



Los Grafos

DRA. KARINA RUBY PÉREZ-DANIEL

FACULTAD DE INGENIERÍA
UNIVERSIDAD PANAMERICANA

ENERO - MARZO, 2020



UNIVERSIDAD
PANAMERICANA

CONTENIDO

- 1 FUNDAMENTOS DE MACHINE LEARNING
- 2 LOS GRAFOS
- 3 LAS REDES SOCIALES
- 4 PRIMEROS PASOS EN NEO4J
- 5 DETALLES TÉCNICOS DE NEO4J



UNIVERSIDAD
PANAMERICANA

SECCIÓN 1

FUNDAMENTOS DE MACHINE LEARNING



UNIVERSIDAD
PANAMERICANA

ENFOQUES DE ML PARA LA RESOLUCIÓN DE PROBLEMAS

Los algoritmos de ML pueden clasificarse de acuerdo al **tipo de estrategias** que usan.

- **Bioinspirados**

- Redes neuronales,
- Algoritmos genéticos,

- **Matemáticos**,

- Grafos y búsquedas,
- Probabilísticos,
- Simbólicos: lógica y Sistemas basados en conocimiento.

- **Mixtos**.

- Redes bayesianas.



UNIVERSIDAD
PANAMERICANA

ÁREAS DE ML

Otra clasificación esta dada por el **tipo de problemas** que ML resuelve.

- Búsquedas
- Representación del conocimiento
 - Simbólico, conexiónista, mixto
- Tratamiento de la incertidumbre,
- Razonamiento,
 - Deductivo, inductivo
- Planificación,
- Toma de decisiones,
- Aprendizaje,
 - Supervisado, no supervisado, semi supervisado, por refuerzo.

¿Qué aplicaciones tiene el ML?

- Juegos, resolución automática de problemas, diagnóstico, predicción, NLP, percepción, robótica, **Minería de datos**.



ALGORITMOS DE ML CON ENFOQUE MATEMÁTICO

La teoría de grafos se encuentra en la categoría de algoritmos de ML con enfoque matemático, en donde se encuentran:

- **Búsqueda en Grafos,**
- Probabilidad (manejo de incertidumbre, probabilidad bayesiana, redes neuronales, aprendizaje),
- Lógicos.



UNIVERSIDAD
PANAMERICANA

BÚSQUEDA EN GRAFOS

Los algoritmos para la *búsqueda en grafos* se pueden clasificar de la siguiente manera:

- **Planificación,**

- Busqueda No informada,
- Búsqueda informada (Heurística),
- Satisfacción de restricciones,

- **Juegos,**

- **Toma de decisiones,**

- **Aprendizaje por refuerzo**



UNIVERSIDAD
PANAMERICANA

BÚSQUEDA EN GRAFOS: PLANIFICACIÓN

No informada

- Anchura,
- Anchura de Costo Uniforme,
- en Profundidad,
- Profundidad Limitada,
- Profundidad Iterativa,
- Bidireccional,
- en Grafos.

UNIVERSIDAD
PANAMERICANA

BÚSQUEDA EN GRAFOS: PLANIFICACIÓN

Informada

- Heurísticas,
- Voraz: Primero el mejor,
- A* y AO*,
- Ascención en colinas,
- Tabú,
- Temple Simulado,
- Haz local,
- Online,
- Algoritmos Genéticos.



BÚSQUEDA EN GRAFOS

Satisfacción de Restricciones

- Búsqueda de vuelta atrás,
- Comprobación hacia adelante,
- Propagación de restricciones,
- Salto atrás dirigido por conflictos,
- Búsqueda local: Mínimos conflictos,
- Acondicionamiento del corte.

UNIVERSIDAD
PANAMERICANA

¿Son relevantes los algoritmos de planificación en grafos en el análisis de redes sociales? ...



UNIVERSIDAD
PANAMERICANA

SECCIÓN 2 | LOS GRAFOS



UNIVERSIDAD
PANAMERICANA

EL MUNDO ES UN GRAFO Y TODO ESTA CONECTADO

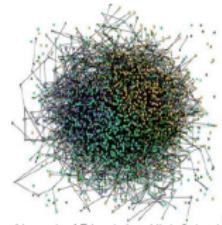
El mundo es un grafo ...

- Gente, los lugares, los eventos
- Las compañías, los mercados
- Los países, la historia, la política
- La ciencia, las artes, la enseñanza
- La tecnología, las redes, las máquinas, las aplicaciones, los usuarios
- El software, el código, las dependencias, la arquitectura de los sistemas, los desarrollos, los deployments
- Los criminales, los estafadores y su comportamiento, etc.

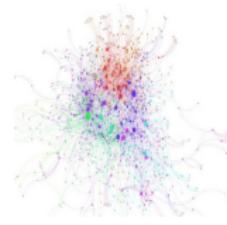


UNIVERSIDAD
PANAMERICANA

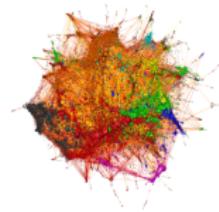
¿CÓMO ES UN GRAFO?



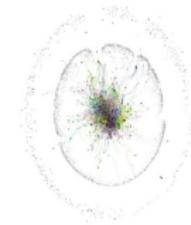
(a) Red de Amigos



(b) Generos musicales y su origen



(c) Relación entre artistas



(d) Relación entre patentes

Figura 1. Ejemplos de algunas redes

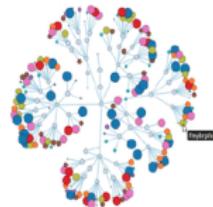
¿Qué tipo de problemas se resuelven mediante los Grafos?

La ruta más corta, flujo máximo por determinada ruta, identifica nodos de interés, matching (tinder), page rank, etc.

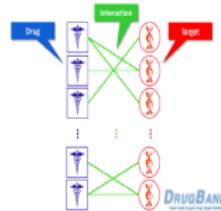


UNIVERSIDAD
PANAMERICANA

¿QUÉ TIPO DE PROBLEMAS SE RESUELVEN MEDIANTE LOS GRAFOS?



(a) Patrones de conexiones



(b) Propiedades topológicas

Figura 2. Problemas que se resuelven mediante grafos

- **Patrones de conexión:** Se realiza para entender como interactua la sociedad evaluada.
- **Propiedades topológicas:** Se usan para identificar los patrones anomalías presentes en un fenómeno.



¿QUÉ TIPO DE PROBLEMAS SE RESUELVEN MEDIANTE LOS GRAFOS?

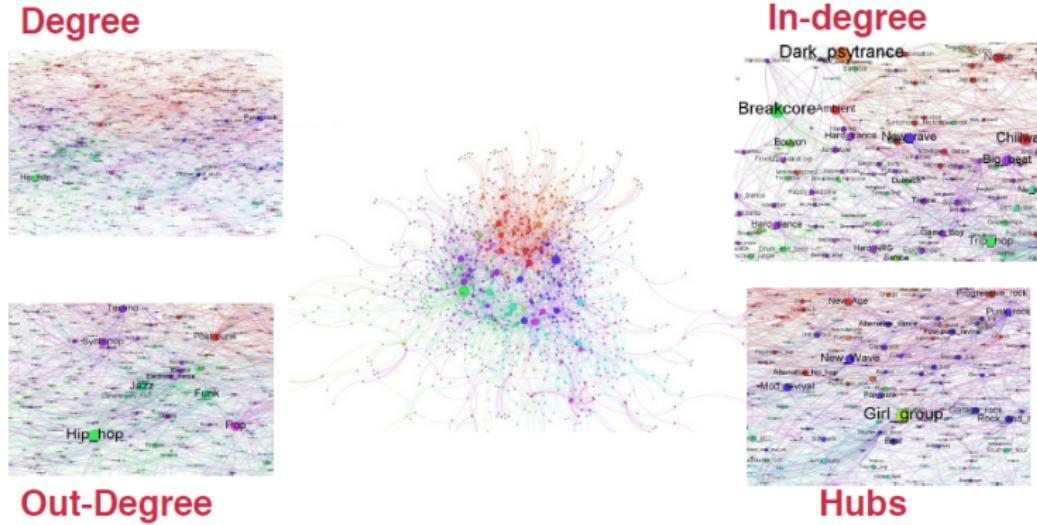


Figura 3. Grados de conexión

La importancia de la información esta basada en las relaciones que se encuentran más que en los individuos.



¿QUÉ TIPO DE PROBLEMAS SE RESUELVEN MEDIANTE LOS GRAFOS?

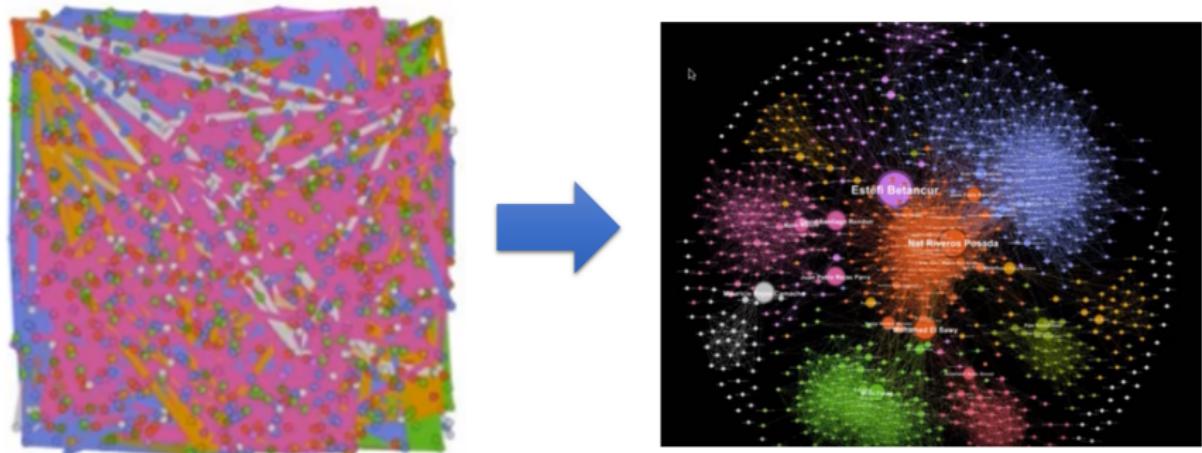


Figura 4. Grados de conexión y la centralidad



UNIVERSIDAD
PANAMERICANA

Volvemos a la pregunta
¿Son relevantes los algoritmos de planificación
en grafos en el análisis de redes sociales? ...



UNIVERSIDAD
PANAMERICANA

SECCIÓN 3

LAS REDES SOCIALES



UNIVERSIDAD
PANAMERICANA

EVOLUCIÓN DE LAS REDES SOCIALES

- En 1934 se publicó el **primer estudio** relacionado al **análisis de las redes sociales** donde se analizó junto a quién se quería sentar cada niño en un salón de clases.
- En los años 50's, el análisis de redes sociales estudió las relaciones entre citas de obras literarias
- En 1977 **se crea Red Internacional para el Análisis de Redes Sociales**,
- En la decada de los 80's llega el **boom** del análisis de redes sociales,
- En 1997 se crea el **primer servicio de red social** (**SNS**), *six degrees*.



Figura 5. Evolución del análisis de redes sociales



UNIVERSIDAD
PANAMERICANA

EVOLUCIÓN DE LAS REDES SOCIALES

- En 1977 se formaliza el análisis de redes sociales (SNA) con la fundación de la Red Internacional para el Análisis de Redes Sociales (INSNA).
- Las redes sociales son un grafo de las relaciones que existen entre personas de una sociedad.
- Los servicios de redes sociales (SNS) son servicios electrónicos que se apoyan en los grafos y en el análisis de redes sociales.

UNIVERSIDAD
PANAMERICANA

LOS SERVICIOS DE REDES SOCIALES

Los Servicios de Redes Sociales (SNS)

- Son representados mediante grafos,
- Utilizan teoría de grafos para su análisis,
- Existen algunas diferencias en la forma en la que se aplican los conceptos correspondientes a la teoría de grafos,
- Existen similitudes con los grafos convencionales.

UNIVERSIDAD
PANAMERICANA

CLASIFICACIÓN DE LAS REDES SOCIALES

- **Tamaño:** Pequeña (sistema de patentes, citas bibliográficas, etc.) y gran escala (correo electrónico, twitter, etc.).
- **Evolución:** Estáticas (no sufre cambios en el tiempo de análisis) y dinámicas.
- **Origen:** Off line (Las relaciones sociales son establecidas sin la intervención de un medio electrónico) y On line.
- **Topología:** Simples y complejas.

UNIVERSIDAD
PANAMERICANA

PROPIEDADES DE LAS REDES SOCIALES

- **Distancia:** Es la longitud del camino más corto de un nodo a otro.
- **Diámetro:** Es la distancia máxima entre 2 nodos en un grafo.
- **Grado:** Es el número de enlaces asociados a un nodo. En el caso de los grafos dirigidos se suele dividir en **In-degree** y en **Out-degree**.
- **Adyacencia:** Se dice que 2 nodos son adyacentes si entre ellos hay un enlace directo.
- **Centralidad:** Determina la importancia relativa de un nodo en un grafo.
- **Vecindad:** Nodos a su alrededor de un determinado nodo.



PROPIEDADES DE LOS GRAFOS

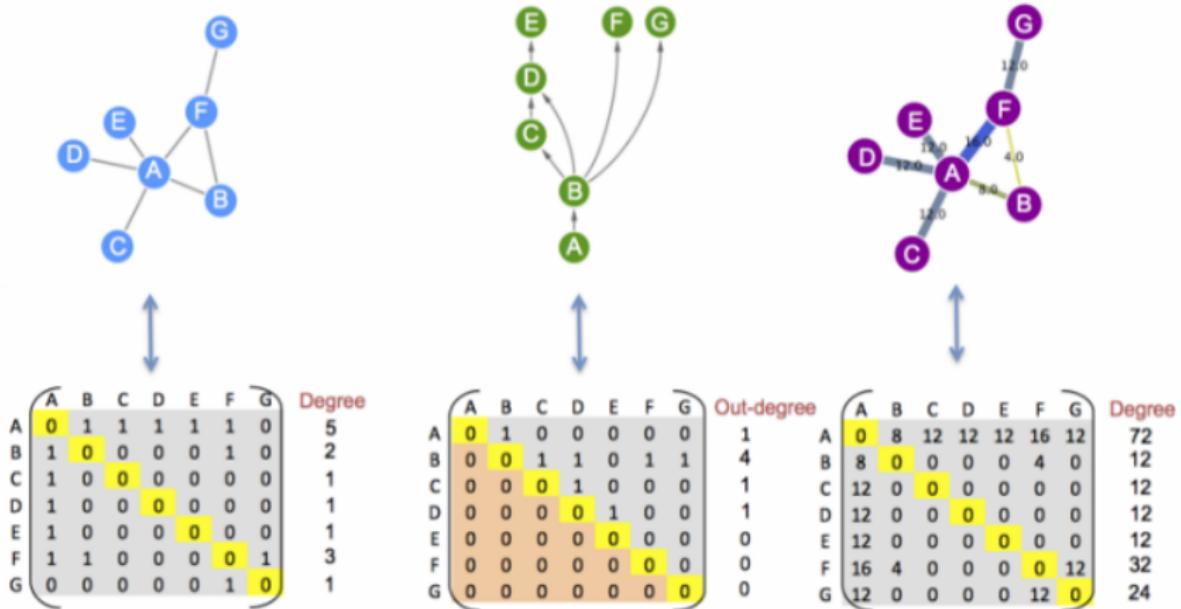


Figura 6. Grado del grafo y su matriz de adyacencia



PROPIEDADES DE LOS GRAFOS

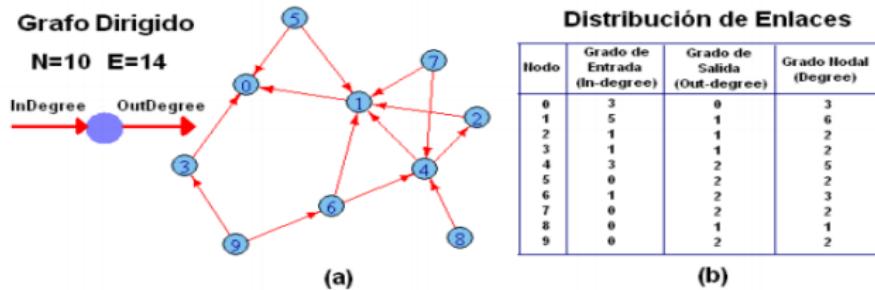


Figura 7. Distribución de enlaces



LOS USUARIOS Y SUS CONEXIONES

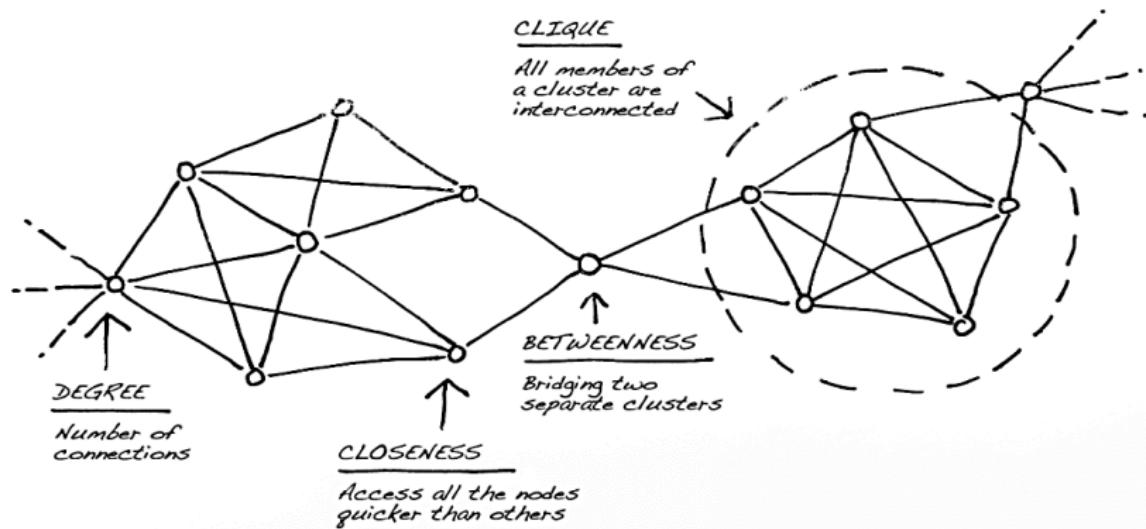


Figura 8. Centralidad en las Redes Sociales



SECCIÓN 4

PRIMEROS PASOS EN NEO4J



UNIVERSIDAD
PANAMERICANA

¿QUÉ ES CYpher?

Neo4j esta basado en el lenguaje Cypher.

Cypher:

- Es un lenguaje declarativo para hacer consultas en bases de datos basadas en grafos.
- Es especialmente eficiente para realizar consultas, actualizaciones y administrar un grafo.
- Esta inspirado en algunas prácticas comúnmente utilizadas en las bases de datos SQL.
- Es intuitivo, rápido (en el desarrollo y ejecución) y ágil que las bases de datos basadas en SQL.



UNIVERSIDAD
PANAMERICANA

LOS BÁSICOS DE NEO4J

Componentes básicos de un grafo en Neo4j.

• Nodos

- Representa un objeto en el grafo
- Puede ser etiquetado

• Relaciones

- Relaciona nodos por *tipo* y *dirección*

• Propiedades

- Pares de tipo *nombre/clave-valor*
- Pueden existir tanto en los nodos como en las relaciones

• Etiquetas

- Agrupan nodos por rol.

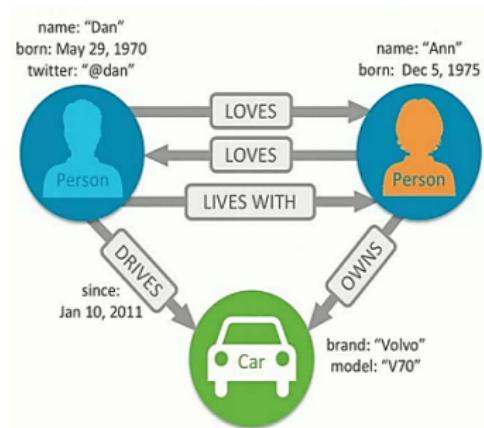


Figura 9. Estructura básica de un grafo en Neo4j



CREANDO NODOS EN NEO4J

¿Cómo creo nodos en Neo4j?

En Neo4j, los nodos se representan con paréntesis ()

CREATE

```
(d:Person{name:"Dan", born: "Mayo 29, 1970", tw:@"dan"}),  
(a:Person{name:"Ann", born: "Dic 5, 1974"})
```



Figura 10. Creación de nodos en Neo4j



UNIVERSIDAD
PANAMERICANA

CREANDO RELACIONES EN NEO4J

¿Qué tipo de relaciones hay en Neo4j y cómo las represento?

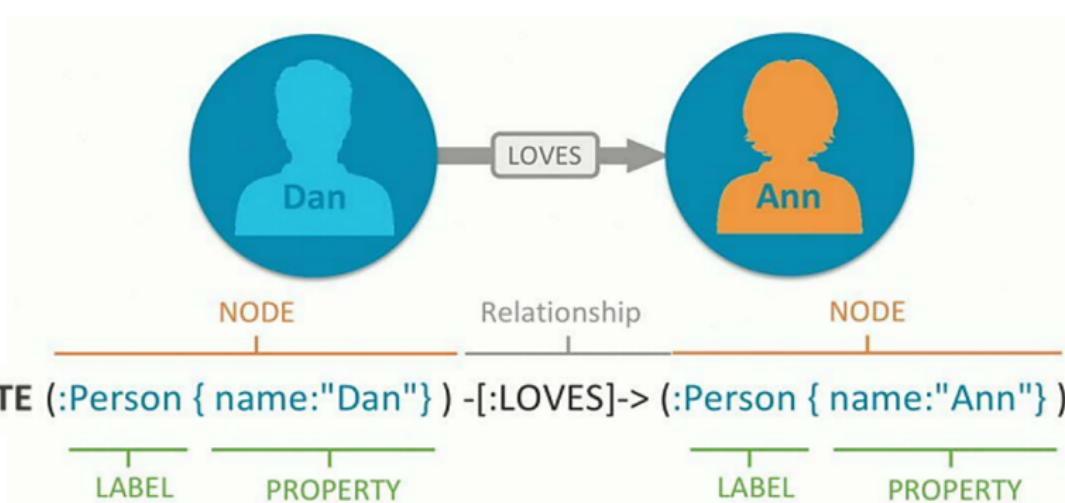
Relaciones

- Las **relaciones** se representan mediante flechas y el detalle de la relación se escribe entre corchetes.
 - ->
 - - [:RELACION] ->
- **Relaciones y su dirección**
 - () - [] -> ()
 - () <- [] - ()
- Unicamente se puede definir una dirección a la vez
- Cypher puede considerar o no la dirección del enlace pero **no soporta** la declaración de enlaces sin dirección.



CREANDO RELACIONES EN NEO4J

¿Cómo creo una relación en un grafo en Neo4j?



Sintáxis.

CREATE

```
(:Person{name:"Dan"})-[:LOVES]->(:Person{name:"Ann"})
```

o bien,

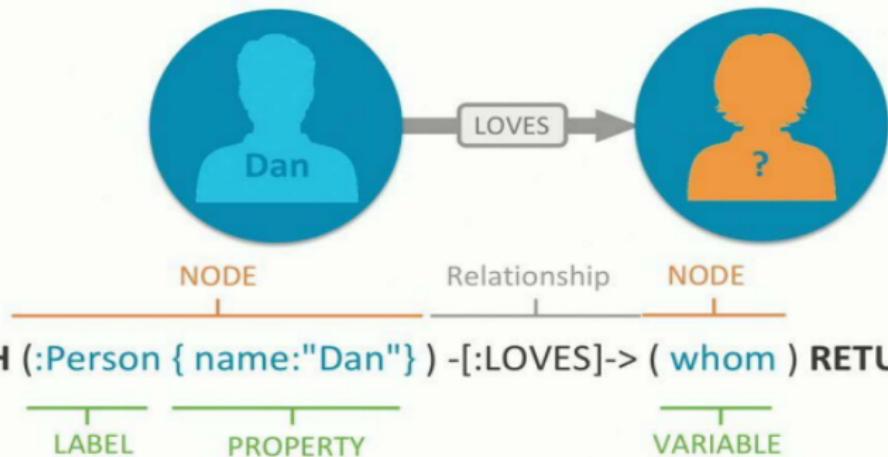
CREATE

(d) - [:LOVES] -> (a)



ENCONTRANDO PATRONES EN GRAFOS EN NEO4J

¿Cómo encuentro los elementos que hacen *match* con determinado patrón?



Sintaxis.

```

MATCH amor =(:Person{name:"Dan"})-[:LOVES]->(pareja)
RETURN amor
  
```

o bien,

```

MATCH amor =(:Person{name:"Dan"})-[:LOVES]->(pareja)
RETURN pareja.name
  
```



SECCIÓN 5

DETALLES TÉCNICOS DE NEO4J



UNIVERSIDAD
PANAMERICANA

DETALLES TÉCNICOS

Elementos sensibles a mayúsculas y minúsculas en Neo4j

- Etiquetas de los nodos (`(:Persona)`)
- Tipos de relaciones (`(:TIENE)`)
- Claves de las propiedades (`{nombre}`)

Naming rules

- **Etiqueta del nodo:** Se recomienda que cada palabra comience con letra mayúscula y continúe con minúsculas.
- **Tipo de Relación:** Se recomienda que se escriba en mayúsculas completamente.
- **Cypher NO es sensible a mayúsculas y minúsculas en sus palabras clave.**

UNIVERSIDAD
PANAMERICANA

PALABRAS RESERVADAS

Clauses

CALL, CREATE, DELETE, DETACH, EXISTS, FOREACH, LOAD,
MATCH, MERGE, OPTIONAL, REMOVE, RETURN, SET, START,
UNION, UNWIND, WITH

Subclauses

LIMIT, ORDER, SKIP , WHERE, YIELD

Modifiers

ASC, ASCENDING, ASSERT, BY, CSV, DESCENDING, ON

Expressions

ALL, CASE, ELSE, END, THEN, WHEN



UNIVERSIDAD
PANAMERICANA

PALABRAS RESERVADAS

Operators

AND, AS, CONTAINS, DISTINCT, ENDS, IN, IS, NOT, OR,
STARTS, XOR

Schema

CONSTRAINT, CREATE, DROP, EXISTS, INDEX, NODE, KEY,
UNIQUE

Literals

false, null, true



UNIVERSIDAD
PANAMERICANA

MÁS DETALLES PARA CREAR UN NODO

- `CREATE (n) , (m)` Crea 2 nodos
- `CREATE (n:Person)` Crea 1 nodo con 1 etiqueta
- `CREATE (n:Person:Mexican)` Crea 1 nodo con 2 etiquetas
- `CREATE (n:Person{name:'Andy', title: 'Developer'})`
Crea 1 nodo con 1 etiqueta y 2 propiedades

● Devuelve un nodo creado:

```
CREATE (a{ name: 'Andy' })
```

```
RETURN a.name
```

Crea 1 nodo con 1 etiqueta y 2 propiedades



UNIVERSIDAD
PANAMERICANA

ELIMINAR UN NODO

- **DELETE**: Se usa para borrar nodos y relaciones.
- Para borrar un nodo es necesario borrar todas las relaciones que lo conectan.
- `MATCH (n:Person { name: 'JOE' })
DELETE n` Borra un único nodo.
- `MATCH (n)
DETACH DELETE n`
Borra todos los nodos y relaciones
- `MATCH (n:Person{name:'Andy'})
DETACH DELETE n` Borra un nodo y todas sus relaciones
- `MATCH (n { name: 'Andy' })-[r:KNOWS]->()
DELETE r` Borra únicamente las relaciones del nodo



CASO PRÁCTICO

TASK

Realiza el siguiente grafo en Neo4j

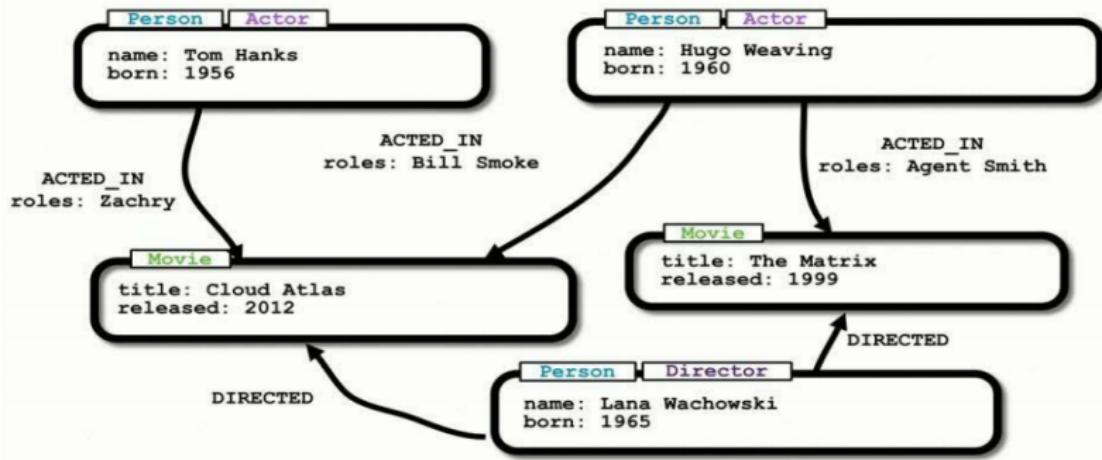


Figura 11. Caso práctico



Gracias por su atención

Fin



UNIVERSIDAD
PANAMERICANA