

Big Data

Tema 6

Big Data



Big Data



Big Data



Big Data

- Índice:
 - Introducción.
 - ¿Qué es Big data? (Definición).
 - Características del Big Data.
 - Aplicaciones del Big data.
 - Fuentes del Big Data.
 - Beneficios y desafíos del Big Data.

Big Data



Introducción



Introducción

- El big data se está convirtiendo en un fenómeno que está cambiando las sociedades y los modelos de negocio de las empresas.
- Aunque no existe unanimidad en la definición de Big Data, sí que es patente la importancia que tiene hoy en día la posibilidad de capturar, almacenar y analizar grandes volúmenes de datos.

Big Data



¿Qué es Big Data?



¿Qué es Big Data?

- En general, el **Big Data** es un término utilizado para describir conjuntos de datos muy grandes y complejos que son difíciles de procesar y analizar utilizando técnicas y herramientas tradicionales.
- Los datos pueden ser de cualquier tipo (datos de transacciones comerciales o financieras, datos de redes sociales, datos de sensores, etc.).

<https://www.oracle.com/es/big-data/what-is-big-data/>

Big Data



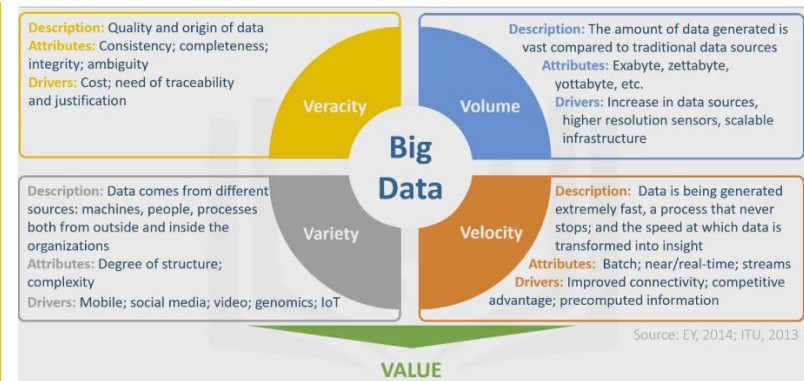
Características del Big Data



Características del Big Data

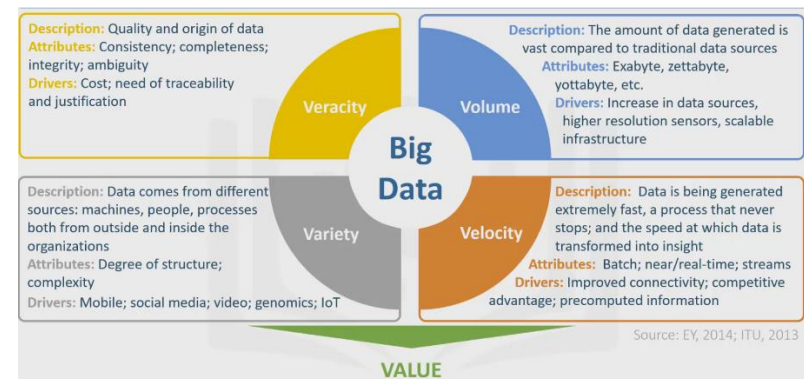
- Los conjuntos de datos de big data se caracterizan por tres atributos o características principales: volumen, variedad y velocidad (las tres V de big data).

- Las tres V de big data:
 - Volumen.
 - Variedad.
 - Velocidad.



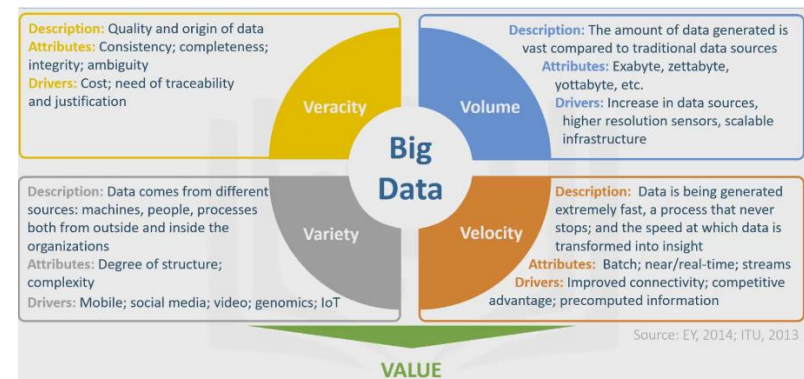
Características del Big Data

- Características del Big data:
 - **Volumen:** Los conjuntos de datos de big data son extremadamente grandes. Hablamos de conjuntos de datos que pueden ir desde decenas de terabytes hasta cientos de petabytes o exabytes.



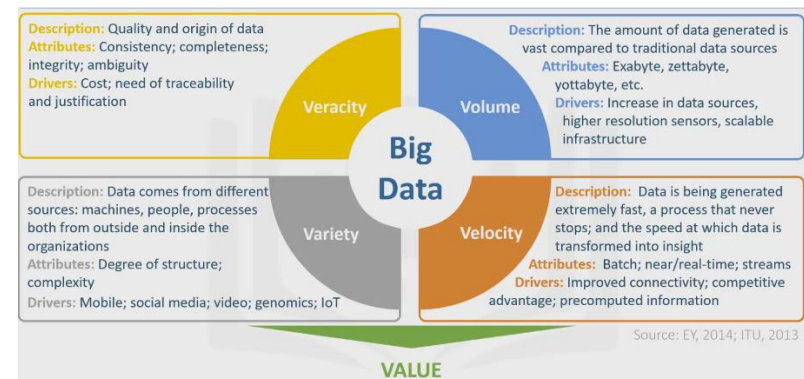
Características del Big Data

- Características del Big data:
 - **Variedad:** Los conjuntos de datos de big data pueden ser muy variados, con datos estructurados (como tablas de bases de datos) y datos no estructurados (como texto, imágenes o audio).



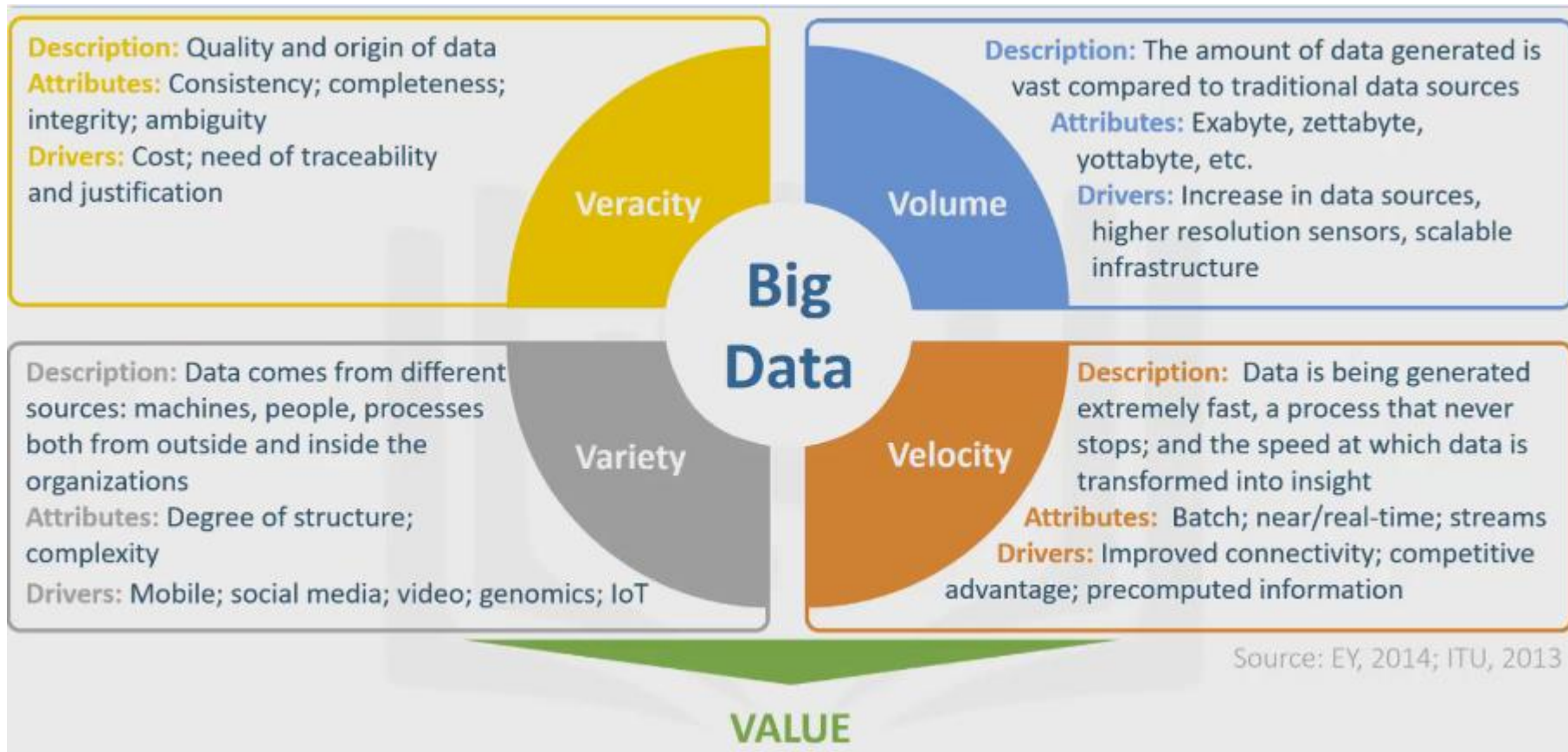
Características del Big Data

- Características del Big data:
 - **Velocidad:** Los conjuntos de datos de big data pueden crecer y cambiar muy rápidamente, lo que significa que es importante poder procesar y analizar los datos de manera eficiente y en tiempo real.



Características del Big Data

- Características del Big data:



Características del Big Data

- El análisis de big data permite a las instituciones, organizaciones gubernamentales y otras entidades tomar decisiones basadas en datos más completos y precisos.

Características del Big Data



Big Data (Tipos de datos)




- La información puede clasificarse en diferentes tipos de datos:
 - Estructurados
 - Tienen un formato definido claro. Por ejemplo, los datos incluidos en una base de datos o aquellos datos de una transacción en que cada campo tiene un claro significado.
 - No estructurados
 - Son datos en crudo que no tienen un formato específico. Estos datos no podemos volcarlos a una base de datos relacional porque no conocemos su tipología. Podemos incluir vídeo, imágenes y voz .
 - Semiestructurados
 - Dato que no se limitan a campos determinados, pues contienen marcadores para separar las diferentes secciones. Son datos que tienen su propio esquema (metadato), ejemplos HTML, XML, Json. También podemos incluir en esta categoría los logs que son una fuente muy importante de información en Big Data.

Big Data (Tipos de datos)

Tipos de datos






Datos estructurados

-  Fichas estandarizadas
-  Formularios web
-  Hojas de cálculo



Datos semiestructurados

-  Lenguaje de marcado XML
-  Bases de datos NoSQL
-  Código HTML



Datos no estructurados

-  Audios
-  Imágenes
-  Vídeos

Big Data

- Para medir la información que generan los miles de millones de usuarios de internet se utilizan unidades de medida de la información del orden de **petabytes** y **exabytes**.
- Todos estos datos se almacenan en los servidores situados en los denominados **centros de datos**. Estas infraestructuras consumen mucha energía y provocan un gran impacto medioambiental.
- El transporte de esta ingente cantidad de información se realiza mediante cables de fibra óptica que cruzan océanos y mares.

Big Data



Fuentes del Big Data



Fuentes del Big Data

- El big data se sustenta en:
 - La capacidad que tienen nuestros dispositivos móviles de generar y captar datos.
 - Las posibilidades que ofrece el internet de las cosas (IoT).
 - El poder actual de las redes sociales.
 - La disponibilidad cada vez mayor de fuentes de datos abiertos (open data).

Big Data



Fuentes del Big Data

El Internet de las cosas (IoT) y otros dispositivos conectados a Internet



Fuentes del Big Data

- El Internet de las cosas (IoT) y otros dispositivos conectados a Internet:
 - En la actividad diaria generamos multitud de datos cuando usamos los dispositivos e Internet. Todos estos datos pueden ser analizados y transformados en información útil para resolver situaciones de nuestra vida.
 - Los dispositivos móviles, gracias a los sensores que llevan incorporados, son los dispositivos capaces de recoger más información. Estos datos se incorporan a los sistemas de información desde los vestidos, los coches inteligentes, los dispositivos médicos, los equipos industriales, etc.
 - Los **weareables** son dispositivos que llevamos encima y que se encargan de recoger datos de nuestra actividad.

Fuentes del Big Data

- El Internet de las cosas (IoT) y otros dispositivos conectados a Internet:
 - El uso del Internet de las cosas y el tratamiento masivo de datos se puede utilizar para conseguir ciudades más eficientes y sostenibles o una gestión y control de la salud de las personas eficaz.

Big Data



Fuentes del Big Data

Las redes sociales



Fuentes del Big Data

- Las redes sociales:
 - Otra fuente de generación de datos son los datos personales provenientes de las interacciones en redes sociales (Facebook, YouTube, Instagram, etc.).
 - Esto incluye datos en forma de imágenes, vídeos, voz, texto y sonido. Suelen ser datos no estructurados o semiestructurados, por lo que plantean un desafío único para su consumo y análisis, ya que para ser de algún valor, los datos deben poder utilizarse. Para ello, hay que disponer de datos limpios (datos relevantes para el cliente y organizados de tal modo que permitan un análisis significativo), cuya preparación que requiere una gran cantidad de trabajo.
 - Nota: Los científicos de datos dedican entre un 50 y un 80 por ciento de su tiempo a seleccionar y preparar los datos antes de que estos puedan utilizarse.

Fuentes del Big Data

- Las redes sociales:
- (Oportunidades y riesgos del tratamiento masivo de datos personales).
 - Hay que ser conscientes del poder actual de las redes sociales, de la cantidad de datos personales que recogen a través de nuestras acciones o cesiones (huella digital), de sus potenciales beneficios para crear una sociedad mejor, pero también de las amenazas que se derivan de su uso incorrecto.
 - Las empresas usan Big Data para analizar las huellas digitales y así ofrecer productos y servicios más personalizados, así como para tomar decisiones empresariales informadas.
 - A pesar de sus beneficios, la huella digital también presenta riesgos para nuestra privacidad. Por eso, es necesario adoptar buenas prácticas para proteger nuestra información personal (utilizar contraseñas seguras, ser cauteloso con la información que se comparte en línea y entender las configuraciones de seguridad de las aplicaciones).

Big Data



Fuentes del Big Data

Los datos abiertos (open data)



Fuentes del Big Data

- Los datos abiertos (open data):
 - Los datos disponibles públicamente en cantidades masivas en fuentes de datos abiertos (open data) permiten que los individuos puedan estar mejor informados y utilizar dichos datos para proponer proyectos innovadores y sostenibles basados en su explotación.
 - Estos datos suelen ser suministrados por las administraciones públicas y existen portales web que se encargan de visualizar estos datos abiertos de forma gráfica.
 - Algunos ejemplos son el data.gov del gobierno de los Estados Unidos, el [WorldFactbook](http://www.worldfactbook.com) de la CIA o el [Open Data Portal](http://open.data.europa.eu) de la Unión Europea.
 - Los datos abiertos sirven a usuarios, desarrolladores y empresas para generar aplicaciones y servicios para el conjunto de la sociedad.

Big Data



Beneficios y desafíos del Big Data



Beneficios y desafíos del Big Data

- Lógicamente, la utilización del tratamiento masivo de datos, gestionados de manera abierta o privativa, lleva aparejada una serie de oportunidades y riesgos que hay que tener en cuenta.

Beneficios y desafíos del Big Data

- Beneficios del big data:
 - El análisis de big data permite a las instituciones, organizaciones gubernamentales y otras entidades tomar decisiones basadas en datos más completos y precisos.
 - El big data permite obtener respuestas más completas, ya que dispone de mayor cantidad de información.

Beneficios y desafíos del Big Data

- **Desafíos de big data:**

Si bien es cierto que el big data promete mucho, también se enfrenta a desafíos:

- El almacenamiento de los datos.
- La preparación de los datos.
- Rápida evolución de las tecnologías de big data.

Beneficios y desafíos del Big Data

- Desafíos de big data
 - **El almacenamiento de los datos:** El big data se caracteriza por su gran tamaño. El volumen de datos duplica su tamaño cada dos años aproximadamente, y, aunque se han desarrollado nuevas tecnologías para el almacenamiento de datos, las organizaciones han de seguir esforzándose por mantener el ritmo de crecimiento de sus datos y por encontrar formas de almacenarlos eficazmente.
 - La preparación de los datos.
 - Rápida evolución de las tecnologías de big data.

Beneficios y desafíos del Big Data

- Desafíos de big data
 - El almacenamiento de los datos.
 - **La preparación de los datos:** No basta con almacenar los datos. Para ser de algún valor, los datos deben poder utilizarse. Disponer de datos limpios (datos relevantes para el cliente y organizados de tal modo que permitan un análisis significativo) requiere una gran cantidad de trabajo.
 - Los científicos de datos dedican entre un 50 y un 80 por ciento de su tiempo a seleccionar y preparar los datos antes de que estos puedan utilizarse.
 - Rápida evolución de las tecnologías de big data.

Beneficios y desafíos del Big Data

- Desafíos de big data
 - El almacenamiento de los datos.
 - La preparación de los datos.
 - **Rápida evolución de las tecnologías de big data:**
La tecnología de big data cambia a un ritmo muy rápido, por lo que mantenerse al día en tecnología de big data supone un desafío constante.

Beneficios y desafíos del Big Data

- Beneficios y desafíos del big data:

Ventajas

- Tomar mejores decisiones
- Mayor rapidez en la toma de decisiones
- Mayor fidelidad de los clientes

Inconvenientes

- Problemas relacionados con el almacenamiento de datos
- Costes de infraestructura y almacenamiento
- Calidad y confiabilidad de los datos
- Privacidad y seguridad de los datos personales
- Brecha digital y desigualdad en el acceso a la información

Big Data



Aplicaciones del Big Data



Aplicaciones del Big Data

- Aplicaciones del big data:
 - Servicios financieros.
 - Cuidado de la salud.
 - Fabricación.
 - Comercio minorista.
 - Telecomunicaciones.
 - Etc.

Aplicaciones del Big Data

Business use cases

There are hundreds of ways big data can give businesses a competitive advantage. Here, we'll explore just a few examples of how industries have used big data to go beyond measuring and counting to being able to predict and understand.



Financial services

Whether it's capturing new market opportunities or reducing fraud, financial services organizations have converted big data into a competitive advantage. With big data, financial institutions can

- Identify patterns that indicate fraud and streamline regulatory reporting
- Gain a better understanding of market trends and customer needs, which can improve decision-making about new products and services
- Detect potential fraud patterns and adhere to regulations



Healthcare

Hospitals, healthcare companies, and researchers produce massive amounts of data. With big data, healthcare professionals can

- Identify disease genes and biomarkers to help patients pinpoint health issues they may face in the future
- Provide better treatment and improve the quality of care without increasing costs
- Detect potential insurance fraud by flagging certain behaviors for further examination



Manufacturing

The digital revolution has empowered manufacturers to find ways to harness all the data they generate. With big data, manufacturers can

- Predict equipment failures
- Assess production processes, proactively respond to customer feedback, and anticipate future demands
- Better understand the flow of production lines and determine the cause of delays



Retail

Big data is used across all stages of the retail process. With big data, retailers can

- Predict customer demand and launch new products
- Identify a company's most loyal customers and target them with special offers
- Use predictive technology to keep shelves stocked and avoid supply chain disruptions



Telecommunications

The popularity of smartphones and other mobile devices has given telecommunications companies tremendous growth opportunities and an ever-expanding volume of data. Big data lets telecommunications companies

- Identify areas with excess capacity and reroute bandwidth as needed
- Predict overall customer satisfaction by analyzing the data they already have about service quality and convenience
- Improve understanding of customer behavior to design new products and features

[Download the Big Data Use Cases Ebook](#)

Big Data



Big data, inteligencia artificial (IA) y aprendizaje automático (ML)



Big data, inteligencia artificial y aprendizaje automático

- Big data, inteligencia artificial y aprendizaje automático
 - La inteligencia artificial (IA) y el aprendizaje automático (ML) son los próximos disruptores o alteradores en la tecnología de big data.
 - Con IA y ML, las computadoras pueden reconocer el contenido de imágenes, traducir lo hablado en un determinado lenguaje, leer textos y comprender el significado de las respuestas de las redes sociales. Donde las plataformas de datos eran al principio herramientas para, simplemente recopilar datos de todo tipo, con AI y ML ahora pueden comprender los datos almacenados y usarlos para iniciar acciones o apoyar decisiones.
 - La integración de AI y ML en plataformas de datos ha permitido muchos nuevos casos de uso que nunca antes se habían podido lograr.

Big Data

- Enlaces:

- Inteligencia artificial y datos abiertos (datos.gob.es)
 - <https://datos.gob.es/es/blog/inteligencia-artificial-y-datos-abiertos-0>
- INTEF (Aula en abierto)
 - <https://formacion.intef.es/aulaenabierto/course/view.php?id=85>
 - <https://formacion.intef.es/aulaenabierto/mod/book/view.php?id=5079>
 - <https://formacion.intef.es/aulaenabierto/course/index.php?categoryid=12>
- <https://www.ufv.es/cuales-son-las-principales-desventajas-del-big-data-preguntas-grados/>
- <https://retos-operaciones-logistica.eae.es/empresas-con-big-data-oportunidad-y-riesgos/>
- <https://www.edsrobotics.com/blog/big-data-que-es-ventajas-desventajas/>
- <https://www.santanderopenacademy.com/es/blog/principales-aplicaciones-big-data.html>

Big Data

Fin

Muchas gracias por su atención