Large-Scale and Multi-Structured Databases

Project Design

JustRecipe

Francesco Campilongo

Daniele Cioffo

Francesco Iemma

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

UNIVERSITÀ DI PISA

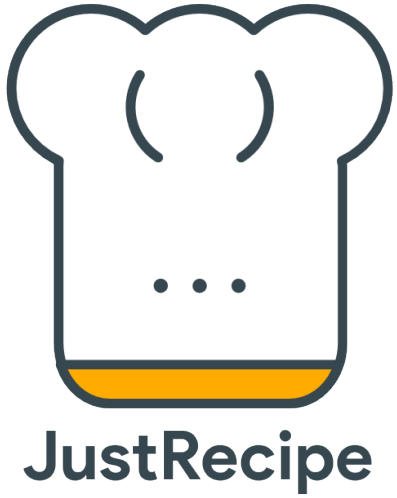
CROSSLAB Innovation for industry 4.0

Application Highlights

The **JustRecipe** application is designed to create a community of cuisine lovers.

The **main features** of the application are:

- Users can browse and add recipes
- User can modify his own recipes
- Social Network features
  - Follow a user
  - Like a recipe
  - Comment a recipe
  - Obtain suggestions



Actors and main supported functionalities

- Unregistered User
  - ❑ Sign-up
- User
  - ❑ Login/Logout
  - ❑ Search a recipe
  - ❑ Browse suggested recipes
  - ❑ Browse recipes of following users
  - ❑ Add a recipe
  - ❑ Edit own recipes
  - ❑ Comment recipes
  - ❑ Follow another user
  - ❑ Like a recipe
- Moderator
  - ❑ Delete comments
- Administrator
  - ❑ Delete users
  - ❑ Delete recipes
  - ❑ Elect moderators

Dataset Description

Source:
https://eightportions.com/datasets/Recipes/
https://www.kaggle.com/hugodarwood/epirecipes?select=full_format_recipes.json

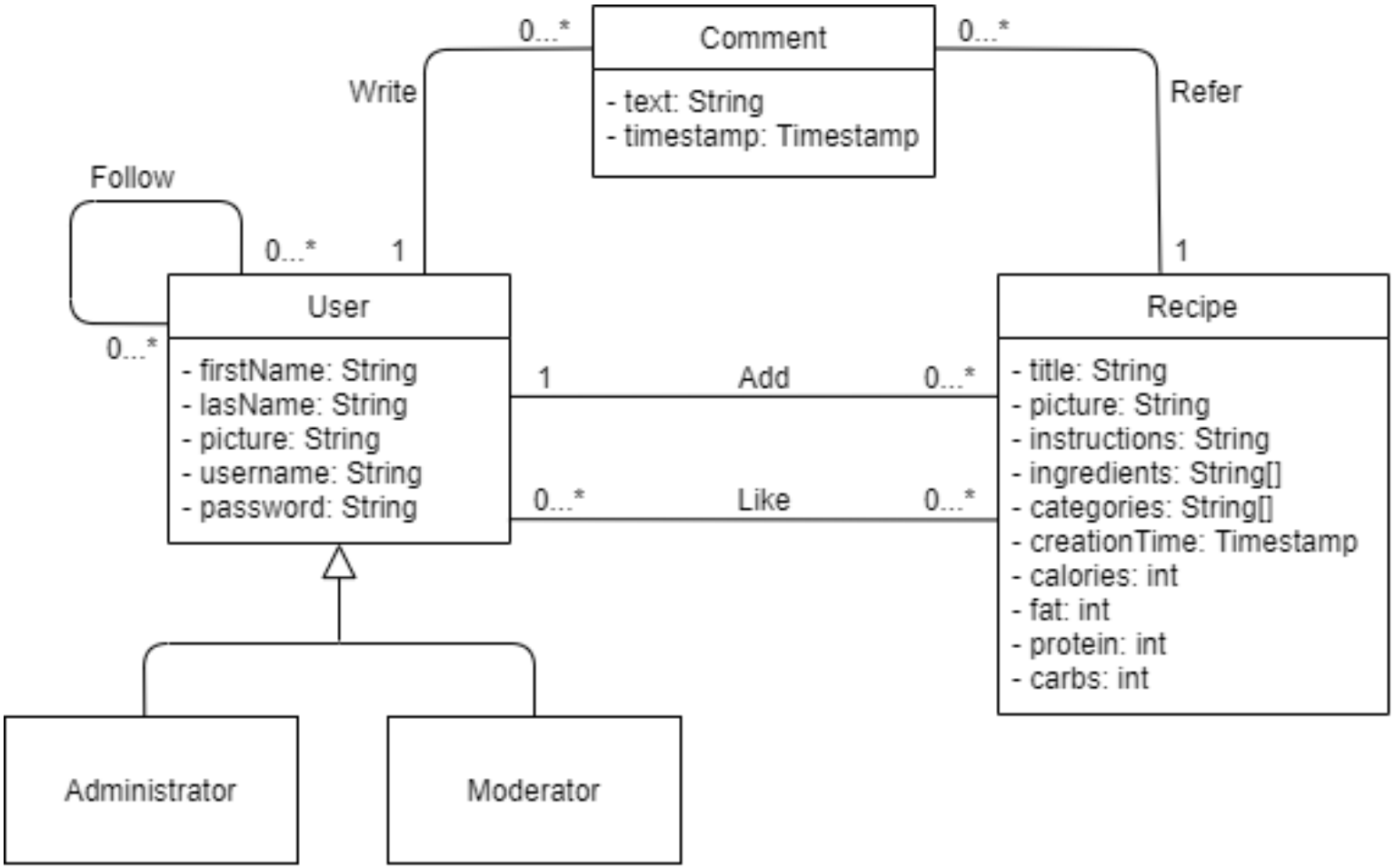
Description:
Datasets containing recipes. Each one of them is in JSON format.

Volume:
45349 recipes (67.8 MB)

Variety:
We use two datasets of recipes, obtained by scraping Foodnetwork.com and Epicurios.com. Both datasets have been cleaned up by removing duplicates.

Velocity/Variability:
Variability and Velocity are ensured by deleting and/or adding recipes and comments

Preliminary UML Class Diagram



Requirements and Entities handled by Document DB

Collections:

- Recipe

High Level Query:

- Search a recipe given certain parameters (recipe's name, recipe's ingredients)
- Low caloric recipes, low fat recipes, low carbs recipes, high protein recipes
- *Search the most common recipe categories*
- *Last written comments (for moderator)*
- *Categories under/over a caloric/carbs/fat/protein average*
- *Show the daily sum of recipe calories added by a user in the last week. A plot will show the weekly trend*
- Last written recipes of the following users
- Fastest recipes

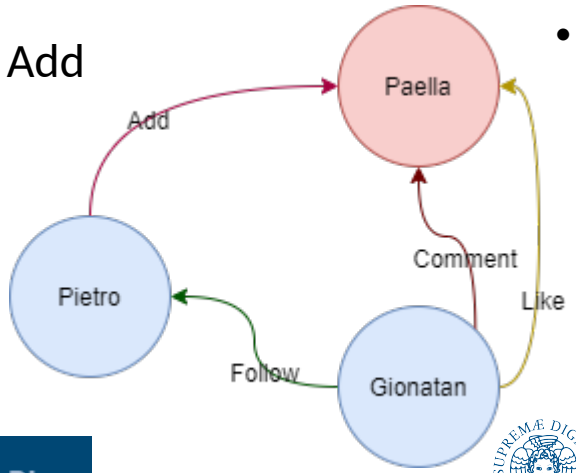
Requirements and Entities handled by Graph DB

Entities:

- Recipe
- User

Relations:

- User – Recipe: Like
- User – User: Follow
- User – Recipe: Comment
- User – Recipe: Add



High Level Query:

- Suggested recipes (considering hops number)
- Best recipes (considering likes)
- Most followed and active user (centrality measure)
- Top Commentators
- Most liked user (considering his/her recipes)

Software Architecture Preliminary Idea

Programming Language:

- JavaFX

DBMSs:

- MongoDB
- Neo4J

Frameworks:

- JUnit
- Log4J
- Maven