

# Big Data - Levante - 28/03/2015

**E**l término *big data* equivale en español a «datos masivos». Es un concepto que cada vez se oye más pero se conoce poco. Últimamente aparecen artículos, ruedas de prensa, dossiers de revistas científicas, como *El profesional de la Información*, número de noviembre-diciembre de 2014, que dedicó un número monográfico a este tema, o en revistas de divulgación como el dossier que aparece en la publicación *Emprendedores* de marzo de 2015, también se están publicando libros que tratan sobre esta materia. La Wikipedia define el *big data* como un término aplicado a un conjunto de datos que superan la capacidad del *software* habitual para ser capturados, gestionados y procesados en un tiempo razonable.

El *big data* es el comienzo de una nueva era, la de los datos, que son el nuevo oro de las empresas. El *big data* es una de las tendencias más actuales en todos los ámbitos de la vida y en muchos sectores, tanto en la empresa como en la administración pública, ya que la generación de contenidos *online* se ha disparado en los últimos años. Los datos son considerados como el oro negro, el nuevo combustible o el motor de los negocios. Los datos son un activo en cualquier institución y se han convertido en una nueva fuente de información y de riqueza con un gran potencial en cuanto a la reutilización y generación de nuevos conocimientos. Ahora bien, el reto es convertir este gran volumen de datos que se generan y almacenan en información útil, para pasar de ser simple información a generar conocimiento. Esto implica cambios no sólo en la gestión de los datos, sino también, y más importante, en su análisis, su explotación y en los procesos de trabajo.

**Viktor Mayer-Schönberger**, profesor de la Universidad de Oxford y uno de los expertos más importantes en el tema de los datos masivos,

## BIG DATA



vos, coautor del libro *Big data, la revolución de los datos masivos*, nos comenta que la mayor cantidad de datos en el mundo está controlada por *Google* y que éste no existiría sin los datos masivos.

En el 2001, **Doug Laney** acuñó la definición de las «3 V» que componen el término *Big Data*. La gestión de los datos masivos implica controlar adecuadamente estas tres variables: volumen de datos, velocidad de transmisión y variedad de datos. Las herramientas *big data* pueden procesar cualquier información, tanto textual como audiovisual, información de sensores de dispositivos automáticos, datos geográficos, información extraída de las redes sociales, blogs, webs... En estos momentos ya se procesa más información de las máquinas que de las personas. Todo ello con una rapidez y con un análisis en tiempo real. A estas tres V de Laney podríamos añadir otras dos: valor de los datos y visualización. El poder recoger, tratar e interpretar los datos permiten a las empresas gestionar y adoptar soluciones a problemas y en cierto modo adelantarse a los acontecimientos y ser por ello más competitivas.

Los usos del *big data* como ya hemos comentado anteriormente se pueden dar en todos los ámbitos y sectores: en las comunicaciones, en el campo de la salud, en la banca, en la agricultura, en el deporte, en logística, o en la administración pública. Por poner un

caso concreto y centrándonos en el campo de la agricultura, que fue la primera donde se aplicó: en el año 1976 en el Departamento de Estadística de la Universidad de Carolina del Norte, Estados Unidos de América, en vistas a mejorar las técnicas de cultivo y conseguir una mejor cosecha, se aplicaron una serie de variables como la temperatura, o la tipología de abonos. Actualmente el estudio de la gestión de datos masivos está muy extendido en este sector. La aplicación del uso masivo de datos permite gestionar fincas o explotaciones agrarias en base a datos recogidos por sensores instalados en el terreno que facilitan información y permiten tomar decisiones, en vistas a un mejor aprovechamiento del riego, abonado, etc. con un consiguiente ahorro de agua y de gastos económicos en la compra de material. En definitiva, se consigue una mejor gestión y una mayor cosecha. Asimismo, también con el uso de los datos se puede anticipar a las enfermedades o plagas de los cultivos y conseguir un mayor aprovechamiento en el tipo de plantación. Otro caso concreto sería la aplicación de herramientas *big data* en la administración pública, ya que es la mayor generadora de datos: dispone de datos fiscales, sociológicos, económicos, etc. de los ciudadanos. La utilización de datos masivos en la administración pública permite cruzar datos, detecta posibles delitos, se conocen con más detalle los datos fiscales, se mejoraran los resultados lo que redundaría en beneficio de los ciudadanos.

¿Quién puede gestionar toda esta información? Como se indica en un artículo de **Ana Torres**, publicado el 16 de febrero de 2015 titulado *No hace falta ser matemático para ser experto en "big data"*, en un futuro no muy lejano harán falta profesionales que traten los datos, apliquen técnicas de explotación y saquen el

máximo valor de los mismos. Prueba de ello es la oferta de cursos de postgrado sobre las técnicas *big data* que están presentando ya en algunas universidades. En diciembre de 2014 se realizaron unas jornadas organizadas por el Instituto Cervantes con el título: *Big data y Bibliotecas: convertir datos en conocimiento*. Siguiendo en esta línea REBIUN (Red de Bibliotecas Universitarias Españolas) ha abierto una brecha con un estudio sobre la aplicación del *big data* en las universidades y el perfil del *big data* librarian como gestor de datos dentro de las propias universidades. El papel de la biblioteca es clave; como profesionales de la información podemos aportar conocimientos y habilidades en aspectos de procesamiento y visualización de los datos para crear valor y nuevos servicios en cuanto al uso y gestión del *big data*. Resulta imprescindible recopilar, organizar y gestionar estos datos masivos previamente para que puedan ser reutilizados y compartidos. El grupo que elabora este informe está formado por: la Universitat de Girona, Universitat de Lleida, Universitat Politècnica de Catalunya, Universidad de Salamanca y Universitat Oberta de Catalunya; se decide por parte de los componentes que sea la Universitat Jaume I la coordinadora, en vistas a presentarlo en noviembre de 2015 en la Universidad de Cantabria, en la Asamblea Anual de REBIUN. En dicho estudio se propone aprovechar sinergias con otros servicios de la Universidad como Informática o la Oficina de Transferencia para los datos de investigación (*research data*).

Como conclusión, la gestión del *big data* es una oportunidad y una estrategia en cualquier organización. Con el estudio y análisis de datos se conseguirán decisiones más acertadas y permitirá a los gestores tener una visión de futuro. Es una nueva forma de gobernanza, en base al análisis de los datos de las instituciones tanto públicas como privadas.