



# WHITE PAPER



*Big Data*  
*El “inicio” de una nueva era*

### AUTORIA



#### **Ricardo Gestal**

8 años de experiencia en Consultoría y TIC  
Manager para Servicios Financieros y Seguros en  
Accenture Technology Consulting  
Ingeniero Superior en Telecomunicación por la UPC  
Linkedin: [www.linkedin.com/pub/ricardo-gestal-ruiz/20/374/2](http://www.linkedin.com/pub/ricardo-gestal-ruiz/20/374/2)



#### **Javier Vidal Plana**

15 años de experiencia en Consultoría en el sector IT Banca  
Arquitecto de Soluciones Analíticas en IBM  
Ingeniero de Telecomunicaciones por la UPC  
Linkedin: [www.linkedin.com/profile/view?id=7615031](http://www.linkedin.com/profile/view?id=7615031)



#### **Juan Carlos Carracedo**

29 años de experiencia en el sector TIC  
Director de Preventa SAP en Seidor  
Licenciado en Matemáticas e Ingeniero Informático  
Linkedin: [www.linkedin.com/in/juancarlos carracedo](http://www.linkedin.com/in/juancarlos carracedo)



#### **Josep Tarruella Guillaumet**

17 años de experiencia en el sector TIC  
Director de Servicios en Lantares  
Ingeniero Informático por la UPC y Máster Executive MBA por IESE  
Linkedin: [www.linkedin.com/in/joseptarruella](http://www.linkedin.com/in/joseptarruella)

## Índice

1. Introducción
2. ¿Qué es el Big Data y qué no es Big Data?
3. Nuevas fuentes de información
  - 3.1. El crecimiento en la complejidad de los datos
  - 3.2. El crecimiento en la complejidad analítica
4. Oportunidades que brinda el Big Data
  - 4.1. Conocer al cliente: la microsegmentación
  - 4.2. Información de geolocalización
  - 4.3. Mejora en los procesos de toma de decisiones
  - 4.4. Nuevos modelos de negocio
5. Impacto del Big Data en las organizaciones
6. Casos de Uso
7. Conclusión

### 1. Introducción

En 1943, Thomas J. Watson, CEO de IBM dijo: “I think there is a world market for maybe five computers”<sup>1</sup>. Desde ese momento hasta la actualidad, donde existen más de 2.400 Millones<sup>2</sup> de personas conectadas a internet con dispositivos que caben en el bolsillo y que son mucho más potentes que los grandes ordenadores de hace apenas 15 años, han pasado muchísimas cosas.

Los sistemas informáticos han sido la gran revolución del siglo XX, y han supuesto tanto la transformación de los modelos de negocio existentes, como la aparición de nuevos modelos impensables anteriormente.

¿Por qué este breve recorrido por el pasado para realizar esta introducción? Pues porque todo apunta a que la irrupción del Big Data, que este artículo cubrirá en detalle, puede llegar a suponer una nueva revolución del mismo orden de magnitud.

Los primeros sistemas informáticos estaban orientados a la automatización de procesos costosos en esfuerzo y poco focalizados en el negocio, como es el caso de los ERP. Éstos permitieron acelerar los procesos y por tanto reducir costes y mejorar el funcionamiento interno de la compañía.

Un poco más tarde, a principio de la década de los 90, Bill Inmon y Ralf Kimball “inventan” el Data Warehouse. El cambio radical es que por primera vez se entiende que los datos de los que dispone la compañía tienen un alto valor por ellos mismos, y se utilizan para generar más valor para el negocio. Es por tanto invertir en los datos para mejorar el negocio, y no invertir en sistemas para ahorrar costes en la gestión interna de la compañía.

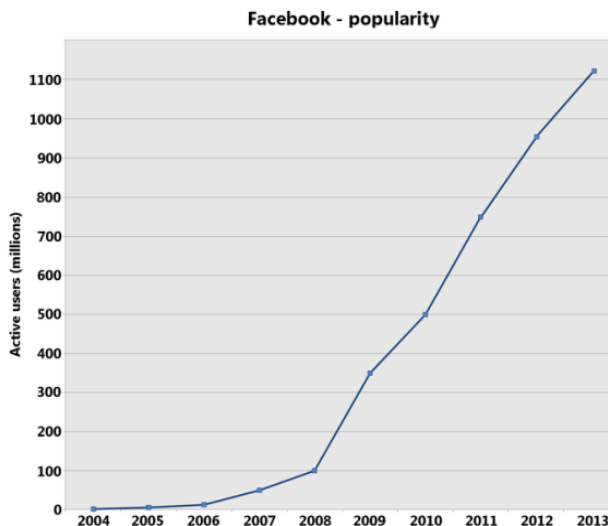
Big Data es el siguiente paso. Ya no hay compañías sin ERP, sin sistemas de soporte a la decisión (DWH, BI) más o menos automatizados, pero estos están orientados básicamente a los datos que se generan internamente. Lo que es Big Data y no es Big Data se tratará en el siguiente punto, pero si se pretende dejar en esta introducción unos primeros elementos de reflexión:

¿Existían las redes sociales hace sólo 10 años? Facebook pasó de 100 millones de usuarios en 2008 a más de 1.100 millones en 2013.

---

<sup>1</sup> Trad: “Creo que existe un mercado mundial para unos cinco ordenadores”

<sup>2</sup> <http://www.internetworldstats.com/stats.htm>



¿Llevaban nuestros “teléfonos” GPS, cámara, giroscopio, brújula, acelerómetro, conexión WIFI, 3G,... que permitiría saber desde nuestra posición en tiempo real hasta si acabamos de tener un accidente con el coche y quién lo conducía?

¿Llevaba nuestro coche, por volver a este tan extendido medio de transporte, miles de sensores que pueden dar información en tiempo real del estado del mismo, estadísticas de conducción, la presión de las ruedas,...?

Está claro que las respuestas son no, no y no. Y efectivamente los sistemas que existen en la actualidad no permiten gestionar estos volúmenes ingentes de información. El mundo está generando información a una velocidad vertiginosa, tanto en entornos virtuales (web, redes sociales, ...) como en el mundo físico, donde la sensorización de TODO está todavía empezando. Ya hoy, el último modelo de Boeing, el 787, genera más de ½ Terabyte de información en cada vuelo. Las ciudades se están sensorizando, cosa que conocemos como Smartcities, donde todo estará conectado, las farolas sabrán apagarse y encenderse según condiciones, los contenedores de basura sabrán cuándo están llenos y generarán órdenes de recogida, ... y así una infinidad de nuevas aplicaciones.

¿Pero dónde estamos? En el inicio. En el inicio del inicio más bien. El reto del Big Data será conseguir dar valor a estos datos, conseguir generar conocimiento útil para las empresas, organizaciones públicas, ciudades,... Como se decía al principio todo apunta hacia una nueva revolución, la de los datos, la que como toda revolución traerá sus propios retos como la seguridad y la privacidad de los mismos, la revolución que generará nuevos modelos de negocio y de nuevo transformará los actuales. En resumen, la revolución del Big Data.

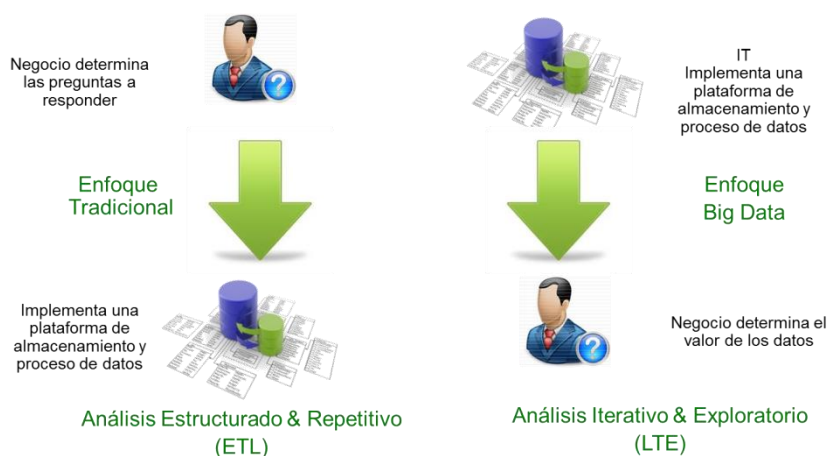
## 2. ¿Qué es el Big Data y qué no es Big Data?

Si bien el término Big Data no tiene una definición precisa, se puede resumir de forma sencilla como un nuevo tipo de soluciones capaces de capturar, administrar y procesar volúmenes de datos que por su tamaño y naturaleza no pueden ser procesados a un tiempo y coste razonables con las soluciones tradicionales de Business Intelligence (BI).

Una definición más formal podría tomarse del grupo de analistas de Gartner, que definen Big Data como un conjunto de activos de información muy variados, de alto volumen y altamente cambiantes que demandan formas innovadoras y eficientes en costes para el procesamiento de la información con el fin de mejorar las capacidades de decisión de los propietarios de los datos.

En este sentido, Big Data es un nuevo paradigma referente a cómo las compañías pueden obtener información relevante para sus estrategias de negocio. Tradicionalmente, se ha asumido de forma implícita que dicha información de valor se encontraba dentro de los repositorios de datos estructurados (bases de datos relacionales).

Como se tratará más adelante durante el resto del artículo, con la evolución de las tecnologías y de los hábitos de los consumidores han aparecido nuevas fuentes de información donde se pueden extraer datos muy valiosos para las compañías, que tradicionalmente se han descartado al considerarse información desestructurada de proceso sin valor para el negocio.



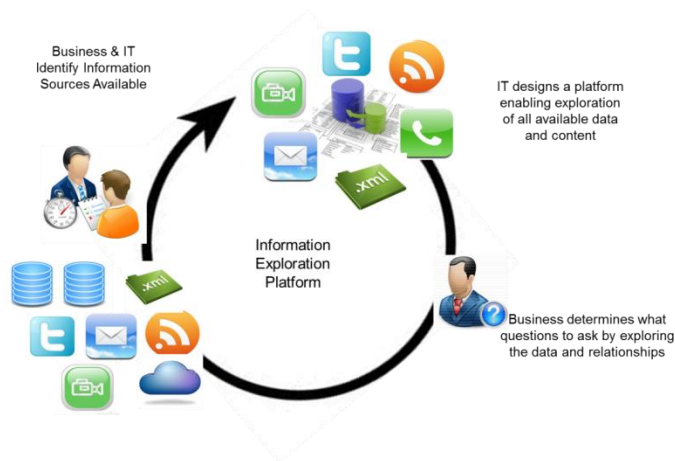
Al considerar el tratamiento de este tipo de datos, es frecuente la confusión de considerar que Big Data no es más que un incremento de la capacidad de almacenamiento para el

software de análisis de datos tradicional del que ya se disponga. Big Data implica cambios más allá de la tecnología dado que desde las áreas de negocio se cambia la concepción respecto a cómo explotar los datos. En este escenario, y de forma simplificada, se puede decir que los datos proponen las preguntas a realizar a negocio en lugar de que negocio haga preguntas predefinidas directamente a los datos estructurados.

En cualquier caso, sí es cierto que este cambio de paradigma impacta en las necesidades de almacenamiento para estas tecnologías, muy por encima de los requerimientos para los repositorios de información estructurada para los escenarios analíticos tradicionales.



De esta manera, la aproximación realizada por Big Data parte de la base de la potencial ventaja competitiva que se puede obtener dentro de los grandes volúmenes de datos no estructurados como pueden ser archivos de log, documentos de texto, tendencias en redes sociales, etc. que pueden contener una información valiosa que las compañías actualmente no están aprovechando.





A diferencia de las instalaciones BI basadas en grandes DataWareHouses (DWH), cuyas estructuras de datos estaban modeladas de acuerdo a los requerimientos de las áreas de negocio, el enfoque que proporciona Big Data está más orientado a ofrecer una visión consolidada de la información estructurada y desestructurada a partir de la cual las unidades de negocio puedan determinar las preguntas a realizar desde la exploración de esta visión de los datos y sus relaciones.

Así pues, Big Data no utiliza un modelo de datos relacional al uso definido de manera formal por los arquitectos de datos. En esencia, Big Data tiene una concepción del mundo tal y como es (no modelada), evitando los procesos de consolidación de datos en los que se estima que se pierde alrededor de un 80% de información explotable que en los escenarios analíticos no se tiene en consideración al ser descartados en este proceso. Al mantener esta percepción invariable del mundo, el contexto original no se ve distorsionado y se mantiene la capacidad de acceder a secciones más detalladas de los datos (drill-down) tras la realización de los análisis iniciales.

Es importante remarcar que Big Data no es un oráculo que vaya a proporcionar unas conclusiones absolutas que marque las pautas estratégicas para el negocio. Más bien debe asumirse que dará indicaciones y abrirá nuevos caminos que deberán ser explorados por las áreas de negocio antes de convertirlos en la estrategia a seguir ya que no todos los caminos llevarán a estrategias adecuadas.

### **3. Nuevas fuentes de información**

Durante muchos años, las empresas han construido un almacén de datos (Data Warehouse) para analizar la actividad comercial y generar conocimiento que ayuden en la toma de decisiones. Estos sistemas de análisis tradicionales se basan a menudo en un modelo clásico donde los datos de múltiples sistemas se capturan, se limpian, transforman y se integran antes de colocarlos en el almacén de datos.

Generalmente, los históricos de actividad del negocio se van acumulando a lo largo de los años permitiendo a las organizaciones utilizar herramientas de Business Intelligence para analizar, comparar e informar sobre la evolución del negocio en el tiempo. Adicionalmente, subconjuntos de datos se extraen frecuentemente en Datamarts, optimizados para realizar análisis multidimensionales específicos para las necesidades de un área de negocio de la empresa.



Sin embargo, aunque los sistemas tradicionales y las tecnologías utilizadas van evolucionando, las empresas ven oportunidades en la incorporación de nuevas fuentes de información a los sistemas de decisión, que quedan fuera de la “zona de confort” de los sistemas tradicionales de Business Intelligence, debido a la complejidad de este nuevo tipo de fuentes de datos.

Algunos ejemplos de este nuevo tipo de fuentes de datos son los siguientes:

- Clientes y potenciales clientes están creando grandes cantidades de nuevos datos en **las redes sociales y en sitios web**. Comentarios y opiniones vertidos en Twitter, Facebook, Pinterest, etc.. Estas redes recogen ingentes cantidades de opiniones, información sobre relaciones entre los usuarios, etc.. que pueden ser de gran valor para las empresas.
- Dentro de las empresas se generan multitud de **logs** cuya información no suele ser explotada (“dark data”). Los logs de los canales web están creciendo a medida que los clientes cambian a los canales online como forma preferida de realizar transacciones e interactuar con las empresas. Un caso específico son los CDR (Call Data Records) de las empresas de telecomunicaciones, que recogen el detalle de todas las llamadas y navegación de los usuarios. Un análisis temprano de estos registros puede mejorar las acciones de marketing y de monitorización de la calidad del servicio.
- **Datos procedentes de sensores** (the internet of things) se están empezando a utilizar para optimizar las operaciones de la empresa. El resultado es la abundancia de nuevos tipos de datos, con un volumen que crece rápidamente y nuevos flujos de datos que necesitan ser analizados en tiempo real. Por ejemplo las compañías eléctricas están empezando a utilizar contadores digitales (Smart meters) capaces de enviar lecturas automáticamente y en intervalos mucho más reducidos, por ejemplo consumo por horas, lo cual permite un seguimiento mucho más fino del consumo, pero que genera una explosión de los datos a almacenar.
- **Datos de geolocalización** procedentes de dispositivos GPS de los clientes o de los empleados. Por ejemplo, las compañías aseguradoras están ofreciendo seguros de automóvil en función del uso, que requiere rastrear la actividad de los vehículos asegurados, recogiendo datos como ubicación, tipo de carretera, velocidad, etc...
- **Datos procedentes de registros de vídeo**. El volumen de información registrada en vídeo está en crecimiento constante dada la gran variedad de dispositivos capaces de registrarlo, y será una tendencia creciente si iniciativas como Google Glass tienen éxito. En las empresas también proliferan los registros de vídeo por ejemplo de cámaras de seguridad. Actualmente existe la posibilidad de utilizar estas fuentes de información con objeto de obtener mejoras operativas. Por ejemplo, utilizar los registros de las cámaras de las oficinas bancarias para identificar patrones de comportamiento específicos, como formación de colas, patrones de movimiento dentro de las oficinas, identificación de clientes, etc...

- **Datos procedentes de registros de voz.** Por ejemplo de llamadas al Call Center. El objetivo puede ser realizar un análisis de la eficiencia de los agentes, identificación de clientes con quejas, etc...
- **Documentos.** En todas las empresas se manejan importantes cantidades de información en forma de documentos que se gestionan manualmente. Las nuevas tecnologías abren nuevas posibilidades de automatizar la extracción y estructuración de esta información. Un caso particular son las Historias Clínicas en Salud, donde existen grandes oportunidades de extracción de conocimiento de todas estas Historias, para agilizar su gestión, y por ejemplo para proveer de una fuente de información importantísima a la industria farmacéutica y de investigación en áreas de salud en general.
- **Open Data.** Existe una importante tendencia de los gobiernos y administraciones públicas para poner toda la información pública que gestionan a disposición de los ciudadanos y de las empresas, con un doble objetivo de por un lado aumentar la transparencia (se enmarca dentro de las tendencias de Gobierno Abierto), y por otra generar valor para el ciudadano y las empresas. Con este objetivo se han creado repositorios de datos públicos como el repositorio del Gobierno de España <http://datos.gob.es/datos/> , que publican conjuntos de datos relacionados con la economía, el tiempo, etc...

### El crecimiento en la complejidad de los datos

En relación con los datos, la complejidad ha aumentado por cuatro aspectos principalmente:

- **La variedad de tipos de datos:** datos no estructurados como texto, grabaciones de vídeo o grabaciones de voz. Un problema importante con estos nuevos tipos de datos es que no tienen un modelo de datos asociado, y debemos explorarla para acabar generando un modelo que tenga valor de negocio. Por esta razón, usualmente es necesario hacer análisis exploratorio en los datos no estructurados antes de poder incorporarlos en los sistemas analíticos tradicionales. En ocasiones estos tipos de datos requieren nuevas técnicas analíticas completamente específicas, como el procesado de lenguaje natural para el tratamiento de textos.
- **El volumen de los datos:** muchos de estos tipos de datos generan volúmenes ingentes de datos difíciles de manejar con los sistemas tradicionales. Ejemplos de estos volúmenes son las redes sociales, o los CDR (Call Data Records) en las empresas de telecomunicaciones, web logs, etc. Estas fuentes de datos rápidamente pueden acumular terabytes o incluso petabytes de información.
- **La velocidad de generación de los datos:** algunos de estos tipos de datos se generan a gran velocidad, lo cual exige sistemas de análisis muy escalables y rápidos, especialmente para la toma decisiones en tiempo real.
- **La veracidad de los datos:** algunas de las nuevas fuentes de datos, como las redes sociales, o los sistemas de geolocalización nos proporcionan información

que puede no ser exacta o correcta. La veracidad de los datos puede comprometer el valor de la información extraída de estos datos, por lo que es necesario desarrollar técnicas que minimicen el posible impacto de la falta de veracidad.

### **El crecimiento en la complejidad analítica**

En términos de complejidad analítica, los nuevos tipos de datos requieren nuevos tipos de algoritmos analíticos para poder extraer información de valor para el negocio.

Por ejemplo, el análisis de información de redes sociales requiere técnicas de tratamiento del lenguaje (Natural Language Processing) para poder identificar interés o sentimientos a partir de los comentarios realizados por los usuarios. El tratamiento de la posición obtenida de un GPS requiere técnicas de georeferenciación para ubicar al usuario en una determinada calle, etc...

Las campañas de marketing en las empresas de Distribución son un buen ejemplo donde las campañas y la selección del momento necesitan optimizarse, especialmente en los canales online en que los clientes están conectados a través de su teléfono móvil en tiempo real. Esto requiere un conocimiento del cliente mucho más detallado:

- Análisis histórico de las características demográficas del cliente y su historial de compras para segmentarlo en función de su perfil de compras
- Análisis de la cesta de la compra para identificar oportunidad de venta cruzada para cada cliente
- Click-stream análisis para entender su comportamiento online y patrones de consulta de productos para generar nuevas ofertas en tiempo real
- Análisis del perfil del cliente en redes sociales para identificar, gustos, comentarios, opiniones sobre productos, etc...
- Análisis en tiempo real de la geolocalización proporcionada por el móvil del cliente, para detectar cuándo está situado cerca de un comercio en el que podamos realizarle ofertas específicas.

Con esto queremos resaltar que, en muchos casos, obtener el conocimiento necesario para resolver una determinada situación de negocio es ahora un proceso que requiere múltiples tipos de análisis con diferentes fuentes de datos y diferentes niveles de complejidad.

Combinar el análisis de información estructurada con información no estructurada se requerirá en múltiples casos de negocio para obtener el conocimiento buscado. La integración de múltiples tipos de información será requerida para generar conocimiento sobre el que tomar decisiones.

A pesar de todas estas complejidades existe una creciente demanda en las empresas para incorporar nuevas fuentes de información en los procesos de decisión que permita generar ventajas competitivas.

### **4. Oportunidades que brinda el Big Data**

A estas alturas ya ha quedado patente que Big Data supone una revolución más que una evolución de los sistemas actuales. Y como toda revolución, viene acompañada de nuevas oportunidades para los que estén preparados llegado el momento.

#### **Conocer al cliente: la microsegmentación**

En los últimos años se ha observado un progresivo alineamiento de las empresas con las necesidades de su cliente. Y para ello es imprescindible conocerlo a fondo. Los primeros intentos fueron los sistemas CRM, pero éstos quedan de nuevo limitados en lo que a análisis de información externa se refiere.

Por tanto, el que mejor use la información actual de su base de clientes, el que mejor la combine con información externa (Redes sociales, geolocalización, ...) será capaz de desarrollar nuevos productos enfocados a cada necesidad.

La capacidad de analizar estos grandes volúmenes de información permitirá prácticamente trabajar a nivel de cliente en lo que a personalización se refiere. Así, se podrán generar campañas específicas para cada persona, teniendo en cuenta cientos de variables cuyas combinaciones hacen a cada cliente único. Por ejemplo, el “Pay as you drive” de los seguros de coche permiten que la prima de cada conductor se calcule según sus hábitos y no sólo teniendo en cuenta la edad, el sexo y la potencia del coche como ocurre ahora. ¿Por qué una persona que sólo usa el coche el sábado y el domingo debo pagar el mismo seguro que otra persona que lo usa todos los días? O un conductor que circula siempre por autopista, donde la estadística dice que hay menos accidentes, ¿por qué debe pagar lo mismo que otro que no la pisa?

Es por tanto la oportunidad de la personalización de los servicios. Conociendo de forma única al cliente es posible ofrecerle exactamente lo que necesita. Y para conocerle a este nivel se debe usar toda la información que existe, que no es poca.

#### **Información de geolocalización**

En los últimos años se está produciendo una explosión en el volumen de datos de geolocalización disponibles. Los móviles equipados con sistemas GPS, la ubicación por redes WIFI, o mediante las celdas de los sistemas de telefonía, son algunas de las tecnologías más comunes que proporcionan información de geolocalización. Esta

tendencia seguirá en aumento ya que se continúa incrementando el número de dispositivos equipados con GPS, como cámaras fotográficas, etc...

Esta información abre un abanico de nuevas posibilidades, de las que intentamos resaltar algunas:

- **Ofertas geolocalizadas:** Es uno de los casos de uso más comunes, la posibilidad de realizar ofertas personalizadas basadas en la ubicación concreta del usuario, de los comercios o servicios que tiene cerca, cruzando sus hábitos de comportamiento con el lugar y el momento ofrece la posibilidad de crear ofertas altamente relevantes para los usuarios.
- **Seguros de automóvil por uso:** La compañía puede conocer con precisión los hábitos de conducción, carreteras, frecuencia de uso, velocidad, etc... permitiendo adaptar los precios a todos estos elementos.
- **Cobro automático de peajes:** de la misma forma que las compañías de seguros pueden utilizar la geolocalización del automóvil para ajustar los precios, las compañías que gestionan las vías de comunicación pueden usarlos para conocer el uso que los automóviles hacen de las vías de comunicación y por lo tanto podría automatizarse en base a esta información el cobro de los peajes.
- **Servicios de emergencia:** los servicios de geolocalización pueden ser muy útiles para facilitar la gestión de situaciones de emergencia, localizando rápidamente a los posibles afectados, y la rápida planificación de los medios necesarios.
- **Planificación urbana:** La información agregada de los movimientos de las personas son una fuente de información muy valiosa para mejorar la planificación urbana, y de vías de comunicación.
- **Planificación de Comercios:** la información agregada del movimiento de las personas también puede ser muy útil para anticipar la posible demanda de una determinada zona comercial, y por lo tanto ayudar a la selección adecuada de las ubicaciones de comercios.

### **Mejora en los procesos de toma de decisiones**

Los procesos actuales de toma de decisiones podrán ser mejorados también con información más cualitativa. Hasta el momento se dispone de información de los sistemas financieros, de las operaciones, marketing, etc. La adición de información sobre el comportamiento de los clientes, sus opiniones, el uso que se hacen de los espacios en las tiendas virtuales, en las físicas,... podrá aportar una mejora en el proceso de toma de las decisiones.

### Nuevos modelos de negocio

Como en cada revolución aparecerán nuevos modelos de negocio que ahora ni se imaginan. El propio Big Data será una fuente de negocio, y de hecho ya lo es. Ya existen los primeros servicios online que permiten, por ejemplo, encontrar la opinión que los consumidores tienen de nuestra marca en redes sociales como Facebook o Twitter. Servicios que hacen que las empresas que no se pueden permitir su construcción interna, puedan acceder a esta valiosa información.

## 5. Impacto del Big Data en las organizaciones

Después de examinar más de 400 grandes compañías, un reciente estudio ha determinado que aquellas empresas que poseen las capacidades analíticas más avanzadas superan, por amplio margen, a sus competidores en varios aspectos:

- Es dos veces más probable que se encuentren como las mejores posicionadas en rendimiento financiero dentro de su industria.
- Es cinco veces más probable que tomen decisiones más rápido que sus competidores.
- Es tres veces más probable que ejecuten las decisiones tal y como lo habían planeado.
- Es dos veces más probable que utilicen datos con frecuencia a la hora de tomar decisiones.

Queda así patente que los pioneros en el uso del Big Data han obtenido una ventaja significativa sobre sus competidores. Esto explica el hecho de que en estos momentos, muchas compañías se pregunten en qué situación se encuentran en el mundo del Big Data respecto a sus rivales y si están obviando una nueva herramienta esencial para ser competitivos.

¿Pero qué es lo que necesita una organización para entrar en el mundo Big Data?

- En primer lugar se necesitan los datos. Grandes cantidades de datos en un formato que nos permita acceder a ellos y analizarlos de forma sencilla. La mayoría de las grandes compañías ya dispone de estos datos, a veces en tal cantidad, que les resulta imposible analizarlos.
- En segundo lugar es necesario disponer de una herramienta avanzada de análisis de datos. En la actualidad existen en el mercado numerosas herramientas tanto propietarias como de código abierto. Lo único que se requiere es disponer del personal con los conocimientos adecuados para poder usarlas.

- Esto nos lleva a la tercera necesidad, la competencia. El análisis avanzado de datos requiere de personal con conocimientos muy amplios. Desde conocimientos estadísticos de análisis hasta el conocimiento de la legalidad vigente sobre la privacidad de datos en diferentes países, pasando por entender en profundidad el negocio y cuáles son las fuentes relevantes de datos.

Pero incluso teniendo cubiertos estos tres puntos puede no conseguirse el objetivo, por la sencilla razón de que Big Data no es solo una nueva tecnología, de hecho es un programa de negocio que requiere del uso experto de una tecnología. Es por eso que no es suficiente con proporcionar la capacidad y conocimientos necesarios, y esperar que el departamento de IT o el de Marketing empiecen a proporcionar resultados. Es necesario introducir Big Data profundamente en la organización. Es la única forma de asegurar que la información se comparte entre las unidades de negocio.

Muchas compañías líderes han comenzado el proceso de introducción de Big Data planteándose objetivos claros: hay que adoptar Big Data como una forma nueva de hacer negocios. Se incorporará la analítica avanzada y las predicciones como elementos clave en la toma de decisiones. La defensa de estas ideas por parte de los líderes de la compañía es clave para que se produzca el cambio cultural necesario, pero es igualmente importante explicar a la organización de qué forma puede ayudar Big Data para mejorar el rendimiento de la compañía, hasta qué punto se pretende llegar y en qué aspectos se va a hacer foco (mejorar los productos o servicios existentes, mejorando los procesos internos, creando nuevos productos, transformando el modelo de negocio,...).

Una vez definidos los objetivos, los líderes del proceso deberán centrarse en desarrollar la capacidad analítica de forma horizontal en la organización. Para ello hay que conseguir vencer la resistencia al cambio y generar tanto los conocimientos como el deseo de usar los datos en toda la organización. Esto es seguramente la tarea más dura ya que las organizaciones no cambian fácilmente, y el valor de Big Data puede no ser evidente para todo el mundo. Es imprescindible incentivar las decisiones tomadas analíticamente, asegurando que poco a poco los datos son incorporados en todos los procesos de toma de decisiones clave.

Los líderes en el proceso pueden crear entonces una nueva unidad en la que centralizar todas las capacidades analíticas de la organización, como por ejemplo un Centro de Excelencia (CoE) dirigido por el jefe de análisis (CAO). Básicamente existen cuatro posibilidades de organización:

- Modelo departamental
- Modelo departamental con soporte central
- Centro de excelencia
- Completamente centralizado



Compañías como AT&T, Google, Amazon o Netflix han seguido alguno de estos modelos.

Obsérvese que el departamento de TI no es el propietario de Big Data en ningún caso. TI juega un papel crítico en proveer y mantener la infraestructura y herramientas necesarias para ejecutar Big Data, pero en ningún caso es propietario o gestiona las capacidades de adopción del Big Data en la compañía.

Construir de cero un CoE precisa de mucho tiempo y no es una tarea fácil. El personal necesario para su formación es uno de los problemas más importantes. Se requieren expertos en estadística, ingenieros de sistemas, gerentes del negocio, expertos legales. Solamente conseguir este personal puede llevar más de un año, y ponerlo en funcionamiento otro año más.

## 6. Casos de Uso

Dependiendo del ámbito de la industria los casos de uso de Big Data pueden variar aunque el fundamento tecnológico sea el mismo. De esta manera, podríamos tener ejemplos como los siguientes:

- Identificación de focos de blanqueo de dinero
- Gestión del fraude
- Gestión de riesgos (créditos, mercados, etc.)
- Análisis de datos de Call Centers
- Segmentación de mercado
- Encuestas digitales
- Análisis de logs

De forma más enfocada a las diferentes industrias, algunos ejemplos de las principales tendencias y aplicaciones donde se podrían aprovechar las ventajas de Big Data serían las siguientes:

Sector	Principales tendencias	Ejemplos de aplicaciones
Servicios Financieros	<ul style="list-style-type: none"> <li>• Colaboración con las áreas de márketing para disponer de información más detallada de los perfiles de los clientes</li> <li>• Unificación de múltiples fuentes de datos para mejorar el conocimiento detallado de los clientes y sus hábitos.</li> <li>• Generación de ofertas de productos personalizadas a través de canales digitales.</li> </ul>	<ul style="list-style-type: none"> <li>• Generación de perfiles de clientes a partir de la combinación de la información residente en redes sociales y sus hábitos de consumo.</li> <li>• Envío de ofertas de productos a través Facebook, Twitter, móvil, etc.</li> <li>• Análisis avanzado de datos para mejorar las ofertas de productos financieros y personalizarlas de acuerdo al perfil del cliente.</li> </ul>
Bienes de consumo	<ul style="list-style-type: none"> <li>• Combinación de datos de localización y redes sociales</li> <li>• Análisis de vídeo contextual</li> </ul>	<ul style="list-style-type: none"> <li>• Datos contextuales combinados para ofrecer productos y experiencia personalizada a través de canales web.</li> <li>• Análisis de los patrones de movimiento en las tiendas para mejorar la comodidad en las compras.</li> </ul>

Sector Sanitario	<ul style="list-style-type: none"> <li>• Automatización de los sistemas de monitorización de pacientes</li> <li>• Sistemas de soporte a la decisiones médicas</li> <li>• Análisis comparativo de tratamientos</li> <li>• Medicina personalizada</li> </ul>	<ul style="list-style-type: none"> <li>• Mejora en los tratamientos y el servicio a los pacientes. Al tener monitorizados todos los pacientes, se evitan las rondas de enfermería para realizar tareas de control.</li> <li>• Las decisiones médicas cada vez requieren el análisis de mayor número de variables y el análisis comparativo de múltiples opciones de tratamiento lo cual hace conveniente disponer de herramientas de apoyo a la decisión.</li> <li>• El análisis comparativo de diferentes tratamientos podría mejorarse mediante el análisis masivo de toda la información que generan los sistemas asistenciales.</li> <li>• Aparecen nuevos tratamientos que requieren complejos análisis personalizados del genoma de cada paciente para la selección del tratamiento más adecuado. Si estos análisis pueden hacerse para cada paciente se abre la puerta a una medicina mucho más personalizada.</li> </ul>
------------------	--	--

Seguros	<ul style="list-style-type: none"> <li>• Ajustar las primas a cada cliente</li> <li>• Detección de fraude</li> </ul>	<ul style="list-style-type: none"> <li>• Entender que el riesgo no es igual para todos, sino que se pueden llegar a crear primas ajustadas según el comportamiento de cada cliente, ya sea para coches, hogar, vida, ...</li> <li>• Nuevas fuentes de información pueden ayudar en la detección del fraude: geolocalización, redes sociales, etc...</li> </ul>
Sector Público	<ul style="list-style-type: none"> <li>• Planificación de recursos</li> <li>• Transparencia y Open Data</li> </ul>	<ul style="list-style-type: none"> <li>• Optimización de redes de transporte mediante la monitorización y actuación en tiempo real.</li> <li>• Mejorar la planificación de recursos públicos mediante nuevas fuentes de información y promoviendo el acceso público a los datos</li> </ul>

Retail	<ul style="list-style-type: none"> <li>• Integración de Social Media en los procesos de planificación</li> </ul>	<ul style="list-style-type: none"> <li>• Optimización de precios en base al análisis de precios de la competencia online</li> <li>• Estimación de la demanda en base a información extraída de internet y redes sociales.</li> </ul>
Fabricación	<ul style="list-style-type: none"> <li>• Integrar Redes Sociales en los procesos de diseño</li> <li>• Mantenimiento preventivo</li> </ul>	<ul style="list-style-type: none"> <li>• Open innovation, facilitando la recogida de ideas para nuevos desarrollos</li> <li>• Diseño de productos en base a conjoint análisis para la selección de los atributos del producto que más valoran los usuarios.</li> <li>• Análisis masivo de logs para realizar mantenimiento preventivo de los productos y reducir los costes de mantenimiento y de no disponibilidad.</li> </ul>
TODOS	<ul style="list-style-type: none"> <li>• Atención al cliente</li> </ul>	<ul style="list-style-type: none"> <li>• Análisis en tiempo real de las comunicaciones con el cliente, ya sea en Callcenter, mail, App, ...</li> </ul>

## 7. Conclusión

En conclusión, la aparición de nuevas fuentes de datos y el incremento exponencial de los volúmenes de información a tratar hacen que las soluciones de Big Data sean capaces de proveer nuevas capacidades sobre las soluciones de BI tradicionales dado que:

- Los tamaños de la información son excesivamente grandes, no siendo posible procesarlos de forma eficiente en tiempo y coste con las soluciones actuales.
- La exploración mediante mecanismos tradicionales de grandes volúmenes de información desestructurada es muy costosa y tiene un alto grado de incertidumbre en referencia al valor recuperado.
- Diferentes unidades de negocio de la compañía con estrategias de gestión de los datos diferenciada que provocan ineficiencias dado que se puede estar explotando el mismo dato de diferentes formas.

Todo esto influirá en cómo las compañías gestionan sus datos y permitirá generar nuevas oportunidades de negocio, que generarán nuevos productos y servicios no imaginados hasta ahora. Afectará con seguridad a la gran mayoría de los sectores que, paulatinamente añadirán Big Data tanto a sus procesos operativos como a los de toma de decisiones.

Veremos cómo la sensorización lo cambiará todo, la manera en que conducimos, la manera en que nuestra compañía de seguros calcula nuestra póliza, la manera en que los ayuntamientos y en general el sector público ahorran dinero en alumbrado, en recogida de residuos y un largo etcétera, a la vez que nos proporcionan nuevos servicios basados en datos. Veremos cómo se podrá ahorrar agua en ciudades y sobretodo en regadíos, o cómo podremos controlar el consumo de cada uno de los electrodomésticos que tenemos en casa y así hacer un consumo mucho más racional.

En definitiva, un nuevo paradigma de realidad basada en datos que será sin duda el siguiente gran paso de la humanidad.

### Sobre ACEC

La Asociación Catalana de Empresas Consultoras, ACEC, es una entidad de carácter privado, sin ánimo de lucro, constituida en el año 1993.

Está vinculada a la AEC (Asociación Española de Empresas de Consultoría) y es miembro de la FEACO (Federación Europea de Asociaciones de Consultoría).

La **ACEC** reúne a las mejores Consultoras con sede o implantadas en Cataluña. Agrupa a **más de 50 empresas de consultoría** de todos los Sectores, representando más del 75% del sector, con un volumen de facturación estimado superior a 1.800 millones de euros anuales y más de 2.000 empleados, siendo su mayoría titulados universitarios.

Los **objetivos** fundamentales de la ACEC son:

- ✓ Promover el máximo nivel de profesionalidad, ética y calidad en la práctica de la Consultoría.
- ✓ Promover el mercado de la Consultoría entre Empresas y Administraciones Públicas en beneficio de las Consultoras Asociadas y de la Sociedad.
- ✓ Contribuir a un mejor conocimiento del mercado de la Consultoría en todos sus ámbitos y la calidad del mismo, facilitando para ello una “acreditación” a las consultoras asociadas (Certificado de Miembro) que sirve de reconocimiento y elemento de diferenciación frente a las consultoras no asociadas.

#### ASSOCIACIÓ CATALANA D'EMPRESES CONSULTORES

Av. Portal de l'Àngel, 36 – 3ª planta

08002 BARCELONA

NIF: G60598687

Tel. 93 492 57 37

Fax. 93 492 57 01

[info@asocat.org](mailto:info@asocat.org)

[www.asocat.org](http://www.asocat.org)



#### ASSOCIACIÓN VINCULADA:



**Asociación Española  
de Empresas de Consultoría (AEC)**

[www.consultoras.org](http://www.consultoras.org)



**EMPRESAS ASOCIADAS**



**EMPRESAS ADHERIDAS**

ATOS  
AYESA  
DELAWARE  
DELOITTE  
ERNST & YOUNG  
GFI Informàtica  
HEWLETT PACKARD  
INSA

IOR CONSULTING  
NEORIS  
OESÍA  
SOPRA GROUP  
TEA-CEGOS  
TELVENT  
UNISYS  
VASS