



Universidad
Tecmilenio®





Fundamentos de big data

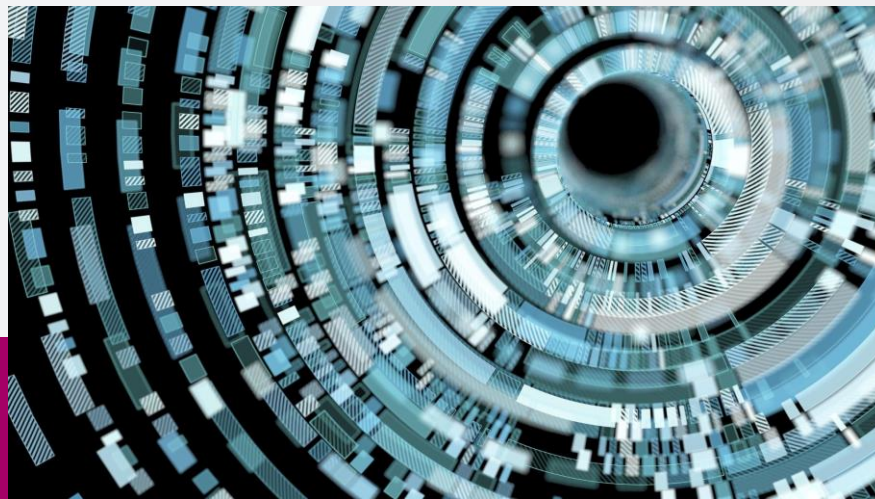
Tema 1. Introducción al big data.





Actualmente en el entorno de la comunicación ha aumentado no solo la **cantidad de información** que existe sino la **velocidad** con la que crece y se divulga, tanto así que se tiene tiempo de analizarla y tomar partido de todo lo que aparece en los diferentes medios.

A esta información **variada, compleja y en grandes cantidades que está disponible en la actualidad, se le denomina “**data**” o **datos**.**



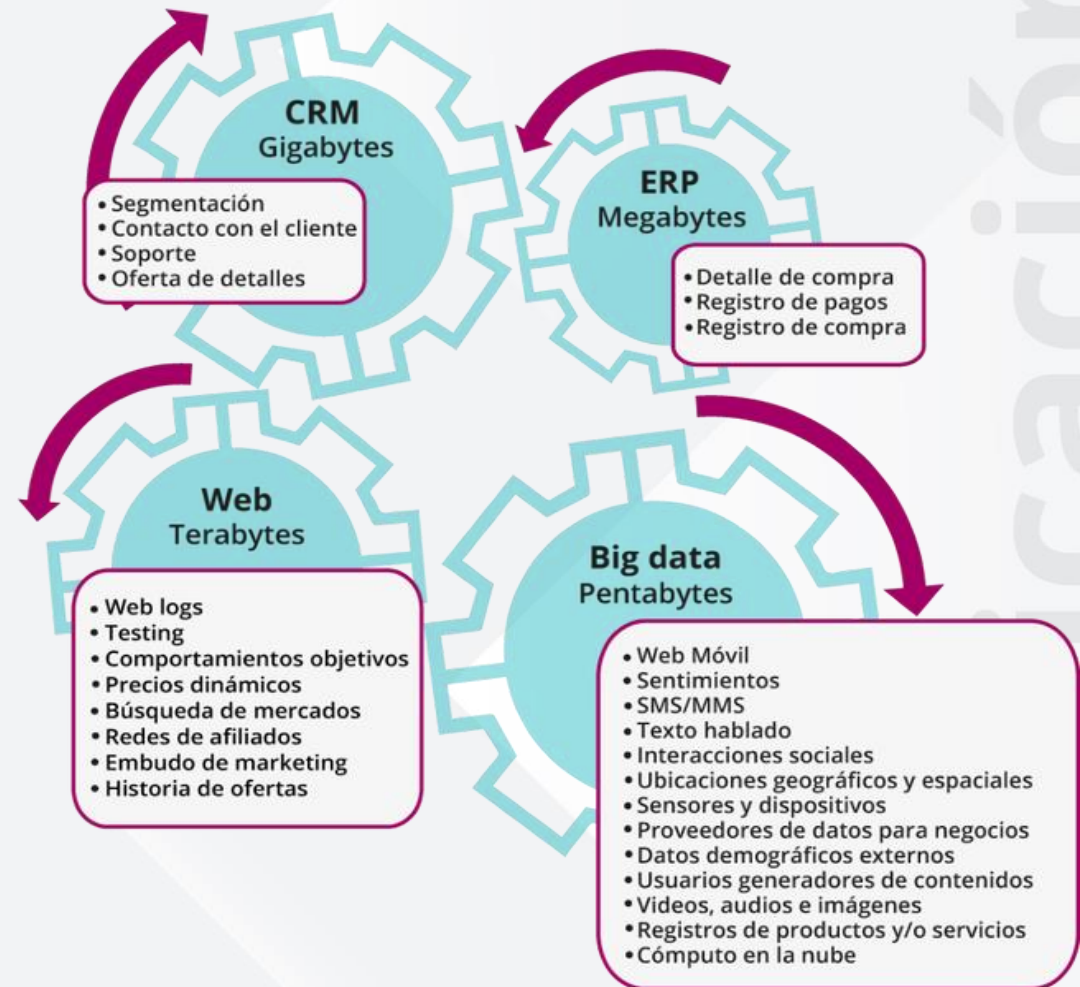


El término **big data**, se refiere a **grandes y complejos volúmenes de datos** de diversas fuentes de información que requieren, a su vez, de diferentes herramientas tecnológicas de apoyo para ser procesados debido a su cantidad y variedad.



Figura 1. Generación de información

Conforme ha ido transcurriendo el tiempo se han incrementado los avances tecnológicos, traducidos en **mayores capacidades de procesamiento y espacio** y al mismo tiempo se ha ido incrementado el uso y explotación de la tecnología, esto ha dado pie a un **incremento exponencial de datos**.



Estas **características** brindan una perspectiva de lo complejo que puede ser el manejo de la información y darle un sentido al propósito de la empresa o departamento.

Fuente: Botello, E. (2022). *Big Data: qué es, para qué sirve y por qué es importante*. Recuperado de



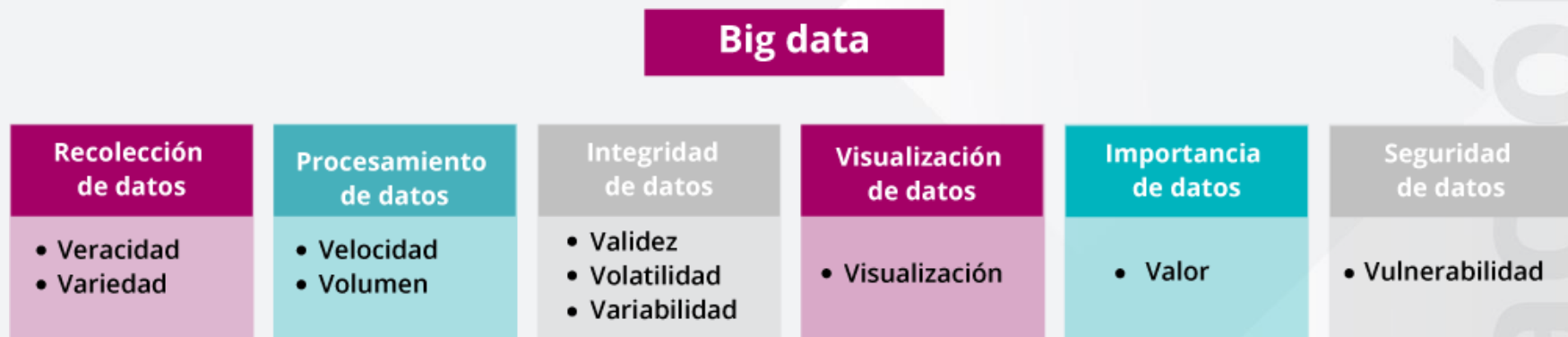
-   <https://www.iebschool.com/blog/valor-big-data/>

Tabla 1. Características del big data

Característica	Descripción	Ejemplo
Volumen	Característica más conocida que hace referencia a la cantidad de datos que se generan.	Cada minuto se suben 300 horas de contenido a YouTube.
Velocidad	Rapidez con la que se genera, produce, y crea la información.	Google procesa en promedio más de 40,000 consultas por segundo
Variedad	Existen muchos tipos de datos sobre todo no estructurados que se están generando.	Audios, videos, imágenes, actualizaciones de estatus en redes sociales. Datos de clics, máquinas, sensores.
Variabilidad	Tres aspectos que afectan: inconsistencias en los datos; multitud de dimensiones de estos; velocidad inconsistente de carga.	En un día un sensor de movimiento puede captar 100,000 movimientos y al siguiente solo 5. *Comportamiento similar en las redes sociales.
Veracidad	A medida que aumentan las demás características, la confianza en los datos disminuye.	Datos abiertos sobre consumo de bebidas. Podemos cuestionar, ¿quién recopiló la información?; ¿de qué tipo de empresa, etc.?
Validez	Se refiere si los datos son correctos y precisos.	Se estima que un científico de datos invierte el 60% de su tiempo en discriminar qué información es válida o no.
Vulnerabilidad	Una violación de seguridad de datos en Big data es de gran impacto.	Un hacker vendiendo contraseñas o información de miles de cuentas de correo y/o redes sociales.
Volatilidad	Antigüedad de los datos para que se consideren relevantes.	Información histórica que no es aplicable hoy en día.
Visualización	Son difíciles de visualizar por la cantidad de datos/variables que se tienen.	Una persona tiene la información de su empresa, de las redes sociales y de sus dispositivos, es complicado concentrar todas estas variables en una sola visualización.
Valor	Solo tiene sentido tener ese dato si nos va a aportar algo al negocio.	Información irrelevante al negocio de los usuarios, es prescindible.

Figura 2. Agrupación de característica del big data.



Fuente: Botello, E. (2022). *Big Data: qué es, para qué sirve y por qué es importante*. Recuperado de <https://www.iebschool.com/blog/valor-big-data/>

Uno de los grandes **beneficios del big data** es el **análisis predictivo**. Los análisis de Big Data ayudan a las empresas a tomar mejores decisiones, al mismo tiempo que maximizan la eficiencia operativa y reducen los riesgos.

Con la ayuda de las herramientas de análisis de big data, las empresas de todo el mundo están mejorando sus estrategias de marketing digital, al utilizar y procesar datos de las plataformas de redes sociales.



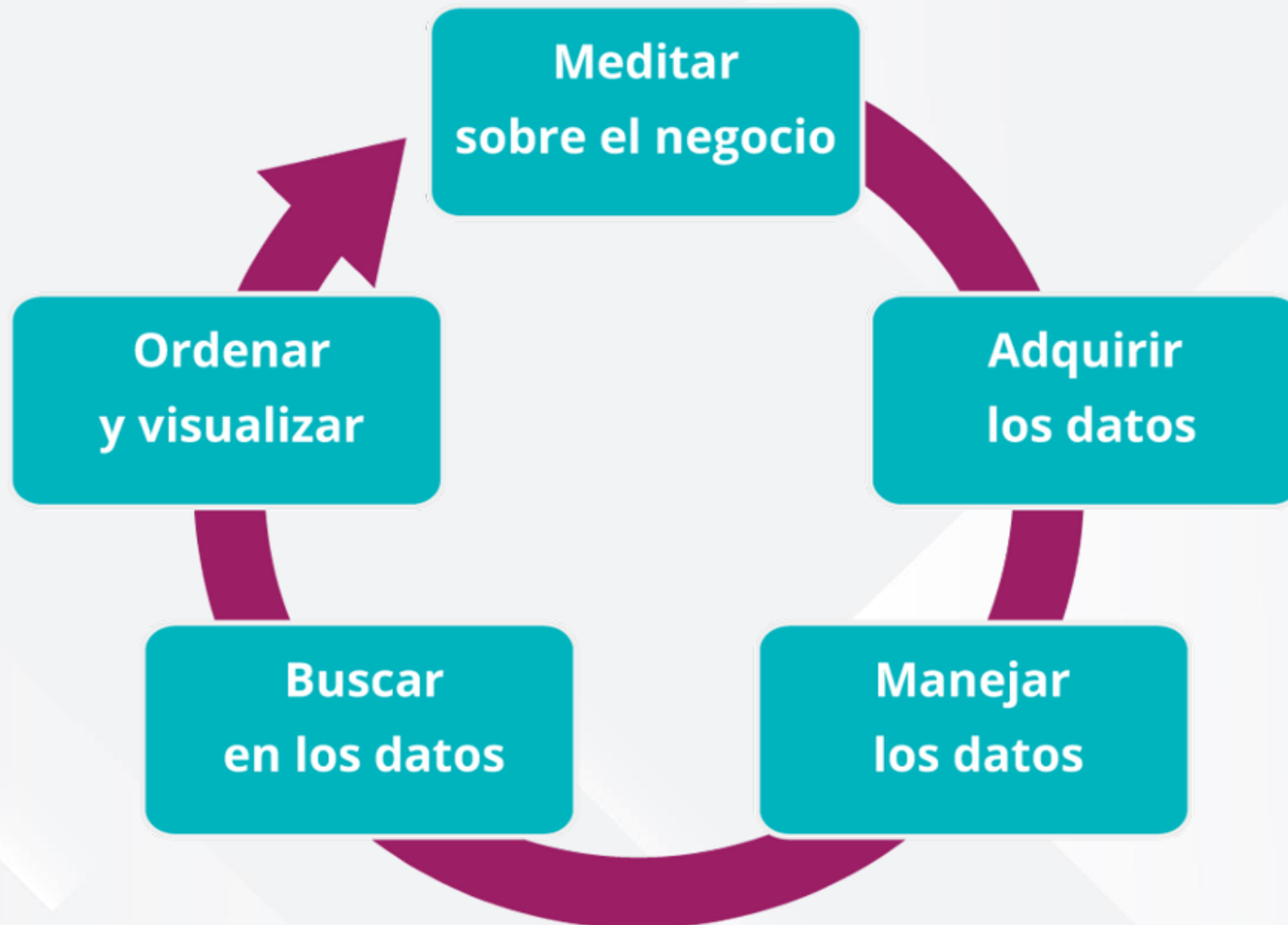
Explicación

Big data combina datos de múltiples fuentes para producir información procesable. **Las empresas pueden ahorrar tiempo y dinero utilizando herramientas de análisis para filtrar datos redundantes.**



Explicación

Figura 3. Metodología para convertir datos en activo del negocio.



Fuente: El arte de medir. (2020). *Big data analytics como activo de negocio*. Recuperado de <https://elartedemedir.com/blog/big-data-analytics-como-activo-de-negocio/>



Tabla 2. Metodología **MAMBO** para integrar el **big data**

Meditar y entender el negocio:

- Su filosofía, características, necesidades, capacidades y recursos.
- Analizar la situación actual del negocio y del mercado.
- Identificar oportunidades del negocio.
- Determinar la medición de objetivos del negocio.
- Alinear la estrategia con el modelo del negocio.

Manejar los datos.

- Entender y describir los datos.
- Explorar e identificar las oportunidades de aprendizaje que aporta valor a los objetivos del negocio.
- Identificar las técnicas y los métodos de análisis más adecuados para apoyar las decisiones estratégicas de la empresa.

Adquirir los datos.

- Identificar las fuentes de los datos disponibles.
- Mapear los datos y sistemas.
- Identificar qué datos extraer para responder a las inquietudes del negocio.
- Capturar los datos.
- Verificar la calidad de los datos.
- Detectar los KPI's del negocio.

Buscar en los datos.

- Analizar los datos. ¿Qué pasó? ¿Por qué? ¿Qué pasará? ¿Cómo podemos cambiar las cosas?
- Comenzar a responder preguntas importantes del negocio.
- Generar productos con impacto en el negocio

Ordenar y visualizar.

- Identificar audiencias, contextos y propósitos para una estrategia de visualización.
- Determinar la forma más relevante de comunicar el conocimiento adquirido mediante los datos.
- Convertir la información en herramientas que faciliten el trabajo de los distintos usuarios.
- Metodología SEVEN.



Tabla. **Ejemplos** de éxito de **análisis de Big Data**

Gracias a metodologías como ésta, la **analítica de big data** permite identificar los casos de éxito dentro de los negocios los cuales van en aumento.

Empresa	Factores de éxito
Airbnb	Recomendaciones según el contexto global. Al sufrir 70% de caída en las reservas, en la pandemia adecuaron su estrategia al ofrecer lo que el público necesitaba: viviendas cómodas, alejadas de las grandes ciudades.
Amazon	Personalización de las recomendaciones. 35% de las ventas anuales son resultado de las sugerencias que lanza Amazon sobre cada compra.
Netflix	Retención de usuarios del 93%, su algoritmo de recomendación hace que el 85% de consumo sea generado a través de él.
Tesla	Vehículos autónomos: Por medio de sensores externos e internos se recopila todo tipo de información, más de 100 millones de datos harán posible la autonomía de los autos.
Starbucks	Con Big Data, eligen en donde ubicar sus tiendas y con su aplicación móvil ha logrado personalizar la experiencia del cliente.
Facebook	Gracias a la analítica han logrado publicidad personalizada y productos más efectivos al enfocarlos al usuario.
ZARA	Predicción de tendencias y estrategia de precios apoyadas en Big Data.
Nike	Se apoya tanto en los dispositivos inteligentes que desarrolla, como en aplicaciones móviles recogiendo datos de las zapatillas. Fomenta la competitividad al compartir los resultados en las redes sociales.
Spotify	Guarda registros de gustos musicales y preferencias de los usuarios, por lo que se le reconoce por las recomendaciones personalizadas. Además, por ello fue capaz de predecir ganadores de premios Grammy acertando en 6 de cada 10.
NBA	Predicción de tendencias y estrategia de precios apoyadas en Big Data.

Fuente: Ibeas, D. (2022). *10 casos de éxito gracias al Big Data*.

Recuperado de <https://contactcenterhub.es/10-casos-de-exito-gracias-al-big-data-2022-21-35300/>





Metodología big data

Objetivo: Elaborar una propuesta que exponga las ventajas y/o beneficios que otorga la aplicación de una metodología de big data para la toma de decisiones.

Instrucciones

- Para realizar esta actividad es importante que consultes el contenido del tema.
- A partir de lo aprendido elabora una propuesta integrando los conceptos más relevantes y significativos del tema en función de mejorar la toma de decisiones en tu empresa o departamento.
- El documento contendrá una breve descripción de la propuesta con una base **MAMBO**, que incluya las ventajas y/o beneficios para la toma de decisiones en tu empresa o departamento.
- Incluye colores, tipografías y material gráfico que respalden y apoyen visualmente tu propuesta.
- Máximo 2 cuartillas.





El **big data** puede apoyar a resolver cuestionamientos que antes no se habían planteado y con el análisis de datos se pueden encontrar soluciones que no eran imaginables y los resultados pueden tener un **desempeño sin precedentes** en la empresa.



Cierre



Botello, E. (2022). *Big data: qué es, para qué sirve y por qué es importante*. Recuperado de <https://www.iebschool.com/blog/valor-big-data/>

Ibeas, D. (2022). *10 casos de éxito gracias al big data*. Recuperado de <https://contactcenterhub.es/10-casos-de-exito-gracias-al-big-data-2022-21-35300/>

El arte de medir. (2020). *Big data analytics como activo de negocio*. Recuperado de <https://elartedemedir.com/blog/big-data-analytics-como-activo-de-negocio/>





Fundamentos de big data

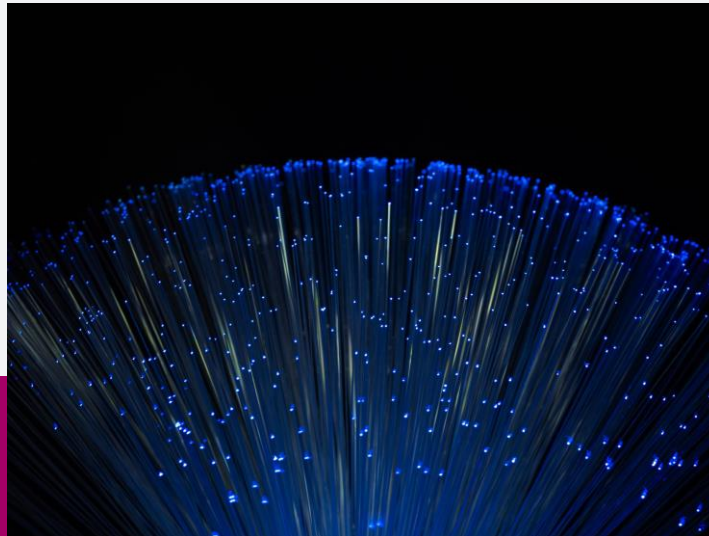
Tema 2. Orígenes de los volúmenes de datos.

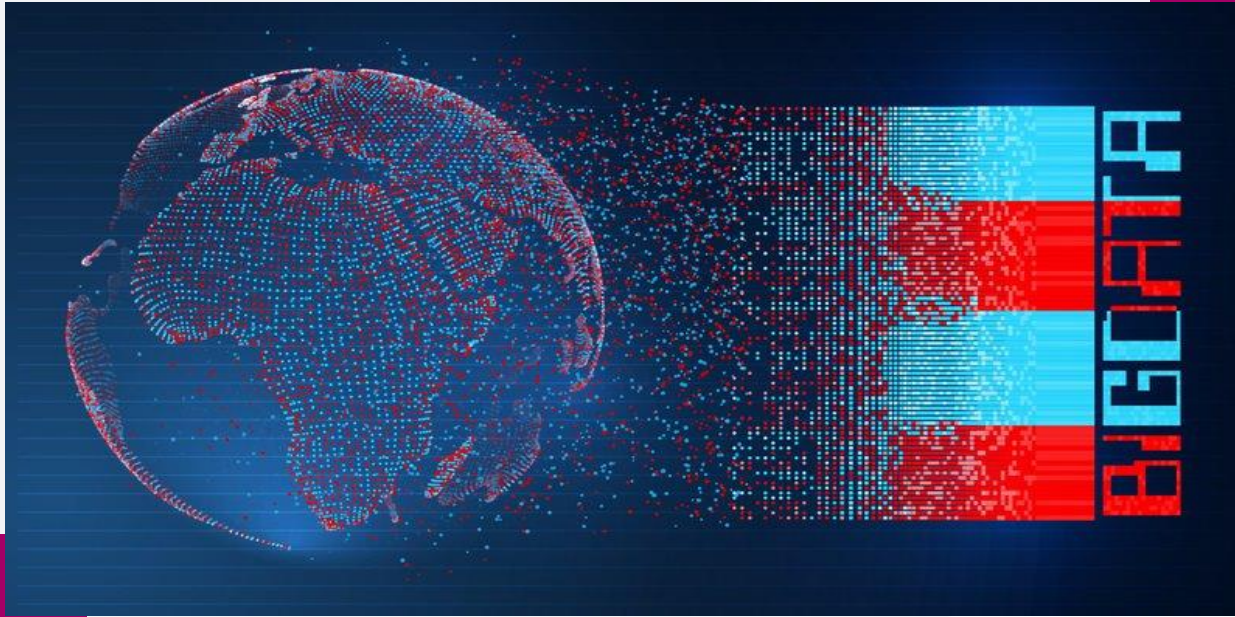




En su verdadera esencia, **big data** no es algo completamente nuevo o solo de la última década. A lo largo del tiempo, las empresas han estado utilizando distintas técnicas de análisis de datos para respaldar el proceso de toma de decisiones.

Es importante **analizar y revisar** de dónde viene toda esta información pues, no toda es veraz; sobre todo actualmente dado que la evolución tecnológica y de sistemas así lo ha propiciado.





En las últimas décadas se ha logrado **almacenar y recolectar** cada vez más rápido más información, lo que ha dado un impulso importante a esta herramienta que brinda mayor apoyo a las decisiones estratégicas del entorno.



Figura 1. Tipos de datos de acuerdo con su origen



Fuente: Nosis Investigación y Desarrollo (2020). *¿Qué es Big Data y por qué es importante?* Recuperado de <https://blog.nosis.com/que-es-big-data/>

Explicación



Datos estructurados: Son cualquier dato que se pueda procesar, acceder y almacenar como un formato fijo.

Datos semiestructurados. Son uno de los tipos de big data según el origen, relacionados con los datos que contienen los formatos a los que se hace referencia.

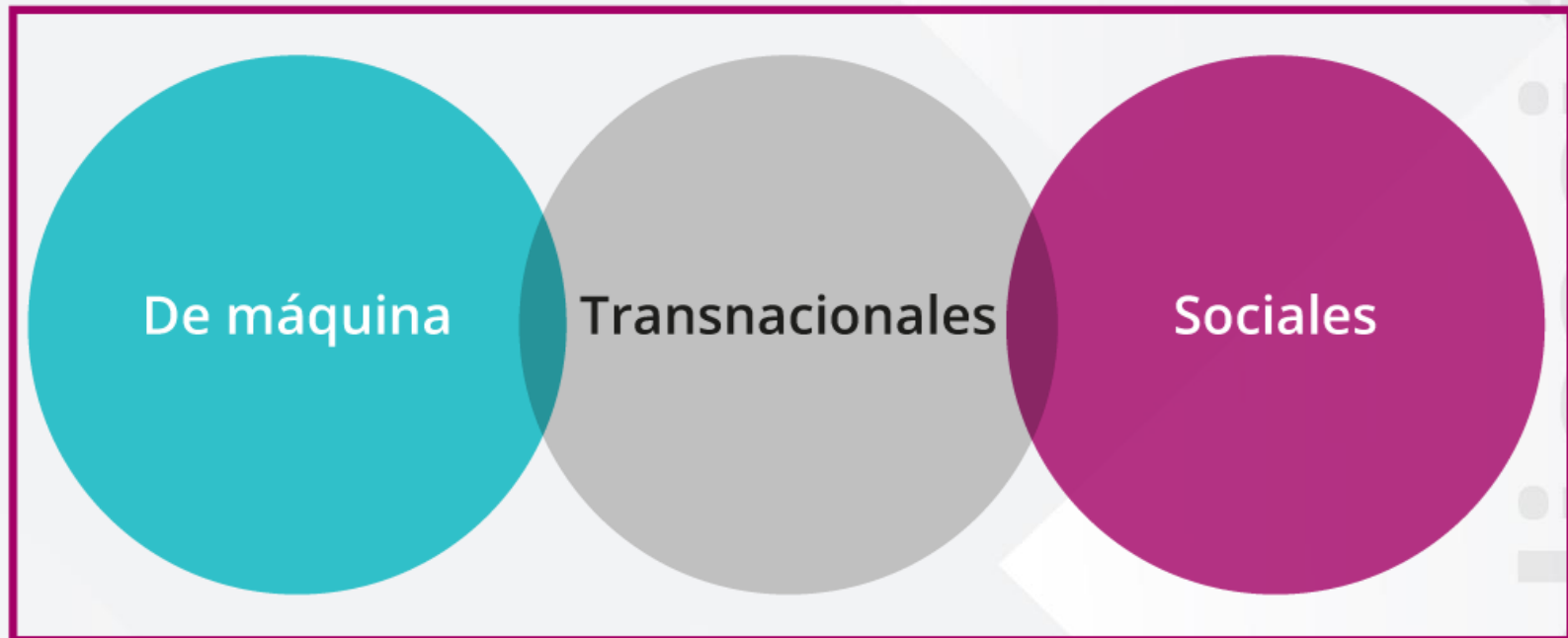
Datos no estructurados. Este es uno de los tipos de big data donde se incorpora el formato de datos de una multitud de archivos.

Subtipos de datos. Existen subtipos de datos que tienen cierto grado de relevancia en el campo de la analítica.





Figura 2. Fuentes primarias de big data.

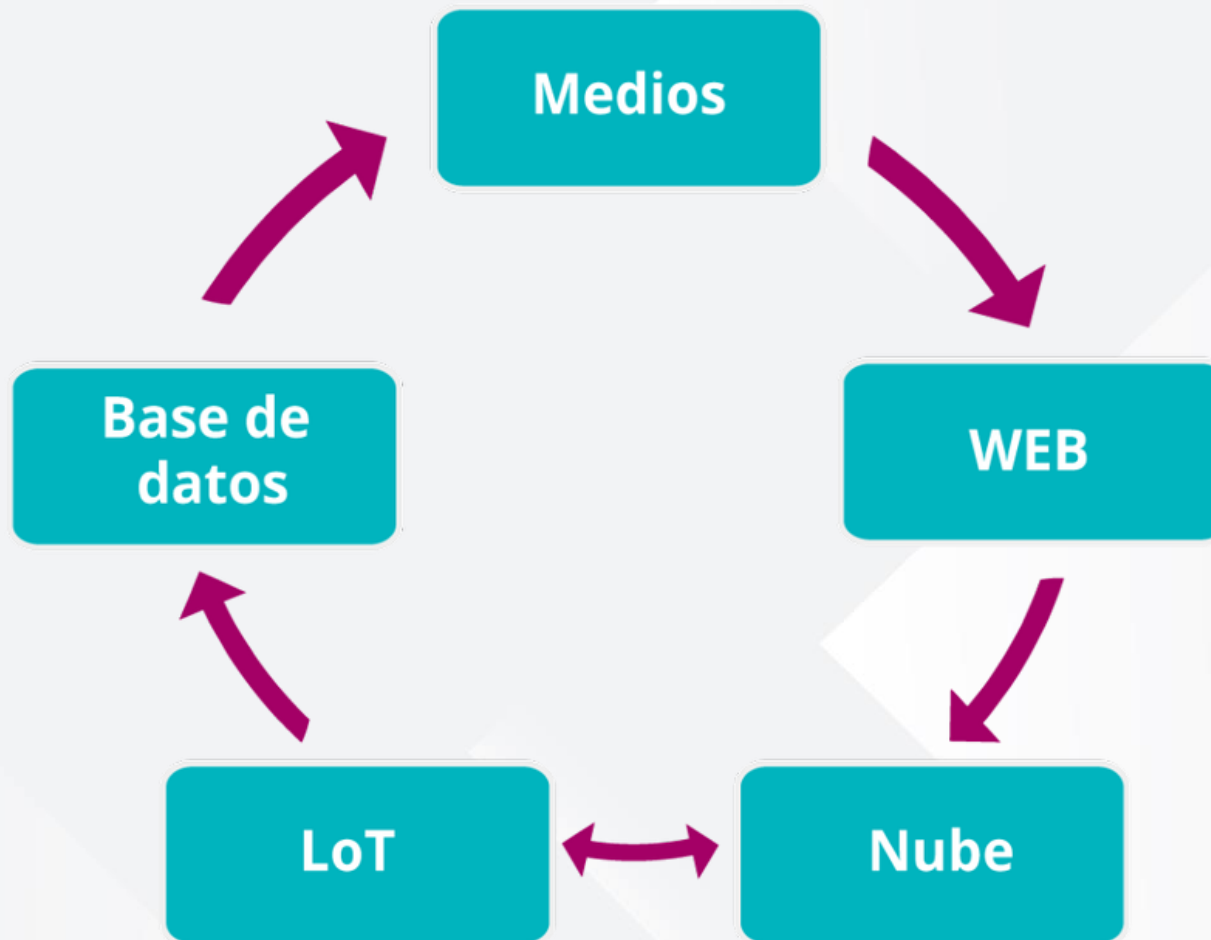


Fuente: Sharma, R. (2021). *Sources of big data: where does it come from?* Recuperado de <https://www.upgrad.com/blog/sources-of-big-data/>



Explicación

Figura 3. Fuentes de información de Big Data.



Fuente: Joshi, N. (2017). *Top 5 sources of big data*. Recuperado de: <https://www.allerin.com/blog/top-5-sources-of-big-data>



Medios. Es la fuente más popular de big data, porque brinda información valiosa sobre las preferencias de los consumidores y las tendencias cambiantes.

Web. La web pública por su enorme tamaño garantiza diversa usabilidad y es especialmente beneficiosa para las empresas emergentes y las PYMES.

Nube. Este almacenamiento tiene datos estructurados y no estructurados, proporcionando información en tiempo real y bajo demanda.

IoT. Se refiere al contenido generado por máquinas que constituyen una fuente valiosa que se origina en sensores conectados a dispositivos electrónicos.

Bases de datos. Las empresas vanguardistas prefieren usar una fusión de datos tradicionales y modernas para adquirir big data relevante.





Fuentes más usadas para **people analytics** son las siguientes:

- Datos de reclutamiento
- Datos demográficos
- Gestión del desempeño
- Gestión del aprendizaje
- Arquitectura de trabajo
- Compensación y beneficios
- Planificación de la sucesión
- Viajes
- Tutoría
- Datos de encuestas de empleados
- Encuesta de compromiso
- Datos de ausencia
- Datos de gestión de la relación con el cliente (CRM)
- Datos financieros
- Datos de gestión de producción
- Datos de ventas





Línea del tiempo

Objetivo: Exponer la diversidad de los tipos de datos, origen de su generación y principales fuentes de big data.

Instrucciones:

- Para realizar esta actividad es importante que consultes el contenido del tema.
- **Elabora una línea del tiempo** que incorpore los contenidos aprendidos en proporción con la generación de los datos, fuentes y cambios significativos de los grandes volúmenes de datos.
- El diseño deberá contener una breve descripción del impacto que cada cambio ha hecho en las organizaciones.
- Recuerda utilizar una plataforma digital que te permita incluir colores, tipografías y material gráfico que respalden y apoyen visualmente tu propuesta.
- Máximo 2 cuartillas.





El **big data** apoya en la resolución de dudas o inquietudes que una empresa pueda tener. Ya sea mediante el **análisis predictivo**, de **diagnóstico**, **descriptivo** o **prescriptivo** se pueden resolver muchas de las dudas que surgen en el día a día organizacional.



Cierre

Joshi, N. (2017). *Top 5 sources of big data*. Recuperado de:
<https://www.allerin.com/blog/top-5-sources-of-big-data>

Nosis Investigación y Desarrollo (2020). *¿Qué es big data y por qué es importante?* Recuperado de <https://blog.nosis.com/que-es-big-data/>

Sharma, R. (2021). *Sources of big data: where does it come from?* Recuperado de:
<https://www.upgrad.com/blog/sources-of-big-data/>



La obra presentada es propiedad de ENSEÑANZA E INVESTIGACIÓN SUPERIOR A.C. (UNIVERSIDAD TECMILENIO), protegida por la Ley Federal de Derecho de Autor; la alteración o deformación de una obra, así como su reproducción, exhibición o ejecución pública sin el consentimiento de su autor y titular de los derechos correspondientes es constitutivo de un delito tipificado en la Ley Federal de Derechos de Autor, así como en las Leyes Internacionales de Derecho de Autor.

El uso de imágenes, fragmentos de videos, fragmentos de eventos culturales, programas y demás material que sea objeto de protección de los derechos de autor, es exclusivamente para fines educativos e informativos, y cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por UNIVERSIDAD TECMILENIO.

Queda prohibido copiar, reproducir, distribuir, publicar, transmitir, difundir, o en cualquier modo explotar cualquier parte de esta obra sin la autorización previa por escrito de UNIVERSIDAD TECMILENIO. Sin embargo, usted podrá bajar material a su computadora personal para uso exclusivamente personal o educacional y no comercial limitado a una copia por página. No se podrá remover o alterar de la copia ninguna leyenda de Derechos de Autor o la que manifieste la autoría del material.

