

BIG DATA

- **INTRODUCCIÓN**

El autor habla sobre diversos conceptos que guardan una estrecha relación con el Big Data, como Business Intelligence, minería de datos u open data.

Asimismo, frente a las denominadas "3V" que constituyen la esencia del Big Data (volumen, variabilidad, velocidad), sugiere una Big Data cuarta V, la de la visualización. Finalmente, tras anticipar la previsible problemática en torno a la gestión de la privacidad de la información, concluye con una reflexión acerca de la dimensión del concepto en Big un contexto de crecimiento exponencial de la información y la seguridad que debe proporcionar la misma.

- **DESARROLLO**

Big Data apareció el pasado año como uno de los términos de moda en todas las revistas de temática científica, sociológica o tecnológica, también en y redes sociales e incluso ya ha dado el salto a las blogs publicaciones económicas y empresariales y las de divulgación más popular.

Tradicionalmente, los principales conceptos agrupados que han definido este nombre han sido las denominadas "3 V": volumen, variabilidad y velocidad.

Macro datos es todo aquello que tiene que ver con grandes Volúmenes de información que se mueven o analizan a alta Velocidad y que pueden presentar una compleja Variabilidad en cuanto a la estructura de su composición.

Siempre me ha parecido que debería añadirse una cuarta uve, la Visualización, ya que no solo forma también parte de ello, sino que muchas de las imágenes que nos traen a la memoria el trabajo con tienen que ver con estas Big Data nuevas formas de 'ver' estos datos.

También al Big Data se lo puede definir como aquel conjunto de datos que excede la capacidad de procesamiento de los sistemas de bases de datos convencionales.

Ventajas y desventajas



Es importante comprender que además de los datos estructurados, aquellos otros que provienen de fuentes de información conocidas y que, por tanto, son fáciles de medir y analizar a través de los sistemas tradicionales, empezamos a poder y querer manejar datos no estructurados: los que llegan de la Web, de las cámaras de los móviles y vídeos, redes sociales, sensores de las ciudades y edificios... La variedad de su origen, además de la rapidez con la que se incrementa su volumen, son algunos de los factores que habían dificultado su análisis hasta ahora. El nuevo y los nuevos modelos permiten software la incorporación a los estudios tanto de un tipo como de otro. Los avances en análisis semántico también permiten estructurar textos escritos por personas de forma automática.

Por ende aquí se pueden mencionar a los atributos principales para considerar un conjunto de datos como Big Data:



- **Volumen:** Cantidad masiva del orden de los cientos de Terabytes.
- **Velocidad:** El flujo de entrada de datos, ingesta, puede resultar tan alto que la información sea muy difícil de procesar y esto condicione la comprensión sobre los datos.

- Variedad: En la estructura de los datos: estructurados, semiestructurados (ficheros con formato conocido tipo Word, Excel) y desestructurados (imágenes, video, audio).
- Valor: La mayoría de los datos tiene poco valor hasta su análisis y procesamiento. Para esta tarea se hacen necesarios sistemas distribuidos y recursos avanzados de análisis de datos.

Tipos de visualización

- Grafos: El análisis de grafos es un aspecto que cada vez adquiere mayor importancia en ámbitos científicos y empresariales.

Tipos de información

- Información geoposicionada: La visualización geoposicionada de los datos es otra de las técnicas ampliamente usadas en Visual Analytics. Mediante el uso de capas añadidas en la cartografía y la inserción de elementos gráficos para representar valores específicos (barras, graficas de tipo tarta, etc.) se pueden representar grandes cantidades de datos a nivel global. Para ello se introducen tecnologías de rendering 3D que permiten representar datos escalares, vectoriales o incluso volúmetricos.
- Información espacio temporal.-El tipo de información espacio-temporal es un tipo de contenido de un amplio uso en infinidad de aplicaciones. La necesidad de analizar la evolución de ciertas variables está en la base de cualquier predicción. De forma clásica, las referencias temporal y espacial han sido tratadas como cualquier otro dato numérico. Sin embargo, en muchos casos esto no permite un análisis válido de los resultados. Mientras que la información espacial puede tener un tratamiento más o menos directo en su representación gráfica (que también es espacial) la información temporal requiere de procesos de abstracción para que pueda ser visualizada,

Campos de aplicación:

Este nuevo mundo está creando nuevos perfiles profesionales siendo el conocido como científico de datos el más citado. Los científicos de datos son profesionales con habilidades en matemáticas, estadística e ingeniería informática, que son capaces de extraer el máximo valor de los datos de la organización, cerrando la brecha entre las necesidades del negocio o la Administración y las Tecnologías de la Información.

El análisis Big Data es aplicable cada vez en más ámbitos y está permitiendo romper los silos de información para ofrecer una visión completa de toda la información relevante para la toma de decisiones, permite además obtener información relevante en un periodo de tiempo donde esta información todavía es válida

Se encuentra en el ámbito empresarial, en el mundo de los negocios, también está relacionado con la minería de datos y con el mundo de la prensa.

Herramientas

- Existe ya una gama amplia de utilidades (principalmente en la ayuda de toma de decisiones en situaciones críticas) es una de las herramientas fundamentales para la creación de valor sobre Big Data. En este sentido, además de las herramientas de procesamiento, son totalmente necesarios sistemas de visualización avanzada e interfaces que permitan interactuar con la propia visualización. La llegada de HTML-5 ha impulsado la aparición de un gran número de herramientas de visualización web adecuadas para la visualización científica y Visual Analytics.
- El problema que posee Big Data es el de la privacidad que es en donde se deben enfatizar así como también en la seguridad, en la propiedad intelectual, e incluso con la responsabilidad, estos son los aspectos que deben ser abordados para que se pueda continuar con el desarrollo de los sistemas de Big Data.

- **CONCLUSIÓN**

Big Data es un concepto con un gran potencial que ya empieza a tener una importante difusión pero aún el potencial del concepto no ha sido todavía explorado en su totalidad. Son muchos los nuevos campos y oportunidades que se abren en torno a este mundo. Es en la capacidad de analizar la gran cantidad de información que nos ofrece el Big Data donde las herramientas avanzadas de visualización son y serán un factor clave en el éxito del desarrollo de estas tecnologías.

MAS INFORMACION

<https://www.ibm.com/developerworks/ssa/local/im/que-es-big-data/>

<http://www.eleconomista.es/tecnologia/noticias/5578707/02/14/La-mod-a-del-Big-Data-En-que-consiste-en-realidad.html>