

# Proyecto Scraping

*DESARROLLO DE APLICACIONES  
PARA CIENCIA DE DATOS*

*CIENCIA E INGENIERÍA DE DATOS*

*UNIVERSIDAD DE LAS PALMAS DE  
GRAN CANARIA*

**LAURA LASSO GARCÍA**  
**13/01/2023**  
**VERSIONES: 1.0-SANPSHOT**

## Resumen

El proyecto se divide en varias funcionalidades; tenemos 3 interfaces;

**HotelDatabase**→ Va a implementar las funciones encargadas de insertar los datos, que están en clase POJO, en las correspondientes tablas de las bases de datos.

**HotelSource**→ Es la interfaz que se va a encargar de deserializar los datos, pasándolos a las correspondientes clases POJO. Cada función que implemente devolverá esos datos deserializados.

**jsonSource**→ Es la interfaz que se encargará de serializar los objetos que se pasen como parámetros a la función que implementa. Esta función servirá para la api a desarrollar.

En general, el funcionamiento sería el siguiente: Cogeríamos los datos de la página web (la clase Scrapping se encarga de ello) y los pasamos a clases POJO. Luego escribiríamos cada objeto de cada clase en las correspondientes tablas de las bases de datos (SqliteHotelDatabase) y por último tenemos la clase que define las peticiones GET de la API (la clase HotelAPI).

## Índice

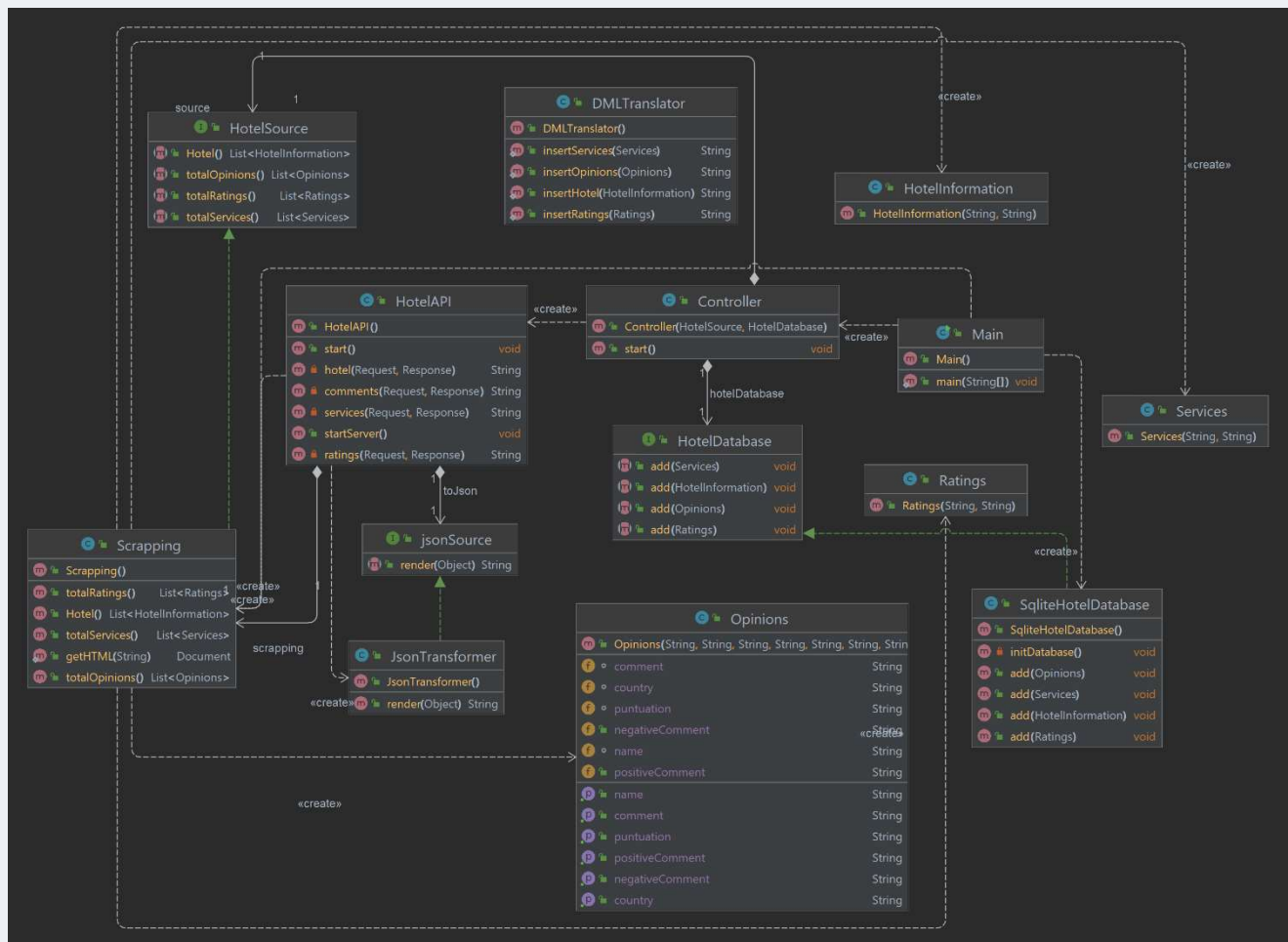
Recursos utilizados.....	5
Diseño.....	5
Conclusiones.....	6
Líneas futuras.....	6
Bibliografía.....	6

## Recursos utilizados

El entorno de desarrollo es IntelliJ e IntelliJ Ultimate y las herramientas de documentación es Word.

## Diseño

Los principios implementados son los SOLID para así mantener un código limpio, ordenado, fácil de mantener de manera que podamos facilitar la nueva entrada de funcionalidades a nuestra aplicación y corrección de errores.



## Conclusiones

Este proyecto se puede enfocar de múltiples formas; no usando las clases POJO o implementar menos módulos. Sin embargo, he optado por usar más clases para que el código sea mucho más fácil de comprender y que de esta forma esté más ordenado y sencillo.

En el caso de no haber usado las clases POJO, lo que se habría hecho es, en vez inicializar nuevos objetos con los valores que se cogen, se pasarían como parámetros a una función que se encargaría de insertar esos datos a las tablas de la base de datos mediante sentencias SQL. Aquí también podemos elegir en usar un módulo que se encargue de traducir esas sentencias SQL y luego pasarlas con los valores a insertar para que se pueda ejecutar en la función, o directamente haber usado un tipo de sentencia SQL llamadas sentencias SQL preparadas pasando los valores que se quieran insertar.

## Líneas futuras

En cuestión de los usos que se pueden dar a este proyecto, es el hecho de poder scrapear una página web y coger la información, ponerla en una base de datos para que pueda ser útil en un futuro. En cuanto al scrapping de una página como booking, tiene múltiples usos, ya que esos datos se pueden usar para poder comparar las opiniones de la gente y así los clientes, puedan elegir un hotel adaptado a sus capacidades. También se podrían comparar las puntuaciones de los servicios para saber cuál es el hotel mejor puntuado. En general, se le puede dar uso en cuando a una aplicación encargada de la comparación, ya sea de opiniones, puntuaciones, etc.

## Bibliografía

[Principios del diseño de clases - Adictos a la Informática \(adictosalainformatica.com\)](http://adictosalainformatica.com)  
[Doce principios de diseño que todo desarrollador debería conocer \(genbeta.com\)](http://genbeta.com)