

13 de enero de 2023

SCRAPER DE BOOKING

Desarrollo de aplicaciones para la
ciencia de datos

Booking.com

- **CURSO:** Segundo.
- **TITULACIÓN:** Grado en Ciencia e Ingeniería de Datos.
- **ESCUELA:** Escuela de Ingeniería Informática.
- **UNIVERSIDAD:** Universidad de Las Palmas de Gran Canaria.

RESUMEN

En este proyecto se ha realizado un web service que ofrece dato de una página web HTML (en este caso la página web de Booking) y se transforman a JSON. Este scraper nos permite recopilar información de manera automática de todos los hoteles de un destino concreto. En mi caso selecciono el destino de Booking Tenerife sur.

El trabajo está dividido en tres paquetes. En primer lugar, tenemos el módulo “scraper”, mediante el cual realizamos scrapping a la página y almacenamos la información de los distintos hoteles. En segundo lugar, el paquete “controller” que se encarga de guardar los datos recolectados mediante el scrapping a la Api. Finalmente, el módulo “api” que forma una conexión localhost implementando una API mediante la cual obtenemos información sobre diversos hoteles obtenida del scraper, como la valoración, localización, comentarios, etc.



ÍNDICE

| | |
|-------------------------------|---|
| 1. Recursos utilizados. ----- | 3 |
| 2. Diseño. ----- | 4 |
| 3. Conclusiones. ----- | 5 |
| 4. Líneas futuras. ----- | 5 |
| 5. Bibliografía. ----- | 5 |

RECURSOS UTILIZADOS

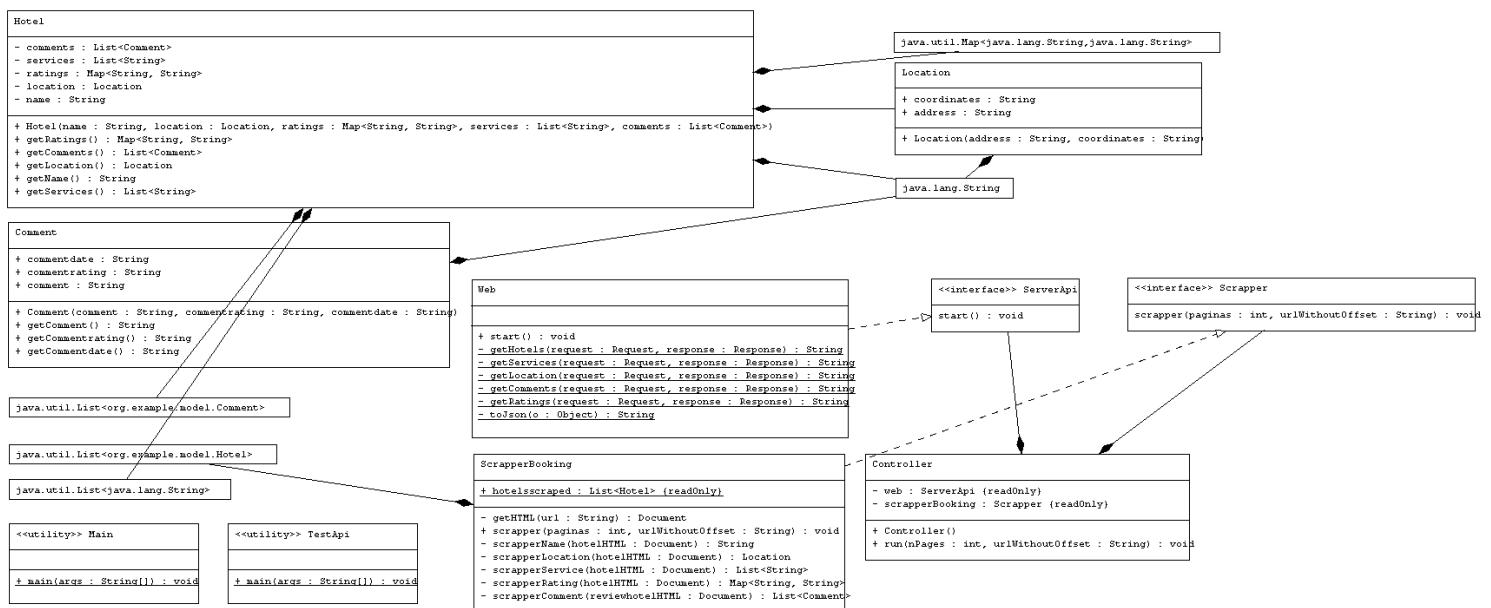
Para este proyecto he utilizado “IntelliJ IDEA Community Edition 2022.2.2” como entorno de desarrollo. Para realizar esta memoria, he utilizado el procesador de texto “Microsoft Word”. Para controlar las versiones usé “Git”, que es un software de control de versiones diseñado por Linus Torvalds, pensando en la eficiencia, la confiabilidad y compatibilidad del mantenimiento de versiones de aplicaciones cuando estas tienen un gran número de archivos de código fuente. También el “GitHub” donde he creado un repositorio donde se encuentra este proyecto.



DISEÑO

Como comenté anteriormente, este proyecto se basa en tres módulos distintos, cada uno con una funcionalidad. El paquete “scrapper” que se encarga de realizar el scrapping para obtener la información sobre los hoteles. El paquete “controller” que almacena esta información en la API, a la que podremos acceder gracias al paquete “api”.

Además, he intentado seguir el diseño MVC exceptuando la parte de View, separando el proyecto en modelo y controlador. Respecto al diseño, he seguido los principios SOLID como el principio de responsabilidad única, es decir, las clases tienen una sola responsabilidad. Además del principio de inversión de dependencias, que se basa en que los módulos de alto nivel no dependan de módulos de bajo nivel.



CONCLUSIONES

Me ha gustado bastante este trabajo ya que me parece algo que tiene mucha utilidad comercial y es muy interesante. Además, también nos resulta muy útil para la asignatura de Fundamentos del Marketing y Comportamiento del Consumidor porque es algo que estudiamos en esa asignatura también.

LÍNEAS FUTURAS

Para comercializar este proyecto tendríamos que comparar los datos de los hoteles y así poder seleccionar el mejor hotel de la zona deseada para recomendar una estancia a los clientes. También habría que realizar una interfaz para que los usuarios puedan encontrar el hotel de manera fácil.

BIBLIOGRAFÍA

<https://stackoverflow.com/>

<https://www.baeldung.com/>



ALEJANDRO BOLAÑOS GARCÍA

13 de enero de 2023