

PFM

HOTEL ADVISOR

MANUAL DE ARQUITECTURA Y DISEÑO DEL SISTEMA

AUTOR: FERNANDO MÁRQUEZ ARROYO



KSCHOOL

ÍNDICE

1. DESCRIPCION DEL SISTEMA.....	2
2. DISEÑO DEL SISTEMA.....	3
2.1 DATOS DE ENTRADA.....	3
2.1.1 HOTELES.....	3
2.1.2 POIS.....	4
2.2 ARQUITECTURA DEL SISTEMA.....	5
2.2.1 DIAGRAMA DE ARQUITECTURA.....	5
2.2.2 CAPA DE PROCESAMIENTO (PROCESSING LAYER).....	6
2.2.3 CAPA DE ALMACENAMIENTO (DATA STORAGE LAYER).....	8
2.2.4 CAPA DE EXPLOTACIÓN (APPLICATION LAYER).....	14
ANEXO: LISTADO DE CATEGORÍAS DE TIPOS DE POI.....	15

1. DESCRIPCION DEL SISTEMA

Me apasiona viajar (soy blogger de viajes) y recientemente al ser partner de la empresa Booking.com he tenido acceso a información de 1.000.000 de hoteles en los 5 continentes (ciudad, país, longitud, latitud, precio por noche, categoría, descripciones en varios idiomas, ...).

Por otro lado, existen datasets con información geolocalizada de alrededor de 14.000.000 de puntos de interés (POIs) de todo el mundo (restaurantes, tiendas, monumentos turísticos, museos, hospitales, paradas de autobús,...) obtenidos a partir de la herramienta Open Street Maps (.osm).

El proyecto consiste en implementar un recomendador de hoteles utilizando técnicas de Big Data para cruzar información procedente de ambos conjuntos de datos, ofreciendo al usuario aquellos hoteles más afines a sus preferencias.

El recomendador dispondrá de varias opciones:

1) Búsqueda tradicional de hoteles basado en diferentes criterios (ciudad, puntuación mínima de hotel, categoría del hotel, precio mínimo por noche, ...).

El recomendador mostrará los hoteles que cumplen los filtros especificados, así como un listado de POIs cercanos a cada uno de los hoteles, de acuerdo a sus preferencias (número máximo de POIs a mostrar por hotel, tipos de POIs, ...)

2) Búsqueda avanzada de hoteles basado en:

1. Búsqueda por características/términos (piscina, sauna, terraza, admite mascotas ...)
2. Búsqueda por textos que aparecen en los comentarios (baño sucio, alejado del centro, ...)
3. Búsqueda de hoteles similares a uno dado

3) Búsqueda de POIs cercanos a un hotel dado. El recomendador devolverá los POIs que se encuentran entre una distancia mínima y máxima del hotel elegido, de acuerdo a las preferencias del usuario.

4) Búsqueda de hoteles a partir de un POI determinado. El recomendador devolverá los hoteles que se encuentran entre una distancia mínima y máxima del POI elegido, de acuerdo a las preferencias del usuario.(p. ej. hoteles a menos de 500 metros del Santiago Bernabeu, hoteles ubicados entre 500 metros y 1 km de las Pirámides de Giza, ...)

5) TOP Hoteles. El sistema mostrará cuáles son los hoteles más caros, con más comentarios, con la puntuación más alta,...

6) Críticas (reviews) de hoteles. El usuario podrá realizar una crítica de un hotel determinado, asignando notas entre 0 y 10 a diversos aspectos (localización, limpieza, confort, instalaciones, personal, wifi, relación calidad/precio), así como un comentario general de texto libre, en el idioma deseado, acerca de su estancia en el hotel. Estas notas afectarán a la puntuación total del hotel para los filtros por puntuación.

Asimismo, cuando se presentan los hoteles recomendados, se mostrará el número de comentarios del hotel, y se tendrá posibilidad de acceder al detalle de cada uno de los comentarios sobre el hotel.

7) Consulta de críticas personales. El usuario tendrá acceso a un listado de sus críticas de hoteles, mostrándose todas las críticas realizadas, así como las notas de los diversos aspectos asociados al hotel.

8) Modificación de preferencias de usuario: El usuario podrá modificar aspectos relativos a sus datos personales (nombre, apellido, password, email, edad, ..) así como sus preferencias (idioma, distancia mínima y máxima de pois, tipos de POIs en los que está interesado,...).

Adicionalmente, dado que el dataset de hoteles no dispone de información acerca de comentarios ni puntuaciones de los mismos, y dado el elevado número de hoteles para escribir las reviews de forma manual, se precisa de un módulo encargado de generar comentarios y notas aleatorias sobre los hoteles, simulando eventos de entrada al recomendador a partir de un sistema externo.

2. DISEÑO DEL SISTEMA

2.1 DATOS DE ENTRADA

El sistema consta de dos conjuntos de datos de entrada:

2.1.1 HOTELES

El dataset de hoteles consta de varios ficheros .TSV (campos separados por tabuladores) por continente, con la siguiente información:

id: Identificador del hotel. Numérico entero. (Ejemplo: 1592340)

name: Nombre del hotel. Texto. (Ejemplo: Sheffield Dreams)

address: Dirección del hotel. Texto (Ejemplo: 47 Gail Road Sheffield Beach)

zip: Código postal del hotel. Texto (Ejemplo: 4420)

city_hotel: Ciudad del hotel. Texto (Ejemplo: Ballito)

cc1: Código del país del hotel. Texto (Ejemplo: za)

ufi: Código UFI del hotel. Numérico (Ejemplo: -1208880)

class: Categoría del hotel. Numérico real. (Ejemplo: 0.0)

currencycode: Moneda del hotel. Texto. (Ejemplo: ZAR)

minrate: Precio mínimo de habitación. Numérico real. (Ejemplo: 2500)

maxrate: Precio máximo de habitación. Numérico real. (Ejemplo: 4000)

preferred: N/A. Numérico entero (Ejemplo: 1)

nr_rooms: Número de habitaciones. Numérico entero. (Ejemplo: 24)

longitude: Coordenada de la longitud del hotel. Numérico real. (Ejemplo: 31.2485182285309)

latitude: Coordenada de la latitud del hotel. Numérico real. (Ejemplo: -29.4874479032936)

public_ranking: Ranking público del hotel. Numérico entero (Ejemplo: 1) (siempre establecido a 1)

hotel_url: URL con la página web de Booking del hotel. Texto
(Ejemplo: <http://www.booking.com/hotel/za/sheffield-dreams.html>)

photo_url: URL con la imagen del hotel. Texto
(Ejemplo: <http://aff.bstatic.com/images/hotel/max500/619/61963454.jpg>)

desc_en: Descripción del hotel en inglés. Texto
(Ejemplo: Sheffield Dreams offers pet-friendly accommodation in Ballito. Providing free private parking, the holiday home is 900 metres from Crocodile Farm)

desc_fr: Descripción del hotel en francés. Texto

desc_es : Descripción del hotel en español. Texto

desc_de: Descripción del hotel en alemán. Texto

desc_nl: Descripción del hotel en holandés. Texto

desc_it: Descripción del hotel en italiano. Texto

desc_pt: Descripción del hotel en portugués. Texto

desc_ja: Descripción del hotel en japonés. Texto

desc_zh: Descripción del hotel en chino. Texto

desc_pl: Descripción del hotel en polaco. Texto

desc_ru: Descripción del hotel en ruso. Texto

desc_sv: Descripción del hotel en sueco. Texto

desc_ar: Descripción del hotel en árabe. Texto

desc_el: Descripción del hotel en griego. Texto

desc_no: Descripción del hotel en noruego. Texto

city_unique: Ciudad única del hotel. Texto (Ejemplo: balitoville)

city_preferred: Ciudad preferida para el hotel. Texto (Ejemplo: Ballito)

continent_id: Identificador del continente del hotel. Numérico entero (Ejemplo: 5)

review_score: Puntuación del hotel. Numérico real (Ejemplo: 1) (siempre establecido a 1)

review_nr: Número de comentarios del hotel. Numérico entero (Ejemplo: 1) (siempre establecido a 1)

El dataset se ha obtenido directamente a través del web site de Booking para su programa de afiliados.

2.1.2 POIS

El dataset de POIS consta de un fichero .CSV (campos separados por el símbolo |) por continente, con la siguiente información:

category_name: Categoría del POI. Numérico entero (Ejemplo: 85)

Ver Anexo: Listado completo de categorías de POI

osmid: Identificador del POI en Open Street Maps (OSM). Texto (Ejemplo: N204648)

latitude: Coordenada de la latitud del POI. Numérico real (Ejemplo: -41.2887639)

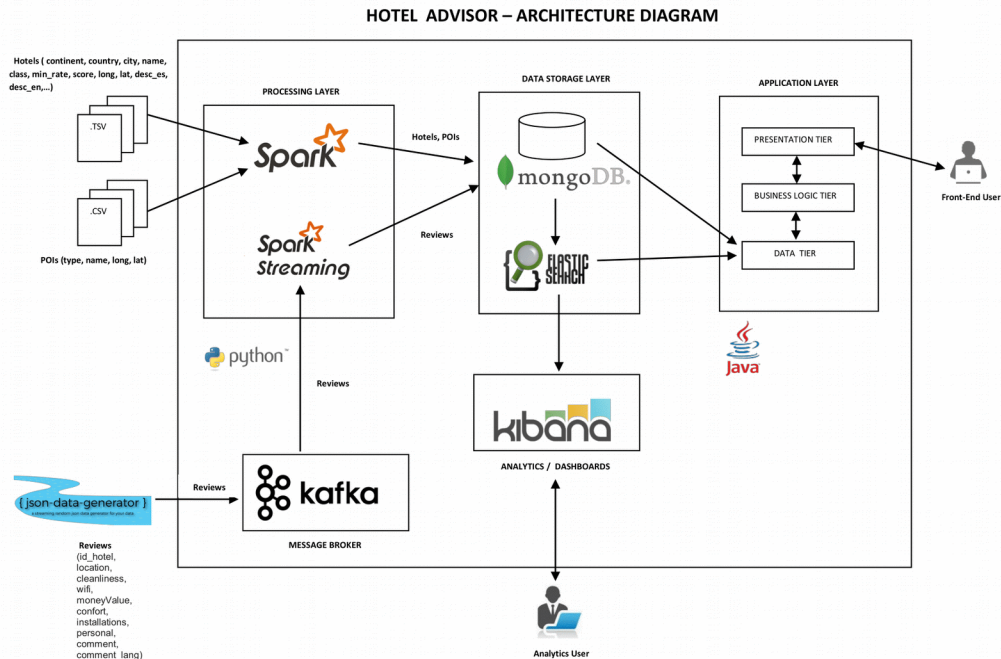
longitude: Coordenada de la longitud del POI. Numérico real (Ejemplo: 174.7772239)

poi_name: Nombre del POI. Texto (Ejemplo: Hawkins Entertainment and Conference Centre)

El dataset de POIS se ha obtenido gracias a la aplicación: <https://github.com/MorbZ/OsmPoisPbf> , la cual extrae todos los POIS de un fichero binario de Open Street Maps (.pbf) y genera un fichero CSV con los resultados.

2.2 ARQUITECTURA DEL SISTEMA

2.2.1 DIAGRAMA DE ARQUITECTURA



2.2.2 CAPA DE PROCESAMIENTO (PROCESSING LAYER)

2.2.2.1 PROCESAMIENTO BATCH (BATCH PROCESSING)

Se trata de la capa encargada del parseo, limpieza y filtrado de los datos de entrada, y posterior almacenamiento en base de datos. El sistema utiliza la tecnología Spark, dada su madurez y altas prestaciones. El lenguaje de programación utilizado para la implementación de esta capa es Python, comunicándose con Spark a través del Spark Python API (PySpark).

Este módulo consta de dos clases:

- * **HotelReader.py:** Encargada de realizar las siguientes acciones:
 - * Accede a los ficheros .tsv con el dataset de los hoteles de cada uno de los continentes.
 - * Extrae la cabecera y realiza el parseo de los ficheros, filtrando aquellos campos de interés para el recomendador, realizando conversión de códigos de país y continente, convirtiendo el valor de la moneda local del hotel a euros, ...
 - * Registra el número de hoteles parseados con éxito y con error, mostrando un listado de los hoteles erróneos de los ficheros.

* Inserta los hoteles en la base de datos de la aplicación (MongoDB) a través del API de Mongo para Python (PyMongo).

* **PoiReader.py:** Encargada de realizar las siguientes acciones:

- * Accede a los ficheros .csv con el dataset de los POIs de cada uno de los continentes.
- * Realiza el parseo de los ficheros, realizando la conversión de tipos de POI a su categoría.
- * Registra el número de POIs parseados con éxito y con error, mostrando un listado de los POIs erróneos de los ficheros.
- * Inserta los POIs en la base de datos de la aplicación (MongoDB) a través del API de Mongo para Python (PyMongo).

2.2.2.2 PROCESAMIENTO REAL TIME (REAL TIME PROCESSING)

Para simular la entrada de comentarios y puntuaciones de hoteles desde un sistema externo, se ha dotado al sistema de un módulo de procesamiento en tiempo real.

Este módulo consta de tres componentes:

2.2.2.2.1 GENERADOR DE EVENTOS - JSON DATA GENERATOR

Los eventos con las notas aleatorias y comentarios de hoteles se generan con el simulador de eventos: Json Data Generator, introduciendo los eventos en un topic de Kafka.

La entrada del Json Data Generator es un fichero json de configuración donde se define el fichero de workflow donde se recogen los campos a generar, así como el nombre del topic de Kafka al que se van a hacer llegar los eventos generados:

* **HotelReviewConfig.json:**

```
{
  "workflows": [{
    "workflowName": "HotelReview",
    "workflowFilename": "HotelReviewWorkflow.json"
  }],
  "producers": [{
    "type": "kafka",
    "broker.server": "localhost",
    "broker.port": 9092,
    "topic": "hotel_reviews",
    "flatten": false,
    "sync": false
  }]
}
```

* **HotelReviewWorkflow.json,** genera eventos con una determinada frecuencia, asignando notas y comentarios a hoteles escogidos de forma aleatoria:

```
{
  "eventFrequency": 100,
  "varyEventFrequency": true,
  "repeatWorkflow": true,
  "timeBetweenRepeat": 50,
  "varyRepeatFrequency": true,
  "steps": [{
    "config": {
      "date": "nowTimestamp()",
      "hotel_id": "integer(1,971478)",
      "age": "integer(18,90)",
      "username": "admin",
      "location": "double(3.0,10.0)",
      "cleanliness": "double(3.0,10.0)",
      "comfort": "double(3.0,10.0)",
      "installations": "double(3.0,10.0)",

```

```

        "personal": "double(3.0,10.0)",
        "wifi": "double(3.0,10.0)",
        "valueMoney": "double(3.0,10.0)",
        "lang": "es",
        "comment": "random('Hotel encantador. Personal muy amable. Baño amplio y cama cómoda. Barrio Tranquilo','Buen hotel. Algo de ruido en las habitaciones. Baño bueno. Desayuno normal. Personal de recepción muy amable y resolvió todas nuestras dudas','Hotel normal. Buena cama pero baño pequeño. Recepcionista diligente.','Hotel mediocre. Cama sucia. Baño pequeño. El barrio da miedo. El tipo de la recepción era muy borde','No es centrico pero hay dos paradas muy cerca que hace que puedas moverte muy rapido a las zonas de interes','El Hotel está muy bien, un poco caro pero cumple con lo pagado. El desayuno buenísimo aunque también algo caro. La única pega es que está alejado del centro y hay que ir en tranvía siempre a la ciudad. En las inmediaciones del Hotel no hay ningún servicio. Tiene parking propio. Si tuviera que volver no elegiría este Hotel por la lejanía del centro. Por el resto el Hotel es fantástico aunque demasiado lujoso para mis intereses','Moderno. Jovial. Limpio. Excelente instalaciones a pasos de un tram que te lleva al centro en unos minutos! El mejor precio y lo mejor que se ofrece! volveria definitivamente')"
```

```

    },
    "duration": 0
  }]
}

```

2.2.2.2.2 BROKER DE MENSAJERIA (MESSAGE BROKER) - KAFKA

El sistema utiliza una cola de mensajes Kafka, para encolar en un topic, los eventos generados mediante el simulador de eventos anterior. Esta tecnología permite gestionar streams de datos, eficientemente y en tiempo real.

2.2.2.2.3 SPARK STREAMING

El sistema se apoya en la tecnología Spark Streaming, para conectarse al topic de Kafka al que llegan las notas y comentarios de los hoteles, y realizar el procesamiento en tiempo real de dicha información.

Este componente consta de una clase:

- * **KafkaHotelReviewStreaming.py**: Encargada de realizar las siguientes acciones:
 - * Crea el contexto de Spark con un segundo de tamaño de batch.
 - * Se conecta al topic de Kafka para recoger los eventos generados aleatoriamente por el Json Data Generator.
 - * Parsea el json que contiene la información del evento (review del hotel).
 - * Obtiene la nota media de la review del hotel, sumando la puntuación de cada uno de los aspectos valorados del hotel.
 - * Añade la review a la lista de reviews del hotel, modifica el número de reviews, así como la nota media del hotel en base a la nueva review, en la base de datos de la aplicación (MongoDB) a través del API de Mongo para Python (PyMongo).

2.2.3 CAPA DE ALMACENAMIENTO (DATA STORAGE LAYER)

Esta capa es la encargada del almacenamiento persistente de la información utilizada en la aplicación.

2.2.3.1 MONGODB

El sistema utiliza MongoDB, una base de datos intuitiva, orientada a documentos y flexible. Ofrece fácil integración tanto con Python como con Java y dispone de un potente módulo para poder lanzar consultas geoposicionadas.

2.2.3.1.1 HOTELES

La información sobre hoteles parseada de los ficheros a través de la capa de procesamiento batch con Spark son almacenados en la colección de MongoDB: "hotels". Como decisión de diseño, se ha optado por almacenar las reviews (críticas) de los hoteles como un objeto embebido "reviews" dentro de la colección de hoteles. Esta información sobre las reviews es almacenada tanto por la capa de procesamiento real time con Spark Streaming, así como por la tier de persistencia del Application Layer cuando el usuario crea una review de forma manual desde la aplicación.

```

{
  "_id" : ObjectId("58c3ba481d41c81521d748c2"),
  "review_nr" : 1,

```



```

"hotel_url" : "http://www.booking.com/hotel/ug/capricon-executive.html",
"city_preferred" : "Kabale",
"id" : 1,
"max_rate_EUR" : null,
"nr_rooms" : 40,
"min_rate_EUR" : 20.6036880601628,
"zip" : "",
"cc1" : "Uganda",
"minrate" : 22.0,
"review_score" : 8.29,
"ufi" : -1728233,
"desc_en" : "Located 300 metres from Kabale Market in Kabale, Capricon Executive Hotel features free WiFi access and free private parking. Guests can enjoy the on-site restaurant. Each room is equipped with a flat-screen TV with satellite channels.",
"currencycode" : "USD",
"preferred" : null,
"address" : "Plot 72 Kabale - Kisoro Road",
"desc_es" : "",
"class" : 3.0,
"photo_url" : "http://aff.bstatic.com/images/hotel/max500/761/76158347.jpg",
"name" : "Capricon Executive Hotel",
"city_unique" : "kabale",
"continent_id" : "Africa",
"coord" : [
  29.9900762178004,
  -1.25695095340165
],
"maxrate" : null,
"city_hotel" : "Kabale",
"public_ranking" : 1,
"reviews" : [
  {
    "username" : "fer",
    "location" : 8.0,
    "cleanliness" : 8.0,
    "comfort" : 8.0,
    "installations" : 8.0,
    "personal" : 9.0,
    "wifi" : 8.0,
    "valueMoney" : 9.0,
    "lang" : "es",
    "comment" : "Bastante buen hotel",
    "score" : 8.29,
    "date" : NumberLong(1489224141392)
  }
]
}

```

Asimismo, se han introducido índices sobre los siguientes campos, para optimizar las consultas que se realizan desde la capa de explotación:

- * coord (coordenadas geográficas): Para las consultas geoposicionadas
- * id (identificador)
- * name (nombre)
- * cc1 (país)
- * city_hotel (ciudad)
- * continent_id (continente)
- * min_rate_EUR (precio mínimo por noche)
- * review_score (puntuación)
- * nr_rooms (número de habitaciones)

2.2.3.1.2 POIS

La información sobre POIs parseada de los ficheros a través de la capa de procesamiento batch con Spark son almacenados en la colección de MongoDB: "pois"

```
{
  "_id" : ObjectId("58e020c31d41c85b2a768e67"),
  "poi_type" : "POI_HAMLET",
  "poi_osm_id" : "N9829766",
  "poi_name" : "Perla Marina",
  "poi_coord" : [
    -70.4625291,
    19.7790987
  ]
}
```

Asimismo, se han introducido índices sobre los siguientes campos, para optimizar las consultas que se realizan desde la capa de explotación:

- * poi_coord (coordenadas geográficas): Para las consultas geoposicionadas
- * poi_name (nombre)
- * poi_type (tipo)
- * poi_osm_id (identificador)

2.2.3.1.3 USUARIOS

El sistema consta de un pequeño módulo de usuarios para controlar el acceso a la aplicación. Esta información se almacena en la colección de MongoDB: "users"

En esta colección se almacena tanto datos personales del usuario (nombre, apellidos, edad, email, ...) como preferencias del usuario (idioma, distancia mínima y máxima de pois, tipos de POIs en los que está interesado (museos, restaurantes, tiendas, ...)). Asimismo, se almacenan todas las reviews (críticas) de hoteles que ha realizado el usuario, incluyendo la fecha en la que fueron realizadas.

```
{
  "_id" : ObjectId("58ecac6e90b04738412ff1cd"),
  "username" : "pedro",
  "password" : "pedro",
  "name" : "pedro",
  "surname" : "lopez",
  "age" : "29",
  "email" : "plopez@gmail.com",
  "language" : "es",
  "poisMinDistance" : 0.0,
  "poisMaxDistance" : 1000.0,
  "poisNumberLimit" : 20,
  "poisTypes" : "0",
  "topHotelsLimit" : 20,
  "reviews" : [
    {
      "hotel_id" : 67997,
      "hotel_name" : "Nam Pueng",
      "location" : 10.0,
      "cleanliness" : 5.0,
      "comfort" : 10.0,
      "installations" : 6.0,
      "personal" : 10.0,
      "wifi" : 10.0,
      "valueMoney" : 10.0,
      "lang" : "es",
      "comment" : "Sucio, pero la wifi iba como un tiro",
      "score" : 8.71,
      "date" : NumberLong(1491905699548)
    },
    {
      "hotel_id" : 28944,
      "hotel_name" : "Nam Pueng Place 1",
      "location" : 10.0,
      "cleanliness" : 9.0,
      "comfort" : 10.0,
      "installations" : 9.0,

```

```

    "personal" : 2.0,
    "wifi" : 10.0,
    "valueMoney" : 10.0,
    "lang" : "es",
    "comment" : "Me lo pasé genial, aunque el tío de recepción era lo peor.",
    "score" : 8.57,
    "date" : NumberLong(1491906211136)
  }
]
}

```

2.2.3.2 ELASTIC SEARCH

El sistema utiliza el motor de búsqueda de Elastic Search para ofrecer al usuario la posibilidad de realizar búsquedas de texto libre sobre la información almacenada en el sistema. La información es cargada en Elastic Search, a través de un conector de MongoDB (Mongo Connector).

Elastic ofrece un API REST por el que se pueden realizar consultas al motor de búsqueda:

Por ejemplo, para buscar hoteles cerca de una isla con piscina, campo de golf y wifi:

```

curl -XGET 'http://localhost:9200/hoteladvisor/hotels/_search?pretty=true' -d '{"query": {"match_phrase": {"desc_es":
{"query": "piscina golf WiFi isla", "slop": 30}}}}'

```

```

{
  "took" : 3,
  "timed_out" : false,
  "_shards" : {
    "total" : 5,
    "successful" : 5,
    "failed" : 0
  },
  "hits" : {
    "total" : 4,
    "max_score" : 0.6833336,
    "hits" : [
      {
        "_index" : "hoteladvisor",
        "_type" : "hotels",
        "_id" : "58c3bb341d41c81521db7edd",
        "_score" : 0.6833336,
        "_source" : {
          "review_nr" : 1,
          "desc_es" : "El Fisher Island Club and Hotel está ubicado en una isla privada y ofrece acceso a un campo de golf y a zonas privadas
de playa privada, piscina al aire libre, spa y conexión WiFi gratuita.",
          "hotel_url" : "http://www.booking.com/hotel/us/fisher-island-club.html",
          "nr_rooms" : 15,
          "cc1" : "United States of America",
          "reviews" : [
            {
              "installations" : 6.6,
              "comfort" : 7.5,
              "lang" : "es",
              "location" : 5.7,
              "valueMoney" : 4.7,
              "personal" : 6.9,
              "date" : 1489326152714,
              "comment" : "un poco caro pero cumple con lo pagado. El desayuno buenísimo aunque también algo caro. La única pega es que
está alejado del centro y hay que ir en tranvía siempre a la ciudad. En las inmediaciones del Hotel no hay ningún servicio. Tiene parking
propio. Si tuviera que volver no elegiría este Hotel por la lejanía del centro. Por el resto el Hotel es fantástico aunque demasiado lujoso
para mis intereses",
              "wifi" : 5.0,
              "cleanliness" : 5.7,
              "username" : "admin",
              "score" : 6.0
            }
          ]
        }
      }
    ]
  }
}

```

```

    ],
    "ufi" : 20023181,
    "preferred" : null,
    "address" : "1 Fisher Island Drive",
    "max_rate_EUR" : null,
    "maxrate" : null,
    "class" : 5.0,
    "photo_url" : "http://aff.bstatic.com/images/hotel/max500/381/38155296.jpg",
    "coord" : [
        -80.1449847221375,
        25.760970770878
    ],
    "review_score" : 6.0,
    "continent_id" : "North America",
    "city_preferred" : "Miami",
    "city_hotel" : "Fisher Island (Florida)",
    "zip" : "FL 33109",
    "id" : 275996,
    "min_rate_EUR" : 1147.2508124408814,
    "city_unique" : "miami",
    "currencycode" : "USD",
    "name" : "Fisher Island Club and Hotel",
    "desc_en" : "The Fisher Island Club and Hotel is located on a private island and features access to a golf course and private beach areas. Guests will also enjoy an outdoor pool and a spa is available. Free WiFi is also available.",
    "minrate" : 1225.0,
    "public_ranking" : 1
  }
},
{
  "_index" : "hoteladvisor",
  "_type" : "hotels",
  "_id" : "58c3ba4c1d41c81521d76fbd",
  "_score" : 0.63576835,
  "_source" : {
    "zip" : "900000",
    "preferred" : null,
    "ufi" : -1999811,
    "review_score" : 0.0,
    "class" : 3.0,
    "coord" : [
        3.31110656261444,
        6.47154105397601
    ],
    "public_ranking" : 1,
    "id" : 9980,
    "address" : "Rafiu Babatunde Tinubu Road, Amuwo Odofin Residential Scheme, Festac",
    "continent_id" : "Africa",
    "currencycode" : "USD",
    "city_preferred" : "Amuwo",
    "city_hotel" : "Lagos ",
    "name" : "Sunfit International",
    "min_rate_EUR" : 64.62065800687414,
    "desc_en" : "Offering outdoor pool and indoor pool, Sunfit International is located in Lagos. The property is 20 km from Ikeja Golf Club and 23 km from Lagos Island. Free Wi-Fi access is available in all areas.",
    "maxrate" : null,
    "hotel_url" : "http://www.booking.com/hotel/ng/sunfit-int-l-ltd.html",
    "nr_rooms" : 42,
    "minrate" : 69.0,
    "max_rate_EUR" : null,
    "cc1" : "Nigeria",
    "review_nr" : 0,
    "photo_url" : "http://aff.bstatic.com/images/hotel/max500/261/26107489.jpg",
    "city_unique" : "amuwo",
    "desc_es" : "El Sunfit International, situado en Lagos, ofrece una piscina al aire libre y otra cubierta. Se encuentra a 20 km del club de golf de Ikeja y a 23 km de la isla de Lagos. La conexión WiFi es gratuita."
  }
}

```

```

    }
  },
  {
    "_index": "hoteladvisor",
    "_type": "hotels",
    "_id": "58c3ba4e1d41c81521d78497",
    "_score": 0.6176454,
    "_source": {
      "zip": "100101",
      "preferred": null,
      "id": 15318,
      "ufi": -2011499,
      "review_score": 6.4,
      "class": 0.0,
      "coord": [
        3.35303485393524,
        6.5854725774934
      ],
      "public_ranking": 1,
      "reviews": [
        {
          "cleanliness": 5.4,
          "lang": "es",
          "installations": 5.7,
          "username": "admin",
          "date": 1489314243638,
          "comment": "Hotel normal. Buena cama pero baño pequeño. Recepcionista diligente.",
          "wifi": 8.5,
          "valueMoney": 8.4,
          "personal": 7.5,
          "confort": 3.1,
          "location": 6.4,
          "score": 6.4
        }
      ],
      "address": "31A Joel Ogunnaike Street, Ikeja",
      "continent_id": "Africa",
      "currencycode": "USD",
      "city_preferred": "Ikeja",
      "city_hotel": "Lagos",
      "name": "GrandBee Suites",
      "min_rate_EUR": 88.97047116888467,
      "desc_en": "Offering outdoor pool, GrandBee Suites is located in Lagos. The property is 25 km from Lagos Island and 2 km from Ikeja Golf Club. Free Wi-Fi access is available in all areas. Units at GrandBee Suites will provide guests with modern amenities.",
      "maxrate": null,
      "hotel_url": "http://www.booking.com/hotel/ng/grandbee-suites.html",
      "nr_rooms": 40,
      "minrate": 95.0,
      "max_rate_EUR": null,
      "cc1": "Nigeria",
      "review_nr": 1,
      "photo_url": "http://aff.bstatic.com/images/hotel/max500/657/65710311.jpg",
      "city_unique": "ikeja",
      "desc_es": "El GrandBee Suites alberga piscina exterior y se encuentra en Lagos. El establecimiento se encuentra a 25 km de la isla de Lagos y a 2 km del club de golf Ikeja. Hay conexión WiFi gratuita en todas las instalaciones."
    }
  },
  {
    "_index": "hoteladvisor",
    "_type": "hotels",
    "_id": "58c3bb3c1d41c81521dbbf1f",
    "_score": 0.43203688,
    "_source": {
      "review_nr": 0,
      "hotel_url": "http://www.booking.com/hotel/us/springhill-suites-cincinnati-northeast.html",

```

```

    "nr_rooms" : 102,
    "cc1" : "United States of America",
    "desc_es" : "Este hotel se encuentra a 4,8 km del campo de golf de Western Row y la isla de Paramount's King, y cuenta con
piscina cubierta y servicio de recepción 24 horas. El establecimiento ofrece suites amplias con conexión WiFi gratuita.",
    "ufi" : 20099585,
    "preferred" : null,
    "address" : "9365 Waterstone Boulevard",
    "max_rate_EUR" : null,
    "maxrate" : null,
    "class" : 3.0,
    "photo_url" : "http://aff.bstatic.com/images/hotel/max500/286/28614722.jpg",
    "coord" : [
      -84.3028,
      39.3009
    ],
    "review_score" : 0.0,
    "continent_id" : "North America",
    "city_preferred" : "Mason",
    "city_hotel" : "Cincinnati (Ohio)",
    "zip" : "OH 45249",
    "id" : 292446,
    "min_rate_EUR" : 70.28667222341889,
    "city_unique" : "mason",
    "currencycode" : "USD",
    "name" : "SpringHill Suites by Marriott Cincinnati Northeast/Mason",
    "desc_en" : "Featuring an indoor pool and a 24-hour front desk, this hotel is located 4.8 km from Western Row Golf Course and
Paramount's King Island. The property offers spacious suites with free Wi-Fi access.",
    "minrate" : 75.05,
    "public_ranking" : 1
  }
}
]
}
}

```

El sistema hace uso del driver de Elastic Search para Java, para poder utilizar el API del motor de búsqueda desde la capa de aplicación. De este modo, se permite al usuario realizar búsquedas textuales, con gran velocidad de respuesta, sobre los campos de descripción de los hoteles, así como de los comentarios de las reviews.

2.2.4 CAPA DE EXPLOTACIÓN (APPLICATION LAYER)

Esta capa se encarga de la explotación de los datos almacenados en el sistema y su visualización por parte del usuario final. El lenguaje de programación utilizado para la implementación de esta capa es Java.

Está compuesta por tres módulos:

2.2.4.1 PRESENTATION TIER

Módulo encargado de realizar la presentación de los datos al usuario final.

Está conformada por una clase:

- * **UserInterface.java:** Encargada de realizar las siguientes acciones:
 - * Mostrar los hoteles recomendados en formato tabular, mediante Java Swing
 - * Mostrar los POIs en formato tabular, mediante Java Swing
 - * Mostrar los comentarios(reviews) en formato tabular, mediante Java Swing

2.2.4.2 BUSINESS LOGIC TIER

Módulo encargado de gestionar la lógica de negocio de la aplicación.

Está conformada por varias clases:

- * **HotelAdvisor.java:** Clase principal de la aplicación, encargada de mostrar los menús y gestionar toda la lógica de la aplicación apoyándose en el resto de clases. Hace uso del Framework de Agregación de MongoDB así como de su API para realizar consultas geoposicionadas. Asimismo, hace uso del API de Elasticsearch para realizar búsquedas textuales sobre su motor de búsqueda.
- * **IO.java:** Encargada de gestionar las operaciones de entrada/salida sobre la consola.
- * **Maths.java:** Encargada de ofrecer algunas funciones matemáticas utilizadas por el resto de módulos.
- * **PoisTypes.java:** Encargada de gestionar los tipos de POIs.
- * **Util.java:** Encargada de ofrecer diversos métodos de utilidad para el resto de módulos.

2.2.4.3 DATA TIER

Módulo encargado de gestionar la persistencia de la aplicación comunicándose con la capa de Data Storage.

Está conformada por varias clases:

- * **DbUtils.java:** Clase para la conexión y creación de objetos en la base de datos MongoDB.
- * **QueryHelper.java:** Clase que proporciona métodos para realizar consultas de usuarios, hoteles y POIs sobre la base de datos MongoDB.
- * **Hotel.java:** POJO que modela la colección de hoteles de la base de datos MongoDB.
- * **HotelQuery.java:** POJO que modela una consulta de hoteles sobre la base de datos MongoDB.
- * **Poi.java:** POJO que modela la colección de pois de la base de datos MongoDB.
- * **ElasticSearch.java:** Clase para la conexión al motor de búsqueda de Elasticsearch.
- * **ElasticSearchQueryHelper.java:** Encapsula métodos de utilidad para realizar consultas sobre el motor de Elasticsearch.

ANEXO: LISTADO DE CATEGORÍAS DE TIPOS DE POI

ACCOMMO_ALPINEHUT,1
ACCOMMO_CAMPING,2
ACCOMMO_CARAVAN,3
ACCOMMO_CHALET,4
ACCOMMO_HOSTEL,5
ACCOMMO_HOTEL,6
ACCOMMO_MOTEL,7
,
AMENITY_COURT,8
AMENITY_FIRESTATION,9
AMENITY_LIBRARY,11
AMENITY_PLAYGROUND,12
AMENITY_POLICE,13
AMENITY_POSTOFFICE,14
AMENITY_PRISON,15
AMENITY_PUBLICBUILDING,16
AMENITY_TOWNHALL,17
,
BARRIER_BLOCKS,18
,
EDUCATION_COLLEGE,19
EDUCATION_NURSERY,20
EDUCATION_SCHOOL,21
EDUCATION_UNIVERSITY,22
,
FOOD_BAR,23
FOOD_BIERGARTEN,24
FOOD_CAFE,25
FOOD_FASTFOOD,26
FOOD_ICECREAM,27
FOOD_PUB,28
FOOD_RESTAURANT,29
,
HEALTH_DENTIST,30
HEALTH_DOCTORS,31
HEALTH_HOSPITALEMERGENCY,32
HEALTH_HOSPITAL,33
HEALTH_PHARMACY,34
HEALTH_VETERINARY,35
,
LANDUSE_ALLOTMENTS,36
LANDUSE_CONIFEROUSDECIDUOUS,37
LANDUSE_CONIFEROUS,38
LANDUSE_DECIDUOUS,39
LANDUSE_GRASS,40
LANDUSE_HILLS,41
LANDUSE_MILITARY,42
LANDUSE_QUARY,43
LANDUSE_SCRUB,44
LANDUSE_SWAMP,45
,
MONEY_BANK,46
MONEY_EXCHANGE,47
,
POW_BAHAI,48
POW_BUDDHIST,49
POW_CHRISTIAN,50
POW_HINDU,51
POW_ISLAMIC,52
POW_JAIN,53
POW_JEWISH,54

POW_SHINTO,55
POW_SIKH,56
POW_UNKOWN,58
,
POI_CAVE,59
POI_CRANE,60
POI_EMBASSY,61
POI_BUNKER,62
POI_MINE,63
POI_PEAK1,64
POI_PEAK,65
POI_CITY,66
POI_HAMLET,67
POI_SUBURB,68
POI_TOWN,69
POI_VILLAGE,70
POI_TOWERCOMMUNICATION,71
POI_TOWERLOOKOUT,72
,
SHOP_ALCOHOL,73
SHOP_BAKERY,74
SHOP_BICYCLE,75
SHOP_BOOK,76
SHOP_BUTCHER,77
SHOP_CARREPAIR,78
SHOP_CAR,79
SHOP_CLOTHES,80
SHOP_COMPUTER,81
SHOP_CONFECTIONERY,82
SHOP_CONVENIENCE,83
SHOP_COPYSHOP,84
SHOP_DEPARTMENTSTORE,85
SHOP_DIY,86
SHOP_FISH,87
SHOP_FLORIST,88
SHOP_GARDENCENTRE,89
SHOP_GIFT,90
SHOP_GREENGROCER,91
SHOP_HAIRDRESSER,92
SHOP_HEARINGAIDS,93
SHOP_HIFI,94
SHOP_JEWELRY,95
SHOP_KIOSK,96
SHOP_LAUNDRETTE,97
SHOP_MARKETPLACE,98
SHOP_PHONE,99
SHOP_MOTORCYCLE,100
SHOP_MUSIC,101
SHOP_NEWSPAPER,102
SHOP_PET,103
SHOP_SHOES,104
SHOP_SUPERMARKET,105
SHOP_TOBACCO,106
SHOP_TOYS,107
SHOP_VENDINGMASCHINE,108
SHOP_VIDEORENTAL,109
,
SPORT_ARCHERY,110
SPORT_BASEBALL,111
SPORT_BASKETBALL,112
SPORT_BOWLING,113
SPORT_CANOE,114
SPORT_CRICKET,115
SPORT_DIVING,116
SPORT_FOOTBALL,117

SPORT_GOLF,118
SPORT_GYM,119
SPORT_GYMNASIUM,120
SPORT_CLIMBING,121
SPORT_HORSE,122
SPORT_ICESKATING,123
SPORT_LEISURECENTER,124
SPORT_MINIATURGOLF,125
SPORT_MOTORRACING,126
SPORT_SHOOTING,127
SPORT_SKATING,128
SPORT_SKIINGDOWNHILL,129
SPORT_SNOOKER,130
SPORT_SOCCER,131
SPORT_STADIUM,132
SPORT_SWIMMING,133
SPORT_TENNIS,134
SPORT_WATERSKI,135
SPORT_SURFING,136
,
TOURIST_ARCHAEOLOGICAL,137
TOURIST_ART,138
TOURIST_ATTRACTION,139
TOURIST_BATTLEFIELD,140
TOURIST_BEACH,141
TOURIST_CASTLE,142
TOURIST_CASTLE2,143
TOURIST_CINEMA,144
TOURIST_FOUNTAIN,145
TOURIST_INFORMATION,146
TOURIST_MEMORIAL,147
TOURIST_MONUMENT,148
TOURIST_MUSEUM,149
TOURIST_NIGHTCLUB,150
TOURIST_RUINS,151
TOURIST_THEATRE,152
TOURIST_THEMEPARK,153
TOURIST_WINDMILL,156
TOURIST_WRECK,157
TOURIST_ZOO,158
,
TRANSPORT_TERMINAL,159
TRANSPORT_AIRPORT,160
TRANSPORT_BUSSTOP,161
TRANSPORT_FUEL,162
TRANSPORT_LIGHTHOUSE,163
TRANSPORT_MARINA,164
TRANSPORT_RENTALCAR,165
TRANSPORT_SUBWAY,166
TRANSPORT_STATION,167
TRANSPORT_TRAMSTOP,168
,
WATER_DAM,169
WATER_TOWER,170
WATER_WEIR,171