



**Es el momento para modernizar sus técnicas
de administración de grandes datos**

Thomas H. Davenport y Ashish Verma

EL ÚLTIMO PERÍODO importante de innovación en la administración de datos fue en los 1980. Las compañías comienzan a darse cuenta entonces que necesitan un lugar permanente para almacenar los datos usados para inteligencia de negocios y análisis. El Wells Fargo Bank recibió, por ejemplo, su primer sistema de almacenamiento de datos de la empresa [enterprise data warehouse (EDW)] a finales de 1983. Este sistema de vanguardia empleó procesamiento en paralelo de datos de bases de datos relacionales, y muchas otras firmas encontraron que ella es una tecnología útil.

Pero la tecnología de administración de datos usada exitosamente en los últimos 30 años no es hoy la tecnología más eficiente y efectiva. Muchas formas de grandes datos, incluyendo imágenes, medios de comunicación social, y datos del sensor, pueden ser difícil de colocar en el formato relacional de filas-y-columnas usualmente requerido por un EDW. Su volumen también los hace costosos para almacenar en una estructura EDW tradicional.

Afortunadamente, en la última década han surgido tecnologías nuevas que de manera radical están cambiando lo que constituye la mejor práctica en las técnicas contemporáneas de administración de datos, incluyendo Hadcoop y otros proyectos de fuente abierta, arquitectura basada-en-la-nube, enfoques para administrar la administración de datos, y los nuevos entornos de hardware para almacenamiento. El precio/desempeño de esas herramientas es sustancialmente mejor que para las tecnologías anteriores, a menudo por uno o más órdenes de magnitud. Incluso la corriente principal de los vendedores de la era anterior de administración de datos ahora está ofreciendo una variedad de productos y servicios que incorporan esas nuevas tecnologías.

Deje que los objetivos de negocio orienten el cambio

Pero la disponibilidad de mejor tecnología está lejos de la única razón para modernizar el entorno de sus datos. Las necesidades del negocio también están llevando a cambio sustancial en el entorno de los datos, y deben ser el orientador último para las iniciativas de modernización. Los objetivos de negocio que podrían motivar un nuevo enfoque para los datos incluyen un énfasis incrementado puesto en entender y predecir mediante analíticas las tendencias del negocio, el deseo por aplicaciones de aprendizaje de máquina e inteligencia artificial en procesos clave basados-en-conocimiento, la necesidad de transmitir

datos provenientes de y hacia máquinas que usan Internet de las Cosas, o preocupaciones incrementadas por seguridad y privacidad. En muchos casos, esas metas simplemente no pueden ser logradas sin modernización de los datos. Un caso de negocios sólido será crítico para las organizaciones que busquen modernizar sus datos; de otra manera, el esfuerzo se sentirá como una abstracción.

En Disney, por ejemplo, el orientador primario de una plataforma de datos modernizada fue la necesidad de mejores analíticas. Los productos de entretenimiento y medios sociales de comunicación tradicionalmente fueron liberados en el mercado con poca capacidad para medir su consumo, pero ahora casi todas las ofertas de los medios de comunicación de hoy pueden ser medidos y sus audiencias analizadas. Para permitir un rango diverso de actividades analíticas, Disney desarrolló una hoja de ruta para una capacidad sofisticada de datos y analíticas, incluyendo un lago de datos, un nuevo conjunto de herramientas analíticas, y un conjunto de casos de uso de negocios para tomar ventaja de las nuevas tecnologías.

Lagunas de datos y su administración

Estos tipos de proyectos típicamente resultan en la implementación de una laguna de datos, o un depósito de datos que permite el almacenamiento de datos en virtualmente cualquier formato. Las lagunas de datos típicamente están basadas en un programa de fuente abierta para servicios distribuidos de archivos, tales como Hadoop. Ellos permiten almacenamientos de datos en gran medida a costo relativamente bajo. Sin embargo, hay múltiples enfoques para las lagunas de datos; por ejemplo, algunas se basan en la nube, algunas en premisa. Los diferentes enfoques de lagunas de datos también proporcionan diferentes niveles de seguridad y gobierno. Por consiguiente, es importante planear cuidadosamente un esfuerzo de modernización antes de implementar cualquier tecnología particular.

Las lagunas de datos también tienen que ser administradas cuidadosamente en orden a no volverse “pantanos de datos” – lagunas con datos de baja calidad, pobemente catalogados, a los cuales no se puede acceder fácilmente. Y en algún momento, la mayoría de datos no-estructurados basados en una laguna de datos necesitarán ser colocados en forma estructurada en orden a ser analizados. Las lagunas de datos, entonces, requieren que los enfoques de administración sean definidos por adelantado para

asegurar calidad, accesibilidad, y las necesarias transformaciones de los datos.

Deloitte le ayudó a una firma global de tecnología, por ejemplo, a hacer la transición desde un almacén de datos de empresa de 600 terabytes hacia una plataforma de laguna de datos. Los datos son usados por 2,800 empleados, de manera que el proceso de conversión necesitó involucrar disruptión mínima. El almacenamiento de laguna todavía usa tecnologías en-la-premisa, pero la compañía ahora tiene un “nivel de consumo” en la nube para acceso fácil y rápido de los usuarios y para procesos automatizados. Y en lugar del tradicional proceso de “extraer, transformar y cargar” [“extract, transform, and load” (ETL)] que consume tiempo, los datos son transformados solo cuando es necesario para el análisis. En otras palabras, es un proceso ETL.

La mayoría de las organizaciones que establecen enfoques de modernización de los datos también intentan no elevar y cambiar los datos existentes en el nuevo entorno de datos. En lugar de ello, intentan al mismo tiempo hacer mejoramientos en los datos, incrementando la integración y la calidad a través de la empresa. Las firmas crecientemente están usando herramientas como aprendizaje de máquina para permitir el emparejamiento probabilístico de los datos; usando este enfoque, los datos que son similares, pero no exactamente los mismos que otros datos, pueden ser emparejados y combinados con poca intervención humana. Este método ascendente de integración de datos algunas veces puede ser más rápido y más efectivo que la mayoría de enfoques descendentes para la integración como la administración maestra de datos [Master Data Management].

La compañía farmacéutica global Glaxo-SmithKline, por ejemplo, usó este enfoque para modernizar e integrar sus datos para investigación y desarrollo. Pudo combinar millones de elementos de datos provenientes de tres dominios diferentes – experimentos, ensayos clínicos, y pruebas genéticas – en una sola laguna de datos basada-en-Hadoop. En solo tres meses la compañía pudo incorporar en la laguna el 100 por ciento de los datos deseados. Para trabajar a través de los tres dominios, el equipo de datos creó un nivel semántico integrado en lo alto de ellos con definiciones y significados estandarizados, y ahora está trabajando en cerca de 20 casos de uso diferentes para los datos dentro de la laguna.¹

Haciendo que suceda

Las compañías que hemos visto son exitosas en la administración de los datos tienen varios atributos comunes. Tales atributos incluyen:

- Involucramiento de ejecutivos clave del negocio – típicamente alguna combinación del CEO, director de información jefe, director financiero jefe, o director de mercadeo jefe – para definir las metas de negocio relacionadas-con-datos y asegurar los compromisos de recursos. Ellos deben ser partes de interés durante todo el proceso de modernización.
- Un proceso bien definido, que típicamente involucra un conjunto de actividades de “Imaginar” para diseñar el nuevo entorno de los datos; actividades de “Implementar” para diseñar, desarrollar, construir, probar, y extender las capacidades modernizadas, y un conjunto de actividades de “Operar” que involucran mantenimiento de aplicaciones e infraestructura, mejoramientos continuos y nuevos lanzamientos, y desarrollo de capacidades adicionales con el tiempo. Cada uno de esos pasos trabaja mejor en una serie de carreras ágiles.
- Constante y consistente entrega de valor; cada carrera debe intentar entregar valor medible. El enemigo más común de la modernización exitosa de datos es la fatiga organizacional y la incapacidad de los ejecutivos para ver el valor, no las limitaciones técnicas.
- Se necesitan nuevas capacidades humanas para implementar y administrar las nuevas tecnologías de datos. Dado que las habilidades para construir y administrar lagunas de datos son de suministro escaso, las compañías necesitan planear temprano cómo serán abastecerán de talento para la administración modernizada de grandes datos.

Las recompensas de negocio están en el almacén para las compañías que tienen éxito en esas iniciativas de modernización de datos. De manera similar, las organizaciones que fallan en emprender o tener éxito en los proyectos de modernización podrían encontrarse a sí mismas en desventaja competitiva proveniente de su incapacidad para implementar modelos y estrategias de negocio intensivos-en-datos.

¹ Thomas H. Davenport and Randy Bean, “Biting the big data management bullet at GlaxoSmithKline,” *Forbes*, January 8, 2018.

Deloitte. Insights

Suscríbase para actualizaciones de Deloitte Insights en www.deloitte.com/insights.



Siga a @DeloitteInsight

Colaboradores

Editorial: Ramani Moses, Abrar Khan, y Blythe Hurley

Promoción: Shraddha Sachdev

Artes: Kevin Weier

Acerca de Deloitte Insights

Deloitte Insights publica artículos originales, reportes y publicaciones periódicas que proporcionan ideas para negocios, el sector público y ONG. Nuestra meta es aprovechar la investigación y experiencia de nuestra organización de servicios profesionales, y la de coautores en academia y negocios, para avanzar la conversación sobre un espectro amplio de temas de interés para ejecutivos y líderes del gobierno. Deloitte Insights es una huella de Deloitte Development LLC.

Acerca de esta publicación

Esta publicación solo contiene información general, y nadie de Deloitte Touche Tohmatsu Limited, sus firmas miembros, o sus afiliados están, por medio de esta publicación, prestando asesoría o servicios de contabilidad, negocios, finanzas, inversión, legal, impuestos, u otros de carácter profesional. Esta publicación no sustituye tales asesoría o servicios profesionales, ni debe ser usada como base para cualquier decisión o acción que pueda afectar sus finanzas o sus negocios. Antes de tomar cualquier decisión o realizar cualquier acción que pueda afectar sus finanzas o sus negocios, usted debe consultar un asesor profesional calificado.

Nadie de Deloitte Touche Tohmatsu Limited, sus firmas miembros, o sus respectivos afiliados serán responsables por cualquier pérdida tenida por cualquier persona que confíe en esta publicación.

About Deloitte

Deloitte se refiere a uno o más de Deloitte Touche Tohmatsu Limited, una compañía privada del Reino Unido limitada por garantía ("DTTL"), su red de firmas miembros, y sus entidades relacionadas. DTTL y cada una de sus firmas miembros son entidades legalmente separadas e independientes. DTTL (también referida como "Deloitte Global") no presta servicios a clientes. En los Estados Unidos, Deloitte se refiere a una o más de las firmas de los Estados Unidos miembros de DTTL, sus entidades relacionadas que operan usando el nombre "Deloitte" en los Estados Unidos y sus respectivas afiliadas. Ciertos servicios pueden no estar disponibles para atestar clientes según las reglas y regulaciones de la contaduría pública. Para aprender más acerca de nuestra red global de firmas miembros por favor vea www.deloitte.com/about.

Copyright © 2018 Deloitte Development LLC. Todos los derechos reservados
Miembro de Deloitte Touche Tohmatsu Limited

Documento original: "**It's time to modernize your big data management techniques**", Deloitte Insights, July 20, 2018. By Tom Davenport and Ashish Verma.

<https://www2.deloitte.com/insights/us/en/topics/analytics/data-management-techniques-approaches-tools.html>

Traducción realizada por Samuel A. Mantilla, asesor de investigación contable de Deloitte & Touche Ltda., Colombia, con la revisión técnica de César Cheng, Socio Director General de Deloitte & Touche Ltda., Colombia.