

# Bases de Datos 2018

## Laboratorio #2

Matias Bordone Carranza, Sergio Canchi, Ramiro Damasi,  
Ezequiel Orbe y Jorge Rafael

### INTRODUCCIÓN

ShowHub la startup de servicios de streaming basado en suscripción ya lleva un tiempo operando y el negocio comienza a despegar. La base de usuarios se triplica semana a semana y el servicio comienza a disputarle market share a Netflix.

Sin embargo, este crecimiento explosivo comienza a desnudar algunos problemas en la infraestructura de la empresa la cual no puede escalar de la forma que el negocio lo requiere.

Uno de los servicios más afectados es el CIS (Content Information System) el cual provee la información sobre los contenidos que **se presenta en pantalla al usuario**.

Este servicio cuenta con una capa de persistencia implementada con una base de datos relacional **fuertemente normalizada**, cuyos contenidos se sirven a través de una API REST.

El equipo a cargo del CIS, ha detectado que el cuello de botella del servicio está dado por la base de datos, lo cual a su vez impacta directamente en la latencia de la API.

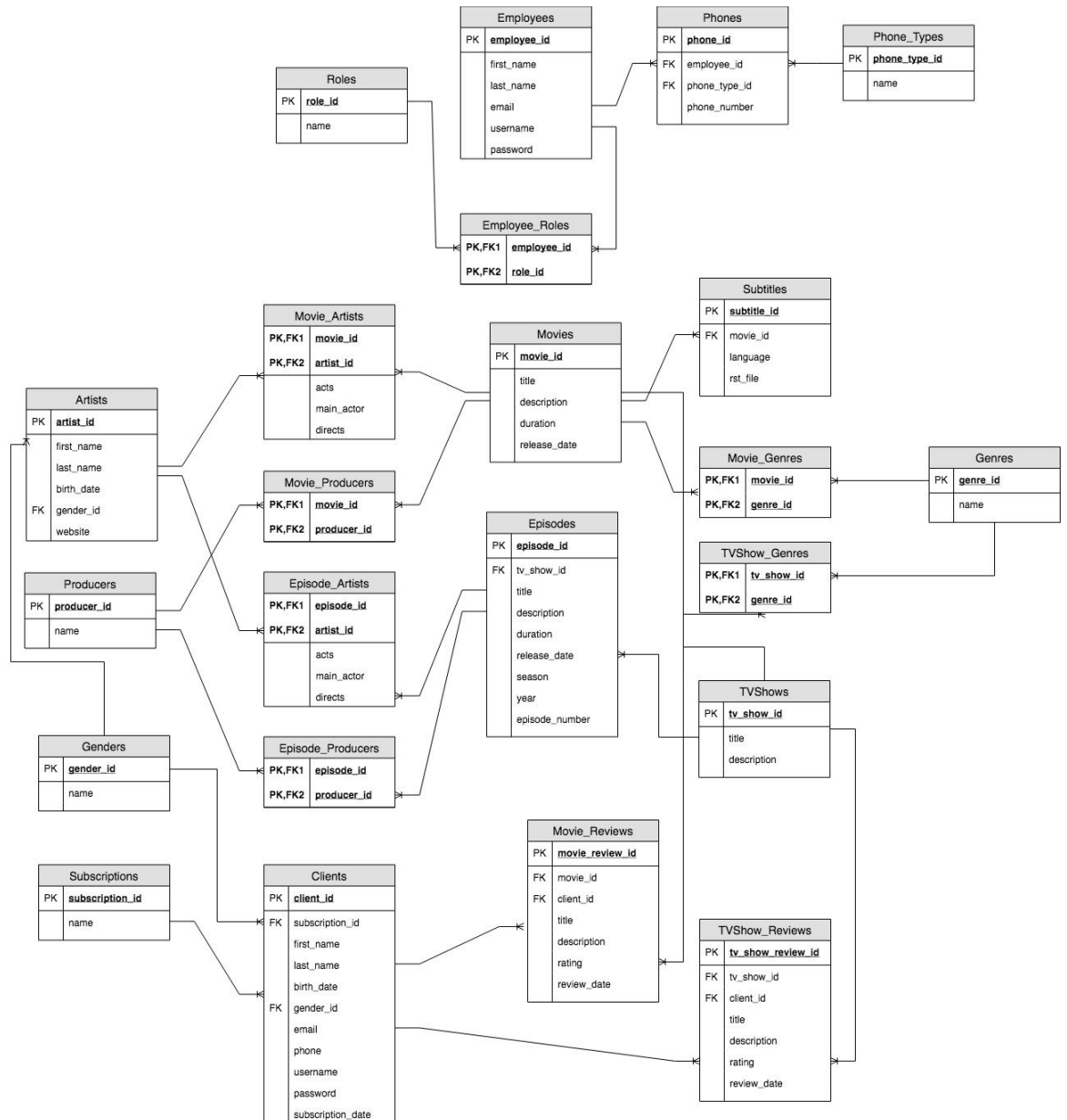
Además de los problemas de escalabilidad, el equipo ha determinado que precisa flexibilizar el esquema de la base de datos para que permita incorporar nueva información que es requerida por los usuarios e iterar rápidamente a medida que se registren nuevas necesidades.

Para resolver estos problemas (mejorar la escalabilidad y ganar flexibilidad en el servicio), el equipo ha decidido reimplementar la capa de persistencia usando MongoDB, una base de datos NoSQL.

Como el equipo no tiene experiencia con este tipo de tecnologías, deciden volver a contratarlos para que realicen el modelado de la nueva base de datos y de esta forma reducir el tiempo de desarrollo.

# Tareas

1. El siguiente diagrama presenta el diseño actual de la base de datos relacional del CIS.



Usando el mismo como referencia, defina las colecciones y los documentos que las mismas contendrán.

2. Genere un dataset con datos para cada una de las colecciones.

3. Usando el dataset generado, responda las siguientes queries usando el shell provisto por Mongo.
- Listar los clientes suscritos al plan PREMIUM en un determinado rango de fechas.
  - Listar los datos de las películas donde el actor 'X' fue protagonista.
  - Listar los episodios correspondientes a un programa de televisión X y un número de temporada N. Listar ordenados por fecha de lanzamiento.
  - Listar los reviews hechos por un cliente X dentro de un rango de fechas.
  - Listar los actores que hayan actuado en películas que pertenecen a géneros de drama o comedia, ordenados por orden alfabético.
  - Dada una película X, calcular para cada estrella el porcentaje de calificaciones recibidas en sus reviews.
  - Dada una película X, calcular su "calificación promedio".
  - Listar las películas mejor calificadas en los últimos 6 meses. Usar la "calificación promedio".
  - Listar las películas dirigidas por dos o más directores femeninos.
  - Listar cantidad de películas por género en orden descendente.
  - Listar los actores que hayan actuado en todos los programas de televisión en donde actuó un actor X.

## Entrega

Fecha de Entrega: 01/11 23:59hs

## Repositorio

En el repositorio provisto por la materia deberán crear la siguiente estructura:

**lab-2/**

**colecciones/**

**coleccion-x.js**

...

**consultas/**

**consulta-a/**

**consulta.js**

...

**consulta-n/**

**Taggear el commit de entrega con el tag: lab-2**