

## El dato se muda al corazón del negocio

La frontera del dato se está moviendo. Los datos, antes una preocupación del área de IT, han pasado a ser una preocupación central de los responsables de negocio de las compañías. Antes se usaban casi exclusivamente con una visión transaccional, mientras que ahora se han convertido en un activo estratégico para las empresas, por el valor que se puede extraer de ellos.

Esto, que ahora está en boca de todos y parece una obviedad, no lo era tanto hace tan solo unos años. Entonces, las personas que tomaban las decisiones de negocio en las compañías, disponían de sistemas de *reporting* unidireccionales y costosos de evolucionar. Y se servían de hojas de cálculo para trabajar con la poca información a la que tenían acceso, con el objetivo siempre de tomar las mejores decisiones. Esta inercia generó silos de información departamentales, que daban lugar a diferentes resultados y puntos de vista sobre aspectos clave del negocio como, por ejemplo, el volumen de ventas, las unidades en *stock*, los consumos incurridos...

¿Qué es lo que ha pasado en este tiempo? ¿Por qué existe una creciente sensación de frustración en las empresas por la dificultad para acceder a la información relevante para la toma de decisiones del día a día?

Lo que ha ocurrido es que ha **irrupido con fuerza en el mercado una nueva generación de herramientas de visualización y análisis de información**, mucho más sencillas de utilizar por las áreas de negocio. Unas herramientas que, además, son capaces de acabar con los silos de información dentro de las compañías. Sin embargo, las unidades de negocio no son capaces de sacar partido de ellas porque las áreas de tecnología de sus empresas no han evolucionado a la misma velocidad a la que aparecen estas nuevas necesidades. En parte, porque con los años, las organizaciones se han cargado con una gran cantidad de sistemas operacionales que son necesarios para la marcha del negocio, y que les absorben una gran parte de sus recursos. Hasta ahora, las necesidades de información no han tenido la misma prioridad que el funcionamiento 24x7 de los sistemas transaccionales, y esto es lo primero que tiene que evolucionar.

El paradigma *big data* está siendo el gran catalizador de este cambio. Nadie se cuestiona ya si la **analítica avanzada** funciona o si debe formar parte del día a día del negocio. La pregunta, ahora, es **cómo sacarle partido antes de que sea demasiado tarde**. El tiempo de las pruebas de concepto ha pasado. Es necesario plantearse cómo se integran estas capacidades en el día a día de la toma de decisiones. Este proceso de integración está obligando a las organizaciones a evolucionar sus modelos operativos actuales, a través de la evolución de los procesos de negocio, de la incorporación de talento, de la modificación de estructuras de gobierno y del desarrollo de líneas estratégicas basadas en la explotación de datos.

Si hacemos un *zoom* al nivel de los datos, veremos que ya se ha desarrollado la tecnología que hace posible almacenar y analizar grandes volúmenes de información. Los algoritmos de inteligencia artificial, que estaban reservados para un pequeño grupo de empresas, se están democratizando. Pero la causa más relevante y, quizá, menos comprendida, que está facilitando que se mueva la frontera del dato, es el denominado *late schema committment*. Esto es, **la capacidad de almacenar los datos en su**

**estructura original y de tratarlos (darles un esquema concreto) justo en el momento de realizar el análisis que se necesite.**

El *late schema commitment* es el gran impulsor de la frontera del dato, que está cada vez más cerca de las áreas de negocio y, por tanto, más cerca de la toma de decisión de las empresas. Hace 15 años, para construir un sistema informacional que proveyera algún tipo de valor añadido al día a día de una compañía, era necesario diseñar y desarrollar un esquema de datos a-priori, sobre el cual se construirían, después, los cubos e informes que se distribuían a las personas que los consumían. El objetivo, que todo el mundo tuviera la misma visión de la información (la única versión de la verdad). Esto funciona muy bien cuando las preguntas de negocio se ciñen a la información almacenada en ese esquema -como por ejemplo “¿cuáles fueron las ventas del mes pasado?”. Pero no tanto cuando las preguntas no respetan esta estructura predefinida -“¿Qué ha pasado con las ventas de la zona norte la semana pasada?”. En la mayoría de los casos, este tipo de información se agrega en periodos determinados, por lo que para responder a la pregunta hay que bucear en los sistemas origen de la mano de un experto en la gestión de datos -no necesariamente conocedor del negocio- que nos ayude a dar con las respuestas. Y por el camino, esperar que lo que entienda uno sea lo mismo que ha dicho el otro. Hoy, los profesionales que reciben la información saben que si les dan acceso a los datos en una forma que ellos entiendan, pueden ser autosuficientes a la hora de buscar respuestas.

El grado de madurez actual de las tecnologías *big data* nos permite mover la frontera del dato hasta el consumidor final, pero la tecnología sola no va a resolver el problema

Estos profesionales están acostumbrados a trabajar con hojas de cálculo, porque les aportan flexibilidad e independencia. La primera y mayor flexibilidad parte del esquema de datos. Cada respuesta puede requerir de su propio esquema y el esquema se debe generar por el propio analista en el momento del análisis. La segunda es la capacidad de generar nuevos indicadores que ayuden a resolver la pregunta en cuestión. Por ello, también necesitan un entorno con capacidad de edición, cosa que supone un reto tecnológico para implementarlo a gran escala.

El grado de madurez actual de las tecnologías *big data* nos permite mover la frontera del dato hasta el consumidor final, pero la tecnología sola no va a resolver el problema por nosotros. Para conseguirlo hay que **tener en cuenta los siguientes factores:**

**La desnormalización de los datos.** Las fuentes de datos de las empresas están estructuradas de forma relacional, para optimizar el uso de almacenamiento. Con la tecnología *big data* se pueden crear los denominados tableros, que no dejan de ser gigantescas hojas de cálculo, donde cada registro tiene toda la información necesaria sobre una parte específica del negocio. De esta forma, podemos filtrar la información a voluntad, sin el riesgo de generar cruces de tablas erróneas que nos lleven a tomar decisiones incorrectas.

Para implementar este concepto es necesario trabajar en la definición de los **esquemas de datos funcionales**, consumibles por los gestores del negocio. Hace unos años, con la irrupción de *big data*, algunos se las prometían muy felices. Parecía que la tecnología iba a solucionar todos los quebraderos de cabeza que daba la gestión de los datos, y que al volcar estos en un *datalake* cualquiera iba a poder acceder a lo que necesitara. Acto seguido, apareció, en contraposición, el término *dataswamp* -en inglés, pantano, o ciénaga de datos-, según el cual, las tareas de gestión de la información para la toma de decisiones no se pueden dejar en manos de cualquiera.

Muchas organizaciones se están dando cuenta ahora de **que los entornos de *datalake* también tienen que estar gobernados** para que la gente los pueda utilizar de forma fiable. No solo eso, sino que un *datalake*, como sistema orientado a la generación de conocimiento, plantea nuevos retos a la hora de gobernar la información, como es el flujo de datos entrante y saliente, los niveles de seguridad del dato, su tiempo de vida, el almacenamiento de los resultados de los análisis para su nuevo consumo, y la adopción de las políticas regulatorias propias de cada sector sobre la información almacenada.

También es importante dotar a los diferentes actores de la organización de las **herramientas** necesarias para el consumo de información, así como el **soporte** necesario para hacer más sofisticados los análisis de forma controlada.

Y, para terminar, lo más importante, que es la **gestión del cambio en la empresa**. Es necesario formar a los receptores de la información y desarrollar nuevos procesos que garanticen la veracidad de los análisis. También es importante eliminar las barreras organizacionales que dificultan el flujo de información entre los sistemas donde se originan y los equipos que los consumen. Y esto requiere de la formación de equipos multidisciplinares y de nuevas estructuras organizacionales que empujen la frontera del dato hacia aquellos que toman las decisiones.