

EL BIG DATA EN LAS ORGANIZACIONES



Universidad Militar Nueva Granada
Facultad de Ciencias Económicas

Alejandra Aguilera Galindo

Código: 2204655

Diplomado Estrategia y Transformación Digital with SAP S/4 HANA
Asesores metodológicos: Eduardo Ruíz Cárdenas y Francisco Fernando Ortega

Bogotá, Colombia

Enero 2021.

Introducción

En el presente trabajo se aborda la continua evolución de los datos, que hacen que las empresas busquen nuevos mecanismos de recolección y almacenaje, fomentando el uso del *Big Data*, su aplicación en el mercado y cómo generar un valor diferencial en el mundo organizacional cambiante, independientemente del tamaño de las empresas.

La información y la tecnología han abierto un nuevo escenario en las organizaciones, donde encontramos enormes cantidades de datos que generan un impacto. El Big Data entra a las empresas como una herramienta idónea, para recolectar mejor los datos disponibles y que las organizaciones estén un paso adelante de sus clientes.

Este documento se desarrolla en un inicio explicando el Big Data en las organizaciones y qué ventajas genera, en una segunda parte se explica los tipos de datos que se encuentran y cómo las organizaciones lo utilizan.

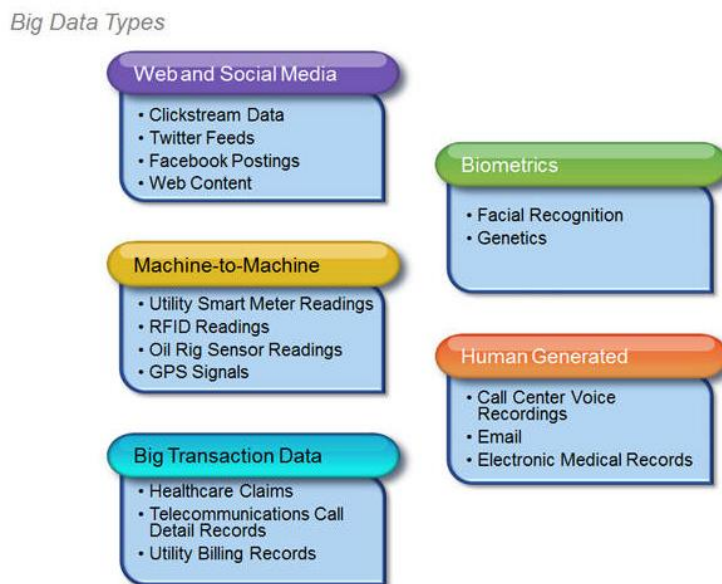
El Big Data en las Organizaciones.

El mundo es cambiante. Cada día se actualiza la tecnología y va generando una evolución donde las empresas deben adaptarse y estar en línea con el cambio y la evolución constante de estas tecnologías. Hoy en día, la empresa que no posee por lo menos una buena base de datos y un buen equipo de tecnología para ofrecer un buen servicio a sus clientes, no es, ni será competitiva en el mercado. Según I (2019) afirma que el internet de las cosas (IoT) supone una total revolución que ha llegado para quedarse y cambiar el mundo. (p.31)

El Big Data es una red donde se almacena y recolecta información por segundos, esta información es recolectada por dispositivos como el celular, el computador, el televisor, entre otros. Estos dispositivos interactúan con el usuario, por medio de las redes sociales, de los correos electrónicos, de las plataformas digitales, y de los sistemas operativos. De acuerdo con

Aguilar (2016) para calcular el volumen de datos que se producen a diario en el mundo es prudente imaginarse la cantidad de personas que acceden a redes sociales como Facebook, Twitter, Instagram, Whatsapp, correo electrónico y demás conversaciones virtuales que pueden realizarse a través del envío y recepción de mensajes de texto, llamadas telefónicas, video llamadas, notas de voz, videos, imágenes y otros (Aguilar 2016, citado por Mariana, E. B., & Margareth Mercado Pérez, 2019, p.19).

Figura 1 *Tipos de Datos de Big Data*



Nota: Fragoso, R. (2012). ¿Qué es Big Data? [Figura]. Recuperado de <https://developer.ibm.com/es/articles/que-es-big-data/>

De acuerdo con la Figura No. 1, los usuarios han pasado de guardar información en carpetas, en libros, en agendas; a guardar todo en un dispositivo electrónico. Estos datos son manejados por el Big Data que posteriormente, permite que empresas a través de las redes sociales vendan esta información a las organizaciones y así tener un acceso más cercano a los usuarios.

Dicho esto, las organizaciones manejan bastante información. Es tanta la información que tienen sobre los usuarios que saben cómo piensan, cómo se comportan, qué gustos tienen, qué hobbies practican, toda la información sobre su familia, empleos, cuentas bancarias, etc. De este modo, toda esta información es administrada estratégicamente.

Importancia de los datos en las organizaciones

Las organizaciones cada día generan datos, algunos son útiles y otros no para la toma de decisiones. Según Marr (2018) existen tres áreas fundamentales en las que los datos son realmente importantes para los negocios: la mejora de la toma de decisiones, la mejora de la actividad y la conversión de los datos en dinero.

En primer lugar, el Big Data permite a las empresas recoger mejores datos sobre el mercado y los clientes. Las empresas están consiguiendo percepciones mucho mejores acerca de lo que quieren los clientes, lo que utilizan (y cómo), la manera en la que compran bienes, y lo que piensan sobre dichos bienes y servicios. Y esta información la pueden utilizar para tomar mejores decisiones en todas las áreas del negocio, desde el diseño de productos y servicios hasta las ventas, el marketing y la asistencia posterior. (p. 8)

El Big Data mejora la productividad de las empresas internamente, no solo porque recolecta información de los clientes y el mercado, sino también información sobre la materia prima, los productos, los almacenes, el inventario, y el tiempo de ensamblaje; además, genera un buen ambiente laboral ya que las empresas llegan a tener una relación de primera con sus trabajadores por medio de la información. Según Marr (2018), el Big Data puede mejorar la eficiencia interna y la actividad de casi todos los tipos de negocio y en muchos departamentos diferentes. Hay empresas que incluso han empezado a utilizar sensores para supervisar los

movimientos de los empleados, su nivel de estrés, la salud e incluso con quién hablan y qué tono de voz utilizan, logrando así mejorar la satisfacción de los empleados y la productividad. (p.8)

La implementación de las nuevas tecnologías permite que las organizaciones mejoren eficientemente y eficazmente. Según Marr (2018) el IoT juega un papel importantísimo a la hora de aumentar el rendimiento operativo; gran parte de este no se centra tanto en los dispositivos inteligentes, sino en sensores. (p.8)

Tipos de Datos

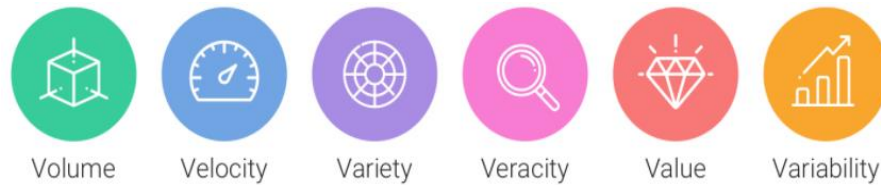
En términos generales, según Fragoso (2012), se define el Big Data como la tendencia tecnológica que ha abierto las puertas hacia un nuevo enfoque de entendimiento y toma de decisiones, la cual es utilizada para describir enormes cantidades de datos (estructurados, no estructurados y semi estructurados) que tomaría demasiado tiempo y sería muy costoso cargarlos a un base de datos relacional para su análisis. (p.11)

Los datos estructurados es la información que se encuentra en las bases de datos. Esta información está organizada por columnas, en filas y con títulos; los datos no estructurados son datos binarios que no tienen estructura interna y no son identificables; los semi estructurados contienen ambas formas de datos (Kyocera,2021).

Las organizaciones deben diferenciar qué tipo de datos están manejando, qué tecnologías deben utilizar para que estos datos sean administrados correctamente y que se afine con los objetivos planteados. Según García, Bustamante, Berlanga, Molina, Padilla y Patricio (2018) actualmente hay un conjunto de herramientas comerciales que llevan a cabo procesos de análisis de datos, como son: Weka, Orange, RapidMiner, Tanagra, KNIME, Oracle Data Mining, IBM SPSS Modeler y SAS Enterprise Miner. (p. 6)

Entonces, la generación de datos estructurados, no estructurados y semi-estructurados son tipos de datos que demanda la aplicación de plataformas tecnológicas que son representadas por el Big Data. (Mariana, E. B., & Margareth Mercado Pérez, 2019, p.28)

Figura 2. *Las grandes V del Big Data*



Nota: Tejedor, B. (2020). Big Data: ¿Qué es y cómo funciona? Recuperado de <https://es.mailjet.com/blog/news/big-data/>

Según Tejedor (2020) el Big Data es una herramienta completa, tal como se evidencia en la figura No 2 la primera V (que es Volumen), hace referencia al manejo de grandes volúmenes de datos que se generan diariamente, sigue la Velocidad con que llegan los datos en tiempo real; la tercera V es la Variedad de datos (los tipos de datos); la Veracidad de los datos representa el hecho de que, son exactos y comprobados a la hora de la toma de decisiones; el Valor depende de las organizaciones, lo cual denota que no todos los datos tienen valor para las empresas; y la última V es la Variabilidad que indica que la información ya está comprobada y lista para utilizarla con distintos fines. El objetivo de las 5 V es analizar los datos de manera visual, rápida, y verídica que permite monitorear y convertir los datos en informes analíticos que es subida a una nube.

Información por segundo

Estamos en la era de la información por segundo. El Big Data genera un problema ético y moral, como la privacidad, la transparencia, la pérdida de identidad, la manipulación emocional, donde hoy en día todo es manejado por influenciadores; por tanto, somos una sociedad que se

deja influenciar. Las organizaciones deben aprender a manejar la información de sus usuarios muy responsablemente, a su vez, deben saber qué temas judiciales existen en la protección de datos, porque si esta información llega a caer en manos no autorizadas puede ser perjudicial. Según Monleón (2015), hoy en día la legislación protege la privacidad de las personas a través de un método de notificación y consentimiento (la empresa informa a los usuarios qué datos se recopilan y cómo se utilizarán posteriormente, y el usuario da su consentimiento).

La cuestión está en que las empresas deben tener principios éticos cuando manejan la información de sus clientes. Linares (2018), expone cuatro principios fundamentales mediante los cuales las organizaciones deben regirse de una ética que evalúe los efectos del poder tecnológico: responsabilidad social, precaución, justicia distributiva y autonomía individual y comunitaria.

Las empresas deben tener en cuenta qué es lo que quieren contar. Si los datos recolectados se afinan con las metas estratégicas, deben definir qué es lo que quieren captar y si es entendible a los ojos de los usuarios; no es recolectar información y que esta no tenga un sentido o no se relacione con los objetivos planteados por la empresa, por eso Martrat (2017), en su modelo del círculo de gestión analógico, señala diferentes elementos a tener en cuenta por las organizaciones donde se centra en 10 pasos. (p.15)

Si se analiza la figura 3, las organizaciones se centran en estos pasos para la recolección de información (El Big Data), esta información es integrada y gestionada en una nube para efectuar el análisis correspondiente y así llegar a la toma de decisiones acertadas para conocer y brindar un buen servicio a los clientes y/o usuarios.

Figura 3 El círculo tradicional o analógico



Nota: Martrat, F. (2017). El círculo tradicional o analógico. En M. Martín Ramo (Ed.), Menos Likes (p. 15). Pearson Educación, S.A.

De acuerdo con Millán (2017), el Big Data en las organizaciones mejora en las áreas de las promociones con campañas mas eficientes y acertadas al público, en el marketing, en las ventas; en materia de compras, ayuda a la mejora de selección de proveedores con productos de alta calidad, en la parte de producción y en recursos humanos, aquí para identificar las características y el comportamiento de sus colaboradores. Los datos son claves para la innovación y el crecimiento. Las organizaciones necesitan un análisis adecuado. (Stubbs,2014, p.10)

Hadoop mapreduce

Una de las tecnologías más relevantes del Big Data es Hadoop. Estos datos son procesados por la arquitectura de la gestión, para generar un análisis de los datos diseñado a partir de las necesidades y estrategias del negocio. Matrat (2017), afirmó que MapReduce se basa en una arquitectura de ordenadores que consiste en un clúster de máquinas clonadas. Es un

modelo de programación y una arquitectura de gestión de la información. La arquitectura particiona las informaciones a procesar en diferentes máquinas, esto permite a los programadores optimizar los recursos de procesos de información. (p.60)

Los procesos de trabajo de Hadoop mapreduce se dividen en tres componentes principales: El driver define la parametrización y el entorno; el segundo es el mapper, que es el responsable de la primera fase, es decir, que se encarga de filtrar y de transformar cada dato original generando un código; por último, la clase reducer se encarga de agregar los resultados producidos por los mappers mediante su clave. (Fernández,2021)

Todos manejamos el Big Data

Desde los inicios se está manejando el tema de generar información, no siempre es necesario tener una base de datos para almacenar dicha información.

Pongamos el ejemplo de una Droguería de barrio, las personas que trabajan en esta droguería tienen información de las personas que viven en el barrio, saben sus gustos, su núcleo familiar, su rutina, sus enfermedades, saben qué medicamentos toman para combatir las enfermedades; toda esta información se almacena en el cerebro y se genera una mini base de datos no visible pero muy funcional.

Figura 4 *Todos Manejamos el Big Data*



Nota: Elaboración Propia Todos Manejamos el Big Data historieta realizada en la pagina https://www.canva.com/design/play?category=tACZCigycaA&referrer=comic-strips-landing-page&utm_source=onboarding

Los niños también están en constante generación y recopilación de datos. Por medio de los programas infantiles, de los programas educativos, de los comerciales televisivos, de otros niños, de sus familiares, entre otros, los niños administran esta información y es manipulada emocional o estratégicamente. Es decir, cuando los niños reciben toda esta información se genera un deseo en ellos de adquirir dichos productos, juguetes, servicios, alimentos o en realizar actividades que ellos vieron y recopilaron en su mini base de datos.

En el día a día estamos generando información nueva y recolectando información que va siendo administrada en nuestra mini base de datos, esta información la poseemos únicamente nosotros y la ponemos en práctica en nuestro vivir.

Según Martra (2017), afirma que estamos en un entorno en el que el cliente no tendrá que buscar las cosas, ellas lo buscarán a él. Las tecnologías de Big Data están en estado incipiente y empiezan a dibujar un entorno en el que serán las empresas las que pondrán productos y servicios a los clientes en función de su comportamiento. Las empresas se esforzarán en detectar lo que uno necesita en cada momento. (p. 24)

Conclusiones

Las organizaciones deben estar alineadas con la nueva era de la tecnología, esto permite que sean competitivas en el mercado, puesto que estamos en un mundo cambiante y necesitan un manejo adecuado de la información para planear sus estrategias y generar valor.

Cada día se genera nueva información, por eso las organizaciones deben implementar el Big Data con buenas herramientas tecnológicas, tener claro qué programas utilizar y qué tipo de datos se está generando para su respectivo análisis y así tomar mejores decisiones.

Las organizaciones deben generar confianza a sus clientes por el manejo de sus datos, para así asegurar clientes fieles y generar nuevos clientes, deben tener en cuenta la parte legal y administrativa.

El Big Data mejora la competitividad de las organizaciones independientemente del tamaño, si se sabe administrar correctamente la información se volverá estratégico, esto lo que permite es una mejora en las áreas de las organizaciones, permite una relación más cercana internamente y externamente.

Una de las tecnologías más relevantes del Big Data es Hadoop, ya que genera una gestión de análisis de los datos que son diseñados a partir de las necesidades y estrategias del negocio, esto permite que las organizaciones en la toma de decisiones sean acertadas y aceptadas por el mercado y los clientes.

Las organizaciones a la hora de manejar los datos deben tener en cuenta qué historia quieren contar a su público, de dónde van a recolectar la información, y si van a utilizar plataformas o si van a hacer encuestas; lo importante es que las organizaciones identifiquen de dónde va salir la información, y que esta información sea analizada y verificada.

La creación de un Big Data se debe realizar dentro de un cluster, esto permite integrar todas las herramientas que se van a utilizar, un ejemplo es el Hadoop.

Todos manejamos el Big Data sea organización o una persona natural, estamos día a día generando datos y recibiendo datos, esta información es administrada y utilizada en nuestro vivir.

En definitiva, el Big Data es la solución de recursos que permiten gestionar, monitorear, analizar, visualizar, generar informes y paneles administrativos fáciles de entender, transformar, integrar y cargar cantidades masivas de datos.

Referencias bibliográficas

- Ana Nieves Millán Reyes. (2017). Big data: El poder de los datos. *Revista General De Información y Documentación*, 27(1), 261-268.
doi:<http://dx.doi.org.ezproxy.umng.edu.co/10.5209/RGID.56583>
- Barranco Fragoso, R. (2012). "¿Qué es **Big Data**?". Recuperado el 19 de Marzo de 2017, de ibm.com: <https://www.ibm.com/developerworks/ssa/local/im/que-es-big-data/>
- Escobar Borja, M., & Mercado Pérez, M. (2019). BIG DATA: UN ANÁLISIS DOCUMENTAL DE SU USO Y APLICACIÓN EN EL CONTEXTO DE LA ERA DIGITAL. (Spanish). *Propiedad Inmaterial*, 28, 273.
- Fernández, O. (02 de enero de 2021). ¿Qué es Hadoop MapReduce?. Aprender BIG DATA. Recuperado de: <https://aprenderbigdata.com/hadoop-mapreduce/>
- Stubbs, E. (2014). *Big data, big innovation : Enabling competitive differentiation through business analytics*. ProQuest Ebook Central <https://search-proquest-com.ezproxy.umng.edu.co>
- García, J, Bustamante, Á, Belanga, A, Molina, J, Padilla, W, Patricio, M (2018). Ciencia de datos. Altaria 1ª Ed. Página: 1. Tomado de <https://www-alfaomegacloud-com.ezproxy.umng.edu.co/reader/ciencia-de-datos?location=20>).
- IBM. (2014). IBM. Obtenido de <http://www.ibm.com/developerworks/ssa/local/im/que-esbig-data/>.
- i, M. (2019). Internet de las cosas. Ediciones de la U. Tomado de <http://www.ebooks7-24.com.ezproxy.umng.edu.co/?il=10048>

- Kyocera. (2017). "Diferencia entre datos estructurados y no estructurados". Recuperado el 25 de Marzo de 2017, de <https://www.kyoceradocumentsolutions.es/es/smarter-workspaces/insights-hub/articles/diferencia-entre-datos-estructurados-y-no-estructurados.html>.
- Linares, J. (2018). Hacia una ética para el mundo tecnológico. *Artefactos*, 7(1), 99-120.
doi:<http://dx.doi.org.ezproxy.umng.edu.co/10.14201/art20187199120>
- Marr, B. (2018). **Data** strategy. Ecoe Ediciones. Página: 1. Tomado de <http://www.ebooks724.com.ezproxy.umng.edu.co/?il=6317&pg=16>).
- Martrat, F. (2017). Menos likes, más leads: Transformación digital en ventas. Pearson Educación. Página: 1. Tomado de <http://www.ebooks724.com.ezproxy.umng.edu.co/?il=6487&pg=65>).
- Mariana, E. B., & Margareth Mercado Pérez. (2019). Big data: Un análisis documental de su uso y aplicación en el contexto de la era digital. *Revista La Propiedad Inmaterial*, 28, 273-293. doi:<http://dx.doi.org.ezproxy.umng.edu.co/10.18601/16571959.n28.10>
- Monleón-Getino, A. (2015). El impacto del big-data en la sociedad de la información. significado y utilidad. *Historia y Comunicación Social*, 20(2), 427-445.
doi:http://dx.doi.org.ezproxy.umng.edu.co/10.5209/rev_HICS.2015.v20.n2.51392
- Tejedor, B. (2020). Big Data: ¿Qué es y cómo funciona? [Figura]. Recuperado de <https://es.mailjet.com/blog/news/big-data/>