



CONCEPTOS FUNDAMENTALES DEL BIG DATA

AUTOR: GASTÓN ADDATI





CONTENIDO

INTRODUCCIÓN.....	3
1. CONCEPTOS FUNDAMENTALES DEL BIG DATA	4
2. UNIDADES DE MEDIDA DE LA INFORMACIÓN.....	9
3. ¿QUÉ ES EL BIG DATA?.....	11
4. CARACTERÍSTICAS DEL BIG DATA.....	13
5. DIFERENCIAS ENTRE BIG DATA Y BUSINESS INTELLIGENCE	15
5.1. ¿Qué es la Inteligencia de Negocios o Business Intelligence?.....	16
6. LA IMPORTANCIA DE LA INTELIGENCIA ARTIFICIAL EN EL BIG DATA	20
7. EMPRESAS QUE UTILIZAN BIG DATA CON INTELIGENCIA ARTIFICIAL	23
 BIBLIOGRAFÍA.....	 25
Referencias	25
Sobre las imágenes.....	25
Notas del autor	25

INTRODUCCIÓN

Los **macrodatos**, también llamados *datos masivos*, inteligencia de datos, datos a gran escala o **big data** (terminología en idioma inglés utilizada comúnmente) es un término que hace referencia a conjuntos de datos tan grandes y complejos que precisan de aplicaciones informáticas no tradicionales de procesamiento de datos para tratarlos adecuadamente.

En el presente libro, profundizaremos acerca de la noción de big data, analizaremos sus principales características, qué es lo que lo diferencia del *Business Intelligence* y, por último, ahondaremos en sus vínculos con la inteligencia artificial.





01

CONCEPTOS FUNDAMENTALES DEL BIG DATA

El crecimiento exponencial de los datos en el mundo, que se ha puesto de manifiesto sobre todo a partir del surgimiento de internet, ha dado lugar a la necesidad de tener que contar con una tecnología que sea capaz de tratar con semejantes volúmenes de datos. Es por eso que comenzó a tratarse y a establecerse el término de big data. Algunos podrán pensar que el surgimiento de las redes sociales, y la mensajería instantánea han contribuido como principales actores a ese crecimiento de información. No obstante, si bien esto es cierto, no es el único; porque miles de millones de usuarios en el mundo usan diversas redes sociales, incluso más de una de manera simultánea, y generan mucha actividad, producen datos, comparten información, y claramente es el principal motivo del crecimiento mencionado, pero no es el único.

Tenemos que mencionar que existen dispositivos que transmiten información tomando ciertas variables de un entorno particular. Si consideramos, por ejemplo, que, mediante diversos sensores de humedad, un especialista en temas de suelos, o un especialista en temas relacionados al agro, podría estar recopilando datos sobre



CONCEPTO

El IoT o Internet de las cosas incluye cualquier objeto –o “cosa”– que pueda conectarse a una red de internet, desde el equipamiento y transporte interior de una fábrica hasta dispositivos móviles y relojes inteligentes.



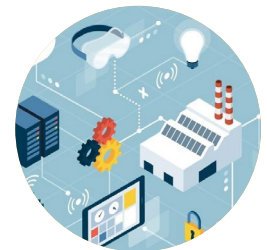
la humedad de un cierto campo previo a realizar una siembra, es muy probable que allí también se utilice el big data. Los datos recopilados por estos sistemas denominados “sistemas de de IoT” (*Internet of Things* – Internet de las cosas) son almacenados y luego procesados.

Al mismo tiempo, tecnologías de RFID también están siendo utilizadas en tiendas donde se vende ropa para autogestionar el stock, automatizar pedidos y obtener datos en tiempo real sobre las ventas realizadas.

Otros dispositivos médicos recolectan información clínica de pacientes, como la cantidad de oxígeno en la sangre, pulsaciones por minuto, forma de la respiración, entre otros. Envían esa información a un repositorio, donde posteriormente se analiza, se procesa y se utiliza para tomar decisiones cruciales de vida o muerte.

En el caso de las ciudades inteligentes, tenemos muchos dispositivos, sensores y elementos que generan información de manera permanente. Es una mezcla de sensores de todo tipo y datos que provienen de diferentes fuentes. Datos que se agruparán para un posterior tratamiento y análisis.

Y así, podríamos seguir dando ejemplos, sobre todo con la tecnología celular, con los dispositivos inteligentes que además de estar conectados a internet tienen GPS, y transmiten información a los satélites para que podamos utilizar la georreferenciación.



En resumen, la explosión de los grandes volúmenes de datos no para de crecer y seguirá en pleno crecimiento. Si bien el punto de inflexión lo podemos identificar en la década de 1990 con el surgimiento de internet, tal como lo conocemos hoy, en verdad es justo mencionar que internet es una parte (importante, por cierto) pero no la única.

Erick Schmidt quien fuera en el año 2010 el CEO de la empresa *Google*, en una conferencia dijo algo sorprendente; afirmó que, hasta el año 2003, la humanidad había generado 5 *Exabytes* de información, a lo largo de toda su historia.



RFID o identificación por radiofrecuencia (del inglés *Radio Frequency Identification*) es un sistema de almacenamiento y recuperación de datos remotos que usa dispositivos denominados etiquetas, tarjetas o transpondedores RFID. El propósito fundamental de la tecnología RFID es transmitir la identidad de un objeto (similar a un número de serie único) mediante ondas de radio. Las tecnologías RFID se agrupan dentro de las denominadas **Auto ID** (*Automatic Identification*, o identificación automática).

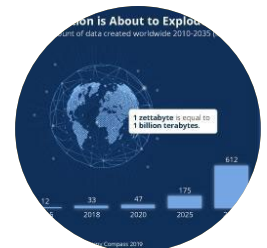


Según otra investigación realizada en el año 2010, se pudo estimar que, en el año 2007, la información creada, capturada, procesada y transmitida en cualquier formato digital, llegó a sobrepasar los 281 Exabytes de información. Para el año 2011 esa cifra fue de 1800 Exabytes y diversos estudios estiman que para el año 2035 se llegarán a cifras que rondan los 2142 Zettabytes...algo así como 2142000 Exabytes de información.

Veamos a continuación cómo se estima y se proyecta el crecimiento de la información en el mundo. Los cálculos pueden variar según estudios, pero en general todos se aproximan a los mismos valores.

Año	Exabytes Generados
2003	5
2007	281
2011	1800
2018	33000
2020	47000
2035	2.142.000

Figura 1. Crecimiento de la información en el mundo (Elaboración propia).



El crecimiento exponencial de los datos en el mundo es claro y contundente. En el año 2012, se publicó un artículo en la prestigiosa revista HBR *Harvard Business Review* que se tituló “*Big data: The management Revolution*”. El artículo fue publicado por Andrew McAfee y Erick Brynjolfsson. Allí, las conclusiones son contundentes: “*La explotación de los nuevos y espectaculares flujos de información pueden mejorar radicalmente el desempeño (rendimiento) de su empresa, sin embargo, será necesario cambiar su cultura de toma de decisiones*”.

La propuesta final de su artículo es concluyente: “*La evidencia es clara: las decisiones controladas por los datos tienden a ser mejores decisiones: los líderes empresariales o bien adoptan esta situación o serán remplazados por otros que lo hagan*”.





Seguramente pueda existir un debate muy complejo en torno a si es conveniente adoptar big data en una empresa o no, del mismo modo que hubo un debate en algún momento en torno a si hay que adoptar las soluciones de Cloud (basadas en la nube) o hay que esperar.

Sin dudas, big data es y será más necesario aún en el futuro. Como todo lo nuevo, y lo desconocido, lleva un tiempo hasta que la tecnología madure y se masifique.

De todas formas, retomando lo que indicaban los autores anteriormente mencionados, no dejan ningún tipo de dudas: *“las empresas que inyectan big data y analítica de big data en sus operaciones muestran tasas de productividad y rentabilidad que son del orden del 5 % al 6 % más altas que aquellas de la competencia o compañías homólogas”* (McAfee, A & Brynjolfsson, E).

Los datos, tal como vimos, crecen exponencialmente y a un ritmo cada vez mayor. A tal punto que existen diversos estudios que se enfocan en cuantificar lo que ocurre en internet durante un minuto de nuestras vidas. Estos estudios son importantes, porque de alguna forma nos permiten comprender cómo va evolucionando el crecimiento de los datos a nivel mundial, sobre todo gracias a las tecnologías, aplicaciones y plataformas que van surgiendo o se van desarrollando año tras año.

Según un informe elaborado por “Visual Capitalist”, lo que se ve en la figura 5 es lo que ocurrió en el año 2020 en un minuto, solo en internet.



Figura 2. Lo que sucede en internet en un minuto (Visual Capitalist).



Aquí podemos detenernos y asombrarnos por los indicadores que se presentan, pero también podemos pensar de qué manera podemos utilizar tantos datos a nuestro favor, para que a partir de ellos podamos generar información valiosa y por qué no, para que las decisiones que podamos tomar con ellas, nos permitan generar alguna ventaja competitiva frente a los demás.



¿Qué sucede en el Internet en 1 Minuto?



02

UNIDADES DE MEDIDA DE LA INFORMACIÓN

Es necesario clasificarlas en medidas, porque ello permite una mejor representación y entendimiento. Por ejemplo, anteriormente mencionábamos que para el año 2003 se habían generado 5 Exabytes de información, por lo que uno podría decir: *¡qué interesante!*, pero *¿a qué equivalen 5 Exabytes de información?* Veamos la figura 6 para comprender la escala de magnitudes en lo que respecta a la medición de la información.

Medida	Simbología	Equivalencia	Equivalente en Bytes
byte	b	8 bits	1 byte
kilobyte	Kb	1024 bytes	1024 bytes
megabyte	MB	1024 KB	1 048 576 bytes
gigabyte	GB	1024 MB	1 073 741 824 bytes
terabyte	TB	1024 GB	1 099 511 627 776 bytes
petabyte	PB	1024 TB	1 125 899 906 842 624 bytes
exabyte	EB	1024 PB	1 152 921 504 606 846 976 bytes
zetabyte	ZB	1024 EB	1 180 591 620 717 411 303 424 bytes
yottabyte	YB	1024 ZB	1 208 925 819 614 629 174 706 176 bytes
brontobyte	BB	1024 YB	1 237 940 039 285 380 274 899 124 224 bytes
geopbyte	GB	1024 BB	1 267 650 600 228 229 401 496 703 205 376 bytes

Figura 3. Escala de magnitudes de la medición de la información
(www.siguanta.com)



CONCEPTO

Las unidades de medida de información son aquellas que se refieren al espacio que ocupa la información en cualquier dispositivo de almacenamiento.



Veamos otro gráfico que nos puede ayudar a entender mejor la escala para la medición de la información.

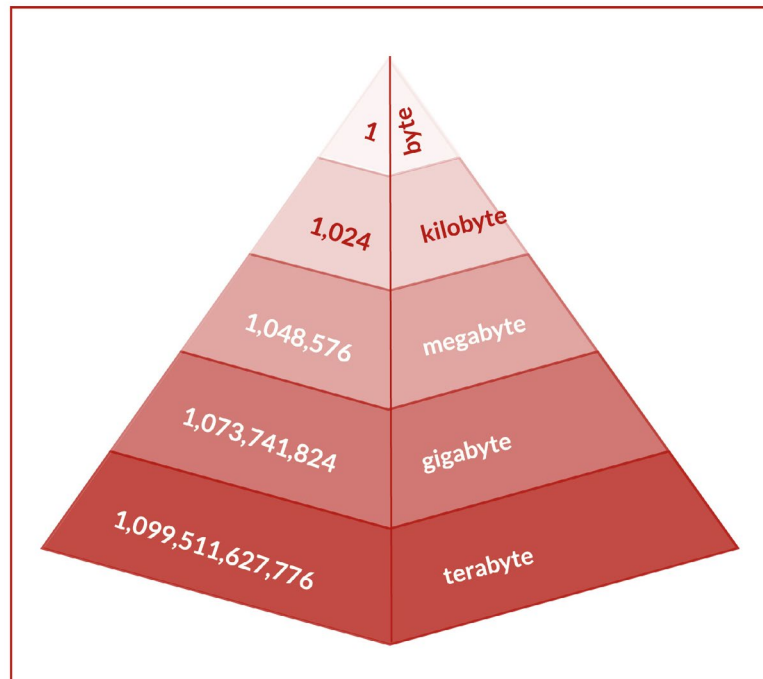


Figura 4. Unidades de medida del byte
(www.siguanta.com)

Será importante comprender las medidas de la información, porque en los proyectos de big data trataremos con grandes volúmenes de datos. Proyectos que pueden ir desde varios *Terabytes* de información hasta varios *PetaBytes* o incluso *Exabytes*. Claro que almacenar tanta información, accederla y procesarla no será fácil, sobre todo si pensamos en modelos de big data que son on premise, es decir, instalados dentro de la organización. Pero de eso nos ocuparemos más adelante.



Podríamos decir y comparar de manera análoga, que 1 Exabyte es el equivalente a 2000000 (dos millones) de computadoras personales, puestas una al lado de la otra.



03

¿QUÉ ES EL BIG DATA?

El término big data surge por primera vez en un artículo de investigación que en el año 1997 publican dos investigadores de la NASA: Michael Cox y David Ellsworth. En ese paper afirman que el gran aumento de los datos se estaba convirtiendo en un problema para los sistemas informáticos actuales. Esto se da a conocer al mundo como “el problema del big data”.

No existe un consenso acerca de lo que es el big data o, mejor dicho, no existe una única definición. Por lo que hay numerosos autores, artículos, documentos y libros que tienen definiciones diferentes, pero que en general se dirigen hacia un mismo concepto, un mismo objetivo.

La consultora Gartner define al big data como “Los activos de información de gran volumen, alta velocidad y/o gran variedad que exigen formas rentables e innovadoras de procesamiento de información que permitan una mejor comprensión, toma de decisiones y automatización de procesos.”¹

Otra definición muy significativa es la que brinda la consultora McKinsey, que define el término del siguiente modo: “big data se refiere a los conjuntos de datos cuyo tamaño está más allá de las capacidades de las herramientas típicas de software de bases de datos para capturar, almacenar, gestionar y analizar”².





Creo que toma una buena dimensión la definición que brinda la consultora internacional IDC 3 sobre el big data: *“El big data es una forma de generación de tecnologías, arquitecturas, y estrategias, diseñadas para capturar y analizar grandes volúmenes de datos, los cuales provienen de múltiples fuentes heterogéneas a una alta velocidad, con el objetivo de extraer u obtener cierto valor económico de ellos”*. Esta definición lo tiene casi todo: habla de tecnologías, estrategias y arquitecturas para poder capturar y analizar los grandes volúmenes de datos. Esto es claramente, porque los sistemas y arquitecturas tecnológicas tradicionales no tienen capacidad de hacerlo o al menos quedan limitados para tal fin.

En lo que respecta a las múltiples fuentes heterogéneas, se refiere a la diversidad de los datos que podemos tener (estructurados y no estructurados) y sobre todo, a la diversidad de donde provienen (internet, redes sociales, IoT, transacciones en un CRM, por citar algunos ejemplos).

En relación con el término de alta velocidad: es cierto, los datos que se generan por minuto hemos visto que son muy grandes, y lo mismo puede ocurrir en nuestras organizaciones. La variación de los precios de una criptomoneda, por citar un ejemplo, puede ser muy cambiante segundo a segundo, minuto a minuto. La velocidad será un factor crítico de éxito para cualquier proyecto de big data. No debemos perder de vista, que el objetivo que persigue el big data es el obtener, a partir de la gran cantidad de datos, información relevante para la toma de decisiones. Decisiones que en el fondo lo que buscarán es ser más eficientes, generar valor agregado y, por supuesto, buscarán obtener mayor rentabilidad.



CONCEPTO

El big data es una forma de generación de tecnologías, arquitecturas, y estrategias, diseñadas para capturar y analizar grandes volúmenes de datos, los cuales provienen de múltiples fuentes heterogéneas a una alta velocidad, con el objetivo de extraer u obtener cierto valor económico de ellos.



04

CARACTERÍSTICAS DEL BIG DATA

Sabemos que una de las características principales del big data son sus conocidas y denominadas 3V que, recordemos, hacen referencia a las principales características: **Volumen, Velocidad y Variedad**. Lo cierto es que con el correr del tiempo y los avances que se van gestando sobre el big data, hay autores que sugieren que en verdad existen otras “V” que se consideran igual de importantes y que por eso analizaremos a continuación.

- **Veracidad:** a medida que la variedad de datos crece, por medio de diversas fuentes heterogéneas de información se vuelve importante el hecho de no perder de vista la veracidad de la información. Resulta bastante lógico pensar que entre tanta información podrían existir datos falsos o que no son ciertos *¿Cómo afectaría eso a una solución de big data?* Es por eso por lo que la veracidad es otro pilar importante al momento de trabajar con big data.
- **Valor:** es claro que los datos por sí solos no aportan nada. Lo interesante y lo que se busca con el big data, es transformar los datos para que a partir de ellos se genere información, porque sabemos que, a partir de la información, obtendremos conocimiento, y ese conocimiento, será crucial para que podamos alcanzar las metas y objetivos organizacionales.



RECUERDA

Características principales del big data:

- Volumen
- Velocidad
- Variedad



- **Viabilidad:** como concepto fundamental se refiere a la viabilidad de disponer de los datos y de la información. Veremos más adelante que no será fácil entre tanta cantidad de datos, disponerlos y conseguirlos en tiempo real (por citar un ejemplo) *¿Es viable para mi compañía utilizar todos los datos en tiempo real para que las decisiones se tomen en cada instante del tiempo?* Veremos para esto que los proyectos de streaming son muy interesantes, pero sobre todo, la capacidad de poder acceder a la información cuando se requiera es la V que conforma la viabilidad.
- **Visualización:** un componente sumamente interesante, necesario, complejo y a la vez complementario a todo lo que venimos hablando sobre big data. Nos referimos al modo en que los datos serán presentados visualmente para que sean legibles, accesibles y sobre todo para encontrar patrones y claves ocultas a investigar. Fenómenos que dentro del big data tienen mucha relevancia y es de especial interés para quienes toman decisiones basadas en datos. Tener un proyecto de big data y no poder visualizar adecuadamente la información no tendría mucho sentido. Es por eso, que esta V es igual de importante que las anteriores.

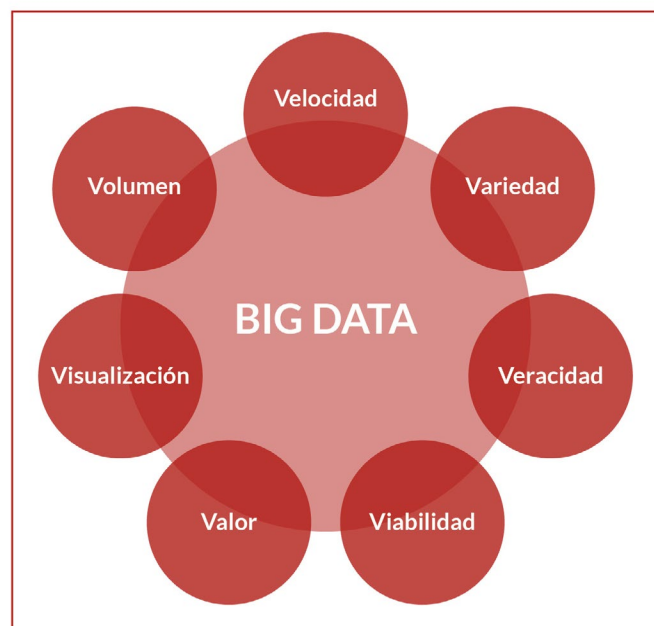


Figura 5. Principales características del big data (Elaboración propia).



05

DIFERENCIAS ENTRE BIG DATA Y BUSINESS INTELLIGENCE

Es normal y hasta cotidiano que muchas organizaciones utilicen sistemas informáticos para conocer los principales indicadores del negocio, o incluso para obtener reportes de diversos tipos y formas, lo que conlleva a que a partir de esa información se tomen decisiones.

A veces, esas decisiones que hay que tomar serán decisiones que hacen a la gestión cotidiana, y otras serán decisiones del negocio a un nivel más estratégico.

Los sistemas de información existen en todo nivel dentro de la empresa. A tal punto, que han sido clasificados y tienen un nombre asignado, tal como puede observarse en la siguiente figura.

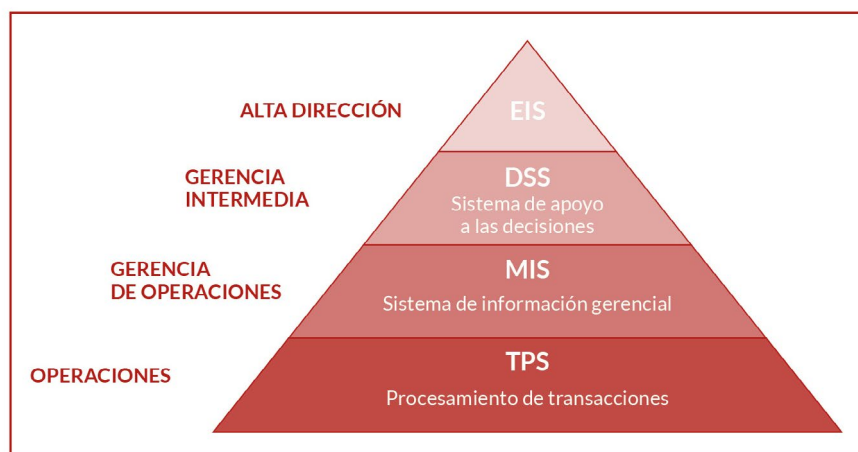


Figura 6. Sistemas de información dentro de la empresa (Elaboración propia).



5.1. ¿QUÉ ES LA INTELIGENCIA DE NEGOCIOS O BUSINESS INTELLIGENCE?

La inteligencia de negocios (o BI por sus siglas en inglés) tiene sus orígenes en la década de 1960 y nació como un simple sistema de intercambio de información entre organizaciones. Recordemos que, por aquellos días, la computación no estaba desarrollada tal como la conocemos en la actualidad. Es por ello que el concepto se desarrolló aún más en la década de 1980, con la aparición del PC, que comenzó a utilizarse para la toma de decisiones, como un proceso de convertir los datos en información. Inicialmente, el BI era algo de lo que se ocupaba el departamento de TI (*Tecnología de la Información*) dentro de las organizaciones, pero luego, dado su crecimiento, se han conformado grupos y áreas específicas que cumplen esta función, separándose así la responsabilidad del área de TI, que en definitiva brinda los medios y recursos técnicos para que los sistemas funcionen.

En la actualidad, las soluciones de BI, se orientan a priorizar los datos gobernados en plataformas confiables, realizando análisis para que los usuarios puedan a gran velocidad, obtener información. Típicamente los sistemas de BI cuentan con herramientas que ayudan a este proceso. Algunas de estas herramientas son las siguientes:

- **Reportes Empresariales.** Con información estadística relevante al negocio. Se analiza cómo fue el resultado de las ventas en cada tienda en los últimos tres meses, por ejemplo.
- **Alertas de negocio.** Estableciendo ciertos criterios se pueden recibir alertas tempranas cuando un proceso ocurra. Por ejemplo, se podría recibir una alerta si un local de comidas no ha tenido movimiento bancario durante las últimas 24 h.
- **Cubos de análisis.** Un cubo es una estructura de datos organizada mediante jerarquías. Cada indicador se puede evaluar en cualquier nivel de jerarquía establecido. Un ejemplo típico es el **cubo de ventas**: para poder visualizar las ventas a nivel diario, mensual, o anual para un cliente determinado, en una zona geográfica determinada, como puede ser un país, o una ciudad. Sin dudas es una herramienta muy utilizada por los sistemas de BI.



CONCEPTO

La Inteligencia de Negocios es un conjunto de datos y aplicaciones que dispone una organización cuyo foco es la generación del conocimiento sobre el negocio, para que a partir de este conocimiento se puedan tomar decisiones.



Estas herramientas, en la actualidad, utilizan tecnologías modernas para presentar, visualizar y operar con la información.



Los sistemas de BI permiten, además, trabajar con Data Mining y análisis estadístico. Estas son herramientas que se utilizan para establecer modelos predictivos o causa efecto entre dos datos.

Por todo esto, podemos decir que los sistemas de BI trabajan con los datos disponibles en una organización, y se centran en el proceso de análisis de dichos datos para convertirlos en Información, que será utilizada para una oportuna toma de decisiones. Por lo general, los datos sobre los que van a trabajar son datos del pasado (sobre lo que ocurrió).

Por otro lado, el big data pondrá foco y se centrará en la captura, el almacenamiento, el procesamiento y transformación de los grandes volúmenes de datos, que pueden llegar a la organización considerando diferentes fuentes a una gran velocidad.

Por supuesto que big data se puede complementar perfectamente con herramientas y tecnologías, como la inteligencia artificial y el propio Business Intelligence para explotar al máximo esos datos, que de manera tradicional y con la infraestructura estándar no podría hacerlo.

Existe otra arista muy importante a considerar que es la llamada Analítica de Negocios o *Business Analytics*.

Existen tres grandes aspectos que conllevan al área del Business Analytics, los cuales explicaremos a continuación.

- **El análisis descriptivo:** me ayudará a entender lo que ocurrió en el pasado dentro de la organización.
- **El análisis predictivo:** me ayudará a establecer modelos que nos sirvan para comprender que puede ocurrir, es decir, predecir el comportamiento futuro.
- **El análisis prescriptivo:** me ayudará a entender qué debo hacer y cómo debo hacerlo en mi organización, para los criterios y reglas del propio negocio.



EJEMPLO

Los días viernes entre las 10 am y las 3 pm se podrá colocar una oferta en un local de comidas, de un producto determinado, porque los sistemas de BI me ayudaron a detectar que, estadísticamente, determinados días, a una hora concreta las probabilidades de compra de dicho producto aumentan por las características de los compradores que se reciben.

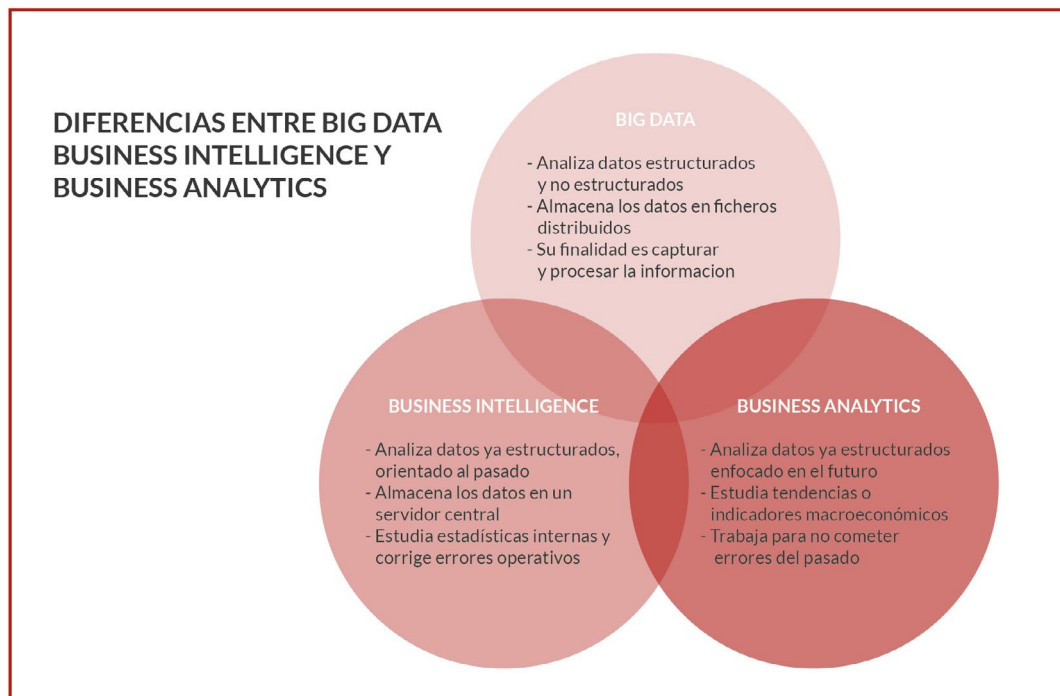


Figura 7. Diferencias entre Big Data, Business Analytics y Business Intelligence



CONCEPTO

Business Analytics es el proceso por el cual se transforman los datos en información de relevancia para el negocio, con el objetivo de tomar decisiones acertadas.

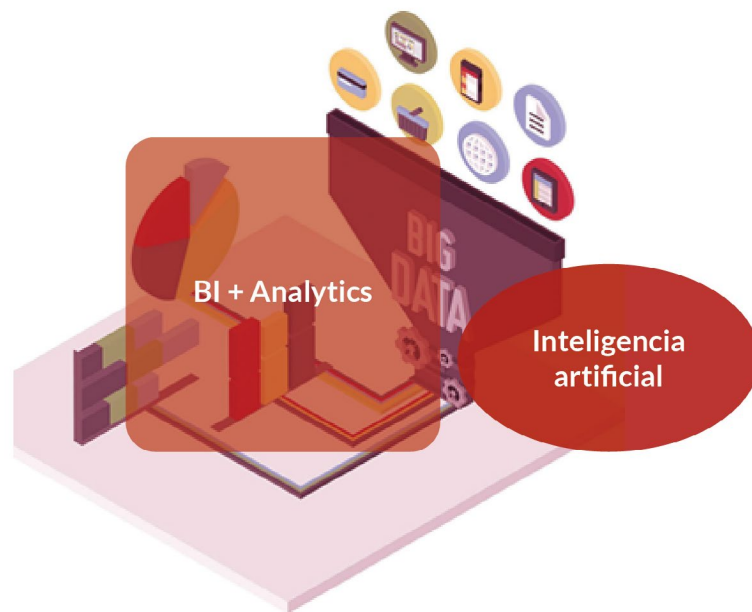


Figura 8. Big data enriquecido (Elaboración propia).

En la figura 8 vemos como el big data es nutrido desde diversas fuentes de datos. Hay un procesamiento y modelado de la información, donde la Inteligencia Artificial interviene de manera directa, para luego, en una etapa posterior, se puedan utilizar los sistemas de BI y aplicar analíticas de datos y así obtener de manera visual los resultados.



CONCLUSIÓN

Como resumen, podemos establecer que el big data por sí solo no tendría mucho sentido. Necesitaremos no solo de la gran cantidad de datos con todas sus particularidades, sino también, las herramientas de BI y de Business Analytics, conjuntamente con la Inteligencia Artificial, para poder explotar al máximo los resultados que esperamos.



06

LA IMPORTANCIA DE LA INTELIGENCIA ARTIFICIAL EN EL BIG DATA

Al igual que sucede con el big data no hay una única definición que se pueda brindar sobre la Inteligencia Artificial.

Según la consultora internacional Gartner, la *inteligencia artificial* (IA) aplica análisis avanzados y técnicas basadas en la lógica, incluido el aprendizaje automático, para interpretar eventos, respaldar y automatizar decisiones y tomar acciones⁴.

La inteligencia artificial tiene muchos ámbitos de aplicación. Podríamos decir que muchas ramas derivan de ella, pero en términos generales, la inteligencia artificial, vista como un conjunto de tecnologías, nos permitirá extraer conocimiento a partir de los datos.

Esta misma IA nos permitirá detectar patrones (por ejemplo, patrones de comportamientos), nos permitirá también extraerlos y sobre todo, nos permitirán hacer que las computadoras aprendan de manera autónoma.



CONCEPTO

Podemos decir que la Inteligencia Artificial es la aplicación de la ciencia informática para resolver problemas de forma inteligente. Es decir, que es algo que, computacionalmente hablando, imita a la inteligencia humana.



Seguramente el lector se estará preguntando: *¿Cómo hace la Inteligencia Artificial para encontrar patrones en los datos?* Recordemos que, sobre todo, cuando hablamos de big data, decimos que los datos pueden ser estructurados o no estructurados y su complejidad es muy alta, su variedad y sobre todo las fuentes de datos son muy variables. La conjunción de las matemáticas y el aprendizaje automático, también conocido como machine learning, es lo que permite identificar patrones y hacer que las computadoras aprendan de manera autónoma. Dependiendo del tipo de modelo de machine learning que se utilice, es probable que cuantos más datos ingresen al sistema (o al modelo), serán mejores o más precisos los resultados.



El aprendizaje automático se puede aplicar a muchos problemas y conjuntos de datos diferentes. Por ejemplo, con el aprendizaje automático o machine learning podemos identificar imágenes de ciertos animales (como un perro, por ejemplo), en grandes colecciones de fotos diversas. También podemos identificar un posible caso de fraude bancario, debido a una secuencia previa y combinaciones de movimientos sospechosos, podemos transformar un archivo de audio y pasarlo a texto, entre otras tantas aplicaciones.

Con relación al big data, la inteligencia artificial es una herramienta indispensable. Indiscutible. Fundamental y muy importante. Si bien su aplicación dependerá de lo que se quiera realizar con ella, es fundamental aplicarla a los grandes volúmenes de datos para que los mismos algoritmos nos ayuden a resolver los problemas, y hasta a tomar decisiones por nosotros mismos.

La inteligencia artificial y, en especial, el machine learning como una rama de esta, es la que se aplica como herramienta para hacer analítica de negocios (*business analytics*), porque gracias al machine learning podríamos hacer análisis descriptivo, predictivo y prescriptivo.



Sin Inteligencia Artificial, el big data no tendría ningún tipo de sentido.

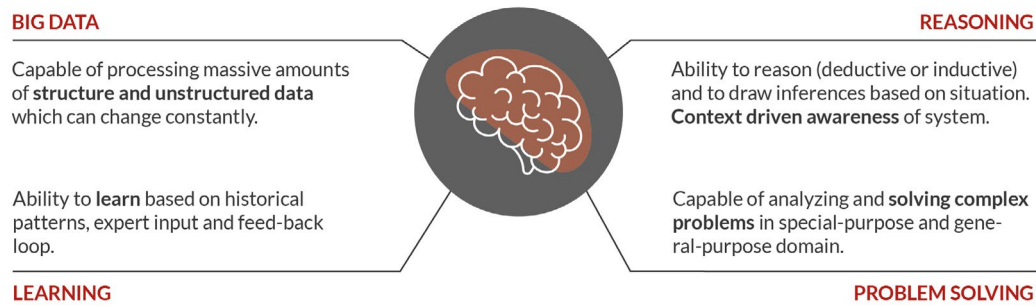


Figura 10. Características del Big Data
(Perspectivas Parte 1: Inteligencia Artificial Definida La terminología más utilizada5).



*Big Data e inteligencia artificial, ¿son el futuro? | Gustavo Arjones
| TEDxEldorado*



07

EMPRESAS QUE UTILIZAN BIG DATA CON INTELIGENCIA ARTIFICIAL

El big data y la inteligencia artificial están presentes en muchos aspectos de nuestras vidas. Aunque muchas veces no nos damos cuenta, cuando accedemos a la plataforma de *Netflix*, notamos que recibimos recomendaciones, nos sugiere contenido, series, películas. Si nos detenemos en las cuestiones técnicas, todas estas sugerencias se realizan gracias a diversos algoritmos de machine learning que combinan nuestros gustos, preferencias y datos históricos.

Del mismo modo, podemos citar el caso de *Spotify*, una de las plataformas de audio más populares que existen. Spotify utiliza sus algoritmos basados en inteligencia artificial (*machine learning*) para que en función de la música que escuchamos habitualmente nos recomiende canciones o géneros musicales acordes a nuestros gustos.

Si vamos a contar ejemplos de big data e Inteligencia Artificial, sin dudas no podemos dejar de mencionar a la empresa *Amazon*, quien es una de las empresas pioneras en el uso de los datos y la Inteligencia artificial. Su sistema de recomendaciones ha servido mucho para aumentar sus ventas, pero sobre todo, han utilizado la Inteligencia artificial para la optimización de los precios, los cuales, son dinámicos dependiendo la hora del día y el tipo de usuario que somos...entre otros factores.

“Amazon.com está basando gran parte de su negocio en sistemas de aprendizaje automático. Sin el aprendizaje automático, Amazon.com no podría ampliar su negocio, mejorar la experiencia y selección de los clientes, y optimizar su velocidad y calidad logística. Amazon.com creó AWS para permitir a otras empresas disfrutar de la misma infraestructura de TI, con ventajas de agilidad



y costes, y ahora sigue democratizando las tecnologías de ML al ponerlas al alcance de todas las empresas.

La estructura de los equipos de desarrollo de Amazon.com y el enfoque en ML para resolver problemas empresariales pragmáticos complicados llevan a Amazon.com y AWS a desarrollar herramientas y dispositivos de ML potentes y sencillos de usar. Estas herramientas se prueban primero en la escala y entorno crítico de Amazon.com, antes de proporcionarse como servicios de AWS para que los usen otras empresas, del mismo modo que otros servicios de TI6”.

Big data también se está empleando para monitorear bebés en la unidad de recién nacidos de un hospital de Toronto (Canadá). El hospital graba y analiza latidos junto al patrón respiratorio de cada Bebé, donde gracias a la Inteligencia artificial, los algoritmos son capaces de predecir infecciones 24 horas antes de que los primeros síntomas aparezcan. De esta forma, los profesionales médicos pueden intervenir y salvar vidas en un entorno sumamente complejo, como lo es un hospital7.

El último caso que mencionaremos en esta semana se refiere al caso de Nike. Una de las empresas más importantes del mundo con relación a productos deportivos. Nike utiliza el big data e inteligencia artificial para personalizar la experiencia de los compradores (clientes) y de los usuarios de sus plataformas.

Esto lo logra gracias a la recolección, análisis y procesamiento de la información que provienen de diferentes atletas de todo el mundo (que utilizan la app de Nike+). Dado que mediante esa Aplicación se pueden registrar datos de todo tipo. Nike aprovecha la Inteligencia artificial para sacar nuevos productos, conocer mejor a su público e incluso, generar ventaja competitiva. Nike aumentó su facturación mundial un 9 % en el año 2019, motivo por el cual, se lanzó a comprar una empresa de big data8.



VIDEO

7 Ejemplos Reales y Casos de Éxito de BIG DATA



BIBLIOGRAFÍA

Holmes, D. E., & Otero-Piñeiro, D. (2018). Big Data: una breve introducción. Barcelona: Antoni Bosch editor.

Ibáñez, D. B. (2015). Big Data y técnicas cuantitativas: una introducción al análisis de contenido informatizado. Revista San Gregorio, 2(8), 106-111.

McAfee, A & Brynjolfsson, E (2012). Big Data: The Management Revolution. Harvard Business Review.

<https://hbr.org/2012/10/big-data-the-management-revolution>

Tascón, M. (2013). Introducción: Big data. Pasado, presente y futuro. Telos: Cuadernos de comunicación e innovación, (95), 47-50.

REFERENCIAS

Sobre las imágenes

Imágenes de portada: Shutterstock

Notas del autor

1. <https://www.gartner.com/en/information-technology/glossary/big-data>
2. <https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation>
3. <https://www.idc.com/>
4. <https://www.gartner.com/en/information-technology/glossary/artificial-intelligence>
5. Extraído de: <https://www2.deloitte.com/cl/es/pages/technology/articles/parte-1-inteligencia-artificial-definida.html>
6. Extraído de: <https://aws.amazon.com/es/machine-learning/what-is-ai/>
7. Ver caso en: <https://www.baoss.es/10-ejemplos-usos-reales-big-data/>
8. Ver caso NIKE-Semana1.
9. <https://www.campusbigdata.com/big-data-blog/item/148-diferencias-entre-big-data-business-analytics-y-business-intelligence>



