Importancia de la utilización de un Data Warehouse (DW) en las empresas.

Autores:
MSc. Emma R. Rizo Rizo
Universidad de Matanzas " Camilo Cienfuegos"
Carretera Varadero Km 3 ½, Matanzas.
emma.rizo@umcc.cu
Ing. Mayté Tápanes Mora
Banco Popular de Ahorro
Calle medio, esquina Jovellanos, Matanzas.
maytetapanes@yahoo.es
Dr. Juan Pedro Febles
Centro Nacional de BioInformática
Industria y San José, Capitolio Nacional, Habana Vieja, Habana
febles@bioinfo.edu.cu
Dra. Vivian Estrada Senti.
Ministerio de Educación Superior
Calle 23 y G, Vedado, Habana.
vestrada@reduniv.edu.cu
Dr. Efraín Sánchez Pérez
Universidad de Matanzas "Camilo Cienfuegos"
Carrertera Varadero Km 3 ½, Matanzas.
Atis_efrain@yahoo.es
Est. Roberto Real Reyes
Universidad de Matanzas "Camilo Cienfuegos"
Carretera Varadero Km 3 ½, Matanzas.
roberto.real@infonet.umcc.cu

Con la informatización de la sociedad y dentro de estas las empresas, ha crecido a nivel mundial la capacidad de generación y almacenamiento de la información, que no puede ser analizada por los métodos tradicionales existentes, mientras mayor es la capacidad para almacenar más y más datos, mayor es la incapacidad para extraer información realmente útil de éstos en las empresas. Mucha información importante, quedaba sepultada y disgregada, y los sistemas existentes no estaban preparados para el nuevo reto.

Mucho se ha hablado de la Era de la Información y sus ventajas; con las nuevas posibilidades se acortan las distancias y crecen los beneficios para quienes tienen acceso al gran caudal de datos. Sin embargo, lo que constituye un valioso recurso para todos, se ha tornado en el gran problema de principios de siglo, *manejar de forma óptima grandes volúmenes de información*. (1)

La competencia en el nuevo ambiente corporativo donde todo está marcado por la información, el conocimiento del mercado y la toma de decisiones, es muy importante saber donde y como se organiza toda la información**.** En este contexto surgen términos que se convierten en conceptos y filosofías de trabajo, como *Data Warehouse***. (2)**

Antes de continuar se hace necesario hacer referencia a diferentes definiciones sobre Data Warehouse, lo cual nos facilitara la comprensión de lo que estamos exponiendo en este artículo.

La más conocida fue propuesta por Inmon **(1)** (considerado el padre de las Bases de Datos) en 1992: “Un DW es una colección de datos orientados a temas, integrados, no-volátiles y variante en el tiempo, organizados para soportar necesidades empresariales”.

En 1993, Susan Osterfeldt **(1)** publica una definición que sin duda acierta en la clave del DW: “Yo considero al DW como algo que provee dos beneficios empresariales reales: Integración y Acceso de datos. DW elimina una gran cantidad de datos inútiles y no deseados, como también el procesamiento desde el ambiente operacional clásico”.

Esta última definición refleja claramente el principal beneficio que el Data Warehouse aporta a la empresa, eliminar aquellos datos que obstaculizan la labor de análisis de información y entregar la información que se requiere en la forma más apropiada, facilitando así el proceso de gestión.

Los *Data Warehouse* surgen con la promesa del manejo y control de la información. Ellos aseguran una vista única de los datos, que pueden provenir de diversas fuentes. Gracias a esto, los usuarios finales no se ven en la necesidad de aprender y utilizar múltiples sistemas de acceso y manipulación de los datos. Un *Data Warehouse* facilita la comprensión de los datos, transformándolos en información útil, teniendo como bandera el apoyo a la Toma de Decisiones. (1)

Por lo anterior se puede comprender el carácter universal de esta herramienta, la que es útil en toda empresa, extendiéndose a partir de aquí por empresa, compañía o negocio a toda organización de producción o de servicios con fines lucrativos o no.

Hay una premisa en el mundo del negocio que plantea que el futuro pertenece a quienes pueden verlo y llegar a él primero. Por tanto, es una recomendación a todas las empresas que tengan automatizados todos o parcialmente sus procesos y cuenten con información acumulada sobre los mismos, la implementación de un Data Warehouse, ya que éste permite no solo comprender lo que está pasando, sino predecir lo que va a suceder.

Como hemos planteado anteriormente los sistemas relacionales no están preparados óptimamente para extraer conocimiento de los datos almacenados, algunos pudieran discrepar con este planteamiento, a pesar de que ya es esta una afirmación dada por cierta en el mundo de hoy. Los sistemas relacionales, por mucho que se les mejore, por mucho que se les quiera adaptar, no son los encargados de esa nueva fase de consulta, para eso ha surgido una nueva forma de organización de los datos, los Data Warehouse, estos son el centro de la arquitectura para los sistemas de información a partir de la década de los '90 hasta la fecha. Soporta el procesamiento informático al proveer una plataforma sólida, a partir de los datos históricos para hacer el análisis. Facilita la integración de sistemas de aplicación no integrados. Organiza y almacena los datos que se necesitan para el procesamiento analítico, informático sobre una amplia perspectiva de tiempo, es aquí donde está el punto de coincidencia de nuestro trabajo con la gestión del conocimiento en la empresa.

La gestión del conocimiento está formada por un conjunto de actividades básicas, siendo las más generales las destinadas a identificar, crear, almacenar, compartir y utilizar el conocimiento. Las organizaciones también pueden incluir otras actividades, como es el caso de filtrar y seleccionar el conocimiento, el Data Warehouse provee el almacenamiento de la información de forma tal que permita el descubrimiento de conocimiento en la misma.

Después de todas los aspectos analizados no nos queda duda de que las empresas necesitan un sistema que permita mediante fotos diarias al negocio extraerle la esencia a los datos, hacer reportes dinámicos, presentarlos gráficamente y de manera cómoda para el usuario, un sistema con registro histórico detallado, preparado para manejar sumarizaciones de millones de record, y no demorarse más que unos minutos o incluso menos. Un sistema pensado especialmente para el proceso de consulta y para cualquier tipo de usuario. Y el Data Warehouse es precisamente el más poderoso de esos otros posibles sistemas. De ahí la importancia del mismo para la toma de decisiones, por eso ese planteamiento de que una empresa sin Data Warehouse, es una empresa en desventaja. Simplemente está, sin él, no tiene la facilidad de conocer a fondo el comportamiento de su negocio como para sacarle el máximo provecho.

El objetivo fundamental de un DW desde nuestro punto de vista, es asistir al ejecutivo en el entendimiento del pasado y contar con los elementos para la planeación del futuro de corto, mediano y largo plazo.

Los ejecutivos y administradores buscan respuestas a preguntas como:

- ¿Qué tipos de proyectos están solicitando nuestros clientes? ¿Qué no están solicitando?
- ¿Qué está haciendo la competencia?
- ¿Cómo están los costos por metros de construcción, comparados con los últimos tres años, en proyectos de igual calidad?
- ¿Qué factores causan incrementos en los costos?
- Los ejecutivos quieren tener respuestas a preguntas cruciales para poder tomar decisiones.

Un Data Warehouse ayuda a resolver estas preguntas de forma eficiente y aporta elementos valiosos de toma de decisiones al personal encargado de las mismas, los cuales pueden ser analistas, ejecutivos o directores, y en forma inmediata, esto es, consultas en línea, sin necesidad de que el usuario final solicite que se elaboren y ejecuten procesos especiales. Por lo tanto, es

importante asegurar que los datos del almacén sean adecuados, suficientes y seleccionados de acuerdo a las necesidades de decisión del negocio que se pretende soportar.

Según Inmon el valor de un DW queda descrito en tres dimensiones(3):

✓ **Mejorar la Entrega de Información:** información completa, correcta, consistente, oportuna y accesible. Información que la gente necesita, en el tiempo que la necesita y en el formato que la necesita.

✓ **Facilitar el Proceso de Toma de Decisiones:** con un mayor soporte de información se obtienen decisiones más rápidas; así también, la gente de negocios adquiere mayor confianza en sus propias decisiones y las del resto, y logra un mayor entendimiento de los impactos de sus decisiones.

✓ **Impacto Positivo sobre los Procesos Empresariales:** cuando a la gente accede a una mejor calidad de información, la empresa puede mejorar:

- Eliminar los retardos de los procesos empresariales que resultan de información incorrecta, inconsistente y/o no existente.
- Integrar y optimizar procesos empresariales a través del uso compartido e integrado de las fuentes de información.
- Eliminar la producción y el procesamiento de datos que no son usados ni necesarios, producto de aplicaciones mal diseñados o ya no utilizados.

Por la importancia que le damos al impacto que puede provocar la utilización del Data Warehouse dentro de una empresa profundizaremos en los mismos, ya que como vimos anteriormente es el tercer valor que le concede Inmon a la utilización de este.

El éxito de un data warehouse no está solo en su construcción, sino en usarlo para mejorar procesos empresariales, operaciones y decisiones. Posicionar un DW para que sea usado efectivamente, requiere entender los impactos de implementación en los siguientes ámbitos(4):

a) Impactos Humanos

Efectos sobre la gente de la empresa:

• **Construcción del DW:** a diferencia del desarrollo de aplicaciones, donde los requerimientos de la empresa logran ser relativamente bien definidos producto de la estabilidad de las reglas de negocio a través del tiempo, construir un DW depende de la realidad de la empresa como de las condiciones que en ese momento existan, las cuáles determinan qué debe contener el DW. Como se dijo anteriormente, la gente de negocios debe participar activamente durante el desarrollo del DW, desde una perspectiva de construcción y creación.

• **Accediendo al DW:** el DW intenta proveer los datos que posibilitan a los usuarios acceder a su propia información cuando ellos la necesitan. Esta aproximación para entregar información tiene varias implicancias:

- La gente de la empresa puede necesitar aprender nuevas destrezas.
- Análisis extensos y demoras de programación para obtener información será eliminada. Como la información estará lista para ser utilizada, las expectativas probablemente aumentarán.
- Nuevas oportunidades pueden existir en la comunidad empresarial para los especialistas de información.
- La gran cantidad de reportes en papel serán reducidas o eliminadas.
- La madurez del DW dependerá del uso activo y retroalimentación de sus usuarios.

• **Usando aplicaciones DSS/EIS:** usuarios de aplicaciones DSS y EIS necesitarán menos experiencia para construir su propia información y desarrollar nuevas destrezas. Es decir, que para los usuarios, el DW extiende el alcance de la información para que puedan acceder directamente en línea, lo que a la vez contribuye en su capacidad para operar con mayor efectividad las tareas diarias relacionadas con la toma de decisiones. Los usuarios del DW pueden acceder a una variada información que puede ser vista de forma multidimensional, presentada como una fuente única confiable y disponible directamente por medio de sus estaciones de trabajo. Como se dijo anteriormente, los usuarios pueden usar sus herramientas familiares, hojas de cálculo, procesadores de textos y software de análisis de datos y análisis estadístico para manipular y evaluar la información obtenida desde el DW.

b) Impactos Empresariales

• **Procesos Empresariales y Decisiones Empresariales.**

Se deben considerar los beneficios empresariales potenciales de los siguientes impactos:

- Los Procesos de Toma de Decisiones pueden ser mejorados mediante la disponibilidad de información. Decisiones empresariales se hacen más rápidas por gente más informada.
- Los procesos empresariales pueden ser optimizados. El tiempo perdido esperando por información que finalmente es incorrecta o no encontrada, es eliminado.
- Conexiones y dependencias entre procesos empresariales se vuelven más claros y entendibles. Secuencias de procesos empresariales pueden ser optimizados para ganar eficiencia y reducir costos.
- Procesos y datos de los sistemas operacionales, así como los datos en el DW, son usados y examinados. Cuando los datos son organizados y estructurados para tener significado empresarial, la gente aprende mucho de los sistemas de información. Pueden quedar expuestos posibles defectos en aplicaciones actuales, siendo posible entonces mejorar la calidad de nuevas aplicaciones.

• **Comunicación e Impactos Organizacionales**.

Apenas el DW comienza a ser fuente primaria de información empresarial consistente, los siguientes impactos pueden comenzar a presentarse:

- La gente tiene mayor confianza en las decisiones empresariales que se toman. Ambos, quienes toman las decisiones como los afectados conocen que está basada en buena información.
- Las organizaciones empresariales y la gente de la cual ella se compone queda determinada por el acceso a la información. De esta manera, la gente queda mejor habilitada para entender su propio rol y responsabilidades como también los efectos de sus contribuciones; a la vez, desarrollan un mejor entendimiento y apreciación con las contribuciones de otros.
- La información compartida conduce a un lenguaje común, conocimiento común, y mejoramiento de la comunicación en la empresa. Se mejora la confianza y cooperación entre distintos sectores de la empresa, viéndose reducida la sectorización de funciones.
- Visibilidad, accesibilidad, y conocimiento de los datos producen mayor confianza en los sistemas operacionales y fomenta aún más su uso.

c) Impactos Técnicos del DW

Considerando las etapas de construcción, soporte del DW y soporte de sistemas operacionales, se tienen los siguientes impactos técnicos:

• **Nuevas destrezas de desarrollo:** cuando se construye el DW, el impacto más grande sobre la gente técnica está dada por la curva de aprendizaje, muchas destrezas nuevas se deben aprender, incluyendo: conceptos y estructura DW.

- El DW introduce muchas tecnologías nuevas (ETT, Carga, Acceso de Datos, Catálogo de Metadatos, Implementación de DSS/EIS), y cambia la manera que nosotros usamos la tecnología existente. Nuevas responsabilidades de soporte, nuevas demandas de recursos y nuevas expectativas, son los efectos de estos cambios.
- Destrezas de diseño y análisis donde los requerimientos empresariales no son posibles de definir de una forma estable a través del tiempo.
- Técnicas de desarrollo incremental y evolutivo.
- Trabajo en equipo cooperativo con gente de negocios como participantes activos en el desarrollo del proyecto.

• **Nuevas responsabilidades de operación:** los cambios sobre los sistemas y datos operacionales deben ser examinados más cuidadosamente para determinar el impacto que estos cambios tienen sobre ellos, y sobre el DW. Para la Dirección de Sistemas, el DW enriquece las capacidades del usuario autosuficiente y hace que la Dirección pueda ofrecer nuevos servicios a los usuarios, sin interferir con las aplicaciones cotidianas de producción, aunque se requiere una asignación de tiempo y personal técnico para el mantenimiento y operación del DW.

Como podemos apreciar el data warehouse tiene un efecto positivo sobre toda la vida de la empresa, va desde el desarrollo de los recursos humanos, mejoras tecnológicas, retos en el dominio de las nuevas tecnologías, etc y por supuesto toma de decisiones más eficientes producto de todo lo anterior.

A pesar de todos los beneficios anteriormente expuestos para lograr una correcta aplicación de esta herramienta presupone un comprometimiento de la alta dirección del negocio, pues este lleva consigo un trabajo colaborativo entre estos y el especialista o los especialistas de informática que van asumir el diseño y creación del mismo, es un trabajo largo y difícil, será necesario por parte de la empresa realizar algunas inversiones que de inmediato no tendrán una respuesta, pero al final facilitará la aplicación de técnicas estadísticas de análisis y modelización para encontrar relaciones ocultas entre los datos del almacén; obteniendo un valor añadido para el negocio, proporcionará la capacidad de aprender de los datos del pasado y de predecir situaciones futuras en diversos escenarios.

También será necesaria una optimización tecnológica y económica en entornos de Centro de Información, estadística o de generación de informes que faciliten la utilización de estas técnicas en la toma de decisiones.

Conclusiones.

Para terminar nuestra exposición podemos decir que los data warehouse no están divorciados de los sistemas operacionales, el primero se nutre de la información que se va acumulando en este ultimo, pero sólo llevando a este aquella información necesaria para la toma de decisiones de la empresa.

Podemos destacar que el DW incluye dos etapas fundamentales, la obtención y transformación de los datos para la carga del DW desde los sistemas operacionales y los procesos y métodos de acceso a los mismos para la

obtención de las diferentes salidas necesitados por el gerente del negocio para su análisis.

El Datawarehousing es la tecnología que mejor permite integrar e soportar el nuevo modelo del negocio, permitiendo ver la información de forma general a lo particular según le interese profundizar al ejecutivo, y que posteriormente permite el desarrollo de estrategias más efectivas y rentables en la toma de decisiones.

Es importante enfatizar en la estrecha relación que debe existir entre el informático que proyectará y mantendrá el sistema y los administrativos del negocio para garantizar un buen diseño y funcionamiento del mismo.

Bibliografía.

1. Vidal LV, Monteagudo MV. Estudio Teórico-Conceptual sobre Data Warehouse [Pregrado]. Ciudad de La Habana.: Universidad de La Habana; 2000.
2. Altuna BLB. Data Mart de Faturación, Data Warehouse Cubacel [Tesis de Pregrado]. Cuidad de La Habana: Universidad de La Habana 2002.
3. Inmon, W. H.; Hackathorn; Richard D. “Using the Data Warehouse”. New York: John Wiley & Sons. ISBN: 0-471-05966. 1994
4. UNED. Implementación de un Data Warehouse para la Universidad Estatal a distancia(UNED). 2002. p. 15.
5. Rizo MER, Cárdenas LES, Pérez DES. Data Warehouse nuevo concepto para informático. UMCC; 2006.
6. Moras IMT, Daniel ILM. Diseño de un data warehouse para la EMPAI. Matanzas. Cuba 2006
7. Education M. Data Warehousing, Data Modeling and Design. MicroStrategy Education. 1996.
8. I.N.E.I. Manual para la construccion de un Data Warehouse(Caracteristicas de un Data Warehouse). 1997 [cited 2005]; Available from: http://www.inei.gob.pe/web/metodologias/attach/lib619/13.htm
9. Nader IJ. Sistema de Apoyo Gerencial Universitario. Buenos Aires: ITBA; 2003.
10. Echeverría DG. Data Warehouse. 1997.
11. Luján-Mora S, Trujillo J. Una método global basado en UML para el diseño de Almacenes de Datos. 2003. p. 11.
12. Sperley E. The Enterprise Data Warehouse:Planning, Building, and Implementation.; 1999.
13. Abelló A, Samos J, Saltor. F. YAM2 (Yet Another Multidimensional Model):An Extension of UML. . Edmonton,Canada: Intl. Database Engineering & Applications Symposium 2002.
14. Cabibbo L, Torlone R. A Logical Approach to Multidimensional Databases. 1998;vol. 1377:183-97.
15. Golfarelli M, Rizzi. S. A methodological Framework for Data Warehouse Design; 1998.
16. Husemann B, Lechtenborger J, Vossen. G. Conceptual Data Warehouse Design. Estocolmo, Suecia; 2000.
17. Sapia C, Blaschka M, Hö ing G, Dinter. B. Extending the E/R Model for the Multidimensional Paradigm. 1998;vol. 1552:105-16.
18. Trujillo J, Palomar M, Gómez J, Song I. Designing Data Warehouses with OO Conceptual Models. IEEE Computer. 2001:66-75.

18. Tryfona N, Busborg F, Christiansen. JG. A Conceptual Model for Data Warehouse Design. Kansas City, USA; 1999.
19. Grady B, James R, Ivar. J. El lenguaje unificado de modelado. . 2000.
20. Rodríguez YD, Roca MC, Oliva RE. GESTOP. Sistema Informático para la Gestión de Operaciones. [Pregrado]. Matanzas: Instituto Superior Politécnico “José Antonio Echeverría”; 2005.
21. Corporation R. Lo nuevo de Rational Rose 2000 2000.
22. Hernández YT. Sistema DataWarehouse del comercio minorista de la corporación CIMEX [pregrado]. Ciudad de La Habana: Universidad de La Habana 2002.
23. Luján-Mora S, Trujillo J. A Data Warehouse Engineering Process. 2004. p. 10.
24. Mazón JN, Trujillo J, Serrano M, Piattini M. Designing Data Warehouses: From Business Requirement Analysis to Multidimensional Modeling. 2000. p. 10.
25. Kimball R, Ross M. The Data Warehouse Toolkit : the complete guide to dimensional modeling. Second Edition ed: John Wiley and Sons, Inc.; 2002.
26. Trujillo J, Luján-Mora S. A UML Based Approach for Modeling ETL Processes in Data Warehouses. 2003. p. 14.