

CALIDAD DE LOS DATOS



RESUMEN EJECUTIVO

No. 12-2006

boletín técnico.

Las empresas de nuestros días, constantemente buscan la manera de mejorar su desempeño, adquirir ventajas competitivas y entender a su mercado reaccionando de manera ágil y proactiva utilizando los datos e información disponible. Esto implica superar la problemática de que los datos utilizados en agrupaciones, sumalizaciones o clasificaciones no representan la realidad, lo anterior sucede debido a que no existen definiciones estándares o integradas de los datos que son claves para realizar análisis estratégicos, si las hay, no son respetadas y por lo general hay errores debido a causas tan simples como "errores de dedo". Lo anterior es resuelto con una adecuada y conveniente implementación de estándares, metodologías y procesos que gestionan la calidad de los datos, donde se aplican principalmente acciones correctivas que permiten tener los datos limpios y de acuerdo a las reglas del negocio.

**COMITÉ TÉCNICO NACIONAL DE
TECNOLOGÍA DE LA INFORMACIÓN**

Por la Lic. Patricia Luna Arredondo

CONSEJO DIRECTIVO NACIONAL 2006

Presidente

Dr. Marco Antonio Cerón Grados

Presidente del Consejo Técnico

C.P.C. Sergio Federico Ruiz Olloqui Vargas

Secretario CDN y Director General IMEF

IQ MBA Juan Carlos Erdozain Rivera

COMITÉ TÉCNICO NACIONAL DE TECNOLOGÍA DE LA INFORMACIÓN

PRESIDENTE

Lic. Miguel Ángel González Granados

INTEGRANTES

Ing. Agustín Pérez

Lic. Alberto Rubio

Lic. Arturo Carvallo

Lic. Arturo Fernández Maldonado

C.P. Carlos Humphrey Pasalagua

Ing. David Goldstein Weitzman

C.P. Ernesto Javier Campos Cervantes

C.P. Isaías Alday Campos

Lic. Carlos Vargas Hernández

Ing. Héctor Rodríguez Álvarez

Ing. Javier Allard Taboada

Lic. Javier Ramírez Mendoza

Ing. José Alfredo Ibarra Almada

Ing. José A. Casas Alatraste Urquiza

Isc. José Gabriel Vilchis P.

Act. José María Alcántara Jiménez

Ing. José Manuel Cano Muñiz

C.P. Luis H. Arredondo Barrera

Lic. Luis Vera Vallejo

Lic. Manuel Osuna y Fernández

Lic. María de los Ángeles Flores

Lic. Mario A. Guerrero Mendoza

Lic. Miguel Ángel González Granados

Lic. Miguel Ángel Iza Martínez

Ing. Flavio Andrés Mora Elizalde

C.P. Guadalupe Castañeda Campos

Ing. Héctor Joel González Rodríguez

Lic. Patricia Luna Arredondo

Dr. Pedro Flores Jiménez

Ing. Ricardo Zermeño González

Lic. Rodolfo González Gatica

C.P. Samuel Servin Espino

Ing. Sergio Ceballos Flores

lic. Maghandi Suárez Pichardo

Ing. José Antonio Lona

Lia. Pedro Reyes Cortés

Coordinador del Comité Técnico

Nacional de Tecnología de la Información

ÍNDICE TEMÁTICO

Introducción

Análisis Técnico

Aspectos generales del uso de los datos

Definiciones fundamentales

Calidad de los datos

Alternativas de Solución

Conclusiones

Bibliografía

INTRODUCCIÓN

Hoy en día las empresas cuentan con herramientas tecnológicas de hardware y software para el registro, procesamiento y administración de sus transacciones u operaciones de negocios.

La frase "si entra basura, basura sale" es muy conocida por los niveles de gerencia y dirección, trasladando esta frase a términos fuertes de la implicación de los datos en los procesos de toma de decisiones, podemos asegurar que si la información contiene basura, entonces las decisiones que se tomen también serán basura.

En el entorno de negocios de nuestros días, no podemos darnos el lujo de tomar decisiones equivocadas, mucho menos de destinar tiempo gerencial o de análisis y toma de decisiones, para validar si los datos en los que sustentamos nuestras decisiones son correctos.

El problema de los datos que no tienen la debida calidad, no es nuevo, es un problema que tiene sus orígenes desde que se iniciaron los primeros sistemas de soporte a decisiones, esto como consecuencia de la implementación de algún sistema tipo ERP (Enterprise Resource Management), mejor conocidos como sistemas operacionales, sistemas aplicativos de negocio o sistemas transaccionales.

Las preguntas de siempre son:

- ¿Los gerentes tienen la información que quieren, necesitan y deben?
- ¿Si los gerentes tienen la información de la pregunta anterior, entonces, su proceso de toma de decisiones puede ser mejor?

Las respuestas por lo general a estas preguntas tienen relación entre dos elementos:

- 1) el proveer información oportuna y precisa
- 2) la habilidad de las personas para usar esta información con el objetivo de contribuir a mejorar el desempeño de la empresa

Estos dos elementos tienen que ver con la forma en la que la empresa principalmente obtiene, genera, procesa, administra y depura sus datos. Estas actividades tienen directa implicación con la calidad de los datos, si los datos tienen la calidad debida, entonces, las decisiones tomadas también tendrán calidad.

ANALISIS TÉCNICO

Aspectos generales del uso de los datos

El conducir una empresa nos habla de un gran proceso en el cual las metas organizacionales deben ser logradas a través de un uso eficiente de los recursos, esto se mide en términos de productividad, es decir, de la relación existente entre los insumos contra los resultados.

El nivel de productividad depende directamente del desempeño de las funciones gerenciales principalmente clasificadas como planeación, organización, integración, dirección, control y retroalimentación. Es sabido por todos que para realizar estas funciones, el gerente o director debe estar siempre tomando decisiones.

Durante años, el proceso de toma de decisiones se había considerado únicamente como un arte puro, un talento que se adquiría por medio de la experiencia. Esto se ha modificado con el paso del tiempo, ahora con la Tecnología de Información y Comunicaciones nos permiten revisar y analizar el desempeño pasado y presente con el fin de pronosticar o definir el futuro de una forma más racional y ágil.

Debido a que el medio ambiente en el que la Alta Dirección se desempeña, es tan cambiante, que de no contar con la ayuda de herramientas automatizadas y vanguardistas, sería imposible entender la complejidad del mundo de negocios de nuestros días en ambos aspectos: interno y externo. Principalmente por que el número disponible de alternativas y variables de las operaciones del negocio es cada vez mayor y de complejidad múltiple.

Por consiguiente, las empresas constantemente buscan la manera de mejorar su desempeño, adquirir ventajas competitivas y entender a su mercado reaccionando de manera ágil y proactiva, utilizando los datos e información disponibles.

Frecuentemente la necesidad de información del negocio es para responder preguntas clave relacionadas con:

- ¿Cómo conducimos el negocio?
- ¿Qué relaciones existen entre productos, clientes, mercado, precios, costos y el tiempo?
- ¿Qué es necesario mejorar, cambiar ó eliminar?
- ¿En que tipo de negocio debo ingresar, mantenerme o salirme?

Principalmente los gerentes toman decisiones basándose en sus sistemas aplicativos de negocio, donde abundan los datos acerca de los clientes, empleados, socios de negocio, los productos, los volúmenes de venta, la facturación, la cobranza, los inventarios, etc. Estos datos, pueden ser utilizados para mejorar el conocimiento del cliente, reducir costos, mejorar productos y mitigar riesgos. Sin embargo, el uso de los sistemas aplicativos para la toma de decisiones es limitado, ya que de manera consistente las aplicaciones han demostrado que son una fuente no fiable de la información para análisis y lentos en los procesos de preparación de información para que puedan ser útiles en análisis de negocios, aunque si es muy

confiable desde el punto de vista de cómo la empresa realiza sus transacciones de negocio.

Esto ocurre en todas y cada una de las empresas, independientemente de su tamaño, giro o sector. La pregunta es, ¿por qué no es fiable, si ahí es donde se operan las transacciones de negocio, donde se realiza la contabilidad y donde está la información de nuestros clientes?, ¿por qué no es fiable si la implementación de estas aplicaciones fue resultado de una gran inversión financiera? La respuesta es muy simple, en el ámbito transaccional la información es correcta, es decir, los datos de qué productos vendimos, en qué cantidades, a qué precio, a qué costo, a quién se las vendimos, cuándo se las vendimos, cuándo se las entregamos, cuánto nos pagó, cuándo nos pagó, que contabilidad se registró, etcétera, es correcta. Pero si queremos llegar al detalle analítico para la toma de decisiones, es muy frecuente que nos encontremos con que la clasificación de los productos tiene errores, que la agrupación de los productos en categorías o familias esté incongruente, que los estándares de conversión de gramajes, kilos, litros, etcétera está mal capturado, que la clasificación del mercado esté equivocada, que la referencia no exista, etcétera; esto no afecta la operación normal del negocio por que son datos que por lo general sólo son utilizados para efectos de análisis de penetración de mercados, de competencia, de rentabilidad de marcas/productos, análisis de mezcla de productos, de costos de distribución por marca/producto, entre otros. En pocas palabras, ver desde otro punto de vista la forma en la que la empresa realiza sus operaciones del negocio.

Lo anterior sucede debido a que no existen definiciones estándares o integradas de los datos que son claves para realizar análisis estratégicos, si las hay, entonces no son totalmente respetadas y por lo general hay errores debido a causas tan simples como "errores de dedo".

Es importante entender que las aplicaciones transaccionales tienen su objetivo de satisfacer los requeri-

mientos de colección y almacenamiento de datos y proveen la infraestructura de software para procesar operaciones de negocio y por lo general cuentan con la validación de los datos necesarios para este procesamiento.

Normalmente ocurre que una empresa que implementó una o varias aplicaciones de negocio, se da cuenta que éstas no cubren sus necesidades de información analítica para la toma de decisiones, en particular para integración de datos de la operación del mes, datos históricos para comparativos, tendencias y realizar proyecciones, datos detallados y datos sumariados principalmente.

Esta problemática se puede resolver mediante la implementación de metodologías y procesos que gestionan la calidad de los datos.

La calidad de datos tiene su importancia estratégica en el valor intangible de los sistemas aplicativos, si los datos en esos sistemas son redundantes, inexactos o incompletos, pueden dañar seriamente la posición competitiva de la empresa al tomar las decisiones equivocadas.

Por otro lado, si los datos no tienen la suficiente calidad, entonces, también tendrán poca calidad y bajo o nulo éxito, las iniciativas estratégicas, tales como e-business, Customer Relationship Management (CRM), Aplicaciones de Soporte a Decisiones (DSA), o Aplicaciones de Inteligencia de Negocios (Business Intelligence). Estas iniciativas dependen de los datos de la empresa que se encuentran en las aplicaciones transaccionales, si la calidad de los datos en esas aplicaciones es pobre, entonces impactará en la confiabilidad, exactitud o precisión, incluso el impacto puede llegar a afectar la efectividad de la implementación de estas iniciativas.

La calidad de los datos es extremadamente importante, porque la calidad determina la utilidad de los datos y la calidad de las decisiones basadas en ellos.

Definiciones fundamentales

Para el claro entendimiento de la calidad de los datos, su importancia, su impacto y de cómo se debe implementar un método o proceso de calidad de los datos, es necesario, primero entender las siguientes definiciones fundamentales.

Dato: Números, letras, símbolos o hechos que describen un objeto, idea, condición, situación u otro factor. El dato es un elemento básico de la información que puede ser procesado. Los datos representan la información.

Elemento de dato: Es el nombre de un conjunto de datos por ejemplo: Meses es el elemento de dato para el contenido de Enero, Febrero, Marzo, Abril, Mayo, Junio, Julio, Agosto, Septiembre, Octubre, Noviembre y Diciembre.

Registro de datos: Es un grupo de elementos de dato que se manipulan como una unidad, por ejemplo, el registro de datos generales de un empleado se integra por los siguientes elementos de dato: número de nómina, nombre, fecha de nacimiento, fecha de contratación, número de puesto, sueldo.

Archivo de datos: Un grupo de registros de datos, por ejemplo, el archivo de datos generales de los empleados.

Base de datos: Conjunto de archivos de datos que integran las relaciones que existen entre los diversos archivos a nivel registro de datos o grupos de registro de datos. En otras palabras, es

una colección de archivos interrelacionados, almacenados en conjunto, procesados, sin redundancias innecesarias y cuya finalidad es la de ser utilizada en por lo menos un sistema aplicativo.

Información: Es el valor o significado que adquieren los datos al ser procesados.

Almacén de Datos (Data Warehouse): Es una colección permanente de información que está aislada de los sistemas aplicativos de operación del negocio, la cual sirve para apoyar el proceso de toma de decisiones de la alta gerencia y dirección de una empresa. El Data Warehouse contiene información que si es adecuadamente extraída, procesada, analizada, combinada y almacenada, entre otras acciones, puede ser una herramienta sumamente importante que incrementa la efectividad en los procesos de toma de decisiones.

Procesamiento de datos: Cualquier operación que el sistema aplicativo o el conjunto de sistemas aplicativos realizan sobre los registros de datos, desde el momento en que son capturados o ingresados en los sistemas hasta el momento en que se depuran o destruyen. Las operaciones de datos abarcan las siguientes acciones: calculo, compilación, unión, selección, búsqueda, sumarización, interpretación, almacenamiento, recuperación, comparación, etc.

Administración de datos: Procesos, instrucciones, reglas, seguridad o rutinas de programas, sistemas o conjunto de sistemas que registran, almacenan y permiten el acceso a la información.

Sistema aplicativo de negocio: Un conjunto de programas, archivos y base de datos que permiten que el negocio realice sus operaciones mediante el procesamiento de datos y su administración. Existen sistemas aplicativos de facturación, contabilidad, compras, etcétera o bien aquellos que integran varias funciones y operaciones de la empresa en un gran sistema aplicativo.

Regla de negocio: Es el significado pre-definido de un proceso de transacciones operativas de la empresa.

Calidad de datos: Es la conveniencia o conformidad de los datos de acuerdo a ciertos requerimientos. Dependiendo de estos requerimientos, una base de datos puede ser buena para una aplicación de negocio, pero no necesariamente para otra. Por lo que quizá sea necesario construir y poblar un almacén de datos o Data Warehouse, por ejemplo, elementos de datos cuyo contenido (dato) esté vacío o tenga un valor incongruente con alguna regla de negocio.

Calidad de los datos

Como se explicó anteriormente, el principio de la toma de decisiones radica en la información que está en los sistemas aplicativos de negocio, mejor conocidos como sistemas operacionales. Es importante recordar que estas soluciones automatizadas han sido diseñadas e implementadas para cubrir las necesidades de operación del negocio, más no son un soporte adecuado para el proceso de toma de decisiones ejecutivas.

Existe una categoría específica de aplicaciones denominada de Inteligencia de Negocios (Business Intelligence), cuya finalidad es un soporte para el análisis de información a los niveles gerenciales y directivos de diversos departamentos, permitiendo la posibilidad de poder afrontar y responder ante la rapidez en la que los cambios del mercado se presentan, además de que apoya en la identificación, anticipación y respuesta a las oportunidades del mercado, obteniendo ventaja de ellas de manera rápida y efectiva.

Esto suele parecer muy sencillo, sin embargo, aún cuando vivimos en la era de la información, tenemos tanta información a nuestra disposición que nuestro problema es comprender su significado. La Tecnología de Información proporciona las herramientas, pero la tecnología por sí misma no puede identificar ni resolver los problemas u oportunidades del negocio.

Las aplicaciones de Inteligencia de Negocios permiten integrar información, gente y tecnología para tomar las mejores decisiones rápidamente, convertir los datos en información útil, relevante y al alcance de la mano para utilizar esto en la gestión empresarial.

No puede existir una plataforma de Inteligencia de Negocios sin ser alimentada con los datos transaccionales de los sistemas operacionales, de hecho deben coexistir; sin embargo, es innegable que existe una brecha entre los datos que los sistemas operacionales tienen y la información que los responsables de la toma de decisiones necesitan, por lo que es importante tomar en cuenta ciertas consideraciones o procesos para convertir los datos en información útil para análisis y toma de decisiones. Esto puede ser muy simple o muy complicado, todo depende de la plataforma tecnológica de software y hardware con que se cuente.

Una plataforma de software común de Inteligencia de Negocios cuenta siempre con un Data Warehouse o Almacén de Datos, el cual se constituye de una serie de procesos y componentes que integran datos de uno o múltiples sistemas operacionales.

El Data Warehouse tiene como objetivo proporcionar información haciendo que los datos independientemente de su volumen, sean correctos, sean consistentes y congruentes para toda la empresa y accesibles para los usuarios finales de una manera eficiente. El Data Warehouse es el centro de información para la toma de decisiones, es un proceso de gestión del flujo de la información dentro de la organización, incluye procesos de extracción, limpieza, transformación, mantenimiento y depuración.

Un componente importante del Data Warehouse es el proceso de integración-transformación, donde ocurren las siguientes funciones que tienen relación directa con las reglas de negocio: conversiones, formato, reestructura, recalcular, definición de valores default, resumir datos, agregar datos, mezclar datos,

resecuenciar, cambiar llaves, añadir elementos de tiempo, semántica de datos, etc.

Estas funciones son necesarias para la calidad de los datos incluyendo su confiabilidad, exactitud, relevancia e integridad.

Para entender la calidad de los datos, partamos de la premisa que solo podemos generar información tan buena como los datos que utilizamos para crearla, cuando trabajemos con calidad de datos debemos de tener presente la expresión: basura entra a basura sale.

Imagínese que no pueda confiar en los datos que se obtienen de los sistemas que utiliza día con día, comúnmente ocurre que los datos que se guardaron en Excel no son los mismos que aparecen en un reporte de su aplicación de negocio, entonces usted o quien es responsable tomar decisiones, tiene que hacer alguna(s) de las siguientes funciones: conciliar, comparar, cuadrar y validar para estar seguro que los datos que va a analizar son confiables para su proceso de toma de decisiones. Esto le toma tiempo, y a veces después de horas e incluso días se da cuenta que el "error" proviene de un dato mal clasificado o un error de dedo.

Lo anterior se evita mediante la implementación de estándares, metodologías y procesos de calidad de los datos donde se aplican principalmente acciones correctivas que permiten tener datos limpios, por lo que es necesario preparar al Data Warehouse con las reglas de negocio para que las acciones que se decida hacer sobre los datos sean las más adecuadas.

La calidad de los datos es una responsabilidad compartida entre los usuarios, las herramientas de software, los administradores de las bases de datos, los usuarios finales de los sistemas aplicativos, etc. ya que puede darse el caso que lo que para unas personas es un error y para otras es una excepción, por ello es muy importante definir claramente las reglas de negocio y los criterios de calidad.

La calidad de los datos no significa el tener los datos perfectos, tampoco implica tener las más caras tecnologías y tampoco significa que no se puede medir.

Es importante que se establezcan normas de calidad de datos en ciertas áreas clave, esto con el fin de evitar que los datos sirvan peligrosamente para tomar decisiones mal fundamentadas ya que el tener datos erróneos puede ser demasiado costoso.

Expertos en el tema de Data Warehouse han dividido en 4 categorías a la Calidad de los Datos:

- Calidad Contextual: Relevancia, valor agregado, oportunidad y completos.
- Calidad Intrínseca: Exactitud, objetividad, credibilidad y reputación.
- Calidad de Accesibilidad: El acceso a los datos y la seguridad de acceso.
- Calidad de Representación: Interpretación, fácil de entender, conciso y consistente.

De manera pragmática, la Calidad de los Datos significa simplemente: tener los datos **conforme a los requerimientos**, por lo tanto, la clave para definir la calidad de los datos recae en determinar esos requerimientos, criterios o indicadores, los más importantes se definen a continuación:

- **Existencia:** Se debe tener una relación de existencia de los datos entre las aplicaciones operacionales de negocio vs la existencia de los datos en el Data Warehouse.
- **Integridad:** Es una medida que garantiza que los datos contenidos en el Data Warehouse son consistentes y congruentes. Es un porcentaje de los registros que tienen elementos de dato no congruentes con las reglas de negocio establecidas, estos pueden ser errores de captura o ingreso de datos ocasionados por pocos o nulos controles de edición y validación, errores de programación de otras aplicaciones, etc. Mientras más baja sea la integridad, la cantidad de trabajo de limpieza es mayor.
- **Propiedad:** El dueño de la información sensible debe estar claro y es quien debe autorizar su acceso, sobre todo a datos sensibles. Es una medida del porcentaje de elementos de dato sensibles sobre el total de elementos de dato.
- **Completo:** Es una medida de contar con el alcance de todos los datos de las transacciones relacionadas. Por lo general la métrica es un porcentaje de los datos completos sobre los datos disponibles.
- **Exactos:** Es una medida de la exactitud o precisión de la información. Por lo general la métrica es un porcentaje de los registros con errores (puede ser por datos o por relaciones) vs el total de registros. También puede tener implicaciones del uso de decimales o escalas en el manejo de números.

- **Confiables:** Que tan viejos son los datos, de que generación de tiempo (cierre de mes, semanal, diario, etc), si los datos son obtenidos directamente de las aplicaciones operacionales o son producto de respaldos u otros procesos.
- **Oportunos:** Es una medida de la disponibilidad de los datos cuando estos son necesarios, significa la periodicidad de la actualización de los mismos y la forma en que son entregados al usuario de una forma en la que se encuentran aceptables. Por lo general es un porcentaje de los datos disponibles en el tiempo especificado.
- **Costo:** Es un factor de calidad, asegura que el valor del dato e información del negocio es considerado. La calidad no es cara, la falta de calidad si. Costo actual vs el costo presupuestado. El costo se calcula mediante una relación entre el tiempo de proceso y el costo de todos los elementos de operación de la infraestructura de hardware y gente destinados para el actualizar el Data Warehouse.

Los requerimientos o indicadores anteriores, para que sean funcionales y se puedan medir, deben acompañarse del establecimiento de estándares y de sus niveles de tolerancia, esto significa que para medir si la calidad de los datos es satisfactoria, el uso de indicadores es clave para entender su valor contra los niveles de tolerancia.

La medida de calidad de datos puede medirse como el nivel de cumplimiento logrado entre el ambiente de datos y las reglas definidas.

ALTERNATIVAS DE SOLUCIÓN

Como se ha explicado en los apartados anteriores, la calidad de los datos es sumamente importante ya que la calidad determina la utilidad de los datos y también la calidad de las decisiones basadas en ellos.

Podemos partir de la implementación de controles en los sistemas operacionales, tales como de validación en captura de datos, de selección de valores predefinidos para elementos de datos, entre otros, principalmente en los sistemas operacionales, sin embargo, esto encarecerá la operación del negocio haciéndola más lenta ya que dependen de la velocidad, disposición y compromiso con la calidad de los datos de las personas que los ingresan y procesan las transacciones, por lo general esto no sucede, ya que son datos que no paran la operación de la empresa e incluso no le interesan al personal responsable de la operación del negocio.

Sin embargo, lo anterior, a la larga se convertirá en un dolor de cabeza, ya que los módulos desarrollados alrededor de sus sistemas transaccionales para garantizar la calidad de los datos, irán encareciendo el mantenimiento y complicando su plataforma de aplicaciones de negocio.

Lo más indicado es hacer lo necesario para garantizar la calidad de los datos en la plataforma de Inteligencia de Negocio, esto, a la larga, le permitirá tener una ágil y sustentable

operación transaccional además de contar con la fortaleza estratégica de datos de una plataforma de Inteligencia de Negocios.

Sabemos que las aplicaciones de Inteligencia de Negocio cuentan con un Data Warehouse, y para resolver los problemas de calidad de los datos, uno de los componentes más importantes es el Metadatos, cuyo significado simplemente es datos acerca de los datos.

Una definición más completa nos dice que los Metadatos son los datos físicos (almacenados en algún medio) y todo el conocimiento de la información acerca de las reglas de negocio, de los datos, de sus restricciones y de su procesamiento, este conocimiento esta almacenado en algún medio e interactúa con el capital intelectual de su empresa.

El metadato le dice al usuario que datos están disponibles, cuales son sus nombres, cuales son sus definiciones, sus relaciones, sus propiedades (tipo, longitud, valores predefinidos, etc), cual es su contenido, de que aplicaciones se tomaron, si hay otras aplicaciones donde existan, quién es el responsable, cómo fue calculado, que criterios de reglas de negocio, de limpieza, de transformación y de calidad son empleados.

El Metadato es una herramienta automatizada de conocimiento y tratamiento de los datos, pero no elimina la labor manual de los propietarios de los datos o de los administradores de los datos de mantener éstos dentro de ciertos criterios de calidad.

Muchas empresas se quedan con la duda acerca de la conveniencia de limpiar o no limpiar los datos mediante procesos complejos o bien de implementar un Metadatos. Por lo general no se detienen a calcular el costo de la calidad de los datos y sobre todo del costo de la calidad de las decisiones, pudiendo ser decisiones buenas o malas, todo radica en función de la calidad de los datos.

Las actividades destinadas para la limpieza y transformación de datos para mantener el Data Warehouse dentro de los criterios de calidad, son fundamentales, independientemente si la empresa tiene el interés y la posibilidad de implementar un Metadatos.

Lo más recomendable es iniciar la implementación de un Metadatos mediante una construcción progresiva, con el paso del tiempo, su empresa podrá contar con un Metadatos que reúna las reglas de negocio y todos los datos acerca de sus datos que le evitarán análisis, validaciones y cuadros innecesarios que son realizados por las personas responsables de tomar las decisiones en su empresa. Por consiguiente y como resultado de este esfuerzo, su empresa tomará las mejores decisiones y de una manera rápida.

CONCLUSIONES

Las decisiones de negocio recaen en el manejo y uso de la información, frecuentemente la información es un activo no administrado, ni siquiera se le da la importancia debida y cuando se requiere sustentar fuertemente alguna decisión, esta se ve retrasada debido a que es necesario validar su sustento, lo que incrementa los costos de la decisión y quizá la pérdida de la oportunidad de ejecución de las acciones derivadas de la decisión.

El mejorar la calidad de los datos a través de ciertas reglas de negocio, procesos o criterios, proveerá a la empresa de una poderosa forma de mejorar el proceso de análisis y toma de decisiones.

Entendiendo las métricas de calidad de los datos e implementando una metodología o proceso de calidad para los datos en las aplicaciones de Inteligencia de negocio, la empresa podrá entonces incrementar la efectividad de sus decisiones e incluso tomarlas con certeza en un tiempo realmente corto.

Lo anterior impactará gratamente en los resultados de su empresa, obtenidos mediante el análisis prospectivo adecuado y proveniente de sus propios datos de negocio.

También será muy claro el valor al negocio de las iniciativas estratégicas que implican la implementación de otro tipo de aplicaciones, tales como CRM, SCM, BPM, etc.

BIBLIOGRAFIA

Building the Customer-Centric Enterprise
Claudia Imhoff. Lisa Loftis. Jonathan G. Geiger.
Editorial Wiley

Building and Managing the Meta Data Repository
David Marco
Editorial Wiley

ESTIMADO SOCIO

boletín técnico.

Cualquier comentario, observación o sugerencia a este Boletín favor de hacerlo llegar directamente a la autora.

Lic. Patricia Luna Arredondo

Socio Director

Calimeria Business Intelligence, S.A. de C.V.

e-mail: patricia@calimeria.com