



CONCEPTOS DE DATOS NO RELACIONALES

CONTENIDOS



- EXPLORE LAS CARACTERÍSTICAS DE LOS DATOS NO RELACIONALES.



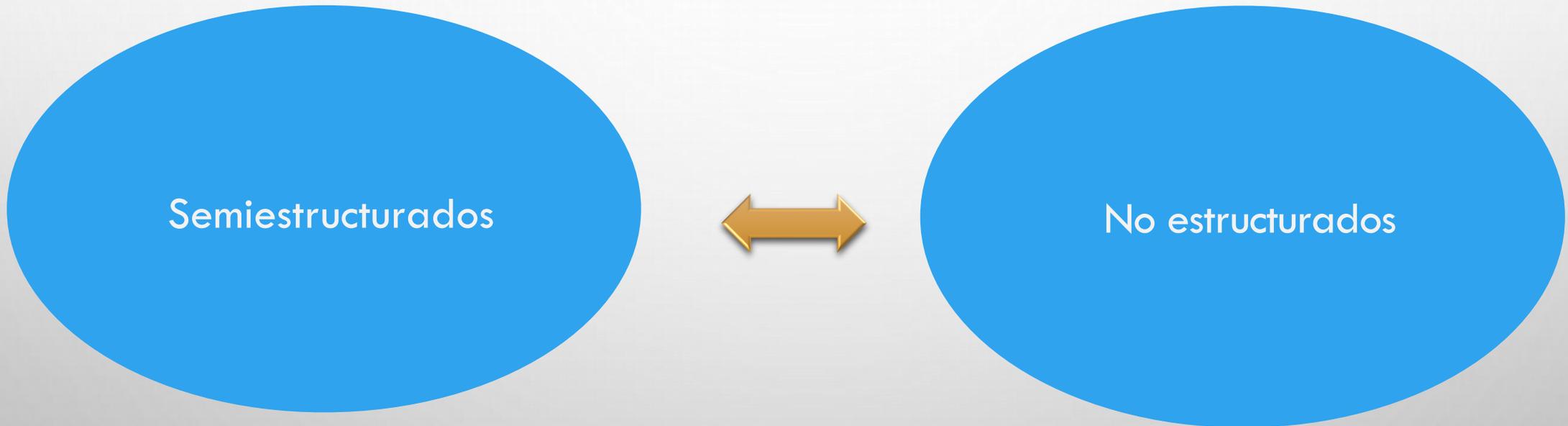
- DEFINA LOS TIPOS DE DATOS NO RELACIONALES.



- DESCRIBA NOSQL Y LOS TIPOS DE BASES DE DATOS NO RELACIONALES.

TIPOS DE DATOS NO RELACIONALES

Los datos no relacionales se dividen, en general, en dos categorías: semiestructurados y no estructurados.



¿QUÉ SON LOS DATOS NO ESTRUCTURADOS?



No contienen campos de forma natural:

Ejemplos: vídeo, audio, streaming de elementos multimedia y documentos.



A menudo se usa para extraer formularios de datos y categorizar o identificar "estructuras".



Se usa con frecuencia en combinación con las capacidades de Machine Learning o Cognitive Services para "extraer datos" mediante:

Text Analytics.

Análisis de sentimiento con API cognitivas.

Visión API.

CARACTERÍSTICAS:

Las colecciones de datos no relacionales pueden tener:

Varias entidades en la misma colección o contenedor con campos diferentes.

Un esquema diferente no tabular.

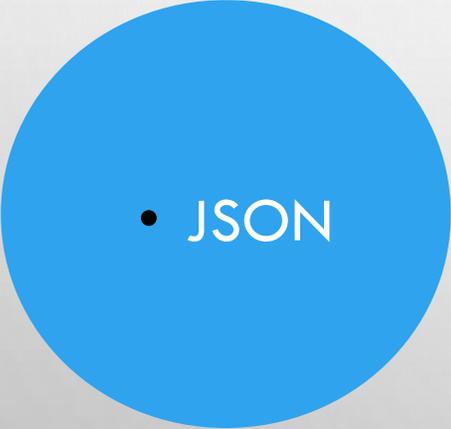
Se suelen definir etiquetando cada campo con el nombre que representan.

- EJEMPLO DE PROGRAMACIÓN DE UNA ENTIDAD

```
## Cliente 1 ID: 1
Nombre: Mark Hanson
Teléfono: [ Domicilio: 1-999-9999999, Empresa: 1-888-8888888, Móvil: 1-777- 7777777 ]
Dirección: [ Domicilio: 121 Main Street, Alguna ciudad, NY, 10110,
             Trabajo: 87 Big Building, Some City, NY, 10111 ]
## Cliente 2 ID: 2
Título: D.
Nombre: Jorge Salgado
Teléfono: [ Domicilio: 0044-1999-333333, Móvil: 0044-17545-444444 ]
Dirección: [ Reino Unido: 86 High Street, Ciudad, Condado, GL8888, Reino Unido,
            EE. UU.: 777 7th Street, Another City, CA, 90111 ]
```

TIPOS DE ARCHIVOS PARA DATOS SEMIESTRUCTURADOS

- Los tipos de formato/archivo incluyen:



• JSON



• AVRO



• ORC



• PARQUET

EJEMPLOS DE USO DE BASES DE DATOS NO RELACIONALES



IoT y telemática:

A menudo requieren ingerir grandes cantidades de datos en ráfagas frecuentes de actividad, los datos son semiestructurados o estructurados, a menudo requieren procesamiento en tiempo real.



Comercio y marketing:

Escenarios comunes para datos distribuidos globalmente, almacenamiento de documentos.



Juegos:

Estadísticas del juego, integración en redes sociales, marcadores, aplicaciones de baja latencia.



Web y móvil:

Se suelen usar con análisis de clics en web y aplicaciones modernas que incluyen bots.

¿Qué es NoSQL?

- Es un término bastante flexible que simplemente significa no relacional.
- Las bases de datos NoSQL (no relacionales) generalmente se dividen en cuatro categorías:
 - almacenes de clave-valor.
 - bases de datos de documentos.
 - bases de datos de familias de columnas.
 - bases de datos de gráficos.

ALMACENES DE CLAVE-VALOR

- Un almacén de clave-valor es el tipo más simple (y a menudo más rápido) de base de datos NOSQL para insertar y consultar datos.
- Cada elemento de datos en un almacén de clave-valor tiene dos elementos, una clave y un valor.
- La clave identifica de forma única el artículo y el valor contiene los datos del artículo.
- El valor es opaco al sistema de administración de bases de datos.
- Los artículos se almacenan por orden de clave.

BASE DE DATOS DE DOCUMENTOS

- Una base de datos de documentos representa el extremo opuesto del espectro NOSQL de un almacén clave-valor.
- Cada documento tiene una identificación única, pero los campos de los documentos son transparentes para el sistema de administración de bases de datos.
- Las bases de datos de documentos normalmente almacenan datos en formato: JSON, XML, YAML, JSON, BSON.
- Los documentos incluso se pueden almacenar como texto sin formato.
- Los campos de los documentos se exponen al sistema de administración de almacenamiento, lo que permite que una aplicación consulte y filtre datos utilizando los valores de estos campos.
- Normalmente, un documento contiene todos los datos de una entidad.
- Los elementos que constituyen una entidad son específicos de la aplicación.
- Por ejemplo, una entidad podría contener los detalles de la clientela, un pedido o una combinación de ambos.

BASES DE DATOS DE FAMILIAS DE COLUMNAS

Organiza los datos en filas y columnas.

El poder real de una base de datos de familias de columnas radica en su enfoque desnormalizado para estructurar datos dispersos.

Cada familia de columnas contiene un conjunto de columnas que están lógicamente relacionadas entre sí.

BASES DE DATOS DE FAMILIAS DE COLUMNAS

- La siguiente imagen muestra una forma de estructurar la misma información que la imagen anterior, mediante el uso de una base de datos de familias de columnas para agrupar los datos en dos familias de columnas que contienen el nombre de la clientela y la información de la dirección.

Clave de la fila	Columna Familias			
IDdeCliente	InfoDeCliente		InfoDeDirección	
1	InfoDeCliente:Título	D.	InfoDeDirección:DirecciónPostal	999 500th Ave
	InfoDeCliente:Nombre	Mark	InfoDeDirección:Ciudad	Bellevue
	InfoDeCliente:Apellido	Hanson	InfoDeDirección:Estado	WA
			InfoDeDirección:CódigoPostal	12345
2	InfoDeCliente:Título	Dña.	InfoDeDirección:DirecciónPostal	888 w. Front St
	InfoDeCliente:Nombre	Lisa	InfoDeDirección:Ciudad	Boise
	InfoDeCliente:Apellido	Andrews	InfoDeDirección:Estado	ID
			InfoDeDirección:CódigoPostal	54321
3	InfoDeCliente:Título	D.	InfoDeDirección:DirecciónPostal	999 500th Ave
	InfoDeCliente:Nombre	Walter	InfoDeDirección:Ciudad	Bellevue
	InfoDeCliente:Apellido	Harp	InfoDeDirección:Estado	WA
			InfoDeDirección:CódigoPostal	12345

BASE DE DATOS DE GRÁFICOS



Las bases de datos de grafos le permiten almacenar entidades, pero el enfoque principal está en las relaciones que estas entidades tienen entre sí.



Además, los bordes pueden tener una dirección que indique la naturaleza de la relación.



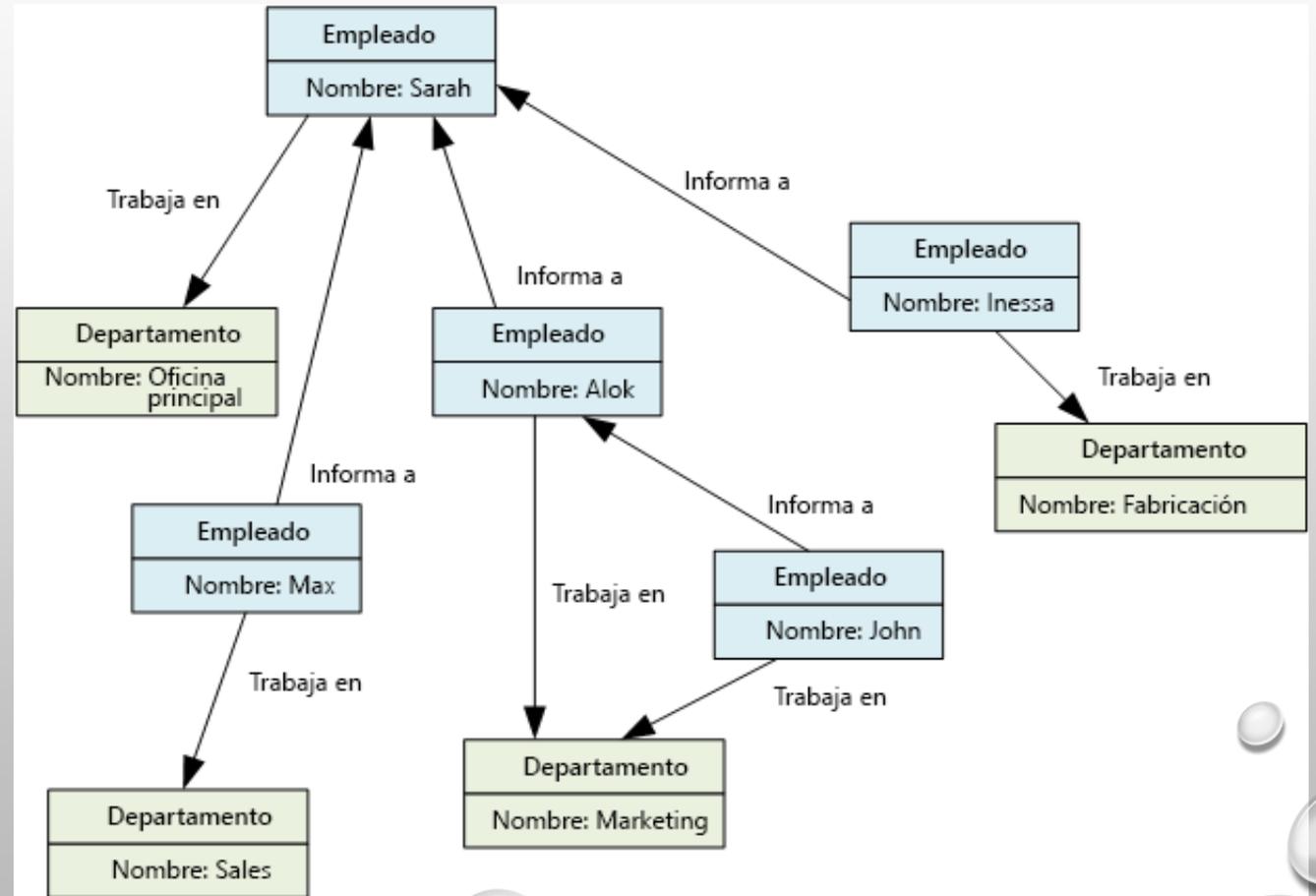
Una base de datos de grafos almacena dos tipos de información: nodos en los que puede pensar como instancias de entidades y bordes, que especifican las relaciones entre los nodos.



El propósito de una base de datos de grafos es permitir que una aplicación realice consultas de manera eficiente que atraviesan la red de nodos y bordes, y analizar las relaciones entre entidades.

BASE DE DATOS DE GRÁFICOS

- La siguiente imagen muestra la base de datos del personal de una organización estructurada como un grafo. Las entidades son los empleados y los departamentos de la organización, y los bordes indican las líneas organizacionales y el departamento en el que trabajan el personal. En este grafo, las flechas en los bordes muestran la dirección de las relaciones.



¿QUÉ APLICACIONES REQUIEREN UNA BASE DE DATOS DE GRAFOS?

- Requisitos empresariales:
- Aplicaciones oltp con alta correlación de datos.
- Fácil actualización de uno o varios objetos.
- Modelado de datos flexible.
- Requisitos de datos que evolucionan.
- Estructuras jerárquicas de datos.

