

**UNIVERSIDAD NACIONAL AUTONOMA
DE MEXICO**

FACULTAD DE CIENCIAS

DATA WAREHOUSE

T E S I S

QUE PARA OBTENER EL TITULO DE

A C T U A R I O

P R E S E N T A :

ANGEL MONTES DE OCA BALTAZARES

DIRECTORA DE TESIS: DRA. AMPARO LOPEZ GAONA



**FACULTAD DE CIENCIAS
UNAM**

2002

**TESIS CON
FALLA DE ORIGEN**





Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MÉXICO

M. EN C. ELENA DE OTEYZA DE OTEYZA
Jefa de la División de Estudios Profesionales de la
Facultad de Ciencias
Presente

Comunicamos a usted que hemos revisado el trabajo escrito:

“DATA WAREHOUSE”

realizado por Angel Montes de Oca Baltazares

con número de cuenta 09051828-9 , quién cubrió los créditos de la carrera de Actuaría

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

Director de Tesis
Propietario

Dra. Amparo López Gaona

Propietario

M. en C. María Guadalupe Elena Ibarguengoitia González

Propietario

Dra. Hanna Oktaba

Suplente

Mat. Salvador López Mendoza

Suplente

M. en I. Maria De Luz Gasca Soto

Amparo
María Guadalupe Elena Ibarguengoitia González
Hanna Oktaba
Salvador López Mendoza
Maria De Luz Gasca Soto

Consejo Departamental de Matemáticas

José Antonio Flores Díaz
M. en C. José Antonio Flores Díaz

FACULTAD DE CIENCIAS
CONSEJO DEPARTAMENTAL
MATEMÁTICAS

A Socorro Baltazares... mi MamÁ...

Apoyo incondicional, guía, ejemplo y amiga

Sin duda alguna por ella he podido llegar a este momento

A Arturo Montes de Oca...

Que más que un hermano es mi mejor amigo ayer, hoy y siempre
Creo que parte del esfuerzo también es tuyo...

A Yolanda Montes de Oca...

En donde estás...

A Alejandra Montes de Oca...

A Lucio Montes de Oca López...

Apoyos y manos cuando los necesité...

A Antonio González (TioTo), José Baeza (Shagi) Cesar
Escobar (Chester) Gustavo Méndez (Gus)...

Compañeros, amigos, hermanos y ocasionalmente botanas

1 - 0

A Martha Gabriela (ejemplo y amiga)

A Circe Cristal (inspiración)... a mua?

A Africa (amor e ilusión)

A todas de las que me enamoré (que a cada paso hay una)

Y a todos los amigos que siempre están en mi, Martha Lucia,
Rosa, Raymundo Rafael, José Armando, Raymundo, Alberto,
Juan, Héctor, Alejandro, Fernando, Graciela, Belén, Beatriz
Elizabeth, Jodie, Laura, Ana, Aída, Wendy, Ventura, Pedro,
Mario, Jesús, Rita, Maria Eugenia, Minerva, Teresa, Cesar,
Manuel, Gabriela, Irene, Carl, Peter, Jaime, Leonardo, Luis,
Carlos, Rodrigo, Adriana, Andre, Denise Elaine, Leonardo,
Alberto, Elizabeth, Alina, Mario, Mónica, Concepción, Jaime,
Arelly, Claudia, Rubén, Salvadory todos aquellos que por olvido
omití...

A Amparo LÓpez

Asesora, quien me supo llevar a la conclusión de este ciclo

A Mary Glazman . . . †

Y a todos los profesores que algo de cada uno de ellos aprendí

Universidad Nacional Autónoma de México

DATA WAREHOUSE

Tesis que para obtener el Título de Actuario
Angel Montes de Oca Baltazares

Contenido

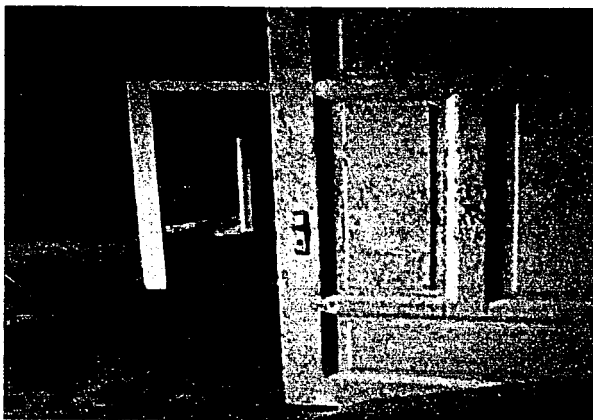
Capítulo	Título	Página
Cero	Introducción	1
Primero	Conceptos Generales	10
Segundo	Administración del Conocimiento (Knowledge Management)	34
Tercero	Medidores de Desempeño (Performance Measurement)	51
Cuarto	El Data Warehouse y la Cadena de Valor	69
Quinto	Equipo Data Warehouse	78
Sexto	El Data Warehouse y la Internet	95
Séptimo	Metodología	109
Octavo	Conclusiones y Recomendaciones	119
	Bibliografía	127

Anexos

- I. Distintas Arquitecturas de Data Warehouses
- II. Modelo Dimensional
- III. Un Tributo a Mary Glazman N. †



Capítulo Cero



Introducción



Introducción

La conclusión final es que sabemos muy poco y sin embargo es asombroso lo mucho que conocemos. Y más asombroso todavía que un conocimiento tan pequeño pueda dar tanto poder
"Anónimo"

En esto descubrieron treinta o cuarenta molinos de viento que hay en aquel campo; y así como Don Quijote los vió, dijo a su escudero:

- ❖ *La ventura va guiando nuestras cosas mejor de los que acertáramos a desear; por que ves allí, amigo treinta o poco más desafortunados gigantes con quien pienso hacer batalla y quitarles a todos las vidas, con cuyos despojos comenzaremos a enriquecer; que ésta es buena guerra, y es gran servicios de Dios quitar tan mala simiente de sobre la faz de la tierra.*
- ❖ *¿Qué gigante?, dijo Sancho Panza*
- ❖ *Aquellos que allí ves, respondió su amo, de los brazos largos que los suelen tener algunos de casi dos leguas.*
- ❖ *Mire vuestra merced, respondió Sancho, que aquellos que allí se parecen no son gigantes, sino molinos de viento, y lo que en ellos parecen brazos son las aspas, que volteadas al viento hacen andar la piedra del molino...*



nalizar las primeras líneas del capítulo donde Don Quijote vive la aventura contra los molinos de viento, nos permite ver que en el mundo real en donde vivimos, algunas veces debemos ser capaces de ver un poco más allá de los que nuestra mirada nos puede mostrar, simplemente "en lo imposible soñar" y lograr hacerlo posible. Todos hablan acerca del desempeño. Pero, ¿cuántas compañías pueden medir su propio desempeño actualmente?. En efecto, las ventajas competitivas dependen de la habilidad de mirar entre líneas para localizar eficiencias ocultas, reducir costos e incrementar la oportunidad de maximizar los ingresos. Y ese tipo de desempeño requiere información sólida,

oportuna, multidimensional¹ acerca de su negocio, sus productos, clientes y mercados. Transformar esa información en conocimiento es la clave para transformar a los negocios en una fuerza competitiva en estos días de un mercado dinámico y global, es ver a los simples molinos en gigantes y encontrar la forma en que se les puede vencer pero desde una perspectiva real.

Muchas de las compañías en México están acostumbradas a tomar decisiones de la manera tradicional, algunas veces ocupando a empleados en la creación de reportes complejos que no satisfacen las necesidades de información y cuando un reporte es entregado a la dirección, los datos con los cuales fue generado ya no reflejan la situación real de la compañía ni de su comportamiento. También, estos reportes muestran la sensibilidad aislada de quien genera los reportes y no reflejan en su totalidad la situación de la organización.

Actualmente el término de Data Warehouse en el ámbito industrial se ha vuelto un tabú, debido principalmente a que el oír de data warehouse los empresarios imaginan grandes inversiones de dinero, ocupación de recursos tanto internos como externos y que al final del proyecto escucharán que necesitarán de más tiempo y dinero para concluir el proyecto de data warehouse.

El presente trabajo tiene la finalidad de mostrar una forma diferente de analizar, primero si en la organización es necesario un data warehouse, o un data mart o simplemente de un operational data store y seleccionar el camino más adecuado a seguir en caso de optar por la implementación de algunos de los anteriores.

Por otro lado, la participación del Actuario en el desarrollo de estas tecnologías se convierte en necesaria, ya que muchas veces se requiere de un análisis con mayor

¹ Para entender del concepto mutidimensional, véase Anexo II Modelo Dimensional de este trabajo.

profundidad sobre diferentes factores de los negocios y que con los fundamentos de estadística y matemáticas aplicadas que tiene un Actuario, él se convierte en una pieza importante de este rompecabezas.

Del concepto original en donde una sola base de datos servía para todos los propósitos de la compañía se ha evolucionado a la arquitectura de datos donde los datos son divididos en un Data Warehouse y una base de datos transaccional. La evolución hacia arquitecturas con rasgos de un Data Warehouse es en respuesta de factores tecnológicos, económicos y organizacionales trabajando simbiótica, simultánea y espontáneamente. La diferencia entre los usuarios de ambientes transaccionales y de Data Warehouse está en el tipo de necesidades de información que requieren para desarrollar sus actividades primordiales dentro de una organización. Al final nos encontramos que los sistemas transaccionales son los que alimentan de información al data warehouse empresarial y a su vez, el sistema data warehouse retroalimenta al sistema transaccional.

Pero, una de las principales preguntas que salta al escuchar el concepto Data Warehouse es ¿Qué es un Data Warehouse?. El presente trabajo busca ir aclarando los puntos más importante de los sistemas data warehouse para al final responder esta y muchas de las preguntas que uno se encuentra al trabajar con este tipo de tecnologías. Obvia y lamentablemente no se pueden resolver en su totalidad las dudas, pero eso es una de las cosas bellas del data warehouse, responder muchas preguntas y dejar otras para analizar y encontrar respuestas que nos lleven sin duda a más preguntas.

Un Data Warehouse es un conjunto de datos integrados orientados a una materia, que varían con el tiempo y que no son transitorios, los cuales soportan el proceso de toma de decisiones de una empresa.

El Data Warehouse es el punto focal de la arquitectura para información de procesos para los sistemas modernos que permite:

- Anticipar en la toma de decisiones
- Mejorar las prácticas de negocios
- Brindar seguridad y confianza a quien la utiliza

Y este proceso tiene como resultado:

- La información deseada
- En el sitio indicado
- En la forma adecuada
- En el momento propicio

Implementar un Data Warehouse es una tarea compleja por la necesidad de interactuar con muchas tecnologías, distribuidores y usuarios finales para entregar resultados que sean eficaces en costo y tiempo.

Los desafíos a vencer en una estrategia de data warehouse son:

- Grandes islas de información
- Estructura de datos
- Calidad de datos / contenido / accesibilidad / fiabilidad
- Múltiples y diversas aplicaciones operativas
- Fuentes de información externa

Estructura del Presente Trabajo

La forma en la que se fue estructurando este trabajo principalmente fue con base en experiencia en el campo de data warehouse y tomando modelos tanto nacionales

como internacionales. La finalidad que tiene cada uno de los principios de capítulo es introducir al lector de una manera amigable al tipo de información que se encontrará en cada uno de ellos, quizá de una manera muy personal y con algunos de los íconos que han enmarcado mi vida.

Los capítulos que conforman este trabajo son nueve, los cuales tienen la finalidad de introducir al posible lector en el tema de data warehouse, a través los principales conceptos, manejados de una forma poco técnica ya que son los usuarios de áreas funcionales los primeros que deben buscar e impulsar este tipo de tecnologías. La descripción de los capítulos es la siguiente:

Introducción

Descripción del propósito del presente trabajo, justificando el porqué se desarrolló y la estructura que se seguirá en cada uno de los capítulos.

Conceptos Generales

Este capítulo se enfoca a los principales conceptos, términos y definiciones que se manejan en el área de data warehouse y busca responder básicamente las preguntas de ¿Qué es un data warehouse? ¿En dónde lo puedo utilizar? y ¿Qué tengo que hacer para desarrollar un data warehouse?, entre otras. También, explicar las diferencias entre los distintos tipos de sistemas tales como OLAP, OLTP, etc.

Administración del Conocimiento (Knowledge Management)

En la parte de Administración del Conocimiento, se muestra cómo el conocimiento se encuentra implícito en toda la organización y cómo pretende ayudar a las organizaciones a identificar tal conocimiento acumulado, tomarlo, depurarlo y transformarlo en datos que a su vez deberá arrojar como resultado un beneficio para

la empresa en materia de conocimiento aplicado a los procesos de toma de decisiones. También como parte importante de este capítulo, se busca que el conocimiento perdure en la organización a través de la transferencia de una persona o personas a otras.

Medidores de Desempeño (Performance Measurement)

Poco a poco el rompecabezas de data warehouse se va uniendo, con los dos capítulos que anteceden a este, se va estructurando la base para el desarrollo de estos sistemas, la importancia que tiene el identificarse cuáles son las medidas en la que se deben enfocar en una organización para saber cómo está trabajando la empresa. Los medidores de desempeño son como los signos vitales de una empresa, el poder identificar con mejor precisión estos signos ayudarán a una empresa a detectar a tiempo anomalías, y establecer nuevas estrategias para solucionar los problemas o identificar los mejores puntos de la empresa y enfocarse mejor a ellos.

El Data Warehouse y la Cadena de Valor

El capítulo del Data Warehouse y la Cadena de Valor pretende mostrar la forma en que se deben ver los procesos de las empresas y concientizar a los usuarios en que una empresa está conformada por engranes, y que al momento de que uno de ellos no funciona correctamente, provoca que el resto de la empresa fracase en su actividad primaria. También, busca que sí se elige un data warehouse, pero no se cuenta con el capital suficiente para implementarlo totalmente, ayudar a definir cuál o cuáles pueden ser las áreas susceptibles de implementar un data mart y construir con éstos un data warehouse por área.

Equipo Data Warehouse

En este capítulo, se muestran las principales características que debe tener un equipo de data warehouse, destacando las principales habilidades que debe mostrar cada uno de los miembros involucrados, este capítulo es sólo ilustrativo y puede tomarse de base para hacer la selección de candidatos. En algunos casos, en especial en México, más de un puesto lo puede tener una sola persona. Los roles y responsabilidades aquí expuestas son las que al menos deberán observarse en un equipo exitoso de data warehouse.

El Data Warehouse y la Internet

La importancia que ha cobrado la Internet hoy en día y la necesidad de dar respuestas rápidas a los mercados cambiantes actuales hace necesario contar con un data warehouse detrás de la Internet principalmente para administrar la información que es recopilada y responder con la mayor exactitud posible.

Además, como se mostró en el capítulo de la cadena de valor, la Internet viene a ser una parte importante en esta cadena, por que amplía la cadena, tanto para los proveedores como para los distribuidores.

Metodología

Considero que hay muchas metodologías en el mercado, esto es debido a que todas parten de una misma, el objetivo que busca presentar este capítulo es el de mostrar cuáles deben ser las etapas mínimas que deben evaluarse para implementar un data warehouse.

Conclusiones

La sección de conclusiones muestra mi opinión sobre los sistemas de data warehouse y su adopción como una herramienta importante para la toma de decisiones, adicionalmente presenta mi opinión de la situación de los sistemas de data warehouse en México y cómo éstos deben ser implementados, aunque a lo largo de este trabajo busque siempre mostrar la adecuación de este concepto a México.

Anexos

En la sección de anexos, se presentan aspectos importantes del tema de data warehouse, pero que por sus características no fue necesario crear un capítulo específico pero sí era importante mostrar el tema para una mejor claridad del concepto de data warehouse, además por que en estos puntos existe mucha bibliografía especializada en el tema.

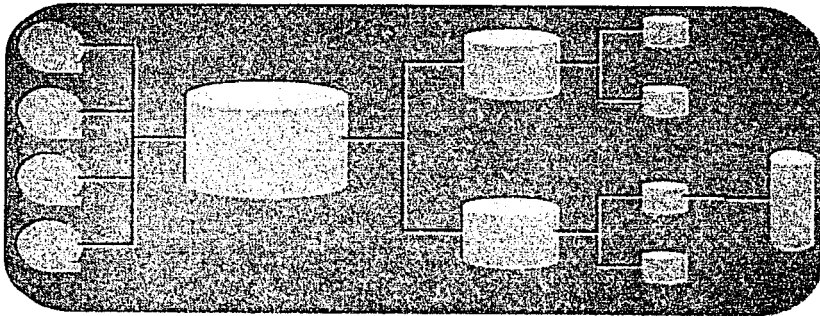
Por último, algunos de los conceptos se manejaron en idioma inglés dado que su traducción se aleja de su significado real o ésta es muy literal.

Así pues, iniciemos esta aventura hacia el Data Warehouse...

Señores, dijo Don Quijote, vámonos poco a poco, pues ya en los nidos de antaño no hay pájaros hogaño. Yo fui loco, y ya soy cuerdo; fui Don Quijote de la Mancha, y soy ahora, como he dicho Alonso Quijano el Bueno. Pueda con vuestras mercedes mi arrepentimiento y mi verdad volverme a la estimación que de mi se tenía, y prosiga adelante señor escribano...



Capítulo Primero



Conceptos Generales



Conceptos Generales

Probablemente se ha oído hablar mucho del término "Data Warehouse" o "Data Warehousing". Evidentemente un concepto muy de moda en el ambiente actual de negocios.

En términos generales y como marco de referencia, "Data Warehouse es una arquitectura de sistemas dedicada a **proporcionar información** que permite anticiparnos en la toma de decisiones, mejorar nuestras prácticas de negocio y brindar seguridad y confianza en quienes la utilizamos".

En uno de sus cuentos cortos, "La Biblioteca de Babel", el escritor Jorge Luis Borges describe una biblioteca infinita. Esto es, un interminable pasillo de habitaciones con estantes para libros, la mayor parte de los libros no tienen un significado y sus títulos son ininteligibles tal como MLO "AXAXXS". La gente se pasea en esta biblioteca hasta que muere, algunos de los eruditos desarrollan las hipótesis de que: "En algún lugar debe haber un catálogo general"; o bien que "todos los libros que uno puede imaginar debieran estar en algún lugar en la biblioteca". Ninguna de estas hipótesis se puede verificar, al final lo único que podemos concluir es que la biblioteca contiene una cantidad infinita de datos pero ninguna información.

Mucho se ha discutido sobre el tema de data warehouse, pero conceptualmente, el data warehouse representa un lugar transparente en donde el conocimiento y la información crítica de los negocios se encuentra almacenada. Los datos son extraídos de diferentes fuentes de información y explotados por los múltiples niveles de las organizaciones para lograr sus necesidades de información ya sea por área o de manera conjunta, estas fuentes pueden ser tanto internas, tales como los sistemas de operación de las compañías, o como externas como competencia, indicadores bursátiles, etc.



La meta del data warehouse es integrar lo mejor de los conceptos de los centros de servicios de información compartidos (Shared Service Data Centers)ⁱ, los sistemas de soporte de decisiones y los sistemas de información ejecutiva para crear una arquitectura de acceso de información integrada, a través de ofrecer la máxima flexibilidad para acceder a la información del negocio y utilizar las herramientas analíticas para procesar la información eficientemente. El enfoque que busca un data warehouse es proveer a las organizaciones más allá de un esfuerzo de crear reportes y de análisis como se ha venido haciendo con las arquitecturas de sistemas transaccionales actuales en donde el proceso de análisis y de creación de reportes se fundamenta en sucesos aislados y de datos generados por el sistema transaccional, que por su naturaleza es necesario esperar hasta el final del día para no afectar el desempeño del equipo central y con esto entorpecer la labor fundamental de las organizaciones.

El data warehouse permite a las organizaciones reducir significativamente tiempo y recursos en la localización, limpieza y acumulación de los datos que se requieren para la toma de decisiones del negocio. Actualmente, las compañías combinan las tecnologías de data warehouse con el negocio analítico clásico, principalmente con los informes de ventas, los muestreos estadísticos, y otros reportes. Los análisis cuantitativos para efficientar y automatizar la toma de decisiones de la misma forma que los paquetes o sistemas de Planificación de Recursos de la Empresa (Enterprise Resource Planning en adelante ERP)ⁱⁱ han transformado las capacidades de procesamiento de las transacciones de las compañías. Usando las técnicas del data warehouse y tecnologías modernas actuales, las compañías son capaces de identificar

ⁱ Un centro de servicios de información compartidos es la instancia en donde se concentran los principales recursos de una compañía con diferentes localidades o subcompañías como puede ser hardware, software e información la cual es accesada por las diferentes compañías de un corporativo, con esto se reducen los costos y se incrementan los niveles de servicio hacia los clientes y se logran mejores acuerdos con los distribuidores.

ⁱⁱ Las aplicaciones Enterprise Resource Planning son conocidas en el mercado por sus siglas en inglés ERP, de los cuales los más conocidos son SAP, PeopleSoft, BAAN, JDEdwards, etc.

las tendencias de ventas, los fracasos o éxitos de los productos, hábitos de consumo de los clientes y los modelos en la forma de vender más rápidamente que no era posible con las arquitecturas de acceso de información anteriores.

Los mercados de data warehouse experimentan un cambio acelerado, el data warehouse se está ligando a la Internet e intranets corporativas como una fuente y manejo de datos. Las diferentes soluciones de data warehouse se están integrando y vendiendo a los clientes con la finalidad de tratar de evitar que cada compañía desarrolle sus propios sistemas, los cuales serán personalizados para cada compañía y además inmaduros ya que no han sido del todo probados. Debido a que la funcionalidad del data warehouse se está integrando con soluciones de ERP's, estos sistemas deben contar con fuerte soporte y actualizaciones periódicas.

Fundamentos para el data warehouse

El concepto básico detrás del data warehouse es el de crear un repositorio transparente en el que se reúnan y organicen los datos críticos del negocio. El data warehouse es un repositorio de datos generado de varias y distintas fuentes, tanto internas como externas. Adecuadamente diseñado, implementado y mantenido, un data warehouse es una herramienta valiosa para la inteligencia de negociosⁱⁱⁱ.

Varios procesos son necesarios para el data warehousing. Estos procesos se clasifican generalmente como: generación del almacén (término acuñado por IDC que incluye el diseño del almacén y la carga de los datos), la administración de datos (almacenaje de datos), el acceso de información (recuperación de los datos) y el análisis de

ⁱⁱⁱ Las herramientas de Inteligencia de Negocios facilitan la toma de decisiones mediante reportes multidimensionales, donde los datos se acceden no como tablas en base de datos, sino en forma conceptual, como objetos de negocio.

información (uso de los datos para apoyar la toma de decisiones). La generación y administración de data warehouses utiliza herramientas para el diseño y el poblado o llenado de los data warehouses. La administración de datos puede utilizar diferentes bases de datos para almacenar la información. Las herramientas de acceso ayudan a los usuarios a que puedan acceder y analizar la información que se encuentra contenida en el data warehouse.

El data warehouse puede ser usado exitosamente por las compañías de diferentes maneras, por ejemplo, para:

- ☞ Proporcionar a los usuarios de negocios una visión enfocada al cliente utilizando datos heterogéneos de la compañía, integrando los datos de ventas, de servicio, de distribución y manufactura y otros sistemas relacionados con el negocio del cliente para modelar la segmentación, planificar la retención de cliente y analizar las ventas cruzadas^{iv}.
- ☞ Eliminar las barreras y funciones entre los distintos departamentos del negocio ofreciendo un camino hacia los datos de forma consistente de las múltiples "islas de información", proporcionando una mirada integral de las actividades del negocio, las funciones y líneas de departamentos que se cruzan.
- ☞ Proporcionar, a nivel macro, un panorama de información de los factores críticos de éxito de la compañía, tales como los impulsores de valor de los accionistas y de los medidores de desempeño de las organizaciones.
- ☞ Aumentar la información financiera tradicional proporcionando un acceso más oportuno y detallado a la información.

^{iv} Un ejemplo de ventas cruzadas es la que se detectó en almacenes donde la venta de pañales tenía cierta relación con la venta de cerveza, ésto obligó a la compañía a redistribuir la ubicación de sus productos.

- ☞ Cuando los data warehouses se empalman con la tecnología de la Internet, los clientes pueden tener acceso a la información específica de los negocios. De cualquier otra manera sería costoso o consumiría mucho tiempo lograr este intercambio de información.
- ☞ Identificar las tendencias globales usando la información para manejar a las unidades en multidivisionales, empresas multinacionales, esto puede incluir tendencias en áreas tales como las de comercialización, planificación de la producción, adquisición y retención de clientes, etc.

Vista general del data warehouse

Por muchos años, las compañías han estado trabajando en mejorar el acceso y análisis de los datos para la toma de decisiones. Sin embargo, estos esfuerzos han pasado a un segundo plano al trabajar en incrementar sus sistemas de procesamiento de transacciones para que les sirvan como punto de captura de datos de la compañía. Actualmente, las compañías redireccionan su atención a la información disponible para los usuarios y sus clientes a través de eficientar y mejorar la administración, organización y acceso a los datos. El enfoque y la tecnología han cambiado de procesamiento de las transacciones del negocio día a día hacia el acceso de la información.

Por su analogía el almacén físico que tiene los bienes^v, el término data warehouse refleja la visión de que los datos son una ventaja competitiva de la compañía. Como otros activos, los datos deben ser identificados, catalogados y almacenados para que

^v Los datos son considerados como uno de los bienes o como los activos más importantes de las organizaciones

los usuarios sean capaces de buscar información y hacer hallazgos interesantes sobre el comportamiento del negocio.

Un factor clave para diseñar un data warehouse es utilizar, en un alto nivel, el modelo de datos de toda la empresa. Un modelo de datos de toda la empresa es una recopilación y selección de los datos usados por todos y cada uno de los departamentos en la empresa, presentados de una forma sólida y consistente. Un modelo de datos de toda la empresa sirve como base para un data warehouse para toda la compañía.

Procesamiento Analítico en Línea (OLAP) vs. Procesamiento de Transacciones en Línea (OLTP)

Una forma de entender qué es el data warehouse es contrastando los componentes del data warehouse. La administración de datos del procesamiento analítico en línea (OLAP), contra las bases de datos operacionales que tienen los datos usados por las aplicaciones de procesamiento de transacción en línea (OLTP). (Véase Tabla 1.1)

Características	OLAP	OLTP
Datos		
Organización	Orientado a área	Orientado a procesos
Alcance	Sistema amplio	Función individual
Horizonte de Tiempo	Datos casi actuales e históricos	Únicamente datos actuales
Holgura	Amplia	Estrecha
Tipo de Datos	Datos resumizados y derivados	Datos operacionales detallados
Precisión	Alta	Alta
Volatilidad	Usualmente no se actualizan directamente	Diseñado para actualizarse directamente
Fuente	Internos y externos	Internos
Modelo de datos	Dimensional	Normalizado

Usuarios		
Número	Algunos usuarios OLTP concurrentes	Muchos usuarios concurrentes
Tipo	Administradores y analistas	Personal operacional y administrativo
Uso		
Proceso de volumen	Bajo volumen de grandes consultas	Alto volumen de grandes transacciones
Acceso típico	Consultas ad hoc con algunos reportes periódicos	Transacciones repetitivas estándar
Impulsores de uso	Análisis de datos	Transacciones
Operación típica	Reportes y análisis	Actualización
Tiempo de respuesta	Desempeño es menos crítico	Desempeño muy sensitivo
Enfocado		
Enfocado a la operación	Flexibilidad	Desempeño y rentabilidad
Enfocado al negocio	Táctico y estratégico	Operacional y táctico; soporta día a día la operación
Enfocado a la administración	Eficacia	Eficiencia

Tabla 1.1: Contraste entre OLAP y OLTP

Los sistemas OLTP soportan las operaciones de negocio día a día capturando los datos acerca de los procesos y acontecimientos del negocio. Ellos son transaccionales por naturaleza: muchos usuarios pueden acceder concurrentemente a una base de datos para agregar, recuperar, actualizar y borrar los registros individuales. Los datos que están siendo usados por un sistema OLTP son almacenados en la base de datos operacional diseñada para los requisitos de las transacciones.

Los sistemas OLAP son de apoyo para la toma de decisiones y es la contraparte de los sistemas OLTP. Los sistemas OLAP permiten que los usuarios analicen los datos que sostienen la toma de decisiones del negocio. Los sistemas OLAP requieren acceso a los datos de una forma ad hoc y también requiere de acceso a datos históricos que representan el comportamiento de la compañía en un período de tiempo determinado. Los datos usados por estos sistemas OLAP se almacenan dentro del data warehouse, a menudo, en una base de datos multidimensional. Estos

sistemas surgen a causa de la incapacidad de los sistemas OLTP para dar respuesta a las consultas ad hoc y al soporte analítico.

Los datos de los sistemas OLAP a menudo se sintetizan fuertemente, no son cambiados o actualizados continuamente y son accedidos sólo como lectura. En un sistema OLAP, una consulta ejecutada varias veces devolverá los mismos resultados hasta que los datos en el data warehouse se hayan actualizado. En un sistema OLTP, los datos están en su mayor nivel de detalle. Una consulta ejecutada varias veces dará resultados diferentes cada vez porque los datos fuente pueden cambiar de minuto a minuto.

En términos simples, los sistemas OLAP ayudan a una compañía a decidir lo que se hace y los OLTP ayudan a una compañía a hacer lo que hace. Los OLTP se enfocan a la administración, mientras que los OLAP se enfocan en las actividades que agregan valor al negocio (decisiones rápidas, mejoras de servicio al cliente e incrementos en la utilidad). Las bases de datos de los sistemas OLTP almacenan los datos necesarios para manejar el negocio en una base diaria. Estos datos también son usados como fuentes de datos para bases de datos de los sistemas OLAP, que consisten de agregaciones, transformaciones, integraciones y colecciones históricas de los datos que originaron en uno o más de los sistemas de OLTP.

Para comprender un poco más sobre las diferencias entre los sistemas OLTP y OLAP, se presenta el siguiente diagrama, en donde se refleja una arquitectura que pretende mostrar las fronteras de ambos sistemas. Véase Figura 1.1



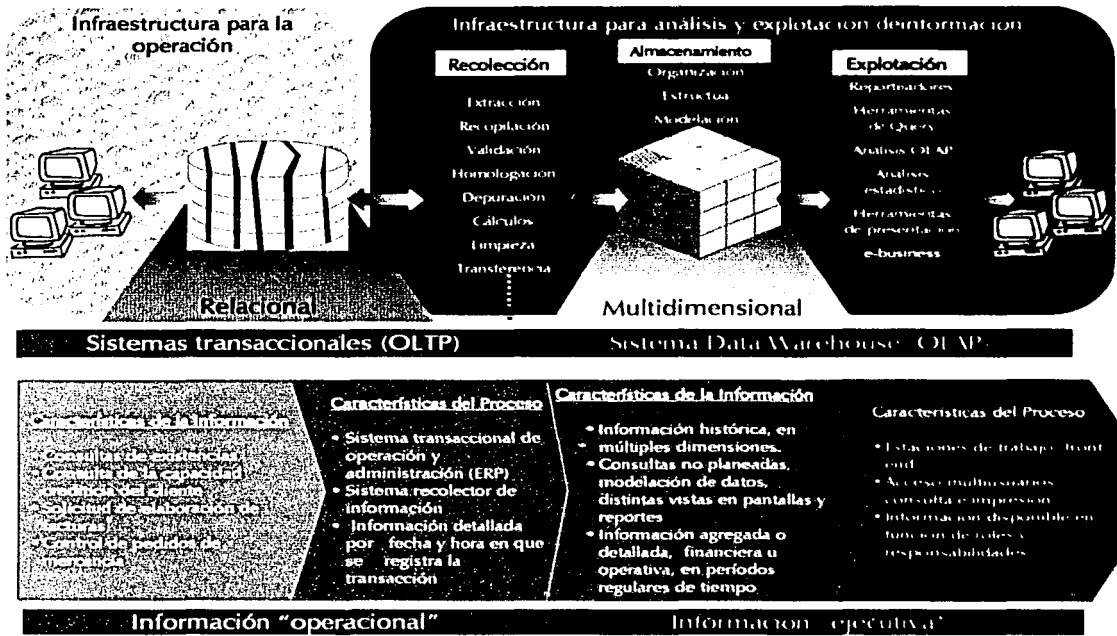


Figura 1.1: Entendimiento de los sistemas OLAP y OLTP

Enfoque del data warehouse

El enfoque del data warehouse va de un escenario sencillo a uno complejo. No hay en el mercado una solución de data warehouse que pueda decir que cuenta con una suite completa de soluciones que un data warehouse requiere. Para tener éxito, la estrategia de data warehouse de una compañía debe alinearse directamente con su estrategia de negocio. Una vez que la estrategia de negocio de data warehouse ha sido acordada, una compañía selecciona un plan estratégico que usa uno o más de los enfoques básicos que se adecue mejor con sus aplicaciones actuales, las arquitecturas de datos y de tecnología. Por ejemplo si la estrategia de una compañía es reducir el costo de fabricación en una base mundial, entonces la decisión de la

compañía estaría enfoca a data warehouses o data marts^{vi} que faciliten o estandaricen los datos con sus distribuidores.

Existen dos enfoques para contar con sistemas de apoyo para la toma de decisiones: los data marts y el data warehouse empresarial. Estos enfoques difieren en la escala y la complejidad. Además de representar estrategias diferentes de data warehouses, estos dos enfoques pueden representar las etapas en la evolución de un data warehouse dentro de una organización. Un tercer enfoque, el Operational Data Store (ODS), también se incorpora en la solución completa del data warehousing.

Data Marts

Un data mart se dirige más hacia cierta área o departamento del negocio. El data mart usualmente está diseñado para soportar los requerimientos de información únicos por departamento o por procesos de negocio. Una compañía puede tener muchos data marts, cada uno dirigido a un subconjunto de la compañía. El incremento en el uso del data mart a menudo es impulsado por usuarios finales y departamentos que quieren construir un sistema enfocado a sus necesidades específicas en vez de esperar a la construcción de un data warehouse para toda la compañía.

Existen dos tipos de Data Marts: dependientes e independientes:

Data Mart – dependiente: Contiene un subconjunto de los datos del data warehouse, los cuales actúan como fuente. Los datos son alimentados del data warehouse central (the hub) al data mart (the spokes), con lo que los data marts están siendo usados para ingresar a los datos – esto se refiere a veces al modelo “data retailing” o

^{vi} En las siguientes líneas se explica a detalle los data marts

“hub and spoke”. El data mart dependiente es básicamente un mecanismo de distribución para datos que se integra ajustadamente al data warehouse de la empresa.

Data Mart – independiente: Deriva sus datos directamente desde los sistemas OLTP operando autónomamente. Un data mart independiente se enfoca en un sujeto o problema y típicamente no es fácil integrarlo con un data warehouse central, particularmente si una empresa tiene múltiples data mart independientes.

El movimiento hacia los data marts es acelerado por la facilidad de reducir los costos en la adquisición de hardware y software.

Derivado de que reduce su alcance, los data marts toman menos tiempo para construirlos, menos costo y son mucho menos complejos que un data warehouse empresarial, es más fácil convenir en la definición de datos para una sola área que para una empresa entera. Por lo tanto, un data mart es recomendable cuando una compañía necesita mejorar el acceso de datos en un área objetivo, tal como el departamento de ventas o el de finanzas, principalmente por que es dónde se pueden concentrar las urgencias de necesidad de información. Los data marts pueden tener también una prueba de eficacia de ciertos productos, proyectos de tecnologías o de negocio. Sin embargo, la introducción indiscriminada de múltiples data marts sin estar ligados uno al otro o un data warehouse empresarial mal planeado, causará grandes problemas a largo plazo.

Debido a que los data marts proliferan rápidamente, pueden subir a una misma “isla de información” problemas que llenen más rápidamente las iniciativas de sistema de información. Data marts aislados son indeseables, en su lugar, deben desarrollarse y desplegarse dentro de la estrategia del data warehouse empresarial para que

eventualmente puedan evolucionar en un ambiente integrado de soporte de decisión.

Data Warehouses

Los data warehouses empresariales proporcionan una visión sólida y consistente de la compañía con terminología estándar de usuarios de negocios y datos que se emplean en toda la organización. La implementación de data warehouses es más compleja que la implementación de data marts porque los data warehouses empresariales son generalmente dirigidos para toda la compañía (en comparación con el enfoque de los data marts) y, por lo tanto, requieren que la empresa establezca una visión centralizada y estructurada de todos sus datos.

En la práctica, pocas organizaciones comienzan implementando un data warehouse empresarial, principalmente porque el costo y el tiempo son bastantes altos y pocas compañías pueden hacer una inversión de este tamaño. En su lugar, las empresas a menudo terminan con una colección de data marts que interactúan entre sí y que además son usados como base para iniciar la construcción de data warehouse empresarial o, también, utilizados como un data warehouse virtual y con el tiempo se convertirá en un data warehouse físico. Las organizaciones pueden encontrar que un solo data warehouse es posible pero no deseable por muchas razones, incluyendo aspectos como geografía, acceso a los datos, diferimiento de los requisitos departamentales y el alcance de las actividades de la organización. Véase la Tabla 1.2, en donde se muestran las principales diferencias y características entre los data marts y los data warehouses.



Características Típicas	Data Marts	Data Warehouses
Esfuerzo		
Alcance	Una sola área	Muchas áreas
Tiempo para construir	Meses	Años
Costo	Miles	Millones
Complejidad	Baja a media	Alta
Datos		
Requerimientos para compartir	Compartido con áreas del negocio	Compartido a toda la compañía
Fuentes	Algunos sistemas operacionales y externos, un data warehouse	Múltiples sistemas operacionales y externos
Tamaño	Megabytes a GigaBytes	Gigabytes a Terabytes
Horizonte de tiempo	Datos actuales e históricos	Datos históricos
Frecuencia de actualización	Optimizado por frecuentes actualizaciones (diario o semanal)	Optimizado por menos actualizaciones frecuentes (semanal o mensual)
Número de instancias	Múltiples data marts distribuidos a través de la compañía	Un solo data warehouse
Tecnología		
Hardware	Abierto	Abierto
Sistemas Operativos	Windows NT, Unix, OS/400	Unix, MVS, OS/2
Bases de Datos	Bases de datos de grupos de trabajo	Grandes bases de datos
Usos		
Número de usuarios concurrentes	Cientos	Cientos de miles
Tipo de usuarios	Analistas de áreas de negocios, Gerentes de área	Analistas corporativos, ejecutivos, directores, Gerentes de procesos
Enfocado al negocio	Optimizando las actividades con las áreas de negocio	Optimización multifuncional y toma de decisiones

Tabla 1.2. Data Marts vs. Data Warehouses

Tienda de Datos Operacionales (Operational Data Store)

Tienda de datos operacionales (Operational Data Store ODS) es un concepto arquitectónico en sistemas usado para proporcionar un adecuado soporte de decisiones y creación de reportes operacionales. Generalmente, un ODS está orientado al sujeto, integrando datos transaccionales de uno o más sistemas

operacionales que han sido transformados para contar con información certera y consistente. Los datos contenidos en un ODS pueden ser datos operacionales detallados y sintetizados^{vii}. Como todos los datos, éstos se modifican en el sistema operacional y una copia de los cambios se queda almacenada en el ODS. Así, el mantenimiento de datos en el ODS está sujeto a frecuentes cambios, tal y como ocurrirían los cambios en el sistema transaccional, a diferencia del data warehouse, que se actualiza de manera periódica y planificada.

El ODS combina algunas características de los sistemas de data warehouse con los de los sistemas OLTP. Esta combinación implica datos de una o más bases de datos operacionales todos éstos combinados en un repositorio al cual los usuarios pueden acceder información granularizada. Así, el ODS es semejante en muchas formas a un sistema de OLAP; sin embargo, un ODS comparte también varias características con los sistemas de OLTP, tales como:

- Contiene datos actuales, detallados y operacionales (típicamente 30 ó 60 días de datos).
- Puede ser ingresado por muchos usuarios concurrentes.
- Es usado para propósitos operacionales.

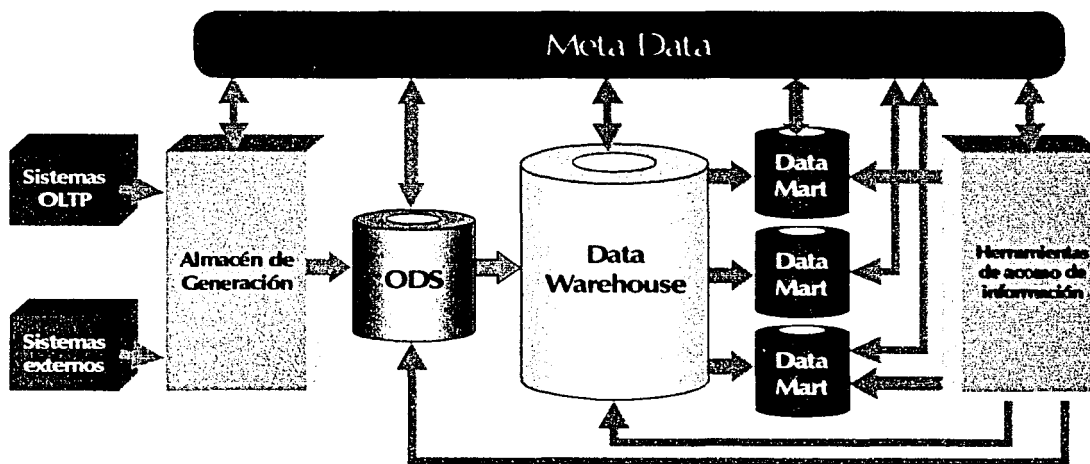
El ODS es particularmente popular como un medio para integrar varios sistemas OLTP, así como también como un soporte operacional secundario y de reporte que ayuda a disminuir el impacto negativo en el desempeño de los sistemas OLTP.

^{vii} El concepto datos sumarizados es más utilizado en este tipo de prácticas, aunque propiamente se debería manejar el concepto de datos resumidos, para los fines de este trabajo se utilizarán como datos sumarizados

Como un componente del proceso de data warehouse, el ODS se integra en el data warehouse como un repositorio para generar un nivel de detalle más atómico para los procesos de extracción. Como tal, las compañías encuentran que el ODS proporciona la fuente de datos granularizados para soportar las aplicaciones de la minería de datos (técnicas de Data Mining)^{viii}. Un ODS se puede utilizar también como un lugar físico en donde los datos se conjuntan y consolidan para ser cargados en el data warehouse y, el ODS puede servir como la fuente de datos para el data warehouse el cual se regenera con el tiempo.

Una solución completa de data warehouse puede incluir un ODS, un data warehouse y data marts. Véase Figura 1.2.

Figura 1.2 Tecnología del data warehouse



^{viii} Las técnicas de data mining se han destacado por ser la herramienta que se utiliza para encontrar relaciones no obvias entre los datos. A las herramientas de data mining se les conoce mejor como los agentes judiciales de los datos que extraen la información más increíble que uno pueda imaginar, en el transcurso de este trabajo se hará mención a estas técnicas, pero debido a su alto nivel de estudio, no se profundizará mucho más en este tema.

Las tecnologías de data warehouse

Los datos de los sistemas operacionales u otros sistemas externos se extraen periódicamente y son transformados (limpiados y resumizados) aunque algunas veces las herramientas de generación del data warehouse son usadas para cargar los datos directamente al data warehouse. Los datos del data warehouse son manejados por herramientas de administración de datos y son recuperados por otras herramientas de acceso de información. Las herramientas de administración del data warehouse ayudan a manipular y administrar todas las operaciones del data warehouse a través de su ciclo operacional de vida. Véase Figura 1.3.

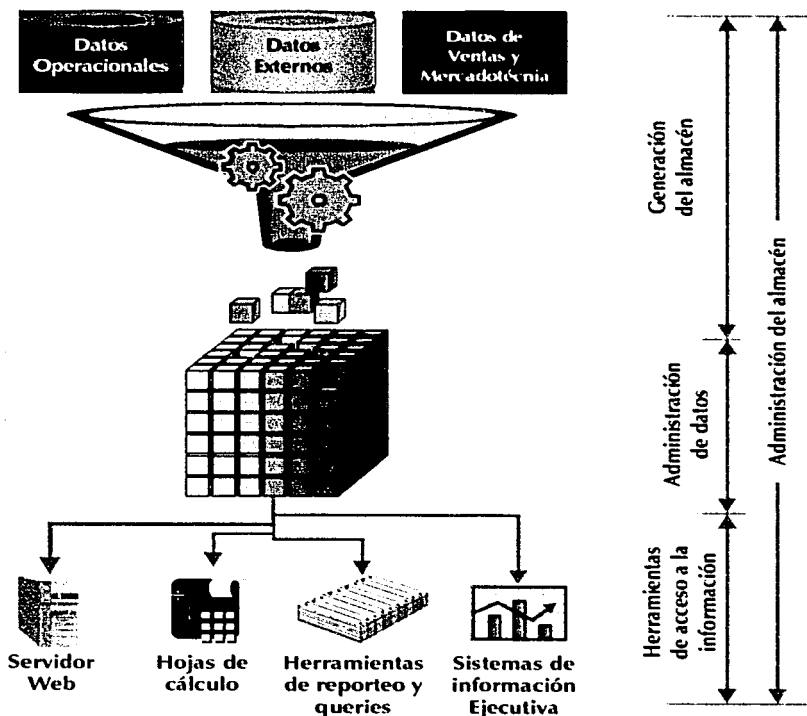


Figura 1.3. Creación el data warehouse



La Generación de data warehouse

La generación de un data warehouse es bastante compleja debido al tiempo utilizado para extraer los datos de los sistemas fuente, transformar los datos (limpieza), mover los datos hacia el servidor en donde se encuentra el data warehouse y cargar los datos en el data warehouse. Generalmente, se estima que de 60 a 80 por ciento del esfuerzo para la construcción de un data warehouse son utilizados en el proceso de generación de almacén.

Preparando datos para el data warehouse

Los sistemas de extracción son ejecutados periódicamente para extraer y recolectar los datos relevantes de los sistemas operacionales. Estos programas especifican ubicación de la información, esto es, información de en donde se encuentran los datos (metadatos)^{ix}, y conocer si es parte de los sistemas operacionales internos o si es de sistemas externos y también controla la información sobre los criterios para la extracción de esas fuentes de información. Algunos programas a menudo realizan la extracción de datos, pero ahora están disponibles para participar en este proceso herramientas de shrink-wrappen (herramientas dinámicas de compactación y descompactación de grandes volúmenes de información).

Los extractores de información de las bases de datos pueden acceder a diferentes formatos de bases de datos y archivos planos, a menudo utilizan los datos que residen en los sistemas operacionales. Las rutinas de la extracción de datos

^{ix} Los meta datos se distinguen por ser los datos acerca de los datos, esto es, imaginemos una tienda de abarrotes, en donde cada uno de los artículos se encuentra localizado con otros productos de características similares, por ejemplo lácteos, ya sea leche, cremas, quesos, etc. de alguna manera, los datos están ubicados por ciertas estructuras de orden en el data warehouse tal y como se encontrarían en la tienda.

interpretan los datos formateados en los sistemas operacionales, identifican los registros alterados y los copian a un archivo intermedio.

Después de que los datos han sido transferidos al archivo intermedio (que puede ser un ODS), los programas o utilerías de transformación son ejecutados para preparar los datos para el data warehouse. El proceso de la transformación incluye los siguientes pasos:

- ⇒ Consolidar los datos de múltiples fuentes de información
- ⇒ Filtrar los datos para eliminar detalles o campos innecesarios
- ⇒ Limpiar los datos para reparar o eliminar datos duplicados o incorrectos
- ⇒ Convertir y traducir los datos operacionales extraídos al formato del data warehouse
- ⇒ Agregar datos
- ⇒ Cargar datos
- ⇒ Asegurar la calidad

Consolidar los datos. El proceso de consolidación de datos implica la integración de datos de múltiples fuentes de información para crear una vista consolidada de los datos. Los registros de varias fuentes de datos se combinan en un archivo, que quizás sea cargado al data warehouse.

g

Filtrar los datos. Filtrar los datos en los archivos consolidados implica seleccionar sólo los datos que el data warehouse necesita e identificar algunos datos innecesarios. Los registros innecesarios son removidos del archivo consolidado. Los registros restantes son filtrados para eliminar atributos innecesarios o duplicados.

Limpiar los datos. Limpiar los datos es un proceso particularmente importante y difícil. Los datos tal como la información del cliente para enviarle correspondencia debe ser exacta, pero a menudo es difícil reconocer que varias entradas representan la misma entidad existente en la fuente de datos (tal como Peter Parker – 123 Park Avenue, Pet Parker – 123 Park Ave., Mr. Parker – 123 Park Avenue, etc). Los datos sin actualizar, redundantes y pobres en calidad deben ser identificados y corregidos antes de introducirse en el data warehouse.

Muchas organizaciones dirigen internamente sus proyectos de implementación de data warehouses pero éstos fallan por no dar la suficiente importancia a la limpieza de los datos. Para poder tener éxito en la implementación de un data warehouse, la limpieza de datos es una de las actividades iterativas fundamentales en el ciclo de vida del data warehouse.

Convertir y traducir los datos. Convertir y traducir los datos de la fuente u origen implica el mapeo de la fuente de datos a la estructura de datos objetivo y convertir los datos de la fuente al nuevo formato de datos. El proceso de la conversión implica reglas que traducen los valores de los datos usados en las aplicaciones fuente u origen en los valores usados en el ambiente de apoyo para la toma de decisiones. Los motores de conversión de datos aceptan típicamente los datos SQL o archivos secuenciales como entrada, estos datos de salida en el formato apropiado del archivo para cargarse en la base objetivo. El proceso de conversión y traducción cubre los siguientes puntos:

- ↳ Define la asociación lógica de campos de datos de los sistemas fuente a campos de datos en la base de datos objetivo (por ejemplo, el campo CUST_N con el formato FIRST-NAME MY LASTNAME en los datos de la fuente es mapeado al

campo del CUSTOMER con el formato LASTNAME FIRSTNAME MY en la base de datos de objetivo).

- ↻ Traduce columnas, campos y los nombres de las tablas en etiquetas más significativas usadas en el data warehouse.
- ↻ Traduce códigos de datos (como 1 por Male) a códigos o descripciones más significativos (como M para Male).
- ↻ Los datos normalizados marcan esta consistencia (por ejemplo, cambiando toda información del valor del salario de semanal o mensual o anual).
- ↻ Cambia la disposición de los datos (tipo de datos, longitud, formato, etcétera) del sistema de fuentes al formato de la base de datos destino.

Agregar Datos. Agregar los datos implica clasificar y sumarizar los datos, además de crear llaves generalizadas para los agregados, así, ellos pueden ser accedados fácilmente. Los datos pueden ser agregados por cualquier dimensión, independientemente de cualquier otra dimensión. Por ejemplo, si los datos atómicos existen por producto, por tienda, por periodos de tiempo, los usuarios pueden agregar productos por categoría de productos, por región o por mes. Es posible crear más niveles de agregación que resulten ser prácticos. El desafío es determinar en que niveles es útil contar con estos niveles de agregación y que a su vez sean útiles para responder a las preguntas y necesidades reales y objetivas de los usuarios.

Cargar datos en el data warehouse. Después que los datos se han transformado en el formato apropiado, éstos se cargan en el data warehouse, que incluye arreglos para el acceso de los datos. El proceso de carga puede ser una carga en bulto, una carga gota

a gota (tal y como ocurre en los sistemas operacionales), o una carga de incremento periódico (actualiza la base de datos con nuevos datos operacionales en una base regular, en promedio, el data warehouse, 50% de los datos son actualizados mensualmente, 25% son actualizados semanalmente y el otro 25% es actualizado diariamente). Los errores y las excepciones del proceso de carga de información deben identificarse y corregirse.

Asegurar la calidad. Los errores de programación y operacionales pueden introducir datos con ciertos problemas durante cualquiera de los varios procesos de extracción de datos de los sistemas fuente para producir los resultados finales utilizando las herramientas de acceso de información. Por lo tanto, un cuadro para identificar, corregir y conciliar errores deberá estar en un lugar y operar cuando el data warehouse se esté creando. La validación y prueba de datos ayuda al monitoreo y resolución de los problemas de calidad de datos identificando y corrigiendo los datos inconsistentes.

Cuándo los nuevos datos se añaden al warehouse, los datos cargados deben ser revisados para asegurarse que la carga tiene éxito y que la calidad de los datos reúne los requisitos mínimos de calidad de los datos.

Esquema de Modelado

Para construir los data warehouse, todas las metodologías en el mercado, utilizan el modelado de datos propuesto por Zachman, el cual lo nombró como el diseño o modelo estrella. Como su nombre lo indica, es un esquema de modelado de información para el data warehouse que se distingue por tener al centro un solo objeto conectado con varios objetos al su alrededor, simulando una estrella básica de cuatro picos, aunque los picos pueden incrementarse. La imagen que más

frecuentemente muestra este singular diseño es la que se muestra a continuación. Véase Figura 1.4:

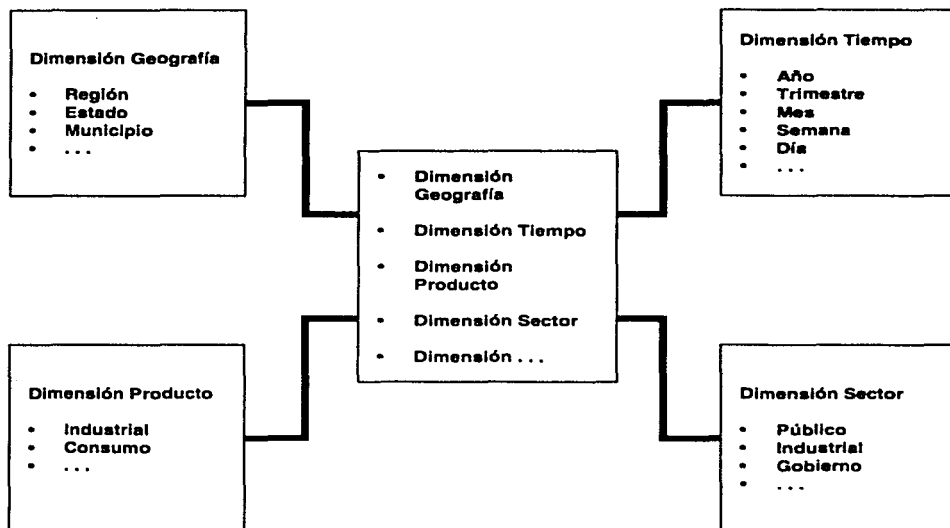


Figura 1.4. Ejemplo de modelo estrella

De las propiedades importantes en el diseño de los modelos estrella es que el objeto central tiene las llaves primarias de los objetos que circundan al objeto central. Un esquema de estrella lógico sencillo consta de una tabla de hechos y varias tablas de dimensión, los esquemas estrella complejos tienen cientos de tablas de hechos y de dimensión. Una tabla de hechos contiene las mediciones básicas de los negocios y consta de millones de hileras. Las tablas de dimensión contienen atributos de negocios que se emplean como criterios de búsqueda. Para definir estos atributos de negocio se emplean técnicas que más adelante se mencionan y que se llaman Balance Scorecards (tableros balanceados).

A veces los modelos estrella dan origen a otros modelos llamados Modelos Copo de Nieve (Snow Flakes), que se definen como una extensión del modelo estrella en donde cada uno de los puntos de la estrella se dividen en más puntos como se puede observarse a continuación. Véase Figura 1.5:

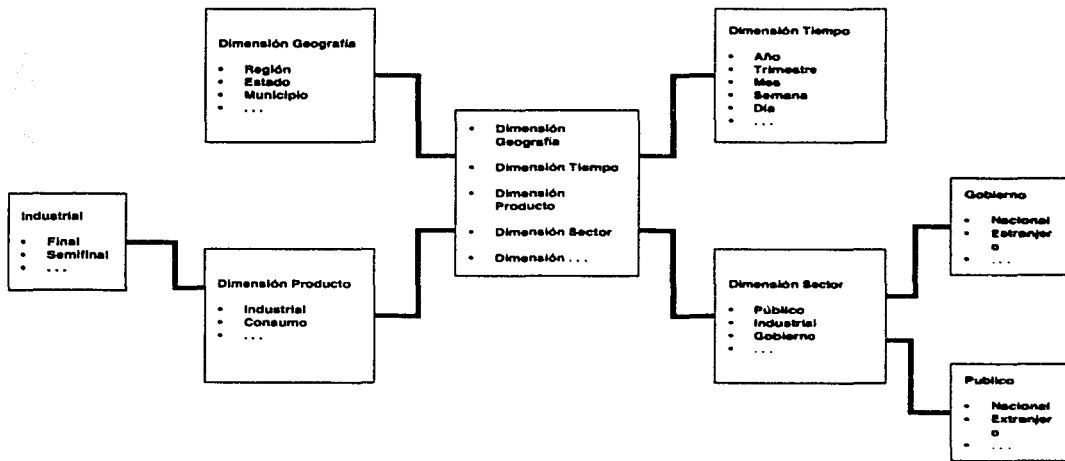
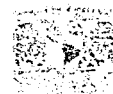
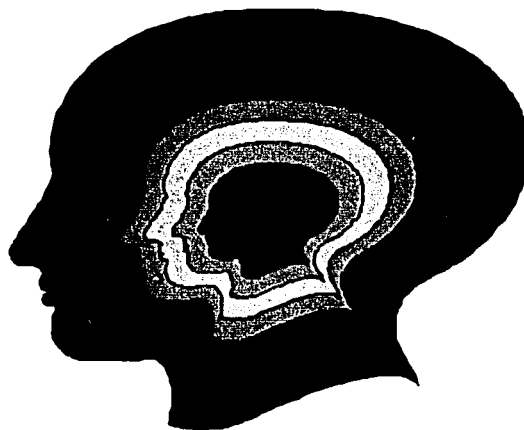


Figura 1.5 Ejemplo de modelo de Copo de Nieve

En esta forma de esquema, las tablas de dimensión del esquema estrella contienen más normas. La ventaja que proporciona el modelo de copo de nieve es que mejora el desempeño de consulta derivado de una utilización mínima en disco duro del equipo y mediante la unión de tablas más pequeñas con ciertas normas establecidas. Asimismo, incrementa la flexibilidad de las aplicaciones debido a la utilización de normas y con esto se reduce la granularidad de las dimensiones.



Capítulo Segundo



Administración del Conocimiento



Administración del Conocimiento

La mente del hombre es capaz de todo, porque todo está contenido en ella, tanto el pasado como el futuro ...
Joseph Conrad

El prefacio del libro Murmullos de la Tierra de Carl Sagan, dice: "El 20 de agosto y el 5 de septiembre de 1997 fueron lanzadas a las estrellas dos extraordinarias naves espaciales. Estos vehículos del espacio, después de haber llevado a cabo una exploración que promete ser detallada y realmente espectacular del sistema solar exterior desde Júpiter hasta Urano entre 1979 y 1986, abandonarán lentamente los sistemas solares convirtiéndose en emisarios de la Tierra al reino de las estrellas. Cada nave Voyager lleva adosado un disco fonográfico de cobre recubierto de oro como mensaje para las posibles civilizaciones extraterrestres que la nave pudiera encontrar en algún lugar y tiempo remotos. Cada disco contiene 118 fotografías de nuestro planeta, de nosotros mismos y nuestras civilizaciones; casi 90 minutos de la mejor música del mundo; un ensayo evolucionario en audio sobre <los sonidos de la Tierra> y saludos en casi sesenta idiomas humanos (y el lenguaje de ballenas) . . ."

Parece que cada vez que surge un nuevo término, éste viene acompañado de múltiples interpretaciones. La administración del conocimiento (Knowledge Management) es un término que se ha convertido en una de las principales iniciativas de la tecnología de información. Como tal, existen distintas definiciones de éste. Algunos dicen que solamente es una nueva forma de administrar documentos y que únicamente incorpora el componente de almacenamiento y un mecanismo de distribución de información. En realidad, la administración del conocimiento es mucho más que eso. Quizás una de las mejores explicaciones de la administración del conocimiento sea la que da Gartner Group¹: "la Administración del Conocimiento promueve un enfoque orientado a identificar, capturar, recuperar, compartir y evaluar las ventajas de información de las empresas. Estos activos, información, pueden incluir bases de datos, documentos, normas y

¹ Firma consultora y evaluadora establecida en Estados Unidos www.gartner.com



procedimientos, así como también la experiencia tácita no capturada o documentada y la experiencia almacenada en la cabeza de cada uno de los trabajadores.

Una de las tendencias de la tecnología es convertir los datos en información y para esto debemos considerar la siguiente fórmula: Datos + Contexto = Información. Derivado de los avances de la tecnología, inclusive la Web, las organizaciones comienzan a orientar sus esfuerzos hacia la integración de: repositorios de información, documentos, groupware, intranet / internet / extranet, workflow y herramientas de búsqueda. Esta integración ayuda no solamente a manejar información sino que también permite tener una forma en la cual la gente pueda aprender y compartir con otras experiencias, de esta manera tenemos que, Información + Experiencia = Conocimiento.

Hay que tener presente que la administración del conocimiento no es del todo nueva. Nosotros usamos una agenda o un libro en donde registramos nuestras actividades y conversaciones diariamente, esto es simplemente una forma de administrar nuestro conocimiento. Con esto, en cualquier momento podemos recurrir a este almacén de conocimiento e identificar las principales actividades. Véase Figura 2.1.



El Data Warehouse es una arquitectura fundamental que apoya a la Administración del Conocimiento

Figura 2.1 Actividades



Beneficios de la Administración del Conocimiento

Si la administración del conocimiento fuera aplicada a toda una empresa, las compañías tendrían la habilidad de responder preguntas complejas en un período de tiempo corto. De esta manera, las compañías se podrían beneficiar de la velocidad de “respuesta del mercado” y de la “calidad en productos y servicios” que resultan cuando el conocimiento se capitaliza. Algunos ejemplos donde la administración del conocimiento puede mostrar beneficios dramáticos son los siguientes¹¹⁵:

- Cuando un nuevo equipo de ingenieros de Ford Motor Co. decidieron dividir el proceso de diseño del Taurus, miembros del equipo encontraron que nadie había generado registros escritos del proceso. Los secretos se perdieron.
- International Harvester, pidió construir una fábrica en Rusia porque 20 años antes se había construido una con mucho éxito, no encontraron a nadie en la organización familiarizada con el primer proyecto ni la documentación.

Si se hubiera utilizado la administración del conocimiento en estos dos ejemplos, las compañías no habrían tenido que asignar tiempo valioso, dinero y recursos humanos para reinventar los procesos que de alguna manera ya habían sido probados en el pasado.

Componentes Involucrados

La función principal de la administración del conocimiento es la de conectar dos componentes “el dueño del conocimiento” y “el explorador del conocimiento”. El conocimiento de uno es transferido a la mente del otro. La administración del

conocimiento proporciona los medios para capturar y almacenar el conocimiento y compartirlo con las personas adecuadas.

Antes de continuar, es importante diferenciar entre dos tipos de conocimiento, el tácito y el explícito.

El conocimiento tácito es más hipotético, subjetivo, personal y substancialmente más difícil de comunicar. El conocimiento tácito es el foco primario de muchas iniciativas de la administración del conocimiento, porque es el repositorio más estratégico de la organización.

El conocimiento explícito es el que se puede expresar clara y formalmente y el desafío primordial es administrar su volumen para asegurarse de su correcta aplicabilidad. El problema que enfrentan las organizaciones es el exceso de información, cuando los niveles de conocimiento explícito llegan a ser tan grandes es difícil filtrarlo adecuadamente. Así, la administración del conocimiento explícito es donde la mayoría de las organizaciones se enfocan.

Bill Inmon, uno de los padres de concepto de data warehousing dice que normalmente las organizaciones durante mucho tiempo acumularon mucho conocimiento en los grandes computadores, pero nunca han explotado ese conocimiento acumulado.¹¹⁶¹

Aplicaciones de la administración del conocimiento

Varios expertos en administración del conocimiento han dividido las funciones claves en cinco aplicaciones, las cuales son las siguientes:

- Intermediar, intermediario de la transferencia del conocimiento entre un proveedor de conocimiento y un explorador de conocimiento para lograr un cruce efectivo de conocimientos.
- Externalizar, transferencia del conocimiento de las mentes de sus poseedores en un repositorio externo para poder compartirlo.
- Internalizar, extracción del conocimiento del repositorio externo y filtrar ese conocimiento para obtener información relevante.
- Entender, la función de los sistemas para la toma de decisiones basado en el conocimiento disponible y la aplicación del conocimiento que se ha transformado.
- Medir, actividades de la administración del conocimiento que miden, mapean y cuantifican el conocimiento y el desempeño de soluciones de la administración de conocimiento.

Para llevar a cabo cada una de las actividades anteriores, es necesario utilizar herramientas, además debemos entender que la administración del conocimiento no es meramente la combinación de las tecnologías existentes actuales. Una solución de administración del conocimiento debe contener las siguientes características:

- Sensibilidad de Contexto. La solución debe ser capaz de "entender" el contexto de la necesidad del conocimiento y presentarlo a la medida.

- ☉ Sensibilidad de Usuario. La solución debe ser capaz de organizar el conocimiento para hacerlo lo más útil posible para el explorador del conocimiento.
- ☉ Flexibilidad. La solución debe ser capaz de manejar y presentar el conocimiento de cualquier forma y con distintas combinaciones.
- ☉ Heurística. La solución debe aprender de los usuarios y del conocimiento que es usado. Con el tiempo, la habilidad de proporcionar conocimiento a los usuarios debe mejorar.
- ☉ Sugestiva. La solución debe ser capaz de deducir cuales son las necesidades de conocimiento de los usuarios y sugerir el conocimiento asociado a sus responsabilidades.

Transferencia de la Administración del Conocimiento

La administración del conocimiento es una disciplina que promueve un enfoque integrado para la creación, captura, organización, acceso y uso de los activos de información de la organización. Estos activos incluyen las bases de datos, los documentos, la experiencia tácita y la experiencia de los individuos de las compañías desarrollando e implementando data warehouses. Un data warehouse contiene muchísimo conocimiento diversificado en sistemas y datos estructurados y en los procesos y reglas de negocio desarrollados por una empresa. Este conocimiento, crucial para realizar retornos importantes de inversión no puede ser accesado por usuarios finales o derivado de sistemas data warehouse.



Para obtener beneficios significativos de este conocimiento, las organizaciones deben desplegar un sistema de apoyo de conocimiento y procesos de administración de conocimiento que se alineen con las necesidades de las organizaciones.

El primer paso para implementar un sistema de apoyo de conocimiento es desarrollar una estructura de transferencia del conocimiento la cual definirá las metas de información de la organización con respecto al uso del data warehouse. Como por ejemplo:

- ⊖ ¿Qué información debe conservarse sobre el ciclo de vida del proyecto?
- ⊖ ¿Cuáles áreas o departamentos de la organización usarán la información?
- ⊖ ¿Cómo se capturará esa información en cada etapa del proyecto?
- ⊖ ¿Cómo se actualiza la información durante la etapa de implementación?

Además, la transferencia del conocimiento proporciona un modelo del ambiente del conocimiento durante el ciclo de vida de un sistema. Esta transferencia ilustra las relaciones e interdependencias de gente, procesos de negocio, acontecimientos y conocimiento. Con esta estructura en el lugar, los administradores de la implementación de warehouses son capaces de anticiparse y planear para los requerimientos de conocimiento de cada una de las fases de implementación y, lo que es más importante, proyectos futuros de data warehousing.

El equipo del proyecto puede desarrollar también mecanismos para transferir o transformar el conocimiento de una fase a otra, negociar exitosamente las “fronteras

de las fases” donde los cambios de personal y la orientación de nuevas tareas puedan tener como resultado pérdida fácil de conocimiento.

Asegurar la transferencia de conocimiento a las fases de las fronteras es muy importante. El conocimiento recopilado durante la construcción y el desarrollo de cada una de las etapas podrá ser necesario después para eficientar data warehouse para nuevas aplicaciones analíticas y de reporte y para ajustar los requerimientos cambiantes del negocio.

Esta transferencia de conocimiento es especialmente importante cuando la compañía utiliza a especialistas externos para ayudarlo con su data warehouse. En estos casos, la meta debe ser transferir el conocimiento y la experiencia al personal que se convertirá en el dueño del data warehouse y preferentemente que esta transferencia sea desde el principio del proyecto.

Ciclo de Vida del Data Warehouse y Herramientas de Administración del Conocimiento

Los proyectos de data warehousing pueden ser caracterizados por varias etapas: Análisis / Diseño / Construcción / Implementación. Cada una de estas etapas implica a diferentes especialistas en tecnología, en negocios, en análisis financiero, en administración y otros que no son especialistas en disciplinas analíticas. Las siguientes tecnologías de administración del conocimiento y procesos son útiles para soportar el análisis/diseño/construcción/implementación del data warehouse:

- Metadatos (reglas de negocio, premisas, extracción de datos y algoritmos de transformación, etc.).
- Un ambiente de administración del conocimiento de trabajo en grupo que proporciona administración de: proyectos, de documento, colaboración / comunicación, experiencia / habilidades del equipo y lecciones aprendidas durante el proceso.
- Un portal de información que proporcione el acceso a fuentes externas de conocimiento e integre la información recibida en el contexto del proyecto de data warehousing y el ambiente de administración del conocimiento.

Las organizaciones que deseen adoptar un sistema data warehouse deben desarrollar un plan para integrar su fase de construcción con un ambiente de conocimiento de tecnología de información. La documentación deberá reflejar los procesos de toma de decisiones de sistema data warehouse, así como también el impacto en cambios en el sistema, se debe identificar e incluir en el modelo de conocimiento.

Parte de las tecnologías de la administración del conocimiento que puede soportar la etapa de productividad son:

- Un controlado y actualizable repositorio para pruebas y entrenamiento así como también "lecciones aprendidas";
- Un repositorio accesible y actualizable para los componentes de conocimiento de tecnología de información, tales como códigos de software, modelos y reglas de negocio;



- Reporteo analítico y herramientas de servicio para acceder y usar reportes basados en el conocimiento.

Durante la fase operacional, un sistema sano de administración del conocimiento es particularmente valioso en las organizaciones que emplean las redes externas, sistemas de negocios electrónicos (e-business), automatización de fuerzas de venta y las cadenas de valor de conocimiento de sistemas extranet. Los sistemas funcionales de la administración del conocimiento deben integrar el conocimiento derivado del almacén con procesos de negocio a través de la empresa. Los elementos claves de estos sistemas de administración de conocimiento son:

- Un solo punto de acceso a la información, normalmente el conocimiento de las organizaciones se encuentra en una gran variedad de bases de datos, tanto externas como internas, la idea central consiste en concentrar en una sola base de datos, el conocimiento más relevante, importante y representativo de la organización.
- Habilidad de notificar o poner sobre aviso a los exploradores del conocimiento cuando los datos de almacén o reportes se actualizan o modifican.
- Integración, acceso a herramientas analíticas especializadas y a aplicaciones basadas en conocimiento.

El desarrollo de un data warehouse se basa en un conjunto complejo de reglas de negocio, algoritmos, suposiciones y una única fuente de datos. El desarrollo de un

data warehouse proporciona a los usuarios acceso fácil a la información, apoyándolos para la toma de decisiones importantes del negocio. Pero esto es sólo posible cuando las organizaciones pueden obtener y entender fácilmente el conocimiento en el que el almacén está fundamentado.

Cultura para la Administración del Conocimiento

¿Qué es lo que hace una organización para descubrir el conocimiento colectivo? En realidad, es menos sobre las cuestiones de tecnología y más sobre los procesos de gente y el negocio. Si los empleados no trabajan en un ambiente colaborativo o si no existe un lugar para concentrar y compartir el conocimiento, ninguna inversión en tecnología cambiará esto. Para que la administración del conocimiento funcione adecuadamente, todos los usuarios deben estar dispuestos a trabajar enfocados a lograr juntos con las metas y expectativas del negocio y la organización.

La administración del conocimiento es colaborativa por naturaleza. Antes de emprender una iniciativa de administración del conocimiento, debe comunicarse a la organización que el nuevo proyecto es importante para la empresa con la finalidad de completar todas las metas establecidas y destacar que es necesario trabajar juntos. Después de todo, el negocio está para hacer dinero, y la administración del conocimiento será implementada con la finalidad de lograr hacerlo.

Así, toma sentido para tratar de unir a la administración del conocimiento con resultados medibles. Esto es un desafío porque el conocimiento es intangible, pero es posible identificar las ganancias en áreas tales como servicio al cliente.

Debemos pensar en los beneficios de la administración del conocimiento ya que puede traer una reducción de tiempo importante al recibir y vender un producto, acelerar el proceso operacional, eliminar la redundancia y la conformidad de regulaciones, por resaltar unas pocas.

Para el empleado que necesita resolver un problema del negocio y que necesita ayuda para hacerlo hay una razón de compartir el conocimiento adquirido. Pero algunos empleados sienten que su valor (y la seguridad del trabajo) será más alto si ellos no comparten lo que saben con otros, así que prefieren acumular el conocimiento. Este tipo de ambiente puede ser el resultado de información combinada en la organización.

La implicación para las organizaciones es que si los empleados no están dispuestos a compartir lo que ellos saben, la tecnología está jalando con el problema. Para que la administración del conocimiento sea verdaderamente eficiente, un ambiente de colaboración debe estar en un lugar en donde los empleados no compitan el uno con el otro.

Tecnología para la Administración del Conocimiento

Para los administradores de la tecnología de información de las organizaciones, no es importante como es definida la administración del conocimiento o si las teorías actuales de la administración del conocimiento abarcan en su totalidad el concepto. La parte importante es cómo usar la tecnología para crear un sistema efectivo para la administración del conocimiento.

Hoy por hoy, no existe en el mercado del software un producto que sea capaz de cubrir por completo las necesidades de la administración del conocimiento. De los

diferentes sistemas que existen, éstos solamente son piezas del gran rompecabezas de administración del conocimiento. Pero, si la finalidad de la administración del conocimiento es la integración, estos productos nunca se usarán para construir una infraestructura adecuada de administración del conocimiento ya que requiere de la integración de diferentes tecnologías, basadas en las necesidades específicas de cada una de las organizaciones.

Desde un punto de vista funcional, hay cinco capacidades claves que deben ser parte de un buen sistema de administración del conocimiento:

1. **Concentrar:** Un sistema de administración del conocimiento debe proporcionar una manera de agregar la información al repositorio del sistema o controlar la ubicación de la información sin necesariamente agregarla al repositorio. Preferentemente, el sistema debe ser capaz de reunir la información de todas las fuentes externas importantes. Además, permitir que todos los usuarios contribuyan a acumular conocimiento o imponer reglas para su acceso y actualización.
2. **Organizar:** Un sistema de administración del conocimiento debe proporcionar un modelo taxonómico o contar con un modelo para organizar y clasificar la información en diferentes formas o niveles de agregación. Preferentemente, el modelo debe ser intuitivo hasta que la gente aprenda a usarlo fácilmente. Este modelo debe ser suficientemente flexible para considerar modificaciones fáciles. Algunos sistemas permiten que los usuarios creen sus propios modelos de organización, proporcionando un contexto personal en cuanto a la aplicabilidad de datos o crear relaciones extraordinarias entre objetos.
3. **Buscar / Distribuir / Entregar:** Un sistema de administración del conocimiento debe proporcionar las maneras de distribuir y entregar información, así como también permitir que los usuarios hagan búsquedas de información dinámicas

de información. Algunos productos entregan la información directamente al usuario o permiten que los usuarios creen "agentes" para notificar sobre nuevos contenidos. Muchos productos ofrecen también la habilidad de buscar en la base de conocimiento para la información específica o vía una búsqueda de solo texto o una búsqueda contextual.

4. Colaborar: Un sistema de administración del conocimiento debe ser colaborativo y debe compartir el conocimiento. Los usuarios necesitan mandar mensajes entre sí, participar en estos flujos de trabajo, informar las rutas a otros usuarios, participar en discusiones, agendar reuniones, etc. Algunos productos proporcionan sus propias capacidades colaborativas.
5. Refinar: Un sistema de administración del conocimiento puede proporcionar las características que mejoran el valor de la información para usos específicos. Algunos productos permiten a los usuarios refinar o analizar el contenido de la base de conocimiento de diferentes maneras.

La Administración del Conocimiento en el Mercado

Una vez que las organizaciones logran un clima o procesos corporativos que promueven el ambiente colaborativo y de compartir el conocimiento entre los empleados, se puede empezar a mirar a la tecnología como el elemento principal para sus procesos de administración del conocimiento. Si las metas del negocio son claras y se tiene un ambiente adecuado para la administración del conocimiento, las tecnologías que se seleccionen serán más fáciles de implementar. La mayoría de las aplicaciones de administraciones del conocimiento son una combinación de muchas tecnologías.



Para desarrollar un sistema valioso de administración del conocimiento, se necesita tener en mente un mapa de toda la situación y entonces desarrollar una infraestructura de tecnologías capaz de apoyar el buen desarrollo y cumplimiento de las metas. Se recomienda comenzar con una aplicación pequeña o hacer primero la implementación en un área específica.

Mientras los vendedores y los tecnólogos abrazan el tema de la administración del conocimiento, los directores de las empresas expresan con escepticismo: "yo no tengo problemas de conocimiento, yo tengo problemas de negocio." ¿La administración del conocimiento seguirá la suerte de la inteligencia artificial como una solución de tecnología en la búsqueda de un problema de negocio? ¿La administración del conocimiento es otra novedad como la "curva experiencia" como un interesante modelo para el análisis pero no para la acción? ¿Cómo deben comprometer las organizaciones a la administración del conocimiento? ¿Cómo lograr que la administración de conocimiento agregue valor al negocio?

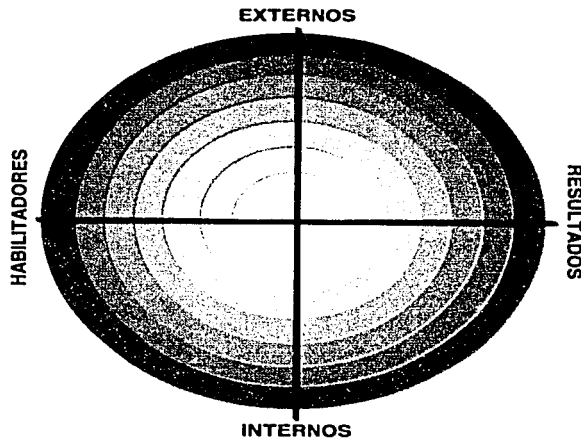
Usando la administración del conocimiento como una herramienta de evaluación combinada con datos tradicionales ayuda a definir los mapas de procesos. En cada tarea o proceso, información y pericia requeridas se definen. La localización de la información o la "localización" del conocimiento es identificada. Este panorama alterno identifica las fuentes del conocimiento y su uso. Este mapa de proceso está orientado a automatizar el acceso al conocimiento.

La administración del conocimiento no puede triunfar si se considera como revolucionaria sobre la tecnología. Cuándo las disciplinas establecidas (administración

de registros) y tecnologías (herramientas de flujos de trabajo y de búsqueda), puede ofrecer nuevas aplicaciones con beneficios más significativos al negocio que no se pueda reconocer como administración del conocimiento, pero sí en las organizaciones mejorarán el rendimiento de su conocimiento crítico, la administración del conocimiento proporcionará una excelente herramienta especialmente cuando se ligue con las estrategias del negocio. Usando el conocimiento de los procesos para mejorar el componente "tácito" de la administración total de calidad y de la mejora de procesos.



Capítulo Tercero



Medidores de Desempeño

Medidores de Desempeño

Los medidores de desempeño son esenciales cuando se maneja el desempeño de la empresa. Frecuentemente, los medidores de desempeño se confunden con el proceso iterativo de la administración del desempeño.

Durante mi estancia en la Facultad de Ciencias de la UNAM, recuerdo una definición que me gustó y que ahora aunque este trabajo no sea necesariamente matemático, si considero un buen ejemplo mencionarla, del libro "Introducción al Cálculo y al Análisis Matemático" de R. Courant y F. John dice "Las nociones básicas del Cálculo son la de derivada e integral: la derivada es una medida de la rapidez de una variación; la integral, una medida del efecto total de un proceso de cambio continuo una comprensión precisa de estos conceptos y de su abrumadora funcionalidad descansa sobre los conceptos de límite y función; los cuales a su vez dependen de un entendimiento del continuo de números. Sólo gradualmente penetrando más y más en la esencia del Cálculo, puede uno apreciar su potencia y belleza"...

La explicación más sencilla que describe los medidores¹ de desempeño es que son monitores métricos aislados del desempeño que explica qué ocurre en una organización. Estos medidores son útiles cuando se usan para compararlos con las metas y medir el grado en que éstas fueron alcanzadas. La medida de los indicadores clave de desempeño (Key Performance Indicators "KPI's") ayuda a los directores a identificar fuerzas y debilidades en la organización; además, proporcionan un punto de partida para el proceso de la administración del desempeño porque demuestra si una organización está en línea con los objetivos estratégicos.

¹ El motivo por el cual se utiliza Medidor en lugar de Medida es que primero se selecciona la herramienta (Medidor), la cual tiene la escala a utilizar (Medida), por ejemplo, un medidor puede estar enfocado a conocer el desempeño de un área en particular, en el área de sistemas puede ser reportes el medidor y la medida los reportes levantados vs. los solucionados. Este mismo medidor no puede ser aplicado a un área de finanzas

Actualmente, muchas organizaciones miden el desempeño con un sistema de análisis sencillo, en lugar de acumular y analizar los KPIs; sin embargo, los medidores de desempeño explican el pasado, no describen la manera en que las organizaciones realizan sus procesos internos. Los resultados finales son derivados de un sistema de administración del desempeño porque enmarcan primordialmente la estrategia de la empresa.

Desgraciadamente, muchos directores o gerentes utilizan KPI's aislados para la toma de decisiones estratégicas. Esto es un desperdicio importante de los medidores de desempeño, ellos deben utilizar los KPI's para determinar si la organización cumple, y en que grado lo hace, con las metas o para identificar las fuerzas y debilidades de la organización. Una consideración es que un sistema de medidores de desempeño aislados conduce campañas de desempeño aisladas y, por lo tanto, la toma de decisiones también es aislada. Muchos han descuidado agregar el elemento estratégico de la administración del desempeño, los directores y gerentes deben darse cuenta que los medidores de desempeño son el primer paso integral en el proceso de la administración del desempeño.

Administración del Desempeño

La administración del desempeño es el proceso iterativo de monitorear el desempeño de la empresa al manejar los procesos internos y realizar el cambio estratégico. Los Tableros Balanceados² (Balance Scorecard) proporcionan una

² El concepto de tablero balanceado se verá con más detalle en las siguientes páginas de este capítulo.

metodología efectiva para la evolución estratégica de la organización. Esta metodología de la administración ofrece un panorama del desempeño de la empresa de arriba hacia abajo con un fuerte enfoque en la visión y estrategia. Para desarrollar un tablero efectivo, uno inicia con la visión de la empresa; es decir, la misión. Después, mientras se tiene en mente la estructura organizacional, los directores y gerentes deben decidir cuáles son los objetivos estratégicos que conducirán al logro exitoso de las metas; entonces, deben decidir cuáles iniciativas estratégicas ayudarán a la empresa a cumplir con los objetivos. Las iniciativas estratégicas se traducen en actividades impulsoras del desempeño, tácticas y específicas. Finalmente, las métricas son establecidas para cada actividad. Estas llegan a ser las medidas del desempeño de la empresa. Ahora la administración tiene un cuadro de medidas de desempeño interrelacionadas.

La administración estratégica de la empresa apoya a las mejoras del proceso que optimizan el desempeño en niveles funcionales y la visión estratégica. El proceso de la administración del desempeño incluye la medida de las métricas clave de la empresa, el análisis de resultados métricos y la planificación y administración de la estrategia de la empresa. Para sobrevivir en el sistema económico global y dinámico actual, la evaluación y corrección estratégicas deben ser posibles. Sólo un sistema de administración de desempeño genera la evolución estratégica.

Para ilustrar, se pueden considerar los siguientes objetivos de desempeño:



➤ **Maximizar Margen vs. Maximizar Utilidad.**

Dependiendo del tipo de producto, se puede ser más inteligente disminuyendo el margen de utilidad para incrementar el total de las utilidades. Estas medidas son inversamente relacionadas, si los directores y gerentes rastrean un valor métrico de manera aislada, no pueden detectar esta relación, con lo que pueden incrementar el margen de utilidad, antes que identificar el margen exacto para utilizarlo al máximo.

➤ **Mejorar el Servicio al Cliente vs. Disminución en el Costo de Inventario.**

Los administradores tradicionalmente luchan contra el dilema de mejorar el servicio al cliente al mantener un costo bajo los inventarios; si el costo de los inventarios es alto, pueden decidir mantener el inventario bajo. Esto puede sacrificar la calidad de servicio al cliente. Si la satisfacción del cliente impacta en servicio y lealtad del cliente, se pueden incrementar los niveles de inventario, si no existe una metodología para detectar relaciones entre el desempeño y los resultados, entonces no hay manera de administrar completamente el desempeño. Las decisiones que se basan en los medidores del desempeño sólo pueden ser una aventura más bien arriesgada. A partir de que en la empresa exista un ambiente dinámico, el sistema de administración del desempeño debe cuantificar las relaciones entre actividades y progreso de las metas estratégicas.



Considerando que los tableros balanceados no son una herramienta de información alrededor de las estrategias y del cuadro del tablero balanceado, las aplicaciones de la administración del desempeño deben ser capaces de lo siguiente:

Medir / Analizar / Planear / Administrar

○ Medir

Este es el proceso por lo cual los usuarios aprenden qué ha pasado, estableciendo métricas que se ligan a las metas estratégicas, los usuarios son capaces de medir la relación entre el desempeño histórico actual y toda la estrategia de la organización. El modelo de la medida contiene los objetivos de desempeño, inclusive mejores prácticas que las de industria para cada uno de las métricas. El desempeño actual se puede comparar entonces contra estos objetivos en una escala del 1 al 10 y ver su evolución. Una aplicación de tablero balanceado debe estar habilitada para:

- **Integrar medidas cuantitativas y cualitativas:** Medidas cuantitativas para ser capaces de ser representadas en términos numéricos: es decir, retorno sobre la inversión y entrega en tiempo. Las medidas cualitativas no tienen valor numérico están generalmente basadas en opiniones, es decir, satisfacción del cliente. Las aplicaciones de tableros balanceados deben permitir la entrada



cliente. Las aplicaciones de tableros balanceados deben permitir la entrada de ambos tipos de medidas, permitiendo el análisis de datos difíciles y de esta forma transformar opiniones en resultados.

- **Establecer objetivos y mejores prácticas:** Evitar evaluar el desempeño como un resultado vacío, las aplicaciones de tableros balanceados deben proporcionar la habilidad de establecer objetivos relativos al desempeño tal como peor / mejor escenario, mejores prácticas de industrias y desempeño histórico.
- **Representar gráficamente el desempeño:** Visualizar los resultados gráficamente, un tablero balanceado debe permitir la representación gráfica del desempeño, esto proporciona los medios para comparar el desempeño actual con los objetivos y mejores prácticas.

○ Analizar

Esto es un ejercicio crítico que habilita a los usuarios a responder ¿por qué?, usando tecnologías y técnicas de inteligencia de negocios, así, los usuarios son capaces de buscar en los datos del desempeño para entender mejor por qué el desempeño es bueno o malo. Llegando en una exacta comprensión, los usuarios pueden manejar efectivamente cambios para corregir el pobre desempeño o perpetuar un buen desempeño. Las siguientes características son imprescindibles para el análisis:



- **Acceder directamente los datos de desempeño:** Entender por qué los resultados del desempeño son buenos o malos, esencial para tener acceso directo a los datos. Las soluciones robustas de tableros balanceados deben proporcionar el acceso a la mayor parte de las bases de datos. Con acceso directo a los datos, los usuarios pueden automatizar el proceso de recolección de datos y realizar a fondo el análisis de resultados de desempeño.
- **Analizar interactivamente los datos:** Con la funcionalidad fundamental de inteligencia de negocios, los usuarios pueden analizar los datos de desempeño de distintas maneras, incluyendo: reporte, análisis multidimensional, tablero de ejecutivos, pronóstico y análisis “qué pasa sí...”. Así, los altos mandos pueden controlar el pulso de la organización con un nivel de visión alta proporcionada en los tableros a los gerentes de operaciones con un panorama detallado de desempeño utilizando la funcionalidad de búsqueda hacia abajo o detallada, gráficos e informes.

⊖ Planear

Este es el proceso en cual los usuarios son capaces de preguntar “qué pasa sí...”. La habilidad de ligar el desempeño a actividades específicas del negocio, detectar relaciones estadísticas que existan entre varias actividades económicas y desarrollar a usuarios para jugar a “qué pasa sí...” con iniciativas estratégicas es el más poderoso entregable de una solución de administración del desempeño. Los usuarios verán claramente cuáles actividades tienen impactos más importantes en la visión organizacional. Este conocimiento los

ayuda a identificar cuáles actividades son sólo un gasto, y cuales no lo son para poder invertir tiempo y dinero en ellas.

El análisis estadístico del modelo de la administración del desempeño facilita a los directores y gerentes a cuantificar la unión entre medidas tácticas y metas estratégicas, probando varios escenarios dentro del modelo de la administración del desempeño se puede identificar cuál estrategia optimiza el desempeño corporativo. La planeación estratégica sugiere remedios para la administración de dilemas y penetrar en resultados de iniciativas hipotéticas. Consecuentemente, los altos mandos habrán establecido la metodología para realizar el cambio estratégico. Las características de la planeación incluyen:

- Validar Estrategias: Estadísticamente validar las hipótesis estratégicas aplicando análisis de regresión a los resultados del desempeño. Los coeficientes de correlación determinarán si las relaciones estadísticas existen entre las iniciativas. Así, se puede determinar la extensión de iniciativas estratégicas de interrelación.
- Predecir el desempeño: Una vez interrelacionadas las iniciativas, estas son cuantificadas, utilizando el análisis "que pasa sí" para predecir los resultados de las hipótesis estratégicas. Esto proporciona la habilidad de probar las hipótesis del negocio antes de la implementación.



➤ Administrar

Una vez que las iniciativas estratégicas han sido descubiertas por la planeación estratégica, los resultados y las iniciativas deben ser comunicadas a toda la organización. Esta tarea trae consigo el proceso de implementar cambios, que requiere de una comunicación extensa. Los usuarios deben ser capaces de comunicar el propósito estratégico y lograr resultados de desempeño compartidos. La evolución estratégica ocurre cuando todos los empleados implementan cambios para alinear las actividades tácticas con la estrategia. La administración estratégica se logra mediante lo siguiente:

- **Diagramas de Causa y Efecto:** Estos son representaciones gráficas de un plan estratégico de la organización. Los diagramas de causa y efecto facilitan a planificadores estratégicos a representar las interdependencias entre iniciativas y áreas funcionales. Por ejemplo, al fortalecer la línea de productos, los gerentes de producción pueden aumentar los recursos dedicados a investigación y desarrollo tecnológico con la finalidad de lograr innovación y mejorar los productos.
- **Conectar el desempeño a la estrategia:** Cada iniciativa estratégica dentro del esquema de causa y efecto debe estar conectada a un tablero subordinado. Con esto, las iniciativas reciben sus resultados del tablero principal. Esto permite que los usuarios pueden combinar directamente la estrategia a las medidas individuales de desempeño.

- **Vincular los planes de acción al desempeño estratégico:** Vinculando los planes de acción a las medidas de desempeño y/o iniciativas estratégicas, los individuos pueden identificar el progreso de proyectos directamente del tablero. Además, los usuarios serán capaces de demostrar cómo varios planes de acción apoyan la estrategia.

La innovación y el cambio económico demandan modificaciones estratégicas. Un sistema de administración del desempeño puede reconocer los cambios en el ambiente de la organización e incrementar la toma de decisiones estratégicas. Por lo tanto, los directores y gerentes estarán tomando un papel activo que reactive el enfoque de desarrollo estratégico, analizando información del desempeño y descubriendo resultados cuantificables para la planeación estratégica.

Tableros Balanceados

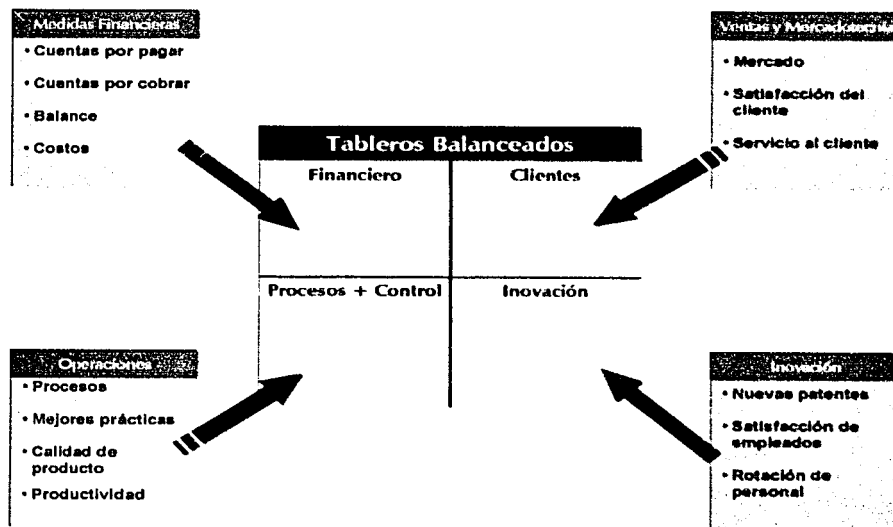
El concepto de tableros balanceados fue acuñado por Robert Kaplan y David Norton en 1992. Los tableros balanceados son una técnica con la cual una compañía puede ver el impacto que las principales actividades que se tienen hoy en la dimensión de críticas del desempeño corporativo.

En un tablero balanceado tradicional se tienen tres dimensiones:

- Financieras;
- Clientes; e



La cuarta dimensión está abierta a dar a la compañía cierta flexibilidad en definir una dimensión primaria del negocio llamada Innovación. Creando la dimensión apropiada del tablero balanceado es donde la estrategia de la organización está vinculada a los sistemas de medidores. Véase Figura 3.1.



Ejemplo de Tablero balanceado

Figura 3.1

Los tableros balanceados en una organización facilitan el conocimiento, uso, incremento e integridad de información como un proceso de valor agregado a la administración. Estos tableros balanceados funcionales ayudan a los usuarios a entender mejor la demanda de los productos y entender su demanda en el mercado.



Evolución del Tablero Balanceado

La llegada de la era de la información ha acelerado el cambio dentro de las empresas. En la mente de muchos, la información ha llegado a ser el activo más importante de la empresa. Armada con información, una firma puede:

- Desarrollar mejores relaciones con clientes
- Descubrir y servir a nuevos mercados
- Introducir productos innovadores
- Producir productos con calidad con ciclos más cortos de desarrollo de producto y reducir costos
- Impulsar la mejora continua
- Desplegar la tecnología para analizar mejor la información

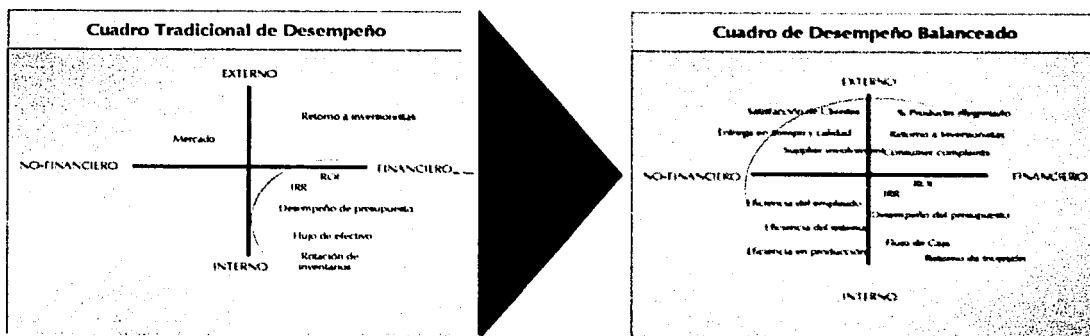


Figura 3.2. Tablero balanceado - Tradicional vs. Medidores de Desempeño Balanceados



El tablero balanceado comprende un panorama de arriba hacia abajo del desempeño de la organización con un fuerte enfoque en la visión y la estrategia. Para desarrollar un tablero efectivo, se inicia con la visión de la organización (en otras palabras, la misión). Después teniendo en mente la estructura de la organización, los directores deben decidir cuáles estrategias encabezarán los objetivos. Estas estrategias se traducen en actividades impulsoras de desempeño, tácticas y específicas. Finalmente, los métricos son establecidos para cada actividad. Estos llegan a ser las medidas del desempeño de la empresa. Así, una vez que la visión y la estrategia se han desarrollado, la métrica individual resultará de manera muy natural.

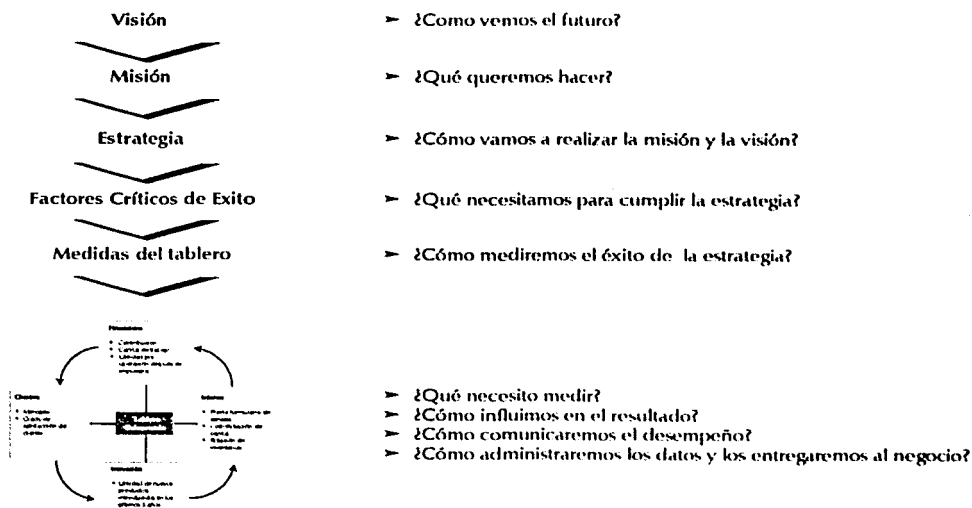


Figura 3.3. Tablero Balanceado Representativo



Volver al Futuro

Entender la situación de la empresa es vital al tomar decisiones y, de aquí en adelante, el éxito se convierte en crítico, aún las técnicas tradicionales de administración de desempeño han discutido que éstas sólo describen la historia de la organización, mientras que una pequeña penetración proporciona tendencias futuras. El tablero balanceado soluciona problemas tales como dilemas históricos proporcionando una herramienta de la administración que no sólo facilita la investigación histórica sino que también se asegura de enfocarse en el futuro. El tablero balanceado está basado en ideas financieras, mientras que un indicador importante del desempeño corporativo tiende a ser algo retrospectivo. En otras palabras, el métrico financiero típicamente nos dice cómo una empresa se ha desempeñado, pero nos da una diminuta indicación de cómo se podría realizar. Un tablero balanceado real debe incluir métricas que proporcionan ambas tanto históricas como futuras. Así, un tablero de la empresa debe estar compuesto de ambos indicadores. Los indicadores delanteros son las medidas que manejan el desempeño, mientras que los indicadores traseros resultados actuales del desempeño de la empresa.

Las Cuatro Perspectivas

Todavía hoy la salud de las organizaciones es estimada comúnmente por medidas aisladas, con mucho énfasis en el área financiera. Para lograr "balance" las metodologías prescriben la evaluación estratégica de cuatro perspectivas:

- Financiera;
- Cliente;
- Interna; y
- Aprendizaje y Crecimiento

Consecuentemente, el Tablero Balanceado fuerza a los directores a abandonar la creencia de que las medidas financieras tradicionales son suficientes para el análisis estratégico de la empresa, el análisis estratégico de la empresa ahora se ha incrementado para incluir dimensiones de cliente, interna y perspectivas de la innovación, así como también la financiera. En vez de analizar un grupo de métricas perdidas, el análisis balanceado de la empresa se enfoca en el análisis de la empresa en las cuatro áreas claves, las cuales se describen a continuación:

- **Financiera.** Históricamente, esta medida ha sido la piedra angular de la medida del desempeño. Como punto inicial es un buen indicador para identificar si la estrategia de la empresa es aplicada y ejecutada. Algunas medidas financieras tradicionales incluyen los ingresos por operación y el retorno de la inversión por empleado.
- **Cliente.** La perspectiva de cliente identifica al cliente y segmento de mercado objetivo para cada unidad de negocio. Los indicadores comunes del desempeño de cliente incluyen el mercado, la satisfacción del cliente y retención de cliente.
- **Interna.** Se enfatiza la importancia de tener dos tipos de productos en la línea de productos: de hoy y de mañana. Analizar la perspectiva interna es clave para diferenciar cuáles productos se colocan en cada categoría. El

diseño de productos, la administración de desarrollos y operaciones maneja el éxito de una línea de producción. La administración interna del desempeño permite que identifiquen cuáles productos deben ser cosechados y cuáles eliminados.

- **Aprendizaje y Crecimiento** – Para sobrevivir, las empresas deben reinventarse a sí mismas para permanecer en el mercado. Las medidas del aprendizaje y crecimiento identifican la infraestructura más apropiada para la firma, son derivadas del recurso humano, sistemas y procedimientos de la organización. Las medidas tradicionales del aprendizaje y crecimiento pueden abarcar: la satisfacción de empleados, la retención y habilidades; desempeño de sistemas de información, los procedimientos de la organización y los estímulos de los empleados.

Comunicación y Realimentación

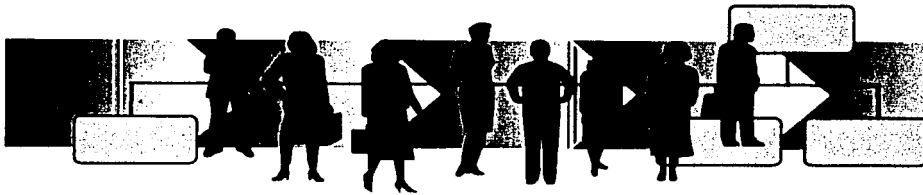
Así como un médico controla la salud de un individuo con el tiempo, los directores rastrean el desempeño de la empresa en intervalos específicos de tiempo para acertar cómo se deben desempeñar con relación a las metas. Más allá de la analogía, los médicos le dicen al paciente como están progresando hacia la meta. Similarmente, los directores deben comunicar primero la visión y la estrategia de la empresa a toda la organización que puede contribuir a su éxito. Esto es alcanzado no sólo proporcionando un panorama general de la visión, sino que también descomponiendo la visión y la estrategia en actividades y medidas específicas. Cada área funcional asumirá la responsabilidad para su métrica asignada. Un buen punto de partida se establece para que cada área funcional contribuya al éxito de la

empresa. Como médicos de la empresa, los directores proporcionan la realimentación del desempeño a los empleados y los ayuda en el ajuste de los objetivos tácticos: procurar que las actividades se alineen con los intereses estratégicos de la empresa.

El Tablero Balanceado ha llegado a ser el principal sistema estratégico de administración del desempeño de las empresas. Utilizándolo, los directores son capaces de traducir los tópicos de la misión en acciones tácticas y así dirigir a la organización hacia el logro de las metas estratégicas. Consecuentemente, el tablero balanceado facilita que directores exploten la información de la compañía y produzcan los resultados económicos esperados.



Capítulo Cuarto



El Data Warehouse y la Cadena de Valor

El Data Warehouse y la Cadena de Valor

Para estructurar un Data Warehouse eficiente hay que considerar que al igual que en una organización existen diferentes áreas, en cada una de las áreas hay diferentes necesidades de información, combinar la cadena de valor y el data warehouse es simplemente un arte.



Al igual que en una obra maestra de música clásica¹, el lograr una combinación entre armonía y pasión, silencio y sentimiento, compases y vivencias, para lograr tener un armonioso Data Warehouse, es necesario no sólo entender la función separada de una cadena de valor en una organización, el Data Warehouse, las áreas involucradas, las herramientas y, al último y no por ello menos importante, la gente, el capital humano.



En general, la definición o percepción de cadena de valor es que se refleja como un incremento teórico del valor sobre y por encima del costo inicial de un bien o un servicio. Generalmente se supone que este valor debe ser superior a los costos acumulados que se han "adicionando" a lo largo de la etapa del proceso de producción. Las actividades de valor agregado real son aquellas que, vistas por el cliente final, son necesarias para proporcionar el bien o el servicio que el cliente está esperando recibir y que está dispuesto a pagar por ello. Hay muchas actividades que la empresa requiere, pero que no agregan valor desde

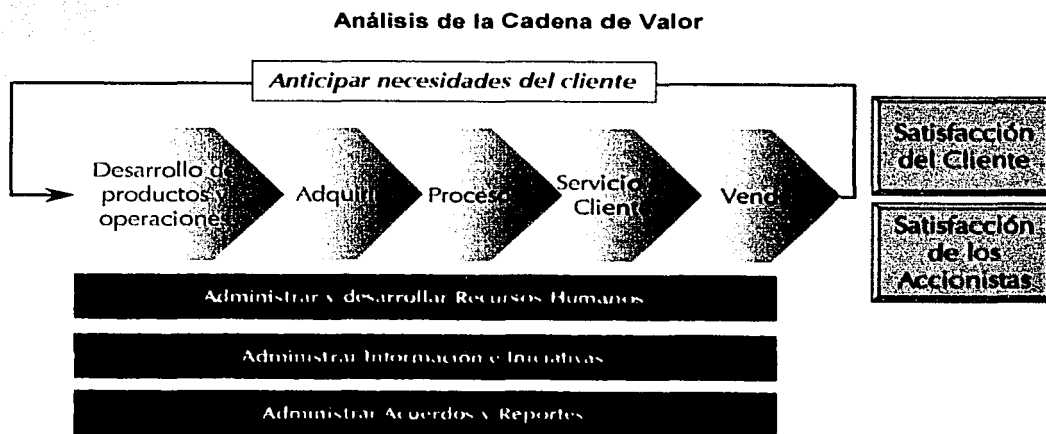
¹ Parte de la partitura de Romanza de Ludwig van Beethoven.

el punto de vista de ventajas para el cliente (actividades de valor agregado en la empresa). Además, existen otras actividades que no agregan valor alguno, por ejemplo, el almacenamiento, pero que son necesarias para el desempeño de la organización.

La técnica de Cadena de Valor, tiene la finalidad de identificar las actividades que se realizan en una organización, las cuales se encuentran inmersas dentro de un sistema denominado sistema de valor, que está conformado por:

- ♪ Cadena de valor de los proveedores
- ♪ Cadena de valor de las unidades del negocio
- ♪ Cadena de valor de los canales de distribución
- ♪ Cadena de valor de los clientes

Un ejemplo de cadena de valor, es el que se muestra en la siguiente gráfica:



La cadena de valor, es un modelo que describe una serie de actividades que adicionan valor a las empresas y a la forma de hacer las cosas en una organización, conectadas del lado del proveedor de la compañía (materiales en bruto, procesos logísticos de entrada y procesos de producción) con su otro lado de la demanda (procesos logísticos de salida, mercadotecnia y ventas).

Mediante el análisis del estado de la cadena de valor, los administradores han sido capaces de rediseñar sus procesos interno y externo para mejorar eficiencia y efectividad. El modelo de cadena de valor, trata la información como un elemento adicional que apoya los procesos de valor agregado y no como una fuente de valor en sí misma. Por ejemplo, los administradores en ocasiones utilizan la información que se genera en los departamentos de inventarios, producción o logística para ayudar a monitorear y controlar estos procesos. Pero muy raramente se utiliza la información en sí misma para crear un nuevo valor para el consumidor final.

Se define el **valor** como la suma de los beneficios percibidos que el cliente recibe menos los costos percibidos por él al adquirir y usar un producto o servicio. La cadena de valor es esencialmente una forma de análisis de la actividad empresarial mediante la cual descomponemos una empresa en sus partes constitutivas, buscando identificar fuentes de ventaja competitiva en aquellas actividades generadoras de valor.

Esa ventaja competitiva se logra cuando la empresa desarrolla e integra las actividades de su cadena de valor de forma menos costosa y mejor diferenciada que sus rivales. Por consiguiente la cadena de valor de una empresa está conformada por

todas sus actividades generadoras de valor agregado y por los márgenes que éstas aportan.

Una cadena de valor genérica está constituida por tres elementos básicos:

- ♪ Actividades Primarias, que son aquellas que tienen que ver con el desarrollo del producto, su producción, las de logística y comercialización y los servicios de post-venta.
- ♪ Actividades de Soporte, a las actividades primarias, como son la administración de los recursos humanos, las de compras de bienes y servicios, desarrollo tecnológico (telecomunicaciones, automatización, desarrollo de procesos e ingeniería, investigación), infraestructura empresarial (finanzas, contabilidad, gerencia de la calidad, relaciones públicas, asesoría legal, gerencia general).
- ♪ Margen, que es la diferencia entre el *valor* total y los costos totales incurridos por la empresa para desempeñar las actividades generadoras de valor.

El Análisis de la Cadena de Valor es una herramienta gerencial para identificar fuentes de Ventaja Competitiva. El propósito de analizar la cadena de valor es identificar aquellas actividades de la empresa que pudieran aportarle una ventaja competitiva potencial. Poder aprovechar esas oportunidades dependerá de la capacidad de la empresa para desarrollar a lo largo de la cadena de valor y mejor que sus competidores, aquellas actividades competitivas cruciales.

De las principales problemáticas a las que se enfrenta uno al construir un Data Warehouse es entender las necesidades de cada una de las áreas del negocio y entender las necesidades específicas de cada uno de los usuarios del data warehouse.

La cadena de valor es una herramienta importante pero que debemos a empezar a combinar identificando el tipo de usuario que hay en cada una de las áreas del negocio, comprender sus necesidades de información y la habilidad que tiene este usuario para buscarla, para esto, se plantea la identificación de usuarios de tipo turista, operarios, granjeros, exploradores y mineros. Sus perfiles serían los siguientes:



Turistas (Generalista)

Directores, Ejecutivos y Usuarios Casuales. Los turistas tienen generalmente una amplia perspectiva del negocio y son experimentados usuarios de Internet, están acostumbrados a una interfaz gráfica bastante sólida eso les proporciona la habilidad de buscar grandes bancos de datos sin necesidad de mucha mecanografía. Los ejecutivos, en particular, necesitan información básica para asegurar el correcto desempeño la compañía.



Operarios (Enfocado)

Soporte al Cliente, Manufactura y Control de Inventarios. Los operarios necesitan día a día información fresca, detallada para completar su trabajo. La mayoría de la información es necesaria en una base planificada, confían fuertemente en consultas

estandarizadas. Ellos esperan transacciones como tiempo de desempeño y de respuesta



Granjeros (Limpia)

Analistas de Ventas, Analistas Financieros, Directores de Mercadotecnia y Analistas de Contabilidad. Los granjeros saben qué datos quieren y en qué medida. A partir de que ellos han trabajado por capturar información sólida y estratégica durante mucho tiempo sin la fábrica corporativa de información, entienden el valor de los datos integrados y limpios.

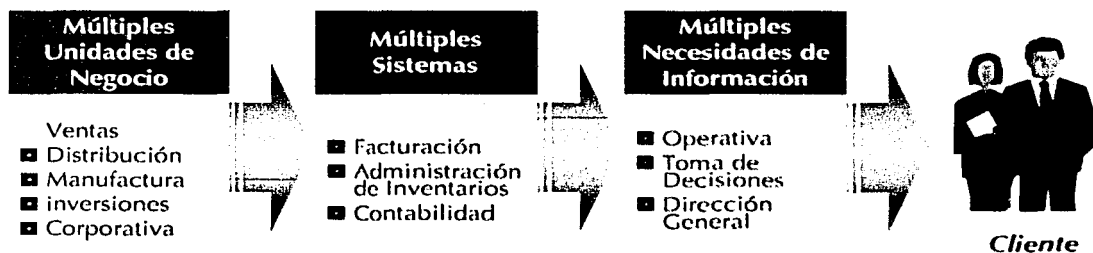


Exploradores (Innovador)

Actuarios, Ingenieros de Control de Procesos, Analistas de Ventas. Los exploradores son los pensadores “fuera de este mundo”. Ellos tratan de encontrar las relaciones entre piezas oscuras de datos y acontecimientos, a menudo tienen como resultado, consultas aleatorias. Los exploradores son a menudo injustos en sus hipótesis. Pero cuando tienen razón, el retorno puede ser tremendo para las corporaciones.

Experto en Mercados, Directores de Riesgo, Especialistas en Logística y Estadísticos. Los mineros buscan detalle a detalle para sostener las tendencias o los modelos. Quieren tanta historia como pueden encontrar. Buscan la “pepita de oro” que dará el eje competitivo a las organizaciones.

Con lo anterior, tanto la cadena de valor, los usuarios y las diferentes necesidades de información, tenemos la combinación para ofrecer a la organización y por ende a los consumidores finales mejor rentabilidad de la empresa, mejores productos y servicios, respectivamente, una forma en la que se puede apreciar mejor esta relación es la que se muestra a continuación.



El análisis que se hace en la organización muchas veces muestra que en más de un departamento se tiene la misma necesidad de información y algunas veces hay usuarios más aptos para obtener la información que a otros le cuesta más trabajo presentar y explotar.

La idea es que una vez que se tiene optimizada la cadena de valor de la organización un siguiente paso razonable es empezar a identificar en cada una de las áreas del negocio a los diferentes usuarios para saber quienes son los turistas, operarios, granjeros exploradores y mineros en la empresa e identificar necesidades reales de información para normalizar los reportes que se necesitarán tanto para la

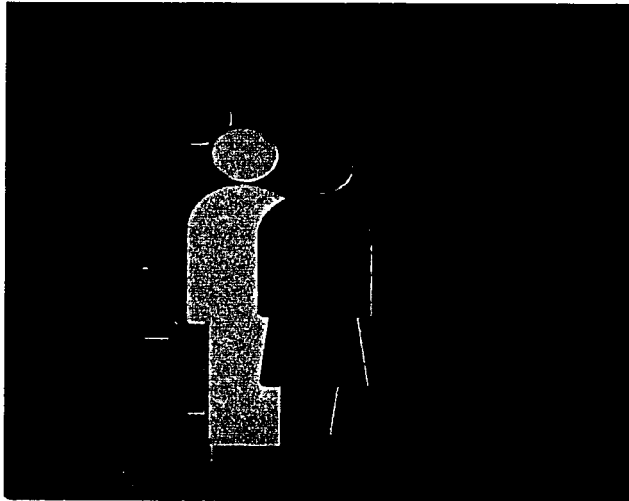


negocio a los diferentes usuarios para saber quienes son los turistas, operarios, granjeros exploradores y mineros en la empresa e identificar necesidades reales de información para normalizar los reportes que se necesitarán tanto para la organización como para cada una de las áreas del negocio, procurando no convertir al data warehouse en una herramienta de reporte sino más bien en una importante herramienta para la toma de decisiones.

Algunas veces resulta complicado que los usuarios se adapten a la nueva forma de los reportes, por lo que una consideración importante es identificar dentro de la organización a los usuarios que cuentan con los mejores reportes, entonces éstos únicamente mejorarlos y personalizarlos para cada una de las áreas, cuando esto sea conveniente, además de que se encuentren alineados a las necesidades de información de la empresa y no solamente a la necesidad personal de usuario.



Capítulo Quinto



Equipo Data Warehouse



Equipo Data Warehouse

Todos para uno y uno para todos... Los tres Mosqueteros

Pensar en un equipo de ensueño para realizar cualquier actividad de la vida humana quizá resulte un poco difícil; ahora bien, imaginemos en un equipo de alto desempeño, como pensaríamos en un equipo de balón cesto, en donde será necesario contar con un armador, el cual es el responsable de repartir el juego para cada uno de los demás miembros, este armador debe saber que tipo de habilidades posee cada uno de los integrantes del equipo y sabrá en que momento utilizar esta habilidad. También, están los tiradores de campo, los cuales son los responsables de lograr la rapidez y exactitud del equipo, y por último y no por ello menos importante están los postes que son los encargados de recuperar los lanzamientos errados y de dar fuerza debajo del aro. Adicionalmente a ellos, se cuenta también con los miembros de la reserva, los cuales pueden entrar en cualquier momento a sustituir a cada uno de los jugadores. Por último, todos ellos son dirigidos por un entrenador, quien es el que sabe qué exigirle a cada uno de sus jugadores y sabe en qué momento debe hacer las adecuaciones necesarias para lograr la victoria; es decir, lograr un alto desempeño ...

Individuos trabajando independientemente no pueden completar exitosamente un proyecto de Data Warehouse. El equipo debe tener la correcta combinación de habilidades y experiencia, pero existen otros factores que hacen a un equipo de data warehouse. Principalmente deben tener metas comunes ya que no es suficiente con crear un modelo o diseñar una base de datos bajo el concepto del modelo estrella¹. Cada uno en el equipo debe ser exitoso en su trabajo y debe ser capaz de incorporar su trabajo en el proyecto. La mayor parte del éxito de un proyecto de data warehouse estriba en la forma en como el equipo trabajará y se esforzará por lograr la meta en común.

¹ El modelo estrella se describió en el Capítulo Primero de Conceptos Generales.



Para un buen desarrollo del proyecto de data warehouse es recomendable un equipo con una estructura de dos capas: la central y la complementaria. (Véase figura 9.1). En donde todos miembros del equipo tienen la responsabilidad del éxito del proyecto.



Figura 5.1 Equipo Central y Equipo Complementario








De esta forma, los miembros que conforman el equipo central son:

- 👤 Super Usuarios y usuarios
- 👤 Administrador del Proyecto
- 👤 Analista del Negocio
- 👤 Administrador de Datos
- 👤 Analista de Calidad de Datos
- 👤 Administrador de Base de Datos
- 👤 Desarrollador de aplicaciones de extracción transformación y carga
- 👤 Desarrollador de aplicaciones de Presentación



 Arquitecto Data Warehouse

Mientras que los miembros que conforman la parte complementaria son:

-  Servicio Técnico
-  Soporte a usuarios
-  Auditor y oficial de seguridad
-  Administrador de herramientas y Web
-  Patrocinador Ejecutivo
-  Otros Participantes
-  Desarrolladores Adicionales

A continuación, una breve descripción de las características y una definición para cada uno de los miembros del equipo de data warehouse

Definición para cada miembro de Data Warehouse

¿Qué tipo de gente debe integrarse para participar en el proyecto de data warehouse? ¿Qué perfil debe tener cada miembro del equipo? Esta selección no es tan solo identificar a una persona en el total de la organización; sino que también es la forma en que a cada uno lo tienen considerado en la organización. La renuencia a asignar niveles y funciones apropiadas puede ser indicio de una organización con falta de compromiso hacia el data warehouse. El nivel y función correcta manda un mensaje al resto de la organización, indica la importancia relativa del proyecto; en algunos casos, el nivel y función compensa el esfuerzo de la organización y de los miembros del equipo. Los niveles y funciones deben escogerse detenidamente, no debe exagerarse la posición pero debe ser descriptivo de las funciones y responsabilidades de desempeñará cada miembro del equipo.

No necesariamente los papeles descritos en este trabajo son los definitivos, es recomendable adecuar los niveles y las funciones de acuerdo al proyecto de data warehouse que se planea implementar. El punto importante es que las responsabilidades listadas para cada papel son asignadas a alguien en el equipo de data warehouse. Las siguientes responsabilidades son obligatorias para un proyecto exitoso.

Patrocinador Ejecutivo

El Patrocinador Ejecutivo es la persona que tiene la necesidad para el Data Warehouseⁱⁱ. Este es generalmente el director del área, el cual reconoce el valor de la toma de decisiones apoyadas en un data warehouse. El patrocinador ejecutivo será el principal impulsor de los nuevos requerimientos, estará enterado de cómo éstos afectarán la planeación, el presupuesto y los requerimientos que ya han sido aceptados. Esta es la persona que proporciona el presupuesto y apoyo político para el proyecto. El patrocinador ejecutivo asegura la disponibilidad de la gente de negocio que puede articular los requisitos y proporcionar apoyo a través de la vida del proyecto. El patrocinador del negocio debe ser también el "Champion" para el proyecto de data warehouse, él debe promover el proyecto cuando tenga oportunidad.

Super Usuarios y Usuarios

Los super usuarios vienen generalmente del lado del negocio. Tienen una relación muy cercana con los usuarios; hablan el mismo idioma, son creíbles y tienen la verdad de los usuarios así. Algunas organizaciones no tienen este papel, entonces este papel lo desempeñan los usuarios. Estos individuos tienen muchísima

ⁱⁱ El Patrocinador ejecutivo no deberá ser una persona del área de sistemas, ya que la necesidad del data warehouse es del las áreas usuarios, esto ayudará a un mayor involucramiento de los usuarios y de la organización para un buen desempeño y logro.

visibilidad completa sobre el ambiente del negocio, ayudan de enlaces entre los usuarios para establecer la verdadera definición de datos y el tipo de datos que se necesitan, tipo de formato necesario y si los datos deben ser detallados o resumidos o ambos. Ellos saben sobre qué seguridad y cuál autorización es necesaria en el maestro de datos.

Los super usuarios monitorean la calidad de los datos, están altamente involucrados en determinar los requisitos de los usuarios y los ayudan a establecer prioridades.

Los super usuarios y usuarios identifican qué datos necesitan, cuan oportuno deben estar y de donde se originan los datos. Los usuarios especifican que información necesitarán y en qué tipo de formato. Determinan la necesidad para datos históricos incluyendo el período y cómo respaldar la información histórica.

Con la ayuda de la Administración de Datos, los usuarios proporcionan la definición del negocio en términos de negocio y no en terminología tecnológica. Estas definiciones pueden almacenarse en un repositorio metadata y deben estar disponibles para que cualquiera usuario corra una consulta o reciba el resultado de algún informe. Muchas organizaciones ponen las definiciones en su Intranet, dando acceso a esas definiciones a toda la organización. El autor de la definición así como de la última fecha de actualización está disponible con la definición. La Administración de datos debe mantener las definiciones al corriente y de desempeñar los mecanismos de actualización.

Los usuarios son dueños de los datos, es decir, que ellos son responsables de determinar quien puede acceder a los datos y en que nivel de detalle. El acceso se puede dar para datos resumidos pero no para detallados. Los dueños son responsables de identificar los dominios (valores válidos) de los datos.

Analista del negocio

El analista del negocio es generalmente del área de tecnología de información pero está muy informado sobre las actividades del negocio tales como sus procesos y datos operacionales que soportan al negocio. El analista del negocio reúne en un amplio rango las habilidades de los usuarios así como sus requisitos tradicionales en la toma de decisiones. Si el analista del negocio es muy bueno, será capaz de ayudar a la gente del negocio a identificar posibilidades adicionales y casar las aplicaciones, lo que dará a la organización una ayuda significativa en la inversión, en la reducción del costo y en la satisfacción del cliente. El analista del negocio facilitará la discusión, guiará a la gente del negocio para anticiparse a las necesidades y oportunidades futuras.

El analista del negocio organizará y promoverá las sesiones con diferentes grupos de usuarios haciendo talleres de trabajo con lluvia de ideas, explorando opciones, discutiendo enfoques, tomando decisiones de cuáles datos y aplicaciones de toma de decisión serán benéficas.

Soporte a Usuarios

El soporte a usuario es la primera línea de la defensa cuando los usuarios tienen problemas, deben estar orientados al cliente, saber que las llamadas de los usuarios no son molestias. Deben estar informados sobre las herramientas, entender la estructura de la base de datos, saber de los datos, estar familiarizado con preguntas e informes ejecutados frecuentemente. También, deben reconocer los problemas de los usuarios y estar familiarizados con preguntas y respuestas frecuentes. Deben tener un perfil de usuario y ser reconocido por los usuarios, ser paciente y sensible y estar informado del progreso de la resolución.

En algunas organizaciones, el personal de soporte a usuarios es el que crea parte de las consultas y de los informes, además administra las bibliotecas que almacenan estas preguntas e informes. La administración incluye el desarrollo e implementación de los estándares para la admisión a la biblioteca de las consultas y el mantenimiento de los reportes. Las consultas y los informes enviados a la biblioteca deben tener un uso importante, no sólo servir como un simple reporte más de la organización, deben estar bien documentados y probados, inclusive probar cualquier modificación. Toda la información debe comunicarse a las partes interesadas con una explicación del propósito de la consulta, como puede ser utilizada, tiempo de respuesta y expectativas de uso de recursos, tiempo de respuesta y que se esperara del reporte o resultado.

También, dependiendo de los recursos y la planeación en la empresa, al área de soporte a usuarios se les asigna el entrenamiento, algunas veces para organizar las clases y otras para administrar a los expertos que imparten los cursos y de auto estudio.

Administrador de Datos

Se puede decir que un administrador de datos obtuvo una segunda oportunidad con el data warehouse. Los papeles y beneficios de la administración de datos no han sido entendidos por la mayoría de las compañías, ahora el papel de administración de datos fue relegado a elaborar documentación que raramente es usada. Con la llegada del data warehouse la administración de datos es ahora una posición clave para el desarrollo de un data warehouse exitoso.

Los administradores de datos modelan los datos del negocio según las normas y reglas de la propia compañía, documentan y mantienen el mapeo entre el modelo

lógico de datos y el diseño físico de la base de datos, entre los datos fuente y los datos objetivo o finales del data warehouse. Los administradores de datos conocen los archivos fuente, saben cuales son los archivos fuente apropiados para la extracción, establecen y mantienen los estándares, comunican la disponibilidad de los datos, además sugieren que departamentos pueden compartir sus datos. El administrador de datos es responsable de administrar las herramientas usadas en el modelado de datos, así como la biblioteca de herramientas, aunque en algunas organizaciones esto es responsabilidad del equipo del desarrollo de aplicaciones.

La administración de los datos es responsable de controlar el repositorio metadata, el cual vendrá de distintas fuentes: Herramientas de modelado, herramientas de extracción-transformación-cargaⁱⁱⁱ, herramientas de consulta, reingeniería de datos de archivos fuente y de herramientas de entrada directa a un repositorio metadata. El repositorio metadata podrá actualizarse, así como también ser accesado por cada uno de estos sistemas fuente. Los controles son necesarios para disminuir la corrupción del metadata. La administración de datos es responsable de controlar nuevas entradas, actualizaciones y borrado de datos.

Desarrollador de Aplicación

El desarrollo de aplicaciones se divide en dos grupos con diferentes habilidades:

1. El aplicaciones de extracción-transformación-carga (ETL) en donde necesitan participar programadores con alta experiencia en lenguajes de programación (Cobol, C + +, etc.)
2. Aplicaciones de presentación final quienes necesitan conocer de lenguajes cliente / servidor y lenguajes más actuales (Visual Basic, Power Play Java,

ⁱⁱⁱ En el mundo del data warehouse son mejor conocidas como herramientas ETL, que significa Extract, Tranform and Load.

etc.), herramientas de la consulta, escritorio de reportes y herramientas OLAP.

Las aplicaciones de presentación final son responsables del proceso de la adquisición que constituye el mayor esfuerzo para cualquier data warehouse, el proceso de extracción-transformación-carga. Se ha estimado que esta tarea toma entre un 70 a 80 por ciento del tiempo y esfuerzos totales para implementar un data warehouse.

Las aplicaciones de presentación final juegan un papel importante en la entrega de datos del data warehouse, generalmente se implican fuertemente con los usuarios.

Auditor de Seguridad

En el ambiente de sistemas, la seguridad es cada vez más importante debido a gente fuera de la compañía tiene acceso a ciertos reportes e informes, e incluso la habilidad de realizar consultas directamente al data warehouse. Este acceso se otorga a los proveedores con la finalidad de mejorar el suministro de la materia prima en tiempo^{iv}. También tienen acceso los clientes comerciales para tenerlos tan de cerca como sea posible de los productos y servicios que la empresa produce. Con el acceso permitido de estas compañías externas, la seguridad debe ser exacta y comprometida^v.

El encargado de seguridad debe trabajar muy de cerca con personas que tienen que ver con la administración de seguridad (administrador Web, administrador de bases de datos, administrador de herramienta de consulta, administrador de

^{iv} En el capítulo de El Data Warehouse y la Cadena de Valor, se hace una descripción más detallada de este punto de suministros.

^v El acceso se realiza a través de la Internet, si se considera que existen los expertos en ingresar a los sistemas de otras compañías.

repositorio, administrador de herramienta CASE y administrador de herramienta de extracción, transformación y carga) para comprender las capacidades de los productos y determinar el enfoque óptimo de seguridad a establecer. Esto podría incluir exposiciones y dificultades en la administración. Este punto no debe convertirse en la pesadilla burocrática de la administración de la productividad.

El principal papel del encargado de seguridad es la de identificar recomendaciones y acciones para tapar hoyos relativos con la seguridad.

Administrador de Base de Datos

El administrador de bases de datos (DBA) es el responsable del aspecto físico del data warehouse. El diseño físico incluirá, desempeño, respaldos y recuperación. Antes de iniciar con el proyecto, el DBA debe entender los requerimientos básicos de los usuarios y de cómo se deberán acceder al data warehouse.

El DBA del data warehouse está generalmente más implicado con los usuarios ya que controla y ajusta constantemente las bases de datos del data warehouse y agrega los índices, sumariza las tablas y monitorean el SQL de las consultas para mejorar el desempeño individual de las consultas.

La administración de la base de datos algunas veces es responsable de planificar de la capacidad y del diseño físico para el buen desempeño, controlando tiempos de respuesta y recursos utilizados (unidad de procesamiento central, el Discos) evaluando el desempeño de los problemas, ajustando las bases de datos y revisando el estatus del SQL escrito por desarrolladores y super usuarios.



Servicio Técnico

El área de servicio técnico, algunas veces llamado administración del sistema, son los responsables de establecer la arquitectura técnica del data warehouse. Esto incluye decisión acerca del tipo hardware, red y sistema operativo; aunque algunas de estas decisiones puedan ser del arquitecto del data warehouse. El servicio técnico debe desarrollar la capacidad de planear el escalamiento en tamaño de data warehouse y evitar que el tamaño llegue a ser más grande de lo que originalmente se planeó, los data warehouses a menudo crecen mucho más de lo que se había planeado. El servicio técnico debe monitorear el desempeño de extracción, transformación y carga, así como de consultas de alto nivel. El servicio técnico desarrollará plan maestro de contingencias en línea con el data warehouse.

Principalmente en México, muchas empresas que invierten fuertemente en tecnología de información no tiene considerados planes de contingencia. Después de los sucesos del 11 de septiembre de 2001 en Nueva York, EUA, muchas compañías se han dado cuenta de la importancia de contar con planes de contingencia en caso de desastres.

Arquitecto Data Warehouse

El arquitecto del data warehouse es la persona que desarrollará la arquitectura que incluiría herramientas de data warehouse y determinará la forma en que trabajarán las interfases entre las distintas herramientas. También, si un data warehouse deberá ser usado, cómo y por quién será utilizado. Si los operational data stores deberán ser utilizados, especificará cómo y por quién serán accesados y cómo este alimentará al data warehouse o al data mart. El arquitecto está comprometido fuertemente con el entendimiento y determinación de los datos fuentes. El

arquitecto determinará cuidadosamente como se integrará el data mart, además, especificará si el data warehouse será de dos o tres niveles.^{vi}

El arquitecto del data warehouse establecerá los estándares y procedimientos para el data warehouse. Una parte importante de estos estándares podrán servir de base para el proceso ETL y determinarán el flujo, tipo de herramientas a utilizar establecerá las responsabilidades para el proceso, inspecciones y verificaciones, procesos de prueba y buscará la aceptación por parte de los usuarios. Los estándares y procedimientos podrán incluirse al uso de datos externos e incorporarlos en el proceso ETL.

Gerente del Proyecto Data Warehouse

El Gerente del proyecto de data warehouse tiene las más completa de las responsabilidades para lograr una implementación exitosa. El Gerente del proyecto define, planea, agenda y controla. El plan del proyecto debe incluir las tareas, entregables y la gente que realizará las tareas, los recursos materiales y de espacio. El Gerente controlará y coordinará las actividades del equipo, además revisará sus entregables. Si se utilizan consultores o especialistas, el gerente del proyecto asignará tareas, actividades de monitores, entregables y se asegurará de la transferencia de conocimiento.

El gerente del proyecto estimará los costos del proyecto, deberá controlar el presupuesto; también, medirá el beneficio y eficiencia del data warehouse.

La tarea más importante para el gerente del proyecto es la selección de la gente correcta, la gente con las habilidades necesarias, que trabajen con los usuarios y con cualquier otro miembro. Para hacer una selección correcta, se requiere de una

^{vi} Véase diagramas de arquitecturas de data warehouse en el apéndice de este trabajo.

comprensión de la importancia del proyecto para un buen proceso de selección. El problema es que la gente buena está asignada a otros proyectos importantes en la compañía, pero un gerente hábil no se rinde en la selección y hasta utiliza recursos externos de ser necesarios. El contratar a recursos externos implica riesgos tales como tiempos por demora, o poco compromiso al proyecto. Un recurso externo necesitará estar familiarizado con la organización, para entender como se hacen las cosas en la empresa, entender el campo de acción y aprender sobre de los datos fuente.

Analista de Calidad de Datos

No hay que menospreciar el puesto de analista de calidad de datos, nadie tiene más responsabilidad para la calidad en el data warehouse que este personaje. La calidad de datos es trabajo de todos, pero es el único dedicado a esta tarea es el analista de calidad de datos. Aunque a menudo los problemas de calidad de datos a se originan en el departamento de tecnología de información, algunas veces el problema proviene de los usuarios. El analista de calidad de datos es responsable de encontrar e informar problemas relacionados a la calidad de los datos, rastrear estos problemas y asegurarse de la resolución. Los problemas encontrados deben ser informados al administrador de datos, en donde se pueden incluir apropiadamente en el modelo lógico de datos. El analista de calidad de datos puede escribir las especificaciones que programan para la transformación lógica que es necesaria durante el proceso de ETL.

El papel más importante para el analista de calidad de datos será evaluando y optimizando los procesos en el data warehouse. Específicamente, en el proceso de ETL. El analista de calidad de datos debe buscar fuentes externas más limpias de datos para reemplazar a las que ni funcionan adecuadamente.

Administrador de Herramientas

El administrador de la herramienta tiene la responsabilidad de tratar con el vendedor de las herramientas y asegurando un soporte excelente. El administrador de herramientas de consultas trata con problemas asociados a las herramientas, inclusive consultas que produce resultados y consultas poco exactas. El administrador de herramienta de consulta tiene la responsabilidad de asignar palabras secretas junto con el responsable de seguridad, cerciorarse de que el ambiente de herramientas de consultas está protegido apropiadamente. También, es el responsable de las actualizaciones.

Administrador Web

El administrador Web es una función nueva en tecnologías de data warehouse y a veces resulta complicado hacer entender a las organizaciones la participación de esta figura.

El administrador Web tiene la responsabilidad de establecer y ambientar un lugar en donde un conjunto de resultados pueda distribuirse, los reportes puedan estar disponibles, en donde las consultas puedan ser puestos en la Web. El administrador Web trabajará en la implementación de herramienta de consultas en la Web, será el responsable de asegurar la seguridad en la Web principalmente del acceso no autorizado.

Consultores y Recursos Externos^{vii}

Los consultores y Recursos Externos a menudo son introducidos para apoyar a la organización al inicio del proyecto para desarrollar la arquitectura, los estándares y

^{vii} Los recursos externos son los profesionales que no pertenecen a una compañía específica y que prestan sus servicios de manera independiente, su costos son más bajos que los de una empresa formal de consultoría

proporcionar habilidades que no se tienen en la organización. Los recursos externos, a menudo son retenidos para soportar una parte de algún software o para entrenar el personal de tecnología de información y a los usuarios en el uso de las herramientas.

Los consultores y los recursos externos son costosos, por lo que las empresas deben usarlos efectivamente. Los consultores deben ser utilizados como mentores, ya que poseen conocimiento y experiencia que las organizaciones carecen. El consultor debe ser utilizado sí en la empresa no se cuentan con muchos recursos o las habilidades son limitadas.

Comités

Los comités deberán tomar decisiones críticas en la dirección del data warehouse. La ausencia de tal comité generalmente significa que no existe un alto nivel de interés en el data warehouse. Normalmente esto implica que el proyecto de data warehouse falle ya que no solamente es un esfuerzo de tiempo; es progresivo.

Comité Técnico

El comité técnico coordina y autoriza los recursos de tecnología de información. Este comité hará las selecciones difíciles acerca de selecciones del personal involucrado en el proyecto. El comité técnico considerará y aprobará cambios en materia de presupuestos y fechas de entrega. Siempre habrá asuntos para ser resueltos entre los grupos de la tecnología de información y siempre habrá proyectos tal como implementación de estándares, conflictos de proyectos, proyectos redundantes, acceso a ciertos datos, asuntos de seguridad y a disponibilidad de equipos y personal.

El comité técnico aprobará la arquitectura del data warehouse, inclusive componentes claves de hardware y software, infraestructura y estándares.

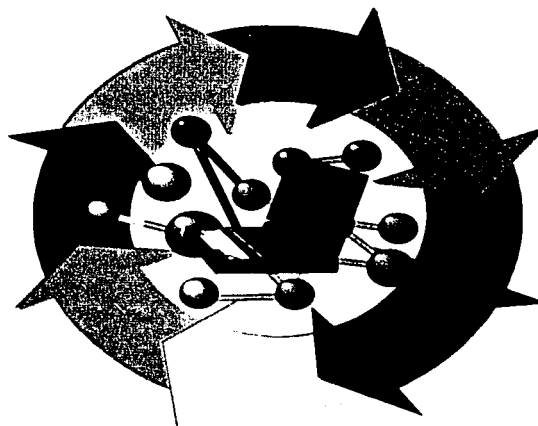
Comité Funcional

El papel más importante de los comités es el de dar prioridad a los proyectos. Frecuentemente se contarán con más candidatos para proyectos de data warehouse que los recursos para implementarlos. Existen algunos proyectos que son obligatorios, tales como los que son establecidos por el gobierno o por ley. La mayor parte de los proyectos de data warehouse no son obligatorios pero se justifican ya que proporcionan una mejora competitiva con el resto de las empresas. El Comité Funcional debe analizar el costo - beneficio y el significado político de cada proyecto.

El comité funcional ayudará a establecer, o al menos a aprobar los objetivos de data warehouse y los medidores críticos de éxito. Aprobarán los presupuestos, revisarán los acuerdos del proyecto de alto nivel, los cambios de esos acuerdos y revisarán los principales momentos y entregables clave del proyecto.



Capítulo Sexto



El Data Warehouse y la Internet



El Data Warehouse y la Internet

Combinar dos tipos diferentes de conceptos resulta interesante, el Internet y el Data Warehouse combinado pueden ofrecer ventajas competitivas a las empresas hoy en día, pero hay que vencer el miedo al Internet y a invertir en un sistema de data warehouse ...

Siempre que veo una telaraña me gusta observar la forma y estructura que tienen, a veces tomo un aspersor de agua y las mojo para poder ver como el rocío del agua me deja ver con mejor claridad la forma que tienen las telarañas.



*L*os ejecutivos deben poner mucha atención en como sus compañías generan valor en ambos mundos, tanto en el mundo físico como en el virtual. Sin embargo, el proceso de generación de valor no es el mismo en ambos mundos. Para entender las diferencias y la interrelación entre el proceso del valor agregado del mundo físico y del mundo de la información, los administradores deben ver más clara y conjuntamente los asuntos estratégicos que enfrentan sus organizaciones.



Administrar dos procesos interactivos de valor agregado en dos reinos mutuamente dependientes implica nuevos conceptos y retos tácticos. Solamente aquellos que entienden cómo dominar a ambos reinos pueden crear y extraer valor de una manera más eficiente y efectiva.

Académicos, consultores y administradores, han descrito largamente el proceso de crear valor en el mundo físico. La mayoría se refiere a estados que envuelven vínculos en una "cadena de valor"¹.

La cadena de valor, es un modelo que describe una serie de actividades que adicionan valor, conectadas del lado del proveedor de la compañía (materiales en bruto, procesos logísticos de entrada y procesos de producción) con su otro lado de la demanda (procesos logísticos de salida, mercadotecnia y ventas).

Mediante el análisis del estado de la cadena de valor, los administradores han sido capaces de rediseñar su proceso tanto interno como externo para mejorar eficiencia y efectividad. El modelo de cadena de valor, trata la información como un elemento adicional que apoya los procesos de valor agregado, no como una fuente de valor en sí misma. Los administradores en ocasiones utilizan la información que ellos capturan en inventarios, producción o logística para ayudar a monitorear y controlar estos procesos. Pero ellos muy raramente, utilizan la información en sí misma para crear un nuevo valor para el consumidor.

Para crear el valor con información, los administradores deben buscar el *marketplace*. Si bien, la cadena de valor del espacio puede proyectar en que lugar compradores y vendedores puedan transferir fondos mediante redes electrónicas exactamente como lo hacen en efectivo. Las compañías emplean sus procesos de valor agregado para

¹ Véase Capítulo Cuarto del Data Warehouse y la Cadena de Valor.

cambiar la ruta de la información hacia un nuevo mercado de servicios y productos únicos para el mundo de la información, en otras palabras las etapas del valor agregado son *virtuales* y ellas están desarrolladas a través y con información.

Crear valor en cualquier etapa de la cadena de valor virtual involucra una secuencia de cinco actividades: integrar, organizar, seleccionar, sintetizar y distribuir información. Exactamente como tomar materiales en bruto y transformarlos en algo útil como la secuencia de tareas que involucra el ensamble de un automóvil o la producción en línea es así como un administrador hoy recolecta información en bruto y le adiciona valor mediante estos pasos.

Adaptación a un Mundo Virtual

Este también es un reflejo de la información de las actividades que tradicionalmente han ocurrido en las etapas del mundo físico una etapa de la cadena de valor virtual paralela a la etapa de la cadena de valor física. Cada actividad es una etapa en la cadena de valor virtual que ocurre a través y con información para proyectarse en las etapas del mundo físico.

En el mundo virtual ésta es una fuente potencial de nuevos ingresos. Por otra parte, la información presenta oportunidades para desarrollar nuevas relaciones con los consumidores a muy bajo costo.

Los administradores de la empresa deben continuar atentos a la cadena de valor física produciendo y vendiendo pero ellos también deben construir y explotar la cadena de valor virtual. Existen estudios que registran una variedad de empresas que intentan hacer negocio en ambos, el lugar y el espacio y hay organizaciones que

hacen dinero de la exitosa explotación de la información de ambos para su cadena de valor.

Más que administrar una serie de procesos de valor agregado, se están administrando dos actualmente. La lógica económica de las dos cadenas es diferente, un entendimiento convencional de la economía de escala y el enfoque no aplica a la cadena de valor virtual en el mismo camino como lo hace en la cadena de valor física; sin embargo, estas dos cadenas deben ser administradas de manera diferente pero en armonía, procurando no descuidar la relación intrínseca que existe entre las cadenas.

Algunas empresas adoptan procesos de valor agregado en tres etapas. En la primera etapa "**Visibilidad**" las empresas requieren una habilidad para ver sus operaciones físicas más eficientemente a través de la información, en esta etapa los administradores utilizan sistemas tecnológicos de información a gran escala para coordinar actividades en sus cadenas de valor físico y los procesos que gobiernan la fundación de una cadena de valor virtual.

En la segunda etapa, "**Proyección de la Capacidad**" las empresas sustituyen actividades virtuales por unas físicas; comienzan a crear una cadena de valor paralela en el mercado. Finalmente los negocios utilizan la información para establecer nuevas relaciones con los consumidores.

En una tercera etapa, los administradores diseñan un flujo de información en sus cadenas de valor virtual para entregar valor a los consumidores en una nueva forma. En efecto, aplican las actividades genéricas de valor agregado a su cadena de valor virtual; sin embargo explotan una "**Matriz de valor**".



De la manera como las empresas se mueven en el mundo de la información para desarrollar sus etapas de valor agregado, el límite de su potencial de crecimiento se incrementa. Cada una de las tres etapas representa una considerable oportunidad para las empresas y su gente.

Visibilidad

Durante años, muchas empresas han invertido en sistemas tecnológicos para permitir a los administradores, coordinar, medir y algunas veces controlar los procesos de negocios. La información recolectada por estos sistemas acerca de las etapas en la cadena de valor, ha ayudado a los administradores a planear, ejecutar y evaluar resultados con gran precisión y rapidez, en otras palabras, la información tecnológica ha permitido a los administradores ver sus operaciones más efectivamente a través del mundo de la información.

Recientemente, los administradores han sido capaces de ganar accesos a la información generada en el curso de las actividades tradicionales. Esta información les ha ayudado en las cadenas de valor físicas como sistemas integrados, más que como un grupo discreto de actividades fuertemente relacionadas. En este sentido, ellos pueden ganar nuevas perspicacias en la administración de las cadenas de valor como un todo más que en una recolección de partes.

En empresas tales como Wal-Martⁱⁱ y Frito Lay han transformado sus tipos de visibilidad en ventajas competitivas. La exitosa utilización de sistemas de información de clase mundial para cada una estas compañías es ahora su conocimiento común.

ⁱⁱ Wal Mart es una de las primeras empresas en implementar sistemas de Data Warehouse para explotar su cadena de valor, uno de los artículos que iniciaron esta revolución es el estudio que se obtuvo de observar que los hombres que normalmente compraban pañales, también se llevaban un paquete de cervezas, por lo que la compañía acomodó los pañales frente a las cervezas y el resultado fue un incremento en ventas de ambos productos.

Sin embargo, consideremos un ejemplo - Frito Lay - desde la perspectiva del *marketspace*. El logro de Frito, con su ampliamente publicitada "revolución informativa" ilustra la necesidad de la iniciativa que deben de tener las empresas para dar sus primeros pasos si ellas están estableciéndose en la explotación de su cadena de valor virtual.

Este es un sistema nervioso central dentro del negocio que comprende: mercadotecnia, ventas, manufactura, logística y finanzas, y también proporciona a los administradores información de los proveedores, clientes y competidores.

Los empleados del área de recolección de información de ventas diarias de Frito, recorriendo tienda por tienda. Los empleados también recolectan información acerca de las ventas y promociones de los productos de los competidores o sobre los nuevos productos lanzados en localidades selectas. La combinación de estos datos con la información de cada etapa de la cadena de valor, permite a los administradores de Frito determinar mejor los niveles de provisión de entradas de materia prima, asignar las actividades de manufactura de la empresa en función de la capacidad de producción y planear rutas para una mayor y más eficiente cobertura de las áreas de mercado. La compañía así puede orientarse a la demanda local con promociones directas, lo que significa la optimización continua de márgenes al enfrentar inventarios de riesgo.

En corto, Frito puede utilizar la información para ver y reaccionar a actividades sólo de su cadena de valor física. La empresa ejecuta acciones en el *marketplace*, mientras ésta monitorea y coordina también acciones en el *marketspace*.

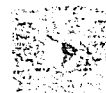


Proyección de la Capacidad

Algunas empresas han establecido la infraestructura necesaria para tener "**Visibilidad**", también hacen más exacto el monitoreo de las etapas de creación de valor agregado. Pueden llegar a manejar su operación y aún instrumentar etapas de valor agregado en el *marketspace* - más rápido, mejor, con más flexibilidad y a más bajo costo. En otras palabras, los administradores pueden comenzar a preguntarse ¿Qué estamos haciendo ahora en el *lugar*, y que podríamos hacer más eficiente o más efectivamente en el *espacio*? ¿Qué etapas del valor agregado actualmente desarrolladas en la cadena de valor físico pueden ser proyectadas al mundo de la cadena de valor virtual? Cuando las compañías mueven actividades del lugar al espacio, comienzan a crear cadenas de valor virtual que paralelamente mejoran las cadenas del valor físico.

Cada administrador sabe que ser competitivo hoy, depende de lograr altos niveles de desarrollo para el cliente, mientras se incurre en bajos costos de investigación, desarrollo y de producción. Tradicionalmente las compañías han explotado más o menos amplias economías de escala en producción, mientras se enfocan a incrementar sus niveles de calidad.

Los fabricantes de automóviles han logrado esta estrategia, entregan productos altamente diferenciados al más bajo costo posible. Cuando las economías de escala no aplican, como en muchos negocios del sector de los servicios, los administradores buscan mejores desarrollos a bajos costos que puedan tener proyección mundial en los cuales, las economías son totalmente diferentes. En la cadena de valor virtual, las empresas pueden encontrar dramáticas disminuciones de costos para entregar extraordinarios resultados de alto valor a los consumidores.



Nuevas Relaciones con los Consumidores

Últimamente, las empresas deben hacer más que crear valor en el espacio, deben sacar valor de él. Muchas veces hacen esto para establecer relaciones de base espacial con los consumidores. Una vez que las compañías llegan a ser expertas en la administración de sus actividades de valor agregado de un lado a otro de las cadenas de valor paralelas, ellas están realmente desarrollando estas nuevas relaciones.

En el mundo de la alta tecnología, los ejemplos de construcciones de relaciones con los consumidores en la cadena de valor virtual abundan. Hoy miles de empresas han establecido lugares en la *World Wide Web* para anunciar productos o para obtener comentarios de los consumidores. Algunas empresas han ido más allá, actualmente han automatizado la interfaz con los consumidores. De esta manera, identifican y satisfacen deseos de los consumidores a más bajo costo.

Una empresa de bases de datos, ahora distribuye sus productos vía la Internet tal y como lo hace a través de sus canales físicos. Esta empresa está vinculándose a las nacientes firmas catalogadas dentro de las de más alta tecnología en los negocios y que han llegado a comercializar vía la Internet.

Otras compañías ven como sus principales retos, el cómo administrar las relaciones individuales con sus clientes tanto en el mercado real como en el mercado virtual. Aquellos que han tenido éxito, tienen la oportunidad de reinventar el valor esencial de sus negocios, y aún más, de la industria entera.



La Matriz de Valor

Las nuevas relaciones que las compañías están desarrollando para sus clientes surgen de una matriz de evaluación de oportunidades. Cada etapa de la cadena de valor virtual como una proyección de la cadena física de valor permite nuevos extractos del flujo de información y cada extracto puede constituir un nuevo producto o servicio. Si los administradores quieren continuar con cualquiera de estas oportunidades, necesitan colocar procesos de análisis de información, organizarla para los consumidores, seleccionar lo más valioso, integrarlo (o sintetizarlo) y distribuirlo los cinco pasos para dar valor agregado al mundo de la información.

Estos pasos de valor agregado, en conjunción con las cadenas de valor virtual hacen una matriz de valor que permite a las compañías identificar más efectivamente los deseos de los clientes y satisfacerlos más eficientemente.

La compañía analizará, organizará, seleccionará, sintetizará y distribuirá información sobre el diseño del producto en proceso de investigación y desarrollo para crear un modelo de simulación en computadora de los consumidores, quiénes podrán entonces entrar en el espacio de diseño virtual y dar retroalimentación lo cual podrá ser usado para dar un valor agregado al diseño abierto de un nuevo producto.

Mientras la información utilizada en tales productos también ayuda a los procesos físicos y alimenta a puntos físicos finales - un automóvil, un disco compacto, una política de seguros, también puede ser el insumo material para nuevas clases de valor.



La industria del periódico es otro ejemplo de como tales procesos pueden cambiar del lugar al espacio. Los ejecutivos pueden aplicar los cinco pasos del valor agregado para cada vínculo de la cadena de valor virtual hacia una visión de una matriz de oportunidades de creación de valor. Por ejemplo, del flujo de información utilizada para apoyar los reportajes y las editoriales, el periódico puede suministrar paquetes de información a lectores con archivos de audio de entrevistas de reporteros, imágenes de sus cuadernos de notas, que no han sido publicadas en papel y aún comentarios editoriales acerca de bosquejos de historias. La matriz de evaluación guía a los administradores hacia como establecer los procesos necesarios para explotar nuevas oportunidades.

Implicaciones para la Administración

Lo que todo esto significa para los administradores, es que ellos deben enfocarse conscientemente a los principios que guían la creación de valor y la extracción a través de las dos cadenas de valor, tanto separada como combinadamente. Estos dos procesos de valor agregado son fundamentalmente diferentes. La cadena de valor física está compuesta de una secuencia lineal de actividades con puntos definidos de entrada y salida. En contraste, la cadena de valor virtual es no lineal, es una matriz de entradas y salidas potenciales que pueden ser ingresadas y distribuidas a lo largo de la amplia variedad de canales.

¿Cómo podemos sensibilizarnos de estas nuevas actividades? - el espacio de información que permite la creación de una cadena de valor virtual y la explotación de una matriz de valor. Para tener éxito en este nuevo ambiente económico, los ejecutivos deben entender las diferencias entre creación y extracción de valor en el *marketplace* y en el *marketspace*; ellos deben manejar ambos efectiva y armónicamente. Más específicamente, los ejecutivos de una compañía deben abarcar

volúmenes de datos guiados por premisas más actualizadas, porque en el *marketspace* muchos de los axiomas de negocios que han guiado a los administradores, ahora ya no aplican.

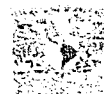
Activos Digitales.

A diferencia de los físicos, los activos digitales no se agotan al consumirse. Las compañías que crean valor con activos digitales pueden ser capaces de volver a cosecharlos a través de un número infinito de transacciones potenciales. Esto cambia la dinámica competitiva de las industrias. Hay proveedores de producto o servicios que deben poner sus precios de acuerdo al tradicional modelo de costos variables basado en el consumo de materiales esenciales habrá tiempos difíciles de competencia contra las compañías que, por explotar sus cadenas de valor virtual pueden ofrecer precios agresivos o mantener márgenes inmóviles.

Nuevas Economías de Escala.

La cadena de valor virtual redefine economías de escala permitiendo que pequeñas compañías logren bajos costos unitarios por productos y servicios en mercados dominados por empresas grandes.

Las nuevas economías de escala hacen posible proveer lo que es necesario, en efecto, frentes de minitiendas para todos y cada uno de los consumidores. Si millones de usuarios requieren del servicio, este puede darse en el momento y exactamente para cada uno.



Nuevos Alcances Económicos.

En el *marketspace*, los negocios pueden redefinir nuevos alcances económicos por el diseño o simplemente por agrupar activos digitales que proporcionan valores a través de muy diferentes y dispares mercados.

Reducción de Costos de Transacción.

Los costos de transacción en la cadena de valor virtual son más bajos que sus contrapartes en la cadena de valor físico, ellos continúan declinando bruscamente tal es el caso de la capacidad de procesamiento por unidad de costo de micro procesos dobles cada 18 meses.

Reequilibrar Provisión y Demanda.

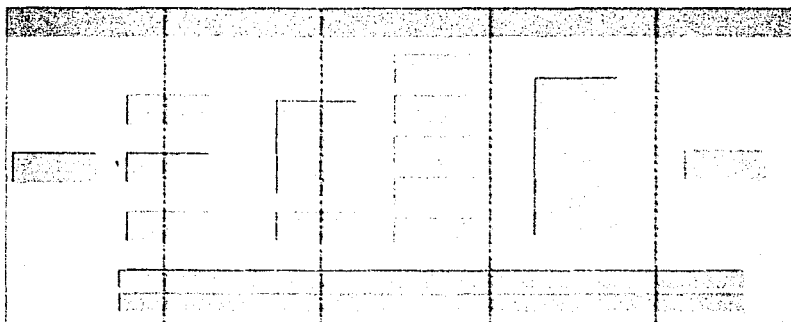
Al combinar estos cuatro axiomas para crear un quinto, el mundo de los negocios aumenta el reclamo de un cambio de pensamiento del lado del proveedor para colocarse del lado del demandante. Como las compañías analizan, organizan, seleccionan, sintetizan, distribuyen información en el *marketspace* mientras administran materias primas y manufacturan bienes en el *marketplace*, ellas tienen la oportunidad de "sentir y responder" a los deseos de los consumidores de manera más simple que manufacturar y vender productos y servicios.

En el mundo de hoy de sobrecapacidad, en el cual la demanda, no-provisión, es escasa, los administradores deben buscar incrementar estrategias del lado de la demanda. La alta administración debe valorar sus negocios - sus fortalezas y debilidades, sus oportunidades y riesgos - a lo largo de las cadenas de valor de ambos

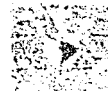
mundos, el virtual y el físico. Hoy, eventos en cualquiera de los dos pueden hacer nacer o quebrar un negocio.



Capítulo Séptimo



Metodología



Metodología

Seguir un sólo camino es retroceder... Igor Stravinsky

HOT CAKES

2 tazas de harina, 2 cucharaditas de Royal

½ cucharadita de sal, 2 huevos

2 cucharaditas de azúcar, 1 cucharadita de Vainilla

Leche la necesaria

Se cierne la harina con el royal, sal y azúcar...

Al igual que en una receta de cocina, una metodología cumple con guiar a la persona en las actividades que debe realizar, y como dice la receta “leche la necesaria” al aplicar cualquier metodología, esta debe adaptarse a las necesidades del proyecto y de la organización, es decir, leche la necesaria.

*P*ara el desarrollo adecuado de un data warehouse es necesario que se seleccione una metodología de implementación adecuada, en el mercado existe una gran variedad de metodologías, a su vez todas se presentan como las mejores para la implementación, aunque no necesariamente sea así. Para iniciar el proceso de evaluación de implementación de data warehouse es necesario que una de las interpretaciones que se tiene en México del Concepto de Data Warehouse es que es mucha la inversión y que los resultados no son tan inmediatos como se esperaría de una inversión de este tamaño. Desarrollar una metodología de implantación no es la finalidad de este trabajo, pero sí lo es dar un poco más de visión para una buena selección de una metodología, por lo que me enfocaré a las principales etapas y que hay que observar en cada una de ellas.

Una buena metodología de data warehouse deberá incluir al menos de las siguientes fases:

Fase Cero	Evaluación de Factibilidad
Fase Uno	Análisis
Fase Dos	Diseño
Fase Tres	Construcción
Fase Cuatro	Implementación
Fase Cinco	Iteración

Algunas metodologías pondrán más fases, o quizá con diferentes nombres que el resto de las otras, pero la esencia debe ser la misma, recordemos que “leche la necesaria” y la leche puede ser de la marca que más le guste al que los prepara.

Cada uno de las fases se describe a continuación:

Fase Cero Evaluación de Factibilidad

Muchas metodologías no incluyen esta etapa cero, ésta es la más importante de cualquier proyecto, ya que es la parte del estudio que busca justificar la inclusión de un data warehouse en una organización. Durante esta etapa se debe cuidar de los siguientes puntos:

- ☉ El costo de esta etapa no debe ser significativo
- ☉ La compañía debe apoyarse en recursos externos, coordinados por una persona de nivel jerárquico adecuado en la organización.
- ☉ Que sea comunicado a toda la organización



- ⊖ Que participen en los talleres y reuniones la gente importante de los departamentos
- ⊖ Que se sujeten a los resultados obtenidos de esta etapa.

Al finalizar esta etapa cero, la dirección de la organización contará con los elementos suficientes para continuar hacia una siguiente etapa, pero debe considerarse y tenerse bien claro que una vez iniciada la siguiente etapa resultará poco práctico finalizarlo y se habrá invertido tiempo dinero y esfuerzo importante para la compañía.

La metodología debe establecer las bases, orientación y utilización para el desarrollo del data warehouse de técnicas especializadas, obteniendo un aprovechamiento óptimo de los recursos, tal como se muestra en el Diagrama 7.1.

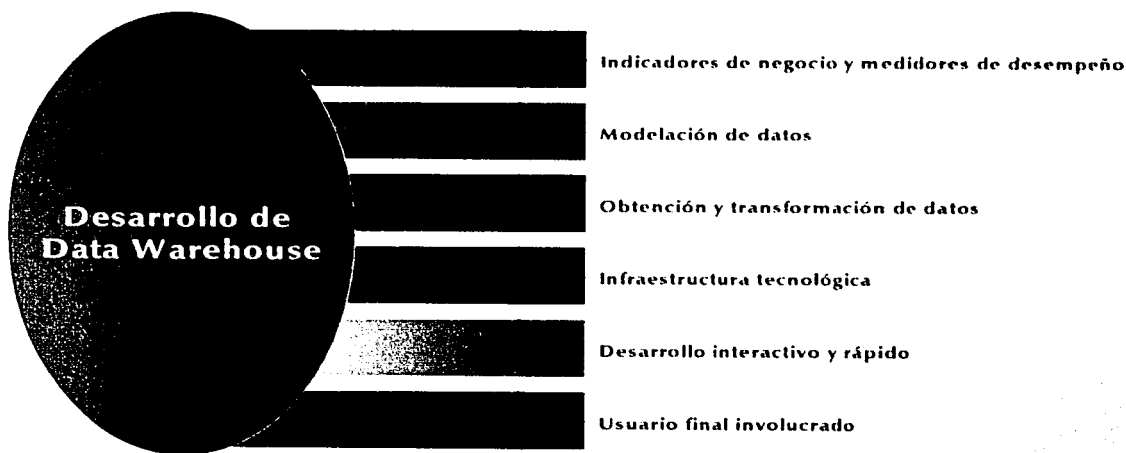


Figura 7.1 Diagrama de Desarrollo del data warehouse



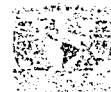
En esta etapa se realizarán talleres y entrevistas para detectar las necesidades de información de cada uno de los Gerentes o Directores de área, y también para detectar a los candidatos para el equipo de data warehouse. Todas las entrevistas deberán manejarse bajo un mismo formato, de manera tal que al momento de concentración de información se pueda contener bajo un mismo esquema y resulte más fácil su interpretación. Los talleres se realizarán con los principales usuarios de cada una de las áreas por separado y después uno con todas las áreas para determinar el grado de integridad entre áreas y poder dimensionar exactamente el trabajo a desarrollar.

El principal resultado de esta etapa, es la factibilidad de un proyecto de data warehouse en la organización, este punto será una guía importante para el resto del proyecto, probablemente el resultado en algunos casos sea que la organización sólo requiere de un proyecto de reporte.

Fase Uno Análisis

La fase de Análisis se debe iniciar con una junta en donde se comuniquen los roles, responsabilidades y fines que se persiguen con el proyecto.

En esta etapa se realizará la abstracción del proyecto, definiendo cual o cuales serán los departamentos que tendrán el data warehouse, cuales serán los fines, analizar las herramientas a utilizar, fuentes de información que servirán de fuente para la base de datos del data warehouse, etc. Posteriormente, se deberá evaluar el ambiente de administración de conocimiento en la organización.



A partir de este punto será necesario tener bien definida la lista de los usuarios que deberá ser capacitado en cada una de herramientas y tener en cuenta el tiempo que se invertirá en cada grupo o persona en materia de capacitación.

En todo momento del proyecto, es recomendable llevar a cabo la documentación, que servirá como la bitácora del proyecto.

Una de las partes importantes de esta fase es el análisis de los procesos, si ya se tienen mapeados los procesos de proyectos anteriores tales como el de la implementación de algún ERP. En esta etapa, se deberán analizar los reportes con los que cuenta la compañía, la forma en que se toman las decisiones, la funcionalidad práctica de los reportes y los resultados que de estos se obtienen e iniciar a definir, desde un punto de vista práctico, si se pueden cambiar, redefinir o definitivamente cambiar por nuevos, es importante considerar que no existen vacas sagradas en este punto.

También, en este punto se deberán definir el proceso analítico de información, es decir, cuales serán las tablas de donde se extraerá la información, cuales serán las reglas para combinar datos, se estimará cual será el crecimiento de la base de datos, etc.

Además, en esta etapa, se tendrán definidas la parte de hardware y software a utilizar y se deberán evaluar al menos a tres diferentes distribuidores por cada una de las herramientas y equipo que se piensa adquirir¹, al igual que en todos los puntos

¹ Resulta importante mencionar en este punto, la selección de la tecnología a utilizar (hardware y software) ya que no deberá depender sólo de factores económicos, existen historias de terror en el mercado de data warehouse nacionales de empresas que seleccionaron juntos hardware, software y servicios de consultoría y a la fecha no han podido arrancar su sistema, esto ha provocado que los líderes cambien, desencadenando en desconfianza en proyectos de este tipo. Es recomendable hacer una selección bastante meticulosa de cada uno de los componentes del data warehouse.

anteriores, se deberán conservar la documentación e integrar un comité en el que participen áreas funcionales y de tecnología de información para la selección de los distintos proveedores de soluciones, incluyendo los de servicios de consultoría.

Al final de este proceso, es recomendable hacer una parada y reunir al equipo de data warehouse y a los distintos proveedores para hacer el análisis de los avances y definir los siguientes pasos a seguir y redireccionar las estrategias pertinentes.

El Diagrama 7.1. anterior, también muestra el desarrollo que deberá observarse para la construcción adecuada de un data warehouse en la organización.

Fase Dos Diseño

En la fase dos deberá continuarse con la documentación del proyecto y los entrenamientos a usuarios.

En esta fase se deberá trabajar fuertemente el proceso de extracción transformación y carga de los datos, primero desde un punto de vista de diseño lógico y después iniciar con el físico, en esta parte del diseño se tendrá que definir claramente toda la cuestión de los metadatos, de tal manera que el sistema sepa a donde ir para extraer la información y que esta información se encuentre exactamente en donde está definida. En este punto, muchas organizaciones realizan el proceso de extracción de información de grandes almacenes de información como puede ser de Main Frames o de repositorios de información. Este punto es delicado por lo que es importante no dejar para después o no prestar el cuidado adecuado a la limpieza de los datos, un principio básico en este paso es "metes basura, obtienes basura", de aquí que surja la idea "pocos datos, mucha información".



Adicionalmente, se debe trabajar en la parte de los nuevos datos que arrojará el sistema, y cuales serán las reglas para obtenerlos, muchas veces los datos final provienen de diferentes datos de procesos, un ejemplo puede ser en la sección de manufactura de una empresa en donde se desee conocer el volumen de producción, en este punto no sólo está implícito el dato final de producción sino que también interviene el desempeño de las máquinas, los turnos, la materia, la oportunidad de entrega de materia prima por parte del distribuidor y hasta factores externos como la posición monetaria.

Otra tarea importante a desarrollar en esta etapa es del diseño de la presentación final para cada uno de los departamentos y principalmente a cada uno de los usuarios del sistema, no verá las misma pantallas el Director General, que el gerente de algunos de los departamentos, recordemos que las decisiones se toman por nivel. Es en esta parte en donde se inicia el juego de las dimensiones de información y el detalle al que se deberá llegar en el proyecto.

Fase Tres Construcción

En esta fase se inicia el proceso de construcción de la transformación de datos, ya es iniciar el poblado de la base de datos final, esta tarea es especialmente tardada, algunas veces se necesitará que el equipo responsable de la base de datos trabaje turnos extra, ya que puede ocurrir que algo impida que se ejecuten los programas y se pierda tiempo valioso.

También por la parte de la presentación final, se deja a un lado el papel y se hace la construcción de las pantallas que se definieron en la fase anterior, algunas veces las pruebas de los resultados se podrán realizar con datos no reales.

Una de las tareas importantes en esta fase es la de pruebas de estrés, en donde se jugarán con los sistemas hasta tratar de encontrar su punto donde el sistema pueda no responder y aplicar las correcciones que se consideren pertinentes.

Al final de esta fase, se deberá hacer una revisión bastante meticulosa de los resultados versus lo originalmente diseñado para medir cuales fueron los ajustes y determinar cuales pueden ser las repercusiones hacia los usuarios y la organización.

Fase Cuatro Implementación

Este es una fase que quita mucho el sueño a todos los miembros del equipo, especialmente por un día antes de dar por liberado el sistema, en este punto se hará un análisis de los resultados que esperan los usuarios, de lo que se realizó para satisfacer las necesidades de información, la utilización de herramientas, los datos, etc. Debido a que en un momento se estará trabajando con datos reales, con reportes reales y estarán conociendo la esencia de la organización, algunas veces los datos que se empiezan a dar resultan sorprendentes para los usuarios y algunas otras no causa ninguna sorpresa.

Es en esta parte en donde los usuarios van a iniciar a explotar los conocimientos que fueron adquiridos a lo largo del proyecto y que sin duda alguna es la parte en donde surgen las dudas sobre ciertos procesos o reportes, pero es una etapa que necesitará madurar a la organización, los usuarios y los sistemas.

A lo largo de todas etapas, será necesario contar con equipos especializados en administración del cambio, principalmente para apresurar el paso de áreas integradas, para que se pueda cobrar conciencia de que una mala introducción de datos, puede provocar algún tipo de conflicto en alguna parte de la cadena de

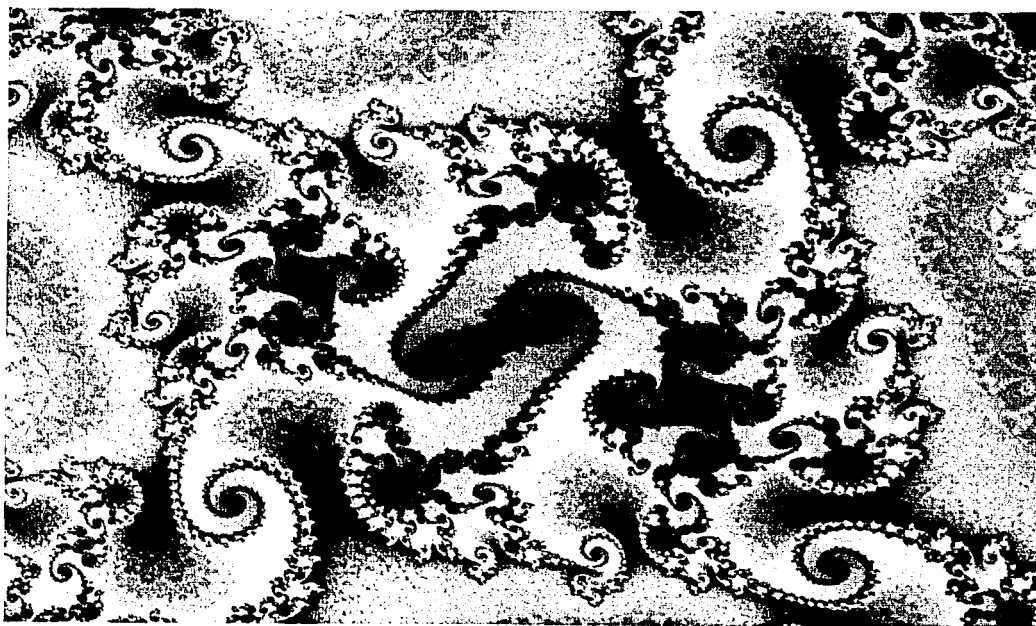
información, el iniciar a los departamentos a dejar de verse como mini compañías en una empresa y verse como una verdadera organización, que persigue metas y logros comunes y que repercutirá en beneficios para todos los involucrados.

Fase Cinco Iteración

Una vez concluido un proyecto de data warehouse, deberá iniciarse otra vez, ya que los usuarios, la empresa, los sistemas y los datos habrán madurado, habrá que definir nuevos y mejores medidores, nuevos y mejores y quizá hasta más prácticos reportes y en el mercado habrá mejores herramientas, lo bello de un data warehouse es que siempre hay algo que mejorarle al sistema, los reportes y los datos.



Capítulo Octavo



Conclusiones y Recomendaciones



Conclusiones y Recomendaciones

La Verdad es Relativa, La Neta es Absoluta...

Adela Micha

Estoy convencido que vivimos en un país hermoso en donde todo se puede, lamentablemente hay muchos retos que vencer todavía. Afortunadamente, estamos viviendo cambios importantes en el ámbito mundial y que impactan en nuestro país, los empresarios comienzan a darse cuenta que la forma de administrar tradicionalmente a sus empresas ya no es tan sencilla, aquellas compañías que inician el proceso de renovación de la noche a la mañana las vemos desaparecer. Creo que México está abriendo los ojos hacia las nuevas tecnologías y sé que hace falta mucho trabajo que hacer, por lo que yo pongo mi pequeño grano de arena, o mejor dicho, estoy dando un primer punto de apoyo...

Una de las razones que me llevó a realizar este trabajo es encontrarme que en México no existe mucha documentación referente al tema de data warehouse, la gran mayoría de los libros están en inglés quizá derivado de que el concepto nace en los Estados Unidos. También, durante mi desarrollo profesional me he encontrado que en México hay mucha renuencia al desarrollo de proyectos de data warehouse y, las compañías que presumen de tener un sistema de este estilo, cuando es analizado se encuentra que le hace falta mucho para cumplir los requisitos mínimos de data warehouse. Existe una empresa que después de una importante inversión de tiempo, recursos humanos y dinero no tenían un data warehouse adecuado, es más el sistema no tenía definido los medidores de desempeño adecuados, creo que este tipo de experiencias son las que han provocado en México el concepto de data warehouse no avance como en otros países.

Como fiel creyente de la teoría de caos, creo que llegará un momento en que esta situación caótica alrededor del concepto de data warehouse cobrará un orden. Y control y poco a poco habrá más compañías adoptando este tipo de tecnologías.

Las conclusiones del trabajo se enfocarán a recomendaciones para el desarrollo de sistema de información para la toma de decisiones en México, las cuales abarcan los puntos más importantes a considerar al momento de seleccionar como estrategia de negocio un data warehouse:

Control de las operaciones en las empresas

Una de las primeras estrategias que debe pensar en desarrollar una empresa en México es la de controlar sus operaciones con los sistemas ERP's (Enterprise Resource Planning), esto es con la finalidad de poder iniciar la integración de sus procesos y contar con un solo criterio para la información interna de la compañía. También es necesario hacer que los usuarios realmente utilicen la funcionalidad de estas herramientas, algunas veces los usuarios se resisten a utilizarlas debido a que es una nueva forma de realizar sus actividades.

Hay que esperar que los usuarios maduren en la utilización del sistema, resultará poco práctico terminar de hacer la implementación e iniciar con el proyecto de data warehouse. También, debe considerarse no realizar la implementación de una estrategia de data warehouse antes de terminar la implementación del sistema ERP, ya que resultará bastante complicado e impráctico.



Selección del Equipo de Trabajo

Para la selección de los equipos de data warehouse debe considerarse que estamos hablando del mercado mexicano, en donde sería deseable contar con una persona para cada una de las características descritas en el capítulo de "Equipo Data Warehouse", pero lamentablemente pocas organizaciones tienen la capacidad de tener a tanta gente en el área de sistemas y resultará aún más complicado asignar a personal de otros departamentos al proyecto sin descuidar sus actividades, por lo que es recomendable considerar tener un equipo de consultores si no de tiempo completo por cuestiones de costo si en partes muy puntuales del proyecto, pero procurando tener siempre a una persona de la compañía integrada al equipo de consultores para asegurar que se cumpla con las necesidades de la empresa y a la vez para que se realice la transferencia de conocimientos.

Inversión en el Proyecto

Si una compañía no puede implementar un sistema data warehouse para toda la organización, es recomendable iniciar un proyecto de data mart por departamento, comenzando en ventas o en finanzas, dependiendo de la naturaleza de la empresa, y después definir otra área para continuar con el proceso de data marts por departamento, pero trazando un plan de acción definiendo prioridades y fechas de implementación para no descuidar la integridad de la información que contendrá el repositorio de información.

Si la empresa pretende construir un data warehouse a través de data marts, debe hacer una evaluación considerando el equipo y el crecimiento de la base de datos para que al momento de la compra y de la utilización, no se corra el riesgo de que no cumpla con las necesidades del proyecto.

Toda inversión que realice la compañía debe estar bien sustentada, analizar las herramientas a utilizar en data warehouse y considerar al momento de la selección del o los distribuidores de soluciones lo siguiente:

- ⊖ Que la compañía tenga representación en México
- ⊖ Que la compañía tenga cartas credenciales en implementaciones exitosas
- ⊖ Que la compañía cuente con soporte técnico de buen nivel
- ⊖ También que las herramientas tengan la capacidad de combinarse con otras herramientas del mercado sin la necesidad de mucha programación
- ⊖ Que cumpla con los requisitos propios de la compañía.

Metodología

A final de cuentas, todas las metodologías existentes en el mercado de data warehouse tienen su origen de la misma fuente, todas tienen casos de éxito como de terror, por lo que para una buena selección de metodología de implementación, se debe considerar que:

- ⊖ Sea madura, es decir, se haya utilizado en otras implementaciones
- ⊖ Muestre claramente las actividades que se deben desarrollar en cada momento
- ⊖ Identifique claramente los puntos de Análisis, Diseño, Construcción o Implementación y Mejora Continua.
- ⊖ Establezca los productos que se deben obtener a cada momento y los específicos al final de cada etapa y los que se tendrán al final del proyecto, de entre los más importantes deberán ser manuales tanto técnicos como operativos, carpeta de configuración, etc.

Debe tenerse en cuenta que una metodología solo es una guía y que se deberá adecuar a las necesidades finales de la empresa y del proyecto. También debe cuidarse de no exceder el tiempo de implementación ya que esto repercutirá en excesos de tiempo, dinero y esfuerzos adicionales.

Justificación de un Data Warehouse

También, presento algunos casos de éxito de usos actuales del data warehouse:

Análisis de Rentabilidad de Clientes

- ☛ Seleccionar los mejores clientes y enfocar directo a compradores potenciales

Ventas y Seguimiento a Embarques

- ☛ Analizar las rutas y detectar mejoras en el reparto

Ventas e Inventarios

- ☛ Analizar millones de cuentas de clientes, creación de perfiles de clientes para mercadotecnia
- ☛ Mejorar la capacidad de análisis de datos y disminuir cargas de trabajo manual"

Control de Créditos

- ☛ Integrar los sistemas de control de créditos y facilitar el análisis de créditos con 57 sistemas alimentadores. Detección oportuna de cartera y reincidencias"

Análisis de Competencia y Penetración

- ☛ Analizar la rentabilidad desde el punto más bajo llamada- para todas las empresas competidoras.
- ☛ Apoyar a la toma de decisiones estratégicas con base a la información generada

Programas de viajero Frecuente



- ⊖ Analizar rutas y destinos y características de los viajeros para incrementar los vuelos

Detección de conductas fraudulentas

- ⊖ Uso indebido de tarjetas robadas

Análisis Financiero

- ⊖ Rentabilidad

Análisis de Competencia / Mercados

- ⊖ Analizar productos y posicionamientos de la competencia y de respuestas de clientes

Análisis de Recursos Humanos

- ⊖ Ocupación
- ⊖ Educación Continua
- ⊖ Desarrollo Profesional, etc.

Para detectar la necesidad de un data warehouse en una organización, habrá que determinar, dependiendo del tipo de industria al que se pertenece, las siguientes situaciones

- ⊖ Falta de información para la toma de decisiones
- ⊖ Ambiente de Alta Competencia
- ⊖ Dificultad para obtener reportes
- ⊖ Información inconsistente y en múltiples plataformas
- ⊖ Problemas de desempeño en los Equipos Centrales
- ⊖ Necesidad de Mejorar el Desempeño del Negocio
- ⊖ Desconocimiento o Incursión en Nuevos Mercados
- ⊖ Procesos de Fusión con otras Empresas
- ⊖ En Proceso de Planeación Estratégica del Negocio / Sistemas
- ⊖ Implantación de un ERP (SAP, PS, Oracle, etc.)
- ⊖ Necesidad de Predecir el Comportamiento de los Clientes
- ⊖ Incursión en Comercio Electrónico
- ⊖ Necesidad de Adquirir o Retener Clientes
- ⊖ Necesidad de Combatir conductas Fraudulentas



Entre otros.

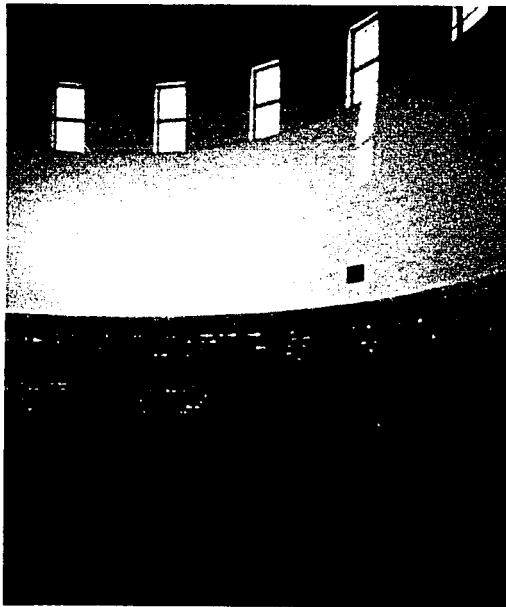
Por último, mi conclusión más clara es que para que nuestras compañías enfrenten las nuevas tendencias de la economía global, es necesario considerar la adopción de sistemas de data warehouse y explotarlo y actualizarlo en cada momento que sean necesarios y buscar una mayor difusión en el ámbito empresarial.

En la medida en que las empresas en México adopten este tipo de tecnologías podrán ser más competitivas y podrán reaccionar a los embates de las nuevas economías, y en la medida en que los usuarios aprendan y se involucren con el sistema el alto costo del data warehouse se verá compensado.

La conclusión final es que al final de este trabajo se presentaron términos, conceptos y definiciones claras para personas no necesariamente técnicas, quizá un poco más funcionales con la finalidad de mostrar un texto amigable para el entendimiento del concepto de Data Warehouse. Hoy por hoy, en el campo de data warehouse existen muchos libros con conceptos principalmente técnicos, que de alguna manera no pueden ser entendidos por la parte funcional y directiva de las empresas.



Bibliografía



Data Warehouse



Bibliografía

Como todos los hombres de la Biblioteca, he viajado en mi juventud; he peregrinado en busca de un libro, acaso del catálogo de catálogos; ahora que mis ojos casi no pueden descifrar lo que escribo.
La Biblioteca de Babel

Una leyenda tendenciosa se extendió durante siglos. Los árabes habrían destruido la célebre biblioteca de Alejandría cuando conquistaron la ciudad en el siglo VII. Los árabes nunca pudieron incendiar la Gran Biblioteca de Alejandría, ni siquiera la Pequeña Biblioteca, sencillamente porque, cuando las tropas llegaron a la ciudad en el 641, ya hacía cientos de años que no existían. Lo que encontraron los árabes fue una ciudad dividida, arruinada y exhausta por siglos de luchas civiles. Tanta belleza no fue destruida por los guerreros árabes que tomaron las ruinas de la ciudad en el 641, sino por los cristianos monofisitas en el año 391. Los cristianos destruyeron e incendiaron violentamente el Serapeum alejandrino. Las llamas arrasaron así la última y fabulosa biblioteca de la Antigüedad. Una parte importante del pensamiento clásico y pagano desaparecía al mismo tiempo.¹

Bibliografía

Para la realización de este trabajo se consultaron diferentes fuentes, la bibliografía es la siguiente:

- 1 Data Warehousing
La integración de Información para la mejor toma de Decisiones
Harjinder S. Gill
Petince Hall Hispanoamérica
1996
ISBN: 968-880-792-3

¹ Síntesis del libro de Pablo de Jevenois, "El fin de la gran biblioteca de Alejandría: La leyenda imposible."
Revista de arqueología (www.zugarlo.es)



- 2 E-business
Roadmap for Success
Ravi Kalakota
Marcia Robinson
Addison Wesley
1999
ISBN: 0201604809
- 3 Measuring Corporate Performance
Peter F. Drucker
Robert G. Eccles
Christopher Meyer
Harvard Business School Press
1998
ISBN: 0-87584-882-6
- 4 The Data Warehouse
Lifecycle Toolkit
Ralph Kimball
Laura Reeves
Margy Ross
Waren Thornthwaite
Addison Wesley
1998
ISBN: 0-471-25547-5
- 5 Using the Data Warehouse
W.H. Inmon
Richard D. Hackathorn
Addison Wesley
1994
ISBN: 0471-05966-8
(Autografiado por el Bill Inmon)
- 6 Advance in Knowledge Discovery an Data Mining
Usama M Fayyad
Gregory Piatetsky-Shapiro
Padhraic Smytth
Ramasamy Uthurusamy
American Association for Artificial Intelligence



- 1996
ISBN: 0-262-56097-6
- 7 The Data Warehouse Toolkit
Ralph Kimball
Addison Wesley
1996
ISBN: 0-471-15337-0
- 8 Understanding and Implementing Successful Data Marts
Douglas Hackney
Addison Wesley
1997
ISBN: 0-201-18380-3
- 9 Data Management
An organizational Perspective
Richard T. Watson
Addison Wesley
1996
ISBN: 0-471-30534-0
- 10 The Data Warehouse Challenge
Taming Data Chaos
Michael H. Brackett
Addison Wesley
1996
ISBN: 0-471-12744-2
- 11 Data Sharing
Using a Common Data Architecture
Michael H. Brackett
Addison Wesley
1994
ISBN: 0-471-30993-1
- 12 Distributed Data Warehousing Using Web Technology
How to Build a More Cost Effective And Flexible Warehouse
R.A. Moeller
2000
American Management Association
ISBN: 0-8144-0588-6

- 13 Mastering Project Management
Applying Advanced Concepts of Systems Thinking
James P. Lewis
Mc Graw Hill
1998
ISBN: 0-7863-188-6
- 14 Data Warehouse Project Management
Sid Adelman
Larissa Terpeluk Moss
Addison Wesley
2000
ISBN: 0-201-61635-1
- 15 Working Knowledge
How Organizations Manage GAT
Thomas H. Davenport
Lauréese Prusak
Addison Wesley
1997
ISBN: 0-201-98763-2

Otras fuentes de información

16. Seminarios de Data Warehouse impartido por Bill Inmon, notas
17. Cursos de Data Warehouse impartidos por PricewaterhouseCoopers, notas
 - ☉ Data Warehouse Methodology
 - ☉ Performance Measurement
 - ☉ Data Resource Management
 - ☉ Cursos Especializados en Herramientas Data Warehouse

También páginas en Internet en lugares especializados del tema de data warehouse como son:

www.datawarehouse.com
www.hb.com

www.dwreview.com
www.dwinstitute.com



Anexo I

Arquitecturas Clásicas de un Data Warehouse

... Hope is a good thing
... maybe the best of things
And no good things ever dies

De la película "Shawshank Redemption"



Sistemas Data Warehouse

A continuación, se muestran algunas de las arquitecturas de data warehouse:

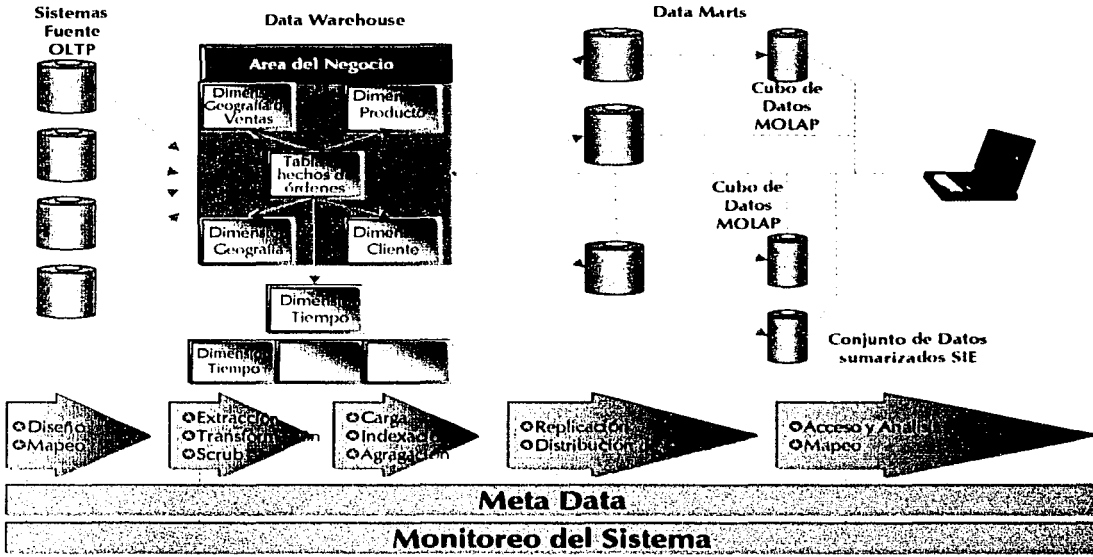


Figura A.1. Arquitectura clásica de un data warehouse



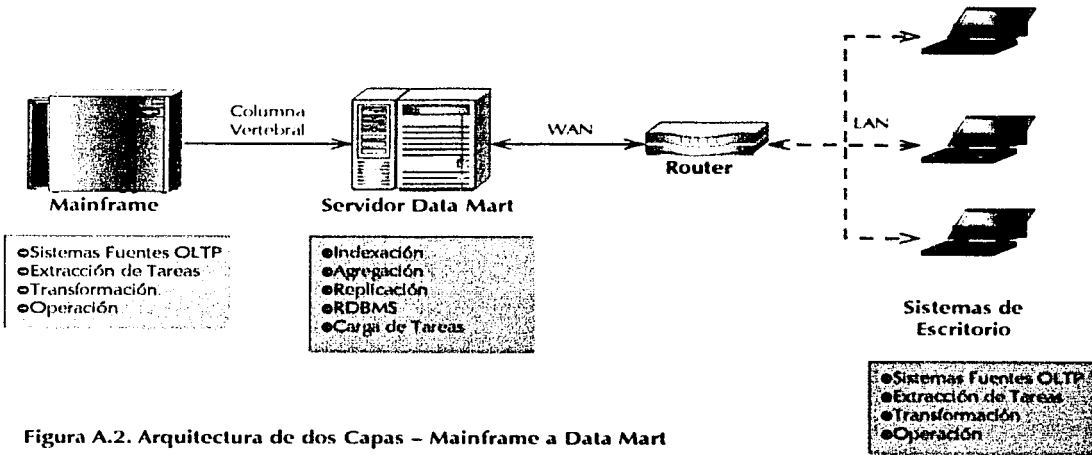


Figura A.2. Arquitectura de dos Capas – Mainframe a Data Mart

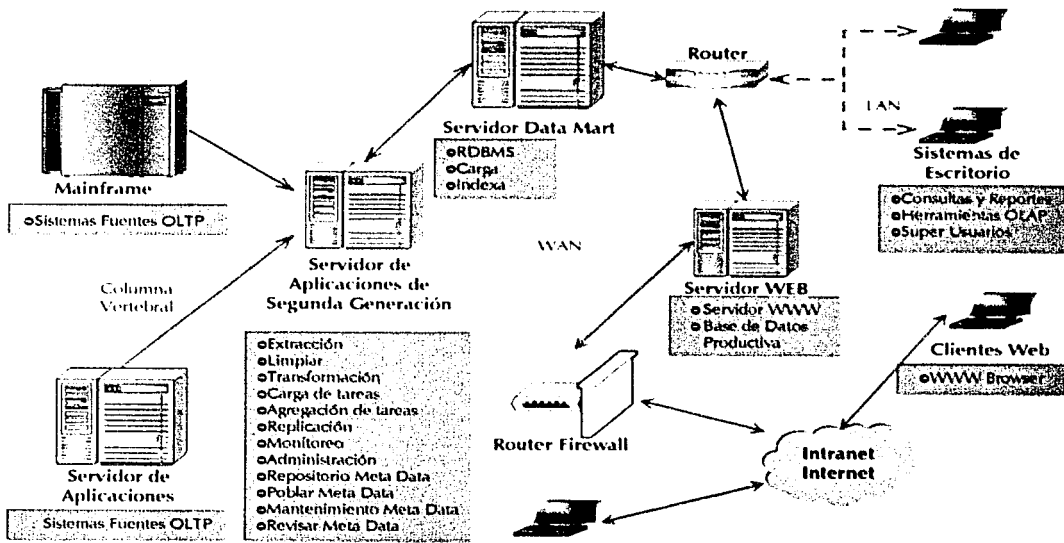


Figura A.3. Arquitectura de Tres Capas – Servidor ROLAP/MOLAP



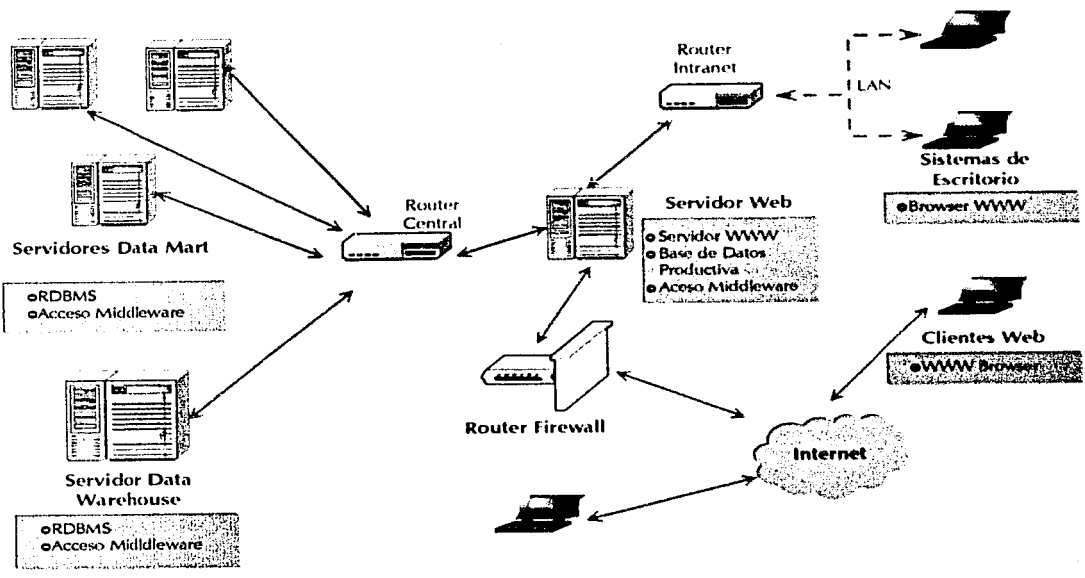


Figura A.4. Topología Intranet / Internet



Anexo II

Modelo Dimensional

El hombre razonable se adapta al mundo; el irrazonable intenta adaptar el mundo a sí mismo... Así pues, el progreso depende del hombre irrazonable

G. Bernard Shaw



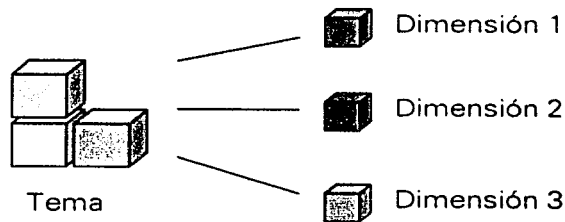
Modelo Dimensional

La técnica de modelado para entender, modelar y analizar consultas empresariales consiste en construir modelos de consultas.

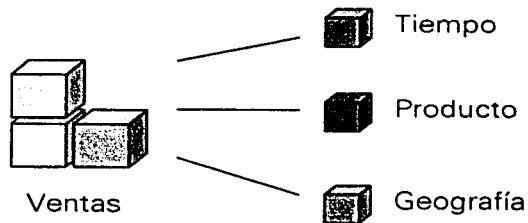
Se inicia con el entendimiento de lo que se desea ver en el reporte, se definen los niveles que se desean observar y analizar y se realizan bosquejos esquemáticos de los informes. Con estas definiciones, se busca en las tablas de hechos y dimensión para el modelo del data warehouse, a continuación se muestra un ejemplo de este modelado.

La forma en que se buscó la interpretación del modelo dimensional es mediante el uso de cubo Rubik.

Modelo de Consulta

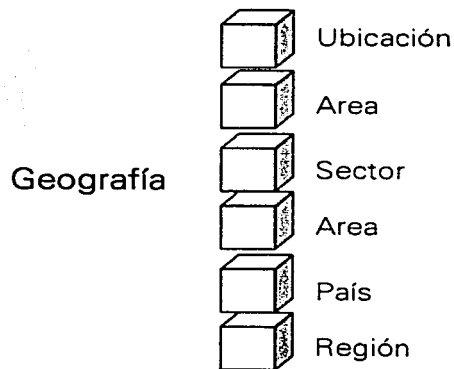
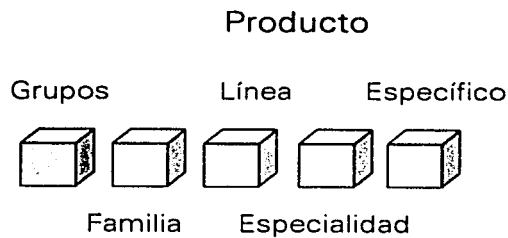
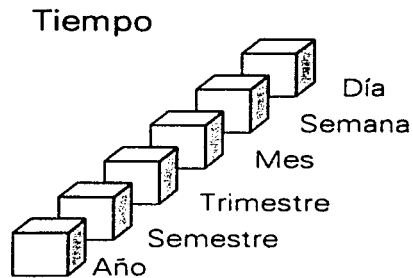


Modelo de consulta para un área de ventas



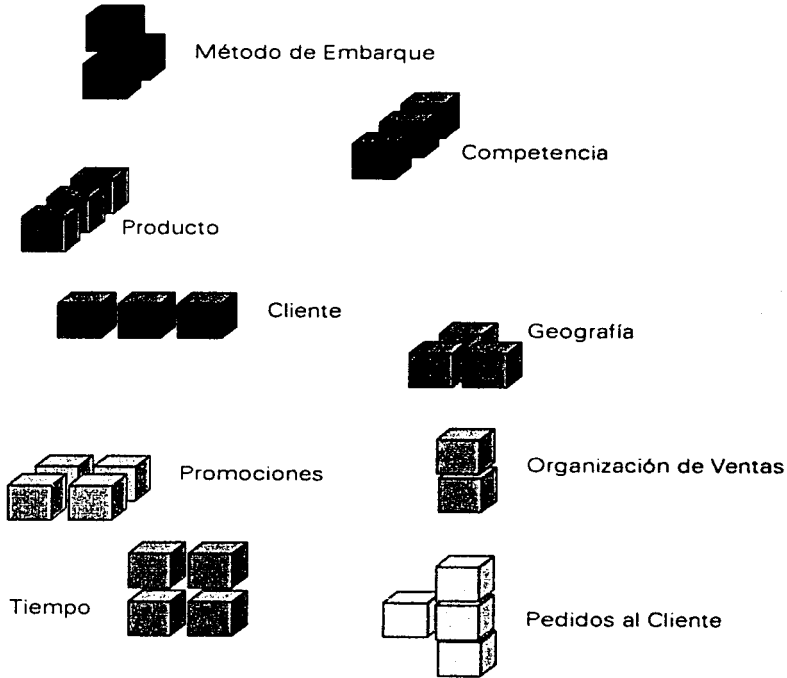
Ahora analicemos por cada una de las dimensiones; es decir, tiempo, producto y geografía

Modelos de tiempo, producto y geografía

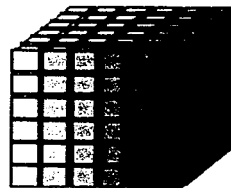


Otras Dimensiones

De esta forma el cubo se va adecuando a las necesidades de información de las compañías, a medida que se analizan las necesidades con mejor detalle, el cubo es cada vez más complejo. Otras dimensiones pueden ser:



Y de esta forma el cubo va tomado su forma final



* * * * *



Anexo III

Un Tributo a Mary Glazman †

Recuerdo que una sesión de trabajo, Mary Glazman † me comentó que ella junto otras amigas de las Facultad son consideradas como las Glorias de la Hipotenusa, seguramente lo sigues siendo.

Mary, gracias por lo que de ti aprendí.



Un Tributo a Mary Glazman †

Hace tiempo cuando inicié la búsqueda de una directora de tesis, platicué con Concepción Ruiz, quién fue profesora durante mucho tiempo y de varias materias en mi estancia en la Facultad de Ciencias. Después de un rato de platicar con Concha, me recomendó que sería más conveniente que otra persona fuera mi directora de tesis, fue entonces cuando caminamos hacia el cubículo de Mary, Concha me presentó con Mary como uno de sus alumnos y así fue como llegué al cubículo de Mary Glazman †.

Francamente, durante mi carrera nunca tomé clases con Mary, por lo que solamente tenía las opiniones y comentarios de otros amigos y compañeros de la Facultad, que sí habían podido ser alumnos de Mary. Por este motivo, al principio me dio algo de miedo trabajar con Mary. Después de un rato de una amena plática con Mary ese miedo fue desapareciendo, al final de nuestra pequeña conversación salí del cubículo de Mary Glazman † con un título para mi Tesis, "Gráficas y Matrices".

Durante mucho tiempo Mary y yo trabajamos en su cubículo, en su casa, por teléfono, por correo electrónico y no sé cuantos medios más, y poco a poco avanzamos en mi Tesis. Principalmente por cuestiones de mi trabajo se fue retrazando mi tesis. En una de nuestras muchas sesiones, platicamos del avance, Mary me comentó esa vez que consideraba que ya teníamos el esqueleto de la tesis, quizá sólo faltaba complementarla con algunas cosas más, pero la parte fuerte ya la teníamos. Tiempo después, cuando volví a comunicarme con Mary me comentó que no nos podíamos ver ya que estaba bajo terapias médicas porque le habían detectado Cáncer. En la medida de lo posible intentaba comunicarme con Mary para saber de su estado de salud, algunas veces podía platicar con ella y otras me decían que estaba descansando. Ese mismo año Mary realizó su último viaje. Mary había fallecido.

Cuando inicié este nuevo trabajo de tesis, Data Warehouse, platicué con la Doctora Amparo López Gaona, sobre mi inquietud de la realización de la tesis. También, le comenté que no quería dejar en un lugar olvidado el trabajo que había hecho con Mary Glazman \dagger sobre Gráficas y Matrices, por lo que pedí a Amparo su autorización de incluir este Anexo, Un Tributo a Mary Glazman \dagger , en el cual sólo incluyo uno de los cuatro capítulos que había podido trabajar con Mary, el capítulo en donde aprendí mucho de Mary, el capítulo en donde conocí a Mary. Así que, con esta sección deo evidencia que tuve la suerte de trabajar, platicar y aprender con Mary Glazman \dagger , sin duda alguna, una de las glorias de la hipotenusa. Simplemente gracias Mary, en donde quiera que estés.

Angel Montes de Oca Baltazares

Gráficas y Matrices

CAPITULO 1

Definición 1.0. Una gráfica es una pareja de conjuntos $[V, A]$, donde V es un conjunto de puntos llamados vértices o nodos y A es un conjunto de líneas que unen a algunos de los vértices. Se denotará como $G = [V, A]$ a una gráfica.

Si los elementos de A tienen dirección, representada con una flecha, en donde se distingue un vértice inicial A y un vértice final B



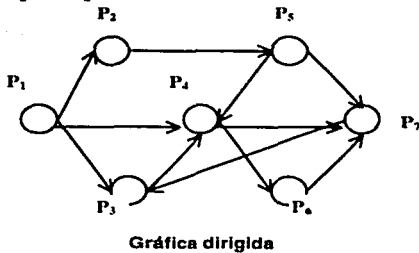
Se dice que la gráfica es orientada o dirigida. En este caso llamaremos arcos a los elementos de A . Si los elementos de A no tienen dirección, los llamaremos aristas y en este caso \mathcal{G} no es dirigida.

Si V y A son conjuntos finitos diremos que $\mathcal{G} = [V, A]$ es una gráfica finita.

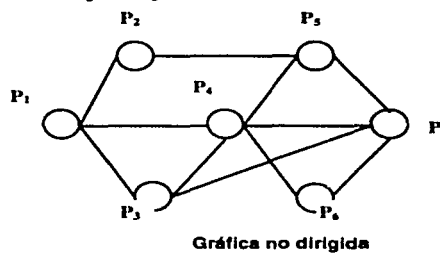
A lo largo de este trabajo \mathcal{G} será una gráfica (dirigida / no dirigida) finita.

Un ejemplo de gráficas dirigidas y no dirigidas son las siguientes:

$\mathcal{G}_1 = [V, A]$



$\mathcal{G}_2 = [V, A]$



Definición 1.1. Sea $\mathcal{G} = [V, A]$ una gráfica no dirigida. Un camino en G es una secuencia de aristas a_1, a_2, \dots, a_n donde a_i está conectado a a_{i-1} por un extremo y a a_{i+1} por el otro.

Si $\mathcal{G} = [V, A]$ es una gráfica dirigida, un camino en \mathcal{G} es una secuencia de arcos a_1, a_2, \dots, a_n en donde el vértice final de a_{i-1} está conectado con el vértice inicial de a_i .

Definición 1.2. Una gráfica no dirigida $\mathcal{G} = [V, A]$ es conexa si para todo par de vértices $i, j \in V$ existe un camino que los une.



Definición 1.3. Un lazo en una gráfica $\mathcal{G} = [V, A]$ es cuando un arco o arista tienen el mismo vértice inicial y final, ejemplo:



Existen distintas formas de asociarle una matriz a una gráfica. Citaremos algunas.

Definición 1.4. Sea $\mathcal{G} = [V, A]$ una gráfica no dirigida. La matriz de adyacencia de \mathcal{G} , que denotaremos $A_{\mathcal{G}}$ es una matriz de $n \times n$ donde $n = |V|$; tal que

$$a_{ij} = \begin{cases} \text{número de aristas entre } i \text{ y } j, \text{ si } i \neq j \\ 2 \text{ veces el número de lazos en } i, \text{ si } i = j \end{cases}$$

Observemos que la matriz de adyacencia de una gráfica no dirigida siempre es no negativa y simétrica.

Así, tenemos que la matriz de adyacencia de la gráfica $\mathcal{G}_2 = [V, A]$ es:

$$A_{\mathcal{G}_2} = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 \end{pmatrix}$$

como puede verse, la matriz de adyacencia $A_{\mathcal{G}_2}$ es no negativa y simétrica.

Definición 1.5. Sea $\mathcal{G} = [V, A]$ una gráfica no dirigida y $A_{\mathcal{G}}$ su matriz de adyacencia. El polinomio característico de \mathcal{G} es $P_{\mathcal{G}}(x) = \det(xI - A_{\mathcal{G}})$.



Definición 1.6. Sea $P \in M_n$ una matriz. P es una matriz de permutación si exactamente una entrada en cada renglón y columna es igual a uno y en todas las demás son cero. En general, al multiplicar una matriz $A \in M_m \times n$ por una matriz de permutación $P \in M_n$, esta última permuta los renglones de A si se realiza la multiplicación por la derecha de A y permuta las columnas de A si se multiplica a la izquierda. El determinante de las matrices de permutación es ± 1 , así que las matrices de permutación son no singulares.

Lema 1.0. Sea $P \in M_n$ una matriz de permutación, entonces

- (a). Para cualquier A , PA puede obtenerse a partir de A permutando los renglones de A exactamente como se permutan los renglones de I para obtener P .
- (b). P es no singular, $P^{-1} = P^T$ tal que $PP^T = P^TP = I$.

Prueba.

- (a). El resultado se sigue de la Definición 1.6. anterior de la multiplicación de matrices de permutación.
- (b). Separemos a P en renglones, con lo que tenemos que

$$\begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix}$$
 son los renglones de e_i de I en algún orden.

Al transponer estos renglones, es decir obtener P^T , tenemos $[r_1^T r_2^T \dots r_n^T]$ que ahora son las columnas de P^T .

Por definición de multiplicación de matrices, implica que el elemento (i, j) de

PP^T es tan sólo $r_i r_j^T$, esto es $\begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$ lo cual implica que $PP^T = I$.

Por lo tanto $PP^T = I$

Ahora trabajando por columnas de P^T

$\begin{bmatrix} c_1^T \\ c_2^T \\ \vdots \\ c_n^T \end{bmatrix}$ que son los renglones de e_i de I en algún orden.

Nuevamente, al transponer estos renglones, es decir, obtener P , obtenemos

$[c_1 c_2 \dots c_n]$ las cuales ahora forman las columnas de P .

Por definición de multiplicación de matrices, implica que el elemento (j, i) de

$P^T P$ es tan sólo $c_j^T c_i$, esto es $\begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$ lo cual implica que $P^T P = I$.

Por lo tanto, $P^T P = I$

Por lo tanto, $PP^T = P^T P = I$

Definición 1.7. Una matriz $A \in M_n$ se dice que es reducible si:

(a) $n = 1$ y $A = 0$; o



(b) $n \geq 2$, existe una matriz de permutación $P \in M_n$ y un entero r con $1 \leq r \leq n-1$ tal que

$$P^T A P = \begin{bmatrix} B & C \\ \mathbf{0} & D \end{bmatrix}$$

Donde $B \in M_r$, $D \in M_{n-r}$ y $C \in M_{r \times n-r}$.

Definición 1.8. Si $A = [a_{ij}] \in M_{mn}$, fijamos $|A| \equiv [a_{ij}]$ y $M(A) \equiv [\mu_{ij}]$, donde;

$$\mu_{ij} = \begin{cases} 1 & \text{si } a_{ij} \neq 0 \\ 0 & \text{si } a_{ij} = 0 \end{cases} \quad \text{la matriz } M(A) \text{ es llamada la matriz indicador de } \Delta.$$

Por ejemplo,

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 1 \\ 4 & 0 & 1 \end{bmatrix} \quad \text{entonces} \quad M(A) = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

Si $|A| > 0$, claramente A es no reducible, y si A es reducible, ésta debe tener al menos $(n-1)$ entradas cero⁽³⁾.

⁽³⁾ Supongamos que deseamos resolver el sistema de ecuaciones lineales $Ax = y$, y supongamos que A es reducible.

Entonces si escribimos $\tilde{A} = P^T A P = \begin{bmatrix} B & C \\ \mathbf{0} & D \end{bmatrix}$, tenemos $Ax = P\tilde{A}P^T x = y$, o $\tilde{A}(P^T x) = P^T y$. Fijamos $P^T x = \tilde{x} = [z^T; \zeta^T]^T$

(desconocida) y $P^T y = \tilde{y} = [w^T; \omega^T]^T$ (conocida), donde $z, w \in C^r$ y $\zeta, \omega \in C^{n-r}$. Entonces el sistema de ecuaciones a

ser resuelto es equivalente a $A \tilde{x} = \tilde{y} = \begin{bmatrix} B & C \\ \mathbf{0} & D \end{bmatrix} \begin{bmatrix} Z \\ \zeta \end{bmatrix} = \begin{bmatrix} w \\ \omega \end{bmatrix}$, esto es, para

$$\begin{aligned} Bz + C\zeta &= w \\ D\zeta &= \omega \end{aligned}$$

Si resolvemos $D\zeta = \omega$ primero para ζ , entonces usamos ζ en la primera ecuación y resolvemos $Bz + C\zeta = w$ para z , tenemos reducido el problema original para dos pequeños problemas, en un principio, ser fáciles de resolver. Esta observación es la que motiva el término de reducibilidad.

Definición 1.9. Una matriz $A \in M_n$ se dice que es irreducible si ésta no es reducible.

Lema 1.1. Sea $A \in M_n$ reducible. Si A tiene un bloque de $(n-r) \times r$ de ceros en la parte inferior izquierda, entonces $|A|^{n-1}$ todas tienen el mismo bloque de $(n-r) \times r$ de ceros en la esquina inferior izquierda.

Prueba.

La demostración se hará por inducción sobre n .

Para $n = 2$, tenemos que el valor para $|A|^{n-1}$ es $|A|^1 = |A|$, por lo que comenzaremos la demostración para las n 's ≥ 2 , así sea $n = 3$.

Supongamos que $A \in M_n$ es de la forma $\begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$ en donde $B \in M_r$, $C \in M_{r \times (n-r)}$,

$D \in M_{n-r}$ y $0 \in M_{(n-r) \times r}$ en donde el bloque 0 está compuesto por ceros en todas sus entradas, así

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1r} & a_{1(r+1)} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2r} & a_{2(r+1)} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rr} & a_{r(r+1)} & \cdots & a_{rn} \\ \hline a_{(r+1)1} & a_{(r+1)2} & \cdots & a_{(r+1)r} & a_{(r+1)(r+1)} & \cdots & a_{(r+1)n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nr} & a_{n(r+1)} & \cdots & a_{nn} \end{bmatrix}$$

de tal forma que $a_{ij} = 0$ si $i > r$ y $j \leq r$, entonces



$$\begin{array}{ccccccc}
 a_{(r+1)1}, & a_{(r+1)2}, & \dots, & a_{(r+1)r} & = & 0 \\
 a_{(r+2)1}, & a_{(r+2)2}, & \dots, & a_{(r+2)r} & = & 0 \\
 \dots & \dots & \dots & \dots & \dots & \dots \\
 a_{n1}, & a_{n2}, & \dots, & a_{nr} & = & 0
 \end{array}$$

sean $i > r$ y $j \leq r$, entonces

$$\begin{aligned}
 \langle A^{j-1} \rangle_{ij} &= \langle A^2 \rangle_{ij} = \langle AA \rangle_{ij} = \sum_{k=1}^n \langle A \rangle_{ik} \langle A \rangle_{kj} = \sum_{k=1}^r \langle A \rangle_{ik} \langle A \rangle_{kj} + \sum_{k=r+1}^n \langle A \rangle_{ik} \langle A \rangle_{kj} \\
 &= \underbrace{\sum_{k=1}^r a_{ik} a_{kj}}_0 + \underbrace{\sum_{k=r+1}^n a_{ik} a_{kj}}_0 = \sum_{k=1}^r \underbrace{a_{ik} a_{kj}}_0 \quad y \quad \sum_{k=r+1}^n \underbrace{a_{ik} a_{kj}}_0
 \end{aligned}$$

Ahora, supongamos que es cierto para $q - 2$, es decir $|A|^{q-2}$ es de la forma

$$\begin{bmatrix} B & C \\ 0 & D \end{bmatrix}, \text{ en donde } a_{ij} = 0 \text{ si } i > r \text{ y } j \leq r, \text{ entonces}$$

Por demostrar que es válido $q - 1$

$$\begin{aligned}
 \langle A^{q-1} \rangle_{ij} &= \langle A^{q-2} A \rangle_{ij} = \sum_{k=1}^n \langle A \rangle^{q-2}_{ik} \langle A \rangle_{kj} = \sum_{k=1}^r \langle A \rangle^{q-2}_{ik} \langle A \rangle_{kj} + \sum_{k=r+1}^n \langle A \rangle^{q-2}_{ik} \langle A \rangle_{kj} \\
 &= \sum_{k=1}^r \langle A \rangle^{q-2}_{ik} a_{kj} + \sum_{k=r+1}^n \langle A \rangle^{q-2}_{ik} a_{kj} \text{ en donde, por hipótesis de inducción}
 \end{aligned}$$

$$\underbrace{\sum_{k=1}^r \langle A \rangle^{q-2}_{ik} a_{kj}}_0$$

y como $a_{ij} = 0$ si $i > r$ y $j \leq r$, entonces

$$\underbrace{\sum_{k=r+1}^n \langle A \rangle^{q-2}_{ik} a_{kj}}_0$$



por lo tanto es cierto para $|A|^{q-1}$

Por lo tanto $|A|^{n-1}$ todas tienen el mismo bloque de $(n-r) \times r$ de ceros

Lema 1.2. Si $P \in M_n$ es una matriz de permutación $|A| = |PAP^T| = P|A|P^T$

Prueba.

Sean $A \in M_{n \times n}$ y $P \in M_n$ una matriz de permutación, entonces

$$|PAP^T| = \langle |PAP^T| \rangle_{ij} = \sum_{k=1}^n \langle PA \rangle_{ik} \langle P^T \rangle_{kj} = \sum_{k=1}^n \sum_{s=1}^n \langle P \rangle_{is} \langle A \rangle_{sk} \langle P^T \rangle_{kj}$$

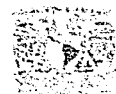
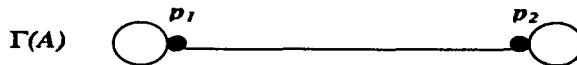
Por la **Definición 1.6** tenemos $|P| = P$ debido a que exactamente una entrada en cada renglón y columna es igual a uno y en todas las demás son cero. Así

$$\sum_{k=1}^n \sum_{s=1}^n \langle P \rangle_{is} \langle A \rangle_{sk} \langle P^T \rangle_{kj} = \sum_{k=1}^n \langle P \rangle_{ik} \langle A \rangle_{kj} = \langle P|A|P^T \rangle_{ij} = P|A|P^T$$

Definición 1.10. La gráfica no dirigida de $A \in M_n$ denotada por $\Gamma(A)$, es la gráfica no dirigida en n nodos p_1, p_2, \dots, p_n tal que existe una arista en $\Gamma(A)$ de p_i a p_j si y sólo si $a_{ij} \neq 0$ ($\mu_{ij} \neq 0$).

Ejemplos

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix};$$



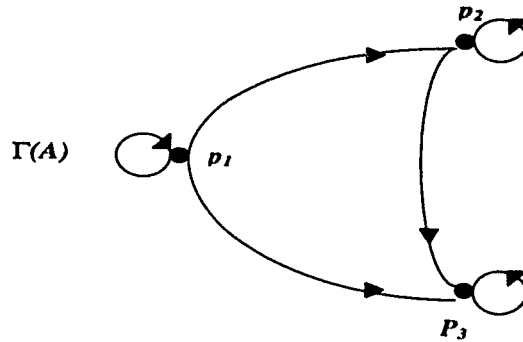
$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix};$$



$$A = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix};$$



$$A = \begin{bmatrix} 3 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$$



Con lo cual tenemos el siguiente,

Teorema 1.1. Una matriz $A \in M_n$ es irreducible si y sólo si $(I + |A|)^{n-1} > 0$ o equivalentemente, si $[I + M(A)]^{n-1} > 0$.

Prueba.



Probaremos que A es reducible si y sólo si $(I + |A|)^{n-1}$ tiene al menos una entrada 0. Supongamos primero que A es reducible, para alguna matriz de permutación P tenemos.

$$A = P \begin{bmatrix} B & C \\ 0 & D \end{bmatrix} P^T = P \tilde{A} P^T$$

Por el **Lema 1.2.** $|A| = |P \tilde{A} P^T| = P |\tilde{A}| P^T$ y por el **Lema 1.1.** $|\tilde{A}|^2, |\tilde{A}|^3, \dots, |\tilde{A}|^{n-1}$ tienen el mismo bloque 0 de $(n-r) \times r$ en la esquina inferior izquierda como la matriz \tilde{A} , entonces

$$\begin{aligned} (I + |A|)^{n-1} &= (I + P |\tilde{A}| P^T)^{n-1} = (P I P^T + P |\tilde{A}| P^T)^{n-1} = (P [I + |\tilde{A}|] P^T)^{n-1} = P (I + |\tilde{A}|)^{n-1} P^T \\ &= P \left[I + (n-1) |\tilde{A}| + \binom{n-1}{2} |\tilde{A}|^2 + \dots + \binom{n-1}{n-1} |\tilde{A}|^{n-1} \right] P^T \end{aligned}$$

y todos los términos del interior de los corchetes tienen bloques 0's de $(n-r) \times r$ en la esquina inferior izquierda, esto es, $(I + |A|)^{n-1}$ es reducible y tiene, por lo tanto entradas cero.

Inversamente, supongamos que para alguna $r \neq s$ la entrada (r, s) de $(I + |A|)^{n-1}$ es igual a cero. Se afirma que no existe un camino dirigido en $\Gamma(A)$ de p_r a p_s de longitud $n-1$. Definiendo el conjunto de nodos de la siguiente manera:

$$S_1 \equiv \{ p_i : p_i = p_s \text{ o éste es un camino en } \Gamma(A) \text{ de } p_i \text{ a } p_s \}$$

y sea S_2 que contiene a todos los nodos de $\Gamma(A)$ que no están en S_1 . Resaltando que $S_1 \cup S_2 = \{ p_1, p_2, \dots, p_n \}$ y $p_s \in S_1 \neq \emptyset$, así $S_2 = \{ p_1, p_2, \dots, p_n \}$. Si existiera un camino de alguno de los nodos p_i de S_2 a algún nodo p_j de S_1 , entonces (por definición de S_1) podría existir un camino de p_i a p_s y así p_j podría estar en S_1 .

Ahora, reetiquetando los nodos tenemos $S_1 = \{\tilde{p}_1, \tilde{p}_2, \dots, \tilde{p}_r\}$ y

$S_2 = \{\tilde{p}_{r+1}, \tilde{p}_{r+2}, \dots, \tilde{p}_n\}$ y notando que

$$\tilde{A} = P^T A P = \begin{bmatrix} B & C \\ \mathbf{0} & D \end{bmatrix}; \quad B \in M_r, \mathbf{0} \in M_{n-r}$$

Así que A es reducible.

El argumento de $[I + M(A)]^{n-1} > \mathbf{0}$ es justamente el mismo.

Definición 1.10. Una matriz $A = [a_{ij}] \in M_n$ se dice que tiene la propiedad de ser fuertemente conexa (**SC**) si para cada par distinto de enteros p, q con $1 \leq p, q \leq n$ existe una secuencia de enteros distintos $k_1 = p, k_2, k_3, \dots, k_{m-1}, k_m = q$, $1 \leq m \leq n$ tal que todas las entradas de la matriz $a_{k_1, k_2}, a_{k_2, k_3}, \dots, a_{k_{m-1}, k_m}$ son distintas de cero,

Definición 1.11. Una gráfica no dirigida Γ es **SC** si entre cualquier par distinto de nodos p_i, p_j en Γ existe aristas de longitudes finitas que inician en p_i y terminan en p_j .

Teorema 1.2. Sea $A \in M_n$. A tiene la propiedad **SC** si y sólo si la gráfica dirigida $\Gamma(A)$ es **SC**.

Prueba.

Supongamos que A es **SC**

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

y como A es SC, entonces para cada entrada (p, q) existe una sucesión $a_{pk_2}, a_{k_2k_3}, \dots, a_{k_{m-1}q} \neq 0$

Sea $\varphi_{pk_2}, \varphi_{k_2k_3}, \dots, \varphi_{k_{m-1}q}$ las flechas asociadas en $\Gamma(A)$

$\gamma = \varphi_{pk_2}, \varphi_{k_2k_3}, \dots, \varphi_{k_{m-1}q}$ es un camino dirigido de longitud finita que inicia en p_p y termina en p_q .

Teorema 1.3. Sea $A \in M_n$ y sean p_i y p_j vértices de $\Gamma(A)$. Existe un camino de longitud m en $\Gamma(A)$ de p_i a p_j si y sólo si $(|A|^m)_{ij} \neq 0$ o equivalentemente $[M(A)^m]_{ij} \neq 0$.

Prueba.

\Rightarrow

La demostración se hará por inducción sobre m

Iniciamos para $m=1$. Supongamos que existe un camino de longitud 1 de p_i a p_j , esto es $|A|_{ij} \neq 0$.

Supongamos que existe γ un camino de longitud 2 de p_i a p_j , es decir, existe

φ_{ik} de p_i a p_k para alguna $k \in \{1, 2, \dots, n\}$ y $k \neq i, j$



Φ_{kj} de p_k a $p_j \Rightarrow |a_{ij}| \neq 0$ y $|a_{ij}| \neq 0$

$$|A|_{ij}^2 = \sum_{s=1}^n |a_{is}| |a_{sj}| = \sum_{s=1}^n |a_{is}| |a_{sj}| + \underbrace{|a_{ik}| |a_{kj}|}_{\neq 0} \neq 0$$

Continuamos para $m=3$

Existe γ un camino de longitud 3 de p_i a p_j , es decir existen flechas

Φ_{ik} de p_i a p_k para alguna $k \in \{1, 2, \dots, n\}$ $k \neq i, j$

Φ_{kl} de p_k a p_l para alguna $l \in \{1, 2, \dots, n\}$ $l \neq i, j$

Φ_{kj} de p_k a p_j de tal manera que $\gamma = \Phi_{ik} \Phi_{kl}$

$\Rightarrow |a_{ik}| \neq 0$ y $|a_{kl}| \neq 0$ $|a_{ij}| \neq 0$

Con lo que

$$|A|_{ij}^3 = \sum_{s=1}^n \sum_{r=1}^n |a_{is}| |a_{sr}| |a_{rj}| = \sum_{s=1}^n \sum_{r=1}^n |a_{is}| |a_{sr}| |a_{rj}| + \underbrace{|a_{ik}| |a_{kl}| |a_{lj}|}_{\neq 0} \neq 0$$

Suponemos que se cumple para $m=q$ y demostraremos que es válido para cuando $m=q+1$

Sea γ un camino de longitud $q+1$ de p_i a p_j , es decir existe

Φ_{ik_1} de p_i a p_{k_1}

$\Phi_{k_1 k_2}$ de p_{k_1} a p_{k_2}

$\Phi_{k_2 k_3}$ de p_{k_2} a p_{k_3}



ésto implica que para alguna $t \in \{1, 2, 3, \dots, n\}$ $a_{it}a_{tj} \Rightarrow a_{it} \neq 0$ y $a_{tj} \neq 0$, lo cual implica que existe una flecha

Φ_{it} de p_i a p_t para alguna $t \in \{1, 2, \dots, n\}$ y $t \neq i, j$ y otra

Φ_{tj} de p_t a $p_j \Rightarrow |a_{tj}| \neq 0$ y $|a_{ij}| \neq 0$

Sea $\gamma = \Phi_{it} \Phi_{tj}$ este es un camino de longitud 2 de p_i a p_j .

Ahora suponemos que se cumple para cuando $m=q$ y por demostrar que es válido para $m=q+1$

$$(|A|^{q+1})_{ij} = (|A|^q |A|)_{ij} = \sum_{t=1}^n (|A|^q)_{it} |a_{tj}| \neq 0$$

Para alguna $t \in \{1, 2, \dots, n\}$ y $t \neq i, j$ $(|A|^q)_{it} |a_{tj}| \neq 0$, con lo que $(|A|^q)_{it} \neq 0$ y $|a_{tj}| \neq 0$

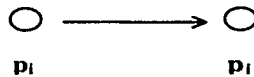
por hipótesis de inducción, existe un camino $\gamma' = \Phi_{ik_1} \Phi_{k_1 k_2} \Phi_{k_2 k_3} \dots \Phi_{k_{q-1} i}$ de longitud q y como $|a_{tj}| \neq 0$, entonces existe una flecha $\gamma = \gamma' \Phi_{tj}$ es un camino de longitud $q+1$ de p_i a p_j .

Teorema 1.4. Sea $A \in M_n$. Los siguientes incisos son equivalentes

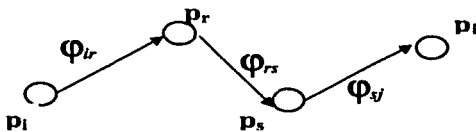
- A es irreducible
- $(I + |A|)^{n-1} > 0$
- $[I + M(A)]^{n-1} > 0$
- $\Gamma(A)$ es fuertemente conexa; y
- A tiene la propiedad **SC**



Definición 1.9a. Sean $\{p_1, p_2, \dots, p_n\}$ vértices de una gráfica. Si existe una flecha de p_i a p_j entonces $a_{ij} \neq 0$.



Denotaremos con φ_{ij} a la flecha que tiene punto inicial p_i y punto final p_j



Definición 1.10. Un camino dirigido γ en una gráfica Γ es una secuencia de flechas $p_{i1}p_{i2}, p_{i2}p_{i3}, p_{i3}p_{i4}, \dots$ en Γ . La lista de nodos en el camino dirigido γ es $p_{i1}p_{i2}, p_{i2}p_{i3}, p_{i3}p_{i4}, \dots$. La longitud de un camino dirigido es el número de arcos o flechas sucesivas en el camino dirigido si este número es finito; de otro modo, el camino dirigido se dice que tiene longitud infinita. Un ciclo es un camino que inicia y termina en el mismo nodo; este nodo tiene lugar exactamente dos veces en el listado ordenado de nodos en el camino, y no otro nodo ocurre más que alguna vez en la lista.

Definición 1.12. La composición de flechas se define de la siguiente manera:

$$\varphi_{in}\varphi_{st} = \begin{cases} \varphi_{in}\varphi_{st} & \text{si } n = s \\ 0 & \text{si } n \neq s \end{cases}$$

Un camino dirigido γ en $\Gamma(A)$ es una composición de flechas

$$\gamma = \varphi_{s_1s_2} \varphi_{s_2s_3} \varphi_{s_3s_4} \varphi_{s_4s_5} \dots$$

Si la composición de flechas es finita, γ es de longitud finita

Teorema 1.5. Una matriz $A \in M_n$ tiene la propiedad SC si y sólo si $|A|$ o $M(A)$, y de aquí ambos, tienen la propiedad SC.

Prueba.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

En donde, para cualquier par (p, q) podemos encontrar una secuencia $a_{k_1 k_2}, a_{k_2 k_3}, \dots, a_{k_{m-1} k_m}$ diferente de cero, así si a está secuencia le aplicamos valor absoluto, obtenemos;

$$|A| = \begin{bmatrix} |a_{11}| & |a_{12}| & \cdots & |a_{1n}| \\ |a_{21}| & |a_{22}| & \cdots & |a_{2n}| \\ \vdots & \vdots & \ddots & \vdots \\ |a_{n1}| & |a_{n1}| & \cdots & |a_{nn}| \end{bmatrix}$$

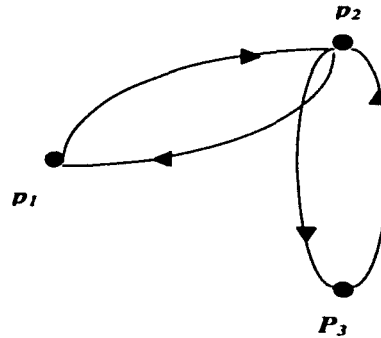
en donde para cualquier (p, q) tenemos la secuencia $|a_{k_1 k_2}|, |a_{k_2 k_3}|, \dots, |a_{k_{m-1} k_m}|$ en donde $|a_{k_i k_j}| \neq 0$.

El concepto de una secuencia de entradas distintas de cero de $A \in M_n$ que se originan al establecer la propiedad SC puede ser resumida en términos de ciertos caminos en una gráfica asociada con A .

Ahora mostraremos que Γ es SC si ésta tiene la propiedad que cualquier par de nodos pertenecen al menos a un ciclo, pero no al revés.

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix};$$

$\Gamma(A)$



$$p_1 = p_1 p_2 p_1 \circ p_1 p_2 p_3 p_2 p_1$$

$$p_3 = p_3 p_2 p_3 \circ p_3 p_2 p_1 p_2 p_3$$

Puede existir más de un camino dirigido entre dos nodos de una gráfica dirigida, pero tales dos caminos con diferentes longitudes pueden no ser esencialmente diferentes; uno puede contener repeticiones de uno o más subcaminos. Esto es claro que si uno siempre visita un nodo dado dos veces haciendo por un camino dirigido, entonces el camino dirigido puede ser reducido (y el punto final podría ser indiferente) al suprimir todos los arcos intermedios entre el primero y el segundo visita al nodo (la subgráfica suprimida es, o contiene, un ciclo).

Observación. Sea Γ una gráfica dirigida en n nodos. Si existe un camino dirigido en Γ entre 2 nodos dados, entonces entre estos nodos existe un camino dirigido que tiene longitud no mayor de $n-1$. ¿Cómo puede uno decir si una matriz dada tiene la propiedad SC?. Si n no es grande o si $M(A)$ tiene una estructura especial, entonces uno puede sólo inspeccionar $\Gamma(A)$ y trazar caminos entre todos los posibles pares de nodos. Sin embargo, generalmente, esto no es práctico, así necesitamos algún método explícito de calcular.

* * * * *



.....
 Φ_{kqj} de p_{kq} a p_j

de tal modo que $\gamma = \Phi_{ik1} \Phi_{k1k2} \Phi_{k2k3} \dots \Phi_{kqj} \Rightarrow$

$a_{ik1}, a_{k1k2}, a_{k2k3}, \dots, a_{kqj} \neq 0$

y como $\gamma' = \Phi_{ik1} \Phi_{k1k2} \Phi_{k2k3} \dots \Phi_{kq-1kq}$ es un camino de longitud q , entonces por

hipótesis de inducción $(|A|^q)_y \neq 0$ así

$$(|A|^{q+1})_y = (|A|^q |A|)_y = \sum_{s=1}^n (|A|^q)_s |a_{sj}| = \sum_{s=1}^n (|A|^q)_s |a_{sj}| + \underbrace{(|A|^q)_{k_q} |a_{k_q j}|}_{\neq 0} \neq 0$$

\therefore si existe un camino de longitud m de p_i a p_j con $(|A|^m)_y \neq 0$

\Leftarrow

Supongamos que $(|A|^m)_y \neq 0$. La demostración se hará por inducción sobre m

Para $m=1$. Supongamos que $|a_{ij}| \neq 0$, ésto implica que existe una flecha $\Phi_{ij} \neq 0$

en donde $\gamma = \Phi_{ij}$ es un camino de longitud 1 de p_i a p_j .

Para $m=2$

$$(|A|^2)_y = \sum_{s=1}^n |a_{is}| |a_{sj}| \neq 0$$

