

## Perspectivas del Big data

### *Big data insights*

- <sup>1</sup> Jesús Alberto Arcos Cobra  <https://orcid.org/0000-0002-4541-7745>  
Universidad de Guayaquil, Guayaquil, Ecuador  
[jesus.arcosc@ug.edu.ec](mailto:jesus.arcosc@ug.edu.ec)
- <sup>2</sup> Lissette Fernanda Alvarado Barrera  <https://orcid.org/0000-0003-2957-9354>  
Instituto Superior Tecnológico Vicente Rocafuerte, Guayaquil, Ecuador  
[lalvarado@istvr.edu.ec](mailto:lalvarado@istvr.edu.ec)
- <sup>3</sup> Kléver Paul Muñoz Sánchez  <https://orcid.org/0000-0002-8241-0649>  
Investigador Independiente, Guayaquil, Ecuador  
[paul\\_sai@hotmail.com](mailto:paul_sai@hotmail.com)



---

#### Artículo de Investigación Científica y Tecnológica

Enviado: 14/12/2021

Revisado: 29/12/2021

Aceptado: 28/01/2022

Publicado: 05/02/2022

DOI: <https://doi.org/10.33262/ap.v4i1.1.178>

---

#### Cítese:

Arcos Cobra, J. A., Alvarado Barrera, L. F., & Muñoz Sánchez, K. P. (2022). Perspectivas del Big data. AlfaPublicaciones, 4(1.1), 514–531. <https://doi.org/10.33262/ap.v4i1.1.178>



**ALFA PUBLICACIONES**, es una Revista Multidisciplinar, **Trimestral**, que se publicará en soporte electrónico tiene como **misión** contribuir a la formación de profesionales competentes con visión humanística y crítica que sean capaces de exponer sus resultados investigativos y científicos en la misma medida que se promueva mediante su intervención cambios positivos en la sociedad. <https://alfapublicaciones.com>

La revista es editada por la Editorial Ciencia Digital (Editorial de prestigio registrada en la Cámara Ecuatoriana de Libro con No de Afiliación 663) [www.celibro.org.ec](http://www.celibro.org.ec)



Esta revista está protegida bajo una licencia Creative Commons Attribution Non Commercial No Derivatives 4.0 International. Copia de la licencia: <http://creativecommons.org/licenses/by-nc-nd/4.0/>

---

**Palabras****claves:**

Big Data,  
tecnologías,  
datos,  
información.

**Keywords:**

Big Data,  
technologies,  
data,  
information.

**Resumen**

En la actualidad, la humanidad está sumergida en un mundo de constantes cambios y evoluciones, presentada por la globalización y los grandes volúmenes de datos que se generan producto de los medios tecnológicos llámense internet, móviles, redes sociales, entre otros, como consecuencia han aparecido las tecnologías del Big Data, a fin de convertir esos datos en información útil para la sociedad y las organizaciones. Siendo el objetivo de la presente investigación indagar y conocer las perspectivas del Big Data, a través de las posiciones de los diferentes autores, para esto se usa una metodología de tipo documental, basada en el diseño bibliográfico. Esta revisión bibliográfica permitió concluir que a pesar de las barreras que se puedan presentar para la adquisición de las nuevas tecnologías, es más valiosa la información que se está generando y que puede ser depurada con las tecnologías del Big Data, siendo los profesionales más idóneos para manipular y gestionar esas tecnologías y sus herramientas, los ingenieros del área de informática o afines, por lo que deben mantenerse actualizados, así como la incorporación de dichas tecnologías en los currículos educativos de los centros universitario, garantizando profesionales de alta gama tan necesarios en la actualidad.

**Abstract**

At present, humanity is immersed in a world of constant changes and evolutions, presented by globalization and the large volumes of data that are generated as a result of technological means, be it the internet, mobile phones, social networks, among others, as a consequence they have appeared. Big Data technologies, in order to convert this data into useful information for society and organizations. Being the objective of the present investigation to research and know the perspectives of Big Data, through the positions of the different authors, for this a documentary-type methodology is used, based on the bibliographic design. This bibliographic review allowed us to conclude that despite the barriers that may arise for the acquisition of new technologies, the information that is being generated is more valuable and that can be refined with Big Data technologies, being the most suitable professionals to manipulate and manage these technologies and their tools, engineers in the area of computer science or related, so they must be kept up to date, as well as the incorporation of these

---

technologies in the educational curricula of the university centers, guaranteeing high-end professionals so necessary in the present.

---

## Introducción

En su proceso evolutivo, la humanidad siempre ha buscado los medios para solventar su subsistencia, convirtiéndose en los primeros tiempos en cazadores, agricultores, constructores de herramientas que les facilitara sus actividades cotidianas, así como mecanismos que les facilitaran algunos procesos.

Indica Toro & Laniado (2019):

... y apareció la necesidad de cuantificar, inicialmente se utilizaron marcas en piedras y árboles. Otra forma de contar era, almacenar en una bolsa una piedra por cada oveja que tuvieran: cuando mataban una, botaban una piedra de la bolsa; cuando compraban otra, agregaban una piedra a la bolsa.

Con el correr del tiempo han surgido nuevos inventos, siempre en pro de obtener mayores beneficios para la sociedad. Uno de esas innovaciones lo marcó el computador, el cual debido a la gran cantidad de información que se maneja hoy por hoy, ha tenido que evolucionar, no solo en las actualizaciones de software y hardware, sino en capacidades de las personas que se hacen uso de ésta y otras herramientas, Ramos (2019) manifiesta que, la tecnología digital se ha convertido una fuente de trabajo y de ingresos económicos. El mismo autor, opina que en la actualidad hay medios que permiten acceder a la información, pero la gran escala de la misma hace que este campo de la tecnología evolucione al igual que las aplicaciones que en ella se apliquen. Así como la utilización de mayores cantidades de información.

Pero, no solo el computador forma parte de esa gama de herramientas tecnológicas que a diario recibe y procesa grandes volúmenes de información (como los registros de páginas web, correos electrónicos, documentos escaneados), sostiene Vega (2020) que, cada vez más dispositivos, herramientas de fabricación y plantas son equipadas con sensores que pueden recolectar inmensas cantidades de datos sobre sí mismos y su entorno, dispositivos inteligentes como los teléfonos móviles o de equipos inteligentes (internet de las cosas), transacciones bancarias, imágenes satelitales, rutas de GPS, datos de redes sociales y muchos otros.

En ese orden de ideas indica Ramos (2019), que existe una gran cantidad de información que pueden ser representados en móviles, computadores, web, sensores robóticos, automovilismo, bancos, anemómetros, que permiten comunicar y medir el

posicionamiento, movimiento, vibración, temperatura, humedad y hasta los cambios químicos que sufre el aire, todas aplicaciones deben tener gran velocidad para medir.

Para Casallas & Jiménez (2017), los datos provienen de diversas fuentes como transacciones bancarias, pagos con tarjeta e incluso búsquedas en internet, también hay datos que nosotros proporcionamos al publicar algo, subimos imágenes o videos entre otros.

Por todo lo anterior, es importante el tratamiento que reciben los datos, a fin de convertirla en información útil para las organizaciones y la humanidad. Toro & Laniado (2019), sostienen que, “quien posea la información práctica para hacer uso de ellas tiene el mundo a sus pies, por la tanto es importante saber uso de la misma, es decir conocer las herramientas y técnicas para procesarla y hacerla útil para el área que se requiera.

A la luz de esa realidad, las sociedades y las organizaciones también debe pasar por ese proceso de actualización y de crecimiento acelerado, a fin de mantenerse a la vanguardia de los mismos. Esta inmensa cantidad de datos disponibles ha implicado el desarrollo de nuevos métodos científicos e ingenieriles para disponer de sistemas y procedimientos capaces de almacenar, procesar y analizar dichos datos, generando de esta manera información y conocimiento (Vega, 2020).

Es decir, el crecimiento en el volumen de datos generados por diferentes sistemas y actividades cotidianas en la sociedad ha forjado la necesidad de modificar, optimizar y generar métodos y modelos de almacenamiento y tratamiento de datos que suplan las falencias que presentan las bases de datos y los sistemas de gestión de datos tradicionales (Hernández et al., 2017).

Como se mencionó anteriormente, la información es poder, por ello, para Flores & Villacís (2017), uno de los desafíos de la información es el de administrar, buscar, almacenar y analizar datos para apoyar a la toma de decisiones; para lo cual, los métodos tradicionales de procesamiento han tenido que evolucionar, con el fin de proporcionar respuestas en tiempo real y considerando el menor costo posible.

Ese proceso de convertir los grandes volúmenes de datos en información útil, puede ser realizado con las tecnologías del Big Data. Y así lo sustenta Flores & Villacís (2017), este fenómeno conocido como Big Data, permite que todos los datos pueden transformarse en información útil y se puedan utilizar para mejorar la toma de decisiones en cualquier organización, mejorar la eficiencia, reducir costos y aumentar ingresos.

Por las razones anteriores, es que el Big Data se ha convertido en una tendencia a nivel mundial y aunque aún no cuenta con un concepto científico o académico consensuado, se augura cada día mayor crecimiento del mercado que lo envuelve y de las áreas de investigación asociadas (Hernández et al., 2017). Por otra parte, el Big Data no busca

sustituir a los sistemas tradicionales, sino construir una nueva tendencia donde se construyan arquitecturas de sistemas que permitan manejar todas las peticiones (Hernández et al., 2017). En definitiva, el Big Data, pretende servir para mejorar la toma de decisiones, mejorar resultados y desempeño, ahorrar costos o generar políticas públicas que impacten de manera positiva a la sociedad (Vega, 2020).

### **Metodología**

La presente investigación está basada en metodología de tipo documental, con base en el diseño bibliográfico. Afirma Rojas (2013), que en la investigación documental se recurre a las fuentes históricas, monografías, información estadística y a todos aquellos documentos que existen sobre el tema para efectuar el análisis del problema.

Por ello, la investigación inicia con un proceso de selección de libros, artículos de revistas, tesis, entre otros, donde prevaleció como criterio, la pertinencia de los mismos con el objeto de estudio, considerando así, solo aquellos que cumplieran con veracidad de la información en fuentes confiables, permitiendo de esta manera sustentar teóricamente el estudio, al conocer y analizar las diferentes posiciones de los autores en referencia a sus perspectivas sobre el Big Data.

### **Resultados y Discusión**

Al analizar los diferentes autores a lo largo de esta investigación, se pudo constatar que los mismo coinciden desde sus apreciaciones del Big Data, que este permite la gestión, administración y procesamiento de los datos para convertirlos en información útil indispensable para la toma de decisiones de las organizaciones, considerando este hecho se desglosan los aspectos que se presentan a continuación.

#### *Acerca del Big Data*

En este apartado, se presentan algunas aproximaciones conceptuales del Big Data, siendo aproximaciones debido que a la fecha no existe una definición tácita, al respecto afirma Ranney (2019), no existe unanimidad en la definición de Big Data, aunque si un cierto consenso en la fuerza disruptiva que suponen los grandes volúmenes de datos y la necesidad de su captura, almacenamiento y análisis.

Por su parte Casallas & Jiménez (2017), afirman que el Big Data es un término que nació del crecimiento constante de volumen o tamaño de la información. Para Vega (2020), el Big Data se puede definir como el conjunto de datos cuyo tamaño supera considerablemente la capacidad de gestión y análisis del software convencional. Sin embargo, el concepto no hace referencia simplemente al tamaño de la información (volumen), sino también a la variedad del contenido y a la velocidad con que se generan los datos, almacenan y analizan.

El término Big Data, suele aplicarse a conjuntos de datos que superan la capacidad del software habitual para ser capturados, gestionados y procesados en un tiempo razonable y por los medios habituales de procesamiento de la información (Pérez Marqués, 2015 citado por Flores & Villacís, 2017).

Indica Ramos (2019), Big Data son herramientas utilizadas para almacenar grandes volúmenes de datos, ya que con las bases de datos tradicionales no es suficiente, la aplicación de las herramientas es creada en la actualidad para almacenar datos, ya sean en las redes sociales, en el gobierno, en las empresas y demás.

Con base a las afirmaciones de los autores anteriores, se puede discernir que una de las herramientas que permite almacenar, procesar y analizar grandes volúmenes de datos se le conoce actualmente como Big Data (cuya traducción equivaldría a grandes datos o datos masivos o macrodatos), esta herramienta a su vez cuenta con una serie de tecnologías que hacen posible generalizar los accesos de los usuarios para desarrollar las búsquedas de datos minimizando en tiempo de ejecución para poder adaptarse a las demandas de tiempo y costo que la demanda.

#### *Del Byte al Big data*

Hablar de byte como la unidad de información de ocho bits, es como remontarse a la era de las cavernas, sin embargo, la utilización del byte históricamente está marcada en el 1947, pero precisamente debido a la revolución tecnológica, en los actuales momentos, resulta obsoleto hablar de megabyte o gigabyte inclusive, especialmente cuando se trata del Big Data.

Casallas & Jiménez (2017), mencionando algunos autores, afirma que para poder saber lo grande que es el Big Data, es importante empezar por el megabyte, que equivale a 1 millón de bytes y nos permitiría almacenar un libro de unas 500 páginas que tuviera sólo texto. En un terabyte, equivalente a un millón de megabytes, podríamos almacenar 2.767 copias de la Enciclopedia Británica, o 16.667 horas de música, o 1.333 horas de video. Se necesitan casi 5 exabytes, o 5 millones de terabytes, para almacenar las secuencias del genoma humano de todas las personas del mundo.

Así mismo afirma el autor que, la compañía americana IDC estima que en 2011 se crearon 1,8 zettabytes de información, es decir, 1.800 exabytes, una cantidad de datos con la que podrían llenarse 57.500 millones de iPads de 32 gigabytes. Con este número de iPads podría construirse una pared dos veces más alta que la muralla china.

Todo lo anterior muestra, como el desbordamiento de datos a ocurrido con los avances tecnológicos, por lo que resulta imperioso conocer las herramientas y técnicas capaces de convertir esos datos en información útil para la humanidad.

### *Dimensión del Big Data*

Es conocido en el mundo del Big Data como las tres V, terminología acuñada en 2001, por Doug Laney, el cual se refiere a volumen, variedad y velocidad. Posteriormente, se agregaron otras dos características: veracidad y valor (Toro & Laniado, 2019).

Tomando como referente a Flores & Villacís (2017), las cinco V quedarían establecidas como:

*Volumen:* Se debe almacenar y trabajar con una gran cantidad de datos, obligando a que se deban utilizar nuevas infraestructuras, las cuales deben ser escalables y distribuidas.

*Velocidad:* Se refiere a la inmediatez para poder utilizarlos más rápidamente, transfórmalos en información útil en el menor tiempo posible, es así que los tiempos necesarios en ser creados, almacenados, analizados y visualizados se deben reducir para poder acercarse al tiempo real.

*Variedad:* Se refiere a la diversidad de los tipos de datos (estructurados, semiestructurados, no estructurados) y sus diferentes fuentes de obtención.

*Visualización:* Poder visualizar la información de los datos de una forma comprensible es otra de las funciones necesarias del Big Data.

*Valor:* El valor económico de los diferentes datos varía significativamente y el desafío es identificar la información útil, transformarla y extraer los datos para su análisis.

Existen otros autores, como el Instituto de Ingeniería del Conocimiento (IIC, 2019), que asegura la existencia de otras dos V, estas son:

*Veracidad de los datos:* Se refiere al grado de fiabilidad de la información recibida, esto es posible usando tecnologías del Big Data que utilice métodos que puedan eliminar datos imprevisibles, es decir es necesario conseguir datos de calidad.

*Viabilidad:* Se trata de la capacidad que tienen las organizaciones en generar un uso eficaz del gran volumen de datos que manejan.

Es así, como la dimensión de las V, permiten ampliar el concepto del Big Data, al afirmar que no solo es la V de Volumen de información, es además un conjunto de características que si se tienen que considerar ya que van a permitir hacer un mejor uso de los datos generados.

### *Impacto del Big Data*

Es evidente que el crecimiento tecnológico ha impactado todas las áreas del quehacer humano, generando grandes volúmenes de datos útiles e inútiles, por lo que se hace

indispensable identificar de acuerdo al área la información pertinente y esto es posible gracias a las herramientas tecnológicas como el Big Data.

Para Casallas & Jiménez (2017), el Big Data es aplicable a múltiples campos, termina siendo importante en ámbitos como economía, tecnologías o en empresas. En ese sentido, también se apunta Toro & Laniado (2019), quien manifiesta que, el procesamiento de datos masivos tiene aplicaciones en todas las áreas del conocimiento, desde la literatura, la salud y la psicología hasta la biología, las matemáticas y la ingeniería. Todos estos datos pueden entregar información valiosa sobre diversos ámbitos, por ejemplo, gestión y optimización de procesos, fallos de equipos, ciclos de mantenimiento, patrones de compra, tendencias del mercado y otros (Vega, 2020). Las áreas más comunes que utilizan Big Data por su gran cantidad de información son: patrones de detección de fraude, medicina, educación, área financiera, turismo, entretenimiento, entre otros (Flores & Villacís, 2017).

A continuación, se describen algunos ejemplos de ámbitos donde el Big Data ha impactado:

#### *En medicina*

Acá toma singular importancia el Big Data, debido a que en esta área se considera la vida y salud de las personas. Algunos de los beneficios que ofrece acá el Big Data, son según Casallas & Jiménez (2017), información poblacional acerca de enfermedades, historial médico, y estado actual del paciente, permitiendo que se pueda predecir una posible epidemia a tiempo mediante un repositorio que genere alertas tempranas, permite detectar enfermedades a corto plazo y genera una mayor organización en el sistema de salud, generar patrones de comportamiento y afectación respecto a los malestares de la persona.

Adiciona el autor anterior que, también se debe considerar la enorme cantidad de datos necesarios para rayos X, pruebas de laboratorio, imágenes médicas, datos de ensayos clínicos, altas, datos a nivel genético, las secuencias genómicas de datos de población, los cuales deben ser de calidad, debido a que las decisiones que se puedan tomar pueden poner o no en riesgo la vida de una persona. Otro punto en la referente a la salud es la del área de la telemedicina que diagnosticar a alguien que no se encuentra presencialmente en el lugar, el poder monitorizar a un paciente sin que éste se encuentre hospitalizado, el intercambio de información desde diferentes partes del mundo, todas estas posibilidades son hoy posibles y mejores gracias al Big Data.

#### *En Marketing y ventas*

Las redes sociales como Facebook, Instagram, YouTube, Twitter o Google están catalogadas como los servicios más utilizados, generando un gran volumen de datos, que si les sabe utilizar pueden constituir un medio ideal para las áreas de marketing y ventas.

Así lo afirman Toro & Laniado (2019), las redes sociales son un recurso invaluable para quienes trabajan con publicidad y mercadeo, pues utilizarlas puede servir para recolectar información y para lanzar campañas, lo cual es importante para publicitar un producto, propagar un mensaje o mejorar la imagen de una compañía.

Otras oportunidades que tiene el uso de los datos masivos en el marketing son la segmentación automática de mercados y el encontrar correlaciones en los comportamientos de los consumidores que puedan ser utilizados para crear nuevas estrategias de mercadeo, para ello el Marketing y las ventas usan la aplicación de Big Data, ya que por medio de importantes aplicaciones pueden llegar a los clientes averiguando su comportamiento, los productos que desean y quieren comprar por medio de las redes sociales y demás (Ramos, 2019).

#### *En las organizaciones*

Para cualquier organización la información es su materia prima más importante, por lo que le resulta indispensable hacer buen uso de ella, por esa razón la organización le abre las puertas al Big Data. Así conseguimos que por ejemplo restaurantes, tiendas, supermercados, aseguradoras, entre otras, usan el Big data permitiéndoles conocer cuándo y cómo vender o prestar un servicio, saber de las tendencias del producto, entre otros.

Indica Ramos (2019), que en el sector de retail utilizan la red social para idear el stock, las tendencias de lo que más se busca en la web y las predicciones meteorológicas para vender ropa y zapato de temporada. Por supuesto, esa información se procesa desde el Big Data. Igualmente, los procesos de recursos humanos están siendo mejorados gracias al análisis del Big Data, desde la detección y adquisición de talento, hasta la medición de la cultura empresarial.

#### *En las máquinas y equipos móviles*

El análisis del Big Data está ayudando a máquinas y dispositivos a ser más inteligentes y autónomos (Ramos, 2019). Según el autor, existen sensores demográficos que se utilizan como tecnología del Big Data, los cuales, permiten mejorar el área de rutas de reparto donde se puede ver en tiempo real los transporte con las mercaderías, otro ejemplo son los vehículos auto pilotados, los cuales usan cámaras, GPS, conexión a internet, y un abanico de computadoras y sensores que permiten al vehículo circular de forma segura por la vía pública sin necesidad de intervención humana.

#### *En la seguridad*

Esta es un área muy importante para la ciudadanía en general, así como para los gobernantes. Por lo tanto, los datos deben ser tratados de manera muy cuidadosa en

especial para evitar posibles espionajes, ciberataques, robos de identidades o sencillamente violar la privacidad del ciudadano. Esto es posible de evitar haciendo uso de tecnologías Big Data, para Ramos (2019), el sistema de Inteligencia Artificial creado en el MIT predice el 85% de los ciberataques; otro ejemplo en el que lleva años empleándose es en la detección de transacciones fraudulentas con tarjetas de crédito.

Otro aspecto de la seguridad que está saliendo fortalecido con el Big Data, son las fuerzas policiales, ya que la información les permite encontrar criminales e inclusive prevenir actividades criminales.

#### *En ciudades Inteligentes*

A pesar de que pareciera muy lejana la existencia de ciudades inteligentes (Smart Cities), ya están acá o van en proceso de convertirse, siendo esto posible con las tecnologías del Big Data. Este tipo de ciudades existen actualmente en China, específicamente en Shenzhen, donde se tienen miles de dispositivos conectados para coleccionar información, administrar remotamente dispositivos y lograr ubicar problemas de seguridad, todo desde el centro de comando de seguridad de la misma ciudad (Galindo, 2020).

#### *Trading Financiero*

Uno de los rasgos de las transacciones de los mercados de capitales es la compra venta de activos, para lo cual es necesario contar con toda la información en tiempo real para la toma rápida de decisiones, por lo cual esta área también hace un uso extensivo del Big Data. Según Ramos (2019), existen una serie de algoritmos para realizar decisiones de compra venta de valores por millones en fracciones de segundo, además de las señales tradicionales que tienen en cuenta los traders humanos como son los análisis técnicos, comportamientos de materias primas, resultados de empresas, sectores, índices, ahora se le añaden noticias en tiempo real, mensajes de redes sociales, foros, declaraciones públicas de personalidades, entre otras informaciones, es decir, se añade un nuevo tipo de datos que anterior al Big Data eran imposible de manejar.

#### *En la educación e Investigación*

La educación tiene como base fundamenta el uso de información pertinente y actualizada para garantizar profesionales de envergadura que se adapte a las necesidades de la sociedad y de las organizaciones. En ese mismo sentido la información alimenta la investigación y la ciencia. Por ello resulta indispensable el buen manejo de dicha información para la evolución y crecimiento de la humanidad.

Tal y como indica Vidal et al. (2019), la investigación en educación se puede beneficiar enormemente del Big Data y la actual revolución informática más aun ahora que la

disponibilidad de recursos en línea relacionados a tecnología educativa amplía las oportunidades de investigación en el aprendizaje.

Por lo anterior resulta imperioso que los centros educativos universitarios destinen dentro de sus enseñanzas las tecnologías del Big Data, y decir como indica Vidal et al. (2019) se deben considerar estos recursos para el proceso de enseñanza - aprendizaje para el desarrollo de competencias de Big Data.

Es así, como actualmente ha surgido la ciencia de los datos que se encarga de validar, comprender y analizar los datos una vez convertidos en información. Afirma Vega (2020), la ciencia de datos es hoy en día una herramienta fundamental para la generación de conocimiento. Indica el autor que entre los objetivos que persigue esta nueva área, se encuentra la búsqueda de modelos que describan patrones y comportamientos a partir de los datos con el fin de comprender los sistemas o hacer predicciones.

Como se indicó anteriormente, estos son solo algunos ejemplos del impacto que actualmente tiene el Big Data en algunas áreas, sin menospreciar otras donde igualmente se está haciendo uso de las misma. Todo esto permite validar la importancia que tiene la utilización de las tecnologías del Big Data en el mundo.

#### *Barreras*

A pesar de todas las ventajas en las diferentes áreas, que representan las tecnologías del Big Data, como se ha evidenciado, se están presentado una serie de barreras o limitaciones para su implantación total.

Los sistemas de computación tradicional, software y hardware, presentan problemas para el trabajar con Big Data que se caracteriza por su volumen de datos, variedad de datos, y velocidad de cambio de datos (Vidal et al. ,2019).

Casallas & Jiménez (2017), divide estas barreras como se describen a continuación:

*Barrera administrativa.* Referida a que tiene que existir un acuerdo entre todas las partes involucradas para llevar a cabo la distribución de información.

*Barrera tecnológica.* Referida a las mejoras en las tecnologías, haciéndola más acorde a las grandes cantidades de datos que deben ser procesado, y que por más que hoy en día haya grandes avances tecnológicos aún falta mejorar en lo que se refiere a Big Data.

*Barrera legal.* Referida a evitar actos corruptos o los riesgos a la privacidad de las personas ya que son factibles por el volumen de datos generados.

Asimismo, existen otras limitaciones o barreras como son, por ejemplo: los escasos de expertos en el área del Big Data; el costo que involucra implementación y actualización

de hardware, software y personal; la incapacidad de hacer que Big Data sea utilizable por usuarios finales y la ausencia de un modelo de negocio convincente y rentable para las organizaciones que hagan uso del mismo.

### *Tecnologías Big data*

Penetrar en el mundo de las tecnologías Big Data, es profundizar en un abanico de alternativas de software que permiten gestionarlas, es hablar de software libre y software propietario, datos estructurados, semiestructurados o no estructurados, en definitiva, es adentrarse en plataformas que involucran conocimientos y destrezas.

Hernández et al. (2017), las definen como tecnologías de Big Data que dan soporte a la captura, transformación, procesamiento y análisis de los datos, ya sean estructurados, semiestructurados o no estructurados.

Big Data posee una gran selección de herramientas como software libre y open source, propietarias y en la nube, las cuales benefician a organizaciones con gran cantidad de datos que requieren de consultas más rápidas y ágiles (Flores & Villacís, 2017).

Según Flores & Villacís (2017) y Hernández et al. (2017), en las tablas 2, 3 y 4 que se presentan a continuación se resumen algunas de las tecnologías del Big Data más utilizadas.

**Tabla 2**  
*Tecnologías Big Data en software libre*

Tecnología	Características
Hadoop	Gestiona una gran cantidad de datos en petabytes, lo que hace posible que el funcionamiento de sus aplicaciones sea el adecuado. Detecta los fallos y vuelve a ejecutar la instrucción, probando con otros caminos de nodos, para que los resultados que se obtienen no sean inconsistentes.
MapReduce	se refiere principalmente a dos procesos independientes, el primero es el map que toma los datos y los convierte en datos agrupados, cada uno de ellos son separados en tuplas; el segundo proceso es reduce que con la salida del map procesa los datos y combina las tuplas en conjuntos más pequeños.
Cassandra	Es un sistema de almacenamiento distribuido que brinda escalabilidad y alta disponibilidad sin comprometer el rendimiento. Algunos de los principales atributos son: Tolerancia a fallos, por medio de la replicación automática de los datos en múltiples nodos; Descentralización, uso de muchos nodos idénticos, sin cuellos de botella en la red; Durable, diseñada para evitar la pérdida de datos; Elasticidad, capacidad de añadir nuevas máquinas para aumentar el rendimiento de lectura y escritura.
Mahout	Tiene como objetivo ofrecer un ambiente para la creación rápida de aplicaciones de aprendizaje máquina escalables y eficientes. El machine learning o aprendizaje máquina es el trasfondo principal de Mahout y corresponde a un subcampo de la

	inteligencia artificial que se centra en el mejoramiento de procesamientos computacionales a partir del análisis de experiencias previas.
Hbase	Algunas de sus principales características son: escalabilidad modular y linear, estricta consistencia de lectura y escritura, facilidad de uso de la app de java para el acceso de clientes, bloqueo de la caché para consultas en tiempo real, soporte de para exportar métricas. Se recomienda su uso cuando se necesita acceso a lectura y escritura de datos en tiempo real sobre Big Data.
Apache Spark	Es una plataforma que está diseñada para mejorar la velocidad y el rendimiento de las aplicaciones Big Data, brindando rapidez a las consultas y el tratamiento de grandes paquetes de datos, esto lo hace mediante la repartición de los procesos a través de la memoria de distintas máquinas.
MongoDB	Es una base de datos NoSQL orientada a documentos. Tiene como el objetivo de brindar un servicio escalable, con gran capacidad de rendimiento y gran disponibilidad, permitiendo que las organizaciones de pequeño, mediano y gran tamaño puedan disminuir precios, permitiendo aumentar el tiempo de comercialización y reducir riesgos.

Fuente: Flores & Villacís (2017) y Hernández et al. (2017)

**Tabla 3**  
*Tecnologías Big Data en software propietario*

Tecnología	Características
Google Cloud Platform	Utiliza la infraestructura base, el análisis de datos y el sistema de aprendizaje automático de Google, para encontrar respuestas con mayor rapidez y crear mejores productos.
Oracle Big Data Appliance	Es una solución de infraestructura que sirve a empresas que poseen grandes proyectos de datos, así como a organizaciones que aumentan sus necesidades a lo largo del tiempo.
IBM Watson Analytics	Ofrece una variedad de soluciones Big Data y analítica para impulsar un negocio, permitiendo a las organizaciones trabajar con datos para responder a las preguntas de negocio más complicadas, descubrir patrones y poner en práctica ideas innovadoras.

Fuente: Flores & Villacís (2017) y Hernández et al. (2017)

En lo que respecta a las herramientas en la nube, estas permiten manejar grandes cantidades de datos mediante el Internet, sin preocuparnos por su capacidad de almacenamiento (Flores & Villacís, 2017). En la tabla. 4, se describirán las características más relevantes que tiene Big Data en cuanto a este tipo de herramientas.

**Tabla 4**
*Tecnologías Big Data en software propietario*

Tecnología	Características
Hadoop on-cloud	Está formado por varios componentes, que permiten aprovechar las características de Big Data en clúster. No requieren un costo de inversión inicial elevada, además puede utilizar máquinas virtuales alquiladas las cuales podrán ser pagadas por medio del tiempo en que estas sean utilizadas.
Microsoft Azure	Es una plataforma que se encuentra en la nube de Microsoft, tiene la ventaja de nube híbrida, es decir, se puede tener todos los centros de datos y la nube pública al mismo tiempo, de esta manera, actúa de forma inmediata. Además, da la accesibilidad de crear servicios y aplicaciones con cualquier dispositivo
Amazon Web Services (AWS) 1	Ofrece potencias de cómputo, almacenamiento de base de datos, entrega de contenido permitiendo a las empresas a escalar y crecer. Es que independiente del lenguaje y del sistema operativo.

Fuente: Flores & Villacís (2017) y Hernández et al. (2017)

*Ingenieros del Big Data*

Debido a que los grandes volúmenes de información se siguen y seguirán generando, por todas las razones anteriormente expresadas, es evidente que este fenómeno crea un abanico de oportunidades para los ingenieros, ya sean informáticos, programadores, de sistemas, de datos, computación o carreras afines, llegando inclusive a la denominación de ingenieros Big Data, es decir la clave para hacer útil todos esos grandes volúmenes de datos para la organización a través de las tecnologías del Big Data es los actuales momento es el Ingeniero.

Para ello es fundamental la preparación y actualización de los ingenieros en el tema del Big Data, indica Mallada (2019), es necesario que los programas de las universidades sigan vivos y se vayan actualizando, sin dejar de lado el aspecto teórico, sobre todo la parte más técnica en la que es necesario ampliar aspectos como la programación o un mayor enfoque práctico sobre sistemas Big Data.

Los ingenieros que asuman la decisión de enfocarse a ser ingeniero Big Data, según Fernández (2022), se encargan de diseñar, construir y gestionar los datos y la infraestructura necesaria para almacenarlos y procesarlos, debe construir la base tecnológica para el análisis de la información, es decir son los responsables de mantener sistemas escalables, con alta disponibilidad y rendimiento, integrando nuevas tecnologías y desarrollando el software necesario.

Conforme a todo lo mencionado, el Ingeniero Big Data tiene su futuro como tecnólogo garantizado, pero para ello, debe estar en constante actualización, porque cada día surgen nuevas tecnologías para el manejo y transformación de los datos necesarios para las

organizaciones y para la sociedad. Igualmente, nuestras universidades deben considerar las actualizaciones de sus pensum de estudios en el área del Big Data.

También es importante señalar que ya existe la ciencia de los datos que se trata de un campo interdisciplinar que combina temas como minería de datos, modelamiento matemático y estadístico, programación e inteligencia artificial. Utiliza técnicas tales como lógica difusa, redes neuronales, machine learning y deep learning, entre otras, mediante el uso de tecnología de gran capacidad (Vega, 2020).

Por todo lo mostrado resulta muy importante para el ingeniero que desea ser parte del Big Data conocer y dominar dichas tecnologías y técnicas.

### **Conclusiones**

¡La información es poder!, esta frase acuñada a Hobbes (filósofo del siglo XVI), aún se mantiene vigente, sin embargo, se debe tener presente que en los actuales momentos esa información debe ser tratada y analizada, a fin de que sea información útil para las organizaciones y la sociedad, y esto es posible con las tecnologías del Big Data.

Razón por la cual, la presente investigación, toma importancia y por ello, una vez revisada la bibliográfica y analizados los diferentes autores, se pudo evidenciar que existe coincidencias en sus aserciones con respecto al significado de la información útil tratada con Big Data, permitiendo así llegar a las siguientes conclusiones:

La generación de grandes volúmenes de datos generados por los diferentes medios, ha demandado de grandes soluciones, y así lo manifiesta Hernández et al, la cantidad de datos ha ido en aumento, lo cual exige que también las técnicas de análisis y tratamiento de datos se hagan más competitivas, el reto no es solo para recoger y gestionar el gran volumen y diferentes tipos de datos, sino también para extraer valor significativo de estos.

Siendo las tecnologías del Big Data la más apropiada para ello, ya que ofrece una gama de software y plataformas tanto libres como propietarias, que permiten adaptarse a cualquier situación y demanda de las organizaciones.

El mayor desafío que existe actualmente es el crear un sistema donde se puede almacenar, buscar, ordenar en tiempo real toda la información que se recibe por medio de celulares, computadores, redes sociales, web para recibir información en tiempo real facilitando la búsqueda de información de datos, a esto se le domina “Big Data” para utilizar a favor de las empresas y demás negocios que no solo facilita los datos, sino también se convierte en una aplicación más económica, más ingresos y menos costo (Flores & Villacís, 2017).

El ingeniero de las diferentes áreas afines de la informática o ingeniero Big Data será la persona más idónea para a través de las tecnologías del Big Data, aprovechar al máximo los datos generados a fin de contribuir con las mejoras sociales y económicas de

organizaciones. Por lo cual, ese ingeniero debe mantenerse a la vanguardia de las últimas actualizaciones en cuanto a software y hardware se refiere.

En definitiva, estamos entrando en una nueva era de la ingeniería impulsada por los datos y hay que estar preparados (Vega, 2020). La función de la persona que gestiona los datos en Big Data tiene la responsabilidad de que estos no sean datos muertos, es decir, que los datos se conviertan en información relevante para la empresa u organización (Flores & Villacís, 2017; Galindo, 2020).

Se hace necesario que los centros educativos también actualicen sus currículos académicos, a fin de garantizar el personal capacitado para dar respuesta a la gran demanda de profesionales necesarios para gestionar y analizar los grandes volúmenes de información que se generan en la actualidad.

Afirma Vidal et al. (2019), que la inclusión de tópicos de Big Data es necesaria para la formación de las nuevas generaciones de ingenieros de las áreas de informática y computación debido a que actualmente la demanda de profesionales especialistas en dicho tópico es cada vez mayor, y el dominio de estos tópicos son absolutamente necesarios para el desarrollo de soluciones a problemáticas actuales de Big Data.

A pesar de las barreras administrativas, tecnológicas o legales, así como a la escasez de expertos en el área del Big Data, el costo que involucra implementación y actualización de hardware, software y personal, entre otras, a la larga son más los beneficios en cuanto a ganancias y crecimiento socio económico que generan el buen uso de los datos para las organizaciones y la sociedad.

### ***Referencias Bibliográficas***

- Casallas, A., & Jiménez, D. (2017). BIG DATA aplicado a la medicina. *Revista Avenir*, 1 (2), 25-28.
- Fernández, O. (2022). *Ingeniero de Datos: Lo que necesitas saber*.
- Flores, P., & Villacís, A. (2017). *Análisis comparativo de las herramientas de BIG DATA en la facultad de ingeniería de la Pontificia Universidad Católica del Ecuador*. Tesis de Grado, Pontificia Universidad Católica de Ecuador, Ingeniería de Sistemas y Computación, Quito.
- Galindo, J. (2020). *Impacto en la sociedad con la implementación masiva de tecnologías móviles de quinta generación 5G. Ensayo de Alta Gerencia*. Bogotá, Colombia.
- Hernández, E., Duque, N., & Moreno, J. (2017). *Big Data: una exploración de investigaciones, tecnologías y casos de aplicación*. Tecnológicas, 20 (39).

- Instituto de Ingeniería del Conocimiento [IIC]. (2019). Las 7 V del Big data: *Características más importantes*.
- Mallada, D. (2019). *Los Ingenieros de Telecomunicación en la Industria 4.0. Bit (211)*, 64-65.
- Ramos, I. (2019). *Análisis comparativo de herramientas de Big Data para la Carrera de Ingeniería en Sistemas Computacionales*. Tesis de Grado, Universidad Estatal del Sur de Manabí, Ingeniería en Sistemas Computacionales.
- Ranney, C. (2019). *Big Data. Análisis de grandes volúmenes de datos en Organizaciones. Comillas*. Universidad Pontificia, Facultad de Ciencias Económicas y Empresariales, Madrid.
- Rojas, R. (2013). *Guía para la realización de investigaciones sociales*. España: Plaza y Valdez.
- Toro, M., & Laniado, H. (2019). Big data: historia, definición, herramientas y aplicaciones en la industria. *Revista Virtual Pro (204)*, 1-5.
- Vega, J. (2020). Datos, Ciencia e Ingeniería. *Ingeniare. Revista chilena de ingeniería*, 28 (1), 2-5.
- Vidal, C., Madariaga, E., Rubio, J., & Urzúa, L. (2019). *Estudio de la Realidad y Viabilidad de la Formación en Big Data en la Academia Chilena*. *Información tecnológica*, 30 (5), 239-248.

El artículo que se publica es de exclusiva responsabilidad de los autores y no necesariamente reflejan el pensamiento de la **Revista Alfa Publicaciones**.



El artículo queda en propiedad de la revista y, por tanto, su publicación parcial y/o total en otro medio tiene que ser autorizado por el director de la **Revista Alfa Publicaciones**.



#### Indexaciones

