Large-Scale and Multi-Structured Databases
Project Design
Itinera

Nicolò Bacherotti

Lorenzo Menchini

Rossana Antonella Sacco

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

UNIVERSITÀ DI PISA

CROSSLAB
Innovation for industry 4.0

Application Highlights

- **Visit Places:**
  - Showing most interesting Hotels, Restaurants and Monuments near a selected city
  - Allowing Admins to add new cities to discover
  - Discover statistics about the city, restaurants and hotels
  - Diving deeper into a place of their interest

- **Review trip:**
  - Joining communities for learning things about the city
  - Review hotels, restaurants and monuments by evaluation
  - Seeing the best places in every city to stay for your trip

- **Moderate communities:**
  - Allowing users to report other users
  - Allow Admin Handling inappropriate behaviors from users
  - Reporting single reviews
  - Report single posts
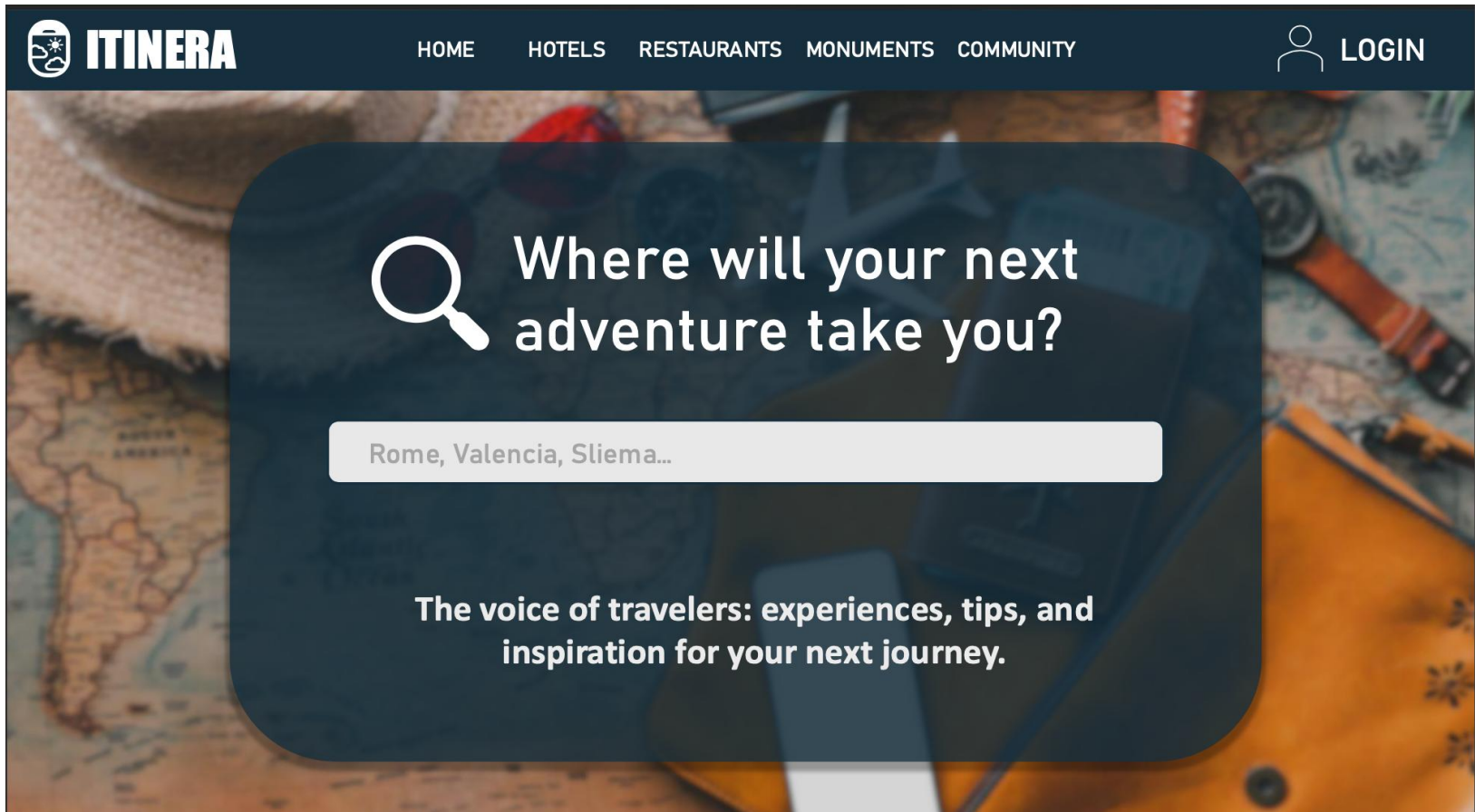 - Allow reply to posts by users inside the communities
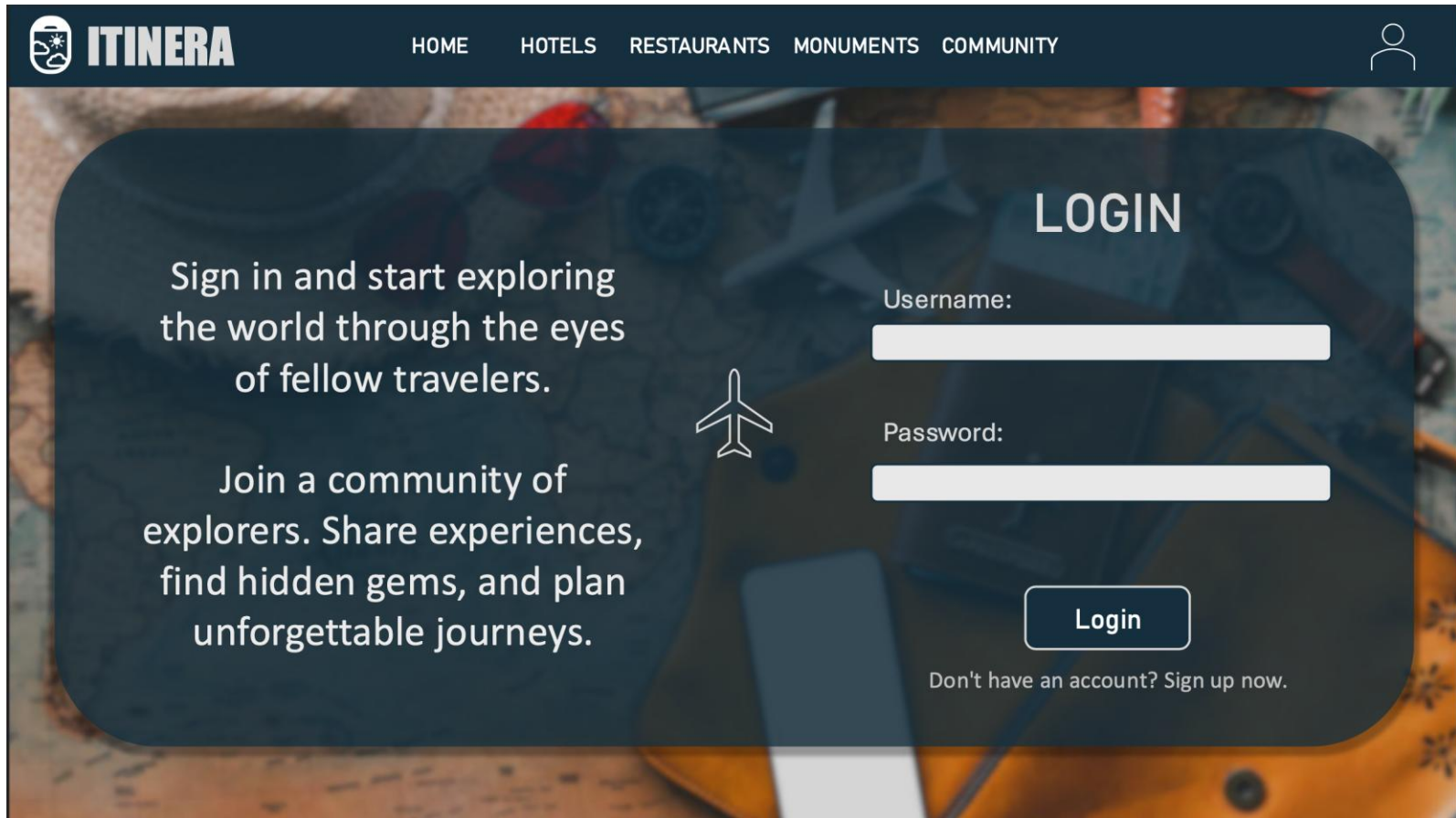
Actors and main mock-ups 1

Main Actors:

Unregistered User

Registered Users

Admins

Actors and main mock-ups 2

Actors and main mock-ups 3

Actors and main mock-ups 4
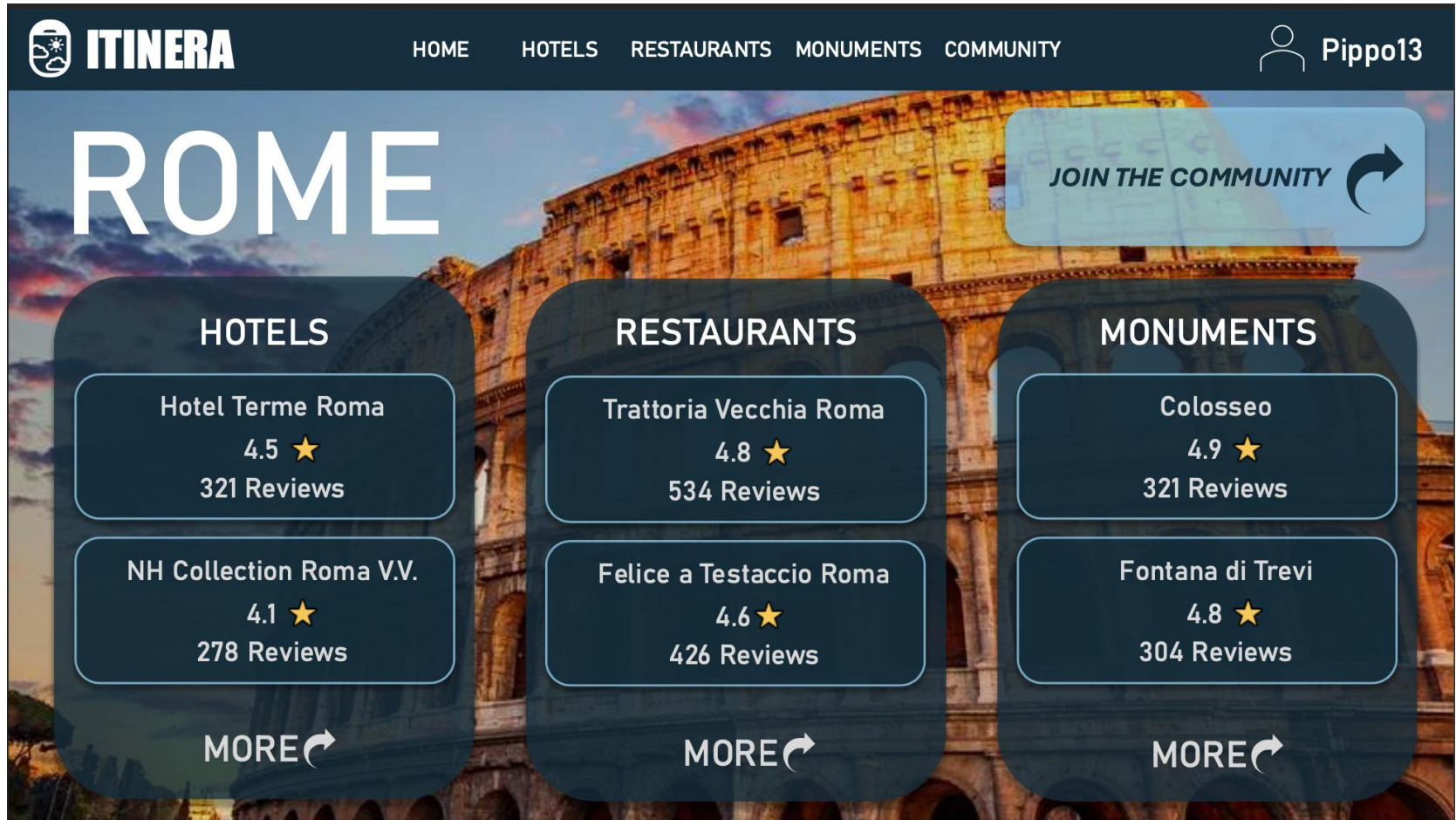
Actors and main mock-ups 5

Actors and main mock-ups 6
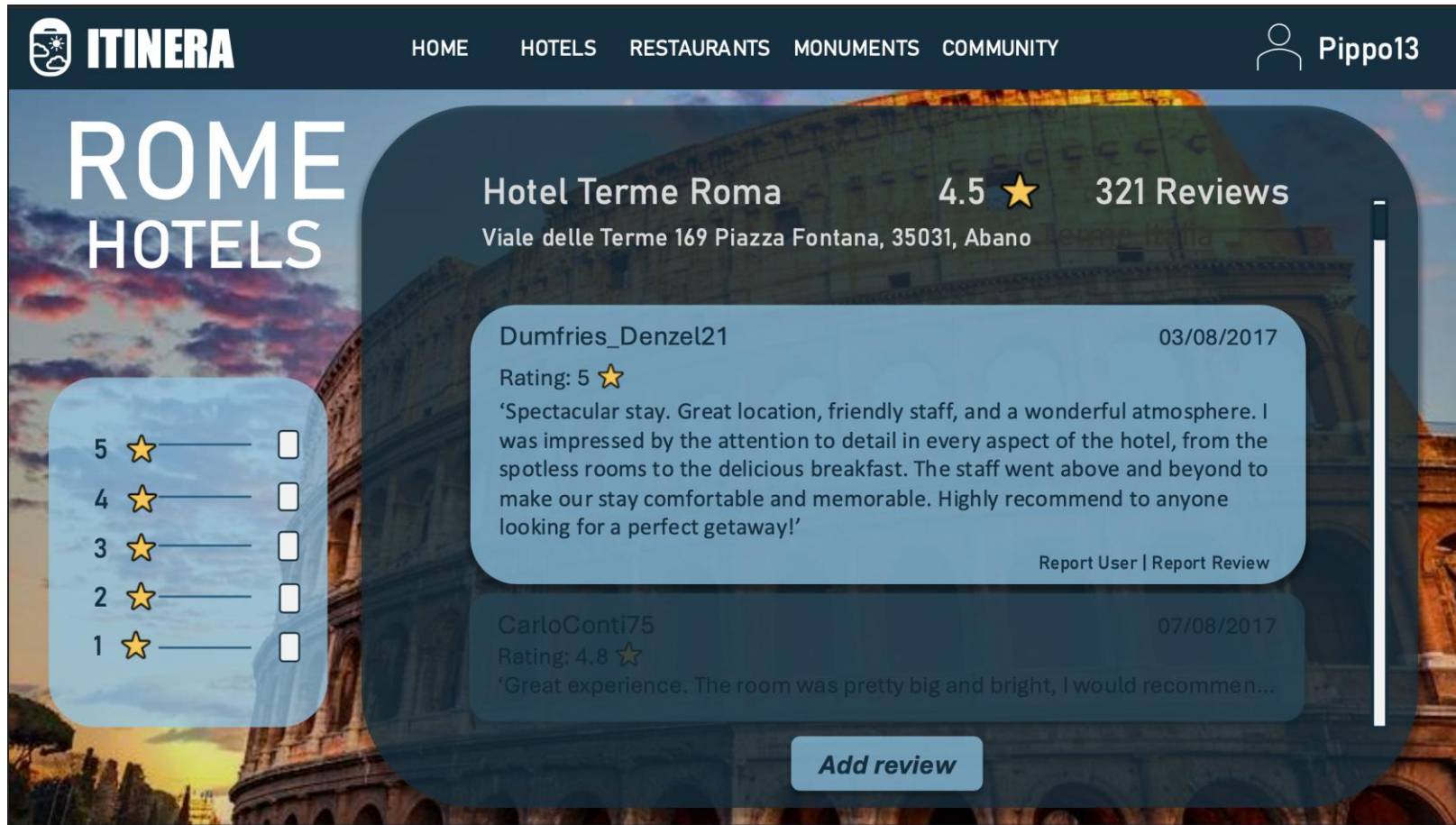
Actors and main mock-ups 7
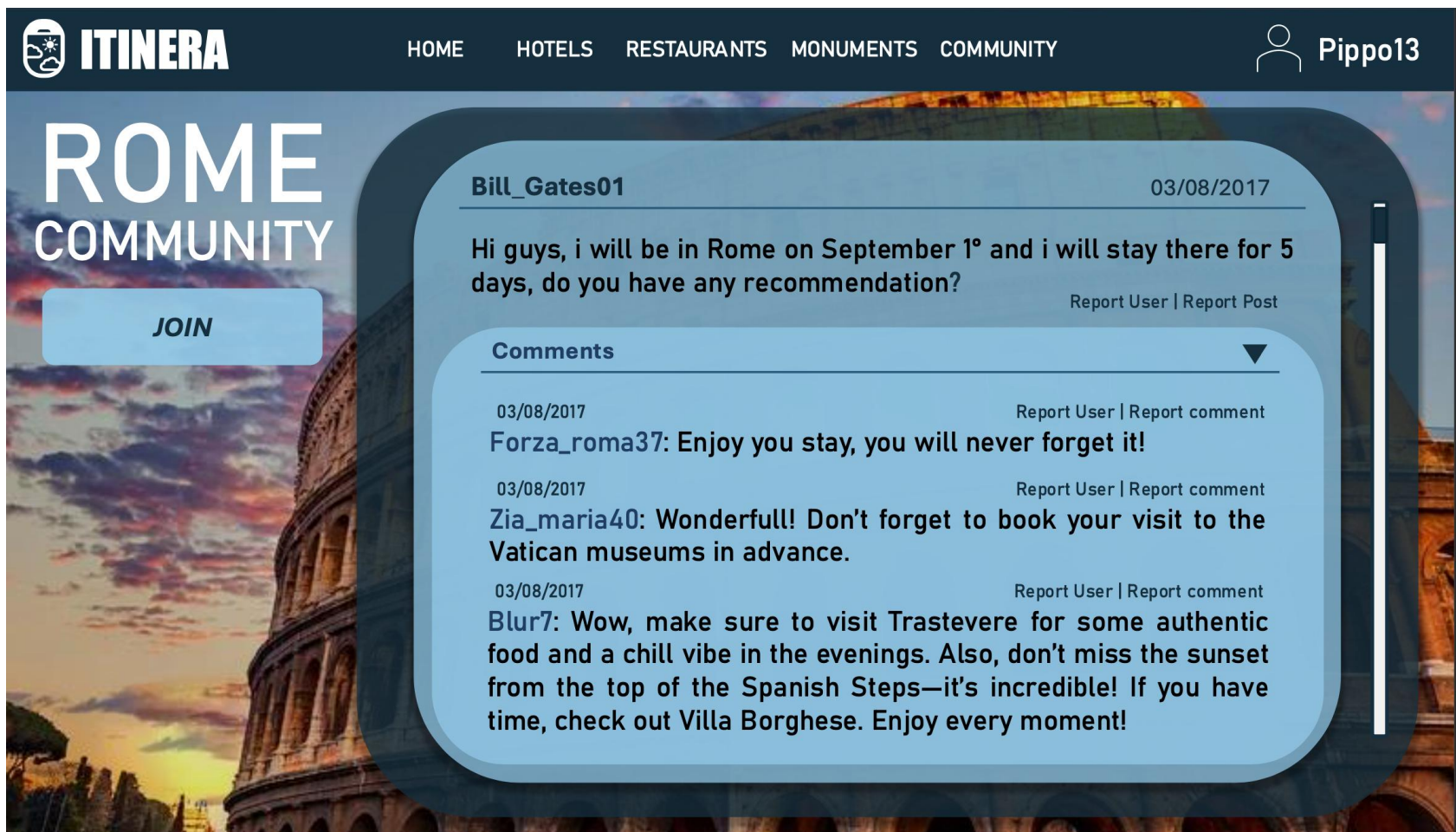
Dataset Description

*Source: reddit API, google places API,
https://www.kaggle.com/datasets/inigolopezrioboo/a-tripadvisor-dataset-for-nlp-tasks?select=London_reviews.csv ,
https://www.kaggle.com/datasets/jiashenliu/515k-hotel-reviews-data-in-europe*

Description: all dataset with google places API collects reviews of restaurants and hotels, with reddit also we obtain comments for communities of cities
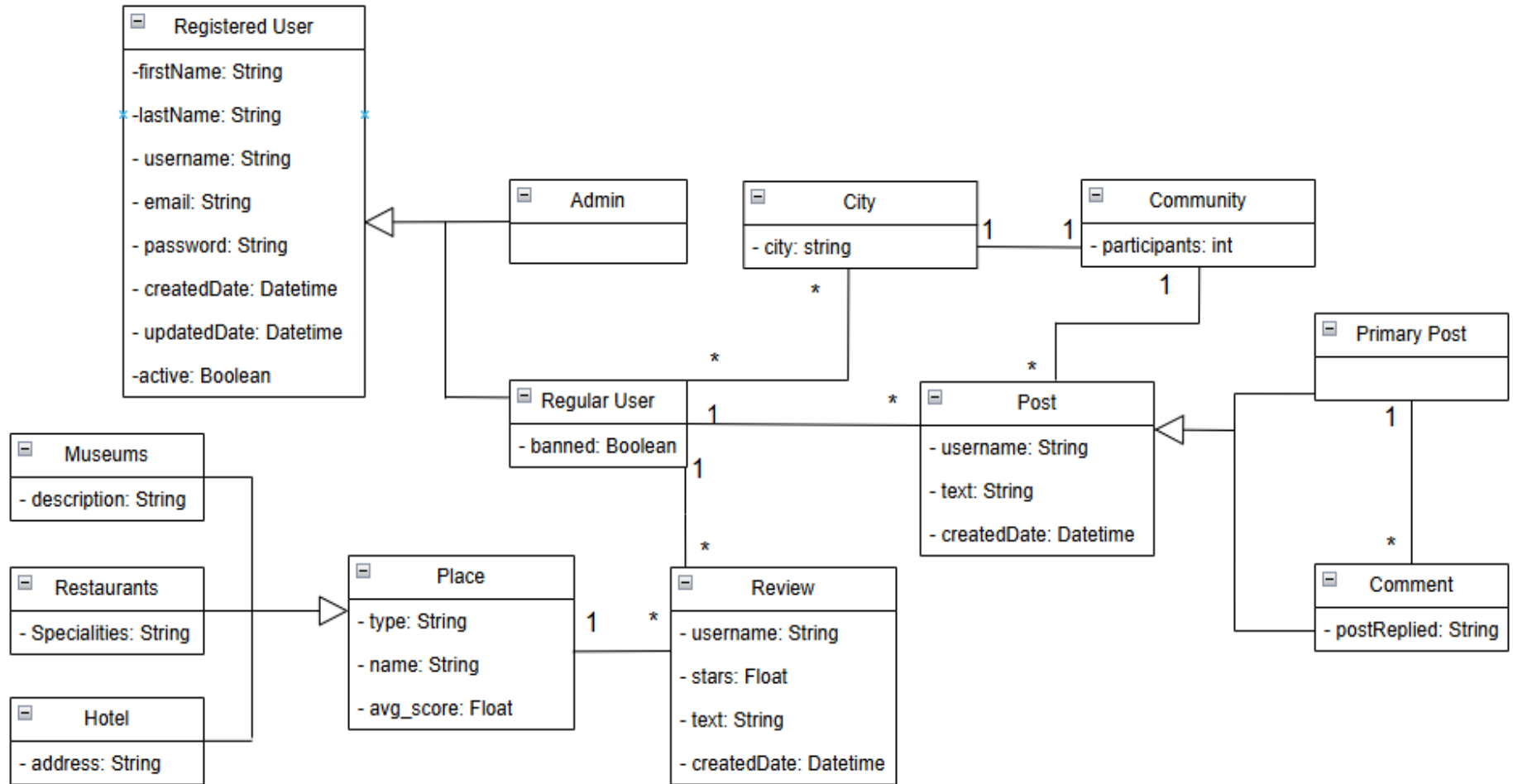
Volume: 238 MB hotel reviews + reddit data for communities + google places reviews of every restaurant, hotel and monument inside 5km radius of the city centre

Variety: variety is achieved by several datasets and APIs with different types of data schema

Preliminary UML Class Diagram

Relevant Queries and Entities handled by Document DB

Entities:
- Users
- Admins
- Posts
- Reviews
- Communities
- Comment
- Places
- City

Queries:
- Shows best reviews for interesting places based on a city
- Shows the most reviewed city by the user
- Shows the number of posts written for each community
- Retrieving user activities and provide them in their page
- Obtain average review value for the selected place
- The places with the highest number of 5 star reviews

Requirements and Entities handled by Graph DB

- Suggests cities to visit.
- Suggests people to follow.
- Suggests posts you might be interested in
- What is the most popular community?
- Most active user in a community.