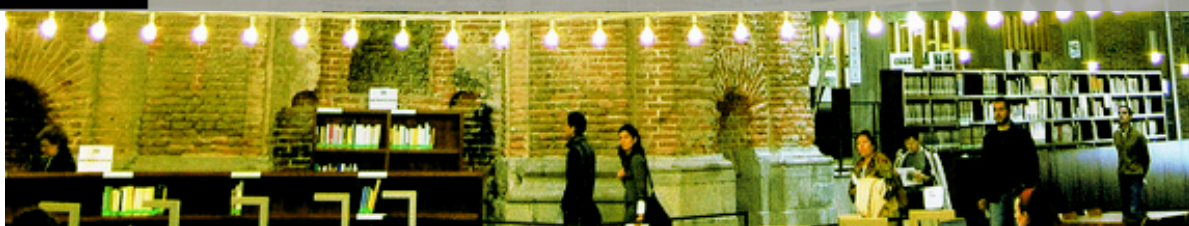


MÁSTER

GESTIÓN/ALMACENAMIENTO DE INFORMACIÓN NO ESTRUCTURADA

PRÁCTICA DEL MÓDULO 5: BASES DE DATOS ORIENTADAS A
GRAFOS



Versión 1.0

Dr. Agustín C. Caminero — Dr. Pablo Ruipérez
Dr. Juan Mario Haut — Dr. Luis Grau

MÁSTER UNIVERSITARIO EN INGENIERÍA Y CIENCIA
DE DATOS

1.Introducción

En este documento se presenta el Trabajo Práctico (TP) del módulo 5 de la asignatura "GESTIÓN/ALMACENAMIENTO DE INFORMACIÓN NO ESTRUCTURADA", del "MÁSTER UNIVERSITARIO EN INGENIERÍA Y CIENCIA DE DATOS" de la UNED.

Este trabajo se realiza de forma individual. En las siguientes secciones se exponen los diferentes ejercicios que es necesario implementar para este trabajo.

La forma de evaluar el trabajo se hará en base a lo siguiente:

- Fichero(s) de texto con el código que responda a las preguntas.
- Una memoria explicativa que incluya pantallazos que demuestren la correcta ejecución de dicho código, así como cualquier detalle o explicación que facilite la corrección.

Se valorará positivamente que la memoria contenga un apartado final donde se explique la opinión del/la estudiante sobre este trabajo, puntos fuertes y/o débiles, recomendaciones para el futuro, así como una valoración general de este módulo.

Este material se deberá incluir en un fichero comprimido y enviado a través del curso virtual dentro de los plazos establecidos para su entrega. El nombre de dicho fichero comprimido deberá tener la estructura *Grafos-ApellidosNombre.zip*, donde *Apellidos* y *Nombre* deben sustituirse por los valores correspondientes para el/la estudiante que realiza el envío. Un nombre correcto es, por ejemplo, el siguiente:

- Grafos-CamineroHerraezAgustin.zip

La puntuación de este trabajo práctico no atenderá únicamente al correcto funcionamiento de las soluciones entregada, sino también a la existencia de una documentación lo suficientemente completa, correctamente escrita y explicativa de todos los aspectos que el estudiante considere importantes (decisiones sobre el código, dificultades, opiniones, etc.).

A continuación se detallan los ejercicios a completar.

2.Ejercicios de evaluación

1. (1 puntos) Crea una base de datos con los datos que se adjuntan en el archivo grafo.txt disponible en el curso virtual. El procedimiento de despliegue del sistema neo4j se detalla en el apartado 3 y la base de datos se explica en el apartado 4.
2. Sobre dicha base de datos, implementa consultas para conseguir la siguiente información:

PRÁCTICA DEL MÓDULO 5: BASES DE DATOS ORIENTADAS A GRAFOS

- 2.1. (2 puntos) Queremos conocer nodos tipo Cliente que hayan emitido una opinión tipo BAD del producto con id=2 (hay 3 clientes con estas características).
- 2.2. (2 puntos) Queremos buscar clientes que formen un "circuito" de relaciones de tipo "APOYA", es decir una secuencia de relaciones, todas en la misma dirección, que acaba y empieza en el mismo Cliente.
- 2.3. (2 puntos) Demetrio ha decidido mostrar su apoyo a Genara. Añadir la relación APOYA entre los nodos correspondientes. No se debe suponer que el nodo existe (ni que no existe), solo se debe asegurar que tras la instrucción ambos existen (sin duplicarlos) y con la relación APOYA indicada.
3. (3 puntos) Crea tu propia base de datos sobre Neo4j, e implementa al menos 2 consultas sobre dichos datos. Se valorará la originalidad y la complejidad de la base de datos y de las consultas creadas.

3. Despliegue de un sistema Neo4j

Para realizar esta práctica se va a utilizar un sistema basado en virtualización ligera con Docker. Para desplegar dicho sistema, deberemos tener Docker instalado en nuestro ordenador, preferiblemente con sistema operativo Ubuntu.

Para proceder al despliegue, ejecutaremos la siguiente orden:

```
docker run \
  --publish=7474:7474 --publish=7687:7687 \
  --volume=$HOME/neo4j/data:/data \
  neo4j
```

Tras ejecutar esta orden, se descargará la imagen oficial de Neo4j [1], y se creará un contenedor accesible desde un navegador a través de la dirección <http://localhost:7474>. Además, se crea un volumen para que la base de datos persista fuera del contenedor.

Podremos acceder al contenedor recién creado utilizando el usuario/clave neo4j/neo4j

Tras el primer login, se obliga al usuario a cambiar dicha clave.

4. Creación de la base de datos

Disponemos de un grafo (archivo grafo.txt) con dos tipos de nodos:

PRÁCTICA DEL MÓDULO 5: BASES DE DATOS ORIENTADAS A GRAFOS

- Producto: Un producto comercializado por nuestra empresa.
- Cliente: un cliente que opina acerca del producto en la página web corporativa.

Además, tenemos 3 tipos de relaciones:

- GOOD: el cliente ha emitido un comentario positivo sobre el producto.
- BAD: el cliente ha emitido un comentario negativo sobre el producto.
- APOYA: Un cliente opina que los comentarios de otro cliente son útiles, y por tanto le muestra su apoyo.

En la Figura 1 se muestra una representación visual de dicha base de datos.

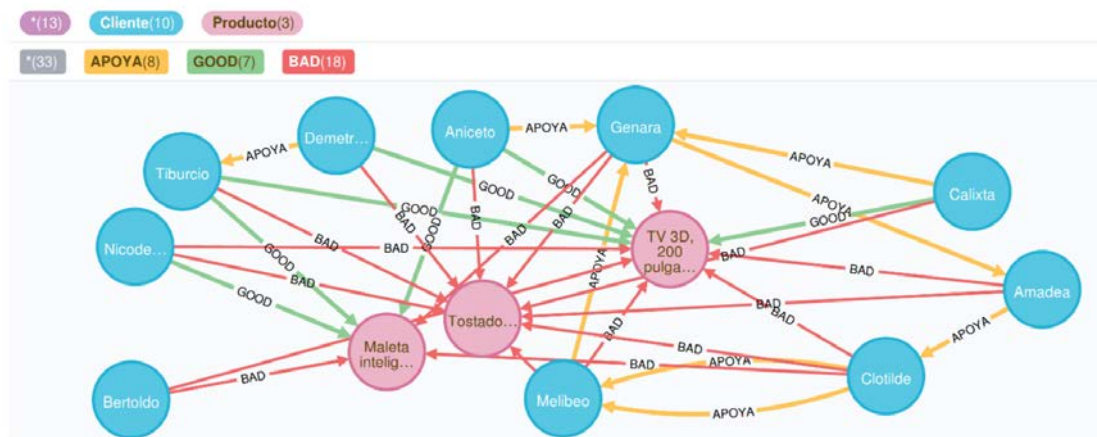


Figura 1. Grafo de la base de datos.

5. Referencias de interés

- [1] Imagen oficial de Neo4j. Disponible en el siguiente enlace: https://hub.docker.com/_/neo4j?tab=description Fecha del último acceso: 16 de marzo de 2020.