

BIG DATA

Universo de oportunidades

Jesús Lares
jesus.lares@ciens.ucv.ve

El mundo globalizado en el cual vivimos, dominado por la Internet y una gama infinita de tecnologías a su alrededor que van desde aplicaciones especializadas de negocios, sitios web, redes sociales, documentos, música, videos, textos; hasta dispositivos móviles inteligentes, sensores, computadores, servidores, entre otros, forman parte de una realidad abrumadora, que cada día genera gran cantidad de datos.

Las dinámicas actuales de flujo de datos han abierto un universo de oportunidades basadas en el conocimiento que es posible obtener; razón por la cual las empresas han comenzado a requerir personal con conocimiento técnico y científico, que les ayude a analizar los grandes volúmenes de datos con la finalidad de generar ideas, estructuras, sistemas, a partir de los datos ya analizados y refinados, que les permita ser competitivas en el tiempo. Para poder llevar a cabo este proceso es necesario entender de Big Data y sus áreas afines.



¿Qué es Big Data?

Big Data es un concepto que se ha puesto de moda en el ámbito tecnológico, es un término de negocios que abarca diferentes elementos informáticos y científicos que en algunos casos suelen ser difíciles de engranar. En un sentido estricto, se refiere al manejo de grandes volúmenes de datos; sin embargo para que se pueda generar información valiosa debe combinarse con la Ciencia de datos.

Pero, ¿Qué son grandes volúmenes de datos? La respuesta es: depende de la empresa u organización. Cada empresa, sobre la base de sus necesidades y la

dinámica de su negocio, genera su propio volumen de datos, el principal problema es: ¿Qué se hace con esos datos? y si dispone de capacidades computacionales que puedan generar información relevante que permita a las organizaciones tomar decisiones acordes con las estrategias y planes de negocios de la organización.

La intervención de las ciencias como: estadística, las matemáticas, la tecnología de la información, la minería de datos, el aprendizaje automático, la inteligencia artificial y un sinfín de áreas, han permitido generar, descubrir y determinar tendencias que ayudan a establecer patrones y hacer predicciones, a partir de los datos. Si no se cuenta con capacidades científicas, técnicas y tecnológicas dentro de la organización para extraer conocimiento de los datos, debemos comenzar a pensar en Big Data como una posible solución.

El Big Data puede caracterizarse desde cinco puntos de vista:

El Volumen de los datos, que describe el tamaño en bytes.

Las 5 V's de Big Data:

La Velocidad con que se generan y se deben procesar los datos.

La Variedad, son las diversas fuentes y tipos de datos, que se almacenará y manipulará.

El Valor que se puede obtener de los datos a partir de que estos sean procesados y modelados, generando resultados que ayuden a la toma de decisiones empresariales.

La Veracidad, que está atada a la confiabilidad de los datos.

¿Por dónde empezar?

Para poder iniciar un proyecto de Big Data lo primero que debemos hacer es identificar el problema que queremos resolver, ejemplos típicos pasan por la identificación de tendencias en las redes sociales, estudio de imágenes satelitales o análisis de comportamiento de sitios web. Hay una infinita gama de situaciones o problemas a analizar que dependerán del negocio u organización.

Una vez identificado el problema se procede a la recolección de los datos desde las fuentes de datos originales y previamente identificadas. En esta etapa es importante hacer un inventario de todas las aplicaciones de la organización, identificando todas las tecnologías involucradas, los sistemas operativos, bases de datos, políticas de seguridad y acceso, confidencialidad de los datos, repositorios digitales o cualquier otra información que sea de interés y necesaria para llevar a buen término esta fase.

La preparación de los datos tiene por objetivo principal obtener datos confiables y de calidad para el éxito de la etapa de análisis, producto de la aplicación de técnicas de limpieza y estandarización de datos. En esta etapa se tiene una idea bastante clara de los volúmenes de datos que debemos almacenar y procesar.

Una vez preparados y almacenados los datos se procede a realizar su análisis a través de paquetes o algoritmos científicos especializados, los cuales servirán de entrada a diferentes procesos y aplicaciones, que permitirán llevar a cabo todas las acciones asociadas a la toma de decisiones.

¿Hacia dónde vamos?



Big Data ha pasado de ser una tendencia a una realidad, es evidente que las tecnologías actuales y futuras generarán volúmenes de datos sin precedentes, datos que de alguna u otra forma estarán allí para ser analizados de manera científica, y así poder determinar tendencias, patrones y comportamientos. Una nueva profesión caracterizada por el conocimiento en Big Data y la Ciencia de Datos está tomando auge a nivel mundial, una profesión que Thomas H. Davenport y D.J. Patil describieron como el trabajo más sexy del siglo 21, el Científico de Datos. ■

