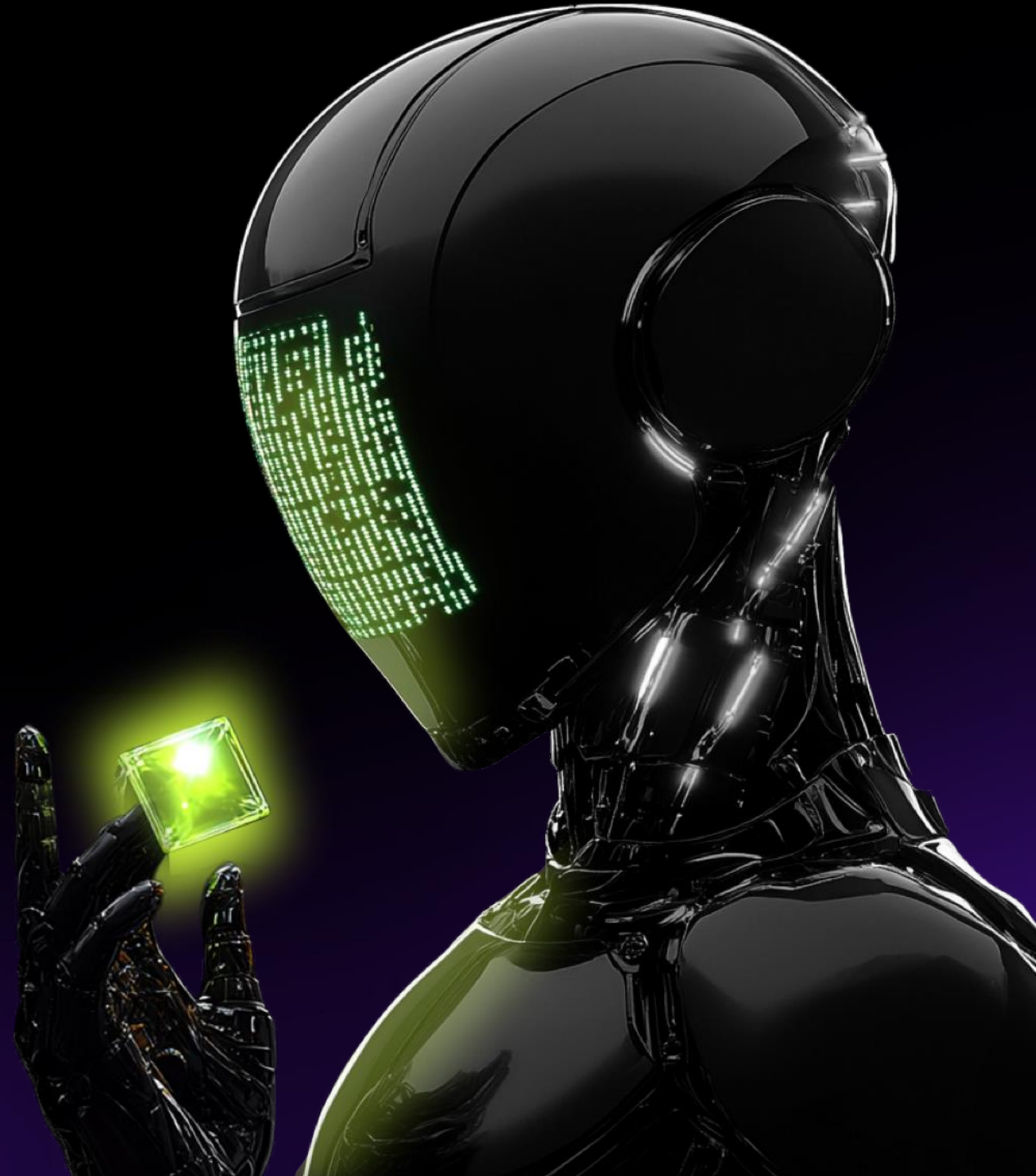


CODING UP MY FUTURE

Bootcamp – Databases

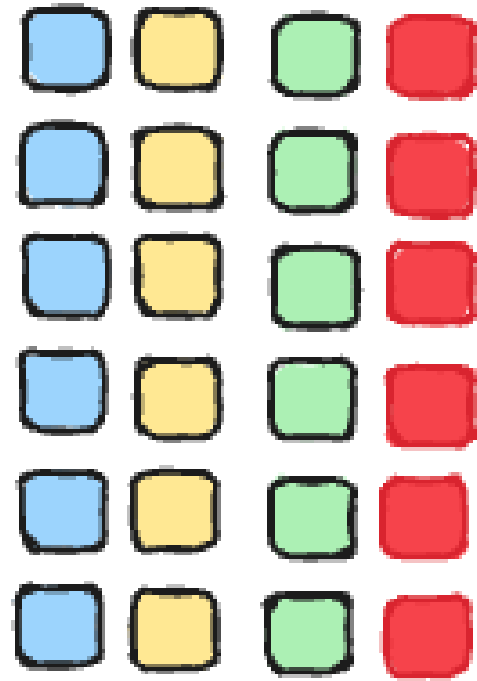
softserve



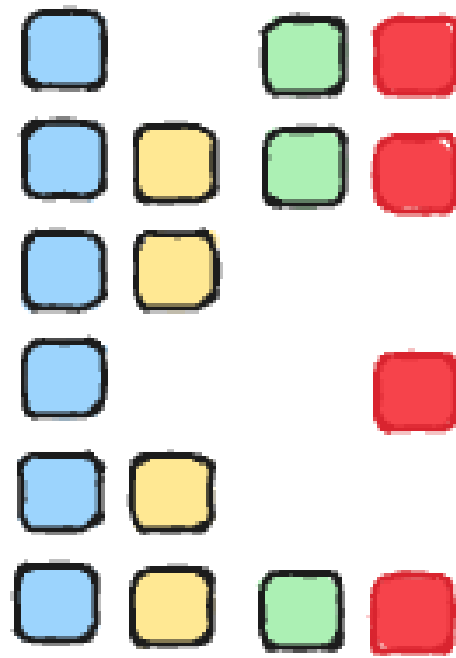


Si ves el ícono de DuckDB, significa que es un contenido que forma parte del curso y debe ser **APRENDIDO**.

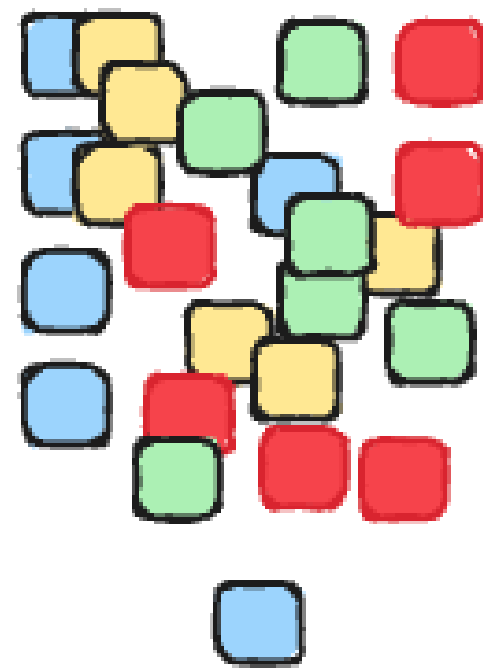
CONCEPTOS ELEMENTALES



Estructurados



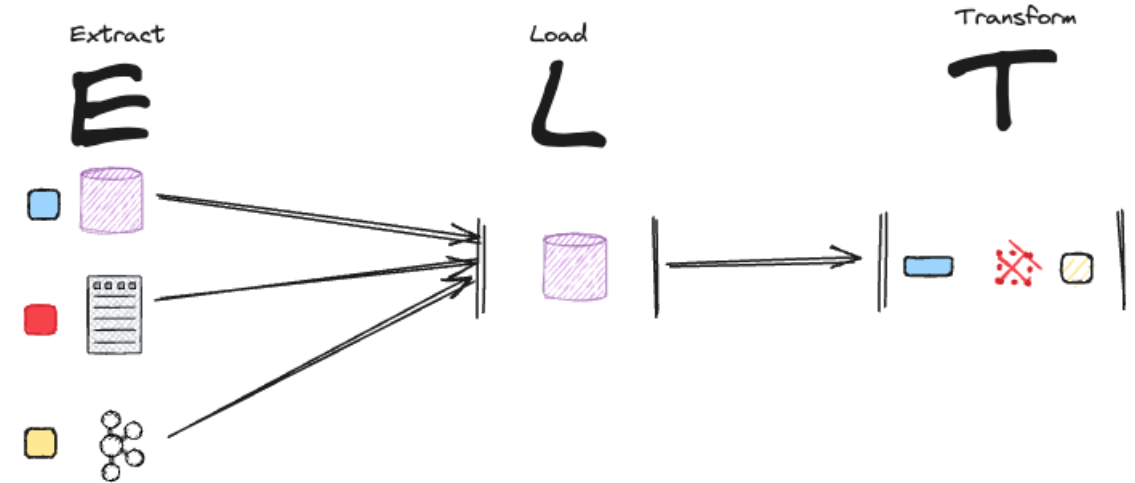
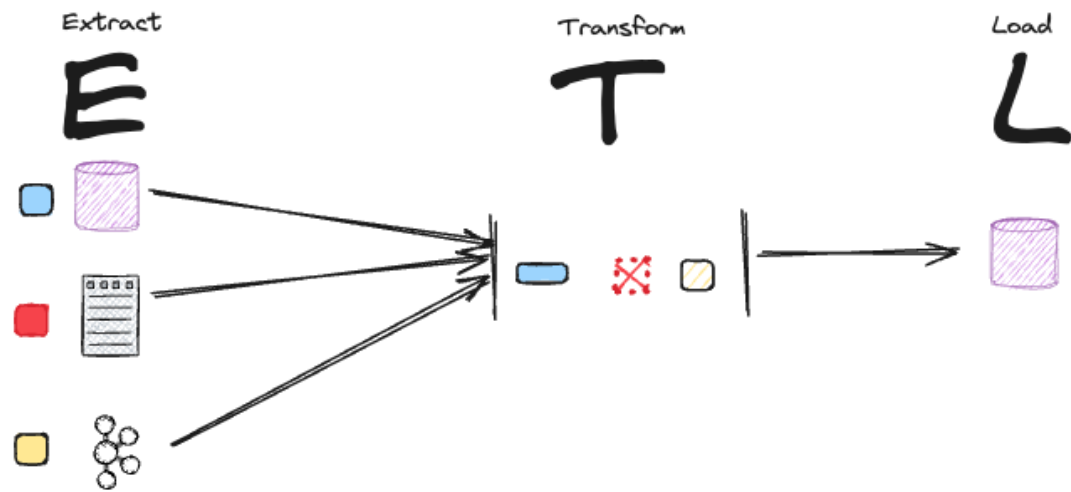
Semi-estructurados



No estructurados



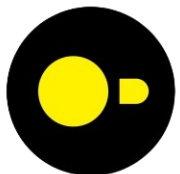
Esquema / Integridad de datos



	ID	NOMBRE	DOC	
✗	111	Pepito	CC	✗
✓	222	Julanito	CC	✗
✗	111	Pepito	CC	✗

Redundancia

softserve



¿CÓMO MANEJAMOS
NORMALMENTE LOS
DATOS?

1 Excel



COLUMNA 1	Columna nro 2	columna # 3	COLUMNA_4
demo	6.317.455	C	Y
50	8.238.761	A	FALSE
22	3.619.942	A	FALSE
31	1.687.379	D	NO
63	6.354.473	D	FALSE
1	8.973.737	B	FALSO
42	3.595.350	B	TRUE
35	6.249.024	D	VERDADERO
3	6.399.078	C	FALSE
boocamp	9.300.696	A	FALSE

2 CSV

```
COLUMNA 1;Columna nro 2;columna # 3;COLUMNA_4
demo;6.317.455;C;Y
50;8.238.761;A;FALSE
22;3.619.942;A;FALSE
31;1.687.379;D;NO
63;6.354.473;D;FALSE
1;8.973.737;B;FALSO
42;3.595.350;B;TRUE
35;6.249.024;D;VERDADERO
3;6.399.078;C;FALSE
boocamp;9.300.696;A;FALSE
```

3 txt

```
COLUMNA 1  Columna nro 2  columna # 3  COLUMNA_4
demo      6.317.455      C      Y
50  8.238.761      A      FALSE
22  3.619.942      A      FALSE
31  1.687.379      D      NO
63  6.354.473      D      FALSE
1   8.973.737      B      FALSO
42  3.595.350      B      TRUE
35  6.249.024      D      VERDADERO
3   6.399.078      C      FALSE
boocamp 9.300.696      A      FALSE
```

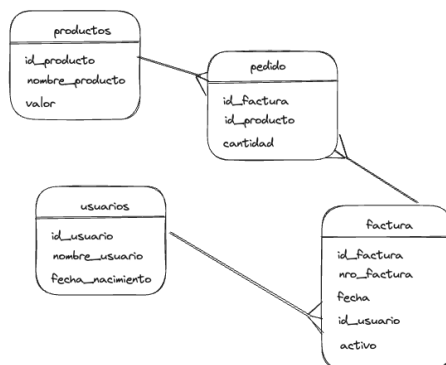
TIPOS DE BASES DE DATOS

BASE DE DATOS RELACIONALES (RDBMS - OLTP)

softserve



Estructurados



Breve descripción

Almacenan datos en tablas con filas y columnas, siguiendo un esquema rígido. Usan SQL para consultas.



Atomicidad (Atomicity): Las transacciones son indivisibles; se completan en su totalidad o no se realizan en absoluto.



Consistencia (Consistency): Garantiza que solo se guardan datos válidos, manteniendo la integridad del sistema.



Aislamiento (Isolation): Cada transacción se ejecuta de manera independiente, sin interferir con otras.



Durabilidad (Durability): Una vez que una transacción se confirma, los datos se guardan de manera permanente, incluso ante fallos del sistema.

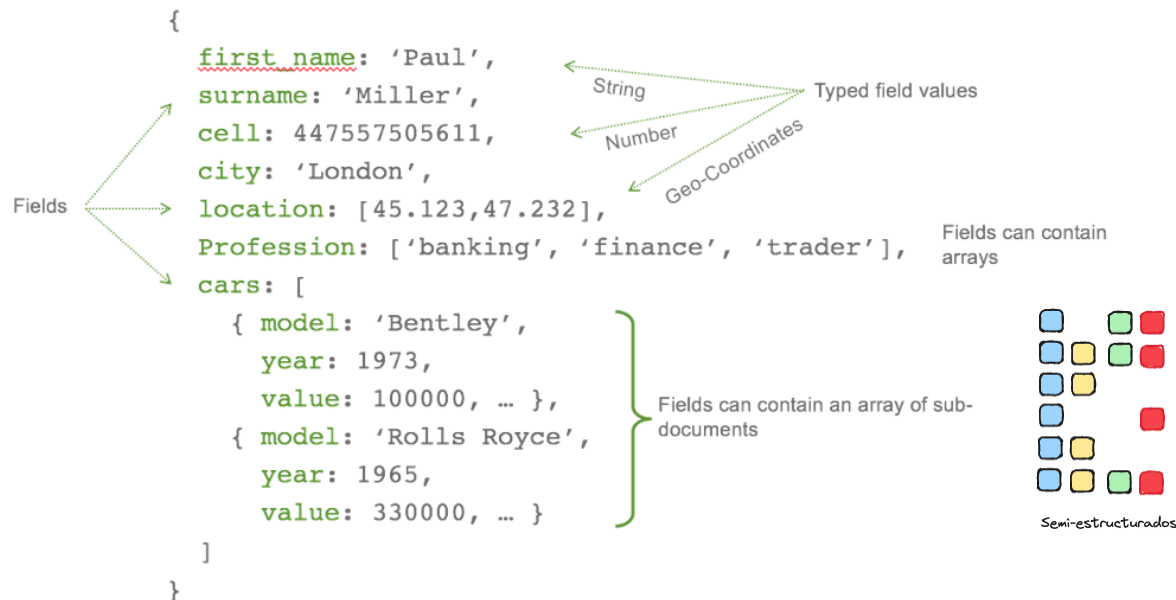
Usabilidad

Perfectas para sistemas que necesitan integridad de datos, transacciones seguras, y relaciones estructuradas, como bancos, ERPs, y aplicaciones financieras, etc.

softserve

BASE DE DATOS NoSQL

softserve



Breve descripción

Base de datos que no usa el modelo relacional de tablas. Permite esquemas flexibles y puede manejar grandes volúmenes de datos.

Usabilidad

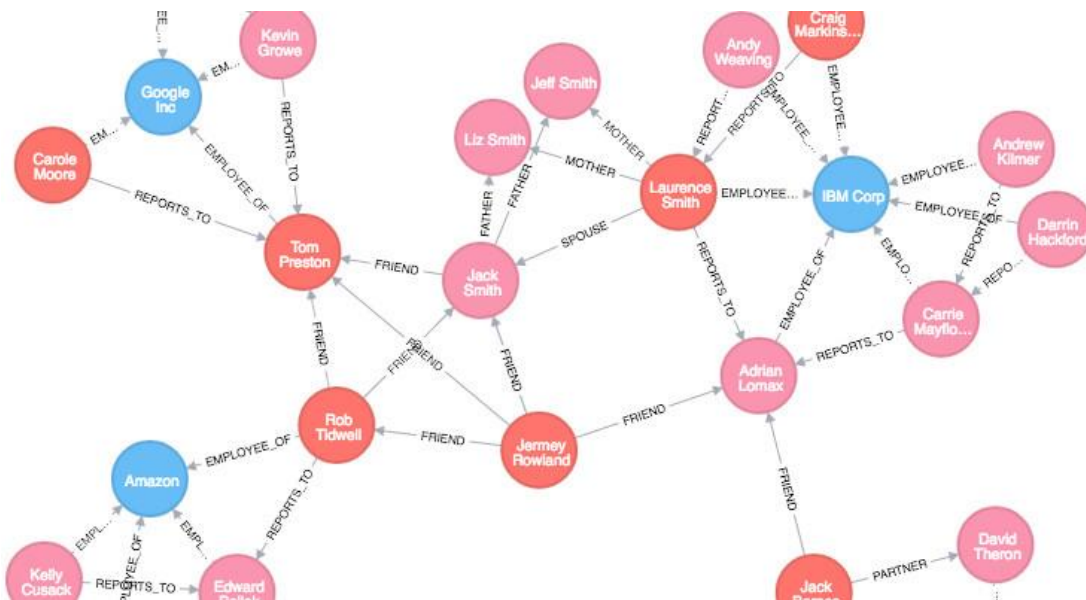
Perfecta para aplicaciones que manejan grandes volúmenes de datos no estructurados o semi-estructurados.



softserve

BASE DE DATOS DE GRAFOS

softserve



Breve descripción

Estructura de datos basada en nodos y aristas para representar relaciones entre entidades.

Usabilidad

Ideal para aplicaciones que necesitan modelar y consultar relaciones complejas entre datos, como redes sociales o sistemas de recomendaciones.



softserve

BASE DE DATOS EN MEMORIA

softserve



```
# 15 €/month DigitalOcean Droplet 3 vCPU 1 GB 15 €/month
# Very fast transactions Very Fast Transaction
$ ./redis-cli REDISQL.CREATE_STATEMENT DB insert
  "INSERT INTO simple VALUES(?1, ?2, ?3);"

$ ./redis-benchmark REDISQL.EXEC_STATEMENT
  DB insert 1 2 3:

      80234.52 requests per second

$ ./redis-cli REDISQL.CREATE_STATEMENT DB insert_multiple
  "INSERT INTO simple VALUES(?1, ?2, ?3), (?4, ?5, ?6);"
$ ./redis-benchmark REDISQL.EXEC_STATEMENT
  DB insert_multiple 1 2 3 4 5 6

      80215.21 requests per second
```

Breve descripción

Almacena datos en la memoria RAM para un acceso y procesamiento extremadamente rápido.

Usabilidad

Ideal para aplicaciones que requieren tiempos de respuesta en milisegundos y operaciones de alta velocidad, como cachés y sistemas de análisis en tiempo real.



softserve

BASE DE DATOS COLUMNARES

softserve

Breve descripción

Almacena datos en columnas en lugar de filas, optimizando las consultas analíticas.

Usabilidad

Perfecta para análisis de grandes volúmenes de datos y consultas de agregación.

ID	Last name	First name	Bonus
53666	Smith	James	8000
50333	Jones	Sam	5000
54673	Taylor	Ann	4000
58930	Burton	Sue	9000

Row-oriented

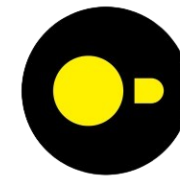
53666	Smith	James	8000
50333	Jones	Sam	5000
54673	Taylor	Ann	4000
58930	Burton	Sue	9000

Column-oriented

53666	50333	54673	58930
Smith	Jones	Taylor	Burton
James	Sam	Ann	Sue
8000	5000	4000	9000

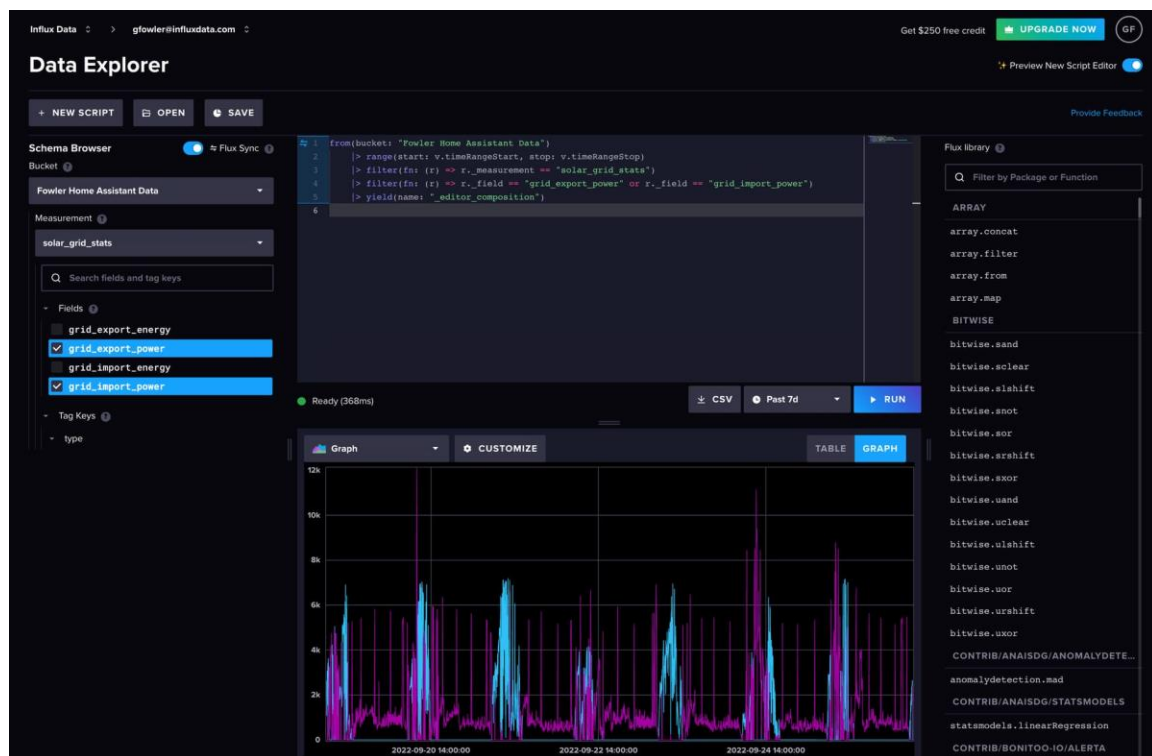


softserve



BASE DE DATOS DE SERIES DE TIEMPO

softserve



Breve descripción

Optimizada para almacenar y consultar datos que están indexados por tiempo.

Usabilidad

Perfecta para el análisis de datos temporales como registros de monitoreo, datos financieros, y datos de sensores IoT.



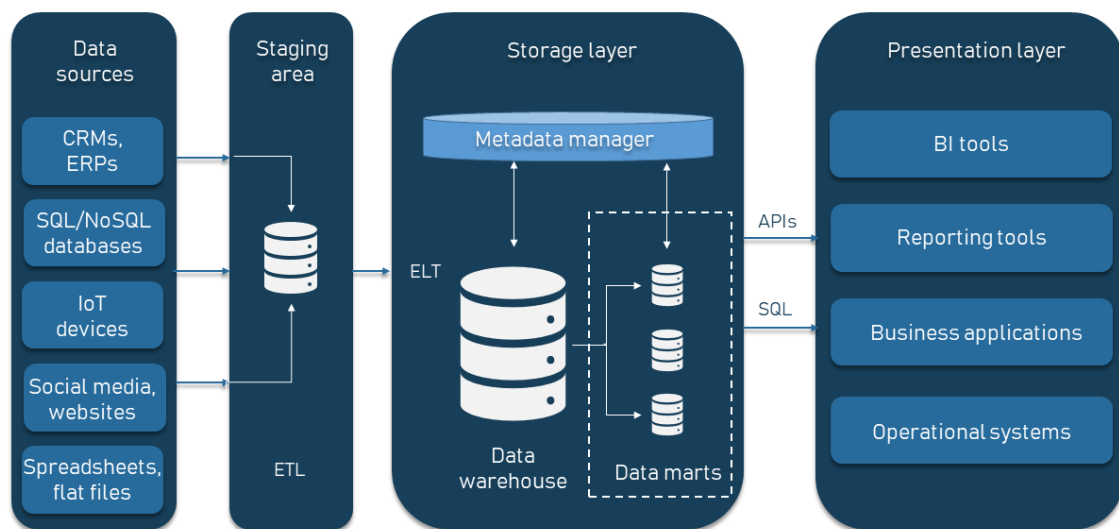
Prometheus

softserve

BASE DE DATOS TIPO WAREHOUSE (OLAP)



ENTERPRISE DATA WAREHOUSE COMPONENTS



Breve descripción

Base de datos centralizada que agrupa datos de diversas fuentes para análisis y reporting.

Usabilidad

Ideal para el análisis de grandes volúmenes de datos históricos y la realización de informes multidimensionales en procesos de BI.



BASE DE DATOS TIPO DATA LAKE

softserve



Breve descripción

Almacén centralizado que guarda datos en su formato bruto, ya sea estructurado, semi-estructurado o no estructurado.

Usabilidad

Perfecto para almacenar grandes volúmenes de datos de diferentes fuentes para su análisis posterior, permitiendo flexibilidad en el tipo de datos almacenados.

softserve



Amazon
S3



Azure Data Lake Storage



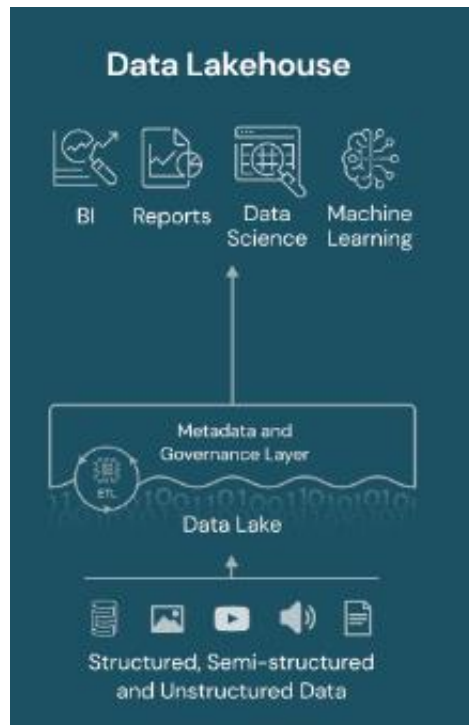
Google Cloud Storage

Microsoft Azure
Blob Storage



BASE DE DATOS TIPO LAKEHOUSE

softserve



Breve descripción

Extensión del Data Lake que proporciona transacciones ACID y manejo de versiones para datos.

Usabilidad

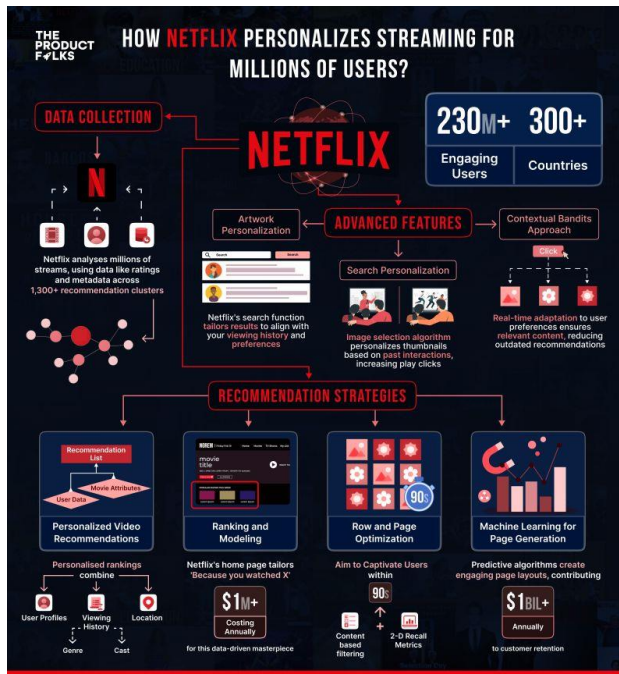
Ideal para gestionar grandes volúmenes de datos con requisitos de consistencia y actualizaciones, permitiendo transacciones fiables y manejo de versiones de datos.



databricks

softserve

BASE DE DATOS TIPO TIEMPO REAL ^{softserve}



Breve descripción

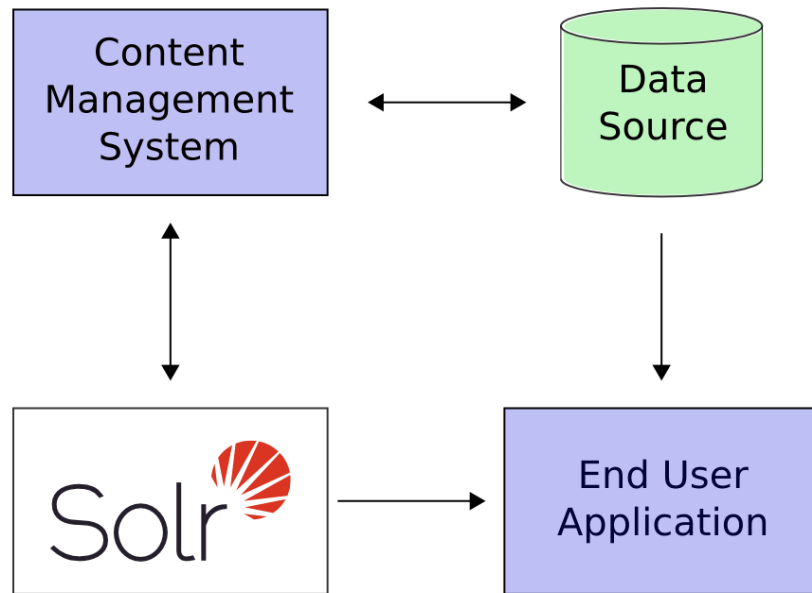
Diseñada para el procesamiento y análisis de datos en tiempo real con baja latencia.

Usabilidad

Perfecta para aplicaciones que requieren procesamiento instantáneo de datos y toma de decisiones en vivo, como detección de fraudes y análisis de eventos en tiempo real.

BASE DE DATOS DE BÚSQUEDA

softserve



Breve descripción

Optimizadas para realizar búsquedas rápidas y análisis de texto completo en grandes volúmenes de datos.

Usabilidad

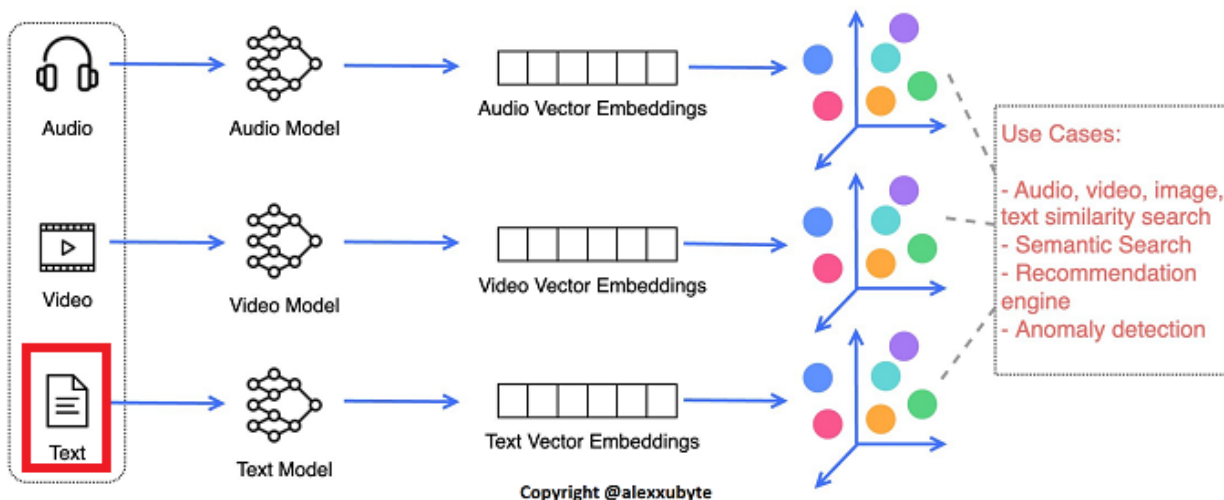
Ideal para aplicaciones que requieren búsquedas eficientes y análisis avanzado de texto, como motores de búsqueda y sistemas de recomendación.



softserve

BASE DE DATOS VECTORIALES

softserve



Breve descripción

Están diseñadas para almacenar y buscar vectores de alta dimensión, que son representaciones numéricas utilizadas en modelos de machine learning y redes neuronales.

Usabilidad

Son ideales para proyectos que involucran búsqueda de similitud en grandes cantidades de datos, como sistemas de recomendación, búsqueda semántica, procesamiento de imágenes y NLP, donde es crucial realizar consultas rápidas y precisas en tiempo real.

softserve



Chroma



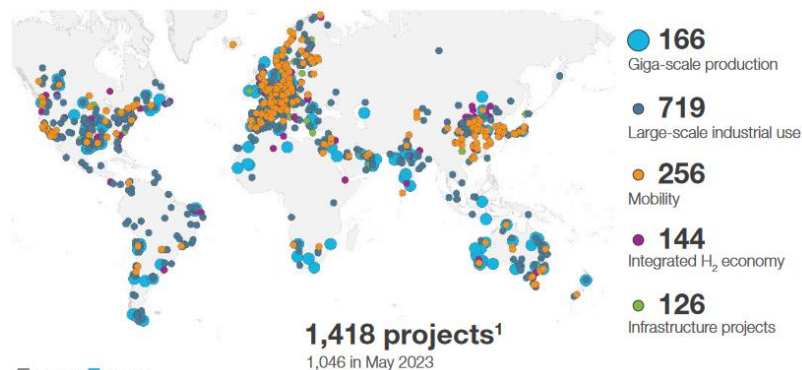
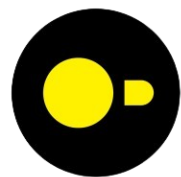
Weaviate



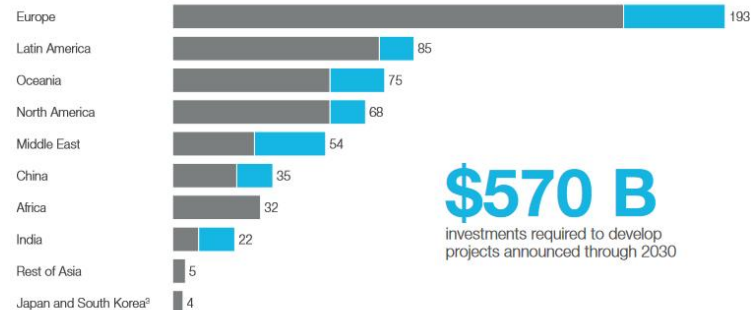
vespa

BASE DE DATOS ESPACIALES

softserve



■ Jan 2023² ■ Oct 2023



\$570 B
investments required to develop
projects announced through 2030

1. Project announcements below 1 MW excluded; includes 7 projects without specified type
2. Jan 2023 values have been updated to most recent capex estimations to keep values comparable
3. Restatement of Jan 2023 data for Japan and South Korea prevents comparison to Oct 2023 data
Source: Project & Investment tracker, as of Oct 2023

Breve descripción

Base de datos diseñada para almacenar y consultar datos geoespaciales y mapas.

Usabilidad

Perfecta para aplicaciones que necesitan gestionar y analizar datos geográficos, como sistemas de información geográfica (GIS), navegación, y análisis espacial.

MySQL[®]
Spatial



softserve

TIPOS DE MOTORES DE BASE DE DATOS RELACIONALES (SGBD)



Breve descripción

Es un motor de base de datos altamente escalable y utilizado en grandes empresas para manejar grandes volúmenes de datos.

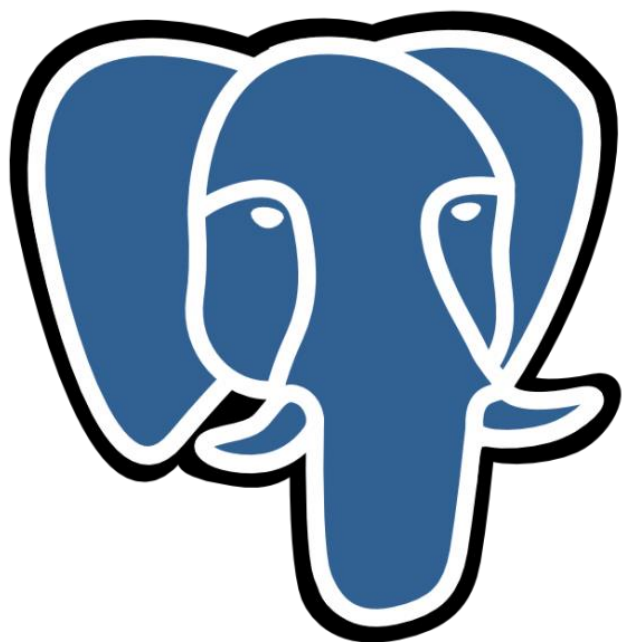
Precio: Pago por licencia. También tiene versiones gratuitas como Oracle Express Edition (XE).

Características

- Soporte para múltiples modelos de datos (relacional, NoSQL, JSON).
- Escalabilidad y alta disponibilidad.
- Funciones avanzadas de seguridad y particionamiento.

Cuando Usarla

En empresas que requieren alta disponibilidad y manejo de grandes volúmenes de transacciones críticas.



PostgreSQL

Breve descripción

Es una base de datos de código abierto conocida por su robustez y soporte a características avanzadas.

Precio: Gratuita (Open-source).

Características

- Soporte para SQL y NoSQL (JSON, XML) Escalabilidad y alta disponibilidad
- Extensibilidad y alta conformidad con los estándares SQL
- Seguridad y rendimiento optimizado

Cuando Usarla

Ideal para desarrollos que requieren flexibilidad, estabilidad y no quieren pagar licencias.



Breve descripción

Es un motor de base de datos ligero que no necesita un servidor, lo que lo hace ideal para aplicaciones pequeñas y embebidas.

Precio: Gratuita (Open-source).

Características

- No requiere configuración de servidor.
- Muy rápida para aplicaciones pequeñas.

Cuando Usarla

Para aplicaciones móviles, sitios web pequeños o cuando se necesita un motor de base de datos ligero y sencillo



MySQLTM

Breve descripción

Es uno de los motores más populares, ampliamente utilizado en aplicaciones web.

Precio: Gratuita (Open-source) con opciones pagas para soporte empresarial

Características

- Rápido y confiable para aplicaciones web.
- Amplio soporte y comunidad.
- Fácil de usar y configurar.

Cuando Usarla

Aplicaciones web y desarrollo donde se requiere rapidez y facilidad de uso, como WordPress, Drupal, etc



Microsoft®
SQL Server®

Breve descripción

Es el motor de bases de datos de Microsoft, utilizado ampliamente en entornos empresariales que trabajan con productos Microsoft.

Precio: Pago por licencia, con opciones gratuitas limitadas como SQL Server Express

Características

- Integración nativa con productos de Microsoft
- Herramientas avanzadas de análisis y reporting
- Seguridad empresarial de alto nivel

Cuando Usarla

En empresas que ya están utilizando productos del ecosistema Microsoft y buscan una solución integrada



Breve descripción

Es el motor de bases de datos de IBM, diseñado para entornos empresariales que necesitan alta disponibilidad y rendimiento

Precio: Pago por licencia, con algunas versiones gratuitas.

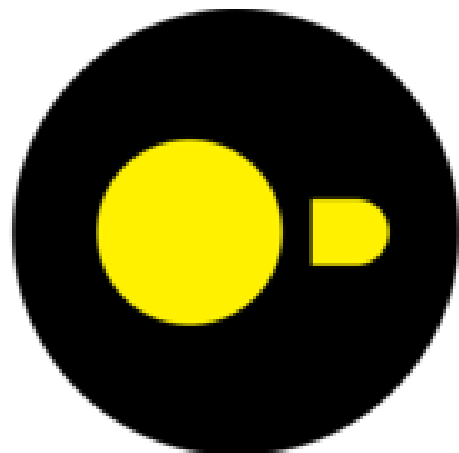
Características

- soporte para bases de datos relacionales y NoSQL
- Alta escalabilidad y rendimiento en sistemas de misión crítica
- Funciones avanzadas de replicación y recuperación ante desastres

Cuando Usarla

Para grandes empresas que requieren un sistema de gestión de bases de datos robusto y con soporte empresarial

¿POR QUÉ USAREMOS DUCKDB?



DuckDB

Breve descripción

Es un motor de bases de datos ligero diseñado especialmente para consultas analíticas embebidas. Se caracteriza por ser altamente eficiente para análisis de datos locales en laptops o servidores pequeños

Precio: Gratuito (Open-source).

Características

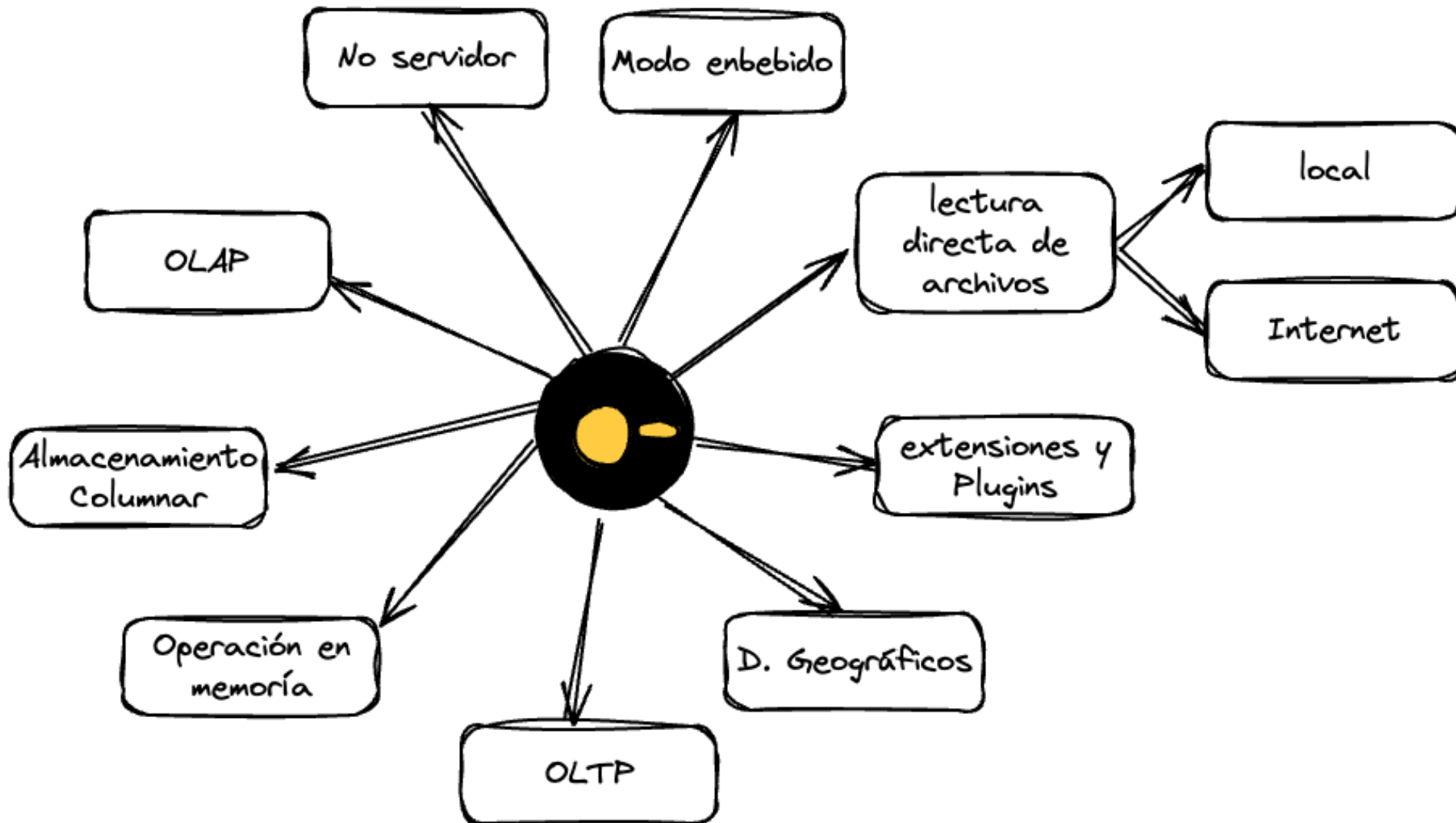
- Alto rendimiento para consultas analíticas sobre grandes volúmenes de datos.
- Sin necesidad de servidor: Se puede utilizar directamente desde una aplicación sin necesidad de un servidor externo.
- Soporta archivos CSV y Parquet, ampliamente utilizados en análisis de datos.

Cuando Usarla

Ideal para proyectos de análisis de datos embebido, donde se necesita rapidez sin desplegar infraestructura compleja, o para trabajar con grandes datasets en laptops o servidores locales. Perfecto para científicos de datos, analistas, y desarrolladores que requieren análisis rápido y eficiente.

TÉCNICAMENTE

softserve



softserve

¿PREGUNTAS?

**THANK
YOU!**



**FOR
THE
FUTURE**