

**EL DATA WAREHOUSE COMO ARQUITECTURA DE
INFORMACION Y SU PROCESO DE ASIMILACION EN UNA
ORGANIZACION**



TESIS PRESENTADA

POR

Jorge Alberto Padilla Balderas

**PRESENTADA A LA DIRECCION ACADEMICA
DE LA UNIVERSIDAD VIRTUAL DEL INSTITUTO
TECNOLOGICO Y DE ESTUDIOS SUPERIORES
DE MONTERREY COMO REQUISITO PARCIAL
PARA OPTAR AL TITULO DE :**

**MAESTRO EN ADMINISTRACION DE TECNOLOGIAS
DE INFORMACION**

DICIEMBRE DE 1997

**MAESTRIA EN ADMINISTRACION DE TECNOLOGIAS
DE INFORMACION**

EL DATA WAREHOUSE COMO ARQUITECTURA DE
INFORMACION Y SU PROCESO DE ASIMILACION EN UNA
ORGANIZACION



Tesis presentada

por

JORGE ALBERTO PADILLA BALDERAS

Presentada a la Dirección Académica de la Universidad Virtual del
Instituto Tecnológico y de Estudios Superiores de Monterrey como
requisito parcial para optar al título de

MAESTRO EN ADMINISTRACIÓN DE TECNOLOGÍAS DE
INFORMACIÓN

DICIEMBRE DE 1997

Maestría en Administración de Tecnologías de Información

EL DATA WAREHOUSE COMO ARQUITECTURA DE
INFORMACION Y SU PROCESO DE ASIMILACION EN UNA
ORGANIZACION

Tesis presentada

por

JORGE ALBERTO PADILLA BALDERAS

Presentada a la Dirección Académica de la Universidad Virtual del
Instituto Tecnológico y de Estudios Superiores de Monterrey como
requisito parcial para optar al título de

MAESTRO EN ADMINISTRACIÓN DE TECNOLOGÍAS DE
INFORMACIÓN


DICIEMBRE DE 1997

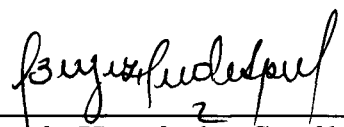
Maestría en Administración de Tecnologías de Información

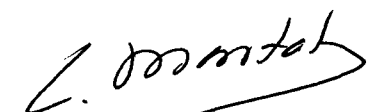
UNIVERSIDAD VIRTUAL DIRECCION ACADEMICA
PROGRAMA DE GRADUADOS EN INGENIERIAS Y TECNOLOGIAS


Los miembros del comité de tesis recomendamos que la presente tesis del Ing. Jorge Alberto Padilla Balderas sea aceptada como requisito parcial para obtener el grado académico de Maestro en Administración de Tecnologías de Información.

Comité de tesis:


Juan Raúl Esparza Martínez, MC
ASESOR PRINCIPAL


Benjamín Hernández Sepúlveda. MC.
SINODAL


Antonio Montalvo Robles, PhD.
SINODAL


María del Socorro Marcos Marcos, PhD.
Directora del Programa de Graduados en
Ingenierías y Tecnologías

Noviembre de 1997

EL DATA WAREHOUSE COMO ARQUITECTURA DE
INFORMACION Y SU PROCESO DE ASIMILACION EN UNA
ORGANIZACION

TESIS

MAESTRIA EN ADMINISTRACION DE TECNOLOGIAS DE
INFORMACION

INSTITUTO TECNOLOGICO Y DE ESTUDIOS SUPERIORES
DE MONTERREY

POR

JORGE ALBERTO PADILLA BALDERAS

DICIEMBRE DE 1997

DEDICATORIA

A Dios, por llenar mi vida de bendiciones y por darme el privilegio de tener el conocimiento y la fe en su hijo Jesucristo.

A mi Padre, Guillermo Padilla Medrano, por brindarme su amor y apoyo en todo momento de mi vida... por ser un ejemplo de honestidad, responsabilidad y disciplina... por apoyarme en cada etapa académica de mi vida y por darme siempre la libertad de emprender lo que más me ha gustado... y sobre todas las cosas, por ser un gran padre, esposo e hijo... muchas gracias.

A mi Madre, Graciela Balderas, por su infinito amor y cariño, por darme los mejores consejos para convertirme en una persona de buenos principios, y que mejor ejemplo que el tuyo para lograrlo... por estar pendiente siempre de mis estudios y de todo lo que ha sido mi vida, no tengo palabras para expresar mi agradecimiento a ti a mi padre por darme la vida, los quiero mucho.

A mi novia, Idalia Tavitas Castillo, por estar siempre a mi lado durante toda la etapa de mi maestría... por compartir conmigo todos los momentos de alegría y de presión en cada materia que cursé, representando un sacrificio que supimos llevar en nuestra relación... por formar parte de mi vida durante los últimos 4 años, en los cuales he recibido amor, comprensión y apoyo en las actividades que he emprendido... doy gracias a Dios por tu vida y por unir nuestros caminos... gracias por todo.

A mis hermanos, Zita y Guillermo, por apoyarme en cada etapa estudiantil de mi vida, y por brindarme su amor y cariño en todo momento... que Dios los bendiga a ustedes y a sus familias.

RECONOCIMIENTOS

Al Ing. Juan Raúl Esparza Martínez, le agradezco el tiempo e interés que mostró hacia el trabajo y por haberme confiado el desarrollo del tema de esta tesis, así también por apoyarme con sus comentarios sobre el escrito y por sus sugerencias en la búsqueda de referencias bibliográficas.

Al Dr. Antonio Montalvo Robles, por sus correcciones y recomendaciones sobre la tesis, por estar siempre dispuesto a realizar las revisiones de este trabajo de investigación y por su participación en la defensa del mismo.

Al Ing. Benjamín Hernández Sepúlveda, por sus comentarios y aportaciones para el desarrollo del tema y por participación en las juntas de revisiones de avances de la tesis.

A los Ingenieros Alfonso Molgado Castañeda, Jorge Chávez Villanueva, Eduardo Anaya Ramírez, Miguel Angel Peña Blanco, Jaime Maysse Muñoz y José Luis Noriega López, por su apoyo tanto en tiempo como en motivación para la realización de esta tesis y en general durante todas las materias de la maestría.

A Iván Cantú Mendieta, por su confianza, apoyo y amistad, por ser una persona en la que puedo contar en cualquier momento, y por ser participe en el inicio de las etapas profesional y de estudios de postgrado.

A mis compañeros de trabajo Gerardo Valencia, Roberto Lara, Héctor López, Lucy Rodríguez, Faustino García, Paty Bujanos, Eduardo Ulloa, Alejandro Castro, Manuel Blanco, David Acosta, y en general al departamento de informática de la Refinería Madero por su apoyo de alguna manera u otra en la realización de esta tesis.

A mis compañeros de maestría Orlando Figón, José Juan Sánchez, Selene Bonilla, Madeleine Stiles, Alba Picos, Ada Delly Silva, María Elena Bolaños, Marianela Vázquez, Juan Carlos Sánchez, Lizet Cantú, de quienes siempre tuve ayuda, buenos consejos y un aliento constante durante los estudios de maestría.

A la Lic. Marisol Saldaña, Lic. Juan Dicostanzo Zaragoza, Lic. Rubén Ramírez, Estelita Meza y al Sr. Mauro de la Universidad Virtual y biblioteca del campus Tampico, por el apoyo y facilidades para la realización de mis estudios de Maestría.

Al Dr. César Balderas García, Sr. Manuel Soto y Juana María Soto, por sus consejos, apoyo y motivación en los momentos más importantes de mi vida.

A mis amistades, Carlos Sarmiento, Mayra Martínez, José Angel Hernández, Alberto Ortega, Irwin Aguilar, Guillermo Mondragón, Alma Villalobos, Juan Carlos Alvarez, por su amistad y confianza, y que aún con el tiempo y la distancia seguimos compartiendo momentos alegres y experiencias.

Al Sr. Gilberto Tavitas, Sra. Elodia Castillo, Verónica Tavitas, Jesús González, Itzamná Tavitas y Narciso Martínez, por su apoyo constante e interés durante cada uno de los cursos de la maestría.

A Pemex-Refinación y al Instituto Tecnológico de Estudios Superiores de Monterrey, 2 grandes Instituciones de nuestro país, por darme la oportunidad de estudiar una maestría, por brindarme todas las facilidades y apoyo económico, así como por ofrecerme la oportunidad de formar parte de ellas y ser parte de su prestigio nacional e internacional.

RESUMEN

En esta tesis se analiza la tecnología de información Data Warehouse, sus características y conceptos generales, se propone visualizarla como una arquitectura de información compuesta por diversas tecnologías de apoyo a la toma de decisiones y por procesos relacionados con la administración de los datos dentro de una organización, asimismo se analiza su introducción y asimilación en una empresa evaluando su grado de aceptación entre los miembros de la misma.

Más que orientar el desarrollo de la tesis hacia un análisis del Data Warehouse desde un punto de vista tecnológico y de procesamiento de datos, se analiza con un enfoque hacia el negocio y hacia las necesidades de información más críticas para la toma de decisiones, para lo cual es necesario conocer a la empresa en sus procesos operativos y organizacionales, visualizar asimismo el verdadero valor de la información y considerar a los datos como un activo más del negocio. Así en esta tesis se utilizan metodologías para poder analizar a una organización de tal manera que puedan obtenerse resultados que apoyen el diseño de un Data Warehouse y que faciliten su introducción en una organización, además se hace énfasis en implantar un Data Warehouse conforme a los objetivos de la empresa, para lo cual se emplea una metodología de alineación estratégica entre la tecnología y el negocio, principalmente en lo que se refiere al soporte del Data Warehouse en el proceso de toma de decisiones en cualquier nivel organizacional.

Debido a que la tecnología Data Warehouse es nueva en las industrias en México, la propuesta principal de esta tesis es presentar un proceso de asimilación tecnológica que permita entender y conocer a dicha innovación, así como identificar los posibles atributos organizacionales, técnicos y humanos que pueden generarse durante su implantación en una empresa. El objetivo final del proceso de asimilación de tecnología es identificar el grado de adopción entre los miembros de un negocio hacia la nueva tecnología, y su aceptación como una herramienta de uso frecuente para la toma de decisiones y para la generación de reportes ejecutivos. Es importante conocer las propuestas de diseño y las etapas de implantación de un Data Warehouse, pero para fines de esta tesis es más relevante determinar los atributos y barreras organizacionales que no permitan a los diseñadores del Data Warehouse implantarla como una arquitectura de información que impacte en las funciones de la alta dirección y en general del personal que toma decisiones. Así durante el desarrollo del proceso de asimilación de la tecnología Data Warehouse se busca equilibrar el conocimiento de la innovación y sus aspectos tecnológicos con los atributos organizacionales que podrían afectar su introducción en una empresa.

TABLA DE CONTENIDO

	Página
Lista de Figuras	xi
Lista de tablas	xiii
Capítulo 1 Introducción	1
1.1. Descripción de la tesis en base al tema investigado.....	1
1.2. Alcance de la investigación.....	4
1.3. Objetivo de la tesis.....	4
1.4. Resultados esperados.....	5
1.5. Descripción de la tesis por capítulos.....	5
1.6. Definiciones, abreviaturas y notas importantes.....	6
Referencias Bibliográficas.....	8
Capítulo 2 Fundamentos teóricos del Data Warehouse.....	9
2.1. Introducción.....	9
2.2. Evolución histórica del procesamiento de datos.....	10
2.3. La nueva propuesta en base de datos.....	17
2.4. Concepto Data Warehouse.....	18
2.4.1. Datos orientados al negocio.....	18
2.4.2. Datos integrados.....	19
2.4.3. Datos que varían en el tiempo.....	19
2.4.4. Datos no volátiles.....	21
2.5. Estructura de un Data Warehouse.....	21
2.6. Sumarización o agregación de datos.....	24
2.7. Metadata.....	25
2.8. Partición de datos.....	26
2.9. El flujo de información en el Data Warehouse.....	29
2.10. Definición del Data Warehouse para fines de esta tesis.....	31
2.11. Conclusiones.....	33
Referencias bibliográficas.....	35
Capítulo 3 Arquitectura de información Data Warehouse.....	37
3.1. Introducción.....	37
3.2. El Data Warehouse como arquitectura de información.....	38
3.3. Tecnología de Base de Datos.....	41

3.4. Herramientas Data Warehouse.....	44
3.5. Data Mining.....	46
3.6. Administración de datos.....	48
3.7. Sistemas de soporte a la decisión.....	52
3.7.1. El Data Warehouse y el proceso de toma de decisiones.....	54
3.8. Sistemas de información ejecutivos.....	58
3.9. Conclusiones.....	61
Referencias Bibliográficas.....	64
Capítulo 4 Análisis de la empresa y su alineación a la arquitectura de información Data Warehouse.....	67
4.1. Introducción.....	67
4.2. Tipos de análisis de la empresa para fines de la arquitectura de información Data Warehouse.....	68
4.3. Flujo de información vertical.....	70
4.3.1. Cadena de Medios-Fines.....	70
4.4. Flujo de información horizontal.....	73
4.4.1. Cadena de valor.....	74
4.5. Relación de Datos-Procesos.....	77
4.5.1. Metodología BSP.....	78
4.6. Alineación de la tecnología Data Warehouse a los objetivos del negocio.....	81
4.6.1. Planeación Informática.....	83
4.7. Proceso de asimilación de tecnología.....	87
4.7.1. Modelo para asimilación de Innovaciones de Meyer y Goes.....	87
4.7.2. Modelo para la asimilación de tecnología de McFarlan y McKenney.....	89
4.7.3. Data Warehouse como innovación tecnológica.....	89
4.8. Conclusiones.....	90
Referencias bibliográficas.....	92
Capítulo 5 Proceso de Asimilación de la Arquitectura de Información Data Warehouse en una organización.....	94
5.1. Introducción.....	94
5.2. Atributos que influyen en la asimilación de la tecnología Data Warehouse	96
5.2.1. Atributos de las innovaciones.....	96
5.2.2. Atributos de los contextos organizacionales.....	97

5.2.3. Atributos procedentes de la interacción de las innovaciones y el contexto organizacional.....	99
5.3. Grado de Adopción de las Innovaciones Tecnológicas.....	100
5.4. Etapas en el diseño de un Data Warehouse.....	101
5.4.1. Propuesta - Ralph Kimball.....	101
5.4.2. Propuesta - William Inmon.....	104
5.4.3. Propuesta - John Ladaga.....	105
5.4.4. Propuesta - A.J. Brown.....	107
5.4.5. Análisis comparativo de las propuestas de diseño de Data Warehouse.....	108
5.5. Proceso de Asimilación de la Arquitectura de información Data Warehouse en una Organización.....	111
Fase I: Investigación.....	115
Fase II: Conocimiento-Persuasión.....	125
Fase III: Evaluación-Decisión.....	139
Fase IV: Introducción-Adopción.....	156
Fase V: Confirmación-Expansión.....	169
5.6. Conclusiones.....	172
Referencias bibliográficas.....	174
 Capítulo 6 Conclusiones y Trabajos Futuros	 177
 Vita	 182

LISTA DE FIGURAS

Figura		Página
2.1	Evolución del Procesamiento de Datos	11
2.2	Evolución natural de los datos (Telaraña de información)	13
2.3	Comparación de datos e información	16
2.4	Integración de datos	20
2.5	Variantes en el tiempo	20
2.6	Datos no volátiles	21
2.7	Estructura de un Data Warehouse	22
2.8	Estrategia de agregación: Sumarización de ventas	24
2.9	Ideas para la partición de datos	27
2.10	Dinámica del flujo de información	30
2.11	Arquitectura de información Data Warehouse	33
3.1	Tipos de Bases de Datos	42
3.2	Ejemplos de los tipos de Bases de Datos	43
3.3	Herramientas claves en un proyecto Data Warehouse	46
3.4	Estructura general de un modelo	56
3.5	Maneras en que el Data Warehouse soporta a un EIS	59
4.1	Cadena de Medios-Fines	70
4.2	Nivel de sumarización en el Data Warehouse	71

4.3	Relación Data Warehouse - Cadena de Medios-fines	73
4.4	Cadena de Valor	75
4.5	Sumarización en ventas para la creación de un Data Mart	76
4.6	Ejemplo de Análisis de Estabilidad	80
4.7	Matriz de relación Datos-Procesos	81
4.8	Estructura de un procedimiento de alineamiento del Plan TI/SI con las estrategias de la empresa	85
5.1	Tabla “fact” y sus dimensiones identificadas	103
5.2	Tendencias en la implantación de un Data Warehouse	134

LISTA DE TABLAS

Tabla	Página
2.1 Diferencias entre las bases de datos operacionales y las bases de datos de información o Data Warehouse	18
2.2 Tipos de consultas de datos	28
2.3 Ejemplo de la partición de datos	28
2.4 Capas que conforman la arquitectura de información Data Warehouse	32
3.1 Analogía entre arquitectura física y arquitectura de información	40
3.2 Marco conceptual de la administración de datos	50
5.1 Tabla comparativa en el diseño del Data Warehouse, punto 1	109
5.2 Tabla comparativa en el diseño del Data Warehouse, punto 2	110
5.3 Tabla comparativa en el diseño del Data Warehouse, punto 3	111
5.4 Fases de la Guía para la Administración del Proceso de Asimilación de Tecnología de Información en una Organización	114
5.5 Grado de Adopción de la innovación Data Warehouse – Fase I	121
5.6 Grado de Adopción de la innovación Data Warehouse – Fase II	135
5.7 Grado de Adopción de la innovación Data Warehouse – Fase III	151
5.8 Grado de Adopción de la innovación Data Warehouse – Fase IV	165

CAPITULO 1

INTRODUCCION

1.1. DESCRIPCION DE LA TESIS EN BASE AL TEMA INVESTIGADO

Si bien es cierto que las Tecnologías de Información han brindado beneficios muy significativos a las organizaciones, también es cierto que dichas tecnologías no han sido tan efectivas para apoyar los procesos de toma de decisiones más importantes en las empresas. "A pesar de todas las inversiones en hardware, software y redes de comunicación en los sistemas de cómputo, mucha de la información clave que necesitan desesperadamente los que toman decisiones esta todavía inaccesible para ellos" [LOEFFEN,1995].

A nivel operativo y táctico los avances en Tecnologías de Información han sido importantes, se han logrado automatizar diversos procesos administrativos y operativos de gran relevancia para las organizaciones, generando ahorros en costos y aumento en la productividad de los usuarios de tales tecnologías. Sin embargo, para apoyar a la toma de decisiones no es suficiente la automatización de procesos administrativos y operativos, también se necesita que los datos dejen de ser sólo números y se conviertan en información con valor estratégico para los altos administradores de una empresa, valor para apoyar el logro de los objetivos del negocio, y valor para hacer que un negocio sea más competitivo en su industria. "En efecto, los administradores y gerentes han sido entrenados para no solicitar información estratégica a los sistemas de información tradicionales" [BURKAN,1991]. "La función de la alta administración no ha sido afectada por la computadora, los problemas administrativos más importantes no han sido resueltos por la automatización" [DEARDEN,1983]. Aún en la actualidad muchos ejecutivos y administradores de empresas no visualizan un impacto directo de la automatización en sus actividades cotidianas.

¿Por qué no ha llegado la tecnología de información a esos niveles de la organización?, ¿Por qué la información es a veces difícil de obtener rápidamente?. Quizá existan muchas respuestas para estas preguntas, pero para propósito de esta tesis se plantean las siguientes respuestas:

- No son administrados estratégicamente los datos y la información
- No existe alineación de los objetivos del negocio con el desarrollo de la tecnología
- Las tecnologías de información no son totalmente asimiladas y aceptadas entre los miembros de la empresa

Estos aspectos ya han sido analizados y estudiados anteriormente, por lo cual no es de interés en esta tesis estudiarlos como tal, sino más bien, enfocarlos y relacionarlos hacia el diseño de una arquitectura de información que soporte adecuadamente el proceso de toma de decisiones en cualquier nivel organizacional, visualizando en el Data Warehouse, tema principal de esta tesis, a esa arquitectura de información. *"Un Data Warehouse es una colección de todos los tipos de datos, con el propósito de soportar el proceso de toma de decisiones en todos los niveles de una organización; estos datos pueden obtenerse de muchas fuentes en la empresa o de fuentes externas"* [LOEFFEN,1995].

No hay consenso general en cuanto a definir un Data Warehouse, en la definición anterior se observan 2 aspectos que son de interés, primero se menciona a 'todos los tipos de datos', esto es, en todos los formatos y definiciones, y después se puntualiza 'todos los niveles de una organización', lo cual puede incluir el apoyo a las decisiones desde un nivel operativo hasta un nivel estratégico. Se considera entonces que el Data Warehouse implica administración de datos, y si soporta la toma de decisiones en todos los niveles organizacionales, significa que el apoyo llega también a la más alta administración de una empresa, por lo que es necesario que la administración de datos se vuelva estratégica en el diseño del Data Warehouse. Asimismo para fines de esta tesis se visualizará al Data Warehouse como una arquitectura de información y no sólo como una colección de datos o un conjunto de herramientas de software.

Las tecnologías de información son capaces de resolver muchos problemas en una empresa, pero si no existe una alineación entre los objetivos del negocio y el desarrollo tecnológico, las soluciones van a estar mal orientadas. "A través de la tecnología tus problemas no están resueltos, el desarrollo de un Data Warehouse debe estar guiado por la misión y visión del negocio y su avance tecnológico, de tal manera que se explote la tecnología ya existente o se invierta adecuadamente en nueva tecnología" [LADAGA,1995]. De la misión y visión se generan las estrategias de una empresa, si el Data Warehouse es planeado para cumplir con estos aspectos, no debe presentarse ningún contratiempo cuando sea necesaria la generación de información que permita la toma de decisiones de una manera rápida y precisa.

La tecnología puede ser muy poderosa y estar al mismo tiempo alineada a las estrategias de una empresa, pero si los usuarios que van a tener contacto con ella no la asimilan correctamente y no la sienten como parte de sus actividades cotidianas,

los resultados que se van a obtener no van a ser los esperados. Para que una tecnología sea implantada exitosamente requiere ser asimilada y aceptada por los miembros de una organización. Si el objetivo de implantar un Data Warehouse en una organización es apoyar el proceso de toma de decisiones, es necesario entonces tener conocimiento de la tecnología, saber alinearla a los objetivos del negocio, y que sea implantada de acuerdo a un proceso de asimilación tecnológica. Bajo estas premisas es el estudio del Data Warehouse en esta tesis, siendo el análisis principal el proceso de asimilación tecnológica y el grado de adopción que va teniendo la tecnología conforme avanza su proceso de implantación en una organización.

El énfasis en el proceso de toma de decisiones al momento de diseñar el Data Warehouse está fundamentado en la importancia de este proceso en el entorno actual de los negocios, tal como se menciona en las siguientes citas:

"Con el ambiente cambiante de los negocios y la necesidad de tener una respuesta directiva mas rápida y correcta, la toma de decisiones estratégicas ha tomado un nuevo sentido de urgencia" [LOEFFEN,1995]. Además, Peter Drucker comenta de "una organización basada en información y conocimiento, con una organización más plana estructuralmente" [DRUCKER,1988], lo cual se complementa con los siguiente, "el incremento del clima competitivo de los negocios y la eliminación de mandos medios, ha llevado a más y más personas a tomar decisiones" [COLE,1995].

Es por esto el énfasis en el proceso de toma de decisiones y la información requerida para llevarlo a cabo adecuadamente. Ahora bien, "El Data Warehouse por si solo no produce resultados mágicamente, sino que se requieren herramientas que presenten la información que necesitan los usuarios" [RADDING,1995]. Entre estas herramientas se pueden mencionar los Sistemas de Soporte a la Decisión y los Sistemas de Información Ejecutivos, los cuales están relacionados con el Data Warehouse, en esta tesis se estudiará tal relación, mas no a profundidad el concepto en si. "Un Data Warehouse es una arquitectura que madura como un nuevo modelo que esta emergiendo como la base para el diseño de las bases de datos de los Sistemas de Soporte a la Decisión" [MORIARTY,1995].

El concepto Data Warehouse es muy reciente, pero va madurando como menciona la cita anterior, algunas empresas de los Estados Unidos ya lo han implantado con éxito. Anteriormente se tenía la idea de lo que es el Data Warehouse, pero no era aplicable del todo debido a que no se contaba con las herramientas actuales de explotación de base de datos e integración de información. Actualmente ya se cuenta con herramientas computacionales y con nuevas tecnologías de información que facilitan el desarrollo e implantación de un Data Warehouse.

1.2. ALCANCE DE LA INVESTIGACION

¿Por qué el estudio de una tecnología como el Data Warehouse para esta tesis?, ¿Qué alcance tiene esta tecnología actualmente y cuál tendrá en el futuro?. El interés en esta tecnología radica principalmente en su naturaleza estratégica para una organización; en su mayoría las Tecnologías de Información (TI), si son alineadas adecuadamente a los objetivos del negocio, toman también una relevancia estratégica, pero el Data Warehouse, como se verá más adelante en su definición básica, está orientado al negocio por si mismo y además, las herramientas que lo conforman tienen como objetivo principal explotar la información de las bases de datos operacionales para apoyar las decisiones más importantes de una organización. Se podría pensar que la tecnología Data Warehouse representa una moda más en el ambiente informático, y podría caer en eso si se visualiza sólo como una solución tecnológica para ciertas necesidades de información, y no se posiciona como una arquitectura de soporte global en la empresa.

El avance tecnológico en Telecomunicaciones, Hardware y Software, el fortalecimiento en algunas tecnologías como Cliente-Servidor, los lenguajes de cuarta generación, SQL, etc. representan una oportunidad para poder diseñar e implantar un Data Warehouse actualmente, y poder lograr en un futuro que cualquier empresa aplique herramientas Data Warehouse a sus procesos de información. En México se han aplicado o se están desarrollando proyectos Data Warehouse, por lo cual es interesante investigar acerca del mismo, la manera de conformar toda la arquitectura de información y de lograr que sea asimilada y aceptada en una organización.

1.3. OBJETIVO DE LA TESIS

Por todo lo anterior, **el objetivo de esta tesis** es estudiar el concepto Data Warehouse, su visualización como una arquitectura de información, su aplicabilidad en las organizaciones, su soporte al proceso de toma de decisiones, y finalmente concluir con una propuesta de análisis que permita su implantación en una empresa mediante un proceso de asimilación tecnológica.

1.4. RESULTADOS ESPERADOS

- 1.- Impulsar la creación de una nueva arquitectura de información que ayude a mejorar la calidad y el manejo de los datos en una empresa.
- 2.- Mostrar los beneficios de un Data Warehouse y principalmente su contribución al proceso de toma de decisiones.
- 3.- Aumentar el conocimiento de la tecnología Data Warehouse y su proceso de implantación.
- 4.- Analizar la asimilación de una tecnología como el Data Warehouse, sus implicaciones organizacionales y su alineación a los objetivos de una empresa.

1.5. DESCRIPCIÓN DE LA TESIS POR CAPÍTULOS

En este capítulo 1 se describe la importancia de la toma de decisiones en las empresas actuales, y sobre todo en aquellas en donde el entorno es muy cambiante y dinámico. Asimismo, se reconoce el valor de la información como parte esencial del proceso de toma de decisiones, así como las tecnologías que pueden hacer que la información aumente la productividad y la calidad de las decisiones no sólo en los niveles operativos, sino también a los niveles estratégicos de las empresas. Asimismo se describe brevemente el tema principal de esta tesis, que es el Data Warehouse, el enfoque que se llevará para estudiarlo, su alcance, los objetivos que se pretenden alcanzar y los resultados esperados de la investigación.

En el capítulo 2 se presenta la evolución del procesamiento de datos, la cual ha provocado un crecimiento sin control en el manejo de la información. Para administrar tal crecimiento se presenta una nueva manera de visualizar a los datos mediante el concepto Data Warehouse, identificando y entendiendo sus procesos y conceptos básicos, para así conformar una arquitectura de información que apoye al proceso de toma de decisiones. Por último se presenta la definición del Data Warehouse para fines de esta tesis, y se propone un esquema para visualizar a la arquitectura de información Data Warehouse relacionada con la pirámide organizacional.

En el capítulo 3 se describe cada componente de la arquitectura de información propuesta, haciendo énfasis en la administración de los datos de una manera estratégica, y en las herramientas de software que hacen posible realizar cada proceso del Data Warehouse; asimismo se analizan los sistemas estratégicos que presentan la información que se genera del Data Warehouse, y que finalmente cubre las necesidades de los usuarios finales.

En el capítulo 4 se presentan diversas metodologías o herramientas de análisis que soportan el estudio de la empresa donde se va a implantar un Data Warehouse, enfocando el análisis en conocer el flujo de la información a través de la organización, y así facilitar el diseño de las estrategias de los diversos procesos del Data Warehouse. Se presenta también un modelo de alineación entre los objetivos del negocio y el desarrollo de la tecnología Data Warehouse, y se establecen las bases del proceso de asimilación de tecnología presentada en el capítulo 5.

En el capítulo 5 se presenta el proceso de asimilación de la arquitectura de información Data Warehouse en una empresa, estableciendo previamente los atributos que influyen en el proceso y las medidas con las que se va a identificar el grado de adopción de la tecnología conforme avanza su desarrollo e implantación. Asimismo se presentan propuestas de diversos autores para el desarrollo del Data Warehouse, con el fin de realizar un análisis comparativo y así entender cada etapa de las propuestas, de tal manera que pueda relacionarse adecuadamente el desarrollo del Data Warehouse con su proceso de asimilación.

Finalmente, en el capítulo 6 se mencionan las conclusiones de la investigación realizada y se especifican áreas de oportunidad para profundizar en el estudio del Data Warehouse y los componentes y procesos que conforman la arquitectura de información.

1.6. DEFINICIONES, ABREVIATURAS Y NOTAS IMPORTANTES

TECNOLOGÍA DE INFORMACIÓN (TI): Término aceptado para representar el rápido crecimiento de equipos (computadoras, dispositivos de almacenamiento de datos, dispositivos de redes y comunicaciones), aplicaciones (sistemas punto de venta, cajeros automáticos, sesiones educativas por multimedia, CAD -computer aided design-, CAM -computer aided manufacturing-, sistemas expertos, robótica, sistemas de reservaciones, etc.) y servicios (herramientas para usuarios finales,

desarrollo de aplicaciones rápidas) usados por las organizaciones para entregar datos, información y conocimiento. [LUFTMAN,1993]

SISTEMAS DE INFORMACION (SI)

SISTEMA DE INFORMACION EJECUTIVOS (EIS)

SISTEMA DE SOPORTE A LA DECISION (DSS)

DATA WAREHOUSE (DW)

MANEJADOR DE BASE DE DATOS (DBMS)

* Algunas veces se menciona al Data Warehouse como arquitectura de información, otras como tecnología, y en otras ocasiones como proyecto, es válida cualquiera de estas definiciones.

* En algunas ocasiones se manejan de manera semejante los términos datos e información, sin embargo en el apartado 2.2 se menciona la diferencia entre ambos conceptos

REFERENCIAS BIBLIOGRAFICAS

- [BURKAN,1991] BURKAN, Wayne C.
Executive Information Systems from proposal through implementation.
Van Nostrand Reinhold, 1991
- [COLE,1995] COLE, Barb.
"Data Warehouse demands straining corporate nets".
Networkworld, vol. 12, num. 35. Agosto 28, 1995.
- [DEARDEN,1983] DEARDEN, John.
"SMR Forum: Will the computer change the job of top management".
Sloan Management Review, vol. 25, num. 1. Otoño 1983
- [DRUCKER,1988] DRUCKER, Peter.
"The coming of the New Organization".
Harvard Business Review, Ene-Feb 1988.
- [LOEFFEN,1995] LOEFFEN, Dorothy.
"The Data Warehouse, a new mission critical resource".
CMA Magazine, Noviembre 1995.
- [LADAGA,1995] LADAGA, John.
"Let business goals drive your data warehouse effort".
Health Management Technology, Octubre 1995.
- [LUFTMAN,1993] LUFTMAN J. N., P. R. Lewis, S. H. Oldach.
"Transforming the enterprise: The alignment of business and information technology strategies".
IBM Systems Journal. vol 32. no. 1, 1993
- [MORIARTY,1995] MORIARTY, Terry.
"Modeling Data Warehouses".
Database Programming and Design. Agosto 1995.
- [RADDING,1995] RADDING, Alan.
"Support decision makers with a Data Warehouse".
Datamation. Marzo 15, 1995.

CAPITULO 2

FUNDAMENTOS TEORICOS DEL DATA WAREHOUSE

2.1 INTRODUCCION

El término Data Warehouse va más allá de ser un simple "almacén de datos" como lo indica su nombre, más bien representa "una nueva arquitectura de administración de datos que permite concentrarlos con el objetivo de ofrecerlos al recurso humano de manera más fácil y extensa" [WALLACE,1994]. Ahora bien esta concentración de datos no se refiere específicamente a agruparlos en un solo lugar físico, como el caso de un servidor en una red de área local (LAN), sino más bien a estructurarlos y relacionarlos de acuerdo a un flujo de información y a un modelo de datos global de una empresa.

El deseo de muchas empresas de tener acceso rápido y flexible a la información viene de varios años atrás, ambientes en donde cada usuario pueda consultar y obtener reportes de manera dinámica y abierta, y no estática y cerrada como se acostumbra en los sistemas transaccionales que actualmente existen en las organizaciones. Es común ver que los usuarios de los sistemas transaccionales soliciten al departamento de Sistemas de Información o Informática reportes con determinada información, la cual puede ser una combinación de datos de diversos sistemas e incluso de datos que se procesan en diferentes áreas funcionales; al solicitar algún reporte con estas características se mantiene una dependencia muy fuerte con el departamento de informática, debido a que los sistemas operacionales que procesan transacciones no fueron diseñados para generar reportes específicos a una necesidad de un usuario, provocando con esto retardos al momento de crear este tipo de reportes. Además de los problemas tecnológicos, existen prácticas organizacionales que dificultan el acceso rápido y flexible a la información, el sentido de pertenencia de los datos es una de ellas, esto es, que ciertos datos pertenezcan a tales áreas o incluso, peor aún, a ciertos usuarios, y que por esta razón no sean compartidos a través de la empresa. La falta de compartición en los datos provoca que los datos no puedan ser concentrados para satisfacer las necesidades de información a un nivel corporativo.

El Data Warehouse implica administrar datos, como lo menciona la definición, y para poder administrarlos se requiere conocer la situación de los mismos dentro de

las organizaciones. Para determinar la situación de los datos se necesita analizar tanto a los sistemas transaccionales que realizan su procesamiento, como también las prácticas organizacionales que determinan su grado de uso e importancia en el negocio. El procesamiento de los datos ha evolucionado históricamente, de esta evolución han surgido aspectos positivos y negativos en el manejo de los datos, lo cual ha forzado la creación de nuevas arquitecturas de información como el Data Warehouse.

2.2. EVOLUCIÓN HISTÓRICA DEL PROCESAMIENTO DE DATOS

La figura 2.1 muestra una reseña histórica de la evolución del procesamiento de los datos, con la cual podemos analizar su crecimiento y los problemas surgidos como consecuencia de una falta de visión y control en dicha evolución, siendo esta una de las causas que motivaron a buscar un cambio en la administración de los datos en las organizaciones.

En la década de los 60's con el boom comercial de las computadoras de procesamiento de datos se crearon archivos maestros en diversas áreas funcionales dentro de la empresa. Al principio era fácil el uso y el mantenimiento de estos archivos, pero a medida de que iban creciendo en número y en cantidad de datos a procesar se fue haciendo más complicada su administración, lo cual provocó entre otras cosas redundancia en los datos y problemas como los siguientes:

- “- La necesidad de sincronizar adecuadamente los datos al momento de realizar alguna actualización.
 - La complejidad en el mantenimiento de los programas
 - La complejidad en el desarrollo de nuevos programas
 - La cantidad de hardware requerido para soportar todos los archivos maestros “
- [INMON, 1996]

En la década de los 70's surge una nueva tecnología de almacenamiento y acceso a datos que son los discos de almacenamiento secundario, mejor conocidos como discos duros y discos flexibles. Esta tecnología permite el acceso directo a los datos con la ayuda de los programas administradores de base de datos (DBMS), siendo ambos impulsores del procesamiento de transacciones conocido hoy en día. Con esta tecnología los datos fueron agrupados en fuentes más sencillas (Base de Datos), pudiendo así solucionar algunos de los problemas mencionados anteriormente. El procesamiento de transacciones en línea (OLTP) surgido en esa

década, le vino a dar otro enfoque a la administración de los datos al fortalecerse el uso del concepto de Base de Datos y la proliferación de terminales y sistemas transaccionales que generalmente se ejecutaban en una base de datos centralizada en equipo mayor, con atributos multiusuario. Algunos de estos sistemas en línea son: automatización de la nómina, control de recursos financieros, sistemas de reservaciones, sistemas punto de venta, sistemas para cajeros automáticos, control de inventarios, etc.

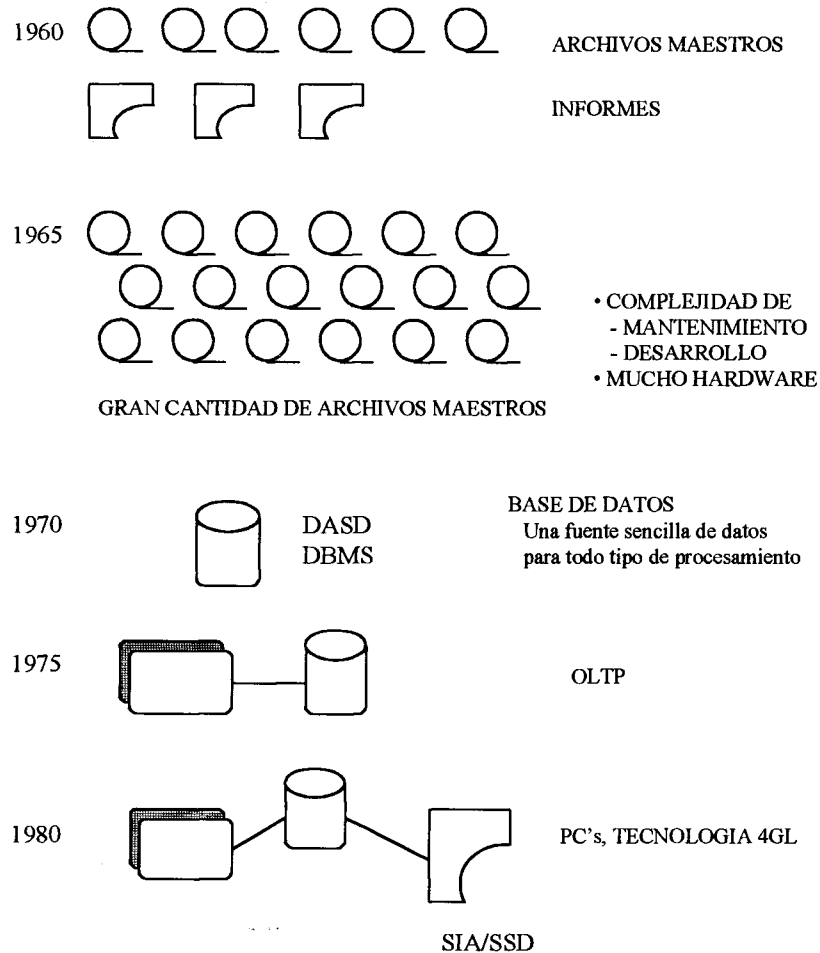


Figura 2.1 - Evolución del Procesamiento de Datos [INMON, 1996]

Los sistemas de transacciones manejan grandes cantidades de datos, permiten accesos rápidos a los registros y generan reportes para soporte operativo en una área funcional. Algunos de estos sistemas se volvieron estratégicos en las empresas y de su buen funcionamiento dependían sus operaciones básicas. Por ejemplo, se tienen

los Sistemas Punto de Venta (POS), estratégicos para el sector comercial, principalmente en tiendas grandes de autoservicio; recientemente estos sistemas ya no dan ventajas competitivas a tales empresas, sino más bien representan una necesidad actual para mantenerse posicionado en los mercados de venta. Lo mismo sucedió con los sistemas de cajeros automáticos, los cuales proporcionaron una ventaja competitiva a los primeros bancos que los aplicaron, más sin embargo ahora es una necesidad tecnológica que debe funcionar adecuadamente para atender eficientemente a los clientes. Algunos otros sistemas como el control de nómina fueron perdiendo importancia estratégica en algunas organizaciones, a tal grado que en muchas de ellas utilizan Outsourcing (servicio externo) para llevarla a cabo.

Para la década de los 80's, nuevas tecnologías -tales como las PC's y los 4GL (lenguajes de cuarta generación)- empezaron a surgir. El usuario final comenzó a asumir un nuevo papel, que es el de controlar los datos y los sistemas; comenzaron a desarrollarse Sistemas de Información administrativos (MIS), Sistemas de Soporte a la toma de decisiones (DSS), con el objetivo de procesar información para soportar decisiones administrativas. Estas tecnologías y sistemas no han sido explotadas eficientemente debido a la evolución natural del procesamiento de datos que ha llevado a muchas empresas a tener una "telaraña de información" como la que se presenta en la figura 2.2. Esta figura nos presenta de manera simbólica el crecimiento sin control de las bases de datos en una empresa. Por una parte se tiene una base de datos centralizada la cual es accesada por muchos usuarios, pero a la vez se han generado diversas bases de datos independientes de menor tamaño para satisfacer necesidades de información de un departamento o área funcional; al no existir un control centralizado en la creación de estas bases de datos provoca problemas como redundancia, inconsistencia y falta de integridad y sincronización en los datos.

"Muchos problemas se han presentado con el crecimiento de esta 'telaraña de información' en las empresas, detectándose principalmente los siguientes:

- a) Falta de credibilidad en los datos
- b) Baja productividad
- c) Falta de capacidad para transformar datos a información" [INMON,1996]

a) Falta de credibilidad en los datos

El crecimiento sin control de los sistemas de información y de las bases de datos independientes ha generado problemas como la redundancia, integridad e inconsistencia en los datos, y esto se refleja cuando la información presentada a la alta administración en forma de reportes por 2 departamentos, no coincide en algunos puntos que pertenecen al mismo proceso de información y del cual deben

surgir los mismos datos. Tomemos el ejemplo de 2 departamentos: Finanzas y capacitación; el reporte generado por el área de Capacitación muestra un listado de los cursos impartidos en software específicos a todos los departamentos de la empresa, entre estos cursos se incluye la capacitación en el software X al área de Finanzas; más sin embargo, el departamento de Finanzas genera un reporte de los cursos que recibió en el año, en el cual no menciona la capacitación en dicho software X; puede parecer un ejemplo muy trivial, pero es representativo de lo que puede suceder al tener bases de datos independientes y sin un control centralizado. Los directivos de la empresa al recibir estos reportes no saben que pensar debido a que éstos fueron generados por un sistema de información en computadora, perdiendo así la credibilidad en los datos que tales sistemas pueden proporcionar. La administración se enfrenta entonces con la necesidad de tomar decisiones basada en políticas y en puestos determinados, y no tanto en la información que tienen en sus manos. Este es un punto crítico en la falta de credibilidad derivada del crecimiento sin control de los sistemas de información y de las bases de datos.

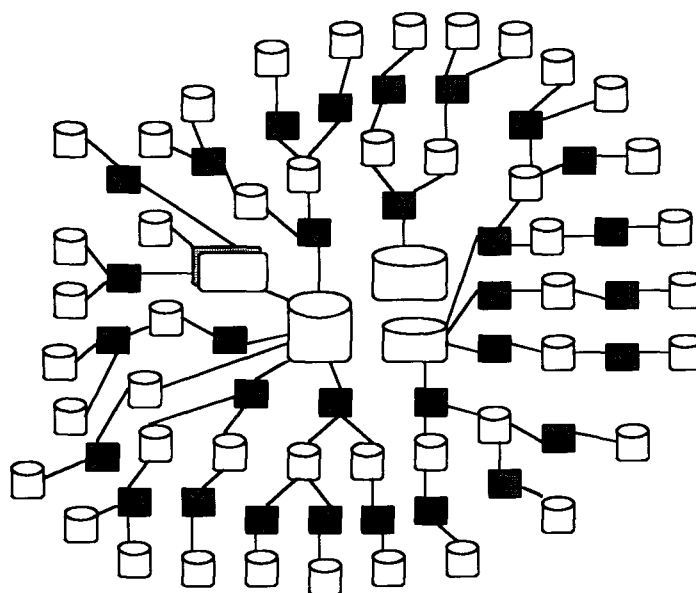


Figura 2.2 - Evolución natural de los datos (Telaraña de información) [INMON,1996]

b) Productividad

La “telaraña de información” también ha provocado una baja productividad en la generación de reportes corporativos. “La administración desea producir un reporte corporativo, usando muchos archivos y colecciones de datos que han sido acumulados en el transcurso de los años. El personal al cual se le asigna la tarea considera que hay 3 cosas que deben llevarse a cabo para producir el reporte:

- Localizar y analizar los datos para el reporte.
- Compilar los datos para el reporte.
- Obtener los recursos informáticos” [INMON,1996]

Esta descripción de Bill Inmon demuestra lo ineficaz que puede ser el producir un reporte corporativo de carácter histórico en una empresa, al requerirse un esfuerzo humano en tiempo y un alto costo por la disponibilidad de los recursos informáticos. La importancia de un reporte corporativo con datos históricos la podemos visualizar desde el punto de vista de la visión que se plantea una organización para los próximos años, lo cual depende de su posicionamiento actual dentro de su industria y de sus logros obtenidos en el pasado; para realizar un plan de negocios que contemple los próximos 5 años se necesita conocer las tendencias y el comportamiento como organización en los años pasados, esto da la pauta para establecer estrategias y objetivos futuros. Y que mejor que los datos históricos de una empresa para mostrar y analizar esas tendencias en el pasado, por lo cual se hace necesario la creación de dichos reportes corporativos de carácter histórico.

Para ejemplificar la baja productividad de los datos se considera el caso de una empresa de manufactura, la cual con nueve plantas que la conforman no puede realizar una negociación favorable en los contratos de compra con sus proveedores más importantes, debido a la dificultad ocasionada por la inconsistencia de los datos, lo cual provoca que las plantas no puedan determinar exactamente los materiales y productos que reciben de cada proveedor.

En caso del diseño de un Data Warehouse la productividad se reflejaría en la calidad de los datos y en la información que se ofrecería a los usuarios de dicha arquitectura. “Son dos factores que pueden impedir entregar datos de calidad que son requeridos por los usuarios de un Data Warehouse: la disparidad de los datos mantenidos por las bases de datos de los sistemas de información actuales (“telaraña de información”), y por otro lado las políticas del negocio hacia la administración de los datos” [MORIARTY,1996]. En el caso de la disparidad de los datos es importante que se busquen estrategias para solucionar o disminuir este problema si se desea implantar un Data Warehouse, ya que uno de los puntos que se consideran de mayor importancia en el diseño de esta arquitectura de información es la integridad de los datos que la organización pueda tener. “El nivel de integración de los datos que una organización puede llevar a cabo y el esfuerzo requerido para alimentar las bases de datos de un Data Warehouse con datos de calidad, depende en gran medida de la disparidad de los datos que la empresa tenga” [MORIARTY,1996].

c) Falta de capacidad para transformar datos en información

Se ha visto que la evolución natural de los datos en una empresa ha provocado problemas que impiden la creación de reportes ejecutivos con calidad, aumentando con esto la incredulidad de los altos ejecutivos hacia los datos. Si se requiere que los datos sean considerados como otro activo mas dentro de una empresa, es importante que estén disponibles y de manera confiable en el momento que se requieran, siendo necesario eliminar esa incapacidad de transformar los datos en información. Primeramente es necesario establecer la diferencia entre datos e información; aquí se presenta la conceptualización por parte de Mc. Fadden y Fred R. en su libro "Data Base Management" [MCFADDEN,1994].

DATOS: Factores que conciernen a cosas como la gente, los objetos o los eventos. Por ejemplo un listado de estudiantes mostrando su matrícula, nombre, carrera, calificaciones, etc.

INFORMACION: Datos que han sido procesados y presentados en forma adecuada para la interpretación humana, frecuentemente con el propósito de revelar tendencias o patrones de comportamiento. En el caso de los estudiantes, podrían ser analizados y sumariados a través del uso de herramientas estadísticas. Los administradores y demás personas en una organización emplean tal información para soportar la toma de decisiones, en este caso la información de los estudiantes podría utilizarse para decidir si introducir nuevas clases o en su defecto contratar a más personal docente. La figura 2.3 nos muestra la diferencia entre lo que son los datos y lo que es la información.

De acuerdo con Bill Inmon existen 2 razones principales que impiden el convertir los datos en información [INMON,1996]:

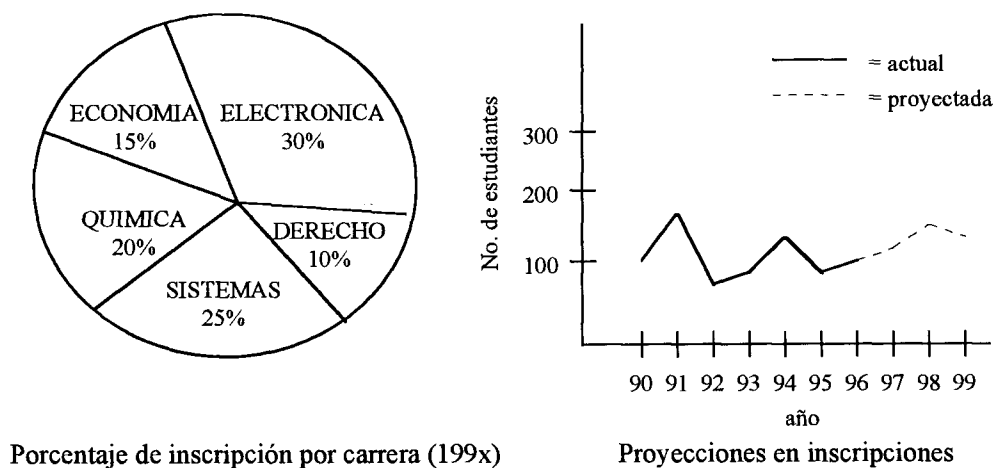
- 1.- La falta de integración en los datos entre las diversas aplicaciones relacionadas con la información que se desea obtener. Por ejemplo: en el caso de una institución educativa, para conocer la cantidad de maestros que han sido asignados a determinadas materias y la cantidad de alumnos asistentes y de que carreras, es necesario que existan relación e integridad en bases de datos como la de Alumnos, Maestros y Materias.
- 2.- Quizá algunas aplicaciones existentes simplemente no tienen la cantidad de datos históricos necesarios para convertir datos en información. Supongamos el caso de una importante tienda de autoservicio donde los altos ejecutivos deseen conocer los patrones de compra por mes de los clientes en los últimos 5 años, considerando aquellas compras motivadas por alguna oferta temporal. La base de datos de ventas dispone de datos de los últimos 5 años, pero quizá no se ha

llevado una política de partición de datos basados en periodos del tiempo; por otro lado las bases de datos de precios y ofertas solo tienen un registro histórico de 2 años, y la base de datos de clientes solo tiene datos del último año. Considerando este ejemplo sería muy difícil elaborar el reporte requerido por los altos ejecutivos, porque nunca se consideró el diseño de las bases de datos de manera histórica para satisfacer requerimientos de información con un enfoque corporativo.

LISTA DE ESTUDIANTES

MATERIA: Matemáticas III		SEMESTRE: Ago-Dic 1996		SALON: 2A	
NOMBRE	MATRICULA	CARRERA	CALIFICACION		
Luis Méndez	763504	ECONOMIA	85		
Carkis Pérez	510951	QUIMICA	90		
Verónica Ruiz	761345	SISTEMAS COMP.	81		
:	:	:	:		

(a) Datos



(b) Información

Figura 2.3 - Comparación de datos e información [MCFADDEN, 1994]

El proceso de convertir los datos en información basado en los requerimientos de información de los que toman decisiones es un objetivo que se persigue al diseñar un Data Warehouse, la conversión de los datos a información puede lograrse con

una adecuada administración de datos, con herramientas de software que realicen los procesos que sean requeridos, y con el uso de sistemas de información que exploten los datos y permitan consultas rápidas y eficientes. Mas que una sola tecnología, se requiere toda una arquitectura de información que tenga la capacidad de controlar y hacer más eficiente el flujo de información en una empresa.

2.3. LA NUEVA PROPUESTA EN BASE DE DATOS

Se ha visto que la evolución natural de los datos y de los sistemas de información ha sido consecuencia de la urgencia para satisfacer determinadas necesidades de información de las diversas áreas funcionales de una empresa; si bien se ha logrado cumplir con tal objetivo auxiliado por los Manejadores de Base de Datos, por el procesamiento de transacciones en líneas, por el uso de diversos lenguajes de programación, entre otros aspectos, las necesidades ahora están cambiando y los usuarios son otros, se requiere tener información corporativa con un tiempo de respuesta bajo para responder rápidamente a los cambios organizacionales; las empresas se están dando cuenta que a través del uso de las tecnologías de información se puede lograr flexibilidad organizacional.

Un Data Warehouse bien diseñado e implantado podría ser la respuesta a ese cambio organizacional y tecnológico en muchas empresas. Se requiere un fuerte esfuerzo para lograrlo, sobre todo si las plataformas tecnológicas en la organización no son las más adecuadas para implantar una tecnología como el Data Warehouse. “Aquéllos factores que eran considerados como positivos al desarrollar las aplicaciones operacionales, son ahora las barreras principales para satisfacer los requerimientos de integración de datos en el almacenamiento de los datos operacionales y en el ambiente Data Warehouse” [MORIARTY,1996]). Estas barreras surgen como consecuencia de la “telaraña de información” explicada anteriormente.

Lo anterior nos lleva a reconocer otro tipo de base de datos diferentes a las base de datos operacionales que actualmente se conocen (ver Tabla 2.1), lo cual viene a representar la nueva propuesta en base de datos en el ambiente Data Warehouse. Las bases de datos son el corazón del Data Warehouse, es importante conocer las características más importantes de su diseño para poder entender el concepto del Data Warehouse; hay que tener en cuenta que se van a satisfacer necesidades de información de los más altos niveles organizacionales, por lo cual el

diseño de las bases de datos en el Data Warehouse va a ser muy diferente al usado con las bases de datos operacionales.

BASE DE DATOS OPERACIONALES VS. DATA WAREHOUSE	
Base de Datos operacionales	Data Warehouse
• Orientada a la aplicación	• Orientada al tópico
• Detallado	• Sumarizado, datos calculados
• Preciso, en el momento del acceso	• Representa valores en el tiempo
• Puede ser actualizado campo por campo	• Accesado y manipulado: no actualizado directamente
• Procesamiento repetitivo	• Procesamiento heurístico y analítico
• Manejo de transacciones	• Manejo de análisis
• Alta disponibilidad	• Disponibilidad moderada a baja
• Estructura estática	• Estructura flexible
• Soporta operaciones diarias	• Soporta necesidades administrativas
• Pocos datos usados en un proceso	• Gran cantidad de datos usados en un proceso
• Tiempo de respuesta de 2 a 3 segundos	• Tiempo de respuesta de varios segundos a minutos

Tabla 2.1. Diferencias entre las bases de datos operacionales y las bases de datos de información o Data Warehouse [FRANCET,1994] [INMON,1996]

2.4. CONCEPTO DATA WAREHOUSE

William Inmon, creador en 1990 del concepto Data Warehouse, lo define como "una colección de datos que soportan las decisiones administrativas, y que están:

- Orientados al negocio
- Integrados
- Variantes en el tiempo
- No volátiles " [HACKATHORN,1995]

2.4.1. Datos orientados al negocio

Cuando se habla de orientados al negocio se refiere al hecho de visualizar a los datos en el entorno mismo de la empresa, en cuanto a sus procesos y áreas

funcionales, más que identificarlos como bytes o bits almacenados. "Estructurados de manera adecuada, los datos orientados al negocio proporcionan una imagen estable de los procesos del negocio, independientemente de los sistemas transaccionales" [HACKATHORN,1995]. Este punto es muy importante cuando se trata de posicionar estratégicamente a la tecnología Data Warehouse en una organización, ya que al tener orientación al negocio permite enfocar su diseño hacia los procesos que son importantes para lograr los objetivos y las estrategias organizacionales.

2.4.2. Datos integrados

Al hablar de datos integrados se refiere al hecho de tener en el Data Warehouse datos de diversos sistemas transaccionales sin importar su formato o estructura de origen, e independientemente de las aplicaciones y la plataforma de hardware. "Un Data Warehouse representa una fuente de datos de naturaleza histórica, integrada y compartida que es diseñada para soportar el proceso de toma de decisiones" [CHASIN,1994]. Uno de los problemas principales que se mencionaron como consecuencia de la evolución natural de los datos (telaraña de información), es la falta de integración de las bases de datos independientes, se mencionó que esto ocasiona dificultades al intentar crear reportes corporativos con carácter histórico. Entre los aspectos básicos de un Data Warehouse, la integración es el más importante de todos, ver figura 2.4.

2.4.3. Datos que varían en el tiempo

Cuando se refiere a datos que varían en el tiempo, se puede explicar directamente con la misma naturaleza histórica del Data Warehouse, "es claro que una de las principales razones para construir un Data Warehouse es separar lo operacional, que son datos transaccionales actualizándose constantemente, de lo histórico, los cuales son datos estáticos requeridos para el análisis de un negocio" [FRANCETT,1995]. Con este enfoque cada dato tiene un valor en el tiempo y puede ser empleado para generar información que es factible a analizarse con herramientas de software, como los sistemas de soporte a la decisión. El alcance histórico de los datos que se analizan en un Data Warehouse es mayor que el que se maneja en las bases de datos operacionales, por lo que el objetivo es poder tener bases de datos históricas que permitan satisfacer las necesidades de información al momento de crear reportes corporativos que requieran datos de varios años atrás, ver figura 2.5.

2.4.4. Datos no volátiles

Con datos no volátiles, Inmon se refiere a la alimentación de datos al Data Warehouse sin reemplazar a los ya existentes, sino más bien la base de datos absorbe nuevos datos, integrándolos con los ya almacenados. La actualización de datos tal como se lleva a cabo en las bases de datos operacionales no se presenta en la tecnología Data Warehouse, aquí la alimentación de la base de datos se realiza en masa, en determinados periodos de tiempo. La integración mencionada en este punto nos da una idea de la importancia de las herramientas de integración de datos para la creación eficiente de un Data Warehouse, ya que además de integrar los datos que provienen de diversas fuentes, se debe analizar la integración a datos históricos ya existentes, ver figura 2.6.

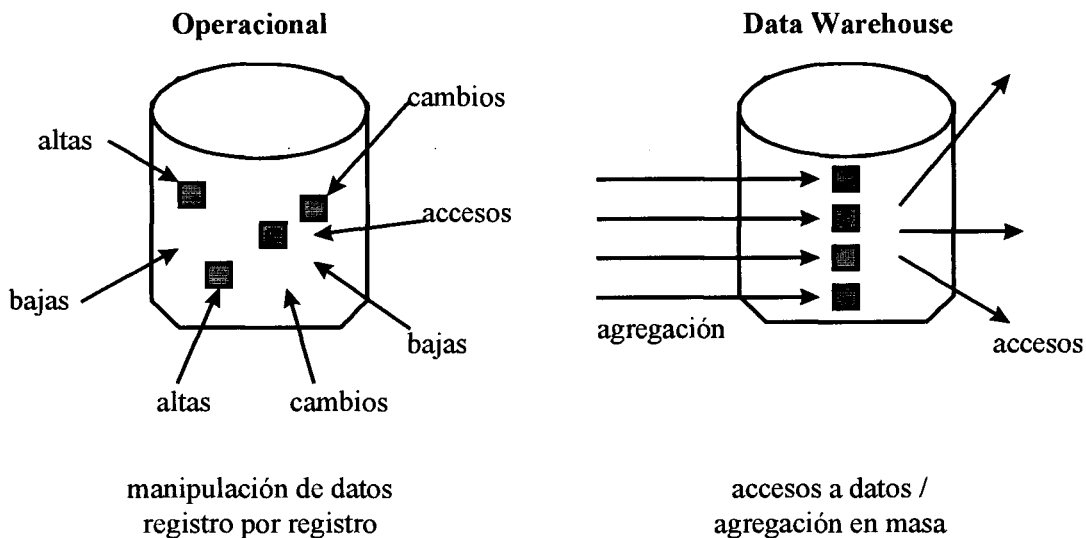


Figura 2.6 - Datos no volátiles [INMON,1996]

2.5. ESTRUCTURA DE UN DATA WAREHOUSE

Se podría pensar que el Data Warehouse viene a ser un manejador de base de datos más, pero con otro nombre; esta arquitectura de datos está fuertemente fundamentada en los conceptos de base de datos y sus sistemas de administración, pero esto es solo una parte de lo que se necesita para crear un Data Warehouse. "La bases de datos son el corazón de cualquier Data Warehouse, y realizar un buen diseño de ellas es tan importante como cuando se lleva a cabo en los sistemas de

procesos transaccionales, además de que son requeridas las mismas habilidades" [LAZAR,1996].

Es importante el diseño de las bases de datos en el Data Warehouse, pero si esto fuera todo se estaría hablando de un sistema de administración clásico de base de datos, y en este caso el enfoque de esta tesis es analizar y presentar al Data Warehouse como una arquitectura de información que apoya al proceso de toma de decisiones. "Mas allá de la base de datos del Data Warehouse, existen otros productos requeridos para tener una infraestructura completa para el soporte a la toma de decisiones" [LAZAR,1996]. La figura 2.7 muestra la estructura de la arquitectura de información Data Warehouse, en donde se visualizan diversos niveles de almacenamiento de datos. Los datos se almacenan en diferentes niveles de detalle conforme al nivel de sumariación requeridos por los usuarios finales.

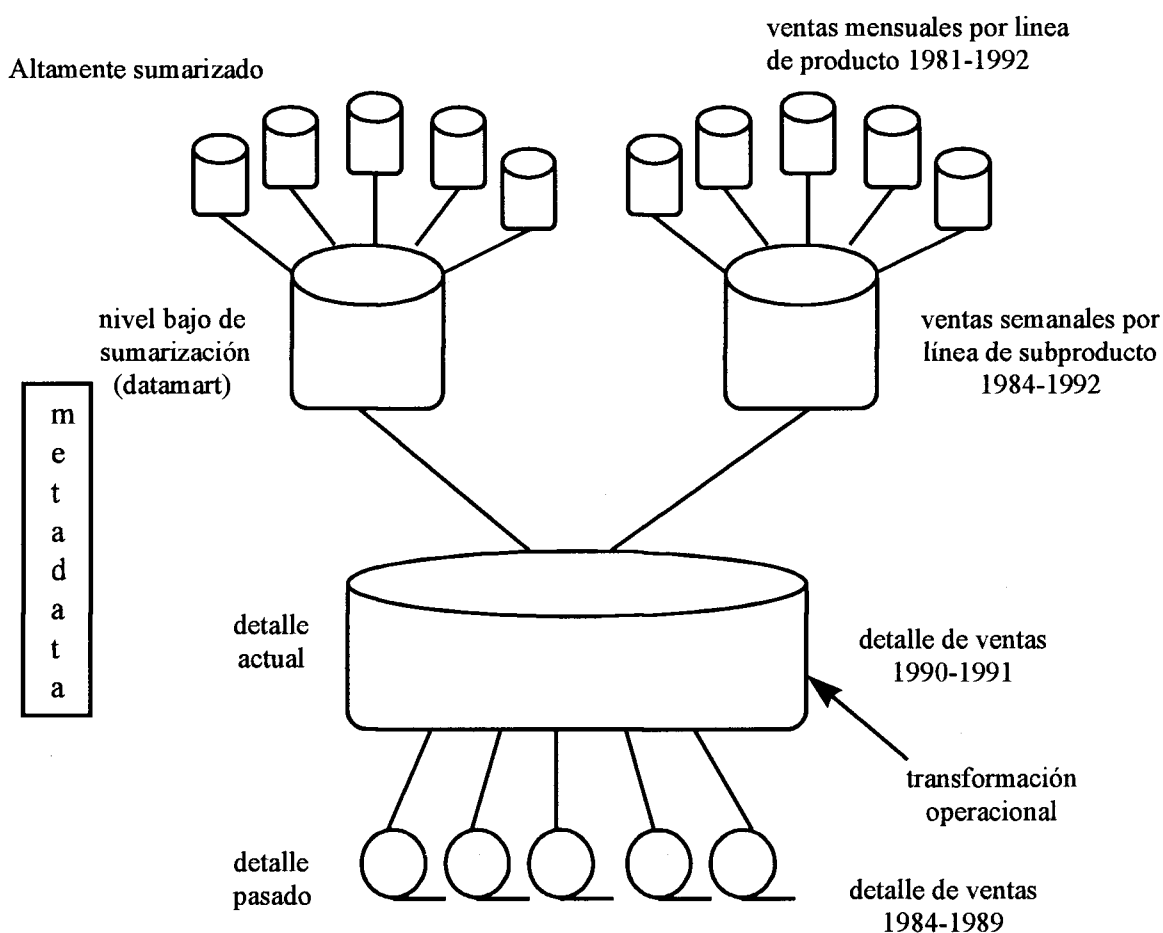


Figura 2.7 - Estructura de un Data Warehouse [INMON,1996]

En el primer nivel correspondiente al detalle pesado, los datos se encuentran generalmente en unidades de almacenamiento secundario, como las cintas magnéticas; los datos son almacenados con un alto nivel de detalle y de diversos años atrás.

El segundo nivel correspondiente al detalle actual, representa las bases de datos operacionales en movimiento, en esta capa el nivel de detalle es igual al guardado en las unidades de almacenamiento secundario, la diferencia está en que estas bases de datos se están actualizando continuamente y soportan lo que es el procesamiento de transacciones en línea. En este nivel de detalle el flujo de la información es hacia arriba y hacia abajo, hacia arriba significa que los datos son sumarizados para alimentar y actualizar bases de datos con un menor nivel de detalle, y es hacia abajo cuando se cumple un determinado período de tiempo donde ya no es necesario mantener dichos datos en las bases de datos operacionales, pasando así a las unidades de almacenamiento secundario.

El tercer nivel de detalle representa los Datamarts. “Un Datamart es una implantación de una estrategia de sumarización y partición de datos que mapea los requerimientos de los que toman decisiones con el fin de optimizar las consultas que ellos más podrían necesitar” [MEREDITH,1996]. En este nivel de detalle los datos están ligeramente sumarizados y generalmente con particiones de datos de acuerdo a ciertos períodos de tiempo, o a través de otros criterios de partición que más adelante se mencionarán; por ejemplo si el período de datos es semanal como lo muestra la figura, es necesario la sumarización semanal de los datos operacionales para alimentar a los Datamarts correspondientes.

Después de los Datamarts viene el nivel más alto de sumarización, en donde el detalle es mínimo y generalmente representa la información que es necesario almacenar en las bases de datos de los Sistemas de Información Ejecutivos o de los Sistemas de Soporte a la toma de decisiones. En este nivel de sumarización la información se actualiza conforme a las necesidades del usuario final, generalmente representado por un tomador de decisiones; asimismo la manera en que se hayan particionado los datos en los niveles anteriores, influye en los procesos de actualización para conformar este nivel de sumarización.

El nivel Data Warehouse comienza cuando los datos operacionales empiezan a particionarse y sumarizarse mediante diversas estrategias de agregación de datos para formar otros niveles de detalle. “Los datos fluyen al Data Warehouse desde el nivel operacional, generalmente ocurre una cantidad fuerte de transformación de datos del nivel operacional al nivel Data Warehouse” [INMON,1996]. En la estructura de la figura 2.7 se muestran solamente dos niveles de sumarización, pero en el diseño del Data Warehouse pueden formarse más niveles dependiendo de las

necesidades de información de la empresa. La estructura organizacional puede influir en el número de niveles de sumariación de los datos, cada nivel jerárquico de dicha estructura puede representar un nivel de sumariación si la importancia de la información que se procesa en ese nivel así lo requiere.

2.6. SUMARIZACIÓN O AGREGACIÓN DE DATOS

“Para manejar y explotar los grandes volúmenes de datos generados por los sistemas de información actuales, los diseñadores del Data Warehouse deben conocer y entender nuevas técnicas de sumariación y partición de los datos” [MEREDITH,1996]. En la estructura del Data Warehouse se mencionó que los Data Marts son datos sumariados de las bases de datos operacionales de una empresa, pero para lograr tal sumariación se requiere un proceso de agregación de datos eficiente y adecuado al diseño del Datamart correspondiente. “La agregación es una sumariación precalculada con el fin de tener un almacenamiento de datos por medio de tablas diversas, para mejorar el rendimiento de las consultas de los usuarios finales” [MADSEN,1996], ver figura 2.8.

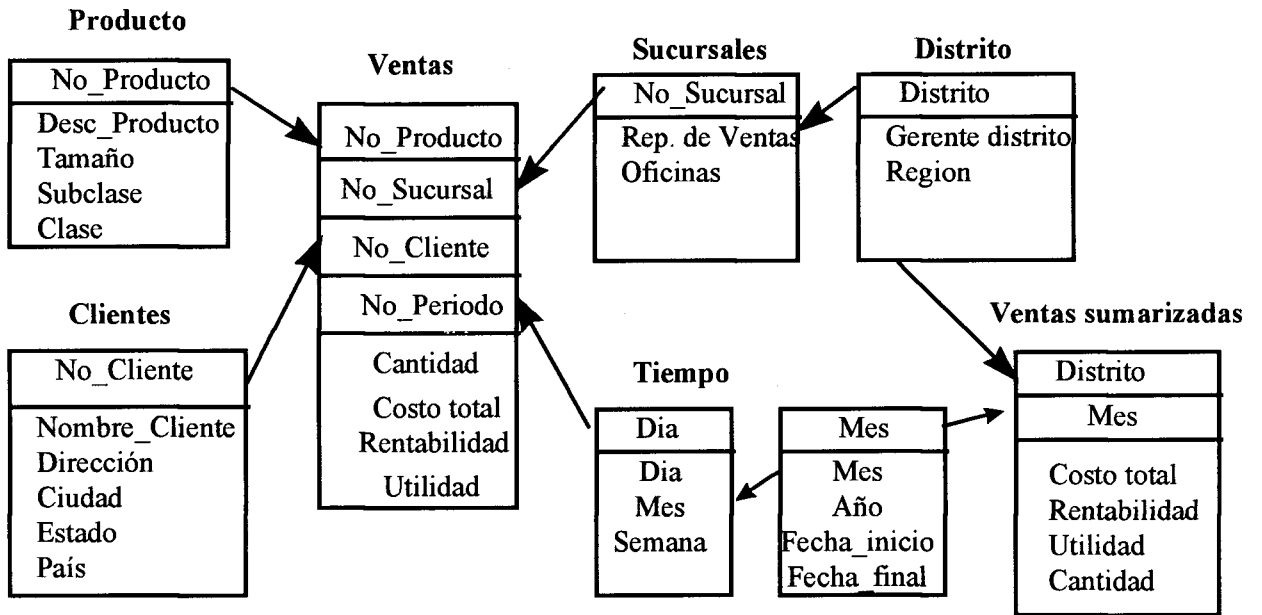


Figura 2.8 - Estrategia de agregación: Sumariación de ventas [MADSEN,1996]

En esta figura se observa que el objetivo que se persigue es el tener las ventas sumariadas por distrito y por mes, donde se muestre tanto el costo de lo vendido, la cantidad, la rentabilidad y la utilidad neta. Para lograr tal sumariación se requiere involucrar diversas tablas de datos, tal es el caso de Productos, Clientes, Sucursales; se necesita tener una orientación al tiempo y una base de datos que se actualice día con día, en este caso la base de datos operacional de Ventas. Las ventas sumariadas representan un Datamart dentro de la arquitectura Data Warehouse y puede ser utilizado en diversos sistemas de nivel gerencial, o bien servir para filtrar datos a otro nivel más alto de sumariación de acuerdo al diseño de la arquitectura.

El movimiento y transformación de los datos de una fuente de almacenamiento a otra puede generar muchos problemas en el proceso de sumariación o agregación de datos, ¿Qué pasaría si un proceso de agregación de datos se repite para el mismo período de tiempo?, o bien ¿Qué repercusiones tendría el sumar datos que posteriormente se descubre que estaban equivocados? ¿Se podrán corregir los datos en el Datamart?. El proceso de sumariación o agregación de datos dentro del Data Warehouse debe realizarse bajo estrictas medidas de control, y normalizado por procedimientos administrativos; además se debe asegurar que las herramientas que se empleen para tal proceso contengan validaciones que minimicen los efectos que pudieran causar este tipo de errores. En la estructura del Data Warehouse de la figura 2.7, se visualiza un elemento, el Metadata, el cual puede convertirse en el medio para mantener el control en el movimiento y transformación de los datos, y a la vez ser un instrumento que permita la flexibilidad para adaptar el diseño del Data Warehouse a nuevos requerimientos de información en la empresa.

2.7. METADATA

“El Metadata representa los datos acerca de los datos. A través del metadata el Data Warehouse puede obtener mejores resultados en sus procesos. El diseño del metadata generalmente contempla los siguientes puntos:

- La estructura de los datos entendible para el programador
- La estructura de los datos entendible para el analista de decisiones
- Las fuentes de datos que alimentan al Data Warehouse
- La transformación de los datos en cada paso de agregación del Data Warehouse
- El modelo de datos
- La relación entre el modelo de datos y el Data Warehouse
- La partición de los datos

- La historia de extracciones y sumalizaciones-agregaciones.
- Un catálogo de información “ [INMON,1996]

El Metadata vendría a ser el análisis del Data Warehouse, independientemente de las herramientas del software que vayan a emplearse durante el proceso de implantación, la creación del Metadata debe realizarse de acuerdo a la administración de los datos que se esté llevando en la empresa. Los analistas de decisiones son muchas veces los que deciden la información que debe procesarse en el Data Warehouse, y para poder determinar adecuadamente los requerimientos de información para realizar un proceso, se requiere de herramientas que ayuden a los analistas de decisiones a entender mejor el ambiente Data Warehouse, siendo ésta la función de un Metadata. “El Metadata es la primera cosa que el analista de decisiones debe mirar al momento de planear el procesamiento analítico de la información para cada tópico del negocio” [INMON,96].

Una razón por lo que el Metadata es importante en el Data Warehouse es que facilita la administración del mapeo entre el medio ambiente operacional y el medio ambiente del Data Warehouse. Los datos sufren una transformación significativa cuando pasan del medio ambiente operacional al medio ambiente del Data Warehouse, la conversión, los filtros, la sumalización, los cambio en las estructuras, son algunas de dichas transformaciones, es importante mantener entonces un monitoreo cuidadoso de las transformaciones de los datos para poder determinar los orígenes de los datos históricos y en algún momento dado, en caso de error, regresar los datos del Data Warehouse nuevamente a las bases de datos operacionales.

2.8. PARTICIÓN DE DATOS

Otro aspecto muy importante en el diseño de un Data Warehouse es la partición de datos. “La partición de datos se refiere a la separación de los datos en unidades físicas que pueden ser manejadas independientemente, en el Data Warehouse el aspecto de partición no se analiza desde un punto de vista si la partición debe ser hecha o no, sino en como llevarla a cabo” [INMON,1996].

Se ha mencionado que el Data Warehouse se diseña con el objetivo de permitir el acceso a los datos de una manera más rápida y confiable para la toma de decisiones. El tener bases de datos estratégicamente particionadas, orientadas al negocio y a las necesidades de información, permitiría que la administración de datos fuera más eficiente y flexible para poder satisfacer cualquier tipo de acceso o

consulta a los datos. “La partición de datos facilitaría el desarrollo de mejores métodos para ejecutar tareas de mantenimiento a la información, tales como:

- Respaldos
- Reestructuración y reorganización de datos
- Indexación
- Monitoreo de datos
- carga de datos” [YOUNG,1996]

Por otro lado, estas ventajas podrían anularse si la partición de datos no es administrada correctamente, lo cual provocaría redundancia y exceso de espacio en disco duro o en cualquier otro dispositivo de almacenamiento. Es precisamente en el Metadata donde se debe llevar el control y administración de los datos para una correcta ejecución y monitoreo de las estrategias de partición. La estrategia de partición debe planearse de acuerdo a las necesidades de información de los usuarios finales, y también con respecto a los tipos de acceso a los datos que se pueden presentar durante el proceso de toma de decisiones. La figura 2.9 nos presenta algunas ideas de partición de datos, la naturaleza histórica de un Data Warehouse, técnicamente hablando, obliga a determinar una llave primaria asignada a un campo que maneje el factor tiempo.

Tener en cuenta que un índice de partición no necesita ser la llave primaria

La llave aleatoria puede ser el número del cliente o el número del producto si no están asignadas secuencialmente

Se puede particionar por departamento o por lugar geográfico

Crear llaves artificiales para dividir los datos por períodos de tiempo. Asegurarse que la partición de los datos no va a provocar dificultad al momento de acceder algunas particiones

Crear una tabla de particiones. Si la llave para un dato es una fecha particular para un lugar geográfico particular colocarla en una partición particular.

Fecha de Inicio	Fecha Final	Lugar geográfico	Número de partición
Enero 1	Enero 31	Costa Este	1
Febrero 1	Febrero 29	Costa Este	22
.....			
Enero 1	Enero 31	Costa del Pacífico	50

Figura 2.9 - Ideas para la partición de datos [YOUNG,1996]

Ahora bien, si la partición de datos persigue en esencia mejorar y hacer más flexible el acceso a la información, debe buscarse entonces crear tales particiones de acuerdo a la cantidad de consultas de datos de los diversos niveles de la organización. Mientras más estratégico sea el pensamiento, mayor es la necesidad de consultas intensivas de datos, ver tabla 2.2.0

Tomadores de decisiones	Tipo de decisión	Tipo de consulta
Ejecutivos	Estratégico	Intensiva
Mandos medios	Estratégico/táctico	Selectiva/Intensiva
Operacional	Táctico	Selectiva

Tabla 2.2 - Tipos de consultas de datos [MEREDITH,1996]

La tabla 2.2 nos permite observar que en los niveles más altos de una organización es donde las consultas a los datos se dan de una manera más intensiva, por lo tanto la partición de datos debe combinarse con alguna estrategia de sumarización o agregación para lograr satisfacer rápidamente las consultas a las bases de datos en ese nivel organizacional, pero a la vez debe evitarse crear muchas particiones de datos que puedan generar otro problema similar a la de la “telaraña de la información”. Por ejemplo, considerando el caso de las ventas de una tienda de autoservicio, el número de particiones que pueden generarse es presentado en la tabla 2.3.

Cliente	Ventas por geografía	Producto	Tiempo
Todos los clientes	Ventas mundiales	Todos los productos	Anual
Un Mercado	Nacional	Una línea de producto	Cuatrimestral
Un Cliente	Regional	Un producto	Mensual
	Por sucursal		Semanal
	Por departamento		Diario
	Por vendedor		
Número total de particiones: $(3*6*3*5) = 270$			

Tabla 2.3 - Ejemplo de la partición de datos [MADSEN,1996]

En este ejemplo la cantidad de particiones es muy alta, por lo cual, para evitar un exceso de particiones en los datos se debe realizar un buen análisis para identificar la información que realmente va a ser requerida en cada medida posible del negocio (cliente, ventas por geografía, producto, tiempo). De dicho análisis pueden establecerse las estrategias de sumariación y partición de los datos que realmente den valor a la información y que soporten de alguna manera la toma de decisiones.

2.9. EL FLUJO DE INFORMACION EN EL DATA WAREHOUSE

La estructura del Data Warehouse de la figura 2.7 nos muestra diversos niveles de sumariación de los datos, los cuales se forman al aplicar diversas estrategias de agregación y partición; un elemento que influye en tales estrategias es el flujo de la información en la empresa, ya sea llevado a cabo de manera manual o automática.

La figura 2.10 nos muestra algunos de los conceptos mencionados como parte del diseño de un Data Warehouse, representando en este caso lo que sería el flujo de la información en una organización típica. “El valor real de un Data Warehouse es la administración del flujo de la información mas que en la colección de datos” [HACKATHORN,1995]. Se puede decir que el Data Warehouse mas que ser un conjunto de productos, representa una arquitectura de información para toda una empresa. Cada uno de los flujos de información que se presentan en la figura son parte importante a considerar cuando se diseñe un Data Warehouse.

El *flujo hacia adentro* facilita la integración de datos, necesaria en un Data Warehouse, sucediendo en el momento de realizar el proceso de validación y reparación de datos, además de que representa el paso de datos del ambiente operacional al ambiente Data Warehouse. El *flujo hacia arriba* es el proceso que le da valor a la información contenida en el Data Warehouse, siendo aquí donde se involucran las estrategias de sumariación o agregación, y partición de datos. El *flujo hacia afuera* representa las posibles consultas en los diferentes niveles organizacionales de la empresa, siendo aquí el usuario el que maneja los datos del Data Warehouse y ya no personal informático. El *flujo hacia abajo* permite mantener la vitalidad del Data Warehouse al enviar los datos antiguos a dispositivos de almacenamiento secundario, siendo éstos los datos que ya no dan valor al negocio.

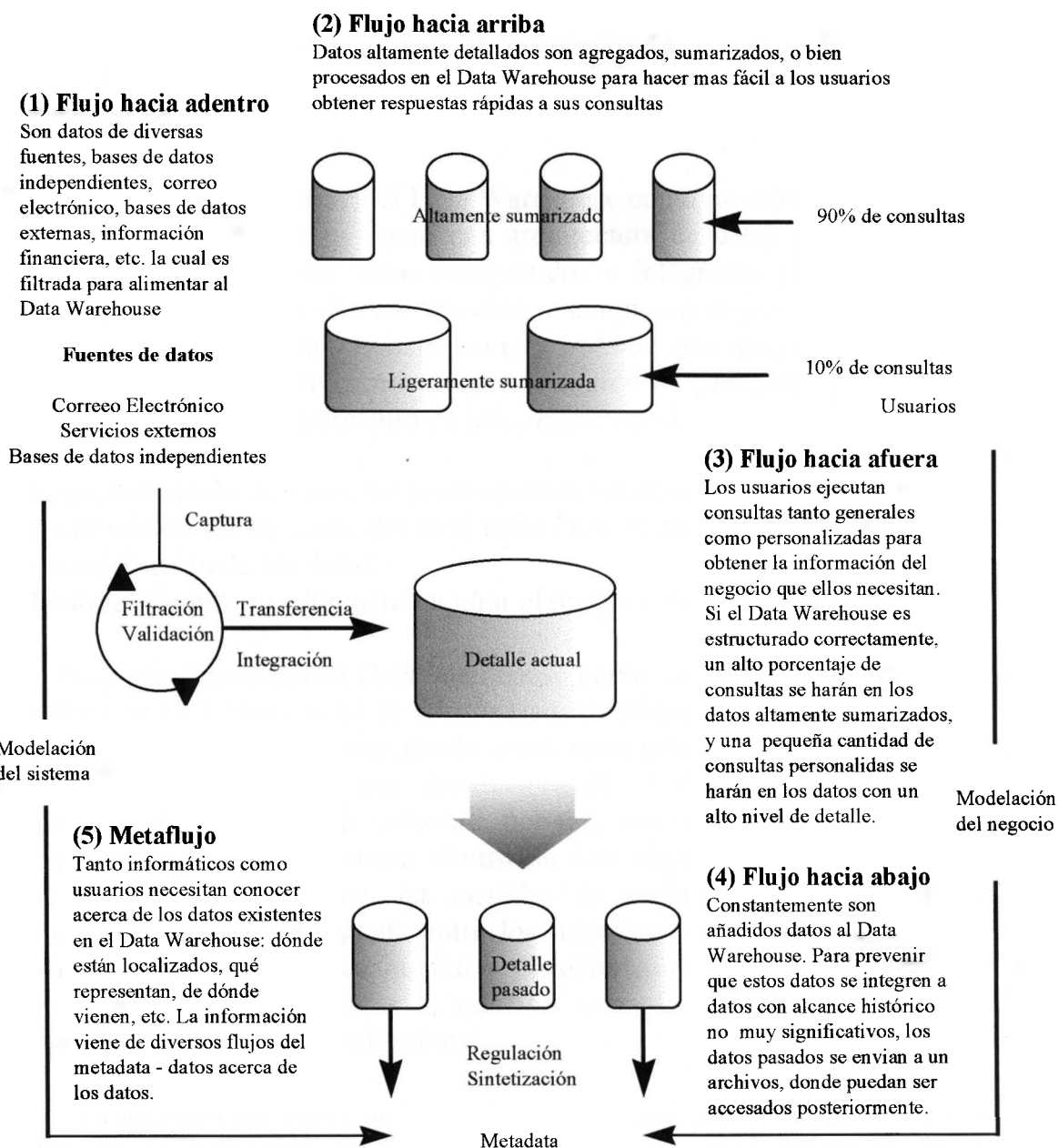


Figura 2.10 - Dinámica del flujo de información [HACKATHORN,1995]

Por último, el *Metaflujo* es el proceso que mueve a los metadata, que en este caso representaría los datos acerca de los otros flujos mencionados. Para responder a los cambios organizacionales, los sistemas de información están constantemente cambiando, y en consecuencia cambian también las estructuras y formatos de los datos, de tal manera que el Data Warehouse requiere una depuración continua que deba basarse en un monitoreo tanto de los sistemas de información, como de los cambios en el medio ambiente de la empresa.

2.10. DEFINICION DEL DATA WAREHOUSE PARA FINES DE ESTA TESIS

Algunos autores definen el Data Warehouse como un administrador de base de datos [INMON,1996], otros como una arquitectura de datos [STRANGE,1995], y otros como una fuente de datos compartidos e integrados [CHASIN,1994]; para fines de esta tesis defino al Data Warehouse como una *arquitectura de información con un enfoque sumariado e histórico en los datos, que apoya al proceso de toma de decisiones en los niveles táctico y estratégico de una organización*. Propongo asimismo dividir esta arquitectura de información en 3 capas principales:

- La transformación de datos del nivel operacional al nivel Data Warehouse
- La administración de datos desde el nivel Data Warehouse hasta el nivel más alto de sumarización de los datos.
- La entrega de información para mejorar el proceso de toma de decisiones

Propongo visualizar al Data Warehouse como una arquitectura de información y su división en 3 capas, a fin de apoyar las actividades que conforman el proceso de asimilación de tecnología presentando como tema principal de esta tesis. Visualizar al Data Warehouse como una arquitectura de información, conformada por un conjunto de herramientas de software, por una serie de conceptos y procesos que hay que entender y saber ubicar dentro de esta arquitectura, permitirá establecer, como se verá más adelante, las medidas de evaluación para determinar si la tecnología está siendo adoptada entre los miembros de la empresa, asimismo los esfuerzos de difusión, entrenamiento y conocimiento de la tecnología van a estar mejor enfocados al identificar la relación entre la arquitectura de información propuesta y el contexto organizacional.

La división por capas de la arquitectura de información Data Warehouse esta basada en la estructura de la figura 2.7; así es como defino la capa de la *transformación de datos* del nivel operacional al nivel Data Warehouse, en donde están involucradas las bases de datos operacionales, todos los procesos de extracción y transformación de datos y la conformación misma de las bases de datos históricas del Data Warehouse. La capa de la *administración de datos* está conformada de todos los procesos de sumarización y partición de datos para la creación de los diversos niveles de detalle, así como la creación de los modelos de datos para entender el flujo de la información en la empresa. Por último, la capa de *entrega de información* consiste en la explotación de las bases de datos altamente sumarizadas para crear análisis, o bien para visualizar la información que soporte los procesos de

toma de decisiones. La tabla 2.4 presenta las capas que conforman la arquitectura de información Data Warehouse que propongo, así como los procesos involucrados, y las herramientas y conceptos relacionados.

CAPA	PROCESOS	CONCEPTOS Y HERRAMIENTAS DE SOFTWARE RELACIONADOS
Transformación de datos	Integración de datos Extracción de datos Creación de datos históricos Limpieza de datos Conversión de datos	OLTP Base de datos Herramientas Data Warehouse (limpieza, modelación, etc.) Data Mining Metadata
Administración de datos	Sumarización de datos Partición de datos Creación de modelos de datos Reorganización de datos	DSS OLAP Herramientas de modelación de datos Data Mining Datamart (Base de Datos) Herramientas como hojas de cálculo y análisis estadístico
Entrega de información	Presentación de Información Análisis de Información	EIS (Base de Datos) DSS Reporteadores y herramientas de análisis

Tabla 2.4 - Capas que conforman la arquitectura de información Data Warehouse

2.11. CONCLUSIONES

En teoría el Data Warehouse se define como una colección de datos que soportan las decisiones administrativas, pero para fines de esta tesis defino al Data Warehouse como una arquitectura de información que esta soportada por una base de datos y una serie de herramientas de software que facilitan el llevar a cabo los procesos principales de esta arquitectura. Se le da ahora un nuevo valor a los datos, que tiene que ver más con la toma de decisiones y no tanto como soporte a determinadas actividades de la empresa.

Tomando como base el concepto de la pirámide organizacional [ANTHONY,1989], y para fines de relacionar el entorno de la empresa con la arquitectura de información Data Warehouse, presento en la figura 2.11 la representación gráfica de esta arquitectura de información. Según Anthony los sistemas de información de una empresa pueden identificarse planteando 3 niveles administrativos: planeación estratégica (nivel estratégico), control administrativo (nivel táctico) y el control operativo (nivel operacional), marcando los requerimientos de información para cada uno de los niveles.

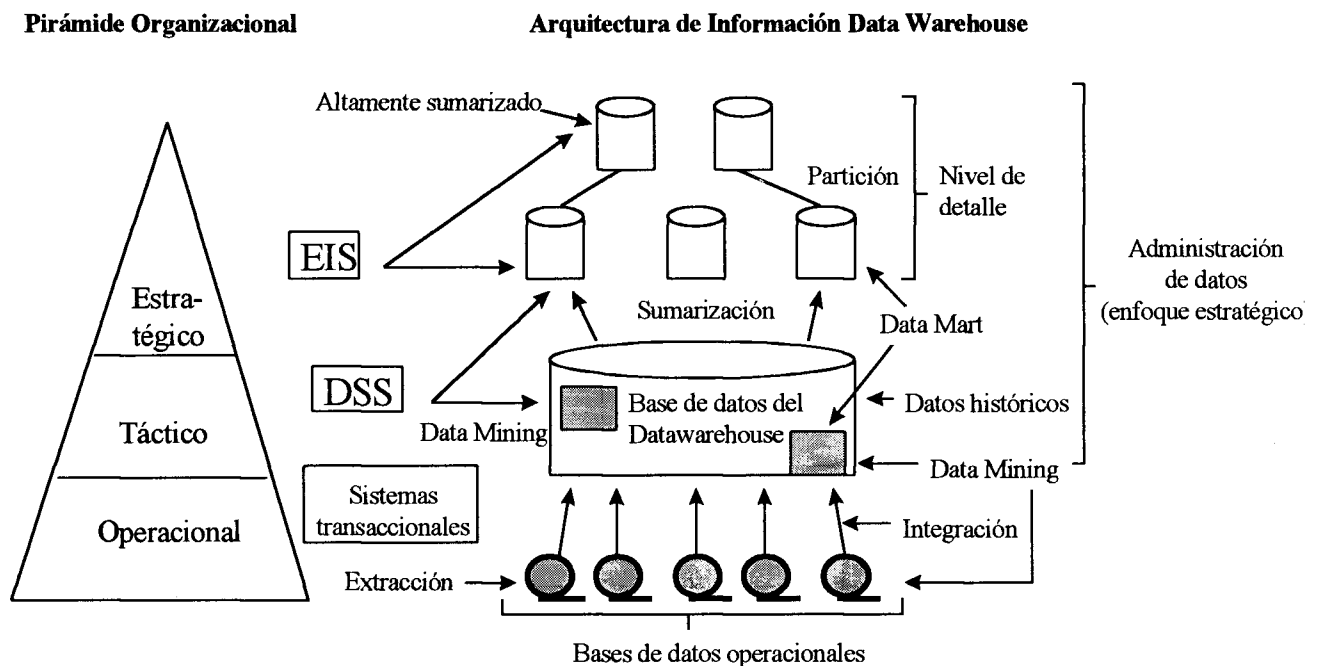


Figura 2.11 - Arquitectura de Información Data Warehouse

Tomando en cuenta el modelo de la pirámide organizacional de Anthony, y la estructura de la arquitectura de información Data Warehouse de la figura 2.7, se establece una relación entre los sistemas de información y los diversos conceptos y procesos involucrados en la arquitectura Data Warehouse mencionados en la tabla 2.4, y así poder ubicar a cada sistema en su nivel organizacional correspondiente, según la pirámide.

En el nivel operacional se identifican los diversos Sistemas transaccionales que soportan las operaciones diarias de la compañía (nómina, control de inventarios, compras, etc.), estos sistemas están directamente relacionados a las bases de datos operacionales de la arquitectura Data Warehouse; de dichas bases de datos se extraen los datos que son “limpiados” e integrados a las bases de datos históricas del Data Warehouse.

En el nivel táctico se identifican los sistemas de información administrativos, los cuales usan datos de los sistemas operacionales para mantener a la administración informada acerca de como está operando la organización (contabilidad, presupuestos, pronósticos); a partir de este nivel comienza el soporte del Data Warehouse para apoyar la toma de decisiones, siendo necesario que la administración de datos se torne estratégica para poder lograr un diseño exitoso de la arquitectura de información. Entre los sistemas de información que pueden explotar la información histórica del Data Warehouse y de algunos Datamarts están los Sistemas de Soporte a la decisión (DSS), los cuales usando modelos estadísticos y matemáticos pueden calcular pronósticos, presupuestos, y demás aplicaciones de nivel táctico.

En el nivel estratégico se identifican los Sistemas de Información Ejecutivos (EIS), los cuales son diseñados para dar a los ejecutivos acceso a las bases de datos externas a la empresa, así como a los niveles más altos de sumarización de los datos internos. Las bases de datos de los EIS pueden crearse de los procesos de sumarización y partición de datos, o bien emplear algún Datamart ya generado con la información requerida.

Algunos de los conceptos representados en la figura 2.11 que conforman la arquitectura de información Data Warehouse son explicados independientemente en el siguiente capítulo de esta tesis.

REFERENCIAS BIBLIOGRAFICAS

- [ANTHONY,1989] ANTONHY, G., Micheal Scott Morton
"A Framework for management information
Systems".
Sloan Management Review, Spring, 1989
- [CHASIN,1994] CHASIN, Ken.
"Designing Data Warehouses using CASE
Technology".
Journal of Systems Management, Mayo, 1994
- [FRANCET,1994] FRANCETT, Barbara.
"Decisions, Decisions: Users take stock of Data
Warehouse Shelves".
Software Magazine, Agosto 1994.
- [FRANCET,1995] FRANCETT, Barbara.
"Database Technologies vie for Data Warehouse
occupancy".
Software Magazine, Abril 1995
- [HACKATHORN,1995] HACKATHORN, Richard.
"Data Warehousing energizes your enterprise".
Datamation, Febrero 1, 1995
- [INMON,1996] INMON, W.H.
Building the Data Warehouse.
John Wiley & Sons, Inc, segunda edición, 1996
- [LAZAR,1996] LAZAR, Bill.
"The data breakthrough".
LAN, Mayo 1996
- [MADSEN,1996] MADSEN, Mark.
"Warehouse design in the aggregate".
Database Programming and Design, Julio 1996.

- [MCFADDEN,1994] Mc. FADDEN, Fred R.
Modern Database Management.
Addison Wesley Publishing, 1994
- [MEREDITH,1996] MEREDITH Mary, Khader Aslam.
"Divide and Aggregate".
Database Programming and Design, Junio 1996
- [MORIARTY,1996] MORIARTY, Terry.
"Barriers to data quality Part I: Disparate Data".
Database Programming and Design, Abril 1996
- [STRANGE,1995] STRANGE, Kevin H.
"The decision maker's gold mine: The Data
Warehouse".
Datamation, Mayo 15, 1995
- [WALLACE,1994] WALLACE, Peggy.
"Building a data warehouse".
Infoworld, Febrero 21, 1994
- [YOUNG,1996] YOUNG, Case.
"So you're not Rumpelstiltskin...".
Database Programming and Design, Julio 1996.

CAPITULO 3

ARQUITECTURA DE INFORMACION DATA WAREHOUSE

3.1. INTRODUCCION

Para fines de esta tesis he definido al Data Warehouse como una arquitectura de información con un enfoque sumariado e histórico en los datos, que apoya el proceso de toma de decisiones en los niveles táctico y estratégico de una organización. La tabla 2.4 y la figura 2.11 nos permiten identificar los componentes principales de esta arquitectura de información. "El término Arquitectura de Información ha sido usado por la comunidad de los informáticos desde antes de los 80's. Es usado como una metáfora que denota un modelo para la creación y flujo de los datos a nivel de toda una corporación. Este modelo y las *metodologías relacionadas* tratan de documentar todas las entidades y sus relaciones en una forma sistemática. Su objetivo principal dentro de una organización es tener un *mapa congruente de los datos* para basándose en él, construir los sistemas de información; el producto final es una estructura que use *tecnologías disponibles* para moldear y forzar al medio ambiente del negocio, de tal modo que determinadas actividades humanas pueden ser hechas más efectivamente". [MCGEE,1993].

Se mencionó la importancia de visualizar al Data Warehouse como una arquitectura de información que permita apoyar las actividades del proceso de asimilación de esta tecnología en una empresa, este apoyo puede presentarse si mediante el diseño de la arquitectura de información Data Warehouse puede generarse un mapa congruente de los datos a nivel de la organización, un mapa que permita eliminar barreras organizacionales en el compartimiento de los datos, el conocimiento del estado de los datos, las interrelaciones entre los procesos de la empresas, y demás aspectos que pueden influir en la grado de adopción o asimilación del Data Warehouse al momento de su implantación. Para el siguiente capítulo se explicará lo referente al flujo de la información de la organización para su análisis en el diseño de la arquitectura de información, y en este capítulo presento las *tecnologías de información disponibles* que considero más importantes analizar para fines de adaptarlas a la arquitectura de información Data Warehouse, de tal manera que se logre una adecuada integración entre la arquitectura y las tecnologías de información que permitan mejorar la productividad en las actividades humanas relacionadas con la toma de decisiones.

En el capítulo anterior se explicaron algunos de los procesos que conforman la arquitectura de información Data Warehouse, la sumarización y la partición de datos estuvieron entre éstos, asimismo se definieron los conceptos Data Mart y Metadata, ahora bien, en este capítulo se explican algunos de los conceptos y herramientas de software que se muestran en la tabla 2.4, principalmente aquéllas que influyen directamente en la comprensión y entendimiento de la arquitectura de información Data Warehouse.

3.2. EL DATA WAREHOUSE COMO ARQUITECTURA DE INFORMACION

Se ha mencionado que el Data Warehouse cumple con la característica de una arquitectura de información en lo que se refiere al flujo de los datos, pero con lo que respecta al objetivo de cualquier arquitectura de información, ¿puede el Data Warehouse crear un mapa congruente de los datos y ser promotor de nuevos sistemas de información?. Estas preguntas no pueden ser directamente contestadas ya que dependen mucho del proceso de implantación del Data Warehouse en cada empresa. Existen diversos factores que influyen en la creación de un mapa congruente de los datos, principalmente en lo que se refiere al apoyo que proporciona la alta dirección para considerar a la administración de datos con un enfoque estratégico, asimismo se tiene la calidad de los datos y las prácticas organizacionales involucradas con el manejo de los mismos, y algo que es muy importante, la difusión que se le dé a la nueva tecnología para que pueda ser asimilada totalmente dentro de la organización, y poder así asegurar la participación del recurso humano en la creación del mapa congruente de los datos, ya que finalmente son ellos los que proponen nuevos sistemas de información de acuerdo a sus necesidades.

Los diversos procesos que conforman la arquitectura Data Warehouse, tal como la sumarización, la extracción, la integración, la partición de datos, etc., pueden ser la base para administrar de manera estratégica a los datos, ya que a través de estos procesos los datos pueden convertirse en información que apoye las decisiones más importantes en una empresa, principalmente aquellas decisiones que afectan a las estrategias del negocio. La administración estratégica de los datos por medio de los procesos Data Warehouse permitiría a los usuarios de la información visualizar a la empresa a través de un modelo de datos más general, más orientado al negocio y más congruente a sus deseos y necesidades.

Ahora bien, los datos necesitan ser explotados y analizados, siendo el medio más importante para lograrlo el uso de sistemas de información y herramientas de software analíticos. “Un indicador muy importante para el éxito del Data Warehouse es si promueve el diseño de nuevos sistemas operacionales” [Hackathorn1995]; además de los sistemas operacionales o transaccionales, la arquitectura Data Warehouse debe promover el desarrollo de nuevos sistemas de información a nivel táctico (DSS) y a nivel estratégico (EIS), que permitan explotar y analizar a los datos.

Hasta aquí hemos identificado diversos componentes de la arquitectura de información Data Warehouse, tal como la administración de datos, los procesos de diseño del Data Warehouse, los DSS y los EIS. Además de estos componentes podemos identificar otros más empleando la analogía entre arquitectura física y arquitectura de información propuesta por James McGee y Laurence Prosak [MCGEE,1993], y de la cual se presenta una adaptación al Data Warehouse. La analogía entre arquitectura física y arquitectura de información esta dividida en tres elementos que interactúan entre sí para alcanzar el objetivo final de una arquitectura física o de información. Los elementos que conforman esta analogía son los siguientes:

- 1) El medio ambiente físico determina los extremos a tomar en cuenta
- 2) Las capacidades y sentidos humanos crean guías y establecen predisposiciones que los arquitectos pueden explotar
- 3) La evolución de las tecnologías estructurales y de edificios introduce nuevas opciones en la solución de espacios, pero también establecen limitantes tecnológicas a tomar en cuenta

La tabla 3.1 presenta la analogía entre arquitectura física y arquitectura de información, identificando ejemplos en cada elemento y su adaptación al Data Warehouse. De esta analogía se identifican otros componentes de la arquitectura de información Data Warehouse, así como algunos conceptos organizacionales y humanos que influyen fuertemente en cualquier proceso de asimilación de tecnología. “Los arquitectos de información necesitan combinar el arte y la tecnología para crear el medio ambiente de información de las organizaciones, deben mantener un balance entre las necesidades de información y las limitaciones tecnológicas, deben localizar en las estrategias del negocio cuál información es importante para la organización y llegar a ser uno de los medios de comunicación para hacer llegar dicha información al resto de la organización. Una buena

arquitectura de información debe ser capaz de representar cualquier tipo de información” [MCGEE,1993]

Elemento	Arquitectura física	Arquitectura de información	Data Warehouse
Medio ambiente físico	Características del clima local, Temperatura máxima y promedio, lluvias, vientos Terreno, Collados, cerros, cuevas Suelo, estabilidad, grietas ocultas	Disponibilidad y calidad en los datos Identificación de competidores y sus actividades Conocimiento actual de las esferas políticas y económicas Conocimiento acerca del desarrollo de tecnologías	Análisis de la situación actual de los datos Conocimiento de la TI que soporte el diseño de un Data Warehouse Conocimiento de la organización y sus estructuras
Capacidades humanas	Tolerancia a los extremos del medio ambiente Límites culturales en cuanto a los espacios físicos personales	Habilidades y fortalezas a largo plazo Procesamiento de información y estilos en la toma de decisiones	Nivel de cultura informática para el uso de nuevas herramientas de software Apoyo al proceso de toma de decisiones Grado de importancia de la información
Evolución y límites tecnológicos	Concreto reforzado Estructuras geofísicas Resistencia de los materiales	Tecnología de Base de Datos Redes de comunicación Tecnologías para trabajo en grupo	Tecnologías de Base de Datos (DBMS) Herramientas de conectividad Herramientas Data Warehouse

Tabla 3.1 - Analogía entre arquitectura física y arquitectura de información [MCGEE,1993]

3.3. TECNOLOGIA DE BASE DE DATOS

La tecnología de base de datos, de acuerdo a la analogía entre arquitectura física y arquitectura de información, se encuentra en los límites tecnológicos a considerar en el diseño de la arquitectura de información Data Warehouse. Las bases de datos bien diseñadas y administradas proporcionan una base sólida para el resto de la arquitectura de información. Como se puede observar en la tabla 2.4, el concepto Base de Datos está incluido en las tres capas que conforman la arquitectura de información Data Warehouse, en cada una de estas capas tiene una función diferente, pero en sí el concepto es el mismo, y hay que entenderlo así primeramente. Desde su creación, el concepto y la aplicación de base de datos ha sido de gran importancia para la administración de datos y en general para algunas tecnologías de información. Desde los primeros manejadores de bases de datos como Dbase hasta los actuales como Oracle e Informix, siempre se ha intentado dar eficiencia a la administración de las bases de datos para generar la información que los usuarios necesitan. Ahora con el Data Warehouse la administración de las bases de datos sigue siendo de gran importancia, pero ahora el enfoque es hacia la administración de decisiones a través de un análisis de lo que contienen las base de datos y no tanto en las consultas a las bases de datos con todo su detalle.

"Una base de datos es una colección de datos que contienen información acerca de una empresa particular, siendo la meta principal de su sistema de administración proporcionar un ambiente que sea conveniente y eficiente para usar la información que se obtiene y se almacena en ella" [KORTH,1991]. Conocer el ambiente que proporciona un DBMS (Data Base Management System) es importante para el análisis de un Data Warehouse ya que a través de éste se conocen las estructuras y los niveles de abstracción de los datos, así también las herramientas de obtención de información, las facilidades en el manejo de los datos, las relaciones entre los mismos, etc. Como definición, "un sistema de administración de base de datos (DBMS) es una colección de archivos interrelacionados y una serie de programas que permiten a los usuarios acceder y modificar esos archivos" [KORTH,1991].

Dentro del estudio de implantación de un Data Warehouse el análisis de las bases de datos son esenciales para llevar a cabo la sumarización, la agregación y la partición de los datos, desde un punto de vista lógico y físico. Como se puede observar en la estructura del Data Warehouse propuesta por Bill Inmon (ver figura 2.7), las bases de datos se encuentran en diversos niveles de detalle y sumarización, desde lo operacional hasta lo altamente sumarizado. Ahora bien, para pasar de un nivel de detalle a otro se requiere conocer adecuadamente las bases de datos y sus

sistemas de administración, para tener de esta manera confiabilidad e integridad en la transformación de los datos del nivel operacional al nivel Data Warehouse o bien, confiabilidad en los procesos de sumarización o agregación de los datos.

“La base de datos es el corazón de cualquier Data Warehouse, y realizar un buen diseño es muy importante en los sistemas OLAP (On Line Analytic Processing) tanto como lo es en los sistemas OLTP (On line Transaction Processing), la diferencia está en que las tablas de las bases de datos en los sistemas OLTP (nivel operacional) están orientadas al detalle, mientras que las tablas en los sistemas OLAP (nivel táctico y estratégico) se conforman de agregaciones, sumarizaciones, y estructuras orientadas al tiempo” [LAZAR,1996]. En la tabla 2.1 se presentaron algunas diferencias entre el diseño de las bases de datos operacionales y las bases de datos a nivel Data Warehouse, y para complementar el entendimiento de la función de las bases de datos en el diseño de la arquitectura de información Data Warehouse es más conveniente visualizarlas de acuerdo al nivel organizacional en donde están posicionadas. La figura 3.1 nos muestra los tipos de bases de datos de acuerdo a esta clasificación.

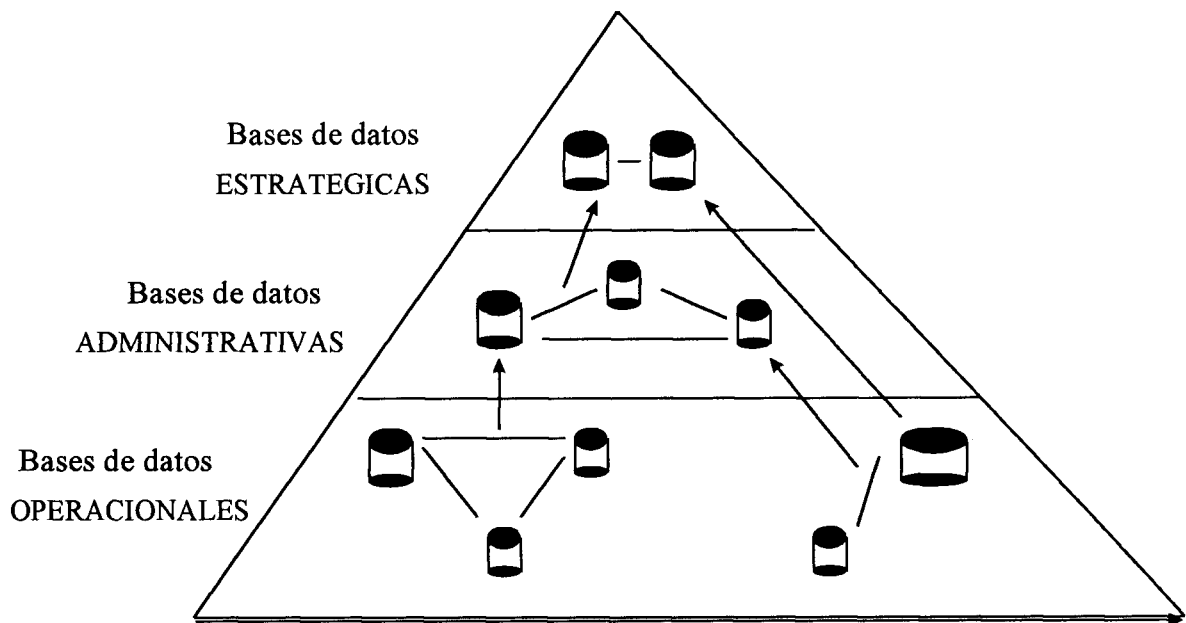


Figura 3.1 - Tipos de Bases de Datos [MCFADDEN,1994]

Con esta clasificación de las bases de datos podemos visualizarlas de acuerdo a las tres capas de la arquitectura de información Data Warehouse, presentadas en la tabla 2.4. Las bases de datos operacionales corresponden a la capa de transformación de los datos, considerando que de dichas bases de datos se generan los procesos

Data Warehouse que corresponden a dicha capa de la arquitectura (extracción, conversión, limpieza de datos, etc.). En la capa de administración de datos se crean los Data Marts, los cuales se conforman principalmente de bases de datos que satisfacen necesidades administrativas o tácticas, cabe aclarar que los Data Marts pueden crearse también a un nivel estratégico dentro de la empresa; en esta capa se involucran las bases de datos administrativas de la figura 3.1. En el nivel más alto de la organización se encuentran las bases de datos estratégicas, las cuales representarían las bases de datos de los Sistemas de Información Ejecutivos (EIS), con lo que nos estaríamos refiriendo a la capa de entrega de información, este tipo de base de datos tienen un alto grado de sumarización.

Se puede decir entonces que un Data Warehouse usa las bases de datos operacionales de una empresa para formar su propia base de datos con un enfoque histórico, y a la vez con la información que genera alimenta a las bases de datos de los sistemas "front-end", los cuales soportan la toma de decisiones de los usuarios finales. De ahí la importancia de entender el concepto de base de datos y sus sistemas de administración. Algunos ejemplos de los tipos de bases de datos se presentan en la figura 3.2.

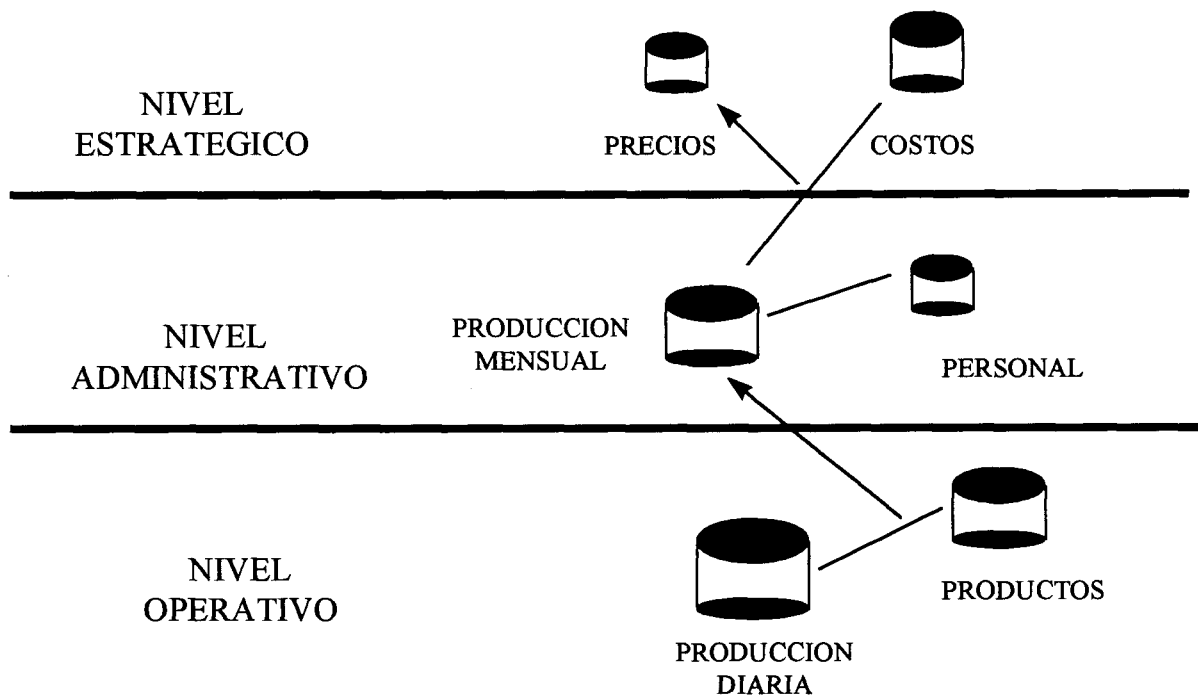


Figura 3.2 - Ejemplos de los tipos de Bases de Datos [MCFADDEN,1994]

3.4. HERRAMIENTAS DATA WAREHOUSE

Hasta hace algún tiempo podría considerarse como imposible diseñar un Data Warehouse tal como se le conoce actualmente, las limitantes tecnológicas en herramientas de software para el manejo de los datos hicieron que la arquitectura de información Data Warehouse se lograra conceptualizar sólo hasta hace algunos años, Bill Inmon definió al Data Warehouse en 1990 (ver 2.4.). En la analogía entre arquitectura física y arquitectura de información se hace mención de la evolución de las tecnologías estructurales y de edificios, siendo las herramientas de software las tecnologías que han estado evolucionando hasta conceptualizar al Data Warehouse. “En los últimos tres años, cientos de herramientas Data Warehouse han sido introducidas o redireccionadas al mercado del Data Warehouse” [ROBINSON,1996].

Los diseñadores del Data Warehouse deben tener la habilidad de evaluar y seleccionar las herramientas más adecuadas para el diseño de su arquitectura de información, más adecuadas en el sentido de la compatibilidad con la tecnología de información que ya existe en la empresa, también con respecto al nivel de cultura informática que tengan los empleados y usuarios para poder usar nuevas herramientas de software.

Para la creación de la arquitectura de información Data Warehouse es importante identificar el estado del medio ambiente físico de la organización, en cuanto a las tecnologías de información existentes, y mediante la evaluación de nuevas herramientas tecnológicas aplicar las que mejor lleven a cabo los procesos Data Warehouse. Es importante dejar claro la diferencia entre las herramientas claves para el desarrollo de la arquitectura Data Warehouse y las herramientas Data Warehouse propiamente dichas; las primeras se refieren a todas las herramientas que hacen posible la creación de la arquitectura Data Warehouse, lógicamente incluyen a las herramientas Data Warehouse, se pueden mencionar por ejemplo, los DBMS y las herramientas de conectividad, estas herramientas se contemplan en las tres capas de la arquitectura de información Data Warehouse; con respecto a las herramientas Data Warehouse podemos identificarlas debido a que realizan procesos de este nuevo concepto, tal como la limpieza de datos (capa de transformación de datos), la modelación de datos (capa de administración de datos), y la generación de reportes analíticos (capa de entrega de información), estas herramientas también se contemplan en todas las capas de la arquitectura de información Data Warehouse.

No es fácil realizar la evaluación de estas herramientas debido a la gran cantidad y diversidad con que se encuentran en el mercado Data Warehouse, por lo

que es conveniente al momento de la evaluación agruparlas de acuerdo a la función que realizan. Teri Robinson [ROBINSON,1996] propone los siguientes ocho componentes o grupos:

- 1) Componentes para la adquisición de datos
Herramientas que capturan, limpian y transportan datos de diversas fuentes a las bases de datos del Data Warehouse.
- 2) Componentes de diseño
Herramientas que definen las bases de datos del Data Warehouse
- 3) Componentes para la administración de datos
Herramientas para el manejo de las bases de datos relacionales
- 4) Componentes para el directorio de información
Herramientas que administran el Metadata
- 5) Componentes para la entrega de datos
Herramientas que distribuyen los datos de las bases de datos del Data Warehouse a los Data Marts o a las bases de datos de sistemas como los DSS o EIS.
- 6) Componentes Middleware
Proporcionan la conectividad entre los clientes y el Data Warehouse
- 7) Componentes de acceso a los datos
Herramientas cliente para consultas, reportes y análisis
- 8) Componentes de administración
Herramientas que manejan tareas operacionales tales como el monitoreo del rendimiento de los datos y las funciones de respaldo y recuperación.

Las herramientas claves de mayor uso en un proyecto Data Warehouse, según consultas aplicadas a 200 jefes de TI, se muestran en la figura 3.3. La agrupación de las herramientas de acuerdo a la función que realizan, tal como lo propone Robinson, además de facilitar la evaluación de las herramientas, permite tener un mejor entendimiento de como poder ubicar a cada herramienta en la arquitectura de información, planear su función y la capacitación que debe tener el recurso humano para poder operarla, estos últimos son aspectos muy importantes en un proceso de asimilación de tecnología en una organización.

3.5. DATA MINING

Entre algunas de las funciones principales del Data Warehouse se encuentran la extracción y el análisis de información valiosa de las bases de datos, sobre todo datos que solucionen problemas semiestructurados o no estructurados. Diseñadores y proveedores de software están creando una tendencia hacia esas herramientas de extracción y análisis de datos, surgiendo con ello un concepto que forma parte y a la vez es independiente al Data Warehouse: El Data Mining.

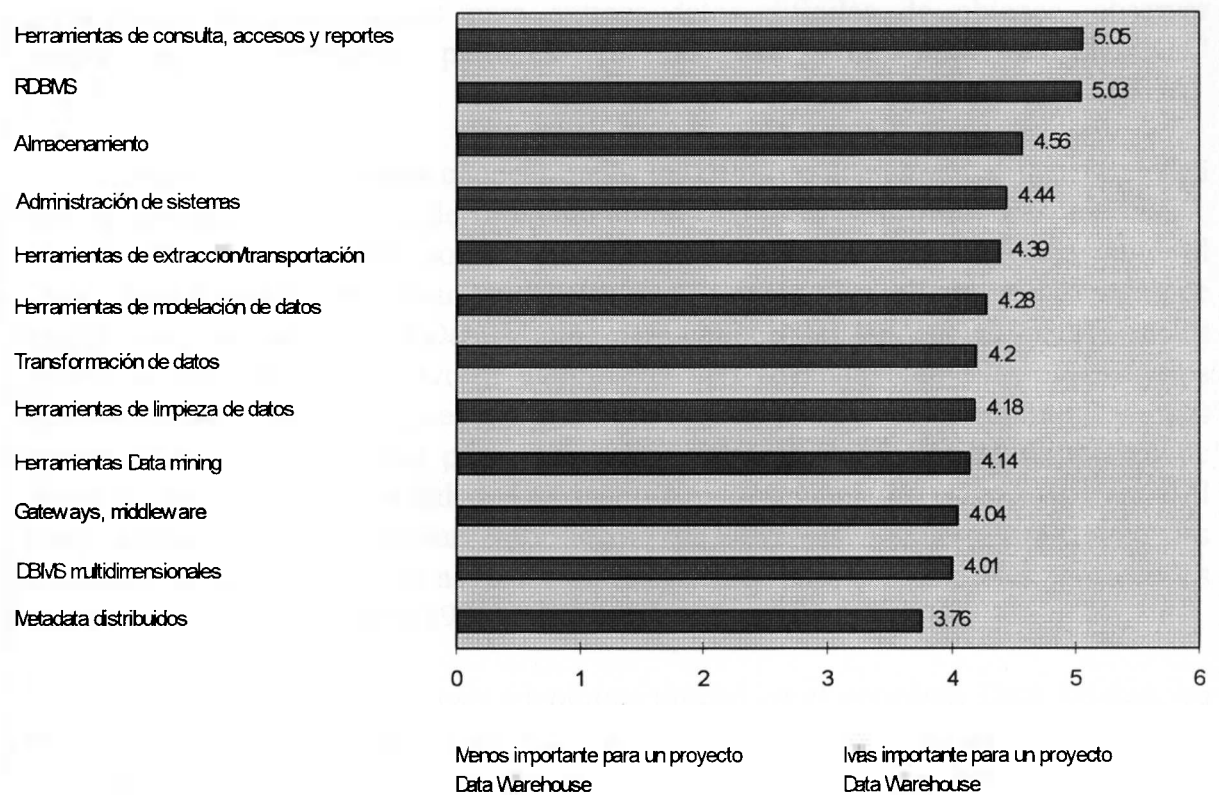


Figura 3.3 - Herramientas claves en un proyecto Data Warehouse [ROBINSON, 1996]

Como se observa en la figura 3.3, las herramientas Data Mining están consideradas como claves para el desarrollo de un Data Warehouse, se estudia aparte en este apartado del capítulo por considerarlo como un componente muy importante en la arquitectura de información Data Warehouse.

Se argumenta [MCGEE,1993] que una de las causas por lo cual no han sido exitosos los esfuerzos en el desarrollo de diversas arquitecturas de información, es la falta de consideración de los datos no estructurados o semiestructurados, así como de los datos externos a la empresa. En cuanto a los datos externos, vemos que el flujo de información en el Data Warehouse (ver figura 2.10 - Flujo hacia adentro -) contempla a este tipo de datos en la integración hacia las bases de datos históricas del Data Warehouse. Con respecto a los datos no estructurados o semiestructurados, las herramientas Data Mining tienen como función analizarlos, ya sea de las bases de datos operacionales o de las bases de datos históricas del Data Warehouse (ver figura 2.11). "El Data Mining es una consecuencia directa del mercado del Data Warehouse; es una nueva clase de accesos y análisis de datos que dan a las empresas capacidades no antes vistas para extraer datos difíciles de obtener, observar tendencias, y reconocer patrones en las bases de datos corporativas" [DEPOMPA,1996].

El Data Mining permite conocer mejor los datos de una empresa, generando así que se conozca más de ella, de su medio ambiente, clientes, proveedores, mercados, etc. Mientras el Data Warehouse permite visualizar a los datos históricamente, el Data Mining permite observar tendencias de los datos para el futuro. "Tiendas de autoservicio quieren usar el Data Mining para conocer las tendencias y preferencias de los clientes, con el objetivo de administrar mas adecuadamente las ofertas y los inventarios en diversos lugares geográficos. Las compañías de telecomunicaciones desean usar el Data Mining para pronosticar patrones de demanda, segmentos de clientes, análisis de rentabilidad. Las instituciones financieras están empleando el Data Mining para consolidar información de diversas fuentes, analizando las tendencias y patrones de los negocios de los clientes para así venderles y ofrecerles mas servicios" [DEPOMPA,1996].

No es objetivo en esta tesis adentrarse mucho en el concepto Data Mining, ya que es muy amplio y variado, pudiendo incluso considerarlo como un tema de tesis a investigar. El objetivo más bien, es analizar la función de las herramientas Data Mining en el diseño de la arquitectura de información Data Warehouse. Las herramientas conocidas como de soporte a la decisión son muy importantes en el contexto de la arquitectura de información Data Warehouse, debido a que son muchas veces la herramienta final de los usuarios de la arquitectura. Dichas herramientas deben estar enfocadas a explotar de la mejor manera posible los datos administrados en el Data Warehouse, ya sean datos históricos, sumariados, calculados, particionados, etc., las herramientas Data Mining pueden ayudar a que la explotación de tales datos sea más eficiente y de acuerdo a un objetivo específico del negocio.

“El Data Mining es usado por las más sofisticadas herramientas de soporte a la decisión. Entre las herramientas más simples se encuentran los Sistemas de Información Ejecutivos, los cuales generan un número limitado de reportes de alto nivel. Entre algunas de las herramientas de mayor complejidad están los Sistemas de Soporte a la Decisión, los cuales permiten a los usuarios seleccionar parámetros para generar un gran cantidad de reportes. Y de las herramientas más complejas que usan el Data Mining están las que generan reportes más personalizados de acuerdo a una necesidad específica del usuario” [KRIVDA,1996].

Se puede observar que el Data Mining puede dar valor agregado a la búsqueda de datos que son administrados en la arquitectura de información Data Warehouse, siendo explotados posteriormente por los diversos Sistemas de Información que apoyan la toma de decisiones, tales como los que se nombraron en la cita anterior, o bien los que se mencionan en la capa de entrega de información de la tabla 2.4. “El Data Mining no es una ‘muestra’ del Data Warehouse, sino más bien representa un segmento reestructurado de datos con un enfoque hacia un objetivo específico de la empresa” [PARSAYE,1996].

La evaluación y aplicación correcta de las herramientas Data Mining pueden contribuir fuertemente en el logro del objetivo de la arquitectura de información Data Warehouse, en cuanto a apoyar el proceso de toma de decisiones en los niveles táctico y estratégico de una organización. Además, el Data Mining puede generar para los usuarios de la información un mejor entendimiento del negocio, generando así una mayor posibilidad de tener un grado de adopción más alto durante el proceso de asimilación de la tecnología en la empresa. "Las herramientas Data Mining encuentran patrones en los datos e infieren reglas para guiar la toma de decisiones, el factor crítico de éxito del Data Mining es entender los problemas del negocio y la tecnología " [EDELSTEIN,1996].

3.6. ADMINISTRACION DE DATOS

Uno de los aspectos de mayor importancia en la arquitectura de información Data Warehouse es la transformación de datos a información, y para que esto suceda se requiere que los datos sean administrados de acuerdo a los tópicos de la empresa y a las decisiones que dan valor al negocio, es decir, que lo presentado a los usuarios finales del Data Warehouse sean ya no datos, sino información procesada con fácil interpretación y con valor agregado para el usuario. El Data Warehouse es una arquitectura orientada a los tópicos del negocio y por consiguiente, su base

tecnológica debe estar encaminada a administrar los datos adecuadamente para alcanzar tal objetivo. Los procesos más importantes en el diseño del Data Warehouse, tal como la sumarización, la partición, la integración de datos, etc., conforman las estrategias del Data Warehouse para poder entregar información que realmente soporte la toma de decisiones, es decir, la administración de los datos en la arquitectura Data Warehouse se basa en tales procesos. La administración de datos en el Data Warehouse tiene un enfoque estratégico debido al alcance que tiene en los más altos niveles organizacionales. Se puede decir entonces, que la administración estratégica de datos es la base de la arquitectura de información Data Warehouse, o bien, que la arquitectura de información Data Warehouse es un producto final de la administración estratégica de los datos.

"Los datos son un recurso compartido a través de toda una organización y muchas las personas involucradas en su manejo, se puede establecer que la falta de algún tipo de coordinación provocara un ambiente de caos que generara un problema global dentro de la organización. Lo anterior puede compararse con la situación de una orquesta sinfónica integrada por músicos de gran talento que no cuenta con un director para combinar su trabajo; el resultado será poco placentero" [GILLENSON,1988]. Tal coordinación es un objetivo que se persigue al diseñar la arquitectura de información Data Warehouse, esto es, convertir al Data Warehouse en el director de la administración de datos en una organización, y así lograr esa sincronía que la orquesta necesita para promover la creación de una empresa basada en información y conocimiento.

Se visualizan algunos objetivos de la administración de datos: "Shah puntualiza que el objetivo de la administración de datos es manejar los datos como un recurso corporativo, al igual que un contralor lo haría con el recurso financiero: el dinero" [SHAH,1984]. "La administración de datos puede satisfacer la necesidad de coordinación y de integración para una efectiva administración desde una perspectiva corporativa" [RAVINDRA,1986]. "Gillenson explica que la administración de datos se refiere a la custodia, la administración, la planeación y la documentación de los datos de una empresa". [GILLENSON,1988]. La administración de datos puede ser analizada desde diversos puntos de vista y conforme a la aplicación que se le va a dar en una empresa. "Un esfuerzo efectivo de administración de datos no presenta un patrón claro y definido en cuanto a las acciones a seguir, y se observa una dependencia muy fuerte en aspectos organizacionales" [GOODHUE,1988].

Para fines de esta tesis la administración de datos se considera como parte de la arquitectura Data Warehouse; desde la transformación de datos operacionales a los datos históricos del Data Warehouse, hasta la creación de los Data Marts o de las bases de datos altamente sumarizadas, todos los procesos implican una

administración estratégica de datos. Además, la administración de datos para una arquitectura de información como el Data Warehouse requiere de un esfuerzo para librar obstáculos y prácticas organizacionales que afectan directamente a la calidad y el manejo de los mismos en una empresa. Una administración estratégica de datos en el Data Warehouse que rompa barreras organizacionales, permitirá una asimilación más exitosa de la arquitectura de información en la empresa.

En la tabla 3.2 se presenta un marco conceptual de la administración de datos dividido en 4 categorías, se analizan brevemente y de manera general para identificar más claramente el papel de la administración de datos en la arquitectura de información Data Warehouse.

Objetivo del negocio	Alcance	Proceso de Planeación de datos	Producto
• Coordinación operacional	• Corporativo	• Modelación estratégica de datos	• Datos operacionales orientados a los tópicos
• Flexibilidad organizacional	• Divisional	• Orientado a una área o proceso del negocio	• Sistemas comunes
• Mejorar la información administrativa	• Funcional	• Ninguno	• Bases de datos de información
• Eficiencia en los sistemas de información			• Servicios de acceso a datos
			• Base para una arquitectura

Tabla 3.2 - Marco conceptual de la administración de datos [GOODHUE, 1988]

Objetivo del negocio

Las acciones concernientes a la administración de datos no son justificadas por argumentos conceptuales o técnicos, sino más bien por necesidades específicas de la organización, como pueden ser: coordinación operacional, flexibilidad organizacional, mejora en la información gerencial, o bien, tener sistemas de

información más eficientes. La arquitectura Data Warehouse debe diseñarse sin perder de vista estos objetivos; en algunas empresas se tienen grandes volúmenes de datos que dificultan apoyar la coordinación de datos en los niveles operacionales, el tiempo que se lleva generar información de esas bases de datos voluminosas es muchas veces muy alto, así como lento el acceso a los datos que pudieran soportar de manera más eficiente dicha coordinación operacional. Aunque en esta tesis el alcance de la arquitectura de información Data Warehouse sea en los niveles táctico y estratégico, el nivel operacional no deja de ser importante, ya que de ahí se extraen los datos que conforman los diversos niveles de sumarización que apoyan la toma de decisiones con información gerencial y ejecutiva. La administración estratégica de datos en el Data Warehouse también puede permitir que los sistemas de información se hagan más eficientes al procesar datos de calidad y orientados a las necesidades del negocio que cada sistema satisface.

Alcance

El alcance de la administración de datos en una empresa está directamente relacionado a los límites de aplicación del Data Warehouse; quizá a corto plazo solamente se diseñe la arquitectura de información para cubrir una sola área funcional de la empresa, ya sea mediante el diseño de Data Marts o bien mediante el uso de sistemas de soporte a la decisión que exploten datos directamente de las bases de datos del Data Warehouse. El objetivo a largo plazo será integrar los Data Marts para crear un Data Warehouse corporativo que apoye cualquier requerimiento de información ejecutiva de la alta administración.

Proceso de planeación de datos

Las acciones referentes a la administración de datos pueden ser guiadas por un modelo estratégico de datos a nivel corporativo, o bien, apoyarse solamente en una representación conceptual de datos de una área funcional o proceso del negocio, o quizá no emplear ningún modelo de datos. En el caso del Data Warehouse, algunos autores [LADAGA,1995], [INMON,1996], recomiendan el desarrollo de un modelo de datos a nivel corporativo y a nivel funcional, que dirijan las estrategias de sumarización y partición de datos y demás procesos del Data Warehouse. Además, la planeación de datos en la arquitectura Data Warehouse es importante para conformar el Metadata, ya que de ahí se deriva el seguimiento a los datos y el empleo de los mismos en cada proceso del Data Warehouse.

Producto

Los productos que aparecen en la tabla 3.2 son todos aplicables a la arquitectura de información Data Warehouse; el Data Warehouse está orientado a los tópicos del negocio y mediante una adecuada estructuración de sus bases de datos puede apoyar la creación de las bases de datos operacionales orientadas a una área o función de la empresa (SADB - Subject Area Data Base); mediante la integración y estandarización de los datos el Data Warehouse puede lograr que diversos sistemas de información compartan los mismos datos sin importar su origen y formato; en la arquitectura Data Warehouse las bases de datos de información se representan con datos altamente resumizados y con valor para la toma de decisiones; los servicios de acceso a los datos deben estar disponibles en cualquier nivel de sumariación de la arquitectura de información, y para tener accesos más rápidos y específicos se deben buscar aplicar estrategias de partición de datos que agrupen a la información de acuerdo a los diversos tipos de consultas y accesos. Por último, la misma arquitectura de información Data Warehouse ya definida y establecida representa un producto de la administración estratégica de los datos en una empresa.

3.7. SISTEMAS DE SOPORTE A LA DECISION

Ya no basta en la actualidad con tomar decisiones por intuición o experiencia, ya que los mercados se han vuelto más competitivos y una decisión mal tomada puede dejar a una empresa fuera de éstos. Como consecuencia de la competencia y la globalización de los mercados, generalmente se dan más alternativas de solución de las que se daban en otros tiempos ante un determinado problema. "Es más difícil tomar decisiones por 2 razones. Primero, el número de alternativas disponibles es mayor ahora que antes, a consecuencia de los avances de la tecnología y los sistemas de comunicación. Segundo, el costo de cometer errores puede ser mucho mayor debido a la complejidad y magnitud de las operaciones, la automatización y la reacción en cadena que puede provocar un error en muchas partes de la organización" [TURBAN,1995].

"La toma de decisiones constituye el núcleo de la actividad gerencial, podemos definir la 'decisión' como la elección de un curso de acción determinado entre varios cursos de acción posibles" [LAZZATI,1990]. La tecnología no solamente ha dado mayores alternativas de solución a un problema, sino también ha ayudado a los administradores a seleccionar las más adecuadas entre ellas. "Los administradores

deben ser mas sofisticados y aprender como usar esas nuevas herramientas y técnicas que están siendo desarrolladas en el campo de la toma de decisiones" [TURBAN,1995]. Una de estas herramientas y quizá la más importante son los sistemas de soporte a la decisión (DSS).

"El concepto de DSS fue pronunciado por primera vez a principios de la década de los 70`s por Scott-Morton bajo el término de Sistema de administración de decisiones, definiéndolo como un sistema interactivo basado en computadora, el cual ayuda a los que toman las decisiones a utilizar datos y modelos para resolver problemas semiestructurados y no estructurados" [TURBAN,1995]. Otra definición que da el mismo autor es la siguiente: "Un sistema de soporte a la decisión acopla los recursos intelectuales de los individuos con las capacidades de la computadora para mejorar la calidad de las decisiones, es un sistema de soporte basado en computadora para los tomadores de decisiones que tratan con problemas semiestructurados" [TURBAN,1995]

En estas definiciones se observan 2 aspectos muy importantes en la creación de la arquitectura de información Data Warehouse, por un lado, la creación de modelos de datos y por otro la tendencia a resolver problemas semiestructurados o no estructurados. Ambas características hacen de los sistemas de soporte a la decisión parte fundamental en el diseño de la arquitectura de información Data Warehouse. Los DSS principalmente se identifican en la capa de administración de datos de la arquitectura Data Warehouse, debido a que influyen directamente en las estrategias de sumariación y partición de datos. De los análisis de decisiones que se realicen de tales sistemas se generan modelos de datos estadísticos o matemáticos que muchas veces requieren el uso de datos de diversas fuentes en la empresa; los DSS pueden obtener sus datos ya sea de las bases de datos operacionales, de las bases de datos históricas del Data Warehouse, o bien, de los Data Marts u otras fuentes de datos ya sumariadas.

El Data Warehouse involucra extraer datos de diversas bases de datos operacionales y así generar una nueva base de datos de soporte a la decisión. "El Data Warehouse es la base de los Sistemas de Soporte a la Decisión" [RADDING,1995]. En esta cita se puede observar la importancia de estudiar estos dos conceptos de manera paralela y relacionarlos óptimamente cuando se llegue al desarrollo de la arquitectura de información, principalmente para que el Data Warehouse y los Sistemas de Soporte a la Decisión contribuyan de manera conjunta a mejorar el proceso de toma de decisiones en una organización.

3.7.1. EL DATA WAREHOUSE Y EL PROCESO DE TOMA DE DECISIONES

Para entender mejor la relación entre el Data Warehouse y los Sistemas de Soporte a la decisión en la creación de la arquitectura de información, se va a usar un modelo de soporte al proceso de toma de decisiones. “El tomar decisiones es un proceso que consiste en seleccionar un curso de acción entre diversas alternativas, con el propósito de lograr uno o varios objetivos” [TURBAN,1995]. Este proceso se presenta en muchas de las actividades organizacionales e incluso personales, en una empresa se tiene como ejemplo la muy importante función administrativa de planeación, la cual involucra una serie de decisiones, principalmente en lo que se refiere a la asignación de recursos, la determinación de responsabilidades, la programación de actividades, etc. Así como la planeación, existen otra serie de funciones administrativas que requieren generar decisiones para llevarlas a cabo. A continuación se presenta un modelo que describe el proceso de toma de decisiones, teniendo como objetivo principal analizarlo para establecer su relación con la tecnología Data Warehouse. [TURBAN,1995]

I) Fase de conocimiento

- Objetivos organizacionales
- Búsqueda y exploración de procedimientos
- Colección de datos
- Identificación del problema
- Clasificación del problema
- Formulación del problema

II) Fase de diseño

- Formulación del modelo
- Establecimiento de criterios de selección
- Búsqueda de alternativas
- Predicción y medición de resultados

III) Fase de selección

- Solución del modelo
- Análisis de sensibilidad
- Selección de la(s) mejor(es) alternativa(s)
- Plan de implantación
- Diseño de un sistema de control

IV) Implantación de la solución

I) Fase de Conocimiento

La fase de conocimiento es importante para la identificación del problema, si es que éste verdaderamente existe, sus síntomas, su magnitud y la formulación del mismo. La identificación de los objetivos organizacionales pertenecientes a esta fase puede estar apoyada por el estudio de una herramienta de análisis conocida como la cadena de medios-fines [LAZATTI,1990], la cual delimita la información y los datos requeridos para apoyar al proceso de toma de decisiones, permitiendo así a los diseñadores del Data Warehouse identificar la información que debe generarse en esta fase del modelo. La cadena de medios-fines será analizada con mayor detalle en el siguiente capítulo de esta tesis.

"La existencia de un problema en una organización puede ser estimada monitoreando y analizando el nivel de productividad de un departamento o de una organización. La medición de la productividad y la construcción del modelo en la fase de diseño están basados en datos, la colección de datos existentes y la estimación de datos futuros es uno de los pasos más difíciles en este análisis" [TURBAN,1995]. En esta fase del modelo los analistas de decisiones pueden obtener de los procesos del Data Warehouse los datos necesarios para identificar, clasificar y formular el problema en análisis, y fomentar a la vez un mecanismo de retroalimentación entre los diseñadores del Data Warehouse y los analistas de decisiones para actualizar eficientemente los procesos existentes en la arquitectura de información, de acuerdo a los cambios en las necesidades de información de la empresa.

II) Fase de diseño

La fase de diseño involucra la generación, el desarrollo y el análisis de posibles cursos de acción, siendo la formulación del modelo el aspecto más importante de esta fase. "La creación del modelo involucra la conceptualización del problema y su abstracción a una forma cuantitativa o cualitativa" [TURBAN,1995].

Formulación del modelo

Una de las características de un Sistema de Soporte a la Decisión es que consta de un modelo (matemático, financiero, etc.), que permite "jugar" con variables de decisión para poder medir y predecir resultados. Las variables de decisión describen los diversos cursos de acción para una situación. La estructura general de un modelo es como la que se muestra en la figura 3.4.

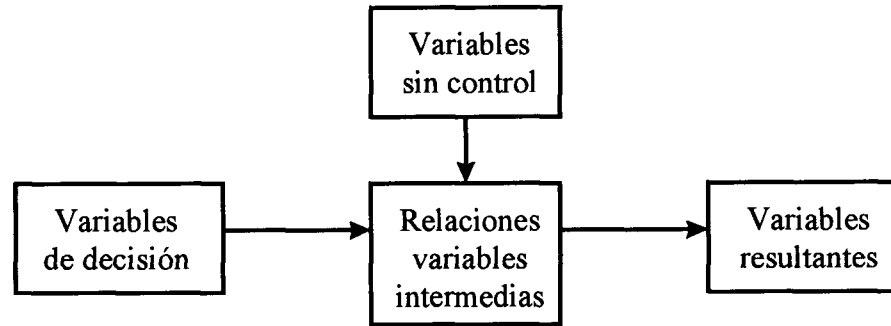


Figura 3.4 Estructura general de un modelo [TURBAN,1995]

Las variables sin control están fuera del alcance del tomador de decisiones y generalmente son variables provenientes del medio ambiente que envuelve a una situación. Las variables intermedias son aquellas que son necesarias para relacionar a las variables de decisión con las variables resultantes. Tomando como ejemplo el mantenimiento a una instalación industrial, tendríamos las siguientes variables:

- Variables de decisión
 - ⇒ Calidad en materiales
 - ⇒ Asignación de recursos materiales y humanos
- Variables intermedias
 - ⇒ Inventario de materiales (existencia en almacén)
 - ⇒ Cantidad de materiales requeridos (planeación)
- Variables sin control
 - ⇒ Precios en materiales
 - ⇒ Capacidad de la maquinaria
- Variables resultantes
 - ⇒ Eficiencia en el mantenimiento
 - ⇒ Costo total.

Ahora bien, los modelos deben ser dinámicos y flexibles para adaptarse al ambiente cambiante de los negocios. "El proceso de toma de decisiones frecuentemente requiere de un rango de datos históricos para análisis comparativos, de tal manera que los usuarios sean capaces de monitorear tendencias y modelos de información en el tiempo" [POE,1995]. La importancia de un buen diseño del Data Warehouse en el aspecto de los datos históricos se refleja en este punto, ya que a través de la orientación hacia lo histórico es posible la detección de tendencias en

los datos que permitan adecuar un modelo ante una problemática determinada. Las herramientas Data Mining pueden emplearse para predecir patrones en los datos, y así conocer con anterioridad el efecto de las variables sin control durante la formulación del modelo. "En un amplio sentido, el Data Mining significa buscar patrones de interés en las bases de datos de la empresa para ejecutar acciones que incrementen sus utilidades, reduzcan sus costos y amplíen sus mercados" [WELDON,1996]. "Las herramientas Data Mining detectan modelos en los datos e infieren reglas para guiar la toma de decisiones. Un Data Mining además mejora la capacidad de análisis en este proceso enfocando su atención a las variables de decisión más importantes" [EDELSTEIN,1996].

III) Fase de selección

Las fronteras entre las fases de diseño y de selección generalmente no están establecidas claramente, debido a que ciertas actividades pueden ser realizadas en cualquiera de las dos fases. La solución del modelo es algo que hace única a esta fase del proceso de toma de decisiones. "La solución del modelo no es lo mismo que la solución del problema que representa dicho modelo, la solución del modelo más bien, produce una recomendación para solucionar el problema" [TURBAN,1995]. Para la solución del modelo pueden utilizarse herramientas analíticas, matemáticas, algoritmos, análisis de sensibilidad, análisis "What-If", análisis Goal Seeking, etc. Los datos en esta fase pueden tomarse de las bases de datos históricas del Data Warehouse o bien de las bases de datos pertenecientes a los DSS.

IV) Implantación de la solución

Ejecutar la decisión basándose en las selecciones hechas en la fase anterior.

El modelo presentado podría representar una utopía en el proceso de toma de decisiones, ya que en su gran mayoría las decisiones se toman urgentemente y sin mucho tiempo para realizar análisis más exhaustivos. La arquitectura de información Data Warehouse puede convertirse en un canal, que mediante la administración de los datos y fundamentado en modelos como el presentado, pueda generar alternativas rápidas de solución a un problema sin perder de vista los objetivos organizacionales y las estrategias de la empresa, y sobre todo, basándose en información histórica de valor y en tendencias o patrones que se pueden obtener de las bases de datos.

3.8. SISTEMAS DE INFORMACION EJECUTIVOS

Generalmente la informática ha proporcionado un soporte muy importante a ciertas actividades de una compañía, pero este soporte ha sido muy bajo o es casi nulo en los niveles más altos de la organización, en donde se toman las decisiones y se realizan las funciones administrativas de mayor relevancia. En los últimos años los esfuerzos en hardware, software y telecomunicaciones se han dirigido a aumentar el nivel de participación de la tecnología de información en asuntos estratégicos de una organización, y a la vez en como cerrar la brecha que aún existe entre los objetivos del negocio y los objetivos de la función de informática. Entre las tecnologías que están enfocadas a este propósito se encuentran los Sistemas de Información Ejecutivos.

En las capas de la arquitectura de información Data Warehouse, los Sistemas de Información Ejecutivos se localizan en la capa de entrega de información. "Los Sistemas de Información Ejecutivos siempre han estado enfocados a proporcionar más información del negocio, a más gente, a través de toda la organización, de una manera electrónica" [HARTLEN,1995]. Ambas tecnologías, el Data Warehouse y los Sistemas de Información Ejecutivos, se necesitan mutuamente para funcionar eficazmente. "Un Data Warehouse por si solo no produce resultados mágicamente, sino que se requiere de herramientas que presenten a los datos como información útil para los usuarios" [RADDING,1995]. Una de estas herramientas son precisamente los Sistemas de Información Ejecutivos. "En su mayoría los EIS son diseñados con un front-end intuitivo y gráfico, con el fin de proporcionar a los ejecutivos el acceso a los almacenes de datos de la empresa" [BARTLETT,1995].

Ahora bien, las tecnologías por si solas dan muy buenos resultados si son llevadas eficazmente en las diversas etapas de su desarrollo e implantación, pero pueden obtenerse mejores y más beneficios de ellas si se tiene la habilidad de relacionarlas e implantarlas en conjunto. La sinergia puede ser aplicada cuando se habla de tecnologías, y lo es aún mas cuando a ésta le son agregados aspectos humanos y organizacionales que tienen que ver con la información que manejan los ejecutivos y gerentes para realizar sus funciones diarias. La arquitectura de información Data Warehouse consiste de una serie de procesos que 'preparan' los datos que los EIS requieren para presentarlos intuitiva y gráficamente. Por otro lado, un EIS requiere de una base de datos propia que contenga los datos requeridos para su funcionamiento. "Los EIS generalmente tienen su propia base de datos, la cual es formada por los datos de los diversos sistemas de información de la empresa" [SHANBAUM,1996]. Esta base de datos no es el Data Warehouse en sí, sino que éste ayuda a crear y formar la base de datos del EIS. "Los EIS pueden crear las

interfaces de una manera fácil, pero las estructuras de los datos son siempre complicadas" [BARTLETT,1995].

No se quiere decir que es más importante la estructura de los datos que la manera en que se presenta la información, más bien, se pretende establecer un equilibrio adecuado entre los EIS y el Data Warehouse para que ambas tecnologías se alineen a los objetivos del negocio. Este equilibrio puede lograrse estableciendo un mecanismo de retroalimentación en ambos sentidos, entre los EIS's y la arquitectura de información Data Warehouse; por un lado la tendencia en los datos permitirá que los analistas de decisiones planeen nuevas formas de análisis de la información que den más valor al proceso de toma de decisiones, y por otro lado, los cambios organizacionales y el medio ambiente competitivo obligan a los ejecutivos a requerir otro tipo de información, lo cual debe reflejarse en el EIS y por consiguiente en los procesos Data Warehouse que preparan la información. La figura 3.5 nos muestra diversas maneras en que el Data Warehouse soporta a un Sistema de Información Ejecutivos.

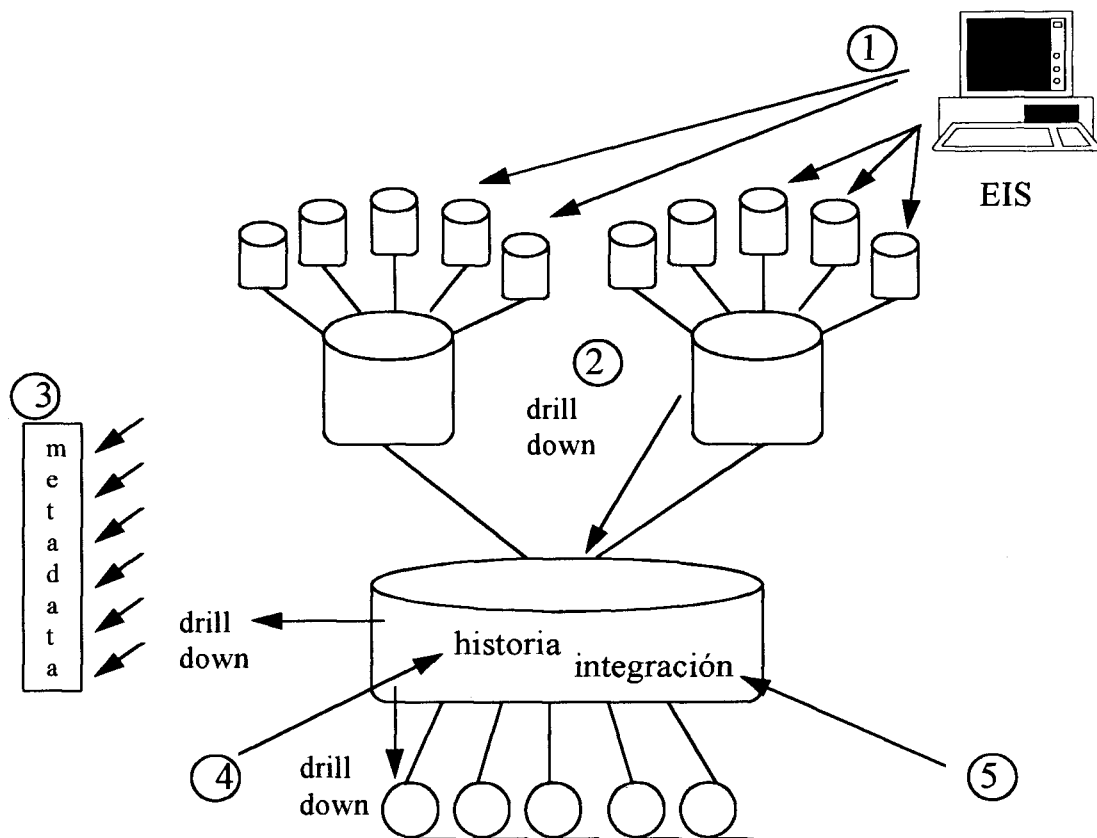


Figura 3.5 - Maneras en que el Data Warehouse soporta a un EIS [INMON,1996]

1) Hay una alta disponibilidad de datos sumariados

Las bases de datos de los Sistemas de Información Ejecutivos se alimentan principalmente de las bases de datos altamente sumariadas de la arquitectura de información Data Warehouse, las cuales se planearon para apoyar la entrega de información en los niveles ejecutivos. Es importante que los procesos del Data Warehouse que generan tal información sumariada se realicen conforme a los tiempos requeridos, para así poder tener siempre actualizadas las bases de datos de los EIS y con ello poder presentar información verídica rápidamente.

2) La estructura de los datos en el Data Warehouse soporta los procesos drill-down

Los niveles de sumariación de los datos en la arquitectura de información Data Warehouse apoyan los procesos drill-down de los EIS, de tal manera que si el ejecutivo desea conocer información con más detalle lo pueda hacer empleando el mismo EIS, mediante módulos sencillos de acceder y que estén directamente enlazados a las bases de datos que manejan determinado nivel de detalle en los datos, siendo transparente este proceso a los ejecutivos.

3) El Metadata permite a los analistas de decisiones planear el desarrollo de nuevas iniciativas

El Metadata contiene los datos acerca de los datos (ver 2.7.), y entre éstos se encuentran las estructuras de los datos entendibles a los analistas de decisiones, los catálogos de información, y todos los procesos Data Warehouse en la arquitectura de información. Al observar el estado de los datos y su alcance, los analistas de decisiones pueden proponer nuevas formas de análisis y reportes, así como otros tipos de accesos que mejoren el proceso de toma de decisiones, o bien, la tendencia en los datos puede permitir a los analistas de decisiones el proponer iniciativas que contribuyan a que determinadas áreas o procesos de la empresa tengan un rendimiento mayor. Dichas propuestas finalmente deben verse reflejadas en el diseño de los Sistemas de Información Ejecutivos.

4) La datos históricos en el Data Warehouse soportan el deseo de la administración de tener nuevas tendencias analíticas en el EIS

La creación de datos históricos en el Data Warehouse permite a los analistas de decisiones y a ejecutivos realizar análisis comparativos de los datos en diversos período de tiempo. De los resultados de esos análisis se pueden generar nuevas

tendencias en el análisis de la información al identificarse patrones en las bases de datos; asimismo se pueden emplear herramientas Data Mining para explotar más eficientemente las bases de datos históricas del Data Warehouse.

5) La integración de los datos antes de introducirse en el Data Warehouse, es lo que la administración necesita para poder mirar a través de la empresa.

Antes de introducirse en el Data Warehouse, los datos están altamente detallados, por lo cual los procesos de integración de los datos deben guardarse en el Meta Data, para que en un momento determinado si un ejecutivo desea conocer cierta actividad de la empresa más profundamente pueda recurrir a los datos detallados como consulta. Conocer de qué manera los datos fueron integrados va a permitir mejorar el tiempo de acceso en este tipo de consultas.

De entre todos los componentes de la arquitectura de información Data Warehouse, los Sistemas de Información Ejecutivos son la herramienta informática que va a estar más cerca de la alta administración, por lo cual, de la eficiencia de estos sistemas dependerá el grado de aceptación de la arquitectura Data Warehouse en los niveles estratégicos de una empresa, y por consiguiente, del apoyo requerido para su implantación en la organización.

3.9. CONCLUSIONES

Se han analizado en este capítulo algunos de los componentes que conforman la arquitectura de información Data Warehouse que propongo en esta tesis, haciendo un énfasis especial en la administración de los datos y en el uso de herramientas de software que permitan explotar la información de las bases de datos del Data Warehouse. Al analizar cada uno de los componentes y su función en la arquitectura de información Data Warehouse se puede visualizar que todos persiguen un objetivo común: administrar estratégicamente los datos para mejorar el flujo de la información en la empresa, y apoyar de esta manera al proceso de toma de decisiones. Al tener este objetivo, el Data Warehouse cumple con la meta de toda arquitectura de información desde un punto de vista tecnológico.

El análisis de cada tecnología de información presentada en este capítulo estuvo dirigido a establecer su relación con el Data Warehouse y principalmente en determinar su función en el diseño de la arquitectura de información. Las

tecnologías presentadas son las que considero deben incluirse en cualquier diseño de un Data Warehouse para realizar los procesos que van desde la transformación de los datos hasta la presentación de la información a los usuarios finales. Hay que tomar en cuenta la velocidad de introducción de nuevas tecnologías, por lo cual resulta importante estar evaluando constantemente el mercado informático para determinar aquellas tecnologías de información que pudieran optimizar el funcionamiento de la arquitectura Data Warehouse.

Tecnológicamente el Data Warehouse comprende aspectos técnicos de base de datos y software, pero como en toda tecnología, se requiere la participación activa del recurso humano para lograr su asimilación en una organización. De los tres elementos que interactúan en la analogía de arquitectura física y arquitectura de información, no se ha contemplado todavía la que corresponde a las capacidades y sentidos humanos que crean patrones y líneas a seguir en el desarrollo de una estructura arquitectónica. Para la creación de la arquitectura de información Data Warehouse propongo la participación de personas con actividades funcionales diversas, es necesario que dicho personal cuente con una cultura de información lo suficientemente alta para aceptar cambios en la manera de realizar sus actividades, sobre todo en lo que se refiere a la compartición de los datos; pero no solamente se espera del recurso humano que sea receptor de las nuevas tecnologías, sino también que sea promotor y participante activo y comprometido con las innovaciones tecnológicas a implantarse en su empresa. De este compromiso por parte de los miembros de la empresa depende mucho la asimilación de la tecnología Data Warehouse, y más aún, cuando a esta tecnología se le visualiza como una arquitectura de información que puede abarcar a toda una organización. A final de cuentas el usuario es el que va a usar la nueva tecnología y si su grado de aceptación es bajo, el rechazo hacia la misma va a impactar negativamente al momento de la implantación.

Precisamente para fines del análisis del proceso de asimilación se presentó el estudio detallado de los componentes de la arquitectura de información Data Warehouse, por un lado para ayudar a visualizar globalmente a la organización mediante el diseño de la arquitectura, y por otro lado mediante el diseño del mapa de datos identificar barreras organizacionales, tecnológicas o humanas que no permitan esa tarea tan esencial en el diseño de un Data Warehouse: la administración de los datos.

Para poder aplicar un proceso de asimilación de tecnología se requiere conocer muy bien tanto a la *empresa* como a la *tecnología*, además se deben identificar las *prácticas organizacionales del recurso humano* que afectan a la operación de ambas. Si se logra conjuntar el conocimiento de la empresa, de la tecnología, su alineación, y las prácticas organizacionales relacionadas con ellas, permitirá que la

aplicación del proceso de asimilación tecnológica del Data Warehouse tenga mayores posibilidades de éxito. El capítulo 2 nos permitió conocer del Data Warehouse y sus aspectos de diseño más importantes, este capítulo nos introdujo en el conocimiento de los componentes del Data Warehouse desde un punto de vista de una arquitectura de información. En el siguiente capítulo propongo analizar a la empresa como un todo, identificando sus estructuras, procesos y flujos de información, y de esta manera estar en condiciones de establecerse ligas entre la arquitectura Data Warehouse y la organización, asimismo el capítulo siguiente va a sentar las bases del proceso de asimilación de tecnología que va permitir proponer un proceso de asimilación del Data Warehouse en una organización y las medidas de evaluación para determinar su grado de adopción como una herramienta de productividad para la toma de decisiones.

REFERENCIAS BIBLIOGRAFICAS

- [BARTLETT,1995] BARTLETT, Jeffrey.
"Executive systems ripe for makeover -white elephants
- emerge as prime target for open systems vendors with
client-server solutions".
CMP Publications, 1995
- [DePOMPA,1996] DePOMPA, Barbara.
"There`s gold in Databases".
Information Week, Enero 8, 1996
- [EDELSTEIN,1996] EDELSTEIN, Herb.
"Mining Data Warehouses".
Information Week, Enero 8, 1996
- [GILLENSON,1988] GILLENSON, Mark L.
Introducción a las bases de datos.
McGraw-Hill. USA. 1988. cap. 13
- [GOODHUE,1988] GOODHUE, Dale L., Judith A. Quillard & John F.
Rockart.
"Managing the Data Resource: A contingency
perspective".
MIS Quartely, vol. 12, no. 3, Septiembre 1988
- [HACKATHORN,1995] HACKATHORN, Richard.
"Data Warehousing energizes your enterprise".
Datamation, Febrero 1, 1995
- [HARTLEN,1995] HARTLEN, Brian.
"A revolution in EIS/DSS".
CMP Publications, Noviembre 1995
- [INMON,1996] INMON, W.H.
Building the Data Warehouse.
John Wiley & Sons, Inc, segunda edición, 1996

- [KORTH,1991] KORTH, Henry F.
Database System Concepts
McGraw-Hill, segunda edición, 1991
- [KRIVDA,1996] KRIVDA, Cheryl D.
"Unearthing underground data".
LAN, Mayo 1996
- [LADAGA,1995] LADAGA, John.
"Let business goals drive your data warehouse effort".
Health Management Technology, Octubre 1995.
- [LAZAR,1996] LAZAR, Bill.
"The Data Breakthrough".
LAN, Mayo 1996
- [LAZZATI,1990] LAZZATI, Santiago.
Claves de la decisión en la empresa: Método y participación.
Ed. Macchi, segunda edición, 1990
- [MCFADDEN,1994] Mc. FADDEN, Fred R.
Modern Database Management.
Addison Wesley Publishing, 1994
- [MCGEE,1993] MCGEE, James & Laurence Prosak.
Managing information strategically.
John Wiley & Sons, 1993
- [PARSAYE,1996] PARSAYE, Kamran.
"Data Mines for Data Warehouses".
Database Programming and Design, suplemento
Septiembre 1996
- [POE,1995] POE, Vidette.
"Data Warehouse architecture is not infrastructure".
Database Programming & Design, Julio 1995
- [RADDING,1995] RADDING, Alan.
"Support Decision makers with a Data Warehouse".
Datamation, Marzo 15, 1995

- [RAVINDRA,1986] RAVINDRA, Palangala S.
"Data administration: An old function adopts to its new role".
Journal of Information systems management. Otoño 1986, pp. 47-51
- [ROBINSON,1996] ROBINSON, Teri.
"To avoid pitfalls, walk before you run".
Software Magazine, Center Market Research, 1996
- [SHAH,1984] SHAH, Arvind D.
"Data administration: it`s crucial".
Datamation, Enero 1984
- [SHANBAUM,1996] SHANBAUM, Suzanne.
"The ABCs of EIS".
Database Programming and Design, Febrero 1996
- [TURBAN,1995] TURBAN, Efraim.
Decision Support and Expert Systems.
Prentice Hall Inc., pp 139-142. Cap. 1, 1995
- [WELDON,1996] WELDON, Jay-Louise.
"Data Mining and visualization".
Database Programming and Design, Mayo 1996

CAPITULO 4

ANALISIS DE LA EMPRESA Y SU ALINEACION A LA ARQUITECTURA DE INFORMACION DATA WAREHOUSE

4.1. INTRODUCCION

Para que una innovación tecnológica se aplique exitosamente en una empresa se requiere su total asimilación y entendimiento entre los miembros de la misma. El conocimiento de la tecnología y de la empresa, y el análisis del grado de interacción y alineación entre ambas, establecen en gran parte las actividades a seguir durante el proceso de asimilación tecnológica del Data Warehouse que propongo en esta tesis. En los capítulos anteriores hice hincapié en entender al Data Warehouse como una arquitectura de información, formada por diversos componentes y procesos, en este capítulo propongo un *análisis de información* dirigido a entender y conocer a la empresa de tal manera que la innovación Data Warehouse se ajuste a sus estructuras y funciones, disminuyendo así la necesidad de realizar cambios en las operaciones del negocio que pudieran afectar negativamente el grado de adopción de la nueva tecnología. El conocimiento de la empresa y su relación con la arquitectura de información Data Warehouse va a permitir identificar la participación del recurso humano en tal relación, siendo la involucración del recurso humano el aspecto más importante de un proceso de asimilación, ya que finalmente son ellos los que van a usar la tecnología, y en su caso los que van a decidir si la aceptan o la rechazan.

“La asimilación de tecnología requiere de una administración innovadora, así como de darse cuenta de que la tecnología no se vende por sí sola. Para que el proceso sea exitoso se requiere: involucrar a los usuarios; fomentar su participación; aplicar la nueva tecnología a problemas piloto; adaptar la tecnología para que sea accesible a los usuarios; proveer entrenamiento y educación formal; dar seguimiento a la efectividad del proceso de difusión; promover las reuniones entre usuarios para el intercambio de ideas; no confiar únicamente en los reportes escritos; proveer recursos como tiempo, dinero, personas, etc. para la realización de la difusión y asimilación de la nueva tecnología; y considerar la transferencia de personas junto con la tecnología” [WOLFF,1989].

Introducir una innovación tecnológica como el Data Warehouse involucra cambiar la manera de hacer ciertas actividades dentro de una organización, y es mediante el proceso de asimilación que se pretende administrar dichos cambios para lograr sensibilizar al recurso humano en la implantación y uso de la nueva tecnología. Desde un punto de vista del Data Warehouse la transferencia y participación del recurso humano en un proyecto de este tipo representa el elemento que faltaba considerar en la analogía entre arquitectura física y arquitectura de información, esto es, lo que corresponde a las capacidades y sentidos humanos que crean patrones y líneas a seguir en el desarrollo de una estructura arquitectónica (ver apartado 3.2.). En el ambiente Data Warehouse, el usuario final representa a los ejecutivos, los analistas de decisiones y otros usuarios finales de la información en los niveles táctico y estratégico de una empresa, es parte importante del diseño del Data Warehouse la retroalimentación entre estos usuarios finales y los diseñadores del Data Warehouse, para determinar las estrategias de los diversos procesos que van a generar la información que soporte el proceso de toma de decisiones, esto es, a través de la retroalimentación el usuario final moldea el diseño del Data Warehouse de acuerdo a sus necesidades y a los objetivos globales de la organización.

Además de fomentar una adecuada retroalimentación entre los usuarios y los diseñadores del Data Warehouse para identificar las necesidades de información, es importante canalizar los esfuerzos en el diseño e implantación del Data Warehouse a través de un plan informático que defina prioridades y establezca la alineación estratégica entre el negocio y el Data Warehouse. Así en este capítulo se presenta un procedimiento de alineación TI-negocio que puede servir como modelo para definir las necesidades de información oportuna que apoyen tanto a las estrategias de la empresa como a las funciones departamentales. Si se logra aplicar adecuadamente un modelo de alineación como el que se presenta en este capítulo, los beneficios que se obtengan de la implantación de un Data Warehouse van a impactar positivamente tanto en los niveles donde se definen las estrategias, como en general en los diversos departamentos o áreas funcionales de una organización, generando finalmente la aceptación entre los miembros de la empresa hacia la nueva tecnología.

4.2. TIPOS DE ANALISIS DE LA EMPRESA PARA FINES DE LA ARQUITECTURA DE INFORMACION DATA WAREHOUSE

El objetivo de realizar un análisis de la empresa en paralelo con el diseño del Data Warehouse consiste en poder crear un mapa congruente de los datos que permita detectar barreras organizacionales en la comunicación interdepartamental, la

compartición de los datos, y demás aspectos que afectan la asimilación de la tecnología Data Warehouse en la organización. “La barrera más común para la asimilación de tecnología es el conflicto entre los diferentes grupos funcionales. La diferencia en actitudes y valores generalmente inhibe la comunicación requerida para una adecuada asimilación de tecnología entre las funciones organizacionales” [BADAWEY,1989].

Los conflictos interfuncionales o interdepartamentales pueden presentarse de manera vertical (entre niveles jerárquicos) o bien, de manera horizontal (lateralmente entre áreas funcionales), resultando importante entender e identificar la naturaleza de ambos para establecer mecanismos que permitan asimilar correctamente la arquitectura Data Warehouse. El recurso humano en muchas ocasiones es el centro de un conflicto interfuncional, siendo el análisis de la empresa un medio para identificar responsabilidades tanto en el manejo de los datos como en la operación de los diversos procesos que vayan a aplicarse en la arquitectura Data Warehouse.

Se identifican 3 tipos de análisis que pueden realizarse en una empresa para fines del Data Warehouse, *el flujo de información en forma vertical*, *el flujo de información en forma horizontal*, y *la relación datos-procesos*. De acuerdo a la definición del término de arquitectura de información presentado en el apartado 3.1, éste constituye un modelo para la creación de un mapa congruente de los datos (relación datos-procesos) que permite el flujo de información en toda la organización. De ahí el interés de relacionar el estudio de esos 3 tipos de análisis con la implantación de la arquitectura de información Data Warehouse, de tal manera que se puedan reflejar los resultados de tales análisis en el proceso de asimilación de esta tecnología de información.

Los 3 tipos de análisis mencionados serán tratados en los siguientes apartados empleando metodologías o herramientas de soporte a las tecnologías de información, haciendo énfasis en el papel del recurso humano y en la importancia de visualizar a la empresa de manera integrada por medio de la información. La información es un medio para lograr que una organización se mantenga integrada en sus procesos o áreas funcionales, es decir, es un medio para lograr que las relaciones interdepartamentales o interfuncionales de la empresa puedan llevarse a cabo sin fricciones, de manera fluida, y enfocadas a los objetivos del negocio. “La información es dinámica, capaz de crear valor y es el pegamento que mantiene a las empresas integradas” [MCGEE,1993].

4.3. FLUJO DE INFORMACION VERTICAL

La arquitectura de información Data Warehouse consta de diversos niveles de sumariación y partición en los datos, creados de acuerdo a cada medida o t3pico del negocio establecido por los analistas de decisiones; cada nivel de sumariación impacta en alg3n nivel jer3rquico de la empresa, por lo que es importante conocer el tipo de informaci3n que necesitan los usuarios en cada nivel organizacional, para as3 definir las estrategias de sumariaci3n y partici3n de los datos m3s adecuadas para cada necesidad. El an3lisis del flujo de informaci3n vertical de una empresa para fines del Data Warehouse consiste precisamente en definir las necesidades de informaci3n en cada nivel jer3rquico de la organizaci3n.

4.3.1. Cadena de Medios-Fines

Para realizar el an3lisis del flujo de informaci3n vertical se emplea la *herramienta de la cadena de medios-fines* [LAZZATI,1990], la cual consiste en identificar y relacionar objetivos de distintos niveles organizacionales, generar alternativas de soluci3n a un problema, pasar de lo general a lo particular, identificar interrelaciones y delimitar la informaci3n que debe reunirse para determinado objetivo. La figura 4.1 presenta un ejemplo de aplicaci3n de esta herramienta.

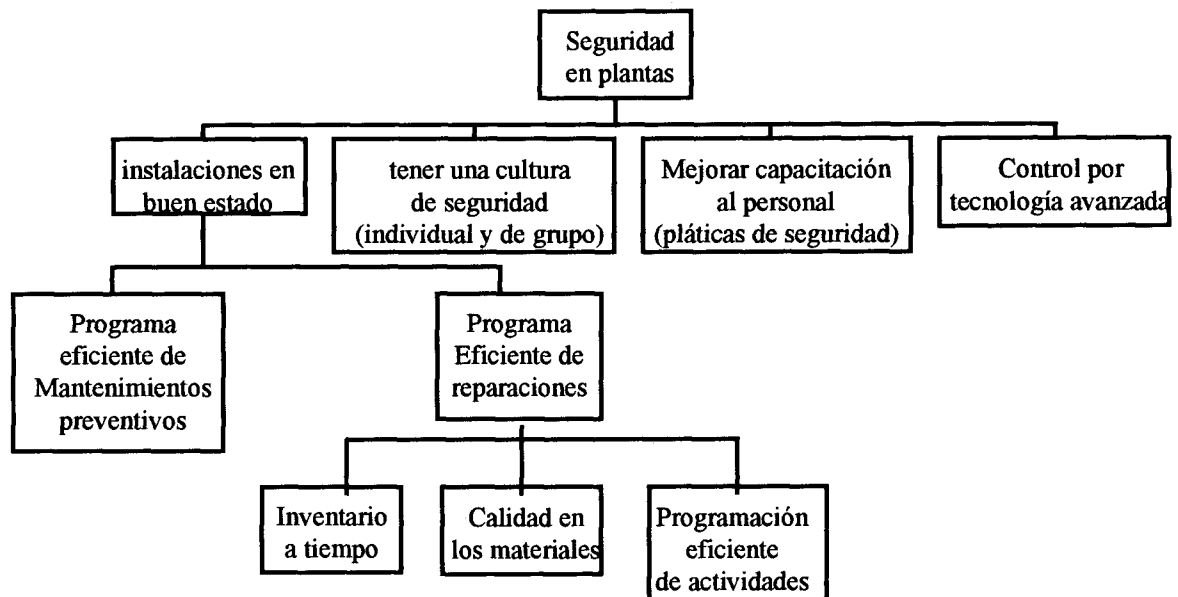


Figura 4.1 - Cadena de Medios-Fines

Como se observa en el ejemplo, toda actividad es un medio para lograr un fin y a la vez es un fin que puede lograrse a través de varios medios. Aunque parece simple este razonamiento y en general esta herramienta, llevarla a cabo con éxito puede convertirla en una herramienta de análisis de gran utilidad en la toma de decisiones. "En síntesis, la elaboración de una cadena de medios-fines se basa en la formulación de 2 preguntas: ¿Para qué? a efectos de ir hacia arriba y ¿Cómo? a efectos de ir hacia abajo, y no es tan importante definir por donde empezamos el análisis, sino donde lo delimitamos tanto arriba como abajo de la cadena" [LAZZATI,1990].

Para fines de esta tesis se propone a la cadena de medios-fines como una herramienta de análisis de información vertical que soporta el diseño de la arquitectura de información Data Warehouse. Se ha comentado que uno de los aspectos más importante en el diseño de un Data Warehouse es la sumarización de los datos, la cual va siendo mayor cuando la información se acerca a los niveles estratégicos de una organización, de tal manera que la sumarización en el nivel operacional es mínima o nula y alto el nivel de detalle de los datos, en este caso se estaría hablando de las bases de datos de los sistemas transaccionales; por otro lado, en el nivel estratégico los datos están altamente sumarizados y es bajo el nivel de detalle, aquí se estaría haciendo referencia a la base de datos de un Sistema de Información Ejecutivo. Tomando como base el ejemplo de la cadena de medios-fines mostrada en la figura 4.1, el nivel de sumarización de los datos podría representarse como se muestra en la figura 4.2.

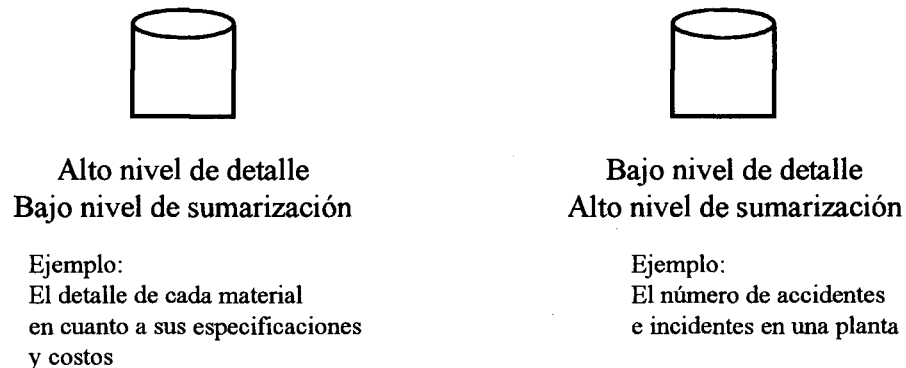


Figura 4.2 - Nivel de sumarización en el Data Warehouse

Un beneficio que puede obtenerse del análisis de la cadena de medios-fines para fines de diseño de la arquitectura de información Data Warehouse es la identificación de la información requerida en las alternativas de solución a un problema, base para la creación del modelo de un Sistema de Soporte a la Decisión. En el ejemplo de la figura 4.1 se observa el desglose del objetivo general de una empresa (seguridad en las plantas) en objetivos más particulares, lo cual permite ir definiendo objetivos secundarios que son generadores de alternativas de solución al momento de tomar una decisión. La generación de alternativas de solución a un problema puede soportar también la aplicación de los procesos Data Mining que se realicen en las bases de datos históricas del Data Warehouse. El análisis del flujo de información en forma vertical empleando la cadena de medios-fines puede apoyar también los procesos drill-down en los Sistemas de Información Ejecutivos (ver apartado 3.8.), siendo este aspecto el más importante del análisis de la cadena de medios-fines al diseñar la arquitectura de información Data Warehouse.

Los procesos drill-down de un Sistema de Información Ejecutivos son base para establecer la relación entre el Data Warehouse y la cadena de medios-fines; esta relación se puede presentar en los niveles de sumarización de la arquitectura Data Warehouse con respecto a los niveles de los objetivos analizados en la cadena de medios-fines. La información que delimita cada objetivo de la organización puede ser la base para establecer las estrategias de sumarización y partición en la arquitectura de información Data Warehouse (Ver Figura 4.3). En esta figura se pueden observar varios de los componentes de la arquitectura de información Data Warehouse, como son los EIS, los Data Marts (base de datos del sistema de planeación de cursos de seguridad), los DSS (sistema de planeación de cursos de seguridad), herramientas Data Warehouse (filtros de pláticas de seguridad), y las bases de datos operacionales (base de datos del sistema de información de pláticas). Con la herramienta de la cadena de medios-fines se pretende por una parte conocer el flujo de información vertical en una empresa, y por otro lado, determinar el diseño de cada nivel de sumarización del Data Warehouse, y los procesos y componentes de la arquitectura de información para la creación de cada nivel de datos.

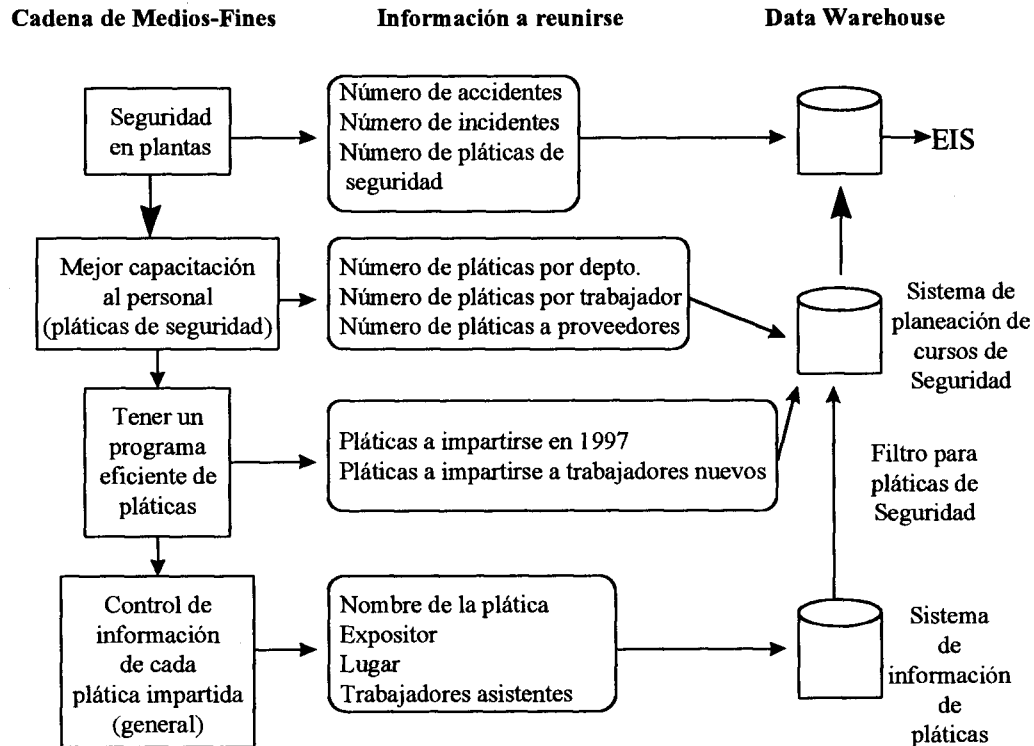


Figura 4.3 - Relación Data Warehouse - Cadena de Medios-Fines

4.4. FLUJO DE INFORMACION HORIZONTAL

Los conflictos interdepartamentales pueden ser causantes directos de una deficiente asimilación de tecnología en una empresa. En el caso de la arquitectura de información Data Warehouse, el efecto de tales conflictos puede ser aún más negativo, debido a la naturaleza propia del diseño de esta arquitectura en lo que se refiere a la compartición de los datos. Pues bien, la información lateral (horizontal) que es compartida entre los diversos departamentos o áreas funcionales de una empresa es muy importante, según se verá en este análisis, para lograr aplicar las estrategias de sumarización, partición e integración de los datos en la arquitectura de información Data Warehouse, estrategias que estén dirigidas a los objetivos del negocio y que den valor a la información que pueda generarse para la toma de decisiones.

Para analizar el flujo de información entre las áreas funcionales de una empresa se hace referencia al modelo de la *cadena de valor* propuesto por Michael Porter [PORTER,1995]. Con el uso de este modelo los diseñadores del Data

Warehouse pueden estar en condiciones de identificar el flujo de información interfuncional para el diseño de la arquitectura de información, así como para poder lograr establecer los canales de comunicación entre las áreas del negocio involucradas con la implantación del proyecto Data Warehouse.

4.4.1. Cadena de valor

“La cadena de valor es un conjunto de actividades o procesos de un negocio de un valor determinado (actividades de valor), que se desarrollan para diseñar, mercadear, transportar y soportar sus productos o servicios. La cadena de valor consiste de actividades primarias, de soporte y el margen o valor agregado. Las actividades primarias son las directamente involucradas en la creación física del producto o servicio, su venta y transporte al comprador, y el soporte después de la venta. Las actividades de soporte complementan las primarias ofreciendo funciones tales como recursos humanos, compras, desarrollo de tecnología y soporte administrativo. El margen o valor agregado es la diferencia entre la suma de los costos de las actividades de valor y la cantidad que los clientes están dispuestos a pagar por el producto o servicio de la organización” [ANDREAU,1991]. La figura 4.4 muestra de manera genérica la representación de la cadena de valor.

No es objetivo de esta tesis analizar la cadena de valor, sino más bien el enfoque es visualizar a una empresa como un conjunto o cadena de actividades interrelacionadas, que puedan ser examinadas independientemente y que puedan ser agrupadas para dar valor a la información de la empresa. Con la cadena de medios-fines los diseñadores del Data Warehouse pueden identificar la información en cada nivel de decisión, lo cual permite establecer los niveles de sumarización en los datos; con la cadena de valor, los diseñadores pueden visualizar y detectar las ligas entre las áreas funcionales de una empresa para establecer las estrategias de integración en los datos que permitan crear tales niveles de sumarización.

CADENA DE VALOR

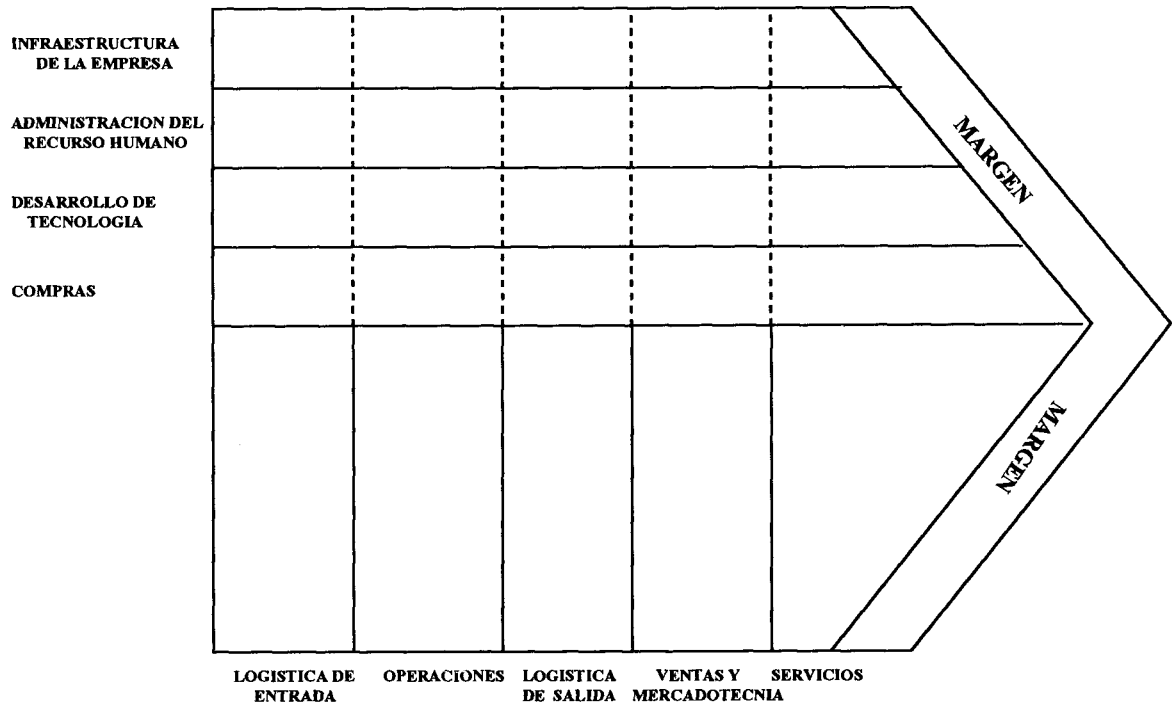


Figura 4.4 - Cadena de valor [ANDREAU,1991]

“La necesidad de integración entre las unidades organizacionales es una manifestación de eslabones. Hay con frecuencia muchos eslabones dentro de la cadena de valor, y la estructura organizacional falla con frecuencia en proporcionar mecanismos para coordinarlas u optimizarlas. La información necesaria para coordinar u optimizar los eslabones casi nunca es recopilada a través de la cadena. Los gerentes de las actividades de apoyo como administración de recursos humanos y desarrollo de tecnología no tienen con frecuencia una vista clara de cómo se relacionan a la posición competitiva general de la empresa, algo que la cadena de valor resalta. La tecnología de sistemas de información es particularmente penetrante en la cadena de valor, ya que cada actividad de valor crea y usa información; además estas tecnologías tienen un papel importante en los eslabones entre las actividades de todo tipo, debido a la coordinación y optimización de los eslabones, que requiere flujo de información entre las actividades” [ANDREAU,1991].

La liga o eslabones en una cadena de valor, las define Porter como las relaciones entre la forma en que una actividad de valor es desarrollada y el costo o desempeño de otra. En estas ligas es donde los diseñadores de la arquitectura de información Data Warehouse pueden apoyarse para definir la información que

defina las estrategias de sumariación, integración y partición de datos que estén orientados al negocio y a los objetivos organizacionales.

Tomemos el ejemplo del desarrollo de un Data Mart en la arquitectura de información Data Warehouse, para tratar de establecer de manera más clara la relación entre el Data Warehouse y la cadena de valor. Para crear un Data Mart se requiere aplicar estrategias de partición y sumariación de datos propias del Data Warehouse, en la figura 4.5 se muestra un ejemplo práctico de sumariación de datos para crear un Data Mart que contenga información sumariada de las ventas de una empresa.

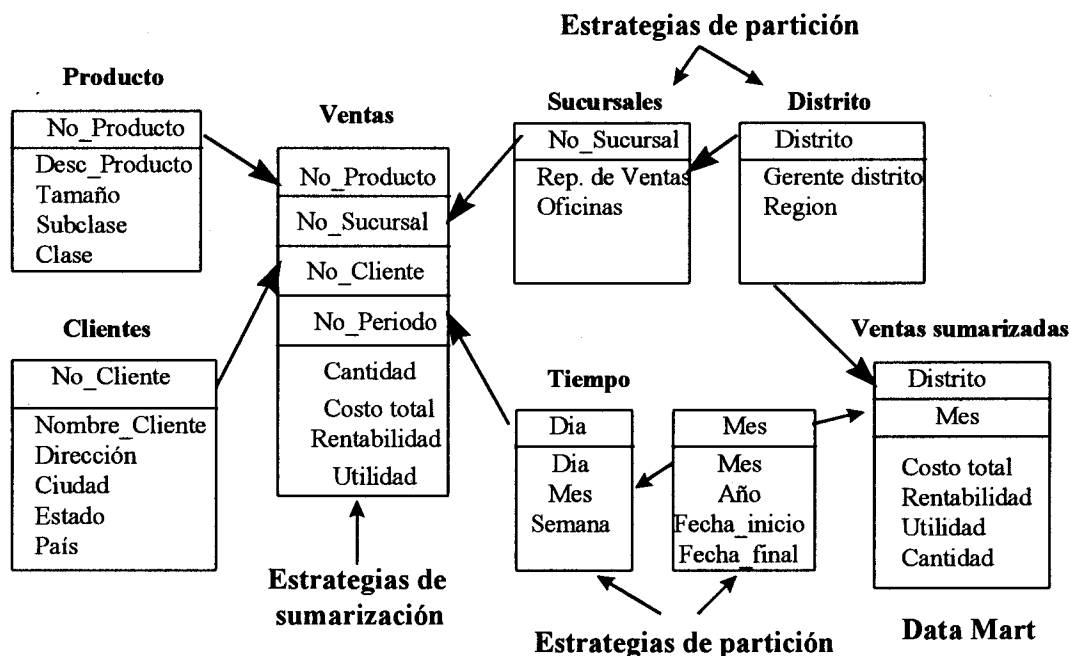


Figura 4.5 - Sumariación en ventas para la creación de un Data Mart

En la figura anterior se puede observar que la tabla VENTAS se forma por la aplicación de diversos procesos de sumariación e integración de datos de diversas fuentes como las tablas de CLIENTES Y PRODUCTOS, asimismo se pueden identificar algunas estrategias de partición de datos por área geográfica como las tablas de SUCURSALES O DISTRITOS, y por unidades de tiempo como la tabla de TIEMPO. La aplicación de todas estas estrategias están orientadas a formar la tabla de VENTAS SUMARIZADAS, la cual representaría el Data Mart para un distrito de ventas específico de la empresa. Pues bien, desde un punto de vista de la cadena de valor, las ventas y la mercadotecnia están directamente relacionadas con las operaciones y con la infraestructura necesaria para transportar los productos de una

sucursal o distrito a otro; de esta manera, podemos visualizar al Data Warehouse desde dos enfoques con respecto a la cadena de valor; un enfoque sería ver al Data Warehouse como una tecnología de información de soporte a las actividades primarias, principalmente en lo que se refiere al mejoramiento del flujo de la información y al apoyo al proceso de toma de decisiones, el otro enfoque sería visualizar al Data Warehouse como parte de las actividades primarias, es decir, la administración de los datos en dichas actividades puede llevarse a cabo de tal manera que inyecte valor a las ligas entre cada actividad primaria, siendo principalmente la administración de datos lo referente a las estrategias de sumarización, partición, integración, filtración, y demás procesos del Data Warehouse.

“Aplicar una efectiva estrategia para desarrollar un Data Mart requiere entender al negocio y su cadena de valor, los procesos del negocio que soporten la cadena de valor, y las decisiones que afecten a la empresa. Combinando este conocimiento con el entendimiento de los datos corporativos ayudará a desarrollar las particiones y sumarizaciones en las tablas de datos, con el fin de optimizar el rendimiento del Data Mart” [MEREDITH,1996].

Así como para el Data Mart, para el diseño de los demás componentes de la arquitectura de información Data Warehouse se requiere conocer al negocio y sus procesos, pudiendo considerar el uso de la cadena de valor como una alternativa para que los diseñadores del Data Warehouse puedan identificar las relaciones interfuncionales que agregan valor a los datos que se comparten en la organización.

4.5. RELACION DATOS-PROCESOS

Además de los análisis hechos anteriormente, resulta conveniente definir los responsables de la información en cada proceso del negocio, por lo que es importante visualizar a una empresa desde un punto de vista de procesos. Además de la identificación de los responsables de la información, el análisis por procesos de una empresa resalta las relaciones interfuncionales al igual que lo hacen las metodologías de la cadena de valor y la cadena de medios-fines. “Un proceso es un conjunto de tareas conectadas lógicamente que normalmente cruzan las barreras funcionales” [MCGEE,1993].

Independientemente de las relaciones interfuncionales en un estudio por procesos a una empresa, lo que se busca para fines de esta tesis es agrupar la información de cada proceso de acuerdo a sus características y similitudes,

identificando los responsables de la información, así como la relación de los datos con cada proceso del negocio. Se considera que el uso de la *metodología BSP* (Business System Planning) [FLAATTEN,1989] puede ser un medio para lograr analizar a una empresa desde un punto de vista de procesos, lo cual permita el establecimiento de la relación datos-procesos de la empresa y el agrupamiento de los procesos de acuerdo a sus características más importantes.

4.5.1. Metodología BSP

Muchas de las metodologías de modelación de negocios han tomado del BSP la manera de identificar los procesos, cada proceso debe ser analizado para determinar y analizar la parte de la organización que es responsable de él. En términos generales la metodología BSP consta de las siguientes etapas:

- 1.- Identificación de procesos en base a productos (servicios) y recursos que los soportan.
- 2.- Identificación de procesos administrativos de control y planeación.
- 3.- Identificación de los procesos operacionales
- 4.- Agrupación o división de los procesos
- 5.- Escritura de la descripción de cada proceso
- 6.- Relación de los procesos con la organización

Para fines del diseño de la arquitectura de información Data Warehouse y para no caer en repeticiones con los análisis de la cadena de medios-fines y la cadena de valor, se analiza solamente de manera específica el punto 4 referente a la agrupación y división de los procesos; asimismo se puede generar un análisis global de los 6 etapas, donde se muestren los requerimientos de los datos y en donde se identifiquen entre otras cosas: entidades, relaciones de datos con procesos, agrupaciones de datos, y el desarrollo de una matriz que relacione a los procesos con los datos. La agrupación de los procesos y el desarrollo de una matriz de la relación datos-procesos son para fines de esta tesis los aspectos a analizarse de la metodología BSP. Lógicamente que para desarrollar la agrupación de los procesos se requiere haber pasado por las primeras 3 etapas de la metodología, y para poder realizar la matriz de relación datos-procesos se necesita haber analizado todas las etapas del BSP; por esto aunque solamente se analicen los aspectos de la agrupación de procesos y la creación de la matriz, se da por hecho que se han visualizado las demás etapas de la metodología.

Agrupación de procesos

Una vez que han sido identificados los procesos de una empresa, estos pueden ser agrupados de tal manera que sean más entendibles por todos los involucrados en su funcionamiento, ya sea para reducir inconsistencias o para combinar procesos que son similares. La importancia de la agrupación de procesos para el diseño de la arquitectura de información Data Warehouse radica principalmente en poder conjuntar la información de los procesos que contengan información relacionada o similar, para así facilitar el acceso a los datos y la creación de reportes específicos a determinados procesos del negocio.

Uno de los aspectos más importantes en el funcionamiento de la arquitectura de información Data Warehouse es el tiempo de respuesta en las consultas y acceso a los datos, por lo que es preciso que la partición de éstos deba realizarse de tal manera que los accesos sean rápidos y adecuados a las necesidades de los usuarios finales. La agrupación de los procesos de una empresa puede ser un determinante para definir las estrategias de partición de los datos, al identificarse similitudes en cuanto a los niveles de consultas y actualizaciones de los datos de un proceso con respecto a los datos de otros. En el diseño de la arquitectura de información Data Warehouse Bill Inmon hace referencia a un término denominado 'Análisis de Estabilidad' [INMON,1996], el cual consiste principalmente en agrupar a los datos de acuerdo a sus tendencias de cambio, siendo el resultado de tal análisis la agrupación de datos con características similares. La figura 4.6 nos muestra un ejemplo del Análisis de Estabilidad en donde se crean 3 tablas derivadas de una tabla de propósito general. El análisis de una empresa por procesos en conjunto con el análisis de estabilidad de los datos puede soportar más adecuadamente la definición de las estrategias de partición de datos dentro de la arquitectura de información Data Warehouse, siendo más factible obtener mejoras en los tiempos de accesos y consultas a la información. La agrupación de los datos como se muestra en la figura 4.6 puede reflejarse posteriormente en el diseño físico de las tablas "fact" y dimensionales del Data Warehouse.

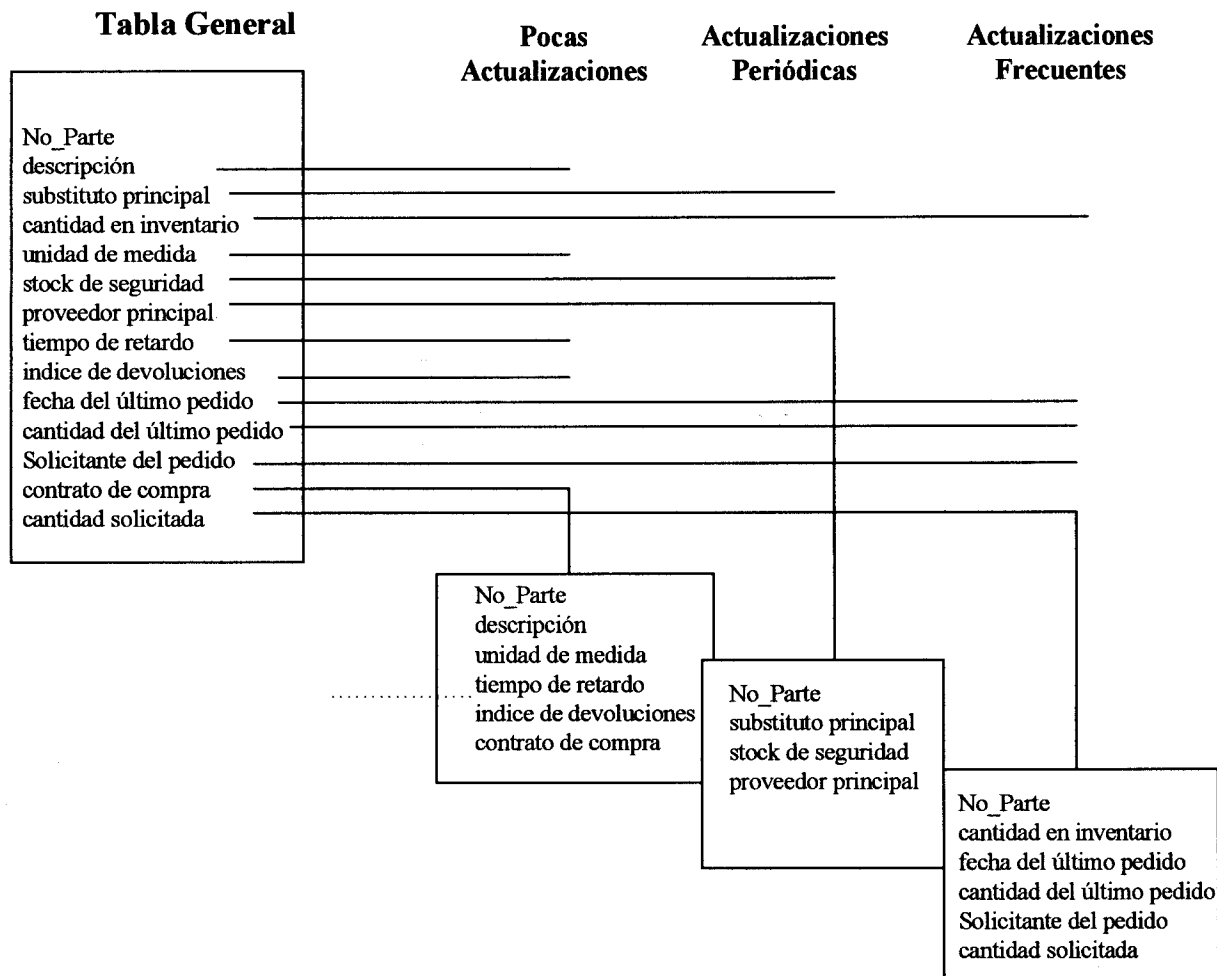


Figura 4.6 - Ejemplo del Análisis de Estabilidad

Del análisis de la metodología BSP puede generarse también una matriz que relacione a los datos con los procesos de una empresa, la figura 4.7 nos muestra la matriz de relación Datos-Procesos para una tienda de autoservicio. En esa matriz se observan los diversos procesos de una empresa de dicho giro y los datos relacionados a cada proceso; como ejemplo podrían determinarse que los procesos Recepción de Mercancías, Pedido de Mercancías y Procesos Contables son factibles a agruparse porque poseen datos en común y porque se encuentran directamente relacionados en algunas de sus actividades, si a tales procesos ya agrupados se les aplica el análisis de estabilidad de acuerdo a sus niveles de actualización y consulta, tendríamos como resultado el diseño de tablas que estén dirigidas a procesos importantes de la organización. La creación de estas tablas es parte del diseño de la arquitectura de información Data Warehouse, ya sea para la generación de bases de

datos sumariadas o para el diseño de un Data Mart que apoye el proceso de toma de decisiones en un área funcional de la empresa.

DATOS PROCESOS	Articulos en Venta	Facturas	Ventas	Pedidos	Descuentos	Inventario Almacen	Cobro de Servicios
Recepción de Mercancías		■		■		■	
Pedido de Mercancías		■		■		■	
Asignación de Precios	■					■	
Actualización Precios POS	■						
Descuentos	■				■		
Ventas Generales	■		■		■		
Ventas Art. Sec	■		■				
Procesos Contables	■	■	■		■		■
Transferencia de Fondos			■				■

Figura 4.7 - Matriz de relación Datos-Procesos

4.6. ALINEACIÓN DE LA TECNOLOGÍA DATA WAREHOUSE A LOS OBJETIVOS DEL NEGOCIO

Teniendo el conocimiento de la tecnología Data Warehouse, su visualización como una arquitectura de información, así como el conocimiento de la empresa y sus datos, se puede considerar que se cuenta con las bases suficientes para lograr un proceso de asimilación tecnológica exitoso. Sin embargo, algunos autores basados en experiencias en el proceso de asimilación de tecnología han constatado que es necesario considerar algunos otros aspectos o requisitos para alcanzar un resultado exitoso en dicho proceso [RAMÍREZ, 1993]:

- Una planeación cuidadosa
- Un proceso participativo
- La participación de agentes de cambio y de “campeones”
- La consideración de la resistencia al cambio

Es precisamente en el primer aspecto a lo que nos referiremos en este apartado de la tesis, esto es, en como realizar una planeación cuidadosa del diseño de la arquitectura de información Data Warehouse, alineada a los objetivos del negocio y fundamentada en un plan informático para la empresa. El aspecto referente al proceso participativo se contempla también en este apartado, ya que se considera inherente en la planeación del Data Warehouse o de cualquier otra tecnología.

Ninguna aplicación de tecnología de información por muy sofisticada que sea, entrega ventajas competitivas sostenibles por si misma, la ventaja se obtiene en el uso de la funcionalidad de la TI de una manera continua y alineada estratégicamente a los objetivos de una empresa. "La alineación estratégica consiste en mantener en armonía la informática con el negocio a través de relacionar adecuadamente sus procesos y lograr ventajas competitivas" [BOAR,1993]. La tecnología Data Warehouse no es la excepción, ya que para que ésta realmente cumpla con las funciones para la cual fue diseñada es necesario que cuente con dicha alineación estratégica.

El éxito del desarrollo e implantación de un Data Warehouse y de cualquier otra tecnología depende de que tenga un enfoque hacia el negocio y no tanto hacia la tecnología. Ahora bien, en muchas organizaciones el papel de las tecnologías de información es apoyar a ciertas funciones de la empresa, colocándola en la estructura organizacional debajo de áreas como Recursos Financieros, Recursos Humanos o de Planeación, lo cual limita la participación de los gerentes de tecnología de información en las decisiones estratégicas de la empresa. "La ventaja estratégica usando tecnología de información no puede llevarse a cabo a menos que la organización vaya más allá de esperar un retorno en la inversión en tecnología y considere el valor de la información como un activo económico por si solo" [LUFTMAN,1993]. Si a la información se le da un carácter de activo económico, es necesario que la tecnología que se emplee para difundirla a través de toda la organización esté fundamentada por una serie de estrategias las cuales deben estar totalmente ligadas a las estrategias del mismo negocio.

4.6.1. Planeación Informática

“Planeación es un proceso continuo para determinar explícita y sistemáticamente cuál debe ser el papel de la informática y la manera de ser evaluada. La planeación de sistemas de información tiene que ver con las actividades necesarias para posicionar de manera estratégica a la informática en una organización” [BOAR,1993]. Existen diversas metodologías para la planeación informática; en esta tesis se presenta una metodología [BOAR,1993], a fin de recalcar la importancia de esta planeación para el éxito del diseño de la arquitectura de información Data Warehouse. El proceso de planeación que a continuación se presenta se divide en 4 etapas, las cuales se describirán brevemente.

I. Análisis Situacional

El objetivo de esta fase es documentar la situación actual de la empresa y la informática, estableciendo las amenazas, oportunidades, fuerzas y debilidades de ambas, para alcanzar, mantener, o aumentar ventajas competitivas. Para poder llevar a cabo este análisis se pueden usar diferentes metodologías, una de ellas es la de los Factores Críticos de Éxito [ROCKART,1979], la cual consiste en identificar “aquellas pocas áreas donde las cosas deben ir muy bien para que el negocio prospere”. El análisis de los Factores Críticos de Éxito tiene dos efectos en los individuos, primero, ayuda a enfocar los esfuerzos en las cosas importantes, y segundo, ayuda a pensar en la información requerida para monitorear los aspectos relevantes, por otra parte el tenerlo a través de una metodología permite analizarlos de manera explícita. La parte de interés de los Factores Críticos de Éxito para el diseño de una Data Warehouse es precisamente la información que se requiere para monitorear tales factores, siendo los Sistemas de Información Ejecutivos parte importante para que esta información llegue de manera confiable y amigable a los altos directivos de una empresa.

Existen algunas otras metodologías que pueden utilizarse para el análisis situacional de una empresa, entre éstas se tiene el análisis SWOT (strengths, weaknesses, opportunities, threats), que permiten determinar las fuerzas, debilidades, oportunidades y amenazas de la organización. El modelo de la cadena de valor y la metodología BSP también pueden emplearse en esta fase de la planeación informática.

II. Liga de la Planeación informática con la Planeación del negocio.

Para fines de esta tesis, ésta etapa del proceso de planeación es la más importante y es en donde se debe poner más énfasis cuando se diseñe un Data Warehouse en una organización. No existe un enfoque único para crear las estrategias o planes informáticos, tampoco para identificar áreas de oportunidad para transformar el negocio. El establecer las ligas entre la planeación informática y el negocio debe ser incluido en cualquier método usado. El enfoque para un método debe ser especificado en función de las circunstancias del negocio y de la tecnología de información con que cuente. El enfoque debe enfatizar el valor de las actividades de la informática.

La importancia de la alineación entre la planeación informática y el negocio para el diseño de la arquitectura de información Data Warehouse, radica en que de esta alineación surgen necesidades de información en todos los niveles y funciones del negocio, información que apoya al logro de las estrategias de la empresa. Algunas de las metodologías que se han visto anteriormente como la cadena de medios-fines y los factores de críticos de éxito apoyan fuertemente en la identificación de tales necesidades de información, pero ahora ubicaremos el análisis como parte de un proceso de planeación informática que puede incluir al Data Warehouse y a otras tecnologías. Rafael Andreau [ANDREAU,1991] presenta un procedimiento de alineamiento del Plan de Tecnologías y Sistemas de Información (TI/SI) con la estrategia del negocio, siendo representado esquemáticamente en la figura 4.8.

Este procedimiento integra las necesidades de información emanadas del cruce de las directrices estratégicas de la empresa con las funciones de negocio de las distintas unidades organizacionales. En las columnas de la matriz de la figura 4.8 se observan las acciones estratégicas que sigue la compañía y en las filas las funciones de negocio de cada uno de los departamentos. De esta matriz surgen las necesidades de información que se integran en el Plan de Tecnologías y Sistemas de Información, el cual posteriormente es base para el desarrollo de un plan informático, que incluiría en su caso el diseño de la arquitectura Data Warehouse. Es necesario entonces, que para el desarrollo de la matriz se utilicen métodos comprobados y aplicados con éxito en algunas empresas; es por eso el énfasis en el análisis de los factores críticos de éxito para apoyar la identificación de las necesidades de información para el logro de las estrategias de la empresa (columnas de la matriz), y en la cadena de medios-fines para la identificación de necesidades de información en lo que se refiere a las funciones de negocio.

Realizando adecuadamente este procedimiento permitirá que los proyectos de tecnología de información que sean desarrollados, por un lado apoyen a las

estrategias de la empresa, y por otro que sirvan a los departamentos en sus actividades cotidianas. “Usar tecnología de información para tener una ventaja estratégica implica la aplicación de tecnología de información para habilitar estrategias que mejoren los factores críticos de éxito o las capacidades básicas, o ambas, de la empresa” [LUFTMAN,1993].

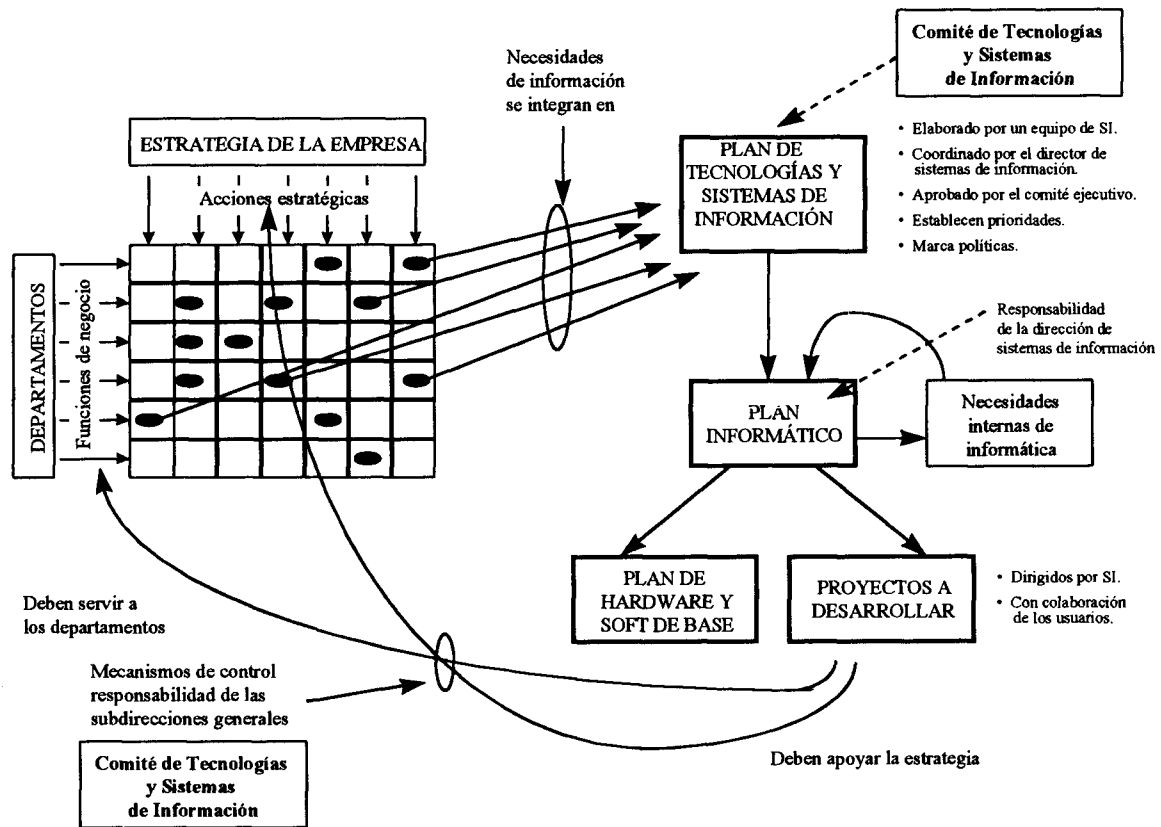


Figura 4.8 - Estructura de un procedimiento de alineamiento del Plan de TI/SI con las estrategias de la empresa [ANDREAU,1991]

III. Planeación de productos de informática

En esta fase debe realizarse el proceso mediante el cual se definan los productos informáticos, los recursos necesarios y forma de organizarlos para cubrir los requerimientos que plantea el tipo de planeación que se está efectuando. En el caso del diseño de la arquitectura de información Data Warehouse, es importante la planeación de las herramientas informáticas requeridas para su implantación,

algunas de estas herramientas fueron analizadas en el capítulo 3 de esta tesis, por ejemplo, los Sistemas de Información Ejecutivos, los Sistemas de Soporte a la Decisión, los Data Marts, las herramientas Data Warehouse, etc. La planeación para el desarrollo de tales herramientas debe ser parte del plan informático, para que de esta manera apoyen tanto a la arquitectura Data Warehouse como a otras tecnologías o sistemas de información de la empresa.

IV. Establecimiento de prioridades estratégicas

Del análisis de las necesidades de información de cada función del negocio y de cada acción estratégica surgen diversos proyectos que no pueden ser tratados al mismo tiempo, por esto se requiere establecer prioridades de ejecución, en función de los objetivos y estrategias del negocio. Para ello se tiene que definir un proceso que debe contar por lo menos con los siguientes pasos:

1. Establecimiento de los criterios de evaluación.
2. Realización del análisis de costo-beneficio de los proyectos.
3. Clasificación de los proyectos de acuerdo a las bases seleccionadas.
4. Realización de la recomendación del orden de ejecución de los proyectos.

“El enfoque en la alineación estratégica es importante para asegurar que las prioridades sean consistentes, que los recursos sean asignados apropiadamente y que la sofisticación de las tecnologías estén de acuerdo con la sofisticación de la organización” [LUFTMAN,1993]. El establecimiento de prioridades permitirá a los diseñadores del Data Warehouse identificar aquellas áreas donde inicialmente debe implantarse esta arquitectura de información. Un aspecto importante al momento de establecer prioridades es considerar la situación de los sistemas de información en la empresa, para determinar el grado de automatización de un proceso y su dependencia de la información.

Se han analizado brevemente las 4 fases para establecer una liga entre el negocio y la tecnología de información, haciendo énfasis en el Data Warehouse por ser una tecnología que tiene un carácter estratégico, por lo cual sin esta liga sería muy difícil su implantación y asimilación en una organización. Por último se recalca el hecho de que “la alineación estratégica no es un evento, sino un camino continuo de transformación y no es por sí sola la tecnología de información la que dará valor a la organización, sino también, por los procesos del negocio que explotan esa tecnología” [LUFTMAN,1993].

4.7. PROCESO DE ASIMILACION DE TECNOLOGIA

Se han tratado ya algunos aspectos relacionados a la arquitectura de información Data Warehouse y sus componentes, asimismo se ha hecho énfasis en la importancia del conocimiento de la empresa desde una dimensión vertical y horizontal en lo que se refiere al flujo de información; así también, la identificación de responsabilidades en cuanto al manejo de la información por parte del recurso humano ha generado la necesidad de identificar la relación entre los procesos y los datos de una empresa. Todos estos puntos han sido analizados especialmente para la arquitectura de información Data Warehouse, siendo el objetivo principal de su estudio comprenderlos y entenderlos totalmente para apoyar el análisis del proceso de asimilación tecnológica del Data Warehouse en una organización, propuesta principal de esta tesis. El análisis del proceso de asimilación de tecnología se presenta en el siguiente capítulo, en este apartado se sientan las bases de dicho proceso y se mencionan propuestas de dos autores, las cuales se emplean para apoyar la propuesta de asimilación tecnológica del siguiente capítulo.

“Se entiende como proceso de Asimilación de Tecnología a las etapas por las cuales atraviesa una organización para introducir un cambio en la esencia de como realiza sus actividades, con la finalidad de hacerlas más eficientes (puede ser desde la reestructuración de tareas hasta la completa automatización de un proceso o tarea), e inicia con la búsqueda de la tecnología, y termina con la total integración en la organización” [RAMÍREZ,1993].

Existen diversos modelos y teorías para lograr una asimilación de una tecnología o innovación tecnológica en una empresa. Brevemente se mencionan a continuación dos de éstos modelos.

4.7.1. Modelo para asimilación de Innovaciones de Meyer y Goes

El modelo propuesto por Meyer y Goes es altamente articulado, trata exclusivamente con innovación tecnológica y reconoce no sólo aspectos económicos y estratégicos de la innovación, sino también sociales, culturales y políticos. Meyer y Goes [MEYER,1988], elaboran su modelo general de innovación tecnológica sugiriendo que tres tipos de factores influyen en el proceso de asimilación de innovaciones en las organizaciones y que esos factores operan dentro del modelo de etapas que proponen:

- 1) Atributos de las innovaciones
Relacionados al grado de riesgo de la tecnología o a qué nivel de habilidades o capacitación se necesita para utilizarla.
- 2) Atributos de los contextos organizacionales
Relacionados con el tamaño de la organización, la complejidad del producto/servicio entregado y las características de los empleados.
- 3) Atributos procedentes de la interacción de las innovaciones y el contexto organizacional
Se relacionan con la compatibilidad de la tecnología con las características de los empleados y la magnitud del apoyo administrativo para la innovación tecnológica.

El modelo de etapas en la asimilación de innovaciones tecnológicas de Meyer y Goes es el siguiente:

ETAPA DE CONOCIMIENTO-CONCIENCIA

1. Compresión: Los individuos saben de la existencia de la innovación.
2. Consideración: Los individuos consideran la conveniencia de la innovación.
3. Discusión: Los individuos entablan conversaciones concernientes a la adopción.

ETAPA DE EVALUACION-SELECCION

4. Propósito de adquisición: La adopción de la tecnología se propone formalmente.
5. Evaluación racional: La inversión propuesta se evalúa de acuerdo a criterios funcionales y financieros.
6. Evaluación política-estratégica: La inversión propuesta se evalúa de acuerdo a criterios políticos y estratégicos.

ETAPA DE ADOPCION-IMPLANTACION

7. Prueba: La tecnología es adquirida pero permanece bajo pruebas de evaluación.
8. Aceptación: La tecnología llega a ser bien aceptada y utilizada frecuentemente.
9. Expansión: La tecnología es expandida, actualizada o reemplazada con una segunda generación del modelo.

4.7.2. Modelo para la Asimilación de tecnología de McFarlan y McKenney

McFarlan y McKenney [MCFARLAN,1982], [MCFARLAN,1983,1], [MCFARLAN,83,2] proponen un modelo incremental para la asimilación de tecnología:

- **Identificación de tecnología e inversión.** El énfasis es en el aprendizaje y en la aplicación de la nueva tecnología. Las habilidades del personal son creadas mediante estudios piloto. En esta fase la experiencia con la tecnología es tan limitada que los participantes no alcanzan a percibir implicaciones a futuro. La falta de atención de la administración, la administración incompetente del proyecto o pocas opciones de los vendedores pueden causar estancamiento e impedir el movimiento hacia la siguiente fase.
- **Experimentación, aprendizaje y adaptación.** El enfoque principal en esta fase es la concientización del usuario sobre los problemas que puede solucionar la nueva tecnología. El éxito en esta fase generalmente lleva a incrementar las peticiones de servicio. Se debe ser muy cuidadoso para permanecer flexible, ya que es esencial para el desarrollo de aplicaciones innovadoras. Puede ocurrir un estancamiento si los administradores no desarrollan y refinan su entendimiento de la tecnología.
- **Racionalización y control administrativo.** Esta fase se caracteriza por la búsqueda de eficiencia a corto plazo. En esta etapa se implanta el control administrativo para monitorear el crecimiento del nuevo recurso. Las actividades en esta fase deben centrarse en aumentar el conocimiento del staff a niveles aceptables; reorganizar para desarrollar nuevos proyectos; y determinar la tecnología adecuada. Si se da mucha estandarización puede ocurrir estancamiento en esta fase.
- **Difusión de la transferencia de tecnología.** Los beneficios y experiencia de la nueva tecnología son diseminados hacia otras unidades de la organización. La base tecnológica está instalada; el aprendizaje está completo; y es enfatizado el análisis y planeación a largo plazo.

4.7.3. Data Warehouse como innovación tecnológica

Se debe tomar en cuenta que la arquitectura Data Warehouse es un concepto relativamente nuevo en el ambiente de las tecnologías de información, su conocimiento en las empresas es poco o nulo tanto para el personal informático

como para los usuarios finales en general. “La creación de un Data Warehouse representa un desafío difícil para el Staffs de TI, que tiene poca o nula experiencia con esta nueva tecnología. Los riesgos están presentes: Un Data Warehouse puede tomar años en desarrollarse, costar una fortuna, y después ser ignorada por el recurso humano de la empresa, los cuales no entienden como usarla apropiadamente” [KAY,1996]. La arquitectura de información Data Warehouse es una innovación tecnológica que involucra cambios importantes en algunos aspectos tecnológicos, humanos y organizacionales, los cuales deben tomarse en cuenta desde la primera etapa del proyecto. Niederman y Brancheau señalan que la mayoría de las tecnologías de información que recientemente se están utilizando pueden ser consideradas como innovaciones [NIEDERMAN,1987].

Definitivamente, lo que se busca al desarrollar un Data Warehouse en una empresa es resolver problemas relacionados con la obtención de información para la generación de reportes corporativos que soporten la toma de decisiones, así como el tener bases de datos con la información que realmente necesitan los ejecutivos en una organización; pero el hecho de considerar al Data Warehouse como una innovación tecnológica implica además el buscar una armonía total entre la tecnología y la misma empresa. Sucede que en muchas ocasiones las soluciones actuales se convierten en los problemas del futuro, y quizá la introducción de una nueva tecnología resuelva algún problema específico del negocio, pero tal vez genera un problema mayor que podría ser de carácter organizacional.

Es así como al adoptar una nueva tecnología no sólo tendrán que administrarse (planear, controlar y evaluar) las actividades relacionadas a las habilidades y conocimientos inherentes y necesarios para utilizar la innovación, sino que además deberán administrarse los cambios necesarios para garantizar el ‘acomodo’ de la nueva tecnología en la organización. Además, es necesario estar conscientes de los problemas organizacionales existentes antes de la adopción de la innovación, y tratar de corregirlos antes o a la par del proceso de implantación, así como controlar y evaluar constantemente los resultados que se van obteniendo conforme se avanza en el proceso.

4.8. CONCLUSIONES

La introducción en una empresa de una innovación tecnológica como el Data Warehouse implica cambios en la manera de hacer las cosas para los usuarios involucrados con esta arquitectura de información, es así que del grado de conocimiento de la empresa en lo referente a sus procesos principales, sus

estructuras organizacionales, funciones, etc. dependerá la calidad y el enfoque del plan de introducción de la tecnología, plan que debe hacer énfasis en la adaptación del Data Warehouse a las estructuras y al ambiente organizacional.

El análisis que propongo para conocer a una empresa a fin de introducir la tecnología Data Warehouse esta basado en herramientas y metodologías que permiten estudiar las estructuras de la organización de tal manera que puedan representarse posteriormente en el diseño de un Data Warehouse; es decir, se buscó relacionar directamente el diseño organizacional con el diseño de la arquitectura de información. Para lograr obtener la realidad del negocio se necesita la participación activa del recurso humano de todos los niveles organizacionales, principalmente los involucrados en los procesos de información, y para lograr un reflejo de esta realidad organizacional en el diseño del Data Warehouse se necesita un conocimiento total de la tecnología, así como un plan informático que guíe la introducción de cada uno de los componentes de la arquitectura de información, a fin de lograr la alineación de cada componente del Data Warehouse con las necesidades de información de la empresa y con los objetivos organizacionales.

Pues bien, para poder alinear la tecnología con el negocio y contar con la participación del recurso humano en tal actividad, se requiere un proceso de asimilación de tecnología que por etapas defina el grado de adopción que va teniendo la tecnología durante sus diversas fases de implantación. Tal proceso de asimilación de la tecnología Data Warehouse debe implicar el conocimiento de la innovación, el conocimiento de la empresa, la participación del recurso humano y la alineación de la tecnología con los objetivos organizacionales. En los capítulos anteriores se trataron los aspectos relacionados con el Data Warehouse y su representación como arquitectura de información, así como componentes que la conforman, en este capítulo se establecieron las bases de un proceso de asimilación de tecnología, mencionando 2 propuestas y 3 atributos que influyen en tal proceso. En el capítulo siguiente tomo como base una guía general de un proceso de asimilación de tecnología y propongo una análisis de cada fase de esta guía alineándola específicamente a la arquitectura Data Warehouse, considerando asimismo los atributos que pueden influir en el grado de adopción de esta arquitectura de información en una organización.

REFERENCIAS BIBLIOGRAFICAS

- [ANDREAU,1991] ANDREAU, Rafael, Joan Ricard & Josep Valor.
Estrategia y Sistemas de Información.
MacGraw-Hill / Interamericana de España, 1991
- [BADAWY,1989] BADAWY, Michael.
“Integration: The fire under technology transfer”.
Industry Week, Junio 19 de 1989. Vol. 238, No. 12
- [BOAR,1993] BOAR, Bernard.
The art of strategic planning for information technology.
John Wiley and Sons, 1993
- [FLAATTEN,1989] FLAATTEN, Per O., Donald J. McCubbrey, P. Declan O’Riordan
& Keith Burgess.
Foundations of Business Systems.
Andersen Consulting, Arthur Andersen & Co., 1989
- [INMON,1996] INMON, W.H.
Building the Data Warehouse.
John Wiley & Sons, Inc, segunda edición, 1996
- [KAY,1996] KAY, Emily.
“A Classic Case of Build vs. Buy”.
Software Magazine, Center Market Research, 1996
- [LAZZATI,1990] LAZZATI, Santiago.
Claves de la decisión en la empresa: Método y participación.
Ed. Macchi, segunda edición, 1990
- [LUFTMAN,1993] LUFTMAN, J.N., P.R. Lewis & S.H. Oldach.
“Transforming the enterprise: The alignment of business and
information technology strategies”.
IBM Systems Journal, vol. 32. No. 1, 1993
- [MCFARLAN,1982] MCFARLAN, F.W. & J.L. McKenney.
“The Information Archipelago-Gaps and Bridges”.
Harvard Business Review, Septiembre-Octubre 1982, vol. 60

- [MCFARLAN,1983,1] MCFARLAN, F.W., J.L. McKenney & P. Pyburn.
“The Information Archipelago-Plotting a Course”.
Harvard Business Review, Enero-Febrero 1983, vol. 61
- [MCFARLAN,1983,2] MCFARLAN, F.W., & J.L. McKenney.
Corporate Information Systems:The Issues Facing Senior Executives.
Richard D. Irwin, Inc. Homewood, Illinois, 1983
- [MCGEE,1993] MCGEE, James & Laurence Prosak.
Managing information strategically.
John Wiley & Sons, 1993
- [MEREDITH,1996] MEREDITH Mary, Khader Aslam.
“Divide and Aggregate”.
Database Programming and Design, Junio 1996
- [MEYER,1988] MEYER, Alan D. & James B. Goes.
“Organizational Assimilation of Innovations: A Multilevel Contextual Analysis”.
Academy of Management Journal, 1988.Vol.31,No. 4,
- [NIEDERMAN,1987] NIEDERMAN, Fred & James C. Brancheau.
“Planning and Managing Technology Change”.
Management Inf. Systems Research Center, Junio 1987
- [PORTER,1995] PORTER, E. Michael.
Competitive Advantage, Creating and sustaining superior performance.
The Free Press, 1ª Edición. EUA, 1995
- [RAMIREZ,1993] RAMIREZ, Rosario.
Administración del proceso de asimilación de tecnología de información en una organización y su aplicación a case
Tesis de maestría ITESM, Monterrey, N.L., 1993
- [ROCKART,1979] ROCKART, John.
“Chief Executives define their own data needs”.
Harvard Business Review, Marzo-Abril 1979.
- [WOLFF,1989] WOLFF, Michael.
“Technology transfer: A GM manager’s strategy”.
Research-Technology Management, Septiembre-Octubre 1989

CAPITULO 5

PROCESO DE ASIMILACION DE LA ARQUITECTURA DE INFORMACIÓN DATA WAREHOUSE EN UNA ORGANIZACION

5.1. INTRODUCCION

Resulta lógico pensar que el diseño de un Data Warehouse dependerá de las características de cada empresa que lo desee implantar, esto es, su tamaño, sus prácticas organizacionales, su tecnología, sus necesidades de información, etc., de tal manera que el proponer una guía detallada para el diseño de un Data Warehouse como una “receta de cocina” no es un objetivo que se persigue en esta tesis, porque resultaría muy complejo aplicar una metodología para diseñar un Data Warehouse y que en diversas empresas se implantara y se lograra su adaptación total a las necesidades de información de cada una de ellas. “Construir un Data Warehouse es probablemente un proyecto que se realiza por primera vez en una empresa, y establecer una serie de pasos para llevar a cabo su implantación de manera exitosa no es algo obvio. Cada ambiente es diferente, de tal modo que es difícil definir una fórmula que garantice su diseño exitoso, sin embargo, algunas reglas estándar deben ser tomadas en cuenta” [BROWN,1995]. Resulta entonces inadecuado proponer en esta tesis algún método o guía de pasos para el diseño de un Data Warehouse, más sin embargo de algunas de las metodologías que ya existen se podrían identificar similitudes en algunas de sus etapas, y mediante un análisis comparativo lograr un entendimiento total de cada una de ellas, es así que previo al proceso de asimilación de la tecnología presento un análisis comparativo de algunas propuestas de diseño de un Data Warehouse de autores diversos, las cuales son mencionadas en este capítulo. El fin que persigo al realizar este comparativo es generalizar la aplicación del proceso de asimilación de tecnología que propongo, de tal manera que pueda ser usado independientemente de la metodología de diseño del Data Warehouse, esto considerando que el proceso de asimilación debe asociarse al diseño del Data Warehouse mediante un análisis en paralelo.

El desarrollo de un Data Warehouse debe realizarse por etapas bien definidas, de tal manera que el análisis de cada etapa permita contemplar los aspectos propios de diseño de esta tecnología, pero a la vez, permita ir integrando paralelamente la

tecnología en la organización mediante un proceso bien establecido de asimilación. Es por eso, que en términos generales lo que propongo en esta tesis es un proceso de análisis de asimilación de la tecnología Data Warehouse en una organización, considerando para esto los siguientes *elementos*:

- Un análisis de diversos atributos que influyen en la adopción y asimilación de la tecnología Data Warehouse en una organización.
- Un proceso de asimilación de tecnología bien definido y entendido para su seguimiento en el diseño de un Data Warehouse.
- Un análisis de cada fase del proceso de asimilación de tecnología mediante la evaluación del grado de adopción de la innovación Data Warehouse, la cual contempla a las etapas en el diseño de la tecnología y a los atributos que influyen en el proceso de asimilación.

Con respecto al primer punto identifiqué los atributos que pueden influir en el proceso de asimilación del Data Warehouse basado en la propuesta de Meyer y Goes [Meyer,1988], los cuales sugieren la existencia de 3 tipos de factores o atributos globales que influyen en un proceso de asimilación. Los atributos identificados forman parte de la evaluación en el grado de adopción que propongo para cada fase del proceso de asimilación.

En el segundo punto propongo una serie de actividades para el análisis de cada fase del proceso de asimilación que presento para el Data Warehouse, se tomó como base una guía de asimilación de tecnologías de información presentado como tesis por Rosario Ramírez en 1993, [Ramírez,1993], en dicha guía se establecen las etapas requeridas para generalizar la aplicación del proceso de asimilación para cualquier tecnología de información.

Referente al tercer elemento propongo una evaluación del grado de adopción que va presentando la arquitectura Data Warehouse conforme avanza su implantación en una organización. Propongo realizar tal evaluación al finalizar cada fase del proceso de asimilación, y en la cual involucro a los atributos que influyen en el proceso de asimilación para cada fase, así como las etapas del diseño del Data Warehouse que puedan estar asociadas a la fase en evaluación. Para llevar a cabo la evaluación presento cinco parámetros que permitan medir el grado de adopción de la tecnología, y para los cuales establezco una serie de preguntas orientadas a determinar dicho grado, de tal manera que al final de cada fase se pueda aplicar un cuestionario a los involucrados con la implantación de la tecnología para conocer el nivel de aceptación o rechazo hacia el Data Warehouse, que finalmente es lo que debe interesar a los promotores de la tecnología para tomar decisiones con respecto al plan de introducción de la arquitectura de información.

Como resultado de este proceso analítico de asimilación tecnológica se deben generar decisiones que fortalezcan los elementos que podrían obstaculizar un diseño exitoso y una adopción adecuada de la tecnología Data Warehouse en una organización. En este capítulo se consideran aspectos importantes del diseño de la arquitectura de información Data Warehouse, algunos de los cuales ya fueron explicados en los capítulos 2 y 3 de esta tesis, y además, se considera lo visto en el capítulo 4 en cuanto al entendimiento del negocio que se requiere tener para el diseño del Data Warehouse, así como la alineación que debe existir entre los objetivos de la empresa y la tecnología Data Warehouse como parte de un plan informático.

5.2. ATRIBUTOS QUE INFLUYEN EN LA ASIMILACION DE LA TECNOLOGIA DATA WAREHOUSE

De acuerdo al modelo para la asimilación de innovaciones de Meyer y Goes (ver 4.7.1), existen tres tipos de atributos que influyen en el proceso de asimilación de una innovación tecnológica en una organización:

- 1) Atributos de la innovación: Características inherentes a las innovaciones tecnológicas
- 2) Atributos del contexto organizacional: Características del ambiente, (organización, dirigentes, etc.)
- 3) Atributos de la interacción de las innovaciones con el contexto organizacional: Características inherentes a la interacción.

Basándonos en los conceptos del Data Warehouse, su visualización como arquitectura de información, y con un enfoque hacia el proceso de asimilación de la tecnología, se identifican los siguientes atributos:

5.2.1. Atributos de las innovaciones.

- Tecnología de información requerida para diseñar el Data Warehouse

El crear un Data Warehouse implica un desafío muy fuerte para el departamento de informática o sistemas de información de una empresa, ya que es poca o nula la experiencia con esta nueva tecnología. Además del

conocimiento propio de la tecnología Data Warehouse, se deben tomar decisiones importantes en hardware, software, conectividad y en una serie de herramientas computacionales que contemplen actividades como la modelación de datos, la extracción, el almacenamiento, la limpieza de datos, etc.

Es importante analizar la situación de la tecnología en una empresa y evaluar la manera de asimilar las variantes tecnológicas del Data Warehouse sin alterar significativamente la plataforma tecnológica que exista en una organización. En tal evaluación se debe considerar la compatibilidad de las nuevas herramientas tecnológicas con la TI existente en la empresa, sobre todo lo referente a las Tecnologías de Base de Datos, herramientas de conectividad y las herramientas propias del Data Warehouse (ver tabla 3.1). Asimismo es necesario analizar las herramientas tecnológicas con las que actualmente cuenta la empresa que podrían soportar los procesos más importantes de la arquitectura de información Data Warehouse.

5.2.2. Atributos de los contextos organizacionales

- Estado actual de los datos

El diseño de un Data Warehouse en una empresa requiere el conocimiento de la situación de sus datos con el fin de determinar las deficiencias y así poder establecer las mejoras que se consideren necesarias para generar la información que los usuarios finales necesitan. La frase “Si basura entra, basura sale” se aplica perfectamente al diseño de un Data Warehouse y se refleja directamente en la información que usan los tomadores de decisiones y los usuarios finales a un nivel ejecutivo. Construir un Data Warehouse es un proceso que consiste en relacionar las necesidades de los usuarios finales con la realidad en la disponibilidad de los datos.

Algunos aspectos a considerar en este atributo son:

- a) Limpieza de los datos
- b) Disponibilidad de los datos
- c) Nivel de integración de los datos

Estos atributos pueden influir de manera importante en el proceso de asimilación tecnológica del Data Warehouse y en el diseño propio de esta innovación, esto debido a que la mayoría de los procesos de la arquitectura de información Data Warehouse están involucrados con la transformación y

administración de los datos para que puedan ser presentados al usuario final de manera clara y precisa a sus necesidades de información.

- **Características del recurso humano requerido**

Para lograr la asimilación total de una tecnología como el Data Warehouse en una organización se requiere que el recurso humano la acepte y la comprenda en su totalidad, principalmente en aquellos aspectos en los cuales está directamente relacionado. Son muchas las personas que pueden participar en el diseño de un Data Warehouse, dependiendo del área funcional que se esté analizando en su momento, es por eso que la administración del recurso humano es un factor importante al momento de diseñar un Data Warehouse en una organización, e incluso pueden hasta crearse nuevas funciones principalmente en lo que se refiere a la administración de los datos.

La asimilación de tecnología requiere de una administración innovadora, así como de darse cuenta de que la tecnología no se vende por sí sola. Para que el proceso sea exitoso se requiere: involucrar a los usuarios, fomentar la participación de los usuarios; aplicar la nueva tecnología a problemas piloto; adaptar la tecnología para que sea accesible a los usuarios; proveer entrenamiento y educación formal; dar seguimiento a la efectividad del proceso de difusión, promover las reuniones entre usuarios para el intercambio de ideas; no confiar únicamente en los reportes escritos; proveer recursos como tiempo, dinero, personas, etc. para la realización de la difusión y asimilación de la nueva tecnología; y considerar la transferencia de personas junto con la tecnología.

Algunos aspectos de este atributo son:

- a) La participación del usuario final
- b) La participación de los analistas de decisiones
- c) La participación de personal informático
- d) Considerar una nueva función: El administrador del Data Warehouse

Además de los aspectos mencionados, hay que tomar en cuenta también el aspecto de las capacidades humanas mencionadas en la analogía entre arquitectura física y arquitectura de información (ver tabla 3.1), las cuales podrían reflejarse en el diseño del Data Warehouse en la habilidad para usar las nuevas herramientas tecnológicas a implantarse, así como en el nivel de participación del recurso humano en la creación de las estrategias de sumarización y partición de los datos al diseñar la arquitectura de información.

5.2.3. Atributos procedentes de la interacción de las innovaciones y el contexto organizacional

- **Prácticas organizacionales**

Al igual que muchas tecnologías y como se mencionó en el atributo anterior, el desarrollo de un Data Warehouse requiere la participación directa y comprometida del recurso humano para alcanzar los objetivos establecidos; y es el recurso humano el que establece las prácticas organizacionales que pudieran facilitar o dificultar la asimilación de una tecnología. Es importante entonces que en el proceso de asimilación se analicen las prácticas organizacionales de una función determinada de la empresa con relación a dicha tecnología, en el caso del Data Warehouse, la función sería la administración de datos para la toma de decisiones.

En la función de la administración de datos, (ver tabla 2.4) el análisis de las prácticas organizacionales estaría dirigido a identificar las barreras u obstáculos que no permitieran la aplicación exitosa de los procesos que intervienen en dicha capa de la arquitectura de información Data Warehouse, procesos en donde intervienen 2 ó más áreas funcionales o departamentales. Los usuarios de una empresa pueden convertirse en los peores enemigos de la misma cuando se intenta administrar a los datos como un activo corporativo. “La barrera más común para la asimilación de tecnología es el conflicto entre los diferentes grupos funcionales. La diferencia en actitudes y valores generalmente inhibe la comunicación requerida para una adecuada asimilación de tecnología entre las funciones organizacionales” [RAMÍREZ,1993].

Algunos aspectos de este tipo de atributo son:

- a) Nivel de pertenencia de los datos
- b) Cambios culturales sobre la compartición de datos
- c) Políticas de acceso a los datos
- d) Documentación de los diversos sistemas de información y sus bases de datos relacionadas
- e) El compromiso de la alta dirección para la implantación de la tecnología.

5.3. GRADO DE ADOPCIÓN DE LAS INNOVACIONES TECNOLÓGICAS

Además de los atributos que influyen en la asimilación de la tecnología Data Warehouse en una organización, es necesario establecer una serie de medidas que permitan observar el avance en el grado de adopción de la tecnología en la empresa. Lo que puede medirse es factible a evaluarse y mejorarse, de ahí la importancia de establecer ciertas medidas en el grado de adopción de una nueva tecnología, para mejorar así los atributos donde se tengan deficiencias y que representen un obstáculo en su asimilación. “Se entiende por *grado de adopción* a la velocidad relativa con la cual una innovación es adoptada por los miembros de un sistema social. Generalmente es medida como el número de individuos que adoptan una nueva idea en un período específico.” [RAMIREZ,1993]

Rogers [ROGERS,1983] indica que existen cinco atributos por los cuales una innovación puede describirse, y que la forma en que los individuos perciben esas características contribuye a su grado de adopción. Estos atributos relacionados a la innovación son los siguientes:

1. *Ventaja relativa* es el grado con la cual una innovación se percibe como mejor que la idea que sustituye o reemplaza.
2. *Compatibilidad* es el grado con el cual una innovación se percibe por los posibles adoptadores como consistente con sus valores existentes, experiencias pasadas y necesidades.
3. *Complejidad* es el grado con el cual una innovación se percibe como relativamente difícil para entenderla o utilizarla.
4. *Experimentabilidad* es el grado con el cual una innovación puede experimentarse bajo bases limitadas.
5. *Observabilidad* es el grado con el cual los resultados de una innovación son visibles a otros.

De manera paralela a la implantación del Data Warehouse se propone analizar estos 5 atributos para determinar el grado de adopción que va teniendo el recurso humano durante cada etapa de implantación de la nueva tecnología.

5.4. ETAPAS EN EL DISEÑO DE UN DATA WAREHOUSE

A fin de entender mejor el proceso de asimilación de la tecnología Data Warehouse en una empresa, se van a mencionar diversas propuestas de diseño de esta tecnología en una empresa. La finalidad de mencionar dichas propuestas es para identificar similitudes o semejanzas que pudieran aumentar el entendimiento del diseño del Data Warehouse, asimismo para poder establecer la relación de cada etapa de las propuestas con cada componente y proceso de la arquitectura de información Data Warehouse. El lograr un entendimiento del diseño de la arquitectura de información Data Warehouse y no tanto conocer una receta de pasos, va a permitir que el proceso de asimilación de la tecnología y los atributos que en él influyen puedan alinearse a la implantación del Data Warehouse. No se pretende determinar cual es la mejor de las propuestas, ni tampoco proponer una nueva que contemple los elementos que se consideren más importantes de cada propuesta, Niederman y Brancheau indican a este respecto que “el entendimiento no puede sustituirse por una receta de pasos” [NIEDERMAN, 1987]

5.4.1 Propuesta - Ralph Kimball

Kimball propone una metodología compuesta de nueve pasos para el diseño de un Data Warehouse exitoso. La metodología [KIMBALL,1996] consiste de los siguientes pasos:

1. Seleccionar los procesos, y como consecuencia identificar las tablas “fact”
2. Definir lo que representa cada tabla “fact”
3. Identificar y conformar las dimensiones de cada tabla “fact”
4. Seleccionar los “facts” o cálculos de la tabla, incluyendo las que son precalculadas
5. Determinar los atributos de cada dimensión con una descripción completa y una terminología propia
6. Establecer una estrategia de monitoreo para determinar cambios de las dimensiones
7. Definir las estrategias de sumarización, las formas de consulta, las prioridades en las consultas, el almacenamiento físico de los datos, etc.
8. Determinar la duración histórica de la base de datos
9. Establecer la urgencia con que los datos van a ser extraídos y cargados en el Data Warehouse

Aunque la propuesta de Kimball pareciera estar muy enfocada al diseño de las bases de datos desde un punto de vista netamente técnico, en realidad está dirigida a satisfacer ampliamente los requerimientos de información de los usuarios, de acuerdo a la realidad existente de los datos en la organización. Esta propuesta está fundamentada en la modelación dimensional de los datos, “la cual permite a los usuarios finales visualizar a los datos desde un punto de vista del negocio y de acuerdo a cada medida posible en la empresa” [KIMBALL,1996]. Esta característica de la modelación dimensional hace más factible la alineación estratégica entre los objetivos de una empresa y su tecnología de información.

El primer paso de la propuesta consiste en identificar los procesos más prioritarios para la organización, para que de esta manera se comiencen a diseñar las tablas “fact” correspondientes, la figura 5.1 nos presenta de manera general el diseño de las tablas según el enfoque de la modelación dimensional; Ventas es el proceso seleccionado.

- 1) La tabla “fact” es la tabla central mayor del proceso donde se almacenan aquellas medidas numéricas posibles del negocio, además de contener las llaves que permiten la relación con las diversas bases de datos dimensionales que se hayan definido.
- 2) La definición de lo que representa la tabla “fact” consiste en determinar el significado de que lo representa cada registro individualmente.
- 3) Las dimensiones representan las medidas posibles del negocio y por ende la diversidad de reportes que pueden ser requeridos por los usuarios finales del Data Warehouse. Ejemplo: La cantidad total de descuentos que se aplicó a una línea de productos en un mes específico. Las dimensiones representan las medidas posibles del negocio y por ende la diversidad de reportes que pueden ser requeridos por los usuarios finales del Data Warehouse. Ejemplo: La cantidad total de descuentos que se aplicó a una línea de productos en un mes específico.
- 4) Los “facts” de cada tabla representan los campos numéricos que pueden ser sumariados, y los cuales deben estar relacionados con el significado de la tabla “fact”, en este caso del proceso de ventas pueden mencionarse: Las ventas_totales, ingresos_totales, total_descuentos, etc.

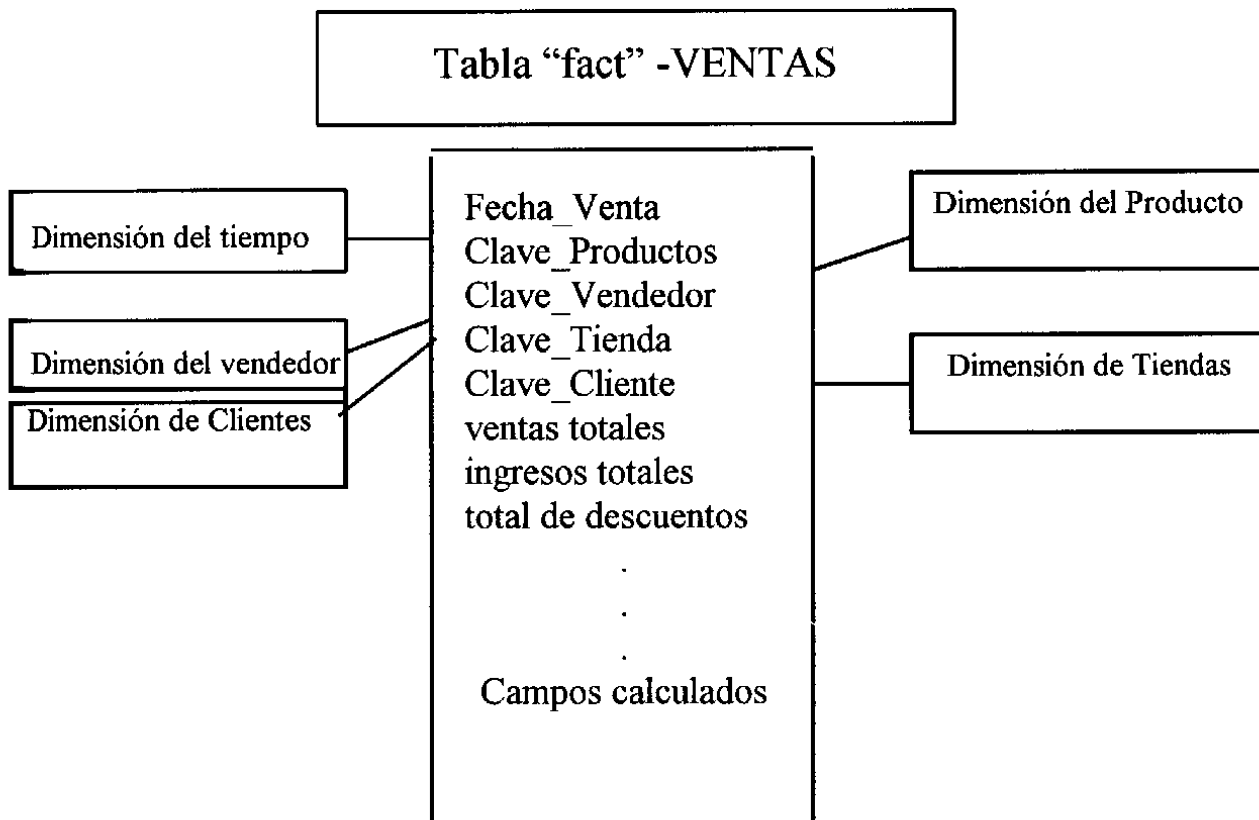


Figura 5.1 – Tabla “fact” y sus dimensiones identificadas [KIMBALL, 1996]

- 5) Los atributos de cada dimensión representan los campos debidamente definidos, guardando una relación con el significado de la tabla “fact”. Por ejemplo si nos referimos a la tabla dimensional de Cliente, un atributo sería la fecha última en que un determinado cliente realizó una compra, el cual estaría relacionado con la llave Fecha_Venta de la tabla “fact”, de tal manera que permita generar las preferencias del cliente y sus fechas de compras.
- 6) Es necesario el monitoreo de los posibles cambios a las tablas dimensionales para cada tabla “fact”
- 7) Las estrategias de sumarización de datos, partición de datos, modos de consulta y otros aspectos físicos del diseño del Data Warehouse son consideradas por Kimball en esta etapa. Cada estrategia de los diversos procesos del Data Warehouse se basa en las tablas “facts” y dimensionales que se hayan definido.

- 8) La duración histórica de las bases de datos se refiere al tiempo contemplado por la tabla “fact” para un determinado proceso, de tal manera que se pueda evitar tener datos inservibles y sin valor histórico para la empresa.
- 9) La urgencia con que los datos son extraídos de las bases de datos operacionales y cargados al Data Warehouse depende del significado de la tabla “fact”, ya que ahí se define el intervalo de tiempo requerido para “refrescar” los datos de dicha tabla.

5.4.2. Propuesta - William Inmon

Inmon [INMON,1996] no menciona una metodología como Kimball, sino más bien se enfoca en los aspectos más importantes relacionados con el diseño del Data Warehouse, y proporciona a la vez un Checklist detallado en forma de cuestionario en donde se identifican y se resuelven los elementos a considerar en dicho diseño.

Según Inmon, el diseño del Data Warehouse debe considerar los siguientes aspectos:

- ⇒ Conocer la situación de los datos operacionales
- Identificar las aplicaciones que serán fuentes de datos para el Data Warehouse
 - Determinar el grado de integración de los datos
 - Conocer la eficiencia en el acceso a los datos de los sistemas de información existentes
 - Considerar la administración del volumen de los datos y el nivel de condensación que debe ser realizado.
- ⇒ Crear un modelo de datos
- Considerar la creación de un modelo de datos corporativo, si no existe, y de ahí derivar un modelo de datos del Data Warehouse
 - Las estructura claves del modelo de datos corporativo deben incluir un elemento en el tiempo
 - El modelo de datos del Data Warehouse debe incluir un análisis de “estabilidad”, esto es, agrupar los datos por su factor de cambio; existen datos que rara vez cambian, los cuales son agrupados con otros de las mismas características, y existen datos que cambian su valor frecuentemente y que son agrupados con otros datos con el mismo ritmo de cambio. El resultado de este

análisis es la creación de grupos con características similares y puede soportar la estrategia de partición de datos. (ver 4.5.1)

- Definir el alcance de integración, es decir, definir las fronteras del modelo de datos de acuerdo al área funcional que se vaya a analizar.
- Relacionar el modelo de datos con el diseño físico de las bases de datos en el ambiente Data Warehouse, impulsados por las estrategias de sumarización y partición de datos que se hayan definido.

⇒ Desarrollar un plan de migración al ambiente Data Warehouse

- Determinar las áreas más críticas de la empresa
- Identificar las relaciones entre las áreas críticas
- Agrupar las áreas críticas por claves y atributos de tal manera que se representen completamente las áreas específicas, debiendo incluir:
 - Los atributos de cada área
 - La claves de cada área
 - Los grupos repetitivos de atributos y claves
 - Las relaciones entre las áreas críticas

El plan de migración va a estar fuertemente influenciado por el modelo de datos

⇒ Determinar el ciclo de los datos

El ciclo de los datos se refiere, según Inmon, al intervalo de tiempo en que un cambio de un dato en el ambiente operacional se ve reflejado en el Data Warehouse. Las estrategias de sumarización, de agregación y de partición de datos determinan en su mayoría el ciclo requerido para los datos.

5.4.3. Propuesta - John Ladaga

Ladaga [LADAGA,1995] propone seis pasos para el diseño de un Data Warehouse, y presenta a esta tecnología como una estrategia administrativa de información para una empresa. Los pasos de esta propuesta son los siguientes:

- 1) Determinar los requerimientos del negocio
- 2) Desarrollar el modelo de datos
- 3) Extraer y convertir datos

- 4) Construir el Data Warehouse
- 5) Desarrollar los Data Marts
- 6) Crear análisis

- 1) Determinar los requerimientos del negocio

Se debe comenzar el diseño de un Data Warehouse identificando los objetivos del negocio o sus factores críticos de éxito para determinar la información que se va a presentar a los usuarios finales. Un error muy común que sucede es el de cargar datos dentro del Data Warehouse solamente porque están accesibles.

- 2) Desarrollar el modelo de datos

Una vez que se definieron los requerimientos del negocio, se debe definir cuales son los datos necesarios para lograr satisfacer dichos requerimientos. Los datos son agrupados lógicamente dentro de un modelo de datos para representar la manera en que el negocio desea realizar sus operaciones

- 3) Extraer y convertir datos

Después de definir los datos necesarios y localizar sus fuentes de origen, el siguiente paso consiste en extraer los datos y convertirlos al ambiente Data Warehouse. La conversión y la limpieza de datos, así como la alineación en la definición de los datos son las actividades que consumen un mayor tiempo en la construcción de un Data Warehouse.

- 4) Construir el Data Warehouse

Básicamente Ladaga nos presenta los siguientes 3 puntos:

- a) Determinar los requerimientos técnicos para la arquitectura de información
- b) Definir la configuración funcional que consta de las herramientas usadas para la consulta de datos, accesos, análisis y reportes
- c) Adquirir, instalar y probar la infraestructura y las herramientas computacionales

- 5) Desarrollar los Data Marts

Desarrollar para cada área funcional de la empresa un Data Mart donde se almacene información sumariada y propia a dicha área funcional. Los usuarios

tienen la libertad de emplear las herramientas que deseen para acceder a un Data Mart específico.

6) Crear análisis

La creación del análisis es la razón por la que se planea el diseño de un Data Warehouse. Aquí se pueden emplear herramientas como los Sistemas de Información Ejecutivos, los Sistemas de Soporte a la Decisión, reportadores diversos, etc. Estas herramientas pueden obtener los datos de los Data Marts o directamente de las bases de datos del Data Warehouse.

5.4.4. Propuesta - A.J. Brown

A.J. Brown propone 10 pasos para una construcción exitosa de un Data Warehouse. Para Brown la meta principal en el diseño de un Data Warehouse es consolidar y a la vez separar los datos usados en los análisis del negocio, pudiendo lograrlo mediante un Manejador de Base de Datos de alto rendimiento que permita un acceso flexible a las bases de datos desde cualquier herramienta de usuario final. Los pasos que propone son los siguientes [BROWN,1995]:

- 1) Definir claramente los objetivos del negocio que se van a perseguir con el diseño del Data Warehouse
- 2) Identificar los criterios de medición que permitan determinar el grado de éxito durante la implantación
- 3) Determinar los datos que van a ser requeridos por el Data Warehouse
- 4) Diseñar el Data Warehouse. Emplear como soporte las técnicas de modelación de datos.
- 5) Estimar el tamaño inicial de las bases de datos en el arranque del proyecto, así como proyectar la posible capacidad en el futuro.
- 6) Determinar la frecuencia con que los nuevos datos van a ser añadidos al Data Warehouse, estimando el tamaño de cada agregación.
- 7) Realizar una prueba piloto con datos reales en alguna área específica de la empresa.

- 8) Dejar a los usuarios trabajar con la prueba piloto por un tiempo determinado, a fin de obtener retroalimentación que mejore el diseño.
- 9) Evaluar una lista corta de vendedores de las diversas herramientas Data Warehouse
- 10) Invertir en programas de capacitación en diversas tecnologías de información relacionadas con el Data Warehouse, tales como: la tecnología Cliente-Servidor, software de conectividad, Sistemas operativos de red, etc.

5.4.5. Análisis comparativo de las propuestas de diseño de un Data Warehouse

Cada una de las propuestas de diseño del Data Warehouse presentan una serie de etapas o aspectos relevantes a tomar en cuenta para implantar la arquitectura de información Data Warehouse; entre las propuestas se pueden visualizar similitudes y enfoques en común que podrían facilitar la complementación de cada propuesta para crear una nueva de mayor alcance. Pero más que buscar la complementación de las propuestas, el enfoque en esta tesis es detectar similitudes basándose en los aspectos de la arquitectura de información Data Warehouse que pudieran facilitar el proceso de asimilación tecnológica en una organización. Los aspectos del Data Warehouse a considerarse como bases para identificar similitudes entre las propuestas fueron tratados en los capítulos anteriores de esta tesis, siendo la meta final lograr un adecuado grado de aceptación de la innovación tecnológica para que pueda ser implantada exitosamente.

Las tablas 5.1, 5.2 y 5.3 presentan la agrupación por similitudes de las propuestas de diseño del Data Warehouse, primeramente se presentan en base a los procesos que conforman la arquitectura de información Data Warehouse (ver tabla 2.4), con el fin de asociar cada una de las etapas con los respectivos procesos de la arquitectura de información que se proponen en esta tesis. Así también se considera el aspecto de la alineación de los objetivos del negocio con los de la tecnología Data Warehouse visto en el apartado 4.6, y por supuesto se contemplan los atributos que influyen en el proceso de asimilación de la tecnología Data Warehouse presentados en el apartado 5.2.

1) Procesos de la arquitectura de información Data Warehouse		
Etapas	Proceso	Autor
Definir lo que representa cada tabla "fact"	Sumarización	Kimball
Identificar y conformar las dimensiones de cada tabla "fact"	Partición	Kimball
Definir las estrategias de sumarización, las formas de consultas, las prioridades en las consultas, el almacenamiento físico de los datos, etc.	Sumarización y partición de los datos	Kimball
Establecer la urgencia y el tamaño con que los datos van a ser extraídos de las bases de datos operacionales y cargados al Data Warehouse	Extracción de datos Conversión de datos	Kimball, Inmon, Brown
Crear un modelo de datos	Creación de modelos de datos	Inmon, Ladaga
Desarrollar los Data Marts	Sumarización y partición de los datos	Ladaga
Crear análisis	Presentación y análisis de información	Ladaga

Tabla 5.1 – Tabla comparativa en el diseño del Data Warehouse, punto 1

En el primer punto comparativo, presentado en la tabla 5.1, se puede identificar que los procesos de sumarización y partición de datos forman parte directa o indirectamente de alguna de las propuestas presentadas, ya sea relacionado con el diseño de las bases de datos o con la creación de conceptos o herramientas de software como los Data Marts o los sistemas de información ejecutivos (EIS). El objetivo de analizar este primer punto es que los diseñadores de un Data Warehouse asocien las etapas de mayor implicación tecnológica (diseño de bases de datos, desarrollo de interfaces para extracción y conversión de datos, desarrollo de sistemas de información, etc.) con los procesos más importantes de la arquitectura de información Data Warehouse.

El enfoque del segundo punto comparativo presentado en la tabla 5.2, esta dirigido a reflejar los objetivos y requerimientos del negocio en el diseño del Data Warehouse, tal como lo dice Kimball al hacer mención de la selección de los procesos más importantes de la empresa para que en base a ellos se diseñen las tablas en los diversos niveles de detalle de la arquitectura de información Data Warehouse. Por lo tanto, la agrupación de las propuestas en este punto se hace con el fin de que al diseñar un Data Warehouse, los objetivos y procesos del negocio sean el punto de partida para establecer los procesos del Data Warehouse y la aplicación de las tecnologías correspondientes para realizarlos.

2) Alineación del Data Warehouse a los objetivos del negocio	
Etapas	Autor
Seleccionar los procesos, y como consecuencia identificar las tablas “fact”	Kimball
Desarrollar un plan de migración al ambiente Data Warehouse determinando las áreas más críticas de la empresa	Inmon
Determinar los requerimientos del negocio	Ladaga
Definir claramente los objetivos del negocio que se van a perseguir con el diseño del Data Warehouse	Brown

Tabla 5.2 – Tabla comparativa en el diseño del Data Warehouse, punto 2

Con respecto a la tabla 5.3, se hace referencia de manera más directa al proceso de asimilación tecnológica del Data Warehouse en una empresa, mediante los atributos que pueden influir en tal proceso; según los autores, el esfuerzo debe estar dirigido a conocer la situación de los datos, así como las herramientas tecnológicas del Data Warehouse, su evaluación y aplicación. La agrupación de estas propuestas se hace con el fin de identificar las etapas en el diseño del Data Warehouse que pueden influir más directamente en la asimilación tecnológica, por supuesto que cada etapa del diseño puede influir en el grado de asimilación de la tecnología, pero quizá no tan directamente como las etapas que se establecen en este grupo. Como se observa en la tabla 5.3 solamente se presentan atributos organizacionales y atributos de la innovación, pero quizá los más importantes en cuanto a su grado de influencia, son los atributos derivados de la interacción de las innovaciones y el contexto organizacional (ver 5.2.3.), debido a que están más relacionados con las prácticas organizacionales en el uso de las tecnologías de información y el manejo de los datos; tales atributos más que presentarse en una etapa del diseño del Data Warehouse, se presentan en la implantación y uso de la tecnología y en cualquier etapa del proceso de asimilación, estando más relacionados con la cultura organizacional y el recurso humano.

Las agrupaciones que se han hecho de las diversas propuestas de desarrollo del Data Warehouse fueron establecidas para apoyar el análisis del proceso de asimilación de la tecnología Data Warehouse que a continuación se presenta. La agrupación fue realizada conforme a las propuestas mencionadas en este capítulo y queda abierta para otras propuestas o metodologías de diseño de un Data Warehouse.

3) Atributos que afectan al proceso de asimilación tecnológica		
Etapas	Atributo	Autor
Conocer la situación de los datos operacionales	Atributo organizacional	Inmon
Determinar los requerimientos técnicos para la arquitectura de información Data Warehouse	Atributo de la innovación	Ladaga
Definir la configuración funcional que consta de las herramientas usadas para la consulta de datos, accesos, análisis y reportes	Atributo de la innovación	Ladaga
Determinar los datos que van a ser requeridos por el Data Warehouse	Atributo organizacional	Brown
Evaluar una lista de vendedores de las diversas herramientas Data Warehouse	Atributo de la innovación	Brown
Invertir en programas de capacitación en diversas tecnologías de información relacionadas con el Data Warehouse	Atributo de la innovación	Brown

Tabla 5.3 – Tabla comparativa en el diseño del Data Warehouse, punto 3

5.5. PROCESO DE ASIMILACION DE LA ARQUITECTURA DE INFORMACIÓN DATA WAREHOUSE EN UNA ORGANIZACION

El análisis que a continuación se presenta está enfocado a analizar detalladamente una guía de asimilación de tecnología de información aplicada a la arquitectura Data Warehouse. En esta tesis se toman como referencia las etapas de dicha guía para definir las actividades a realizar y los aspectos a considerar específicamente para el Data Warehouse. En cada actividad y aspecto a considerar se involucran todos los procesos, componentes, atributos y demás aspectos de la arquitectura de información Data Warehouse que se han estado mencionando en los capítulos anteriores; asimismo se identifican las etapas de diseño del Data Warehouse basándose en las agrupaciones realizadas en el apartado 5.4; lo que se pretende es enriquecer cada etapa de la guía de asimilación de tecnología y al final de cada fase evaluar el grado de adopción de la tecnología Data Warehouse con ciertas medidas previamente establecidas.

Nos podríamos preguntar si realmente es necesario analizar la asimilación del Data Warehouse durante su proceso de implantación en una empresa, o quizá puede resultar un esfuerzo infructuoso, y que con solo planear adecuadamente la introducción de la tecnología puede resultar suficiente para garantizar el éxito de la implantación. Pues bien, la respuesta a esta cuestión dependería del concepto que se tenga del Data Warehouse en una empresa y de su aplicación como tecnología, quizá se puede visualizar al Data Warehouse como una herramienta de software asociada a un reporteador, o bien se puede ver como un conjunto de bases de datos conectadas e integradas; sin embargo, para fines de esta tesis se propone al Data Warehouse como una arquitectura de información con un enfoque sumariado e histórico que apoya al proceso de toma de decisiones en los niveles táctico y estratégico de una organización; tomando como base esta definición se podría considerar al Data Warehouse como una innovación tecnología que involucra la aplicación de nuevas maneras de administrar los datos y la información de un negocio. Lo anterior puede implicar que se presenten cambios tecnológicos y organizacionales en la empresa para que el Data Warehouse pueda ser implantado, y es precisamente con un proceso de asimilación de tecnología que tales cambios pueden administrarse más adecuadamente para evitar un impacto negativo de los mismos en las operaciones de la organización; y por otro lado que la tecnología pueda ser aceptada e involucrada en los procesos del negocio desde su diseño, implantación y puesta en marcha.

Ahora bien, basarse en una guía de asimilación de tecnología puede implicar perder de vista el orden de las etapas de diseño del Data Warehouse, y con ello la pérdida de control en la implantación, por lo que es necesario que se relacione adecuadamente la guía de asimilación de tecnología con las etapas de diseño del Data Warehouse, por lo que en la evaluación que se propone al final de cada fase se haga hincapié en dicha relación. Se observa en las diversas propuestas del diseño de un Data Warehouse una inclinación hacia aspectos técnicos muy específicos, tal como el diseño de las bases de datos y sus tablas correspondientes; para el diseño en sí del Data Warehouse, el conocimiento y aplicación de tales actividades son esenciales para alcanzar el objetivo final que se desea, pero con esto no se asegura que la organización acepte totalmente a la tecnología Data Warehouse y la sienta como una herramienta tecnológica de uso cotidiano para realizar sus funciones, es así que los diseñadores del Data Warehouse deben medir el avance en la implantación de la tecnología pero también el avance en el grado de adopción de la misma en forma paralela.

La guía que sirve como base en la realización del análisis de asimilación Data Warehouse [RAMÍREZ,1993] fue seleccionada por estar enfocada específicamente a las tecnologías de información, así como por estar fundamentada en los modelos de etapas de Meyer y Goes [Meyer,1988], complementándose con otros estudios realizados y modelos propuestos ya probados (Rogers, McFarlan y McKenney, etc.). Las fases de esta guía se presentan en la tabla 5.4.

Durante el análisis del proceso de asimilación se presenta con letra *itálica* lo que corresponde a la guía general de asimilación tecnológica y con letra estándar se presentan las propuestas de esta tesis específicamente para la innovación Data Warehouse.

Al final de cada fase se presenta una evaluación del Data Warehouse, y una serie de medidas que podrían ser indicativas del grado de adopción de la tecnología para cada fase del proceso. La evaluación consiste en identificar los atributos que influyen en el proceso de asimilación del Data Warehouse (ver 5.2.), así como el establecer las etapas de diseño del Data Warehouse que puedan tener relación directa con las actividades de la fase en análisis. Y como parámetros del grado de adopción se contemplan cinco atributos que son medidos de acuerdo a una serie de preguntas que contemplan a los conceptos, procesos y componentes de la arquitectura de información Data Warehouse. La finalidad es aplicar estas preguntas a los miembros de la organización relacionados con la implantación del Data Warehouse, y determinar así el grado de adopción de la tecnología en la empresa, lo cual a su vez puede permitir soportar la toma de acciones para mantener o cambiar el rumbo del plan de implantación de la innovación tecnológica. Los miembros de la empresa que sean consultados pueden responder las preguntas en base a tres medidas de aceptación de la tecnología, estas son, alta, media y baja, y de esta manera al final de la aplicación de las encuestas, se puedan sumar los valores y determinar así el grado de adopción de cada medida de evaluación. Después de la evaluación se describe textualmente la selección de las etapas de diseño del Data Warehouse, así como el objetivo que se pretende con cada medida de evaluación.

	Fases		Etapas	Tipo de atributo que influye
I	Investigación	I.1.	Búsqueda de tecnología disponible	Atributos de la innovación
II	Conocimiento-Persuasión	II.1.	Conocimiento de la tecnología	Atributos de la innovación y de la organización
		II.2.	Establecimiento de las necesidades de la organización	
		II.3.	Consideración y persuasión de la adopción de la tecnología	
		II.4.	Discusión para la adquisición de la tecnología	
III	Evaluación-Decisión	III.1.	Proposición de adquisición de la tecnología	Atributos de la innovación y de la organización
		III.2.	Evaluación racional-financiera	
		III.3.	Evaluación política-estratégica	
		III.4.	Evaluación de aptitud organizacional	
		III.5.	Pruebas o experimentación de la tecnología	
		III.6.	Decisión de adopción, rechazo o posposición del uso de la tecnología	
IV	Introducción-Adopción	IV.1.	Redefinición, reestructuración y planeación de la introducción de la tecnología	Atributos de la interacción de la innovación y la organización
		IV.2.	Aprendizaje y clarificación del uso de la tecnología	
		IV.3.	Difusión y reconocimiento de logros del uso de la tecnología	
		IV.4.	Adopción, aceptación y rutinización del uso de la tecnología	
V	Confirmación-Expansión	V.1.	Evaluación del uso de la tecnología	Atributos de la interacción de la innovación y la organización
		V.2.	Confirmación del uso de la tecnología	
		V.3.	Expansión e institucionalización del uso de la tecnología	

Tabla 5.4 – Fases de la Guía para la Administración del Proceso de Asimilación de Tecnología de Información en una Organización [RAMÍREZ, 1993]

FASE I: INVESTIGACION

El objetivo de esta fase es que la organización esté en contacto con el ambiente tecnológico externo a ella.

ETAPAS

I.1. BÚSQUEDA DE TECNOLOGÍA DISPONIBLE

El objetivo es tener contacto con la TI disponible, para seleccionar la más apropiada para satisfacer las necesidades de la organización.

Una organización en un ambiente competitivo debe estar buscando siempre nuevas tecnologías con las cuales pudiera mejorar su posición dentro de su rama industrial o bien, permitirle mantener su ubicación exitosamente. Dentro del ambiente de las tecnologías tenemos a las tecnologías de información (TI), que en los últimos años han dado un soporte muy importante a las empresas en el logro de ventajas competitivas mediante el mejoramiento de sus procesos. Entre las tecnologías de información se encuentra el Data Warehouse como una alternativa para solucionar problemas relacionados al manejo de la información. Para determinar si se evalúa al Data Warehouse como una tecnología que pudiera solucionar las necesidades de información de una empresa, primeramente es necesario conocer sus áreas de aplicación.

“Se han identificado cuatro ambientes que son considerados como serios candidatos para un Data Warehouse:

- Datos sucios
- Alto nivel de análisis
- Tiempo real
- Gran volumen de datos” [MYERS,1995]

Si una organización tiene alguno o varios de estos ambientes, es necesario que evalúe la posibilidad de empezar un proyecto Data Warehouse como una solución a los problemas que pueden derivarse de dichos ambientes, (ver apartado 2.2):

- Falta de credibilidad en los datos
- Baja productividad
- Falta de capacidad para transformar datos en información

Es importante aplicar los conceptos de alineación de la tecnología Data Warehouse a los objetivos del negocio, mencionados en el apartado 4.6, y mediante un plan

informático determinar las necesidades más importantes en una empresa y las alternativas tecnológicas para satisfacerlas, permitiendo así evaluar con mayor claridad las tecnologías de información disponibles. La evaluación de la tecnología disponible en esta etapa debe realizarla el departamento de informática o sistemas de información de la empresa, ya que por medio de ellos se da el contacto principal con las innovaciones en el área de las tecnologías de información, además cuentan con un conocimiento de la empresa y de sus necesidades más importantes.

Actividades a realizar y aspectos a considerar para el Data Warehouse

- La TI disponible, así como la TI que pronto va a estar disponible

Además del Data Warehouse se deben evaluar algunas tecnologías como Cliente-Servidor, redes de computo, sistemas manejadores de bases de datos relacionales (RDBMS), etc. “Puede decirse que existen 3 componentes técnicos básicos para cada Data Warehouse:

- Los métodos de extracción y movimiento de datos
- El software de conectividad para una integración cliente-servidor
- Un RDBMS de alto rendimiento” [BROWN,1995]

⇒ Los métodos de extracción y movimiento de datos son herramientas propias de un Data Warehouse; es necesario conocer lo que proporciona cada herramienta para determinar su grado de aplicación en el diseño de esta tecnología. A través de estas herramientas de software se logra extraer datos de las bases de datos operacionales, se “limpian” y posteriormente se mueven a las bases de datos históricas del Data Warehouse. Los coordinadores del proyecto deben conocer las herramientas que ya existen en el mercado para poder seleccionar, o bien crear internamente, las que mejor pudieran adaptarse a la situación del ambiente operacional de los datos, y a la vez que presenten compatibilidad con el RDBMS que se tenga en la empresa.

La búsqueda de las herramientas de software para la extracción y movimiento de los datos en esta primera fase va a permitir establecer con bases más sólidas las estrategias de extracción y conversión de datos como procesos del Data Warehouse. En la tabla 5.1 podemos observar que estos procesos del Data Warehouse son considerados por tres autores cuando hablan de definir la urgencia, frecuencia y el tamaño con que los datos van a ser extraídos de las bases de datos operacionales y convertidos a un formato Data Warehouse. La evaluación previa de estas herramientas permitiría conocer la capacidad que tienen para soportar las cargas en el procesamiento de los datos y los tiempos que

pueden emplear en cada uno de los procesos; además, el conocimiento previo de estas herramientas de software va a permitir apoyar no solamente los procesos de conversión del nivel operacional al Data Warehouse, sino también en cada nivel de sumariación o detalle en donde se requiere algún proceso de extracción y conversión de datos, hay que tomar en cuenta que la conversión de datos es una de las actividades que consumen mayor tiempo en la operación de un Data Warehouse.

⇒ Software de conectividad para una integración cliente-servidor. El objetivo que se persigue al hablar de conectividad en el Data Warehouse es el de proporcionar a los usuarios finales el acceso rápido a las bases de datos de dicha tecnología desde su propia PC o estación de trabajo. Es importante establecer entonces, los protocolos de conectividad que permitan la actualización de datos de un Data Mart o de una base de datos sumariada, directamente de las bases de datos operacionales o de las bases de datos del Data Warehouse, ya que posiblemente un Data Mart se conforme por datos que se encuentran físicamente distantes, y en donde se requiere tener una conectividad eficiente para lograr su integración.

La arquitectura Data Warehouse se caracteriza por tener una infraestructura que soporta el flujo de la información en una empresa, la dinámica en el flujo de la información presentada en la figura 2.10 sería muy difícil sin las herramientas de conectividad adecuadas. Es necesario tener conocimiento de los protocolos y software de comunicación, de las herramientas middleware, ODBC, etc., que pudieran adaptarse a la TI (hardware, software, comunicaciones) de la empresa. El estudio de las herramientas de conectividad permitiría soportar cualquier proceso del Data Warehouse involucrado con la dinámica del flujo de información de esta arquitectura, principalmente en los procesos correspondientes a las capas de transformación y administración de datos.

⇒ Un RDBMS de alto rendimiento

“Tal vez la elección más crítica cuando se construye un Data Warehouse es la selección de un RDBMS donde residirán los datos que permitirán la toma de decisiones; el software seleccionado debe contemplar los siguientes puntos:

- Un buen rendimiento en consultas complejas
- Capacidad de realizar procesamientos analíticos
- Habilidad para manejar grandes volúmenes de datos” [BROWN,1995].

Los encargados de evaluar el RDBMS para el Data Warehouse deben contemplar lo que existe en el mercado y lo que ofrecen los vendedores de tal tecnología, sobre todo conocer oportunamente las mejoras o adaptaciones de los Manejadores de Bases de Datos más conocidos, principalmente en lo que se

refiere a ejecutarse en ambientes más analíticos que operacionales. La arquitectura Data Warehouse involucra el concepto de base de datos en todos sus procesos, es por eso la importancia de seleccionar el manejador de base de datos que mejor soporte a cada uno de los procesos y conceptos, enfocando el esfuerzo de evaluación a determinar el manejador que pueda soportar un mejor procesamiento analítico de las bases de datos operacionales. La evaluación también puede basarse en las características de las bases de datos del Data Warehouse presentadas en la tabla 2.1.

- El grupo responsable de la búsqueda de tecnología debe tener conocimientos e intereses no sólo técnicos de TI, sino también del negocio, para identificar cuáles tareas son factibles de automatizarse.

Ladaga en la primera etapa de su propuesta nos menciona la importancia de definir los requerimientos del negocio a través de la identificación de sus objetivos o factores críticos de éxito antes de comenzar un Data Warehouse. El estudio de los factores críticos de éxito [ROCKART,1979] permite entender la situación del negocio y tiene dos efectos en los individuos, que en este caso es el grupo responsable de la búsqueda de tecnología; primero, ayuda a enfocar los esfuerzos en las cosas importantes, segundo, ayuda a pensar en la información requerida para monitorearlos. Con este análisis se gana conocimiento del negocio y permite evaluar la tecnología que pueda soportar el monitoreo de la información inherente a los factores críticos del mismo.

- Mantener un panorama amplio de la organización como un sistema interconectado.

Al momento de evaluar una tecnología y en este caso el Data Warehouse, es recomendable visualizar a la empresa como un todo y no como un conjunto de islas que trabajan independientemente y con metas propias. El Data Warehouse es una arquitectura de información basada en diversos niveles de detalle de los datos; mientras menor sea el detalle que exista en los datos, significa que se han aplicado mayor cantidad de filtros y estrategias de sumarización de diversas fuentes de datos para conformar la información que se requiere para determinado análisis en la empresa. En las estrategias de sumarización y en el diseño de un Data Mart se requiere entender al negocio como un sistema interconectado. Para diseñar una estrategia que permita la creación de un Data Mart se debe entender el negocio y su cadena de valor, los procesos del negocio que soportan la cadena de valor y las decisiones que afectan a la organización. El conocimiento de esto, combinado con

un entendimiento de un modelo corporativo de datos, ayudarán a desarrollar las particiones en las tablas y las agregaciones que más puedan optimizar el rendimiento del Data Warehouse.

Visualizar a la empresa utilizando la cadena de valor o desde un punto de vista de procesos, permitiría evaluar más adecuadamente las tecnologías que se requieran para soportar las necesidades globales de la organización. En el capítulo 4 se identifican 3 tipos de análisis que pueden realizarse en una empresa para fines del Data Warehouse, estos son, el flujo de información en forma vertical, el flujo de información en forma horizontal, y la relación datos-procesos; estos tipos de análisis pueden ser la base para visualizar a una empresa como un sistema interconectado.

- Pensar en términos no sólo de las actividades actuales de la empresa, sino también en nuevas formas posibles de operar.

La tecnología que se vaya a seleccionar debe tener la suficiente flexibilidad para adaptarse a los cambios organizacionales que se presenten en el negocio. El Data Warehouse, tal como lo definió Inmon, está orientado hacia los tópicos del negocio, y Kimball con su modelación dimensional propone que cada tabla dimensional sea una medida posible de la empresa; es decir, en ambos diseños se refleja la forma de operar de una empresa en el diseño del Data Warehouse. La creación de una nueva forma de operar dentro del negocio sería una nueva dimensión en el diseño de las bases de datos, y se contemplaría también la modificación en las estrategias de sumarización y partición de los datos. Una correcta evaluación para identificar las herramientas Data Warehouse más flexibles a cambios organizacionales permitirá que al momento de diseñar las dimensiones de las tablas “fact”, estas sean diseñadas con una estructura flexible y dinámica. La administración de un Data Warehouse debe involucrar un proceso de mejora continua de monitoreo a los cambios en los sistemas operacionales y en el medio ambiente cambiante del negocio.

- Realizar la búsqueda de oportunidades en las siguientes áreas: TI emergente, productos en el mercado, relaciones interorganizacionales y dentro de su propio ambiente organizacional.

Se observa un crecimiento muy acelerado en la creación de herramientas de software en diversas tecnologías de información, y en el mercado del Data Warehouse no ha sido la excepción. “En los últimos 3 años, cientos de herramientas Data Warehouse han sido introducidas o redireccionadas hacia el mercado informático. Más sin embargo, existen procesos del Data Warehouse que todavía se

llevan a cabo manualmente debido a la inmadurez de dichas herramientas” [ROBINSON,1996,1]. Es un verdadero reto la búsqueda y selección adecuada de las herramientas de software que nos van a ayudar a resolver un problema en la organización y más aún tratándose de la tecnología Data Warehouse. “La selección de los productos apropiados que no solamente hagan bien su función, sino además que trabajen en conjunto, es un desafío muy fuerte y requiere mucho tiempo y recursos para evaluarlos” [STRANGE, 1995].

Evaluar una lista de vendedores de las diversas herramientas Data Warehouse es una etapa de diseño propuesta por Brown, siendo también un atributo que afecta a este proceso de asimilación, ya que de la adecuada identificación de los diversos proveedores de software y sus productos se pueden tener respuestas más rápidas durante el diseño e implantación del Data Warehouse. Es decir, es tan importante conocer la tecnología disponible como a los proveedores de dicha tecnología. La gráfica de la figura 3.3 nos muestra los “ingredientes” claves en el diseño de un Data Warehouse, y para lo cual cada uno de ellos requiere un software específico. Dicha gráfica del capítulo 3 agrupa a las herramientas tecnológicas de acuerdo a la función que realizan, generando con esto que se pueda llevar a cabo una mejor evaluación de las mismas, así como para tener un mejor entendimiento del papel que juega cada herramienta en la arquitectura de información Data Warehouse, y las habilidades que debe tener el recurso humano para poder operarlas.

GRADO DE ADOPCION DE LA INNOVACION DATA WAREHOUSE

FASE I: INVESTIGACION

Atributos que influyen en el proceso de asimilación	Ventaja relativa			Compatibilidad			Complejidad			Experimentabilidad			Observabilidad		
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
<p>* Investigación de la Tecnología de información requerida para diseñar el Data Warehouse</p> <p align="center">↓ ↓ ↓ ↓</p> <p>* Definir los requerimientos del negocio</p> <p>* Evaluar una lista de vendedores de las diversas herramientas Data Warehouse</p> <p>* Identificar y conformar las dimensiones de cada tabla "fact"</p> <p>* Establecer la urgencia y el tamaño con que los datos van a ser extraídos de las bases operacionales y cargados al Data Warehouse</p> <p>Etapas del Data Warehouse involucradas</p>				<p>- ¿Apoya el DW a los objetivos del negocio o a sus Factores Críticos de Éxito?</p> <p>- ¿Los usuarios necesitan mejores formas de acceder la información?</p> <p>- ¿Es factible a mejorar el proceso de toma de decisiones de la empresa, empleado el Data Warehouse?</p>	<p>- ¿La tecnología evaluada es compatible a las plataformas de TI existentes?</p> <p>- ¿La tecnología evaluada se adapta a los procesos del DW que se planean implantar?</p>	<p>- ¿Es difícil la búsqueda, evaluación y adquisición de los productos Data Warehouse existentes en el mercado?</p> <p>- ¿Es una tecnología flexible a cambios organizacionales?</p>	No aplica			No aplica					

Grado de Adaptación

- 1 - Bajo
- 2 - Medio
- 3 - Alto

Grado de Adaptación (Complejidad)

- 1 - Alta
- 2 - Media
- 3 - Baja

Tabla 5.5

DESCRIPCIÓN DE LA TABLA 5.5 – GRADO DE ADOPCIÓN, FASE I

Debido al poco conocimiento y experiencia que se tiene del Data Warehouse, puede resultar compleja la *investigación* de las tecnologías de información requeridas para realizar su diseño. Así este atributo de la innovación puede influir fuertemente en la asimilación de la tecnología durante su implantación, identificándose en las siguientes etapas del diseño del Data Warehouse una influencia más directa.

a) Definición de los requerimientos del negocio

La definición de los requerimientos del negocio por medio de los objetivos o factores críticos de éxito, puede orientar la búsqueda de las tecnologías hacia aquéllas que satisfagan directamente tales requerimientos. La falta de una definición correcta de los requerimientos del negocio para los cuales se planea desarrollar el Data Warehouse, puede provocar que la evaluación de las tecnologías sea mal enfocada y sin una alineación real a la empresa. Esto último generaría incremento en los costos y retardos en tiempo, y principalmente en una falta de aceptación de las herramientas tecnológicas en la organización. Si por otro lado se definen adecuadamente los requerimientos del negocio, es factible establecer los objetivos del Data Warehouse orientados a satisfacer dichos requerimientos, y basándose en los objetivos, determinar la información que se necesita para lograrlos, permitiendo que la evaluación de las tecnologías de información se oriente a los procesos de la arquitectura de información Data Warehouse que generarían dicha información.

b) Evaluar una lista de vendedores de la diversas herramientas Data Warehouse

Pueden haber muchos proveedores de herramientas Data Warehouse en el mercado informático, y muchas soluciones para el diseño de esta tecnología, pero para que la evaluación no consuma mucho tiempo y recursos, es importante identificar a proveedores que por su probada capacidad y experiencia ofrezcan soluciones “completas” para el diseño del Data Warehouse, y que a la vez sean compatibles a la plataforma tecnológica existente en la organización. De esta manera se puede lograr que la evaluación de los proveedores y sus herramientas Data Warehouse se realice más rápidamente y conforme a los tiempos establecidos de planeación.

c) Identificar y conformar las dimensiones de cada tabla “fact”

Esta etapa del diseño del Data Warehouse esta enfocada principalmente hacia aspectos muy técnicos como el diseño de las bases de datos, pero es importante recalcar que al momento de establecer las dimensiones de las tablas “fact” se están definiendo los tópicos del negocio en donde el Data Warehouse va a apoyar para la toma de decisiones. Es así, que la evaluación de las herramientas Data Warehouse debe realizarse tomando en cuenta las operaciones de la empresa y también las nuevas formas posibles de operar, o bien los tópicos futuros del negocio, de tal manera que se orienten los esfuerzos de evaluación hacia las herramientas que ofrezcan mayor flexibilidad en el diseño de las tablas.

d) Establecer la urgencia y el tamaño con que los datos van a ser extraídos de las bases de datos operacionales y cargados al Data Warehouse.

La definición de los procesos de extracción y conversión dentro de la arquitectura de información Data Warehouse va a permitir establecer la capacidad necesaria de procesamiento en los datos, en cuanto al tiempo de respuesta en cada proceso, la compatibilidad en los formatos de los datos para realizar las conversiones o integraciones, etc. Se requiere entonces establecer las características que deben tener las herramientas Data Warehouse para que puedan aplicarse eficientemente en los procesos de la arquitectura de información. Dichas características pueden considerarse como un filtro más para evaluar a los proveedores del Data Warehouse.

Medidas de Evaluación para el grado de adopción

Ventaja relativa

En esta fase de investigación se intenta establecer si las herramientas Data Warehouse que se están evaluando, apoyan realmente la creación de una arquitectura de información que soporte el proceso de toma de decisiones en los niveles táctico y estratégico de una empresa. Esto implica que los accesos a la información relevante se lleven a cabo más eficientemente, ya sea para consultas o para la creación de todo tipo de reportes ejecutivos.

Compatibilidad

Con esta medida de evaluación se intenta definir si las herramientas Data Warehouse que se estén investigando son totalmente compatibles a las tecnologías de información existente en la empresa. Además de los aspectos tecnológicos,

también es posible detectar si las herramientas son adaptables a los procesos planeados para aplicarse en la arquitectura de información Data Warehouse.

Complejidad

A través de esta medida se puede determinar si la evaluación de las herramientas Data Warehouse esta resultando muy compleja y costosa en tiempo y en recursos. Debe enfatizarse en la selección de los proveedores a evaluarse y en la búsqueda de una tecnología que se adapte a las condiciones actuales y futuras de la empresa.

Experimentabilidad

Esta medida de evaluación no aplica en esta fase del proceso de asimilación, debido a que la innovación solamente se está investigando y evaluando y no existe ninguna prueba experimental realizada hasta el momento.

Observabilidad

Esta medida no aplica en esta fase de investigación debido a que no se han realizado pruebas que permitan visualizar resultados.

FASE II: CONOCIMIENTO-PERSUASION

El objetivo de esta fase es que los miembros de la organización obtengan conocimiento de la existencia y del uso potencial de la TI en su industria; que lleguen a formarse una actitud hacia la innovación; y comparen la posible correspondencia entre la innovación y las necesidades de la organización.

La búsqueda de la tecnología de información disponible de la fase anterior estaba a cargo principalmente por el departamento de informática o sistemas de información de una empresa y no se involucraba al usuario de la tecnología directamente. En esta segunda fase la participación de los usuarios relacionados con los procesos de información es importante para alcanzar un conocimiento de la tecnología y su posible aplicación. Sobre la base de este conocimiento el usuario puede empezar a formarse una actitud de aceptación o rechazo hacia la innovación, lo cual puede facilitar a los promotores del Data Warehouse tomar las acciones necesarias en cada caso. Asimismo, el conocimiento que tienen los usuarios del área de negocio en que se encuentran puede ser aprovechada por los diseñadores del Data Warehouse para crear la arquitectura de información conforme a las necesidades de la organización.

ETAPAS

II.1 CONOCIMIENTO DE LA TECNOLOGÍA

El objetivo de esta etapa es que los miembros de la organización obtengan entendimiento de las funciones de la innovación.

Los usuarios de las tecnologías de información generalmente desean soluciones rápidas a sus problemas más comunes, de tal manera que al evaluarse una nueva tecnología lo que buscan en ella es que pueda resolver su situación problemática en el menor tiempo posible. De esta manera, es importante que la difusión de la tecnología Data Warehouse a través de una organización sea guiada hacia los beneficios que puede generar tanto para la empresa en sí como para cada usuario en específico. Para esto, es necesario conocer las necesidades de información de la empresa y la manera de interactuar de cada usuario con ella, siendo considerado este aspecto en la siguiente etapa de esta fase del proceso de asimilación. Asimismo, el entendimiento del Data Warehouse como una arquitectura de información, su filosofía en el manejo de los datos, su orientación a los objetivos de la empresa, su relación con los niveles organizacionales y demás aspectos considerados en la figura

2.11, pueden dar a los diseñadores del Data Warehouse una idea clara de la relación entre esta tecnología, el negocio y el usuario final. Esta etapa y la siguiente de esta fase pueden analizarse en paralelo para fines de la asimilación de la tecnología Data Warehouse.

Actividades a realizar y aspectos a considerar para el Data Warehouse

- Asistir a seminarios o clases, diplomados, conferencias, etc. para conocer las características generales de las funciones de la nueva tecnología.

Un proveedor de la tecnología Data Warehouse puede mostrar casos prácticos y reales en empresas donde se esté implantando o se haya implantado esta tecnología, los problemas más comunes que se encontraron y el nivel de rechazo de parte de los usuarios del negocio. El conocimiento de la tecnología puede también incrementarse a través de la investigación en diversas fuentes bibliográficas, tales como libros, artículos de revistas reconocidas, y más recientemente por medio de Internet, donde incluso se pueden localizar centros de información exclusivos para el Data Warehouse. El personal del departamento de informática o sistemas de información puede asistir a seminarios en donde le muestren los beneficios y funciones de esta innovación tecnológica, para después presentar el conocimiento adquirido ya adaptado a la estructura organizacional y a las necesidades de la empresa. Para lograr lo anterior, la persona o grupo de personas que asista a tales seminarios debe contar con un conocimiento bien fundamentado tanto de la tecnología de información como del negocio, y sobre todo con una percepción de la alineación entre ambas, así también debe tener clara la visión del departamento de sistemas de información a través del conocimiento del plan informático con que se cuenta.

- Instituir un programa de educación continua

El Data Warehouse es una tecnología que se va desarrollando por etapas y en cada una de ellas se involucran usuarios de distintas áreas funcionales de la empresa, por lo tanto, la capacitación a los usuarios debe ser continua y siempre con un enfoque global de la empresa, de tal manera que los usuarios conozcan y entiendan las funciones de otros usuarios y la relación que existe o puede existir entre ellos. En términos generales, se tratará de capacitar a los usuarios para que conozcan como la tecnología Data Warehouse puede influir en sus tareas diarias, en la relación con otros usuarios y en la comprensión de cómo dicha tecnología es capaz de satisfacer las necesidades más importantes de la empresa. A manera de que aumenta la cultura

informática en una organización, mayor es la participación de los usuarios y su conocimiento en las herramientas para la obtención de información. El programa de mejora continua debe incluir capacitación en el Data Warehouse a nivel usuario, así como de otras tecnologías de información; con esto los usuarios van a tener la capacidad de identificar mejoras a un proceso y efectuar la retroalimentación requerida para una óptima operación de la arquitectura Data Warehouse.

Los resultados que se obtengan de este programa de educación continua pueden ser un indicativo del entendimiento del Data Warehouse que se va teniendo a nivel usuario, esto puede influir en el establecimiento de las prioridades de implantación del Data Warehouse en una empresa, ya que si en un área funcional no se tiene todavía la capacidad para asimilar la tecnología, tal vez no convenga su aplicación inmediata en ella. La decisión de implantarse o no en un área funcional finalmente la toman los diseñadores del Data Warehouse basándose en un plan informático, el cual debe incluir todo lo referente a las necesidades estratégicas de la empresa, la TI existente, etc., y podría complementarse con el nivel de entendimiento del usuario con respecto a la nueva tecnología. El tema de la planeación informática se mencionó con mayor detalle en el apartado 4.6.1. de esta tesis. “En el corazón del éxito en el desarrollo del Data Warehouse está la retroalimentación entre la arquitectura de datos y los usuarios finales o analistas de decisiones; esta retroalimentación permite modificar constantemente el Data Warehouse añadiendo otros datos a las bases de datos, siendo esto constante para todo el ciclo de vida del Data Warehouse” [INMON,1996].

II.2. ESTABLECIMIENTO DE LAS NECESIDADES DE LA ORGANIZACION

El objetivo de esta etapa es asegurar que las necesidades del negocio dirijan la introducción de la nueva tecnología a la organización.

Esta etapa del proceso de asimilación puede ser realizada en paralelo o complementarse con la etapa anterior referente al conocimiento de la tecnología. Se ha mencionado en diversas partes de esta tesis la importancia de alinear a la tecnología con el negocio, y desde el punto de vista del proceso de asimilación esto se puede lograr con el análisis complementario de estas dos primeras etapas de la segunda fase del proceso de asimilación. El esfuerzo Data Warehouse debe ser guiado por el negocio y apoyado por la tecnología, reforzándose el hecho de emplear un plan informático que dirija las iniciativas de nuevos proyectos de tecnología de información.

Actividades a realizar y aspectos a considerar para el Data Warehouse

- Enfatizar en la búsqueda de la alineación de las estrategias y necesidades del negocio con la TI.

“El reconocimiento del papel estratégico que una tecnología de información pueda tener asegura su alineamiento al negocio” [LUFTMAN,1993]. Tomando como base este punto, los diseñadores de la arquitectura de información Data Warehouse deben difundir a esta tecnología con un carácter estratégico y orientado a los tópicos del negocio, algo que es propio al diseño mismo del Data Warehouse. El plan de alineación de la TI/SI con las estrategias de la empresa presentada en la figura 4.8, puede ser un soporte muy importante en el logro de esta actividad del proceso de asimilación, al obtenerse las necesidades de información del análisis de la matriz entre las acciones estratégicas y las funciones del negocio se garantiza que el plan informático apoye tanto a las estrategias de la empresa como sirva a los departamentos en sus actividades cotidianas.

Otra ventaja que se puede obtener de un plan informático es facilitar la evaluación de las herramientas de software requeridas para el diseño de la arquitectura de información Data Warehouse; al generarse el plan de software base derivado del plan informático, la evaluación de las herramientas estará mejor enfocada a las necesidades del negocio y a los proyectos de TI como el Data Warehouse. En la fase I de este proceso de asimilación puede resultar complejo evaluar las herramientas Data Warehouse que mejor pudieran adaptarse a la arquitectura de información, debido a la gran diversidad de herramientas de software dentro del mercado Data Warehouse, además de estas herramientas, se tiene que establecer adecuadamente la relación entre el Data Warehouse y los diversos sistemas de información, ya sea a un nivel operacional (transaccionales), a un nivel táctico (Sistemas de Soporte a la decisión) o a un nivel estratégico (Sistemas de Información Ejecutivos). Como se puede observar existe una gran diversidad de sistemas de información que son necesarios administrar para conformar la arquitectura de información Data Warehouse, y sin una adecuado plan informático y una alineación estratégica a los objetivos de la empresa podrían surgir errores que generen costos mayores y retardos en tiempo en la implantación del proyecto. “Enfocarse a una alineación estratégica es importante para asegurar que las prioridades sean consistentes, que los recursos sean asignados apropiadamente y que la sofisticación de la tecnologías vayan de acuerdo a la sofisticación de la organización” [LUFTMAN,1993].

- Prestar atención a las necesidades de los usuarios de la TI dentro de la organización.

Los usuarios de la TI son los que finalmente usan la tecnología y son los que generan las retroalimentaciones que adecuan la tecnología al medio ambiente cambiante de una empresa. En el caso del Data Warehouse “los requerimientos para el Data Warehouse no pueden ser conocidos hasta que es parcialmente creado y usado por los usuarios finales o analistas de toma de decisiones, de tal manera que esta tecnología no puede ser diseñada de acuerdo a la metodología clásica guiada por los requerimientos de los usuarios, pero a la vez sería error no anticiparse a dichos requerimientos, por lo cual un balance intermedio sería lo más adecuado. Dicho balance intermedio consistiría en identificar la información que es crítica en un proceso de la empresa a través de entrevistas a los usuarios, no tanto para conocer sus necesidades específicas, sino más bien para determinar la información que ellos utilizan y que se considera crítica para el proceso en el cual participan. El medio ideal para lograr obtener esa información es a través de entrevistas a los usuarios finales; Ralph Kimball [KIMBALL,1996] propone una entrevista formal estructurada. “Las entrevistas a los usuarios finales es el primer paso más importante en el diseño de un Data Warehouse, y persigue principalmente dos propósitos:

1. Proporciona a los diseñadores del Data Warehouse una visión de las necesidades y expectativas de los usuarios finales, permitiendo un acercamiento inicial entre el equipo de desarrollo del Data Warehouse y los aspectos del negocio.
2. Permite a los diseñadores alcanzar un nivel de conciencia junto con el usuario final del Data Warehouse, que se espera obtener, ajustando y corrigiendo algunas de las expectativas de los usuarios. “

Algo muy importante que se puede obtener de estas entrevistas es el grado de aceptación de los usuarios finales hacia alguna limitante en la información que van a obtener del Data Warehouse, provocado quizá por la falta de disponibilidad en los datos que se requieren para lograr tal información. Por otro lado los usuarios finales en el ambiente Data Warehouse son diferentes a los usuarios finales de un ambiente operacional; generalmente los usuarios del Data Warehouse se encuentran en un nivel táctico o estratégico en la organización. En el desarrollo de un sistema transaccional de carácter operativo el usuario final es sin duda el elemento principal para definir los requerimientos de información durante la etapa de análisis del sistema; sin embargo para el diseño de un Data Warehouse, además del usuario se encuentran los analistas para el soporte a la toma de decisiones, los cuales van a ser los responsables de definir las estrategias de sumarización, integración y presentación de datos necesarios para soportar una decisión de un administrador o ejecutivo en un momento dado. Es primordial entonces identificar a los usuarios de

un Data Warehouse en una organización, a fin de conocer sus necesidades de información más importantes; Inmon presenta una definición del usuario final de un Data Warehouse:

“El usuario de un Data Warehouse es una persona que participa en algún proceso de toma de decisiones en la organización, conoce más del negocio que de la tecnología, y su principal función es la definición y el descubrimiento de la información que es usada para realizar alguna decisión corporativa” [INMON,1996].

II.3. CONSIDERACION Y PERSUASION DE LA ADOPCION DE LA TECNOLOGIA

El objetivo de esta etapa es que los miembros de la organización comiencen a considerar la conveniencia de la innovación.

“La persuasión se realiza cuando un individuo se forma una actitud favorable o desfavorable acerca de la innovación, que lo conduzca a una toma de decisión para la adopción o rechazo de la misma” [ROGERS,1983]. La correcta aplicación de las dos etapas anteriores es fundamental para que el individuo se forme una actitud favorable hacia el Data Warehouse. El interés del usuario puede venir directamente de la visualización de los beneficios que puede obtenerse con el Data Warehouse, o bien del conocimiento de la tecnología a través de seminarios, o del interés mostrado por parte de los diseñadores del Data Warehouse en las entrevistas para identificar las necesidades de información; sea cual fuera la causa, en esta etapa el usuario debe sentir cierta inclinación hacia la innovación tecnológica, y si así no fuera, es necesario aplicar las actividades que a continuación se mencionan y que también influyen en el grado de adopción del Data Warehouse. La responsabilidad de controlar esta etapa es del grupo que investiga las nuevas tecnologías, pero deben tener coordinación y apoyo de la alta administración, para asegurar la continuidad del proceso de asimilación.

Actividades a realizar y aspectos a considerar para el Data Warehouse

- Reconocer la importancia de la información para el logro de los fines de la organización.

“Desde hace 30 ó 40 años cuando aparecieron las primeras herramientas de procesamiento de datos, los administradores de la empresas han sobrestimado y desestimado la importancia de la información en la organización” [DRUCKER,1995].

La información requiere ser considerada como un activo dentro de la empresa para que proyectos de tecnologías de información como el Data Warehouse puedan ser adoptados y asimilados en la misma. En el capítulo 2 se definió al Data Warehouse como una arquitectura de información que se basa principalmente en la manera en cómo fluye la información en todos los niveles de una empresa, por otro lado, una de las causas por las cuales un Data Warehouse se necesita en una organización es por la falta de confiabilidad en la información que es presentada a las personas que toman decisiones, o bien a los ejecutivos y directivos de una empresa; si tales decisiones se toman sin un uso adecuado de la información, entonces el procesamiento de los datos para generar información es irrelevante y el Data Warehouse no tendrían razón ser; pero si por otro lado, la empresa le da valor a la información y soporta su proceso de toma de decisiones por medio del uso de la misma, entonces el nivel de importancia de la información en esa empresa es suficiente para considerar el desarrollo de un Data Warehouse como una alternativa de mejora en el manejo de la información.

“En la ‘era de la información’ las organizaciones competirán sobre la base de su habilidad de adquirir, manejar, interpretar y usar la información en forma efectiva. La información es dinámica, capaz de crear valor y es el pegamento que mantiene a las empresas integradas” [MCGEE,1993]. Si se reconoce que lo que da la ventaja a una empresa reside en el uso de la información, el valor de la tecnología de informática estará en hacer accesible esa información a los diferentes usuarios en el momento y en el lugar que la requieran para tomar decisiones.

- Establecer las bases para la formación de “campeones” de la tecnología

“Un campeón es una persona con una fuerte creencia en la tecnología, y que será el líder para la exploración y desarrollo de la misma” [BUSWICK,1990]. Este “campeón” debe contar con el apoyo decidido y total de parte de la alta dirección o gerencia de la empresa para conseguir los recursos necesarios para el diseño exitoso del Data Warehouse; además del conocimiento de la tecnología, debe ser una persona con capacidad de relacionarse con cualquier miembro de la organización, de tal manera que pueda lograr el apoyo necesario para la obtención de información de cada departamento o área funcional; es decir, hablar en lenguaje técnico y a la vez, hablar en el lenguaje del negocio.

En el caso de la tecnología Data Warehouse, el tener conocimiento de conceptos como el de Bases de Datos, de los Sistemas de información como los EIS, DSS, etc., establece que el “campeón” debe salir de algún departamento dónde fluya o exista grandes cantidades de datos e información y en donde tales conceptos sean perfectamente conocidos y comprendidos. Hasta esta fase, el departamento de

informática se ha encargado de desarrollar el diseño del Data Warehouse, así como las primeras etapas de este proceso de asimilación con el apoyo de alta gerencia y de los usuarios relacionados a los procesos de información; a partir de esta etapa el campeón va a estar involucrado a cada proceso relacionado con el diseño de la arquitectura de información Data Warehouse y va a ser el intermediario entre la alta administración y el grupo de desarrollo de la innovación tecnológica. El campeón debe tener un conocimiento del plan informático y de la situación de la tecnología de información en la empresa.

- Identificar posibles agentes de cambio.

La arquitectura de información Data Warehouse requiere de personas que la promuevan en todas las áreas funcionales de la organización, que influyan en la decisión de adoptarla y que vayan estableciendo las bases para su implantación en cada área específica. Los agentes de cambio para la tecnología Data Warehouse pueden ser los usuarios finales de información de cada área funcional de la empresa, son ellos los que conocen los problemas en la generación de la información de soporte a la toma de decisiones y son también los que van a establecer las estrategias de sumariación, consultas y partición de datos para un diseño exitoso de un Data Warehouse; además influirán, una vez implantada la tecnología, en los procesos de retroalimentación para hacer que el Data Warehouse refleje lo que el negocio necesita en su momento.

- Solicitar demostraciones y presentaciones a varios proveedores de la tecnología

La evaluación de la lista de proveedores de las diversas herramientas Data Warehouse es una etapa propuesta por Brown en el diseño del Data Warehouse, y una de las actividades que debe llevarse a cabo como parte de dicha evaluación es la solicitud de demostraciones y presentaciones de los productos de cada proveedor. Debido a la gran cantidad de herramientas relacionadas con esta innovación tecnológica, se debe buscar a los proveedores que ofrezcan soluciones totales o que agrupen diversas tareas en un solo producto, de tal manera que esta actividad no se prolongue en tiempo, y sea más fácil la selección de las herramientas para el diseño de la arquitectura de información Data Warehouse. La figura 3.3 de esta tesis presenta las herramientas claves de un proyecto Data Warehouse y su agrupación de acuerdo la función que realizan; la información que ahí se muestra puede ser útil en la identificación de los proveedores y sus respectivas herramientas de software. “La industria de productos de información, que puede incluir productos basados en datos, información y conocimiento, es extremadamente dinámica en términos de su

crecimiento en la introducción de nuevos productos” [MEYER,1996]. Ante este crecimiento de productos de información y proveedores, se debe seleccionar adecuadamente los proveedores que van a ser invitados para realizar demostraciones en la empresa. Las presentaciones deben enfocarse a promover las ventajas de un Data Warehouse con base a las herramientas propias de cada proveedor y su posible compatibilidad con otras herramientas de otros proveedores de la industria. Es importante dejar claro el nivel de estandarización que se tiene en la industria Data Warehouse para que se considere como un punto a favor en la persuasión de la adopción de la tecnología.

“Aunque algunos vendedores de los Manejadores de Base de Datos, herramientas de extracción y herramientas de consultas han incorporado algunas facilidades necesarias para realizar la partición y sumarización de los datos, ningún vendedor actualmente tiene las respuestas a todos los problemas, por lo tanto, los desarrolladores del Data Warehouse deben agrupar las piezas para crear una solución total de diferentes proveedores mediante la integración de sus herramientas” [MEREDITH,1996].

II.4. DISCUSION PARA LA ADQUISICION DE LA INNOVACION

El objetivo de esta etapa consiste en que los individuos inicien conversaciones concernientes a la adopción de la tecnología.

Actividad a realizar para el Data Warehouse

- Promover reuniones para discutir la propuesta de creación o adquisición de la tecnología para realizar pruebas.

Al conocer los beneficios de la tecnología y lo que ofrece cada proveedor, la organización debe decidir en el diseño de un Data Warehouse entre tres posibles propuestas [KAY,1996]:

a) Construir su propio Data Warehouse

El crear un Data Warehouse implica un desafío muy fuerte para el departamento de informática o sistemas de información de una empresa, ya que es poca o nula la experiencia con esta nueva tecnología. Además del conocimiento de la tecnología se deben tomar decisiones importantes en hardware, software, conectividad y en una serie de tareas que incluyen a la modelación, la extracción, el almacenamiento y la limpieza de datos, entre otras.

b) Utilizar Outsourcing

Esta propuesta es la menos empleada por las organizaciones que ya han implantado un Data Warehouse, debido que al considerarse a esta tecnología de carácter estratégico, el usar un Outsourcing podría generar la pérdida de control sobre los datos críticos de la empresa.

c) Contratar los servicios de una consultoría externa

Esta es quizá la propuesta más interesante de las tres, ya que a través de una consultoría externa se pueden cubrir los aspectos tecnológicos más complejos y que consumen más tiempo en el ciclo del desarrollo del Data Warehouse, sin perder con esto el control sobre los datos y los procesos de información de la compañía.

La decisión de construir o comprar un Data Warehouse depende del tamaño de la compañía, los recursos, las habilidades técnicas y administrativas y la urgencia en la implantación del proyecto. Cualquiera que sea la decisión, debe definirse adecuadamente el equipo de trabajo interno de la empresa para el desarrollo del Data Warehouse, teniendo ya identificados los agentes de cambio para cada área funcional de la misma. La figura 5.2 nos muestra la tendencia en este rubro del Data Warehouse en un comparativo entre los años 1996 y lo que se espera en el año 2000.

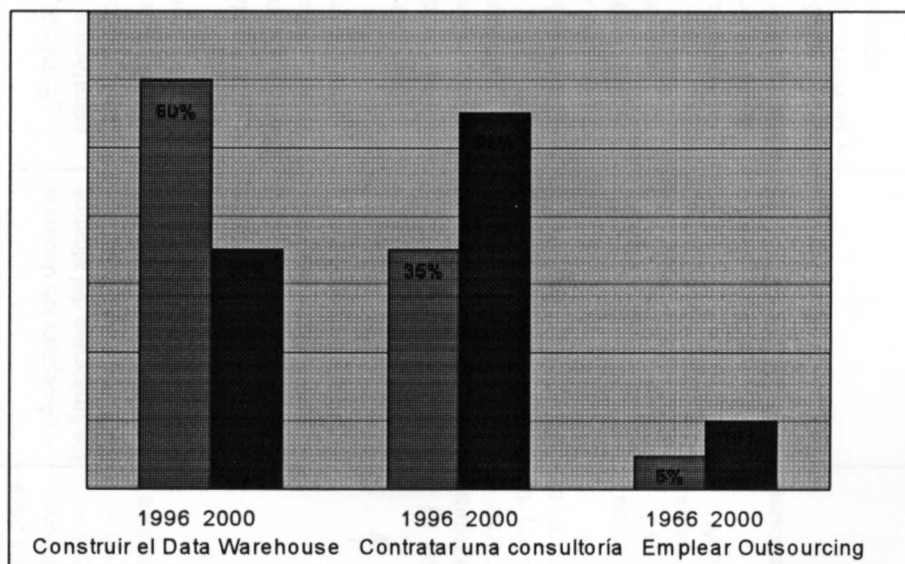


Figura 5.2 – Tendencias en la implantación de un Data Warehouse

FASE II: CONOCIMIENTO - PERSUASION

Atributos que influyen en el proceso de asimilación	Ventaja relativa			Compatibilidad			Complejidad			Experimentabilidad			Observabilidad		
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
<p>*Conocimiento de la tecnología de información requerida para diseñar el Data Warehouse</p> <p>* Identificación del recurso humano participante en el desarrollo de la arquitectura de información</p> <ul style="list-style-type: none"> - Desarrolladores - Usuarios - Promotores <p style="text-align: center;">↓ ↓ ↓ ↓</p> <p>* Invertir en programas de capacitación en la TI relacionada con el Data Warehouse</p> <p>* Definir claramente los objetivos del negocio que se van a perseguir con el DW</p> <p>* Evaluar a los proveedores de las herramientas DW</p> <p>* Crear un modelo de datos corporativo para el DW</p> <p>Etapas del Data Warehouse involucradas</p>				<ul style="list-style-type: none"> - ¿Aceptan los usuarios las limitantes del alcance del Data Warehouse para satisfacer las necesidades de información? - ¿Es factible mejorar el flujo de información que apoya al proceso de toma de decisiones? 	<ul style="list-style-type: none"> - ¿Existe alineación estratégica entre las tecnologías de información y los objetivos de la empresa? -¿Los proyectos de Tecnologías de Información están guiados por un plan informático? -¿Están identificados los analistas de toma de decisiones en la empresa? -¿Es considerada la información como un activo dentro de la empresa? -¿Han surgido personas fuera del equipo de desarrollo del Data Warehouse que promuevan y se interesen en la tecnología? 	<ul style="list-style-type: none"> - ¿Existe estandarización en cuanto a términos por parte de los proveedores del DW? - ¿Son accesibles las fuentes de información para conocer más sobre la tecnología? -¿Pueden integrarse las herramientas que presentan los proveedores para crear una solución total del DW? -¿Cuenta el personal informático con las habilidades técnicas necesarias para aplicar o desarrollar nuevas herramientas de software para cada proceso del DW? -¿Cuál es el grado de cultura informática en la empresa para aceptar y usar la tecnología? -¿Han sido comprendidos todos los conceptos y procesos del DW? 				No aplica				<ul style="list-style-type: none"> -¿Se conocen casos exitosos de empresas en donde se haya implantado un Data Warehouse? 	

Grado de Adaptación

- 1 - Bajo
- 2 - Medio
- 3 - Alto

Grado de Adaptación (Complejidad)

- 1 - Alta
- 2 - Media
- 3 - Baja

Tabla 5.6

DESCRIPCIÓN DE LA TABLA 5.6 – GRADO DE ADOPCIÓN, FASE II

En esta fase del proceso de asimilación se busca principalmente aumentar el conocimiento de la tecnología Data Warehouse entre los miembros de la empresa, de tal manera que se vaya formando una actitud positiva hacia la tecnología. Si no se logra difundir a la tecnología adecuadamente, el personal de la empresa puede justificar el no operar las herramientas Data Warehouse ante la falta de conocimiento y entendimiento de ellas, afectando de esta manera la asimilación de la innovación. Así se consideran dos atributos que influyen en el proceso de asimilación, desde el punto de vista de la innovación se tiene el conocimiento de las tecnologías de información para diseñar el Data Warehouse, y desde el punto de vista del contexto organizacional se considera la identificación del recurso humano participante en el desarrollo de esta tecnología, siendo éstos a los que hay que enfocar los programas de capacitación. Las siguientes etapas de diseño del Data Warehouse pueden ser influenciadas por tales atributos.

- a) Invertir en programas de capacitación en diversas tecnologías de información relacionadas con el Data Warehouse.

Para planear el programa de capacitación se debe establecer primeramente que tipo de conocimiento de la arquitectura Data Warehouse se necesita adquirir, y por supuesto determinar a que miembros de la empresa se les dará la capacitación. Un buen programa de capacitación orientado al personal correcto puede facilitar el aumento del conocimiento requerido para implantar el Data Warehouse en una empresa, además de permitir elevar la motivación del recurso humano y con esto ir formando una actitud de aceptación y asimilación hacia la tecnología.

- b) Definir claramente los objetivos del negocio que se van a perseguir con el diseño del Data Warehouse

Los objetivos del negocio dan la pauta para determinar las necesidades de información que van a ser cubiertas por la arquitectura Data Warehouse. Del análisis de la información se definen las áreas involucradas con su proceso de generación, y una vez identificadas dichas áreas se pueden establecer los cursos de capacitación requeridos para las mismas.

- c) Evaluar a los proveedores de las herramientas Data Warehouse

En la fase anterior la evaluación de los proveedores consistió en identificar a los que podrían apoyar en el diseño del Data Warehouse en la empresa, en esta fase se debe evaluar la manera en la cual los proveedores pueden participar en la

difusión del conocimiento de la innovación tecnológica. Para esto, debe existir estandarización en cuanto a términos y conceptos del Data Warehouse se refiere, así como compatibilidad entre las tecnologías de los proveedores; una vez lograda la estandarización y la compatibilidad, es más factible que los miembros de la organización entiendan el concepto Data Warehouse y sus ventajas y beneficios para la empresa.

d) Crear un modelo de datos corporativo para el Data Warehouse

Un modelo de datos de la empresa permitiría identificar la participación de cada área funcional del negocio en el diseño del Data Warehouse, y por consiguiente, la identificación del recurso humano a involucrarse en dicho diseño. Los usuarios son muy importantes en el Data Warehouse, generalmente son los que toman decisiones, conocen al negocio y establecen los requerimientos de información para realizar alguna decisión corporativa, en el modelo de datos se pueden definir fronteras que representarían el alcance de la arquitectura de información Data Warehouse en la organización, por lo que es de interés difundir el conocimiento del Data Warehouse de manera realista para no generar falsas expectativas entre los miembros de la organización, lo cual ocasionaría inconformidades y falta de aceptación hacia la innovación tecnológica.

Medidas de Evaluación para el grado de adopción

Ventaja relativa

Con esta medida de evaluación se pretende determinar si los usuarios al tener conocimiento de la tecnología, consideren factible que ésta pueda mejorar el proceso en que ellos participan para la toma de decisiones, con las limitantes y el alcance con que se diseñará la arquitectura de información Data Warehouse.

Compatibilidad

Con esta medida se pretende determinar si la arquitectura Data Warehouse se adapta a las condiciones existentes de la empresa, tanto desde un punto de vista organizacional como tecnológico. Se considera principalmente el aspecto de la importancia que se le da a información en la empresa, lo cual puede justificar la aceptación o rechazo de la tecnología Data Warehouse. Con el conocimiento que se tenga del Data Warehouse puede determinarse si la información es tan importante como para invertir en un proyecto de esta naturaleza, y sobre todo si la alta

administración es capaz de crear funciones que no existían, tal como los analistas de decisiones.

Complejidad

A través de esta medida de evaluación se pretende determinar si los miembros de la organización son capaces de entender y comprender los conceptos y procesos de la arquitectura de información Data Warehouse, además es necesario definir la función de cada componente y la participación del personal involucrado con la aplicación de cada proceso. Por otro lado, desde el aspecto técnico, es importante determinar la complejidad para integrar las herramientas que presentan los proveedores Data Warehouse y su adaptación a la plataforma tecnológica de la organización.

Experimentabilidad

Esta medida de evaluación no aplica debido a que no se ha realizado ningún tipo de prueba par el diseño del Data Warehouse.

Observabilidad

El estudio de casos de aplicaciones exitosas del Data Warehouse permite generar una capacidad de análisis para adaptar las características del caso con las condiciones actuales de la empresa. Lo importante de esta medida de evaluación es que se difundan en el negocio los resultados de la aplicación exitosa del Data Warehouse en otras organizaciones y se presenten las características globales que hicieron de la implantación un éxito.

FASE III: EVALUACION-DECISION

Los objetivos de esta fase son que la adquisición de la tecnología se proponga formalmente; se evalúe la inversión de acuerdo a criterios funcionales, financieros, políticos y estratégicos; se evalúe la aptitud organizacional; la tecnología se ponga a prueba; y se llegue a la decisión de adopción, rechazo o posposición.

III.1. PROPOSICION DE ADQUISICION DE LA TECNOLOGIA

El objetivo de esta etapa es realizar una proposición formal de adquisición de la innovación para fines de pruebas o experimentación.

Actividad a realizar para el Data Warehouse

- Documentar formalmente el inicio de la etapa de evaluación, siendo aprobado por la alta administración.

Esta actividad es solamente documentativa y debe incluir aspectos de las 2 primeras fases de este proceso de asimilación, entre lo más destacado se puede mencionar:

- Las condiciones de la empresa por las cuales es necesario un Data Warehouse
- Los beneficios y el alcance de la innovación tecnológica
- La alineación de la tecnología con los objetivos del negocio
- Los responsables del diseño del Data Warehouse
- La evaluación de la TI necesaria para implantar la innovación
- Los proveedores de la TI y sus soluciones

III.2. EVALUACION RACIONAL-FINANCIERA

Con esta etapa se pretende conocer el valor real de la adquisición de la tecnología, lo cual implica saber el costo de la innovación, así como de los recursos y ajustes necesarios para poderla operar dentro de la organización, y comparar estos costos con el beneficio potencial que brinda para satisfacer las necesidades existentes en la organización.

Aspectos a considerar para el Data Warehouse

- Los costos de todos los recursos necesarios para el uso de la tecnología.

El costo de implantar un Data Warehouse en una organización es generalmente alto, y puede variar dependiendo de su alcance y objetivos de implantación. “Desafortunadamente, la arquitectura Data Warehouse requiere fuertes inversiones tanto en tiempo como en dinero” [WATTERSON,1996]. Al ser una innovación tecnológica dentro de la empresa, el Data Warehouse tiene costos inherentes a los recursos necesarios para crear la infraestructura necesaria para su implantación; podemos mencionar los recursos informáticos como las herramientas de software en general, la necesidad de equipos como computadoras, dispositivos periféricos y de comunicación; también tenemos costos en el recurso humano como el personal requerido para su desarrollo y para su operación, aquí se involucran costos de entrenamiento y capacitación.

Además de los costos mencionados anteriormente, existen otros que son propios al diseño del Data Warehouse, los cuales pueden generarse al aplicarse los diversos procesos de esta arquitectura de información, entre los que podemos mencionar la sumarización, extracción y conversión, partición de los datos, etc., los cuales pueden influir en el incremento o disminución del costo global del diseño de la innovación tecnológica. Estos últimos costos del Data Warehouse los podemos asociar a los medios ambientes mencionados en la etapa I.1 de este proceso de asimilación.

⇒ Datos “sucios”

Para el diseño de la arquitectura Data Warehouse se requieren usar datos de diversas bases de datos de la empresa para generar información que soporte la toma de decisiones o bien, para la generación de un reporte corporativo específico. Para lograr tal proceso de integración de datos se requiere primero localizarlos y tenerlos disponibles, y después hacer que todos se integren en una base común; los datos “sucios” en ambos casos provocan que se hagan más complejas tales tareas y por consiguiente aumenten los costos en el diseño del Data Warehouse. “Los datos ‘sucios’ aumentan los costos de implantación de una Data Warehouse, convirtiendo a los analistas y desarrolladores de esta tecnología en ‘arqueólogos’ que buscan datos disponibles de diversas fuentes para encontrar los que puedan contribuir a satisfacer los requerimientos de información de los usuarios finales” [MORIARTY,1996,1].

Los datos “sucios” pueden existir en cualquier empresa, y para minimizar el efecto en los costos que pudieran tener en el desarrollo del Data Warehouse es necesario identificar los datos que realmente van a ser usados en la generación de información, y no seleccionar los datos solamente porque están disponibles, ya que esto provocaría costos en la limpieza de datos que al final no dan valor al proceso de información del Data Warehouse.

⇒ Alto nivel de análisis

Como se ha dicho anteriormente, la arquitectura Data Warehouse tiene como uno de sus objetivos principales apoyar el proceso de toma de decisiones en una organización, y es mediante al análisis de información como tal apoyo puede llevarse a cabo. Ahora bien, mientras más procesos analíticos de información existan en una empresa mayor va a ser la cantidad de datos que son necesarios transformar, y por consiguiente, mayor el costo de diseño del Data Warehouse. “Según Alan Paller, director del instituto de educación e investigación del Data Warehouse, del 70 al 80 por ciento del tiempo y dinero gastado en el Data Warehouse es por la transformación de datos” [ROBINSON,1996,2].

Tal como se mencionó en la etapa I.1 de este proceso de asimilación, los métodos de extracción y movimiento de datos se llevan a cabo mediante herramientas de software que actualmente están disponibles en el mercado del Data Warehouse, por lo cual la evaluación que se haya hecho de los proveedores para conocer sus productos es muy importante para administrar los costos que van a generarse al integrar las diversas herramientas entre sí, y con la tecnología de información existente en la empresa. “Los principales costos para implantar un Data Warehouse se generan de la aplicación de los recursos necesarios para integrar los productos que permitan crear la arquitectura de información” [STRANGE,1995].

⇒ Tiempo real

El procesamiento de información analítica en el Data Warehouse, a pesar de ser heurística, también requiere estar actualizada al momento de las consultas, por lo cual lo mencionado en la etapa I.1 referente a las herramientas de conectividad, es importante para tener dicha actualización. Las herramientas cliente-servidor pueden ser usadas en conjunto con las herramientas Data Warehouse para hacer que la información fluya de manera más eficiente en la empresa, dando la impresión de estar en tiempo real, tal como sucede con las bases de datos operacionales. El costo para presentar la información del Data Warehouse en tiempo real puede elevarse si no se cuenta con las herramientas necesarias de conectividad y Cliente-Servidor para la creación de esta arquitectura de información, además de que los diversos procesos

del Data Warehouse sean aplicados correctamente en el tiempo establecido y con los parámetros necesarios

⇒ Gran volumen de datos

Mientras mayor cantidad de datos exista en una empresa, mayor es la dificultad para identificarlos, accederlos y crear análisis con ellos. Es por eso que las estrategias de partición de datos son importantes en el diseño del Data Warehouse para hacer que los datos sean accesibles de manera más fácil y eficiente. Además de las estrategias de partición de datos, la aplicación del concepto de análisis de estabilidad basado en la agrupación de procesos (ver figura 4.6), puede mejorar los tiempos de accesos y consultas a la información, generando con esto ahorro en costos en tiempo y en procesos de integración y transformación de datos que serían necesarios aplicarse de otro modo. Aunado a lo anterior, el Data Warehouse almacena datos históricos sumariados de muchos años atrás, lo cual hace que las aplicaciones Data Warehouse manejen datos en el rango de los Gbytes o Terabytes. No sólo la tecnología propia al Data Warehouse y su eficiencia son aspectos a considerar en el diseño de dicha tecnología, sino también se deben evaluar los costos de almacenamiento y procesamiento de datos.

Desde el punto de vista de hardware, para evitar altos costos de almacenamiento y la inversión en plataformas de hardware con mayor capacidad se necesita que las estrategias de sumariación y partición de datos sean aplicadas teniendo en consideración la minimización del espacio de almacenamiento y el procesamiento de datos requerido para llevarlas a cabo, porque mientras mayor sea el procesamiento de la información, mayor es la memoria que se requiere para realizarlos y cumplir con los procesos de actualización para presentar la información en tiempo real. Existen dos aspectos que pueden ser considerados en el diseño de un Data Warehouse para reducir los costos de procesamiento y almacenamiento, la compactación de datos [INMON,1996], y el paralelismo en el procesamiento y almacenamiento de los datos [ALUR,1996],[INMON,1996].

- El Análisis Costo-Beneficio de la tecnología.

Los resultados del análisis costo-beneficio influyen fuertemente en la decisión de adquirir o rechazar una nueva tecnología. En lo que se refiere a las tecnologías de información, los beneficios que se obtienen son muchas veces intangibles, y no se reflejan directamente en el Retorno sobre la Inversión (ROI) que espera la alta administración de una organización. Uno de los aspectos interesantes del Data Warehouse es que la justificación en costo para dicha tecnología generalmente no se

basa en el análisis del Retorno sobre la inversión, ya que tal análisis requiere que los beneficios del Data Warehouse sean conocidos antes de empezar su construcción, y en muchos de los casos los beneficios reales del Data Warehouse no son conocidos ni anticipados al momento de diseñar la tecnología.

Esto último refuerza la importancia de reconocer el valor de la información y el soporte que puede dar al proceso de toma de decisiones críticas para el éxito de una empresa; de tal manera que se justifique por sí misma la implantación de un Data Warehouse al mejorar los procesos de información dentro de una organización. Si los diseñadores del Data Warehouse “venden” correctamente la idea con respecto al arquitectura de información que planean desarrollar, y muestran a los directivos del negocio los beneficios “intangibles” que pueden obtenerse, la adopción de la tecnología será una realidad. La justificación costo-beneficio para una Data Warehouse puede decirse que es directamente proporcional a la importancia que se le dé a la información en la empresa.

A pesar de lo difícil que es justificar el costo de implantación de un Data Warehouse, - ¿cómo podría asignarse un valor monetario a la satisfacción de un cliente?, o bien, ¿cómo estimar el costo asociado a una oportunidad de mercado desaprovechada por reaccionar lentamente? -, Jack Sweeney, uno de los pioneros del desarrollo del Data Warehouse llama a esta tecnología como una “mina de oro” [STRANGE,1995], por el impacto que puede tener en el Retorno sobre la Inversión. Lo traduce en términos del negocio de la siguiente manera:

- ◇ Retención de un cliente mediante un servicio más personalizado basado en información
- ◇ Mejoramiento en la toma de decisiones en inversiones con alto grado de riesgo
- ◇ Análisis de datos históricos para determinar tendencias en diversas funciones del negocio
- ◇ Detección de fraudes en instituciones financieras
- ◇ Aumento en la productividad de los usuarios finales de los sistemas de información de la empresa, al hacer a estos últimos más eficientes
- ◇ etc.

III.3. EVALUACION POLITICA-ESTRATEGICA

El objetivo de esta etapa es cuestionar y comprobar si es que el uso de la tecnología beneficia o apoya a las estrategias del negocio, así como a las políticas existentes en la organización.

Actividad a realizar para el Data Warehouse

- Establecer la liga entre los planes estratégicos existentes de la organización y la adquisición de la TI.

Se ha dado mucho énfasis a la alineación de las estrategias de la empresa con los objetivos de la Tecnología de Información, y es en esta actividad del proceso de asimilación donde se debe comprobar que el uso de la innovación tecnológica beneficia o apoya a las estrategias del negocio. En el capítulo 4 de esta tesis se mostraron algunas metodologías que analizan a una empresa en su manejo de la información, así como su relación con el Data Warehouse, estas fueron la cadena de valor, la cadena de medios-fines y el BSP. Para fines de esta actividad del proceso de asimilación se propone a la herramienta de análisis de la cadena de medios-fines para que de manera gráfica y documentativa se explique el soporte que va a dar la arquitectura de información Data Warehouse a las estrategias del negocio.

Tomemos el ejemplo presentado en la figura 4.3 del capítulo anterior, en el cual se muestran diversos objetivos organizacionales (fines) y los medios para lograrlo; un medio para lograr seguridad en las plantas de proceso químico en una empresa es el mejoramiento de la capacitación al personal en aspectos de seguridad, y un medio para lograr esto último es teniendo un programa eficiente de pláticas de seguridad, y así sucesivamente para cada objetivo o fin organizacional. Como se puede observar, el análisis de la cadena de medios fines permite identificar alternativas de decisión o medios para alcanzar una meta específica, y para los diseñadores de un Data Warehouse esto permitiría reunir la información que se requiere en cada nivel de decisión de la empresa, y así planear el desarrollo e implantación de los componentes de la arquitectura de información Data Warehouse, en el caso del ejemplo se presenta a los sistemas de información tanto operacionales, tácticos y estratégicos involucrados con dicha arquitectura y su relación con la información que fue reunida de cada nivel de decisión.

III.4. EVALUACION DE APTITUD ORGANIZACIONAL

Con esta etapa se pretende realizar una evaluación para conocer si las condiciones dentro de la organización, así como las características propias de sus individuos, propician la introducción de la nueva tecnología.

Actividades a realizar y aspectos a considerar para el Data Warehouse

- Evaluar el grado con el cual la innovación afectará el balance de poder entre las personas y grupos funcionales dentro de la organización; muchas veces causado por la redistribución de la información y la toma de decisiones.

A raíz del surgimiento y crecimiento de las PC's, de las herramientas de software de bases de datos que en estos equipos se procesan, y el aumento de la cultura informática en algunos usuarios finales de la TI, se originó un crecimiento en el número de base de datos independientes a los sistemas transaccionales. Esto generó que el usuario final fuera el que llevara el control de los datos, provocando así un sentido de pertenencia hacia ellos, no permitiendo su compartición a través de la organización. "Muchas áreas funcionales de un negocio se consideran a si mismo los dueños de los datos de sus sistemas de información, y se sienten con el derecho de usarlos en la manera que ellos quieran para satisfacer sus necesidades" [MORIARTY,1996,2].

La creación de la arquitectura de información Data Warehouse requiere la integración de datos de diversas fuentes de la empresa, y es a través de la compartición de los datos la manera de lograr tal integración. Es un hecho que el diseño de un Data Warehouse implica "obligar" al personal a compartir sus datos, además cada usuario debe permitir que "su" información sea accesada para consultas con un mayor nivel de detalle, ya que al tener información sumariada de diversos tópicos del negocio más personas desean acceder los datos desde donde se genera dicha información, y así tomar decisiones con una base más sólida; estos son de los aspectos más difíciles en que los diseñadores del Data Warehouse se van a enfrentar al momento de implantar la innovación tecnológica. Ahora bien, si desde el comienzo se involucró a la alta administración en el proyecto y se difundieron ampliamente los beneficios que esta tecnología puede generar, el impacto que pueda presentarse por el rechazo hacia la nueva tecnología por parte del usuario final debe ser menor, al existir un compromiso general de la organización hacia la nueva tecnología. De ahí la importancia de que todas las etapas previas de este proceso de asimilación se hayan aplicado adecuadamente, así también que la evaluación del grado de adopción se haya realizado con los parámetros de medida correctos de acuerdo al contexto organizacional, es decir, tratar de que en ninguna actividad del proceso haya existido insatisfacción o rechazo hacia la innovación.

Para esta actividad es importante que durante el diseño de la arquitectura de información Data Warehouse se establezcan las estrategias de los procesos que van a generar información para la toma de decisiones, y de esta manera poder determinar los datos que van a ser requeridos en cada proceso, y una vez identificados los datos que se van a necesitar analizar las barreras que pueden evitar que éstos sean facilitados para la creación del Data Warehouse. Es decir, que para evaluar el grado en que la innovación afectará el balance de poder entre las personas y grupos

funcionales en una organización, es necesario que se hayan definido los objetivos y el alcance del Data Warehouse y se hayan empezado a definir las tablas de datos y la información que es necesario reunir en cada nivel de decisión.

- La flexibilidad organizacional para aceptar y difundir el cambio.

La frustración que existe hacia la TI en muchas organizaciones, ha generado un rechazo a cualquier cambio que pueda surgir de tales tecnologías. Es importante identificar aquellas áreas de la organización donde la resistencia al cambio es mayor, quizá debido a que se han utilizado los mismos métodos de trabajo por muchos años, o porque exista un rechazo natural hacia cualquier nueva tecnología, o quizá simplemente porque en determinadas áreas no creen en la tecnología y consideran que con sus métodos sus problemas van a seguir siendo resueltos. “Cuando es necesario aplicar un nuevo conocimiento o realizar un nuevo proceso, muchos usuarios del negocio hacen todo lo posible por resolver dicha situación por si mismos, más que dejarlo al ciclo de cambio normal de cualquier Tecnología de Información” [MORIARTY,1996,2].

La resistencia al cambio puede generarse también por la renuencia de los miembros de una organización a modificar las estructuras establecidas y los grupos de poder, por lo cual el coordinador del Data Warehouse debe actuar con un estilo de liderazgo tipo “vendedor”, y con características que lo hagan orientarse fuertemente a cada tarea y a las relaciones interpersonales. La resistencia al cambio debe analizarse junto con la actividad anterior de esta etapa, referente al balance de poder entre las personas y grupos funcionales, ya que los grupos de poder son el factor organizacional que más puede influir en el grado de adopción del Data Warehouse en una empresa. Es importante identificar las áreas donde la resistencia al cambio sea menor, y junto con el plan informático y el análisis de la información de la empresa, poder desarrollar el plan de migración al ambiente Data Warehouse.

- El grado de interconexión de la organización

La interconexión en este caso no es a través de medios electrónicos, sino más bien, al grado en que la información, automatizada o no, es compartida entre las diversas áreas funcionales de la empresa. La información puede fluir ya sea horizontalmente, es decir, entre los diversos departamentos, o bien, verticalmente hacia niveles organizacionales más altos o más bajos. En ambos casos se deben detectar las “barreras” que impidan el libre fluir de la información, y que puedan afectar posteriormente la implantación de la arquitectura de información Data

Warehouse. El valor real del Data Warehouse es el manejo del flujo de la información, más que la colección de datos.

Representar a la organización desde un punto de vista de procesos puede facilitar la identificación del grado de interconexión en la organización. Al definirse y analizarse cada proceso del negocio se puede determinar que parte de la organización es responsable del mismo, la información que maneja, la que necesita y la que proporciona a otros procesos. La metodología BSP y la cadena de valor son herramientas de apoyo para analizar a una empresa desde un punto de vista de procesos. En el apartado 4.5. del capítulo anterior se menciona la relación datos-procesos que puede surgir al analizar a una empresa a través de los procesos y su influencia en el diseño de la arquitectura de información Data Warehouse, siendo el BSP la metodología recomendada para realizar dicho análisis. Es importante detectar las prácticas organizacionales que puedan afectar el intercambio de información entre los procesos, el sentido de pertenencia de los datos es una de estas prácticas; pero también se presentan otras prácticas organizacionales que los diseñadores del Data Warehouse deben eliminar o administrar, entre éstas tenemos las siguientes:

⇒ Todos quieren ser usuarios finales, pero nadie quiere ser fuente de información.

En el diseño del Data Warehouse, un nivel de sumarización de datos está dirigido a un grupo específico de usuarios, al crearse un nivel más alto los usuarios pueden ser otros, y así sucesivamente para cada nivel de sumarización o detalle de la información. El mismo diseño del Data Warehouse obliga a toda persona de la empresa involucrada con la información a ser usuario en su nivel de sumarización y a ser fuente para los niveles más altos.

⇒ La falta de estandarización de los datos

La integración de los datos es una de las actividades básicas de diseño del Data Warehouse, pero para lograrlo se requiere la estandarización de los datos de las diversas fuentes de información. Si las bases de datos se encuentran en diferentes formatos es necesario aplicar procesos de conversión del nivel operacional al nivel Data Warehouse, tal como se muestra en la estructura de la arquitectura de información Data Warehouse presentada en la figura 2.11.

⇒ Un tópico del negocio puede variar en su definición de un departamento a otro

El grado de interconexión entre dos departamentos de una empresa puede verse afectado si el concepto de un tópico que ambos comparten es distinto y con objetivos no en común. La creación del modelo de datos, como proponen diversos

autores dentro de sus etapas de diseño del Data Warehouse, puede eliminar esta práctica organizacional al definirse claramente las relaciones entre los datos de cada área funcional o proceso de la empresa, asimismo los usuarios de la información pueden entender y visualizar a cada tópico del negocio desde un punto de vista global o corporativo, eliminando con esto inconsistencias en la definición de los datos y redundancia en el almacenamiento de los mismos.

III.5. PRUEBAS O EXPERIMENTACION DE LA TECNOLOGIA

El objetivo de esta etapa es poner a prueba la nueva tecnología, para verificar cuál es el grado de acoplamiento a las necesidades reales y específicas de la organización.

Actividades a realizar y aspectos a considerar para el Data Warehouse

- Realizar las pruebas a través de una aplicación real de la organización.

La aplicación real del Data Warehouse que vaya a realizarse debe ser lo suficientemente pequeña para terminarse rápidamente, pero a la vez debe ser lo suficientemente importante para impactar en la organización. Dicha aplicación debe incluir las características más importantes en el diseño del Data Warehouse, tal como la sumarización, partición, acceso a datos, etc., y debe también enfocarse a un tópico específico de la empresa.

Esta aplicación real de prueba puede servir para identificar los aspectos de diseño que pueden representar obstáculos para la implantación exitosa de la tecnología, asimismo se pueden ir analizando las herramientas de software que proporcionan los diversos proveedores de Data Warehouse, con el fin de identificar las herramientas que mejor se adapten a las plataformas tecnológicas y a las necesidades de información de la organización, es decir, determinar los requerimientos técnicos para la arquitectura de información Data Warehouse. Otro beneficio que se puede obtener de las pruebas del Data Warehouse es el establecimiento de mecanismos de retroalimentación entre los usuarios y los diseñadores de la tecnología, ya que desde las mismas pruebas se puede ir creando un ambiente de trabajo donde prevalezca la comunicación y la retroalimentación entre los involucrados en el proyecto, aspectos tan importantes en la implantación de cualquier nueva tecnología, y en especial con el Data Warehouse, en donde se requieren actualizaciones constantes a los datos que ahí se almacenan, de acuerdo a las necesidades cambiantes de la organización.

- Realizar presentaciones de los resultados de las pruebas a los interesados.

Las presentaciones que se lleven a cabo de los resultados de la aplicación de prueba del Data Warehouse deben estar enfocadas a tres aspectos:

- Mostrar los beneficios potenciales de la tecnología, si los resultados fueron positivos y se obtuvieron en un corto plazo, con esto se puede justificar en parte el análisis costo-beneficio realizado previamente.
- Identificar y presentar los obstáculos que surgieron tanto tecnológicos, humanos como organizacionales, de tal manera que puedan establecerse compromisos de mejora, siendo partícipes de tales compromisos, tanto la alta administración, como los coordinadores del Data Warehouse y el personal del área en donde se realizó la aplicación. El análisis de dichos obstáculos puede ser útil en la prevención de otros que pudieran presentarse en diferentes áreas funcionales de la empresa, por lo que es importante conocer el grado de interconexión de un área funcional con el resto de la organización para determinar en cierto momento el nivel de afectación que pueda ocasionar las barreras organizacionales de dicha área a las demás.
- Presentar la evaluación del grado de adopción que tuvo la tecnología Data Warehouse durante la implantación de las pruebas, analizando los parámetros de medición que se hayan establecido para así definir acciones que mejoren o mantengan el rendimiento del proceso de asimilación.

- Necesidad de cierto nivel inicial de soporte, capacitación y educación.

Como uno de los obstáculos de carácter humano es lo referente a la falta de capacitación del personal para entender y aplicar de manera eficiente las nuevas herramientas de software, se necesita que los mismos proveedores de la tecnología, así como miembros del grupo de desarrollo del Data Warehouse, impartan capacitación principalmente en el uso de las herramientas, y a la vez, que fomenten en la gente una cultura de información, de análisis y conocimiento. Promover una cultura de información dentro de la empresa es un aspecto muy importante para lograr que la tecnología Data Warehouse sea aceptada y asimilada en los procesos de toma de decisiones de la organización, ya que si el recurso humano visualiza a la información como un activo, así también verá a las tecnologías que hagan más eficiente el manejo de la misma. En esta actividad se deben establecer las bases para un programa de concienciación en el uso de la información y en las ventajas competitivas que puede proporcionar a una organización.

III.6. DECISION DE ADOPCION, RECHAZO O POSPOSICION

El objetivo de esta etapa es tomar la decisión de la adquisición y uso de la nueva tecnología.

Actividad a realizar para el Data Warehouse

- Sobre la base de los resultados de las pruebas realizadas, en el entendimiento de la nueva tecnología y en la identificación de los beneficios potenciales que justifiquen su costo, la alta administración va a decidir si adoptar, rechazar o posponer el diseño de la arquitectura Data Warehouse en la organización. En caso de adopción se debe documentar formalmente las razones o justificaciones de la decisión. Dicho documento debe estar aprobado por la alta administración. Por otro lado, si se decidió posponer la introducción de la innovación, se debe llevar a cabo una evaluación de aquellos aspectos que resultaron bajos en su grado de adopción, a fin de concentrarse en ellos y buscar más adelante una respuesta de adopción de la tecnología.

Atributos que influyen en el proceso de asimilación	Ventaja relativa			Compatibilidad			Complejidad			Experimentabilidad			Observabilidad		
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
<p>* Estado actual de los datos</p> <ul style="list-style-type: none"> - Limpieza - Disponibilidad - Integración <p>* Determinar las herramientas Data Warehouse para la transformación de los datos</p> <p style="text-align: center;">↓ ↓ ↓ ↓</p> <p>* Creación de análisis</p> <p>* Conocer la situación de los datos operacionales</p> <p>* Determinar los datos que van a ser requeridos por el DW</p> <p>* Definir lo que representa cada tabla "fact"</p> <p>* Desarrollar un plan de migración al ambiente DW determinando las áreas más críticas de la empresa</p> <p>* Determinar los requerimientos técnicos para la arquitectura de información DW</p> <p>* Invertir en programas de capacitación en diversas tecnologías de información DW</p> <p>Etapas del Data Warehouse involucradas</p>				<ul style="list-style-type: none"> - ¿Va a soportar la nueva tecnología las necesidades de análisis de información más importantes de la empresa? - ¿Son positivos los resultados del análisis costo-beneficio en la evaluación del Data Warehouse? - ¿Se percibe a la tecnología Data Warehouse como una herramienta que permita aumentar la productividad del usuario final? 	<ul style="list-style-type: none"> - ¿Las estrategias de la organización pueden apoyarse en la información que el Data Warehouse va a procesar? - ¿Existe en la empresa una cultura de compartición de datos? - ¿Cuál es el sentir de los individuos de la empresa hacia las tecnologías de información y su capacidad de mejorar los métodos de trabajo que siempre han usado? - ¿Existe un compromiso de la alta dirección para solucionar los problemas y obstáculos presentados en el desarrollo de la aplicación de prueba del Data Warehouse? - ¿Existen grupos de poder en la organización? 	<ul style="list-style-type: none"> - ¿Se tienen las facilidades para "limpiar" los datos sucios de la empresa? - ¿Es factible integrar las nuevas herramientas de transformación de datos a las herramientas de software ya existentes? - ¿Se manejan grandes volúmenes de datos en la empresa? - ¿Existen muchas bases de datos independientes en la organización? - ¿Pueden derribarse las barreras organizacionales que afectan el intercambio de información entre los procesos? - ¿Se tienen las herramientas DW adecuadas para el procesamiento de los datos? 	<ul style="list-style-type: none"> - ¿Es posible diseñar un Data Warehouse para solo una área funcional de la empresa, donde se incluyan los aspectos más importantes de la arquitectura de información Data Warehouse? - ¿Se aplicaron todos los procesos DW en la aplicación prueba sin el apoyo directo de otras áreas funcionales de la empresa? 	<ul style="list-style-type: none"> - ¿Los resultados obtenidos de la aplicación de prueba del Data Warehouse fueron positivos y alcanzados en corto tiempo? - ¿Ha aumentado en la empresa el interés en el uso eficiente de los datos, la información y el conocimiento? - ¿Cumplió la aplicación prueba del DW con los objetivos establecidos y difundidos? 							

Grado de Adaptación
 1 - Bajo
 2 - Medio
 3 - Alto

Grado de Adaptación (Complejidad)
 1 - Alta
 2 - Media
 3 - Baja

Tabla 5.7

DESCRIPCIÓN DE LA TABLA 5.7 – GRADO DE ADOPCIÓN, FASE III

En esta fase del proceso de asimilación se decide si se adopta, se rechaza o se pospone el uso de la tecnología Data Warehouse en la empresa, para lo cual son analizados diversos aspectos organizacionales, funcionales, políticos y financieros, los cuales en un momento dado influyen en la decisión final. Para el Data Warehouse, se visualizan 2 atributos que influyen en el proceso de asimilación y que afectan a los aspectos antes mencionados: el estado de los datos en la empresa y la selección de las herramientas Data Warehouse para la transformación de los datos, esto es, desde la extracción hasta la conversión. Las etapas del diseño del Data Warehouse que se son influenciadas por estos atributos son las que se mencionan a continuación.

a) Creación de Análisis

En esta fase del proceso de asimilación se determina la cantidad de procesamiento analítico que se llevará a cabo en la arquitectura de información Data Warehouse, mientras mayores sean los requerimientos de análisis, mayor va a ser la cantidad de datos a transformarse del ambiente operacional al ambiente Data Warehouse. Para realizar los procesos de transformación de datos (extracción, conversión, integración) se necesita conocer con anticipación la situación de los datos y las herramientas Data Warehouse que se adapten a tales procesos, por lo cual tanto el conocimiento de los datos como de la tecnología son importantes para satisfacer las necesidades de análisis de la empresa. La eficiencia en la definición de las estrategias de transformación de datos puede influir en la evaluación racional-financiera que se haga del Data Warehouse.

b) Conocer la situación de los datos operacionales

Esta etapa del diseño del Data Warehouse es base para determinar el nivel de estandarización de los datos a nivel operacional, lo cual es fundamental para lograr el nivel deseado de interconexión en la organización durante el diseño de la arquitectura de información Data Warehouse, siendo esto último una actividad en la evaluación de la aptitud organizacional.

c) Determinar los datos que van a ser requeridos por el Data Warehouse

La implantación de un Data Warehouse en una empresa generalmente implica altos costos en almacenamiento en disco y en hardware requerido para el procesamiento de los datos. Para evitar que estos costos sean muy elevados, se necesita que las estrategias de sumarización y partición de los datos sean

aplicadas solamente con la información que realmente va a soportar la toma de decisiones o el análisis de algún tópico importante del negocio, esto reduciría el espacio en disco necesario para el almacenamiento de los datos históricos y el procesamiento de los datos para realizar los procesos del Data Warehouse, influyendo finalmente en la evaluación financiera de esta tecnología.

d) Definir lo que representa cada tabla “fact”

Una de las prácticas organizacionales que más puede influir en la evaluación de la aptitud organizacional es el sentido de pertenencia de los datos por parte de los usuarios finales, es entonces importante identificar dentro del contexto organizacional las barreras que puedan impedir el adecuado compartimiento de los datos. Para poder identificar estas barreras organizacionales es necesario primero establecer que datos son necesarios integrar de cada área funcional, por lo cual un adecuado diseño de las tablas “fact” en cualquier nivel de la arquitectura Data Warehouse puede permitir identificar tales datos, de acuerdo a la información a sumarizar de cada tabla.

e) Desarrollar un plan de migración al ambiente Data Warehouse determinando las áreas más críticas de la empresa.

El plan de migración al ambiente Data Warehouse debe tomar en cuenta el nivel de resistencia al cambio que existe en cada área funcional de la empresa, a fin de evitar retrasos en el proyecto, o peor aún, una falta de aceptación hacia la nueva tecnología. La resistencia al cambio puede presentarse por varias razones, entre ellas tenemos la afectación que puede tener una innovación tecnológica en los grupos de poder de la empresa, asimismo por que la implantación de la nueva tecnología va a cambiar radicalmente la manera de hacer las cosas por parte de los miembros del negocio, sean éstas u otras razones, hay que tomarlas en cuenta al momento de realizar el plan de migración al ambiente Data Warehouse.

f) Determinar los requerimientos técnicos para la arquitectura de información Data Warehouse.

Durante las pruebas de la tecnología Data Warehouse deben establecerse claramente los requerimientos técnicos para la creación de la arquitectura de información, es decir, las pruebas de una aplicación real en la organización puede servir como base en menor escala para determinar los procesos más importantes de la arquitectura Data Warehouse y las herramientas de software necesarias para su aplicación.

- g) Invertir en programas de capacitación en diversas tecnologías de información relacionadas con el Data Warehouse.

A través de las pruebas de la tecnología Data Warehouse pueden identificarse deficiencias en la operación de las herramientas de software, la dificultad en el uso de las herramientas Data Warehouse puede generar una actitud de rechazo hacia la innovación tecnológica. Así los programas de capacitación deben orientarse a lograr dos objetivos principales, el primer es que el usuario aprenda a operar eficientemente cada herramienta de software, y el segundo es que a través de tales herramientas el personal acepte y asimile la arquitectura de información Data Warehouse en su totalidad.

Medidas de Evaluación para el grado de adopción

Ventaja relativa

Con esta medida de evaluación se pretende determinar si la tecnología Data Warehouse se percibe como una herramienta que pueda ser usada de manera cotidiana para aumentar la productividad del usuario final, y desde un punto de vista de arquitectura de información, si es capaz de motivar la compartición de los datos entre los miembros de la organización para lograr así una interconexión total de la empresa. Además, se intenta medir si los resultados del análisis costo-beneficio son positivos para la aceptación de la innovación Data Warehouse.

Compatibilidad

A través de esta medida de evaluación se pretende determinar si la aptitud organizacional es la adecuada para aceptar a la tecnología Data Warehouse, principalmente en lo que se refiere a la existencia de grupos de poder que presenten obstáculos en la compartición de los datos. Asimismo es importante evaluar con esta medida, el grado en el cual los procesos Data Warehouse pueden adaptarse a los métodos de trabajo usados en la empresa, y así evitar cambios radicales en las actividades de los miembros del negocio.

Complejidad

El objetivo de esta medida de evaluación es definir si la situación de los datos en la empresa es muy complicada para lograr que la tecnología Data Warehouse se implante en los tiempos planeados. Otro aspecto a considerar en esta medida es la facilidad en la operación de las diversas herramientas Data Warehouse y la

capacidad de éstas para soportar los niveles de procesamiento de datos que pueden presentarse en algunos procesos de la arquitectura de información Data Warehouse

Experimentabilidad

Con esta medida se pretende determinar si la arquitectura de información Data Warehouse puede implantarse por etapas en diversas áreas de la empresa, es decir, si los procesos de esta arquitectura pueden realizarse sin la necesidad de contar con procesos completos Data Warehouse en otras áreas funcionales relacionadas al área de aplicación.

FASE IV: INTRODUCCION-ADOPCION

Esta fase involucra las decisiones, estrategias, eventos, acciones y actividades necesarias que permitan poner a la innovación en uso rutinario en la organización.

En esta fase de la asimilación tecnológica se pretende que los procesos de la arquitectura de información Data Warehouse sean aceptados como actividades rutinarias en la empresa, asimismo que las herramientas Data Warehouse sean operadas y aceptadas como herramientas de productividad por los miembros de la organización. El plan de introducción de la tecnología debe adaptarse a los parámetros observados durante la evaluación del grado de adopción de la tecnología en las primeras fases de este proceso de asimilación, de tal manera que todos los aspectos tecnológicos, herramientas, estructura en los datos, diseño de las tablas “fact” y dimensionales, se lleven a cabo sobre la base de los resultados de las evaluaciones del grado de adopción, para así lograr finalmente la asimilación de la tecnología Data Warehouse en la empresa.

IV.1. REDEFINICION, REESTRUCTURACION Y PLANEACION DE LA INTRODUCCION DE LA TECNOLOGIA

El objetivo de esta etapa es que la innovación se modifique o se adapte para acoplarse a la situación particular de la organización.

Actividades a realizar y aspectos a considerar para el Data Warehouse

- Reconocer que la estructura de la organización y su contexto influyen en el flujo de la información relacionada con la nueva tecnología.

Se ha mencionado que el valor real de un Data Warehouse es el control del flujo de la información de una empresa y no tanto la colección de datos; en el caso del diseño del Data Warehouse, mientras más niveles jerárquicos existan en un negocio, mayor es la complejidad para que esta tecnología controle el flujo de información organizacional. Recordemos que el Data Warehouse es una arquitectura de información compuesta de diversos niveles de abstracción o sumarización de datos, y que cada nivel representa una medida o tópico posible del negocio y generalmente se relaciona a un nivel organizacional dentro de la empresa; basándose en esto es que se afirma que mientras más niveles existan en una estructura organizacional mayor es la cantidad de niveles de sumarización o abstracción en los datos, así como mayor la cantidad de estrategias de sumarización, partición, filtros y conversión de datos.

Peter Drucker [DRUCKER,1988] menciona que las empresas ahora están buscando tener estructuras organizacionales más horizontales, en donde la información pueda fluir más fácilmente y donde más personas tomen decisiones; además menciona que las tecnologías de información deben ser promotoras en la creación de empresas basadas en información y en conocimiento, donde haya más especialistas y donde las estructuras organizacionales “planas” permitan que cada miembro de la empresa tenga la información que necesite para realizar su trabajo.

La idea de Drucker de tener empresas basadas en información, en conocimiento y con estructuras organizacionales planas, sería lo ideal para implantar la arquitectura de información Data Warehouse, pero se debe aceptar la situación de una empresa y hacer que la tecnología se adapte al contexto organizacional y a sus procesos. Este proceso de asimilación de tecnología es una herramienta de análisis para soportar la adaptación de la tecnología Data Warehouse a la empresa.

- Realizar un plan formal para introducir la tecnología.

A partir de esta fase del proceso de asimilación se hace referencia al diseño del Data Warehouse como un proyecto aceptado por la alta administración. En primer lugar, se deben establecer los objetivos o metas de la nueva tecnología, es decir, determinar de manera más precisa lo que se pretende obtener con esta innovación tecnológica; los objetivos deben ser medibles y delimitados en el tiempo para que puedan ser factibles a evaluarse y mejorarse.

En el análisis costo-beneficio de la etapa III.2 se mencionaron diversos beneficios que se pueden obtener al implantar un Data Warehouse. Probablemente algunos de estos beneficios se relacionan a algún objetivo que se pretenda alcanzar en nuestra organización, por ejemplo, si hablamos de la detección de fraudes, la organización puede establecer como objetivo del Data Warehouse, el disminuir en un 50% las acciones fraudulentas en 1998 al mejorar el control de las transacciones financieras en la empresa.

Los objetivos que se establezcan para la arquitectura de información Data Warehouse deben ser claros, realistas a la situación particular de la empresa, estar bien orientados, y con sus alcances debidamente delimitados. El análisis previo de la organización es muy importante tanto desde el punto de vista del negocio, como tecnológico y humano, para el adecuado establecimiento de los objetivos del Data Warehouse, siendo esto último fundamental para elaborar un plan de introducción de la tecnología donde se plasmen las estrategias necesarias para alcanzar dichos objetivos.

De los objetivos se pueden derivar las áreas o aspectos del negocio en donde primero se implantaría el Data Warehouse, por ejemplo, si el objetivo es aumentar el número de clientes de una determinada línea de producto, las áreas del negocio directamente involucradas serían Ventas y Mercadotecnia; es necesario identificar también los procesos de información relacionados a dichas áreas del negocio y al objetivo que se pretende alcanzar, ya que el plan de introducción del Data Warehouse debe incluir aspectos técnicos como el diseño de las tablas de datos asociadas a cada proceso de información. La identificación de las áreas del negocio en donde se va a implantar la tecnología es una actividad prioritaria en este plan de introducción.

El plan de introducción debe contemplar por otro lado, la elaboración de un programa de capacitación a los posibles usuarios de la nueva tecnología, dependiendo por supuesto, de las áreas del negocio donde se vaya a implantar inicialmente; definir la reorganización del trabajo dentro de esas áreas; identificar o crear el flujo de información mediante procedimientos bien elaborados, donde se especifique las fuentes y los usuarios de la información de un determinado proceso; identificar los responsables del proceso de información para cada área del negocio y los procesos relacionados y sus responsables. El plan de introducción del Data Warehouse debe también incluir la asignación de recursos humanos, financieros y el tiempo para cada fase del proyecto. Al hablar del recurso humano nos referimos al personal involucrado con la información en el área donde se va a implantar el Data Warehouse, así como el personal técnico informático “especialista” de dicha área. Las estrategias de sumarización y partición de datos que implican costos en cuanto a la adquisición o creación de herramientas de software, así como el hardware requerido para el almacenamiento de datos durante el proceso de partición, dependerán del presupuesto que se asigne en el plan, por lo cual es necesario que se evalúe adecuadamente el alcance que va a tener el Data Warehouse para que se refleje en el presupuesto que se vaya a autorizar.

- Aprobación del primer proyecto a realizar (piloto)

El proyecto piloto debe ser impactante en cuanto a la visibilidad de sus resultados, es decir, los promotores de la tecnología y la alta administración deben visualizar los beneficios que pueden generarse con esta tecnología.

Bill Inmon [INMON,1996] menciona que la prueba piloto del Data Warehouse puede ser hecha rápidamente y con muy poco presupuesto, y que a partir de esta prueba, los analistas de decisiones podrán empezar a explorar las posibilidades y las ventajas del Data Warehouse más eficientemente. Asimismo, menciona que la

prueba piloto de un Data Warehouse generalmente se aplica en las siguientes áreas del negocio:

- Finanzas
- Mercadotecnia
- Ventas

Un aspecto muy importante en el diseño del Data Warehouse es identificar y establecer las métricas o medidas de éxito del proceso de implantación para su evaluación y mejora. Este proyecto piloto nos va a permitir reafirmar las medidas de éxito que se hayan establecido previamente, o bien identificar algunas otras que los analistas de decisiones consideren importante para realizar una evaluación más eficiente.

- Considerar el posible rediseño del trabajo y las reestructuraciones de los equipos de trabajo de la organización.

Si la introducción de la tecnología Data Warehouse en una organización implica cambios fuertes en la manera en como se trabaja y en la manera en como están formados los grupos de trabajo, se necesita realizar antes de introducir la nueva tecnología, una programa amplio de educación y capacitación al recurso humano, enfatizar no solo en “cómo hacerlo”, sino también en el “porqué” se está haciendo. Si el recurso humano entiende el porqué de la tecnología y sus beneficios, el rediseño del trabajo y los cambios en los equipos se realizarán sin muchos problemas. Los usuarios finales deben conocer como usar y obtener la información, y entender a la vez, los diversos pasos que conforman el proceso que genera la información que ellos utilizan. Los usuarios que son a la vez fuentes de información, deben conocer como acceder la información que ellos necesitan y como filtrar dicha información a otros niveles de abstracción, por lo cual deben entender la correspondencia de los datos que ellos usan con la información que requieren en otros niveles organizacionales. De aquí la importancia de no solamente capacitar en el conocimiento del Data Warehouse, sino también educar en el entendimiento de esta tecnología.

IV.2. APRENDIZAJE Y CLARIFICACION DEL USO DE LA TECNOLOGIA

Con esta etapa se pretende que los miembros de la organización tengan un contacto más directo con la tecnología, y su significado llegue a ser completamente comprendido.

Actividades a realizar y aspectos a considerar para el Data Warehouse

- Llevar a cabo la realización del proyecto piloto, poniendo especial énfasis en el aprendizaje y aplicación de la nueva tecnología

Una vez definida el área donde se va a aplicar el proyecto piloto, el siguiente paso es comenzar con el diseño técnico de la arquitectura de información Data Warehouse. Hay semejanzas en las propuestas de diseño de un Data Warehouse presentadas en el apartado 5.4, la mayoría de los autores coinciden en el desarrollo de un *modelo de datos* que permita agruparlos lógicamente para que representen la operación de una empresa. Ahora bien, el proyecto piloto contempla solamente un proceso o área funcional del negocio y esto debe verse reflejado cuando se definan las fronteras en el modelo de datos. Es importante que al llegar a esta etapa del proceso de asimilación ya se haya desarrollado o actualizado, si ya existía, el modelo de datos, para que al momento de realizar el proyecto piloto solamente se delimite el alcance de dicho modelo y se continúe con el siguiente paso del diseño del Data Warehouse.

Del área funcional o proceso en donde se aplica el proyecto, pueden derivarse una o más medidas posibles del negocio que los analistas de decisiones requieran en primera instancia para explorar la información del Data Warehouse, la *modelación dimensional* propuesta por Kimball (ver 5.4.1.), permite que los usuarios finales visualicen a los datos desde un punto de vista del negocio y de acuerdo a cada medida posible de la empresa. Una vez identificados los principales requerimientos de información de los usuarios o analistas de decisiones de una área funcional, y posteriormente reflejado en el modelo de datos, se hace más fácil la tarea de determinar las *fuentes de datos* que van a emplearse para conformar la arquitectura de información Data Warehouse.

El proyecto piloto debe ser pequeño pero impactante, de tal manera que la información que sea producto final del Data Warehouse se forme gracias a un proceso que involucró a la mayoría de los aspectos del diseño de un Data Warehouse. Quizá nada más se definan tres niveles de abstracción de datos y se requieran pocos procesos de filtración, conversión, sumarización o formateo de datos, pero lo importante es que ahí estén y que funcionen, y que integren a la vez diversas fuentes de datos de esa misma área funcional o quizá fuentes de datos de diversas áreas. La integración de los datos entre diversas fuentes va a permitir a los diseñadores del Data Warehouse identificar prácticas organizacionales que no permitan que la información fluya adecuadamente, así también visualizar conflictos entre el recurso humano, y por supuesto, probar las herramientas de conectividad y las redes informáticas aplicadas en el proyecto.

Si ya se tienen las fuentes y sus atributos identificados, se pueden *diseñar las tablas "fact"* del Data Warehouse, así como también las tablas dimensionales que representen a cada requerimiento o medida del negocio. Con la modelación de datos se identifican las relaciones existentes entre las tablas diseñadas, y de dichas relaciones pueden surgir las estrategias de sumarización y partición de datos que permitan ir formando niveles de detalle en la información; en cada nivel de información deben diseñarse las tablas principales y dimensionales que representen a los datos fielmente.

Si estamos haciendo referencia a una área funcional de la empresa, podemos entonces crear un Data Mart para dicha área, así la información que se vaya generando en el Data Warehouse puede almacenarse en un servidor de datos único que represente al Data Mart; este servidor posteriormente estará enlazado a otros servidores de la empresa conteniendo un Data Mart, y de esta manera se estaría creando un Data Warehouse corporativo. Pero en esta fase del proceso de asimilación estamos hablando de un Data Mart para un proyecto piloto, por lo cual se hace necesario explotar la información contenida en el mismo y que de valor a todo el proceso de información que se haya diseñado. La explotación de la información estará a cargo de las herramientas de análisis que previamente se hayan evaluado, pueden ser reportadores, macros en hojas de cálculo, o bien, desarrollar un sistema de soporte a la toma de decisión o un sistema de información ejecutivo que funcione para tales propósitos, con todas estas herramientas se pueden generar diversos tipos de análisis, a excepción de los EIS, ya que tienen como función solo acceder y visualizar información que los ejecutivos necesitan.

El proyecto piloto, aunque pequeño, debe contemplar los aspectos más importantes en el diseño del Data Warehouse, y puede servir tanto para evaluar todos los recursos aplicados al proyecto prueba, como para visualizar tendencias que pueden presentarse en otras áreas funcionales de la organización.

- Seleccionar el personal adecuado para participar en el equipo del proyecto introductorio y mantener mucha comunicación entre todos los miembros del grupo de desarrollo.

Del área funcional de aplicación del proyecto piloto deben seleccionarse personas con experiencia y con conocimientos de la información que ahí se procesa, ya sea de manera automática o manual, para que participen en el diseño del proyecto y junto con los analistas de decisiones determinen los tópicos y los requerimientos de información más importantes de su área. El personal seleccionado debe tener ya conocimiento del Data Warehouse y de los elementos que conforman esta arquitectura de información, así también conocer el plan introductorio de esta

tecnología y el plan informático desde donde se derivan las prioridades de implantación del Data Warehouse en la organización.

- Recordar que es necesario darle el poder al empleado para realizar el trabajo.

La motivación y el darle poder al empleado para tomar decisiones son factores importantes para lograr la asimilación exitosa de la tecnología Data Warehouse en una empresa. La capacitación y la educación continua contribuyen a elevar la motivación del personal al tener conocimiento de la importancia que tiene su participación en el proceso global de información del Data Warehouse. Dentro de los procesos del Data Warehouse existen algunos en los cuales se requiere tener medidas de seguridad para llevarlos a cabo, por ejemplo, algún proceso de filtración o sumalización de datos que afecte a las tablas históricas de otro nivel de información, o bien procesos de descarga de datos hacia medios de almacenamiento secundarios para su respaldo, al no ser ya de valor para seguir almacenados en las bases de datos del Data Warehouse. Generalmente estos procesos se realizan en intervalos de tiempo ya definidos, pero en otras ocasiones se solicitan de urgencia para satisfacer alguna demanda de información en la empresa, como por ejemplo algún reporte corporativo. La repetición de algún proceso de este tipo puede ocasionar duplicidad en la información de algunas tablas “fact”, afectando posteriormente la información de otro nivel de sumalización más alto.

Si bien, deben existir medidas de seguridad que evite este tipo de errores, lo mejor sería darle seguridad al personal que realiza estos procesos y el poder para que tome la decisión de ejecutarlos, que esté consciente de la responsabilidad que tiene, pero a la vez que no sienta presión por lo que pudiera suceder si se comete algún error.

IV.3. DIFUSION Y RECONOCIMIENTO DE LOGROS DEL USO DE LA TECNOLOGIA

El objetivo de esta etapa es reconocer el esfuerzo realizado por los miembros de la organización para adoptar la nueva tecnología.

Aspectos a considerar para el Data Warehouse:

- Deben reconocerse los esfuerzos realizados por el personal para mejorar el proceso de información de la empresa

Los resultados que se hayan obtenido de la implantación de la prueba piloto del Data Warehouse deben difundirse a través de la organización, haciendo énfasis en las mejoras que se hayan obtenido en el flujo de la información en el área de aplicación, así como en la participación del personal para lograr que todos los procesos del Data Warehouse se lograrán llevar a cabo. Por otro lado, especificar que el proyecto no está terminado, sino que requiere una retroalimentación continua para mejorar su rendimiento, y por supuesto presentar el plan de adaptación del Data Warehouse a otras áreas funcionales o procesos de la organización.

- **Identificar a los impulsores de la tecnología**

El reconocimiento y los premios para los que impulsan la tecnología Data Warehouse y defienden los cambios organizacionales que conlleva, deben hacerse públicos para que la organización conozca al personal que hace posible que la nueva tecnología se implante exitosamente en la organización. Hay que recordar que el Data Warehouse no es proyecto con un inicio y un final, sino más bien, es un proceso que requiere una retroalimentación continua que permita lograr que las bases de datos del Data Warehouse reflejen realmente las necesidades de información de la empresa. La retroalimentación necesita de la participación activa del personal, y mientras más enterados estén de los beneficios de la nueva tecnología, que conozcan el proceso que se está llevando a cabo, a los participantes, y demás, mayor va a ser la motivación para seguir contribuyendo en el proyecto y mejor va a ser el proceso de retroalimentación que refleje los nuevos cambios que requiera la tecnología para adaptarse a las necesidades del negocio.

IV.4. ADOPCION, ACEPTACION Y RUTINIZACION DEL USO DE LA TECNOLOGIA

El objetivo que persigue esta etapa es que la innovación pierda su identidad de separación y llegue a ser un elemento en las actividades rutinarias de la organización

Aspecto a considerar para el Data Warehouse

- **Evaluar la motivación y satisfacción de los miembros de la organización al realizar su trabajo utilizando la nueva tecnología.**

Mediante encuestas a los usuarios del Data Warehouse se puede determinar el grado de satisfacción hacia la nueva tecnología. Aparte de los procesos de retroalimentación que se realicen, es importante conocer lo que piensa cada usuario, conocer los obstáculos que se presentaron al cambiar su manera de trabajar y las dificultades que se tuvieron al utilizar las herramientas Data Warehouse. Una medida de evaluación para determinar la satisfacción del usuario con respecto al Data Warehouse es el grado de retroalimentación que proporcionen del diseño y de los componentes de esta arquitectura de información, del interés que muestren por la información que se esté generando de esta tecnología, y por las propuestas que realicen para alcanzar nuevos tópicos del negocio o necesidades de información que se consideren estratégicos en un área funcional o en toda la empresa.

Un objetivo que debe alcanzarse durante la operación del Data Warehouse es que el departamento de sistemas de información de la empresa se involucre en lo más mínimo en las actividades de los usuarios finales para ejecutar los procesos que conforman la arquitectura Data Warehouse; una vez que el usuario no necesite el soporte del personal informático, la tecnología Data Warehouse se convertirá en una herramienta rutinaria en las actividades de la organización, siendo esto último, el propósito principal que se persigue con el diseño de la arquitectura de información Data Warehouse empleando este proceso de asimilación de tecnología.

GRADO DE ADOPCION DE LA INNOVACION DATA WAREHOUSE

FASE IV: INTRODUCCION - ADOPCION

Atributos que influyen en el proceso de asimilación	Ventaja relativa			Compatibilidad			Complejidad			Experimentabilidad			Observabilidad		
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
<p>* La tecnología de información requerida para desarrollar el proyecto piloto</p> <p>* Las prácticas organizacionales involucradas en el proyecto piloto</p> <ul style="list-style-type: none"> - Compartición de los datos - Políticas de acceso a los datos - Aceptación de la tecnología - Nivel de pertenencia de los datos <p style="text-align: center;">↓ ↓ ↓ ↓</p> <p>* Definir claramente los objetivos del negocio que se van a perseguir con el diseño del DW</p> <p>* Seleccionar los procesos, y como consecuencia identificar las tablas "fact"</p> <p>* Definir las estrategias de sumarización, las formas de consulta, las prioridades en las consultas, el almacenamiento físico de los datos, etc.</p> <p>* Determinar los requerimientos técnicos para la arquitectura de información DW</p> <p>Etapas del Data Warehouse involucradas</p>															
	- ¿Han mejorado los procesos de información a raíz de la implantación del proyecto piloto del Data Warehouse?			- ¿Existen muchos niveles jerárquicos en la estructura organizacional de la empresa?			- ¿Está bien definido el flujo de información en la empresa para establecer las estrategias de sumarización y partición de los datos?			- ¿El proyecto piloto es lo suficientemente impactante para clarificar el uso, las características y los beneficios de la tecnología Data Warehouse?			- ¿Permite el proyecto piloto visualizar algunas tendencias en el diseño del DW que pudieran presentarse en otras áreas funcionales de la organización?		
							- ¿Ha sido entendida la arquitectura Data Warehouse para la realización del proyecto piloto?			- ¿Se cuenta con el personal adecuado para desarrollar el proyecto piloto, tanto personal informático como gente del negocio?			- ¿Ha sido aceptada la tecnología Data Warehouse hasta convertirse en una herramienta rutinaria de trabajo?		
							- ¿Son identificables las medidas posibles del negocio para explorar la información del Data Warehouse?								
							- ¿Se presentaron muchos problemas técnicos al implantarse el proyecto piloto?								
							- ¿Se presentaron muchos problemas organizacionales al implantarse el proyecto piloto?								

Grado de Adaptación

- 1 - Bajo
- 2 - Medio
- 3 - Alto

Grado de Adaptación (Complejidad)

- 1 - Alta
- 2 - Media
- 3 - Baja

Tabla 5.8

DESCRIPCIÓN DE LA TABLA 5.8 – GRADO DE ADOPCIÓN, FASE IV

En esta fase del proceso de asimilación es donde más se reflejan las diversas etapas de desarrollo del Data Warehouse, debido a que en esta fase se contemplan actividades que van desde la introducción hasta la rutinización y mejora continua de esa tecnología. Así cualquiera de las actividades que conforman esta fase del proceso de asimilación puede abarcar todo el ciclo de diseño propuesto para el Data Warehouse. La realización del proyecto piloto es la actividad esencial de esta fase del proceso de asimilación, más sin embargo, para fines de evaluar el grado de adopción de la tecnología, la actividad a evaluarse en esta fase es la elaboración del plan de introducción del Data Warehouse, considerando para esto las diversas etapas de las propuestas de diseño del Data Warehouse mencionadas en este capítulo. Las actividades que a continuación se presentan son las que se consideran que tienen una relación más directa con el plan de introducción y adopción de la tecnología Data Warehouse en la organización.

- a) Definir claramente los objetivos del negocio que se van a perseguir con el diseño del Data Warehouse.

La identificación de los objetivos del negocio que se pretendan alcanzar con la implantación de la arquitectura de información Data Warehouse, puede ser una base para enfocar el plan de introducción de esta arquitectura hacia las áreas o tópicos del negocio que tengan una relación directa con las actividades necesarias para el logro de dichos objetivos.

- b) Seleccionar los procesos, y como consecuencia identificar las tablas “fact”.

Además de identificar las áreas o tópicos del negocio para alcanzar ciertos objetivos de la empresa, es necesario seleccionar los procesos de información inherentes a cada área del negocio, y de los cuales puedan surgir los datos que van a satisfacer determinada necesidad de información de un tópico de la organización. Los procesos de información seleccionados van a ser la base para diseñar las tablas “fact” y otros aspectos tecnológicos de diseño incluidos en el plan de introducción del Data Warehouse.

- c) Definir las estrategias de sumarización, las formas de consulta, las prioridades en las consultas, el almacenamiento físico de los datos, etc.

Dentro del plan de introducción del Data Warehouse deben especificarse claramente los recursos que van a aplicarse en el diseño de la arquitectura de información, recursos tanto humanos, en tiempo, como financieros. Los

parámetros para medir la asignación de estos recursos pueden derivarse de la aplicación de las estrategias de los procesos más comunes del Data Warehouse, como son la sumarización y partición de los datos, los tipos de consulta, el almacenamiento físico de los datos, etc. El análisis de las estrategias de los procesos del Data Warehouse puede soportar el alcance del plan de introducción y la asignación de los recursos para su implantación.

- d) Determinar los requerimientos técnicos para la arquitectura de información Data Warehouse.

La realización de la evaluación de las herramientas Data Warehouse en las fases iniciales de este proceso de asimilación sirvió para identificar aquellas herramientas que pudieran adaptarse al diseño de la arquitectura de información Data Warehouse, ahora bien, en esta actividad del proceso de asimilación deben especificarse de las herramientas Data Warehouse seleccionadas, aquellas que se acoplen a la tecnología de información existente en las áreas de aplicación propuestas en el plan de introducción; hay que tomar en cuenta entre otras cosas, la diversidad de formatos de base de datos, plataformas de hardware, software existente, etc. Es importante entonces definir la configuración funcional para cada área de aplicación del Data Warehouse, en el cual deben incluirse principalmente las herramientas de usuario final, como para la consulta de datos, accesos, análisis y reportes.

Medidas de Evaluación para el grado de adopción

Ventaja relativa

Con esta medida de evaluación se pretende determinar la percepción de los usuarios finales con respecto al Data Warehouse como una tecnología que ha mejorado el manejo de la información en la empresa, y específicamente en las actividades que cada usuario realiza.

Compatibilidad

A través de esta medida de evaluación se pretende conocer la compatibilidad de la arquitectura de información Data Warehouse con la estructura organizacional y los procesos de la empresa, además resulta importante conocer la opinión de los miembros del negocio respecto a las medidas que se estén tomando para adaptar la tecnología Data Warehouse a la situación particular del negocio.

Complejidad

Como resultado de esta medida se deben identificar todos los obstáculos presentados durante la implantación del proyecto piloto del Data Warehouse, obstáculos como las barreras organizacionales para el intercambio de los datos, los problemas tecnológicos en la aplicación de algunos procesos, inconsistencia en la definición de las estrategias de sumarización y partición de los datos, problemas con la operación de las herramientas Data Warehouse, etc.

Observabilidad

El objetivo de esta medida de evaluación es conocer si los resultados del proyecto piloto del Data Warehouse puedan ser difundidos y visualizados positivamente en toda la organización, tomando en cuenta que el proyecto piloto se implantó solamente en un área funcional de la empresa. Además es importante conocer si los beneficios obtenidos y los beneficios potenciales del Data Warehouse para otras áreas del negocio, puedan ser percibidos adecuadamente en la organización.

Experimentabilidad

Con esta medida se pretende determinar el impacto de la prueba piloto del Data Warehouse, tomando en cuenta las limitantes existentes durante su implantación, como puede ser la delimitación hacia un área funcional de la empresa, la aplicación de pocos procesos de sumarización y partición de los datos, entre otras.

FASE V: CONFIRMACION - EXPANSION

El objetivo de esta fase es evaluar los resultados de la utilización de la innovación, así como de la implantación realizada.

V.1. EVALUACION DEL USO DE LA TECNOLOGIA

La finalidad de esta etapa es identificar los resultados tangibles del uso de la nueva tecnología, del valor proporcionado al negocio, y del proceso de implantación realizado

Aspectos a considerar para el Data Warehouse

- Comunicar a todos los miembros de la organización los resultados de la evaluación y mostrar los resultados obtenidos contra los esperados.

Los objetivos establecidos para la implantación del Data Warehouse deben ser revisados para determinar el porcentaje de cumplimiento de cada uno de ellos; además es importante la comparación de los resultados esperados contra los reales. Asimismo, es necesario analizar los efectos del uso de la tecnología en la organización: medir la efectividad y la eficiencia resultantes de la utilización de las herramientas Data Warehouse, para lograrlo se deben evaluar diversos aspectos, entre los cuales se mencionan los siguientes:

- Los ahorros financieros (costos) por el flujo más eficiente de la información
- Tiempo de respuesta en los accesos a la información
- Rapidez en la creación de un reporte ejecutivo
- Calidad en la información entregada
- Cantidad de alternativas en el análisis de la información para la toma de decisiones

La evaluación de estos aspectos y otros más, consistiría básicamente en la realización de un comparativo entre su efectividad y eficiencia, antes y después de la utilización del Data Warehouse. Los resultados de la evaluación deben ser difundidos a través de la organización, principalmente en el área o áreas involucradas con la implantación de la prueba piloto. Para fines de presentar resultados con objetividad, sería importante presentar medidas de evaluación más específicas y los beneficios que generan para la organización, por ejemplo, si se está hablando del rendimiento en la creación de reportes ejecutivos, se pueden presentar ahorros de tiempo concretos para integrar la información, sumarla y presentarla

en ambientes gráficos amigables, comparado a la manera en como se hacía anteriormente, y por supuesto, presentar los beneficios que puede generar un reporte ejecutivo a tiempo para la toma de decisiones.

- Evaluar la implantación realizada y el proceso de cambio, reconocer los ajustes necesarios que deben hacerse y realizarlos.

La evaluación de la implantación de la tecnología Data Warehouse podría enfocarse principalmente a los tiempos de realización del proyecto, comparando los tiempos que se tenían previstos con los tiempos reales de implantación. El análisis del grado de adopción que va teniendo la tecnología durante su implantación puede ser un soporte para identificar causas de retardos o inconformidades, así como para definir acciones a tomar para reducir tiempos y costos en el diseño de la arquitectura de información.

Implantar la tecnología Data Warehouse requiere de muchos recursos tanto económicos como en tiempo, por lo cual la evaluación de los retardos de tiempo en las diversas etapas del proceso de implantación deben analizarse concienzudamente para determinar los factores que ocasionaron tales retardos, y así poder realizar los ajustes que sean necesarios para disminuir los tiempos en cada etapa del proceso.

V.2. CONFIRMACION DEL USO DE LA TECNOLOGIA

En esta etapa se pretende obtener la confirmación del uso de la innovación, que ya había sido anteriormente tomada. Si no se obtiene, entonces se suspende el uso de la innovación en la organización.

En esta etapa la alta administración y los diseñadores del Data Warehouse deben formalmente confirmar el uso de la tecnología, revisando tanto el plan informático como el plan propio de la arquitectura de información, y así de esta manera realizar las adecuaciones que sean necesarias para alinear a la tecnología con el negocio, y para definir acciones complementarias derivadas del proceso de asimilación.

V.3. EXPANSION E INSTITUCIONALIZACION DEL USO DE LA TECNOLOGIA

El objetivo de esta etapa es que la tecnología se expanda, actualice o reemplace con una segunda generación del modelo de la innovación.

Aspecto a considerar para el Data Warehouse

- La posible transferencia de personas para difundir la nueva tecnología, así como la experiencia de su utilización

La planeación para el diseño del Data Warehouse debe estar completamente alineado al plan informático, debido a que éste refleja las acciones estratégicas de una empresa y las actividades principales de la misma, y por otro lado, mediante el plan informático el proceso de expansión e institucionalización de la tecnología va a estar bien soportado al conocerse la situación de la TI en la empresa y los proyectos que se estén llevando a cabo. El conocimiento de la tecnología y de la empresa en general ayudará a que las actualizaciones que se hagan al Data Warehouse vayan de acuerdo al entorno cambiante del negocio, y a las plataformas tecnológicas sobre las cuales los diversos procesos de la arquitectura de información Data Warehouse se lleven a cabo.

Evaluación del grado de adopción de la tecnología Data Warehouse

Tomando en cuenta que esta fase del proceso de asimilación esta enfocada a la confirmación y expansión de la tecnología, no se considera viable aplicar la evaluación a una tecnología que ya ha sido adoptada en la empresa, por lo cual el esfuerzo de las actividades de esta fase debe enfocarse a lograr que la tecnología sea expandida y asimilada en toda la organización, y confirmada su uso como una herramienta cotidiana de productividad.

5.6. CONCLUSIONES

Así como las propuestas de diseño del Data Warehouse deben adaptarse a las características y necesidades de una organización, así también el proceso de asimilación de la tecnología Data Warehouse que propongo debe ser flexible y abierto para aplicarse en cualquier tipo de empresa. Así las actividades que establezco en cada una de las fases del proceso de asimilación son recomendaciones generales que podrían aplicarse en cualquier situación organizacional, más sin embargo los diseñadores de la arquitectura de información pueden agregar nuevas actividades que personalicen su análisis de acuerdo las necesidades más específicas en su organización.

De la misma manera, las evaluaciones para determinar el grado de adopción de la tecnología Data Warehouse al final de cada fase del proceso de asimilación se realizan conforme a los atributos que influyen en el proceso y en las etapas de las propuestas de diseño del Data Warehouse presentadas en este capítulo, a lo cual los diseñadores del Data Warehouse pueden agregar nuevos atributos que consideren influyan en la asimilación de la tecnología, y basarse ya sea en otras propuestas o bien en una sola que consideren ser la más viable para usarla en el diseño e implantación de la arquitectura de información. Asimismo, las preguntas que se proponen para evaluar a cada medida de adopción son consideradas como básicas para tener los suficientes parámetros de evaluación en cada medida, de tal manera que al aplicar estas preguntas más las que los diseñadores del Data Warehouse agreguen, puedan dar una tendencia del grado de asimilación y adopción de la innovación tecnológica.

En este proceso de asimilación es importante la asociación de conceptos y la aplicación de metodologías, identificar la relación entre las etapas en el diseño del Data Warehouse y las fases del proceso de asimilación es sin duda el aspecto de asociación más importante que hay que establecer. Así también el entendimiento del Data Warehouse como arquitectura de información y la relación entre sus componentes es factor relevante para desarrollar algunas de las actividades que propongo en el proceso de asimilación, y por otro lado se tiene el conocimiento de la empresa y su alineación con el Data Warehouse, lo cual se puede obtener con el uso de metodologías de análisis como la cadena de medios-fines y la cadena de valor y que pueden facilitar la percepción de la implantación del Data Warehouse a nivel corporativo.

Es importante remarcar que la finalidad de aplicar este proceso de asimilación en una organización es lograr que la tecnología Data Warehouse sea aceptada y

asimilada tanto por el personal que va a operar las herramientas de software, como por los que van a aplicar los procesos que conformarían la arquitectura de información, así como por aquellas personas involucradas con la información que se va a generar, y por supuesto que la tecnología sea vista como una herramienta de soporte a la toma de decisiones entre los niveles ejecutivos y tácticos de la empresa. El esfuerzo de este proceso de asimilación debe estar dirigido a lograr esto, independientemente de las propuestas de diseño, y de las herramientas tecnológicas a emplearse para su implantación.

REFERENCIAS BIBLIOGRAFICAS

- [ALUR,1996] ALUR, Nagraj.
"Data Warehousing & Parallel Technology".
Database Programming & Design, Enero 1996
- [BROWN,1995] BROWN, A.J.
"What is a Data Warehouse?"
Unix Review, Agosto 1995
- [BUSWICK,1990] BUSWICK, Ted.
"Champions of technology innovation".
Training & Development, Febrero 1990.
- [DRUCKER,1988] DRUCKER, Peter.
"The coming of the New Organization".
Harvard Business Review, Enero-Febrero 1988.
- [DRUCKER,1995] DRUCKER, Peter F.
"The information executives truly need".
Harvard Business Review, Enero-Febrero 1995
- [INMON,1996] INMON, W.H.
Building the Data Warehouse.
John Wiley & Sons, Inc, segunda edición, 1996
- [KAY,1996] KAY, Emily.
"A Classic Case of Build vs. Buy"
Software Magazine, Center Market Research, 1996
- [KIMBALL,1996] KIMBALL, Ralph.
The Data Warehouse Toolkit: Practical Techniques for
building dimensional Data Warehouses.
John Wiley & Sons, 1996
- [LADAGA,1995] LADAGA, John.
"Let business goals drive your data warehouse effort".
Health Management Technology, Octubre 1995

- [LUFTMAN,1993] LUFTMAN, J.N., P.R. Lewis & S.H. Oldach.
“Transforming the enterprise: The alignment of business
and information technology strategies”.
IBM Systems Journal, vol. 32. No. 1, 1993
- [MCGEE,1993] MCGEE, James & Laurence Prosak.
Managing information strategically.
John Wiley & Sons, 1993
- [MEREDITH,1996] MEREDITH, Mary, Khader Aslam.
“Divide and Aggregate”.
Database Programming and Design, Junio 1996
- [MEYER,1996] MEYER, Marc H., Michael H. Zack
“The Design and Development of Information Products”.
Sloan Management Review, spring 1996.
- [MYERS,1995] MYERS, Marc
“Do you really need a data warehouse?”
Network World, diciembre 18, 1995
- [MEYER,1988] MEYER, Alan D. & James B. Goes.
“Organizational Assimilation of Innovations: A Multilevel
Contextual Analysis”.
Academy of Management Journal, 1988. Vol.31, No. 4,
- [MORIARTY,1996,1] MORIARTY, Terry.
“Barriers to data quality Part I: Disparate Data”.
Database Programming and Design, Abril 1996
- [MORIARTY,1996,2] MORIARTY, Terry.
“Barriers to data quality Part II: Business Practices”.
Database Programming and Design, Abril 1996
- [NIEDERMAN,1987] NIEDERMAN, Fred & James C. Brancheau.
“Planning and Managing Technology Change”.
Management Information Systems Research Center,
Junio 1987

- [RAMIREZ,1993] RAMIREZ, Rosario.
Administración del proceso de asimilación de tecnología de información en una organización y su aplicación a case
Tesis de maestría ITESM, Monterrey, N.L., 1993
- [ROBINSON,1996,1] ROBINSON, Teri.
“To avoid pitfalls, walk before you run”.
Software Magazine, Center Market Research, 1996
- [ROBINSON,1996,2] ROBINSON, Teri.
“It All Starts with Good, Clean Data”.
Software Magazine, Center Market Research, 1996
- [ROCKART,1979] ROCKART, John.
“Chief Executives define their own data needs”.
Harvard Business Review, Marzo-Abril 1979.
- [ROGERS,1983] ROGERS, Everett M.
Difussion of innovaciones.
The Free Press, tercera edición, 1983.
- [STRANGE,1995] STRANGE, Kevin H.
“The decision maker’s gold mine: The Data Warehouse”.
Datamation, Mayo 15, 1995
- [WATTERSON,1996] WATTERSON, Karen.
“Build your own Warehouse”.
SunExpert Magazine, Octubre 1996

CAPITULO 6

CONCLUSIONES Y TRABAJOS FUTUROS

Las organizaciones requieren actualmente tomar decisiones urgentemente y con bases bien establecidas, y es a través de la información como pueden generarse alternativas de decisión de manera eficiente, rápida y confiable para responder al dinamismo de los nuevos mercados. La evaluación y el estudio de las tecnologías de información debe convertirse en una actividad prioritaria dentro de las empresas, para conocer la funcionalidad de las innovaciones tecnológicas y los beneficios que pueden obtenerse al implantarse. Entre las tecnologías de información de punta se tiene al Data Warehouse, la cual representa una innovación tecnológica con una visualización diferente en la administración de las bases de datos y en la generación de información.

Se ha definido al Data Warehouse en esta tesis como una arquitectura de información con un enfoque sumariado e histórico en los datos, que apoya al proceso de toma de decisiones en los niveles táctico y estratégico de una organización. Esta definición se estableció en base a los siguientes puntos:

➤ Arquitectura de información

El diseño del Data Warehouse implica analizar el flujo de información de una empresa, automatizada o no, se generan asimismo niveles de abstracción de los datos para ser explotados como información de valor para la toma de decisiones, todo esto basado en un modelo de datos corporativo, en donde se definen los procesos del negocio y sus atributos y relaciones en lo que respecta a los datos. Además el Data Warehouse requiere el uso de sistemas de información y herramientas de software para el procesamiento de datos que permitan presentar la información de acuerdo a los tópicos del negocio de manera amigable y clara.

➤ Bases de datos sumarizadas e históricas

La administración de las bases de datos en el diseño del Data Warehouse está orientado a soportar el proceso de toma de decisiones en los niveles táctico y estratégico de una empresa, lo cual hace necesario que el diseño de las bases de datos sean guiadas por las reglas del negocio y los requerimientos de información en tales niveles organizacionales. Esta nueva propuesta en el diseño

de bases de datos requiere aplicar diversos procesos que permitan sumarizar los datos de diversas fuentes de la empresa para generar los niveles de abstracción que se hayan establecido durante el diseño de la arquitectura de información, y por otro lado, para poder generar los reportes corporativos que permitan conocer la situación de la empresa en un período de tiempo específico, se necesita que los datos sean administrados en base a ciertas dimensiones en el tiempo requeridos por los analistas de decisiones del negocio.

➤ Soporte al proceso de toma de decisiones

Para poder tomar una decisión de un tópico del negocio se necesita conocer los antecedentes del mismo y sus posibles alternativas de decisión, siendo la información sobre el tópico el medio más eficiente para lograrlo. El diseño del Data Warehouse debe orientarse para obtener información de cada tópico de la empresa, y en base a un modelo de datos establecido por los analistas de decisiones, poder apoyar mediante información en la generación y selección de las alternativas de decisión.

➤ Aplicación en los niveles táctico y estratégico de una organización

En los niveles estratégico y táctico de una empresa se requiere consultar información sumariada de diversos tópicos del negocio relacionados con los procesos de toma de decisiones más relevantes, a diferencia del nivel operativo que tiene que ver más con datos diarios del estado de la empresa y su entorno. Mientras mayor es el nivel organizacional, menor es el nivel de detalle de los datos y mayor el nivel de sumariación, lo cual debe reflejarse en el diseño de la arquitectura de información Data Warehouse.

Diseñar la arquitectura de información Data Warehouse requiere del conocimiento de la empresa en donde va a implantarse, principalmente en lo que respecta a sus procesos y su flujo de la información. Este conocimiento del negocio permitirá a los diseñadores del Data Warehouse diseñar esta arquitectura de información con una orientación a la organización y sus procesos, y si además este diseño forma parte de un plan informático, la alineación de la tecnología Data Warehouse con los objetivos del negocio será más estrecha y eficiente para lograr asimilar la tecnología exitosamente.

Con el estudio del proceso de asimilación del Data Warehouse en una empresa visto en esta tesis, se pretende lograr una efectiva transferencia de la nueva tecnología a utilizar y de la habilitación en aspectos de cambio en los procesos y reglas del negocio, así como que el recurso humano comprenda y se comprometa a

los cambios asociados a la implantación de la nueva tecnología y lograr así su implantación total en la organización. Durante cada fase del proceso de asimilación se proponen una serie de actividades que los diseñadores de un Data Warehouse pueden contemplar para crear un compromiso y sensibilización hacia la aceptación y convencimiento de las nuevas formas de trabajo que puedan generarse con el Data Warehouse, más allá incluso de la exitosa implantación de la propia tecnología.

El proceso de asimilación de la tecnología Data Warehouse que aquí se presenta se pretende que sea abierto a las necesidades de cada empresa, de tal manera que las actividades que se proponen en cada fase del proceso sirvan de base para determinar otras más que complementen el estudio de asimilación. Y para darle un carácter más global a este proceso de asimilación se tomaron atributos que influyen en este proceso y que son factibles a aplicarse en cualquier organización, los cuales pueden resumirse en los siguientes puntos:

- Evaluar de la tecnología de información Data Warehouse y sus diversos proveedores
- Realizar programas de capacitación para dotar al personal de los conocimientos y habilidades para operar la nueva tecnología.
- Llevar a cabo un plan de difusión del Data Warehouse entre los miembros de la organización para aumentar el conocimiento y entendimiento de la tecnología, sus beneficios, así como para establecer compromisos y responsabilidades.
- Analizar los procesos de cambio, realineamiento y cultura de información, identificando los agentes de cambio y los promotores de la tecnología para aumentar la sensibilización hacia ésta.
- Alinear las etapas de implantación del Data Warehouse con las fases del proceso de asimilación, para poder evaluar en forma paralela ambos procesos y determinar más claramente su grado de adopción entre los miembros de la organización.
- Enfatizar en la búsqueda de la alineación de las estrategias y necesidades del negocio con la implantación del Data Warehouse.
- Analizar la situación de los datos en la empresa.
- Poner énfasis en el análisis Costo-Beneficio de la nueva tecnología.

- Identificar barreras organizacionales y grupos de poder que obstaculicen el adecuado compartimiento de los datos.
- Analizar el grado de interconexión de la organización y sus diferentes flujos de información.
- Establecer el proceso de implantación del Data Warehouse conforme a un plan informático global en la empresa.
- Planear adecuadamente el proyecto piloto, poniendo especial énfasis en el aprendizaje de la nueva tecnología y en la aplicación de los procesos básicos del Data Warehouse.
- Difundir y comunicar los avances en la implantación de la innovación tecnológica, así como los resultados obtenidos en cada una de las etapas.

Los atributos mencionados están relacionados a la innovación tecnológica, a aspectos organizacionales, y a la interacción entre la tecnología y la organización. Estos atributos pueden formar parte de cualquier proceso de asimilación de tecnología, pero en esta tesis se consideraron para su análisis por el grado de influencia que pueden tener en el diseño e implantación de la arquitectura de información Data Warehouse.

TRABAJOS FUTUROS

Uno de los resultados esperados de esta tesis fue presentar al Data Warehouse como una arquitectura de información que ayude a mejorar la calidad y el manejo de los datos en una empresa. Del análisis del Data Warehouse como una arquitectura de información, presentado en el capítulo 3, pueden derivarse áreas de oportunidad para desarrollar e investigar nuevos temas relacionados al concepto Data Warehouse, entre las cuales se pueden mencionar:

- El desarrollo de Sistemas de Soporte a la Decisión empleando bases de datos generadas con los procesos de la arquitectura de información Data Warehouse. En los modelos estadísticos que se usan en un DSS poder analizar y establecer las relaciones con las estrategias de los procesos Data Warehouse almacenadas en el Metadata, por lo cual sería de interés establecer mecanismos de comunicación entre el Metadata y los modelos de apoyo a la toma de decisiones.

- El desarrollo de Sistemas de Información Ejecutivos empleando los procesos Data Warehouse como back-end en la generación de las bases de datos a nivel ejecutivo. Quizá el aspecto de mayor interés es el soporte que puede proporcionar la arquitectura de información Data Warehouse con los procesos drill-down de un EIS, es decir que a través de un mismo sistema de información los datos puedan visualizarse desde el más mínimo de detalle hasta la consulta de un registro de una base de datos transaccional.
- El diseño de la arquitectura de información Data Warehouse sobre una plataforma cliente-servidor, identificando en ella los componentes back-end, middle-ware y front-end, así como los procesos del Data Warehouse a nivel servidor y a nivel cliente.
- Analizar el diseño del Data Warehouse soportado por herramientas work-flow, orientado hacia la administración bases de datos documentales y en la integración de las herramientas de productividad para la automatización de oficinas.
- Investigar la relación entre el Data Warehouse y el Data Mining para poder equilibrar el análisis de las bases de datos históricas con la identificación de las tendencias y patrones de la información más importantes de una empresa.
- Proponer el diseño e implantación de un Data Warehouse empleando WEB servers y navegadores de internet, para crear una intranet apoyada por los procesos Data Warehouse.
- Estudiar al Data Warehouse haciendo énfasis en el análisis dimensional de la información, y en la administración de los datos orientados al negocio y a las necesidades de los ejecutivos y analistas de decisiones.

