

## TP4- Big Data

### 1) Describa con sus palabras el concepto de Big Data

Big Data es un conjunto de datos cuyo tipo o tamaño supera la capacidad de una base de datos relacional para gestionar los datos.

Algunas características que incluyen son:

- El alto volumen
- La gran velocidad
- Gran variedad

Las fuentes de datos son más complejas, ya que, son impulsadas por IA redes sociales, IoT, etc.

### 2) Enumerar ejemplos de fuentes de datos que generen gran cantidad de información

Algunos ejemplos de fuentes de datos son:

- **Redes Sociales:** Plataformas como Facebook, Twitter, Instagram, LinkedIn generan enormes cantidades de datos a través de publicaciones, interacciones, comentarios y perfiles de usuarios.
- **Sensores IoT:** Dispositivos conectados a Internet como medidores inteligentes, cámaras de vigilancia, sensores ambientales y dispositivos de salud generan datos constantemente.
- **Aplicaciones y Sitios Web:** Las aplicaciones y sitios web recopilan datos de usuarios, como comportamientos de navegación, transacciones, búsquedas y registros de actividad.
- **Datos Geoespaciales:** Sistemas de posicionamiento global (GPS), imágenes satelitales y sistemas de información geográfica (SIG) generan grandes cantidades de datos sobre ubicaciones y geografía.
- **Archivos de Internet:** Los motores de búsqueda y repositorios de datos como Wikipedia almacenan información significativa y constantemente actualizada.

- 3) Mencione algunas tecnologías y herramientas comunes para gestionar y analizar Big Data.

Algunos ejemplos de herramientas para gestionar datos son:

- **Hadoop**: es una herramienta open source que permite gestionar los grandes volúmenes de datos, analizarlos y procesarlos. Hadoop implementa MapReduce, un modelo de programación que da soporte a la computación paralela sobre grandes colecciones de datos.
- **NoSQL**: son sistemas que no usan SQL como lenguaje de consultas, lo que, aunque no puede garantizar la integridad de los datos, les permite obtener ganancias significativas en escalabilidad y rendimiento al trabajar con Big Data. Una de las bases de datos NoSQL más populares es MongoDB.
- **Spark**: es un framework de computación que permite procesar los datos de forma rápida. Permite escribir aplicaciones en Java, Scala, Python, R y SQL y funciona tanto sobre Hadoop o Apache como de forma independiente o en la nube. Te permite acceder a centenares de fuentes de datos.

- 4) ¿Cuáles son algunos ejemplos de aplicaciones del Big Data en la vida cotidiana?

Algunos ejemplos del uso de Big Data son:

- **Compras en Línea**: Sitios como Amazon, eBay y Alibaba utilizan Big Data para recomendaciones de productos basadas en el historial de compras y comportamientos de navegación de los usuarios.
- **Servicios de Streaming**: Plataformas como Netflix, Spotify y YouTube utilizan análisis de Big Data para personalizar recomendaciones de contenido y mejorar la experiencia del usuario.
- **Servicios de Transporte**: Aplicaciones como Uber y Didi utilizan Big Data para optimizar rutas, estimar tiempos de llegada y calcular tarifas dinámicas basadas en la demanda del usuario.
- **Asistentes Virtuales**: Tecnologías como Siri de Apple, Google Assistant y Amazon Alexa utilizan Big Data para comprender mejor los comandos de voz, responder preguntas y mejorar la precisión en las respuestas.

