

Temas presentación II

La idea de las presentaciones es que en grupos de 4 personas exploren en más profundidad tópicos relacionados con este curso. En esta ronda de presentaciones los temas serán propuestos en torno a los temas de Bases de Datos de Grafos y *Graph Analytics*. La idea es que investiguen en torno a un tema de los que se detallan a continuación y realicen un video de **entre 06 a 12 minutos** que van a subir a YouTube donde expongan los resultados de la investigación. Además, cada grupo deberá brindar *feedback* a las presentaciones de cada otro grupo.

1. Temas

Los temas para esta ronda son los siguientes.

1.1. Neo4J vs SQL

En este tema se busca entender los casos de uso donde conviene usar Neo4J y los casos de uso donde conviene usar SQL. Para esto, se espera que muestren casos concretos, resultados de experimentos y/o papers que se hayan hecho la respecto, entre otros.

1.2. OrientDB

Para este tema se busque que averigüen sobre la Base de Datos de Grafos **OrientDB**. Se espera que hagan un tutorial que incluya la instalación, la forma de cargar datos y el lenguaje de consultas con algunos ejemplos. Además, se espera que discutan brevemente posibles ventajas y desventajas respecto a Neo4J.

1.3. ArangoDB

Para este tema se busque que averigüen sobre la Base de Datos de Grafos **ArangoDB**. Se espera que hagan un tutorial que incluya la instalación, la forma de cargar datos y el lenguaje de consultas con algunos ejemplos. Además, se espera que discutan brevemente posibles ventajas y desventajas respecto a Neo4J.

1.4. Funcionalidades adicionales de Neo4J

Para este tema se espera que busquen funcionalidades adicionales de Neo4J que no hayamos visto en clases y que sean interesantes. Por ejemplo, ¿hay formas de exigir un esquema para los datos?

1.5. El lenguaje para BDs de grafos Gremlin

Además de Cypher, existe otro popular lenguaje para Bases de Datos de Grafos llamado **Gremlin**. Este lenguaje no tiene un motor asociado, sino que utiliza otros motores como *backend*. A diferencia de SPARQL y Cypher, este lenguaje más que en patrones se basa en *instrucciones de navegación*. La idea de este tema es que hagan un tutorial sobre este lenguaje, incluyendo cómo correr consultas sobre algún motor de grafos.

1.6. Bases de Datos RDF

En este tema se espera que escojan un sistema de Bases de Datos RDF y hagan un tutorial de su instalación y que muestren ejemplos de cómo correr consultas en el sistema. Los sistemas más populares son **Openlink Virtuoso** y **Apache Jena**¹.

¹El profesor recomienda más intentar con Apache Jena.

1.7. RDFLib

Existe una librería de Python llamada RDFLib, que permite trabajar con datos RDF, incluso haciendo consultas a Endpoints. La idea de este tema es que hagan un tutorial de uso de esta librería, incluyendo la importación a Python de documentos RDF y la ejecución de consultas de forma local y en *endpoints* remotos.

1.8. Semantic Web en la práctica

La idea de este tema es que investiguen y presenten casos de estudios en los que se hayan aplicado tecnologías basadas en la Web Semántica en la práctica de forma exitosa.

1.9. Datos orientados a grafos en la industria

Los datos orientados a grafos han tomado mucha fuerza en la industria, sobre todo bajo el tópico de nombre *Knowledge Graphs*. La idea de este tema es que investiguen en torno a este tópico y que presenten cómo se han adoptado las tecnologías basadas en manejo de grafos. Además de Neo4J, puede ver el caso de Amazon Neptune, una base de datos de grafos en la nube.

1.10. Tutorial de Python+Neo4J

En clases vimos que existía una librería para trabajar con Neo4J desde Python llamada `neo4j` (no confundir con `py2neo`). La idea de este tema es construir un tutorial completo para esta librería, que incluya de cómo cargar datos, cómo hacer consultas, y otros temas similares.

1.11. Tutorial del módulo de Neo4J GDS

En clases vimos que existe un módulo de Neo4J llamado GDS que permite ejecutar procedimientos de *Graph Analytics*. En este tema se espera que hagan un tutorial de este módulo, además de explicar a grandes rasgos los algoritmos de *Graph Analytics* que ejecuten.

1.12. GraphQL

GraphQL es un lenguaje que nos permite hacer consultas a APIs en la Web pensando que los datos se ven como un grafo. La idea de este tema es hacer un tutorial completo del lenguaje GraphQL, además de explicar a alto nivel cómo se puede montar una API que se exponga con este lenguaje, por ejemplo, utilizando Apollo (<https://www.apollographql.com/>).

1.13. Otro tema

¿Tienes otro tema para esta ronda de presentaciones? ¡Publicítelo! Pueden proponer su propio tema como grupo, con previa autorización del profesor. La idea es que vaya en la misma línea de los otros temas planteados.

2. Feedback

Luego de que todos los videos estén publicados, cada grupo tendrá una semana para entregar *feedback* para cada uno de los videos de todos los otros grupos. Esto vale un 15 % de la nota.

3. Detalles Académicos

Cada grupo deberá enviar un correo al profesor con su preferencia para dos temas hasta el jueves 13 de mayo. Idealmente, no puede haber dos grupos que hagan el mismo tema, por lo que el tema de cada grupo se asignará por orden de llegada. Para que puedan trabajar en grupo para esta actividad, el lunes 31 de mayo no habrá clases.

Importante: El plazo para esta entrega asume que **no van a trabajar durante la semana de pausa**.

El link del video se debe entregar por un formulario antes del miércoles 09 de junio a las 20:00 horas. Pueden realizar la tarea en grupos de hasta cuatro personas. La fecha límite del *feedback* será informada después.