

Integridad de Datos y BI

El Valor Estratégico de la Integración



Tabla de contenidos

- 0. Introducción
- 1. Calidad de datos vs mundo real
- 2. Integración de datos ante el permanente cambio en la dinámica del negocio
- 3. En busca de la integridad de los datos
- 4. Información: valor y riesgo
- 5. Criticidad de la integración de datos en proyectos de Business Intelligence
- 6. Integración de datos y BI: el marco de trabajo ideal en 4 pasos
- 7. Siete factores críticos de éxito para la integración de datos y Business Intelligence
- 8. Importancia de la calidad de datos para la toma de decisiones
- 9. Integridad de datos interna y externa
- 10. Conclusiones

0. Introducción

Business Intelligence no es fruto únicamente de un componente tecnológico, sino que supone el cúmulo de procesos, personas y tecnología abocados a la obtención de información de alto valor para la toma de decisiones. De hecho, las empresas bien informadas pueden orientar sus acciones hacia los objetivos prefijados de una forma más directa, por lo que las soluciones de Business Intelligence son vitales en la competitividad del negocio y críticas a la hora de encontrar y reforzar las ventajas competitivas.

Y todo esto se apoya en datos. Datos que no provienen del área de TI, ya que allí no se crean la información (sólo se proporcionan mecanismos para almacenar, recuperar y archivar los datos con precisión y de forma segura), sino que se obtienen de todas partes y al mismo tiempo, en un flujo continuo. Esto significa que, muchas veces, las distintas fuentes que proporcionan datos pueden ser incompatibles entre sí.

Una empresa que invierte tiempo y dinero en una iniciativa de BI:

- Ha construido y fundamentado cierto retorno sobre su inversión.
- Ha trabajado por la reducción de desperdicios y la maximización de beneficios, tal y como dicta el clima económico actual.
- Cuenta con un marco sólido de desarrollo de proyectos que comprende una buena arquitectura de datos, gobierno de datos de calidad, las herramientas adecuadas y el liderazgo eficaz.

Pero siempre verá limitada la fiabilidad del análisis de negocio que realice a la calidad de los datos que tenga. En definitiva, puede decirse que, a pesar de toda la inversión, si la calidad de los datos y las pruebas de integridad de datos son pobres, el negocio estará sumido en un escenario de alto riesgo, que se multiplica cuando intervienen además procesos de integración.

1. Calidad de datos vs mundo real

El mundo real es lo suficientemente complejo y dinámico, en cuanto al volumen y características de la información que, proveniente de él, se procesa diariamente. Considerando su estrecha relación con la información almacenada en la base de datos, es de importancia crítica el poder dedicar un esfuerzo constante a lograr que el conjunto de datos empresarial generado mantenga un alto grado de fidelidad que le sea equiparable.

Es por ello que se debe minimizar la posible desviación existente entre los hechos almacenados en las bases de datos, que representan el comportamiento pasado, presente o esperado de nuestro negocio; y el mundo real que se trata de capturar. La dificultad reside en que la realidad del negocio cambia muchas más veces y mucho más rápido de lo que es posible digerir en términos de datos e información. Al mismo tiempo, la forma de adaptarse a estos cambios puede afectar de forma negativa a la calidad de los datos, si no se garantiza la seguridad en los procesos.

El desafío para las empresas

El reto para muchas organizaciones es, poder seguir el ritmo que la realidad del negocio impone y, a la vez, mantenerse competitivas y preparadas para la única constante, el cambio. Así, en términos de información, son requeridas las siguientes acciones:

- Reducir la brecha existente entre la frecuencia de producción y acceso a los datos y su procesamiento.
- Mantener estos niveles controlados en el tiempo.
- Establecer procesos para garantizar que una base de datos siga siendo el fiel reflejo del universo que se está representando.
- Tratar de aproximarse a un modelo en el que la actualización continua y la gestión automática de datos sean una realidad.
- Poder hacerlo sin perder las cualidades del dato: integridad, calidad y veracidad; que son indispensables para su correcta interpretación.
- Garantizar que los procesos de integración de datos no se ven afectados de forma negativa en ningún momento.

2. Integración de datos ante el permanente cambio en la dinámica del negocio

La dinámica del negocio está sujeta a muchos cambios, tanto de origen interno, como procedentes del entorno; y por ello se hace imprescindible el emplear un tiempo en analizar esas transformaciones y su posible impacto sobre la empresa. Sólo de esta forma es posible estar preparados a la hora de tomar acciones que aseguren que la calidad de los datos permanece intacta y que los procesos de integración de datos no se han visto afectados de manera negativa.

Cómo evaluar el cambio

Existen dos formas de abordar esta evaluación. Ambas opciones se complementan, garantizando la capacidad de análisis perseguida en el nivel adecuado de control que se precisa:

- Evaluar la frecuencia y magnitud del cambio: consiste en identificar aquellos aspectos que detonan el cambio y el esfuerzo del negocio, para abordarlo midiendo su impacto. Así se hace posible ajustar esfuerzos futuros en base a las propias capacidades.
- Identificar variables endógenas y exógenas que impactan al proceso de integración de datos: concretamente, habría que tomar en consideración todos esos factores agrupados bajo el denominador común "variables de crecimiento", y entre los que se encuentran:
 - Apoyo y patrocinio comprometido por parte de la alta dirección (dirigido por la estrategia del negocio).
 - Visión y modelo de negocio correctamente establecidos.
 - Composición equilibrada de equipos de expertos (SME's).
 - Calidad y cantidad de datos controlables.
 - Aplicación de tecnologías escalables, flexibles y amigables.
 - Metas y objetivos de la organización claramente establecidos.
 - Gestión eficaz de proyectos.

- Participación activa de las partes interesadas (stakeholders).
- Precisión e integridad de los datos y la información.
- Infraestructura y uso de sistemas legados.
- Fiabilidad y flexibilidad del sistema.
- Acompañamiento del cambio y las tendencias tecnológicas.
- Selección de tecnologías que conjuguen habilidad de ejecución con visión evolutiva.

Afianzando el autoconocimiento

Una vez analizado el entorno, hace falta concentrarse en los propios sistemas para conocer dónde se encuentra la empresa, y para ello sería oportuno:

- Analizar requerimientos de flujo de datos (unidireccional vs bidireccional): determinando si la petición de movimiento de datos es unidireccional o bidireccional; la única manera de poder alcanzar un nivel de comprensión óptimo sobre las necesidades del negocio y sus procesos, que más adelante se traducirá en mecanismos de flujo de datos. De esta forma, es posible plantear procesos ETL de una forma más generalizada, donde el movimiento de datos no deba fluir en un único sentido necesariamente.
- Considerar la diversidad y tamaño de las fuentes y tipos de datos: manejando un número mucho más grande de datos de diferentes tipologías (datos estructurados y semiestructurados) a la vez que una gran variedad de fuentes (tanto internas como externas) que garanticen que se cumplen:
 - Requisitos de baja latencia para apoyar la toma de decisiones en tiempo real (relación de entrada y salida de datos casi imperceptible).
 - Ciclos de actualización de datos veloz.
 - Mayor número de conexiones entre Data Marts.
 - Necesidad de disponibilidad de los sistemas para toma de decisiones 24x7.

La integración de datos

Por último, queda centrarse en la integración de datos en sí misma, atendiendo tanto a su papel en relación a la estrategia de la empresa, como a la calidad del propio dato, que debería ser la base de cualquier proceso. Y para ello, se necesita:

- Alinear los requerimientos del negocio con los procesos de integración de datos: la alineación de requerimientos del negocio con las capacidades de TI, sumado a la alineación de procesos y

tecnología, es uno de los aspectos más complejos en cualquier organización debido a la incertidumbre tecnológica que existe.

- Establecer métricas para evaluar la calidad de los procesos de integración de datos. La vía más efectiva para monitorizar estos procesos garantizando:

- Confiabilidad (determinar que el trabajo es realizado en el tiempo y forma esperados).
- Mantenimiento (en lo referente a niveles de servicio).
- Vigencia (proporcionar latencia adecuada en la gestión de datos).
- Escalabilidad (capacidad de manejo de mayor volumen de datos).
- Disponibilidad (garantizar que los recursos físicos del sistema estarán disponibles cuando sea necesario).
- Flexibilidad (capacidad para adaptarse a nuevas necesidades o cambios del negocio).
- Consistencia (asegurar que los datos que pueblan el almacén de datos son correctos y cumplen con restricciones de integridad).
- Trazabilidad (capacidad de rastreo de la procedencia de los datos y sus cambios o linaje).
- Auditoría (con la finalidad de proteger la privacidad y la seguridad de datos orientadas al cumplimiento legal).

3. En busca de la integridad de los datos

Muchas veces, los principios del manejo de los datos empresariales asumen la calidad de los mismos y su idoneidad a la hora de tomar decisiones partiendo de su procesamiento. Sin embargo, lejos de actuarse de manera acertada, esta vía especula sobre la fiabilidad de los datos, que quizás no son tan saludables y aptos como se piensa, como para asumir riesgos desde planteamientos estratégicos a favor del negocio.

Es por ello que, el paradigma de Business Intelligence, procura establecer un proceso claramente diferenciado en la maduración de los datos existentes que pasa por:

- Transformar datos en bruto en información relevante y oportuna.
- Traducir la información en conocimiento estratégico y operacional.
- Explotar dicho conocimiento para entender dónde existe valor dentro del negocio de tal forma que se puedan enfocar la capacidad y los recursos disponibles.

Hoy día los negocios se mantienen en pie y son altamente competitivos gracias a sus decisiones acertadas, pero también fracasan a partir de aquellos juicios que se basan únicamente en conclusiones intuitivas empapadas en subjetividad, que quedan muy lejos de un aprovechamiento lógico del alto potencial implícito contenido en sus datos.

Ante un inminente crecimiento, tanto en la forma como en la dinámica del mercado global actual, las empresas deberán afinar sus estrategias para lograr procesar cada vez mayores cantidades de datos que vendrán en diferentes formas y procederán de fuentes muy diversas. Todo ello asegurando la calidad del dato, que debe considerarse la base de cualquier proceso.

Integridad de datos: garantías implícitas

La integridad de datos hace referencia a tres cualidades que se deben asegurar para poder partir de una base de fiabilidad de la información que se maneja:

- Precisión de los datos: es la profundidad del conocimiento determinada por los datos de que se

dispone. Puede evaluarse de distintas formas, como por ejemplo en función del nivel de desagregación de los indicadores de desempeño que se están utilizando para medir el del negocio.

- Confiabilidad de los datos: que determinará el grado de riesgo que se está dispuesto a asumir a la hora de tomar decisiones con base en la información procesada.

- Completitud de los datos: en referencia al factor crítico que permite obtener una visual inclusiva, en donde la información pueda validarse desde diferentes perspectivas, sin perder precisión y maximizando la confiabilidad. Estrechamente relacionado con las fuentes de datos, en relación a si contienen o no toda la información necesaria para satisfacer la demanda empresarial actual y futura.

4. Información: valor y riesgo

Todos los negocios se sustentan en base a datos y, hoy en día, la mayoría de las empresas los recogen, gestionan y utilizan con dos finalidades básicas:

- Operativa, necesaria para el normal funcionamiento de la empresa: sería la finalidad que hay detrás de los datos operacionales.
- Interpretativa. El uso analítico de los datos respalda el objetivo de mejorar el negocio y sus perspectivas.

Cuando los aspectos analíticos logran permear en el nivel de operación del negocio, la empresa es capaz de explotar el conocimiento generado y crear valor que proporcione retroalimentación. Procurar este flujo continuo de datos y velar por su seguimiento supone minimizar desviaciones y afianzar puntos fuertes.

Ambos escenarios de uso, operativo y analítico, necesitan contar con una alta calidad de los datos manejados. Esto sugiere la necesidad de establecer procesos que garanticen que estos niveles son alcanzados y las expectativas satisfechas, como:

- Los procesos de operación, análisis y creación de valor.
- El propio ciclo de vida del dato.
- Los distintos desafíos que atraviesan las organizaciones en torno a la información y su gestión.

Variables que influyen directamente en el valor de la información

La ley de conservación de la energía establece que “la energía no se crea ni se destruye, sólo se transforma”. Aplicando esta premisa al mundo empresarial y del manejo de datos, es fácil observar que el dato tampoco se genera por sí mismo, sino que es producto de la conjunción de elementos y hechos tales como las ventas, las compras, etc. procesos en los que intervienen diferentes variables (endógenas y exógenas) que afectan invariablemente a su creación.

Una transacción o acontecimiento en el mundo real da origen a un dato, formado por características y elementos cualitativos y/o cuantitativos. Este dato, una vez concebido, recorrerá su propio ciclo de

vida y se adaptará a los diferentes requerimientos de información según sea necesario para el negocio, viviendo tanto tiempo como su relevancia dicte.

Durante esta trayectoria se mezclará con otros datos, interactuando para generar información de valor sin perder su esencia original, pero nunca morirá en el tiempo. Un dato podrá destruirse pero no sin antes haber afectado los hechos y entorno del contexto empresarial que lo originó. Por ello también puede decirse que el dato no se autogenera ni se destruye, sino que se transforma.

El gran desafío de las organizaciones de hoy es lograr que sus datos mantengan la integridad desde su origen, garantizando que sean precisos y confiables. Un dato que no contenga dichas cualidades podrá mezclarse con otros para generar información que, a su vez, carecerá de fiabilidad y por lo tanto, resultará en una solución de Business Intelligence de alto riesgo para la toma de decisiones.

Los tres factores que más habitualmente pueden dar origen a un dato con baja completitud o integridad son:

- El factor humano: que durante el registro, copia o transferencia de datos puede inyectar errores.
- Inconsistencia en las operaciones del negocio: por problemas de desfase o debido a la forma empleada para procesar y/o actualizar los datos.
- Defectos a la hora de integrar y conciliar datos provenientes de diferentes fuentes: debido a su incompatibilidad original o por errores posteriores durante su proceso de integración.

Es por esto que, a la hora de garantizar la integridad de los datos, es primordial asegurar las bases para una correcta integración, orientada al Business Intelligence.

5. Criticidad de la integración de datos en proyectos de Business Intelligence

Hoy día, pueden encontrarse multitud de soluciones de Business Intelligence en el mercado. Sin embargo, la diferencia entre unas y otras está en su utilidad, fiabilidad y perspectiva. De hecho, muchas de ellas presentan una importante falta de interoperabilidad con los sistemas de BI, y se limitan a transferir datos y generar reportes y/o dashboards, aportando prácticamente nada a la visión del negocio.

Cuando esto ocurre, generalmente se debe a que estas soluciones de Business Intelligence carecen de la infraestructura básica que permita garantizar la integración de datos en los procesos de extracción, transformación y carga. Y este problema, que se inicia al darse esta falta de solidez, se propaga por los sistemas de BI, ya que éstos son tan confiables como los datos que presentan, por lo que, cuando se basan en información obsoleta, incompleta o inexacta, comienzan a dar lugar a lagunas de fiabilidad que traen consigo graves repercusiones.

La visibilidad y el enfoque, claves de la Business Intelligence

Para tomar decisiones enfocadas a acciones es necesario partir de una base informativa que lo permita. Quienes ostentan esta responsabilidad precisan de ciertos niveles de calidad y actualización en los datos que jamás podrían obtenerse desde un enfoque lineal tradicional, como el que venía empleándose hasta ahora, y que impulsaba el análisis de la información histórica. Hoy existe una nueva demanda de información accionable que se sustenta en determinados factores críticos.

Para lograr esta necesaria visibilidad a través de Business Intelligence y de una plataforma unificada de integración de datos hace falta asegurar:

- Movimiento optimizado de datos.
- Calidad de los datos.

- Datos en tiempo real con alta consistencia.
- Interoperabilidad entre sistemas heterogéneos.
- Una BI que detone acción.

Las tendencias en la integración de datos para los proyectos de BI en la nueva era de la información

Al hablar de las tendencias en integración de datos nos hallamos ante la dicotomía ¿ETL o EL-T? Es cierto que las tecnologías ETL soportan la extracción, transformación y la integración de datos desde múltiples fuentes. Sin embargo, las tecnologías EL-T comienzan a demostrar ventajas significativas sobre las ETL. Y, de hecho, algunas compañías (las mejor equipadas y más avanzadas tecnológicamente), han comenzado a realizar procesos de reingeniería para pasar de soluciones ETL a EL-T.

Los motivos son sus beneficios, como el que la extracción y el proceso de carga puedan aislarse del proceso de transformación. Pero optar por EL-T tiene otras ventajas:

- El proceso de cambio implica menor costo, tiempo y riesgo.
- Permite extraer datos desde las diferentes fuentes y cargarlos en el Data Warehouse para luego transformarlos estando ya allí almacenados.
- Aporta significativos beneficios, en términos de desempeño, optimización para appliances y otros aspectos especialmente interesantes para las nuevas soluciones altamente demandantes de Business Intelligence.

6. Integración de datos y BI: el marco de trabajo ideal en 4 pasos

La información es el activo empresarial que mayor valor aporta a la organización, independientemente de su tamaño, actividad, proyecto u objetivos. El dato se revela como el eje sobre el que giran el resto de variables que afectan al día a día de una empresa y, por ello, para establecer un marco de integración apropiado, es fundamental considerar tanto procesos, como estándares, personas y herramientas dentro de la estrategia de posicionamiento informacional adoptada.

Mediante el uso de un marco de integración de datos, es posible transformar los datos en información consistente y de alta calidad para que los responsables de la toma de decisiones logren, en cada momento y cada vez que sea preciso, medir y dar estrecho seguimiento a los indicadores de desempeño.

Cuáles son los componentes del marco de trabajo en la integración de datos

Tal y como se mencionaba en líneas anteriores, para establecer un marco de trabajo adecuado a la integración de datos y que propicie la calidad que se exige, es necesaria una perspectiva en cuatro dimensiones que comprenda:

- Procesos (de dos tipos):

- Procesos de identificación de requerimientos de los datos y de la propia solución.
- Procesos de recolección física de datos y su correspondiente transformación orientada a la toma de decisiones.

- Estándares: mediante los cuales se logra obtener consistencia y precisión en los datos, asegurando al mismo tiempo su integridad y validez. Ejemplo de áreas de aplicación serían la gestión de

proyecto, la tecnología o la arquitectura.

- Herramientas: dentro de un marco de integración de datos pueden encontrarse herramientas de modelado (físico y lógico), de perfilamiento (almacenamiento y formato), de preparación (limpieza), propagación (agregación-sumarización y de formulación. Estas herramientas serían empleadas, por ejemplo, en el momento de ejecutar la transferencia de datos del Data Warehouses a los Data Marts y también cuando se requiera el manejo de metadatos, relacionado con las distintas entidades del negocio.
- Personas y habilidades: un entorno de trabajo adecuado requiere de gente con experiencia, personas que puedan demostrar un completo entendimiento, tanto de la práctica como de los procesos que se originan en torno a Data Warehouse y Business Intelligence

7. Siete factores críticos de éxito para la integración de datos y Business Intelligence

Cuando se busca alcanzar una integración de datos y Business Intelligence exitosa, es necesario atender a una serie de variables que determinarán el buen curso del proceso. Algunas de ellas se resumen a continuación:

- Balance entre requerimientos locales y globales: un imprescindible, ya que la falta de este equilibrio podría conducir a la generación de silos de información. Para conseguirlo hace falta tener en cuenta ambos tipos de requerimientos y para ello se hará necesario aplicar bottom-up y top-down.
- Enfoque entre calidad de datos y control: el aspecto más complejo en la implementación de una gran solución de Business Intelligence es la calidad de los datos y, por eso, para conseguir alcanzar un enfoque que mantenga su alineación con la directrices corporativas, hay que garantizar una estructura de gestión de datos rigurosamente aplicada. Para lograrlo, se hace necesario en primer término el identificar a los propietarios de los datos para la creación de datos maestros específicos.
- Flexibilidad en el modelo de datos: este aspecto es crítico ya que hay que tener en cuenta la escalabilidad en cualquier modelo de datos, especialmente si se entiende comprendida en un entorno sometido a un continuo cambio. Sólo aplicando esta premisa se puede llevar a buen puerto una iniciativa de Business Intelligence.
- Estandarización, reusabilidad y automatización: en toda iniciativa de automatización aplicada a la integración de datos, la estandarización es un aspecto clave. En concreto, la implantación de estándares globales sin un arsenal de componentes reutilizables, conlleva un gran desafío que pocas organizaciones pueden asumir.
- Procesos: es imprescindible asegurar que el equipo de ejecución comprende los procesos de negocio y los datos subyacentes, ya que el impacto de realizar un cambio en un punto posterior en el tiempo daría lugar a sobreesfuerzos, al haberse perdido claridad en los procesos del negocio.

- Manejo de dependencias: el grado de dependencia en una iniciativa de Business Intelligence es alta y lo es mucho más en iniciativas de carácter global, por eso es importante que las interdependencias sean gestionadas muy de cerca. Para ello se necesita identificar los posibles cambios y/o mejoras que se han programado, evaluando el impacto potencial de los mismos sobre el sistema en conjunto.

- Desafíos operativos (24x7): tanto la secuenciación como la planificación de recuperación ante fallos son aspectos muy importantes a tener en cuenta. En el caso de una infraestructura global este desafío se incrementa, ya que las cargas de datos y el reporting se producen casi de manera simultánea.

Cuando la integración de datos se ha realizado de forma impecable, el resultado es una BI útil que dará lugar a una toma de decisiones más precisa, donde el riesgo queda minimizado. Sin embargo, para que las acciones que se desprenden de esas resoluciones estén encaminadas hacia los objetivos de forma efectiva, hay que asegurar en todo momento la calidad de los datos que son sometidos a procesamiento.

8. Importancia de la calidad de datos para la toma de decisiones

Garantizar la calidad de los datos no es tarea trivial

Garantizar la calidad de los datos no es sencillo. Está claro que lo deseable sería contar con datos precisos, actualizados y completos pero, por desgracia, el mundo real está lejos de ser ideal. Lograr datos de alta calidad obliga a tener una comprensión muy clara del significado, el contexto y la intención de los mismos, donde no existan las ambigüedades y además, a ser posible, se cuente con definiciones estandarizadas que puedan servir de base para la futura toma de decisiones. En este proceso sirve de gran ayuda el poder contar con:

- Conocimiento de las fuentes de origen:
- Control de la trayectoria del dato.
- Glosario de negocio.

¿Cuándo se descubren los errores en los datos?

En el mejor de los casos, la mala información se reconoce de inmediato y es excluida del proceso de toma de decisiones. En el peor, los datos defectuosos, incompletos o poco fiables no se reconocen, y provocan una toma de decisiones errónea.

En la mayoría de las veces, por desgracia, la mala información se descubre al final del proceso de transformación de los datos, momento en que el flujo de información ha llegado a manos de su consumidor final y las acciones han sido llevadas a cabo.

Este tipo de información tiende a ser generada utilizando datos que se han almacenado durante algún tiempo, y repararla puede suponer un proyecto complejo si no se cuenta con los procesos, la capacidad y la tecnología adecuados para ello.

Aspectos fundamentales para una solución de Business Intelligence

La información de alta calidad es esencial en todos los aspectos del ecosistema empresarial hoy en día. La mejora de la calidad de los datos y, por tanto, de la información derivada de éstos y el conocimiento que generan, son aspectos fundamentales durante el proceso de implementación de una solución de Business Intelligence. y también como paso previo a esta etapa.

Por lo general, se producen muchos errores e incidentes adversos en los procesos de toma de decisiones estratégicas, que no son achacables a la solución de BI en sí misma, sino que son el resultado de datos e información pobres. Esto conlleva:

- Un impacto directo al grado de competitividad de la empresa.
- Un aumento en los costes de los procesos.
- La inhibición del intercambio de información, el análisis, la investigación y las iniciativas de medición del desempeño.

9. Integridad de datos interna y externa

Hay varias formas en las que los problemas de calidad de datos pueden afectar al proceso de concepción de una solución analítica de Business Intelligence y cada una de ellas estaría relacionada con una cualidad diferente del dato comprendida dentro del concepto de integridad.

Integridad de datos interna

Esta condición tendría su reflejo en problemas de tipo referencial dentro de las bases de datos, es decir, en el modo en que los datos encajan entre sí.

- Ejemplo práctico: podrían ocasionarse problemas con la integridad de datos interna si una unidad de negocio se eliminase del sistema, pero los contactos que incluye no fuesen suprimidos a la par. En este caso, los contactos pasarían a ser "huérfanos" al perder toda relación con la entidad de negocio que los gestó.
- Problema: dependiendo del tipo de análisis que se esté haciendo, una situación de este tipo podría derivar en ciertos problemas en el estudio de la información.
- Solución: las políticas que se gestan e implementan desde el data Governance son las encargadas de velar por esta integridad de los datos.

Integridad de datos externa

La integridad de datos externa atañe a la consistencia de los datos entre sistemas heterogéneos.

- Ejemplo práctico: si se fusionan los datos entre Ventas, donde se contiene al cliente XYZ, y Contabilidad, donde se ha registrado al mismo cliente como XYZ S.A..
- Problema: no existe match tras fusionar datos supuestamente "idénticos" entre dos o más sistemas.

- Solución: mantener la integridad externa significa asegurar la consistencia de los datos a través de las diferentes aplicaciones fuente.

10. Conclusiones

Debido a su auge en los últimos años, la inversión actual en tecnologías de Business Intelligence y herramientas analíticas ha ido en aumento. Pero es importante asegurarse también de que la calidad en los datos sea proporcional, ya que ésta es la única manera de poder alcanzar las metas planteadas.

El desembolso en herramientas de BI no necesariamente generará un retorno en la inversión por el mero hecho de adquirirlas e implementarlas. Es necesario hacer, de forma previa, un minucioso análisis en cuanto a calidad de datos.

Los riesgos a los que la empresa deberá enfrentarse cuando no garantiza unos mínimos en cuanto a integridad de los datos, pasan por la imposibilidad misma de procesarlos y, aún en mayor medida, de convertirlos en información relevante para la toma de decisiones. El resultado es una falta de perspectiva absoluta que acrecienta las desviaciones, separando cada vez más a la organización de la consecución de sus objetivos estratégicos.

Integridad de datos y toma de decisiones, los desafíos

Cuando el volumen de la información a gestionar va en aumento y el tiempo para hacerlo disminuye cada vez más rápido hay que plantearse:

- Que en los años venideros habrá que enfrentarse a tipos de datos más complejos de los que se han venido manejando hasta el momento (donde los datos estructurados han sido el estándar de uso en los sistemas operacionales por norma general).
- Los mayores desafíos vendrán de la mano de Big Data y los datos semiestructurados y no estructurados, como resultado de la difusión de uso de las redes sociales.
- La voz de cliente se perfila como un factor de significativa relevancia a la hora de resolver muchas cuestiones que conciernen la proyección de la empresa y su competitividad.
- Business Intelligence ha dejado de ser únicamente una forma de estudiar el comportamiento del negocio en el pasado, para entenderse como una práctica que permite abordar iniciativas analíticas

en tiempo real (BI operacional o cercano a tiempo real) apoyándose en las nuevas tecnologías.

- Las diferentes necesidades del negocio dictarán la forma en que se explotará la información, ya sea para estudiar el comportamiento pasado o entender en tiempo real qué está ocurriendo, adquiriendo así la capacidad de tomar decisiones conforme las situaciones o eventos se vayan produciendo.