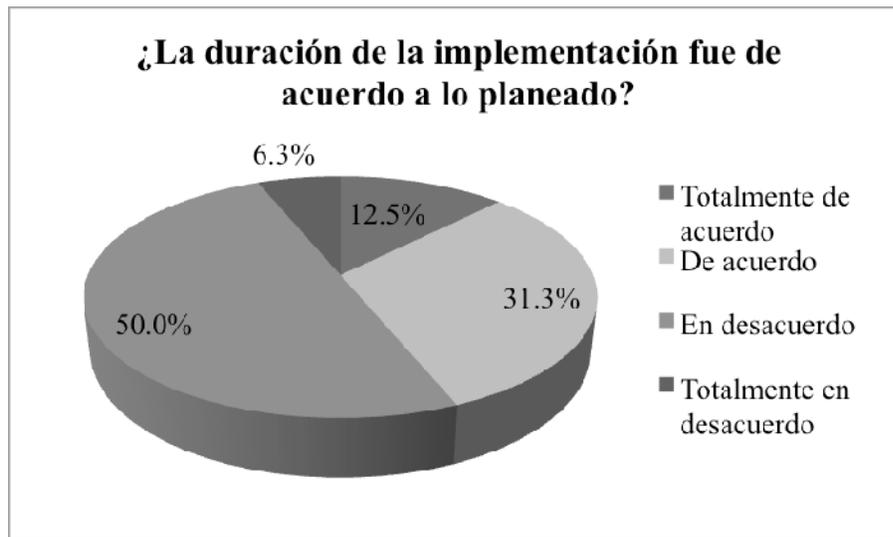


*Figura 8.8 Grado de conformidad del costo de la implementación*

Los resultados obtenidos muestran un alto grado de seriedad por parte de las financieras con respecto a costos. Sin duda para ellas es importante terminar de acuerdo al costo planeado, y así lo afirmaron el 75% de las empresas encuestadas.

## ***Tiempo***

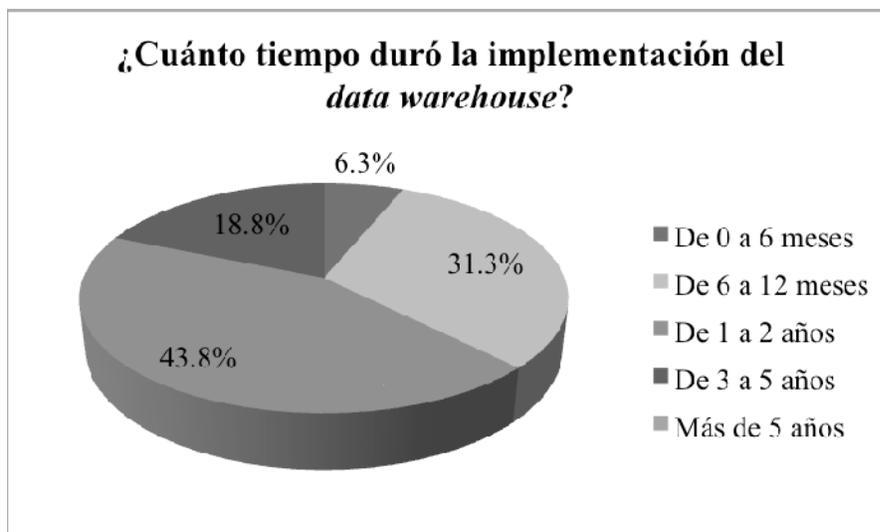
El recurso humano requiere estar disponible para distintas operaciones del negocio, siendo de gran impacto una eficaz administración del tiempo. Se les preguntó su opinión sobre la duración de la implementación, en términos de planeación. Los resultados se presentan en la figura 8.9.



*Figura 8.9 Duración de la implementación conforme a lo planeado*

Este punto muestra una inadecuada práctica de planeación del tiempo por parte de las financieras. La mayoría de ellas desconoce la complejidad de este tipo de proyectos, ya que al momento de hacer la planeación y estimación de las actividades, se les trata como cualquier proyecto de informática.

Una vez definido el grado de conformidad de la duración de la implementación, se procede a analizar en específico cuánto duró dicho proceso. En las empresas encuestadas se obtuvieron los porcentajes que se muestran en la figura 8.10.

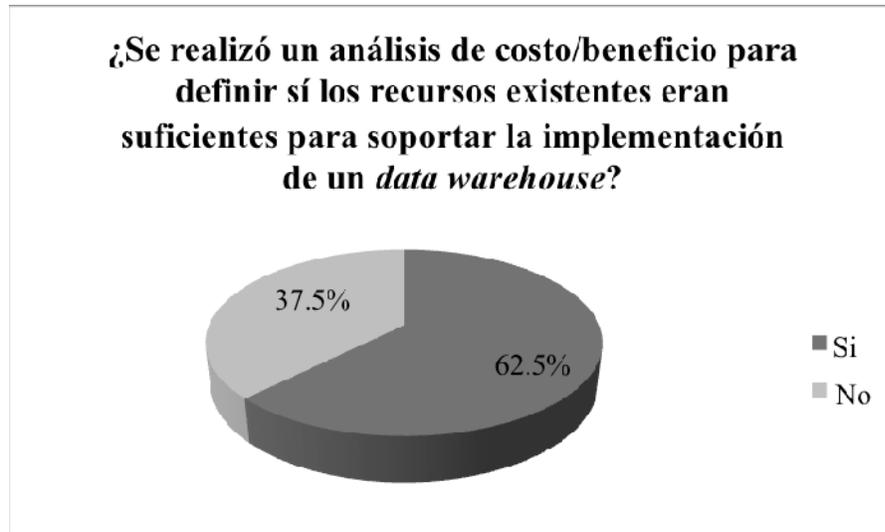


*Figura 8.10 Tiempo que duró la implementación del data warehouse*

En base a la gráfica, es posible observar que la mayoría los proyectos de *data warehouse* tienen una duración considerable. Lo anterior es indicador de problemas de complejidad técnica y afecta a diversos procesos del negocio, alargando los proyectos a un lapso de uno a dos años.

### ***Estudio de factibilidad***

Con un estudio de factibilidad es posible definir en qué medida es redituable un proyecto en particular, razón por la cual se les cuestionó sobre la realización de dicho estudio para el caso del *data warehouse*. Las empresas contestaron tal como se muestra en la figura 8.11.



*Figura 8.11 Análisis de costo/beneficio para definir la suficiencia de los recursos para soportar la implementación*

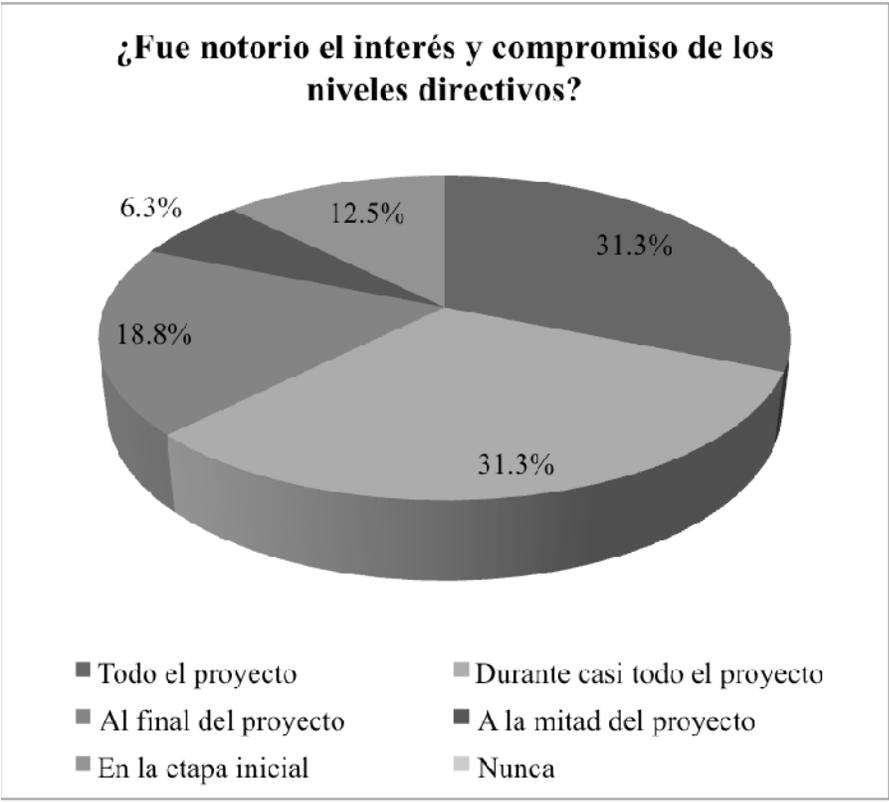
Para la mayoría de los encuestados fue importante la aplicación de un estudio de factibilidad, pero no existió una diferencia determinante en comparación con los que no lo hicieron. Tomando en cuenta lo anterior, se puede inferir la existencia de una falta de conocimiento de sus propios recursos y/o el desconocimiento de la complejidad de esta tecnología, lo que lleva al equipo a comprometerse con un proyecto que tendrá probablemente dificultades.

### **8.3. Aspectos sociales**

Para este punto se evaluaron los aspectos de apoyo al proyecto por parte de la administración, elaboración de expectativas realistas, involucramiento de los usuarios y de personas con experiencia, beneficios recibidos, y resistencia al cambio. En seguida se muestra el detalle de cada uno.

#### ***Apoyo al proyecto por parte de la administración***

Los directores de las empresas financieras están presentes en todas las decisiones que involucran a la organización. Al ser ellos también jugadores clave en los proyectos de informática, se cuestionó a los encuestados sobre su interés y compromiso durante el proyecto de *data warehouse*, obteniendo la composición mostrada en la figura 8.12.



*Figura 8.12 Interés y compromiso de los niveles directivos*

De acuerdo con lo anterior, para la mayoría de los directivos está clara la importancia de su compromiso con el proyecto. Sin duda establecer un vínculo de confianza entre la administración y el equipo del proyecto disminuye el escepticismo del personal, y por ende facilita su incorporación a las operaciones del negocio.

Tratando otro aspecto relacionado, se requiere que los directivos presten atención en las actividades de capacitación, seguridad de la información, requerimientos de

infraestructura tecnológica, adaptación de la herramienta, estándares de calidad, y el aspecto humano. Ante esto, se les preguntó sobre el apoyo recibido en dichas actividades en el proyecto de *data warehouse*, obteniendo las respuestas ilustradas en la figura 8.13.

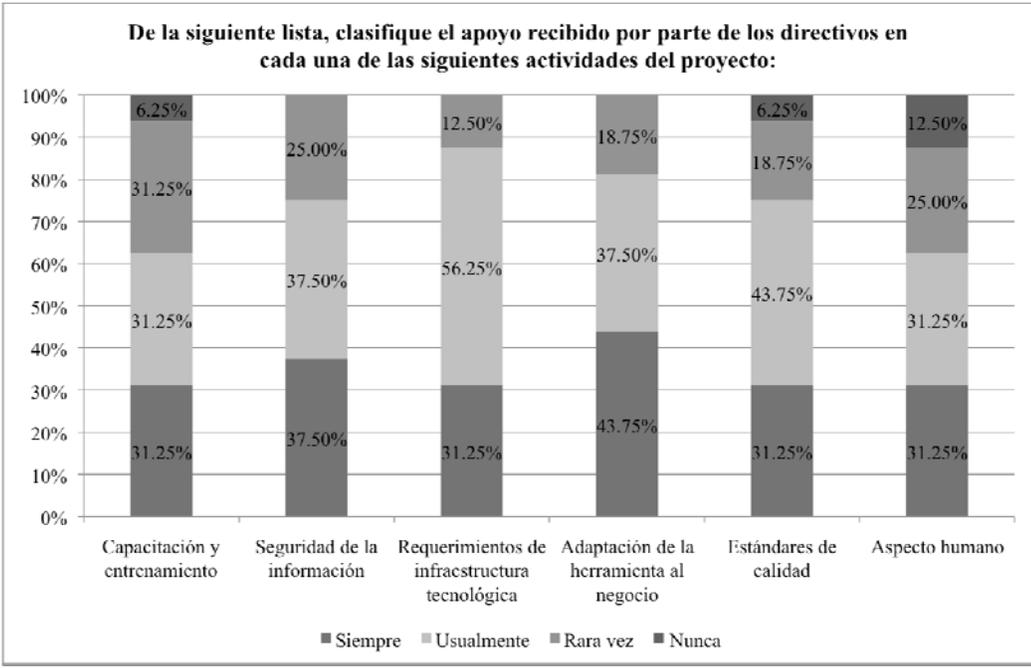


Figura 8.13 Apoyo recibido de la administración en las actividades del proyecto

Se puede observar que todas las actividades son de interés para la administración, además de mostrarse cuáles tienen mayor importancia. Según los resultados, los puntos a destacar son tres.

Primero, la definición de requerimientos de infraestructura tecnológica es la actividad más relevante para las empresas financieras. La administración sabe que sin el contexto tecnológico adecuado, no es posible obtener todos los beneficios del *data warehouse*.

Segundo, para la administración tiene la misma relevancia los aspectos de seguridad de información y estándares de calidad. Estos aspectos son diferenciadores en el Sector Financiero y su ausencia representa una desventaja con respecto a los competidores.

Concluyendo las observaciones, el aspecto humano es el de menor prioridad para la administración. No se tiene una idea clara de la influencia de los empleados en el éxito o fracaso del proyecto de *data warehouse*.

Retomando la temática, se les cuestionó sobre el apoyo recibido en las etapas de elaboración de la iniciativa del proyecto, asignación del presupuesto, elaboración del plan, selección y evaluación de *software*, implementación del *data warehouse*, y capacitación. Las respuestas obtenidas se pueden visualizar en la figura 8.14.

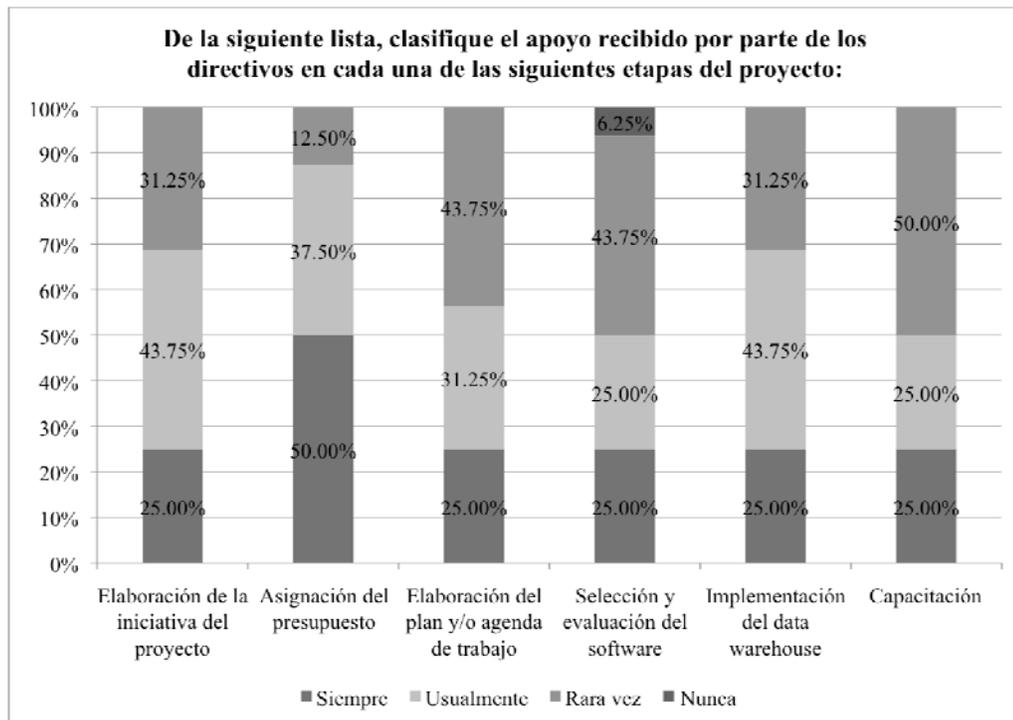


Figura 8.14 Apoyo recibido de la administración en las etapas del proyecto

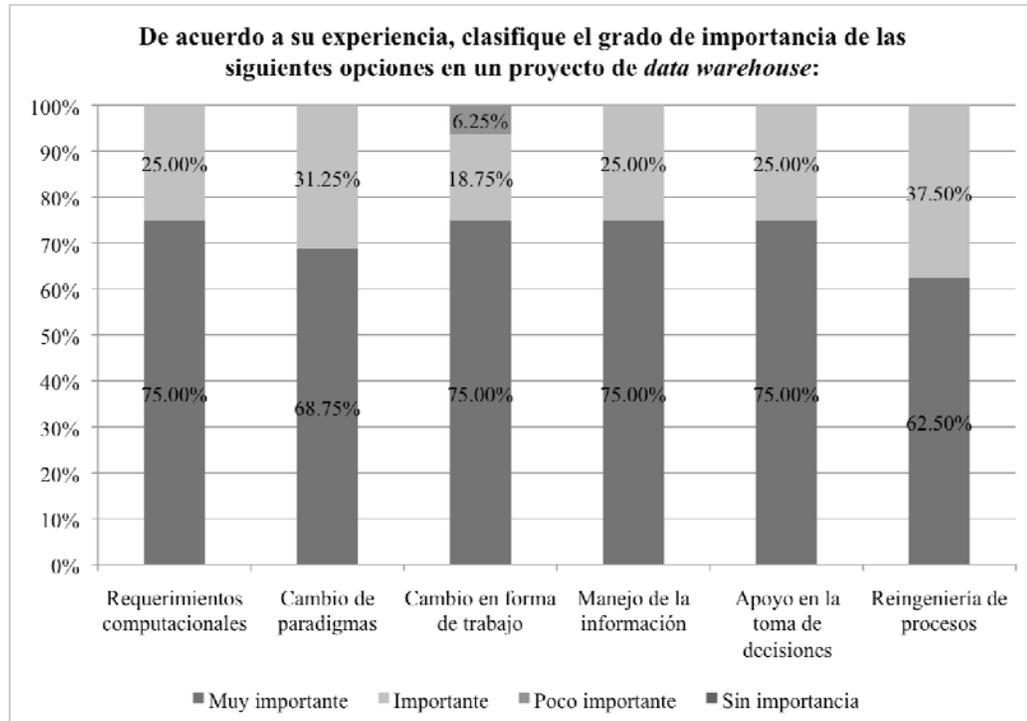
De los resultados ilustrados se observa que en general son aspectos de un grado de interés medio para la administración. Sobresalen dos observaciones.

En primer lugar, no es de sorprenderse que la asignación de presupuestos es la etapa en la que más estuvo involucrada la administración, dejando la elaboración de la iniciativa del proyecto y la implementación como cuestiones de segunda prioridad. En las tres actividades debe haber participación la administración, porque es la encargada de definir los objetivos de cada una de las actividades. El hecho de contar con un apoyo remarcable en la etapa de asignación de presupuestos, es reflejo de un cuidado especial en lo referente a la inversión de un proyecto de informática.

Y por último, la etapa de selección y evaluación de *software* resultó ser dónde se tiene menor soporte por parte de la administración. Al tratarse de una etapa que exige un amplio conocimiento técnico y metodológico, no es de interés para la administración, delegando la tarea al equipo del *data warehouse*.

### ***Elaboración de expectativas realistas***

Cada proyecto de informática exige al capital humano un cambio en sus actividades, y cada grupo de usuarios tiene sus propias expectativas. A los encuestados se les preguntó por las actividades que definen a un proyecto de *data warehouse*, tales como la elaboración de los requerimientos computacionales, cambio de paradigmas, cambio en forma de trabajo, manejo de la información, apoyo en la toma de decisiones, y reingeniería de procesos. Los resultados se representan en la figura 8.15.



*Figura 8.15 Grado de importancia en las actividades para elaborar expectativas realistas*

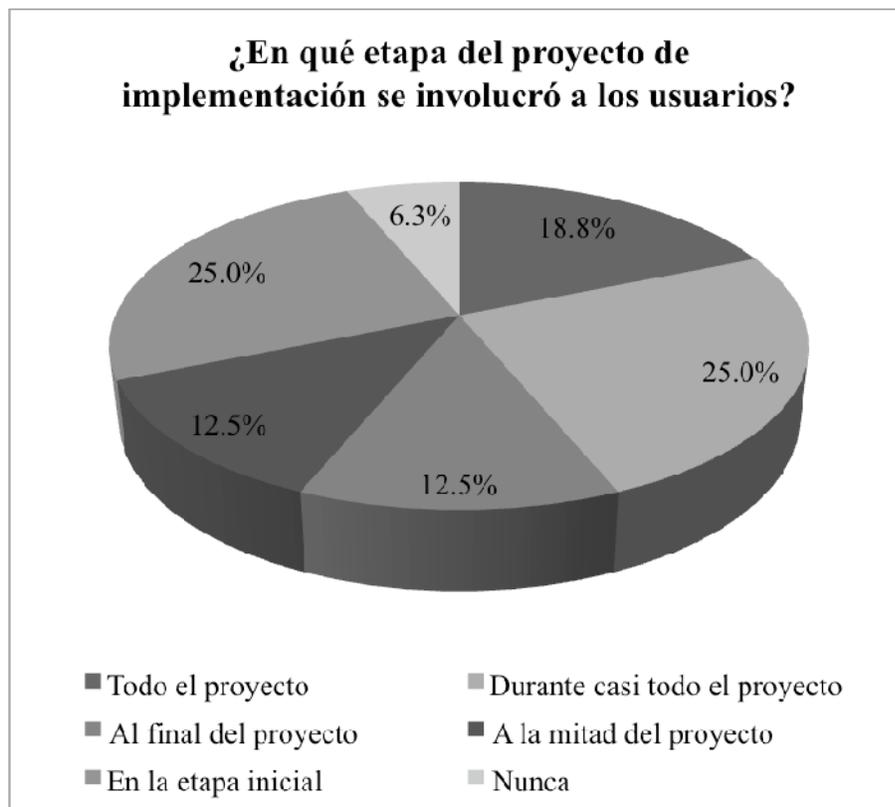
Resulta obvio notar que para todas las empresas involucradas en un proyecto de *data warehouse*, son muy importantes las actividades mostradas en la figura 8.15. Debido a la baja dispersión de los datos, sólo es relevante observar las opciones que ocupan el primer y último lugar de importancia para las empresas financieras.

El apoyo en la toma de decisiones, el manejo de la información, y los requerimientos computacionales son las opciones de mayor relevancia para las empresas financieras. No es de sorprenderse que las dos primeras sean importantes debido a que se tratan de características distintivas del *data warehouse*, y la tercera es muy conocida por el fuerte impacto que causa esta tecnología a una infraestructura tecnológica.

Concluyendo, la reingeniería de procesos es la opción de menor relevancia para los financieros. Ellos saben que existirá un cambio en los procesos de negocio al momento de la implementación, pero se toman a la ligera todas estas modificaciones.

### ***Involucramiento de los usuarios***

La participación de los usuarios ayuda a la identificación de las características de un producto. En seguida se les preguntó sobre la etapa en la que se involucró a los usuarios del proyecto de implementación, obteniendo la gráfica representada en la figura 8.16.

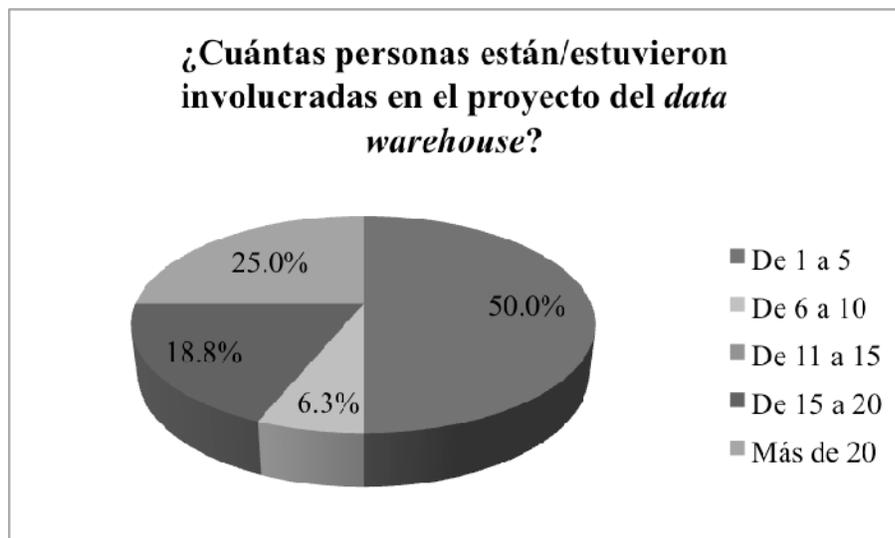


*Figura 8.16 Etapa del proyecto de implementación en la que se involucró a los usuarios*

Sin duda se puede observar una alta dispersión de datos entre las respuestas, lo que indica una falta de claridad sobre esa etapa de involucramiento. A medida que una

tecnología madura, se tiene mayor información del impacto de los usuarios en determinadas etapas, permitiendo elaborar un juicio más personalizado en este punto.

Ya mencionado el aspecto de la participación de los usuarios, es aconsejable considerar el detalle del personal involucrado. Con respecto a lo anterior, se les cuestionó sobre la cantidad de personas involucradas en el proyecto de *data warehouse*, obteniendo las respuestas ilustradas en la figura 8.17.



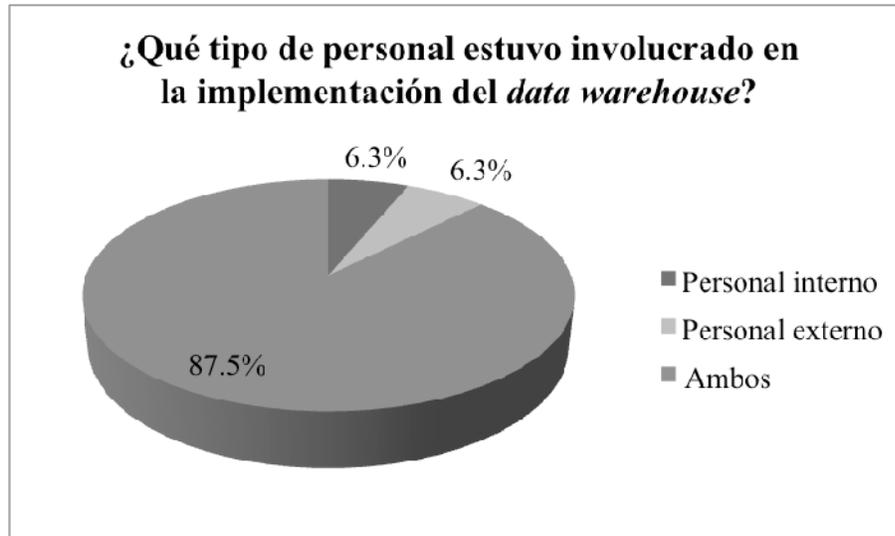
*Figura 8.17 Cantidad de personas involucradas en el proyecto del data warehouse*

La estadística muestra que los proyectos de *data warehouse* están en su mayoría a cargo de pequeños grupos de trabajo. Cabe mencionar que esta tecnología ya se encuentra consolidada en el mercado, facilitando esta tarea a un reducido equipo de trabajo.

#### ***Involucramiento de personas con experiencia***

No todo el personal tiene el conocimiento y la experiencia en la tecnología de *data warehouse*. Ante esto, se les cuestionó sobre el tipo de personal que estuvo involucrado

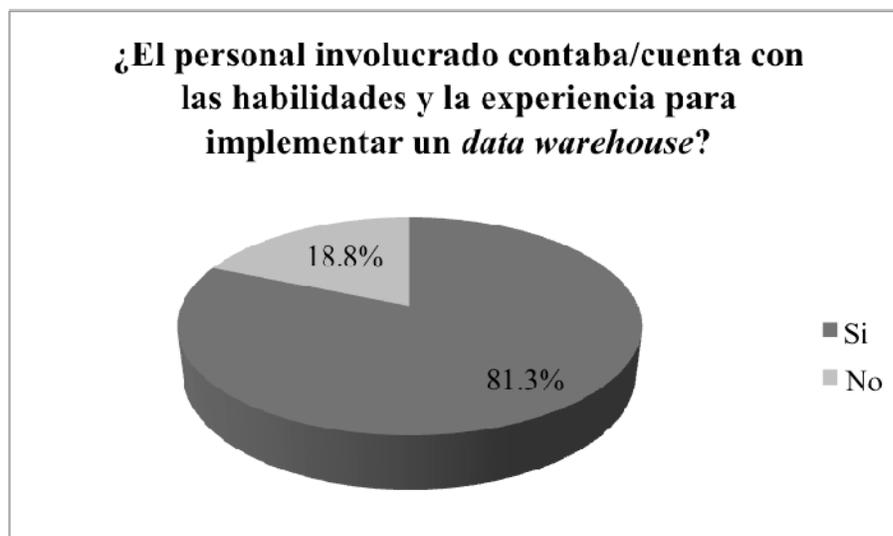
en la implementación del *data warehouse*, obteniendo la distribución mostrada en la figura 8.18.



*Figura 8.18 Tipo de personal involucrado en la implementación del data warehouse*

De acuerdo a las respuestas, es necesario formar un equipo de especialistas integrado por los dos tipos de personal. Se puede inferir que existe una fuerte interdependencia entre el personal interno y externo, debido al nivel de acoplamiento de la tecnología con los procesos del negocio.

En adición a la pregunta anterior, se cuestionó sobre la suficiencia de habilidades y experiencia del equipo del proyecto. Las respuestas que se expresan en la figura 8.19.

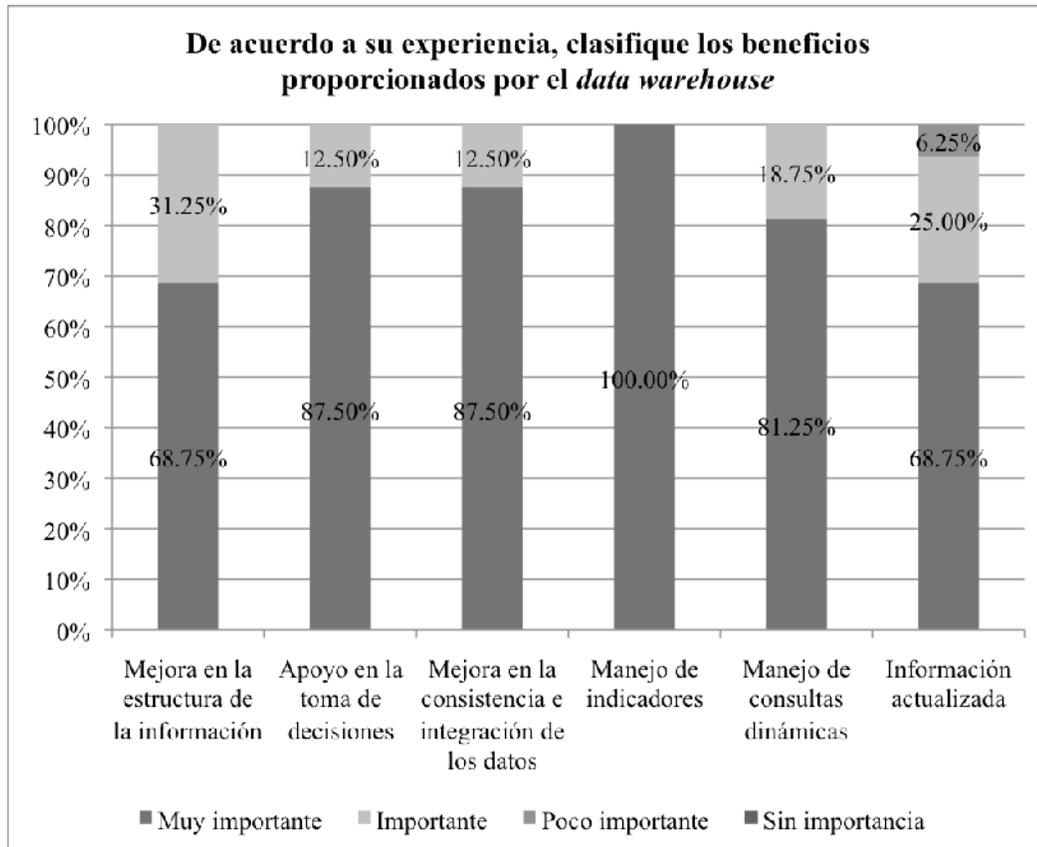


*Figura 8.19 Habilidades y la experiencia del personal involucrado en la implementación del data warehouse*

Se infiere que para las empresas financieras es crucial que su personal cuente con las habilidades necesarias. Los directivos saben que el éxito del proyecto es determinado por las capacidades de su capital humano, teniendo especial cuidado en este aspecto.

### ***Beneficios recibidos***

El nivel de impacto de un proyecto está en función de los beneficios percibidos en comparación con la inversión realizada. Se solicitó el apoyo de los encuestados para calificar los beneficios relacionados con estructuras de información, toma de decisiones, consistencia e integración de datos, indicadores, consultas dinámicas, e información actualizada. Los resultados obtenidos se ilustran en la figura 8.20.



*Figura 8.20 Beneficios proporcionados por el data warehouse*

Analizando al detalle los aspectos señalados, se reconocen favorablemente los beneficios de un proyecto de *data warehouse*. Considerando la baja dispersión de datos de los resultados, las observaciones relevantes a mencionar son las que ocupan el primer y último lugar en grado de importancia.

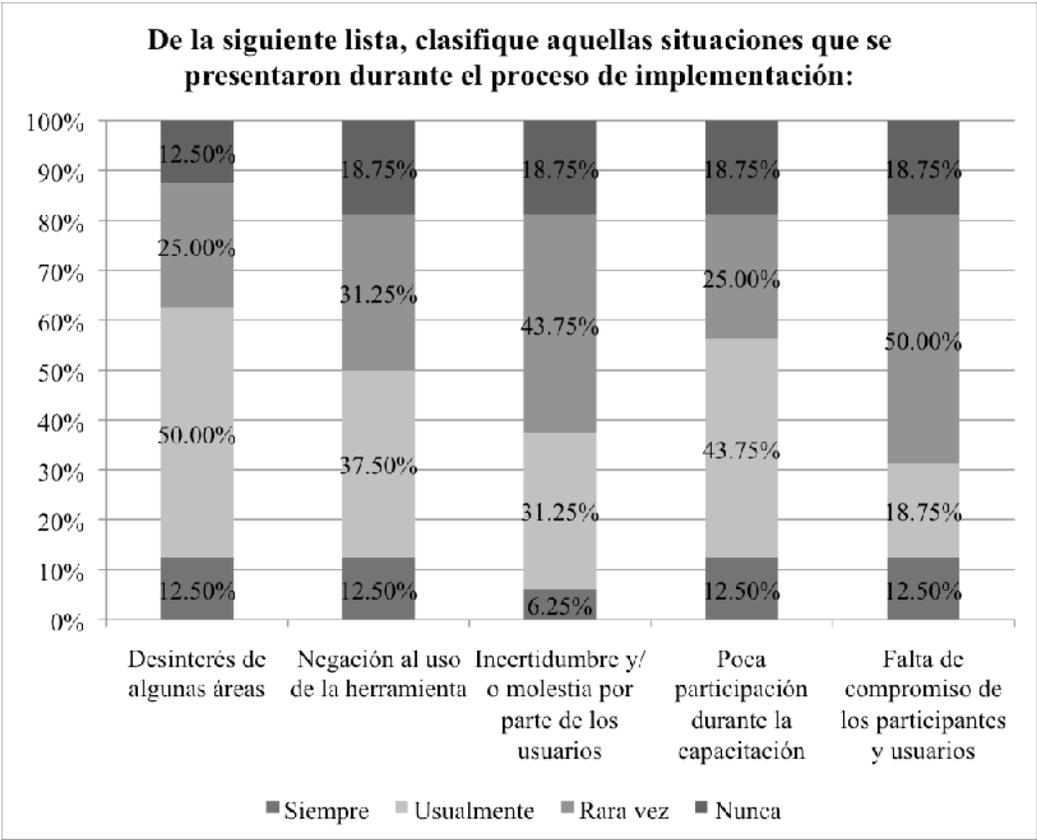
El manejo de indicadores es el beneficio más importante para las empresas financieras. A éstas les interesa alcanzar un mayor grado de control, siendo el *data warehouse* la alternativa para mejorar la administración de sus procesos de negocio, y así acceder a una nueva perspectiva estratégica.

Por último, la información actualizada es el beneficio menos relevante para las empresas financieras. En la mayoría de las organizaciones el *data warehouse* se utiliza

para registrar la historia de la organización, perdiendo de esa forma su verdadero potencial.

**Resistencia al cambio**

La gente odia los cambios por el esfuerzo involucrado para adecuarse a un nuevo estado. En cuanto a situaciones relacionadas con la resistencia al cambio, se les preguntó sobre el desinterés, negación al uso de la herramienta, incertidumbre, molestia por parte de los usuarios, poca participación, y falta de compromiso para el caso del *data warehouse*. Así lo señalan los resultados de la figura 8.21.



*Figura 8.21 Situaciones relacionadas con la resistencia al cambio durante la implementación*

Los problemas de desinterés de algunas áreas, la poca participación durante la capacitación, y la negación del uso de la herramienta fueron los problemas de mayor presencia. En contraste, los problemas relacionados con la incertidumbre y molestia por parte de los usuarios, y la falta de compromiso son los que tuvieron menos peso en las empresas financieras.

Los tres primeros problemas se relacionan con una mala percepción de la tecnología de *data warehouse* por parte de los empleados, ya que no se logra visualizar su utilidad y por consecuencia se tienen dificultades durante la capacitación.

Cerrando el tema, los equipos de proyectos de *data warehouse* conocen la forma de manejar los dos problemas menos relevantes. Se puede observar que las empresas financieras saben mitigar los riesgos relacionados con esas cuestiones, y se reconoce el esfuerzo adicional del capital humano por efecto de la modificación de sus actividades.

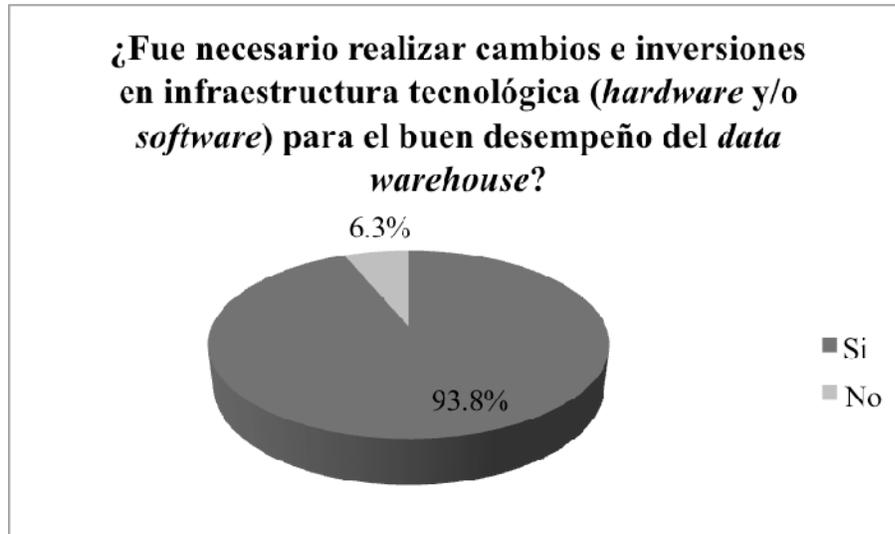
#### **8.4. Aspectos tecnológicos**

Para los aspectos tecnológicos se evaluaron las siguientes variables: *software* y *hardware* necesario, alineación de la tecnología con el negocio, compatibilidad con el resto de los sistemas, estrategia de implementación, y proveedores. El análisis de cada una se muestra a continuación.

##### ***Software y hardware necesario***

Una adecuada infraestructura tecnológica permite utilizar al máximo todos los componentes del *data warehouse*, y por tanto dar acceso a nuevas capacidades. Ante esto se preguntó sobre la necesidad de invertir en infraestructura tecnológica para

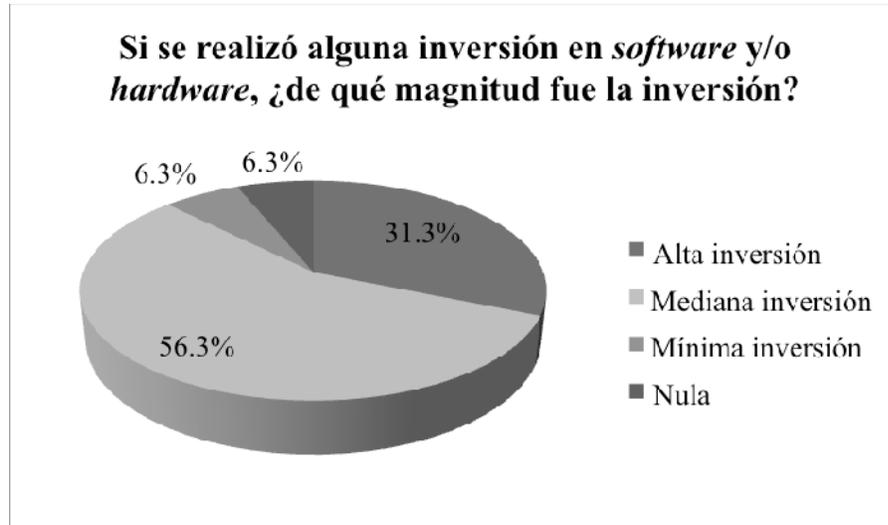
asegurar el buen desempeño del *data warehouse*, ilustrándose las repuestas en la figura 8.22.



*Figura 8.22 Cambios e inversiones en infraestructura tecnológica para el buen desempeño del data warehouse*

Sin duda una inversión en infraestructura fue una exigencia para las empresas financieras. En las respuestas se refleja el alto grado de complejidad en los requerimientos tecnológicos de un *data warehouse*, evidenciando que en casi la totalidad de los casos no se cuenta con sistemas lo suficientemente robustos.

El segundo tema abordado fue definir la magnitud de la inversión en *software* y *hardware*. La tendencia se ilustra en las respuestas de la figura 8.23.

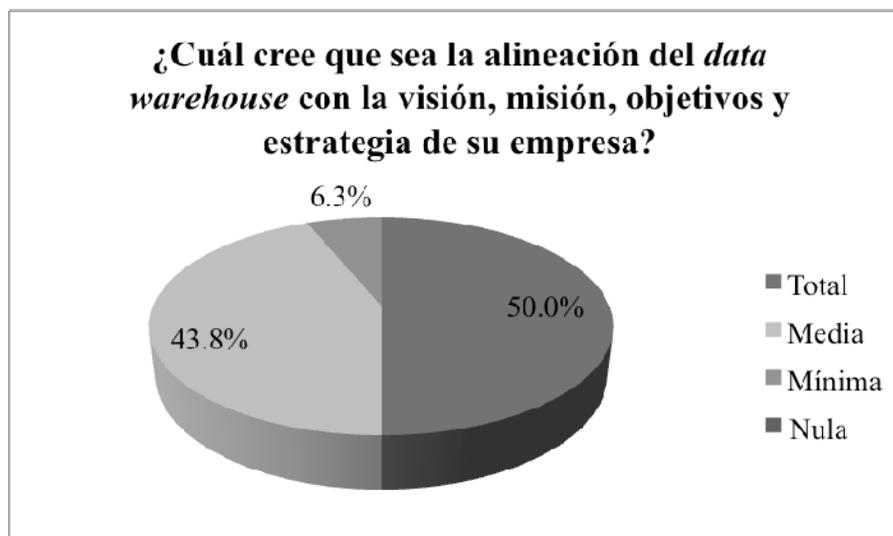


*Figura 8.23 Magnitud de la inversión en software y/o hardware*

Se puede apreciar que en general, los proyectos de *data warehouse* necesitan de una inversión considerable en *hardware* y/o *software*; esta clase de proyectos revolucionan la infraestructura tecnológica de las organizaciones, y la administración debe estar preparada para enfrentar el cambio.

#### ***Alineación de la tecnología con los objetivos estratégicos***

Para todo proyecto de informática es crítico conocer su alineación con los objetivos de la organización. Se les cuestionó sobre el caso del *data warehouse*, obteniendo los resultados que se muestran en la figura 8.24.



*Figura 8.24 Alineación del data warehouse con la visión, misión, objetivos y estrategias*

Resulta interesante observar que para la mitad de las empresas es claro que una alineación total proporciona valor a la organización. Se puede intuir el desconocimiento de cómo el *data warehouse* ayuda a las empresas financieras, solamente aprovechando las características distintivas.

#### ***Compatibilidad con el resto de los sistemas***

Fallar al definir la compatibilidad de una herramienta tecnológica es uno de los errores comunes en los proyectos de informática. Cuestionando sobre la capacidad de integración del *data warehouse* con el resto de los sistemas, se presentan los resultados en la figura 8.25.



*Figura 8.25 Compatibilidad del data warehouse con el resto de los sistemas de la organización*

Definitivamente se observa un esfuerzo por parte del equipo del *data warehouse* por conseguir integrarlo con el resto de sus sistemas. Es común que una organización tenga muchos sistemas de información de distintos proveedores, lo cual dificulta su integración, pero las empresas financieras han mostrado tener buenas políticas de implementación de nuevos sistemas, disminuyendo considerablemente este riesgo.

### ***Estrategia de implementación***

Una estrategia se define como un conjunto de tácticas para alcanzar un objetivo en específico. En cuanto a la estrategia de implementación, se les preguntó sobre cuál fue la utilizada en el proyecto de *data warehouse*. Las respuestas aparecen en la figura 8.26.

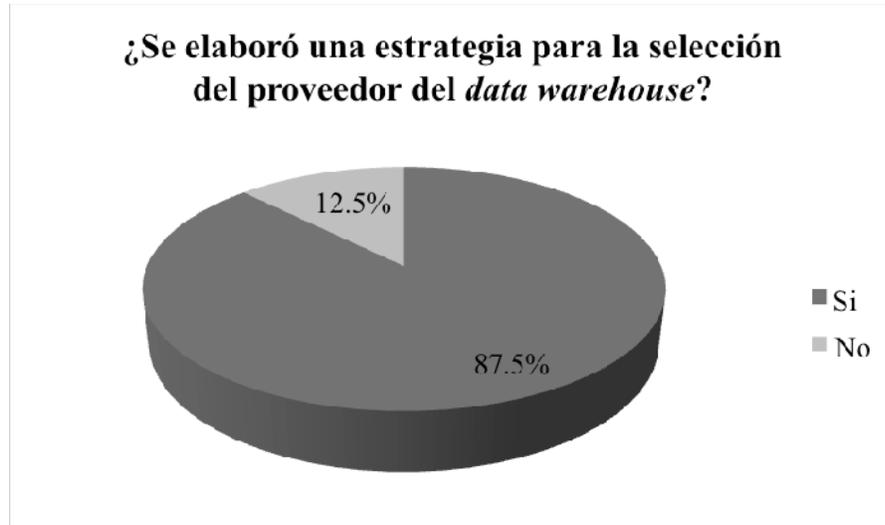


*Figura 8.26 Estrategia de implementación del data warehouse*

Las opiniones de los encuestados señalan que la estrategia de implementación predilecta resultó ser la gradual, dejando como segundas alternativas la estrategia por prototipo y directa. Las empresas financieras saben de antemano que el proyecto de *data warehouse* tiene una duración considerable, dificultando el control de las actividades y aumentando la presión ejercida sobre el equipo implementador. Para la Dirección es importante evitar los riesgos relacionados al proyecto, y las bondades de la estrategia de implementación gradual la convierten en una alternativa adecuada.

### ***Proveedores***

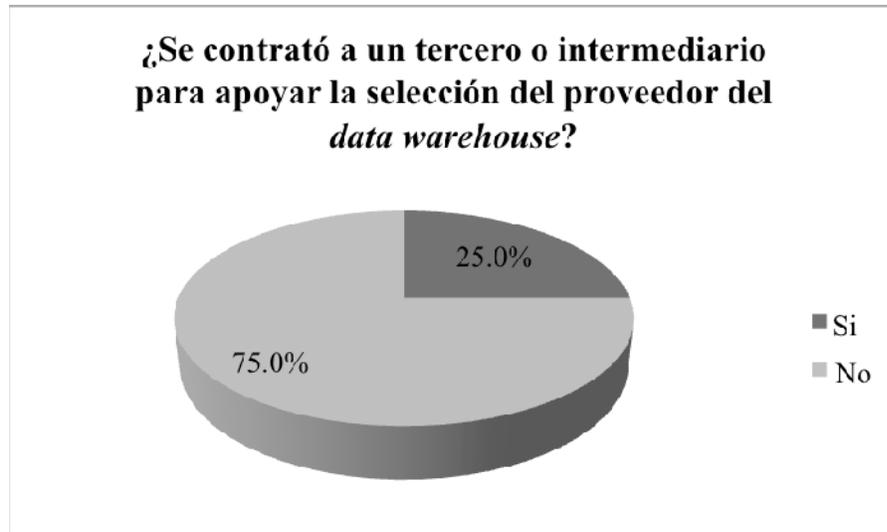
El proceso de selección de proveedor involucra una inspección minuciosa de precios, reputación, capacidades y compromisos con sus clientes. Al cuestionarles a los encuestados al respecto para el caso del *data warehouse*, se observan los resultados de la figura 8.27.



*Figura 8.27 Presencia de una estrategia para la selección del proveedor del data warehouse*

Definitivamente las empresas financieras saben que es importante elaborar una estrategia de selección de proveedor para evitar problemas posteriores. La magnitud de la inversión y el grado en que serán afectados los sistemas de información son de constante preocupación de las empresas financieras, por lo que hay una especial atención en la selección del proveedor.

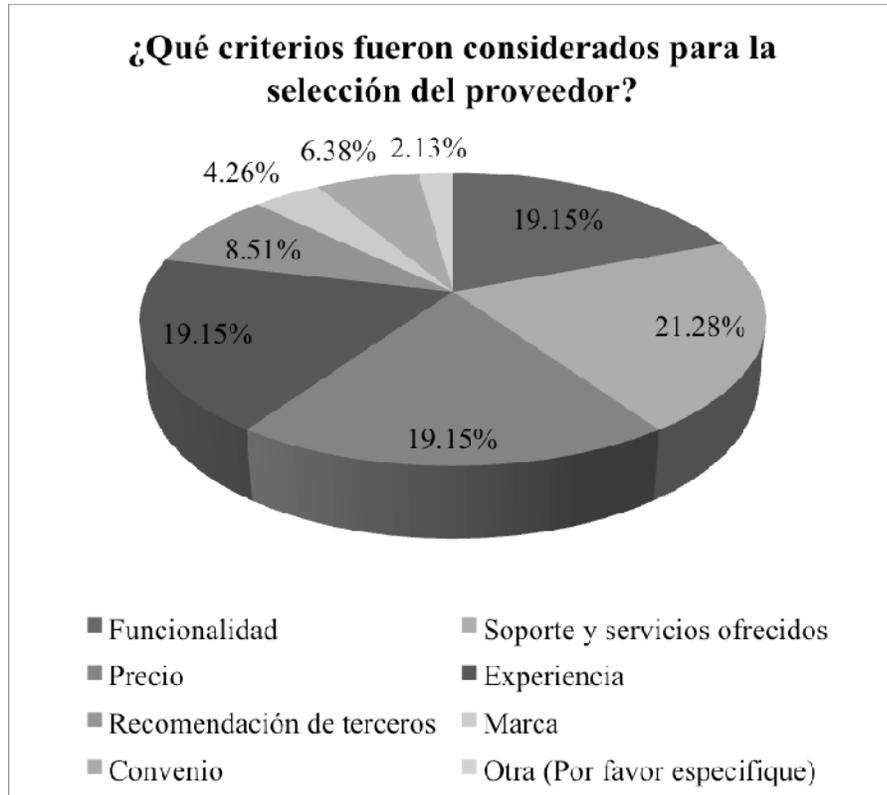
Retomando el tema de proveedores, hay que considerar la intervención de un tercero o intermediario para tener una perspectiva más clara en el proceso de selección. Al cuestionar con respecto al proyecto de *data warehouse*, se obtuvieron las respuestas reflejadas en la figura 8.28.



*Figura 8.28 Presencia de un tercero o intermediario para apoyar al proceso de selección del proveedor del *data warehouse**

Es posible observar un alto nivel de confianza en el equipo del proyecto al momento de la selección del proveedor. Se infiere que la intervención de terceros solo sucede cuando el personal no tiene los conocimientos suficientes para juzgar correctamente a los proveedores.

Una vez determinada la intervención de terceros lo siguiente fue analizar cuáles fueron los criterios utilizados para su selección. Los resultados al respecto en el caso del *data warehouse* se muestran en la figura 8.29.



*Figura 8.29 Criterios para la selección del proveedor de data warehouse*

Lo anterior indica que el criterio principal para la selección del proveedor es el soporte y servicios ofrecidos, siendo criterios secundarios el precio, la funcionalidad y la experiencia. La marca, recomendaciones, convenios y otros criterios no son bien apreciados por las empresas financieras, por lo que se puede decir que el servicio post-venta es el cimiento para la construcción de relaciones duraderas con sus proveedores.

Como cierre del tema hay que considerar el servicio de los proveedores proporcionado a las empresas financieras. Los resultados se pueden observar en la figura 8.30.



*Figura 8.30 Atención/servicio del proveedor de data warehouse*

Resulta obvio notar los proveedores del *data warehouse* han prestado un buen servicio a las empresas financieras. El bajo porcentaje de las empresas que calificaron como excelente el servicio es indicador de prudencia por parte de las financieras. Es de destacarse que la construcción de vínculos de confianza con los clientes es determinante para los proveedores, y así lo indicaron los resultados obtenidos.

### **8.5. Factores críticos para la implementación del *data warehouse* en el Sector Financiero**

Una vez analizado a detalle las respuestas de la encuesta, es necesario definir un factor crítico. Un factor crítico es algo que debe ocurrir para conseguir una meta en particular. Para este estudio la meta es lograr implementar exitosamente un *data warehouse* en el Sector Financiero de Monterrey. Considerando que todas las respuestas son originadas de proyectos exitosos, los factores encontrados se muestran a continuación.

### ***Factores críticos desde la perspectiva económica***

Aunque las organizaciones cada vez utilizan más las tecnologías de información para alcanzar mejores niveles de eficiencia en sus operaciones, muchos proyectos arrancan sin tener un entendimiento claro de los costos y riesgos involucrados. Con la finalidad de conocer a detalle esta faceta económica, se cuestionó sobre el enfoque al negocio, recursos necesarios, costo, tiempo y estudio de factibilidad. Los factores económicos identificados fueron:

- 1. La información obtenida del *data warehouse* utilizada para la toma de decisiones afecta su implementación.** Una adecuada adopción proporciona una amplia variedad de elementos de valor para los tomadores de decisiones.
- 2. La información obtenida del *data warehouse* utilizada para realizar análisis afecta su implementación.** Al adoptar la tecnología se posibilita a los usuarios realizar análisis de grandes volúmenes de datos y manipularlos de diferentes formas.
- 3. La información obtenida del *data warehouse* utilizada para realizar consultas afecta su implementación.** La facilidad de realizar consultas complejas de información estimula la adopción de la tecnología.
- 4. La información obtenida del *data warehouse* utilizada para planificar afecta su implementación.** La adopción de la tecnología provee de los elementos necesarios para hacer una planeación consolidada.
- 5. La información obtenida del *data warehouse* utilizada para elaborar reportes afecta su implementación.** Con la tecnología se pueden realizar de forma automática diversos reportes, lo que influencia positiva o negativamente en la adopción.

6. **La información obtenida del *data warehouse* utilizada para elaborar estadísticas afecta su implementación.** La facilidad de aplicación de técnicas estadísticas que proporciona la tecnología estimula su adopción.
7. **La información obtenida del *data warehouse* utilizada para elaborar pronósticos afecta su implementación.** La adopción permite a los usuarios tener una perspectiva más amplia para elaborar mejores pronósticos.
8. **La inversión en capacitación afecta la implementación del *data warehouse*.** La adopción de la tecnología incurre en grandes gastos de capacitación, siendo de importancia durante el proceso.

Casi la totalidad de los factores encontrados se refieren a capacidades adquiridas al momento de la implementación.

#### ***Factores críticos desde la perspectiva social***

Las organizaciones están formadas por la gente y la implementación de una nueva tecnología afecta la forma en cómo trabajan, y por lo tanto modifica las interrelaciones de la empresa. Dada la importancia del aspecto social, fue necesario cuestionar sobre el apoyo recibido por la administración, expectativas realistas, participación de usuarios y personas experimentadas, beneficios, y la resistencia al cambio. Los factores encontrados del aspecto social fueron:

1. **Los requerimientos computacionales afectan la implementación del *data warehouse*.** A medida que los requerimientos computacionales son más complejos, dificultan la adopción de la tecnología.

2. **El cambio en la forma de trabajo afecta la implementación del *data warehouse*.**  
La severidad con la que cambia el estilo de trabajo del capital humano estimula positiva o negativamente la adopción de la tecnología.
3. **El manejo de la información afecta la implementación del *data warehouse*.** La adopción de la tecnología mediante la manipulación de la información permite identificar y capitalizar nuevas oportunidades.
4. **La participación del personal interno y externo en conjunto afecta la implementación del *data warehouse*.** La alta volatilidad de la tecnología exige que el personal tenga conocimiento del negocio y de los detalles técnicos involucrados, para así lograr una adopción exitosa.
5. **Las habilidades y experiencia del personal involucrado afecta la implementación del *data warehouse*.** La presencia de personal capacitado estimula e influencia directamente la adopción de la tecnología.
6. **La mejora en consistencia e integración de datos influencia en la implementación del *data warehouse*.** La adopción de la tecnología incrementa el valor de la información de las empresas financieras.
7. **La facilidad de manejar indicadores influencia en la implementación del *data warehouse*.** Las empresas financieras necesitan conocer el contexto del negocio, siendo relevante el beneficio de poder manejar más indicadores.
8. **La posibilidad de realizar consultas dinámicas influencia en la implementación del *data warehouse*.** Al implementar la tecnología, los ejecutivos financieros sin necesidad de conocimientos informáticos pueden acceder a la información de cualquier sección de la organización.

Al respecto se obtuvo que para las empresas financieras es crucial manejar temas relacionados con el personal y sus habilidades, además de que la tecnología de *data warehouse* es conocida por sus beneficios relacionados con datos, indicadores y consultas dinámicas.

### ***Factores críticos desde la perspectiva tecnológica***

Las especificaciones técnicas representan las condiciones a cumplir para alcanzar el éxito, pero lo más importante es el soporte a los objetivos y la estrategia del negocio. En este punto se cuestionó sobre el *software* y *hardware*, alineación de la tecnología con el negocio, compatibilidad, estrategia de implementación, y proveedores. Los factores identificados del aspecto tecnológico fueron:

- 1. La inversión en infraestructura tecnológica afecta la implementación del *data warehouse*.** Una adopción exitosa de la tecnología es influenciada por las capacidades de la infraestructura de *hardware* y *software* de la organización.
- 2. La capacidad de integración del *data warehouse* con el resto de los sistemas afecta su adopción.** Si las empresas financieras buscan reducir costos, mejorar la productividad, y compartir la información clave, la compatibilidad de la tecnología es vital para lograr estos cometidos.
- 3. Una estrategia de selección de proveedor afecta la implementación del *data warehouse*.** La existencia de una estrategia al momento de la selección de proveedor ayuda a obtener el mayor valor al menor precio.
- 4. El servicio prestado por el proveedor afecta la implementación del *data warehouse*.** El servicio del proveedor está interconectado con la productividad de la tecnología, por lo que es un aspecto de importancia para los ejecutivos financieros.

En cuanto a la perspectiva tecnológica, se comprueba que no solo son importantes los detalles técnicos, sino que además es clave la integración y los aspectos relacionados con los proveedores.

## **8.6. Recomendaciones para la implementación exitosa del *data warehouse* en el Sector Financiero**

El proyecto de *data warehouse* es una inversión que consume tiempo, dinero y recursos, con la esperanza de recibir un mayor valor a cambio. Las cifras mostradas en los resultados de la encuesta muestran una necesidad de tomar acciones para alcanzar los beneficios esperados, listándose en seguida recomendaciones para mejorar el proceso de implementación en las financieras.

- **Elaborar un plan de monitoreo.** Es considerado como una herramienta que sirve para el diseño, implementación, administración y evaluación de las actividades de monitoreo de la información. Mediante esta planeación es posible promover esta actividad.
- **Entrenar al personal en análisis multidimensional.** El entrenamiento permite a todos los usuarios no familiarizados con la tecnología, adquirir las capacidades para obtener todos sus beneficios. Al tener los conocimientos, los usuarios podrán realizar análisis multidimensionales sin mayor problema.
- **Elaborar un *project charter*.** Un *project charter* es un plan táctico para ejecutar un proyecto de informática, que define la coordinación de recursos. Funciona como un acuerdo con los patrocinadores del proyecto en donde se especifican cuáles recursos

serán requeridos, y quién será el responsable de proveerlos. Por lo tanto, al hacer explícitos los recursos necesarios se evitan confusiones o malos entendidos.

- **Definir un *total cost of ownership*.** Conocido también como “costo total de propiedad”, permite a la administración conocer todos los costos asociados a la inversión en un proyecto de información. Al elaborar el documento es posible tener una estimación de costos de mayor exactitud.
- **Elaborar un *work breakdown structure*.** Abreviado en inglés como *WBS*, es una estructura jerárquica que permite visualizar todas las actividades que necesitan ser realizadas para completar un proyecto. El *WBS* es una herramienta efectiva que ayuda a terminar un proyecto a tiempo.
- **Utilizar la técnica *Delphi*.** La técnica *Delphi* es un método de estimación de tiempo que integra la opinión de un conjunto de expertos sobre una temática en particular. Tomando en consideración la naturaleza del *data warehouse*, un consenso sobre el tiempo de implementación es una buena alternativa para tener un reducido margen de error.
- **Utilizar la técnica *PERT*.** En inglés, *Program Evaluation and Review Technique (PERT)*, es un método de estimación que utiliza una distribución estadística para calcular el tiempo requerido para llevar a cabo las actividades de un proyecto.

Por cada actividad se hacen estimaciones de tiempo optimistas, reales, y pesimistas. Una vez obtenidas, se utiliza la siguiente fórmula en cada una de las actividades para obtener la estimación *PERT*:

Estimación: Tiempo optimista + (4 x Tiempo real) + Tiempo pesimista

Sumando el tiempo de todas las actividades se obtiene el total esperado del proyecto. El método PERT en conjunto con la técnica Delphi son de gran utilidad para el proyecto de *data warehouse*.

- **Definir el impacto en el margen de utilidad bruta.** Esta razón financiera indica el porcentaje ganado por cada unidad monetaria invertida en el negocio. La forma de calcularla es:

$$\text{Margen de Utilidad Bruta: } \frac{\text{Utilidad bruta}}{\text{Ingresos}}$$

El *data warehouse* es un proyecto que aumenta el valor de los productos de una organización por la aplicación de una estrategia más fundamentada, e impacta directamente en las ventas netas de una organización.

- **Definir el *total benefits of ownership*.** En español “beneficios totales de propiedad”, es un documento que incluye todos los beneficios directos, en transcurso, e indirectos de un proyecto en particular. Al conocer los beneficios esperados, se puede definir el impacto económico tangible del *data warehouse*.
- **Calcular el período de retorno de la inversión.** El método de retorno de la inversión (en inglés, *payback period*) determina el tiempo de recuperación de la inversión inicial de un proyecto. La fórmula para calcularlo es la siguiente:

$$\text{Payback Period: } \frac{\text{Inversión inicial}}{\text{Flujo de efectivo neto}}$$

Aunque tiene la desventaja de no considerar el dinero generado después del período de retorno de la inversión, es un buen método para definir el riesgo de la inversión. Para el proyecto de *data warehouse* proporciona evidencia tangible.

- **Calcular el punto de equilibrio.** El método de punto de equilibrio, similar al *payback period*, intenta determinar el punto en el cual el proyecto empieza a recuperar su inversión inicial. La fórmula para calcularlo se muestra a continuación:

$$\text{Punto de equilibrio: } \frac{\text{Inversión inicial}}{\text{Margen de utilidad neta}}$$

Es un indicador fácil de calcular, sirve para medir el riesgo del proyecto, además de ser un soporte al momento de presentarlo ante la administración.

- **Calcular el retorno de inversión.** Conocido también en inglés como *Return of Investment (ROI)*, es un indicador financiero que mide el valor esperado de un proyecto en particular. Se calcula de la forma mostrada en seguida:

$$\text{ROI del Proyecto: } \frac{\text{Total de beneficios esperados} - \text{Total de costos esperados}}{\text{Total de costos esperados}}$$

Calculando el ROI se muestra el valor del proyecto en términos de inversión, tiempo y recursos, siendo de utilidad para comparar los beneficios y costos del *data warehouse* con otras alternativas tecnológicas.

- **Involucrar a los usuarios mediante un grupo de enfoque.** En inglés *focus group*, es una técnica informal para evaluar las necesidades de los usuarios antes y después de arrancar un proyecto. Dicho grupo es dirigido por un moderador que interactúa con un conjunto de usuarios, dando origen a una serie de ideas y observaciones para

mejorar el proyecto. Resultaría favorable utilizar un grupo de enfoque para el *data warehouse*, con el fin de mostrar la importancia de involucrar a más personas de otras áreas de la organización, aclarar todas las inquietudes de los usuarios, y aumentar su interés.

- **Certificar al personal en tecnología de *data warehouse*.** No cabe duda que tener un equipo con lo último en conocimientos tecnológicos es clave para conseguir una adecuada implementación. Con el entrenamiento derivado de la certificación, se logra que el personal desarrolle las habilidades y el conocimiento para tener un nivel adecuado de eficiencia y efectividad.
- **Elaborar un plan de administración del riesgo.** El *Project Management Institute* (2008) define a la administración del riesgo como un proceso sistemático de identificación, análisis, y respuesta a los riesgos de un proyecto. Los riesgos amenazan el plan del proyecto, ocasionando problemas potenciales en presupuesto, calendarización, personal, y recursos. El equipo de *data warehouse* debe estar preparado para manejar toda la incertidumbre que se pueda generar en la organización.
- **Elaborar un plan de administración del cambio.** La introducción del *data warehouse* a la organización implica un cambio significativo, por lo que un plan es clave para lograr una transición efectivamente. Marchewka (2006) propone cuatro pasos para la elaboración del plan.
  1. Evaluación de la voluntad, disposición y habilidad al cambio
  2. Desarrollar o adaptar una estrategia para el cambio
  3. Implementar el plan de administración del cambio y darle seguimiento

#### 4. Evaluar la experiencia y las lecciones aprendidas

Al elaborar el plan se administra la transición del cambio, y los problemas relacionados al proyecto disminuyen considerablemente debido a la seriedad con la que se toma este aspecto.

- **Establecer la gobernanza de TI.** Sharma, Stone y Ekinici (2009) definen a la gobernanza de TI como el proceso o estructura para asegurar que las inversiones en tecnologías de información se están llevando adecuadamente. Por tratarse del Sector Financiero, es común que no haya claridad en como la tecnología de *data warehouse* se relaciona con la estrategia de la organización. Al establecer una gobernanza de TI se mantiene la confianza en la administración y se alinean los objetivos estratégicos de la organización.
  
- **Contratar a un tercero para la selección de proveedor.** Considerando el alto costo de un *data warehouse*, el proceso de selección de proveedor no debe ser tomado a la ligera para evitar problemas en tiempo, alcance y presupuesto, o peor aún, en la cancelación del proyecto. La administración debe considerar contratar a un tercero por las siguientes razones:
  1. La actividad de selección de proveedor es muy grande y compleja
  2. No se puede dedicar muchos recursos de tiempo completo a la actividad
  3. El proceso de selección por un tercero es más transparente y estructurado
  4. Se asegura que las respuestas de los proveedores sean comparables
  5. Se obtiene el mejor valor y precio

Sin duda se establecerá una relación a largo plazo con el proveedor del *data warehouse*, y para asegurar el adecuado funcionamiento de la herramienta es de utilidad contar con una perspectiva renovada.

- **Elaborar una solicitud de propuesta.** Después de que la organización ha seleccionado un proyecto, el cliente es el responsable de elaborar un documento que sirve como guía para los proveedores al momento de hacer su oferta a la organización.

Una solicitud de propuesta (en inglés, *Request for Proposal*, abreviado *RFP*), es un documento que contiene información que es muy importante para el proyecto.

Las partes que componen el RFP son:

1. Resumen de las necesidades y propuesta de acción
2. Estatuto de trabajo detallando del alcance y los entregables principales
3. Requerimientos y especificaciones de los entregables
4. Responsabilidades del proveedor y el cliente
5. Calendario del proyecto
6. Calendario de costos y pagos
7. Tipo de contrato
8. Experiencia y personal
9. Criterios de evaluación

Al tener detalle en cada una de sus partes se pueden evitar conflictos, malos entendidos, reclamaciones legales entre el proveedor y el cliente, y problemas de satisfacción del servicio.

- **Elaboración de un estudio comparativo.** Spendolini (2005) define a un estudio comparativo (en inglés, *benchmark*) como un proceso sistemático para la evaluación de productos y servicios, con la finalidad de realizar mejoras a la organización. Un análisis detallado de los proveedores para el proyecto de *data warehouse* ayuda a seleccionar la alternativa que se adecua a la estrategia de la organización, aprovecha de mejor manera la inversión en infraestructura tecnológica, y reduce los problemas relacionados con la compatibilidad.

Por último, las recomendaciones de aplicarse ayudarán a mejorar la calidad y precisión del proceso de implementación, promoviendo la tecnología del *data warehouse* en el sector.

## 9. Conclusiones y trabajos futuros

### 9.1. Conclusiones Generales

Desde tiempo atrás, los directivos enfrentan diariamente el reto de tomar decisiones cruciales que dictan el rumbo de la organización, teniendo como consecuencia un incremento exponencial en peticiones de información cada vez más complejas. Indudablemente las tecnologías de información son de gran apoyo para responder con eficacia a las demandas del mercado, sin embargo, la calidad de la información representa un obstáculo. Resulta frecuente contar con información parcial, ya que se encuentra diseminada a través de toda la organización, representando un reto su acceso, integración y organización. Turban, Sharda, Aronson y King (2008) mencionan que el *data warehouse* es una tecnología con la capacidad de integrar las bases de datos de la organización para generar una sola versión de la verdad. La ruta a seguir parece clara, pero la implementación de esta tecnología tiene algunas particularidades, que al fallar se afectara notablemente el tiempo, alcance, y presupuesto del proyecto.

Para efectos de esta tesis se decidió identificar los factores críticos para la implementación del *data warehouse* en el Sector Financiero. Aplicando una encuesta a dieciséis empresas del giro financiero de Monterrey, se cuestionaron aspectos de índole económica, social y tecnológica; con el fin de dar a conocer cuáles deben ser consideradores para futuras implementaciones, además de una serie de recomendaciones que impactarán favorablemente este proceso. A su vez, fue seleccionado el Sector Financiero por sus altas demandas de información, adicional al protagónico de los sistemas para el apoyo a la toma de decisiones observado en las empresas de este giro.

Debido al extraordinario desarrollo tecnológico de los últimos tiempos, las empresas financieras se han visto obligadas a contar con sistemas de apoyo a la toma de decisiones. Los resultados de la encuesta son prometedores, por la consolidación de la tecnología del *data warehouse* en este sector. En la mayoría de las empresas se observó que la tecnología se encuentra en la etapa de ejecución, asignada a operar en conjunto con los sistemas de CRM y ERP de la organización. Es buen momento para que la administración de las financieras reflexione sobre la importancia de tener información acertada y precisa; al contar con parcialidades se puede acarrear un problema complejo que impacte directamente en el valor del negocio.

A su vez, la consolidación de la tecnología indica el principio de procesos de implementación más sencillos. La composición de los resultados mostró que esta tarea en su mayoría se asigna a pequeños grupos de especialistas, integrados por personal tanto dentro como fuera de la organización. Sin embargo, ¿la administración se ha cuestionado sobre la correlación del equipo del proyecto, con los problemas de resistencia al cambio?, porque la evidencia muestra lo contrario. Siendo su responsabilidad, los problemas de desinterés y poca participación predominan derivados de una deficiente capacitación. Hay que dejar de delegar esta tarea solo al equipo del *data warehouse*, porque se trata de una tecnología que afectara considerablemente las operaciones de la organización, y debe ser adoptado por todos, incluyendo a la administración.

Mencionando a la administración, el apoyo recibido en el proyecto es un área de oportunidad en la que hay que continuar trabajando. ¿Arrancar un proyecto en dónde no se tiene la certeza sobre la suficiencia de recursos?, ¿un apoyo mediático en las actividades del proyecto?, ¿sólo la mitad de los proyectos alineados totalmente con la

estrategia del negocio?, estamos hablando de catalizadores que ocasionaran problemas posteriormente. En el contexto de las financieras, la administración no ha logrado percibir los riesgos involucrados en la inadecuada implementación del *data warehouse*; ya que aunque se perciben los beneficios de la tecnológica con claridad, dependerá de la calidad del proceso de implementación la conversión de éstos, a un valor económico.

En materia de puntos a mejorar, el desfase en tiempo de implementación es una constante en esta clase de proyectos. Por tratarse de un sector que no necesita adquirir nuevos sistemas de información para mantener el paso de la competencia, sino lo importante es garantizar su funcionamiento en las peores situaciones, los proyectos estratégicos, como los del *data warehouse*, pasan a segundo plano. El líder es el encargado de mantener el equilibrio entre el tiempo, alcance y presupuesto en cualquier proyecto, y para el caso concreto de las financieras se valen del tiempo, para castigar lo menos posible el presupuesto estimado. Al parecer no se han dado cuenta es la interdependencia entre esas variables, ya que al alargar el tiempo del proyecto se incurre en muchos costos indirectos derivados de la implementación, tal es el caso de los costos de consultaría de los proveedores. Resulta necesario cambiar la mentalidad de las empresas financieras con respecto a los sistemas de información, y de lograrlo aumentaría el valor de sus productos o servicios, por el valor agregado de la información de calidad.

Bajo el punto de vista de profesionista de tecnologías, la cultura del empresario mexicano le da un rol secundario a los sistemas de información. Es común oír de planes de expansión de sucursales bajo la premisa “entre más sucursales tenga, más ganancias tengo”, pero proyectos para segmentar a los clientes realmente redituables, son sólo cuestiones en las agendas de los grandes corporativos. Quedó demostrado que un

proyecto de *data warehouse* tiene implicaciones de naturaleza económica, tecnológica y social, siendo sus efectos cruciales para una implementación exitosa.

Finalmente, esta investigación resulta importante por la evidencia tangible descubierta con respecto a la tecnología del *data warehouse*, en el caso concreto del Sector Financiero. Durante el proceso de recolección de datos se contactaron a muchas empresas, antes de tener los resultados de los candidatos que cumplían los requisitos, fue asombroso encontrarse con la ignorancia del tema en varias de ellas. El cometido de la investigación se cumplió, aumentando además el conocimiento sobre tecnologías de información en México. Siempre existen estudios de esta naturaleza en países de primer mundo, pero si queremos investigaciones precisas del contexto mexicano, hay que tomar a sus organizaciones como referencia.

## 9.2. Trabajos Futuros

El *data warehouse* es una tecnología fundamental para otros sistemas de información. En esta investigación se abordó un tema de utilidad para el Sector Financiero, sin embargo esta tecnología en el contexto mexicano ofrece grandes oportunidades de investigación. A continuación se exponen algunos temas de interés para futuras investigaciones.

- **Realizar un estudio exploratorio para conocer el impacto del *cloud computing* en la adopción del *data warehouse*.** El *cloud computing* es una tendencia de época que está evolucionando los productos y servicios de *software*, no sería sorprendente ver al *data warehouse* bajo este nuevo esquema.
- **Elaborar un estudio del *data warehouse* bajo el enfoque de gobierno de tecnologías de información (TI).** Se identificaron deficiencias en el soporte

proporcionado por la administración, pero las políticas empleadas en los aspectos de TI pudieran ser las causantes de los problemas de implementación.

- **Llevar a cabo un estudio en otro sector siguiendo los lineamientos de esta investigación.** Descubrir las particularidades de otra industria, sería un buen lineamiento para una investigación comparativa.
- **Desarrollar una investigación orientada exclusivamente a conocer los detalles de la metodología de implementación.** El estudio mostró que la implementación es un proceso distintivo del *data warehouse*.
- **Hacer una investigación orientada a conocer el rol que juega el *data warehouse* en la organizaciones.** Usualmente existen deficiencias para mostrar el impacto económico tangible de un proyecto estratégico de TI.
- **Desarrollar un estudio exploratorio de la adopción del *data warehouse* bajo el esquema de código libre.** Los proveedores de *data warehouse* en formato de código libre están madurando a un ritmo acelerado, sólo es cuestión de tiempo para que se conviertan en una opción viable para las empresas.
- **Análisis a detalle sobre el uso que le dan las organizaciones al *data warehouse*.** Realizar una auditoría a detalle del *data warehouse* puede ayudar a identificar los problemas raíz de la implementación.
- **Elaborar una investigación para conocer mayor detalle el involucramiento de los usuarios en el proyecto de *data warehouse*.** Durante el estudio se encontraron distintas opiniones con respecto a la participación del usuario, cuando ellos determinan el éxito o fracaso de la herramienta en la organización.

La demanda en información más confiable y en un mejor formato continuará creciendo exponencialmente, por lo que esta investigación es referencia a las vertientes que le deparan en el futuro al *data warehouse*.

## Apéndices

### Apéndice 1. Formato de la encuesta

#### Introducción

El objetivo de esta encuesta es recolectar los datos necesarios para la investigación de campo de la tesis titulada “Factores Críticos para la Implementación del *Data Warehouse* en el Sector Financiero de Monterrey”, la cual es requisito para obtener el grado de Maestro en Administración de Tecnologías de Información del Tecnológico de Monterrey. La investigación tiene únicamente fines académicos, por lo cual la información recabada será tratada con absoluta confidencialidad.

De antemano agradezco su apoyo.

Ing. Isaías Velázquez Reyes

#### Instrucciones

- No hay respuestas correctas o incorrectas. Éstas simplemente reflejan su opinión personal.
- Marque con claridad la opción elegida con una cruz. Recuerde: NO se deben marcar dos opciones.
- Si tiene dudas con respecto a la pregunta, por favor cuestione a la persona que le entregó la encuesta.

#### Preguntas

1. ¿Cuenta su organización con alguno de los siguientes sistemas de información?

ERP - Enterprise Resource Planning	CRM - Customer Relationship Management	SCM - Supply Chain Management	Ninguno
O	O	O	O

2. ¿Cuál es el nombre de la herramienta de *data warehouse* que utilizan?

- a) Teradata DBMS
- b) Oracle DBMS
- c) IBM InfoSphere Balanced Warehouse
- d) Netezza DBMS
- e) Microsoft SQL Server
- f) Desarrollo propio
- g) Otro. Especifique \_\_\_\_\_

3. ¿En qué fase de implementación se encuentra la herramienta de *data warehouse*?

- a) Definición
- b) Planeación
- c) Ejecución
- d) Cierre
- e) Evaluación

4. ¿Considera importante y/o necesario para su empresa la implementación de un *data warehouse*?

Si	No
----	----

O	O
---	---

5. Clasifique el grado de importancia que tiene la información obtenida del *data warehouse* en las siguientes actividades

	Muy importante	Importante	Poco importante	Sin importancia
Monitoreo de la información	O	O	O	O
Análisis de la información	O	O	O	O
Consulta de la información	O	O	O	O
Investigación	O	O	O	O
Toma de decisiones estratégicas	O	O	O	O
Planeación	O	O	O	O
Elaboración de reportes	O	O	O	O
Elaboración de estadísticas	O	O	O	O
Análisis multidimensional	O	O	O	O
Pronósticos	O	O	O	O

6. ¿Fue necesario invertir en capacitación para el personal al momento de la implementación del *data warehouse*?

Si	No
O	O

7. “El costo de la implementación fue de acuerdo a lo planeado”

Totalmente de acuerdo	De acuerdo	En desacuerdo	Totalmente en desacuerdo
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

8. Al inicio de la implementación del *data warehouse*, ¿se contó con los recursos necesarios para realizar el proyecto?

Siempre	Casi siempre	Ocasionalmente	Nunca
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

9. ¿La duración de la implementación fue de acuerdo a lo planeado?

Totalmente de acuerdo	De acuerdo	En desacuerdo	Totalmente en desacuerdo
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

10. ¿Cuánto tiempo duró la implementación del *data warehouse*?

De 0 a 6 meses	De 6 a 12 meses	De 1 a 2 años	De 3 a 5 años	Más de 5 años
<input type="radio"/>				

11. ¿Se realizó un análisis de costo/beneficio para definir si los recursos existentes eran suficientes para soportar la implementación de un *data warehouse*?

Si	No
O	O

12. ¿Fue notorio el interés y compromiso de los niveles directivos?

- a) Todo el proyecto
- b) Durante casi todo el proyecto
- c) Al final del proyecto
- d) A la mitad del proyecto
- e) En la etapa inicial
- f) Nunca

13. De la siguiente lista, clasifique el apoyo recibido por parte de los directivos en cada una de las siguientes actividades del proyecto:

	Siempre	Usualmente	Rara vez	Nunca
Capacitación y entrenamiento	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Seguridad de la información	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Requerimientos de infraestructura tecnológica	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Adaptación de la herramienta al negocio	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Estándares de calidad	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Aspecto humano	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

14. De la siguiente lista, clasifique el apoyo recibido por parte de los directivos en cada una de las siguientes etapas del proyecto:

	Siempre	Usualmente	Rara vez	Nunca
Elaboración de la iniciativa del proyecto	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Asignación del presupuesto	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Elaboración del plan y/o agenda de trabajo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Selección y evaluación del <i>software</i>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Implementación del <i>data warehouse</i>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Capacitación	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

15. De acuerdo a su experiencia, clasifique el grado de importancia de las siguientes opciones en un proyecto de *data warehouse*

	Muy importante	Importante	Poco importante	Sin importancia
Requerimientos computacionales	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Cambio de paradigmas	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Cambio en forma de trabajo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Manejo de la información	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Apoyo en la toma de decisiones	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Reingeniería de procesos	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
--------------------------	-----------------------	-----------------------	-----------------------	-----------------------

16. ¿En qué etapa del proyecto de implementación se involucró a los usuarios?

Todo el proyecto	Durante casi todo el proyecto	Al final del proyecto	A la mitad del proyecto	En la etapa inicial	Nunca
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

17. ¿Cuántas personas están/estuvieron involucradas en el proyecto de *data warehouse*?

De 1 a 5	De 6 a 10	De 11 a 15	De 15 a 20	Más de 20
<input type="radio"/>				

18. ¿Qué tipo de personal estuvo involucrado en la implementación del *data warehouse*?

Personal interno	Personal externo	Ambos
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

19. ¿El personal involucrado contaba/cuenta con las habilidades y la experiencia para implementar un *data warehouse*?

Si	No
<input type="radio"/>	<input type="radio"/>

20. De acuerdo a su experiencia clasifique los beneficios proporcionados por el *data warehouse*

	Muy importante	Importante	Poco importante	Sin importancia
Mejora en la estructura de la información	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Apoyo en la toma de decisiones	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mejora en la consistencia e integración de los datos	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Manejo de indicadores	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Manejo de consultas dinámicas	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Información actualizada	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

21. De la siguiente lista, clasifique aquellas situaciones que se presentaron durante el proceso de implementación

	Siempre	Usualmente	Rara vez	Nunca
Desinterés de algunas áreas	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Negación al uso de la herramienta	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Incertidumbre y/o molestia por parte de los usuarios	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Poca participación durante la capacitación	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Falta de compromiso de los participantes y usuarios	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
---	-----------------------	-----------------------	-----------------------	-----------------------

22. ¿Fue necesario realizar cambios e inversiones en infraestructura tecnológica (*hardware* y/o *software*) para el buen desempeño del *data warehouse*?

<input type="radio"/> Si	<input type="radio"/> No
<input type="radio"/>	<input type="radio"/>

23. Si se realizó alguna inversión en *software* y/o *hardware*, ¿de qué magnitud fue la inversión?

<input type="radio"/> Alta inversión	<input type="radio"/> Mediana inversión	<input type="radio"/> Mínima inversión	<input type="radio"/> Nula
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

24. ¿Cuál cree que sea la alineación del *data warehouse* con la visión, misión, objetivos y estrategia de su empresa?

<input type="radio"/> Total	<input type="radio"/> Media	<input type="radio"/> Mínima	<input type="radio"/> Nula
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

25. ¿El *data warehouse* puede integrarse a los otros sistemas de su empresa?

Si	No
<input type="radio"/>	<input type="radio"/>

26. ¿Qué estrategia se utilizó para la implementación del *data warehouse*? (N/A: No Aplica)

Directa	Paralela	Gradual	Prototipo	N/A
<input type="radio"/>				

27. ¿Se elaboró una estrategia para la selección del proveedor del *data warehouse*?

Si	No
<input type="radio"/>	<input type="radio"/>

28. ¿Se contrató a un tercero o intermediario para apoyar la selección del proveedor del *data warehouse*?

Si	No
<input type="radio"/>	<input type="radio"/>

29. ¿Qué criterios fueron considerados para la selección del proveedor?

- Funcionalidad
- Soporte y servicio ofrecidos

- Precio
- Experiencia
- Recomendación de terceros
- Marca
- Convenio
- Otro. Especifique \_\_\_\_\_

30. ¿Cómo es la relación con su proveedor de *data warehouse*?

Excelente	Buena	Mala	Pésima
○	○	○	○

***Contacto***

Si usted está interesado en conocer los resultados de esta encuesta, por favor, introduzca su dirección de correo electrónico. Los resultados se le harán llegar en la brevedad posible.

---

Muchas gracias por su colaboración

## Referencias

- Benaroch, M. (2002). Managing information technology risk: a real options perspective. *Journal of Management Information Systems*, 19(2), 43–84.
- Beyer, M. A. (2009). Overview of Data Warehouse Project Delivery in 2009. *Gartner Research*.
- Beyer, Feinberg, & Berg. (2008). Key Issues for Delivering a Data Warehouse Project. *Gartner Research*.
- Bharadwaj, A., Keil, M., & Mähring, M. (2009). Effects of information technology failures on the market value of firms. *The Journal of Strategic Information Systems*, In Press, Corrected Proof.
- Bitterer, A. (2005). Steer Clear of Common Data Warehousing Pitfalls. *Gartner Research*.
- Blechar, M., & Thompson, J. (2009). Applying Data Mart and Data Warehousing Concepts to Metadata Management. *Gartner Research*.
- Crandall, R., & Crandall, W. (2008). *New Methods of Competing in the Global Marketplace: Critical Success Factors from Service and Manufacturing* (1st ed.). Boca Raton, FL: CRC Press.
- Feinberg, D. (2009a). *Latin America Scenario: The Current State and Future Direction of IT*. Paper presented at the Gartner Enterprise Integration Summit 2009.
- Feinberg, D. (2009b). *The Future of DBMS for Data Warehousing*. Paper presented at the Gartner Enterprise Integration Summit 2009.

Feinberg, D. (2009c). *The Truth About Open-Source BI/DI/DW*. Paper presented at the Gartner Enterprise Integration Summit 2009.

Feinberg, D., & Beyer, M. A. (2008). Magic Quadrant for Data Warehouse Database Management Systems. *Gartner Research*.

Goldfinch, S. (2007). Pessimism, Computer Failure, and Information Systems Development in the Public Sector. *Public Administration Review*, 67(5), 917-929.

Han, J., & Kamber, M. (2006). *Data Mining: Concepts and Techniques* (2nd ed.). San Francisco, CA: Morgan Kaufmann.

Hawkins, S. (2008). Magic Quadrants and MarketScopes: How Gartner Evaluates Vendors Within a Market. *Gartner Research*.

Hernández, R., Fernández, C., & Baptista, P. (2006). *Metodología de la Investigación* (4th ed.). México, D.F.: McGraw-Hill Interamericana.

Hwang, H. G., Ku, C. -Y., Yen, D. C., & Cheng, C. -C. (2004). Critical factors influencing the adoption of data warehouse technology: a study of the banking industry in Taiwán. *Decision Support Systems*, 37(1), 1-21.

Imhoff, C., Gallemmo, N., & Geiger, J. G. (2003). *Mastering Data Warehouse Design: Relational and Dimensional Techniques* (1st ed.). Indianapolis, IN: Wiley.

Inmon, W. H. (2005). *Building the data warehouse* (4th ed.). Indianapolis, IN: Wiley.

Inmon, W. H., Strauss, D., & Neushloss, G. (2008). *DW 2.0: The Architecture for the Next Generation of Data Warehousing* (1st ed.). Burlington, MA: Morgan Kaufmann Publishers Inc.

Kimball, R., & Caserta, J. (2004). *The Data Warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data* (1st ed.). Boca Raton, FL: Wiley.

Kimball, R., & Ross, M. (2002). *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling* (2nd ed.). New York, NY: Wiley.

Lee, J. C., & Myers, M. D. (2004). Dominant actors, political agendas, and strategic shifts over time: a critical ethnography of an enterprise systems implementation. *The Journal of Strategic Information Systems*, 13(4), 355-374.

Lehman, J. (2008). Magic Quadrants and MarketScopes: How Gartner Evaluates Vendors Within a Market. *Gartner Intraweb*.

Litan, R. E., & Rivlin, A. M. (2001). *Beyond the Dot-coms: The Economic Promise of the Internet* (1st ed.). Washington, DC: Brookings Institution Press.

Marchewka, J. T. (2006). *Information Technology Project Management* (2nd ed.). Danvers, MA: Wiley.

Marcos, M. d. S. (1998). *Manual para la elaboración de tesis* (1st ed.). México, D.F.: Trillas.

Mundy, J., Thornthwaite, W., & Kimball, R. (2006). *The Microsoft Data Warehouse Toolkit: With SQL Server 2005 and the Microsoft Business Intelligence Toolset* (1st ed.). Indianapolis, IN: Wiley.

Pressman, R. S. (2005). *Ingeniería del software: Un enfoque práctico* (6th ed.). México, D.F.: McGraw Hill.

Project Management Institute (2008). *A Guide to the Project Management Body of*

*Knowledge* (4th ed.). New York, NY: Project Management Institute.

Ponniah, P. (2001). *Data Warehousing Fundamentals: A Comprehensive Guide for IT Professionals* (1st ed.). New York, NY: Wiley-Interscience.

Rainardi, V. (2008). *Building a Data Warehouse: With Examples in SQL Server* (1st ed.). New York: Apress.

Raval, A. (2008). *Data Warehousing To Drive Innovation At Bank Of Baroda*

[Página de BizTech2.com]. 28 de julio de 2009

<<http://biztech2.in.com/india/casestudies/enterprise-solutions/data-warehousing-to-drive-innovation-at-bank-of-baroda/35071/0>>.

Rivard, S., Raymond, L., & Verreault, D. (2006). Resource-based view and competitive strategy: An integrated model of the contribution of information technology to firm performance. *The Journal of Strategic Information Systems*, 15(1), 29-50.

"scientific method n." The New Oxford American Dictionary, second edition. Ed. Erin McKean. Oxford University Press, 2005. Oxford Reference Online. Oxford University Press. ITESM Campus Monterrey. 18 January 2010 <<http://0-www.oxfordreference.com.millenium.itesm.mx/views/ENTRY.html?subview=Main&entry=t183.e68621>>

Sharma, D., Stone, M., & Ekinici, Y. (2009). IT governance and project management: A qualitative study. *Journal of Database Marketing & Customer Strategy Management*, 16, 29-50.

Silvers, F. (2008). *Building and Maintaining a Data Warehouse* (1st ed.). Danvers, MA: Auerbach Publications.

- Spendolini, M. (2005). *Benchmarking* (1st ed.). México D.F.: Grupo Editorial Norma.
- Standish Group (2001). *CHAOS* (1st ed.). West Yarmouth, MA: The Standish Group.
- Tarafdar, M., & Gordon, S. R. (2007). Understanding the influence of information systems competencies on process innovation: A resource-based view. *The Journal of Strategic Information Systems*, 16(4), 353-392.
- Thoo, E., Friedman, T., Beyer, M. A., Feinberg, D., Bitterer, A., Bell, T., et al. (2008). Hype Cycle for Data Management, 2008. *Gartner Research*.
- Turban, E., Sharda, R., Aronson, J. E., & King, D. (2008). *Business Intelligence: A Managerial Approach* (1st ed.). Saddle River, NJ: Pearson.
- Watson, H. J., Goodhue, D. L., & Wixom, B. H. (2002). The benefits of data warehousing: why some organizations realize exceptional payoffs. *Information & Management*, 39(6), 491-502.
- Wixom, B. H., & Watson, H. J. (2001). An Empirical Investigation of the Factors Affecting Data Warehousing Success. *MIS Quarterly*, 25(1), 17-41.