

Big Data: Marketing o Realidad?

Por. Walter Urrea León
CEO

La imparable avalancha de desarrollos tecnológicos irremediabilmente ha convergido en algún punto para dar origen a lo que los alemanes presentaron en 2011 como industria 4.0. Y aunque fue solo hasta el 2013 cuando se hizo oficial por el gobierno alemán mediante un documento redactado por la ACATECH (Academia nacional de ciencia e ingeniería), con el apoyo de el DFKI (Centro de investigación alemán para la inteligencia artificial) y de grandes empresas del país germano en el marco de la Feria tecnológica de Hannover -Messe; La posición tomada fue clara:

Automatización de la fábrica al punto que hoy en día se conoce cómo fabrica inteligente (*Smart Factory*), la cual se apalanca sobre algunas tecnologías disruptivas como lo son: IoT (*internet de las cosas*), la ciberseguridad, La nube (*Cloud Computing*) y Big Data.

Aunque estamos ya finalizando el 2019, los efectos de dicha transformación siguen siendo insospechados, lo que genera preocupación en muchos sectores de la sociedad; expertos de muchas áreas han creado un ambiente perfecto para el debate. Pues cada una de estas tecnologías tiene sus defensores y detractores.

Hablemos de los *Wearables* por ejemplo, que forman parte del IoT (internet de las cosas), son considerados por muchos como accesorios necesarios en un mundo cada vez más interconectado, sin embargo, sus detractores creen que estos dispositivos son invasivos del tiempo y espacio vital de las personas y que ponen en riesgo su seguridad. Y es que la ciberseguridad garantiza no solo la protección de los Sistemas industriales y las infraestructuras críticas sino también la de los ciudadanos de a pie.

Dentro de los defensores acérrimos se encuentran los nativos digitales (los nacidos entre 1995 y 2015). Esto debido a que dichos dispositivos encuadran perfectamente dentro de las características propias de esta generación como por ejemplo el querer recibir la información de forma ágil e inmediata, el manejo de múltiples tareas y procesos paralelos, la educación a través de gráficos y no de textos, el trabajo en línea y el amor por la recompensa inmediata por su desempeño.

Ahora bien, el Big Data. ¿Realmente las empresas en Colombia y el mundo están trabajando ya con esta tecnología? ¿O es algo producto del marketing tecnológico en el que vivimos? Recordemos que se debe hablar de este termino

cuando se cumplen ciertas condiciones o premisas que en sí conforman lo que los expertos llaman las cinco “V” conceptuales del Big Data. Estas son: Volumen, Velocidad, Variedad, Veracidad y Valor del Dato.

Si analizamos únicamente la “V” inicial, la que da más sentido al término, es decir, Volumen - gran dato o gran cantidad de datos. Se supone que es Big Data cuando las herramientas normalizadas de bases de datos actuales (*Oracle, SQL, MySQL, PostgreSQL, etc*) no pueden almacenar un volumen de información, porque esta desborda la capacidad técnica de la herramienta. Aunque debemos tener presente que este concepto es variable y no fijo. Por ejemplo, hace unos años las bases de datos podían almacenar y procesar 1GB de información - por decir cualquier cifra; En la siguiente versión, el fabricante de dicha base de datos ampliaba su capacidad y entonces ya podía almacenar y procesar 2,3,4 o 10GB, etc. Lo que implica que ya no estaríamos hablando de Big Data, pues ninguna empresa hoy en día puede decir que su información no es capaz de ser almacenada y procesada por ningún sistema actual de base de datos o almacenamiento en la nube para considerar otra de las tecnologías mencionadas, la nube (*Cloud Computing*). Este concepto por supuesto es variable según esto.

Ahora, revisemos el concepto Variedad, el más crítico de cumplir. Indica que todos los datos deben ser heterogéneos, no de un solo tipo. Analicémoslo desde el siguiente ejemplo: una empresa tiene una base de datos de 100.000 registros en su contabilidad. Esto no puede ser considerado Big Data, pues a pesar del tamaño y/o cantidad de información, incumple el concepto de variedad. Es decir, no hay más que datos contables allí. La información que recogen las empresas está restringida al uso de esa empresa y no de múltiples compañías. La posibilidad de que una empresa pueda llegar a una gran variedad de datos es prácticamente cero, pues la información que recibe toda empresa se limita a lo que esta recoge para su propósito, uso y razón de ser.

De otro lado, si consideramos el concepto Velocidad tenemos que entre más se requiera, más costos se deben asumir. Y si tenemos un enorme volumen de información, pero no tenemos como procesarla en tiempos óptimos, no sirve de nada. El costo de procesar a gran velocidad grandes volúmenes de datos es difícilmente alcanzable por las empresas debido a la fuerte inversión que demandan los equipos para ello. Ni aun con la reducción de costos que tenemos hoy producto del Cloud Computing.

Veracidad, este concepto es muy importante y difícil de cumplir. En el mejor escenario, supongamos que podríamos contar con un alto volumen de datos de múltiples fuentes y estos fueran de varios tipos. No podemos garantizar que estos sean veraces. Siempre existirán datos que no correspondan a la realidad o verdad dentro de la información. Muchos administradores de bases de datos

son conscientes que la información que procesan tiene en un alto porcentaje errores; Y hoy en día internet ha ayudado a la proliferación de todo tipo de información errónea o falsa lo que limita las posibilidades de alcanzar esta premisa.

Finalmente, Valor del Dato. La información que poseemos pierde valor en la medida que esta quieta. Sino la utilizamos para procesarla con más información, así esta sea muy valiosa, pierde valor en el tiempo. A más información almacenada en estado de reposo, más dinero perdemos.

Así las cosas, difícilmente hoy en día estamos trabajando con BigData, pues para hacerlo se deben cumplir con estas premisas. Junto con lo anterior solo algunos pocos podrían acercarse al término Big Data en la práctica. Tal vez una empresa como Facebook o Google por mencionar actores privados, aunque la veracidad de la información que procesan pone en entredicho esta afirmación pues por sus servidores circulan perfiles falsos y mucha información fake. De otro lado el gobierno de China si tenemos en cuenta a un estado como actor público.

Evidentemente, el marketing y la tecnología tienen una estrecha relación; ambos se benefician del otro y hacen que el ecosistema funcione. El punto está en nosotros, los usuarios y consumidores somos quienes debemos poder discernir entre estas tecnologías que hoy en día nos venden con la promesa de que tendremos a dos manos el cielo tecnológico o simplemente buenas tecnologías que hoy están a nuestro alcance y cumplen con un propósito específico, no creyendo que estamos tocando la cúspide de la pirámide tecnológica como no lo quieren hacer ver. Recordemos que esto es una industria y por supuesto los fabricantes quieren vender sus productos.