Texto

Descripción generada automáticamente con confianza media

UNIVERSIDAD UNIMINUTO

FACULTADO DE INGENIERIA

ESPECIALIZACIÓN TECNOLÓGICA EN HERRAMIENTAS PARA LA

GESTIÓN DE DATOS BIG DATA

**ESTADO DEL ARTE DEL BIGDATA**

Integrantes

GUSTAVO ANDRES DIAZ PRIETO (000792780)

Bogotá. 2023

Contenido

[1. INTRODUCCIÓN 3](#_Toc128660066)

[2. BIG DATA 3](#_Toc128660067)

[2.1 Historia del Big Data 3](#_Toc128660068)

[2.2 Impacto del Big Data en el mundo 3](#_Toc128660069)

[2.3 Tecnologías de análisis de datos 3](#_Toc128660070)

[2.4 Herramientas para la gestión de datos 3](#_Toc128660071)

[2.5 Beneficios del Big Data 3](#_Toc128660072)

[3. VISUALIZACIÓN DE DATOS 3](#_Toc128660073)

[3.1 PowerBi 3](#_Toc128660074)

[3.2 DataStudio 3](#_Toc128660075)

[4. CONCLUSIONES Y RETOS PARA EL BIG DATA 3](#_Toc128660076)

1. INTRODUCCIÓN

La capacidad actual para producir información (**datos**) ha crecido exponencialmente respecto a

años anteriores. La enorme cantidad de datos a disposición hace necesario el desarrollo de herramientas que permitan el procesado y análisis de estos, para identificar y extraer la información relevante; hoy en día las compañías no saben qué hacer con el gran volumen de datos e información almacenada en diferentes medios o bases de datos, los cuales pueden ser de gran importancia, principalmente en la toma de decisiones.

La eficiente captura y análisis de los datos redundará en la competitividad de las empresas, desde la fabricación, cadena de suministro, hasta los sistemas de gestión, procesos que pueden ser mejorados a través del uso de Big Data y entrar en una era de interoperabilidad entre entidades que puedan suministrar información de calidad que aporten para el crecimiento en todos los sectores.

Para dar sentido al término Big Data y que la industria disponga de mecanismos para la toma de mejores decisiones, conocer mejor su negocio, generar posibles oportunidades de negocio y verificar o refutar teorías y modelos existentes, es necesario incorporar dos términos más, necesarios, para el desarrollo de Big Data:

1. Big Data Analytics: Ciencia de examinar datos en bruto con el propósito de sacar conclusiones sobre esa información. Implica aplicar un proceso algorítmico o mecánico para obtener conocimiento. Por ejemplo, aplicar un proceso para buscar correlaciones significativas entre varias series de datos.
2. Cloud Computing: Modelo para permitir el acceso conveniente y bajo demanda a un conjunto compartido de recursos computacionales configurables (por ejemplo, redes, servidores, almacenamiento, aplicaciones y servicios), que se pueden aprovisionar y liberar rápidamente con un esfuerzo mínimo de gestión o una interacción entre el proveedor de servicios [[1]](#footnote-1)
3. BIG DATA

Cuando hablamos de Big Data nos referimos a conjuntos de datos o combinaciones de conjuntos de datos cuyo tamaño (volumen), complejidad (variabilidad) y velocidad de crecimiento (velocidad) dificultan su captura, gestión, procesamiento o análisis mediante tecnologías y herramientas convencionales, tales como bases de datos relacionales y estadísticas convencionales o paquetes de visualización, dentro del tiempo necesario para que sean útiles. Por esta razón se describen a continuación una serie de aspectos que se deben tener en cuenta para conocer lo que es Big Data.

* 1. Historia del Big Data
  2. Impacto del Big Data en el mundo
  3. Tecnologías de análisis de datos

Dada la magnitud y complejidad de los desarrollos, métodos, técnicas, etc. que actualmente se incluyen bajo la denominación Big Data, dividiremos las tendencias en Tecnologías de análisis de datos y las Herramientas para la gestión de datos a Gran Escala:

* **Microsoft Power BI**: Líder de 2020 como plataforma de analítica de datos e inteligencia de negocio. Gracias a sus dos versiones Power BI Desktop (gratuita) y, Power BI Pro (precio por usuario), permiten al usuario crear potentes cuadros de mando (Dashboard) de forma muy sencilla e intuitiva.
* **Tableau**: Se trata de una plataforma de análisis extremo a extremo (end to end), ya que permite preparar los datos antes de su posterior procesado con la herramienta Prep. También incorpora un sistema de arrarstrar y soltar en la hoja de trabajo (Tableau Desktop) para generar potentes gráficas dinámicas, sobre las que podemos resaltar datos en concreto, hacer zoom, etc
* **Qlik:** Qlik View y Qlik Sense, siendo esta última la más moderna y que, de hecho, supone la transición de View a Sense. Qlik Sense, se utiliza como un gestor de transformación, una plataforma que utiliza un novedoso motor asociativo de análisis (incorporando Inteligencia Artificial) para sugerir conocimientos y, automatizar procesos.
* **ThoughtSpot:** Permite realizar operaciones de computación paralela, es posible procesar 100 billones de filas en 100 segundos. Aprende sobre las analíticas que realiza el usuario para ofrecer insights, ocultos para mejorar la toma de decisiones a través de cuadros de mando. Los precios por uso de la aplicación están sujetos por la capacidad de procesamiento de datos.
* **MicroStrategy:** Como plataforma de análisis de datos, proporciona todas las funciones que necesita cualquier compañía para desarrollar aplicaciones que permitan a través del dato, conocer la información de la empresa y acelerar los negocios. Además de la capacidad de generar informes personalizados o, cuadros de mando, también incluye aplicaciones de aprendizaje automático (Machine Learning), Big Data (conexión nativa con Hadoop), Cloud computing, etc
* **Looker:** Enfocada para extraer el máximo valor de los datos, Looker ofrece dentro de sus soluciones herramientas BI para realizar análisis modernos, realizando cuadros de mando en tiempo real, fácil acceso a los datos, flujos de trabajo (workflow) basados en los datos, es capaz de conectarse con múltiples fuentes de datos como Redshift, Snowflake, BigQuery y, otras 50 herramientas SQL
* **Tibco Spotfire:** Esta herramienta, incluye IA basada en el manejo de los datos, para incluir funciones de analítica avanzada, como, por ejemplo, sistemas de recomendación en búsqueda. Incluye módulos para realizar: Location Analytics, análisis en tiempo real a través de Data Streams, Data Wranling o, analítica productiva utilizando funciones de Python, R, SAS y Matlab para realizar ciencia de datos. Su precio por licencia es de 1.300$ por año y usuario[[2]](#footnote-2)
  1. Herramientas para la gestión de datos
  2. Beneficios del Big Data

1. VISUALIZACIÓN DE DATOS
   1. PowerBi
   2. DataStudio
2. CONCLUSIONES Y RETOS PARA EL BIG DATA

* La transformación digital está impulsando la tecnología en todas las direcciones, a medida que profundizamos en el internet de las cosas, el aprendizaje automático y el Big Data. En 2020 existían 9700 millones de dispositivos conectados a internet; en 2030 se espera que esa cifra alcance los 29.400 millones. La inmensa cantidad de datos generados por ellos hará que el Big Data tenga un papel absolutamente esencial. Dentro de los mas relevante para el uso de Big Data tenemos:
  + Business Intelligence
  + Tecnología de nube
  + Los datos como servicio (DaaS)
  + Tecnología y bienestar (sector salud)
  + Vehículos sin conductor
  + Big data para ayudar en la investigación del cambio climático
  + El análisis en tiempo real gana tracción tras la pandemia
  + Más procesamiento del lenguaje natural (PLN)
  + Automatización del análisis del big data[[3]](#footnote-3)

1. https://www.atiga.es/web/wp-content/uploads/2017/03/Estado-del-Arte-Big-Data-CC-DA.pdf [↑](#footnote-ref-1)
2. https://blogs.imf-formacion.com/blog/tecnologia/las-10-mejores-herramientas-de-analisis-de-datos-202006/ [↑](#footnote-ref-2)
3. https://www.cyberclick.es/numerical-blog/tendencias-en-data-y-analytics-para-2023 [↑](#footnote-ref-3)