

Integración Aparente de Datos Utilizando Intranets

Tesis presentada

por

Francisco Alfonso Aguilar Valenzuela

**Presentada ante la Dirección Académica de la Universidad Virtual del
Instituto Tecnológico de Estudios Superiores de Monterrey
como requisito parcial para optar
al Título de**

MAESTRO EN ADMINISTRACION DE SISTEMAS DE INFORMACION

Diciembre de 1999.

Maestría en Administración de Sistemas de información

Integración Aparente de Datos Utilizando Intrenets

Tesis presentada

por

Francisco Alfonso Aguilar Valenzuela

**Presentada ante la Dirección Académica de la Universidad Virtual del
Instituto Tecnológico de Estudios Superiores de Monterrey
como requisito parcial para optar
al Título de**

MAESTRO EN ADMINISTRACION DE SISTEMAS DE INFORMACION

Diciembre de 1999.

Maestría en Administración de Sistemas de información

DEDICATORIA

A Jesús: Mi guía y ejemplo a seguir.
A mi esposa Rosalba: Por ser mi gran amor, compañera y amiga.
A Jaquelyn y Alfonso: Mis grandes tesoros.
A mis padres: Por su ejemplo y amor.

AGRADECIMIENTOS

Antes que a nadie, agradezco a Dios el haberme dado la fuerza, salud y apoyo para salir adelante en todo el proceso que significó el estudio de mi programa de maestría y la elaboración de este trabajo de tesis.

A mi director de tesis Dr. José I. Icaza le agradezco su gran apoyo, paciencia y acertada dirección, que sin ellos no me habría sido posible realizar este trabajo de tesis.

A mis sinodalès Yael Robles, Benjamín Parra y Agustín González por su apoyo y consejos.

Agradezco también al Centro de Investigación en Alimentación y Desarrollo, A.C., especialmente al Dr. Inocencio Higuera Ciapara por el apoyo brindado, impulsándome siempre a seguirme superando.

Al Ing. Alfonso Pompa, director del Campus Sonora Norte, por su apoyo económico mediante una beca de complemento de colegiatura.

A todos mis compañeros y amigos del Departamento de Servicios de Apoyo del C.I.A.D.,A.C. por el respaldo que me brindaron para facilitarme la asistencia a los cursos y realización de trabajos.

Al personal del ITESM Campus Sonora Norte, de la Biblioteca Edilia Ballesteros, de la UV, Edith Encinas y José Luis.

A mis padres, por la educación que me dieron y el deseo de superación y esfuerzo que me inculcaron.

A mi esposa e hijos por el tiempo que les quité con el fin de dar este paso en mi vida.

RESUMEN

INTEGRACION APARENTE DE DATOS UTILIZANDO INTRANETS

DICIEMBRE 1999

FRANCISCO ALFONSO AGUILAR VALENZUELA

INGENIERO INDUSTRIAL EN ELECTRONICA
INSTITUTO TECNOLOGICO DE HERMOSILLO

MAESTRO EN ADMINISTRACION DE SISTEMAS DE INFORMACION
INSTITUTO TECNOLOGICO DE ESTUDIOS SUPERIORES DE MONTERREY
CAMPUS SONORA NORTE

DIRIGIDO POR EL PROFESOR DR. JOSE I. ICAZA ACERETO

En este trabajo de tesis, se introduce el concepto de integración aparente de datos, que consiste en presentar a los usuarios una vista integrada de los datos empresariales, lograda a partir de copias y transformaciones de los datos operativos, los cuales no se encuentran integrados. La utilización de la metodología propuesta para lograr un almacén de datos (data warehouse), constituye una de las formas para lograr la integración aparente, entre otras que se presentan. Además, se consideran las ventajas que ofrece la tecnología de intranet en el manejo de los datos empresariales, poniendo especial atención en los atributos que facilitan la implementación de la integración aparente.

Los resultados se ilustran con un caso práctico, llevado a cabo en el Centro de Investigación en Alimentación y Desarrollo, A.C. (CIAD), en el que se aprecia la evolución de la Tecnología de Información en dicho centro y las causas que provocaron la desintegración de los datos, así como también los problemas relacionados por dicha desintegración.

Por último se hace un análisis de las necesidades de los usuarios con respecto a la integración de datos y se diseña la estrategia de integración, apoyados en la metodología del almacén de datos y utilizando la tecnología de la intranet.

INDICE

| | |
|---|-----|
| "INTEGRACION APARENTE DE DATOS UTILIZANDO INTRANETS"..... | i |
| Integración Aparente de Datos Utilizando Intranets. Página de Firmas..... | ii |
| DEDICATORIA | iii |
| AGRADECIMIENTOS | iv |
| RESUMEN | v |
| INDICE | vi |
| INDICE DE CUADROS | ix |
| INDICE DE FIGURAS | x |
| INTRODUCCION | 1 |
| Metodología | 3 |
| CAPITULO I: LA INFORMACION..... | 5 |
| El valor de la información..... | 5 |
| Datos e Información..... | 5 |
| Calidad de la información..... | 7 |
| Exacta..... | 7 |
| Oportuna..... | 8 |
| Actual..... | 8 |
| Accesible | 8 |
| Enfocada..... | 8 |
| Dentro del Rango..... | 9 |
| Integrada..... | 9 |
| Segura | 9 |
| La Información y la Tecnología de Información | 10 |
| CAPITULO II: ADMINISTRACION DE LOS DATOS | 13 |
| Fuentes de Información | 14 |
| Las 5 Fuerzas Competitivas de Porter..... | 15 |
| La cadena de valor de Porter..... | 16 |
| Factores Críticos de Exito..... | 17 |
| Integración de datos..... | 18 |
| Beneficios de la integración de los datos..... | 21 |
| Integración de datos. ¿Cómo se puede lograr? | 24 |
| Acciones arriba-abajo | 25 |
| Acciones de abajo-arriba | 27 |
| Integración aparente de los datos | 28 |
| Grado de integración de datos..... | 28 |
| Incertidumbre y Ambigüedad | 29 |
| Diferencia entre el Administrador de Datos y el Administrador de Base de Datos | 32 |
| El mayordomo de los datos (data steward)..... | 33 |
| El área de trabajo de un Mayordomo de los Datos..... | 34 |
| Conclusión | 34 |
| CAPITULO III. INTRANETS | 36 |
| INTRANET: Definiciones..... | 36 |
| Historia de la Intranet y el World Wide Web..... | 37 |

| | |
|--|-----------|
| Intranets: concepto | 41 |
| Características principales de una intranet | 44 |
| Facilidad de Uso | 44 |
| Creación rápida de prototipos..... | 46 |
| Escalable | 46 |
| El concepto de cliente universal | 47 |
| Integración de estrategias de cómputo distribuido..... | 49 |
| Aprovechamiento de fuentes de información actual..... | 49 |
| Acceso a Bases de Datos..... | 50 |
| Nuevas Posibilidades..... | 51 |
| Aspectos importantes para crear una intranet..... | 51 |
| Planeación | 51 |
| Diseño..... | 52 |
| Acceso a Datos..... | 52 |
| Herramientas | 53 |
| Integración | 53 |
| Seguridad | 53 |
| Elevar la velocidad de acceso a usuarios remotos..... | 54 |
| Formar el Equipo -- Incluir a los Generadores de la Información..... | 54 |
| Promoción..... | 55 |
| Desarrollar una Política de Publicación en la Intranet..... | 55 |
| CAPITULO IV. LA TECNOLOGIA DE INTRANET Y LA INTEGRACION DE | |
| LOS DATOS..... | 57 |
| El impacto de la tecnología de intranet y las metodologías tradicionales para lograr la integración de la información | 57 |
| Características principales de una Intranet y su relación con la integración de datos..... | 58 |
| Acceso a información en diferentes sistemas operativos..... | 59 |
| Facilidad de hacer cambios y modificaciones a la información desde cualquier plataforma de hardware y software..... | 59 |
| Menor costo de software..... | 59 |
| Menor costo de entrenamiento y menor tiempo de aprendizaje..... | 60 |
| Menor costo de mantenimiento..... | 60 |
| Análisis de la tecnología de intranet como una herramienta para la integración de los datos en las empresas..... | 60 |
| Factores que influyen en la integración de la información | 62 |
| El volumen de información..... | 62 |
| El ambiente en el cual se desarrolla la empresa | 64 |
| El nivel de interdependencia entre las unidades..... | 66 |
| El grado de homogeneidad del software y hardware utilizado..... | 67 |
| El factor económico..... | 70 |
| CAPITULO V: ALMACEN DE DATOS (DATA WAREHOUSE) | 72 |
| Colección de datos orientada a temas..... | 73 |
| Colección de datos Integrada | 73 |
| Colección de datos Volátil..... | 74 |
| Colección de datos Variable en el Tiempo..... | 74 |
| La Estructura del Almacen de Datos..... | 75 |

| | |
|---|-----|
| Granularidad | 78 |
| Partición..... | 81 |
| Metadatos | 81 |
| Modelos de Datos y de Procesos y el Almacen de Datos..... | 82 |
| El Almacen de Datos y los Modelos de Datos | 83 |
| Data mining | 84 |
| Datamart | 85 |
| La Tecnología del Web y el Almacen de Datos | 87 |
| CAPITULO VI. Caso práctico CIAD..... | 90 |
| ¿Qué es el CIAD?..... | 90 |
| Historia de los sistemas en el CIAD..... | 92 |
| El sistema actual..... | 94 |
| Planteamiento del problema del CIAD ocasionado por la desintegración de los datos..... | 95 |
| Proceso de adquisiciones | 96 |
| Caso Práctico. ¿Cómo puede ayudar la tecnología de intranet a resolver el problema de desintegración de los datos?..... | 99 |
| Requerimientos de Información | 101 |
| SOLICITANTE..... | 101 |
| Compras..... | 103 |
| Presupuestos | 103 |
| Almacen | 104 |
| Contabilidad | 105 |
| Directivos del Centro..... | 106 |
| Fuentes de Información | 108 |
| Extracción de datos del ambiente operativo | 112 |
| Seguridad..... | 112 |
| Metadatos | 113 |
| La tecnología de intranet como apoyo para crear el diccionario de datos .. | 114 |
| CONCLUSIONES..... | 116 |
| BIBLIOGRAFIA | 118 |
| CURRICULUM VITAE | 120 |

INDICE DE CUADROS

| | | |
|-------------|--|-----|
| Cuadro 2.1 | Ejemplo de desintegración de la Información | 19 |
| Cuadro 2.2 | El impacto de la integración de los datos en ambientes cambiantes..... | 31 |
| Cuadro 4.1 | Importancia de la calidad de los datos de entrada en los resultados de la calidad de la información de salida de los sistemas | 61 |
| Cuadro 4.2 | Impacto del volumen de información de los sistemas en relación con la integración de los datos y el impacto de la tecnología de intranet. Considerando que no existe planeación estratégica de datos | 63 |
| Cuadro 4.3 | Impacto del ambiente en el que se desenvuelve la empresa en relación con la integración de los datos y el impacto de la tecnología de intranet | 64 |
| Cuadro 4.4 | Impacto de la dependencia entre las unidades en relación con la integración de los datos y el impacto de la tecnología de intranet | 66 |
| Cuadro 4.5 | Impacto del ambiente en el cual estan implantados los sistemas en relación con la integración de los datos y el impacto de la tecnología de intranet. Considerando el mismo número de usuarios | 68 |
| Cuadro 5.1 | Medios de almacenamiento de datos, su velocidad y precios..... | 80 |
| Cuadro 6.1 | Muestra de Algunas Partidas, del Catálogo Presupuestal del CIAD | 97 |
| Cuadro 6.2 | Situación de las requisiciones solicitadas por el usuario | 101 |
| Cuadro 6.3 | Situación del Presupuesto con Información de requisiciones en tránsito | 102 |
| Cuadro 6.4 | Situación de los proveedores: tiempos de entrega y pago | 103 |
| Cuadro 6.5 | Reporte de pagos efectuados por contabilidad | 104 |
| Cuadro 6.6 | Requisiciones asociadas a las ordenes de compra y su estatus | 105 |
| Cuadro 6.7 | Descripción del contenido de la orden de compra | 105 |
| Cuadro 6.8 | Reporte de articulos recibidos en almacen y precios establecidos por compras para comparación con factura | 106 |
| Cuadro 6.9 | Requerimientos de información de los altos ejecutivos (con datos falsos, debido a la naturaleza de la información presentada) | 107 |
| Cuadro 6.10 | Ejemplo de un artículo en el diccionario de datos o Metadatos | 114 |

INDICE DE FIGURAS

| | | |
|------------|--|-----|
| Figura 2.1 | Modelo de Porter de las Fuerzas Competitivas | 15 |
| Figura 2.2 | Cadena de Valor de Porter | 17 |
| Figura 2.3 | Diferentes manifestaciones de desintegración de datos (Inmon, 1996) | 21 |
| Figura 3.1 | Planes corporativos para el uso de Intranet | 43 |
| Figura 4.1 | Sistema de procesamiento de datos | 61 |
| Figura 5.1 | La estructura del almacén de datos (Inmon, 1996) | 75 |
| Figura 5.2 | Data Marts (Miramón, 1996) | 86 |
| Figura 5.3 | Almacén de Datos Virtual | 88 |
| Figura 6.1 | El CIAD,A.C. y sus unidades foráneas | 91 |
| Figura 6.2 | Distribución del personal del CIAD,A.C. | 92 |
| Figura 6.3 | Presiones que enfrenta el sistema actual del CIAD. | 95 |
| Figura 6.4 | Diagrama de Flujo del Proceso de Adquisiciones en el CIAD | 98 |
| Figura 6.5 | Fuente de Información y transformaciones para cubrir las necesidades de los tomadores de decisiones | 100 |
| Figura 6.6 | Flujo y Transformación de los Datos desde el Ambiente Operativo | 109 |
| Figura 6.7 | Ejemplo de algunos archivos fuente y sus transformaciones | 111 |

INTRODUCCION

En los últimos 25 años, el mundo ha evolucionado de una economía industrial a una economía de información, y en las siguientes décadas, la información más que la tierra o el capital dirigirá la creación del bienestar y la prosperidad. En esta clase de economía, el éxito está determinado más por el saber que por el tener y la competencia se basa en la habilidad de adquirir, manipular, interpretar y usar la información en forma efectiva (McGee et al, 1993).

El problema de la desintegración de los datos, es un problema real y se observa en gran cantidad de empresas. Existe una amplia variedad de causas que pueden dar origen a la desintegración de los datos, como por ejemplo, los intereses de manejo de información independiente, mostrados por las diferentes áreas de la empresa, enfocados principalmente, a solucionar sus necesidades. Otras causas son el crecimiento no planeado de la empresa, el diseño de sistemas que promueve la creación de islas de información, sin considerar a la empresa como un todo. Asimismo, la utilización de tecnología que no permite un flujo de información efectivo, la cual depende en gran medida de los sistemas operativos y los equipos utilizados.

Existen varias formas para obtener la integración de los datos dentro de la empresa, pero este trabajo se enfocará principalmente en mostrar como la tecnología utilizada en intranet, puede servir como una herramienta para lograr lo que se conoce como "*integración aparente*".

La tecnología de intranet, nos ofrece la posibilidad de lograr la integración aparente de los datos empresariales, pero como toda tecnología, ésta debe de considerarse como una herramienta para lograr un determinado trabajo. Es preciso apoyar esta tecnología con una planeación de datos que asegure su buen manejo, optimización e integración para lograr el objetivo de brindar información de calidad a los tomadores de decisiones a cualquier nivel dentro de la empresa.

En este trabajo se muestra la importancia que tiene la información y la forma en la cual la integración de los datos ayuda en gran medida a generar

información de calidad, como apoyo para la toma de decisiones dentro de una empresa.

Se definirá el concepto de integración aparente y se ilustrará un método para lograrla, combinando la tecnología de intranets con la metodología del almacén de datos (data warehouse). Este método se ejemplificará por medio del análisis de la situación que se presenta en el Centro de Investigación en Alimentación y Desarrollo, A.C.

En el primer capítulo, se verá principalmente la importancia que la información ha adquirido dentro de las empresas y la evolución y cambios que éstas han tenido que hacer con el fin de manejarla como un recurso, también se verá la diferencia que existe entre la información y los datos, los que a menudo son considerados como lo mismo. También en este capítulo se verán los principales aspectos de calidad que la información debe poseer con el fin de convertirse en una herramienta confiable para la toma de decisiones y por último en este capítulo también se analizará la diferencia entre la información y la Tecnología de información.

Una vez que los ejecutivos se dan cuenta de la importancia que la información juega en sus negocios, éstos deben de implementar una serie de estrategias con el fin de aprovechar al máximo las bondades de éste recurso. El capítulo II analiza algunos aspectos que se deben de tomar en cuenta para obtener información estratégica, para la empresa, es decir, el análisis de las diferentes fuentes de información. También, en este capítulo se ve a más detalle el concepto de integración de los datos e integración aparente, y por último se hace un pequeño análisis de los actores principales que deben llevar a cabo este tipo de trabajos encaminados a mejorar la calidad de la información.

La tecnología de intranet, ha revolucionado el mundo de la información empresarial, ofreciendo una serie de ventajas para un manejo más fluido de información a través de la empresa. En el capítulo III se verán algunas de las ventajas que nos ofrece dicha tecnología.

En el capítulo IV, debido al enfoque de este trabajo, es preciso hacer un análisis de las ventajas de la tecnología de intranet, pero en relación con el manejo de la información y poniendo principal atención en la forma en la cual esta tecnología se puede convertir en una herramienta confiable para lograr la integración total, parcial o aparente de los datos empresariales.

El capítulo V analiza la metodología del almacén de datos, su estructura y principales componentes y su relación con la integración aparente de datos.

En el capítulo VI, se hará el análisis de la situación de una empresa con datos desintegrados, las problemáticas derivadas de dicha desintegración y las ventajas que podrían obtenerse de una acción de integración aparente de datos, partiendo de datos que se encuentran realmente desintegrados, aprovechando las ventajas de la tecnología de intranet y apoyándose en la metodología del almacén de datos, con el fin de aprovechar la inversión actual tanto de hardware como de software y lograr una fuente confiable de información para apoyo en la toma de decisiones.

Metodología

La metodología utilizada en este trabajo de tesis se basa principalmente en métodos cualitativos. Los primeros 5 capítulos de este trabajo están apoyados en una revisión bibliográfica con el fin de sentar las bases de la importancia de la información dentro de las empresas, mostrar las características principales de la tecnología de intranet y la metodología del almacén de datos.

En el último capítulo se estudia el caso particular de una empresa y se hace un análisis del grado de desintegración de datos en la misma y la forma en la cual dicha empresa se podría ver beneficiada al optar por el uso de la tecnología de intranet y la metodología del almacén de datos.

Para este caso, la recolección de la información se obtuvo principalmente de entrevistas tanto de los usuarios de los servicios administrativos como del personal administrativo. Con estas entrevistas se logró hacer una evaluación de los problemas relacionados con la desintegración de la información en el CIAD.

También se utilizó la observación para entender el funcionamiento de la empresa, sin participar directamente en las actividades, pero teniendo la posibilidad de libre movimiento en las diferentes áreas y teniendo la facilidad para hacer preguntas con el fin de satisfacer los requerimientos de información de apoyo para este trabajo.

CAPITULO I: LA INFORMACION

Actualmente nos encontramos sumergidos en la era de la información, por lo tanto, el valor que ésta representa para los ejecutivos de las empresas, es cada vez mayor. Ante el nivel competitivo actual, los ejecutivos de las empresas se enfrentan a decisiones con un alto nivel de complejidad, precisión y repercusión, por lo que es indispensable apoyar dichas decisiones con información de calidad.

Para que la información se considere de calidad, ésta debe reunir ciertas características, las cuales, se analizarán a continuación en este capítulo.

El valor de la información

La información representa el producto principal de cualquier sistema de información. En estos días se habla mucho sobre los grandes beneficios derivados del uso de la tecnología de información y la aplicación de grandes sistemas como es el caso de los Sistemas de Soporte a la Toma de Decisiones (Decisión Support Systems - DSS) y los Sistemas de Información Ejecutivos (Executive Information Systems - EIS), entre otros. También se ha observado o confirmado las ventajas competitivas que con el apoyo de dicha tecnología, han logrado empresas como WalMart, Federal Express, Frito Lay, American Airlines, etc. Pero no hay que perder de vista el pilar principal sobre el cual están basados estos sistemas, "**la información**", y no hay que olvidar que "si alimentas a tu sistema con basura obtendrás como resultado basura" (Garbage In Garbage Out - GIGO).

Datos e Información

Antes de profundizar en el tema, es necesario comprender la diferencia que existe entre los datos y la información. Existen varios conceptos en relación a sus significados, como se muestra a continuación.

La información es un recurso valioso para la empresa, pero los datos por ellos mismos representan solo costos (adquisición, almacenamiento, etc). Hasta que se convierten en información, los datos no tienen un beneficio directo.

Es común que se utilicen los términos datos e información de manera indistinta, pero la gente relacionada con sistemas de información hace una distinción muy cuidadosa entre estos conceptos. Se ha llegado a la conclusión de que los datos guardados o impresos en reportes regularmente no proporcionan una buena fuente de información para la empresa. Los datos no son mas que hechos y figuras grabados; información es conocimiento derivado de los datos. Ningún dato por si solo puede ser información. Por ejemplo el número 36 no es mas que un simple número. Aun si supiera que dicho número representa la edad de una persona, este seguiría siendo un dato, sin ningún uso o propósito en general, a menos que lo pusiéramos junto con el nombre de la persona (otro dato). Aún en este punto, esta unión no tendría mucho valor como información, porque requeriría de otros datos para poder obtener la información deseada. También debe ser algo que nos ayude a aprender, algo previamente desconocido u olvidado, de otra forma, sería solo una repetición inútil. La información se deriva de los datos y los datos por si mismos son solo la materia prima para la creación de información (Verstraete,1998)

Otra definición de Información es la que ofrecen McGee et al (1993). Ellos definen a la información como una colección de datos organizados, ordenados, con un significado y contexto. La información debe informar, mientras que los datos no tienen ese objetivo. La información debe tener límites, mientras que los datos pueden ser ilimitados. Los datos se pueden considerar en una forma aislada, mientras que la información debe estar enfocada a los usuarios específicos. La información es datos en uso y el uso implica a un usuario.

La palabra dato se refiere a hechos crudos concernientes a gente, objetos, eventos u otras entidades, como puede ser una colección de respuestas en una encuesta de mercado. Una organización colecciona y guarda una vasta cantidad de datos de diferentes fuentes tanto internas como externas, una cantidad que

posiblemente puede ser inmensa si se dejara en este estado. La información se deriva de los datos que han sido procesados y presentados en forma adecuada para la interpretación humana, y a menudo con el propósito de revelar tendencias o patrones. Por ejemplo los datos de la encuesta anteriormente mencionada se pueden analizar y resumir a través del uso de pruebas estadísticas tales como promedios, desviaciones estándar, tendencias, modas, etc.

Calidad de la información

La información si se considera como un recurso importante para la organización, tiene un valor y a medida que la calidad de ésta aumenta, también aumenta su valor. Pero este valor siempre es relativo al tomador de decisiones y al contexto en el cual se toma la decisión.

Para que la información sea de calidad, ésta debe presentar ciertas características deseables, como por ejemplo: que sea exacta, oportuna, actual, accesible, enfocada, dentro del rango, integrada, y segura. Dado el gran número de características necesarias para lograr la calidad de la información, en varias ocasiones es difícil obtenerla, pero el nivel de calidad deseado dependerá en gran medida del tomador de decisiones o usuario y la importancia de las decisiones que tome.

Exacta

La exactitud es una característica muy importante que debe tener la información, pero ésta a su vez depende de otras características como la actualidad. Por ejemplo, si se necesita información referente al valor del dólar americano en pesos mexicanos, la característica de exactitud puede estar relacionada con el número de dígitos requeridos por el usuario, sin embargo, conociendo que ésta información varía en el transcurso del día, es necesario acompañar dicha información con los datos de fecha y hora.

Oportuna

La oportunidad de la información consiste en obtener la información dentro de los límites de tiempo establecidos. De nada nos sirve la información un día, una hora o un minuto después.

Actual

La información debe ser actual, información al día. Para muchas empresas, esta característica es muy importante, ya que puede significar grandes sumas de dinero, como en el caso de los bancos, empresas petroleras, etc. las cuales compran o venden de acuerdo al valor actual de las divisas o precios de los artículos. Esta característica se relaciona mucho con la exactitud, ya que si la información no esta actualizada, entonces la exactitud no tendrá significado.

Accesible

La accesibilidad se define como la efectividad con que puede responderse a solicitudes de consultas y reportes sobre datos diversos. La accesibilidad se puede lograr ya sea por medios impresos o computacionales. La accesibilidad es la facilidad de acceso a la información deseada, que puede presentarse en manuales, catalogos, información variante, reportes, consultas, etc.

Enfocada

Enfocada al usuario final. La información puede ser correcta, exacta, oportuna, integrada, accesible pero, si el formato en el cual se le presenta al usuario final no cumple con sus necesidades o no es entendible, de nada sirve el trabajo invertido en obtenerla. Es necesario que el usuario la entienda. No es lo mismo información dirigida a un alto ejecutivo que para un obrero, o para dos altos ejecutivos con especialidades diferentes.

Dentro del Rango

La información debe provenir de las fuentes adecuadas, de las cuales se pueda obtener información detallada. De nada sirve obtener información de exportación de productos pesqueros de América del Norte, cuando necesitamos información específica de México.

Integrada

Icaza (1993) define que la integración total de la información se logra, cuando se reduce a cero la redundancia lógica y se cuenta con mínima redundancia física, totalmente documentada y controlada. Existe redundancia lógica cuando un mismo dato es nombrado o definido de diversas maneras en varias partes o por varias personas de la empresa. Hay redundancia física cuando existen varias copias de un mismo dato físicamente almacenadas en diversos sitios, o bien cuando se almacenan datos derivados que pueden calcularse a partir de otros datos. "Además, se habla de integración real, cuando los datos se encuentran completamente integrados lógicamente y físicamente, y de integración aparente cuando se crean vistas de la información para crear la impresión de que los datos se encuentran integrados." Dada la importancia que esta característica de la información representa para este trabajo, en el capítulo siguiente se hará una descripción detallada de ella.

Segura

La seguridad es un aspecto muy importante que no se debe dejar a un lado, y más cuando se han invertido grandes cantidades de recursos por lograr un producto de información de alta calidad. Para ello es preciso implementar sistemas de seguridad tales como respaldos, seguridad contra intrusos, etc., los cuales permitan que ese valioso recurso siga siendo un activo dentro de la empresa. Un aspecto muy importante relacionado con el tema de seguridad lo representa el hecho de que si la empresa ha logrado todos los aspectos de

calidad de la información, ésta se vuelve más vulnerable al robo. La información ya procesada, accesible, etc., es un blanco muy fácil de poder sustraer, ya que su misma calidad y accesibilidad ofrecen ventajas mas amplias para su adquisición.

Como se puede ver estas son algunas las características que la información debe tener con el fin de poderse considerar una herramienta esencial para la toma de decisiones, pero también hay que tomar en cuenta que esta información puede ser utilizada por personal en el ambiente operativo y convertirla en una ventaja competitiva para la empresa.

La Información y la Tecnología de Información

McGee et al (1993) señalan que uno de los principales errores que se han cometido en esta era, es creer que una competencia basada en información es un sinónimo de inversiones en tecnología de información. Varias organizaciones han invertido grandes cantidades en tecnologías de información para resolver sus problemas, pero no todas ellas han tenido el éxito esperado.

Las inversiones en tecnología de información, no crean ventajas o mayor productividad por si solas, como lo puede hacer una maquinaria o herramienta moderna. No es solo la tecnología, sino la tecnología en uso la que crea el valor. El valor de la tecnología de información depende de la información y el papel que ésta juegue en las organizaciones (McGee et al, 1993).

La información puede crear un valor significativo para las organizaciones, creando nuevos productos y servicios, mejorando así, la calidad de la toma de decisiones a través de la organización.

Las computadoras no son necesariamente un componente de un sistema de información, pero ellas son regularmente una parte de cualquier sistema moderno y ayudan a mejorar la eficacia y efectividad de la información que se utiliza.

Hasta hace poco tiempo, todos los sistemas de información eran manuales. Hace miles de años, los fenicios desarrollaron un sistema de registro muy

complejo que les ofrecía la información necesaria para administrar procesos muy complejos de comercio. Por dos mil años, las técnicas de procesamiento de datos no tuvieron una evolución considerable en sus métodos básicos, y se puede decir que los negocios se manejaron exitosamente sin la ayuda de las computadoras. Pero es seguro que nuestros antecesores, hubieran agradecido la ayuda de las computadoras, si hubieran tenido acceso a ellas, así como también, es preciso señalar que el volumen de información que se manejaba era mucho menor comparado con lo que se maneja en las grandes empresas de la actualidad.

La computarización puede mejorar grandemente el proceso interno de la organización. Existen dos formas básicas en las que cualquier proceso organizacional puede ser mejorado.

La primera, es mejorar su eficiencia. Las computadoras calculan rápidamente y con exactitud, ahorrando una gran cantidad de horas/ hombre. También pueden guardar y recuperar instantáneamente datos y procesos los cuales pueden convertirse en salidas de información, tales como reportes y resúmenes. La “oficina sin papel” puede reducir en forma significativa el almacenamiento de papel y los costos y tiempos asociados con su manejo. Las computadoras pueden ayudar a los negocios a tomar los datos o la información derivada de ellos y comunicarla rápidamente a los tomadores de decisiones y otros empleados claves a través de la organización, no importa cuán esparcida este en el mundo. La eficiencia es por lo tanto obtenida en términos del procesamiento.

La segunda forma en la cual las computadoras pueden ser usadas, es para mejorar la efectividad de los procesos de la organización, en tal forma que impacten las ganancias del negocio. Los sistemas de información, pueden utilizarse estratégicamente para cambiar la forma en la cual compiten las compañías. Compañías tales como Frito-Lay y Federal Express entre otras, han introducido tecnologías basadas en computadoras, que les han dado la oportunidad para hacer una reingeniería de sus procesos básicos. Esto a su vez,

les ha permitido ofrecer una mayor variedad de servicios a sus clientes y apoyo para enfocarse mejor al nicho de mercado deseado.

Los empleados de los negocios, actualmente son llamados a menudo los trabajadores del conocimiento, porque mucho de su trabajo involucra operaciones que dependen de sistemas de información computarizados. Ellos trabajan más eficientemente y toman mejores decisiones a todos los niveles de la organización con el apoyo de diversos sistemas de información. Hoy, todos los negocios, no importa lo pequeños sean, pueden disfrutar de dichas ventajas si sus empleados tienen el conocimiento para hacerlo.

Como se dijo anteriormente, la tecnología de información, es un elemento básico en los sistemas actuales y se pueden lograr una gran cantidad de ventajas al utilizarla. La tecnología, brinda la posibilidad de solucionar algunos aspectos de la calidad de la información, pero existen otras características de la calidad de la información, como es el caso de la integración de los datos, que requieren el apoyo de una metodología, para asegurar su obtención.

En el siguiente capítulo se hablará sobre algunos procesos tradicionales, que sirven de ayuda para lograr el atributo de integración de los datos.

CAPITULO II: ADMINISTRACION DE LOS DATOS

En el capítulo anterior se habló de la información y las características que ésta debe reunir para ser considerada como un apoyo para la toma de decisiones, pero ¿Cómo se puede lograr información de calidad?. Es preciso trabajar desde los orígenes de ésta, es decir, desde las fuentes de datos que consideramos la materia prima que alimentará a nuestros sistemas de procesamiento de datos, con el fin de lograr el resultado esperado "información".

Una vez que los ejecutivos de la empresa han reconocido el valor que representa la información para afrontar los retos y presiones de la competencia y la globalización de los mercados, es necesario enfocar los esfuerzos para crear la infraestructura adecuada para la administración de la información. En este capítulo se verán algunas formas para lograr información estratégica para la empresa, partiendo desde los orígenes o fuentes de datos, pasando por los diferentes procesos que nos pueden conducir a lograr la integración de los datos y finalizando en el logro de una herramienta valiosa para la toma de decisiones.

Lograr información de calidad requiere tomar en cuenta varios aspectos. Una característica de calidad que se considera muy importante para este trabajo, es la integración de los datos, que consiste en lograr un lenguaje común entre las diferentes áreas de la empresa, con el fin de lograr un flujo de información efectivo. Para lograr la integración de los datos existen varias alternativas, las cuales van desde una redefinición total de todos los datos empresariales hasta soluciones de integración aparentes.

El objetivo de crear información de calidad es ofrecer bases sólidas que permitan tomar decisiones correctas con márgenes de error mínimo, lo cual se ha convertido en una exigencia en la actualidad. Las empresas y los empresarios no se pueden dar el lujo de "adivinar" o basar sus decisiones en suposiciones o presentimientos, es necesario crear la infraestructura necesaria para obtener bases sólidas de información.

Actualmente muchas empresas invierten grandes cantidades de dinero en equipos y software para administrar su información, pero esto no siempre es suficiente, ya que como se mencionó anteriormente, la tecnología de información crea muy pocos beneficios por si sola.

A la información se le puede considerar como un recurso, por lo tanto, debe de administrarse como tal, es decir, se le deben dedicar al menos los mismos cuidados que a los recursos tradicionales. Es necesario que tenga un administrador, que se encargue de su buen uso, mantenimiento, generación, seguridad, etc. A esta persona se le conoce como el Administrador de los Datos.

La administración de los datos, tiene un amplio rango de funciones, entre las cuales se encuentran: la planeación de las bases de datos, su diseño, implementación, mantenimiento y cuidado, el desempeño de la misma y ofrecer el entrenamiento y soporte a los usuarios.

Es muy importante, que el administrador de los datos tenga en mente la estrategia de la empresa, con el fin de apoyarla, proporcionando información de calidad a todos los usuarios de este recurso.

Por lo anterior, la información que se genera con los sistemas debe estar basada y fundamentada en fuentes confiables, que permitan ofrecer un producto de gran calidad. El administrador de los datos, debe ser una persona capaz de buscar las fuentes de información críticas y de calidad, las cuales se convertirán en la materia prima, que a su vez, se transformará en información de calidad.

Fuentes de Información

El análisis de las fuentes y orígenes de los datos, son aspectos muy importantes en el desarrollo de los sistemas, debido a que representan la materia prima que entrará a los mismos. Hay que recordar, que la calidad de estos depende directamente de las fuentes de información y el impacto hacia nuestra empresa.

La globalización de los mercados, ha cambiado la forma en la cual las empresas deben fijar sus estrategias para sobrevivir en un ambiente competitivo.

Estas estrategias ya no se pueden limitar a investigar a la empresa y su competencia dentro de la misma industria, sino que deben tener mayores alcances, con el fin de que los sistemas que se generen, basados en tales estrategias, arrojen información con bases más sólidas.

La información acerca del ambiente competitivo de la empresa, ayuda a los ejecutivos a identificar las amenazas y oportunidades para la compañía, y establece el área para el diseño de una respuesta competitiva más efectiva. La información, también funciona como un recurso crítico para el diseño de estrategias alternas.

Las 5 Fuerzas Competitivas de Porter

El profesor Michael Porter de la escuela de negocios de Harvard, desarrolló un sistema, el cual es el más aceptado para el diseño de estrategias competitivas para las empresas. El sugiere que una estrategia efectiva, debe tomar en cuenta no solo las acciones y reacciones de los rivales directos, sino también, los papeles de los clientes, proveedores, y productos alternos que satisfagan la misma necesidad básica, así como también, la amenaza de nuevos entrantes como se puede apreciar en la figura 2.1.

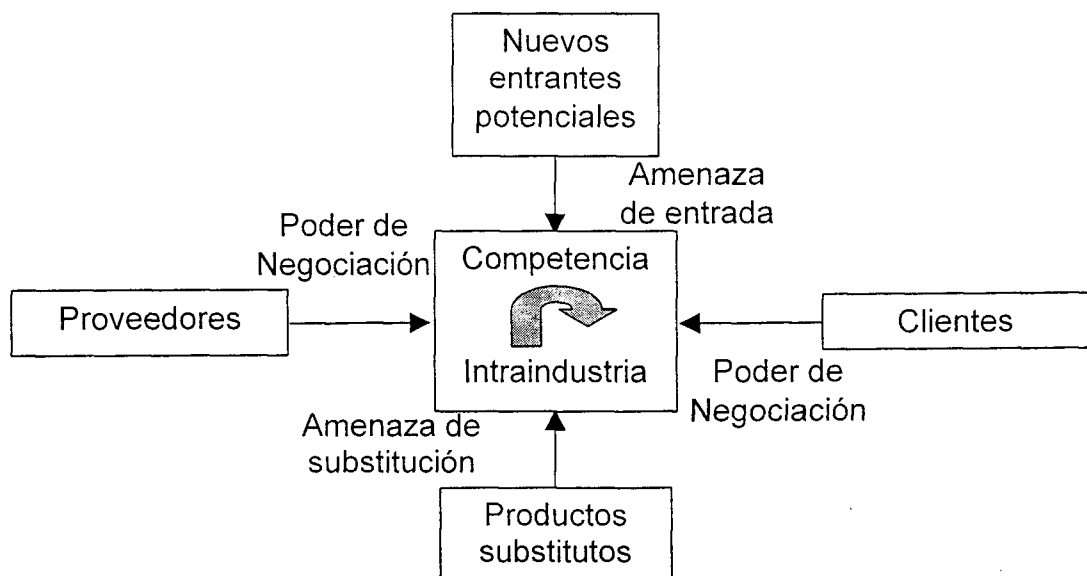


Figura 2.1 Modelo de Porter de las Fuerzas Competitivas

El valor del modelo de Porter, desde una perspectiva de información, es que éste, estimula a los ejecutivos a considerar un rango más amplio de información estratégica, que lo que anteriormente se utilizaba en la mayoría de las empresas.

El modelo anteriormente descrito, se complica aun más debido a que éste a su vez es una parte de un esquema mayor, en el cual la empresa se convierte en un proveedor de otra empresa o su cliente, y tanto sus proveedores como clientes son a su vez eslabones de otras cadenas. Así pues, tenemos que una fábrica de aluminio, vende hojas de aluminio a una empresa fabricante de envases, que a su vez, vende sus productos a una empresa de bebidas, la cual tiene su modelo individual.

Es importante mencionar, que a menudo, la información estratégica más crítica para nuestra empresa, puede encontrarse en uno o dos niveles fuera del modelo de la industria actual. Por ejemplo, el productor de aluminio, puede obtener información más valiosa para su empresa, al observar a la industria de los refrescos, más que a su propia industria. El, puede analizar los posibles efectos derivados de la introducción de envases de plástico en el mercado y de su aceptación por parte del consumidor. Esta información que se encuentra dos niveles fuera de la cadena de valor de la empresa, puede enviar señales valiosas, que sirvan para modificar la estrategia actual y no permitir que nuestra industria, se vea seriamente afectada de forma repentina.

La cadena de valor de Porter

La cadena de valor de Porter, es otro recurso valioso, el cual, también puede utilizarse como medio para obtener las necesidades de información en la empresa.

En su análisis, Porter señala las partes dentro de la empresa en las cuales los productos o servicios generados por la compañía van ganando valor (ver figura 2.2). Esta forma de analizar a la empresa, puede servir para detectar las necesidades de información en cada área y su relación con las otras y, que de esta forma, el desarrollo de los sistemas tome en cuenta las necesidades y

deficiencias individuales, con el fin de mejorar el flujo y disponibilidad de información clave, para hacer más eficiente el proceso. Hay que tener en cuenta, que la cadena de Porter analiza muy poco el ambiente externo de la empresa, a diferencia de las 5 Fuerzas Competitivas que se vieron anteriormente.

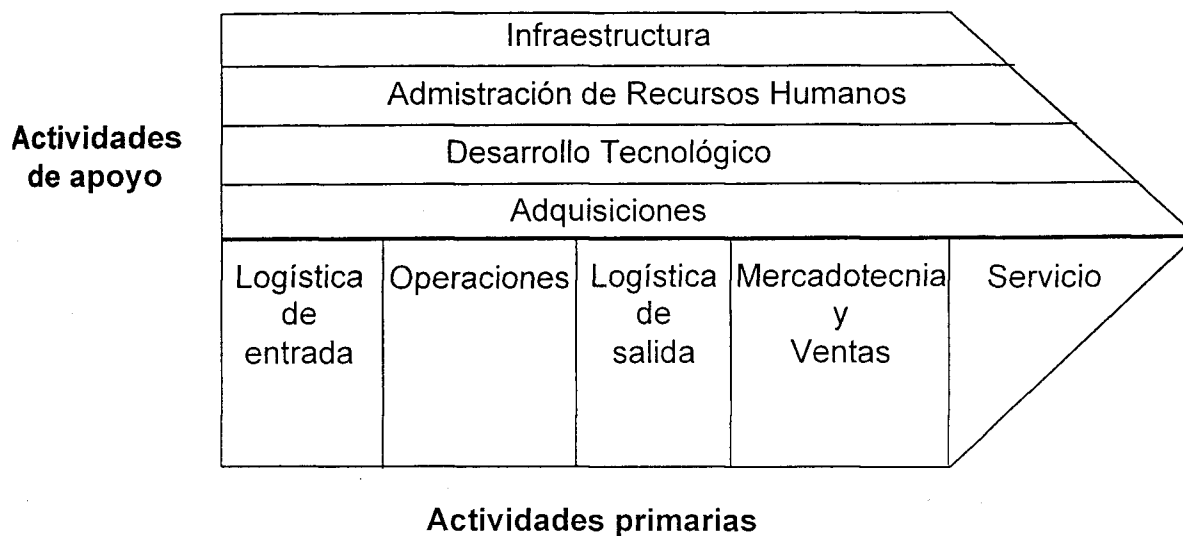


Figura 2.2 Cadena de Valor de Porter.

Factores Críticos de Éxito

Los factores críticos de éxito, son otra herramienta que se puede utilizar, con el fin de obtener las bases de información o la materia prima para alimentar nuestros sistemas con información valiosa y de alto impacto para nuestra empresa. Los factores críticos del éxito, son una serie de elementos en los cuales se basa el éxito de la empresa, estos factores, pueden incluir entre otros, el servicio personalizado al cliente, puntualidad, mejor tiempo de respuesta que la competencia, etc. Entonces, la estrategia de los sistemas de la empresa, se pueden basar en diseños que ayuden a mantener estos factores de éxito. También, existen los factores críticos de fracaso, y éstos actúan de forma inversa, es vital para la empresa tenerlos bien vigilados, con el fin de evitar que se desarrollen y se conviertan en un problema real para la empresa. Los factores

críticos del éxito, son los que hacen que la empresa sobresalga de la competencia, los factores críticos de fracaso, son los que hacen que la empresa se mantenga en el mercado.

Es posible que ninguna de las tres opciones anteriormente expuestas sirvan para las necesidades propias de nuestra empresa, pero se pueden utilizar, como base para buscar fuentes de información necesarias para el desarrollo de las estrategias de la empresa. También, se puede pensar en utilizar una mezcla de ellas, con el fin de detectar gran parte de las posibles fuentes de información. Posiblemente, lo anterior resulte bastante complejo, y arroje una gran cantidad de resultados, pero esto, permitirá seleccionar aquellas fuentes con más impacto para nuestra empresa.

Integración de datos.

La integración, es una palabra que por sí sola no representa nada o puede causar confusión, a las personas ajenas a la administración de la información, pero se puede considerar, como una de las características más importantes de la calidad de la misma. Esta característica de calidad de la información, se considerara una de las más difíciles de lograr o detectar su falta, ya que se le puede considerar el enemigo oculto. Esto se debe principalmente a que la desintegración de la información, da como resultado información errónea, la cual aparentemente esta correcta.

Antes de seguir adelante con la integración de los datos, se presentarán algunas definiciones, con el fin de conocer más a detalle su significado.

- La integración de datos generalmente significa la estandarización de las definiciones y estructuras de los datos por medio del uso de esquemas conceptuales comunes de las diferentes fuentes de datos (Heimbindingner, 1985, Litwin, 1990).
- Un requisito indispensable para lograr una buena comunicación entre las diferentes unidades de una empresa y más aún cuando éstas se encuentran

separadas es contar con un lenguaje común de comunicación. En el mundo de los sistemas de información computacionales una forma que puede tomar este lenguaje común son los datos integrados o datos con definiciones estándar o comunes (Goodhue, 1992).

La integración de los datos asegura que los datos tienen el mismo significado y uso a través del tiempo y los usuarios, haciendo que los datos sean consistentes o lógicamente compatibles en los diferentes sistemas o bases de datos (Martin, 1986).

Estos tres autores se refieren a una integración real de los datos. Para lograrla en una empresa, pudiera ser necesario llegar a definiciones comunes y rehacer físicamente datos y sistemas, lo cual por lo general resulta en un proceso bastante largo y gravoso. Una alternativa más económica y viable para muchas empresas esta representada por la integración aparente de los datos, la cual presenta una vista de datos integrada para el usuario, partiendo de datos que en la realidad se encuentran desintegrados. Más adelante se presentará un poco mas a detalla dicha técnica.

La desintegración de la información, se presenta en una infinidad de empresas. Aunque existe un gran número de causas que pueden dar origen a este problema, éste, se deriva principalmente de las malas prácticas de diseño y desarrollo de sistemas, en los cuales, la atención se centra solo en el área beneficiada por dicho sistema, sin tomar en cuenta a la empresa como un sistema total. En el siguiente cuadro, se muestra un ejemplo de una empresa con problemas de desintegración.

Cuadro 2.1 Ejemplo de desintegración de la información.

| Descripción | Con Datos Integrados | | Sin Datos Integrados | |
|----------------|----------------------|------------------|----------------------|------------------|
| | Banca Hipotecaria | Tarjetas Crédito | Banca Hipotecaria | Tarjetas Crédito |
| Nombre | Ana | Ana | Ana | Pérez C., Ana |
| Cve del area | 1245 | 5943 | 1245 | 5943 |
| Cve Id Interna | 58488 | 58488 | 58488 | 23588 |
| Ingr. Mensual | 8,000 | 8,000 | 8,000 | 5,000 |

Como se puede observar, la desintegración de la información en esta institución no permite a sus funcionarios realizar trabajos coordinados entre sus diferentes áreas. Si el presidente del banco con datos desintegrados desea saber cuantos clientes tiene en total, sumaría sus clientes de tarjetas de crédito a los de banca hipotecaria, lo cual generaría un error, por que es muy probable que existan clientes en ambas partes. Otro problema que trae consigo la desintegración, es la reducción del número de maniobras que se pueden hacer entre unidades. También pueden existir problemas graves cuando una empresa se fusiona con otra y al fusionarse, sus sistemas, darán como resultado, un volumen de datos enorme. Si a esto se le añade el aspecto de la desintegración en alguna de ellas o en ambas, el problema se tornará aún más difícil.

A todo lo anterior se le puede llamar "parálisis causada por la desintegración de la información", ya que el funcionario se verá imposibilitado para crear o mejorar servicios o productos hacia sus clientes, los cuales pueden ser la base de la diferenciación que permita distinguir a dicha empresa entre sus competidores.

Entre las múltiples facetas que los datos desintegrados pueden tomar estan los que se muestran en la figura 2.3. En los diferentes programas de la empresa, la codificación para el dato de sexo puede tomar varias formas, entre las que se encuentran: masculino/femenino, m/f, 1/0, etc.

Otro ejemplo de desintegración, se puede ver en el siguiente caso. El campo "tuberías", el cual se maneja con el mismo nombre a través de toda la empresa, puede contener información en diferentes unidades de medidas como: centímetros, pulgadas, metros, pies, etc.

La transformación de campos, es otro aspecto relacionado con la integración. La misma información, existe en las diferentes aplicaciones, pero con nombres diferentes.

También, otro aspecto de integración, consiste en tener definiciones diferentes de los campos, como puede ser: diferentes tipos de campos, uno puede ser tipo caracter, otro numérico, etc. Igualmente, puede haber diferencia entre dos

campos del mismo tipo, pero con diferente definición, como ejemplo podemos poner un campo caracter con 10 espacios y otro con 12, 16, 30, etc.

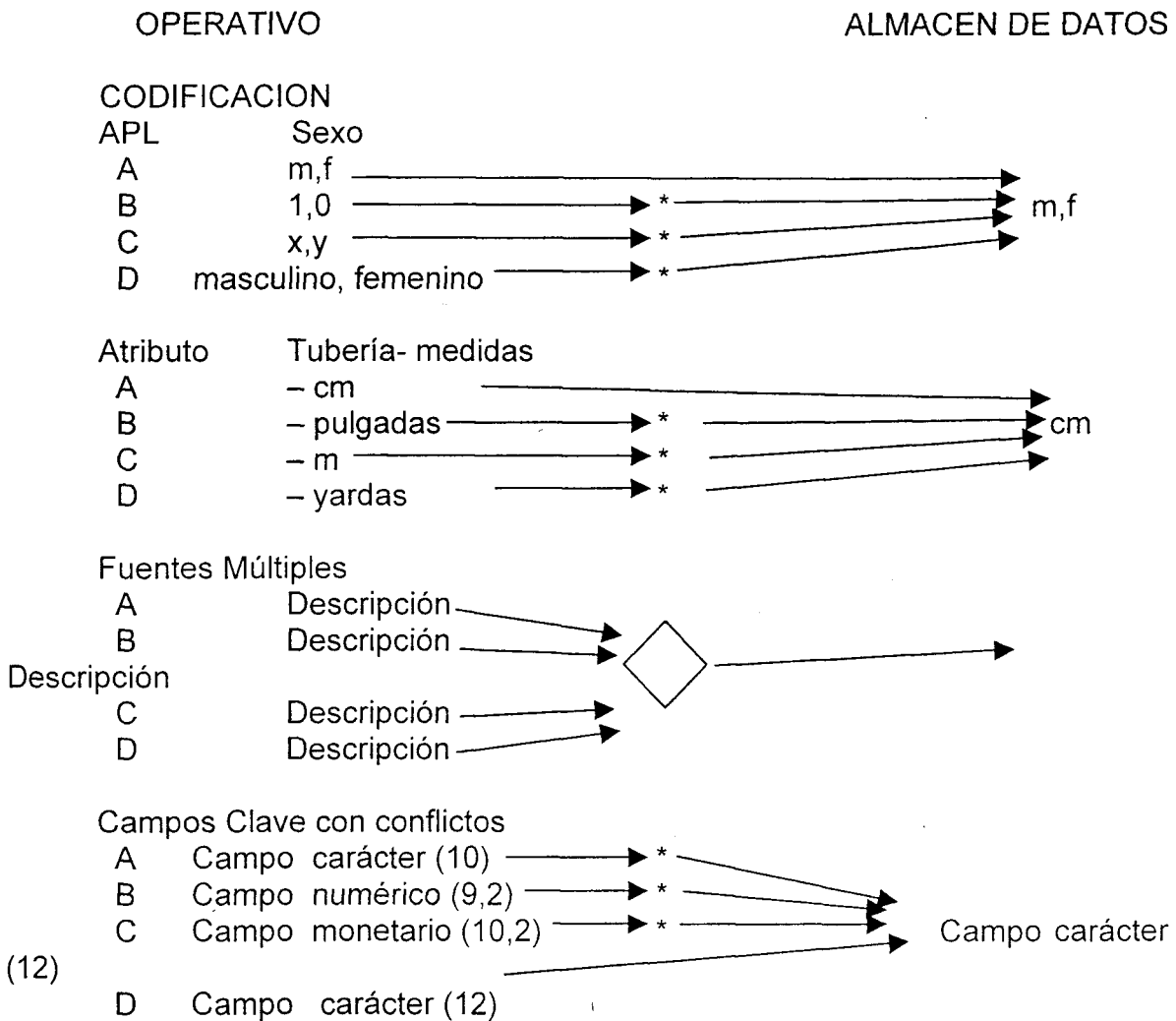


Figura 2.3 Diferentes manifestaciones de desintegración de datos (Inmon, 1996).

Beneficios de la integración de los datos

¿Qué beneficio nos representa contar con datos integrados en nuestra empresa?. En primer lugar y quizá el más importante, es que crea la base necesaria, para el diseño e implementación de sistemas que generarán

información confiable. Otro beneficio consiste, en la habilidad para compartir o resumir información de las diferentes áreas de la organización. La necesidad para compartir información es mayor, cuando las acciones o eventos a través de dichas unidades son interdependientes, esto es, donde las acciones de una subunidad, afectan a las acciones o productos de otra subunidad

Como consecuencia de los beneficios de la integración de los datos, es posible plantear una nueva gama de productos y servicios, los cuales, se pueden convertir en el punto diferenciador, que haga más atractiva nuestra empresa a los clientes o clientes potenciales.

Asimismo, ofrece la oportunidad de brindar a los altos ejecutivos, más herramientas para la toma de decisiones al presentar una gama distinta y más amplia, de apoyos informáticos.

También, permite mayores oportunidades de acceder información generada en cualquier área de la empresa y así apoyar para que la misma presente una cara común a sus clientes.

Inmon (1983) señala que una ventaja de la integración de los datos, es la habilidad que adquiere la empresa para hacerle frente a los cambios del mundo que la rodea y de sus propios usuarios. A medida que la integración sea menor, menor será la capacidad de la empresa para reaccionar a dichos cambios.

Se pueden encontrar al menos dos impactos diferentes, que la integración de los datos pudiera traer a la organización: (1) Una administración de la información mejorada, para una mejor comunicación a través de la empresa. (2) Una coordinación operacional entre unidades interdependientes de la organización (Goodhue et al, 1988)

En estos momentos de crisis financieras, en las que algunas empresas recurren a la fusión con otras empresas o bien, grandes instituciones adquieren o absorben empresas pequeñas, la integración de la información puede facilitar en gran medida, los procesos de unificación de los sistemas de dichas empresas.

Por otra parte, la globalización de los mercados ha traído grandes ventajas al consumidor, ofreciendo una gama de productos y servicios más amplia. Pero

también, esta nueva corriente ha afectado a varias empresas del país, obligándolas a ser más eficientes y mejorar la calidad de sus productos o servicios. Hoy en día, es muy importante para una empresa atacar ciertos nichos de mercado, los cuales, pueden estar representados por edades de las personas, nivel de ingresos, etc. Este tipo de enfoque, requiere que los empresarios cuenten con herramientas de información, que les permitan tener mejor control y conocimiento de sus usuarios o clientes, así como también, realizar estrategias dirigidas a atraer ese tipo de clientes.

En estos momentos, es indispensable que una empresa tenga la capacidad de ofrecer productos o servicios con cierto grado de diferenciación, que atraigan al cliente para que seleccione nuestra empresa, o que ayuden a conservar los clientes actuales. Dichos diferenciadores, pueden estar representados por: descuentos atractivos, regalos, etc. a cierto tipo de clientes.

Para lograr lo anterior, se requiere de un esfuerzo previo para mejorar la calidad de la información, el cual permita, a los ejecutivos de la empresa, realizar sus iniciativas sin temor a cometer errores que perjudiquen a la empresa.

Un ejemplo actual, se puede tomar de una compañía aseguradora, la cual le ofrece descuentos extras a sus clientes por asegurar más de un automóvil, otro descuento más si no tuvo reclamaciones el año anterior, etc. Estos descuentos que ofrece esta compañía, los cuales pueden ser insignificantes, muchas veces se convierten en el factor de decisión, que induce al cliente a resolver por una u otra compañía.

Pero esta compañía de seguros, puede ir un poco más adelante, ya que sólo esta tomando en cuenta el área de seguros de automóviles. Imagine la ventaja competitiva que pudiera tener esta compañía sobre sus rivales, si pudiera hacerle un ofrecimiento al cliente de otro descuento, si accede a tomar el seguro de vida de la misma empresa, y así descuentos posteriores si utiliza algún otro de los servicios de la misma. Para la empresa, esto representa una gran cantidad de clientes potenciales, pero para lograr este tipo de esfuerzos sin temor a perjudicar a la empresa, es necesario que exista una comunicación confiable, entre las

diferentes unidades de la empresa, y lo anterior, solo se puede lograr efectivamente, con el apoyo de la integración de los datos empresariales.

Después de haber analizado el ejemplo anterior, y ver las ventajas que representa la integración de los datos, entonces ¿por qué las empresas no lo hacen? o ¿cuáles son los principales problemas para lograr la integración?

Se han hecho algunos estudios, en los cuales varias empresas han fracasado en sus intentos por lograr la integración de sus datos, y las principales causas detectadas, han sido las siguientes: errores en su implementación, como falta de apoyo de la alta administración (Goodhue, 1988; Hoefler, 1989), defectos en la metodología utilizada (Goodhue¹, 1992), Posibilidad de que los beneficios obtenidos no sean lo suficientemente buenos comparados con la inversión (Goodhue², 1992). También otro aspecto muy importante y más relacionado con las empresas mexicanas es, el tiempo requerido para lograr la integración, ya que las crisis por las que han atravesado las empresas mexicanas, han provocado que los directivos sean muy cautelosos, en lo que respecta a inversiones a largo plazo, debido a que el período de vida muchas empresas, es inferior al tiempo requerido para lograr la integración de sus datos.

Integración de datos. ¿Cómo se puede lograr?

Icaza (1993), señala que una empresa puede tomar diversas acciones para promover la integración de sus datos. Estas acciones pueden clasificarse en dos grandes grupos: acciones de arriba-abajo y acciones de abajo-arriba.

Las primeras están basadas en un enfoque estratégico, que promueve la creación de modelos de datos que soporten la estrategia empresarial –“El Deber Ser de la Empresa”. Usualmente, requieren de un gran esfuerzo cuyos resultados se aprecian a largo plazo. Las segundas suelen ser más modestas en su alcance, y buscan aprovechar la inversión de la empresa en datos y sistemas ya existentes “El Es de la Empresa”.

Acciones arriba-abajo

Entre las acciones de "arriba-abajo" se encuentra la planeación estratégica de datos, y la Ingeniería de Información (Martin et al, 1989; Martin, 1990), las cuales especifican el desarrollo de modelos de datos empresariales, bajo los cuales debe quedar integrada toda la información y que sirven como arquitectura para el desarrollo de sistemas.

Goodhue et al (1992), definen a la Planeación Estratégica de Datos (PED) como un enfoque formalizado de planeación centrado en los datos que construye un modelo de la empresa, sus funciones, procesos, y sus datos como una base para identificar o implantar un conjunto integrado de sistemas de información que contribuirán a satisfacer las necesidades estratégicas del negocio.

Como se ha expuesto anteriormente, la falta de integración lógica de los datos a través de los sistemas de información hace muy difícil o imposible responder a búsquedas interdepartamentales o interfuncionales. La PED es una metodología que ataca estos problemas. Apoyándose en que asume que un grupo relativamente estable de entidades de datos se encuentra en el centro de las necesidades de procesamiento de información de la organización. La PED es un planteamiento de planeación formalizado tipo arriba-abajo, centrado en los datos, que construye un modelo de la empresa, sus funciones y sus datos importantes como una base para identificar e implementar un grupo integrado de sistemas de información.

La principal función de la PED consiste en producir una arquitectura que guíe el desarrollo de los sistemas. Esto involucra la creación de nuevas aplicaciones y la revisión y modificación de algunos sistemas actuales con el fin de que los sistemas empresariales emigren gradualmente hacia un conjunto de aplicaciones y bases de datos integradas.

Uno de los factores críticos para el éxito de la PED, lo constituye la integración de datos, la cual debe ser la meta estratégica de la organización. Es importante lograr un balance adecuado entre la integración de datos a nivel empresa y la flexibilidad de las unidades. Una crítica que señala Icaza (1993) a

esta metodología, es que no toma muy en cuenta la enorme inversión que la empresa ya tiene en archivos y bases de datos existentes.

El proceso de la PED es largo y costoso y por lo general, una vez que se ha logrado la jerarquía total de modelos empresariales, existe una gran diferencia entre los modelos y los sistemas actuales.

Otro aspecto muy importante, que se debe tener en cuenta antes de tomar la decisión de iniciar un proceso de PED, consiste en analizar aparte del tiempo de desarrollo y el costo de los mismos (en un ambiente dinámico esto puede ser considerado un elemento bastante crítico), la capacitación de los usuarios, el costo del software involucrado, las modificaciones al hardware actual, alineación de los sistemas con las metas de la empresa, la inercia al cambio, etc.

Se ha encontrado en repetidas ocasiones, que el trabajo y tiempo requeridos para la conversión de los datos y sistemas existentes a los nuevos modelos y bases de datos, son por lo general muy grandes y costosos, por lo cual muchas veces podría terminar por desecharse.

La PED es un método para ayudar a mejorar la integración de los datos de la empresa, aunque se han encontrado más problemas que éxitos de iniciativas de PED (Goodhue²). Es indispensable entonces para el Administrador de Sistemas de Información adquirir experiencia basada en problemas o soluciones que hayan logrado otras empresas en el pasado y no cometer el mismo error; analizar los casos y seleccionar las ideas o metodologías a las cuales se les pueda sacar provecho. En la literatura se puede observar que la administración de los datos no es algo fácil de lograr y no existe un método que pueda servir para todas las organizaciones, ya que cada una tiene sus características particulares. Es importante conformar un resumen con las sugerencias, así como las advertencias de intentos anteriores que puedan servir como base, para lograr una buena administración de la información.

Acciones de abajo-arriba

Debido a la inversión en tiempo y capital requeridos para una Planeación Estratégica de Datos completa, los riesgos asociados y la dinámica de los mercados actuales, son muchas las empresas que han decidido buscar alternativas, que les permitan lograr la Integración de sus datos. Estas alternativas, no se basan en las metas y objetivos de la empresa, sino en los datos ya existentes con los que se cuentan, y parten de ellos con el fin de lograr la integración y otros aspectos relacionados con la mejora de la calidad de la información actual.

Icaza (1993), en su artículo de Implantación gradual de la administración de datos, desarrolla una metodología que va encaminada a mejorar la calidad de los datos, sin tener que enfrascarse en un proceso largo y costoso de PED, el cual iría dando frutos con cada acción realizada. Entre las acciones propuestas por Icaza (1993), se encuentran:

- Documentación de datos. Levantamiento de un diccionario de datos.
- Resolución de homónimos y sinónimos.
- Estandarización de formatos de almacenamiento
- Normalización de archivos
- Estandarización de códigos.
- Ingeniería de reversa - de los datos a los modelos de datos.
- Conversión de sistemas para que utilicen un manejador de base de datos.
- Programas "puentes" (convertidores, traductores).
- Programas filtro (reductores) y bases de información. ("Information Warehouses") (Goodhue, 1988).
- Procesadores humanos de consultas (Goodhue, 1988).

El principal peligro de las acciones abajo-arriba, es que los datos actuales reflejan el "Ser" actual de la organización, no su "Deber Ser". Si se llevan a cabo exclusivamente éste tipo de acciones, se pierde la oportunidad de rehacer a la empresa, eliminar datos innecesarios, introducir nuevos datos requeridos, entre otros.

Integración aparente de los datos

Como una alternativa a la integración real de los datos, ya sea utilizando una acción “arriba – abajo” o “abajo – arriba”, se presenta en seguida la integración aparente por medio de un ejemplo.

Hay varias formas de lograr una integración aparente de los datos. Como ejemplo, si se tienen dos archivos de “máquinas” que utilizan códigos diferentes para identificar a las mismas máquinas, se podría crear un programa que utilice una tabla de mapeo de códigos para combinar registros de ambos archivos. El programa presentaría una pantalla que combinaría ambos archivos, “para el usuario de éste programa, los datos aparentarían venir de una fuente única o integrada”.

Otra alternativa esta cubierta por toda una metodología que se denomina “data warehouse” y sus tecnologías asociadas. En ella, se crea un caché o almacén intermedio de datos integrados, que proceden de datos operativos los cuales pueden o no estar integrados. Para todos lo usuarios del almacen de datos, los datos aparentan estar integrados, ofreciéndoles así, contar con una herramienta muy valiosa para la toma de decisiones sin tener que invertir en tiempo y capital en un proceso de integracion real.

Una tercera alternativa consiste en utilizar el mecanismo de “vistas”, las cuales son parte de las características de algunos manejadores de bases de datos heterogéneas distribuídas, para lograr la integración aparente de diversas bases de datos definidas en forma independiente y no totalmente compatibles.

Grado de integración de datos

Es preciso el determinar el nivel de integración deseado, ya que al tratar de integrar toda la información de una empresa, puede resultar en un trabajo costoso, el cual nunca llegue a su fin. Es por esto, que es necesario buscar un nivel que permita lograr la perfecta comunicación entre los sistemas y que a la vez, no consuma gran cantidad de recursos financieros y tiempo. Es necesario que este

tipo de decisiones, se tomen en forma centralizada, ya que de nada serviría que una unidad trabajase en la integración de su información en forma independiente.

La integración de los datos, es un ejemplo de un lenguaje altamente formalizado para describir los eventos dentro de los límites de la organización. El alcance de la integración de datos, es el punto al cual este lenguaje es usado a través de las diferentes áreas o subunidades dentro de la organización. Entonces, el grado de integración de los datos de un sistema de información computarizado, puede ser visto como una característica de diseño, importante en el manejo de la incertidumbre organizacional. Para entender cuando la integración de datos es especialmente necesaria, debemos analizar profundamente las fuentes de incertidumbre que enfrenta la empresa.

Incetidumbre y Ambigüedad

Existe una diferencia entre la incertidumbre y la ambigüedad. Esta diferencia puede ayudarnos a entender cuando los sistemas de información computarizados (con o sin integración de sus datos), pueden ser de poca ayuda en el cumplimiento de las necesidades de información de la empresa. Incertidumbre, es la ausencia de información específica solicitada. Por ejemplo, un gerente de ventas de un supermercado pudiera necesitar la información referente a la baja de un 50% en las ventas, de los productos lácteos en el último trimestre. La pregunta es clara, y al colocar los datos en las variables específicas, la incertidumbre acabará o disminuirá (Goodhue², 1992).

La ambigüedad, significa que existen interpretaciones múltiples de la misma situación, las cuales, por lo general son conflictivas, es decir falta información más precisa. Cuando la ambigüedad se presenta, la información necesaria para resolver la situación no es clara. Por ejemplo basándonos en el ejemplo anterior, el gerente podría recibir diferentes tipos de información que pudieran ser apropiados, como puede ser : el comportamiento del personal de ventas, acciones de la competencia, situación económica, época del año, etc. (Goodhue²,1992)

En general, la incertidumbre puede reducirse con la cantidad de información suficiente, mientras que la ambigüedad puede reducirse con información suficientemente rica.

Goodhue² (1992), en su artículo *The Impact of Data Administration in the Cost and Benefits of Information Systems*, concluye que la integración total de los datos, es una propuesta bastante cara y por lo tanto que no es redituable integrar todos los datos de la organización, el propone buscar la integración parcial, buscando grandes beneficios y evitando grandes costos.

Otra propuesta para buscar la integración parcial consiste en pedir a todas las subunidades, que utilicen el mismo tipo de aplicaciones a nivel empresa para el manejo de las funciones mas generales, como pudiera ser: nóminas, órdenes de compra, órdenes de servicio, control de inventarios, contabilidad, etc. Este tipo de aplicaciones, ayudarían a que la información que se maneje en estos sistemas, tenga la misma estructura, etc. Con esto, se lograría la integración de ciertos datos críticos a nivel empresa. Estas aplicaciones a nivel empresa, incluyen un modelo de datos estándar y por lo tanto una definición de datos común.

Otra acción, consiste en desarrollar bases de datos empresariales (clientes, productos, etc.) y hacer que todas las aplicaciones que usen dichos datos, tengan acceso de lectura, escritura y modificación de los datos contenidos en dichas bases de datos.

También, se puede identificar una selección de datos críticos y obligar la definición común de dichos elementos a nivel empresa. Estas definiciones estándar, pueden ser obligatorias en todos los desarrollos de nuevos sistemas.

Goodhue² (1992) nos muestra los resultados de su estudio del impacto de la integración de los datos en ambientes cambiantes (ver cuadro 2.2).

Es importante señalar que a medida que la interdependencia entre las unidades se incrementa, los beneficios derivados de la integración de los datos, también se incrementan, por lo tanto, la cantidad de integración de datos en una organización debería también incrementarse.

Cuadro 2.2 El impacto de la integración de los datos en ambientes cambiantes

| En presencia de | Resultará en |
|--------------------------------------|--|
| Interdependencia entre subunidades | Mejor comunicación Mejor coordinación |
| Diferenciación entre subunidades | Más compromiso o más costos de diseño, más retrasos burocráticos |
| Diferenciación y ambiente turbulento | Aún más compromiso o Aún más costos de diseño |

Cuando las subunidades son independientes, se tiene que buscar un equilibrio entre la estandarización organizacional y la autonomía departamental. Hay que tomar en cuenta que la integración de los datos, puede involucrar no solo la pérdida de la autonomía local, sino también una pérdida en la efectividad local. Por esto, es necesario hacer un análisis del grado de integración requerido y lograr que se cumplan los requisitos de tener un lenguaje común para cierta información global y flexibilidad local, para desarrollar los sistemas para enfrentar las necesidades particulares.

Otro aspecto importante que se debe tomar en cuenta, es el tiempo requerido para reaccionar a cambios en el ambiente local. Para lograr una integración empresarial, es necesario crear un área central encargada del manejo de los datos estándar dentro de la empresa, esto obliga a las unidades a depender de un área central. En ambientes tranquilos, puede que no represente un gran esfuerzo para las unidades el cumplir con las disposiciones centrales, pero en ambientes cambiantes, la burocracia puede tener un gran impacto a nivel local, es decir, si el ambiente local cambia, las necesidades de información también cambian. Entonces, es necesario solicitar modificaciones, las cuales deben ser estudiadas y aprobadas por el organismo central, el cual a su vez, tiene que ver si estos cambios no afectarán a otras unidades. Esta burocracia, puede hacer que la unidad enfrente problemas en el tiempo de respuesta, lo cual afectará directamente su desempeño local.

Diferencia entre el Administrador de Datos y el Administrador de Base de Datos

Smith² (1997), comenta que varias organizaciones no ven una diferencia esencial entre la administración de datos y la administración de base de datos. Como un resultado, existe confusión sobre los papeles de dichas funciones y sus responsabilidades. Cada disciplina es necesaria para la buena administración del recurso informático de la empresa, pero Smith² (1997), recomienda que dichas actividades nunca deben de combinarse en una persona o un subgrupo. Cada disciplina requiere diferentes habilidades, entrenamiento y talentos, por lo que la mayoría de las veces cuando se transfiere una persona de una disciplina a otra no lo hace exitosamente. La administración de datos y sus subdisciplinas (modelación de datos, definición de datos, planeación y análisis), son relativamente nuevas en el campo del procesamiento de datos. Es hasta hace 10-15 años, que la industria ha dado una consideración seria a la administración lógica y control de la información como un recurso corporativo. Existe una falta de entendimiento de los propósitos y objetivos de la administración de datos, aun entre los profesionales experimentados del procesamiento de datos.

Es fácil hasta cierto punto, que un administrador de base de datos o programador de sistemas pueda migrar de una industria a otra, con un mínimo de reentrenamiento, siempre y cuando la tecnología permanezca constante; mientras que un administrador de datos, tiene mucho que aprender en una nueva industria, la cual no le es familiar, con el fin de ser efectivo. Para tener un impacto en el diseño de datos y la administración de la información, se requiere un entendimiento completo de las metas, objetivos y tácticas de la organización y la industria en la cual se manejan (aseguradoras, farmacéutica, banco, etc.). Un papel de la administración de datos es abocar la planeación y coordinación de los recursos informáticos, a través de aplicaciones relacionadas y áreas del negocio. Al hacer lo anterior, se incrementa la cantidad de datos compartidos, y se disminuye la cantidad de datos redundantes.

Según Smith² (1997), "el principal objetivo de la administración de datos es integrar y administrar los recursos informáticos de toda la empresa".

El mayordomo de los datos (data steward)

Con más frecuencia, las empresas, están demandando mejores y mejores fuentes de datos. El crecimiento explosivo de los almacenes de datos y la sofisticación de las herramientas de acceso, han probado que los datos son uno de los bienes más críticos que las compañías poseen. Los datos transformados a información, deben ser entregados a los tomadores de decisiones de forma rápida, concisa y principalmente exacta.

Imhoff (1997), define en su artículo (data stewardship – finally a process for achieving data integrity), a un mayordomo de datos (data stewardship), como la persona que tiene como objetivo principal, la administración del recurso dato de la empresa para mejorar su rehuso, accesibilidad y calidad. La responsabilidad de un mayordomo de datos, es aprobar los estándares de nomenclatura del negocio, desarrollar definiciones de datos consistentes, determinar alias de datos, desarrollar cálculos estándares y derivados, documentar las reglas del negocio, monitorear la calidad de los datos en el almacén de datos, definir los requerimientos de seguridad, entre otros.

El almacén de datos, es un mecanismo excelente para reunir información y ponerla en las manos de los tomadores de decisiones. Sin embargo, esta información será tan buena como los datos de donde se obtenga. Un mayor esfuerzo se debe de hacer, al definir, integrar y sincronizar los datos, los cuales provienen de una infinidad de lugares y de sistemas a través de la empresa.

Los metadatos (datos acerca de los datos), son desarrollados por los mayordomos de datos, los cuales pueden ser usados por los trabajadores del conocimiento de la empresa en sus análisis diarios, para determinar que comparaciones se deben hacer, que tendencias son significativas y que las "manzanas de hecho, sean comparadas con manzanas", etc.

A medida que crece la demanda de los almacenes de datos con información de calidad, la necesidad de la función de mayordomía de los datos también ha crece. Cada vez son más las compañías que se han dado cuenta del

papel tan crítico que esta función ofrece en la búsqueda de alta calidad y disponibilidad de los datos.

El área de trabajo de un Mayordomo de los Datos

Según Imhoff (1997), una función típica a nivel corporativo de mayordomía de datos, debería tener un mayordomo de los datos asignado para cada área principal. Estas áreas, consisten de entidades de datos críticas o entidades como clientes, ordenes, productos, segmentos de mercado, empleados, inventarios, etc. Usualmente, existen de 15 a 20 áreas principales en cualquier empresa. Como un ejemplo, un mayordomo de datos debería ser responsable del área cliente y otro de productos.

El mayordomo de los datos responsable de un área, generalmente trabaja con un grupo selecto de empleados representando todos los aspectos de la empresa para esa área. También Imhoff (1997) señala que este grupo, encabezado por el mayordomo de los datos, es responsable de resolver problemas de integración en cada área. Los resultados de dicho grupo de empleados, se pasan a las funciones de administración de datos y administración de bases de datos para la implementación dentro de los modelos de datos corporativos, repositorio de metadatos y finalmente, la construcción del almacén de datos. Es recomendable que exista un líder de los mayordomos de los datos, el cual será el responsable del trabajo de los otros mayordomos, determinando y controlando su dominio.

Conclusión

En este capítulo, se ha hablado a grandes rasgos de las metodologías para alcanzar la integración de la información y de algunos de las labores del personal involucrado en esta actividad. Ahora bien, en la actualidad, los volúmenes de datos que se generan en las empresas son cada vez mayores, y esto obliga a que dichas metodologías anteriormente expuestas se apoyen con la tecnología, con el fin de hacer más eficiente y productivo el proceso de búsqueda de información. En

los últimos años, la tecnología ha dado un gran paso para facilitar el acceso a la información, y un ejemplo muy palpable lo constituye la tecnología utilizada en los servidores WWW. Esta tecnología se usa mundialmente y se conoce con el nombre de internet, pero cuando se utiliza dentro de los límites de la empresa se le conoce como intranet. El siguiente capítulo hablará sobre dicha tecnología y algunas sus ventajas.

CAPITULO III. INTRANETS

La tecnología de la intranet, se ha convertido en una herramienta de gran valor para las empresas en relación con el acceso y distribución de la información. Si a la internet se le considera la causa principal de la revolución de la información, a la intranet se le puede considerar la causa principal de la revolución de la información empresarial.

En este capítulo se verán aspectos generales de la intranet y las ventajas que esta tecnología ofrece.

INTRANET: Definiciones

Mientras que los principios de los años 90's fueron claramente los "años de internet", el año de 1996 se ha proclamado como el "año de la intranet", pero nos podemos preguntar ¿Qué es una intranet?, existen varias definiciones para describir a la intranet y a continuación se presentan algunas de ellas.

- La tecnología del Web dedicada al desarrollo de aplicaciones normales dentro de las compañías. Esos Web internos son llamados "intranets", y su crecimiento exponencial está apoyando al crecimiento de un mercado nuevo de herramientas y técnicas para facilitar su mismo desarrollo (Bickel, 1996).
- Redes corporativas internas que usan la tecnología del WEB (Gurley, 1996).
- Intranet es el último término que describe cuando los protocolos de internet y sus aplicaciones se usan no solo para acceder el vasto recurso de la internet, sino para mover información dentro de los límites de la organización (Haynal, 1996).
- Una intranet es un sistema de información interno basado en la tecnología de internet, servicios web, protocolos de comunicación TCP/IP y HTTP y publicación HTML. La internet es una tecnología que permite a una organización definirse a ella misma como una entidad completa, un grupo o familia donde todos conocen sus papeles y todos están trabajando para mejorar la organización (Hinrichs, 1977).

- Intranet es una red interna de comunicaciones corporativa que utiliza tecnologías de internet o Web como TCP/IP, HTML y Java, para ofrecer a las empresas la misma funcionalidad que el Web hace en la internet. El servidor Web que incluye la intranet reside dentro del control de acceso (firewall) de la compañía (el acceso a ellos es detenido en los límites de la empresa)
- Intranet son “internets internamente enfocadas” que representan ambientes necesarios para el apoyo a la infraestructura de comunicaciones, con el fin de encontrar ideas y recursos y contactar comunidades alejadas.
- La intranet consiste de servidores Web corporativos internos, disponibles para los empleados a través de la Red de área local, o a través de conexiones telefónicas privadas.

Historia de la Intranet y el World Wide Web

Como se pudo apreciar en las definiciones anteriores, la intranet es en pocas palabras, la aplicación de la tecnología utilizada por internet, pero con un enfoque corporativo. Por lo tanto, para hablar de la historia de la intranet, es necesario partir de la historia de internet.

Internet inició en los años 60's como un proyecto de la Agencia de Proyectos de Investigación Avanzada de la Defensa (Defense Advanced Research Projects Agency, DARPA) (Parker, 1995). La creciente importancia de las computadoras dio inicio a muchos retos, como compartir información entre sitios diversos y redes, así como mantener el flujo de información intacto, a pesar de posibles problemas que pudieran presentarse en alguno de los sitios (Telleen, 1995). Esta característica, representaba un aspecto muy importante para el departamento de la defensa de los Estados Unidos, ya que en esos años, el problema de la guerra fría y de la posibilidad de un ataque nuclear a sus instalaciones militares, así como a las ciudades principales, eran inminentes. Así, el departamento de defensa buscó la forma de que sus comunicaciones se mantuvieran en funcionamiento, aun en el caso de que se presentara cualquier clase de contratiempo.

En base a esto, la internet esta fundamentada en un conjunto de protocolos desarrollados para permitir que redes distribuidas puedan enrutar y pasar información entre ellas independientemente, de tal manera que si algún sitio esta fuera de servicio, la información se pueda enrutar a través de trayectorias alternas hacia su destino final. El protocolo desarrollado para este propósito se le llamó protocolo de internet (Internetworking Protocol - IP) (Telleen, 1995).

Parker (1995), menciona que el Protocolo Internet es responsable de mover a través de las redes, los paquetes de datos ensamblados, por medio del protocolo de control de transmisión (Transfer Control Protocol - TCP). A fin de determinar enrutamientos y destinos, utiliza un conjunto de direcciones únicas para cada dispositivo en la red.

El IP, se expandió rápidamente en la comunidad militar, como una forma para compartir información entre los investigadores militares. Debido a que los militares estaban trabajando en un gran número de proyectos de investigación en diferentes universidades alrededor del país, y a que el protocolo IP ofrecía una forma efectiva para mover la información a través de diversas redes, éste, se difundió rápidamente fuera de la comunidad militar. También, alcanzó instituciones de investigación y universidades europeas (Telleen, 1995)

El TCP, es un protocolo de comunicaciones, que proporciona transferencia confiable de datos. Es responsable de ensamblar datos de aplicaciones en capas superiores a paquetes estándar y asegurarse que los datos se transfieren correctamente (Parker, 1995)

Internet fue propuesta originalmente por la precursora de DARPA, la Agencia de Proyectos de Investigación Avanzada (Advanced Research Projects Agency, ARPA), como una forma de probar la viabilidad de las redes de conmutación de paquetes. (Cuando el enfoque de ARPA se volvió de naturaleza militar, se cambio el nombre). Durante su estadía en el proyecto, ARPA previó una red de líneas rentadas conectadas por nodos de conmutación. La red se denominó ARPAnet y los nodos de conmutación se conocieron como

Procesadores de Mensajes de Internet (Internet Messages Processors, IMPs) (Parker, 1995).

Parker (1995), menciona que ARPAnet inicialmente estaba formada por cuatro IMPs, localizados en la Universidad de California en Los Angeles, La Universidad de California en Santa Barbara, el Instituto de Investigación de Stanford y la Universidad de Utah.

A través de los años se tuvieron una serie de problemas con los protocolos y servicios de internet, pero se fueron resolviendo por medio de reuniones y así se pulieron los protocolos y servicios ofrecidos por internet. A finales de los años 80's, surgió nuevamente un problema en internet. En un principio, la información consistió principalmente de correo electrónico y archivos de datos. Algunos protocolos de correo y de transferencia de archivos han evolucionado, para tratar con dichos requerimientos. Sin embargo, empezaron a surgir varias clases de archivos nuevos en internet, como archivos de multimedia, los cuales contienen no solo sonido e imágenes, sino también hiperlazos que permiten a los usuarios saltar o navegar dentro de los archivos en una forma aleatoria, o saltar a otros archivos que contienen información relacionada (Telleen, 1995)

Telleen (1995) comenta que en 1989, el European Particle Physics Laboratory (CERN), inició un proyecto interno muy exitoso que desencadenó la creación de estándares para transmitir esta nueva clase de información a través de internet. Los componentes básicos consistieron de un estándar, para crear archivos multimedia y de hiperlazos, y otro estándar, para enviar estos archivos al solicitante. Al estándar de archivo se le llamó Lenguaje de Señalización de HiperLazos (HyperText Markup Language o HTML) y al estándar para el servidor, Protocolo de Transferencia de HiperLazos (HyperText Transfer Protocol o HTTP). Un servidor que este ejecutando HTTP, enviará archivos HTML a través de la red a los clientes que lo soliciten.

Estos dos estándares, proporcionan las bases para una nueva clase de acceso a información computacional. Crear archivos multimedia en una forma estándar, permite que se pueda construir el software del cliente para que no solo

pueda recuperar archivos de un servidor HTTP, sino también abrirlos y mostrarlos como parte de su requerimiento. Y puesto que el archivo contiene hiperlazos a otros archivos (aun cuando estos residen en otras computadoras), el usuario tendrá la facilidad de navegar con una interface muy similar a los documentos impresos. Esta tecnología, elimina la complejidad de acceder información en computadoras distribuidas (Telleen, 1995).

A los archivos de multimedia que se envían o se reciben de esta forma se les conocen comúnmente como "páginas". La información enviada a la máquina cliente que solicitó la información, se le conoce como página. Esto se debe a que la información que normalmente debería ser un documento en papel, a menudo se divide en unidades individuales más pequeñas que están hiperlazadas. Esta separación, ofrece al usuario la flexibilidad de decidir lo que desea ver, y además le ahorra tiempo, así como también se disminuye el ancho de banda requerido debido a que no es necesario transferir información innecesaria.

La mayoría de los navegadores, se pueden programar para abrir una página específica cuando se inicia una sesión. Generalmente esta página contiene un índice con hiperlazos a otras páginas que el usuario solicita frecuentemente. Esta página se le conoce como página principal (home page), y si el navegador tiene un botón llamado "home page", esta será la página que se presentará. Los documentos de hiperlazo mencionados en los párrafos anteriores, generalmente tienen una página que es el equivalente a la tabla de contenidos, o al menos ofrece un punto de inicio común para un conjunto de información. A esta página, también se le conoce como la página principal. Este tipo de páginas principales son muy usadas en las intranets como el primer contacto entre la intranet y el usuario, por lo general son muy fáciles de usar y facilitan la integración del usuario en el nuevo ambiente.

Al igual que internet, en la cual opera también el web, una página específica puede consultarse por diferentes rutas. Esta colección de páginas y de hiperlazos en la internet se le conoce con el nombre de World Wide Web (WWW o W3). Está completamente distribuida y un autor o publicador de información

general, no tiene idea de todas las páginas que tienen apuntadores hacia su página, por lo general otro publicador que haya consultado anteriormente esa página y le haya llamado la atención por alguna causa, es muy probable que incluya un lazo en su diseño hacia esa dirección.

Anteriormente, la información era muy difícil de obtener o de ofrecer, ya que se tenía un conjunto limitado de información para trabajar. La tecnología de internet ha ampliado enormemente esta base de información, ahora el problema consiste en encontrar información de calidad en internet.

Intranets: concepto

Actualmente las compañías se encuentran bajo una gran presión ocasionada principalmente por la globalización de los mercados, este tipo de competencia exige la utilización eficiente de la información interna así como también los flujos desde y hacia la empresa. Es por esto, que los administradores de los datos, necesitan crear instrumentos de información, basados en la utilización de tecnologías actuales, con el fin de ofrecer las herramientas que sus empleados necesitan, para lograr el éxito en este ambiente competitivo.

Los empleados necesitan mejorar el acceso a sus recursos de información, sin que esto signifique para la empresa una gran inversión en capital o tiempo. Las intranets ofrecen una plataforma de comunicaciones altamente efectiva, la cual está actualizada y es extensible. Una intranet básica puede iniciarse en horas o días y puede servir como un centro de información para toda la compañía.

En los años recientes, cientos de desarrolladores han creado una gran cantidad de programas, procedimientos y herramientas para internet. Las poderosas tecnologías que han resultado, incluyendo los navegadores y servidores de Web, el HTML y el HTTP, han facilitado el proceso de compartir toda clase de información en diversos sitios de internet (Derfler, 1996).

Pero el uso de las tecnologías de internet, no se tiene que limitar a una plataforma de información mundial. Cada día, más organizaciones se están dando cuenta que las mismas tecnologías utilizadas en internet se pueden usar también

para crear un sistema poderoso de información y colaboración corporativa, popularmente llamado intranet. De acuerdo a Zona Research (Derfler, 1996), una empresa de investigación de mercados de Redwood City, California, estima que para 1998 las ganancias en la venta de servidores Web para intranet, serán cuatro veces mayores que aquellas para internet.

Al darse cuenta de la gran cantidad de posibilidades que representaba el uso de la tecnología del Web dentro de las empresas, las grandes compañías desarrolladoras de software como Microsoft, Netscape, Lotus, etc., han enfocado gran parte de sus esfuerzos al desarrollo de aplicaciones para aprovechar ese potencial. Este esfuerzo, aunado a la competencia inter-empresa ha dado como resultado una carrera en el desarrollo de software, la cual en cuestión de meses ha generado cientos de productos diferentes. Inclusive, dentro de las mismas empresas, podemos distinguir versiones del mismo producto, que se actualizan varias veces al año.

El gran auge que han tenido las intranets, se puede advertir en casi cualquier revista de informática, en las cuales, es un hecho encontrar artículos y propaganda de las empresas desarrolladora, ofreciendo sus últimos productos y resaltando las características nuevas, ya sea sobre la competencia o la versión anterior. Pero no solo las revistas de informática hablan acerca de intranet, este término, desde hace tiempo sobrepasó los límites de la informática, para adentrarse en el ámbito de los negocios.

Sullivan (1996), comenta que la revolución que ha originado el mercado de las intranets, tanto en software como en hardware, ha ocasionado que el diseño y planeación del desarrollo de una red se convierta en un proceso bastante complicado. Si recordamos, hasta hace algunos años era muy fácil poder predecir la forma en la cual se vería la red interna de una compañía dentro de un año, pero la enorme cantidad de cambios en tecnologías emergentes relacionadas con internet-intranet ha convertido esas predicciones en algo inútil o en un proceso que se tiene que analizar en el corto plazo.

Debido al impacto de intranet, en 1996, Forrester Research (1996) entrevistó a 50 de las 500 compañías de Fortune y encontró que dos terceras partes ya tenían o estaban considerando en involucrarse con aplicaciones de intranet, ver figura 3.1. Estas compañías han identificado a la intranet como un mecanismo poderoso para mejorar el acceso a la información.

Planes Corporativos para el uso de Intranet

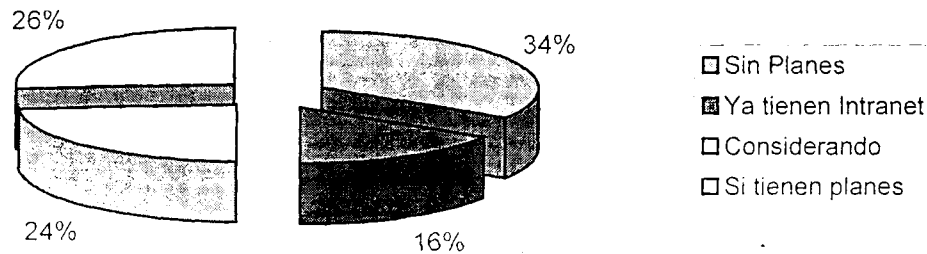


Figura 3.1 Planes corporativos par el uso de Intranet.

A que se debe su gran éxito?. Bueno, existen una infinidad de características que convierten a una intranet en una tecnología que ha atraído la atención de los empresarios y el público en general.

El auge que ha tenido la intranet, se debe principalmente a sus características, entre las que destacan su facilidad de uso, su costo, su tiempo de implantación, etc.

La ventaja mas buscada por las empresas, es la capacidad de tener acceso a información a través de casi cualquier plataforma. Lo anterior, se puede considerar como una gran ventaja para las compañías que tienen gran cantidad de sistemas operativos o equipos diferentes, y esto, es la base para buscar la integración aparente de los datos dentro de la empresa.

También, existe gran variedad de aplicaciones comunes para el usuario como email, navegadores, entre otros. En una gran variedad de plataformas, esto permite que el usuario, no tenga que invertir mucho tiempo en capacitación, al estar frente a diferentes equipos de la empresa.

Características principales de una intranet

Facilidad de Uso

Si la dificultad para acceder información, excede el valor de la misma, entonces el usuario no la utilizará o encontrará una forma alterna de obtenerla.

El navegador es el software del cliente, que recupera y muestra los archivos HTML. El primer navegador gráfico fue Mosaic, de la Universidad de Illinois. Muchos de los navegadores de hoy, están basados en Mosaic. Sin embargo, puesto que la información es estándar, no importa que navegador se use, siempre y cuando soporte los estándares (Telleen, 1995)

Los usuarios se han acostumbrado a recuperar información a través de sus navegadores, por lo tanto, el navegador se ha convertido en la interface universal para todos los tipos de información, ya sea que los recursos residan fuera en la internet, o en la computadora personal o en la red interna de la empresa.

De la misma forma, el servidor de web, se ha convertido en el anfitrión universal o coordinador de este tipo de información. Hoy, miles de servidores web ofrecen un fácil acceso a un rango increíble de recursos de información con sólo apuntar y hacer click. Esta característica, es una gran ventaja, ya que el lenguaje HTML, el cual lleva al usuario de la mano, permite un aprovechamiento máximo, con un mínimo de tiempo requerido para el entrenamiento.

Otra de las grandes ventajas que proporcionan las intranets, es que la interface con las aplicaciones es el navegador, como Netscape y Microsoft Explorer, los cuales facilitan el acceso a la información. El navegador es muy sencillo de usar y mucha gente ya ha tenido contacto con este tipo de

herramientas al acceder internet. Ahora, esta tecnología se está aplicando hacia el interior de las empresas, su facilidad de uso ofrece una gama muy amplia de posibilidades y herramientas para los encargados de presentar y generar información hacia todas las áreas de la empresa.

Al comparar la facilidad de uso de los navegadores, con los sistemas desarrollados o groupware, se puede notar una gran diferencia. El navegador es muy sencillo de usar, con una capacitación de una hora se puede lograr un gran avance, cosa que no pasa con los sistemas actuales en la mayoría de las empresas, los cuales, son utilizados en su mayoría por el personal intermedio de la empresa.

El HTML, también provee un mecanismo poderoso para la integración de sistemas de información y aplicaciones de base de datos que pueden verse en un documento con hiperlazos (los lazos de hipertexto simplifican la navegación y la recuperación de información), motores de búsqueda y formatos. Por ejemplo los manuales de procedimientos y políticas de una compañía, pueden estar principalmente en texto, pero hay ciertas secciones que pueden ser tomadas de una base de datos tales como los rangos de sueldos , puestos, diagramas de la organización, beneficios, planes de pensión, etc., (Bickel, 1996).

Es muy importante también, señalar que los navegadores actuales incluyen otro tipo de herramientas además del navegador común, como son: editores de HTML, manejadores de correo electrónico, grupos de noticias, búsquedas, etc.

Como se mencionó anteriormente, los lazos de hipertexto simplifican la navegación y la recuperación de información. Es este último principio, lo que ha generado un incremento muy fuerte en el uso de internet en los últimos años. Los usuarios simplemente seleccionan y hacen clic para navegar a través de la red mundial. Hoy, el tráfico en la red es por concepto de páginas HTML, es mucho mayor que otro tipo de tráfico en la internet y el uso de gráficas, audio, video y otros tipos de datos en los servidores continuará incrementando el tráfico.

Creación rápida de prototipos

La red internet, ofrece una arquitectura maravillosa para la próxima generación de aplicaciones cliente servidor con alcances inter e intra empresa, permitiendo a las compañías desarrollar aplicaciones interplataformas fáciles y económicas. Esto es por lo cual, muchas compañías líderes ya están usando o tienen en prototipo el uso de intranets.

Los costos iniciales para implementar una intranet son bajos. Se puede iniciar con una PC, agregarle memoria, un buen disco duro, instalar el software del servidor de Web, la tarjeta de red y así iniciar el proceso. Pero esto, puede ser un poco engañoso, principalmente basado en estudios que se han hecho en algunas compañías que ya tienen su intranet, y se debe al crecimiento exponencial que experimentan. En un principio, posiblemente se tenga información general de la empresa, pero debido a lo fácil, económico y ventajoso que resulta publicar en este medio, en pocos meses se rebasaran los límites del servidor inicial, el cual tendrá que crecer en varios aspectos como velocidad de procesamiento, capacidad de almacenamiento, memoria RAM, ancho de banda, personal administrador y de apoyo, aplicaciones de mayor calidad como conectividad con base de datos, multimedia, etc. Debido a lo anterior, la supuesta economía de crear una intranet se convierte en un gancho hacia inversiones mayores. Pero al poner en una balanza las ventajas que se obtienen contra la inversión, un esfuerzo de crear una intranet si es redituable.

Escalable

Como se mencionó anteriormente, la intranet puede iniciarse con una máquina pequeña y con software gratuito e ir escalando tanto en hardware como aplicaciones más poderosas enfocadas a las necesidades de la empresa.

Esta característica, permite que la empresa realice pruebas de dicha tecnología, sin la necesidad de arriesgar recursos financieros.

El concepto de cliente universal

Anteriormente, el problema para compartir información entre equipos de diferentes marcas, sistemas operativos, etc., consistía principalmente en que éstos eran propietarios. No se podía tener comunicación interna de información a través de ellos sin un grupo de programadores y software extra, para cada nueva necesidad. Con la tecnología de intranet, se tiene la posibilidad de acceso a más información, aplicaciones, datos, conocimiento, procesos, etc., disponibles a través de la misma ventana o el mismo navegador. No es necesario hacer conversiones entre los diferentes formatos, esperar a que los programadores codifiquen los nuevos sistemas, y el largo proceso de filtrado a través de los diferentes consultores.

Existen navegadores para la mayoría de los sistemas que soportan gráficos, estos incluyen Windows, Macintosh, X-Windows y OS/2. Existen también navegadores para sistemas no gráficos los cuales muestran las porciones de texto de los documentos consultados. La disponibilidad de navegadores a través de todas estas plataformas, es significativa. Los ambientes operativos del autor, el servidor y el visualizador del cliente, son independientes uno de otro. Los documentos construidos usando el HTML y estándares relacionados y entregados por un servidor HTTP, se pueden consultar y ver desde cualquier cliente, sin importar el sistema operativo en el cual se hayan creado o al cual sean enviados (Telleen, 1995).

El lenguaje HTML, permite el desarrollo de formas y tiene la capacidad de recibir información. Esto significa que la interface de usuario es capaz de solicitar así como enviar información. Así, se abre la posibilidad de escribir aplicaciones cliente servidor, donde el autor no se tiene que preocupar de la codificación del lado del cliente. De hecho, las aplicaciones se crean asumiendo que el cliente es el navegador.

Aún cuando esta tecnología es todavía muy joven, parece tener un efecto análogo en la administración de la información, como lo fue la invención de la

electrónica de microprocesadores de estado sólido en el hardware. Esto es, modularizando funciones y simplificando aplicaciones para permitir el cambio a un nuevo nivel de integración, que esta más cercano a las funciones del negocio de la empresa.

Uno de los usos mas exitosos del Web en los negocios es el apoyo en las consultas de información, las cuales se habrían tenido que llevar a cabo por los representantes de servicio al cliente. Las opciones como: revisar las cuentas bancarias y transacciones recientes vía una transacción encriptada (Wells Fargo Bank), o investigar sobre la localización y estado de un paquete específico enviado por mensajería (Federal Express), o encontrar oportunidades de trabajo, ver los horarios de los cines, teatros, etc. Son unas cuantas de las múltiples aplicaciones de las intranets en los negocios (Telleen, 1995).

Lo anterior, no sería posible o estaría limitado a una cantidad específica de usuarios, si no se contara con la característica de acceso a la información, desde diferentes plataformas, ofrecida por los servidores Web.

El presentar una cara común (front end) al usuario, permite que la cultura y el esparcimiento del conocimiento de los sistemas generados para intranets, tengan asesores los cuales pueden tener conocimiento en diferentes sistemas operativos, es decir, que un experto en Unix, Windows, etc. puede asesorar a un usuario final sin importar el sistema que este usando, ya que el front end, es una plataforma fácil de usar e independiente del programa que este atrás (back end). Esto, también permite que los expertos en los diferentes sistemas, puedan aportar ideas con una gran variedad de puntos de vista, situación que anteriormente no se podía aprovechar. De la misma manera, cada experto puede sugerir aplicaciones pensando en las ventajas del sistema en el es experto y así lograr un sistema de información heterogéneo pero con mayor capacidad.

Es un hecho, las intranets trabajan y trabajarán en un gran número de plataformas. Las páginas, se pueden consultar con navegadores en plataformas UNIX, Macs, Windows, etc. Pero una vez que se inicie el uso de aplicaciones y tecnologías mas avanzadas, es muy probable que se tengan que hacer ajustes

detallados, para que dichas aplicaciones sean capaces de soportar ciertos navegadores. Los desarrolladores están trabajando para resolver los conflictos y establecer estándares, pero es bien sabido como funciona esto, inclusive se puede ver que en lugar de tender hacia la estandarización, que sería ideal desde el punto de vista de los usuarios, parece ser que existe una guerra entre los dos desarrolladores principales de navegadores, (Microsoft y Netscape) los cuales, están realizando nuevos desarrollos en los que se pretende que las ventajas presentadas solo se puedan observar en sus propios navegadores.

Integración de estrategias de cómputo distribuido

Las intranets permiten integrar estrategias de cómputo distribuido, es decir que los autores del contenido tienen el recurso para publicar, más a la mano, no es necesario que se muevan de sus oficinas para poder modificar alguna página de información.

También cualquier error detectado por los usuarios de dicha página, puede corregirse al momento, así como el mismo usuario puede hacer sugerencias para mejorar el contenido o hacerlo más agradable o asimilable (a diferencia de información impresa en papel). Con este tipo de distribución, se tiene la posibilidad de que la información se pueda actualizar, sin que esto signifique un costo adicional.

Este tipo de publicación es más dinámico y existe una mayor interacción entre el generador y el receptor de la información logrando con esto mejorar la calidad de la información y un diálogo más abierto entre ambas partes.

Aprovechamiento de fuentes de información actual.

La serie de protocolos que maneja intranet, permiten que se pueda acceder información existente en las bases de datos, documentos en procesadores de texto, hojas electrónicas, presentaciones, etc.

Esta característica, permite que se pueda integrar rápidamente la información actual y darle un valor agregado, al tenerla disponible para su acceso,

en un formato más fácil y con posibilidades de llegar a mayor número de usuarios. En lugar de que la información esté guardada en archiveros o computadoras personales y que al momento de necesitarla, el usuario tenga que solicitarla, sacar copias, etc. La información puede estar disponible a cualquier hora.

Acceso a Bases de Datos

Una de las aplicaciones más actuales en intranet, es el acceso dinámico a bases de datos, el que permite, que por medio del navegador, se pueda obtener información a través de consultas, y de la misma manera, el usuario pueda introducir información a dicha base de datos y así mejorar y facilitar el proceso de adquisición de los datos.

Es muy importante, antes de decidir cual aplicación se va a seleccionar para el acceso a bases de datos a través del internet, que se analice a detalle todas las ventajas que proporciona el distribuidor del software, ya que muchas veces, solo se compra publicidad. La mayoría de los productos nuevos, aparentemente cumplen con todas las necesidades, pero muchos de ellos, todavía permanecen en versiones beta como lo menciona Rob Bilson, un webmaster de Amkor Electronics, Inc. en West Chester, Pensilvania. El agrega que el problema principal, no lo origina la conversión o la interface, sino el trabajo de limpiar los resultados, y recomienda asegurarse de obtener de los proveedores, todas las limitaciones de los productos.

Es difícil ser el primero en el campo del desarrollo de nuevos productos, muchos artículos, hablan de los fracasos o de las limitaciones que tuvieron algunas compañías al tratar de crear sus propias intranets, hace 3 años, ¿por qué?. El motivo principal, fue que existía muy poca infraestructura de software capaz de poder apoyar dichos esfuerzos. Actualmente, tan solo 3 años después, la situación es muy diferente, pero es preciso tomar en cuenta las recomendaciones anteriores, todavía existen muchas empresas desarrolladoras que hablan mucho de lo que sus productos incluirán en el futuro, pero muchas

veces, dichas características no se cumplen o se cumplen a un nivel muy por debajo de las necesidades de la empresa.

Nuevas Posibilidades

Otra gran ventaja de la tecnología de intranet, es que tienen la posibilidad de manejar otros medios, como audio y video, con lo cual, las aplicaciones se pueden hacer más atractivas e ilustrativas, mejorando así la comunicación.

La oportunidad de incluir este tipo de medios, amplía en gran medida las posibilidades de los desarrolladores y aumenta la gama de opciones, para los usuarios.

Aspectos importantes para crear una intranet.

Planeación

El proceso de establecer una intranet, aunque puede ser sencillo en un principio, demanda una adecuada planeación para evitar o disminuir al máximo los problemas derivados del crecimiento que seguramente experimentará.

Se recomienda, desarrollar un plano de la empresa para la intranet, que tipo de información se ofrecerá en un principio y hacia donde se pretende llegar en el futuro cercano. Es conveniente recalcar que como esta tecnología es muy dinámica, entonces no es recomendable planear a largo plazo.

El proceso de planeación, deberá involucrar estrategias y aplicaciones de la intranet, la revisión de hardware y software actuales y las modificaciones o adquisiciones necesarias, también se requerirá la elaboración de un presupuesto, el cual incluya todas y cada una de las inversiones necesarias para implementar la intranet.

Es muy probable que se pretenda ofrecer este servicio a todo el personal de la institución, lo cual ocasiona la actualización de muchos equipos o la

adquisición de unidades nuevas, también es necesario verificar si actualmente se esta utilizando el protocolo TCP/IP.

Es necesario, hacer una investigación realista de las necesidades de información, ya que de esto dependerán en gran medida las características del equipo que la soportará, así como también, el software necesario.

Diseño

Después de la planeación, vendrá el diseño de los elementos que configurarán la intranet, así como también, las aplicaciones con las que se manejará la información organizacional.

Se procederá a la instalación de la infraestructura para la intranet, instalando el servidor, tarjetas de red, extendiendo la red a todo el equipo que se haya planeado, instalación del software de TCP/IP verificando su buen funcionamiento, etc.

Se debe planear el uso de un servidor dedicado, para manejar el volumen de información. "Tan pronto como se empieza a poner información en el Web de la empresa, aun si eso sea solo un manual personal al que casi a nadie pudiera interesar, otros en la compañía sesearán publicar sus necesidades" comenta Clay Ryder, un analista de Inter/Intranet de Zona Research. A mayor cantidad de datos que se publiquen, mayor cantidad de consultas que se harán. Y a mayor número de consultas, mayor cantidad de usuarios esperando una respuesta (McCarthy, 1996).

Susan Goldner, administradora de Tecnologías de Internet de Federal Express, trabaja con grandes cantidades de tráfico. "Con 30,000 empleados en tres localidades diferentes están planeando contar con un servidor de intranet en cada localidad, aparte del servidor para acceso a internet" (McCarthy, 1996).

Acceso a Datos

Es muy importante crear las políticas de restricciones para los usuarios, ya que una intranet puede contener información confidencial, por lo tanto es

necesario hacer la administración correspondiente de los usuarios y establecer los criterios y controles del acceso a la información.

Herramientas

Seleccionar las herramientas de intranet más apropiadas para el desarrollo de las páginas de la organización. Existe una gran variedad de herramientas con características diferentes, por lo cual, es necesario seleccionar cuidadosamente la herramienta que se adapte a las necesidades de la empresa. Es necesario saber los requerimientos de información, los cuales pueden ser muy amplios, dependiendo del usuario al que se pretende llegar. Hay que recordar, que la información también debe estar enfocada al receptor de la misma, por lo que, no todos los formatos pueden ser útiles para todos los usuarios.

Integración

Muchas de las aplicaciones actuales podrán migrarse al ambiente de intranet, lo que requiere de un estudio de cuales seran las elegidas para pasar a dicho ambiente, asi como también todos los procesos que involucren dicha transición.

También, es necesario hacer el análisis de las aplicaciones que no pueden integrarse al ambiente de intranet y cuales pueden ser las posibles soluciones para incorporar cierta información hacia la intranet, como lo puede ser, el almacen de datos.

Seguridad

Procedimientos para los usuarios. Es indispensable que toda red, por pequeña que sea, cuente al menos con los mínimos sistemas de seguridad como el control de acceso a usuarios a la red. Esta primera barrera, consiste en evitar que usuarios externos o internos entren a los servidores de la empresa con cuentas diferentes a las de ellos. Para esto, es necesario crear una cultura de la

seguridad, la cual muestre los posibles problemas que se pueden presentar al no tener cuidado con las cuentas de usuarios y de sus respectivas palabras claves.

Una "pared de fuego" (firewall), es la forma más aceptada para asegurar que personas ajenas puedan tener acceso a datos confidenciales. Muchos proveedores de servidores en la Red, incluyendo Microsoft, Netscape y Process Software, ya han integrado firewalls a sus productos de intranets. Sin embargo, analistas y usuarios, sugieren que los firewalls, pueden ser solo la primer línea de defensa en el desarrollo de una intranet corporativa.

A medida que el tráfico aumenta en la intranet, se puede ver, como Goldner de Federal Express lo ha hecho, que los firewalls, exigen un precio al disminuir el tiempo de respuesta. Un servidor dedicado no ayudará mucho al desempeño, tanto como se podría esperar, si gran parte del tráfico de la intranet pasa primero por el firewall (principalmente tráfico de empleados que no están en las oficinas centrales) (McCarthy, 1996).

Elevar la velocidad de acceso a usuarios remotos.

Aun sin firewalls, la intranet será una devoradora de ancho de banda. Los usuarios internos de las oficinas principales, donde se tiene un backbone de fibra y equipo de 10/100 Mbps, puede que no adviertan la disminución de velocidad por los procesos de seguridad, pero los usuarios remotos si lo notarán.

Formar el Equipo -- Incluir a los Generadores de la Información.

El siguiente paso, consiste en asegurarse de que el desarrollador y su equipo, no terminen en el negocio de publicación del Web.

Para evitar lo anterior, los usuarios y analistas sugieren que se escriba una lista de información, que la compañía necesitará dentro del Web interno, después, encontrar que departamento es responsable de la creación de dicha información, de tal forma que se pueda reclutar gente, para ayudar a la captura y mantenimiento del Web interno.

La tecnología es un aliado, actualmente, la mayoría de las casas de procesadores de texto, incluyen sus productos con convertidores a formatos de HTML.

Promoción

“Ponlo a funcionar y la gente vendrá a utilizarlo”, por muy buena que sean las páginas o la voluntad de informar, siempre será necesario difundir y promocionar el uso de la intranet dentro de la empresa. Es necesario que en las reuniones de ejecutivos se anuncien los beneficios de la intranet y que se busque el mecanismo para involucrar al personal.

Una de las grandes ventajas que tiene la tecnología moderna de la intranet es que evita o disminuye el proceso de volver a capturar información. Mucha de la información que aparece en las páginas, fácilmente se puede recuperar y utilizarse en otro documento, disminuyendo así el proceso de recaptura.

De esta manera, los usuarios de la intranet, se pueden ahorrar bastante tiempo en la generación de reportes, presentaciones, etc.

Desarrollar una Política de Publicación en la Intranet.

“El Web es un mundo de información, pero también es cierto que, es difícil poder obtener información de calidad, ya que ésta, se ha introducido sin ninguna política, lo que ha ocasionado que al hacer algún tipo de búsqueda, se obtenga mucha basura o no exactamente lo que se está buscando. Esto, se debe principalmente a que existen páginas de todos tipos: comerciales, personales, o que contienen la palabra clave buscada, pero por muchas razones, ésta no está relacionada con el tema de búsqueda.

Antes de permitir que cualquier persona en la compañía produzca su material, se deben considerar una serie de aspectos, con el fin de uniformizar y estandarizar el contenido y forma, en la cual lucirán las páginas y la clase de información que tendrán dentro de la intranet. Para esto, es necesario crear una serie de políticas a lo largo de la compañía, con el fin de estandarizar dichos

procesos. Es necesario también, establecer políticas para determinar que enlaces a otros documentos deben incluirse en algunos materiales, quién tendrá acceso a este material, y que material de la intranet estará disponible para internet.

Es importante, elaborar estándares para homogenizar las páginas y facilitar la navegación dentro de las mismas. Esto, se debe hacer con el fin de disminuir el tiempo de capacitación, así como el tiempo requerido por el usuario para encontrar lo buscado. Muchas veces, los usuarios pasan de una página a otra dentro de la misma intranet y se encuentran con un ambiente totalmente diferente, y esto ocasiona pérdidas de tiempo y muchas veces de interés.

Al homogenizar las pantallas, el usuario navegará libre y rápidamente entre las diferentes páginas del web interno, y encontrará, con mayor facilidad, lo que busca.

En este capítulo, se habló de las principales ventajas de la tecnología de intranet, en referencia con el manejo de la información, pero para este trabajo, es necesario hacer un análisis de las ventajas que ésta tecnología puede ofrecer para mejorar la integración de los datos. El siguiente capítulo, está dedicado a ver en que forma las ventajas de dicha tecnología apoyarán al proceso de integración de los datos.

CAPITULO IV. LA TECNOLOGIA DE INTRANET Y LA INTEGRACION DE LOS DATOS

La tecnología de intranet, ofrece una serie de ventajas con respecto al manejo de la información dentro de las empresas, pero el objetivo de este trabajo, es el analizar la forma en la cual dicha tecnología apoya a la integración de los datos empresariales.

A continuación se harán una serie de análisis comparativos de las ventajas que brinda dicha tecnología, en diferentes escenarios.

El impacto de la tecnología de intranet y las metodologías tradicionales para lograr la integración de la información

Anteriormente, la ingeniería de información y la planeación estratégica de datos, ofrecían las herramientas esenciales para lograr un ambiente integrado dentro de la empresa, actualmente la tecnología de intranet, ofrece un apoyo a dichas metodologías, para mejorar o ampliar las opciones, en referencia con la integración de la información.

Hasta antes de la aparición de la tecnología de intranet, como Icaza (1993) menciona en su reporte, la redundancia física usualmente no era posible ni conveniente eliminarla por completo, debido a limitaciones tecnológicas. Este tipo de limitaciones, se referían principalmente, a que a través de la empresa, había una serie de sistemas en diferentes plataformas de hardware y software, entre las cuales era muy difícil o imposible la comunicación, provocando que la información se tuviera que repetir en los diferentes sistemas, en esos días, era necesario contar con un buen registro de la localización de los datos redundantes, para llevar a cabo procedimientos manuales o automatizados y así controlar la consistencia de las copias.

Actualmente, las limitaciones tecnológicas se han disminuido drásticamente, debido a la aparición de la tecnología de los WWW, la cual permite un ambiente multiplataformas, atributo que se puede considerar como uno de los

principales, desde el punto de vista de la integración de los datos. Esta nueva tecnología, nos permite disminuir la redundancia física de los datos, puesto que ya es posible tener la información de la empresa en cualquier base de datos, desde casi cualquier sistema operativo y tener acceso a ella, desde también, un número mayor de sistemas, lo cual anteriormente era imposible.

Icaza (1993) señala que la redundancia física no era conveniente eliminarla por completo, pero esto estaba analizado desde el punto de vista de las limitaciones tecnológicas que se tuvieron hasta mediados de los 90's. Actualmente, también la redundancia física posiblemente no sea conveniente eliminarla por completo, pero esto será desde el punto de vista estratégico, como podría considerarse el hecho de tener varios sitios con la misma información (lo cual sería un tipo de redundancia física), con el fin de ofrecer varias opciones de acceso para los usuarios .

Dado que la tecnología de intranet usa los mismos protocolos que internet, y por lo tanto, una intranet puede conectarse a la red mundial de internet, se puede considerar el tener redundancia física de los datos, con el fin de que los servicios de información empresariales estén disponibles en varios lugares a través del mundo, y que estos se actualicen vía la red interna o internet, dependiendo del grado de confidencialidad de la información o la seguridad requerida. Esto permitirá acceso a la información, aún si por alguna causa un sitio se encuentra fuera de servicio, lo cual puede suceder por un sinnúmero de causas (incendios, inundaciones, problemas con el suministro eléctrico, problemas con equipos, líneas de comunicación, etc.), las cuales muchas veces están fuera de nuestras manos.

Características principales de una Intranet y su relación con la integración de datos.

La tecnología de intranet tiene muchas características que pueden utilizarse para mejorar la calidad de los datos, aquí, se analizarán, aquellas que tiene más relación con la integración de los datos.

Acceso a información en diferentes sistemas operativos

Esta característica de la tecnología de intranet, nos ofrece la oportunidad de disminuir la redundancia física en los sistemas ya que no es necesario tener información repetida en diferentes bases de datos. Se puede tener un servidor o sistema con la información relevante y hacer que todos los usuarios consulten, y puedan modificar la información del mismo. Esto también apoya la definición de datos estándar reduciendo la redundancia lógica.

Facilidad de hacer cambios y modificaciones a la información desde cualquier plataforma de hardware y software

Esta característica, esta muy relacionada con la anterior, debido a que permite también que cada usuario desde diferentes sistemas operativos y diferentes localidades tengan acceso a modificar y consultar la información única. Esto evita la necesidad de repetir información única y por lo tanto reducir la redundancia física, con esta característica se brinda la oportunidad de que las aplicaciones también estén centralizadas favoreciendo así a reducir la redundancia lógica.

Menor costo de software

El costo relativamente bajo permite hacer pruebas piloto y movimientos para facilitar el acceso a los datos en diferentes sistemas dando así la oportunidad de eliminar aquellos sistemas redundantes. También ofrece la oportunidad de tener varios sistemas interconectados con la misma información. Aunque lo anterior representaría un tipo de redundancia física, esta estaría más fácilmente documentada y permitiría ofrecer la información en distintos puntos de tal manera que el usuario pudiera elegir el sitio que le ofrezca un mejor tiempo de respuesta o el más económico.

Si el sistema esta conectado a internet, el usuario podría elegir el sitio con tiempo de respuesta menor, si el sistema no esta conectado a internet, el usuario

podría elegir al sitio más cercano y disminuir el costo asociado con la larga distancia.

Menor costo de entrenamiento y menor tiempo de aprendizaje

Estos aspectos quizás no tengan directamente un efecto importante sobre la integración de la información, pero indirectamente, al disminuir el costo y el tiempo de aprendizaje y ofrecer una plataforma estandar de interfaz con el usuario esto permite que el nuevo personal cuente más rápido con información relacionada con los datos estándar de la empresa. El usuario aprendería en menor tiempo donde encontrar las definiciones de los datos estándar y su repercusión en la empresa.

Menor costo de mantenimiento

Debido a que la información tiene la posibilidad de estar integrada en un equipo, los cambios o modificaciones que se le hagan a ésta también se verán disminuidos dado que se harán a un solo sistema. Los costos y tiempos de mantenimiento se reducen y de la misma manera las posibilidades de generar redundancia lógica. Solo se tiene que trabajar en un sistema y de esta manera es menos probable que se tenga un dato con diferentes nombres o que en un nombre tenga información bastante diferente.

Análisis de la tecnología de intranet como una herramienta para la integración de los datos en las empresas.

Mucho se ha hablado de la importancia de que el desarrollo de los sistemas de información de las empresas estén alineados con los objetivos y metas de la organización. En verdad esta es una de las bases principales para que un sistema cumpla con el objetivo de ser un apoyo para la toma de decisiones.

También es cierto que la calidad de la información que entra a los sistemas debe considerarse como una de las principales características en el diseño de los mismos, como se puede apreciar en la figura 4.1 la materia prima con la que

trabajan los sistemas son los datos y si éstos son de mala calidad el resultado que los sistemas mostraran será también de mala calidad y baja confiabilidad. Hay que recordar la frase de que si se alimenta a un sistema con basura se obtendrá como respuesta basura (Garbage In Garbage Out).

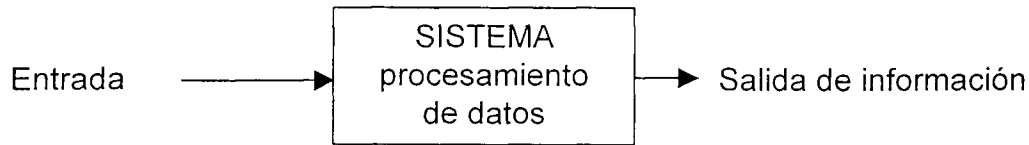


Figura 4.1 Sistema de Procesamiento de datos

Esta frase anterior, representa una gran verdad y un gran reto para los administradores y desarrolladores de sistemas. El cuadro siguiente, nos muestra que los sistemas por más bien diseñados que se encuentren, no podrán ofrecer resultados confiables, si la base sobre la que parten es de mala calidad. También es importante recalcar, que la tecnología por muy buena y sofisticada que sea, hasta el momento no es capaz de generar información partiendo de basura. Debido a lo anterior, el administrador de los datos de la empresa, debe de luchar por transmitir la importancia que tiene el proceso de captura o introducción de los datos (materia prima), para la generación de información de calidad.

Cuadro 4.1 Importancia de la calidad de los datos de entrada en los resultados de la calidad de la información de salida de los sistemas.

| Calidad de los datos de entrada al sistema | Calidad del sistema de procesamiento | Calidad de la información de salida |
|--|--------------------------------------|-------------------------------------|
| Baja calidad | Baja | Baja calidad |
| Baja calidad | Alta | Baja calidad |
| Alta calidad | Baja | Mediana calidad |
| Alta calidad | Alta | Alta calidad |

El diseñador de sistemas, debe de buscar y utilizar todas las herramientas, que le permitan llevar a las personas encargadas de capturar la información, a través de caminos bien definidos, los cuales, reduzcan al máximo la posibilidad de introducir información errónea. También, es necesario que los capturistas entiendan la labor tan importante que están realizando, en pro del buen funcionamiento del sistema y las repercusiones de su trabajo.

Factores que influyen en la integración de la información

Existen muchos factores que afectan o hacen más evidentes los problemas relacionados con la desintegración de la información. Entre ellos se encuentran: el volumen de información que se maneja, el ambiente en el cual se desarrolla la empresa, la interdependencia entre las unidades, el grado de homogeneidad del software y hardware utilizado, el factor económico, etc.

El volumen de información.

La desintegración de la información depende mucho del volumen de información que se maneja en la empresa, entre más grande sea éste así será el impacto debido a la desintegración (ver cuadro 4.2). Por lo general, las empresas que han crecido paulatinamente, tardan más en darse cuenta de este problema, ya que por lo regular, al principio, los volúmenes de información que manejan son pequeños y permiten a los encargados de generar los reportes, la realización de los mismos. Este proceso, por lo general, requiere de mucha intervención manual, y se pueden considerar verdaderos rompecabezas, que a medida que crece el volumen de información, se vuelven más complicados. Al principio, estas empresas solucionan este tipo de problemas, destinando más personal para dichas labores, pero llega el momento en el cual, el tiempo y personal requerido para lograr los reportes, sobrepasan los tiempos de respuesta y presupuesto destinado para esta labor y en esos momentos se hace completamente evidente,

la falta de integración de la información, y la necesidad de implementar las acciones necesarias para lograrla.

Cuadro 4.2 Impacto del volumen de información de los sistemas en relación con la integración de los datos y el impacto de la tecnología de intranet. Considerando que no existe planeación estratégica de datos.

| Volumen de Información | Redundancia Física y Lógica | Impacto de la Tecnología de intranet |
|------------------------|-----------------------------|--------------------------------------|
| Pequeño | Mínima (1) | Mínimo |
| Mediano | Media (2) | Mínimo |
| Grande | Gran impacto (3) | Mínimo |

1. En sistemas los cuales manejan un volumen de información pequeño, el impacto de la integración de los datos es mínimo, la redundancia física será mínima debido a que no se tendrán muchos sistemas o si se tienen, éstos manejarán un volumen muy pequeño de datos, lo que derivará en una desintegración leve y fácil de controlar. En cuanto a la redundancia lógica, el impacto es mínimo y se puede solucionar con un procesador de información manual, el cual se encargará de documentar la información requerida. En este caso, el impacto de la tecnología de intranet es mínimo.

2. En sistemas en los que se maneja un volumen de información intermedio, el impacto de la desintegración de la información empieza a tornarse en un problema serio para la empresa. Dado que el volumen de información es más complejo, los mandos medios y superiores empiezan a sentir los efectos de la desintegración de la información, cada vez es más difícil y tardado el proceso de obtención de información, el área encargada de proporcionar información dedicará horas, días y hasta meses para lograr su trabajo y cada vez requerirá de más personal. En este caso, es indispensable apoyarse en una metodología para

lograr la integración de los datos, ya que el uso de la tecnología de intranet por si misma no proporcionaría ninguna ventaja extra.

3. En este caso cuando se manejan grandes volúmenes de información, el impacto de la desintegración de los datos es contundente, debido a que es imposible contar con un equipo de procesamiento de información humano, pues éste tendría que ser enorme. En estos casos, el apoyo de la tecnología de intranet por si solo también es mínimo, pero si se apoya dicha tecnología con alguna metodología de integración de datos, se convertirá en una gran herramienta para generar información de calidad.

El ambiente en el cual se desarrolla la empresa .

El ambiente en el que se desarrolla la empresa es otro ingrediente que se debe de analizar, como se puede apreciar en el cuadro 4.3. En ambientes estables se pueden hacer planes a largo plazo e iniciar un proceso de integración de la información, basados en alguna metodología de PED, el cual puede tomar años su implementación. En ambientes turbulentos, no se cuenta con el tiempo necesario para lograr un trabajo completo de integración de la información, entonces, sera necesario optar por una acción abajo-arriba, en la cual se puede implementar una estrategia de integración aparente de los datos.

Cuadro 4.3 Impacto del ambiente en el que se desenvuelve la empresa en relación con la integración de los datos y el impacto de la tecnología de intranet.

| Ambiente | Redundancia Física y Lógica | Impacto de la Tecnología de Intranet |
|-----------------|------------------------------------|---|
| Estable | Baja | Alto |
| Intermedio | Media | Alto |
| Inestable | Alta | Alto |

1. En ambientes estables, se tiene la oportunidad de implementar una planeación estratégica de datos, la cual como se ha señalado anteriormente es un proceso largo. El ambiente estable permite que este trabajo de planeación no se vea alterado a través del tiempo requerido para su desarrollo e implementación, entonces se puede decir que el ambiente estable, favorece a la reducción de la redundancia física y lógica. En este tipo de ambiente el impacto de la tecnología de intranet es bastante considerable, ya que permitiría aprovechar al máximo las ventajas de apoyo a la integración de datos.

2. En ambientes intermedios, es preciso tener en cuenta el tiempo de duración de un estudio tendiente a disminuir las redundancias física y lógica, es necesario seleccionar el nivel deseado de integración de la información y el grado en el cual se aplicará, y el tipo de acción que se utilizará. En este ambiente, el impacto de la tecnología de intranet, será de acuerdo a las al grado de integración deseado, pero independientemente de este aspecto, la tecnología de intranet se puede considerar una herramienta de gran apoyo para mejorar la integración de los datos, ya que puede utilizarse como un instrumento para la creación de un diccionario de datos institucional, o para crear una integración parcial de los datos.

3. En ambientes inestables, el tiempo requerido para realizar un desarrollo e implementación de la PED no permite que se utilice este tipo de metodologías, debido principalmente a que al cambiar las condiciones externas a la empresa también necesitarán cambiar el enfoque de los sistemas. En estos ambientes, es cuando tecnología de intranet tiene su mayor impacto, ya que la flexibilidad y rapidez que ofrece, permite hacerle frente a los constantes cambios en el ambiente externo. En estos ambientes, la tecnología de intranet puede ayudar, como en el caso anterior, a crear el diccionario de datos institucional y también, representa una excelente plataforma para iniciar una acción de integración aparente.

El nivel de interdependencia entre las unidades.

El cuadro 4.4, muestra otro aspecto que se debe tomar en cuenta, el nivel de interdependencia entre las unidades. Entre más alto sea el nivel de interdependencia entre las unidades se favorecerá más una iniciativa de integración de datos ya que es muy probable que la información que se maneje en estas unidades sea muy parecida y los beneficios de lograr un flujo efectivo de información superarán a los costos de implementación del proceso de integración.

Cuadro 4.4 Impacto de la dependencia entre las unidades en relación con la integración de los datos y el impacto de la tecnología de intranet.

| Dependencia entre Las unidades | Redundancia Física y Lógica | Impacto de la Tecnología de Intranet |
|---|--|---|
| Alta | Minima (1) | Medio |
| Intermedia | Media (2) | Alto |
| Baja | Alta (3) | Alto |

1. Cuando existe una dependencia alta entre las unidades, esto quiere decir que existe mucho contacto y un gran flujo de información entre ellas. En estos casos, las empresas por lo general tienen hardware y software muy parecido entre las unidades, dado que por lo general se maneja el mismo tipo de información. En estos casos, la redundancia física y lógica tienden a ser bajas (aunque depende de muchos otros factores) y el impacto de una PED es grande. Aquí, el apoyo de la tecnología de intranet se le puede considerar medio, ya que solo apoyaría en los casos en los que existieran diferencias en los sistemas operativos o equipos que se manejen a través de la empresa.

2. En estos casos, la redundancia física y lógica son intermedias, es decir, existen algunas unidades o áreas de la empresa que van a tener sistemas

incompatibles. En estos casos la redundancia física y lógica empiezan a generar problemas de comunicación entre las áreas. Aquí, el impacto de la tecnología de intranet para mejorar la integración de la información se convierte en un gran apoyo.

3. Cuando la interdependencia es mínima o baja, Goodhue no recomienda una integración global, ya que esto mermaría la flexibilidad a las unidades para hacer frente a los problemas propios de sus áreas, al hacer más lentos los procesos de cambios y adecuaciones de los sistemas, debido al tiempo que se tiene que destinar a hacer trámites burocráticos para mantener una estandarización. Por otra parte, la baja interdependencia entre las unidades favorece a una desintegración de la información, ya que por lo general cada unidad se enfocará a atacar sus problemas específicos, y al hacer esto, es posible que se cuente con hardware y software apropiados para hacerle frente a sus necesidades, el cual, es casi seguro que será diferente entre las áreas.

También, en el caso del diseño de sus sistema, éstos estarán encaminados principalmente a resolver sus problemas y es casi seguro que también sean diferentes al de las otras áreas. Lo anterior favorece en gran medida a la desintegración de los datos. Aquí, existen muchas alternativas que se pueden utilizar con el fin de obtener información a nivel empresa. En estos casos, debido a la gran diversidad, tanto de software como de hardware a lo largo de la empresa, el impacto de la tecnología de intranet es bastante grande, ya que permite la creación de un lenguaje común, y facilita la creación de sistemas que tomen información de las diferentes áreas y la presente como si ésta estuviera integrada (integración aparente).

El grado de homogeneidad del software y hardware utilizado.

Es importante el tipo de hardware y software utilizado en la empresa, ya que éste es un factor muy importante que afecta a la integración de la información, como se puede apreciar en el siguiente cuadro.

Cuadro 4.5 Impacto del ambiente en el cual estan implantados los sistemas en relación con la integración de los datos y el impacto de la tecnología de intranet. Considerando el mismo número de usuarios.

| Ambiente | Redundancia Física | Redundancia Lógica |
|---------------------|----------------------|-----------------------------|
| Monousuario | Elevada(1) | Depende de ped (4) baja |
| Redes de Area Local | Baja (2) | Depende de ped (5) mínima |
| Intranet | Mínima casi cero (3) | Depende de ped (6) ≥ 0 |

1. El problema en un ambiente monousuario en referencia a la redundancia física, estriba principalmente en que la información de las bases de datos, se tiene que repetir en las máquinas principales del sistema, es decir, se tiene una copia en aquellas máquinas, que requieren información de la base de datos o que graban información a ésta.

2. En un ambiente multiusuario, el problema de la redundancia física se ve considerablemente reducido, debido a que por lo general se tiene una máquina principal o un servidor, donde se almacenan todas las bases de datos y los usuarios, tienen acceso a esa información, a través de una red local, lo que disminuye en gran medida, la necesidad de tener información repetida. El problema, se presenta cuando se trata de una empresa grande, la cual tiene varios sistemas operativos, o equipos no compatibles. Esto favorecerá a que se tenga que repetir información en los diferentes equipos o sistemas, lo cual creará redundancia física.

3. En el caso de una intranet, el problema de la redundancia física se puede considerar como estratégico o si se quiere nulo. Estratégico, debido a que si se desea repetir información, esto se hará con el fin de dar mayor rapidez de respuesta a los usuarios, es decir, si se tiene una intranet con acceso a la información empresarial a nivel mundial, podría ser deseable tener varios servidores con la misma información alrededor del mundo, de tal manera, que los usuarios pudieran escoger cual de ellos esta menos saturado, o también como

nos hemos podido dar cuenta, que muchas veces por problemas técnicos de los proveedores de acceso o de los mismos servidores, estos no se encuentran disponibles.

4. En cuanto a la redundancia lógica, en un sistema monousuario, dependerá en gran medida de la organización de la información en la empresa, es decir, si se cuenta con el apoyo de una planeación estratégica de datos. Por supuesto, en este ambiente es mucho mayor el trabajo que se tendrá que realizar para tratar de mantener en un nivel bajo la redundancia lógica. Dependerá mucho de los sistemas propios de cada una de las unidades y de que tipo de información se hará común entre ellas. Al encontrarse un poco más libres, será un poco más difícil conseguir el compromiso de las áreas para cumplir con los estándares fijados. También, esto requerirá de un adecuado sistema de comunicación y coordinación entre las áreas el cual se encargue de enviar a cada área el diccionario de datos común para la empresa y tendrá que encargarse de evaluar y conseguir el consenso entre las unidades, cuando se presenten cambios. En ambientes turbulentos, esto puede considerarse como una traba o un freno ya que cualquier cambio propuesto, principalmente aquellos originados por los ajustes necesarios, para hacerle frente a cambios inesperados, tendrá que ser autorizado por la coordinación de área, lo cual es un trámite, que por lo regular toma tiempo.

5. En un ambiente multiusuario, se contarán con más herramientas para mantener una redundancia lógica mínima y dependerá en gran medida también de la planeación de los datos, el único problema podría derivarse de tener información en varias plataformas, lo que complicaría el proceso de mantener bajo el nivel de redundancia lógica. También aquí como en el caso anterior se tendría que llegar a un consenso entre las áreas para cualquier cambio.

6. En ambientes de intranet, la redundancia lógica tiene la posibilidad de ser menor, debido a que se cuenta con las herramientas necesarias para apoyar este tipo de iniciativas, como es el caso de un diccionario de datos el cual puede estar al alcance de todas las unidades y en el cual no importa la plataforma en la cual se este trabajando, ofreciendo con esto un diccionario el cual puede estar

actualizado al momento. También, los sistemas basados en esta tecnología, no tendrían ningún problema, ya que la tecnología de intranet hace las veces de un sistema operativo universal con hardware universal. Esto es, la tecnología de intranet nos permite tener un sistema heterogéneo tanto en hardware como en software como si se tuviera todo el sistema en una red local.

El factor económico.

No se puede dejar a un lado el factor económico y menos en un país como México, el cual se ha tenido que enfrentar a varios problemas financieros. El factor económico, tiene un gran peso desde el punto de vista integración de la información, ya que muchas empresas tienen distintos equipos, de distintas marcas, de distintos sistemas operativos, los cuales son incompatibles entre ellos, es decir, la comunicación entre ellos es casi nula, lo cual favorece en gran medida a la desintegración de la información, principalmente a la redundancia física, ya que existirán muchas bases de datos repetidas a través de los diferentes equipos, también esto impide que exista una comunicación fluida y ágil, entre las diferentes áreas o los diferentes sistemas de la organización. Además, como Icaza (1993) señala, hay muchas empresas que tienen sus sistemas funcionando y están conformes con el desempeño de éstos, quizás, su principal problema sea que éstos no se pueden comunicar entre sí.

El proceso de homogenizar una plataforma de hardware y software, conlleva a una serie de desembolsos, los cuales no muchas empresas están dispuestas o no tienen la capacidad para realizarlos. El cambio de una plataforma de software, por ejemplo, muchas veces llevará también a la actualización del hardware, a la capacitación del personal, etc.

En estos casos, la tecnología de intranet, se convierte en un gran aliado para estas empresas, ya que su característica de poder instalarse en casi cualquier sistema operativo y hardware, y poder enviar y recibir información hacia y desde dichas plataformas, la convierten en una gran herramienta para mejorar o

apoyar cualquier iniciativa de estandarización de la información y para la depuración o la generación de bases de datos comunes dentro de la empresa.

Es verdad que se tendrá que invertir en la capacitación del personal, pero el tiempo requerido para lograrlo es mínimo, debido a las facilidades que ofrece esta tecnología, por la variedad de tipos de información que se pueden agregar.

La tecnología de intranet, permite a la empresa contar con una variedad más amplia de herramientas para hacerle frente al problema de la desintegración de los datos. El factor económico como se mencionó anteriormente, es un aspecto que debe tomarse muy en cuenta en nuestro país. Entonces, una alternativa confiable y económica, para integrar los datos empresariales, consiste en utilizar la tecnología de intranet en cualquiera de las opciones que se mencionaron a través de este capítulo. Como son, creación de diccionarios de datos institucionales, creación de almacenes de datos virtuales con integración aparente, etc.

Como se puede ver, la tecnología de intranet presenta una serie de ventajas estratégicas y relativamente económicas para apoyar al proceso de integración de la información empresarial.

Recientemente, el tema de los almacenes de datos ha vuelto a tomar importancia, principalmente debido a las grandes ventajas que se pueden obtener, en materia de recursos informáticos, con la aplicación de dicha metodología. El almacén de datos representa una nueva opción para los ejecutivos de la empresa, en relación con el apoyo a la toma de decisiones. En el siguiente capítulo, se hablará más a detalle de esta metodología y sus ventajas con respecto a la información que ofrecen; además, se explicará cómo esta metodología constituye una manera de lograr la integración aparente, presentada en el capítulo II.

CAPITULO V: ALMACEN DE DATOS (DATA WAREHOUSE)

En los últimos años se ha empezado a oír con más frecuencia el término de data warehouse (almacen de datos), de hecho, éste término se acuñó desde principios de los años 80's, el cual fue concebido como una forma factible de transformar la gran cantidad de datos operativo, en una herramienta inteligente de apoyo a la toma de decisiones, comenta Teresko. En aquellos tiempos, William Inmon, notó que dichos silos de información, podrían organizarse en activos de la empresa y a esto le llamó, almacen de datos. Desde entonces, la metodología se han ido refinando y se ha beneficiado en gran medida por los adelantos tecnológicos, los cuales, ofrecen un mayor espectro de oportunidades, para convertir esa gran cantidad de datos, los que anteriormente tenían un uso común, en una herramienta sólida, para brindar un apoyo confiable a los mandos superiores de la empresa.

En estos momentos, los altos ejecutivos, necesitan tener información del funcionamiento de la empresa, y ¿Qué mejor lugar para obtener dicha información, que los sistemas que apoyan la operación diaria de la misma?.

Las aplicaciones operativas de la empresa, generan una gran cantidad de datos diarios que son parte de la materia prima del almacen de datos. Es por esto, que gran cantidad de ejecutivos, sienten una gran atracción de iniciar el desarrollo de un almacen de datos. Pero el principal problema inicia aquí, al pensar que la creación de un almacen de datos, consiste en extraerlos del ambiente operativo y meterlos al almacen de datos. Nada más alejado de la verdad comenta Inmon (1996).

Inmon (1996), señala que el almacen de datos, es el centro de la arquitectura de los sistemas de información de los años 90's y por lo tanto, la base del procesamiento de los Sistemas de Soporte a la Toma de Decisiones (DSS- Decision Support Systems), y lo define como, "una colección de datos orientada a temas o especializada, integrada, no volátil y variante con el tiempo, la cual, apoya a la toma de decisiones a nivel ejecutivo".

Los datos que entran al almacén de datos, provienen del ambiente operativo, en casi todos los casos. Como Inmon (1996) afirma, "El almacén de datos, es siempre un almacén de datos transformado, físicamente separado de los datos de las aplicaciones que se encuentran en el ambiente operativo"

Colección de datos orientada a temas

Se dice que el almacén de datos es una colección de datos orientada a temas o especializada, por que apoya a los ejecutivos en la toma de decisiones de temas específicos. Las aplicaciones clásicas están organizadas alrededor de las operaciones de la misma compañía. Por ejemplo, en un banco, las aplicaciones operativas, pudieran ser: préstamos, ahorros, tarjetas de crédito, etc. Para este mismo ejemplo, los temas especializados pudieran ser clientes, productos, actividades (Inmon,1996).

El ambiente operativo se preocupa tanto por el diseño de la base de datos como del diseño del proceso. El ambiente del almacén de datos se enfoca en la modelación de los datos y el diseño de la base de datos.

Colección de datos Integrada

El proceso de convertir los datos de las aplicaciones operativas en un almacén de datos conlleva una serie de pasos entre los que destacan la estandarización de las diferentes fuentes de información, es decir la integración de los datos a nivel empresarial.

De todos los aspectos relacionados con el almacén de datos éste es el más importante ya que constituye la base para la calidad en la toma de decisiones. Como se ha dicho anteriormente, en la mayoría de las empresas, las aplicaciones que se han desarrollado a través de los años, con el fin de satisfacer objetivos particulares y sin tomar en cuenta las necesidades de intercambio de información futuras.

Una característica muy importante, del ambiente creado por la arquitectura del almacén de datos, es el proceso de integración de los mismos, que ocurre a

través de dicha arquitectura. A medida que los datos pasan del ambiente operativo, al ambiente del almacén de datos, éstos son integrados.

No existe forma en la cual, se traigan datos del ambiente operativo al ambiente del almacén de datos sin integrarlos. Si estos llegan en un estado desintegrado, no podrán ser usados para lograr un apoyo a la toma de decisiones empresarial, por lo tanto, una integración corporativa de los datos, es la esencia del ambiente generado por el almacén de datos.

Colección de datos Volátil

La tercer característica importante de un almacén de datos, es que la no volatilidad. Por lo regular, los datos del ambiente operativo se accesan y manipulan regularmente un registro a la vez. La actualización de los datos, se hace en el ambiente operativo. Pero los datos del almacén de datos, muestran un conjunto muy diferente de características. Los datos del almacén de datos, se cargan (regularmente en volumen) y se accesan. Pero la actualización de los datos (en términos generales), no ocurre en el ambiente del almacén de datos.

Colección de datos Variable en el Tiempo

La última característica importante que destaca Inmon (1996) en el almacén de datos, es la variabilidad en el tiempo:

- El horizonte de tiempo para el almacén de datos, es significativamente mayor que el de las aplicaciones operativas. Un horizonte de tiempo de 30 a 90 días, es normal para las aplicaciones operativas; un horizonte de tiempo de 5 a 10 años, es normal para el almacén de datos.
- Las bases de datos de las aplicaciones operativas, contienen el valor actual de los datos cuya exactitud es válida en el momento del acceso, el cual también puede ser actualizado. Los datos del almacén de datos, no son nada más que una serie sofisticada de fotografías tomadas en cierto momento en el tiempo.

- La estructura clave de los datos operativos, puede o no tener elementos de tiempo, tales como año, mes día, etc. La estructura clave del almacén de datos, siempre contiene algún elemento de tiempo.

La Estructura del Almacén de Datos

El almacén de datos, esta estructurado como se muestra en la figura 5.1. La figura 5.1, muestra que existen diferentes niveles de detalle en el almacén de datos. Existe un nivel más viejo de detalle, un nivel de detalle actual, un nivel de detalle ligeramente resumido (datamart), y un nivel de datos altamente resumido.

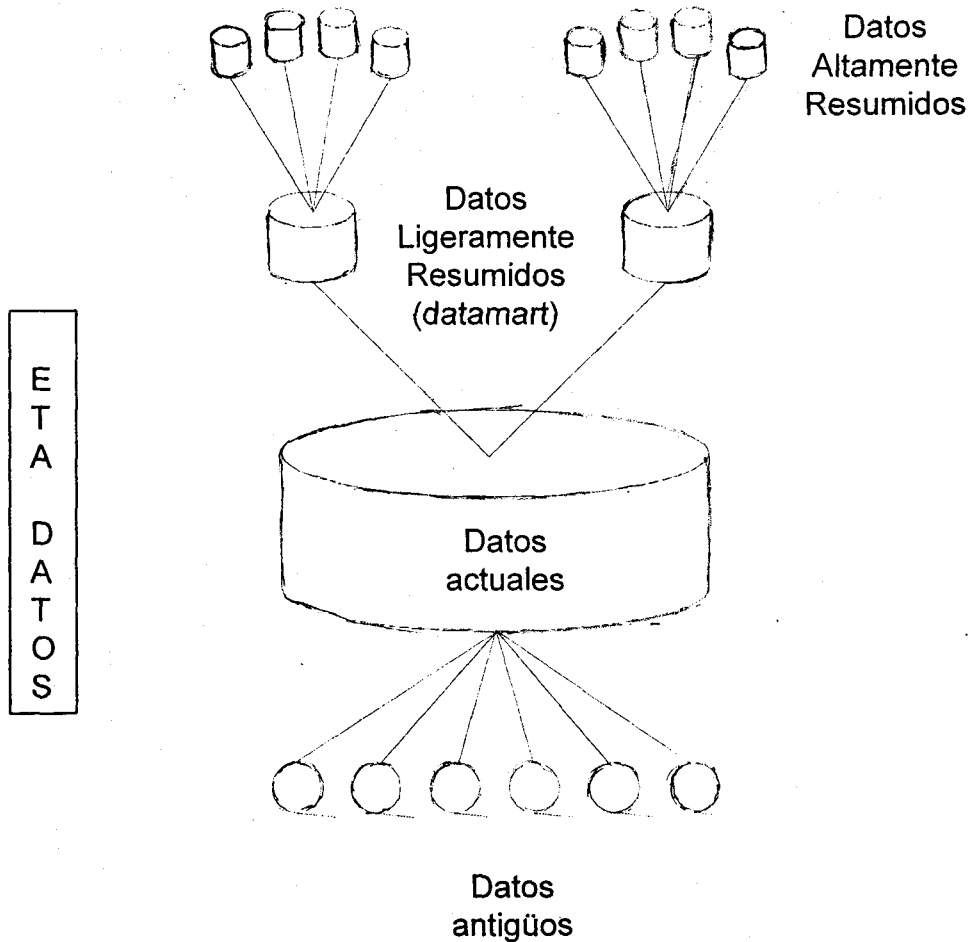


Figura 5.1 La estructura del almacén de datos (Inmon, 1996).

Los datos fluyen desde el ambiente operativo, hacia el almacén de datos y por lo regular, los datos experimentan una cantidad de transformaciones.

Una vez que los datos envejecen, pasan del detalle actual, al detalle antiguo. Al resumirse los datos, pasan del detalle actual, al ligeramente resumido y del ligeramente resumido, al altamente resumido.

La creación de un almacén de datos, no es un proceso sencillo de extracción de datos del ambiente operativo de la empresa, como ya se ha mencionado, éste, involucra una serie de acciones, que requiere de grandes inversiones de tiempo, diseño, infraestructura, etc. De nada serviría crear un gran almacén de datos que no ofrezca información confiable.

La integración de la información que entra al almacén de datos, es el aspecto más importante que se debe tomar en cuenta. Este proceso, involucra un gran esfuerzo de estandarización y modificación de la fuente de datos, proveniente del ambiente operativo.

De la integración de los datos y de la importancia de la misma ya se ha hablado a través de este trabajo, por lo tanto, aquí solo se mencionará que esta característica, es la más importante dentro del almacén de datos.

Por otra parte, la integración o la falta de ella en los sistemas actuales, no es la única dificultad en la transformación de los datos del ambiente operativo, al ambiente del almacén de datos. Otro problema bastante fuerte, lo constituye la eficiencia en el acceso a los datos sistemas actuales. ¿Cómo sabe el sistema que adquiere información de los sistemas del ambiente operativo, si ya adquirió información de dicho sistema?

Para formar el almacén de datos, se necesita de tres tipos de adquisición de datos: la carga de datos archivados (históricos), la carga de datos del ambiente operativo y la carga de los cambios o modificaciones desde la última actualización.

El primer tipo, no representa un gran reto, ya que se hace una sola vez o pocas veces, es decir, se toma información que se considera reelevant de los datos históricos almacenados. Datos de más de un año de antigüedad.

En esta misma línea, se encuentran los datos que se cargan de los sistemas actuales, es decir también se hace una sola vez. Esta información puede ser de los datos del presente año.

El reto principal, lo constituye la carga de los datos actuales a medida que se efectúan los cambios en el sistema, es decir, los cambios recientes. Los problemas por lo regular se tienen, por que dichos datos, muchas veces no cuentan con un atributo de fecha, que permita tomar la información en cierto período de tiempo.

Debido a lo anterior, existen varias técnicas que ayudan a poder seleccionar dichos cambios en la información o la información más reciente. Una técnica, consiste en tomar información de un archivo delta. Este archivo, es creado por la aplicación y graba los cambios hechos a la información.

Otra técnica, consiste en tomar datos de un archivo de auditoria o archivo log, el cual, contiene esencialmente la misma información de un archivo delta.

También, se puede cambiar el código de los sistemas actuales, con el fin de poder agregarle un elemento de tiempo, que facilite la obtención de los datos.

Inmon (1996), menciona estas técnicas, las cuales considera se deben analizar para seleccionar la más adecuada, ya que cada una de ellas tienen sus ventajas y desventajas, y por lo tanto, requieren de un análisis, el cual involucra muchos factores como pueden ser: las aplicaciones actuales de la empresa, presupuesto, tiempo, etc.

Además de la integración y el desempeño, Inmon (1996) menciona otros aspectos importantes en el proceso de creación del almacén de datos: los cambios que experimentan los datos a través del tiempo y el volumen de los mismos.

En el núcleo de un ambiente estructurado Inmon (1996) ha identificado dos clases de datos: primitivos y derivados. Los datos primitivos, son datos que sirven para llevar a cabo las operaciones diarias de la compañía. Los datos derivados, son aquellos que están en forma resumida o a los que se les han efectuado cálculos para cumplir con las necesidades de los ejecutivos de la compañía. Los

datos primitivos, se pueden actualizar, mientras que los derivados no. Los datos primitivos son actuales y los derivados son históricos. Los datos operativos son primitivos y los datos de los Sistemas de Apoyo a la Toma de Decisiones, son derivados. Los datos primitivos, apoyan a las funciones cotidianas, mientras que los datos derivados, apoyan a las funciones administrativas.

Los datos más actualizados, son los primitivos y se encuentran en las aplicaciones que soportan la actividad diaria de la empresa. Cuando se extraen datos de este ambiente hacia el almacén de datos, se convierten en datos derivados, los cuales, ya no se pueden actualizar, sin embargo, todos sabemos que a diario, se hacen modificaciones a los datos primitivos, producto de errores, cambios, devoluciones, etc., originando así, diferencias con la información que contiene el almacén de datos. Este error, permanecerá ahí, hasta que se vuelva a hacer la actualización de la información en el almacén.

Derivado de lo anterior, es imprescindible que la información en el almacén de datos, cuente con un elemento de tiempo (fecha, hora) ligada a dicha información, la que haga referencia al momento en el cual se obtuvo la información, es decir, "esta información" es válida, para esta fecha.

Es importante también, establecer cada cuando se va a actualizar la información del almacén de datos, debido a que entre más seguido se realice dicho proceso, mayor serán los recursos necesarios para poder soportarlo.

Granularidad

El aspecto más importante en el diseño de un almacén de datos, es la granularidad. Granularidad, se refiere a el nivel de detalle o resumen que tienen las unidades de datos, en el almacén de datos. A mayor nivel de detalle, es el menor nivel de granularidad.

La granularidad de los datos, ha sido siempre un elemento de diseño. En las primeras aplicaciones operativas que se construyeron, la granularidad fue dada por un hecho. Cuando los datos se actualizan, es casi un hecho que se

almacenarán al menor nivel de granularidad. Pero en el ambiente del almacén de datos, la granularidad no se puede dar por hecho.

Inmon (1996) señala, que la razón por la cual la granularidad es el principal elemento de diseño, en el ambiente de almacén de datos, es que ésta, afecta profundamente la cantidad de datos que se almacenan y a la riqueza de las búsquedas, que se pueden realizar. El volumen de datos en el almacén de datos, irá en relación directa, al nivel de detalle deseado en las búsquedas.

El aspecto de granularidad, se puede ejemplificar con lo siguiente: suponga que se tiene una compañía de telefonía celular, la cual graba los datos relacionados con cada llamada que realizan sus clientes. Este tipo de almacenamiento de información a gran nivel de detalle, se le denomina como un bajo nivel de granularidad. Al final del mes, cada cliente tiene un promedio de 200 registros (uno por cada uno de los datos que se almacenan con cada llamada efectuada, a lo largo del mes), que requiere alrededor de 40 KB de espacio. Las ventajas de este nivel de granularidad, es la capacidad de poder responder a casi cualquier pregunta que se desee, con respecto a dicho cliente, las desventajas, están relacionadas con el nivel de almacenamiento requerido, para tener el registro de todos los clientes y el tiempo necesario, para encontrar la información requerida entre tantos datos.

Por otro lado, un alto grado de granularidad, podría consistir en que la compañía grabara un resumen de las actividades de sus clientes mes con mes. En este caso, el volumen requerido para almacenar los datos de sus clientes sería mínimo, las búsquedas rápidas, pero las desventajas, derivarían principalmente de la incapacidad de contestar preguntas más específicas de los clientes.

Entonces, el diseñador del almacén de datos debe buscar el equilibrio entre tener la capacidad para contestar cualquier pregunta y que el volumen de información sea tan grande, que las búsquedas sean lentas y que los costos de almacenamiento sean demasiado grandes. Por otra parte, debe estar consciente de que esta información, crecerá mes con mes, pero también, la tecnología tendrá lo suyo, en cuanto a adelantos relacionados con

almacenamiento de mayores volúmenes de información a menor precio, mayores velocidades de acceso, etc.

Como una forma de resolver este problema, Inmon (1996) señala que muchas empresas han adoptado sistemas de doble o múltiple granularidad. Uno que permita que se almacene una gran cantidad de datos, los requeridos para contestar las preguntas de los tomadores de decisiones, el cual deberá estar en los equipos más modernos y rápidos y tendrá la información más actual y otro u otros, que estarán en medios más lentos pero económicos, con bajos niveles de detalle.

Es decir, se almacenan en cintas los datos históricos y en discos duros los datos más actuales, así como los resúmenes de datos históricos más inmediatos. Hay que tener en cuenta, que los medios de almacenamiento tienen ventajas y desventajas, se tienen medios económicos de gran capacidad, pero tienen la desventaja que, por lo regular son lentos y requieren de apoyo manual (montar la cinta con la información de hace 5 años, etc.) y los medios de almacenamiento rápidos, por lo general son más caros, ver cuadro 5.1.

Cuadro 5.1 Medios de almacenamiento de datos, su velocidad y precios.

| TIPO DE MEDIO | VELOCIDAD | PRECIO |
|-------------------|------------|----------|
| Memoria Principal | muy rápida | muy cara |
| Memoria expandida | muy rápida | cara |
| Memoria cache | muy rápida | cara |
| Discos duros | rápidos | moderado |
| Discos ópticos | no lentos | no caros |
| Cinta | lenta | barata |

Partición

Otro aspecto importante en el diseño de un almacén de datos, es la partición. La Partición, se refiere a la división de los datos en unidades físicas separadas, que pueden ser manejadas independientemente.

División de los datos en unidades pequeñas, más fáciles de manejar. Dichas unidades de datos, pertenecen a una y solo una partición. Estas divisiones, pueden ser por fecha, línea de negocios, geográfica, unidad organizacional o mezcla de las anteriores. Sin embargo, en el ambiente de el almacén de datos, casi siempre es obligatorio que uno de los criterios para particionar los datos, sea la fecha.

A menudo, se dice que si la granularidad y la partición de los datos se hace de una manera correcta, entonces, casi todos los otros aspectos del diseño e implementación del almacén de datos, se lograrán fácilmente.

Metadatos

Por una variedad de razones, los metadatos llegan a ser aún más importantes en el ambiente del almacén de datos, que en el ambiente operativo. En el ambiente operativo, los metadatos son relegados al mismo nivel de importancia que la documentación. Sin embargo, los metadatos, en el ambiente del almacén de datos son esenciales, debido a que los datos, sufren una transformación muy significativa a medida que pasan desde el ambiente operativo, al ambiente del almacén de datos.

Entre los cambios que se pueden destacar están: la conversión, filtrado, resumen, cambios estructurales, etc. Existe pues, una necesidad de mantener un registro adecuado de los cambios y el lugar ideal para hacerlo en el almacén de datos, son los metadatos.

La importancia de mantener este cuidadoso registro de la transformación, se ve reforzada por los eventos que ocurren cuando un administrador necesita rastrear los datos desde su origen (el proceso de drill - down).

Smith¹ (1977), en su escrito de cómo implementar una estrategia de metadatos, señala que: "Metadatos son datos referentes a los datos". Es la colección de información acerca de los datos, reunidos en una aplicación o base de datos, como ejemplos de metadatos podemos incluir: la definición del elemento dato, nombres de los elementos del negocio, abreviación de dichos elementos, el tipo de datos (alfanumérico, decimal, fecha, etc.) y el tamaño del elemento, fuente, localización, transformaciones, etc. Todas estas piezas de metadatos, son de interés para varios miembros de la organización; algunos son de interés solo de ciertos miembros del staff de IS, mientras otras piezas podrían ser muy útiles para la gente de la empresa que desea navegar en el Almacén de Datos Corporativo.

Una estrategia de metadatos, puede apoyar en el logro de las metas orientadas a los datos, proporcionando un enfoque, para la compartición de los recursos informáticos de una organización (Smith, 1977¹). Esto ofrece un reconocimiento del valor de los datos, sus componentes y uso dentro y a través de la organización. Una estrategia de metadatos, puede proporcionar un mapa, para administrar los requerimientos expansivos de información, que el negocio coloca en el ambiente de los sistemas de información. Una estrategia de metadatos, resalta la importancia de un departamento central de administración de datos, para las organizaciones que están preocupadas por la calidad de los mismos, su integridad y rehusos. Por último, el desarrollo e implementación de una estrategia de metadatos, ofrece a la organización, un medio para iniciar la medición del valor de los recursos informáticos bajo su control.

Modelos de Datos y de Procesos y el Almacén de Datos

Antes de tratar de aplicar cualquier técnica de diseño, Inmon (1996) sugiere, que el diseñador debe entender la aplicabilidad y los límites de las mismas. El modelo de procesos, aplica solo al ambiente operativo, mientras que el modelo de datos, se puede aplicar a ambos: al ambiente operativo y al ambiente del almacén de datos.

Un modelo de procesos, consiste típicamente de lo siguiente:

- Descomposición funcional
- Diagrama de contexto nivel cero
- Diagrama de flujo de datos
- Estructura organizacional
- Diagrama de transición de estados
- Pseudocódigo

Existen muchos contextos y ambientes en los cuales un modelo de procesos es invaluable, pero al construir un almacén de datos, es un obstáculo.

El Almacén de Datos y los Modelos de Datos

Como se mencionó anteriormente, el modelo de datos se puede aplicar tanto a los sistemas actuales como al ambiente del almacén de datos. Entonces, si se desarrolla un modelo de datos corporativo, éste se hará sin importar diferencias entre los sistemas de apoyo a la operación y el almacén de datos. Por su naturaleza, el modelo de datos corporativo, contiene solo los datos primitivos. Para construir un modelo de datos para cualquier ambiente, es necesario que el punto de partida sea el modelo de datos corporativo. Sin embargo, este nuevo modelo debe de incluir factores de desempeño, a medida que el modelo es transportado al ambiente de los sistemas actuales. Por lo general, son pocos los cambios que se le hacen al modelo de datos corporativo al transportarlo al ambiente operativo.

Sin embargo, Inmon (1996) determina que existen un número considerable de cambios al modelo de datos corporativo, cuando éste se aplica al ambiente del almacén de datos. El primer paso, consiste en eliminar aquellos datos que solo se usan en el ambiente operativo. Paso siguiente, las estructuras claves del modelo de datos corporativo, son reforzadas con el indispensable elemento de tiempo. También, se agregan datos derivados al modelo de datos corporativo, de forma que los datos derivados, sean calculados solo una vez, y no cada vez que se

utilizan. Definitivamente, esto se puede considerar como una falta grave de integración de datos (redundancia física), pero aquí, se hace por motivos de desempeño y ésta, debe estar perfectamente documentada.

Por último, se hace un análisis de estabilidad, entre el modelo de datos corporativo y el modelo de datos del almacén de datos. El análisis de estabilidad, consiste en agrupar atributos de datos, basados en la tendencia al cambio.

Es decir, los datos que cambian rara vez, se agrupan con otros datos similares, los datos que cambian a menudo, son agrupados con datos que cambian a menudo. El resultado del análisis de estabilidad, es agrupar datos con características similares.

A grandes rasgos, estos serían los aspectos principales que relacionan al almacén de datos con los modelos de datos empresariales.

Data mining

Otro término muy relacionado con el almacén de datos, es el data mining (explotación de datos). El data mining, es el análisis de las relaciones entre los datos, que previamente no se habían descubierto. Por ejemplo, los registros de ventas de una marca particular de tenis, podrían, si es que son suficientemente analizadas y relacionadas a otro mercado de datos, revelar una correlación de temporada, con la compra de ropa deportiva de la misma marca .

Los resultados del data mining incluyen:

- Asociaciones, o cuando un evento puede estar correlacionado a otro evento (consumidores que compran cerveza, también compran botanas en un cierto porcentaje de las veces)
- Secuencias, o un evento que lleva a un evento posterior (la compra de una alfombraseguido por la compra de cortinas)
- Clasificación, la identificación de modelos o patrones y una nueva organización de datos resultante (por ejemplo, perfiles de clientes que hacen un particular tipo de compras)

- Agrupamiento, o encontrar y visualizar grupos de hechos que no se conocían anteriormente.
- Predicciones, o simplemente descubrir patrones en los datos que pueden llevar a predicciones futuras.

Datamart

Un datamart, es un depósito o almacén de datos obtenidos de datos operativo y de otras fuentes, que está diseñado para servir a una comunidad específica de tomadores de decisiones. Dependiendo de su alcance, los datos pueden provenir de una base de datos empresarial o de un almacén de datos. El énfasis de un datamart, consiste en cumplir las demandas específicas de un grupo particular de usuarios del conocimiento, en términos de análisis, contenido, presentación y facilidad de uso.

En la práctica, los términos datamart y almacén de datos, cada uno tiende a implicar de alguna forma, la presencia del otro. Sin embargo, muchos escritores que usan los términos, parecen estar de acuerdo que el diseño de un datamart, tiende a iniciar desde un análisis de las necesidades del usuario y el data warehouse tiende a iniciar del análisis de los datos existentes y como éstos, pueden ser colectados en forma tal, que se puedan usar posteriormente. Un datamart, es un depósito de datos que puede derivar de un almacén de datos o no y que se enfoca principalmente a la facilidad de acceso y uso para un propósito particular. En general, un almacén de datos, tiende a ser un concepto estratégico muy general y un datamart, tiende a ser táctico y enfocado a apoyar una necesidad inmediata y específica.

Cuando el datamart proviene de un almacén de datos corporativo, su objetivo consiste en crear pequeños almacenes satélites para atender las necesidades específicas de una unidad geográfica o departamental, sin tener que conectarse al almacén central, esto, desde luego, requiere una organización de datos distribuída que debe mantenerse en sincronía absoluta. La estrategia de datamart, no solo ha sido usada como un complemento al almacén de datos

central, sino como una estrategia sustituto, esto quiere decir que, en lugar de crear un un gran almacén de datos corporativo, se implantan pequeños almacenes que satisfacen necesidades particulares, y que extraen la información directamente de los sistemas de operación, como se puede apreciar en la figura 5.2 Este enfoque permite crecer modularmente, sin la necesidad de realizar grandes inversiones iniciales, obteniendo resultados en tiempos relativamente cortos. Por otro lado, en este tipo de arquitectura, nos enfrentamos al problema de no contar con un modelo de datos corporativo y centralizado, lo que podría traer como consecuencia problemas de integración.

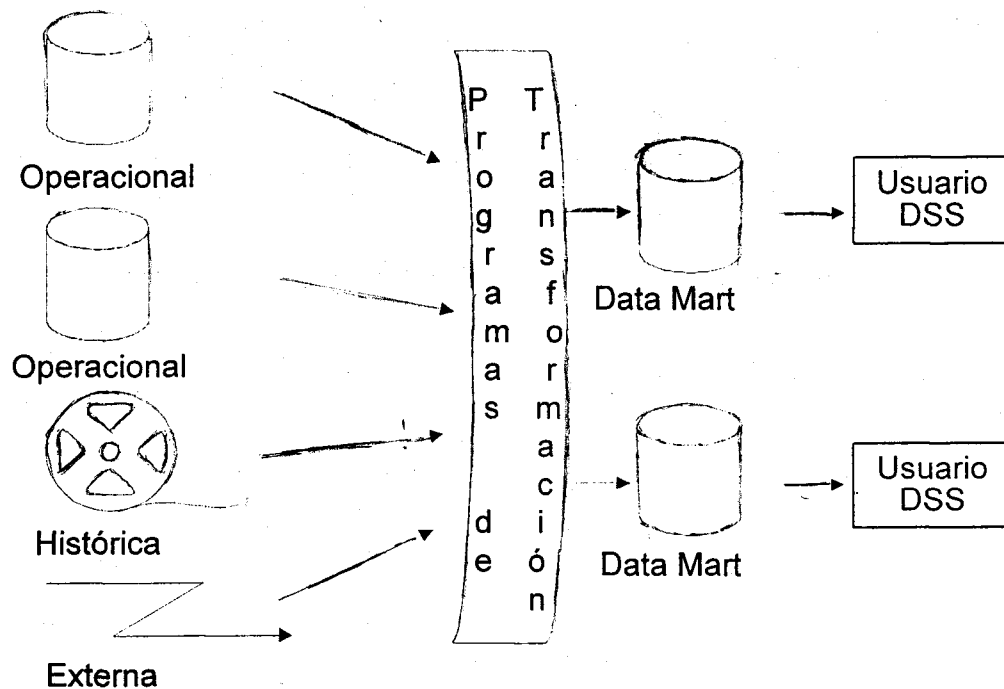


Figura 5.2 Data Marts (Miramón,1996).

La Tecnología del Web y el Almacen de Datos

Actualmente, muchas compañías están haciendo los movimientos necesarios para unir herramientas de información, tales como las almacen de datos y los grandes bancos de datos, utilizando la tecnología del Web, con el fin de crear Portales de Información Empresariales (Enterprise Information Portals - EIP) que permitan a los usuarios, acceder datos grabados en cualquiera de dichas aplicaciones. EIP es una estrategia de manejo de datos que provee una ventana hacia el conocimiento empresarial, reviviendo datos anteriormente inaccesibles o en estado letárgico, para que estos puedan ser comparados, analizados y compartidos por cualquier usuario en la organización. Dichas compañías que han adoptado los EIP, han adquirido rápidamente una ventaja competitiva, encabezada por bajos precios, incremento en ventas, un mejor despliegue de recursos y mejoras en la productividad interna.

La tecnología Web, permite al usuario tener acceso a la información de ambientes de cómputo distribuidos, como una unidad estándar, sin preocuparse de su localización física o el formato en que se encuentra.

Este enfoque, ha sido popularizado por Information Builders, a través de la comercialización del middleware EDA/SQL, que permite el acceso a múltiples bases de datos o estructuras de archivos en diversas plataformas de hardware, el almacen de datos virtual, ver figura 5.3, brinda a los tomadores de decisiones, además del acceso, una vista relacional y uniforme de la información de los sistemas, que soportan la operación diaria de la empresa. La implantación, se fundamenta en la instalación de un repositorio centralizado de metadatos, donde se manejan las reglas del negocio y los apuntadores a donde se encuentra localizada físicamente la información, ya que no existe un almacen de datos físico.

Esta aproximación, facilita el acceso de datos de manera corporativa, permitiendo la incorporación de diversas aplicaciones, sin la necesidad de invertir en la infraestructura de hardware para soportar su operación, sin embargo, este

enfoque es limitado, ya que cuando se manejan grandes volúmenes de información, puede llegar a afectar el desempeño de los sistemas de transacción

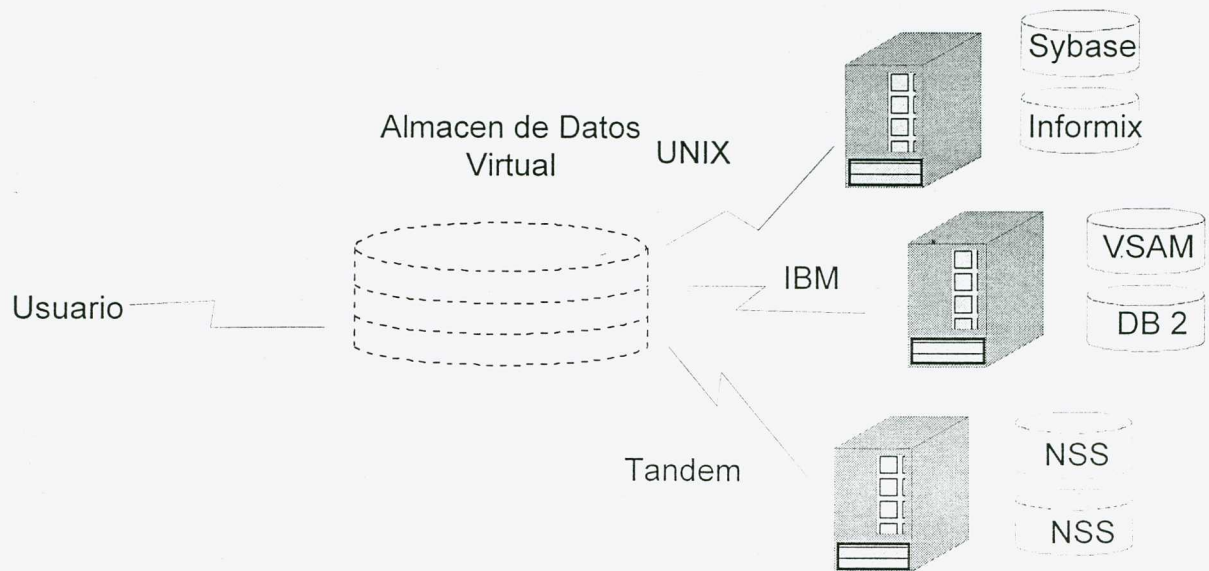


Figura 5.3 Almacén de Datos Virtual.

en línea (OLTP siglas en inglés). Aunado a lo anterior, esta alternativa también presenta el problema de la información histórica, ya que como sabemos, los sistemas en el ambiente operativo por lo regular, no guardan información más allá de un año, así como la creación de resúmenes de información, aspectos que son manejados en el almacén de datos.

En este capítulo, se ha hablado sobre la metodología del almacén de datos y sobre temas afines. Queda claro que esta metodología constituye una forma de lograr la integración aparente de datos, ya que los usuarios del Almacén, tienen impresión de estar accediendo datos realmente integrados, cuando éstos en realidad se encuentran en un repositorio intermedio (el almacén) alimentado con datos que pueden o no estar integrados. Esta metodología presenta, junto con la

tecnología de intranet, una serie de ventajas con respecto a la creación y manejo de nuevas fuentes de información. El capítulo siguiente, es un caso práctico, en el cual, se analizarán las ventajas que pueden ofrecer dichas metodologías, con el fin de lograr la integración aparente de los datos, en el CIAD, el cual, es un centro de investigación, que como varias empresas, tiene problemas de desintegración de la información.

CAPITULO VI. CASO PRÁCTICO CIAD

En este trabajo de tesis, se decidió incluir una empresa, que tuviera algunas características, que pudieran ser utilizadas para exponer las ventajas de la tecnología de intranet, como una herramienta para lograr la integración aparente de los datos empresariales. En este caso práctico, se tomó como ejemplo al CIAD, una institución que como muchas otras, han pasado de ser una empresa pequeña a una mediana. Sus problemas relacionados con los sistemas de información internos, los cuales fueron desarrollados con el objetivo de satisfacer las necesidades propias de cada departamento y el problema de desintegración de los datos que generaron. En este capítulo, también se analizan posibles soluciones, para lograr una integración aparente, basados en la tecnología de la intranet y los beneficios que se pudieran obtener al aplicar dicha tecnología.

¿Qué es el CIAD?

El CIAD, es el Centro de Investigación en Alimentación y Desarrollo, A.C., es una asociación civil, dedicada a la investigación científica y tecnológica. Fue fundado en la Ciudad de Hermosillo, Sonora el 16 de marzo de 1982 por acuerdo de la Secretaría de Programación y Presupuesto, publicado en el Diario Oficial de la Federación el 22 de Septiembre de 1981. Los organismos asociados que constituyen el CIAD, A.C. desde su fundación, son: La Secretaría de Educación Pública, el Consejo Nacional de Ciencia y Tecnología, el Gobierno del Estado de Sonora, la Universidad Autónoma de México y el Instituto Politécnico Nacional. Los objetivos del CIAD son la realización de investigación científica y tecnológica, la formación de recursos humanos a nivel posgrado y la vinculación con el sector productivo.

Actualmente, el CIAD en Hermosillo cuenta con las direcciones de Nutrición, Ciencia de los Alimentos y Desarrollo Regional, Tecnología de Alimentos de Origen Animal y Tecnología de Alimentos de Origen Vegetal, también cuenta con 4 unidades foráneas: la primer unidad foránea, se creó en

Mazatlán, Sin., en agosto de 1993 y su especialidad es la Acuicultura y Manejo Ambiental, en octubre de 1993 se creó la unidad Culiacán y su especialidad es la Fisiología y Tecnología Poscosecha de Productos Hortícolas, la unidad Guaymas, cuyo objetivo es realizar investigación aplicada y brindar servicios de apoyo y asesoría a la industria pesquera y la unidad más reciente, en Cd. Cuauhtemoc, Chihuahua con la especialidad en Tecnología Hortofrutícola (ver figura 6.1).

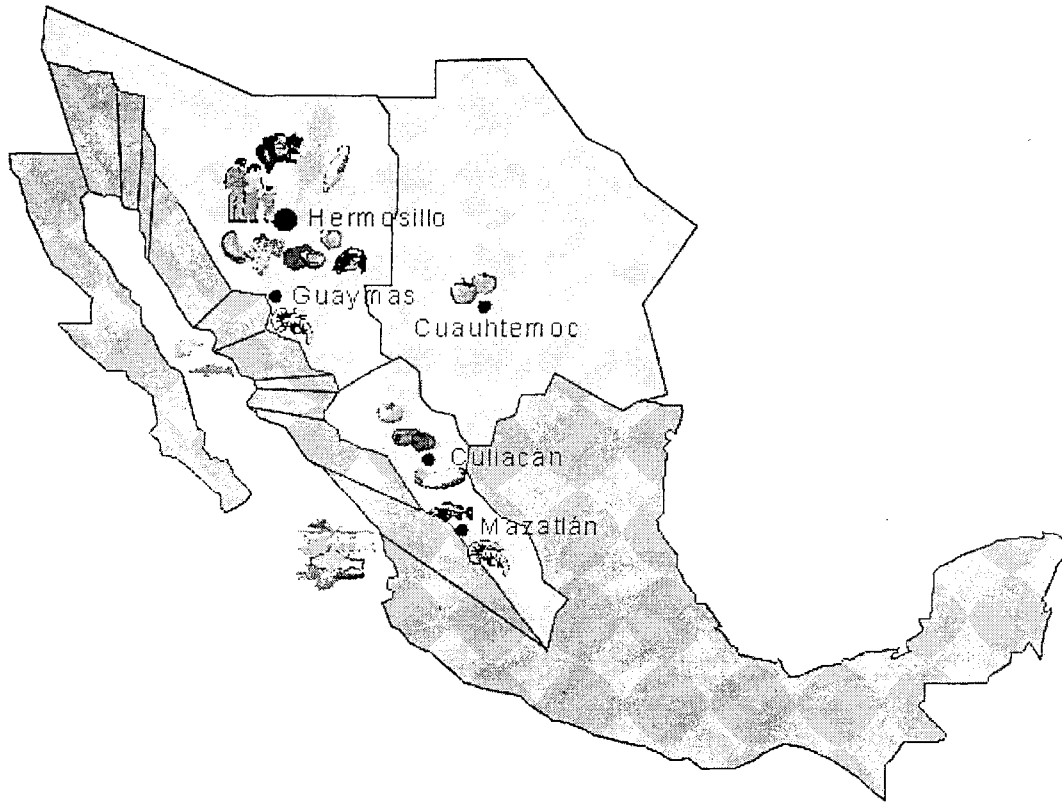


Figura 6.1 El CIAD, A.C. y sus unidades foráneas

El CIAD en la actualidad, tiene aproximadamente 303 empleados, de los cuales 99 son investigadores, 135 técnicos académicos y 69 personal administrativo, ver figura 6.2.

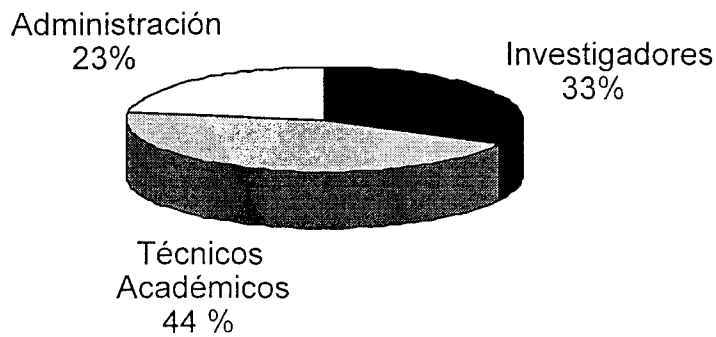


Figura 6.2 Distribución del personal del CIAD, A.C.

Historia de los sistemas en el CIAD

El CIAD como anteriormente se mencionó, inició siendo una institución bastante pequeña y los primeros años, la mayor parte del manejo de información se hacía de forma manual, las razones eran varias: una era el volúmen de información tan pequeño que se manejaba, otra, el presupuesto inicial era también reducido y se destinaba principalmente al equipamiento de los laboratorios y recordando que a principios de los años 80's, no eran tan fácil la adquisición de equipos personales de cómputo, por que apenas iniciaba la era de las computadoras personales.

A finales de 1983, se adquirió el primer equipo de cómputo del CIAD, una máquina TRS-80 de Radio Shack, en la cual se crearon los primeros programas del Centro.

Al año siguiente se adquirieron 2 Computadoras personales, los principales usos eran, programas estadísticos y de bases de datos, así como programación en BASIC, para ciertas labores administrativas.

Los primeros programas que se hicieron, estaban limitados principalmente por las capacidades de los equipos, y el volumen de información que se manejaba, era muy pequeño y hacia que cualquier problema que pudiera surgir debido a la desintegración de los datos, fuera casi imperceptible. Aunado a lo anterior, en esas épocas, se podían tener todos los sistemas administrativos, en una sola máquina, por lo tanto, los problema de desintegración, derivaban

principalmente de fallas en el diseño. También, hay que recordar, que en esas épocas no se ponía mucha atención a los problemas de la desintegración de los datos, eran los inicios de la computación, a un nivel accesible para las empresas pequeñas en México.

En el cuadro 4.2, se analizan los problemas relacionados con la desintegración de los datos, en empresas que manejan volúmenes de información pequeños, como fué el caso de los inicios del CIAD. En esta tabla, se puede ver que el impacto en relación con la desintegración de los datos es pequeña.

A mediados de los años 80's, la cantidad de equipos del Centro empezó a crecer, así como también el área administrativa, se empezó a especializar. Al principio, dos o tres personas se encargaban de todo el proceso administrativo, como la contabilidad, presupuestos, recursos humanos, almacén, etc. Estas personas, conformaban toda la administración, pero a medida que el personal empezó a crecer, se tuvieron que crear departamentos especializados, para manejar estas funciones administrativas, de forma exclusiva.

Estos departamentos, con el tiempo también empezaron a crecer, así como sus volúmenes de información. A finales de los años 80's, cada departamento contaba con su propio sistema, el cual resolvía sus necesidades en cuanto al manejo de su información interna.

Como se puede ver, esto propició, que a través del Centro se tuvieran una serie de islas de información. En aquellos tiempos, las solicitudes de información por parte del órgano rector del CIAD, no significaban una presión muy fuerte, y dichas solicitudes, por lo general, se integraban por una persona, que interpretaba, mezclaba y resumía la información proveniente de los diferentes departamentos, para obtener el informe requerido. Este proceso, se podía realizar en uno o dos días y como las exigencias no eran tan fuertes, se podía decir, que se cumplía dentro del rango de tiempo y calidad requerida.

A partir de 1993, el CIAD tuvo un crecimiento muy fuerte, principalmente debido a la creación de la primer unidad foránea. Esta unidad, empezó con un grupo de investigadores que en los primeros años no representaron una carga

muy pesada para ser absorbida por la administración desde Hermosillo, pero esta nueva unidad marcó el inicio de un crecimiento muy acelerado de la institución, un crecimiento que no estaba planeado y que se fue dando a petición de diferentes instancias. A los dos años de haberse creado la Unidad Mazatlán, la Unión de Productores de Hortalizas del vecino estado de Sinaloa apoyada por el Gobernador del Estado, solicitó al CIAD la creación de una unidad en Culiacán, con el fin de poder hacer frente a una serie de problemas que estaban padeciendo los productores de hortalizas de dicho estado, los cuales, representaban una serie de pérdidas en producción o cierre de mercados debido a la mala calidad de los mismos, y así surgió la Unidad Culiacán. El mismo año, se creó en Guaymas, Son. la unidad Guaymas, la cual tiene como principal labor, los estudios e inspección de plantas empacadoras de productos marinos. Finalmente, hace dos años se creó la Unidad Chihuahua, también a petición de los productores de manzana del Estado de Chihuahua. Toda esta evolución del Centro, así como también la evolución de sus programas de Maestría iniciado en 1983 y su programa de Doctorado iniciado en 1996, han traído un crecimiento tanto de personal, necesidades de información, y relaciones con diferentes instituciones del sector público y privado.

El sistema actual

El sistema administrativo que se utiliza actualmente en el CIAD, es el mismo que se ha usado desde finales de los 80's, este sistema, se desarrollo para ambientes monousuario. Los motivos por los cuales se han mantenido dichos sistemas, son varios, entre ellos, se puede mencionar que los sistemas habían cumplido con las necesidades de información de las áreas, y el personal que los desarrolló, hace las modificaciones que se le solicitan. Los problemas empezaron cuando dicho personal se vió involucrado en otras actividades y solo se dedicó a darle mantenimiento del sistema, sin tener tiempo para diseñar el mismo, con un manejador de bases de datos más moderno y que ofreciera la oportunidad de utilizar las ventajas de la red interna.

Como se mencionó anteriormente, el sistema que se utiliza actualmente cumplía con los requerimientos de información hasta hace relativamente poco tiempo, pero una serie de factores, como la desintegración de la información, el crecimiento acelerado de los últimos años (manejo de un mayor volumen de información con el mismo personal), las exigencias mayores por parte del órgano de gobierno del CIAD, la relación creciente con el sector productivo, las necesidades de información interdepartamental para lograr una administración adecuada de los recursos, así como las necesidades de las nuevas áreas del Centro empezaron a deteriorar el desempeño de los sistemas, requiriendo cada vez más, el apoyo de mayor número de personas y la utilización de herramientas alternas como hojas de cálculo, etc. (ver figura 6.3.)

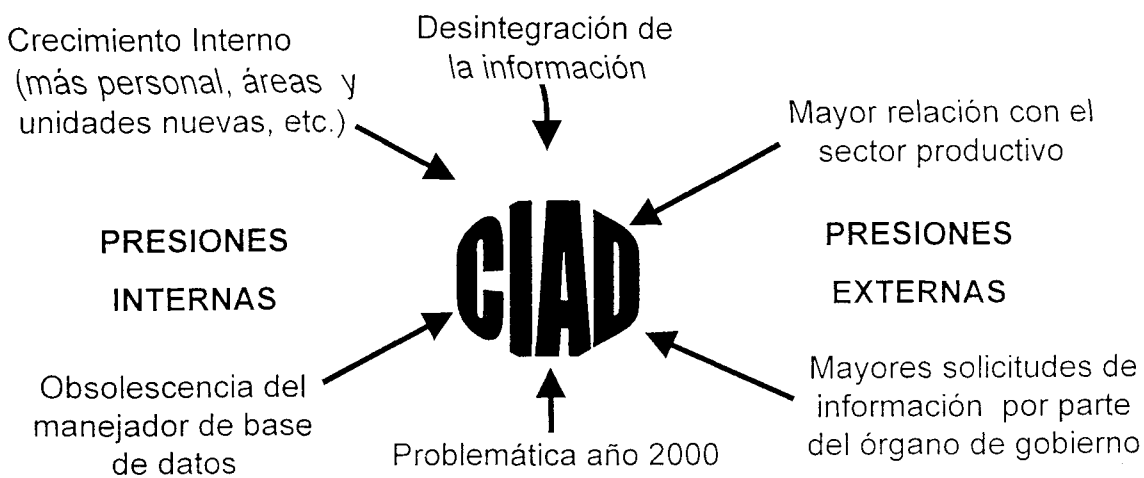


Figura 6.3 Presiones que enfrenta el sistema actual del CIAD.

**Planteamiento del problema del CIAD ocasionado
por la desintegración de los datos**

Como se puede ver, el sistema se encuentra en crisis y una gran parte de este problema lo representa la desintegración de la información.

En el área administrativa del CIAD, existen 4 áreas principales, las cuales manejan un alto porcentaje del volumen de información interno del CIAD, estas

cuatro áreas son: adquisiciones, presupuestos, contabilidad y almacén. En estas áreas, se tienen sistemas muy independientes entre sí, aunque existe una gran relación entre dichas áreas, lo que sugeriría según Goodhue² (1992), que se buscara la integración de los datos a gran escala.

Los problemas relacionados con la desintegración son varios y entre ellos podemos mencionar: inconsistencias entre los reportes que presentan las áreas por separado, la conciliación entre el área de presupuestos, compras y contabilidad es un trabajo bastante laborioso y que consume mucho tiempo, el departamento de presupuestos no puede ofrecer al investigador, ni a los altos ejecutivos, información que les permita tomar decisiones adecuadas respecto a las adquisiciones, por que los presupuestos no están al día y por lo tanto su información no es confiable, etc.

Como se puede ver, la desintegración de la información es un freno para el desarrollo institucional, ya que entorpece el proceso de toma de decisiones y ocasiona que éstas, no estén bien fundamentadas, generando una serie de problemas.

Proceso de adquisiciones

En el CIAD, cada investigador tiene uno o varios proyectos, cada uno de los cuales, cuenta con su presupuesto anual para su desarrollo, estos presupuestos, pueden estar apoyados por diferentes fuentes de financiamiento, como son: proyectos con el sector productivo, con universidades, centros de investigación, proyectos apoyados por CONACyT o SEP, etc. A cada presupuesto, se le asignan cantidades para la adquisición de diferentes productos o servicios, los cuales están clasificados en partidas presupuestales, como se pueden ver en el cuadro 6.1.

El investigador para adquirir algún bien o servicio necesita llenar una requisición y entregarla al departamento de compras. Este a su vez cotiza los productos solicitados y los envía al departamento de presupuestos para aprobación.

Cuadro 6.1 Muestra de Algunas Partidas, del Catálogo Presupuestal del CIAD.

| Capítulo 2000 Materiales y Suministros | | |
|---|---|---|
| 2100 | Materiales y Utiles de Administración | |
| | 2101 | Materiales y Utiles de Oficina |
| | 2102 | Material de Limpieza |
| | 2103 | Material Didáctico y de Apoyo Informativo |
| | 2104 | Material Estadístico y Geográfico |
| | 2105 | Materiales y Utiles de Impresión y Reproducción |
| | 2106 | Materiales y Utiles para el Procesamiento en Equipos Informáticos |
| 2200 | Alimentos y Utensilios | |
| | 2201 | Alimentación de Personas |
| | 2203 | Utensilios para el Servicio de Alimentación |
| 2400 | Materias Primas y Materiales de Producción | |
| | 2401 | materiales de Construcción |
| | 2402 | Estructuras y Manufacturas |
| | 2403 | Materiales Complementarios |
| | 2404 | Material Eléctrico |
| 2500 | Productos Químicos, Farmacéuticos y de Laboratorio | |
| 2600 | Combustibles, Lubricantes y Aditivos | |
| 2700 | Vestuarios, Blancos, Prendas de Protección y Art. Deportivos | |
| | | |
| Capítulo 3000 Servicios Generales | | |
| | | |
| 3100 | Servicios Básicos | |
| | 3101 | Servicio Postal |
| 3200 | Servicios de Arrendamiento | |
| | 3201 | Arrendamiento de Edificios y Locales |
| 3300 | Servicios de Asesoría, Informáticos, Estudios e Investigaciones | |
| | 3301 | Asesoría |
| | 3302 | Capacitación |
| | 3303 | Servicios de Informática |
| 3400 | Servicios Comercial y Bancario | |
| 3500 | Servicios de Mantenimiento, Conservación e Instalación | |
| 3600 | Servicios de Difusión e Información | |
| 3700 | Servicios de Traslado e Instalación | |
| 3800 | Servicios Oficiales | |

El departamento de presupuestos verifica que exista dinero en el presupuesto, de ser así, compromete la cantidad solicitada. Al ser aprobados por el departamento de presupuestos, éste regresa la autorización al departamento de

compras, quien genera una orden de compra y la envia al proveedor, el proveedor entrega los artículos al almacen, éste verifica que el proveedor este entregando la cantidad y el artículo solicitado y genera la autorización para que el proveedor entregue la factura al departamento de contabilidad. El departamento de contabilidad hace los asientos contables correspondientes y genera el cheque al proveedor.

Después el departamento de presupuestos verifica el pago hecho por contabilidad y cambia en el presupuesto del investigador, el concepto de comprometido a ejercido. El investigador recibe el producto después de que el almacen recibe la mercancía. La figura 6.4 muestra el flujo de información en el proceso de adquisiciones.

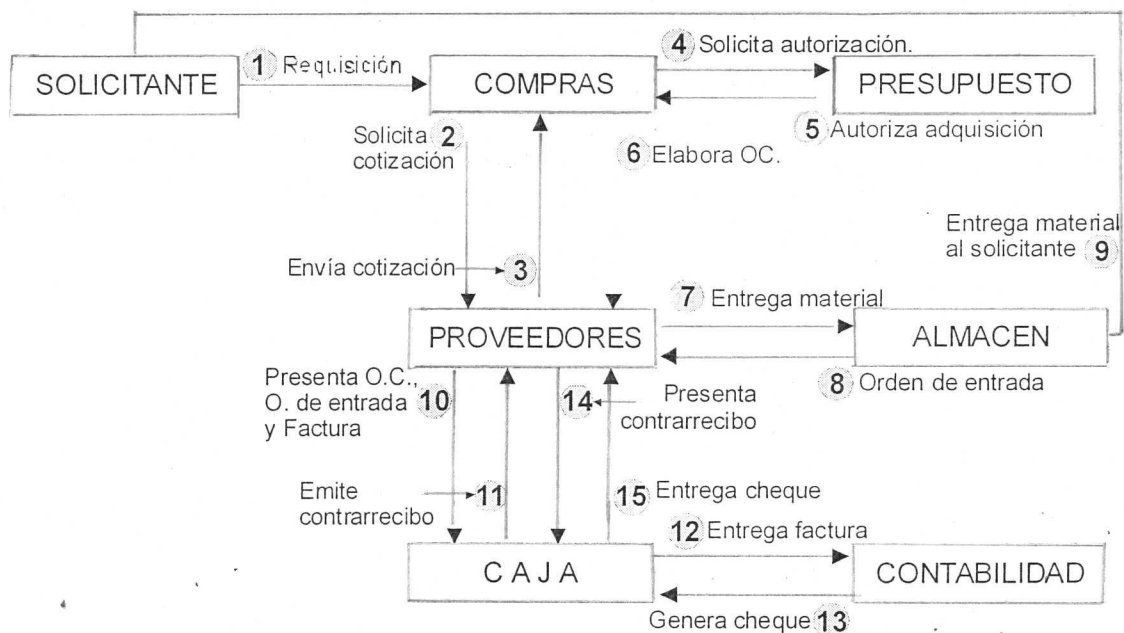


Figura 6.4 Diagrama de Flujo del Proceso de Adquisiciones en el CIAD.

Este proceso, aparentemente sencillo, tiene un alto grado de complejidad, debido a diferentes detalles que ocurren dentro del proceso, como son: compras de artículos en moneda extranjera, tiempos de entrega, entregas parciales de mercancías, facturación con precios o productos diferentes a los solicitados, devoluciones, pagos anticipados, cancelaciones, cierres presupuestales y contables, etc.

Caso Práctico. ¿Cómo puede ayudar la tecnología de intranet a resolver el problema de desintegración de los datos?

Viendo el problema de desintegración de datos tan grande que afecta al Centro, existen una gran cantidad de áreas, que pueden verse beneficiadas por un trabajo de integración de datos. Pero es necesario, seguir una serie de pasos antes de pensar en la utilización de la tecnología, que en este trabajo, esta representada por la tecnología de intranet.

El primer paso, hacia la integración de la información, consiste en seleccionar la estrategia a seguir, si va a ser una acción arriba-abajo, basada en modelos de datos que soporten la estrategia de la empresa o si se partirá de la información y los sistemas actuales, para buscar la integración aparente, acción "abajo - arriba".

Esta última opción, la cual aprovecha gran parte de la inversión en software de la empresa, tiene la ventaja de obtener frutos a corto, mediano y largo plazo. En este proceso, se parte de la situación actual de la empresa, "el es de la empresa" y trata de lograr una integración paulatina de los datos, utilizando una serie de estrategias para lograrlo.

Debido a la naturaleza de este trabajo, se seleccionó esta última opción, ya que se pretende utilizar parte de la metodología del almacén de datos, la cual selecciona datos del ambiente operativo y ofrecer información para la toma de decisiones de los ejecutivos de la empresa.

El segundo paso, consiste en fijar el alcance de integración deseado, hasta donde y entre que áreas, se requiere tener un flujo efectivo y eficiente de información. Hay que tener en cuenta, que si se decide integrar datos en solo una parte de la empresa, los beneficios se obtendrán, solo en esta área, pero si a futuro, se pretende incluir otras áreas, será necesario volver a realizar un estudio de las áreas involucradas y los datos asociados. Posiblemente, habrá áreas que no se incluyan en este trabajo, debido a su independencia de las otras, ya que es

preciso recordar, que un esfuerzo de integración empresarial, posiblemente le quite flexibilidad a ciertas áreas, para poder enfrentar sus retos particulares.

El alcance de la integración deseado, parte fundamentalmente de las necesidades de información de los usuarios.

Es muy importante, tener en cuenta que, tanto el factor económico, como el tiempo disponible, marcan la pauta con respecto al camino que se seguirá. Para este caso en particular, el alcance se fijó entre los siguientes departamentos administrativos: compras, presupuestos, almacén y contabilidad y para efectos de ejemplificar el caso, se tomará como base el proceso de adquisiciones, como se pudo ver en la figura 6.4.

En resumen, este trabajo se basará en la metodología del almacén de datos, tomando la información del ambiente operativo, por tanto, se considera una acción abajo-arriba de integración aparente y el alcance comprenderá a los departamentos anteriormente descritos.

En primera instancia, se analizarán los requerimientos de información de los usuarios y se hará un estudio de los datos en el ambiente operativo, para ver si éstos, pueden cubrir las necesidades de los usuarios y después, se hará el proceso de integración de los mismos, como se muestra en la figura 6.5.

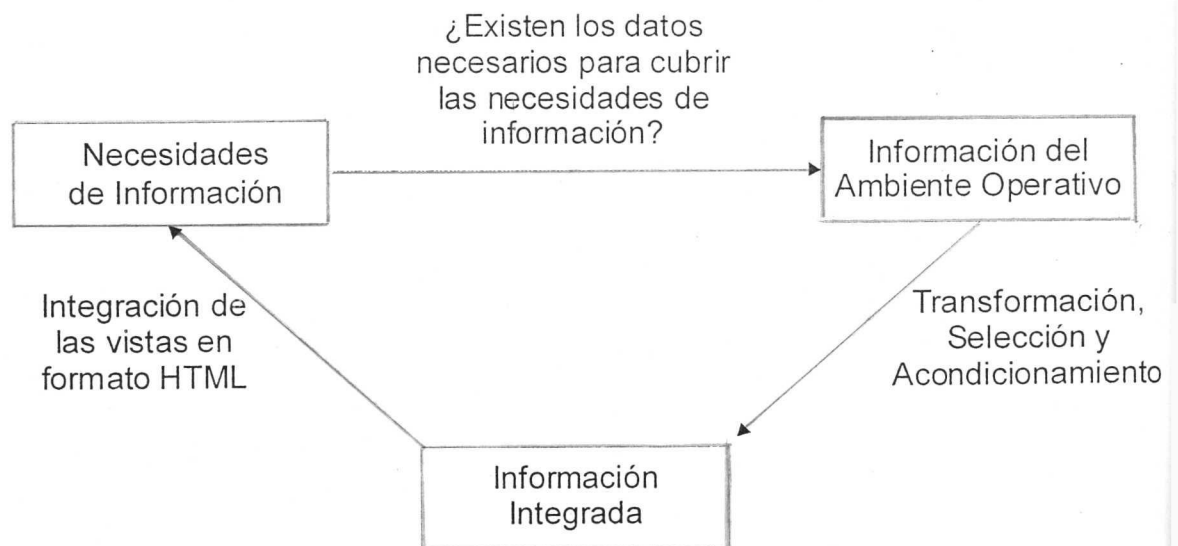


Figura 6.5 Fuente de Información y transformaciones para cubrir las necesidades de los tomadores de decisiones.

Requerimientos de Información

SOLICITANTE

El solicitante necesita saber, cuantas requisiciones ha hecho en el año, así como, en que etapa del proceso de adquisiciones, se encuentran sus pedidos, como se puede ver en el cuadro 6.2.

Cuadro 6.2 Situación de las requisiciones solicitadas por el usuario.

| No. Req | FECHAS | | | | | |
|---------|-------------------|-----------------------|---------------------|----------------------|-----------------|-------------------|
| | Recepción Compras | Envío para Cotización | Recepción art. cot. | Autorizac. Presupues | Orden de Compra | Recepción Almacén |
| 1421 | 12-Ene-99 | 13-Ene-99 | 21-Ene-99 | 21-Ene-99 | 22-Ene-99 | 04-Mar-99 |
| 1649 | 15-Ene-99 | 18-Ene-99 | 22-Ene-99 | 25-Ene-99 | *** | *** |
| 1785 | 21-Ene-99 | 22-Ene-99 | 04-Feb-99 | 08-Feb-99 | *** | *** |
| 1789 | 15-Ene-99 | 18-Ene-99 | *** | *** | *** | *** |
| 1823 | 15-Feb-99 | 15-Feb-99 | 22-Feb-99 | 23-Feb-99 | 23-Feb-99 | 25-Feb-99 |
| 1849 | 23-Feb-99 | 24-Feb-99 | 02-Mar-99 | *** | 03-Mar-99 | 03-Mar-99 |
| 1959 | 29-Mar-99 | 30-Mar-99 | *** | *** | *** | *** |
| 1960 | 31-Mar-99 | 31-Mar-99 | *** | *** | *** | *** |

Otro requerimiento de información muy importante para el solicitante, consiste en conocer la situación real de su presupuesto. Entre los problemas actuales, podemos mostrar el siguiente ejemplo: imagine que el investigador solicita su presupuesto, y le informan que tiene disponible \$ 50,000.00, entonces, basándose en dicha información, hace una requisición por un equipo, pensando que cuenta con dicha cantidad. Probablemente, a los días le comuniquen, que su proyecto no cuenta con los fondos suficientes para efectuar dicha adquisición. Existen varias fuentes de error, pero la más común, es que haya efectuado algunas requisiciones, las cuales no estaban cotizadas en el momento que solicitó su presupuesto y que, posteriormente fueron enviadas a presupuestos, para su aprobación, disminuyendo por consiguiente, su saldo disponible.

Entonces, el investigador requiere que la información de su presupuesto, contenga una listado de las requisiciones que están en proceso, ya sea sin cotizar o sin haberse autorizado, como se puede ver en el cuadro 6.3.

Cuadro 6.3 Situación del Presupuesto con Información de requisiciones en tránsito.

6215 COMPUTO
Fuente de financiamiento: 1
Reporte al 11/10/96

| PARTIDAS | Autorizado | Ejercido | Comprom | Remanente |
|---------------------------------|---------------|---------------|--------------|---------------|
| 2101 Mat. y Utiles de Oficina | 3,500 | 225 | 1,500 | 1,775 |
| 2102 Mat. de Limpieza | 6,000 | 0 | 3,000 | 3,000 |
| 2103 Mat. Didáctico y de Apoyo | 7,000 | 1,750 | 500 | 4,750 |
| 2104 Mat. Estadíst. y Geográf | 1,500 | 300 | 100 | 1,100 |
| 2105 Mat. y Utiles de Impresión | 3,500 | 3,000 | 100 | 400 |
| 2106 Mat. y Utiles para el Pr | 1,000 | 500 | 0 | 500 |
| 2404 Mat. Eléctrico | 250 | 100 | 0 | 150 |
| 2601 Comb., Lubric. y Aditivos | 500 | 400 | 100 | 0 |
| 2701 Vestuarios, Blancos, Pren | 750 | 100 | 0 | 650 |
| 2000 Mat. y Suministros Total | 24,000 | 6,375 | 5,300 | 12,325 |
| 3101 Servicio Postal | 750 | 500 | 0 | 250 |
| 3301 Asesoría | 5,000 | 0 | 0 | 5,000 |
| 3302 Capacitación | 7,500 | 1,200 | 1,000 | 5,300 |
| 3303 Servicios de Informática | 5,000 | 3,500 | 0 | 1,500 |
| 3706 Viáticos en el Extranjero | 8,000 | 5,000 | 1,000 | 2,000 |
| 3803 Congresos, convenciones | 3,000 | 1,200 | 1,500 | 300 |
| 3000 Serv. Generales Total | 29,250 | 11,400 | 3,500 | 14,350 |
| | | | | |
| T O T A L E S | 53,250 | 17,775 | 8,800 | 26,675 |

| | |
|---|--------|
| Remanente Total del Proyecto | 26,675 |
| # Requisiciones en proceso, (sin cotizar) | |
| | 1789 |
| | 1959 |
| | 1960 |
| # Requisiciones en proceso, (sin autorizar) | |
| | 1849 |

Compras

El departamento de compras, requiere saber, cuales proveedores estan cumpliendo con sus tiempos de entrega y si estan entregando el total de las mercancías solicitadas. También, necesita saber, cuales órdenes de compra, no han sido surtidas en un período razonable de tiempo, para decidir, si cancela el pedido con dicho proveedor, lo adquiere con otro o simplemente lo cancela.

También, necesita enterarse, si el Centro esta cumpliendo con los compromisos de pago. Como se puede ver, en el cuadro 6.4 se muestran las necesidades propias de este departamento.

Cuadro 6.4 Situación de los proveedores: tiempos de entrega y pago.

| Proveedor: Compuprovedores, S.A. | | | | | | |
|----------------------------------|-----------|-----------|----------------|------------------|-----------|-------------|
| Orden de Compra | Solicitud | Recepción | tiempo entrega | Entrega Completa | Pago | tiempo pago |
| 3 | 23-Ene-99 | 31-Ene-99 | 8 | si | 07-Feb-99 | 7 |
| 8 | 25-Ene-99 | 06-Feb-99 | 12 | si | 21-Feb-99 | 15 |
| 15 | 27-Ene-99 | 05-Feb-99 | 9 | si | 19-Feb-99 | 14 |
| 45 | 06-Feb-99 | 09-Feb-99 | 3 | si | 15-Feb-99 | 6 |
| 79 | 15-Feb-99 | 27-Feb-99 | 12 | No | 08-Mar-99 | 9 |
| Promedio en días | | | 9 | | | 10 |

Presupuestos

La necesidad principal del departamento de presupuestos, consiste en saber exactamente, cuando se efectúan los pagos de las órdenes de compras, las cuales, están ligadas a las requisiciones que previamente aprobó, esto, con el fin de grabar la cantidad exacta de la compra y hacer los ajustes pertinentes a los presupuestos, originado por, diferencias entre lo presupuestado y lo pagado. Hay que tener en cuenta que, el departamento de presupuestos, compromete del presupuesto del investigador una cantidad X, para cada requisición, la cual, puede

tener variaciones en precios, en el proceso de compra (desde que se ordena, hasta que se recibe la factura), ocasionados por, variaciones en el tipo de cambio, mercado inestable, etc. El cuadro 6.5, muestra las principales necesidades, en forma de reporte, de este departamento.

Cuadro 6.5 Reporte de pagos efectuados por contabilidad.

| Orden de Compra | Pago Efectuado | Fecha | Proveedor |
|-----------------|----------------|------------|----------------------|
| 1 | 2,000.00 | 5/Ene/1999 | Compuproveedores |
| 2 | 1,500.75 | 5/Ene/1999 | Accesorios PC |
| 4 | 3,750.35 | 5/Ene/1999 | Muebles de Oficina |
| 5 | 500.00 | 5/Ene/1999 | Papeleria Hermosillo |
| 6 | 635.60 | 6/Ene/1999 | PC Siglo XXI |
| 8 | 1,895.95 | 6/Ene/1999 | Imprenta ROSA |
| 9 | 100.00 | 6/Ene/1999 | PC Refacciones |
| 10 | 1,632.25 | 6/Ene/1999 | Reactivos HPLC |
| 11 | 8,500.00 | 6/Ene/1999 | INGL-Gases |
| 13 | 300.00 | 7/Ene/1999 | Plomeria La Carreta |
| 15 | 175.50 | 7/Ene/1999 | Ductos del Noroeste |
| 20 | 45.00 | 7/Ene/1999 | Agua Purif. |
| 35 | 1,000.00 | 7/Ene/1999 | CIPSA |

También, otro requerimiento de información, consiste en saber cuando una orden de compra se ha cancelado y la requisición o requisiciones asociadas a ella, como se puede ver en en el cuadro 6.6. Esta necesidad, es con el fin, de poner nuevamente a disposición del investigador, el dinero que estaba comprometido y que finalmente no se usó.

Almacen

El almacen, necesita saber exactamente, las cantidades y descripciones de los artículos que va a recibir (este aspecto si se maneja en el sistema actual) y la orden de compra que lo esta solicitando, como se puede apreciar en el cuadro

6.7. Esta información, es una réplica de la que el departamento de compras genera, cuando hace una orden de compra.

Cuadro 6.6 Requisiciones asociadas a las ordenes de compra y su estatus

| Orden de Compra | REQUISICIONES ASOCIADAS | | | | Proveed | Estatus |
|-----------------|-------------------------|----|----|----|---------|-----------|
| | 1 | 7 | 8 | 12 | | |
| 1 | 1 | 7 | 8 | 12 | 45 | |
| 2 | 3 | 6 | | | 9 | |
| 3 | 2 | 5 | | | 19 | |
| 4 | 4 | | | | 32 | Cancelada |
| 5 | 9 | 10 | 15 | 18 | 23 | |
| 6 | 11 | 14 | 17 | | 5 | |
| 7 | 13 | 16 | | | 26 | |
| 8 | 19 | 20 | 35 | | 2 | |
| 9 | 21 | | | | 8 | |
| 10 | 36 | | | | 9 | |
| 11 | 22 | 23 | 28 | | 2 | Cancelada |

Cuadro 6.7 Descripción del contenido de la orden de compra.

Orden de Compra # 8
Proveedor # 2 Papelería LOS ROSALES

| Cant | Unidad | Descripción | Precio Unit. | Importe |
|------|---------|------------------------------------|--------------|----------|
| 1 | Pieza | Cuaderno 100 Hojas Tamaño Carta | 11.95 | 11.95 |
| 5 | Pieza | Lapiz mirado # 3 | 2.20 | 11.00 |
| 10 | Pieza | Pluma bic punto mediano color azul | 3.80 | 38.00 |
| 2 | Paquete | Folder tamaño carta | 43.45 | 86.90 |
| 5 | Paquete | HP Glossy Paper 50 hojas | 300.00 | 1,500.00 |
| | | | SUB-TOTAL | 1,647.85 |
| | | | IVA | 247.18 |
| | | | TOTAL | 1,895.03 |

Contabilidad

Contabilidad, requiere contar con un instrumento, que efectúe una conciliación, para verificar que coincidan los costos que fueron cotizados, contra los precios presentados en la factura y así, poder asegurar que los pagos, están

de acuerdo a lo pactado anteriormente, por el departamento de compras. También, necesita un reporte de las cantidades recibidas en almacén, como se muestra en el cuadro 6.8 para cotejar con la factura.

Cuadro 6.8 Reporte de artículos recibidos en almacén y precios establecidos por compras para comparación con factura.

Orden de Compra # 8
Proveedor # 2 Papelería LOS ROSALES

| Cantidad | | Unidad | Descripción | Precio Unit. | Importe |
|----------|-------|---------|------------------------------------|------------------|-----------------|
| Solicit | Recib | | | | |
| 1 | 1 | Pieza | Cuaderno 100 Hojas Tamaño Carta | 11.95 | 11.95 |
| 5 | 5 | Pieza | Lapiz mirado # 3 | 2.20 | 11.00 |
| 10 | 10 | Pieza | Pluma bic punto mediano color azul | 3.80 | 38.00 |
| 2 | 2 | Paquete | Folder tamaño carta | 43.45 | 86.90 |
| 5 | 5 | Paquete | HP Glossy Paper 50 hojas | 300.00 | 1,500.00 |
| | | | | SUB-TOTAL | 1,647.85 |
| | | | | IVA | 247.18 |
| | | | | TOTAL | 1,895.03 |

Directivos del Centro

Los altos ejecutivos de la institución, también tienen una serie de requerimientos de información. Por lo regular, requieren información resumida a nivel dirección o institución, principalmente de los departamentos de presupuestos y contabilidad. Esta información, como se mencionó anteriormente, es solicitada constantemente, para llenar formatos y encuestas del órgano de gobierno, o para conocer la forma en la cual se está ejerciendo el presupuesto, etc. El cuadro 6.9 muestra una de las muchas necesidades de información de los directivos de la institución.

Cuadro 6.9 Requerimientos de información de los altos ejecutivos (con datos falsos, debido a la naturaleza de la información presentada)

SITUACION DEL PRESUPUESTO INSTITUCIONAL
FUENTE DE FINANCIAMIENTO: 1
REPORTE AL 10/NOV/1997

| PARTIDAS | Autoriz. | Ejercido | Comprom | Reman. |
|--------------------------------------|-----------|-----------|-----------|-----------|
| 2101 Materiales y Utiles de Oficina | 437,500 | 245,311 | 187,500 | 4,689 |
| 2102 Material de Limpieza | 750,000 | 325,012 | 375,000 | 49,988 |
| 2103 Material Didáctico y de Apoyo | 875,000 | 623,210 | 62,500 | 189,290 |
| 2104 Material Estadístico y Geográf. | 187,500 | 37,500 | 12,500 | 137,500 |
| 2105 Materiales y Utiles de Impresi | 437,500 | 375,000 | 55,332 | 7,168 |
| 2106 Materiales y Utiles para el Pr | 125,000 | 78,632 | 15,275 | 31,093 |
| 2404 Material Eléctrico | 31,250 | 12,500 | 12,123 | 6,627 |
| 2601 Combustibles, Lubric. y Adit. | 62,500 | 48,225 | 12,500 | 1,775 |
| 2701 Vestuarios, Blancos, Prendas | 93,750 | 12,500 | 1,500 | 79,750 |
| | | | | |
| 2000 Materiales y Suministros Total | 3,000,000 | 1,757,890 | 734,230 | 507,880 |
| | | | | |
| 3101 Servicio Postal | 93,750 | 62,500 | 22,372 | 8,878 |
| 3301 Asesoría | 625,000 | 422,388 | 183,246 | 19,366 |
| 3302 Capacitación | 937,500 | 478,190 | 125,000 | 334,310 |
| 3303 Servicios de Informática | 625,000 | 437,500 | 119,983 | 67,517 |
| 3706 Viáticos en el Extranjero | 1,000,000 | 825,446 | 125,000 | 49,554 |
| 3803 Congresos, convenciones | 375,000 | 169,441 | 187,500 | 18,059 |
| | | | | |
| 3000 Servicios Generales Total | 3,656,250 | 2,395,465 | 763,101 | 497,684 |
| | | | | |
| T O T A L E S | 6,656,250 | 4,153,355 | 1,497,331 | 1,005,564 |

| | |
|-----------------|-----------|
| Remanente Total | 1,005,564 |
|-----------------|-----------|

| | |
|--|------|
| # Requisiciones en proceso (sin cotizar) | |
| | 1789 |
| | 1960 |
| | : |
| # Requisiciones en proceso (sin autorizar) | |
| | 1849 |
| | 2016 |
| | : |

Fuentes de Información

Una vez analizados los requerimientos de información de las diferentes áreas, el siguiente paso, consiste en identificar las fuentes de datos en el ambiente operativo, de donde se obtendrá la información requerida.

Las fuentes de datos en el ambiente operativo, provienen de 2 orígenes distintos, en referencia al manejador de base de datos. La información que se maneja en el departamento de presupuestos y contabilidad, está en el formato del manejador de base de datos para computadoras personales, Knowledge Man para DOS, este formato, no permite que la información pueda accesarse directamente, desde otro paquete comercial, pero ofrece la posibilidad, de generar un archivo con información en formato ASCII. A pesar de que la información en estos departamentos, tiene el mismo formato y la gran interrelación que existe entre ellos, los sistemas, fueron desarrollados con el fin de satisfacer las necesidades particulares de cada departamento.

La información que utiliza el departamento de compras y el almacén, esta en formato del manejador de base de datos para computadoras personales, FoxPro2 para DOS. Entre estas dos áreas, existe un grado de integración y flujo de información, pero el problema, es que ésta información, no se puede compartir con la información de las otras áreas. La figura 6.6, muestra los orígenes y transformaciones que los datos experimentan desde el ambiente operativo, hasta que se convierten en información valiosa para los tomadores de decisiones.

Una vez identificadas las fuentes de información, las cuales tienen orígenes diferentes, es necesario convertirlas a un formato común, que permita la visualización de la información en el mismo lenguaje. El siguiente paso, consiste en seleccionar los datos, que sirven para responder a las necesidades de los usuarios. Después, se inicia el proceso de integración y estandarización de dichos datos, y se crean (siguiendo la metodología del almacén de datos) las tablas, desde las que se obtendrá directamente la información, las cuales, sufrirán una última transformación de formato, para que puedan ser consultadas por los

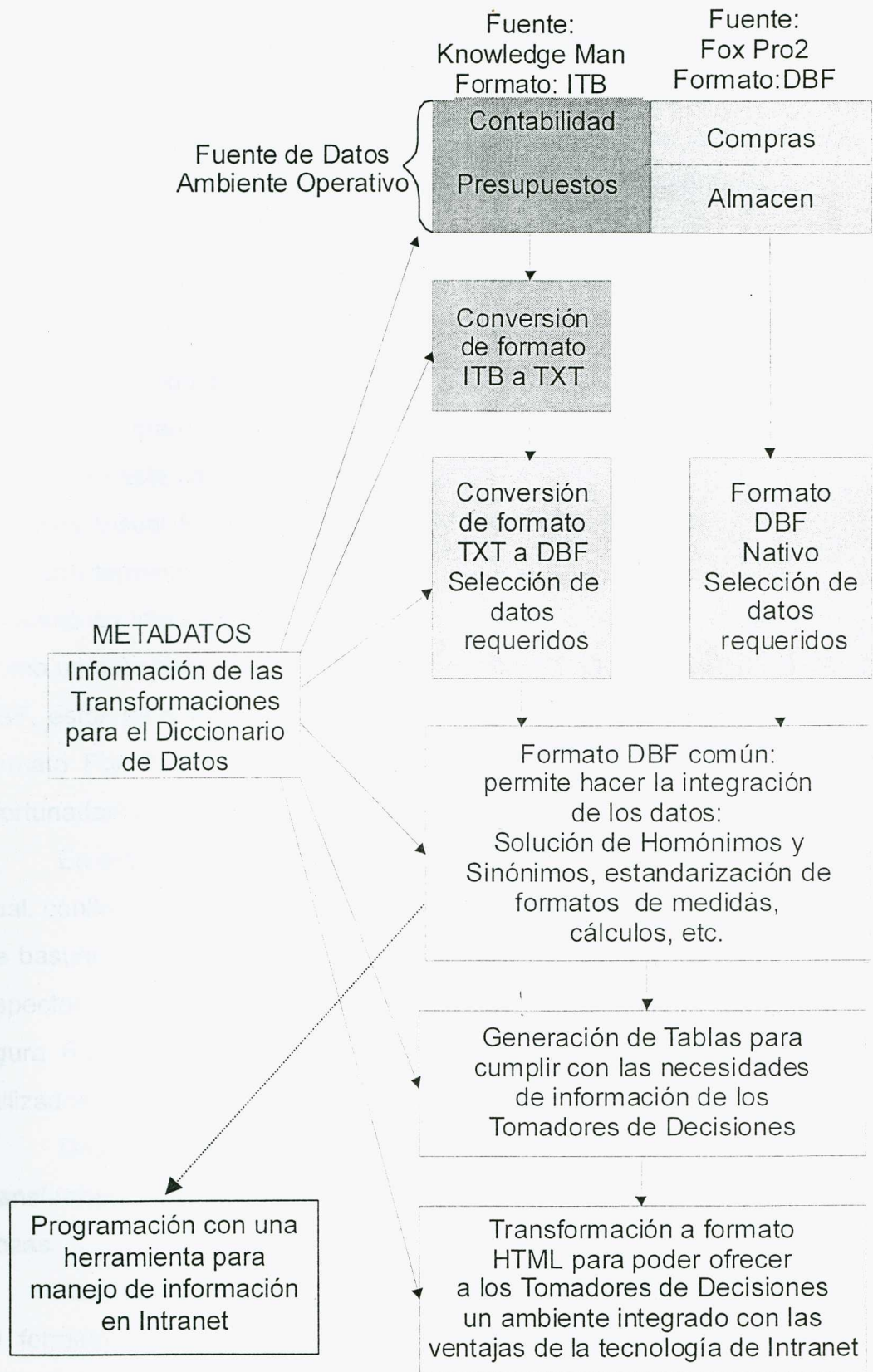


Figura 6.6 Flujo y Transformaciones de los Datos desde el Ambiente Operativo.

usuarios. Esta información, estará disponible para el usuario, a través de su navegador preferido, y para él, esta información estará integrada. Aunque en la realidad la integración de los datos será aparente, los beneficios para esta aplicación serán equivalentes a que si la información fuente estuviera totalmente integrada.

Una variante que se puede utilizar, es la que se aprecia en la figura 6.6 con la línea punteada. Aquí, se toman los datos directamente de las tablas con un formato adecuado para la extracción y el proceso de integración se hace por medio de programación, con la herramienta de intranet.

Para este caso en particular, se hicieron pruebas con el manejador de base de datos Visual FoxPro2. Primero, se tomaron las tablas del DBMS Knowledge Man con terminación ITB y se convirtieron a formato ASCII. El segundo paso, consistió en importar dichos archivos ASCII, a tablas de Visual FoxPro 2 (DBF). Como un dato técnico, al pasar los datos del formato ITB al ASCII, e importarlos a DBF, estos se importan como tipo caracter. También, las tablas que estaban en formato Fox Pro 2 para DOS, se importaron al formato de Visual FoxPro 2, afortunadamente, aquí, el formato se conserva.

En este ambiente común, se inició el proceso de integración de los datos, el cual, conlleva una serie de pasos entre los que se pueden mencionar: eliminación de basura y de datos innecesarios, solución de homónimos y sinónimos, y otros aspectos, como los que se ejemplificaron en la figura 2.3, entre otros. También, la figura 6.7 muestra un ejemplo de las tablas origen y destino de los datos utilizados.

Después de identificar los datos origen que se utilizarán y las transformaciones necesarias para lograr la integración, se pueden hacer dos cosas.

Una opción es, crear tablas nuevas, las cuales después de transformarse al formato HTML puedan responder directamente a las necesidades de los usuarios anteriormente planteadas. Esta opción, estaría muy relacionada con la

metodología del almacen de datos. Las ventajas y desventajas de esta opción, serían las mismas que las del almacen de datos.

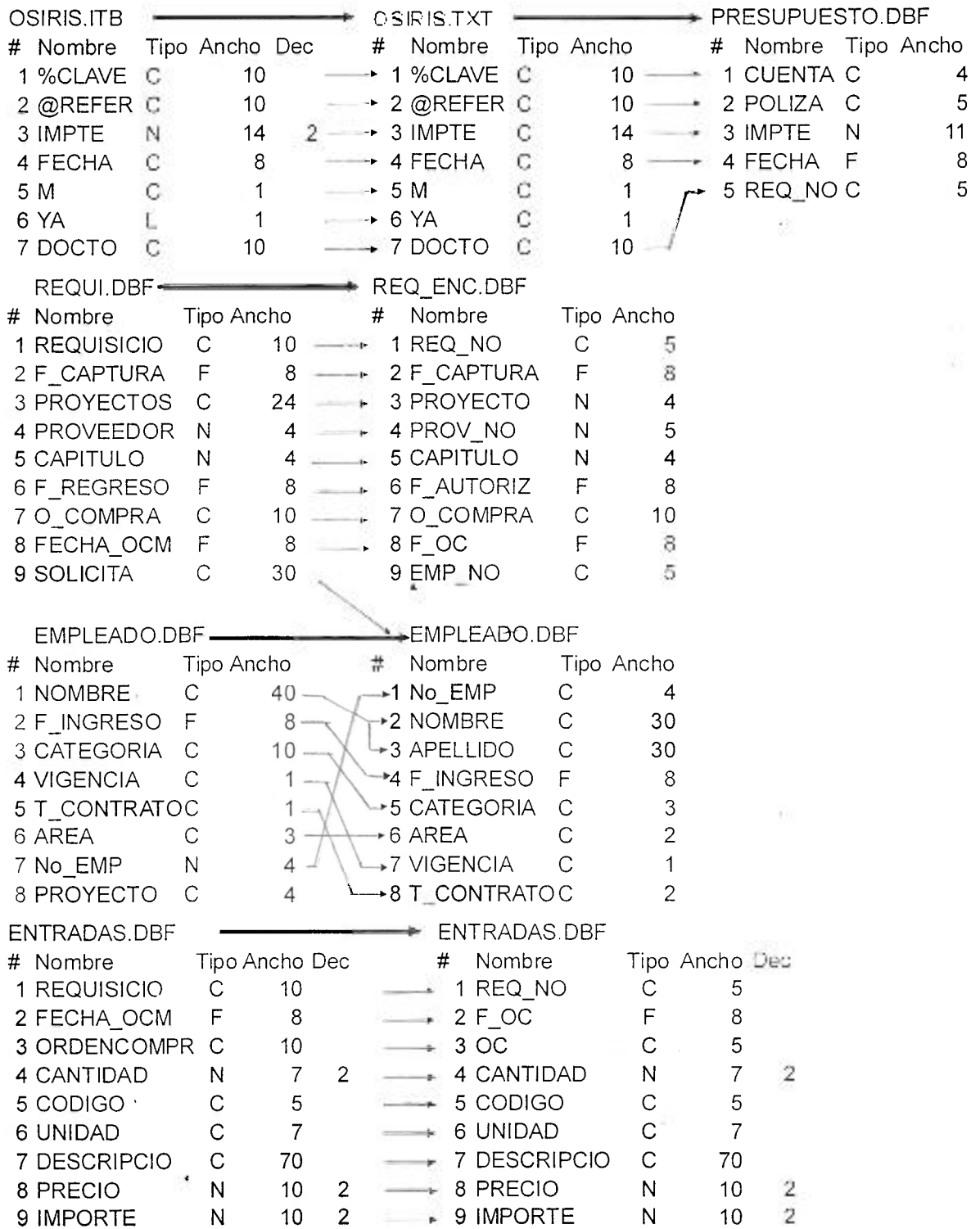


Figura 6.7 Ejemplo de algunos archivos fuente y sus transformaciones

Otra opción, consiste en elaborar programas con la tecnología de intranet y directamente, por medio de programación, realizar el proceso de integración, tomando los datos directamente del ambiente operativo. Esta opción se le puede llamar como el almacén de datos virtual, ya que no se crearían copias de los datos seleccionados sino que, se extraerían del ambiente operativo. Una de las limitantes de esta opción, es que por lo general, en el ambiente operativo no se trabaja con datos de más de un año y también, al meter una carga extra de trabajo, ésta puede afectar al desempeño del sistema.

Extracción de datos del ambiente operativo

La extracción de datos del ambiente operativo, es un proceso que requiere de atención especial. El primer paso, consiste en grabar en el nuevo sistema, toda la información histórica que se utilizará, después se introduce la información de presente ciclo y por último, se tiene que cargar la información diaria. Este último paso, que parece tan sencillo, puede convertirse en un grave problema si no se cuenta con una forma de saber, si ya se cargo cierta información o desde donde se debe iniciar la nueva carga de información.

Afortunadamente, la información en las diferentes tablas del ambiente operativo para este caso práctico, esta acompañada por un campo de fecha, el cual permite que la nueva adquisición de información, se haga con referencia a dicha fecha.

Seguridad

Para este caso no se plantea niveles de seguridad al momento de accesar la información, pero es necesario, de acuerdo con las políticas institucionales, crear un sistema de seguridad, que permita identificar al usuario y los derechos de acceso a la información.

Todo este proceso de mejorar la calidad de los datos, con el fin de convertirlos en una herramienta para la toma de decisiones, es un proceso muy caro y que involucra un alto nivel de esfuerzo.

La información, en estos momentos, debe considerarse como un elemento de gran valor. Como se dijo anteriormente, ahora la información está bien depurada, seleccionada, con un alto grado de calidad y en un formato que se puede considerar universal.

Pues bien, toda esta serie de atributos, también tienen su desventaja, debido a que puede ser un blanco fácil, para personas ajenas a la empresa. Más aún, si se decide utilizar internet como un medio de comunicación entre las diferentes unidades de la empresa, hay que estar preparados contra "Hackers", "Crakers" y virus, los cuales abundan en internet.

En pocas palabras, es necesario fijar altos niveles de seguridad, para evitar el robo o la pérdida de este recurso, ahora tan importante para la empresa.

Metadatos

Debido a la gran importancia que representa el conocer la fuente de los datos y la serie de transformaciones que éstos experimentan desde el ambiente operativo hasta que se le presentan al tomador de decisiones, es necesario la creación de los metadatos o diccionario de datos. La información que tendrá este diccionario de datos será información de los datos que intervengan en el nuevo sistema, así como una perfecta descripción de todos y cada uno de los cambios sufridos desde su origen. Un diccionario de datos debe de almacenar al menos la información señalada en el cuadro 6.10. La cantidad de información que se utilice en el diccionario de datos, permitirá ofrecer una mejor administración y control de los sistemas actuales y de los sistemas futuros.

Cuadro 6.10 Ejemplo de un artículo en el diccionario de datos o Metadatos.

| | |
|---------------------------------------|--|
| NOMBRE DEL CAMPO | REQ_NO |
| EJEMPLO | 00125 |
| TIPO | CARÁCTER |
| ANCHO | 5 |
| DECIMALES | 0 |
| INDICE | SI |
| TABLAS INTEGRADAS DONDE SE UTILIZA | SITUACION.DBF REQ_EN_OC.DBF |
| ENCARGADO ORIGEN | MDR |
| ENCARGADO DESTINO | JMPR |
| ORIGEN | DEPTO PRESUPUESTOS |
| FORMATO DE ORIGEN | KNOWLEDGE MAN |
| ARCHIVO | OSIRIS.ITB |
| NOMBRE DEL CAMPO | DOCTO |
| EJEMPLO | RE-00125__ |
| TIPO | C |
| ANCHO | 10 |
| DECIMALES | 0 |
| INDICE | NO |
| TRANSFORMACION | EL ARCHIVO OSIRIS.ITB SE CONVIERTE A OSIRIS.TXT, DESPUES AL CAMPO DOCTO SE LE ELIMINAN LOS 3 PRIMEROS CARACTERES "RE-" Y LOS ESPACIOS EN BLANCO AL FINAL. DESPUES SE IMPORTA AL ARCHIVO REQ_NO CON LAS CARACTERISTICAS ARRIBA DESCRITAS. |

La tecnología de intranet como apoyo para crear el diccionario de datos

Una forma en la cual, la tecnología de intranet se convierte en gran ayuda en el proceso de creación del almacén de datos, consiste en que ésta tecnología, brinda la oportunidad de crear un archivo común, el cual, puede ser accedido y modificado desde cualquier máquina en la red. Esto permite, que la creación del diccionario de datos, sea un proceso ágil y centralizado, con aportaciones de diferentes áreas. Este diccionario de datos en el ambiente de intranet, permitirá consultar la información de los datos que ya se han integrado, con el fin de que los desarrollos futuros, estén basados en datos que cumplan con los requisitos y

acuerdos aquí establecidos. El grupo de integración, puede estar trabajando en diferentes áreas y en diferentes sistemas y estar consultando y agregando información a un archivo común, solicitar o plantear modificaciones a los datos ya existentes, etc. En resumen, la tecnología de intranet, brinda un ambiente excelente para lograr un diccionario de datos organizacional. Los usuarios de este diccionario, fácilmente pueden identificar los datos que usarán, así como también, pueden agregar los nuevos programas que hacen uso de ellos. Se puede decir que en este caso, el diccionario de datos sería uno de los pilares, donde se apoyaría el proceso de integración, en cuanto a su definición se refiere.

CONCLUSIONES

En este trabajo de tesis, se presentó el concepto "integración aparente de datos", el cual ofrece a las empresas una alternativa rápida y económica para poner a disposición de sus ejecutivos, una herramienta de información indispensable para el proceso de toma de decisiones.

Como se pudo ver en este trabajo, la integración aparente de datos se puede lograr de varias formas, pero la tecnología de intranet ofrece una serie de características, que son de gran ayuda para la administración de información. Para este trabajo en particular, la característica más importante es que la tecnología de intranet ofrece el equivalente a un lenguaje universal, el cual permite un flujo e intercambio efectivo de información entre aplicaciones y equipos de cómputo, anteriormente incompatibles, ofreciendo así, acceso a un sinnúmero de nuevas fuentes de datos.

Este acceso a nuevas fuentes de datos, provenientes de diferentes sistemas y plataformas nos permite implementar con mayor facilidad y eficiencia la integración aparente de datos, pero también es importante mencionar, que esta facilidad ofrecida por la tecnología de intranet, nos genera un nuevo problema, más datos desintegrados. Como se puede ver, la tecnología nos facilita el manejo y acceso a los datos pero por sí sola, no es capaz de generar información integrada. Es necesario apoyar a la tecnología con una metodología que nos guíe a través de procesos y transformaciones para ofrecer a los usuarios datos integrados los cuales pueden provenir de fuentes desintegradas.

En este trabajo se vieron las características y ventajas que la metodología del almacén de datos ofrece a las empresas. Esta metodología por lo regular toma datos del ambiente operativo, los transforma, resume y crea nuevas bases de datos para ofrecer información integrada y de calidad a sus usuarios.

Aquí se plantea utilizar la unión de las ventajas de la metodología del almacén de datos y de la tecnología de intranet, para implementar la integración aparente de datos.

En el caso práctico, se plantea tomar los datos (desintegrados) provenientes de las diferentes aplicaciones y generar un almacén de datos virtual

apoyado en la tecnología de intranet, evitando así, la necesidad de crear nuevas bases de datos.

Esta combinación ofrece una alternativa bastante atractiva para los ejecutivos de ésta y otras empresas, los cuales ahora tienen la posibilidad de estar frente a un ambiente integrado (aparentemente), y resolver gran parte de sus problemas de comunicación y de manejo de información, sin la necesidad de hacer grandes inversiones en tiempo y capital.

También, la tecnología de intranet nos ofrece la posibilidad de crear un diccionario de datos empresarial, el cual es indispensable en un proceso de integración de datos, ya sea integración real o aparente.

El grado de beneficios que una empresa pueda obtener al utilizar el concepto de integración aparente apoyándose con la tecnología de intranet, dependerá de muchos factores, como pueden ser: el ambiente competitivo en el cual se desarrolla la empresa, el grado de interdependencia entre las unidades, el grado de homogeneidad de los sistemas y equipos actuales de la empresa, etc. Pero independientemente de los factores anteriormente expuestos, la búsqueda de la integración aparente, apoyada con la tecnología de intranet, se convierten en una herramienta valiosa y económica que casi cualquier empresa puede intentar con el fin de ofrecer a sus ejecutivos información de calidad.

BIBLIOGRAFIA

- Bickel, Robert. "Building Intranets: Internal Webs give companies a new solution to an old problem".NET CONNECTIONS. Internet World Vol. 7 No. 3 March 1996.
- Derfler, Frank J. Jr., "The intranet Platform: A Universal Client?", PC Magazine Online, April 23, 1996.
- Goodhue, Dale L., Quillard, Judith A., Rockart, John F., "Managing the Data Resource: A Contingency Perspective", MIS Quarterly, Vol 12, No. 3, September 1988. Center for Information Systems Research, Sloan School of Management Massachusetts Institute of Technology.
- Goodhue¹, Dale L. March 1992. Center for Information Systems Research, Sloan School of Management Massachusetts Institute of Technology.
- Goodhue², Dale L., Wybo, Michael D., Kirsch, Laurie J., The impact of data integration on the costs and benefits of information systems. MIS Quarterly / September 1992 p. 293-311.
- Gurley, William. "It's the end of the net as we know it". FORTUNE. April 29, 1996.
- Haynal, Russ. Undersantdign the Intranet. Enterprise Reengineering, March 1996.
- Heimbigner, D. and McLeod, D. "A Federated Architecture for Information Management,"ACM Transactions on Office Information Systems (3 :3), July 1985, pp. 253-278.
- Hinrichs, Randy J., "Intranets: What's The Bottom Line?", SunSoft/Prentice Hall, 1997.
- Icaza, José I., "Implantación gradual de la administración de datos", Reporte técnico CII93-05. Centro de Investigación en Informática. Instituto Tecnológico de Estudios Superiores de Monterrey, N.L., México, 1993,Campus Monterrey.
- Imhoff, Claudia. "Data Stewardship - Finally a Process for Achieving Data Integrity". The Data Administration Newsletter, Robert S. Seiner – Publisher and Editor –Sep. 1997. <http://www.tdan.com/i002fe03.htm>
- Inmon, W.H. "Building the Data Warehouse", second edition, John Wiley & Sons, Inc., U.S.A. 1996, 401 p.
- Inmon , W.H. "Management Control of Data Processing: Preventing Management By Crisis", Prentice-Hall, Inc., Englewood Cliffs, U.S.A., 1983, p. 45-56

- Litwin, W. Mark. L., and Roussopoulos, N. "Interoperability of Multiple Autonomous Databases," ACM Computer Surveys (22 :3), September, 1990, p 267-293.
- Martin, J. Information Engineering, Savant Research Studies, Canforth, Lancashire, England, 1986.
- Martin, J., Leben, J. "Strategic Information Planning Methodologies, 2nd. Ed., Prentice Hall, 1989.
- Martin, J., "Information Engineering". Vol 3., Prentice Hall, 1990.
- Miramon Commons, Bernardo. "Data Warehousing: Estrategias Generales de Implantación", Revista Soluciones Avanzadas.No. 94. 15 de Junio de 1996.
- McCarthy, Vance. "Jump Start Your I-Nets", Datamation, Feb. 1996.
- McGee, James V. and Prusak, Laurence "Managing Information Strategically", The Ernst & Young Center for Information Technology and Strategy., John Wiley & Sons, Inc. 1993 p. 1,21,22
- Niccolai, James, "AT&T Internet, intranet package promises high speed, reliable service", Computerworld. Online News Story, 10/09/97. <http://www2.computerworld.com/home/online9697.nsf/All/971009at&t18F2E>
- Parker, Timothy, "Aprendiendo TCP/IP en 14 días". Prentice Hall Hispanoamericana, S.A., 1995 p. 30-36.
- Smith¹, Anne Marie. 1997, "How to Implement a meta Data Strategy". The Data Administration Newsletter, Robert S. Seiner – Publisher and Editor – <http://www.sgi.net/tdan/i002fe06.htm>,
- Smith², Anne Marie. 1997. "IRM Data Administration vs. Database Administration". The Data Administration Newsletter, Robert S. Seiner – Publisher and Editor – <http://www.sgi.net/tdan/i002fe07.htm>,
- Sullivan, Eamonn. 1996, " The safest bet in a changing world". InterSights. PCWEEK Online, May 20,.
- Telleen, Steven L., "Intranet Methodology (tm): Concepts and Rationale" Intranet planning & management. Amdahl Corporation. 1997. <http://www.amdahl.com/doc/products/bsg/intra/concepts.html>.
- Verstraete, Anthony A., Ph. D. AAV@PSUVM.psu.edu, Senior Lecturer in MIS Management Science and Information Systems Smeal College of Business Administration. 1998. <http://www.smeal.psu.edu/misweb/infosys/ibisinfo.html>

