



# BIG DATA

El término Big Data hace referencia a la recolección, análisis y acumulación recurrente de gran cantidad de datos, incluidos datos personales, de una variedad de fuentes, que son sujetos a un procesamiento automático mediante algoritmos informáticos y procesamiento de datos avanzado, utilizando tanto datos almacenados, como transmitidos, para generar ciertas correlaciones, tendencias y patrones (gran análisis de datos).

Hoy en día, el Big Data se define comúnmente como datos que proceden de una mayor variedad de fuentes, que llega en volúmenes cada vez mayores y con una velocidad cada vez mayor. El crecimiento, la velocidad y la complejidad de los datos son impulsados por el despliegue de miles de millones de sensores inteligentes y dispositivos que están transmitiendo datos (popularmente llamados el *Internet de las Cosas*) y por otras fuentes de datos semi-estructurados y estructurados.

El gran potencial del Big Data reside en que, al ser los datos recopilados y analizados de forma continua, se puede extraer de ellos una información de gran valor que ayuda a orientar a un negocio en la toma de decisiones.

A pesar de que se reconoce que convertir los datos en información constituye una gran ventaja, cuestiones como el concepto de *Analytics* en tiempo real, veracidad de las diferentes fuentes de datos y la falta de presupuesto económico y capacidad tecnológica en la mayoría de los casos, son, sin duda, barreras que ralentizan la puesta en marcha de proyectos de Big Data.

No obstante, el concepto de Big Data no es nuevo ya que muchas empresas llevan ya manejando grandes volúmenes de datos desde hace tiempo. Lo que marca la diferencia entre las aplicaciones analíticas y de gestión y los nuevos conceptos de Big Data son cuatro palabras de referencia para empresas pioneras en este terreno: Volumen, Velocidad, Variedad y Valor.

## El BIG DATA definido por las cuatro “uves”

<b>Volumen</b>	<b>La cantidad de datos.</b> Hablamos de Big Data cuando los volúmenes de datos no pueden ser manejados por las herramientas tradicionales de gestión de bases de datos. Big data requiere el procesamiento de altos volúmenes de datos de baja densidad, datos de Hadoop <sup>1</sup> no estructurados, es decir, datos de valores desconocidos ( <i>mensajes de Twitter, flujos de clics en páginas web y aplicaciones móviles, tráfico de red, equipos con sensores que capturan datos a la velocidad de la luz, etc</i> ). El objetivo del análisis de Big data es convertir estos datos de Hadoop en información útil.
<b>Velocidad</b>	<b>El ritmo al que los datos se reciben y, posiblemente con el que se aplica alguna acción.</b> Los datos se reciben, cambian y evolucionan a una gran velocidad. Con los sistemas tradicionales es imposible analizar de forma inmediata los grandes volúmenes de datos que llegan y en muchas ocasiones son necesarios los análisis y las respuestas en tiempo real. Desde un punto de vista operativo, las experiencias de aplicaciones móviles tienen una gran base de usuarios, mayor tráfico de red y la expectativa de respuesta inmediata. Algunas aplicaciones del

<sup>1</sup> Software de código abierto que sirve para almacenar y analizar cantidades masivas de datos, tanto estructurados como no estructurados: terabytes o más de correo electrónico, lecturas de sensores, registros de servidor, fuentes de Twitter, señales de GPS..., cualquier tipo de datos que se pueda imaginar. Con Hadoop se pueden procesar grandes conjuntos de datos desordenados y obtener conocimiento y respuestas a partir de ellos.



	Internet de las cosas (IoT, por sus siglas en inglés) tienen ramificaciones de estado y seguridad que requieren acciones y evaluaciones en tiempo real. Otros productos inteligentes preparados para Internet funcionan prácticamente en tiempo real.
<b>Variedad</b>	<b>Nuevos tipos de datos no estructurados.</b> Este concepto hace referencia a la inclusión de otros tipos de fuentes de datos, diferentes a las tradicionales. Los tipos de datos no estructurados o semiestructurados, como el texto, audio o vídeo, requieren procesamiento adicional para obtener significado y metadatos de apoyo. Una vez que se comprenden, los datos no estructurados tienen los mismos requisitos que los datos estructurados, como el resumen, el alineamiento, la trazabilidad para auditorías y privacidad.
<b>Valor</b>	<b>Los datos tienen un valor intrínseco que debe descubrirse.</b> Existe una amplia gama de técnicas cuantitativas y de investigación para extraer valor de un gran conjunto de datos, sin necesidad de tener que usar solo una muestra de los mismos. Sin embargo, encontrar este valor también requiere nuevos procesos de descubrimiento en el que participen analistas, usuarios y ejecutivos con conocimientos. El verdadero reto de big data es el componente humano, es decir, aprender a formular las preguntas adecuadas, reconocer patrones, realizar suposiciones fundamentadas y predecir comportamientos.

### Oportunidades y aplicaciones para el sector financiero

En servicios financieros (banca y seguros), existen oportunidades significativas de obtención de beneficios mediante la aplicación de tecnologías y metodologías de Big Data y Analytics.

La principal área de aplicación de Big Data y Analytics la constituye la segmentación avanzada de clientes. La incorporación de nuevas fuentes de datos a las segmentaciones tradicionales posibilita la obtención de una visión y un entendimiento más completos de las necesidades de los clientes, lo que lleva aparejada la definición de propuestas de valor más adaptadas a los diferentes perfiles de la demanda. Conocer mejor a los clientes permite realizar un *targeting* más ajustado, y esto se traduce en mayor eficiencia y rendimiento de las acciones comerciales. Además, el Big Data permite definir las acciones apropiadas de clientes para conseguir una mayor fidelización de los mismos.

Otras aplicaciones del Big Data son la gestión de la "omnicanalidad"– tan importante en un contexto en el que las nuevas tecnologías ponen a disposición de los clientes un mayor número de canales de comunicación mejorados y alta capacidad de transacción con las empresas – y la definición de estrategias de *pricing* dinámico por segmento de cliente, en un contexto que se acerca al *real time*.

Documentos de interés	
ORACLE	<a href="#">Guide to Big Data</a>
EY	<a href="#">Big Data en el sector financiero español</a>