



## Tarea de Laboratorio 2: MongoDB INF-325 Bases de Datos Avanzadas

Profesor: Mauricio Figueroa Colarte

11 de octubre de 2024

### 1 CONTEXTO

Se requiere diseñar, poblar y consultar una base de datos en **MongoDB** que almacena los datos de juegos<sup>1</sup> de la plataforma **Steam**<sup>2</sup>, para lo cual se dispone del *dataset* llamado **games.json**. Esta fuente de datos es un JSON que recopila 85.103 documentos con los datos específicos de videojuegos para PC.

La estructura de los documentos que almacena la colección de documentos games es la siguiente:

```
{
  "_id": "ObjectId('6631d344cba7716b183e534f')",
  "name": "DOOM Eternal: The Ancient Gods - Part Two",
  "release_date": "Mar 18, 2021",
  "required_age": 0,
  "price": 19.99,
  "dlc_count": 0,
  "detailed_description": "The Ancient Gods - Part Two is the epic conclusion to the DOOM Slayer...",
  "about_the_game": "The Ancient Gods - Part Two is the epic conclusion to the DOOM Slayer...",
  "short_description": "The Ancient Gods - Part Two is the epic conclusion to the DOOM Slayer...",
  "reviews": "",
  "header_image": "https://cdn.akamai.steamstatic.com/steam/apps/1098293/header.jpg?t=169...",
  "website": "https://slayersclub.bethesda.net/en",
  "support_url": "http://help.bethesda.net/",
  "support_email": "",
  "windows": true,
  "mac": false,
  "linux": false,
  "metacritic_score": 0,
  "metacritic_url": "",
  "achievements": 0,
  "recommendations": 3860,
  "notes": "Similar to DOOM from 2016, DOOM Eternal will contain Blood and Gore, I...",
  "supported_languages": Array,
  "full_audio_languages": Array,
  "packages": Array
}
```

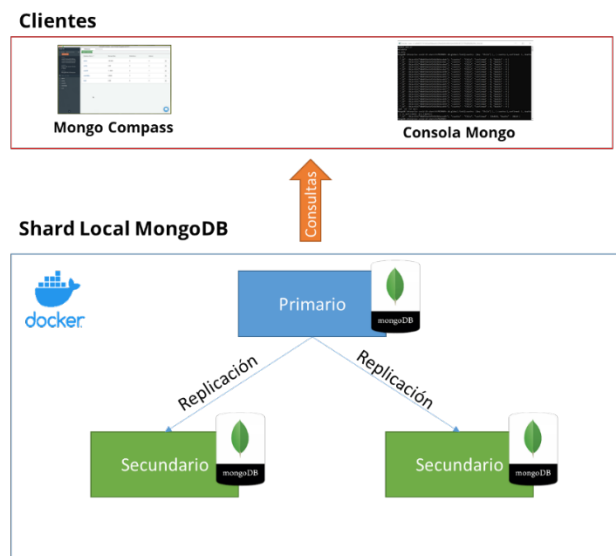
**Figura 1:** Estructura de documento de juego Steam

<sup>1</sup> Dataset y explicación de los datos en : <https://www.kaggle.com/datasets/fronkongames/steam-games-dataset?resource=download>

<sup>2</sup> <https://store.steampowered.com/?l=spanish>

## 2 REQUISITOS

1. Implementar la arquitectura de *Cluster* con un set de 3 réplicas. Una de las máquinas será el primario y las otras dos serán los secundarios. Esta configuración deberá estar operativa en su máquina local, donde se levantarán las 3 instancias de **Docker** con cada uno de los servicios de **MongoDB** con la redundancia de la Base de Datos llamada **"Steam"** que contendrá la colección de documentos llamada **"Games"**. En la siguiente imagen se muestra la arquitectura requerida.



**Figura 2:** Arquitectura General

2. En base al dataset **games.json** entregado, implementar las siguientes consultas requeridas para resolver las preguntas del negocio utilizando el cliente **MongoDB Compass** :
  - a. Devolver los nombres y fechas de lanzamiento de juegos que se lanzaron en el año 2018. Usa una proyección para mostrar solo esos campos y limita los resultados a los primeros 20 juegos.
  - b. Devolver los nombres y el número de contenidos descargables de juegos que tengan un precio entre \$10 y \$50 y al menos un contenido descargable. Ordena los resultados por el número de DLC de manera descendente y limita a los primeros 10.
  - c. Devolver el nombre, la edad requerida y la calificación de los usuarios (si existe) para todos los juegos aptos para mayores de 12 años. Incluye una condición para que solo aparezcan aquellos juegos que tienen una calificación de usuarios mayor de 80. Limita los resultados a 15 juegos.



3. En base al dataset **games.json** entregado, implementar las siguientes consultas requeridas para resolver las preguntas del negocio utilizando **línea de comando de MongoDB**:
  - a. Identifica los 10 juegos más caros lanzados después del año 2015 para analizar las tendencias de precios en lanzamientos recientes.
  - b. Determina los 5 juegos con mejor calificación de usuarios entre los que son aptos para todas las edades (`required_age = 0`), para entender qué características valoran más los usuarios en estos juegos.
  - c. Analiza la relación entre el precio de los juegos y la calificación de los usuarios, clasificando los juegos en tres categorías de precio: bajo (hasta \$10), medio (entre \$11 y \$30), y alto (más de \$30). Muestra los 5 juegos con mejores calificaciones en cada categoría para determinar si el precio influye en la percepción de calidad del usuario.
4. Una vez realizados los puntos anteriores, se pide mostrar evidencia objetiva que permita demostrar la alta **disponibilidad** del clúster, por ejemplo, bajando el servidor primario para que responda alguno de los secundarios **sin que el usuario se percate del problema.**