

## Texto y enlaces ocultos

Ocultar texto o enlaces en el contenido para manipular el posicionamiento en los resultados de búsqueda de Google se puede considerar un comportamiento engañoso e infringe las directrices para webmasters de Google. Se puede ocultar contenido de texto (como un número excesivo de palabras clave) de distintas maneras. Por ejemplo:

- Usar texto blanco con fondo blanco
- Incluir texto detrás de una imagen
- Usar CSS para incluir texto fuera de la pantalla
- Seleccionar el valor 0 como tamaño de fuente
- Ocultar un enlace incluyendo un hipervínculo en solo un carácter (por ejemplo, en una coma en mitad de un párrafo)

Al evaluar tu sitio para ver si incluye texto o enlaces ocultos, busca todo lo que un usuario que visite tu sitio no pueda leer con facilidad. ¿Hay texto o enlaces que son solo para buscadores y no para usuarios?

No todo el texto oculto se considera engañoso. Por ejemplo, si en tu sitio se usan tecnologías a las que los buscadores no pueden acceder fácilmente, como JavaScript, imágenes o archivos Flash, usar texto descriptivo en estos elementos

puede mejorar su accesibilidad. Recuerda que hay muchos usuarios que utilizan lectores de pantalla, navegadores móviles, navegadores sin conectores y conexiones lentas, por lo que no podrán ver ese contenido y aprovecharán también el texto descriptivo. Para probar la accesibilidad de tu sitio, desactiva JavaScript, Flash y las imágenes en tu navegador o utiliza un navegador de solo texto como Lynx. Consulta estas sugerencias para que tu sitio sea accesible:

- Imágenes: utiliza el atributo "alt" para incluir texto descriptivo. Además, recomendamos acompañar la imagen de un pie y de un texto descriptivo que puedan leer los usuarios. Consulta más consejos sobre la publicación de imágenes.
- JavaScript: coloca el mismo contenido de JavaScript en una etiqueta <noscript>. Si utilizas este método, debes comprobar que el contenido es exactamente idéntico al de JavaScript y que se muestra a los usuarios que no tienen JavaScript habilitado en su navegador.
- Vídeos: incluye un texto descriptivo sobre el vídeo en HTML. También puedes añadir transcripciones. Consulta más consejos sobre la publicación de vídeos.

## **Texto y enlaces ocultos**

Ocultar texto o enlaces en el contenido para manipular el posicionamiento en los resultados de búsqueda de Google se puede considerar un comportamiento engañoso e infringe las directrices para webmasters de Google. Se puede ocultar contenido de texto (como un número excesivo de palabras clave) de distintas maneras. Por ejemplo:

- Usar texto blanco con fondo blanco.
- Incluir texto detrás de una imagen
- Usar CSS para incluir texto fuera de la pantalla
- Seleccionar el valor 0 como tamaño de fuente.
- Ocultar un enlace incluyendo un hipervínculo en solo un carácter (por ejemplo, en una coma en mitad de un párrafo)

Al evaluar tu sitio para ver si incluye texto o enlaces ocultos, busca todo lo que un usuario que visite tu sitio no pueda leer con facilidad. ¿Hay texto o enlaces que son solo para buscadores y no para usuarios?

No todo el texto oculto se considera engañoso. Por ejemplo, si en tu sitio se usan tecnologías a las que los buscadores no pueden acceder fácilmente, como

JavaScript, imágenes o archivos Flash, usar texto descriptivo en estos elementos puede mejorar su accesibilidad. Recuerda que hay muchos usuarios que utilizan lectores de pantalla, navegadores móviles, navegadores sin conectores y conexiones lentas, por lo que no podrán ver ese contenido y aprovecharán también el texto descriptivo. Para probar la accesibilidad de tu sitio, desactiva JavaScript, Flash y las imágenes en tu navegador o utiliza un navegador de solo texto como Lynx. Consulta estas sugerencias para que tu sitio sea accesible:

- Imágenes: utiliza el atributo "alt" para incluir texto descriptivo. Además, recomendamos acompañar la imagen de un pie y de un texto descriptivo que puedan leer los usuarios. Consulta más consejos sobre la publicación de imágenes.
- JavaScript: coloca el mismo contenido de JavaScript en una etiqueta <noscript>. Si utilizas este método, debes comprobar que el contenido es exactamente idéntico al de JavaScript y que se muestra a los usuarios que no tienen JavaScript habilitado en su navegador.
- Vídeos: incluye un texto descriptivo sobre el vídeo en HTML. También puedes añadir transcripciones. Consulta más consejos sobre la publicación de vídeos.

## **Texto y enlaces ocultos**

Ocultar texto o enlaces en el contenido para manipular el posicionamiento en los resultados de búsqueda de Google se puede considerar un comportamiento engañoso e infringe las <u>directrices para webmasters</u> de Google. Se puede ocultar contenido de texto (como un número excesivo de palabras clave) de distintas maneras. Por ejemplo:

- Usar texto blanco con fondo blanco
- Incluir texto detrás de una imagen
- Usar CSS para incluir texto fuera de la pantalla
- Seleccionar el valor 0 como tamaño de fuente
- Ocultar un enlace incluyendo un hipervínculo en solo un carácter (por ejemplo, en una coma en mitad de un párrafo)

- Imágenes: utiliza el atributo "alt" para incluir texto descriptivo. Además, recomendamos acompañar la imagen de un pie y de un texto descriptivo que puedan leer los usuarios. Consulta más consejos sobre la publicación de imágenes.
- JavaScript: coloca el mismo contenido de JavaScript en una etiqueta <noscript>. Si utilizas este método, debes comprobar que el contenido es exactamente idéntico al de JavaScript y que se muestra a los usuarios que no tienen JavaScript habilitado en su navegador.
- Vídeos: incluye un texto descriptivo sobre el vídeo en HTML. También puedes añadir transcripciones. Consulta más consejos sobre la publicación de vídeos.

## Texto y enlaces ocultos

Ocultar texto o enlaces en el contenido para manipular el posicionamiento en los resultados de búsqueda de Google se puede considerar un comportamiento engañoso e infringe las directrices para webmasters de Google. Se puede ocultar contenido de texto (como un número excesivo de palabras clave) de distintas maneras. Por ejemplo:

- Usar texto blanco con fondo blanco
- Incluir texto detrás de una imagen
- Usar CSS para incluir texto fuera de la pantalla
- Seleccionar el valor 0 como tamaño de fuente
- Ocultar un enlace incluyendo un hipervínculo en solo un carácter (por ejemplo, en una coma en mitad de un párrafo)

- Imágenes: utiliza el atributo "alt" para incluir texto descriptivo. Además, recomendamos acompañar la imagen de un pie y de un texto descriptivo que puedan leer los usuarios. Consulta más consejos sobre la publicación de imágenes.
- JavaScript: coloca el mismo contenido de JavaScript en una etiqueta <noscript>. Si utilizas este método, debes comprobar que el contenido es exactamente idéntico al de JavaScript y que se muestra a los usuarios que no tienen JavaScript habilitado en su navegador.
- Vídeos: incluye un texto descriptivo sobre el vídeo en HTML. También puedes añadir transcripciones. Consulta más consejos sobre la publicación de vídeos.

## Texto y enlaces ocultos

Ocultar texto o enlaces en el contenido para manipular el posicionamiento en los resultados de búsqueda de Google se puede considerar un comportamiento engañoso e infringe las <u>directrices para webmasters</u> de Google. Se puede ocultar contenido de texto (como un número excesivo de palabras clave) de distintas maneras. Por ejemplo:

- Usar texto blanco con fondo blanco
- Incluir texto detrás de una imagen
- Usar CSS para incluir texto fuera de la pantalla
- Seleccionar el valor 0 como tamaño de fuente
- Ocultar un enlace incluyendo un hipervínculo en solo un carácter (por ejemplo, en una coma en mitad de un párrafo)

- Imágenes: utiliza el atributo "alt" para incluir texto descriptivo. Además, recomendamos acompañar la imagen de un pie y de un texto descriptivo que puedan leer los usuarios. Consulta más consejos sobre la publicación de imágenes.
- JavaScript: coloca el mismo contenido de JavaScript en una etiqueta <noscript>. Si utilizas este método, debes comprobar que el contenido es exactamente idéntico al de JavaScript y que se muestra a los usuarios que no tienen JavaScript habilitado en su navegador.
- Vídeos: incluye un texto descriptivo sobre el vídeo en HTML. También puedes añadir transcripciones. Consulta más consejos sobre la publicación de vídeos.

## Texto y enlaces ocultos

Ocultar texto o enlaces en el contenido para manipular el posicionamiento en los resultados de búsqueda de Google se puede considerar un comportamiento engañoso e infringe las <u>directrices para webmasters</u> de Google. Se puede ocultar contenido de texto (como un número excesivo de palabras clave) de distintas maneras. Por ejemplo:

- Usar texto blanco con fondo blanco
- Incluir texto detrás de una imagen
- Usar CSS para incluir texto fuera de la pantalla
- Seleccionar el valor 0 como tamaño de fuente
- Ocultar un enlace incluyendo un hipervínculo en solo un carácter (por ejemplo, en una coma en mitad de un párrafo)

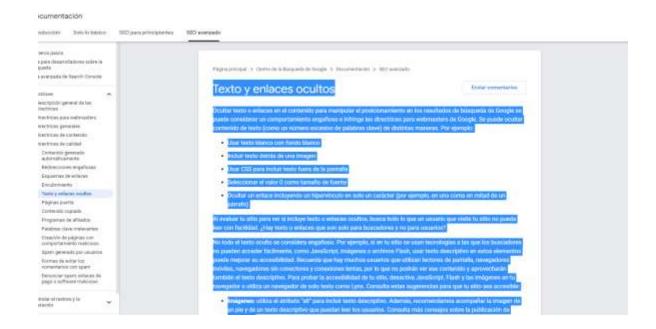
- Imágenes: utiliza el atributo "alt" para incluir texto descriptivo. Además, recomendamos acompañar la imagen de un pie y de un texto descriptivo que puedan leer los usuarios. Consulta más consejos sobre la publicación de imágenes.
- JavaScript: coloca el mismo contenido de JavaScript en una etiqueta <noscript>. Si utilizas este método, debes comprobar que el contenido es exactamente idéntico al de JavaScript y que se muestra a los usuarios que no tienen JavaScript habilitado en su navegador.
- Vídeos: incluye un texto descriptivo sobre el vídeo en HTML. También puedes añadir transcripciones. Consulta más consejos sobre la publicación de vídeos.

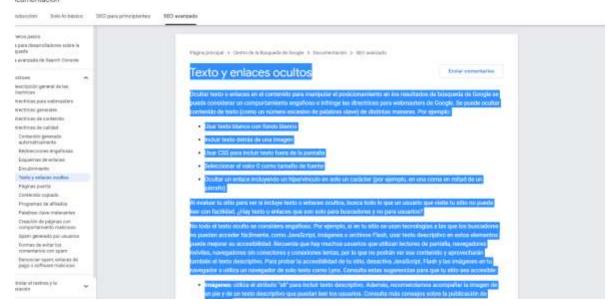
## **Texto y enlaces ocultos**

Ocultar texto o enlaces en el contenido para manipular el posicionamiento en los resultados de búsqueda de Google se puede considerar un comportamiento engañoso e infringe las directrices para webmasters de Google. Se puede ocultar contenido de texto (como un número excesivo de palabras clave) de distintas maneras. Por ejemplo:

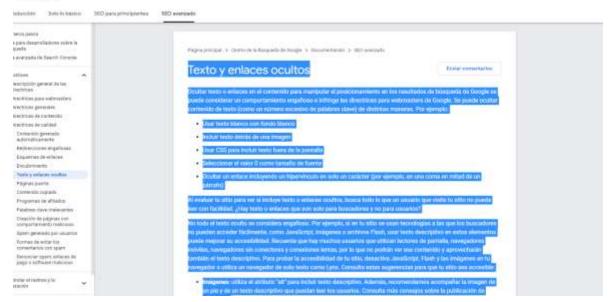
- Usar texto blanco con fondo blanco
- Incluir texto detrás de una imagen
- Usar CSS para incluir texto fuera de la pantalla
- Seleccionar el valor 0 como tamaño de fuente
- Ocultar un enlace incluyendo un hipervínculo en solo un carácter (por ejemplo, en una coma en mitad de un párrafo)

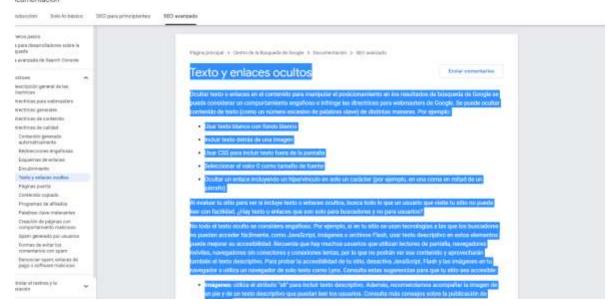
- Imágenes: utiliza el atributo "alt" para incluir texto descriptivo. Además, recomendamos acompañar la imagen de un pie y de un texto descriptivo que puedan leer los usuarios. Consulta más consejos sobre la publicación de imágenes.
- JavaScript: coloca el mismo contenido de JavaScript en una etiqueta <noscript>. Si utilizas este método, debes comprobar que el contenido es exactamente idéntico al de JavaScript y que se muestra a los usuarios que no tienen JavaScript habilitado en su navegador.
- **Vídeos:** incluye un texto descriptivo sobre el vídeo en HTML. También puedes añadir transcripciones. Consulta <u>más consejos sobre la publicación de vídeos</u>.



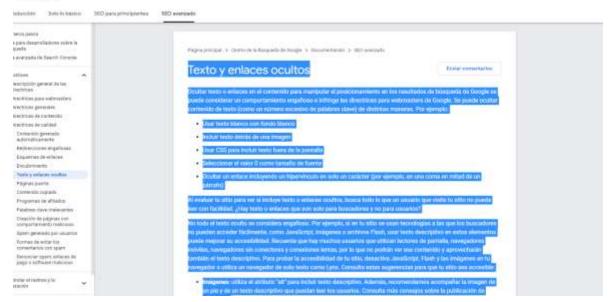


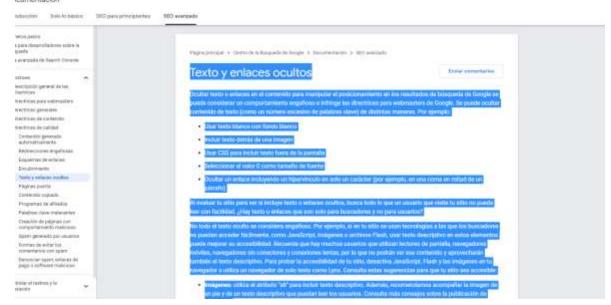
#### cumentacion

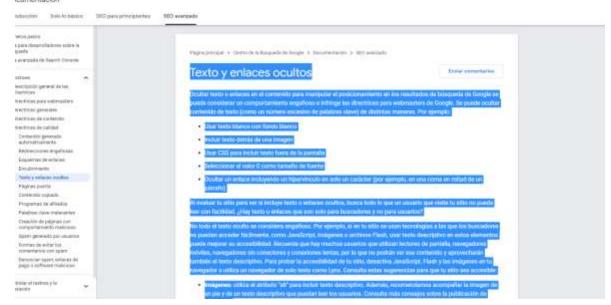


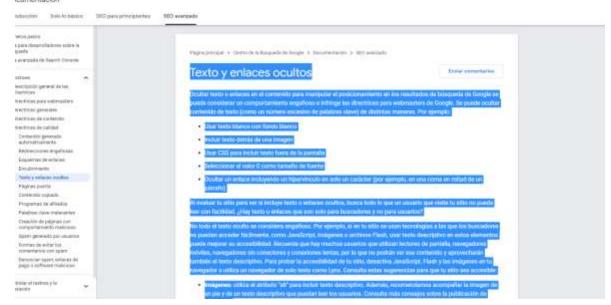


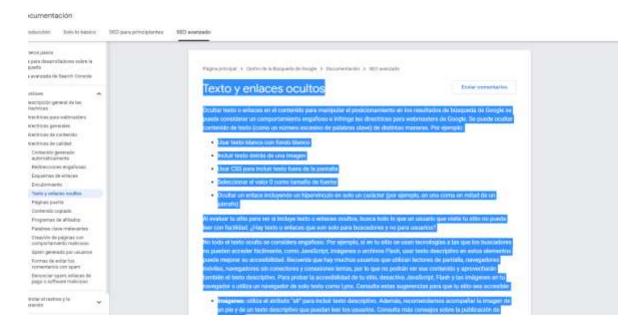
#### cumentacion











### 2. Pasado y presente del Data Warehouse

Históricamente, los data warehouses se habían formado utilizando datos repetitivos estructurados que eran filtrados antes de entrar en el data warehouse. Sin embargo, en los últimos años, el data warehouse ha evolucionado debido a información contextual que ahora se puede adjuntar a los datos no estructurados y que también puede ser almacenada.

Aquellos primeros datos relacionales estructurados no podían ser mezclados y emparejados para temas analíticos con datos textuales no estructurados. Pero con el advenimiento de la contextualización, estos tipos de análisis ahora sí pueden hacerse de forma naturales y fácil.

En el data warehouse, datos no repetitivos, como los comentarios en una encuesta, correos electrónicos y conversaciones, se tratan de forma diferente a las ocurrencias repetitivas de datos, como el flujo de clics, mediciones o el procesamiento máquina o analógico. Los datos no repetitivos son datos basados en textos que fueron generados por la palabra escrita o hablada, leída y reformateada y, lo que es más importante, ahora puede ser contextualizada. Con el fin de extraer cualquier sentido de los datos no repetitivos para su uso en el Data Warehouse, deben tener el contexto de los datos establecidos.

En muchos casos, el contexto de los datos no repetitivos es más importante que los datos en sí. En cualquier caso, los datos no repetitivos no pueden utilizarse para la toma de decisiones hasta que se haya establecido el contexto.

Data Lakes y Data Warehouses: ¿mutuamente exclusivos o partners perfectos?

Los data lakes han surgido en el paisaje de Data Management en los últimos años, sin embargo, data lake no es necesariamente un reemplazo del data warehouse. En cambio, complementan los esfuerzos existentes y dan soporte al descubrimiento de

**nuevas preguntas**. Una vez que se descubren esas preguntas se optimizan las respuestas. Y optimizar puede significar moverse fuera del data lake para ir a un data mart o al data warehouse.

<u>:Resuelve tus dudas sobre el Data Lake</u> con uno de nuestros expertos!

Estas son algunas diferencias clave entre data lake y data warehouse:

- Datos: Un data warehouse sólo almacena datos que han sido modelados o estructurados, mientras que un Data Lake no hace acepción de datos. Lo almacena todo, estructurado, semiestructurado y no estructurado.
- **Procesamiento**: Antes de que una empresa pueda cargar datos en un data warehouse, primero debe darles forma y estructura, es decir, los datos deben ser modelados. Eso se llama schema-on-write. Con un data lake, sólo se cargan los datos sin procesar, tal y como están, y cuando esté listo para usar los datos, es cuando se le da forma y estructura. Eso se llama schema-on-read. Dos enfoques muy diferentes.
- Almacenamiento: Una de las principales características de las tecnologías de big data, como Hadoop, es que el coste de almacenamiento de datos es relativamente bajo en comparación con el de un data warehouse. Hay dos razones principales para esto: en primer lugar, Hadoop es software de código abierto, por lo que la concesión de licencias y el soporte de la comunidad es gratuito. Y segundo, Hadoop está diseñado para ser instalado en hardware de bajo coste.
- Agilidad: Un almacén de datos es un repositorio altamente estructurado, por definición. No
  es técnicamente difícil cambiar la estructura, pero puede tomar mucho tiempo dado todos
  los procesos de negocio que están vinculados a ella. Un data lake, por otro lado, carece de
  la estructura de un data warehouse, lo que da a los desarrolladores y a los científicos de
  datos la capacidad de configurar y reconfigurar fácilmente y en tiempo real sus modelos,
  consultas y aplicaciones.
- Seguridad: La tecnología del data warehouse existe desde hace décadas, mientras que la tecnología de big data (la base de un Data Lake) es relativamente nueva. Por lo tanto, la capacidad de asegurar datos en un data warehouse es mucho más madura que asegurar datos en un data lake. Cabe señalar, sin embargo, que se está realizando un importante esfuerzo en materia de seguridad en la actualidad en la industria de Big Data.

El data warehouse continúa evolucionando

Hay todavía un gran potencial en la evolución del data warehouse y su uso sigue adelante.

El procesamiento analítico clásico de datos basado en transacciones se realiza en el data warehouse como siempre se ha hecho. Nada ha cambiado ahí. Pero **ahora puede hacerse** analítica sobre datos contextualizados, y esa forma de análisis es nueva. La mayoría de las organizaciones, hasta ahora no habían podido basar su toma de decisiones en datos textuales no estructurados. Y ahora hay una nueva forma de análisis posible en un Data

Warehouse: la posibilidad de la mezcla de análisis. El análisis combinado se realiza mediante una combinación de datos transaccionales estructurados y datos contextuales no estructurados.

Existen muchas otras formas de análisis que también son posibles. Estas formas incluyen análisis predictivo y prescriptivo, así como tecnologías de machine learning y otras que están cambiando la forma en que los datos son administrados y analizados. El data warehouse ha sido durante mucho tiempo un elemento básico de la arquitectura de datos empresarial y sigue teniendo tiene un gran futuro en el nuevo mundo de Big Data y la analítica avanzada.

#### 3. Data Warehouse en la nube

Los data warehouses están atravesando actualmente dos transformaciones muy importantes que tienen el potencial de impulsar niveles significativos de innovación empresarial:

- La primera área de transformación es el impulso para aumentar la agilidad general. La gran mayoría de los departamentos de TI están experimentando un rápido aumento de la demanda de datos. Los directivos quieren tener acceso a más y más datos históricos, mientras que al mismo tiempo, los científicos de datos y los analistas de negocios están explorando formas de introducir nuevos flujos de datos en el almacén para enriquecer el análisis existente, así como impulsar nuevas áreas de análisis. Esta rápida expansión de los volúmenes y fuentes de datos significa que los equipos de TI necesitan invertir más tiempo y esfuerzo asegurando que el rendimiento de las consultas permanezca constante y necesitan proporcionar cada vez más entornos para equipos individuales para validar el valor comercial de los nuevos conjuntos de datos.
- La segunda área de transformación gira en torno a la necesidad de mejorar el control de costes. Existe una creciente necesidad de hacer más con cada vez menos recursos, al mismo tiempo que se garantiza que todos los datos sensibles y estratégicos estén completamente asegurados, a lo largo de todo el ciclo de vida, de la manera más rentable.

Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

La nube está demostrando ser un facilitador dominante. Permite a las organizaciones enfrentarse de forma activa a los desafíos que presentan estas dos transformaciones clave.

¿Por qué migrar a la nube?

Las 3 tecnologías que ya tienen y seguirán teniendo un gran impacto en la transformación de negocios son: la nube, datos y análisis, e Internet de las Cosas. Estas tres tecnologías son parte clave del ecosistema del data warehouse actual. Por lo tanto, es posible sacar la conclusión de que el data warehouse en la nube tiene un gran potencial para generar un impacto significativo en los negocios.

La importancia de la nube para el data warehouse está directamente vinculado a tres factores clave:

- Mayor agilidad
- Mejor control de costes
- Co-localización

#### Mejora de la agilidad

Muchos data warehouses están embarcados en una fase de actualización. Con mucho trabajo puesto ahora en torno al big data, las empresas buscan aprovechar los nuevos flujos de datos y nuevos tipos de análisis más ricos para apoyar e impulsar nuevas áreas, tales como:analítica de clientes de 360°, analísis predictivo, detección de fraude, análisis de IoT y el establecimiento de los datos como centro de beneficio. Muchos de estos proyectos requieren el aprovisionamiento de nuevos entornos de hardware y el despliegue de software. Es más rápido, más fácil y más eficiente iniciar estos nuevos proyectos centrados en los datos utilizando los servicios en la nube.

#### Mayor control de costes

Muchos equipos de TI están buscando maneras de consolidar los data marts existentes, cada uno ejecutándose en hardware dedicado o incluso en hardware propietario, en un único entorno integrado. Los servicios en la nube de calidad ofrecen la oportunidad perfecta para iniciar este tipo de proyectos. El cumplimiento legal no puede ser visto como un extra opcional cuando se planifica un traslado a la nube. Los activos de datos deben estar protegidos a lo largo de todo el ciclo de vida. Los servicios en la nube deben facilitar todo esto, siendo más rentables ya que todas las características de seguridad se pueden habilitar de forma predeterminada y mejorada, y actualizada de forma transparente.

#### Co-Localización para una carga más rápida

La mayoría de los data warehouses generan datos directamente desde aplicaciones clave, como entrada de pedidos, ventas, finanzas y fabricación. Por lo tanto, **tiene mucho sentido ubicar conjuntamente el data warehouse junto con los sistemas fuente que ya se estén ejecutándose en la nube**. La ubicación conjunta ofrece una carga de datos más rápida, lo que significa que los usuarios obtienen un acceso más oportuno a sus datos.

Principales ventajas de mover el data warehouse a la nube

Existen **tres ventajas principales** para mover un data warehouse a la nube, y estas están directamente vinculados a los tres controladores clave enumerados anteriormente:

- Más fácil consolidación y racionalización
- Monetización más rápida de los datos en la nube
- La nube ofrece mejor protección

# 4. Data Warehouse vs Big Data vs BI: cuales son las diferencias

En este apartado compararemos brevemente los conceptos Data Warehouse (DW), Big Data y Business Intelligence (BI).

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para tener una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

Se trata de tres conceptos completamente diferentes que tienen en común una nueva manera de lidiar con los datos, siempre teniendo en cuenta la existencia de una gran volumen de información en varios formatos que contribuyen, de forma estructurada o no estructurada, a la toma de decisiones estratégicas. El objetivo final de cualquiera de estas tecnologías es ofrecer una ventaja competitiva a las empresas, pero la forma en que se utiliza es la que marcará la diferencia.

Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

#### Big Data

Llamamos big data a un gran volumen de datos con una variedad, complejidad y velocidad de crecimiento enorme y que además tienen la característica de no ser estructurados. Eso significa que no son relacionales, estando además fuera del entorno corporativo. Es un tipo de tecnología que te permite analizar los datos en tiempo real y puede provenir de diferentes fuentes y formas, tales como mensajería instantánea, redes sociales, registros de grabaciones, imágenes, mensajes de correo electrónico, etc.

# Guía gratuita: "Del Bit... al Big Data", descubre todo lo que necesitas saber sobre Big Data y mucho más.

Para tener una idea de la importancia del big data, este mercado está creciendo cada año, alrededor de un 40%. La paradoja reside en el hecho de que cuanto mayor es la evolución del big data, mayor es la escasez de profesionales cualificados para satisfacer esa demanda. La expectativa es que en unos pocos años, haya un crecimiento aún más significativo para el uso de big data y por lo tanto un aumento en la demanda mano de obra especializada y de partners tecnológicos especializados en esta materia.

#### Data Warehouse

Por otro lado, un data warehouse almacena datos consolidados de diversas fuentes o sistemas de la empresa. Se trata de datos estructurados, que tiene como objetivo principal ser precisos y de alta calidad para de esta forma poder dar soporte a la toma de decisiones de la empresa. Se trata de conseguir tener todos los datos juntos para después poder dividirlos para hacer un análisis de determinados sectores o estrategias.

#### **Business Intelligence**

Un Business Intelligence (BI) es una especie de "cuello de botella" de los datos recogidos del data warehouse, que llegan de forma exacta y útil para ayudar a la toma de decisiones. Business Intelligence transforma los datos en información útil para analizar no sólo los negocios, sino también las principales estrategias corporativas.

# Guía gratuita "Integridad de datos y BI", descubre cómo sacarle el máximo partido a tu proyecto de Integridad de datos

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

### 5. Material complementario

Soluciones y recursos para Data Warehouse

- Informatica Cloud para Microsoft Azure SQL Data Warehouse
- Test Data Warehouse
- MDM y calidad de datos para el data warehouse

#### Guías

- Data Lake: Superando las limitaciones del Data Warehouse
- El data warehouse inteligente

#### Artículos

- ¿Qué es warehouse y qué debe contemplar un data warehouse moderno?
- Beneficios y soluciones para un Data Warehouse en la nube
- Lo que necesitas saber acerca de la arquitectura de un datawarehouse
- Data warehousing: 3 ocasiones de hacer las cosas bien
- ¿Qué es un Data Warehouse y cómo saber cuándo lo necesitas implementar?
- Cómo optimizar la arquitectura de un data warehouse
- Data warehouse concepts: tu glosario de almacén de datos
- Data Warehouse, el primer paso para una solución de BI
- Tutoriales sobre DataWarehouse en pdf
- Cómo gestiona Informatica los nuevos medios de almacenamiento de datos
- ¿Qué es un Data Warehouse para un director de ventas?
- Bl en línea: el reto actual de los data warehouses
- El poder de data warehouse: Informatica PowerCenter

### 2. Pasado y presente del Data Warehouse

Históricamente, los data warehouses se habían formado utilizando datos repetitivos estructurados que eran filtrados antes de entrar en el data warehouse. Sin embargo, en los últimos años, el data warehouse ha evolucionado debido a información contextual que ahora se puede adjuntar a los datos no estructurados y que también puede ser almacenada.

Aquellos primeros datos relacionales estructurados no podían ser mezclados y emparejados para temas analíticos con datos textuales no estructurados. Pero con el advenimiento de la contextualización, estos tipos de análisis ahora sí pueden hacerse de forma naturales y fácil.

En el data warehouse, datos no repetitivos, como los comentarios en una encuesta, correos electrónicos y conversaciones, se tratan de forma diferente a las ocurrencias repetitivas de datos, como el flujo de clics, mediciones o el procesamiento máquina o analógico. Los datos no repetitivos son datos basados en textos que fueron generados por la palabra escrita o hablada, leída y reformateada y, lo que es más importante, ahora puede ser contextualizada. Con el fin de extraer cualquier sentido de los datos no repetitivos para su uso en el Data Warehouse, deben tener el contexto de los datos establecidos.

En muchos casos, el contexto de los datos no repetitivos es más importante que los datos en sí. En cualquier caso, los datos no repetitivos no pueden utilizarse para la toma de decisiones hasta que se haya establecido el contexto.

Data Lakes y Data Warehouses: ¿mutuamente exclusivos o partners perfectos?

Los data lakes han surgido en el paisaje de Data Management en los últimos años, sin embargo, data lake no es necesariamente un reemplazo del data warehouse. En cambio, complementan los esfuerzos existentes y dan soporte al descubrimiento de nuevas preguntas. Una vez que se descubren esas preguntas se optimizan las respuestas. Y optimizar puede significar moverse fuera del data lake para ir a un data mart o al data warehouse.

<u>Resuelve tus dudas sobre el Data Lake</u> <u>con uno de nuestros expertos!</u>

Estas son algunas diferencias clave entre data lake y data warehouse:

- Datos: Un data warehouse sólo almacena datos que han sido modelados o estructurados, mientras que un Data Lake no hace acepción de datos. Lo almacena todo, estructurado, semiestructurado y no estructurado.
- **Procesamiento**: Antes de que una empresa pueda cargar datos en un data warehouse, primero debe darles forma y estructura, es decir, los datos deben ser modelados. Eso se llama schema-on-write. Con un data lake, sólo se cargan los datos sin procesar, tal y como están, y cuando esté listo para usar los datos, es cuando se le da forma y estructura. Eso se llama schema-on-read. Dos enfoques muy diferentes.

- Almacenamiento: Una de las principales características de las tecnologías de big data, como Hadoop, es que el coste de almacenamiento de datos es relativamente bajo en comparación con el de un data warehouse. Hay dos razones principales para esto: en primer lugar, Hadoop es software de código abierto, por lo que la concesión de licencias y el soporte de la comunidad es gratuito. Y segundo, Hadoop está diseñado para ser instalado en hardware de bajo coste.
- Agilidad: Un almacén de datos es un repositorio altamente estructurado, por definición. No
  es técnicamente difícil cambiar la estructura, pero puede tomar mucho tiempo dado todos
  los procesos de negocio que están vinculados a ella. Un data lake, por otro lado, carece de
  la estructura de un data warehouse, lo que da a los desarrolladores y a los científicos de
  datos la capacidad de configurar y reconfigurar fácilmente y en tiempo real sus modelos,
  consultas y aplicaciones.
- Seguridad: La tecnología del data warehouse existe desde hace décadas, mientras que la tecnología de big data (la base de un Data Lake) es relativamente nueva. Por lo tanto, la capacidad de asegurar datos en un data warehouse es mucho más madura que asegurar datos en un data lake. Cabe señalar, sin embargo, que se está realizando un importante esfuerzo en materia de seguridad en la actualidad en la industria de Big Data.

El data warehouse continúa evolucionando

Hay todavía un gran potencial en la evolución del data warehouse y su uso sigue adelante.

El procesamiento analítico clásico de datos basado en transacciones se realiza en el data warehouse como siempre se ha hecho. Nada ha cambiado ahí. Pero ahora puede hacerse analítica sobre datos contextualizados, y esa forma de análisis es nueva. La mayoría de las organizaciones, hasta ahora no habían podido basar su toma de decisiones en datos textuales no estructurados. Y ahora hay una nueva forma de análisis posible en un Data Warehouse: la posibilidad de la mezcla de análisis. El análisis combinado se realiza mediante una combinación de datos transaccionales estructurados y datos contextuales no estructurados.

Existen muchas otras formas de análisis que también son posibles. Estas formas incluyen análisis predictivo y prescriptivo, así como tecnologías de machine learning y otras que están cambiando la forma en que los datos son administrados y analizados. El data warehouse ha sido durante mucho tiempo un elemento básico de la arquitectura de datos empresarial y sigue teniendo tiene un gran futuro en el nuevo mundo de Big Data y la analítica avanzada.

### 3. Data Warehouse en la nube

Los data warehouses están atravesando actualmente dos transformaciones muy importantes que tienen el potencial de impulsar niveles significativos de innovación empresarial:

- La primera área de transformación es el impulso para aumentar la agilidad general. La gran mayoría de los departamentos de TI están experimentando un rápido aumento de la demanda de datos. Los directivos quieren tener acceso a más y más datos históricos, mientras que al mismo tiempo, los científicos de datos y los analistas de negocios están explorando formas de introducir nuevos flujos de datos en el almacén para enriquecer el análisis existente, así como impulsar nuevas áreas de análisis. Esta rápida expansión de los volúmenes y fuentes de datos significa que los equipos de TI necesitan invertir más tiempo y esfuerzo asegurando que el rendimiento de las consultas permanezca constante y necesitan proporcionar cada vez más entornos para equipos individuales para validar el valor comercial de los nuevos conjuntos de datos.
- La segunda área de transformación gira en torno a la necesidad de mejorar el control de costes. Existe una creciente necesidad de hacer más con cada vez menos recursos, al mismo tiempo que se garantiza que todos los datos sensibles y estratégicos estén completamente asegurados, a lo largo de todo el ciclo de vida, de la manera más rentable.

Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

La nube está demostrando ser un facilitador dominante. Permite a las organizaciones enfrentarse de forma activa a los desafíos que presentan estas dos transformaciones clave.

¿Por qué migrar a la nube?

Las 3 tecnologías que ya tienen y seguirán teniendo un gran impacto en la transformación de negocios son: la nube, datos y análisis, e Internet de las Cosas. Estas tres tecnologías son parte clave del ecosistema del data warehouse actual. Por lo tanto, es posible sacar la conclusión de que el data warehouse en la nube tiene un gran potencial para generar un impacto significativo en los negocios.

La importancia de la nube para el data warehouse está directamente vinculado a tres factores clave:

- Mayor agilidad
- Mejor control de costes
- Co-localización

Mejora de la agilidad

Muchos data warehouses están embarcados en una fase de actualización. Con mucho trabajo puesto ahora en torno al big data, las empresas buscan aprovechar los nuevos flujos de datos y nuevos tipos de análisis más ricos para apoyar e impulsar nuevas áreas, tales como:analítica de clientes de 360°, analisis predictivo, detección de fraude, análisis de loT y el establecimiento de los datos como centro de beneficio. Muchos de estos proyectos requieren el aprovisionamiento de nuevos entornos de hardware y el despliegue de software. Es más rápido, más fácil y más eficiente iniciar estos nuevos proyectos centrados en los datos utilizando los servicios en la nube.

#### Mayor control de costes

Muchos equipos de TI están buscando maneras de consolidar los data marts existentes, cada uno ejecutándose en hardware dedicado o incluso en hardware propietario, en un único entorno integrado. Los servicios en la nube de calidad ofrecen la oportunidad perfecta para iniciar este tipo de proyectos. El cumplimiento legal no puede ser visto como un extra opcional cuando se planifica un traslado a la nube. Los activos de datos deben estar protegidos a lo largo de todo el ciclo de vida. Los servicios en la nube deben facilitar todo esto, siendo más rentables ya que todas las características de seguridad se pueden habilitar de forma predeterminada y mejorada, y actualizada de forma transparente.

Co-Localización para una carga más rápida

La mayoría de los data warehouses generan datos directamente desde aplicaciones clave, como entrada de pedidos, ventas, finanzas y fabricación. Por lo tanto, tiene mucho sentido ubicar conjuntamente el data warehouse junto con los sistemas fuente que ya se estén ejecutándose en la nube. La ubicación conjunta ofrece una carga de datos más rápida, lo que significa que los usuarios obtienen un acceso más oportuno a sus datos.

Principales ventajas de mover el data warehouse a la nube

Existen **tres ventajas principales** para mover un data warehouse a la nube, y estas están directamente vinculados a los tres controladores clave enumerados anteriormente:

- Más fácil consolidación y racionalización
- Monetización más rápida de los datos en la nube
- La nube ofrece mejor protección

# 4. Data Warehouse vs Big Data vs BI: cuales son las diferencias

En este apartado compararemos brevemente los conceptos Data Warehouse (DW), Big Data y Business Intelligence (BI).

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para tener una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

Se trata de tres conceptos completamente diferentes que tienen en común una nueva manera de lidiar con los datos, siempre teniendo en cuenta la existencia de una gran volumen de información en varios formatos que contribuyen, de forma estructurada o no estructurada, a la toma de decisiones estratégicas. El objetivo final de cualquiera de estas tecnologías es ofrecer una ventaja competitiva a las empresas, pero la forma en que se utiliza es la que marcará la diferencia.

Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

#### Big Data

Llamamos big data a un gran volumen de datos con una variedad, complejidad y velocidad de crecimiento enorme y que además tienen la característica de no ser estructurados. Eso significa que no son relacionales, estando además fuera del entorno corporativo. Es un tipo de tecnología que te permite analizar los datos en tiempo real y puede provenir de diferentes fuentes y formas, tales como mensajería instantánea, redes sociales, registros de grabaciones, imágenes, mensajes de correo electrónico, etc.

# Guía gratuita: "Del Bit... al Big Data", descubre todo lo que necesitas saber sobre Big Data y mucho más.

Para tener una idea de la importancia del big data, este mercado está creciendo cada año, alrededor de un 40%. La paradoja reside en el hecho de que cuanto mayor es la evolución del big data, mayor es la escasez de profesionales cualificados para satisfacer esa demanda. La expectativa es que en unos pocos años, haya un crecimiento aún más significativo para el uso de big data y por lo tanto un aumento en la demanda mano de obra especializada y de partners tecnológicos especializados en esta materia.

#### Data Warehouse

Por otro lado, un data warehouse almacena datos consolidados de diversas fuentes o sistemas de la empresa. Se trata de datos estructurados, que tiene como objetivo principal ser precisos y de alta calidad para de esta forma poder dar soporte a la toma de decisiones de la empresa. Se trata de conseguir tener todos los datos juntos para después poder dividirlos para hacer un análisis de determinados sectores o estrategias.

#### Business Intelligence

Un Business Intelligence (BI) es una especie de "cuello de botella" de los datos recogidos del data warehouse, que llegan de forma exacta y útil para ayudar a la toma de decisiones. Business Intelligence transforma los datos en información útil para analizar no sólo los negocios, sino también las principales estrategias corporativas.

# Guía gratuita "Integridad de datos y BI", descubre cómo sacarle el máximo partido a tu proyecto de Integridad de datos

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

### 5. Material complementario

Soluciones y recursos para Data Warehouse

Informatica Cloud para Microsoft Azure SQL Data Warehouse

- Test Data Warehouse
- MDM y calidad de datos para el data warehouse

#### Guías

- Data Lake: Superando las limitaciones del Data Warehouse
- El data warehouse inteligente

#### Artículos

- ¿Qué es warehouse y qué debe contemplar un data warehouse moderno?
- Beneficios y soluciones para un Data Warehouse en la nube
- Lo que necesitas saber acerca de la arquitectura de un datawarehouse
- Data warehousing: 3 ocasiones de hacer las cosas bien
- ¿Qué es un Data Warehouse y cómo saber cuándo lo necesitas implementar?
- Cómo optimizar la arquitectura de un data warehouse
- Data warehouse concepts: tu glosario de almacén de datos
- Data Warehouse, el primer paso para una solución de BI
- Tutoriales sobre DataWarehouse en pdf
- Cómo gestiona Informatica los nuevos medios de almacenamiento de datos
- ¿Qué es un Data Warehouse para un director de ventas?
- Bl en línea: el reto actual de los data warehouses
- El poder de data warehouse: Informatica PowerCenter

### 2. Pasado y presente del Data Warehouse

Históricamente, los data warehouses se habían formado utilizando datos repetitivos estructurados que eran filtrados antes de entrar en el data warehouse. Sin embargo, en los últimos años, el data warehouse ha evolucionado debido a información contextual que ahora se puede adjuntar a los datos no estructurados y que también puede ser almacenada.

Aquellos primeros datos relacionales estructurados no podían ser mezclados y emparejados para temas analíticos con datos textuales no estructurados. Pero con el advenimiento de la contextualización, estos tipos de análisis ahora sí pueden hacerse de forma naturales y fácil.

En el data warehouse, datos no repetitivos, como los comentarios en una encuesta, correos electrónicos y conversaciones, se tratan de forma diferente a las ocurrencias repetitivas de datos, como el flujo de clics, mediciones o el procesamiento máquina o analógico. Los datos no repetitivos son datos basados en textos que fueron generados por la palabra escrita o hablada, leída y reformateada y, lo que es más importante, ahora puede ser contextualizada. Con el fin de extraer cualquier sentido de los datos no repetitivos para su uso en el Data Warehouse, deben tener el contexto de los datos establecidos.

En muchos casos, el contexto de los datos no repetitivos es más importante que los datos en sí. En cualquier caso, los datos no repetitivos no pueden utilizarse para la toma de decisiones hasta que se haya establecido el contexto.

Data Lakes y Data Warehouses: ¿mutuamente exclusivos o partners perfectos?

Los data lakes han surgido en el paisaje de Data Management en los últimos años, sin embargo, data lake no es necesariamente un reemplazo del data warehouse. En cambio, complementan los esfuerzos existentes y dan soporte al descubrimiento de nuevas preguntas. Una vez que se descubren esas preguntas se optimizan las respuestas. Y optimizar puede significar moverse fuera del data lake para ir a un data mart o al data warehouse.

¡Resuelve tus dudas sobre el Data Lake con uno de nuestros expertos!

Estas son algunas diferencias clave entre data lake y data warehouse:

- Datos: Un data warehouse sólo almacena datos que han sido modelados o estructurados, mientras que un Data Lake no hace acepción de datos. Lo almacena todo, estructurado, semiestructurado y no estructurado.
- **Procesamiento**: Antes de que una empresa pueda cargar datos en un data warehouse, primero debe darles forma y estructura, es decir, los datos deben ser modelados. Eso se llama schema-on-write. Con un data lake, sólo se cargan los datos sin procesar, tal y como están, y cuando esté listo para usar los datos, es cuando se le da forma y estructura. Eso se llama schema-on-read. Dos enfoques muy diferentes.
- Almacenamiento: Una de las principales características de las tecnologías de big data, como Hadoop, es que el coste de almacenamiento de datos es relativamente bajo en comparación con el de un data warehouse. Hay dos razones principales para esto: en primer lugar, Hadoop es software de código abierto, por lo que la concesión de licencias y el soporte de la comunidad es gratuito. Y segundo, Hadoop está diseñado para ser instalado en hardware de bajo coste.
- Agilidad: Un almacén de datos es un repositorio altamente estructurado, por definición. No
   es técnicamente difícil cambiar la estructura, pero puede tomar mucho tiempo dado todos
   los procesos de negocio que están vinculados a ella. Un data lake, por otro lado, carece de
   la estructura de un data warehouse, lo que da a los desarrolladores y a los científicos de
   datos la capacidad de configurar y reconfigurar fácilmente y en tiempo real sus modelos,
   consultas y aplicaciones.
- Seguridad: La tecnología del data warehouse existe desde hace décadas, mientras que la tecnología de big data (la base de un Data Lake) es relativamente nueva. Por lo tanto, la capacidad de asegurar datos en un data warehouse es mucho más madura que asegurar datos en un data lake. Cabe señalar, sin embargo, que se está realizando un importante esfuerzo en materia de seguridad en la actualidad en la industria de Big Data.

El data warehouse continúa evolucionando

Hay todavía un gran potencial en la evolución del data warehouse y su uso sigue adelante.

El procesamiento analítico clásico de datos basado en transacciones se realiza en el data warehouse como siempre se ha hecho. Nada ha cambiado ahí. Pero ahora puede hacerse analítica sobre datos contextualizados, y esa forma de análisis es nueva. La mayoría de las organizaciones, hasta ahora no habían podido basar su toma de decisiones en datos textuales no estructurados. Y ahora hay una nueva forma de análisis posible en un Data Warehouse: la posibilidad de la mezcla de análisis. El análisis combinado se realiza mediante una combinación de datos transaccionales estructurados y datos contextuales no estructurados.

Existen muchas otras formas de análisis que también son posibles. Estas formas incluyen análisis predictivo y prescriptivo, así como tecnologías de machine learning y otras que están cambiando la forma en que los datos son administrados y analizados. El data warehouse ha sido durante mucho tiempo un elemento básico de la arquitectura de datos empresarial y sigue teniendo tiene un gran futuro en el nuevo mundo de Big Data y la analítica avanzada.

### 3. Data Warehouse en la nube

Los data warehouses están atravesando actualmente dos transformaciones muy importantes que tienen el potencial de impulsar niveles significativos de innovación empresarial:

- La primera área de transformación es el impulso para aumentar la agilidad general. La gran mayoría de los departamentos de TI están experimentando un rápido aumento de la demanda de datos. Los directivos quieren tener acceso a más y más datos históricos, mientras que al mismo tiempo, los científicos de datos y los analistas de negocios están explorando formas de introducir nuevos flujos de datos en el almacén para enriquecer el análisis existente, así como impulsar nuevas áreas de análisis. Esta rápida expansión de los volúmenes y fuentes de datos significa que los equipos de TI necesitan invertir más tiempo y esfuerzo asegurando que el rendimiento de las consultas permanezca constante y necesitan proporcionar cada vez más entornos para equipos individuales para validar el valor comercial de los nuevos conjuntos de datos.
- La segunda área de transformación gira en torno a la necesidad de mejorar el control de costes. Existe una creciente necesidad de hacer más con cada vez menos recursos, al mismo tiempo que se garantiza que todos los datos sensibles y estratégicos estén completamente asegurados, a lo largo de todo el ciclo de vida, de la manera más rentable.

Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

La nube está demostrando ser un facilitador dominante. Permite a las organizaciones enfrentarse de forma activa a los desafíos que presentan estas dos transformaciones clave.

#### ¿Por qué migrar a la nube?

Las 3 tecnologías que ya tienen y seguirán teniendo un gran impacto en la transformación de negocios son: la nube, datos y análisis, e Internet de las Cosas. Estas tres tecnologías son parte clave del ecosistema del data warehouse actual. Por lo tanto, es posible sacar la conclusión de que el data warehouse en la nube tiene un gran potencial para generar un impacto significativo en los negocios.

La importancia de la nube para el data warehouse está directamente vinculado a tres factores clave:

- Mayor agilidad
- Mejor control de costes
- Co-localización

#### Mejora de la agilidad

Muchos data warehouses están embarcados en una fase de actualización. Con mucho trabajo puesto ahora en torno al big data, las empresas buscan aprovechar los nuevos flujos de datos y nuevos tipos de análisis más ricos para apoyar e impulsar nuevas áreas, tales como:analítica de clientes de 360°, analisis predictivo, detección de fraude, análisis de IoT y el establecimiento de los datos como centro de beneficio. Muchos de estos proyectos requieren el aprovisionamiento de nuevos entornos de hardware y el despliegue de software. Es más rápido, más fácil y más eficiente iniciar estos nuevos proyectos centrados en los datos utilizando los servicios en la nube.

#### Mayor control de costes

Muchos equipos de TI están buscando maneras de consolidar los data marts existentes, cada uno ejecutándose en hardware dedicado o incluso en hardware propietario, en un único entorno integrado. Los servicios en la nube de calidad ofrecen la oportunidad perfecta para iniciar este tipo de proyectos. El cumplimiento legal no puede ser visto como un extra opcional cuando se planifica un traslado a la nube. Los activos de datos deben estar protegidos a lo largo de todo el ciclo de vida. Los servicios en la nube deben facilitar todo esto, siendo más rentables ya que todas las características de seguridad se pueden habilitar de forma predeterminada y mejorada, y actualizada de forma transparente.

#### Co-Localización para una carga más rápida

La mayoría de los data warehouses generan datos directamente desde aplicaciones clave, como entrada de pedidos, ventas, finanzas y fabricación. Por lo tanto, **tiene mucho sentido ubicar conjuntamente el data warehouse junto con los sistemas fuente que ya se estén ejecutándose en la nube**. La ubicación conjunta ofrece una carga de datos más rápida, lo que significa que los usuarios obtienen un acceso más oportuno a sus datos.

Principales ventajas de mover el data warehouse a la nube

Existen **tres ventajas principales** para mover un data warehouse a la nube, y estas están directamente vinculados a los tres controladores clave enumerados anteriormente:

- Más fácil consolidación y racionalización
- Monetización más rápida de los datos en la nube
- La nube ofrece mejor protección

# 4. Data Warehouse vs Big Data vs BI: cuales son las diferencias

En este apartado compararemos brevemente los conceptos Data Warehouse (DW), Big Data y Business Intelligence (BI).

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para tener una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

Se trata de tres conceptos completamente diferentes que tienen en común una nueva manera de lidiar con los datos, siempre teniendo en cuenta la existencia de una gran volumen de información en varios formatos que contribuyen, de forma estructurada o no estructurada, a la toma de decisiones estratégicas. El objetivo final de cualquiera de estas tecnologías es ofrecer una ventaja competitiva a las empresas, pero la forma en que se utiliza es la que marcará la diferencia.

Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

Big Data

Llamamos big data a un gran volumen de datos con una variedad, complejidad y velocidad de crecimiento enorme y que además tienen la característica de no ser estructurados. Eso significa que no son relacionales, estando además fuera del entorno corporativo. Es un tipo de tecnología que te permite analizar los datos en tiempo real y puede provenir de diferentes fuentes y formas, tales como mensajería instantánea, redes sociales, registros de grabaciones, imágenes, mensajes de correo electrónico, etc.

# Guía gratuita: "Del Bit... al Big Data", descubre todo lo que necesitas saber sobre Big Data y mucho más.

Para tener una idea de la importancia del big data, este mercado está creciendo cada año, alrededor de un 40%. La paradoja reside en el hecho de que cuanto mayor es la evolución del big data, mayor es la escasez de profesionales cualificados para satisfacer esa demanda. La expectativa es que en unos pocos años, haya un crecimiento aún más significativo para el uso de big data y por lo tanto un aumento en la demanda mano de obra especializada y de partners tecnológicos especializados en esta materia.

#### Data Warehouse

Por otro lado, un data warehouse almacena datos consolidados de diversas fuentes o sistemas de la empresa. Se trata de datos estructurados, que tiene como objetivo principal ser precisos y de alta calidad para de esta forma poder dar soporte a la toma de decisiones de la empresa. Se trata de conseguir tener todos los datos juntos para después poder dividirlos para hacer un análisis de determinados sectores o estrategias.

#### **Business Intelligence**

Un Business Intelligence (BI) es una especie de "cuello de botella" de los datos recogidos del data warehouse, que llegan de forma exacta y útil para ayudar a la toma de decisiones. Business Intelligence transforma los datos en información útil para analizar no sólo los negocios, sino también las principales estrategias corporativas.

### Guía gratuita "Integridad de datos y BI", descubre cómo sacarle el máximo partido a tu proyecto de Integridad de datos

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

### 5. Material complementario

Soluciones y recursos para Data Warehouse

- Informatica Cloud para Microsoft Azure SQL Data Warehouse
- Test Data Warehouse
- MDM y calidad de datos para el data warehouse

#### Guías

- Data Lake: Superando las limitaciones del Data Warehouse
- El data warehouse inteligente

#### Artículos

- ¿Qué es warehouse y qué debe contemplar un data warehouse moderno?
- Beneficios y soluciones para un Data Warehouse en la nube
- Lo que necesitas saber acerca de la arquitectura de un datawarehouse
- Data warehousing: 3 ocasiones de hacer las cosas bien
- ¿Qué es un Data Warehouse y cómo saber cuándo lo necesitas implementar?
- Cómo optimizar la arquitectura de un data warehouse
- Data warehouse concepts: tu glosario de almacén de datos
- Data Warehouse, el primer paso para una solución de BI
- Tutoriales sobre DataWarehouse en pdf
- Cómo gestiona Informatica los nuevos medios de almacenamiento de datos

- ¿Qué es un Data Warehouse para un director de ventas?
- Bl en línea: el reto actual de los data warehouses
- El poder de data warehouse: Informatica PowerCenter

### 2. Pasado y presente del Data Warehouse

Históricamente, los data warehouses se habían formado utilizando datos repetitivos estructurados que eran filtrados antes de entrar en el data warehouse. Sin embargo, en los últimos años, el data warehouse ha evolucionado debido a información contextual que ahora se puede adjuntar a los datos no estructurados y que también puede ser almacenada.

Aquellos primeros datos relacionales estructurados no podían ser mezclados y emparejados para temas analíticos con datos textuales no estructurados. Pero con el advenimiento de la contextualización, estos tipos de análisis ahora sí pueden hacerse de forma naturales y fácil.

En el data warehouse, datos no repetitivos, como los comentarios en una encuesta, correos electrónicos y conversaciones, se tratan de forma diferente a las ocurrencias repetitivas de datos, como el flujo de clics, mediciones o el procesamiento máquina o analógico. Los datos no repetitivos son datos basados en textos que fueron generados por la palabra escrita o hablada, leída y reformateada y, lo que es más importante, ahora puede ser contextualizada. Con el fin de extraer cualquier sentido de los datos no repetitivos para su uso en el Data Warehouse, deben tener el contexto de los datos establecidos.

En muchos casos, el contexto de los datos no repetitivos es más importante que los datos en sí. En cualquier caso, los datos no repetitivos no pueden utilizarse para la toma de decisiones hasta que se haya establecido el contexto.

Data Lakes y Data Warehouses: ¿mutuamente exclusivos o partners perfectos?

Los data lakes han surgido en el paisaje de Data Management en los últimos años, sin embargo, data lake no es necesariamente un reemplazo del data warehouse. En cambio, complementan los esfuerzos existentes y dan soporte al descubrimiento de nuevas preguntas. Una vez que se descubren esas preguntas se optimizan las respuestas. Y optimizar puede significar moverse fuera del data lake para ir a un data mart o al data warehouse.

¡Resuelve tus dudas sobre el Data Lake con uno de nuestros expertos!

Estas son algunas diferencias clave entre data lake y data warehouse:

 Datos: Un data warehouse sólo almacena datos que han sido modelados o estructurados, mientras que un Data Lake no hace acepción de datos. Lo almacena todo, estructurado, semiestructurado y no estructurado.

- **Procesamiento**: Antes de que una empresa pueda cargar datos en un data warehouse, primero debe darles forma y estructura, es decir, los datos deben ser modelados. Eso se llama schema-on-write. Con un data lake, sólo se cargan los datos sin procesar, tal y como están, y cuando esté listo para usar los datos, es cuando se le da forma y estructura. Eso se llama schema-on-read. Dos enfoques muy diferentes.
- Almacenamiento: Una de las principales características de las tecnologías de big data, como Hadoop, es que el coste de almacenamiento de datos es relativamente bajo en comparación con el de un data warehouse. Hay dos razones principales para esto: en primer lugar, Hadoop es software de código abierto, por lo que la concesión de licencias y el soporte de la comunidad es gratuito. Y segundo, Hadoop está diseñado para ser instalado en hardware de bajo coste.
- Agilidad: Un almacén de datos es un repositorio altamente estructurado, por definición. No
  es técnicamente difícil cambiar la estructura, pero puede tomar mucho tiempo dado todos
  los procesos de negocio que están vinculados a ella. Un data lake, por otro lado, carece de
  la estructura de un data warehouse, lo que da a los desarrolladores y a los científicos de
  datos la capacidad de configurar y reconfigurar fácilmente y en tiempo real sus modelos,
  consultas y aplicaciones.
- Seguridad: La tecnología del data warehouse existe desde hace décadas, mientras que la tecnología de big data (la base de un Data Lake) es relativamente nueva. Por lo tanto, la capacidad de asegurar datos en un data warehouse es mucho más madura que asegurar datos en un data lake. Cabe señalar, sin embargo, que se está realizando un importante esfuerzo en materia de seguridad en la actualidad en la industria de Big Data.

El data warehouse continúa evolucionando

Hay todavía un gran potencial en la evolución del data warehouse y su uso sigue adelante.

El procesamiento analítico clásico de datos basado en transacciones se realiza en el data warehouse como siempre se ha hecho. Nada ha cambiado ahí. Pero ahora puede hacerse analítica sobre datos contextualizados, y esa forma de análisis es nueva. La mayoría de las organizaciones, hasta ahora no habían podido basar su toma de decisiones en datos textuales no estructurados. Y ahora hay una nueva forma de análisis posible en un Data Warehouse: la posibilidad de la mezcla de análisis. El análisis combinado se realiza mediante una combinación de datos transaccionales estructurados y datos contextuales no estructurados.

Existen muchas otras formas de análisis que también son posibles. Estas formas incluyen análisis predictivo y prescriptivo, así como tecnologías de machine learning y otras que están cambiando la forma en que los datos son administrados y analizados. El data warehouse ha sido durante mucho tiempo un elemento básico de la arquitectura de datos empresarial y sigue teniendo tiene un gran futuro en el nuevo mundo de Big Data y la analítica avanzada.

#### 3. Data Warehouse en la nube

Los data warehouses están atravesando actualmente dos transformaciones muy importantes que tienen el potencial de impulsar niveles significativos de innovación empresarial:

- La primera área de transformación es el impulso para aumentar la agilidad general. La gran mayoría de los departamentos de TI están experimentando un rápido aumento de la demanda de datos. Los directivos quieren tener acceso a más y más datos históricos, mientras que al mismo tiempo, los científicos de datos y los analistas de negocios están explorando formas de introducir nuevos flujos de datos en el almacén para enriquecer el análisis existente, así como impulsar nuevas áreas de análisis. Esta rápida expansión de los volúmenes y fuentes de datos significa que los equipos de TI necesitan invertir más tiempo y esfuerzo asegurando que el rendimiento de las consultas permanezca constante y necesitan proporcionar cada vez más entornos para equipos individuales para validar el valor comercial de los nuevos conjuntos de datos.
- La segunda área de transformación gira en torno a la necesidad de mejorar el control
  de costes. Existe una creciente necesidad de hacer más con cada vez menos recursos, al
  mismo tiempo que se garantiza que todos los datos sensibles y estratégicos estén
  completamente asegurados, a lo largo de todo el ciclo de vida, de la manera más rentable.

Descubre cómo ampliar tu negocio en la nube. Contacta con un experto en CLOUD!

La nube está demostrando ser un facilitador dominante. Permite a las organizaciones enfrentarse de forma activa a los desafíos que presentan estas dos transformaciones clave.

¿Por qué migrar a la nube?

Las 3 tecnologías que ya tienen y seguirán teniendo un gran impacto en la transformación de negocios son: la nube, datos y análisis, e Internet de las Cosas. Estas tres tecnologías son parte clave del ecosistema del data warehouse actual. Por lo tanto, es posible sacar la conclusión de que el data warehouse en la nube tiene un gran potencial para generar un impacto significativo en los negocios.

La importancia de la nube para el data warehouse está directamente vinculado a tres factores clave:

- Mayor agilidad
- Mejor control de costes
- Co-localización

#### Mejora de la agilidad

Muchos data warehouses están embarcados en una fase de actualización. Con mucho trabajo puesto ahora en torno al big data, las empresas buscan aprovechar los nuevos flujos de datos y nuevos tipos de análisis más ricos para apoyar e impulsar nuevas áreas, tales como:analítica de clientes de 360°, analisis predictivo, detección de fraude, análisis de IoT y el establecimiento de los datos como centro de beneficio. Muchos de estos proyectos requieren el aprovisionamiento de nuevos entornos de hardware y el despliegue de software. Es más rápido, más fácil y más eficiente iniciar estos nuevos proyectos centrados en los datos utilizando los servicios en la nube.

#### Mayor control de costes

Muchos equipos de TI están buscando maneras de consolidar los data marts existentes, cada uno ejecutándose en hardware dedicado o incluso en hardware propietario, en un único entorno integrado. Los servicios en la nube de calidad ofrecen la oportunidad perfecta para iniciar este tipo de proyectos. El cumplimiento legal no puede ser visto como un extra opcional cuando se planifica un traslado a la nube. Los activos de datos deben estar protegidos a lo largo de todo el ciclo de vida. Los servicios en la nube deben facilitar todo esto, siendo más rentables ya que todas las características de seguridad se pueden habilitar de forma predeterminada y mejorada, y actualizada de forma transparente.

#### Co-Localización para una carga más rápida

La mayoría de los data warehouses generan datos directamente desde aplicaciones clave, como entrada de pedidos, ventas, finanzas y fabricación. Por lo tanto, tiene mucho sentido ubicar conjuntamente el data warehouse junto con los sistemas fuente que ya se estén ejecutándose en la nube. La ubicación conjunta ofrece una carga de datos más rápida, lo que significa que los usuarios obtienen un acceso más oportuno a sus datos.

Principales ventajas de mover el data warehouse a la nube

Existen **tres ventajas principales** para mover un data warehouse a la nube, y estas están directamente vinculados a los tres controladores clave enumerados anteriormente:

- Más fácil consolidación y racionalización
- Monetización más rápida de los datos en la nube
- La nube ofrece mejor protección

# 4. Data Warehouse vs Big Data vs BI: cuales son las diferencias

En este apartado compararemos brevemente los conceptos Data Warehouse (DW), Big Data y Business Intelligence (BI).

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para tener una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

Se trata de tres conceptos completamente diferentes que tienen en común una nueva manera de lidiar con los datos, siempre teniendo en cuenta la existencia de una gran volumen de información en varios formatos que contribuyen, de forma estructurada o no estructurada, a la toma de decisiones estratégicas. El objetivo final de cualquiera de estas tecnologías es ofrecer una ventaja competitiva a las empresas, pero la forma en que se utiliza es la que marcará la diferencia.

Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

#### Big Data

Llamamos big data a un gran volumen de datos con una variedad, complejidad y velocidad de crecimiento enorme y que además tienen la característica de no ser estructurados. Eso significa que no son relacionales, estando además fuera del entorno corporativo. Es un tipo de tecnología que te permite analizar los datos en tiempo real y puede provenir de diferentes fuentes y formas, tales como mensajería instantánea, redes sociales, registros de grabaciones, imágenes, mensajes de correo electrónico, etc.

# Guía gratuita: "Del Bit... al Big Data", descubre todo lo que necesitas saber sobre Big Data y mucho más.

Para tener una idea de la importancia del big data, este mercado está creciendo cada año, alrededor de un 40%. La paradoja reside en el hecho de que cuanto mayor es la evolución del big data, mayor es la escasez de profesionales cualificados para satisfacer esa demanda. La expectativa es que en unos pocos años, haya un crecimiento aún más significativo para el uso de big data y por lo tanto un aumento en la demanda mano de obra especializada y de partners tecnológicos especializados en esta materia.

#### Data Warehouse

Por otro lado, un data warehouse almacena datos consolidados de diversas fuentes o sistemas de la empresa. Se trata de datos estructurados, que tiene como objetivo principal ser precisos y de alta calidad para de esta forma poder dar soporte a la toma de decisiones de la empresa. Se trata de conseguir tener todos los datos juntos para después poder dividirlos para hacer un análisis de determinados sectores o estrategias.

#### **Business Intelligence**

Un Business Intelligence (BI) es una especie de "cuello de botella" de los datos recogidos del data warehouse, que llegan de forma exacta y útil para ayudar a la toma de decisiones. Business Intelligence transforma los datos en información útil para analizar no sólo los negocios, sino también las principales estrategias corporativas.

Guía gratuita "Integridad de datos y BI", descubre cómo sacarle el máximo partido a tu proyecto de Integridad de datos

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

### 5. Material complementario

Soluciones y recursos para Data Warehouse

- Informatica Cloud para Microsoft Azure SQL Data Warehouse
- Test Data Warehouse
- MDM y calidad de datos para el data warehouse

#### Guías

- Data Lake: Superando las limitaciones del Data Warehouse
- El data warehouse inteligente

#### Artículos

- ¿Qué es warehouse y qué debe contemplar un data warehouse moderno?
- Beneficios y soluciones para un Data Warehouse en la nube
- Lo que necesitas saber acerca de la arquitectura de un datawarehouse
- Data warehousing: 3 ocasiones de hacer las cosas bien
- ¿Qué es un Data Warehouse y cómo saber cuándo lo necesitas implementar?
- Cómo optimizar la arquitectura de un data warehouse
- Data warehouse concepts: tu glosario de almacén de datos
- Data Warehouse, el primer paso para una solución de BI
- Tutoriales sobre DataWarehouse en pdf
- Cómo gestiona Informatica los nuevos medios de almacenamiento de datos
- ¿Qué es un Data Warehouse para un director de ventas?
- Bl en línea: el reto actual de los data warehouses
- El poder de data warehouse: Informatica PowerCenter

### 2. Pasado y presente del Data Warehouse

Históricamente, los data warehouses se habían formado utilizando datos repetitivos estructurados que eran filtrados antes de entrar en el data warehouse. Sin embargo, en los últimos años, el data warehouse ha evolucionado debido a información contextual que ahora se puede adjuntar a los datos no estructurados y que también puede ser almacenada.

Aquellos primeros datos relacionales estructurados no podían ser mezclados y emparejados para temas analíticos con datos textuales no estructurados. Pero con el advenimiento de la contextualización, estos tipos de análisis ahora sí pueden hacerse de forma naturales y fácil.

En el data warehouse, datos no repetitivos, como los comentarios en una encuesta, correos electrónicos y conversaciones, se tratan de forma diferente a las ocurrencias repetitivas de datos, como el flujo de clics, mediciones o el procesamiento máquina o analógico. Los datos no repetitivos son datos basados en textos que fueron generados por la palabra escrita o hablada, leída y reformateada y, lo que es más importante, ahora puede ser contextualizada. Con el fin de extraer cualquier sentido de los datos no repetitivos para su uso en el Data Warehouse, deben tener el contexto de los datos establecidos.

En muchos casos, el contexto de los datos no repetitivos es más importante que los datos en sí. En cualquier caso, los datos no repetitivos no pueden utilizarse para la toma de decisiones hasta que se haya establecido el contexto.

Data Lakes y Data Warehouses: ¿mutuamente exclusivos o partners perfectos?

Los data lakes han surgido en el paisaje de Data Management en los últimos años, sin embargo, data lake no es necesariamente un reemplazo del data warehouse. En cambio, complementan los esfuerzos existentes y dan soporte al descubrimiento de nuevas preguntas. Una vez que se descubren esas preguntas se optimizan las respuestas. Y optimizar puede significar moverse fuera del data lake para ir a un data mart o al data warehouse.

<u>:Resuelve tus dudas sobre el Data Lake</u> con uno de nuestros expertos!

Estas son algunas diferencias clave entre data lake y data warehouse:

- Datos: Un data warehouse sólo almacena datos que han sido modelados o estructurados, mientras que un Data Lake no hace acepción de datos. Lo almacena todo, estructurado, semiestructurado y no estructurado.
- **Procesamiento**: Antes de que una empresa pueda cargar datos en un data warehouse, primero debe darles forma y estructura, es decir, los datos deben ser modelados. Eso se llama schema-on-write. Con un data lake, sólo se cargan los datos sin procesar, tal y como están, y cuando esté listo para usar los datos, es cuando se le da forma y estructura. Eso se llama schema-on-read. Dos enfoques muy diferentes.
- Almacenamiento: Una de las principales características de las tecnologías de big data, como Hadoop, es que el coste de almacenamiento de datos es relativamente bajo en comparación con el de un data warehouse. Hay dos razones principales para esto: en primer lugar, Hadoop es software de código abierto, por lo que la concesión de licencias y el soporte de la comunidad es gratuito. Y segundo, Hadoop está diseñado para ser instalado en hardware de bajo coste.
- Agilidad: Un almacén de datos es un repositorio altamente estructurado, por definición. No
  es técnicamente difícil cambiar la estructura, pero puede tomar mucho tiempo dado todos
  los procesos de negocio que están vinculados a ella. Un data lake, por otro lado, carece de
  la estructura de un data warehouse, lo que da a los desarrolladores y a los científicos de
  datos la capacidad de configurar y reconfigurar fácilmente y en tiempo real sus modelos,
  consultas y aplicaciones.
- **Seguridad**: La tecnología del data warehouse existe desde hace décadas, mientras que la tecnología de big data (la base de un Data Lake) es relativamente nueva. Por lo tanto, la

capacidad de asegurar datos en un data warehouse es mucho más madura que asegurar datos en un data lake. Cabe señalar, sin embargo, que se está realizando un importante esfuerzo en materia de seguridad en la actualidad en la industria de Big Data.

El data warehouse continúa evolucionando

Hay todavía un gran potencial en la evolución del data warehouse y su uso sigue adelante.

El procesamiento analítico clásico de datos basado en transacciones se realiza en el data warehouse como siempre se ha hecho. Nada ha cambiado ahí. Pero ahora puede hacerse analítica sobre datos contextualizados, y esa forma de análisis es nueva. La mayoría de las organizaciones, hasta ahora no habían podido basar su toma de decisiones en datos textuales no estructurados. Y ahora hay una nueva forma de análisis posible en un Data Warehouse: la posibilidad de la mezcla de análisis. El análisis combinado se realiza mediante una combinación de datos transaccionales estructurados y datos contextuales no estructurados.

Existen muchas otras formas de análisis que también son posibles. Estas formas incluyen análisis predictivo y prescriptivo, así como tecnologías de machine learning y otras que están cambiando la forma en que los datos son administrados y analizados. El data warehouse ha sido durante mucho tiempo un elemento básico de la arquitectura de datos empresarial y sigue teniendo tiene un gran futuro en el nuevo mundo de Big Data y la analítica avanzada.

### 3. Data Warehouse en la nube

Los data warehouses están atravesando actualmente dos transformaciones muy importantes que tienen el potencial de impulsar niveles significativos de innovación empresarial:

- La primera área de transformación es el impulso para aumentar la agilidad general. La gran mayoría de los departamentos de TI están experimentando un rápido aumento de la demanda de datos. Los directivos quieren tener acceso a más y más datos históricos, mientras que al mismo tiempo, los científicos de datos y los analistas de negocios están explorando formas de introducir nuevos flujos de datos en el almacén para enriquecer el análisis existente, así como impulsar nuevas áreas de análisis. Esta rápida expansión de los volúmenes y fuentes de datos significa que los equipos de TI necesitan invertir más tiempo y esfuerzo asegurando que el rendimiento de las consultas permanezca constante y necesitan proporcionar cada vez más entornos para equipos individuales para validar el valor comercial de los nuevos conjuntos de datos.
- La segunda área de transformación gira en torno a la necesidad de mejorar el control de costes. Existe una creciente necesidad de hacer más con cada vez menos recursos, al mismo tiempo que se garantiza que todos los datos sensibles y estratégicos estén completamente asegurados, a lo largo de todo el ciclo de vida, de la manera más rentable.

### Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

La nube está demostrando ser un facilitador dominante. Permite a las organizaciones enfrentarse de forma activa a los desafíos que presentan estas dos transformaciones clave.

¿Por qué migrar a la nube?

Las 3 tecnologías que ya tienen y seguirán teniendo un gran impacto en la transformación de negocios son: la nube, datos y análisis, e Internet de las Cosas. Estas tres tecnologías son parte clave del ecosistema del data warehouse actual. Por lo tanto, es posible sacar la conclusión de que el data warehouse en la nube tiene un gran potencial para generar un impacto significativo en los negocios.

La importancia de la nube para el data warehouse está directamente vinculado a tres factores clave:

- Mayor agilidad
- Mejor control de costes
- Co-localización

Mejora de la agilidad

Muchos data warehouses están embarcados en una fase de actualización. Con mucho trabajo puesto ahora en torno al big data, las empresas buscan aprovechar los nuevos flujos de datos y nuevos tipos de análisis más ricos para apoyar e impulsar nuevas áreas, tales como:analítica de clientes de 360°, analisis predictivo, detección de fraude, análisis de IoT y el establecimiento de los datos como centro de beneficio. Muchos de estos proyectos requieren el aprovisionamiento de nuevos entornos de hardware y el despliegue de software. Es más rápido, más fácil y más eficiente iniciar estos nuevos proyectos centrados en los datos utilizando los servicios en la nube.

Mayor control de costes

Muchos equipos de TI están buscando maneras de consolidar los data marts existentes, cada uno ejecutándose en hardware dedicado o incluso en hardware propietario, en un único entorno integrado. Los servicios en la nube de calidad ofrecen la oportunidad perfecta para iniciar este tipo de proyectos. El cumplimiento legal no puede ser visto como un extra opcional cuando se planifica un traslado a la nube. Los activos de datos deben estar protegidos a lo largo de todo el ciclo de vida. Los servicios en la nube deben facilitar todo esto, siendo más rentables ya que todas las características de seguridad se pueden habilitar de forma predeterminada y mejorada, y actualizada de forma transparente.

#### Co-Localización para una carga más rápida

La mayoría de los data warehouses generan datos directamente desde aplicaciones clave, como entrada de pedidos, ventas, finanzas y fabricación. Por lo tanto, **tiene mucho sentido ubicar conjuntamente el data warehouse junto con los sistemas fuente que ya se estén ejecutándose en la nube**. La ubicación conjunta ofrece una carga de datos más rápida, lo que significa que los usuarios obtienen un acceso más oportuno a sus datos.

Principales ventajas de mover el data warehouse a la nube

Existen **tres ventajas principales** para mover un data warehouse a la nube, y estas están directamente vinculados a los tres controladores clave enumerados anteriormente:

- Más fácil consolidación y racionalización
- Monetización más rápida de los datos en la nube
- La nube ofrece mejor protección

## 4. Data Warehouse vs Big Data vs BI: cuales son las diferencias

En este apartado compararemos brevemente los conceptos Data Warehouse (DW), Big Data y Business Intelligence (BI).

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para tener una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

Se trata de tres conceptos completamente diferentes que tienen en común una nueva manera de lidiar con los datos, siempre teniendo en cuenta la existencia de una gran volumen de información en varios formatos que contribuyen, de forma estructurada o no estructurada, a la toma de decisiones estratégicas. El objetivo final de cualquiera de estas tecnologías es ofrecer una ventaja competitiva a las empresas, pero la forma en que se utiliza es la que marcará la diferencia.

Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

#### Big Data

Llamamos big data a un gran volumen de datos con una variedad, complejidad y velocidad de crecimiento enorme y que además tienen la característica de no ser estructurados. Eso significa que no son relacionales, estando además fuera del entorno corporativo. Es un tipo de tecnología que te permite analizar los datos en tiempo real y puede provenir de diferentes fuentes y formas, tales como mensajería instantánea, redes sociales, registros de grabaciones, imágenes, mensajes de correo electrónico, etc.

## Guía gratuita: "Del Bit... al Big Data", descubre todo lo que necesitas saber sobre Big Data y mucho más.

Para tener una idea de la importancia del big data, este mercado está creciendo cada año, alrededor de un 40%. La paradoja reside en el hecho de que cuanto mayor es la evolución del big data, mayor es la escasez de profesionales cualificados para satisfacer esa demanda. La expectativa es que en unos pocos años, haya un crecimiento aún más significativo para el uso de big data y por lo tanto un aumento en la demanda mano de obra especializada y de partners tecnológicos especializados en esta materia.

#### Data Warehouse

Por otro lado, un data warehouse almacena datos consolidados de diversas fuentes o sistemas de la empresa. Se trata de datos estructurados, que tiene como objetivo principal ser precisos y de alta calidad para de esta forma poder dar soporte a la toma de decisiones de la empresa. Se trata de conseguir tener todos los datos juntos para después poder dividirlos para hacer un análisis de determinados sectores o estrategias.

#### Business Intelligence

Un Business Intelligence (BI) es una especie de "cuello de botella" de los datos recogidos del data warehouse, que llegan de forma exacta y útil para ayudar a la toma de decisiones. Business Intelligence transforma los datos en información útil para analizar no sólo los negocios, sino también las principales estrategias corporativas.

## Guía gratuita "Integridad de datos y BI", descubre cómo sacarle el máximo partido a tu proyecto de Integridad de datos

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

## 5. Material complementario

Soluciones y recursos para Data Warehouse

- Informatica Cloud para Microsoft Azure SQL Data Warehouse
- Test Data Warehouse
- MDM y calidad de datos para el data warehouse

#### Guías

- Data Lake: Superando las limitaciones del Data Warehouse
- El data warehouse inteligente

#### Artículos

- ¿Qué es warehouse y qué debe contemplar un data warehouse moderno?
- Beneficios y soluciones para un Data Warehouse en la nube
- Lo que necesitas saber acerca de la arquitectura de un datawarehouse
- Data warehousing: 3 ocasiones de hacer las cosas bien
- ¿Qué es un Data Warehouse y cómo saber cuándo lo necesitas implementar?
- Cómo optimizar la arquitectura de un data warehouse
- Data warehouse concepts: tu glosario de almacén de datos
- Data Warehouse, el primer paso para una solución de BI
- Tutoriales sobre DataWarehouse en pdf
- Cómo gestiona Informatica los nuevos medios de almacenamiento de datos
- ¿Qué es un Data Warehouse para un director de ventas?
- Bl en línea: el reto actual de los data warehouses
- El poder de data warehouse: Informatica PowerCenter

## 2. Pasado y presente del Data Warehouse

Históricamente, los data warehouses se habían formado utilizando datos repetitivos estructurados que eran filtrados antes de entrar en el data warehouse. Sin embargo, en los últimos años, el data warehouse ha evolucionado debido a información contextual que ahora se puede adjuntar a los datos no estructurados y que también puede ser almacenada.

Aquellos primeros datos relacionales estructurados no podían ser mezclados y emparejados para temas analíticos con datos textuales no estructurados. Pero con el advenimiento de la contextualización, estos tipos de análisis ahora sí pueden hacerse de forma naturales y fácil.

En el data warehouse, datos no repetitivos, como los comentarios en una encuesta, correos electrónicos y conversaciones, se tratan de forma diferente a las ocurrencias repetitivas de datos, como el flujo de clics, mediciones o el procesamiento máquina o analógico. Los datos no repetitivos son datos basados en textos que fueron generados por la palabra escrita o hablada, leída y reformateada y, lo que es más importante, ahora puede ser contextualizada. Con el fin de extraer cualquier sentido de los datos no repetitivos para su uso en el Data Warehouse, deben tener el contexto de los datos establecidos.

En muchos casos, el contexto de los datos no repetitivos es más importante que los datos en sí. En cualquier caso, los datos no repetitivos no pueden utilizarse para la toma de decisiones hasta que se haya establecido el contexto.

Data Lakes y Data Warehouses: ¿mutuamente exclusivos o partners perfectos?

Los data lakes han surgido en el paisaje de Data Management en los últimos años, sin embargo, data lake no es necesariamente un reemplazo del data warehouse. En cambio, complementan los esfuerzos existentes y dan soporte al descubrimiento de nuevas preguntas. Una vez que se descubren esas preguntas se optimizan las respuestas. Y optimizar puede significar moverse fuera del data lake para ir a un data mart o al data warehouse.

## ¡Resuelve tus dudas sobre el Data Lake con uno de nuestros expertos!

Estas son algunas diferencias clave entre data lake y data warehouse:

- Datos: Un data warehouse sólo almacena datos que han sido modelados o estructurados, mientras que un Data Lake no hace acepción de datos. Lo almacena todo, estructurado, semiestructurado y no estructurado.
- **Procesamiento**: Antes de que una empresa pueda cargar datos en un data warehouse, primero debe darles forma y estructura, es decir, los datos deben ser modelados. Eso se llama schema-on-write. Con un data lake, sólo se cargan los datos sin procesar, tal y como están, y cuando esté listo para usar los datos, es cuando se le da forma y estructura. Eso se llama schema-on-read. Dos enfoques muy diferentes.
- Almacenamiento: Una de las principales características de las tecnologías de big data, como Hadoop, es que el coste de almacenamiento de datos es relativamente bajo en comparación con el de un data warehouse. Hay dos razones principales para esto: en primer lugar, Hadoop es software de código abierto, por lo que la concesión de licencias y el soporte de la comunidad es gratuito. Y segundo, Hadoop está diseñado para ser instalado en hardware de bajo coste.
- Agilidad: Un almacén de datos es un repositorio altamente estructurado, por definición. No
  es técnicamente difícil cambiar la estructura, pero puede tomar mucho tiempo dado todos
  los procesos de negocio que están vinculados a ella. Un data lake, por otro lado, carece de
  la estructura de un data warehouse, lo que da a los desarrolladores y a los científicos de
  datos la capacidad de configurar y reconfigurar fácilmente y en tiempo real sus modelos,
  consultas y aplicaciones.
- Seguridad: La tecnología del data warehouse existe desde hace décadas, mientras que la tecnología de big data (la base de un Data Lake) es relativamente nueva. Por lo tanto, la capacidad de asegurar datos en un data warehouse es mucho más madura que asegurar datos en un data lake. Cabe señalar, sin embargo, que se está realizando un importante esfuerzo en materia de seguridad en la actualidad en la industria de Big Data.

El data warehouse continúa evolucionando

Hay todavía un gran potencial en la evolución del data warehouse y su uso sigue adelante.

El procesamiento analítico clásico de datos basado en transacciones se realiza en el data warehouse como siempre se ha hecho. Nada ha cambiado ahí. Pero ahora puede hacerse analítica sobre datos contextualizados, y esa forma de análisis es nueva. La mayoría de las organizaciones, hasta ahora no habían podido basar su toma de decisiones en datos textuales no estructurados. Y ahora hay una nueva forma de análisis posible en un Data Warehouse: la posibilidad de la mezcla de análisis. El análisis combinado se realiza mediante una combinación de datos transaccionales estructurados y datos contextuales no estructurados.

Existen muchas otras formas de análisis que también son posibles. Estas formas incluyen análisis predictivo y prescriptivo, así como tecnologías de machine learning y otras que están cambiando la forma en que los datos son administrados y analizados. El data warehouse ha sido durante mucho tiempo un elemento básico de la arquitectura de datos empresarial y sigue teniendo tiene un gran futuro en el nuevo mundo de Big Data y la analítica avanzada.

#### 3. Data Warehouse en la nube

Los data warehouses están atravesando actualmente dos transformaciones muy importantes que tienen el potencial de impulsar niveles significativos de innovación empresarial:

- La primera área de transformación es el impulso para aumentar la agilidad general. La gran mayoría de los departamentos de TI están experimentando un rápido aumento de la demanda de datos. Los directivos quieren tener acceso a más y más datos históricos, mientras que al mismo tiempo, los científicos de datos y los analistas de negocios están explorando formas de introducir nuevos flujos de datos en el almacén para enriquecer el análisis existente, así como impulsar nuevas áreas de análisis. Esta rápida expansión de los volúmenes y fuentes de datos significa que los equipos de TI necesitan invertir más tiempo y esfuerzo asegurando que el rendimiento de las consultas permanezca constante y necesitan proporcionar cada vez más entornos para equipos individuales para validar el valor comercial de los nuevos conjuntos de datos.
- La segunda área de transformación gira en torno a la necesidad de mejorar el control de costes. Existe una creciente necesidad de hacer más con cada vez menos recursos, al mismo tiempo que se garantiza que todos los datos sensibles y estratégicos estén completamente asegurados, a lo largo de todo el ciclo de vida, de la manera más rentable.

Descubre cómo ampliar tu negocio en la nube. Contacta con un experto en CLOUD!

La nube está demostrando ser un facilitador dominante. Permite a las organizaciones enfrentarse de forma activa a los desafíos que presentan estas dos transformaciones clave.

¿Por qué migrar a la nube?

Las 3 tecnologías que ya tienen y seguirán teniendo un gran impacto en la transformación de negocios son: la nube, datos y análisis, e Internet de las Cosas. Estas tres tecnologías son parte clave del ecosistema del data warehouse actual. Por lo tanto, es posible sacar la conclusión de que el data warehouse en la nube tiene un gran potencial para generar un impacto significativo en los negocios.

La importancia de la nube para el data warehouse está directamente vinculado a tres factores clave:

- Mayor agilidad
- Mejor control de costes
- Co-localización

#### Mejora de la agilidad

Muchos data warehouses están embarcados en una fase de actualización. Con mucho trabajo puesto ahora en torno al big data, las empresas buscan aprovechar los nuevos flujos de datos y nuevos tipos de análisis más ricos para apoyar e impulsar nuevas áreas, tales como:analítica de clientes de 360°, analísis predictivo, detección de fraude, análisis de IoT y el establecimiento de los datos como centro de beneficio. Muchos de estos proyectos requieren el aprovisionamiento de nuevos entornos de hardware y el despliegue de software. Es más rápido, más fácil y más eficiente iniciar estos nuevos proyectos centrados en los datos utilizando los servicios en la nube.

#### Mayor control de costes

Muchos equipos de TI están buscando maneras de consolidar los data marts existentes, cada uno ejecutándose en hardware dedicado o incluso en hardware propietario, en un único entorno integrado. Los servicios en la nube de calidad ofrecen la oportunidad perfecta para iniciar este tipo de proyectos. El cumplimiento legal no puede ser visto como un extra opcional cuando se planifica un traslado a la nube. Los activos de datos deben estar protegidos a lo largo de todo el ciclo de vida. Los servicios en la nube deben facilitar todo esto, siendo más rentables ya que todas las características de seguridad se pueden habilitar de forma predeterminada y mejorada, y actualizada de forma transparente.

#### Co-Localización para una carga más rápida

La mayoría de los data warehouses generan datos directamente desde aplicaciones clave, como entrada de pedidos, ventas, finanzas y fabricación. Por lo tanto, **tiene mucho sentido ubicar conjuntamente el data warehouse junto con los sistemas fuente que ya se estén ejecutándose en la nube**. La ubicación conjunta ofrece una carga de datos más rápida, lo que significa que los usuarios obtienen un acceso más oportuno a sus datos.

Principales ventajas de mover el data warehouse a la nube

Existen **tres ventajas principales** para mover un data warehouse a la nube, y estas están directamente vinculados a los tres controladores clave enumerados anteriormente:

- Más fácil consolidación y racionalización
- Monetización más rápida de los datos en la nube
- La nube ofrece mejor protección

## 4. Data Warehouse vs Big Data vs BI: cuales son las diferencias

En este apartado compararemos brevemente los conceptos Data Warehouse (DW), Big Data y Business Intelligence (BI).

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para tener una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

Se trata de tres conceptos completamente diferentes que tienen en común una nueva manera de lidiar con los datos, siempre teniendo en cuenta la existencia de una gran volumen de información en varios formatos que contribuyen, de forma estructurada o no estructurada, a la toma de decisiones estratégicas. El objetivo final de cualquiera de estas tecnologías es ofrecer una ventaja competitiva a las empresas, pero la forma en que se utiliza es la que marcará la diferencia.

Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

#### Big Data

Llamamos big data a un gran volumen de datos con una variedad, complejidad y velocidad de crecimiento enorme y que además tienen la característica de no ser estructurados. Eso significa que no son relacionales, estando además fuera del entorno corporativo. Es un tipo de tecnología que te permite analizar los datos en tiempo real y puede provenir de diferentes fuentes y formas, tales como mensajería instantánea, redes sociales, registros de grabaciones, imágenes, mensajes de correo electrónico, etc.

## Guía gratuita: "Del Bit... al Big Data", descubre todo lo que necesitas saber sobre Big Data y mucho más.

Para tener una idea de la importancia del big data, este mercado está creciendo cada año, alrededor de un 40%. La paradoja reside en el hecho de que cuanto mayor es la evolución del big data, mayor es la escasez de profesionales cualificados para satisfacer esa demanda. La expectativa es que en unos pocos años, haya un crecimiento aún más significativo para el uso de big data y por lo tanto un aumento en la demanda mano de obra especializada y de partners tecnológicos especializados en esta materia.

#### Data Warehouse

Por otro lado, un data warehouse almacena datos consolidados de diversas fuentes o sistemas de la empresa. Se trata de datos estructurados, que tiene como objetivo principal ser precisos y de alta calidad para de esta forma poder dar soporte a la toma de decisiones de la empresa. Se trata de conseguir tener todos los datos juntos para después poder dividirlos para hacer un análisis de determinados sectores o estrategias.

#### **Business Intelligence**

Un Business Intelligence (BI) es una especie de "cuello de botella" de los datos recogidos del data warehouse, que llegan de forma exacta y útil para ayudar a la toma de decisiones. Business Intelligence transforma los datos en información útil para analizar no sólo los negocios, sino también las principales estrategias corporativas.

## Guía gratuita "Integridad de datos y BI", descubre cómo sacarle el máximo partido a tu proyecto de Integridad de datos

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

## 5. Material complementario

Soluciones y recursos para Data Warehouse

- Informatica Cloud para Microsoft Azure SQL Data Warehouse
- Test Data Warehouse
- MDM y calidad de datos para el data warehouse

#### Guías

- Data Lake: Superando las limitaciones del Data Warehouse
- El data warehouse inteligente

#### Artículos

- ¿Qué es warehouse y qué debe contemplar un data warehouse moderno?
- Beneficios y soluciones para un Data Warehouse en la nube
- Lo que necesitas saber acerca de la arquitectura de un datawarehouse
- Data warehousing: 3 ocasiones de hacer las cosas bien
- ¿Qué es un Data Warehouse y cómo saber cuándo lo necesitas implementar?
- Cómo optimizar la arquitectura de un data warehouse
- Data warehouse concepts: tu glosario de almacén de datos
- Data Warehouse, el primer paso para una solución de BI
- Tutoriales sobre DataWarehouse en pdf
- Cómo gestiona Informatica los nuevos medios de almacenamiento de datos
- ¿Qué es un Data Warehouse para un director de ventas?
- Bl en línea: el reto actual de los data warehouses
- El poder de data warehouse: Informatica PowerCenter

### 2. Pasado y presente del Data Warehouse

Históricamente, los data warehouses se habían formado utilizando datos repetitivos estructurados que eran filtrados antes de entrar en el data warehouse. Sin embargo, en los últimos años, el data warehouse ha evolucionado debido a información contextual que ahora se puede adjuntar a los datos no estructurados y que también puede ser almacenada.

Aquellos primeros datos relacionales estructurados no podían ser mezclados y emparejados para temas analíticos con datos textuales no estructurados. Pero con el advenimiento de la contextualización, estos tipos de análisis ahora sí pueden hacerse de forma naturales y fácil.

En el data warehouse, datos no repetitivos, como los comentarios en una encuesta, correos electrónicos y conversaciones, se tratan de forma diferente a las ocurrencias repetitivas de datos, como el flujo de clics, mediciones o el procesamiento máquina o analógico. Los datos no repetitivos son datos basados en textos que fueron generados por la palabra escrita o hablada, leída y reformateada y, lo que es más importante, ahora puede ser contextualizada. Con el fin de extraer cualquier sentido de los datos no repetitivos para su uso en el Data Warehouse, deben tener el contexto de los datos establecidos.

En muchos casos, el contexto de los datos no repetitivos es más importante que los datos en sí. En cualquier caso, los datos no repetitivos no pueden utilizarse para la toma de decisiones hasta que se haya establecido el contexto.

Data Lakes y Data Warehouses: ¿mutuamente exclusivos o partners perfectos?

Los data lakes han surgido en el paisaje de Data Management en los últimos años, sin embargo, data lake no es necesariamente un reemplazo del data warehouse. En cambio, complementan los esfuerzos existentes y dan soporte al descubrimiento de nuevas preguntas. Una vez que se descubren esas preguntas se optimizan las respuestas. Y optimizar puede significar moverse fuera del data lake para ir a un data mart o al data warehouse.

¡Resuelve tus dudas sobre el Data Lake con uno de nuestros expertos!

Estas son algunas diferencias clave entre data lake y data warehouse:

- Datos: Un data warehouse sólo almacena datos que han sido modelados o estructurados, mientras que un Data Lake no hace acepción de datos. Lo almacena todo, estructurado, semiestructurado y no estructurado.
- **Procesamiento**: Antes de que una empresa pueda cargar datos en un data warehouse, primero debe darles forma y estructura, es decir, los datos deben ser modelados. Eso se llama schema-on-write. Con un data lake, sólo se cargan los datos sin procesar, tal y como están, y cuando esté listo para usar los datos, es cuando se le da forma y estructura. Eso se llama schema-on-read. Dos enfoques muy diferentes.

- Almacenamiento: Una de las principales características de las tecnologías de big data, como Hadoop, es que el coste de almacenamiento de datos es relativamente bajo en comparación con el de un data warehouse. Hay dos razones principales para esto: en primer lugar, Hadoop es software de código abierto, por lo que la concesión de licencias y el soporte de la comunidad es gratuito. Y segundo, Hadoop está diseñado para ser instalado en hardware de bajo coste.
- Agilidad: Un almacén de datos es un repositorio altamente estructurado, por definición. No
  es técnicamente difícil cambiar la estructura, pero puede tomar mucho tiempo dado todos
  los procesos de negocio que están vinculados a ella. Un data lake, por otro lado, carece de
  la estructura de un data warehouse, lo que da a los desarrolladores y a los científicos de
  datos la capacidad de configurar y reconfigurar fácilmente y en tiempo real sus modelos,
  consultas y aplicaciones.
- Seguridad: La tecnología del data warehouse existe desde hace décadas, mientras que la tecnología de big data (la base de un Data Lake) es relativamente nueva. Por lo tanto, la capacidad de asegurar datos en un data warehouse es mucho más madura que asegurar datos en un data lake. Cabe señalar, sin embargo, que se está realizando un importante esfuerzo en materia de seguridad en la actualidad en la industria de Big Data.

El data warehouse continúa evolucionando

Hay todavía un gran potencial en la evolución del data warehouse y su uso sigue adelante.

El procesamiento analítico clásico de datos basado en transacciones se realiza en el data warehouse como siempre se ha hecho. Nada ha cambiado ahí. Pero ahora puede hacerse analítica sobre datos contextualizados, y esa forma de análisis es nueva. La mayoría de las organizaciones, hasta ahora no habían podido basar su toma de decisiones en datos textuales no estructurados. Y ahora hay una nueva forma de análisis posible en un Data Warehouse: la posibilidad de la mezcla de análisis. El análisis combinado se realiza mediante una combinación de datos transaccionales estructurados y datos contextuales no estructurados.

Existen muchas otras formas de análisis que también son posibles. Estas formas incluyen análisis predictivo y prescriptivo, así como tecnologías de machine learning y otras que están cambiando la forma en que los datos son administrados y analizados. El data warehouse ha sido durante mucho tiempo un elemento básico de la arquitectura de datos empresarial y sigue teniendo tiene un gran futuro en el nuevo mundo de Big Data y la analítica avanzada.

### 3. Data Warehouse en la nube

Los data warehouses están atravesando actualmente dos transformaciones muy importantes que tienen el potencial de impulsar niveles significativos de innovación empresarial:

- La primera área de transformación es el impulso para aumentar la agilidad general. La gran mayoría de los departamentos de TI están experimentando un rápido aumento de la demanda de datos. Los directivos quieren tener acceso a más y más datos históricos, mientras que al mismo tiempo, los científicos de datos y los analistas de negocios están explorando formas de introducir nuevos flujos de datos en el almacén para enriquecer el análisis existente, así como impulsar nuevas áreas de análisis. Esta rápida expansión de los volúmenes y fuentes de datos significa que los equipos de TI necesitan invertir más tiempo y esfuerzo asegurando que el rendimiento de las consultas permanezca constante y necesitan proporcionar cada vez más entornos para equipos individuales para validar el valor comercial de los nuevos conjuntos de datos.
- La segunda área de transformación gira en torno a la necesidad de mejorar el control
  de costes. Existe una creciente necesidad de hacer más con cada vez menos recursos, al
  mismo tiempo que se garantiza que todos los datos sensibles y estratégicos estén
  completamente asegurados, a lo largo de todo el ciclo de vida, de la manera más rentable.

Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

La nube está demostrando ser un facilitador dominante. Permite a las organizaciones enfrentarse de forma activa a los desafíos que presentan estas dos transformaciones clave.

¿Por qué migrar a la nube?

Las 3 tecnologías que ya tienen y seguirán teniendo un gran impacto en la transformación de negocios son: la nube, datos y análisis, e Internet de las Cosas. Estas tres tecnologías son parte clave del ecosistema del data warehouse actual. Por lo tanto, es posible sacar la conclusión de que el data warehouse en la nube tiene un gran potencial para generar un impacto significativo en los negocios.

La importancia de la nube para el data warehouse está directamente vinculado a tres factores clave:

- Mayor agilidad
- Mejor control de costes
- Co-localización

Mejora de la agilidad

Muchos data warehouses están embarcados en una fase de actualización. Con mucho trabajo puesto ahora en torno al big data, las empresas buscan aprovechar los nuevos flujos de datos y nuevos tipos de análisis más ricos para apoyar e impulsar nuevas áreas, tales como:analítica de clientes de 360°, analísis predictivo, detección de fraude, análisis de IoT y el establecimiento de los datos como centro de beneficio. Muchos de estos proyectos requieren el aprovisionamiento de nuevos entornos de hardware y el despliegue de software. Es más rápido, más fácil y más eficiente iniciar estos nuevos proyectos centrados en los datos utilizando los servicios en la nube.

#### Mayor control de costes

Muchos equipos de TI están buscando maneras de consolidar los data marts existentes, cada uno ejecutándose en hardware dedicado o incluso en hardware propietario, en un único entorno integrado. Los servicios en la nube de calidad ofrecen la oportunidad perfecta para iniciar este tipo de proyectos. El cumplimiento legal no puede ser visto como un extra opcional cuando se planifica un traslado a la nube. Los activos de datos deben estar protegidos a lo largo de todo el ciclo de vida. Los servicios en la nube deben facilitar todo esto, siendo más rentables ya que todas las características de seguridad se pueden habilitar de forma predeterminada y mejorada, y actualizada de forma transparente.

Co-Localización para una carga más rápida

La mayoría de los data warehouses generan datos directamente desde aplicaciones clave, como entrada de pedidos, ventas, finanzas y fabricación. Por lo tanto, tiene mucho sentido ubicar conjuntamente el data warehouse junto con los sistemas fuente que ya se estén ejecutándose en la nube. La ubicación conjunta ofrece una carga de datos más rápida, lo que significa que los usuarios obtienen un acceso más oportuno a sus datos.

Principales ventajas de mover el data warehouse a la nube

Existen **tres ventajas principales** para mover un data warehouse a la nube, y estas están directamente vinculados a los tres controladores clave enumerados anteriormente:

- Más fácil consolidación y racionalización
- Monetización más rápida de los datos en la nube
- La nube ofrece mejor protección

# 4. Data Warehouse vs Big Data vs BI: cuales son las diferencias

En este apartado compararemos brevemente los conceptos Data Warehouse (DW), Big Data y Business Intelligence (BI).

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para tener una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

Se trata de tres conceptos completamente diferentes que tienen en común una nueva manera de lidiar con los datos, siempre teniendo en cuenta la existencia de una gran volumen de información en varios formatos que contribuyen, de forma estructurada o no estructurada, a la toma de decisiones estratégicas. El objetivo final de cualquiera de estas tecnologías es ofrecer una ventaja competitiva a las empresas, pero la forma en que se utiliza es la que marcará la diferencia.

Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

#### Big Data

Llamamos big data a un gran volumen de datos con una variedad, complejidad y velocidad de crecimiento enorme y que además tienen la característica de no ser estructurados. Eso significa que no son relacionales, estando además fuera del entorno corporativo. Es un tipo de tecnología que te permite analizar los datos en tiempo real y puede provenir de diferentes fuentes y formas, tales como mensajería instantánea, redes sociales, registros de grabaciones, imágenes, mensajes de correo electrónico, etc.

## Guía gratuita: "Del Bit... al Big Data", descubre todo lo que necesitas saber sobre Big Data y mucho más.

Para tener una idea de la importancia del big data, este mercado está creciendo cada año, alrededor de un 40%. La paradoja reside en el hecho de que cuanto mayor es la evolución del big data, mayor es la escasez de profesionales cualificados para satisfacer esa demanda. La expectativa es que en unos pocos años, haya un crecimiento aún más significativo para el uso de big data y por lo tanto un aumento en la demanda mano de obra especializada y de partners tecnológicos especializados en esta materia.

#### Data Warehouse

Por otro lado, un data warehouse almacena datos consolidados de diversas fuentes o sistemas de la empresa. Se trata de datos estructurados, que tiene como objetivo principal ser precisos y de alta calidad para de esta forma poder dar soporte a la toma de decisiones de la empresa. Se trata de conseguir tener todos los datos juntos para después poder dividirlos para hacer un análisis de determinados sectores o estrategias.

#### Business Intelligence

Un Business Intelligence (BI) es una especie de "cuello de botella" de los datos recogidos del data warehouse, que llegan de forma exacta y útil para ayudar a la toma de decisiones. Business Intelligence transforma los datos en información útil para analizar no sólo los negocios, sino también las principales estrategias corporativas.

## Guía gratuita "Integridad de datos y BI", descubre cómo sacarle el máximo partido a tu proyecto de Integridad de datos

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

### 5. Material complementario

Soluciones y recursos para Data Warehouse

Informatica Cloud para Microsoft Azure SQL Data Warehouse

- Test Data Warehouse
- MDM y calidad de datos para el data warehouse

#### Guías

- Data Lake: Superando las limitaciones del Data Warehouse
- El data warehouse inteligente

#### Artículos

- ¿Qué es warehouse y qué debe contemplar un data warehouse moderno?
- Beneficios y soluciones para un Data Warehouse en la nube
- Lo que necesitas saber acerca de la arquitectura de un datawarehouse
- Data warehousing: 3 ocasiones de hacer las cosas bien
- ¿Qué es un Data Warehouse y cómo saber cuándo lo necesitas implementar?
- Cómo optimizar la arquitectura de un data warehouse
- Data warehouse concepts: tu glosario de almacén de datos
- Data Warehouse, el primer paso para una solución de BI
- Tutoriales sobre DataWarehouse en pdf
- Cómo gestiona Informatica los nuevos medios de almacenamiento de datos
- ¿Qué es un Data Warehouse para un director de ventas?
- Bl en línea: el reto actual de los data warehouses
- El poder de data warehouse: Informatica PowerCenter

### 2. Pasado y presente del Data Warehouse

Históricamente, los data warehouses se habían formado utilizando datos repetitivos estructurados que eran filtrados antes de entrar en el data warehouse. Sin embargo, en los últimos años, el data warehouse ha evolucionado debido a información contextual que ahora se puede adjuntar a los datos no estructurados y que también puede ser almacenada.

Aquellos primeros datos relacionales estructurados no podían ser mezclados y emparejados para temas analíticos con datos textuales no estructurados. Pero con el advenimiento de la contextualización, estos tipos de análisis ahora sí pueden hacerse de forma naturales y fácil.

En el data warehouse, datos no repetitivos, como los comentarios en una encuesta, correos electrónicos y conversaciones, se tratan de forma diferente a las ocurrencias repetitivas de datos, como el flujo de clics, mediciones o el procesamiento máquina o analógico. Los datos no repetitivos son datos basados en textos que fueron generados por la palabra escrita o hablada, leída y reformateada y, lo que es más importante, ahora puede ser contextualizada. Con el fin de extraer cualquier sentido de los datos no repetitivos para su uso en el Data Warehouse, deben tener el contexto de los datos establecidos.

En muchos casos, el contexto de los datos no repetitivos es más importante que los datos en sí. En cualquier caso, los datos no repetitivos no pueden utilizarse para la toma de decisiones hasta que se haya establecido el contexto.

Data Lakes y Data Warehouses: ¿mutuamente exclusivos o partners perfectos?

Los data lakes han surgido en el paisaje de Data Management en los últimos años, sin embargo, data lake no es necesariamente un reemplazo del data warehouse. En cambio, complementan los esfuerzos existentes y dan soporte al descubrimiento de nuevas preguntas. Una vez que se descubren esas preguntas se optimizan las respuestas. Y optimizar puede significar moverse fuera del data lake para ir a un data mart o al data warehouse.

¡Resuelve tus dudas sobre el Data Lake con uno de nuestros expertos!

Estas son algunas diferencias clave entre data lake y data warehouse:

- Datos: Un data warehouse sólo almacena datos que han sido modelados o estructurados, mientras que un Data Lake no hace acepción de datos. Lo almacena todo, estructurado, semiestructurado y no estructurado.
- **Procesamiento**: Antes de que una empresa pueda cargar datos en un data warehouse, primero debe darles forma y estructura, es decir, los datos deben ser modelados. Eso se llama schema-on-write. Con un data lake, sólo se cargan los datos sin procesar, tal y como están, y cuando esté listo para usar los datos, es cuando se le da forma y estructura. Eso se llama schema-on-read. Dos enfoques muy diferentes.
- Almacenamiento: Una de las principales características de las tecnologías de big data, como Hadoop, es que el coste de almacenamiento de datos es relativamente bajo en comparación con el de un data warehouse. Hay dos razones principales para esto: en primer lugar, Hadoop es software de código abierto, por lo que la concesión de licencias y el soporte de la comunidad es gratuito. Y segundo, Hadoop está diseñado para ser instalado en hardware de bajo coste.
- Agilidad: Un almacén de datos es un repositorio altamente estructurado, por definición. No
  es técnicamente difícil cambiar la estructura, pero puede tomar mucho tiempo dado todos
  los procesos de negocio que están vinculados a ella. Un data lake, por otro lado, carece de
  la estructura de un data warehouse, lo que da a los desarrolladores y a los científicos de
  datos la capacidad de configurar y reconfigurar fácilmente y en tiempo real sus modelos,
  consultas y aplicaciones.
- Seguridad: La tecnología del data warehouse existe desde hace décadas, mientras que la tecnología de big data (la base de un Data Lake) es relativamente nueva. Por lo tanto, la capacidad de asegurar datos en un data warehouse es mucho más madura que asegurar datos en un data lake. Cabe señalar, sin embargo, que se está realizando un importante esfuerzo en materia de seguridad en la actualidad en la industria de Big Data.

El data warehouse continúa evolucionando

Hay todavía un gran potencial en la evolución del data warehouse y su uso sigue adelante.

El procesamiento analítico clásico de datos basado en transacciones se realiza en el data warehouse como siempre se ha hecho. Nada ha cambiado ahí. Pero ahora puede hacerse analítica sobre datos contextualizados, y esa forma de análisis es nueva. La mayoría de las organizaciones, hasta ahora no habían podido basar su toma de decisiones en datos textuales no estructurados. Y ahora hay una nueva forma de análisis posible en un Data Warehouse: la posibilidad de la mezcla de análisis. El análisis combinado se realiza mediante una combinación de datos transaccionales estructurados y datos contextuales no estructurados.

Existen muchas otras formas de análisis que también son posibles. Estas formas incluyen análisis predictivo y prescriptivo, así como tecnologías de machine learning y otras que están cambiando la forma en que los datos son administrados y analizados. El data warehouse ha sido durante mucho tiempo un elemento básico de la arquitectura de datos empresarial y sigue teniendo tiene un gran futuro en el nuevo mundo de Big Data y la analítica avanzada.

### 3. Data Warehouse en la nube

Los data warehouses están atravesando actualmente dos transformaciones muy importantes que tienen el potencial de impulsar niveles significativos de innovación empresarial:

- La primera área de transformación es el impulso para aumentar la agilidad general. La gran mayoría de los departamentos de TI están experimentando un rápido aumento de la demanda de datos. Los directivos quieren tener acceso a más y más datos históricos, mientras que al mismo tiempo, los científicos de datos y los analistas de negocios están explorando formas de introducir nuevos flujos de datos en el almacén para enriquecer el análisis existente, así como impulsar nuevas áreas de análisis. Esta rápida expansión de los volúmenes y fuentes de datos significa que los equipos de TI necesitan invertir más tiempo y esfuerzo asegurando que el rendimiento de las consultas permanezca constante y necesitan proporcionar cada vez más entornos para equipos individuales para validar el valor comercial de los nuevos conjuntos de datos.
- La segunda área de transformación gira en torno a la necesidad de mejorar el control de costes. Existe una creciente necesidad de hacer más con cada vez menos recursos, al mismo tiempo que se garantiza que todos los datos sensibles y estratégicos estén completamente asegurados, a lo largo de todo el ciclo de vida, de la manera más rentable.

Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

La nube está demostrando ser un facilitador dominante. Permite a las organizaciones enfrentarse de forma activa a los desafíos que presentan estas dos transformaciones clave.

#### ¿Por qué migrar a la nube?

Las 3 tecnologías que ya tienen y seguirán teniendo un gran impacto en la transformación de negocios son: la nube, datos y análisis, e Internet de las Cosas. Estas tres tecnologías son parte clave del ecosistema del data warehouse actual. Por lo tanto, es posible sacar la conclusión de que el data warehouse en la nube tiene un gran potencial para generar un impacto significativo en los negocios.

La importancia de la nube para el data warehouse está directamente vinculado a tres factores clave:

- Mayor agilidad
- Mejor control de **costes**
- Co-localización

#### Mejora de la agilidad

Muchos data warehouses están embarcados en una fase de actualización. Con mucho trabajo puesto ahora en torno al big data, las empresas buscan aprovechar los nuevos flujos de datos y nuevos tipos de análisis más ricos para apoyar e impulsar nuevas áreas, tales como:analítica de clientes de 360°, analisis predictivo, detección de fraude, análisis de IoT y el establecimiento de los datos como centro de beneficio. Muchos de estos proyectos requieren el aprovisionamiento de nuevos entornos de hardware y el despliegue de software. Es más rápido, más fácil y más eficiente iniciar estos nuevos proyectos centrados en los datos utilizando los servicios en la nube.

#### Mayor control de costes

Muchos equipos de TI están buscando maneras de consolidar los data marts existentes, cada uno ejecutándose en hardware dedicado o incluso en hardware propietario, en un único entorno integrado. Los servicios en la nube de calidad ofrecen la oportunidad perfecta para iniciar este tipo de proyectos. El cumplimiento legal no puede ser visto como un extra opcional cuando se planifica un traslado a la nube. Los activos de datos deben estar protegidos a lo largo de todo el ciclo de vida. Los servicios en la nube deben facilitar todo esto, siendo más rentables ya que todas las características de seguridad se pueden habilitar de forma predeterminada y mejorada, y actualizada de forma transparente.

#### Co-Localización para una carga más rápida

La mayoría de los data warehouses generan datos directamente desde aplicaciones clave, como entrada de pedidos, ventas, finanzas y fabricación. Por lo tanto, **tiene mucho sentido ubicar conjuntamente el data warehouse junto con los sistemas fuente que ya se estén ejecutándose en la nube**. La ubicación conjunta ofrece una carga de datos más rápida, lo que significa que los usuarios obtienen un acceso más oportuno a sus datos.

Principales ventajas de mover el data warehouse a la nube

Existen **tres ventajas principales** para mover un data warehouse a la nube, y estas están directamente vinculados a los tres controladores clave enumerados anteriormente:

- Más fácil consolidación y racionalización
- Monetización más rápida de los datos en la nube
- La nube ofrece mejor protección

## 4. Data Warehouse vs Big Data vs BI: cuales son las diferencias

En este apartado compararemos brevemente los conceptos Data Warehouse (DW), Big Data y Business Intelligence (BI).

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para tener una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

Se trata de tres conceptos completamente diferentes que tienen en común una nueva manera de lidiar con los datos, siempre teniendo en cuenta la existencia de una gran volumen de información en varios formatos que contribuyen, de forma estructurada o no estructurada, a la toma de decisiones estratégicas. El objetivo final de cualquiera de estas tecnologías es ofrecer una ventaja competitiva a las empresas, pero la forma en que se utiliza es la que marcará la diferencia.

Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

Big Data

Llamamos big data a un gran volumen de datos con una variedad, complejidad y velocidad de crecimiento enorme y que además tienen la característica de no ser estructurados. Eso significa que no son relacionales, estando además fuera del entorno corporativo. Es un tipo de tecnología que te permite analizar los datos en tiempo real y puede provenir de diferentes fuentes y formas, tales como mensajería instantánea, redes sociales, registros de grabaciones, imágenes, mensajes de correo electrónico, etc.

## Guía gratuita: "Del Bit... al Big Data", descubre todo lo que necesitas saber sobre Big Data y mucho más.

Para tener una idea de la importancia del big data, este mercado está creciendo cada año, alrededor de un 40%. La paradoja reside en el hecho de que cuanto mayor es la evolución del big data, mayor es la escasez de profesionales cualificados para satisfacer esa demanda. La expectativa es que en unos pocos años, haya un crecimiento aún más significativo para el uso de big data y por lo tanto un aumento en la demanda mano de obra especializada y de partners tecnológicos especializados en esta materia.

#### Data Warehouse

Por otro lado, un data warehouse almacena datos consolidados de diversas fuentes o sistemas de la empresa. Se trata de datos estructurados, que tiene como objetivo principal ser precisos y de alta calidad para de esta forma poder dar soporte a la toma de decisiones de la empresa. Se trata de conseguir tener todos los datos juntos para después poder dividirlos para hacer un análisis de determinados sectores o estrategias.

#### **Business Intelligence**

Un Business Intelligence (BI) es una especie de "cuello de botella" de los datos recogidos del data warehouse, que llegan de forma exacta y útil para ayudar a la toma de decisiones. Business Intelligence transforma los datos en información útil para analizar no sólo los negocios, sino también las principales estrategias corporativas.

### Guía gratuita "Integridad de datos y BI", descubre cómo sacarle el máximo partido a tu proyecto de Integridad de datos

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

### 5. Material complementario

Soluciones y recursos para Data Warehouse

- Informatica Cloud para Microsoft Azure SQL Data Warehouse
- Test Data Warehouse
- MDM y calidad de datos para el data warehouse

#### Guías

- Data Lake: Superando las limitaciones del Data Warehouse
- El data warehouse inteligente

#### Artículos

- ¿Qué es warehouse y qué debe contemplar un data warehouse moderno?
- Beneficios y soluciones para un Data Warehouse en la nube
- Lo que necesitas saber acerca de la arquitectura de un datawarehouse
- Data warehousing: 3 ocasiones de hacer las cosas bien
- ¿Qué es un Data Warehouse y cómo saber cuándo lo necesitas implementar?
- Cómo optimizar la arquitectura de un data warehouse
- Data warehouse concepts: tu glosario de almacén de datos
- Data Warehouse, el primer paso para una solución de BI
- Tutoriales sobre DataWarehouse en pdf
- Cómo gestiona Informatica los nuevos medios de almacenamiento de datos

- ¿Qué es un Data Warehouse para un director de ventas?
- Bl en línea: el reto actual de los data warehouses
- El poder de data warehouse: Informatica PowerCenter

## 2. Pasado y presente del Data Warehouse

Históricamente, los data warehouses se habían formado utilizando datos repetitivos estructurados que eran filtrados antes de entrar en el data warehouse. Sin embargo, en los últimos años, el data warehouse ha evolucionado debido a información contextual que ahora se puede adjuntar a los datos no estructurados y que también puede ser almacenada.

Aquellos primeros datos relacionales estructurados no podían ser mezclados y emparejados para temas analíticos con datos textuales no estructurados. Pero con el advenimiento de la contextualización, estos tipos de análisis ahora sí pueden hacerse de forma naturales y fácil.

En el data warehouse, datos no repetitivos, como los comentarios en una encuesta, correos electrónicos y conversaciones, se tratan de forma diferente a las ocurrencias repetitivas de datos, como el flujo de clics, mediciones o el procesamiento máquina o analógico. Los datos no repetitivos son datos basados en textos que fueron generados por la palabra escrita o hablada, leída y reformateada y, lo que es más importante, ahora puede ser contextualizada. Con el fin de extraer cualquier sentido de los datos no repetitivos para su uso en el Data Warehouse, deben tener el contexto de los datos establecidos.

En muchos casos, el contexto de los datos no repetitivos es más importante que los datos en sí. En cualquier caso, los datos no repetitivos no pueden utilizarse para la toma de decisiones hasta que se haya establecido el contexto.

Data Lakes y Data Warehouses: ¿mutuamente exclusivos o partners perfectos?

Los data lakes han surgido en el paisaje de Data Management en los últimos años, sin embargo, data lake no es necesariamente un reemplazo del data warehouse. En cambio, complementan los esfuerzos existentes y dan soporte al descubrimiento de nuevas preguntas. Una vez que se descubren esas preguntas se optimizan las respuestas. Y optimizar puede significar moverse fuera del data lake para ir a un data mart o al data warehouse.

¡Resuelve tus dudas sobre el Data Lake con uno de nuestros expertos!

Estas son algunas diferencias clave entre data lake y data warehouse:

- Datos: Un data warehouse sólo almacena datos que han sido modelados o estructurados, mientras que un Data Lake no hace acepción de datos. Lo almacena todo, estructurado, semiestructurado y no estructurado.
- Procesamiento: Antes de que una empresa pueda cargar datos en un data warehouse, primero debe darles forma y estructura, es decir, los datos deben ser modelados. Eso se

llama schema-on-write. Con un data lake, sólo se cargan los datos sin procesar, tal y como están, y cuando esté listo para usar los datos, es cuando se le da forma y estructura. Eso se llama schema-on-read. Dos enfoques muy diferentes.

- Almacenamiento: Una de las principales características de las tecnologías de big data, como Hadoop, es que el coste de almacenamiento de datos es relativamente bajo en comparación con el de un data warehouse. Hay dos razones principales para esto: en primer lugar, Hadoop es software de código abierto, por lo que la concesión de licencias y el soporte de la comunidad es gratuito. Y segundo, Hadoop está diseñado para ser instalado en hardware de bajo coste.
- Agilidad: Un almacén de datos es un repositorio altamente estructurado, por definición. No
  es técnicamente difícil cambiar la estructura, pero puede tomar mucho tiempo dado todos
  los procesos de negocio que están vinculados a ella. Un data lake, por otro lado, carece de
  la estructura de un data warehouse, lo que da a los desarrolladores y a los científicos de
  datos la capacidad de configurar y reconfigurar fácilmente y en tiempo real sus modelos,
  consultas y aplicaciones.
- Seguridad: La tecnología del data warehouse existe desde hace décadas, mientras que la tecnología de big data (la base de un Data Lake) es relativamente nueva. Por lo tanto, la capacidad de asegurar datos en un data warehouse es mucho más madura que asegurar datos en un data lake. Cabe señalar, sin embargo, que se está realizando un importante esfuerzo en materia de seguridad en la actualidad en la industria de Big Data.

El data warehouse continúa evolucionando

Hay todavía un gran potencial en la evolución del data warehouse y su uso sigue adelante.

El procesamiento analítico clásico de datos basado en transacciones se realiza en el data warehouse como siempre se ha hecho. Nada ha cambiado ahí. Pero ahora puede hacerse analítica sobre datos contextualizados, y esa forma de análisis es nueva. La mayoría de las organizaciones, hasta ahora no habían podido basar su toma de decisiones en datos textuales no estructurados. Y ahora hay una nueva forma de análisis posible en un Data Warehouse: la posibilidad de la mezcla de análisis. El análisis combinado se realiza mediante una combinación de datos transaccionales estructurados y datos contextuales no estructurados.

Existen muchas otras formas de análisis que también son posibles. Estas formas incluyen análisis predictivo y prescriptivo, así como tecnologías de machine learning y otras que están cambiando la forma en que los datos son administrados y analizados. El data warehouse ha sido durante mucho tiempo un elemento básico de la arquitectura de datos empresarial y sigue teniendo tiene un gran futuro en el nuevo mundo de Big Data y la analítica avanzada.

#### 3. Data Warehouse en la nube

Los data warehouses están atravesando actualmente dos transformaciones muy importantes que tienen el potencial de impulsar niveles significativos de innovación empresarial:

- La primera área de transformación es el impulso para aumentar la agilidad general. La gran mayoría de los departamentos de TI están experimentando un rápido aumento de la demanda de datos. Los directivos quieren tener acceso a más y más datos históricos, mientras que al mismo tiempo, los científicos de datos y los analistas de negocios están explorando formas de introducir nuevos flujos de datos en el almacén para enriquecer el análisis existente, así como impulsar nuevas áreas de análisis. Esta rápida expansión de los volúmenes y fuentes de datos significa que los equipos de TI necesitan invertir más tiempo y esfuerzo asegurando que el rendimiento de las consultas permanezca constante y necesitan proporcionar cada vez más entornos para equipos individuales para validar el valor comercial de los nuevos conjuntos de datos.
- La segunda área de transformación gira en torno a la necesidad de mejorar el control
  de costes. Existe una creciente necesidad de hacer más con cada vez menos recursos, al
  mismo tiempo que se garantiza que todos los datos sensibles y estratégicos estén
  completamente asegurados, a lo largo de todo el ciclo de vida, de la manera más rentable.

Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

La nube está demostrando ser un facilitador dominante. Permite a las organizaciones enfrentarse de forma activa a los desafíos que presentan estas dos transformaciones clave.

¿Por qué migrar a la nube?

Las 3 tecnologías que ya tienen y seguirán teniendo un gran impacto en la transformación de negocios son: la nube, datos y análisis, e Internet de las Cosas. Estas tres tecnologías son parte clave del ecosistema del data warehouse actual. Por lo tanto, es posible sacar la conclusión de que el data warehouse en la nube tiene un gran potencial para generar un impacto significativo en los negocios.

La importancia de la nube para el data warehouse está directamente vinculado a tres factores clave:

- Mayor agilidad
- Mejor control de **costes**
- Co-localización

#### Mejora de la agilidad

Muchos data warehouses están embarcados en una fase de actualización. Con mucho trabajo puesto ahora en torno al big data, las empresas buscan aprovechar los nuevos flujos de datos y nuevos tipos de análisis más ricos para apoyar e impulsar nuevas áreas, tales como:analítica de clientes de 360°, analisis predictivo, detección de fraude, análisis de IoT y el establecimiento de los datos como centro de beneficio. Muchos de estos proyectos requieren el aprovisionamiento de nuevos entornos de hardware y el despliegue de software. Es más rápido, más fácil y más eficiente iniciar estos nuevos proyectos centrados en los datos utilizando los servicios en la nube.

#### Mayor control de costes

Muchos equipos de TI están buscando maneras de consolidar los data marts existentes, cada uno ejecutándose en hardware dedicado o incluso en hardware propietario, en un único entorno integrado. Los servicios en la nube de calidad ofrecen la oportunidad perfecta para iniciar este tipo de proyectos. El cumplimiento legal no puede ser visto como un extra opcional cuando se planifica un traslado a la nube. Los activos de datos deben estar protegidos a lo largo de todo el ciclo de vida. Los servicios en la nube deben facilitar todo esto, siendo más rentables ya que todas las características de seguridad se pueden habilitar de forma predeterminada y mejorada, y actualizada de forma transparente.

#### Co-Localización para una carga más rápida

La mayoría de los data warehouses generan datos directamente desde aplicaciones clave, como entrada de pedidos, ventas, finanzas y fabricación. Por lo tanto, tiene mucho sentido ubicar conjuntamente el data warehouse junto con los sistemas fuente que ya se estén ejecutándose en la nube. La ubicación conjunta ofrece una carga de datos más rápida, lo que significa que los usuarios obtienen un acceso más oportuno a sus datos.

Principales ventajas de mover el data warehouse a la nube

Existen **tres ventajas principales** para mover un data warehouse a la nube, y estas están directamente vinculados a los tres controladores clave enumerados anteriormente:

- Más fácil consolidación y racionalización
- Monetización más rápida de los datos en la nube
- La nube ofrece mejor protección

## 4. Data Warehouse vs Big Data vs BI: cuales son las diferencias

En este apartado compararemos brevemente los conceptos Data Warehouse (DW), Big Data y Business Intelligence (BI).

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para tener una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

Se trata de tres conceptos completamente diferentes que tienen en común una nueva manera de lidiar con los datos, siempre teniendo en cuenta la existencia de una gran volumen de información en varios formatos que contribuyen, de forma estructurada o no estructurada, a la toma de decisiones estratégicas. El objetivo final de cualquiera de estas tecnologías es ofrecer una ventaja competitiva a las empresas, pero la forma en que se utiliza es la que marcará la diferencia.

Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

#### Big Data

Llamamos big data a un gran volumen de datos con una variedad, complejidad y velocidad de crecimiento enorme y que además tienen la característica de no ser estructurados. Eso significa que no son relacionales, estando además fuera del entorno corporativo. Es un tipo de tecnología que te permite analizar los datos en tiempo real y puede provenir de diferentes fuentes y formas, tales como mensajería instantánea, redes sociales, registros de grabaciones, imágenes, mensajes de correo electrónico, etc.

## Guía gratuita: "Del Bit... al Big Data", descubre todo lo que necesitas saber sobre Big Data y mucho más.

Para tener una idea de la importancia del big data, este mercado está creciendo cada año, alrededor de un 40%. La paradoja reside en el hecho de que cuanto mayor es la evolución del big data, mayor es la escasez de profesionales cualificados para satisfacer esa demanda. La expectativa es que en unos pocos años, haya un crecimiento aún más significativo para el uso de big data y por lo tanto un aumento en la demanda mano de obra especializada y de partners tecnológicos especializados en esta materia.

#### Data Warehouse

Por otro lado, un data warehouse almacena datos consolidados de diversas fuentes o sistemas de la empresa. Se trata de datos estructurados, que tiene como objetivo principal ser precisos y de alta calidad para de esta forma poder dar soporte a la toma de decisiones de la empresa. Se trata de conseguir tener todos los datos juntos para después poder dividirlos para hacer un análisis de determinados sectores o estrategias.

#### Business Intelligence

Un Business Intelligence (BI) es una especie de "cuello de botella" de los datos recogidos del data warehouse, que llegan de forma exacta y útil para ayudar a la toma de decisiones. Business Intelligence transforma los datos en información útil para analizar no sólo los negocios, sino también las principales estrategias corporativas.

Guía gratuita "Integridad de datos y BI", descubre cómo sacarle el máximo partido a tu proyecto de Integridad de datos

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

### 5. Material complementario

Soluciones y recursos para Data Warehouse

- Informatica Cloud para Microsoft Azure SQL Data Warehouse
- Test Data Warehouse
- MDM y calidad de datos para el data warehouse

#### Guías

- Data Lake: Superando las limitaciones del Data Warehouse
- El data warehouse inteligente

#### Artículos

- ¿Qué es warehouse y qué debe contemplar un data warehouse moderno?
- Beneficios y soluciones para un Data Warehouse en la nube
- Lo que necesitas saber acerca de la arquitectura de un datawarehouse
- Data warehousing: 3 ocasiones de hacer las cosas bien
- ¿Qué es un Data Warehouse y cómo saber cuándo lo necesitas implementar?
- Cómo optimizar la arquitectura de un data warehouse
- Data warehouse concepts: tu glosario de almacén de datos
- Data Warehouse, el primer paso para una solución de BI
- Tutoriales sobre DataWarehouse en pdf
- Cómo gestiona Informatica los nuevos medios de almacenamiento de datos
- ¿Qué es un Data Warehouse para un director de ventas?
- Bl en línea: el reto actual de los data warehouses
- El poder de data warehouse: Informatica PowerCenter

## 2. Pasado y presente del Data Warehouse

Históricamente, los data warehouses se habían formado utilizando datos repetitivos estructurados que eran filtrados antes de entrar en el data warehouse. Sin embargo, en los últimos años, el data warehouse ha evolucionado debido a información contextual que ahora se puede adjuntar a los datos no estructurados y que también puede ser almacenada.

Aquellos primeros datos relacionales estructurados no podían ser mezclados y emparejados para temas analíticos con datos textuales no estructurados. Pero con el advenimiento de la contextualización, estos tipos de análisis ahora sí pueden hacerse de forma naturales y fácil.

En el data warehouse, datos no repetitivos, como los comentarios en una encuesta, correos electrónicos y conversaciones, se tratan de forma diferente a las ocurrencias

repetitivas de datos, como el flujo de clics, mediciones o el procesamiento máquina o analógico. Los datos no repetitivos son datos basados en textos que fueron generados por la palabra escrita o hablada, leída y reformateada y, lo que es más importante, ahora puede ser contextualizada. Con el fin de extraer cualquier sentido de los datos no repetitivos para su uso en el Data Warehouse, deben tener el contexto de los datos establecidos.

En muchos casos, el contexto de los datos no repetitivos es más importante que los datos en sí. En cualquier caso, los datos no repetitivos no pueden utilizarse para la toma de decisiones hasta que se haya establecido el contexto.

Data Lakes y Data Warehouses: ¿mutuamente exclusivos o partners perfectos?

Los data lakes han surgido en el paisaje de Data Management en los últimos años, sin embargo, data lake no es necesariamente un reemplazo del data warehouse. En cambio, complementan los esfuerzos existentes y dan soporte al descubrimiento de nuevas preguntas. Una vez que se descubren esas preguntas se optimizan las respuestas. Y optimizar puede significar moverse fuera del data lake para ir a un data mart o al data warehouse.

<u>:Resuelve tus dudas sobre el Data Lake</u> con uno de nuestros expertos!

Estas son algunas diferencias clave entre data lake y data warehouse:

- Datos: Un data warehouse sólo almacena datos que han sido modelados o estructurados, mientras que un Data Lake no hace acepción de datos. Lo almacena todo, estructurado, semiestructurado y no estructurado.
- **Procesamiento**: Antes de que una empresa pueda cargar datos en un data warehouse, primero debe darles forma y estructura, es decir, los datos deben ser modelados. Eso se llama schema-on-write. Con un data lake, sólo se cargan los datos sin procesar, tal y como están, y cuando esté listo para usar los datos, es cuando se le da forma y estructura. Eso se llama schema-on-read. Dos enfoques muy diferentes.
- Almacenamiento: Una de las principales características de las tecnologías de big data, como Hadoop, es que el coste de almacenamiento de datos es relativamente bajo en comparación con el de un data warehouse. Hay dos razones principales para esto: en primer lugar, Hadoop es software de código abierto, por lo que la concesión de licencias y el soporte de la comunidad es gratuito. Y segundo, Hadoop está diseñado para ser instalado en hardware de bajo coste.
- Agilidad: Un almacén de datos es un repositorio altamente estructurado, por definición. No
  es técnicamente difícil cambiar la estructura, pero puede tomar mucho tiempo dado todos
  los procesos de negocio que están vinculados a ella. Un data lake, por otro lado, carece de
  la estructura de un data warehouse, lo que da a los desarrolladores y a los científicos de
  datos la capacidad de configurar y reconfigurar fácilmente y en tiempo real sus modelos,
  consultas y aplicaciones.
- **Seguridad**: La tecnología del data warehouse existe desde hace décadas, mientras que la tecnología de big data (la base de un Data Lake) es relativamente nueva. Por lo tanto, la capacidad de asegurar datos en un data warehouse es mucho más madura que asegurar

datos en un data lake. Cabe señalar, sin embargo, que se está realizando un importante esfuerzo en materia de seguridad en la actualidad en la industria de Big Data.

El data warehouse continúa evolucionando

Hay todavía un gran potencial en la evolución del data warehouse y su uso sigue adelante.

El procesamiento analítico clásico de datos basado en transacciones se realiza en el data warehouse como siempre se ha hecho. Nada ha cambiado ahí. Pero ahora puede hacerse analítica sobre datos contextualizados, y esa forma de análisis es nueva. La mayoría de las organizaciones, hasta ahora no habían podido basar su toma de decisiones en datos textuales no estructurados. Y ahora hay una nueva forma de análisis posible en un Data Warehouse: la posibilidad de la mezcla de análisis. El análisis combinado se realiza mediante una combinación de datos transaccionales estructurados y datos contextuales no estructurados.

Existen muchas otras formas de análisis que también son posibles. Estas formas incluyen análisis predictivo y prescriptivo, así como tecnologías de machine learning y otras que están cambiando la forma en que los datos son administrados y analizados. El data warehouse ha sido durante mucho tiempo un elemento básico de la arquitectura de datos empresarial y sigue teniendo tiene un gran futuro en el nuevo mundo de Big Data y la analítica avanzada.

### 3. Data Warehouse en la nube

Los data warehouses están atravesando actualmente dos transformaciones muy importantes que tienen el potencial de impulsar niveles significativos de innovación empresarial:

- La primera área de transformación es el impulso para aumentar la agilidad general. La gran mayoría de los departamentos de TI están experimentando un rápido aumento de la demanda de datos. Los directivos quieren tener acceso a más y más datos históricos, mientras que al mismo tiempo, los científicos de datos y los analistas de negocios están explorando formas de introducir nuevos flujos de datos en el almacén para enriquecer el análisis existente, así como impulsar nuevas áreas de análisis. Esta rápida expansión de los volúmenes y fuentes de datos significa que los equipos de TI necesitan invertir más tiempo y esfuerzo asegurando que el rendimiento de las consultas permanezca constante y necesitan proporcionar cada vez más entornos para equipos individuales para validar el valor comercial de los nuevos conjuntos de datos.
- La segunda área de transformación gira en torno a la necesidad de mejorar el control
  de costes. Existe una creciente necesidad de hacer más con cada vez menos recursos, al
  mismo tiempo que se garantiza que todos los datos sensibles y estratégicos estén
  completamente asegurados, a lo largo de todo el ciclo de vida, de la manera más rentable.

### Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

La nube está demostrando ser un facilitador dominante. Permite a las organizaciones enfrentarse de forma activa a los desafíos que presentan estas dos transformaciones clave.

¿Por qué migrar a la nube?

Las 3 tecnologías que ya tienen y seguirán teniendo un gran impacto en la transformación de negocios son: la nube, datos y análisis, e Internet de las Cosas. Estas tres tecnologías son parte clave del ecosistema del data warehouse actual. Por lo tanto, es posible sacar la conclusión de que el data warehouse en la nube tiene un gran potencial para generar un impacto significativo en los negocios.

La importancia de la nube para el data warehouse está directamente vinculado a tres factores clave:

- Mayor agilidad
- Mejor control de costes
- Co-localización

Mejora de la agilidad

Muchos data warehouses están embarcados en una fase de actualización. Con mucho trabajo puesto ahora en torno al big data, las empresas buscan aprovechar los nuevos flujos de datos y nuevos tipos de análisis más ricos para apoyar e impulsar nuevas áreas, tales como:analítica de clientes de 360°, analisis predictivo, detección de fraude, análisis de IoT y el establecimiento de los datos como centro de beneficio. Muchos de estos proyectos requieren el aprovisionamiento de nuevos entornos de hardware y el despliegue de software. Es más rápido, más fácil y más eficiente iniciar estos nuevos proyectos centrados en los datos utilizando los servicios en la nube.

Mayor control de costes

Muchos equipos de TI están buscando maneras de consolidar los data marts existentes, cada uno ejecutándose en hardware dedicado o incluso en hardware propietario, en un único entorno integrado. Los servicios en la nube de calidad ofrecen la oportunidad perfecta para iniciar este tipo de proyectos. El cumplimiento legal no puede ser visto como un extra opcional cuando se planifica un traslado a la nube. Los activos de datos deben estar protegidos a lo largo de todo el ciclo de vida. Los servicios en la nube deben facilitar todo esto, siendo más rentables ya que todas las características de seguridad se pueden habilitar de forma predeterminada y mejorada, y actualizada de forma transparente.

#### Co-Localización para una carga más rápida

La mayoría de los data warehouses generan datos directamente desde aplicaciones clave, como entrada de pedidos, ventas, finanzas y fabricación. Por lo tanto, tiene mucho sentido ubicar conjuntamente el data warehouse junto con los sistemas fuente que ya se estén ejecutándose en la nube. La ubicación conjunta ofrece una carga de datos más rápida, lo que significa que los usuarios obtienen un acceso más oportuno a sus datos.

Principales ventajas de mover el data warehouse a la nube

Existen **tres ventajas principales** para mover un data warehouse a la nube, y estas están directamente vinculados a los tres controladores clave enumerados anteriormente:

- Más fácil consolidación y racionalización
- Monetización más rápida de los datos en la nube
- La nube ofrece mejor protección

## 4. Data Warehouse vs Big Data vs BI: cuales son las diferencias

En este apartado compararemos brevemente los conceptos Data Warehouse (DW), Big Data y Business Intelligence (BI).

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para tener una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

Se trata de tres conceptos completamente diferentes que tienen en común una nueva manera de lidiar con los datos, siempre teniendo en cuenta la existencia de una gran volumen de información en varios formatos que contribuyen, de forma estructurada o no estructurada, a la toma de decisiones estratégicas. El objetivo final de cualquiera de estas tecnologías es ofrecer una ventaja competitiva a las empresas, pero la forma en que se utiliza es la que marcará la diferencia.

Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

#### Big Data

Llamamos big data a un gran volumen de datos con una variedad, complejidad y velocidad de crecimiento enorme y que además tienen la característica de no ser estructurados. Eso significa que no son relacionales, estando además fuera del entorno corporativo. Es un tipo de tecnología que te permite analizar los datos en tiempo real y puede provenir de diferentes fuentes y formas, tales como mensajería instantánea, redes sociales, registros de grabaciones, imágenes, mensajes de correo electrónico, etc.

## Guía gratuita: "Del Bit... al Big Data", descubre todo lo que necesitas saber sobre Big Data y mucho más.

Para tener una idea de la importancia del big data, este mercado está creciendo cada año, alrededor de un 40%. La paradoja reside en el hecho de que cuanto mayor es la evolución del big data, mayor es la escasez de profesionales cualificados para satisfacer esa demanda. La expectativa es que en unos pocos años, haya un crecimiento aún más significativo para el uso de big data y por lo tanto un aumento en la demanda mano de obra especializada y de partners tecnológicos especializados en esta materia.

#### Data Warehouse

Por otro lado, un data warehouse almacena datos consolidados de diversas fuentes o sistemas de la empresa. Se trata de datos estructurados, que tiene como objetivo principal ser precisos y de alta calidad para de esta forma poder dar soporte a la toma de decisiones de la empresa. Se trata de conseguir tener todos los datos juntos para después poder dividirlos para hacer un análisis de determinados sectores o estrategias.

#### Business Intelligence

Un Business Intelligence (BI) es una especie de "cuello de botella" de los datos recogidos del data warehouse, que llegan de forma exacta y útil para ayudar a la toma de decisiones. Business Intelligence transforma los datos en información útil para analizar no sólo los negocios, sino también las principales estrategias corporativas.

## Guía gratuita "Integridad de datos y BI", descubre cómo sacarle el máximo partido a tu proyecto de Integridad de datos

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

## 5. Material complementario

Soluciones y recursos para Data Warehouse

- Informatica Cloud para Microsoft Azure SQL Data Warehouse
- Test Data Warehouse
- MDM y calidad de datos para el data warehouse

#### Guías

- Data Lake: Superando las limitaciones del Data Warehouse
- El data warehouse inteligente

#### Artículos

- ¿Qué es warehouse y qué debe contemplar un data warehouse moderno?
- Beneficios y soluciones para un Data Warehouse en la nube
- Lo que necesitas saber acerca de la arquitectura de un datawarehouse
- Data warehousing: 3 ocasiones de hacer las cosas bien
- ¿Qué es un Data Warehouse y cómo saber cuándo lo necesitas implementar?
- Cómo optimizar la arquitectura de un data warehouse
- Data warehouse concepts: tu glosario de almacén de datos
- Data Warehouse, el primer paso para una solución de BI
- Tutoriales sobre DataWarehouse en pdf
- Cómo gestiona Informatica los nuevos medios de almacenamiento de datos
- ¿Qué es un Data Warehouse para un director de ventas?
- Bl en línea: el reto actual de los data warehouses
- El poder de data warehouse: Informatica PowerCenter

### 2. Pasado y presente del Data Warehouse

Históricamente, los data warehouses se habían formado utilizando datos repetitivos estructurados que eran filtrados antes de entrar en el data warehouse. Sin embargo, en los últimos años, el data warehouse ha evolucionado debido a información contextual que ahora se puede adjuntar a los datos no estructurados y que también puede ser almacenada.

Aquellos primeros datos relacionales estructurados no podían ser mezclados y emparejados para temas analíticos con datos textuales no estructurados. Pero con el advenimiento de la contextualización, estos tipos de análisis ahora sí pueden hacerse de forma naturales y fácil.

En el data warehouse, datos no repetitivos, como los comentarios en una encuesta, correos electrónicos y conversaciones, se tratan de forma diferente a las ocurrencias repetitivas de datos, como el flujo de clics, mediciones o el procesamiento máquina o analógico. Los datos no repetitivos son datos basados en textos que fueron generados por la palabra escrita o hablada, leída y reformateada y, lo que es más importante, ahora puede ser contextualizada. Con el fin de extraer cualquier sentido de los datos no repetitivos para su uso en el Data Warehouse, deben tener el contexto de los datos establecidos.

En muchos casos, el contexto de los datos no repetitivos es más importante que los datos en sí. En cualquier caso, los datos no repetitivos no pueden utilizarse para la toma de decisiones hasta que se haya establecido el contexto.

Data Lakes y Data Warehouses: ¿mutuamente exclusivos o partners perfectos?

Los data lakes han surgido en el paisaje de Data Management en los últimos años, sin embargo, data lake no es necesariamente un reemplazo del data warehouse. En cambio, complementan los esfuerzos existentes y dan soporte al descubrimiento de nuevas preguntas. Una vez que se descubren esas preguntas se optimizan las respuestas. Y optimizar puede significar moverse fuera del data lake para ir a un data mart o al data warehouse.

### <u>:Resuelve tus dudas sobre el Data Lake</u> <u>con uno de nuestros expertos!</u>

Estas son algunas diferencias clave entre data lake y data warehouse:

- Datos: Un data warehouse sólo almacena datos que han sido modelados o estructurados, mientras que un Data Lake no hace acepción de datos. Lo almacena todo, estructurado, semiestructurado y no estructurado.
- **Procesamiento**: Antes de que una empresa pueda cargar datos en un data warehouse, primero debe darles forma y estructura, es decir, los datos deben ser modelados. Eso se llama schema-on-write. Con un data lake, sólo se cargan los datos sin procesar, tal y como están, y cuando esté listo para usar los datos, es cuando se le da forma y estructura. Eso se llama schema-on-read. Dos enfoques muy diferentes.
- Almacenamiento: Una de las principales características de las tecnologías de big data, como Hadoop, es que el coste de almacenamiento de datos es relativamente bajo en comparación con el de un data warehouse. Hay dos razones principales para esto: en primer lugar, Hadoop es software de código abierto, por lo que la concesión de licencias y el soporte de la comunidad es gratuito. Y segundo, Hadoop está diseñado para ser instalado en hardware de bajo coste.
- Agilidad: Un almacén de datos es un repositorio altamente estructurado, por definición. No
  es técnicamente difícil cambiar la estructura, pero puede tomar mucho tiempo dado todos
  los procesos de negocio que están vinculados a ella. Un data lake, por otro lado, carece de
  la estructura de un data warehouse, lo que da a los desarrolladores y a los científicos de
  datos la capacidad de configurar y reconfigurar fácilmente y en tiempo real sus modelos,
  consultas y aplicaciones.
- Seguridad: La tecnología del data warehouse existe desde hace décadas, mientras que la tecnología de big data (la base de un Data Lake) es relativamente nueva. Por lo tanto, la capacidad de asegurar datos en un data warehouse es mucho más madura que asegurar datos en un data lake. Cabe señalar, sin embargo, que se está realizando un importante esfuerzo en materia de seguridad en la actualidad en la industria de Big Data.

El data warehouse continúa evolucionando

Hay todavía un gran potencial en la evolución del data warehouse y su uso sigue adelante.

El procesamiento analítico clásico de datos basado en transacciones se realiza en el data warehouse como siempre se ha hecho. Nada ha cambiado ahí. Pero ahora puede hacerse analítica sobre datos contextualizados, y esa forma de análisis es nueva. La mayoría de las organizaciones, hasta ahora no habían podido basar su toma de decisiones en datos textuales no estructurados. Y ahora hay una nueva forma de análisis posible en un Data Warehouse: la posibilidad de la mezcla de análisis. El análisis combinado se realiza mediante una combinación de datos transaccionales estructurados y datos contextuales no estructurados.

Existen muchas otras formas de análisis que también son posibles. Estas formas incluyen análisis predictivo y prescriptivo, así como tecnologías de machine learning y otras que están cambiando la forma en que los datos son administrados y analizados. El data

warehouse ha sido durante mucho tiempo un elemento básico de la arquitectura de datos empresarial y sigue teniendo tiene un gran futuro en el nuevo mundo de Big Data y la analítica avanzada.

## 3. Data Warehouse en la nube

Los data warehouses están atravesando actualmente dos transformaciones muy importantes que tienen el potencial de impulsar niveles significativos de innovación empresarial:

- La primera área de transformación es el impulso para aumentar la agilidad general. La gran mayoría de los departamentos de TI están experimentando un rápido aumento de la demanda de datos. Los directivos quieren tener acceso a más y más datos históricos, mientras que al mismo tiempo, los científicos de datos y los analistas de negocios están explorando formas de introducir nuevos flujos de datos en el almacén para enriquecer el análisis existente, así como impulsar nuevas áreas de análisis. Esta rápida expansión de los volúmenes y fuentes de datos significa que los equipos de TI necesitan invertir más tiempo y esfuerzo asegurando que el rendimiento de las consultas permanezca constante y necesitan proporcionar cada vez más entornos para equipos individuales para validar el valor comercial de los nuevos conjuntos de datos.
- La segunda área de transformación gira en torno a la necesidad de mejorar el control de costes. Existe una creciente necesidad de hacer más con cada vez menos recursos, al mismo tiempo que se garantiza que todos los datos sensibles y estratégicos estén completamente asegurados, a lo largo de todo el ciclo de vida, de la manera más rentable.

Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

La nube está demostrando ser un facilitador dominante. Permite a las organizaciones enfrentarse de forma activa a los desafíos que presentan estas dos transformaciones clave.

¿Por qué migrar a la nube?

Las 3 tecnologías que ya tienen y seguirán teniendo un gran impacto en la transformación de negocios son: la nube, datos y análisis, e Internet de las Cosas. Estas tres tecnologías son parte clave del ecosistema del data warehouse actual. Por lo tanto, es posible sacar la conclusión de que el data warehouse en la nube tiene un gran potencial para generar un impacto significativo en los negocios.

La importancia de la nube para el data warehouse está directamente vinculado a tres factores clave:

Mayor agilidad

- Mejor control de costes
- Co-localización

## Mejora de la agilidad

Muchos data warehouses están embarcados en una fase de actualización. Con mucho trabajo puesto ahora en torno al big data, las empresas buscan aprovechar los nuevos flujos de datos y nuevos tipos de análisis más ricos para apoyar e impulsar nuevas áreas, tales como:analítica de clientes de 360°, analísis predictivo, detección de fraude, análisis de loT y el establecimiento de los datos como centro de beneficio. Muchos de estos proyectos requieren el aprovisionamiento de nuevos entornos de hardware y el despliegue de software. Es más rápido, más fácil y más eficiente iniciar estos nuevos proyectos centrados en los datos utilizando los servicios en la nube.

## Mayor control de costes

Muchos equipos de TI están buscando maneras de consolidar los data marts existentes, cada uno ejecutándose en hardware dedicado o incluso en hardware propietario, en un único entorno integrado. Los servicios en la nube de calidad ofrecen la oportunidad perfecta para iniciar este tipo de proyectos. El cumplimiento legal no puede ser visto como un extra opcional cuando se planifica un traslado a la nube. Los activos de datos deben estar protegidos a lo largo de todo el ciclo de vida. Los servicios en la nube deben facilitar todo esto, siendo más rentables ya que todas las características de seguridad se pueden habilitar de forma predeterminada y mejorada, y actualizada de forma transparente.

## Co-Localización para una carga más rápida

La mayoría de los data warehouses generan datos directamente desde aplicaciones clave, como entrada de pedidos, ventas, finanzas y fabricación. Por lo tanto, tiene mucho sentido ubicar conjuntamente el data warehouse junto con los sistemas fuente que ya se estén ejecutándose en la nube. La ubicación conjunta ofrece una carga de datos más rápida, lo que significa que los usuarios obtienen un acceso más oportuno a sus datos.

Principales ventajas de mover el data warehouse a la nube

Existen **tres ventajas principales** para mover un data warehouse a la nube, y estas están directamente vinculados a los tres controladores clave enumerados anteriormente:

- Más fácil consolidación y racionalización
- Monetización más rápida de los datos en la nube
- La nube ofrece mejor protección

# 4. Data Warehouse vs Big Data vs BI: cuales son las diferencias

En este apartado compararemos brevemente los conceptos Data Warehouse (DW), Big Data y Business Intelligence (BI).

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para

tener una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

Se trata de tres conceptos completamente diferentes que tienen en común una nueva manera de lidiar con los datos, siempre teniendo en cuenta la existencia de una gran volumen de información en varios formatos que contribuyen, de forma estructurada o no estructurada, a la toma de decisiones estratégicas. El objetivo final de cualquiera de estas tecnologías es ofrecer una ventaja competitiva a las empresas, pero la forma en que se utiliza es la que marcará la diferencia.

Descubre cómo ampliar tu negocio en la nube. Contacta con un experto en CLOUD!

## Big Data

Llamamos big data a un gran volumen de datos con una variedad, complejidad y velocidad de crecimiento enorme y que además tienen la característica de no ser estructurados. Eso significa que no son relacionales, estando además fuera del entorno corporativo. Es un tipo de tecnología que te permite analizar los datos en tiempo real y puede provenir de diferentes fuentes y formas, tales como mensajería instantánea, redes sociales, registros de grabaciones, imágenes, mensajes de correo electrónico, etc.

## Guía gratuita: "Del Bit... al Big Data", descubre todo lo que necesitas saber sobre Big Data y mucho más.

Para tener una idea de la importancia del big data, este mercado está creciendo cada año, alrededor de un 40%. La paradoja reside en el hecho de que cuanto mayor es la evolución del big data, mayor es la escasez de profesionales cualificados para satisfacer esa demanda. La expectativa es que en unos pocos años, haya un crecimiento aún más significativo para el uso de big data y por lo tanto un aumento en la demanda mano de obra especializada y de partners tecnológicos especializados en esta materia.

#### **Data Warehouse**

Por otro lado, un data warehouse almacena datos consolidados de diversas fuentes o sistemas de la empresa. Se trata de datos estructurados, que tiene como objetivo principal ser precisos y de alta calidad para de esta forma poder dar soporte a la toma de decisiones de la empresa. Se trata de conseguir tener todos los datos juntos para después poder dividirlos para hacer un análisis de determinados sectores o estrategias.

## **Business Intelligence**

Un Business Intelligence (BI) es una especie de "cuello de botella" de los datos recogidos del data warehouse, que llegan de forma exacta y útil para ayudar a la toma de decisiones. Business Intelligence transforma los datos en información útil para analizar no sólo los negocios, sino también las principales estrategias corporativas.

## Guía gratuita "Integridad de datos y BI", descubre cómo sacarle el máximo partido a tu proyecto de Integridad de datos

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

## 5. Material complementario

Soluciones y recursos para Data Warehouse

- Informatica Cloud para Microsoft Azure SQL Data Warehouse
- Test Data Warehouse
- MDM y calidad de datos para el data warehouse

#### Guías

- Data Lake: Superando las limitaciones del Data Warehouse
- El data warehouse inteligente

#### Artículos

- ¿Qué es warehouse y qué debe contemplar un data warehouse moderno?
- Beneficios y soluciones para un Data Warehouse en la nube
- Lo que necesitas saber acerca de la arquitectura de un datawarehouse
- Data warehousing: 3 ocasiones de hacer las cosas bien
- ¿Qué es un Data Warehouse y cómo saber cuándo lo necesitas implementar?
- Cómo optimizar la arquitectura de un data warehouse
- Data warehouse concepts: tu glosario de almacén de datos
- Data Warehouse, el primer paso para una solución de BI
- Tutoriales sobre DataWarehouse en pdf
- Cómo gestiona Informatica los nuevos medios de almacenamiento de datos
- ¿Qué es un Data Warehouse para un director de ventas?
- Bl en línea: el reto actual de los data warehouses
- El poder de data warehouse: Informatica PowerCenter

## 2. Pasado y presente del Data Warehouse

Históricamente, los data warehouses se habían formado utilizando datos repetitivos estructurados que eran filtrados antes de entrar en el data warehouse. Sin embargo, en los últimos años, el data warehouse ha evolucionado debido a información contextual que ahora se puede adjuntar a los datos no estructurados y que también puede ser almacenada.

Aquellos primeros datos relacionales estructurados no podían ser mezclados y emparejados para temas analíticos con datos textuales no estructurados. Pero con el advenimiento de la contextualización, estos tipos de análisis ahora sí pueden hacerse de forma naturales y fácil.

En el data warehouse, datos no repetitivos, como los comentarios en una encuesta, correos electrónicos y conversaciones, se tratan de forma diferente a las ocurrencias repetitivas de datos, como el flujo de clics, mediciones o el procesamiento máquina o analógico. Los datos no repetitivos son datos basados en textos que fueron generados por la palabra escrita o hablada, leída y reformateada y, lo que es más importante, ahora puede ser contextualizada. Con el fin de extraer cualquier sentido de los datos no repetitivos para su uso en el Data Warehouse, deben tener el contexto de los datos establecidos.

En muchos casos, el contexto de los datos no repetitivos es más importante que los datos en sí. En cualquier caso, los datos no repetitivos no pueden utilizarse para la toma de decisiones hasta que se haya establecido el contexto.

Data Lakes y Data Warehouses: ¿mutuamente exclusivos o partners perfectos?

Los data lakes han surgido en el paisaje de Data Management en los últimos años, sin embargo, data lake no es necesariamente un reemplazo del data warehouse. En cambio, complementan los esfuerzos existentes y dan soporte al descubrimiento de nuevas preguntas. Una vez que se descubren esas preguntas se optimizan las respuestas. Y optimizar puede significar moverse fuera del data lake para ir a un data mart o al data warehouse.

<u>:Resuelve tus dudas sobre el Data Lake</u> con uno de nuestros expertos!

Estas son algunas diferencias clave entre data lake y data warehouse:

- Datos: Un data warehouse sólo almacena datos que han sido modelados o estructurados, mientras que un Data Lake no hace acepción de datos. Lo almacena todo, estructurado, semiestructurado y no estructurado.
- **Procesamiento**: Antes de que una empresa pueda cargar datos en un data warehouse, primero debe darles forma y estructura, es decir, los datos deben ser modelados. Eso se llama schema-on-write. Con un data lake, sólo se cargan los datos sin procesar, tal y como están, y cuando esté listo para usar los datos, es cuando se le da forma y estructura. Eso se llama schema-on-read. Dos enfoques muy diferentes.
- Almacenamiento: Una de las principales características de las tecnologías de big data, como Hadoop, es que el coste de almacenamiento de datos es relativamente bajo en comparación con el de un data warehouse. Hay dos razones principales para esto: en primer lugar, Hadoop es software de código abierto, por lo que la concesión de licencias y el soporte de la comunidad es gratuito. Y segundo, Hadoop está diseñado para ser instalado en hardware de bajo coste.
- Agilidad: Un almacén de datos es un repositorio altamente estructurado, por definición. No
  es técnicamente difícil cambiar la estructura, pero puede tomar mucho tiempo dado todos
  los procesos de negocio que están vinculados a ella. Un data lake, por otro lado, carece de
  la estructura de un data warehouse, lo que da a los desarrolladores y a los científicos de
  datos la capacidad de configurar y reconfigurar fácilmente y en tiempo real sus modelos,
  consultas y aplicaciones.
- **Seguridad**: La tecnología del data warehouse existe desde hace décadas, mientras que la tecnología de big data (la base de un Data Lake) es relativamente nueva. Por lo tanto, la

capacidad de asegurar datos en un data warehouse es mucho más madura que asegurar datos en un data lake. Cabe señalar, sin embargo, que se está realizando un importante esfuerzo en materia de seguridad en la actualidad en la industria de Big Data.

El data warehouse continúa evolucionando

Hay todavía un gran potencial en la evolución del data warehouse y su uso sigue adelante.

El procesamiento analítico clásico de datos basado en transacciones se realiza en el data warehouse como siempre se ha hecho. Nada ha cambiado ahí. Pero ahora puede hacerse analítica sobre datos contextualizados, y esa forma de análisis es nueva. La mayoría de las organizaciones, hasta ahora no habían podido basar su toma de decisiones en datos textuales no estructurados. Y ahora hay una nueva forma de análisis posible en un Data Warehouse: la posibilidad de la mezcla de análisis. El análisis combinado se realiza mediante una combinación de datos transaccionales estructurados y datos contextuales no estructurados.

Existen muchas otras formas de análisis que también son posibles. Estas formas incluyen análisis predictivo y prescriptivo, así como tecnologías de machine learning y otras que están cambiando la forma en que los datos son administrados y analizados. El data warehouse ha sido durante mucho tiempo un elemento básico de la arquitectura de datos empresarial y sigue teniendo tiene un gran futuro en el nuevo mundo de Big Data y la analítica avanzada.

## 3. Data Warehouse en la nube

Los data warehouses están atravesando actualmente dos transformaciones muy importantes que tienen el potencial de impulsar niveles significativos de innovación empresarial:

- La primera área de transformación es el impulso para aumentar la agilidad general. La gran mayoría de los departamentos de TI están experimentando un rápido aumento de la demanda de datos. Los directivos quieren tener acceso a más y más datos históricos, mientras que al mismo tiempo, los científicos de datos y los analistas de negocios están explorando formas de introducir nuevos flujos de datos en el almacén para enriquecer el análisis existente, así como impulsar nuevas áreas de análisis. Esta rápida expansión de los volúmenes y fuentes de datos significa que los equipos de TI necesitan invertir más tiempo y esfuerzo asegurando que el rendimiento de las consultas permanezca constante y necesitan proporcionar cada vez más entornos para equipos individuales para validar el valor comercial de los nuevos conjuntos de datos.
- La segunda área de transformación gira en torno a la necesidad de mejorar el control de costes. Existe una creciente necesidad de hacer más con cada vez menos recursos, al mismo tiempo que se garantiza que todos los datos sensibles y estratégicos estén completamente asegurados, a lo largo de todo el ciclo de vida, de la manera más rentable.

## Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

La nube está demostrando ser un facilitador dominante. Permite a las organizaciones enfrentarse de forma activa a los desafíos que presentan estas dos transformaciones clave.

¿Por qué migrar a la nube?

Las 3 tecnologías que ya tienen y seguirán teniendo un gran impacto en la transformación de negocios son: la nube, datos y análisis, e Internet de las Cosas. Estas tres tecnologías son parte clave del ecosistema del data warehouse actual. Por lo tanto, es posible sacar la conclusión de que el data warehouse en la nube tiene un gran potencial para generar un impacto significativo en los negocios.

La importancia de la nube para el data warehouse está directamente vinculado a tres factores clave:

- Mayor agilidad
- Mejor control de costes
- Co-localización

Mejora de la agilidad

Muchos data warehouses están embarcados en una fase de actualización. Con mucho trabajo puesto ahora en torno al big data, las empresas buscan aprovechar los nuevos flujos de datos y nuevos tipos de análisis más ricos para apoyar e impulsar nuevas áreas, tales como:analítica de clientes de 360°, analisis predictivo, detección de fraude, análisis de IoT y el establecimiento de los datos como centro de beneficio. Muchos de estos proyectos requieren el aprovisionamiento de nuevos entornos de hardware y el despliegue de software. Es más rápido, más fácil y más eficiente iniciar estos nuevos proyectos centrados en los datos utilizando los servicios en la nube.

Mayor control de costes

Muchos equipos de TI están buscando maneras de consolidar los data marts existentes, cada uno ejecutándose en hardware dedicado o incluso en hardware propietario, en un único entorno integrado. Los servicios en la nube de calidad ofrecen la oportunidad perfecta para iniciar este tipo de proyectos. El cumplimiento legal no puede ser visto como un extra opcional cuando se planifica un traslado a la nube. Los activos de datos deben estar protegidos a lo largo de todo el ciclo de vida. Los servicios en la nube deben facilitar todo esto, siendo más rentables ya que todas las características de seguridad se pueden habilitar de forma predeterminada y mejorada, y actualizada de forma transparente.

## Co-Localización para una carga más rápida

La mayoría de los data warehouses generan datos directamente desde aplicaciones clave, como entrada de pedidos, ventas, finanzas y fabricación. Por lo tanto, **tiene mucho sentido ubicar conjuntamente el data warehouse junto con los sistemas fuente que ya se estén ejecutándose en la nube**. La ubicación conjunta ofrece una carga de datos más rápida, lo que significa que los usuarios obtienen un acceso más oportuno a sus datos.

Principales ventajas de mover el data warehouse a la nube

Existen **tres ventajas principales** para mover un data warehouse a la nube, y estas están directamente vinculados a los tres controladores clave enumerados anteriormente:

- Más fácil consolidación y racionalización
- Monetización más rápida de los datos en la nube
- La nube ofrece mejor protección

# 4. Data Warehouse vs Big Data vs BI: cuales son las diferencias

En este apartado compararemos brevemente los conceptos Data Warehouse (DW), Big Data y Business Intelligence (BI).

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para tener una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

Se trata de tres conceptos completamente diferentes que tienen en común una nueva manera de lidiar con los datos, siempre teniendo en cuenta la existencia de una gran volumen de información en varios formatos que contribuyen, de forma estructurada o no estructurada, a la toma de decisiones estratégicas. El objetivo final de cualquiera de estas tecnologías es ofrecer una ventaja competitiva a las empresas, pero la forma en que se utiliza es la que marcará la diferencia.

Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

#### Big Data

Llamamos big data a un gran volumen de datos con una variedad, complejidad y velocidad de crecimiento enorme y que además tienen la característica de no ser estructurados. Eso significa que no son relacionales, estando además fuera del entorno corporativo. Es un tipo de tecnología que te permite analizar los datos en tiempo real y puede provenir de diferentes fuentes y formas, tales como mensajería instantánea, redes sociales, registros de grabaciones, imágenes, mensajes de correo electrónico, etc.

## Guía gratuita: "Del Bit... al Big Data", descubre todo lo que necesitas saber sobre Big Data y mucho más.

Para tener una idea de la importancia del big data, este mercado está creciendo cada año, alrededor de un 40%. La paradoja reside en el hecho de que cuanto mayor es la evolución del big data, mayor es la escasez de profesionales cualificados para satisfacer esa demanda. La expectativa es que en unos pocos años, haya un crecimiento aún más significativo para el uso de big data y por lo tanto un aumento en la demanda mano de obra especializada y de partners tecnológicos especializados en esta materia.

#### Data Warehouse

Por otro lado, un data warehouse almacena datos consolidados de diversas fuentes o sistemas de la empresa. Se trata de datos estructurados, que tiene como objetivo principal ser precisos y de alta calidad para de esta forma poder dar soporte a la toma de decisiones de la empresa. Se trata de conseguir tener todos los datos juntos para después poder dividirlos para hacer un análisis de determinados sectores o estrategias.

## Business Intelligence

Un Business Intelligence (BI) es una especie de "cuello de botella" de los datos recogidos del data warehouse, que llegan de forma exacta y útil para ayudar a la toma de decisiones. Business Intelligence transforma los datos en información útil para analizar no sólo los negocios, sino también las principales estrategias corporativas.

# Guía gratuita "Integridad de datos y BI", descubre cómo sacarle el máximo partido a tu proyecto de Integridad de datos

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

## 5. Material complementario

Soluciones y recursos para Data Warehouse

- Informatica Cloud para Microsoft Azure SQL Data Warehouse
- Test Data Warehouse
- MDM y calidad de datos para el data warehouse

### Guías

- Data Lake: Superando las limitaciones del Data Warehouse
- El data warehouse inteligente

#### Artículos

- ¿Qué es warehouse y qué debe contemplar un data warehouse moderno?
- Beneficios y soluciones para un Data Warehouse en la nube
- Lo que necesitas saber acerca de la arquitectura de un datawarehouse
- Data warehousing: 3 ocasiones de hacer las cosas bien
- ¿Qué es un Data Warehouse y cómo saber cuándo lo necesitas implementar?
- Cómo optimizar la arquitectura de un data warehouse
- Data warehouse concepts: tu glosario de almacén de datos
- Data Warehouse, el primer paso para una solución de BI
- Tutoriales sobre DataWarehouse en pdf
- Cómo gestiona Informatica los nuevos medios de almacenamiento de datos
- ¿Qué es un Data Warehouse para un director de ventas?
- Bl en línea: el reto actual de los data warehouses
- El poder de data warehouse: Informatica PowerCenter

## 2. Pasado y presente del Data Warehouse

Históricamente, los data warehouses se habían formado utilizando datos repetitivos estructurados que eran filtrados antes de entrar en el data warehouse. Sin embargo, en los últimos años, el data warehouse ha evolucionado debido a información contextual que ahora se puede adjuntar a los datos no estructurados y que también puede ser almacenada.

Aquellos primeros datos relacionales estructurados no podían ser mezclados y emparejados para temas analíticos con datos textuales no estructurados. Pero con el advenimiento de la contextualización, estos tipos de análisis ahora sí pueden hacerse de forma naturales y fácil.

En el data warehouse, datos no repetitivos, como los comentarios en una encuesta, correos electrónicos y conversaciones, se tratan de forma diferente a las ocurrencias repetitivas de datos, como el flujo de clics, mediciones o el procesamiento máquina o analógico. Los datos no repetitivos son datos basados en textos que fueron generados por la palabra escrita o hablada, leída y reformateada y, lo que es más importante, ahora puede ser contextualizada. Con el fin de extraer cualquier sentido de los datos no repetitivos para su uso en el Data Warehouse, deben tener el contexto de los datos establecidos.

En muchos casos, el contexto de los datos no repetitivos es más importante que los datos en sí. En cualquier caso, los datos no repetitivos no pueden utilizarse para la toma de decisiones hasta que se haya establecido el contexto.

Data Lakes y Data Warehouses: ¿mutuamente exclusivos o partners perfectos?

Los data lakes han surgido en el paisaje de Data Management en los últimos años, sin embargo, data lake no es necesariamente un reemplazo del data warehouse. En cambio, complementan los esfuerzos existentes y dan soporte al descubrimiento de nuevas preguntas. Una vez que se descubren esas preguntas se optimizan las respuestas. Y optimizar puede significar moverse fuera del data lake para ir a un data mart o al data warehouse.

## <u>:Resuelve tus dudas sobre el Data Lake</u> con uno de nuestros expertos!

Estas son algunas diferencias clave entre data lake y data warehouse:

- Datos: Un data warehouse sólo almacena datos que han sido modelados o estructurados, mientras que un Data Lake no hace acepción de datos. Lo almacena todo, estructurado, semiestructurado y no estructurado.
- **Procesamiento**: Antes de que una empresa pueda cargar datos en un data warehouse, primero debe darles forma y estructura, es decir, los datos deben ser modelados. Eso se llama schema-on-write. Con un data lake, sólo se cargan los datos sin procesar, tal y como están, y cuando esté listo para usar los datos, es cuando se le da forma y estructura. Eso se llama schema-on-read. Dos enfoques muy diferentes.
- Almacenamiento: Una de las principales características de las tecnologías de big data, como Hadoop, es que el coste de almacenamiento de datos es relativamente bajo en comparación con el de un data warehouse. Hay dos razones principales para esto: en primer lugar, Hadoop es software de código abierto, por lo que la concesión de licencias y el soporte de la comunidad es gratuito. Y segundo, Hadoop está diseñado para ser instalado en hardware de bajo coste.
- Agilidad: Un almacén de datos es un repositorio altamente estructurado, por definición. No
  es técnicamente difícil cambiar la estructura, pero puede tomar mucho tiempo dado todos
  los procesos de negocio que están vinculados a ella. Un data lake, por otro lado, carece de
  la estructura de un data warehouse, lo que da a los desarrolladores y a los científicos de
  datos la capacidad de configurar y reconfigurar fácilmente y en tiempo real sus modelos,
  consultas y aplicaciones.
- Seguridad: La tecnología del data warehouse existe desde hace décadas, mientras que la tecnología de big data (la base de un Data Lake) es relativamente nueva. Por lo tanto, la capacidad de asegurar datos en un data warehouse es mucho más madura que asegurar datos en un data lake. Cabe señalar, sin embargo, que se está realizando un importante esfuerzo en materia de seguridad en la actualidad en la industria de Big Data.

El data warehouse continúa evolucionando

Hay todavía un gran potencial en la evolución del data warehouse y su uso sigue adelante.

El procesamiento analítico clásico de datos basado en transacciones se realiza en el data warehouse como siempre se ha hecho. Nada ha cambiado ahí. Pero ahora puede hacerse analítica sobre datos contextualizados, y esa forma de análisis es nueva. La mayoría de las organizaciones, hasta ahora no habían podido basar su toma de decisiones en datos textuales no estructurados. Y ahora hay una nueva forma de análisis posible en un Data Warehouse: la posibilidad de la mezcla de análisis. El análisis combinado se realiza mediante una combinación de datos transaccionales estructurados y datos contextuales no estructurados.

Existen muchas otras formas de análisis que también son posibles. Estas formas incluyen análisis predictivo y prescriptivo, así como tecnologías de machine learning y otras que están cambiando la forma en que los datos son administrados y analizados. El data

warehouse ha sido durante mucho tiempo un elemento básico de la arquitectura de datos empresarial y sigue teniendo tiene un gran futuro en el nuevo mundo de Big Data y la analítica avanzada.

## 3. Data Warehouse en la nube

Los data warehouses están atravesando actualmente dos transformaciones muy importantes que tienen el potencial de impulsar niveles significativos de innovación empresarial:

- La primera área de transformación es el impulso para aumentar la agilidad general. La gran mayoría de los departamentos de TI están experimentando un rápido aumento de la demanda de datos. Los directivos quieren tener acceso a más y más datos históricos, mientras que al mismo tiempo, los científicos de datos y los analistas de negocios están explorando formas de introducir nuevos flujos de datos en el almacén para enriquecer el análisis existente, así como impulsar nuevas áreas de análisis. Esta rápida expansión de los volúmenes y fuentes de datos significa que los equipos de TI necesitan invertir más tiempo y esfuerzo asegurando que el rendimiento de las consultas permanezca constante y necesitan proporcionar cada vez más entornos para equipos individuales para validar el valor comercial de los nuevos conjuntos de datos.
- La segunda área de transformación gira en torno a la necesidad de mejorar el control de costes. Existe una creciente necesidad de hacer más con cada vez menos recursos, al mismo tiempo que se garantiza que todos los datos sensibles y estratégicos estén completamente asegurados, a lo largo de todo el ciclo de vida, de la manera más rentable.

Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

La nube está demostrando ser un facilitador dominante. Permite a las organizaciones enfrentarse de forma activa a los desafíos que presentan estas dos transformaciones clave.

¿Por qué migrar a la nube?

Las 3 tecnologías que ya tienen y seguirán teniendo un gran impacto en la transformación de negocios son: la nube, datos y análisis, e Internet de las Cosas. Estas tres tecnologías son parte clave del ecosistema del data warehouse actual. Por lo tanto, es posible sacar la conclusión de que el data warehouse en la nube tiene un gran potencial para generar un impacto significativo en los negocios.

La importancia de la nube para el data warehouse está directamente vinculado a tres factores clave:

Mayor agilidad

- Mejor control de costes
- Co-localización

## Mejora de la agilidad

Muchos data warehouses están embarcados en una fase de actualización. Con mucho trabajo puesto ahora en torno al big data, las empresas buscan aprovechar los nuevos flujos de datos y nuevos tipos de análisis más ricos para apoyar e impulsar nuevas áreas, tales como:analítica de clientes de 360°, analísis predictivo, detección de fraude, análisis de loT y el establecimiento de los datos como centro de beneficio. Muchos de estos proyectos requieren el aprovisionamiento de nuevos entornos de hardware y el despliegue de software. Es más rápido, más fácil y más eficiente iniciar estos nuevos proyectos centrados en los datos utilizando los servicios en la nube.

## Mayor control de costes

Muchos equipos de TI están buscando maneras de consolidar los data marts existentes, cada uno ejecutándose en hardware dedicado o incluso en hardware propietario, en un único entorno integrado. Los servicios en la nube de calidad ofrecen la oportunidad perfecta para iniciar este tipo de proyectos. El cumplimiento legal no puede ser visto como un extra opcional cuando se planifica un traslado a la nube. Los activos de datos deben estar protegidos a lo largo de todo el ciclo de vida. Los servicios en la nube deben facilitar todo esto, siendo más rentables ya que todas las características de seguridad se pueden habilitar de forma predeterminada y mejorada, y actualizada de forma transparente.

## Co-Localización para una carga más rápida

La mayoría de los data warehouses generan datos directamente desde aplicaciones clave, como entrada de pedidos, ventas, finanzas y fabricación. Por lo tanto, tiene mucho sentido ubicar conjuntamente el data warehouse junto con los sistemas fuente que ya se estén ejecutándose en la nube. La ubicación conjunta ofrece una carga de datos más rápida, lo que significa que los usuarios obtienen un acceso más oportuno a sus datos.

Principales ventajas de mover el data warehouse a la nube

Existen **tres ventajas principales** para mover un data warehouse a la nube, y estas están directamente vinculados a los tres controladores clave enumerados anteriormente:

- Más fácil consolidación y racionalización
- Monetización más rápida de los datos en la nube
- La nube ofrece mejor protección

# 4. Data Warehouse vs Big Data vs BI: cuales son las diferencias

En este apartado compararemos brevemente los conceptos Data Warehouse (DW), Big Data y Business Intelligence (BI).

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para

tener una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

Se trata de tres conceptos completamente diferentes que tienen en común una nueva manera de lidiar con los datos, siempre teniendo en cuenta la existencia de una gran volumen de información en varios formatos que contribuyen, de forma estructurada o no estructurada, a la toma de decisiones estratégicas. El objetivo final de cualquiera de estas tecnologías es ofrecer una ventaja competitiva a las empresas, pero la forma en que se utiliza es la que marcará la diferencia.

Descubre cómo ampliar tu negocio en la nube. Contacta con un experto en CLOUD!

## Big Data

Llamamos big data a un gran volumen de datos con una variedad, complejidad y velocidad de crecimiento enorme y que además tienen la característica de no ser estructurados. Eso significa que no son relacionales, estando además fuera del entorno corporativo. Es un tipo de tecnología que te permite analizar los datos en tiempo real y puede provenir de diferentes fuentes y formas, tales como mensajería instantánea, redes sociales, registros de grabaciones, imágenes, mensajes de correo electrónico, etc.

## Guía gratuita: "Del Bit... al Big Data", descubre todo lo que necesitas saber sobre Big Data y mucho más.

Para tener una idea de la importancia del big data, este mercado está creciendo cada año, alrededor de un 40%. La paradoja reside en el hecho de que cuanto mayor es la evolución del big data, mayor es la escasez de profesionales cualificados para satisfacer esa demanda. La expectativa es que en unos pocos años, haya un crecimiento aún más significativo para el uso de big data y por lo tanto un aumento en la demanda mano de obra especializada y de partners tecnológicos especializados en esta materia.

#### **Data Warehouse**

Por otro lado, un data warehouse almacena datos consolidados de diversas fuentes o sistemas de la empresa. Se trata de datos estructurados, que tiene como objetivo principal ser precisos y de alta calidad para de esta forma poder dar soporte a la toma de decisiones de la empresa. Se trata de conseguir tener todos los datos juntos para después poder dividirlos para hacer un análisis de determinados sectores o estrategias.

## **Business Intelligence**

Un Business Intelligence (BI) es una especie de "cuello de botella" de los datos recogidos del data warehouse, que llegan de forma exacta y útil para ayudar a la toma de decisiones. Business Intelligence transforma los datos en información útil para analizar no sólo los negocios, sino también las principales estrategias corporativas.

## Guía gratuita "Integridad de datos y BI", descubre cómo sacarle el máximo partido a tu proyecto de Integridad de datos

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

## 5. Material complementario

Soluciones y recursos para Data Warehouse

- Informatica Cloud para Microsoft Azure SQL Data Warehouse
- Test Data Warehouse
- MDM y calidad de datos para el data warehouse

#### Guías

- Data Lake: Superando las limitaciones del Data Warehouse
- El data warehouse inteligente

#### Artículos

- ¿Qué es warehouse y qué debe contemplar un data warehouse moderno?
- Beneficios y soluciones para un Data Warehouse en la nube
- Lo que necesitas saber acerca de la arquitectura de un datawarehouse
- Data warehousing: 3 ocasiones de hacer las cosas bien
- ¿Qué es un Data Warehouse y cómo saber cuándo lo necesitas implementar?
- Cómo optimizar la arquitectura de un data warehouse
- Data warehouse concepts: tu glosario de almacén de datos
- Data Warehouse, el primer paso para una solución de BI
- Tutoriales sobre DataWarehouse en pdf
- Cómo gestiona Informatica los nuevos medios de almacenamiento de datos
- ¿Qué es un Data Warehouse para un director de ventas?
- Bl en línea: el reto actual de los data warehouses
- El poder de data warehouse: Informatica PowerCenter

## 2. Pasado y presente del Data Warehouse

Históricamente, los data warehouses se habían formado utilizando datos repetitivos estructurados que eran filtrados antes de entrar en el data warehouse. Sin embargo, en los últimos años, el data warehouse ha evolucionado debido a información contextual que ahora se puede adjuntar a los datos no estructurados y que también puede ser almacenada.

Aquellos primeros datos relacionales estructurados no podían ser mezclados y emparejados para temas analíticos con datos textuales no estructurados. Pero con el advenimiento de la contextualización, estos tipos de análisis ahora sí pueden hacerse de forma naturales y fácil.

En el data warehouse, datos no repetitivos, como los comentarios en una encuesta, correos electrónicos y conversaciones, se tratan de forma diferente a las ocurrencias repetitivas de datos, como el flujo de clics, mediciones o el procesamiento máquina o analógico. Los datos no repetitivos son datos basados en textos que fueron generados por la palabra escrita o hablada, leída y reformateada y, lo que es más importante, ahora puede ser contextualizada. Con el fin de extraer cualquier sentido de los datos no repetitivos para su uso en el Data Warehouse, deben tener el contexto de los datos establecidos.

En muchos casos, el contexto de los datos no repetitivos es más importante que los datos en sí. En cualquier caso, los datos no repetitivos no pueden utilizarse para la toma de decisiones hasta que se haya establecido el contexto.

Data Lakes y Data Warehouses: ¿mutuamente exclusivos o partners perfectos?

Los data lakes han surgido en el paisaje de Data Management en los últimos años, sin embargo, data lake no es necesariamente un reemplazo del data warehouse. En cambio, complementan los esfuerzos existentes y dan soporte al descubrimiento de nuevas preguntas. Una vez que se descubren esas preguntas se optimizan las respuestas. Y optimizar puede significar moverse fuera del data lake para ir a un data mart o al data warehouse.

<u>:Resuelve tus dudas sobre el Data Lake</u> con uno de nuestros expertos!

Estas son algunas diferencias clave entre data lake y data warehouse:

- Datos: Un data warehouse sólo almacena datos que han sido modelados o estructurados, mientras que un Data Lake no hace acepción de datos. Lo almacena todo, estructurado, semiestructurado y no estructurado.
- **Procesamiento**: Antes de que una empresa pueda cargar datos en un data warehouse, primero debe darles forma y estructura, es decir, los datos deben ser modelados. Eso se llama schema-on-write. Con un data lake, sólo se cargan los datos sin procesar, tal y como están, y cuando esté listo para usar los datos, es cuando se le da forma y estructura. Eso se llama schema-on-read. Dos enfoques muy diferentes.
- Almacenamiento: Una de las principales características de las tecnologías de big data, como Hadoop, es que el coste de almacenamiento de datos es relativamente bajo en comparación con el de un data warehouse. Hay dos razones principales para esto: en primer lugar, Hadoop es software de código abierto, por lo que la concesión de licencias y el soporte de la comunidad es gratuito. Y segundo, Hadoop está diseñado para ser instalado en hardware de bajo coste.
- Agilidad: Un almacén de datos es un repositorio altamente estructurado, por definición. No
  es técnicamente difícil cambiar la estructura, pero puede tomar mucho tiempo dado todos
  los procesos de negocio que están vinculados a ella. Un data lake, por otro lado, carece de
  la estructura de un data warehouse, lo que da a los desarrolladores y a los científicos de
  datos la capacidad de configurar y reconfigurar fácilmente y en tiempo real sus modelos,
  consultas y aplicaciones.
- **Seguridad**: La tecnología del data warehouse existe desde hace décadas, mientras que la tecnología de big data (la base de un Data Lake) es relativamente nueva. Por lo tanto, la

capacidad de asegurar datos en un data warehouse es mucho más madura que asegurar datos en un data lake. Cabe señalar, sin embargo, que se está realizando un importante esfuerzo en materia de seguridad en la actualidad en la industria de Big Data.

El data warehouse continúa evolucionando

Hay todavía un gran potencial en la evolución del data warehouse y su uso sigue adelante.

El procesamiento analítico clásico de datos basado en transacciones se realiza en el data warehouse como siempre se ha hecho. Nada ha cambiado ahí. Pero ahora puede hacerse analítica sobre datos contextualizados, y esa forma de análisis es nueva. La mayoría de las organizaciones, hasta ahora no habían podido basar su toma de decisiones en datos textuales no estructurados. Y ahora hay una nueva forma de análisis posible en un Data Warehouse: la posibilidad de la mezcla de análisis. El análisis combinado se realiza mediante una combinación de datos transaccionales estructurados y datos contextuales no estructurados.

Existen muchas otras formas de análisis que también son posibles. Estas formas incluyen análisis predictivo y prescriptivo, así como tecnologías de machine learning y otras que están cambiando la forma en que los datos son administrados y analizados. El data warehouse ha sido durante mucho tiempo un elemento básico de la arquitectura de datos empresarial y sigue teniendo tiene un gran futuro en el nuevo mundo de Big Data y la analítica avanzada.

## 3. Data Warehouse en la nube

Los data warehouses están atravesando actualmente dos transformaciones muy importantes que tienen el potencial de impulsar niveles significativos de innovación empresarial:

- La primera área de transformación es el impulso para aumentar la agilidad general. La gran mayoría de los departamentos de TI están experimentando un rápido aumento de la demanda de datos. Los directivos quieren tener acceso a más y más datos históricos, mientras que al mismo tiempo, los científicos de datos y los analistas de negocios están explorando formas de introducir nuevos flujos de datos en el almacén para enriquecer el análisis existente, así como impulsar nuevas áreas de análisis. Esta rápida expansión de los volúmenes y fuentes de datos significa que los equipos de TI necesitan invertir más tiempo y esfuerzo asegurando que el rendimiento de las consultas permanezca constante y necesitan proporcionar cada vez más entornos para equipos individuales para validar el valor comercial de los nuevos conjuntos de datos.
- La segunda área de transformación gira en torno a la necesidad de mejorar el control de costes. Existe una creciente necesidad de hacer más con cada vez menos recursos, al mismo tiempo que se garantiza que todos los datos sensibles y estratégicos estén completamente asegurados, a lo largo de todo el ciclo de vida, de la manera más rentable.

## Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

La nube está demostrando ser un facilitador dominante. Permite a las organizaciones enfrentarse de forma activa a los desafíos que presentan estas dos transformaciones clave.

¿Por qué migrar a la nube?

Las 3 tecnologías que ya tienen y seguirán teniendo un gran impacto en la transformación de negocios son: la nube, datos y análisis, e Internet de las Cosas. Estas tres tecnologías son parte clave del ecosistema del data warehouse actual. Por lo tanto, es posible sacar la conclusión de que el data warehouse en la nube tiene un gran potencial para generar un impacto significativo en los negocios.

La importancia de la nube para el data warehouse está directamente vinculado a tres factores clave:

- Mayor agilidad
- Mejor control de costes
- Co-localización

Mejora de la agilidad

Muchos data warehouses están embarcados en una fase de actualización. Con mucho trabajo puesto ahora en torno al big data, **las empresas buscan aprovechar los nuevos flujos de datos y nuevos tipos de análisis más ricos para apoyar e impulsar nuevas áreas**, tales como:analítica de clientes de 360°, analísis predictivo, detección de fraude, análisis de loT y el establecimiento de los datos como centro de beneficio. Muchos de estos proyectos requieren el aprovisionamiento de nuevos entornos de hardware y el despliegue de software. **Es más rápido, más fácil y más eficiente iniciar estos nuevos proyectos centrados en los datos utilizando los servicios en la nube**.

Mayor control de costes

Muchos equipos de TI están buscando maneras de consolidar los data marts existentes, cada uno ejecutándose en hardware dedicado o incluso en hardware propietario, en un único entorno integrado. Los servicios en la nube de calidad ofrecen la oportunidad perfecta para iniciar este tipo de proyectos. El cumplimiento legal no puede ser visto como un extra opcional cuando se planifica un traslado a la nube. Los activos de datos deben estar protegidos a lo largo de todo el ciclo de vida. Los servicios en la nube deben facilitar todo esto, siendo más rentables ya que todas las características de seguridad se pueden habilitar de forma predeterminada y mejorada, y actualizada de forma transparente.

## Co-Localización para una carga más rápida

La mayoría de los data warehouses generan datos directamente desde aplicaciones clave, como entrada de pedidos, ventas, finanzas y fabricación. Por lo tanto, **tiene mucho sentido ubicar conjuntamente el data warehouse junto con los sistemas fuente que ya se estén ejecutándose en la nube**. La ubicación conjunta ofrece una carga de datos más rápida, lo que significa que los usuarios obtienen un acceso más oportuno a sus datos.

Principales ventajas de mover el data warehouse a la nube

Existen **tres ventajas principales** para mover un data warehouse a la nube, y estas están directamente vinculados a los tres controladores clave enumerados anteriormente:

- Más fácil consolidación y racionalización
- Monetización más rápida de los datos en la nube
- La nube ofrece mejor protección

# 4. Data Warehouse vs Big Data vs BI: cuales son las diferencias

En este apartado compararemos brevemente los conceptos Data Warehouse (DW), Big Data y Business Intelligence (BI).

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para tener una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

Se trata de tres conceptos completamente diferentes que tienen en común una nueva manera de lidiar con los datos, siempre teniendo en cuenta la existencia de una gran volumen de información en varios formatos que contribuyen, de forma estructurada o no estructurada, a la toma de decisiones estratégicas. El objetivo final de cualquiera de estas tecnologías es ofrecer una ventaja competitiva a las empresas, pero la forma en que se utiliza es la que marcará la diferencia.

Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

Big Data

Llamamos big data a un gran volumen de datos con una variedad, complejidad y velocidad de crecimiento enorme y que además tienen la característica de no ser estructurados. Eso significa que no son relacionales, estando además fuera del entorno corporativo. Es un tipo de tecnología que te permite analizar los datos en tiempo real y puede provenir de diferentes fuentes y formas, tales como mensajería instantánea, redes sociales, registros de grabaciones, imágenes, mensajes de correo electrónico, etc.

## Guía gratuita: "Del Bit... al Big Data", descubre todo lo que necesitas saber sobre Big Data y mucho más.

Para tener una idea de la importancia del big data, este mercado está creciendo cada año, alrededor de un 40%. La paradoja reside en el hecho de que cuanto mayor es la evolución del big data, mayor es la escasez de profesionales cualificados para satisfacer esa demanda. La expectativa es que en unos pocos años, haya un crecimiento aún más significativo para el uso de big data y por lo tanto un aumento en la demanda mano de obra especializada y de partners tecnológicos especializados en esta materia.

#### Data Warehouse

Por otro lado, un data warehouse almacena datos consolidados de diversas fuentes o sistemas de la empresa. Se trata de datos estructurados, que tiene como objetivo principal ser precisos y de alta calidad para de esta forma poder dar soporte a la toma de decisiones de la empresa. Se trata de conseguir tener todos los datos juntos para después poder dividirlos para hacer un análisis de determinados sectores o estrategias.

## Business Intelligence

Un Business Intelligence (BI) es una especie de "cuello de botella" de los datos recogidos del data warehouse, que llegan de forma exacta y útil para ayudar a la toma de decisiones. Business Intelligence transforma los datos en información útil para analizar no sólo los negocios, sino también las principales estrategias corporativas.

# Guía gratuita "Integridad de datos y BI", descubre cómo sacarle el máximo partido a tu proyecto de Integridad de datos

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

## 5. Material complementario

Soluciones y recursos para Data Warehouse

- Informatica Cloud para Microsoft Azure SQL Data Warehouse
- Test Data Warehouse
- MDM y calidad de datos para el data warehouse

### Guías

- Data Lake: Superando las limitaciones del Data Warehouse
- El data warehouse inteligente

### Artículos

- ¿Qué es warehouse y qué debe contemplar un data warehouse moderno?
- Beneficios y soluciones para un Data Warehouse en la nube
- Lo que necesitas saber acerca de la arquitectura de un datawarehouse
- Data warehousing: 3 ocasiones de hacer las cosas bien
- ¿Qué es un Data Warehouse y cómo saber cuándo lo necesitas implementar?
- Cómo optimizar la arquitectura de un data warehouse
- Data warehouse concepts: tu glosario de almacén de datos
- Data Warehouse, el primer paso para una solución de BI
- Tutoriales sobre DataWarehouse en pdf
- Cómo gestiona Informatica los nuevos medios de almacenamiento de datos
- ¿Qué es un Data Warehouse para un director de ventas?
- Bl en línea: el reto actual de los data warehouses
- El poder de data warehouse: Informatica PowerCenter

## 2. Pasado y presente del Data Warehouse

Históricamente, los data warehouses se habían formado utilizando datos repetitivos estructurados que eran filtrados antes de entrar en el data warehouse. Sin embargo, en los últimos años, el data warehouse ha evolucionado debido a información contextual que ahora se puede adjuntar a los datos no estructurados y que también puede ser almacenada.

Aquellos primeros datos relacionales estructurados no podían ser mezclados y emparejados para temas analíticos con datos textuales no estructurados. Pero con el advenimiento de la contextualización, estos tipos de análisis ahora sí pueden hacerse de forma naturales y fácil.

En el data warehouse, datos no repetitivos, como los comentarios en una encuesta, correos electrónicos y conversaciones, se tratan de forma diferente a las ocurrencias repetitivas de datos, como el flujo de clics, mediciones o el procesamiento máquina o analógico. Los datos no repetitivos son datos basados en textos que fueron generados por la palabra escrita o hablada, leída y reformateada y, lo que es más importante, ahora puede ser contextualizada. Con el fin de extraer cualquier sentido de los datos no repetitivos para su uso en el Data Warehouse, deben tener el contexto de los datos establecidos.

En muchos casos, el contexto de los datos no repetitivos es más importante que los datos en sí. En cualquier caso, los datos no repetitivos no pueden utilizarse para la toma de decisiones hasta que se haya establecido el contexto.

Data Lakes y Data Warehouses: ¿mutuamente exclusivos o partners perfectos?

Los data lakes han surgido en el paisaje de Data Management en los últimos años, sin embargo, data lake no es necesariamente un reemplazo del data warehouse. En cambio, complementan los esfuerzos existentes y dan soporte al descubrimiento de nuevas preguntas. Una vez que se descubren esas preguntas se optimizan las respuestas. Y optimizar puede significar moverse fuera del data lake para ir a un data mart o al data warehouse.

## <u>:Resuelve tus dudas sobre el Data Lake</u> con uno de nuestros expertos!

Estas son algunas diferencias clave entre data lake y data warehouse:

- Datos: Un data warehouse sólo almacena datos que han sido modelados o estructurados, mientras que un Data Lake no hace acepción de datos. Lo almacena todo, estructurado, semiestructurado y no estructurado.
- **Procesamiento**: Antes de que una empresa pueda cargar datos en un data warehouse, primero debe darles forma y estructura, es decir, los datos deben ser modelados. Eso se llama schema-on-write. Con un data lake, sólo se cargan los datos sin procesar, tal y como están, y cuando esté listo para usar los datos, es cuando se le da forma y estructura. Eso se llama schema-on-read. Dos enfoques muy diferentes.
- Almacenamiento: Una de las principales características de las tecnologías de big data, como Hadoop, es que el coste de almacenamiento de datos es relativamente bajo en comparación con el de un data warehouse. Hay dos razones principales para esto: en primer lugar, Hadoop es software de código abierto, por lo que la concesión de licencias y el soporte de la comunidad es gratuito. Y segundo, Hadoop está diseñado para ser instalado en hardware de bajo coste.
- Agilidad: Un almacén de datos es un repositorio altamente estructurado, por definición. No
  es técnicamente difícil cambiar la estructura, pero puede tomar mucho tiempo dado todos
  los procesos de negocio que están vinculados a ella. Un data lake, por otro lado, carece de
  la estructura de un data warehouse, lo que da a los desarrolladores y a los científicos de
  datos la capacidad de configurar y reconfigurar fácilmente y en tiempo real sus modelos,
  consultas y aplicaciones.
- Seguridad: La tecnología del data warehouse existe desde hace décadas, mientras que la tecnología de big data (la base de un Data Lake) es relativamente nueva. Por lo tanto, la capacidad de asegurar datos en un data warehouse es mucho más madura que asegurar datos en un data lake. Cabe señalar, sin embargo, que se está realizando un importante esfuerzo en materia de seguridad en la actualidad en la industria de Big Data.

El data warehouse continúa evolucionando

Hay todavía un gran potencial en la evolución del data warehouse y su uso sigue adelante.

El procesamiento analítico clásico de datos basado en transacciones se realiza en el data warehouse como siempre se ha hecho. Nada ha cambiado ahí. Pero ahora puede hacerse analítica sobre datos contextualizados, y esa forma de análisis es nueva. La mayoría de las organizaciones, hasta ahora no habían podido basar su toma de decisiones en datos textuales no estructurados. Y ahora hay una nueva forma de análisis posible en un Data Warehouse: la posibilidad de la mezcla de análisis. El análisis combinado se realiza mediante una combinación de datos transaccionales estructurados y datos contextuales no estructurados.

Existen muchas otras formas de análisis que también son posibles. Estas formas incluyen análisis predictivo y prescriptivo, así como tecnologías de machine learning y otras que están cambiando la forma en que los datos son administrados y analizados. El data

warehouse ha sido durante mucho tiempo un elemento básico de la arquitectura de datos empresarial y sigue teniendo tiene un gran futuro en el nuevo mundo de Big Data y la analítica avanzada.

## 3. Data Warehouse en la nube

Los data warehouses están atravesando actualmente dos transformaciones muy importantes que tienen el potencial de impulsar niveles significativos de innovación empresarial:

- La primera área de transformación es el impulso para aumentar la agilidad general. La gran mayoría de los departamentos de TI están experimentando un rápido aumento de la demanda de datos. Los directivos quieren tener acceso a más y más datos históricos, mientras que al mismo tiempo, los científicos de datos y los analistas de negocios están explorando formas de introducir nuevos flujos de datos en el almacén para enriquecer el análisis existente, así como impulsar nuevas áreas de análisis. Esta rápida expansión de los volúmenes y fuentes de datos significa que los equipos de TI necesitan invertir más tiempo y esfuerzo asegurando que el rendimiento de las consultas permanezca constante y necesitan proporcionar cada vez más entornos para equipos individuales para validar el valor comercial de los nuevos conjuntos de datos.
- La segunda área de transformación gira en torno a la necesidad de mejorar el control de costes. Existe una creciente necesidad de hacer más con cada vez menos recursos, al mismo tiempo que se garantiza que todos los datos sensibles y estratégicos estén completamente asegurados, a lo largo de todo el ciclo de vida, de la manera más rentable.

Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

La nube está demostrando ser un facilitador dominante. Permite a las organizaciones enfrentarse de forma activa a los desafíos que presentan estas dos transformaciones clave.

¿Por qué migrar a la nube?

Las 3 tecnologías que ya tienen y seguirán teniendo un gran impacto en la transformación de negocios son: la nube, datos y análisis, e Internet de las Cosas. Estas tres tecnologías son parte clave del ecosistema del data warehouse actual. Por lo tanto, es posible sacar la conclusión de que el data warehouse en la nube tiene un gran potencial para generar un impacto significativo en los negocios.

La importancia de la nube para el data warehouse está directamente vinculado a tres factores clave:

Mayor agilidad

- Mejor control de costes
- Co-localización

## Mejora de la agilidad

Muchos data warehouses están embarcados en una fase de actualización. Con mucho trabajo puesto ahora en torno al big data, las empresas buscan aprovechar los nuevos flujos de datos y nuevos tipos de análisis más ricos para apoyar e impulsar nuevas áreas, tales como:analítica de clientes de 360°, analisis predictivo, detección de fraude, análisis de IoT y el establecimiento de los datos como centro de beneficio. Muchos de estos proyectos requieren el aprovisionamiento de nuevos entornos de hardware y el despliegue de software. Es más rápido, más fácil y más eficiente iniciar estos nuevos proyectos centrados en los datos utilizando los servicios en la nube.

## Mayor control de costes

Muchos equipos de TI están buscando maneras de consolidar los data marts existentes, cada uno ejecutándose en hardware dedicado o incluso en hardware propietario, en un único entorno integrado. Los servicios en la nube de calidad ofrecen la oportunidad perfecta para iniciar este tipo de proyectos. El cumplimiento legal no puede ser visto como un extra opcional cuando se planifica un traslado a la nube. Los activos de datos deben estar protegidos a lo largo de todo el ciclo de vida. Los servicios en la nube deben facilitar todo esto, siendo más rentables ya que todas las características de seguridad se pueden habilitar de forma predeterminada y mejorada, y actualizada de forma transparente.

## Co-Localización para una carga más rápida

La mayoría de los data warehouses generan datos directamente desde aplicaciones clave, como entrada de pedidos, ventas, finanzas y fabricación. Por lo tanto, tiene mucho sentido ubicar conjuntamente el data warehouse junto con los sistemas fuente que ya se estén ejecutándose en la nube. La ubicación conjunta ofrece una carga de datos más rápida, lo que significa que los usuarios obtienen un acceso más oportuno a sus datos.

Principales ventajas de mover el data warehouse a la nube

Existen **tres ventajas principales** para mover un data warehouse a la nube, y estas están directamente vinculados a los tres controladores clave enumerados anteriormente:

- Más fácil consolidación y racionalización
- Monetización más rápida de los datos en la nube
- La nube ofrece mejor protección

# 4. Data Warehouse vs Big Data vs BI: cuales son las diferencias

En este apartado compararemos brevemente los conceptos Data Warehouse (DW), Big Data y Business Intelligence (BI).

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para

tener una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

Se trata de tres conceptos completamente diferentes que tienen en común una nueva manera de lidiar con los datos, siempre teniendo en cuenta la existencia de una gran volumen de información en varios formatos que contribuyen, de forma estructurada o no estructurada, a la toma de decisiones estratégicas. El objetivo final de cualquiera de estas tecnologías es ofrecer una ventaja competitiva a las empresas, pero la forma en que se utiliza es la que marcará la diferencia.

Descubre cómo ampliar tu negocio en la nube. Contacta con un experto en CLOUD!

## Big Data

Llamamos big data a un gran volumen de datos con una variedad, complejidad y velocidad de crecimiento enorme y que además tienen la característica de no ser estructurados. Eso significa que no son relacionales, estando además fuera del entorno corporativo. Es un tipo de tecnología que te permite analizar los datos en tiempo real y puede provenir de diferentes fuentes y formas, tales como mensajería instantánea, redes sociales, registros de grabaciones, imágenes, mensajes de correo electrónico, etc.

## Guía gratuita: "Del Bit... al Big Data", descubre todo lo que necesitas saber sobre Big Data y mucho más.

Para tener una idea de la importancia del big data, este mercado está creciendo cada año, alrededor de un 40%. La paradoja reside en el hecho de que cuanto mayor es la evolución del big data, mayor es la escasez de profesionales cualificados para satisfacer esa demanda. La expectativa es que en unos pocos años, haya un crecimiento aún más significativo para el uso de big data y por lo tanto un aumento en la demanda mano de obra especializada y de partners tecnológicos especializados en esta materia.

#### **Data Warehouse**

Por otro lado, un data warehouse almacena datos consolidados de diversas fuentes o sistemas de la empresa. Se trata de datos estructurados, que tiene como objetivo principal ser precisos y de alta calidad para de esta forma poder dar soporte a la toma de decisiones de la empresa. Se trata de conseguir tener todos los datos juntos para después poder dividirlos para hacer un análisis de determinados sectores o estrategias.

## **Business Intelligence**

Un Business Intelligence (BI) es una especie de "cuello de botella" de los datos recogidos del data warehouse, que llegan de forma exacta y útil para ayudar a la toma de decisiones. Business Intelligence transforma los datos en información útil para analizar no sólo los negocios, sino también las principales estrategias corporativas.

## Guía gratuita "Integridad de datos y BI", descubre cómo sacarle el máximo partido a tu proyecto de Integridad de datos

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

## 5. Material complementario

Soluciones y recursos para Data Warehouse

- Informatica Cloud para Microsoft Azure SQL Data Warehouse
- Test Data Warehouse
- MDM y calidad de datos para el data warehouse

#### Guías

- Data Lake: Superando las limitaciones del Data Warehouse
- El data warehouse inteligente

#### Artículos

- ¿Qué es warehouse y qué debe contemplar un data warehouse moderno?
- Beneficios y soluciones para un Data Warehouse en la nube
- Lo que necesitas saber acerca de la arquitectura de un datawarehouse
- Data warehousing: 3 ocasiones de hacer las cosas bien
- ¿Qué es un Data Warehouse y cómo saber cuándo lo necesitas implementar?
- Cómo optimizar la arquitectura de un data warehouse
- Data warehouse concepts: tu glosario de almacén de datos
- Data Warehouse, el primer paso para una solución de BI
- Tutoriales sobre DataWarehouse en pdf
- Cómo gestiona Informatica los nuevos medios de almacenamiento de datos
- ¿Qué es un Data Warehouse para un director de ventas?
- Bl en línea: el reto actual de los data warehouses
- El poder de data warehouse: Informatica PowerCenter

## 2. Pasado y presente del Data Warehouse

Históricamente, los data warehouses se habían formado utilizando datos repetitivos estructurados que eran filtrados antes de entrar en el data warehouse. Sin embargo, en los últimos años, el data warehouse ha evolucionado debido a información contextual que ahora se puede adjuntar a los datos no estructurados y que también puede ser almacenada.

Aquellos primeros datos relacionales estructurados no podían ser mezclados y emparejados para temas analíticos con datos textuales no estructurados. Pero con el advenimiento de la contextualización, estos tipos de análisis ahora sí pueden hacerse de forma naturales y fácil.

En el data warehouse, datos no repetitivos, como los comentarios en una encuesta, correos electrónicos y conversaciones, se tratan de forma diferente a las ocurrencias repetitivas de datos, como el flujo de clics, mediciones o el procesamiento máquina o analógico. Los datos no repetitivos son datos basados en textos que fueron generados por la palabra escrita o hablada, leída y reformateada y, lo que es más importante, ahora puede ser contextualizada. Con el fin de extraer cualquier sentido de los datos no repetitivos para su uso en el Data Warehouse, deben tener el contexto de los datos establecidos.

En muchos casos, el contexto de los datos no repetitivos es más importante que los datos en sí. En cualquier caso, los datos no repetitivos no pueden utilizarse para la toma de decisiones hasta que se haya establecido el contexto.

Data Lakes y Data Warehouses: ¿mutuamente exclusivos o partners perfectos?

Los data lakes han surgido en el paisaje de Data Management en los últimos años, sin embargo, data lake no es necesariamente un reemplazo del data warehouse. En cambio, complementan los esfuerzos existentes y dan soporte al descubrimiento de nuevas preguntas. Una vez que se descubren esas preguntas se optimizan las respuestas. Y optimizar puede significar moverse fuera del data lake para ir a un data mart o al data warehouse.

<u>:Resuelve tus dudas sobre el Data Lake</u> con uno de nuestros expertos!

Estas son algunas diferencias clave entre data lake y data warehouse:

- Datos: Un data warehouse sólo almacena datos que han sido modelados o estructurados, mientras que un Data Lake no hace acepción de datos. Lo almacena todo, estructurado, semiestructurado y no estructurado.
- **Procesamiento**: Antes de que una empresa pueda cargar datos en un data warehouse, primero debe darles forma y estructura, es decir, los datos deben ser modelados. Eso se llama schema-on-write. Con un data lake, sólo se cargan los datos sin procesar, tal y como están, y cuando esté listo para usar los datos, es cuando se le da forma y estructura. Eso se llama schema-on-read. Dos enfoques muy diferentes.
- Almacenamiento: Una de las principales características de las tecnologías de big data, como Hadoop, es que el coste de almacenamiento de datos es relativamente bajo en comparación con el de un data warehouse. Hay dos razones principales para esto: en primer lugar, Hadoop es software de código abierto, por lo que la concesión de licencias y el soporte de la comunidad es gratuito. Y segundo, Hadoop está diseñado para ser instalado en hardware de bajo coste.
- Agilidad: Un almacén de datos es un repositorio altamente estructurado, por definición. No
  es técnicamente difícil cambiar la estructura, pero puede tomar mucho tiempo dado todos
  los procesos de negocio que están vinculados a ella. Un data lake, por otro lado, carece de
  la estructura de un data warehouse, lo que da a los desarrolladores y a los científicos de
  datos la capacidad de configurar y reconfigurar fácilmente y en tiempo real sus modelos,
  consultas y aplicaciones.
- **Seguridad**: La tecnología del data warehouse existe desde hace décadas, mientras que la tecnología de big data (la base de un Data Lake) es relativamente nueva. Por lo tanto, la

capacidad de asegurar datos en un data warehouse es mucho más madura que asegurar datos en un data lake. Cabe señalar, sin embargo, que se está realizando un importante esfuerzo en materia de seguridad en la actualidad en la industria de Big Data.

El data warehouse continúa evolucionando

Hay todavía un gran potencial en la evolución del data warehouse y su uso sigue adelante.

El procesamiento analítico clásico de datos basado en transacciones se realiza en el data warehouse como siempre se ha hecho. Nada ha cambiado ahí. Pero ahora puede hacerse analítica sobre datos contextualizados, y esa forma de análisis es nueva. La mayoría de las organizaciones, hasta ahora no habían podido basar su toma de decisiones en datos textuales no estructurados. Y ahora hay una nueva forma de análisis posible en un Data Warehouse: la posibilidad de la mezcla de análisis. El análisis combinado se realiza mediante una combinación de datos transaccionales estructurados y datos contextuales no estructurados.

Existen muchas otras formas de análisis que también son posibles. Estas formas incluyen análisis predictivo y prescriptivo, así como tecnologías de machine learning y otras que están cambiando la forma en que los datos son administrados y analizados. El data warehouse ha sido durante mucho tiempo un elemento básico de la arquitectura de datos empresarial y sigue teniendo tiene un gran futuro en el nuevo mundo de Big Data y la analítica avanzada.

## 3. Data Warehouse en la nube

Los data warehouses están atravesando actualmente dos transformaciones muy importantes que tienen el potencial de impulsar niveles significativos de innovación empresarial:

- La primera área de transformación es el impulso para aumentar la agilidad general. La gran mayoría de los departamentos de TI están experimentando un rápido aumento de la demanda de datos. Los directivos quieren tener acceso a más y más datos históricos, mientras que al mismo tiempo, los científicos de datos y los analistas de negocios están explorando formas de introducir nuevos flujos de datos en el almacén para enriquecer el análisis existente, así como impulsar nuevas áreas de análisis. Esta rápida expansión de los volúmenes y fuentes de datos significa que los equipos de TI necesitan invertir más tiempo y esfuerzo asegurando que el rendimiento de las consultas permanezca constante y necesitan proporcionar cada vez más entornos para equipos individuales para validar el valor comercial de los nuevos conjuntos de datos.
- La segunda área de transformación gira en torno a la necesidad de mejorar el control de costes. Existe una creciente necesidad de hacer más con cada vez menos recursos, al mismo tiempo que se garantiza que todos los datos sensibles y estratégicos estén completamente asegurados, a lo largo de todo el ciclo de vida, de la manera más rentable.

## Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

La nube está demostrando ser un facilitador dominante. Permite a las organizaciones enfrentarse de forma activa a los desafíos que presentan estas dos transformaciones clave.

¿Por qué migrar a la nube?

Las 3 tecnologías que ya tienen y seguirán teniendo un gran impacto en la transformación de negocios son: la nube, datos y análisis, e Internet de las Cosas. Estas tres tecnologías son parte clave del ecosistema del data warehouse actual. Por lo tanto, es posible sacar la conclusión de que el data warehouse en la nube tiene un gran potencial para generar un impacto significativo en los negocios.

La importancia de la nube para el data warehouse está directamente vinculado a tres factores clave:

- Mayor agilidad
- Mejor control de costes
- Co-localización

Mejora de la agilidad

Muchos data warehouses están embarcados en una fase de actualización. Con mucho trabajo puesto ahora en torno al big data, las empresas buscan aprovechar los nuevos flujos de datos y nuevos tipos de análisis más ricos para apoyar e impulsar nuevas áreas, tales como:analítica de clientes de 360°, analisis predictivo, detección de fraude, análisis de IoT y el establecimiento de los datos como centro de beneficio. Muchos de estos proyectos requieren el aprovisionamiento de nuevos entornos de hardware y el despliegue de software. Es más rápido, más fácil y más eficiente iniciar estos nuevos proyectos centrados en los datos utilizando los servicios en la nube.

Mayor control de costes

Muchos equipos de TI están buscando maneras de consolidar los data marts existentes, cada uno ejecutándose en hardware dedicado o incluso en hardware propietario, en un único entorno integrado. Los servicios en la nube de calidad ofrecen la oportunidad perfecta para iniciar este tipo de proyectos. El cumplimiento legal no puede ser visto como un extra opcional cuando se planifica un traslado a la nube. Los activos de datos deben estar protegidos a lo largo de todo el ciclo de vida. Los servicios en la nube deben facilitar todo esto, siendo más rentables ya que todas las características de seguridad se pueden habilitar de forma predeterminada y mejorada, y actualizada de forma transparente.

## Co-Localización para una carga más rápida

La mayoría de los data warehouses generan datos directamente desde aplicaciones clave, como entrada de pedidos, ventas, finanzas y fabricación. Por lo tanto, **tiene mucho sentido ubicar conjuntamente el data warehouse junto con los sistemas fuente que ya se estén ejecutándose en la nube**. La ubicación conjunta ofrece una carga de datos más rápida, lo que significa que los usuarios obtienen un acceso más oportuno a sus datos.

Principales ventajas de mover el data warehouse a la nube

Existen **tres ventajas principales** para mover un data warehouse a la nube, y estas están directamente vinculados a los tres controladores clave enumerados anteriormente:

- Más fácil consolidación y racionalización
- Monetización más rápida de los datos en la nube
- La nube ofrece mejor protección

# 4. Data Warehouse vs Big Data vs BI: cuales son las diferencias

En este apartado compararemos brevemente los conceptos Data Warehouse (DW), Big Data y Business Intelligence (BI).

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para tener una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

Se trata de tres conceptos completamente diferentes que tienen en común una nueva manera de lidiar con los datos, siempre teniendo en cuenta la existencia de una gran volumen de información en varios formatos que contribuyen, de forma estructurada o no estructurada, a la toma de decisiones estratégicas. El objetivo final de cualquiera de estas tecnologías es ofrecer una ventaja competitiva a las empresas, pero la forma en que se utiliza es la que marcará la diferencia.

Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

### Big Data

Llamamos big data a un gran volumen de datos con una variedad, complejidad y velocidad de crecimiento enorme y que además tienen la característica de no ser estructurados. Eso significa que no son relacionales, estando además fuera del entorno corporativo. Es un tipo de tecnología que te permite analizar los datos en tiempo real y puede provenir de diferentes fuentes y formas, tales como mensajería instantánea, redes sociales, registros de grabaciones, imágenes, mensajes de correo electrónico, etc.

## Guía gratuita: "Del Bit... al Big Data", descubre todo lo que necesitas saber sobre Big Data y mucho más.

Para tener una idea de la importancia del big data, este mercado está creciendo cada año, alrededor de un 40%. La paradoja reside en el hecho de que cuanto mayor es la evolución del big data, mayor es la escasez de profesionales cualificados para satisfacer esa demanda. La expectativa es que en unos pocos años, haya un crecimiento aún más significativo para el uso de big data y por lo tanto un aumento en la demanda mano de obra especializada y de partners tecnológicos especializados en esta materia.

#### Data Warehouse

Por otro lado, un data warehouse almacena datos consolidados de diversas fuentes o sistemas de la empresa. Se trata de datos estructurados, que tiene como objetivo principal ser precisos y de alta calidad para de esta forma poder dar soporte a la toma de decisiones de la empresa. Se trata de conseguir tener todos los datos juntos para después poder dividirlos para hacer un análisis de determinados sectores o estrategias.

## Business Intelligence

Un Business Intelligence (BI) es una especie de "cuello de botella" de los datos recogidos del data warehouse, que llegan de forma exacta y útil para ayudar a la toma de decisiones. Business Intelligence transforma los datos en información útil para analizar no sólo los negocios, sino también las principales estrategias corporativas.

# Guía gratuita "Integridad de datos y BI", descubre cómo sacarle el máximo partido a tu proyecto de Integridad de datos

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

## 5. Material complementario

Soluciones y recursos para Data Warehouse

- Informatica Cloud para Microsoft Azure SQL Data Warehouse
- Test Data Warehouse
- MDM y calidad de datos para el data warehouse

### Guías

- Data Lake: Superando las limitaciones del Data Warehouse
- El data warehouse inteligente

### Artículos

- ¿Qué es warehouse y qué debe contemplar un data warehouse moderno?
- Beneficios y soluciones para un Data Warehouse en la nube
- Lo que necesitas saber acerca de la arquitectura de un datawarehouse
- Data warehousing: 3 ocasiones de hacer las cosas bien
- ¿Qué es un Data Warehouse y cómo saber cuándo lo necesitas implementar?
- Cómo optimizar la arquitectura de un data warehouse
- Data warehouse concepts: tu glosario de almacén de datos
- Data Warehouse, el primer paso para una solución de BI
- Tutoriales sobre DataWarehouse en pdf
- Cómo gestiona Informatica los nuevos medios de almacenamiento de datos
- ¿Qué es un Data Warehouse para un director de ventas?
- Bl en línea: el reto actual de los data warehouses
- El poder de data warehouse: Informatica PowerCenter

## 2. Pasado y presente del Data Warehouse

Históricamente, los data warehouses se habían formado utilizando datos repetitivos estructurados que eran filtrados antes de entrar en el data warehouse. Sin embargo, en los últimos años, el data warehouse ha evolucionado debido a información contextual que ahora se puede adjuntar a los datos no estructurados y que también puede ser almacenada.

Aquellos primeros datos relacionales estructurados no podían ser mezclados y emparejados para temas analíticos con datos textuales no estructurados. Pero con el advenimiento de la contextualización, estos tipos de análisis ahora sí pueden hacerse de forma naturales y fácil.

En el data warehouse, datos no repetitivos, como los comentarios en una encuesta, correos electrónicos y conversaciones, se tratan de forma diferente a las ocurrencias repetitivas de datos, como el flujo de clics, mediciones o el procesamiento máquina o analógico. Los datos no repetitivos son datos basados en textos que fueron generados por la palabra escrita o hablada, leída y reformateada y, lo que es más importante, ahora puede ser contextualizada. Con el fin de extraer cualquier sentido de los datos no repetitivos para su uso en el Data Warehouse, deben tener el contexto de los datos establecidos.

En muchos casos, el contexto de los datos no repetitivos es más importante que los datos en sí. En cualquier caso, los datos no repetitivos no pueden utilizarse para la toma de decisiones hasta que se haya establecido el contexto.

Data Lakes y Data Warehouses: ¿mutuamente exclusivos o partners perfectos?

Los data lakes han surgido en el paisaje de Data Management en los últimos años, sin embargo, data lake no es necesariamente un reemplazo del data warehouse. En cambio, complementan los esfuerzos existentes y dan soporte al descubrimiento de nuevas preguntas. Una vez que se descubren esas preguntas se optimizan las respuestas. Y optimizar puede significar moverse fuera del data lake para ir a un data mart o al data warehouse.

## <u>:Resuelve tus dudas sobre el Data Lake</u> <u>con uno de nuestros expertos!</u>

Estas son algunas diferencias clave entre data lake y data warehouse:

- Datos: Un data warehouse sólo almacena datos que han sido modelados o estructurados, mientras que un Data Lake no hace acepción de datos. Lo almacena todo, estructurado, semiestructurado y no estructurado.
- **Procesamiento**: Antes de que una empresa pueda cargar datos en un data warehouse, primero debe darles forma y estructura, es decir, los datos deben ser modelados. Eso se llama schema-on-write. Con un data lake, sólo se cargan los datos sin procesar, tal y como están, y cuando esté listo para usar los datos, es cuando se le da forma y estructura. Eso se llama schema-on-read. Dos enfoques muy diferentes.
- Almacenamiento: Una de las principales características de las tecnologías de big data, como Hadoop, es que el coste de almacenamiento de datos es relativamente bajo en comparación con el de un data warehouse. Hay dos razones principales para esto: en primer lugar, Hadoop es software de código abierto, por lo que la concesión de licencias y el soporte de la comunidad es gratuito. Y segundo, Hadoop está diseñado para ser instalado en hardware de bajo coste.
- Agilidad: Un almacén de datos es un repositorio altamente estructurado, por definición. No
  es técnicamente difícil cambiar la estructura, pero puede tomar mucho tiempo dado todos
  los procesos de negocio que están vinculados a ella. Un data lake, por otro lado, carece de
  la estructura de un data warehouse, lo que da a los desarrolladores y a los científicos de
  datos la capacidad de configurar y reconfigurar fácilmente y en tiempo real sus modelos,
  consultas y aplicaciones.
- Seguridad: La tecnología del data warehouse existe desde hace décadas, mientras que la tecnología de big data (la base de un Data Lake) es relativamente nueva. Por lo tanto, la capacidad de asegurar datos en un data warehouse es mucho más madura que asegurar datos en un data lake. Cabe señalar, sin embargo, que se está realizando un importante esfuerzo en materia de seguridad en la actualidad en la industria de Big Data.

El data warehouse continúa evolucionando

Hay todavía un gran potencial en la evolución del data warehouse y su uso sigue adelante.

El procesamiento analítico clásico de datos basado en transacciones se realiza en el data warehouse como siempre se ha hecho. Nada ha cambiado ahí. Pero ahora puede hacerse analítica sobre datos contextualizados, y esa forma de análisis es nueva. La mayoría de las organizaciones, hasta ahora no habían podido basar su toma de decisiones en datos textuales no estructurados. Y ahora hay una nueva forma de análisis posible en un Data Warehouse: la posibilidad de la mezcla de análisis. El análisis combinado se realiza mediante una combinación de datos transaccionales estructurados y datos contextuales no estructurados.

Existen muchas otras formas de análisis que también son posibles. Estas formas incluyen análisis predictivo y prescriptivo, así como tecnologías de machine learning y otras que están cambiando la forma en que los datos son administrados y analizados. El data

warehouse ha sido durante mucho tiempo un elemento básico de la arquitectura de datos empresarial y sigue teniendo tiene un gran futuro en el nuevo mundo de Big Data y la analítica avanzada.

## 3. Data Warehouse en la nube

Los data warehouses están atravesando actualmente dos transformaciones muy importantes que tienen el potencial de impulsar niveles significativos de innovación empresarial:

- La primera área de transformación es el impulso para aumentar la agilidad general. La gran mayoría de los departamentos de TI están experimentando un rápido aumento de la demanda de datos. Los directivos quieren tener acceso a más y más datos históricos, mientras que al mismo tiempo, los científicos de datos y los analistas de negocios están explorando formas de introducir nuevos flujos de datos en el almacén para enriquecer el análisis existente, así como impulsar nuevas áreas de análisis. Esta rápida expansión de los volúmenes y fuentes de datos significa que los equipos de TI necesitan invertir más tiempo y esfuerzo asegurando que el rendimiento de las consultas permanezca constante y necesitan proporcionar cada vez más entornos para equipos individuales para validar el valor comercial de los nuevos conjuntos de datos.
- La segunda área de transformación gira en torno a la necesidad de mejorar el control de costes. Existe una creciente necesidad de hacer más con cada vez menos recursos, al mismo tiempo que se garantiza que todos los datos sensibles y estratégicos estén completamente asegurados, a lo largo de todo el ciclo de vida, de la manera más rentable.

Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

La nube está demostrando ser un facilitador dominante. Permite a las organizaciones enfrentarse de forma activa a los desafíos que presentan estas dos transformaciones clave.

¿Por qué migrar a la nube?

Las 3 tecnologías que ya tienen y seguirán teniendo un gran impacto en la transformación de negocios son: la nube, datos y análisis, e Internet de las Cosas. Estas tres tecnologías son parte clave del ecosistema del data warehouse actual. Por lo tanto, es posible sacar la conclusión de que el data warehouse en la nube tiene un gran potencial para generar un impacto significativo en los negocios.

La importancia de la nube para el data warehouse está directamente vinculado a tres factores clave:

Mayor agilidad

- Mejor control de costes
- Co-localización

## Mejora de la agilidad

Muchos data warehouses están embarcados en una fase de actualización. Con mucho trabajo puesto ahora en torno al big data, las empresas buscan aprovechar los nuevos flujos de datos y nuevos tipos de análisis más ricos para apoyar e impulsar nuevas áreas, tales como:analítica de clientes de 360°, analísis predictivo, detección de fraude, análisis de loT y el establecimiento de los datos como centro de beneficio. Muchos de estos proyectos requieren el aprovisionamiento de nuevos entornos de hardware y el despliegue de software. Es más rápido, más fácil y más eficiente iniciar estos nuevos proyectos centrados en los datos utilizando los servicios en la nube.

## Mayor control de costes

Muchos equipos de TI están buscando maneras de consolidar los data marts existentes, cada uno ejecutándose en hardware dedicado o incluso en hardware propietario, en un único entorno integrado. Los servicios en la nube de calidad ofrecen la oportunidad perfecta para iniciar este tipo de proyectos. El cumplimiento legal no puede ser visto como un extra opcional cuando se planifica un traslado a la nube. Los activos de datos deben estar protegidos a lo largo de todo el ciclo de vida. Los servicios en la nube deben facilitar todo esto, siendo más rentables ya que todas las características de seguridad se pueden habilitar de forma predeterminada y mejorada, y actualizada de forma transparente.

## Co-Localización para una carga más rápida

La mayoría de los data warehouses generan datos directamente desde aplicaciones clave, como entrada de pedidos, ventas, finanzas y fabricación. Por lo tanto, tiene mucho sentido ubicar conjuntamente el data warehouse junto con los sistemas fuente que ya se estén ejecutándose en la nube. La ubicación conjunta ofrece una carga de datos más rápida, lo que significa que los usuarios obtienen un acceso más oportuno a sus datos.

Principales ventajas de mover el data warehouse a la nube

Existen **tres ventajas principales** para mover un data warehouse a la nube, y estas están directamente vinculados a los tres controladores clave enumerados anteriormente:

- Más fácil consolidación y racionalización
- Monetización más rápida de los datos en la nube
- La nube ofrece mejor protección

# 4. Data Warehouse vs Big Data vs BI: cuales son las diferencias

En este apartado compararemos brevemente los conceptos Data Warehouse (DW), Big Data y Business Intelligence (BI).

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para

tener una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

Se trata de tres conceptos completamente diferentes que tienen en común una nueva manera de lidiar con los datos, siempre teniendo en cuenta la existencia de una gran volumen de información en varios formatos que contribuyen, de forma estructurada o no estructurada, a la toma de decisiones estratégicas. El objetivo final de cualquiera de estas tecnologías es ofrecer una ventaja competitiva a las empresas, pero la forma en que se utiliza es la que marcará la diferencia.

Descubre cómo ampliar tu negocio en la nube. Contacta con un experto en CLOUD!

## Big Data

Llamamos big data a un gran volumen de datos con una variedad, complejidad y velocidad de crecimiento enorme y que además tienen la característica de no ser estructurados. Eso significa que no son relacionales, estando además fuera del entorno corporativo. Es un tipo de tecnología que te permite analizar los datos en tiempo real y puede provenir de diferentes fuentes y formas, tales como mensajería instantánea, redes sociales, registros de grabaciones, imágenes, mensajes de correo electrónico, etc.

## Guía gratuita: "Del Bit... al Big Data", descubre todo lo que necesitas saber sobre Big Data y mucho más.

Para tener una idea de la importancia del big data, este mercado está creciendo cada año, alrededor de un 40%. La paradoja reside en el hecho de que cuanto mayor es la evolución del big data, mayor es la escasez de profesionales cualificados para satisfacer esa demanda. La expectativa es que en unos pocos años, haya un crecimiento aún más significativo para el uso de big data y por lo tanto un aumento en la demanda mano de obra especializada y de partners tecnológicos especializados en esta materia.

#### **Data Warehouse**

Por otro lado, un data warehouse almacena datos consolidados de diversas fuentes o sistemas de la empresa. Se trata de datos estructurados, que tiene como objetivo principal ser precisos y de alta calidad para de esta forma poder dar soporte a la toma de decisiones de la empresa. Se trata de conseguir tener todos los datos juntos para después poder dividirlos para hacer un análisis de determinados sectores o estrategias.

## **Business Intelligence**

Un Business Intelligence (BI) es una especie de "cuello de botella" de los datos recogidos del data warehouse, que llegan de forma exacta y útil para ayudar a la toma de decisiones. Business Intelligence transforma los datos en información útil para analizar no sólo los negocios, sino también las principales estrategias corporativas.

## Guía gratuita "Integridad de datos y BI", descubre cómo sacarle el máximo partido a tu proyecto de Integridad de datos

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

## 5. Material complementario

Soluciones y recursos para Data Warehouse

- Informatica Cloud para Microsoft Azure SQL Data Warehouse
- Test Data Warehouse
- MDM y calidad de datos para el data warehouse

#### Guías

- Data Lake: Superando las limitaciones del Data Warehouse
- El data warehouse inteligente

#### **Artículos**

- ¿Qué es warehouse y qué debe contemplar un data warehouse moderno?
- Beneficios y soluciones para un Data Warehouse en la nube
- Lo que necesitas saber acerca de la arquitectura de un datawarehouse
- Data warehousing: 3 ocasiones de hacer las cosas bien
- ¿Qué es un Data Warehouse y cómo saber cuándo lo necesitas implementar?
- Cómo optimizar la arquitectura de un data warehouse
- Data warehouse concepts: tu glosario de almacén de datos
- Data Warehouse, el primer paso para una solución de BI
- Tutoriales sobre DataWarehouse en pdf
- Cómo gestiona Informatica los nuevos medios de almacenamiento de datos
- ¿Qué es un Data Warehouse para un director de ventas?
- Bl en línea: el reto actual de los data warehouses
- El poder de data warehouse: Informatica PowerCenter

## 2. Pasado y presente del Data Warehouse

Históricamente, los data warehouses se habían formado utilizando datos repetitivos estructurados que eran filtrados antes de entrar en el data warehouse. Sin embargo, en los últimos años, el data warehouse ha evolucionado debido a información contextual que ahora se puede adjuntar a los datos no estructurados y que también puede ser almacenada.

Aquellos primeros datos relacionales estructurados no podían ser mezclados y emparejados para temas analíticos con datos textuales no estructurados. Pero con el advenimiento de la contextualización, estos tipos de análisis ahora sí pueden hacerse de forma naturales y fácil.

En el data warehouse, datos no repetitivos, como los comentarios en una encuesta, correos electrónicos y conversaciones, se tratan de forma diferente a las ocurrencias repetitivas de datos, como el flujo de clics, mediciones o el procesamiento máquina o analógico. Los datos no repetitivos son datos basados en textos que fueron generados por la palabra escrita o hablada, leída y reformateada y, lo que es más importante, ahora puede ser contextualizada. Con el fin de extraer cualquier sentido de los datos no repetitivos para su uso en el Data Warehouse, deben tener el contexto de los datos establecidos.

En muchos casos, el contexto de los datos no repetitivos es más importante que los datos en sí. En cualquier caso, los datos no repetitivos no pueden utilizarse para la toma de decisiones hasta que se haya establecido el contexto.

Data Lakes y Data Warehouses: ¿mutuamente exclusivos o partners perfectos?

Los data lakes han surgido en el paisaje de Data Management en los últimos años, sin embargo, data lake no es necesariamente un reemplazo del data warehouse. En cambio, complementan los esfuerzos existentes y dan soporte al descubrimiento de nuevas preguntas. Una vez que se descubren esas preguntas se optimizan las respuestas. Y optimizar puede significar moverse fuera del data lake para ir a un data mart o al data warehouse.

<u>:Resuelve tus dudas sobre el Data Lake</u> con uno de nuestros expertos!

Estas son algunas diferencias clave entre data lake y data warehouse:

- Datos: Un data warehouse sólo almacena datos que han sido modelados o estructurados, mientras que un Data Lake no hace acepción de datos. Lo almacena todo, estructurado, semiestructurado y no estructurado.
- **Procesamiento**: Antes de que una empresa pueda cargar datos en un data warehouse, primero debe darles forma y estructura, es decir, los datos deben ser modelados. Eso se llama schema-on-write. Con un data lake, sólo se cargan los datos sin procesar, tal y como están, y cuando esté listo para usar los datos, es cuando se le da forma y estructura. Eso se llama schema-on-read. Dos enfoques muy diferentes.
- Almacenamiento: Una de las principales características de las tecnologías de big data, como Hadoop, es que el coste de almacenamiento de datos es relativamente bajo en comparación con el de un data warehouse. Hay dos razones principales para esto: en primer lugar, Hadoop es software de código abierto, por lo que la concesión de licencias y el soporte de la comunidad es gratuito. Y segundo, Hadoop está diseñado para ser instalado en hardware de bajo coste.
- Agilidad: Un almacén de datos es un repositorio altamente estructurado, por definición. No
  es técnicamente difícil cambiar la estructura, pero puede tomar mucho tiempo dado todos
  los procesos de negocio que están vinculados a ella. Un data lake, por otro lado, carece de
  la estructura de un data warehouse, lo que da a los desarrolladores y a los científicos de
  datos la capacidad de configurar y reconfigurar fácilmente y en tiempo real sus modelos,
  consultas y aplicaciones.
- **Seguridad**: La tecnología del data warehouse existe desde hace décadas, mientras que la tecnología de big data (la base de un Data Lake) es relativamente nueva. Por lo tanto, la

capacidad de asegurar datos en un data warehouse es mucho más madura que asegurar datos en un data lake. Cabe señalar, sin embargo, que se está realizando un importante esfuerzo en materia de seguridad en la actualidad en la industria de Big Data.

El data warehouse continúa evolucionando

Hay todavía un gran potencial en la evolución del data warehouse y su uso sigue adelante.

El procesamiento analítico clásico de datos basado en transacciones se realiza en el data warehouse como siempre se ha hecho. Nada ha cambiado ahí. Pero ahora puede hacerse analítica sobre datos contextualizados, y esa forma de análisis es nueva. La mayoría de las organizaciones, hasta ahora no habían podido basar su toma de decisiones en datos textuales no estructurados. Y ahora hay una nueva forma de análisis posible en un Data Warehouse: la posibilidad de la mezcla de análisis. El análisis combinado se realiza mediante una combinación de datos transaccionales estructurados y datos contextuales no estructurados.

Existen muchas otras formas de análisis que también son posibles. Estas formas incluyen análisis predictivo y prescriptivo, así como tecnologías de machine learning y otras que están cambiando la forma en que los datos son administrados y analizados. El data warehouse ha sido durante mucho tiempo un elemento básico de la arquitectura de datos empresarial y sigue teniendo tiene un gran futuro en el nuevo mundo de Big Data y la analítica avanzada.

### 3. Data Warehouse en la nube

Los data warehouses están atravesando actualmente dos transformaciones muy importantes que tienen el potencial de impulsar niveles significativos de innovación empresarial:

- La primera área de transformación es el impulso para aumentar la agilidad general. La gran mayoría de los departamentos de TI están experimentando un rápido aumento de la demanda de datos. Los directivos quieren tener acceso a más y más datos históricos, mientras que al mismo tiempo, los científicos de datos y los analistas de negocios están explorando formas de introducir nuevos flujos de datos en el almacén para enriquecer el análisis existente, así como impulsar nuevas áreas de análisis. Esta rápida expansión de los volúmenes y fuentes de datos significa que los equipos de TI necesitan invertir más tiempo y esfuerzo asegurando que el rendimiento de las consultas permanezca constante y necesitan proporcionar cada vez más entornos para equipos individuales para validar el valor comercial de los nuevos conjuntos de datos.
- La segunda área de transformación gira en torno a la necesidad de mejorar el control de costes. Existe una creciente necesidad de hacer más con cada vez menos recursos, al mismo tiempo que se garantiza que todos los datos sensibles y estratégicos estén completamente asegurados, a lo largo de todo el ciclo de vida, de la manera más rentable.

#### Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

La nube está demostrando ser un facilitador dominante. Permite a las organizaciones enfrentarse de forma activa a los desafíos que presentan estas dos transformaciones clave.

¿Por qué migrar a la nube?

Las 3 tecnologías que ya tienen y seguirán teniendo un gran impacto en la transformación de negocios son: la nube, datos y análisis, e Internet de las Cosas. Estas tres tecnologías son parte clave del ecosistema del data warehouse actual. Por lo tanto, es posible sacar la conclusión de que el data warehouse en la nube tiene un gran potencial para generar un impacto significativo en los negocios.

La importancia de la nube para el data warehouse está directamente vinculado a tres factores clave:

- Mayor agilidad
- Mejor control de costes
- Co-localización

Mejora de la agilidad

Muchos data warehouses están embarcados en una fase de actualización. Con mucho trabajo puesto ahora en torno al big data, las empresas buscan aprovechar los nuevos flujos de datos y nuevos tipos de análisis más ricos para apoyar e impulsar nuevas áreas, tales como:analítica de clientes de 360°, analisis predictivo, detección de fraude, análisis de IoT y el establecimiento de los datos como centro de beneficio. Muchos de estos proyectos requieren el aprovisionamiento de nuevos entornos de hardware y el despliegue de software. Es más rápido, más fácil y más eficiente iniciar estos nuevos proyectos centrados en los datos utilizando los servicios en la nube.

Mayor control de costes

Muchos equipos de TI están buscando maneras de consolidar los data marts existentes, cada uno ejecutándose en hardware dedicado o incluso en hardware propietario, en un único entorno integrado. Los servicios en la nube de calidad ofrecen la oportunidad perfecta para iniciar este tipo de proyectos. El cumplimiento legal no puede ser visto como un extra opcional cuando se planifica un traslado a la nube. Los activos de datos deben estar protegidos a lo largo de todo el ciclo de vida. Los servicios en la nube deben facilitar todo esto, siendo más rentables ya que todas las características de seguridad se pueden habilitar de forma predeterminada y mejorada, y actualizada de forma transparente.

#### Co-Localización para una carga más rápida

La mayoría de los data warehouses generan datos directamente desde aplicaciones clave, como entrada de pedidos, ventas, finanzas y fabricación. Por lo tanto, **tiene mucho sentido ubicar conjuntamente el data warehouse junto con los sistemas fuente que ya se estén ejecutándose en la nube**. La ubicación conjunta ofrece una carga de datos más rápida, lo que significa que los usuarios obtienen un acceso más oportuno a sus datos.

Principales ventajas de mover el data warehouse a la nube

Existen **tres ventajas principales** para mover un data warehouse a la nube, y estas están directamente vinculados a los tres controladores clave enumerados anteriormente:

- Más fácil consolidación y racionalización
- Monetización más rápida de los datos en la nube
- La nube ofrece mejor protección

# 4. Data Warehouse vs Big Data vs BI: cuales son las diferencias

En este apartado compararemos brevemente los conceptos Data Warehouse (DW), Big Data y Business Intelligence (BI).

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para tener una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

Se trata de tres conceptos completamente diferentes que tienen en común una nueva manera de lidiar con los datos, siempre teniendo en cuenta la existencia de una gran volumen de información en varios formatos que contribuyen, de forma estructurada o no estructurada, a la toma de decisiones estratégicas. El objetivo final de cualquiera de estas tecnologías es ofrecer una ventaja competitiva a las empresas, pero la forma en que se utiliza es la que marcará la diferencia.

Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

#### Big Data

Llamamos big data a un gran volumen de datos con una variedad, complejidad y velocidad de crecimiento enorme y que además tienen la característica de no ser estructurados. Eso significa que no son relacionales, estando además fuera del entorno corporativo. Es un tipo de tecnología que te permite analizar los datos en tiempo real y puede provenir de diferentes fuentes y formas, tales como mensajería instantánea, redes sociales, registros de grabaciones, imágenes, mensajes de correo electrónico, etc.

## Guía gratuita: "Del Bit... al Big Data", descubre todo lo que necesitas saber sobre Big Data y mucho más.

Para tener una idea de la importancia del big data, este mercado está creciendo cada año, alrededor de un 40%. La paradoja reside en el hecho de que cuanto mayor es la evolución del big data, mayor es la escasez de profesionales cualificados para satisfacer esa demanda. La expectativa es que en unos pocos años, haya un crecimiento aún más significativo para el uso de big data y por lo tanto un aumento en la demanda mano de obra especializada y de partners tecnológicos especializados en esta materia.

#### Data Warehouse

Por otro lado, un data warehouse almacena datos consolidados de diversas fuentes o sistemas de la empresa. Se trata de datos estructurados, que tiene como objetivo principal ser precisos y de alta calidad para de esta forma poder dar soporte a la toma de decisiones de la empresa. Se trata de conseguir tener todos los datos juntos para después poder dividirlos para hacer un análisis de determinados sectores o estrategias.

#### Business Intelligence

Un Business Intelligence (BI) es una especie de "cuello de botella" de los datos recogidos del data warehouse, que llegan de forma exacta y útil para ayudar a la toma de decisiones. Business Intelligence transforma los datos en información útil para analizar no sólo los negocios, sino también las principales estrategias corporativas.

# Guía gratuita "Integridad de datos y BI", descubre cómo sacarle el máximo partido a tu proyecto de Integridad de datos

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

## 5. Material complementario

Soluciones y recursos para Data Warehouse

- Informatica Cloud para Microsoft Azure SQL Data Warehouse
- Test Data Warehouse
- MDM y calidad de datos para el data warehouse

#### Guías

- Data Lake: Superando las limitaciones del Data Warehouse
- El data warehouse inteligente

#### **Artículos**

- ¿Qué es warehouse y qué debe contemplar un data warehouse moderno?
- Beneficios y soluciones para un Data Warehouse en la nube
- Lo que necesitas saber acerca de la arquitectura de un datawarehouse
- Data warehousing: 3 ocasiones de hacer las cosas bien
- ¿Qué es un Data Warehouse y cómo saber cuándo lo necesitas implementar?
- Cómo optimizar la arquitectura de un data warehouse
- Data warehouse concepts: tu glosario de almacén de datos
- Data Warehouse, el primer paso para una solución de BI
- Tutoriales sobre DataWarehouse en pdf
- Cómo gestiona Informatica los nuevos medios de almacenamiento de datos
- ¿Qué es un Data Warehouse para un director de ventas?
- Bl en línea: el reto actual de los data warehouses
- El poder de data warehouse: Informatica PowerCenter

### 2. Pasado y presente del Data Warehouse

Históricamente, los data warehouses se habían formado utilizando datos repetitivos estructurados que eran filtrados antes de entrar en el data warehouse. Sin embargo, en los últimos años, el data warehouse ha evolucionado debido a información contextual que ahora se puede adjuntar a los datos no estructurados y que también puede ser almacenada.

Aquellos primeros datos relacionales estructurados no podían ser mezclados y emparejados para temas analíticos con datos textuales no estructurados. Pero con el advenimiento de la contextualización, estos tipos de análisis ahora sí pueden hacerse de forma naturales y fácil.

En el data warehouse, datos no repetitivos, como los comentarios en una encuesta, correos electrónicos y conversaciones, se tratan de forma diferente a las ocurrencias repetitivas de datos, como el flujo de clics, mediciones o el procesamiento máquina o analógico. Los datos no repetitivos son datos basados en textos que fueron generados por la palabra escrita o hablada, leída y reformateada y, lo que es más importante, ahora puede ser contextualizada. Con el fin de extraer cualquier sentido de los datos no repetitivos para su uso en el Data Warehouse, deben tener el contexto de los datos establecidos.

En muchos casos, el contexto de los datos no repetitivos es más importante que los datos en sí. En cualquier caso, los datos no repetitivos no pueden utilizarse para la toma de decisiones hasta que se haya establecido el contexto.

Data Lakes y Data Warehouses: ¿mutuamente exclusivos o partners perfectos?

Los data lakes han surgido en el paisaje de Data Management en los últimos años, sin embargo, data lake no es necesariamente un reemplazo del data warehouse. En cambio, complementan los esfuerzos existentes y dan soporte al descubrimiento de nuevas preguntas. Una vez que se descubren esas preguntas se optimizan las respuestas. Y optimizar puede significar moverse fuera del data lake para ir a un data mart o al data warehouse.

#### <u>:Resuelve tus dudas sobre el Data Lake</u> con uno de nuestros expertos!

Estas son algunas diferencias clave entre data lake y data warehouse:

- Datos: Un data warehouse sólo almacena datos que han sido modelados o estructurados, mientras que un Data Lake no hace acepción de datos. Lo almacena todo, estructurado, semiestructurado y no estructurado.
- **Procesamiento**: Antes de que una empresa pueda cargar datos en un data warehouse, primero debe darles forma y estructura, es decir, los datos deben ser modelados. Eso se llama schema-on-write. Con un data lake, sólo se cargan los datos sin procesar, tal y como están, y cuando esté listo para usar los datos, es cuando se le da forma y estructura. Eso se llama schema-on-read. Dos enfoques muy diferentes.
- Almacenamiento: Una de las principales características de las tecnologías de big data, como Hadoop, es que el coste de almacenamiento de datos es relativamente bajo en comparación con el de un data warehouse. Hay dos razones principales para esto: en primer lugar, Hadoop es software de código abierto, por lo que la concesión de licencias y el soporte de la comunidad es gratuito. Y segundo, Hadoop está diseñado para ser instalado en hardware de bajo coste.
- Agilidad: Un almacén de datos es un repositorio altamente estructurado, por definición. No
  es técnicamente difícil cambiar la estructura, pero puede tomar mucho tiempo dado todos
  los procesos de negocio que están vinculados a ella. Un data lake, por otro lado, carece de
  la estructura de un data warehouse, lo que da a los desarrolladores y a los científicos de
  datos la capacidad de configurar y reconfigurar fácilmente y en tiempo real sus modelos,
  consultas y aplicaciones.
- Seguridad: La tecnología del data warehouse existe desde hace décadas, mientras que la tecnología de big data (la base de un Data Lake) es relativamente nueva. Por lo tanto, la capacidad de asegurar datos en un data warehouse es mucho más madura que asegurar datos en un data lake. Cabe señalar, sin embargo, que se está realizando un importante esfuerzo en materia de seguridad en la actualidad en la industria de Big Data.

El data warehouse continúa evolucionando

Hay todavía un gran potencial en la evolución del data warehouse y su uso sigue adelante.

El procesamiento analítico clásico de datos basado en transacciones se realiza en el data warehouse como siempre se ha hecho. Nada ha cambiado ahí. Pero ahora puede hacerse analítica sobre datos contextualizados, y esa forma de análisis es nueva. La mayoría de las organizaciones, hasta ahora no habían podido basar su toma de decisiones en datos textuales no estructurados. Y ahora hay una nueva forma de análisis posible en un Data Warehouse: la posibilidad de la mezcla de análisis. El análisis combinado se realiza mediante una combinación de datos transaccionales estructurados y datos contextuales no estructurados.

Existen muchas otras formas de análisis que también son posibles. Estas formas incluyen análisis predictivo y prescriptivo, así como tecnologías de machine learning y otras que están cambiando la forma en que los datos son administrados y analizados. El data

warehouse ha sido durante mucho tiempo un elemento básico de la arquitectura de datos empresarial y sigue teniendo tiene un gran futuro en el nuevo mundo de Big Data y la analítica avanzada.

#### 3. Data Warehouse en la nube

Los data warehouses están atravesando actualmente dos transformaciones muy importantes que tienen el potencial de impulsar niveles significativos de innovación empresarial:

- La primera área de transformación es el impulso para aumentar la agilidad general. La gran mayoría de los departamentos de TI están experimentando un rápido aumento de la demanda de datos. Los directivos quieren tener acceso a más y más datos históricos, mientras que al mismo tiempo, los científicos de datos y los analistas de negocios están explorando formas de introducir nuevos flujos de datos en el almacén para enriquecer el análisis existente, así como impulsar nuevas áreas de análisis. Esta rápida expansión de los volúmenes y fuentes de datos significa que los equipos de TI necesitan invertir más tiempo y esfuerzo asegurando que el rendimiento de las consultas permanezca constante y necesitan proporcionar cada vez más entornos para equipos individuales para validar el valor comercial de los nuevos conjuntos de datos.
- La segunda área de transformación gira en torno a la necesidad de mejorar el control de costes. Existe una creciente necesidad de hacer más con cada vez menos recursos, al mismo tiempo que se garantiza que todos los datos sensibles y estratégicos estén completamente asegurados, a lo largo de todo el ciclo de vida, de la manera más rentable.

Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

La nube está demostrando ser un facilitador dominante. Permite a las organizaciones enfrentarse de forma activa a los desafíos que presentan estas dos transformaciones clave.

¿Por qué migrar a la nube?

Las 3 tecnologías que ya tienen y seguirán teniendo un gran impacto en la transformación de negocios son: la nube, datos y análisis, e Internet de las Cosas. Estas tres tecnologías son parte clave del ecosistema del data warehouse actual. Por lo tanto, es posible sacar la conclusión de que el data warehouse en la nube tiene un gran potencial para generar un impacto significativo en los negocios.

La importancia de la nube para el data warehouse está directamente vinculado a tres factores clave:

Mayor agilidad

- Mejor control de costes
- Co-localización

#### Mejora de la agilidad

Muchos data warehouses están embarcados en una fase de actualización. Con mucho trabajo puesto ahora en torno al big data, las empresas buscan aprovechar los nuevos flujos de datos y nuevos tipos de análisis más ricos para apoyar e impulsar nuevas áreas, tales como:analítica de clientes de 360°, analísis predictivo, detección de fraude, análisis de loT y el establecimiento de los datos como centro de beneficio. Muchos de estos proyectos requieren el aprovisionamiento de nuevos entornos de hardware y el despliegue de software. Es más rápido, más fácil y más eficiente iniciar estos nuevos proyectos centrados en los datos utilizando los servicios en la nube.

#### Mayor control de costes

Muchos equipos de TI están buscando maneras de consolidar los data marts existentes, cada uno ejecutándose en hardware dedicado o incluso en hardware propietario, en un único entorno integrado. Los servicios en la nube de calidad ofrecen la oportunidad perfecta para iniciar este tipo de proyectos. El cumplimiento legal no puede ser visto como un extra opcional cuando se planifica un traslado a la nube. Los activos de datos deben estar protegidos a lo largo de todo el ciclo de vida. Los servicios en la nube deben facilitar todo esto, siendo más rentables ya que todas las características de seguridad se pueden habilitar de forma predeterminada y mejorada, y actualizada de forma transparente.

#### Co-Localización para una carga más rápida

La mayoría de los data warehouses generan datos directamente desde aplicaciones clave, como entrada de pedidos, ventas, finanzas y fabricación. Por lo tanto, tiene mucho sentido ubicar conjuntamente el data warehouse junto con los sistemas fuente que ya se estén ejecutándose en la nube. La ubicación conjunta ofrece una carga de datos más rápida, lo que significa que los usuarios obtienen un acceso más oportuno a sus datos.

Principales ventajas de mover el data warehouse a la nube

Existen **tres ventajas principales** para mover un data warehouse a la nube, y estas están directamente vinculados a los tres controladores clave enumerados anteriormente:

- Más fácil consolidación y racionalización
- Monetización más rápida de los datos en la nube
- La nube ofrece mejor protección

# 4. Data Warehouse vs Big Data vs BI: cuales son las diferencias

En este apartado compararemos brevemente los conceptos Data Warehouse (DW), Big Data y Business Intelligence (BI).

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para

tener una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

Se trata de tres conceptos completamente diferentes que tienen en común una nueva manera de lidiar con los datos, siempre teniendo en cuenta la existencia de una gran volumen de información en varios formatos que contribuyen, de forma estructurada o no estructurada, a la toma de decisiones estratégicas. El objetivo final de cualquiera de estas tecnologías es ofrecer una ventaja competitiva a las empresas, pero la forma en que se utiliza es la que marcará la diferencia.

Descubre cómo ampliar tu negocio en la nube. Contacta con un experto en CLOUD!

#### Big Data

Llamamos big data a un gran volumen de datos con una variedad, complejidad y velocidad de crecimiento enorme y que además tienen la característica de no ser estructurados. Eso significa que no son relacionales, estando además fuera del entorno corporativo. Es un tipo de tecnología que te permite analizar los datos en tiempo real y puede provenir de diferentes fuentes y formas, tales como mensajería instantánea, redes sociales, registros de grabaciones, imágenes, mensajes de correo electrónico, etc.

## Guía gratuita: "Del Bit... al Big Data", descubre todo lo que necesitas saber sobre Big Data y mucho más.

Para tener una idea de la importancia del big data, este mercado está creciendo cada año, alrededor de un 40%. La paradoja reside en el hecho de que cuanto mayor es la evolución del big data, mayor es la escasez de profesionales cualificados para satisfacer esa demanda. La expectativa es que en unos pocos años, haya un crecimiento aún más significativo para el uso de big data y por lo tanto un aumento en la demanda mano de obra especializada y de partners tecnológicos especializados en esta materia.

#### **Data Warehouse**

Por otro lado, un data warehouse almacena datos consolidados de diversas fuentes o sistemas de la empresa. Se trata de datos estructurados, que tiene como objetivo principal ser precisos y de alta calidad para de esta forma poder dar soporte a la toma de decisiones de la empresa. Se trata de conseguir tener todos los datos juntos para después poder dividirlos para hacer un análisis de determinados sectores o estrategias.

#### **Business Intelligence**

Un Business Intelligence (BI) es una especie de "cuello de botella" de los datos recogidos del data warehouse, que llegan de forma exacta y útil para ayudar a la toma de decisiones. Business Intelligence transforma los datos en información útil para analizar no sólo los negocios, sino también las principales estrategias corporativas.

## Guía gratuita "Integridad de datos y BI", descubre cómo sacarle el máximo partido a tu proyecto de Integridad de datos

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

### 5. Material complementario

Soluciones y recursos para Data Warehouse

- Informatica Cloud para Microsoft Azure SQL Data Warehouse
- Test Data Warehouse
- MDM y calidad de datos para el data warehouse

#### Guías

- Data Lake: Superando las limitaciones del Data Warehouse
- El data warehouse inteligente

#### **Artículos**

- ¿Qué es warehouse y qué debe contemplar un data warehouse moderno?
- Beneficios y soluciones para un Data Warehouse en la nube
- Lo que necesitas saber acerca de la arquitectura de un datawarehouse
- Data warehousing: 3 ocasiones de hacer las cosas bien
- ¿Qué es un Data Warehouse y cómo saber cuándo lo necesitas implementar?
- Cómo optimizar la arquitectura de un data warehouse
- Data warehouse concepts: tu glosario de almacén de datos
- Data Warehouse, el primer paso para una solución de BI
- Tutoriales sobre DataWarehouse en pdf
- Cómo gestiona Informatica los nuevos medios de almacenamiento de datos
- ¿Qué es un Data Warehouse para un director de ventas?
- Bl en línea: el reto actual de los data warehouses
- El poder de data warehouse: Informatica PowerCenter

## 2. Pasado y presente del Data Warehouse

Históricamente, los data warehouses se habían formado utilizando datos repetitivos estructurados que eran filtrados antes de entrar en el data warehouse. Sin embargo, en los últimos años, el data warehouse ha evolucionado debido a información contextual que ahora se puede adjuntar a los datos no estructurados y que también puede ser almacenada.

Aquellos primeros datos relacionales estructurados no podían ser mezclados y emparejados para temas analíticos con datos textuales no estructurados. Pero con el advenimiento de la contextualización, estos tipos de análisis ahora sí pueden hacerse de forma naturales y fácil.

En el data warehouse, datos no repetitivos, como los comentarios en una encuesta, correos electrónicos y conversaciones, se tratan de forma diferente a las ocurrencias repetitivas de datos, como el flujo de clics, mediciones o el procesamiento máquina o analógico. Los datos no repetitivos son datos basados en textos que fueron generados por la palabra escrita o hablada, leída y reformateada y, lo que es más importante, ahora puede ser contextualizada. Con el fin de extraer cualquier sentido de los datos no repetitivos para su uso en el Data Warehouse, deben tener el contexto de los datos establecidos.

En muchos casos, el contexto de los datos no repetitivos es más importante que los datos en sí. En cualquier caso, los datos no repetitivos no pueden utilizarse para la toma de decisiones hasta que se haya establecido el contexto.

Data Lakes y Data Warehouses: ¿mutuamente exclusivos o partners perfectos?

Los data lakes han surgido en el paisaje de Data Management en los últimos años, sin embargo, data lake no es necesariamente un reemplazo del data warehouse. En cambio, complementan los esfuerzos existentes y dan soporte al descubrimiento de nuevas preguntas. Una vez que se descubren esas preguntas se optimizan las respuestas. Y optimizar puede significar moverse fuera del data lake para ir a un data mart o al data warehouse.

<u>:Resuelve tus dudas sobre el Data Lake</u> con uno de nuestros expertos!

Estas son algunas diferencias clave entre data lake y data warehouse:

- Datos: Un data warehouse sólo almacena datos que han sido modelados o estructurados, mientras que un Data Lake no hace acepción de datos. Lo almacena todo, estructurado, semiestructurado y no estructurado.
- **Procesamiento**: Antes de que una empresa pueda cargar datos en un data warehouse, primero debe darles forma y estructura, es decir, los datos deben ser modelados. Eso se llama schema-on-write. Con un data lake, sólo se cargan los datos sin procesar, tal y como están, y cuando esté listo para usar los datos, es cuando se le da forma y estructura. Eso se llama schema-on-read. Dos enfoques muy diferentes.
- Almacenamiento: Una de las principales características de las tecnologías de big data, como Hadoop, es que el coste de almacenamiento de datos es relativamente bajo en comparación con el de un data warehouse. Hay dos razones principales para esto: en primer lugar, Hadoop es software de código abierto, por lo que la concesión de licencias y el soporte de la comunidad es gratuito. Y segundo, Hadoop está diseñado para ser instalado en hardware de bajo coste.
- Agilidad: Un almacén de datos es un repositorio altamente estructurado, por definición. No
  es técnicamente difícil cambiar la estructura, pero puede tomar mucho tiempo dado todos
  los procesos de negocio que están vinculados a ella. Un data lake, por otro lado, carece de
  la estructura de un data warehouse, lo que da a los desarrolladores y a los científicos de
  datos la capacidad de configurar y reconfigurar fácilmente y en tiempo real sus modelos,
  consultas y aplicaciones.
- **Seguridad**: La tecnología del data warehouse existe desde hace décadas, mientras que la tecnología de big data (la base de un Data Lake) es relativamente nueva. Por lo tanto, la

capacidad de asegurar datos en un data warehouse es mucho más madura que asegurar datos en un data lake. Cabe señalar, sin embargo, que se está realizando un importante esfuerzo en materia de seguridad en la actualidad en la industria de Big Data.

El data warehouse continúa evolucionando

Hay todavía un gran potencial en la evolución del data warehouse y su uso sigue adelante.

El procesamiento analítico clásico de datos basado en transacciones se realiza en el data warehouse como siempre se ha hecho. Nada ha cambiado ahí. Pero ahora puede hacerse analítica sobre datos contextualizados, y esa forma de análisis es nueva. La mayoría de las organizaciones, hasta ahora no habían podido basar su toma de decisiones en datos textuales no estructurados. Y ahora hay una nueva forma de análisis posible en un Data Warehouse: la posibilidad de la mezcla de análisis. El análisis combinado se realiza mediante una combinación de datos transaccionales estructurados y datos contextuales no estructurados.

Existen muchas otras formas de análisis que también son posibles. Estas formas incluyen análisis predictivo y prescriptivo, así como tecnologías de machine learning y otras que están cambiando la forma en que los datos son administrados y analizados. El data warehouse ha sido durante mucho tiempo un elemento básico de la arquitectura de datos empresarial y sigue teniendo tiene un gran futuro en el nuevo mundo de Big Data y la analítica avanzada.

### 3. Data Warehouse en la nube

Los data warehouses están atravesando actualmente dos transformaciones muy importantes que tienen el potencial de impulsar niveles significativos de innovación empresarial:

- La primera área de transformación es el impulso para aumentar la agilidad general. La gran mayoría de los departamentos de TI están experimentando un rápido aumento de la demanda de datos. Los directivos quieren tener acceso a más y más datos históricos, mientras que al mismo tiempo, los científicos de datos y los analistas de negocios están explorando formas de introducir nuevos flujos de datos en el almacén para enriquecer el análisis existente, así como impulsar nuevas áreas de análisis. Esta rápida expansión de los volúmenes y fuentes de datos significa que los equipos de TI necesitan invertir más tiempo y esfuerzo asegurando que el rendimiento de las consultas permanezca constante y necesitan proporcionar cada vez más entornos para equipos individuales para validar el valor comercial de los nuevos conjuntos de datos.
- La segunda área de transformación gira en torno a la necesidad de mejorar el control de costes. Existe una creciente necesidad de hacer más con cada vez menos recursos, al mismo tiempo que se garantiza que todos los datos sensibles y estratégicos estén completamente asegurados, a lo largo de todo el ciclo de vida, de la manera más rentable.

#### Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

La nube está demostrando ser un facilitador dominante. Permite a las organizaciones enfrentarse de forma activa a los desafíos que presentan estas dos transformaciones clave.

¿Por qué migrar a la nube?

Las 3 tecnologías que ya tienen y seguirán teniendo un gran impacto en la transformación de negocios son: la nube, datos y análisis, e Internet de las Cosas. Estas tres tecnologías son parte clave del ecosistema del data warehouse actual. Por lo tanto, es posible sacar la conclusión de que el data warehouse en la nube tiene un gran potencial para generar un impacto significativo en los negocios.

La importancia de la nube para el data warehouse está directamente vinculado a tres factores clave:

- Mayor agilidad
- Mejor control de costes
- Co-localización

Mejora de la agilidad

Muchos data warehouses están embarcados en una fase de actualización. Con mucho trabajo puesto ahora en torno al big data, **las empresas buscan aprovechar los nuevos flujos de datos y nuevos tipos de análisis más ricos para apoyar e impulsar nuevas áreas**, tales como:analítica de clientes de 360°, analísis predictivo, detección de fraude, análisis de loT y el establecimiento de los datos como centro de beneficio. Muchos de estos proyectos requieren el aprovisionamiento de nuevos entornos de hardware y el despliegue de software. **Es más rápido, más fácil y más eficiente iniciar estos nuevos proyectos centrados en los datos utilizando los servicios en la nube**.

Mayor control de costes

Muchos equipos de TI están buscando maneras de consolidar los data marts existentes, cada uno ejecutándose en hardware dedicado o incluso en hardware propietario, en un único entorno integrado. Los servicios en la nube de calidad ofrecen la oportunidad perfecta para iniciar este tipo de proyectos. El cumplimiento legal no puede ser visto como un extra opcional cuando se planifica un traslado a la nube. Los activos de datos deben estar protegidos a lo largo de todo el ciclo de vida. Los servicios en la nube deben facilitar todo esto, siendo más rentables ya que todas las características de seguridad se pueden habilitar de forma predeterminada y mejorada, y actualizada de forma transparente.

#### Co-Localización para una carga más rápida

La mayoría de los data warehouses generan datos directamente desde aplicaciones clave, como entrada de pedidos, ventas, finanzas y fabricación. Por lo tanto, **tiene mucho sentido ubicar conjuntamente el data warehouse junto con los sistemas fuente que ya se estén ejecutándose en la nube**. La ubicación conjunta ofrece una carga de datos más rápida, lo que significa que los usuarios obtienen un acceso más oportuno a sus datos.

Principales ventajas de mover el data warehouse a la nube

Existen **tres ventajas principales** para mover un data warehouse a la nube, y estas están directamente vinculados a los tres controladores clave enumerados anteriormente:

- Más fácil consolidación y racionalización
- Monetización más rápida de los datos en la nube
- La nube ofrece mejor protección

# 4. Data Warehouse vs Big Data vs BI: cuales son las diferencias

En este apartado compararemos brevemente los conceptos Data Warehouse (DW), Big Data y Business Intelligence (BI).

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para tener una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

Se trata de tres conceptos completamente diferentes que tienen en común una nueva manera de lidiar con los datos, siempre teniendo en cuenta la existencia de una gran volumen de información en varios formatos que contribuyen, de forma estructurada o no estructurada, a la toma de decisiones estratégicas. El objetivo final de cualquiera de estas tecnologías es ofrecer una ventaja competitiva a las empresas, pero la forma en que se utiliza es la que marcará la diferencia.

Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

Big Data

Llamamos big data a un gran volumen de datos con una variedad, complejidad y velocidad de crecimiento enorme y que además tienen la característica de no ser estructurados. Eso significa que no son relacionales, estando además fuera del entorno corporativo. Es un tipo de tecnología que te permite analizar los datos en tiempo real y puede provenir de diferentes fuentes y formas, tales como mensajería instantánea, redes sociales, registros de grabaciones, imágenes, mensajes de correo electrónico, etc.

## Guía gratuita: "Del Bit... al Big Data", descubre todo lo que necesitas saber sobre Big Data y mucho más.

Para tener una idea de la importancia del big data, este mercado está creciendo cada año, alrededor de un 40%. La paradoja reside en el hecho de que cuanto mayor es la evolución del big data, mayor es la escasez de profesionales cualificados para satisfacer esa demanda. La expectativa es que en unos pocos años, haya un crecimiento aún más significativo para el uso de big data y por lo tanto un aumento en la demanda mano de obra especializada y de partners tecnológicos especializados en esta materia.

#### Data Warehouse

Por otro lado, un data warehouse almacena datos consolidados de diversas fuentes o sistemas de la empresa. Se trata de datos estructurados, que tiene como objetivo principal ser precisos y de alta calidad para de esta forma poder dar soporte a la toma de decisiones de la empresa. Se trata de conseguir tener todos los datos juntos para después poder dividirlos para hacer un análisis de determinados sectores o estrategias.

#### Business Intelligence

Un Business Intelligence (BI) es una especie de "cuello de botella" de los datos recogidos del data warehouse, que llegan de forma exacta y útil para ayudar a la toma de decisiones. Business Intelligence transforma los datos en información útil para analizar no sólo los negocios, sino también las principales estrategias corporativas.

# Guía gratuita "Integridad de datos y BI", descubre cómo sacarle el máximo partido a tu proyecto de Integridad de datos

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

## 5. Material complementario

Soluciones y recursos para Data Warehouse

- Informatica Cloud para Microsoft Azure SQL Data Warehouse
- Test Data Warehouse
- MDM y calidad de datos para el data warehouse

#### Guías

- Data Lake: Superando las limitaciones del Data Warehouse
- El data warehouse inteligente

#### Artículos

- ¿Qué es warehouse y qué debe contemplar un data warehouse moderno?
- Beneficios y soluciones para un Data Warehouse en la nube
- Lo que necesitas saber acerca de la arquitectura de un datawarehouse
- Data warehousing: 3 ocasiones de hacer las cosas bien
- ¿Qué es un Data Warehouse y cómo saber cuándo lo necesitas implementar?
- Cómo optimizar la arquitectura de un data warehouse
- Data warehouse concepts: tu glosario de almacén de datos
- Data Warehouse, el primer paso para una solución de BI
- Tutoriales sobre DataWarehouse en pdf
- Cómo gestiona Informatica los nuevos medios de almacenamiento de datos
- ¿Qué es un Data Warehouse para un director de ventas?
- Bl en línea: el reto actual de los data warehouses
- El poder de data warehouse: Informatica PowerCenter

### 2. Pasado y presente del Data Warehouse

Históricamente, los data warehouses se habían formado utilizando datos repetitivos estructurados que eran filtrados antes de entrar en el data warehouse. Sin embargo, en los últimos años, el data warehouse ha evolucionado debido a información contextual que ahora se puede adjuntar a los datos no estructurados y que también puede ser almacenada.

Aquellos primeros datos relacionales estructurados no podían ser mezclados y emparejados para temas analíticos con datos textuales no estructurados. Pero con el advenimiento de la contextualización, estos tipos de análisis ahora sí pueden hacerse de forma naturales y fácil.

En el data warehouse, datos no repetitivos, como los comentarios en una encuesta, correos electrónicos y conversaciones, se tratan de forma diferente a las ocurrencias repetitivas de datos, como el flujo de clics, mediciones o el procesamiento máquina o analógico. Los datos no repetitivos son datos basados en textos que fueron generados por la palabra escrita o hablada, leída y reformateada y, lo que es más importante, ahora puede ser contextualizada. Con el fin de extraer cualquier sentido de los datos no repetitivos para su uso en el Data Warehouse, deben tener el contexto de los datos establecidos.

En muchos casos, el contexto de los datos no repetitivos es más importante que los datos en sí. En cualquier caso, los datos no repetitivos no pueden utilizarse para la toma de decisiones hasta que se haya establecido el contexto.

Data Lakes y Data Warehouses: ¿mutuamente exclusivos o partners perfectos?

Los data lakes han surgido en el paisaje de Data Management en los últimos años, sin embargo, data lake no es necesariamente un reemplazo del data warehouse. En cambio, complementan los esfuerzos existentes y dan soporte al descubrimiento de nuevas preguntas. Una vez que se descubren esas preguntas se optimizan las respuestas. Y optimizar puede significar moverse fuera del data lake para ir a un data mart o al data warehouse.

#### <u>:Resuelve tus dudas sobre el Data Lake</u> con uno de nuestros expertos!

Estas son algunas diferencias clave entre data lake y data warehouse:

- Datos: Un data warehouse sólo almacena datos que han sido modelados o estructurados, mientras que un Data Lake no hace acepción de datos. Lo almacena todo, estructurado, semiestructurado y no estructurado.
- **Procesamiento**: Antes de que una empresa pueda cargar datos en un data warehouse, primero debe darles forma y estructura, es decir, los datos deben ser modelados. Eso se llama schema-on-write. Con un data lake, sólo se cargan los datos sin procesar, tal y como están, y cuando esté listo para usar los datos, es cuando se le da forma y estructura. Eso se llama schema-on-read. Dos enfoques muy diferentes.
- Almacenamiento: Una de las principales características de las tecnologías de big data, como Hadoop, es que el coste de almacenamiento de datos es relativamente bajo en comparación con el de un data warehouse. Hay dos razones principales para esto: en primer lugar, Hadoop es software de código abierto, por lo que la concesión de licencias y el soporte de la comunidad es gratuito. Y segundo, Hadoop está diseñado para ser instalado en hardware de bajo coste.
- Agilidad: Un almacén de datos es un repositorio altamente estructurado, por definición. No
  es técnicamente difícil cambiar la estructura, pero puede tomar mucho tiempo dado todos
  los procesos de negocio que están vinculados a ella. Un data lake, por otro lado, carece de
  la estructura de un data warehouse, lo que da a los desarrolladores y a los científicos de
  datos la capacidad de configurar y reconfigurar fácilmente y en tiempo real sus modelos,
  consultas y aplicaciones.
- Seguridad: La tecnología del data warehouse existe desde hace décadas, mientras que la tecnología de big data (la base de un Data Lake) es relativamente nueva. Por lo tanto, la capacidad de asegurar datos en un data warehouse es mucho más madura que asegurar datos en un data lake. Cabe señalar, sin embargo, que se está realizando un importante esfuerzo en materia de seguridad en la actualidad en la industria de Big Data.

El data warehouse continúa evolucionando

Hay todavía un gran potencial en la evolución del data warehouse y su uso sigue adelante.

El procesamiento analítico clásico de datos basado en transacciones se realiza en el data warehouse como siempre se ha hecho. Nada ha cambiado ahí. Pero ahora puede hacerse analítica sobre datos contextualizados, y esa forma de análisis es nueva. La mayoría de las organizaciones, hasta ahora no habían podido basar su toma de decisiones en datos textuales no estructurados. Y ahora hay una nueva forma de análisis posible en un Data Warehouse: la posibilidad de la mezcla de análisis. El análisis combinado se realiza mediante una combinación de datos transaccionales estructurados y datos contextuales no estructurados.

Existen muchas otras formas de análisis que también son posibles. Estas formas incluyen análisis predictivo y prescriptivo, así como tecnologías de machine learning y otras que están cambiando la forma en que los datos son administrados y analizados. El data

warehouse ha sido durante mucho tiempo un elemento básico de la arquitectura de datos empresarial y sigue teniendo tiene un gran futuro en el nuevo mundo de Big Data y la analítica avanzada.

#### 3. Data Warehouse en la nube

Los data warehouses están atravesando actualmente dos transformaciones muy importantes que tienen el potencial de impulsar niveles significativos de innovación empresarial:

- La primera área de transformación es el impulso para aumentar la agilidad general. La gran mayoría de los departamentos de TI están experimentando un rápido aumento de la demanda de datos. Los directivos quieren tener acceso a más y más datos históricos, mientras que al mismo tiempo, los científicos de datos y los analistas de negocios están explorando formas de introducir nuevos flujos de datos en el almacén para enriquecer el análisis existente, así como impulsar nuevas áreas de análisis. Esta rápida expansión de los volúmenes y fuentes de datos significa que los equipos de TI necesitan invertir más tiempo y esfuerzo asegurando que el rendimiento de las consultas permanezca constante y necesitan proporcionar cada vez más entornos para equipos individuales para validar el valor comercial de los nuevos conjuntos de datos.
- La segunda área de transformación gira en torno a la necesidad de mejorar el control de costes. Existe una creciente necesidad de hacer más con cada vez menos recursos, al mismo tiempo que se garantiza que todos los datos sensibles y estratégicos estén completamente asegurados, a lo largo de todo el ciclo de vida, de la manera más rentable.

Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

La nube está demostrando ser un facilitador dominante. Permite a las organizaciones enfrentarse de forma activa a los desafíos que presentan estas dos transformaciones clave.

¿Por qué migrar a la nube?

Las 3 tecnologías que ya tienen y seguirán teniendo un gran impacto en la transformación de negocios son: la nube, datos y análisis, e Internet de las Cosas. Estas tres tecnologías son parte clave del ecosistema del data warehouse actual. Por lo tanto, es posible sacar la conclusión de que el data warehouse en la nube tiene un gran potencial para generar un impacto significativo en los negocios.

La importancia de la nube para el data warehouse está directamente vinculado a tres factores clave:

Mayor aqilidad

- Mejor control de costes
- Co-localización

#### Mejora de la agilidad

Muchos data warehouses están embarcados en una fase de actualización. Con mucho trabajo puesto ahora en torno al big data, las empresas buscan aprovechar los nuevos flujos de datos y nuevos tipos de análisis más ricos para apoyar e impulsar nuevas áreas, tales como:analítica de clientes de 360°, analisis predictivo, detección de fraude, análisis de IoT y el establecimiento de los datos como centro de beneficio. Muchos de estos proyectos requieren el aprovisionamiento de nuevos entornos de hardware y el despliegue de software. Es más rápido, más fácil y más eficiente iniciar estos nuevos proyectos centrados en los datos utilizando los servicios en la nube.

#### Mayor control de costes

Muchos equipos de TI están buscando maneras de consolidar los data marts existentes, cada uno ejecutándose en hardware dedicado o incluso en hardware propietario, en un único entorno integrado. Los servicios en la nube de calidad ofrecen la oportunidad perfecta para iniciar este tipo de proyectos. El cumplimiento legal no puede ser visto como un extra opcional cuando se planifica un traslado a la nube. Los activos de datos deben estar protegidos a lo largo de todo el ciclo de vida. Los servicios en la nube deben facilitar todo esto, siendo más rentables ya que todas las características de seguridad se pueden habilitar de forma predeterminada y mejorada, y actualizada de forma transparente.

#### Co-Localización para una carga más rápida

La mayoría de los data warehouses generan datos directamente desde aplicaciones clave, como entrada de pedidos, ventas, finanzas y fabricación. Por lo tanto, tiene mucho sentido ubicar conjuntamente el data warehouse junto con los sistemas fuente que ya se estén ejecutándose en la nube. La ubicación conjunta ofrece una carga de datos más rápida, lo que significa que los usuarios obtienen un acceso más oportuno a sus datos.

Principales ventajas de mover el data warehouse a la nube

Existen **tres ventajas principales** para mover un data warehouse a la nube, y estas están directamente vinculados a los tres controladores clave enumerados anteriormente:

- Más fácil consolidación y racionalización
- Monetización más rápida de los datos en la nube
- La nube ofrece mejor protección

# 4. Data Warehouse vs Big Data vs BI: cuales son las diferencias

En este apartado compararemos brevemente los conceptos Data Warehouse (DW), Big Data y Business Intelligence (BI).

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para

tener una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

Se trata de tres conceptos completamente diferentes que tienen en común una nueva manera de lidiar con los datos, siempre teniendo en cuenta la existencia de una gran volumen de información en varios formatos que contribuyen, de forma estructurada o no estructurada, a la toma de decisiones estratégicas. El objetivo final de cualquiera de estas tecnologías es ofrecer una ventaja competitiva a las empresas, pero la forma en que se utiliza es la que marcará la diferencia.

Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

#### Big Data

Llamamos big data a un gran volumen de datos con una variedad, complejidad y velocidad de crecimiento enorme y que además tienen la característica de no ser estructurados. Eso significa que no son relacionales, estando además fuera del entorno corporativo. Es un tipo de tecnología que te permite analizar los datos en tiempo real y puede provenir de diferentes fuentes y formas, tales como mensajería instantánea, redes sociales, registros de grabaciones, imágenes, mensajes de correo electrónico, etc.

## Guía gratuita: "Del Bit... al Big Data", descubre todo lo que necesitas saber sobre Big Data y mucho más.

Para tener una idea de la importancia del big data, este mercado está creciendo cada año, alrededor de un 40%. La paradoja reside en el hecho de que cuanto mayor es la evolución del big data, mayor es la escasez de profesionales cualificados para satisfacer esa demanda. La expectativa es que en unos pocos años, haya un crecimiento aún más significativo para el uso de big data y por lo tanto un aumento en la demanda mano de obra especializada y de partners tecnológicos especializados en esta materia.

#### Data Warehouse

Por otro lado, un data warehouse almacena datos consolidados de diversas fuentes o sistemas de la empresa. Se trata de datos estructurados, que tiene como objetivo principal ser precisos y de alta calidad para de esta forma poder dar soporte a la toma de decisiones de la empresa. Se trata de conseguir tener todos los datos juntos para después poder dividirlos para hacer un análisis de determinados sectores o estrategias.

#### **Business Intelligence**

Un Business Intelligence (BI) es una especie de "cuello de botella" de los datos recogidos del data warehouse, que llegan de forma exacta y útil para ayudar a la toma de decisiones. Business Intelligence transforma los datos en información útil para analizar no sólo los negocios, sino también las principales estrategias corporativas.

## Guía gratuita "Integridad de datos y BI", descubre cómo sacarle el máximo partido a tu proyecto de Integridad de datos

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

### 5. Material complementario

Soluciones y recursos para Data Warehouse

- Informatica Cloud para Microsoft Azure SQL Data Warehouse
- Test Data Warehouse
- MDM y calidad de datos para el data warehouse

#### Guías

- Data Lake: Superando las limitaciones del Data Warehouse
- El data warehouse inteligente

#### **Artículos**

- ¿Qué es warehouse y qué debe contemplar un data warehouse moderno?
- Beneficios y soluciones para un Data Warehouse en la nube
- Lo que necesitas saber acerca de la arquitectura de un datawarehouse
- Data warehousing: 3 ocasiones de hacer las cosas bien
- ¿Qué es un Data Warehouse y cómo saber cuándo lo necesitas implementar?
- Cómo optimizar la arquitectura de un data warehouse
- Data warehouse concepts: tu glosario de almacén de datos
- Data Warehouse, el primer paso para una solución de BI
- Tutoriales sobre DataWarehouse en pdf
- Cómo gestiona Informatica los nuevos medios de almacenamiento de datos
- ¿Qué es un Data Warehouse para un director de ventas?
- Bl en línea: el reto actual de los data warehouses
- El poder de data warehouse: Informatica PowerCenter

## 2. Pasado y presente del Data Warehouse

Históricamente, los data warehouses se habían formado utilizando datos repetitivos estructurados que eran filtrados antes de entrar en el data warehouse. Sin embargo, en los últimos años, el data warehouse ha evolucionado debido a información contextual que ahora se puede adjuntar a los datos no estructurados y que también puede ser almacenada.

Aquellos primeros datos relacionales estructurados no podían ser mezclados y emparejados para temas analíticos con datos textuales no estructurados. Pero con el advenimiento de la contextualización, estos tipos de análisis ahora sí pueden hacerse de forma naturales y fácil.

En el data warehouse, datos no repetitivos, como los comentarios en una encuesta, correos electrónicos y conversaciones, se tratan de forma diferente a las ocurrencias repetitivas de datos, como el flujo de clics, mediciones o el procesamiento máquina o analógico. Los datos no repetitivos son datos basados en textos que fueron generados por la palabra escrita o hablada, leída y reformateada y, lo que es más importante, ahora puede ser contextualizada. Con el fin de extraer cualquier sentido de los datos no repetitivos para su uso en el Data Warehouse, deben tener el contexto de los datos establecidos.

En muchos casos, el contexto de los datos no repetitivos es más importante que los datos en sí. En cualquier caso, los datos no repetitivos no pueden utilizarse para la toma de decisiones hasta que se haya establecido el contexto.

Data Lakes y Data Warehouses: ¿mutuamente exclusivos o partners perfectos?

Los data lakes han surgido en el paisaje de Data Management en los últimos años, sin embargo, data lake no es necesariamente un reemplazo del data warehouse. En cambio, complementan los esfuerzos existentes y dan soporte al descubrimiento de nuevas preguntas. Una vez que se descubren esas preguntas se optimizan las respuestas. Y optimizar puede significar moverse fuera del data lake para ir a un data mart o al data warehouse.

<u>:Resuelve tus dudas sobre el Data Lake</u> con uno de nuestros expertos!

Estas son algunas diferencias clave entre data lake y data warehouse:

- Datos: Un data warehouse sólo almacena datos que han sido modelados o estructurados, mientras que un Data Lake no hace acepción de datos. Lo almacena todo, estructurado, semiestructurado y no estructurado.
- **Procesamiento**: Antes de que una empresa pueda cargar datos en un data warehouse, primero debe darles forma y estructura, es decir, los datos deben ser modelados. Eso se llama schema-on-write. Con un data lake, sólo se cargan los datos sin procesar, tal y como están, y cuando esté listo para usar los datos, es cuando se le da forma y estructura. Eso se llama schema-on-read. Dos enfoques muy diferentes.
- Almacenamiento: Una de las principales características de las tecnologías de big data, como Hadoop, es que el coste de almacenamiento de datos es relativamente bajo en comparación con el de un data warehouse. Hay dos razones principales para esto: en primer lugar, Hadoop es software de código abierto, por lo que la concesión de licencias y el soporte de la comunidad es gratuito. Y segundo, Hadoop está diseñado para ser instalado en hardware de bajo coste.
- Agilidad: Un almacén de datos es un repositorio altamente estructurado, por definición. No
  es técnicamente difícil cambiar la estructura, pero puede tomar mucho tiempo dado todos
  los procesos de negocio que están vinculados a ella. Un data lake, por otro lado, carece de
  la estructura de un data warehouse, lo que da a los desarrolladores y a los científicos de
  datos la capacidad de configurar y reconfigurar fácilmente y en tiempo real sus modelos,
  consultas y aplicaciones.
- **Seguridad**: La tecnología del data warehouse existe desde hace décadas, mientras que la tecnología de big data (la base de un Data Lake) es relativamente nueva. Por lo tanto, la

capacidad de asegurar datos en un data warehouse es mucho más madura que asegurar datos en un data lake. Cabe señalar, sin embargo, que se está realizando un importante esfuerzo en materia de seguridad en la actualidad en la industria de Big Data.

El data warehouse continúa evolucionando

Hay todavía un gran potencial en la evolución del data warehouse y su uso sigue adelante.

El procesamiento analítico clásico de datos basado en transacciones se realiza en el data warehouse como siempre se ha hecho. Nada ha cambiado ahí. Pero ahora puede hacerse analítica sobre datos contextualizados, y esa forma de análisis es nueva. La mayoría de las organizaciones, hasta ahora no habían podido basar su toma de decisiones en datos textuales no estructurados. Y ahora hay una nueva forma de análisis posible en un Data Warehouse: la posibilidad de la mezcla de análisis. El análisis combinado se realiza mediante una combinación de datos transaccionales estructurados y datos contextuales no estructurados.

Existen muchas otras formas de análisis que también son posibles. Estas formas incluyen análisis predictivo y prescriptivo, así como tecnologías de machine learning y otras que están cambiando la forma en que los datos son administrados y analizados. El data warehouse ha sido durante mucho tiempo un elemento básico de la arquitectura de datos empresarial y sigue teniendo tiene un gran futuro en el nuevo mundo de Big Data y la analítica avanzada.

### 3. Data Warehouse en la nube

Los data warehouses están atravesando actualmente dos transformaciones muy importantes que tienen el potencial de impulsar niveles significativos de innovación empresarial:

- La primera área de transformación es el impulso para aumentar la agilidad general. La gran mayoría de los departamentos de TI están experimentando un rápido aumento de la demanda de datos. Los directivos quieren tener acceso a más y más datos históricos, mientras que al mismo tiempo, los científicos de datos y los analistas de negocios están explorando formas de introducir nuevos flujos de datos en el almacén para enriquecer el análisis existente, así como impulsar nuevas áreas de análisis. Esta rápida expansión de los volúmenes y fuentes de datos significa que los equipos de TI necesitan invertir más tiempo y esfuerzo asegurando que el rendimiento de las consultas permanezca constante y necesitan proporcionar cada vez más entornos para equipos individuales para validar el valor comercial de los nuevos conjuntos de datos.
- La segunda área de transformación gira en torno a la necesidad de mejorar el control de costes. Existe una creciente necesidad de hacer más con cada vez menos recursos, al mismo tiempo que se garantiza que todos los datos sensibles y estratégicos estén completamente asegurados, a lo largo de todo el ciclo de vida, de la manera más rentable.

### Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

La nube está demostrando ser un facilitador dominante. Permite a las organizaciones enfrentarse de forma activa a los desafíos que presentan estas dos transformaciones clave.

¿Por qué migrar a la nube?

Las 3 tecnologías que ya tienen y seguirán teniendo un gran impacto en la transformación de negocios son: la nube, datos y análisis, e Internet de las Cosas. Estas tres tecnologías son parte clave del ecosistema del data warehouse actual. Por lo tanto, es posible sacar la conclusión de que el data warehouse en la nube tiene un gran potencial para generar un impacto significativo en los negocios.

La importancia de la nube para el data warehouse está directamente vinculado a tres factores clave:

- Mayor agilidad
- Mejor control de costes
- Co-localización

Mejora de la agilidad

Muchos data warehouses están embarcados en una fase de actualización. Con mucho trabajo puesto ahora en torno al big data, las empresas buscan aprovechar los nuevos flujos de datos y nuevos tipos de análisis más ricos para apoyar e impulsar nuevas áreas, tales como:analítica de clientes de 360°, analisis predictivo, detección de fraude, análisis de IoT y el establecimiento de los datos como centro de beneficio. Muchos de estos proyectos requieren el aprovisionamiento de nuevos entornos de hardware y el despliegue de software. Es más rápido, más fácil y más eficiente iniciar estos nuevos proyectos centrados en los datos utilizando los servicios en la nube.

Mayor control de costes

Muchos equipos de TI están buscando maneras de consolidar los data marts existentes, cada uno ejecutándose en hardware dedicado o incluso en hardware propietario, en un único entorno integrado. Los servicios en la nube de calidad ofrecen la oportunidad perfecta para iniciar este tipo de proyectos. El cumplimiento legal no puede ser visto como un extra opcional cuando se planifica un traslado a la nube. Los activos de datos deben estar protegidos a lo largo de todo el ciclo de vida. Los servicios en la nube deben facilitar todo esto, siendo más rentables ya que todas las características de seguridad se pueden habilitar de forma predeterminada y mejorada, y actualizada de forma transparente.

#### Co-Localización para una carga más rápida

La mayoría de los data warehouses generan datos directamente desde aplicaciones clave, como entrada de pedidos, ventas, finanzas y fabricación. Por lo tanto, **tiene mucho sentido ubicar conjuntamente el data warehouse junto con los sistemas fuente que ya se estén ejecutándose en la nube**. La ubicación conjunta ofrece una carga de datos más rápida, lo que significa que los usuarios obtienen un acceso más oportuno a sus datos.

Principales ventajas de mover el data warehouse a la nube

Existen **tres ventajas principales** para mover un data warehouse a la nube, y estas están directamente vinculados a los tres controladores clave enumerados anteriormente:

- Más fácil consolidación y racionalización
- Monetización más rápida de los datos en la nube
- La nube ofrece mejor protección

# 4. Data Warehouse vs Big Data vs BI: cuales son las diferencias

En este apartado compararemos brevemente los conceptos Data Warehouse (DW), Big Data y Business Intelligence (BI).

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para tener una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

Se trata de tres conceptos completamente diferentes que tienen en común una nueva manera de lidiar con los datos, siempre teniendo en cuenta la existencia de una gran volumen de información en varios formatos que contribuyen, de forma estructurada o no estructurada, a la toma de decisiones estratégicas. El objetivo final de cualquiera de estas tecnologías es ofrecer una ventaja competitiva a las empresas, pero la forma en que se utiliza es la que marcará la diferencia.

Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

#### Big Data

Llamamos big data a un gran volumen de datos con una variedad, complejidad y velocidad de crecimiento enorme y que además tienen la característica de no ser estructurados. Eso significa que no son relacionales, estando además fuera del entorno corporativo. Es un tipo de tecnología que te permite analizar los datos en tiempo real y puede provenir de diferentes fuentes y formas, tales como mensajería instantánea, redes sociales, registros de grabaciones, imágenes, mensajes de correo electrónico, etc.

## Guía gratuita: "Del Bit... al Big Data", descubre todo lo que necesitas saber sobre Big Data y mucho más.

Para tener una idea de la importancia del big data, este mercado está creciendo cada año, alrededor de un 40%. La paradoja reside en el hecho de que cuanto mayor es la evolución del big data, mayor es la escasez de profesionales cualificados para satisfacer esa demanda. La expectativa es que en unos pocos años, haya un crecimiento aún más significativo para el uso de big data y por lo tanto un aumento en la demanda mano de obra especializada y de partners tecnológicos especializados en esta materia.

#### Data Warehouse

Por otro lado, un data warehouse almacena datos consolidados de diversas fuentes o sistemas de la empresa. Se trata de datos estructurados, que tiene como objetivo principal ser precisos y de alta calidad para de esta forma poder dar soporte a la toma de decisiones de la empresa. Se trata de conseguir tener todos los datos juntos para después poder dividirlos para hacer un análisis de determinados sectores o estrategias.

#### Business Intelligence

Un Business Intelligence (BI) es una especie de "cuello de botella" de los datos recogidos del data warehouse, que llegan de forma exacta y útil para ayudar a la toma de decisiones. Business Intelligence transforma los datos en información útil para analizar no sólo los negocios, sino también las principales estrategias corporativas.

# Guía gratuita "Integridad de datos y BI", descubre cómo sacarle el máximo partido a tu proyecto de Integridad de datos

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

## 5. Material complementario

Soluciones y recursos para Data Warehouse

- Informatica Cloud para Microsoft Azure SQL Data Warehouse
- Test Data Warehouse
- MDM y calidad de datos para el data warehouse

#### Guías

- Data Lake: Superando las limitaciones del Data Warehouse
- El data warehouse inteligente

#### Artículos

- ¿Qué es warehouse y qué debe contemplar un data warehouse moderno?
- Beneficios y soluciones para un Data Warehouse en la nube
- Lo que necesitas saber acerca de la arquitectura de un datawarehouse
- Data warehousing: 3 ocasiones de hacer las cosas bien
- ¿Qué es un Data Warehouse y cómo saber cuándo lo necesitas implementar?
- Cómo optimizar la arquitectura de un data warehouse
- Data warehouse concepts: tu glosario de almacén de datos
- Data Warehouse, el primer paso para una solución de BI
- Tutoriales sobre DataWarehouse en pdf
- Cómo gestiona Informatica los nuevos medios de almacenamiento de datos
- ¿Qué es un Data Warehouse para un director de ventas?
- Bl en línea: el reto actual de los data warehouses
- El poder de data warehouse: Informatica PowerCenter

### 2. Pasado y presente del Data Warehouse

Históricamente, los data warehouses se habían formado utilizando datos repetitivos estructurados que eran filtrados antes de entrar en el data warehouse. Sin embargo, en los últimos años, el data warehouse ha evolucionado debido a información contextual que ahora se puede adjuntar a los datos no estructurados y que también puede ser almacenada.

Aquellos primeros datos relacionales estructurados no podían ser mezclados y emparejados para temas analíticos con datos textuales no estructurados. Pero con el advenimiento de la contextualización, estos tipos de análisis ahora sí pueden hacerse de forma naturales y fácil.

En el data warehouse, datos no repetitivos, como los comentarios en una encuesta, correos electrónicos y conversaciones, se tratan de forma diferente a las ocurrencias repetitivas de datos, como el flujo de clics, mediciones o el procesamiento máquina o analógico. Los datos no repetitivos son datos basados en textos que fueron generados por la palabra escrita o hablada, leída y reformateada y, lo que es más importante, ahora puede ser contextualizada. Con el fin de extraer cualquier sentido de los datos no repetitivos para su uso en el Data Warehouse, deben tener el contexto de los datos establecidos.

En muchos casos, el contexto de los datos no repetitivos es más importante que los datos en sí. En cualquier caso, los datos no repetitivos no pueden utilizarse para la toma de decisiones hasta que se haya establecido el contexto.

Data Lakes y Data Warehouses: ¿mutuamente exclusivos o partners perfectos?

Los data lakes han surgido en el paisaje de Data Management en los últimos años, sin embargo, data lake no es necesariamente un reemplazo del data warehouse. En cambio, complementan los esfuerzos existentes y dan soporte al descubrimiento de nuevas preguntas. Una vez que se descubren esas preguntas se optimizan las respuestas. Y optimizar puede significar moverse fuera del data lake para ir a un data mart o al data warehouse.

#### <u>:Resuelve tus dudas sobre el Data Lake</u> <u>con uno de nuestros expertos!</u>

Estas son algunas diferencias clave entre data lake y data warehouse:

- Datos: Un data warehouse sólo almacena datos que han sido modelados o estructurados, mientras que un Data Lake no hace acepción de datos. Lo almacena todo, estructurado, semiestructurado y no estructurado.
- **Procesamiento**: Antes de que una empresa pueda cargar datos en un data warehouse, primero debe darles forma y estructura, es decir, los datos deben ser modelados. Eso se llama schema-on-write. Con un data lake, sólo se cargan los datos sin procesar, tal y como están, y cuando esté listo para usar los datos, es cuando se le da forma y estructura. Eso se llama schema-on-read. Dos enfoques muy diferentes.
- Almacenamiento: Una de las principales características de las tecnologías de big data, como Hadoop, es que el coste de almacenamiento de datos es relativamente bajo en comparación con el de un data warehouse. Hay dos razones principales para esto: en primer lugar, Hadoop es software de código abierto, por lo que la concesión de licencias y el soporte de la comunidad es gratuito. Y segundo, Hadoop está diseñado para ser instalado en hardware de bajo coste.
- Agilidad: Un almacén de datos es un repositorio altamente estructurado, por definición. No
  es técnicamente difícil cambiar la estructura, pero puede tomar mucho tiempo dado todos
  los procesos de negocio que están vinculados a ella. Un data lake, por otro lado, carece de
  la estructura de un data warehouse, lo que da a los desarrolladores y a los científicos de
  datos la capacidad de configurar y reconfigurar fácilmente y en tiempo real sus modelos,
  consultas y aplicaciones.
- Seguridad: La tecnología del data warehouse existe desde hace décadas, mientras que la tecnología de big data (la base de un Data Lake) es relativamente nueva. Por lo tanto, la capacidad de asegurar datos en un data warehouse es mucho más madura que asegurar datos en un data lake. Cabe señalar, sin embargo, que se está realizando un importante esfuerzo en materia de seguridad en la actualidad en la industria de Big Data.

El data warehouse continúa evolucionando

Hay todavía un gran potencial en la evolución del data warehouse y su uso sigue adelante.

El procesamiento analítico clásico de datos basado en transacciones se realiza en el data warehouse como siempre se ha hecho. Nada ha cambiado ahí. Pero ahora puede hacerse analítica sobre datos contextualizados, y esa forma de análisis es nueva. La mayoría de las organizaciones, hasta ahora no habían podido basar su toma de decisiones en datos textuales no estructurados. Y ahora hay una nueva forma de análisis posible en un Data Warehouse: la posibilidad de la mezcla de análisis. El análisis combinado se realiza mediante una combinación de datos transaccionales estructurados y datos contextuales no estructurados.

Existen muchas otras formas de análisis que también son posibles. Estas formas incluyen análisis predictivo y prescriptivo, así como tecnologías de machine learning y otras que están cambiando la forma en que los datos son administrados y analizados. El data

warehouse ha sido durante mucho tiempo un elemento básico de la arquitectura de datos empresarial y sigue teniendo tiene un gran futuro en el nuevo mundo de Big Data y la analítica avanzada.

#### 3. Data Warehouse en la nube

Los data warehouses están atravesando actualmente dos transformaciones muy importantes que tienen el potencial de impulsar niveles significativos de innovación empresarial:

- La primera área de transformación es el impulso para aumentar la agilidad general. La gran mayoría de los departamentos de TI están experimentando un rápido aumento de la demanda de datos. Los directivos quieren tener acceso a más y más datos históricos, mientras que al mismo tiempo, los científicos de datos y los analistas de negocios están explorando formas de introducir nuevos flujos de datos en el almacén para enriquecer el análisis existente, así como impulsar nuevas áreas de análisis. Esta rápida expansión de los volúmenes y fuentes de datos significa que los equipos de TI necesitan invertir más tiempo y esfuerzo asegurando que el rendimiento de las consultas permanezca constante y necesitan proporcionar cada vez más entornos para equipos individuales para validar el valor comercial de los nuevos conjuntos de datos.
- La segunda área de transformación gira en torno a la necesidad de mejorar el control de costes. Existe una creciente necesidad de hacer más con cada vez menos recursos, al mismo tiempo que se garantiza que todos los datos sensibles y estratégicos estén completamente asegurados, a lo largo de todo el ciclo de vida, de la manera más rentable.

Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

La nube está demostrando ser un facilitador dominante. Permite a las organizaciones enfrentarse de forma activa a los desafíos que presentan estas dos transformaciones clave.

¿Por qué migrar a la nube?

Las 3 tecnologías que ya tienen y seguirán teniendo un gran impacto en la transformación de negocios son: la nube, datos y análisis, e Internet de las Cosas. Estas tres tecnologías son parte clave del ecosistema del data warehouse actual. Por lo tanto, es posible sacar la conclusión de que el data warehouse en la nube tiene un gran potencial para generar un impacto significativo en los negocios.

La importancia de la nube para el data warehouse está directamente vinculado a tres factores clave:

Mayor aqilidad

- Mejor control de costes
- Co-localización

#### Mejora de la agilidad

Muchos data warehouses están embarcados en una fase de actualización. Con mucho trabajo puesto ahora en torno al big data, las empresas buscan aprovechar los nuevos flujos de datos y nuevos tipos de análisis más ricos para apoyar e impulsar nuevas áreas, tales como:analítica de clientes de 360°, analísis predictivo, detección de fraude, análisis de loT y el establecimiento de los datos como centro de beneficio. Muchos de estos proyectos requieren el aprovisionamiento de nuevos entornos de hardware y el despliegue de software. Es más rápido, más fácil y más eficiente iniciar estos nuevos proyectos centrados en los datos utilizando los servicios en la nube.

#### Mayor control de costes

Muchos equipos de TI están buscando maneras de consolidar los data marts existentes, cada uno ejecutándose en hardware dedicado o incluso en hardware propietario, en un único entorno integrado. Los servicios en la nube de calidad ofrecen la oportunidad perfecta para iniciar este tipo de proyectos. El cumplimiento legal no puede ser visto como un extra opcional cuando se planifica un traslado a la nube. Los activos de datos deben estar protegidos a lo largo de todo el ciclo de vida. Los servicios en la nube deben facilitar todo esto, siendo más rentables ya que todas las características de seguridad se pueden habilitar de forma predeterminada y mejorada, y actualizada de forma transparente.

#### Co-Localización para una carga más rápida

La mayoría de los data warehouses generan datos directamente desde aplicaciones clave, como entrada de pedidos, ventas, finanzas y fabricación. Por lo tanto, tiene mucho sentido ubicar conjuntamente el data warehouse junto con los sistemas fuente que ya se estén ejecutándose en la nube. La ubicación conjunta ofrece una carga de datos más rápida, lo que significa que los usuarios obtienen un acceso más oportuno a sus datos.

Principales ventajas de mover el data warehouse a la nube

Existen **tres ventajas principales** para mover un data warehouse a la nube, y estas están directamente vinculados a los tres controladores clave enumerados anteriormente:

- Más fácil consolidación y racionalización
- Monetización más rápida de los datos en la nube
- La nube ofrece mejor protección

# 4. Data Warehouse vs Big Data vs BI: cuales son las diferencias

En este apartado compararemos brevemente los conceptos Data Warehouse (DW), Big Data y Business Intelligence (BI).

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para

tener una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

Se trata de tres conceptos completamente diferentes que tienen en común una nueva manera de lidiar con los datos, siempre teniendo en cuenta la existencia de una gran volumen de información en varios formatos que contribuyen, de forma estructurada o no estructurada, a la toma de decisiones estratégicas. El objetivo final de cualquiera de estas tecnologías es ofrecer una ventaja competitiva a las empresas, pero la forma en que se utiliza es la que marcará la diferencia.

Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

#### Big Data

Llamamos big data a un gran volumen de datos con una variedad, complejidad y velocidad de crecimiento enorme y que además tienen la característica de no ser estructurados. Eso significa que no son relacionales, estando además fuera del entorno corporativo. Es un tipo de tecnología que te permite analizar los datos en tiempo real y puede provenir de diferentes fuentes y formas, tales como mensajería instantánea, redes sociales, registros de grabaciones, imágenes, mensajes de correo electrónico, etc.

## Guía gratuita: "Del Bit... al Big Data", descubre todo lo que necesitas saber sobre Big Data y mucho más.

Para tener una idea de la importancia del big data, este mercado está creciendo cada año, alrededor de un 40%. La paradoja reside en el hecho de que cuanto mayor es la evolución del big data, mayor es la escasez de profesionales cualificados para satisfacer esa demanda. La expectativa es que en unos pocos años, haya un crecimiento aún más significativo para el uso de big data y por lo tanto un aumento en la demanda mano de obra especializada y de partners tecnológicos especializados en esta materia.

#### Data Warehouse

Por otro lado, un data warehouse almacena datos consolidados de diversas fuentes o sistemas de la empresa. Se trata de datos estructurados, que tiene como objetivo principal ser precisos y de alta calidad para de esta forma poder dar soporte a la toma de decisiones de la empresa. Se trata de conseguir tener todos los datos juntos para después poder dividirlos para hacer un análisis de determinados sectores o estrategias.

#### **Business Intelligence**

Un Business Intelligence (BI) es una especie de "cuello de botella" de los datos recogidos del data warehouse, que llegan de forma exacta y útil para ayudar a la toma de decisiones. Business Intelligence transforma los datos en información útil para analizar no sólo los negocios, sino también las principales estrategias corporativas.

## Guía gratuita "Integridad de datos y BI", descubre cómo sacarle el máximo partido a tu proyecto de Integridad de datos

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

## 5. Material complementario

Soluciones y recursos para Data Warehouse

- Informatica Cloud para Microsoft Azure SQL Data Warehouse
- Test Data Warehouse
- MDM y calidad de datos para el data warehouse

#### Guías

- Data Lake: Superando las limitaciones del Data Warehouse
- El data warehouse inteligente

#### **Artículos**

- ¿Qué es warehouse y qué debe contemplar un data warehouse moderno?
- Beneficios y soluciones para un Data Warehouse en la nube
- Lo que necesitas saber acerca de la arquitectura de un datawarehouse
- Data warehousing: 3 ocasiones de hacer las cosas bien
- ¿Qué es un Data Warehouse y cómo saber cuándo lo necesitas implementar?
- Cómo optimizar la arquitectura de un data warehouse
- Data warehouse concepts: tu glosario de almacén de datos
- Data Warehouse, el primer paso para una solución de BI
- Tutoriales sobre DataWarehouse en pdf
- Cómo gestiona Informatica los nuevos medios de almacenamiento de datos
- ¿Qué es un Data Warehouse para un director de ventas?
- Bl en línea: el reto actual de los data warehouses
- El poder de data warehouse: Informatica PowerCenter

## 2. Pasado y presente del Data Warehouse

Históricamente, los data warehouses se habían formado utilizando datos repetitivos estructurados que eran filtrados antes de entrar en el data warehouse. Sin embargo, en los últimos años, el data warehouse ha evolucionado debido a información contextual que ahora se puede adjuntar a los datos no estructurados y que también puede ser almacenada.

Aquellos primeros datos relacionales estructurados no podían ser mezclados y emparejados para temas analíticos con datos textuales no estructurados. Pero con el advenimiento de la contextualización, estos tipos de análisis ahora sí pueden hacerse de forma naturales y fácil.

En el data warehouse, datos no repetitivos, como los comentarios en una encuesta, correos electrónicos y conversaciones, se tratan de forma diferente a las ocurrencias repetitivas de datos, como el flujo de clics, mediciones o el procesamiento máquina o analógico. Los datos no repetitivos son datos basados en textos que fueron generados por la palabra escrita o hablada, leída y reformateada y, lo que es más importante, ahora puede ser contextualizada. Con el fin de extraer cualquier sentido de los datos no repetitivos para su uso en el Data Warehouse, deben tener el contexto de los datos establecidos.

En muchos casos, el contexto de los datos no repetitivos es más importante que los datos en sí. En cualquier caso, los datos no repetitivos no pueden utilizarse para la toma de decisiones hasta que se haya establecido el contexto.

Data Lakes y Data Warehouses: ¿mutuamente exclusivos o partners perfectos?

Los data lakes han surgido en el paisaje de Data Management en los últimos años, sin embargo, data lake no es necesariamente un reemplazo del data warehouse. En cambio, complementan los esfuerzos existentes y dan soporte al descubrimiento de nuevas preguntas. Una vez que se descubren esas preguntas se optimizan las respuestas. Y optimizar puede significar moverse fuera del data lake para ir a un data mart o al data warehouse.

<u>:Resuelve tus dudas sobre el Data Lake</u> con uno de nuestros expertos!

Estas son algunas diferencias clave entre data lake y data warehouse:

- Datos: Un data warehouse sólo almacena datos que han sido modelados o estructurados, mientras que un Data Lake no hace acepción de datos. Lo almacena todo, estructurado, semiestructurado y no estructurado.
- **Procesamiento**: Antes de que una empresa pueda cargar datos en un data warehouse, primero debe darles forma y estructura, es decir, los datos deben ser modelados. Eso se llama schema-on-write. Con un data lake, sólo se cargan los datos sin procesar, tal y como están, y cuando esté listo para usar los datos, es cuando se le da forma y estructura. Eso se llama schema-on-read. Dos enfoques muy diferentes.
- Almacenamiento: Una de las principales características de las tecnologías de big data, como Hadoop, es que el coste de almacenamiento de datos es relativamente bajo en comparación con el de un data warehouse. Hay dos razones principales para esto: en primer lugar, Hadoop es software de código abierto, por lo que la concesión de licencias y el soporte de la comunidad es gratuito. Y segundo, Hadoop está diseñado para ser instalado en hardware de bajo coste.
- Agilidad: Un almacén de datos es un repositorio altamente estructurado, por definición. No
  es técnicamente difícil cambiar la estructura, pero puede tomar mucho tiempo dado todos
  los procesos de negocio que están vinculados a ella. Un data lake, por otro lado, carece de
  la estructura de un data warehouse, lo que da a los desarrolladores y a los científicos de
  datos la capacidad de configurar y reconfigurar fácilmente y en tiempo real sus modelos,
  consultas y aplicaciones.
- **Seguridad**: La tecnología del data warehouse existe desde hace décadas, mientras que la tecnología de big data (la base de un Data Lake) es relativamente nueva. Por lo tanto, la

capacidad de asegurar datos en un data warehouse es mucho más madura que asegurar datos en un data lake. Cabe señalar, sin embargo, que se está realizando un importante esfuerzo en materia de seguridad en la actualidad en la industria de Big Data.

El data warehouse continúa evolucionando

Hay todavía un gran potencial en la evolución del data warehouse y su uso sigue adelante.

El procesamiento analítico clásico de datos basado en transacciones se realiza en el data warehouse como siempre se ha hecho. Nada ha cambiado ahí. Pero ahora puede hacerse analítica sobre datos contextualizados, y esa forma de análisis es nueva. La mayoría de las organizaciones, hasta ahora no habían podido basar su toma de decisiones en datos textuales no estructurados. Y ahora hay una nueva forma de análisis posible en un Data Warehouse: la posibilidad de la mezcla de análisis. El análisis combinado se realiza mediante una combinación de datos transaccionales estructurados y datos contextuales no estructurados.

Existen muchas otras formas de análisis que también son posibles. Estas formas incluyen análisis predictivo y prescriptivo, así como tecnologías de machine learning y otras que están cambiando la forma en que los datos son administrados y analizados. El data warehouse ha sido durante mucho tiempo un elemento básico de la arquitectura de datos empresarial y sigue teniendo tiene un gran futuro en el nuevo mundo de Big Data y la analítica avanzada.

### 3. Data Warehouse en la nube

Los data warehouses están atravesando actualmente dos transformaciones muy importantes que tienen el potencial de impulsar niveles significativos de innovación empresarial:

- La primera área de transformación es el impulso para aumentar la agilidad general. La gran mayoría de los departamentos de TI están experimentando un rápido aumento de la demanda de datos. Los directivos quieren tener acceso a más y más datos históricos, mientras que al mismo tiempo, los científicos de datos y los analistas de negocios están explorando formas de introducir nuevos flujos de datos en el almacén para enriquecer el análisis existente, así como impulsar nuevas áreas de análisis. Esta rápida expansión de los volúmenes y fuentes de datos significa que los equipos de TI necesitan invertir más tiempo y esfuerzo asegurando que el rendimiento de las consultas permanezca constante y necesitan proporcionar cada vez más entornos para equipos individuales para validar el valor comercial de los nuevos conjuntos de datos.
- La segunda área de transformación gira en torno a la necesidad de mejorar el control de costes. Existe una creciente necesidad de hacer más con cada vez menos recursos, al mismo tiempo que se garantiza que todos los datos sensibles y estratégicos estén completamente asegurados, a lo largo de todo el ciclo de vida, de la manera más rentable.

#### Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

La nube está demostrando ser un facilitador dominante. Permite a las organizaciones enfrentarse de forma activa a los desafíos que presentan estas dos transformaciones clave.

¿Por qué migrar a la nube?

Las 3 tecnologías que ya tienen y seguirán teniendo un gran impacto en la transformación de negocios son: la nube, datos y análisis, e Internet de las Cosas. Estas tres tecnologías son parte clave del ecosistema del data warehouse actual. Por lo tanto, es posible sacar la conclusión de que el data warehouse en la nube tiene un gran potencial para generar un impacto significativo en los negocios.

La importancia de la nube para el data warehouse está directamente vinculado a tres factores clave:

- Mayor agilidad
- Mejor control de costes
- Co-localización

Mejora de la agilidad

Muchos data warehouses están embarcados en una fase de actualización. Con mucho trabajo puesto ahora en torno al big data, las empresas buscan aprovechar los nuevos flujos de datos y nuevos tipos de análisis más ricos para apoyar e impulsar nuevas áreas, tales como:analítica de clientes de 360°, analisis predictivo, detección de fraude, análisis de IoT y el establecimiento de los datos como centro de beneficio. Muchos de estos proyectos requieren el aprovisionamiento de nuevos entornos de hardware y el despliegue de software. Es más rápido, más fácil y más eficiente iniciar estos nuevos proyectos centrados en los datos utilizando los servicios en la nube.

Mayor control de costes

Muchos equipos de TI están buscando maneras de consolidar los data marts existentes, cada uno ejecutándose en hardware dedicado o incluso en hardware propietario, en un único entorno integrado. Los servicios en la nube de calidad ofrecen la oportunidad perfecta para iniciar este tipo de proyectos. El cumplimiento legal no puede ser visto como un extra opcional cuando se planifica un traslado a la nube. Los activos de datos deben estar protegidos a lo largo de todo el ciclo de vida. Los servicios en la nube deben facilitar todo esto, siendo más rentables ya que todas las características de seguridad se pueden habilitar de forma predeterminada y mejorada, y actualizada de forma transparente.

#### Co-Localización para una carga más rápida

La mayoría de los data warehouses generan datos directamente desde aplicaciones clave, como entrada de pedidos, ventas, finanzas y fabricación. Por lo tanto, tiene mucho sentido ubicar conjuntamente el data warehouse junto con los sistemas fuente que ya se estén ejecutándose en la nube. La ubicación conjunta ofrece una carga de datos más rápida, lo que significa que los usuarios obtienen un acceso más oportuno a sus datos.

Principales ventajas de mover el data warehouse a la nube

Existen **tres ventajas principales** para mover un data warehouse a la nube, y estas están directamente vinculados a los tres controladores clave enumerados anteriormente:

- Más fácil consolidación y racionalización
- Monetización más rápida de los datos en la nube
- La nube ofrece mejor protección

# 4. Data Warehouse vs Big Data vs BI: cuales son las diferencias

En este apartado compararemos brevemente los conceptos Data Warehouse (DW), Big Data y Business Intelligence (BI).

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para tener una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

Se trata de tres conceptos completamente diferentes que tienen en común una nueva manera de lidiar con los datos, siempre teniendo en cuenta la existencia de una gran volumen de información en varios formatos que contribuyen, de forma estructurada o no estructurada, a la toma de decisiones estratégicas. El objetivo final de cualquiera de estas tecnologías es ofrecer una ventaja competitiva a las empresas, pero la forma en que se utiliza es la que marcará la diferencia.

Descubre cómo ampliar tu negocio en la nube. ¡Contacta con un experto en CLOUD!

#### Big Data

Llamamos big data a un gran volumen de datos con una variedad, complejidad y velocidad de crecimiento enorme y que además tienen la característica de no ser estructurados. Eso significa que no son relacionales, estando además fuera del entorno corporativo. Es un tipo de tecnología que te permite analizar los datos en tiempo real y puede provenir de diferentes fuentes y formas, tales como mensajería instantánea, redes sociales, registros de grabaciones, imágenes, mensajes de correo electrónico, etc.

## Guía gratuita: "Del Bit... al Big Data", descubre todo lo que necesitas saber sobre Big Data y mucho más.

Para tener una idea de la importancia del big data, este mercado está creciendo cada año, alrededor de un 40%. La paradoja reside en el hecho de que cuanto mayor es la evolución del big data, mayor es la escasez de profesionales cualificados para satisfacer esa demanda. La expectativa es que en unos pocos años, haya un crecimiento aún más significativo para el uso de big data y por lo tanto un aumento en la demanda mano de obra especializada y de partners tecnológicos especializados en esta materia.

#### Data Warehouse

Por otro lado, un data warehouse almacena datos consolidados de diversas fuentes o sistemas de la empresa. Se trata de datos estructurados, que tiene como objetivo principal ser precisos y de alta calidad para de esta forma poder dar soporte a la toma de decisiones de la empresa. Se trata de conseguir tener todos los datos juntos para después poder dividirlos para hacer un análisis de determinados sectores o estrategias.

#### Business Intelligence

Un Business Intelligence (BI) es una especie de "cuello de botella" de los datos recogidos del data warehouse, que llegan de forma exacta y útil para ayudar a la toma de decisiones. Business Intelligence transforma los datos en información útil para analizar no sólo los negocios, sino también las principales estrategias corporativas.

# Guía gratuita "Integridad de datos y BI", descubre cómo sacarle el máximo partido a tu proyecto de Integridad de datos

Los tres conceptos están interconectados y la perspectiva es que, cada vez más, la mayoría de empresas utilicen el análisis generado por este tipo de tecnologías para una visión más analítica de su negocio y así poder tomar las mejores decisiones para crecer.

## 5. Material complementario

Soluciones y recursos para Data Warehouse

- Informatica Cloud para Microsoft Azure SQL Data Warehouse
- Test Data Warehouse
- MDM y calidad de datos para el data warehouse

#### Guías

- Data Lake: Superando las limitaciones del Data Warehouse
- El data warehouse inteligente

#### Artículos

- ¿Qué es warehouse y qué debe contemplar un data warehouse moderno?
- Beneficios y soluciones para un Data Warehouse en la nube
- Lo que necesitas saber acerca de la arquitectura de un datawarehouse
- Data warehousing: 3 ocasiones de hacer las cosas bien
- ¿Qué es un Data Warehouse y cómo saber cuándo lo necesitas implementar?
- Cómo optimizar la arquitectura de un data warehouse
- Data warehouse concepts: tu glosario de almacén de datos
- Data Warehouse, el primer paso para una solución de BI
- Tutoriales sobre DataWarehouse en pdf
- Cómo gestiona Informatica los nuevos medios de almacenamiento de datos
- ¿Qué es un Data Warehouse para un director de ventas?
- Bl en línea: el reto actual de los data warehouses
- El poder de data warehouse: Informatica PowerCenter



