



Evaluación: Gestión de Grandes Volúmenes de Datos

- **Consideraciones:**

El examen deberá ser entregado en un documento gdoc con el desarrollo de cada uno de los ejercicios, printscreens, scripts de consultas y justificaciones.

A- MongoDB: CRUD - Replicación - Sharding

Resolver cada uno de los siguientes puntos en el motor MongoDB:

- Para resolver este ejercicio se utilizará el 'mongo shell' de MongoDB.
- Debe conectarse a una instancia activa de mongod, se puede utilizar una base de datos llamada mongo_exercise.

1- Documentos a Insertar

Insertar los siguientes documentos en una colección llamada movies.

```
title : Fight Club
writer : Chuck Palahniuk
year : 1999
actors : [
  Brad Pitt
  Edward Norton
]
```

```
title : Pulp Fiction
writer : Quentin Tarantino
year : 1994
actors : [
  John Travolta
  Uma Thurman
]
```

```
title : Inglorious Basterds
writer : Quentin Tarantino
year : 2009
actors : [
  Brad Pitt
  Diane Kruger
  Eli Roth
]
```

title : The Hobbit: An Unexpected Journey
writer : J.R.R. Tolkein
year : 2012
franchise : The Hobbit

title : The Hobbit: The Desolation of Smaug
writer : J.R.R. Tolkein
year : 2013
franchise : The Hobbit

title : The Hobbit: The Battle of the Five Armies
writer : J.R.R. Tolkein
year : 2012
franchise : The Hobbit
synopsis : Bilbo and Company are forced to engage in a war against an array of combatants and keep the Lonely Mountain from falling into the hands of a rising darkness.

title : Pee Wee Herman's Big Adventure

title : Avatar

2- Consultas / Buscar documentos

Realizar las siguientes consultas en la colección movies:

1. Obtener todos los documentos
2. Obtener documentos con writer igual a "Quentin Tarantino"
3. Obtener documentos con actors que incluyan a "Brad Pitt"
4. Obtener documentos con franchise igual a "The Hobbit"
5. Obtener todas las películas de los 90s.
6. Obtener las películas estrenadas entre el año 2000 y 2010.

3- Actualizar Documentos

1. Agregar sinopsis a "The Hobbit: An Unexpected Journey" : "A reluctant hobbit, Bilbo Baggins, sets out to the Lonely Mountain with a spirited group of dwarves to reclaim their mountain home - and the gold within it - from the dragon Smaug."

4- Búsqueda por Texto / Text Search

1. Encontrar las películas que en la sinopsis contengan la palabra "Bilbo".

5- Eliminar Documentos

1. Eliminar la película "Pee Wee Herman's Big Adventure".

6- Alta Disponibilidad: Replicación y Sharding

1. Diseñar un cluster con tres nodos (primary, secondary, arbiter) utilizando la colección cargada en el punto anterior.
2. Configurar el cluster con replicación. Mostrar con printscreens y distintas consultas vistas en clase el correcto funcionamiento del cluster diseñado.
3. Justificar adecuadamente cuál podría ser la sharding key correspondiente si nos solicitaran aplicar sharding.
4. Mostrar el funcionamiento del failover.
5. Establezca si las siguientes afirmaciones son verdaderas o falsas. Justifique:
 - a. Aplicar sharding implica dotar a una infraestructura de HA (High Availability).
 - b. Una arquitectura que presenta escalamiento vertical permite poder ejecutar procesos Map-Reduce en paralelo.
 - c. La presencia de bases NoSQL en una solución son sinónimo de que estamos ante una problemática de Big Data.
 - d. La base de config no es requisito mandatorio que cuente con un esquema de HA.
 - e. MongoDB soporte únicamente un esquema de replicación Master-Slave.

B- Neo4j: CRUD

Dado el requerimiento detallado a continuación:

Se precisa diseñar un blog de noticias donde los usuarios registrados puedan publicar sus comentarios:

- Cada autor tiene un nombre, un nombre de usuario (único) y una cuenta de Twitter (única). Además, de forma opcional, los usuarios pueden tener un teléfono de contacto.
- Las noticias tienen un título, un cuerpo y una fecha de publicación. Son publicadas por un autor. Además, los usuarios pueden puntuar las noticias publicadas por otros usuarios.
- Las noticias reciben comentarios, quedando registrado la persona que lo escribió, el comentario escrito y el momento en el que lo hizo.

1- Diseño del Modelo

1. Confeccionar y especificar el modelo de datos necesario para resolver el requerimiento.
2. Gráfico del modelo.

2- CRUD

1. Cargar los datos que usted genere según el modelo definido en el punto 1.
2. Luego, resolver las siguientes consultas. Colocar printscreen de cada una de ellas + resultado de la misma:
 - a. Devolver las noticias cuyo título termine en "o".
 - b. Devolver los usuarios en los que exista el campo "teléfono".
 - c. Devolver los usuarios que, o bien tengan número de teléfono o bien, se llamen Miguel.
 - d. Devolver las noticias publicadas en marzo o abril.
 - e. Devolver las noticias publicadas entre los años 2019 y 2020.
 - f. Devolver los datos de todas las relaciones "PUNTUA" con puntuación igual a 4.
 - g. Devolver los datos de las relaciones "REDACTO" de las noticias publicadas en 2020 cuyo autor se llame "Laro".
 - h. Devolver las cuenta de twitter de todos los usuarios que hayan puntuado noticias de 2018.