

Sitio: **Formación Profesional a Distancia**

Curso: **36018173. CSIFC90. MP5075. Big data aplicado. 2023-2024. (Grupo A) (132_36018173_ZSIFC90_MP5075_A)**

Glosario: **Glosario Big Data**

B

Batch

Estrategia de procesamiento en la que se acumulan grupos de datos y se procesan en diferido,

Big Data

- Todo aquello que esté relacionado con la gestión del almacenamiento y uso de una cantidad tan grande de datos que no puede ser realizada por sistemas y métodos tradicionales.

C

Cluster

Conjunto hardware y software que permite nodos interconectados por red que funcionan proporcionando la suma de sus recursos como si fueran una única máquina.

Commodity hardware

Son equipos domésticos aunque potentes que se usan como nodos en clusters de big data.

D

Dato estructurado

Son aquellos datos compuestos por campos fijos y bien definidos. Ejemplos típicos son los que podemos encontrar como tablas en una base de datos relacional tradicional como "Clientes", "Productos", "Movimientos", "Almacén"

En cliente tendremos campos como "Nombre", "Apellido", "Dirección", "Fecha_nacimiento"

Los datos estructurados forman tablas rectangulares sin muchos huecos.

Dato no estructurado

Son aquellos datos que tienen una estructura interna nativa y que no se corresponde con los campos tradicionales de texto, número, fecha o booleano. Un ejemplo es un audio, un vídeo o, en general, cualquier archivo multimedia.

E

Escalado horizontal

- Consiste en sumar las capacidades de almacenamiento y proceso de un conjunto de máquinas llamadas nodos como una única máquina lógica.

Escalado vertical

- Ampliación hardware de un equipo informático mediante nuevos y más potentes componentes.

M

Metadatos

Es toda aquella información que acompaña al dato y tiene como misión describir el contexto en el que se generó ese dato. Aquí encontraremos información como la fecha de creación o el usuario propietario pero también otra específica del dato.

Por ejemplo, en un archivo de imagen podemos encontrar como metadato la marca y modelo de la cámara que se usó para sacar esa foto. En peticiones web podemos encontrar metadatos relacionados con el navegador que ha usado el cliente.

O

OLTP

Procesamiento de transacciones on line es el conjunto de tecnologías hardware y software enfocadas a tratar movimientos típicos de las bases de datos como altas, bajas, modificaciones y lecturas.

S

Streaming

Estrategia de procesamiento consistente en procesar desde memoria fragmentos de un flujo a medida que van llegando.

ho

T

Tiempo real

Estrategia de procesamiento que procesa transacciones tan pronto como se producen dando urgentemente una respuesta que no siempre tiene por que exacta en términos de actualización. Para conseguir más rapidez se suele trabajar en memoria.

Transaccional

Estrategia de procesamiento en la que se procesan las transacciones lo más rápido posible. Un ejemplo de transacción es una operación con un registro de la base de datos. Queremos que una actualización de un registro se haga rápido y esté disponible para todos inmediatamente después.