Feyzan Colak, Flavio Messina, Noemi Cherchi
Large-scale and multi-structured databases project
2023-2024

Contents

Introduction

MangaVerse is a web application project developed for the Large-scale and multi-structured databases course of the University of Pisa. This web application aims to provide users with a comprehensive platform to explore, search, interact with a vast collection of manga and anime and interact with the other users.

The website is accessible without login providing a limited number of functionalities. Once a user logs in, the platform offers a wide range of features to personalize the user experience, in particular the social features. The application manages user and media content suggestions based on interactions, preferences and user information. Beside the user roles, the web application also has managerial roles. MangaVerse provides an analytics dashboards to track media contents and user activities also managing media contents and user accounts. These features allow manager to add, update, or remove manga and anime entries and monitor trends and rating.

Through its comprehensive set of features, MangaVerse aims to provide a community of manga and anime enthusiasts with a platform to explore, share, and engage with their favorite content. This platform enhances the user experience and facilitates deep engagement with both the content and the community.

Analysis

Actors

- **Unregistered User**: A visitor who has not logged in on the platform.
- **Registered User**: A user who has created an account on the platform.
- **Manager**: A registered user with administrative privileges.

Requirements

Unregistered User:

- Register/Login:
  - Create a new account to access additional features.
  - Use valid credentials (email and password) to log into the account.
- Browse Media Contents.
- Search and Filter Media Contents:
  - Find specific manga or anime by title.
  - Utilize basic filtering options to refine the media content list.
- View Media Content Trends.
- View Media Content:
  - View limited information about each media content.
- View Media Content Details:
  - View detailed information about each media content.
  - View reviews and ratings for each media content.
  - View number of likes for each media content.
- Browse Users.
- Search Users by Username.
- View User:
  - View limited information about each user.
- View User Details:
  - View detailed information about each user.
  - View anime and manga liked by the user.
  - View followers and following of the user.

Registered User:

- Logout.
- Browse Media Contents.
- Search and Filter Media Contents:

  - Find specific manga or anime by title.
  - Utilize basic filtering options to refine the media content list.
- View Media Content Trends.
- View Media Content:
  - View limited information about each media content.
- View Media Content Details:
  - View detailed information about each media content.
  - View reviews and ratings for each media content.
  - View number of likes for each media content.
- Browse Users.
- Search Users by Username.
- View User:
  - View limited information about each user.
- View User Details:
  - View detailed information about each user.
  - View anime and manga liked by the user.
  - View followers and following of the user.
- Profile Management:
  - Edit and update personal information (e.g., profile picture, bio).
  - Delete own profile.
- Like/Unlike Media Contents.
- Follow/Unfollow Users.
- Review Media Contents:
  - Add comment and rating to manga and anime.
  - Edit/Delete own reviews.
- Advanced Recommendations:
  - Receive media content suggestions based on user interactions and personal information.
  - Receive users suggestions based on user interactions.

Manager(Registered User with Administrative Features):

- Logout.
- Browse Media Contents.
- Search and Filter Media Contents:
  - Find specific manga or anime by title.
  - Utilize basic filtering options to refine the media content list.
- View Media Content Trends.
- View Media Content:
  - View limited information about each media content.

- View Media Content Details:
  - View detailed information about each media content.
  - View reviews and ratings for each media content.
  - View number of likes for each media content.
- Browse Users.
- Search Users by Username.
- View User:
  - View limited information about each user.
- View User Details:
  - View detailed information about each user.
  - View anime and manga liked by the user.
  - View followers and following of the user.
- Analytics Dashboard:
  - View user analytics (distribution and app rating).
  - View manga analytics (trends and average rating).
  - View anime analytics (trends and average rating).
- Content Management:
  - Add new media content (manga and anime).
  - Update/Remove existing media content.

Non Functional Requirements

Performance

- Response Time: The system should have low latency, with pages loading within an acceptable timeframe.
- Scalability: The system should be able to handle an increasing number of users and data without significant degradation in performance.
- Concurrency: The application should support multiple users simultaneously without performance bottlenecks. For very high traffic scenarios, acceptable delays may be introduced.
- Availability: The system should be available 24/7, with minimal downtime for maintenance.
- Replication: The system should have data replication to ensure data availability and fault tolerance.

Security

- Controlled User Operations: Users should only be able to perform operations that they are authorized to do.

Data Integrity

- Data Consistency: The system should maintain data consistency across all components and databases.

User Interface

- Responsiveness: The user interface should be responsive, providing a consistent and seamless experience across various devices and screen sizes.
- Intuitiveness: The interface should be user-friendly, with clear navigation and easily understandable features.

UML use case diagram

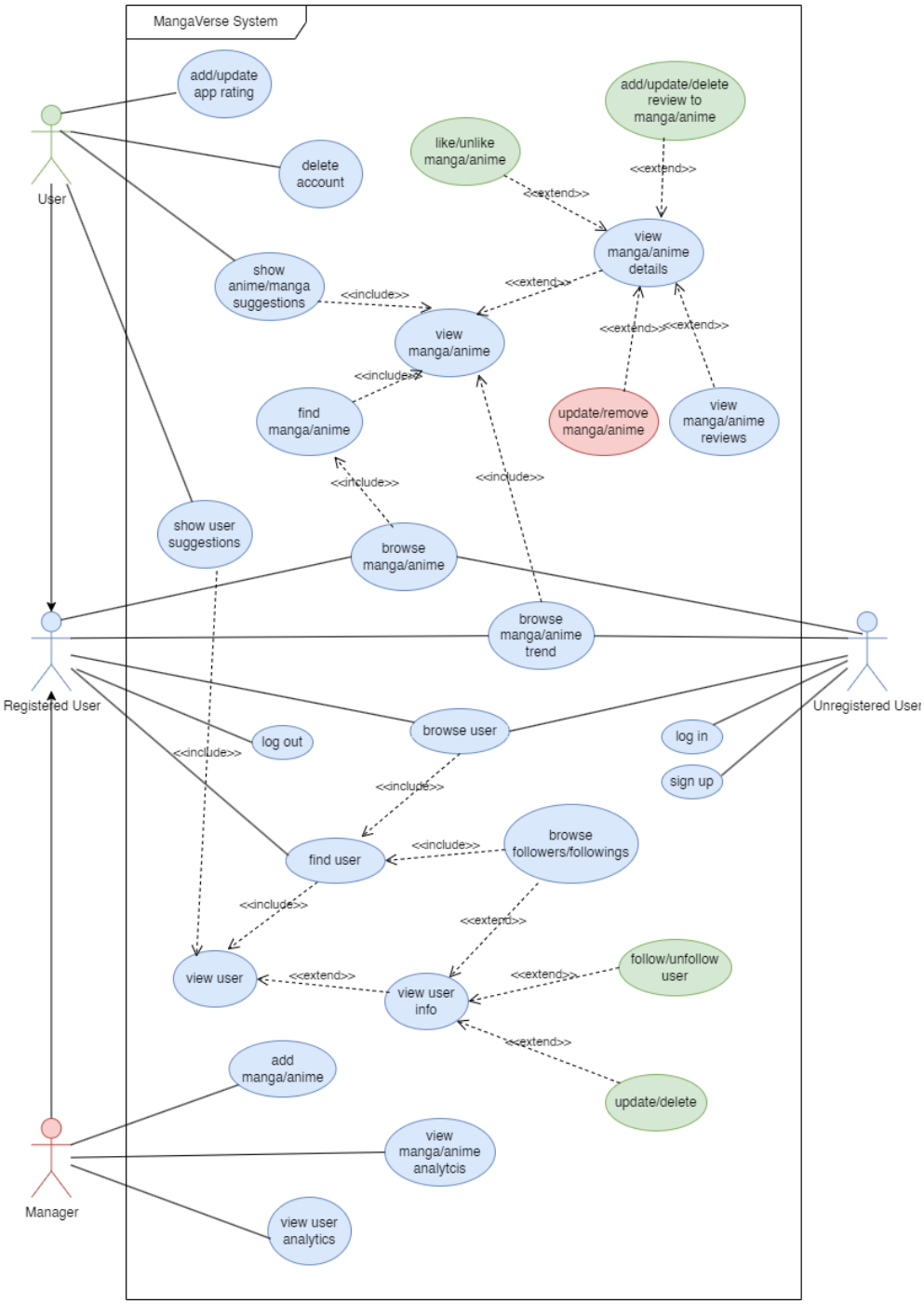


Figure 2.1: UML Use Case Diagram

UML class diagram

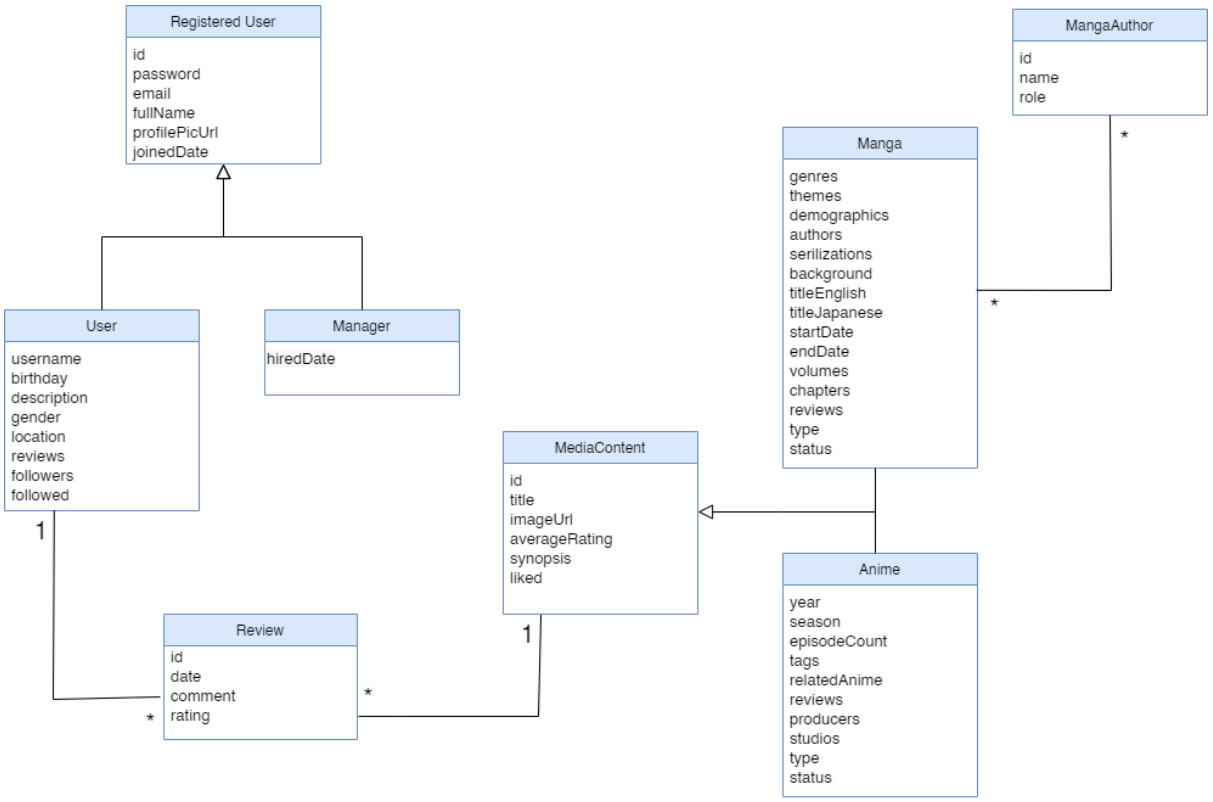


Figure 2.2: UML Class Diagram

Data Modeling

Data Collection

Sources: https://www.kaggle.com/datasets/dbdmobile/myanimelist-dataset?select=users-score-2023.csv, MyAnimeList.net, anilist.com, kitsu.io, livechart.me, anime-planet.com, nofity.moe, anisearch.com, anidb.net

Description: Manga, users and scores datasets were collected from MyAnimeList.net site using the official API and another unofficial API (Jikan). The anime datasets were collected from all the sources.

Variety: The datasets contain a variety of data types, including text, numbers, and dates. Anime are collected from 8 different sources. All the information is collected in 4 different csv files.

Volume: The datasets contain a large volume of data, with thousands of entries for anime, manga, users, and scores. The total size of the datasets is around 3 GB.

Data Cleaning and Preprocessing

Python scripts were used to clean and preprocess the data. The following steps were performed: reviews were created by merging the users and scores datasets, and creating comments about the media contents; the anime dataset was created by putting togethere the diffrerent sources; the manga dataset was created from MyAnimeList.net; the users dataset was cleaned and missing information, like email and password, was added.

Design

The web application needs to handle a big amount of data, so we decided to use a combination of different databases to store and manage the data. We will use a document database to store users, media contents and reviews data, and a graph database to store relationships between users and media content. This will allow us to efficiently store and retrieve data, as well as handle complex relationships between data.

Document Database

For the document database, we will use MongoDB. MongoDB is a NoSQL database that stores data in flexible, JSON-like documents. It is a popular choice for applications that require flexibility and scalability. These documents are flexible, meaning they can have different fields and structures. This makes MongoDB a good choice for applications that require flexibility in their data model. MongoDB is also a scalable database, meaning it can handle large amounts of data and traffic. It is designed to scale out, meaning you can add more servers to handle more traffic.

Collections The database will have the following collections:

- Anime: This collection will store information about anime, such as titles, tags, and synopsis.
- Manga: This collection will store information about manga, such as titles, genres, and authors.
- Reviews: This collection will store user ratings and comments for media content.
- Users: This collection will store user data, such as usernames, passwords, email addresses, gender and location.

MongoDB document example

Anime:


```
{
  "_id": "65789bb52f5d29465d0abcfb",
  "title": "0",
  "type": "SPECIAL",
  "episodes": 1,
  "status": "FINISHED",
  "picture": "https://cdn.myanimelist.net/images/anime/12/81160.jpg",
  "tags": [
    "drama",
    "female protagonist",
    "indefinite",
    "music",
    "present"
  ],
  "producers": "Sony Music Entertainment",
  "studios": "Minakata Laboratory",
  "synopsis": "This music video tells how a shy girl with a secret love and curiosity...",
  "latest_reviews": [
    {
      "id": "657b301306c134f18884924c",
      "date": "2023-10-03T22:00:00.000+00:00",
      "rating": 4,
      "user": {
        "id": "6577877ce68376234760745c",
        "username": "Tolstij_Trofim",
        "picture": "https://thypix.com/wp-content/uploads/2021/10/manga-profile-picture
    -10..."
      }
    },
  ],
  "anime_season": {
    "season": "FALL",
    "year": 2013
  },
  "average_rating": 6.7,
  "avg_rating_last_update": true,
  "likes": 4
}
```

Manga:


```
{
  "_id": "657ac61bb34f5514b91ea223",
  "title": "Berserk",
  "type": "MANGA",
  "status": "ONGOING",
  "genres": [
    "Action",
    "Adventure",
    "Award Winning",
    "Drama",
    "Fantasy",
    "Horror",
    "Supernatural"
  ],
  "themes": [
    "Gore",
    "Military",
    "Mythology",
    "Psychological"
  ],
  "demographics": [
    "SEINEN"
  ],
  "authors": [
    {
      "id": 1868,
      "role": "Story & Art",
      "name": "Kentarou Miura"
    },
    {
      "serializations": "Young Animal"
    }
  ],
  "synopsis": "Guts, a former mercenary now known as the \"Black Swordsman,\" is out fo...
    ",
  "title_english": "Berserk",
  "start_date": "1989-08-25T00:00:00.000+00:00",
  "picture": "https://cdn.myanimelist.net/images/manga/1/157897l.jpg",
  "average_rating": 3.33,
  "latest_reviews": [
    {
      "user": {
        "id": "6577877be683762347605ce7",
        "username": "calamity_razes",
        "picture": "https://imgbox.com/7MaTkBQR"
      },
      "date": "2012-12-15T00:00:00.000+00:00",
      "comment": "An insult to the art of manga; avoid at all costs.",
      "id": "657b302206c134f18886f5ef"
    },
  ],
  "anime_season": {
    "season": "FALL",
    "year": 2013
  },
  "average_rating": 6.7,
  "avg_rating_last_update": true,
  "likes": 4
}

```

Reviews:


```
{
  "_id": "657b300806c134f18882f2f1",
  "user": {
    "id": "6577877be68376234760596d",
    "username": "Dragon_Empress",
    "picture": "images/account-icon.png",
    "location": "Columbus, Georgia",
    "birthday": "1987-07-29T00:00:00.000+00:00",
    "rating": 7
  },
  "anime": {
    "id": "65789bbc2f5d29465d0b18b7",
    "title": "Slayers Revolution",
    "date": "2023-07-23T06:27:54.000+00:00",
    "comment": "Above-average quality in animation and soundtrack."
  }
}

```


Users:


```
{
  "_id": "6577877be683762347605859",
  "email": "xdavis@example.com",
  "password": "290cb38a679d5eb68d11b9ea1e21f48234eba6de19f95612dbcb70ce0c7e4e78",
  "description": "Liberating the mind from stress with the power of anime zen.",
  "picture": "https://thypix.com/wp-content/uploads/2021/10/manga-profile-picture-44",
  "username": "Xinil",
  "gender": "Male",
  "birthday": "1985-03-04T00:00:00.000+00:00",
  "location": "Libya",
  "joined_on": "2014-05-29T00:00:00.000+00:00",
  "app_rating": 5,
  "followed": 40,
  "followers": 29
}

```


The field `"app_rating"` is used to know the general satisfaction of the user with the application.

CRUD operations

- Create: This operation will allow users to create new documents in the database. For example, users can create new reviews for anime and manga.
- Read: This operation will allow users to read documents from the database. For example, users can read information about anime and manga and about other users.
- Update: This operation will allow users to update documents in the database. For example, users can update their reviews for anime and manga, they can also update their own profile, the manager can update media contents.
- Delete: This operation will allow users to delete documents from the database. For example, users can delete their reviews for anime and manga, the manager can delete media contents.

Graph Database

For the graph database, we will use Neo4j. Neo4j is a graph database that stores data in nodes and relationships. It is a popular choice for applications that require complex relationships between data. Neo4j is a graph database, which means it stores data in nodes and relationships. Nodes represent entities, such as users or products, and relationships represent connections between nodes. This makes Neo4j a good choice for applications that require complex relationships between data. Neo4j is also a scalable database, meaning it can handle large amounts of data and traffic. It is designed to scale out, meaning you can add more servers to handle more traffic. This makes Neo4j a good choice for applications that need to scale quickly.

Nodes

The database will have the following nodes:

- User: This node will store information about users, such as id, usernames, and picture.
- Anime: This node will store information about anime, such as id, titles and picture.
- Manga: This node will store information about manga, such as id, titles and picture.

Relationships

The database will have the following relationships:

- LIKE: This relationship will connect users to anime and manga nodes. It will store the date when the user liked the media content.
- FOLLOW: This relationship will connect users to other users.

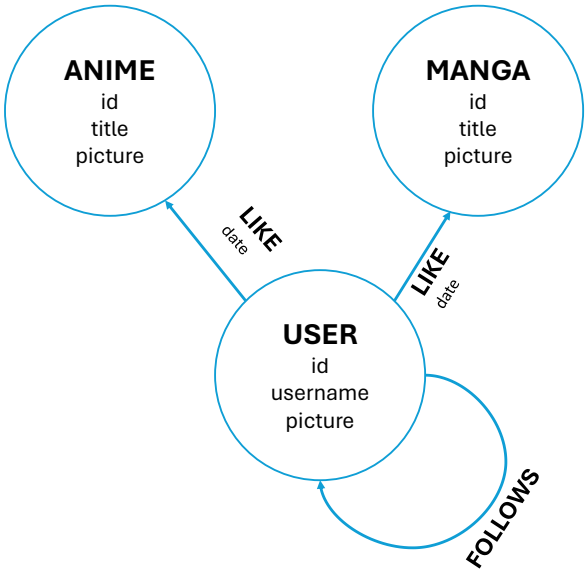


Figure 3.1: GraphDB

CRUD operations

- Create: This operation will allow users to create new nodes and relationships in the database. For example, users can create new relationships between users and media content:

  A user can LIKE a media content:

```
    MATCH (u:User {id: $userId}), (a:Anime {id: $animeId})
    WHERE NOT (u)-[:LIKE]->(a)
    CREATE (u)-[r:LIKE {date: $date} ]->(a)
    RETURN r

```

Listing 3.1: Create Like Relationship

  A user can FOLLOW another user:

```
    MATCH (u:User {id: $userId}), (f:User {id: $followedUserId})
    WHERE NOT (u)-[:FOLLOWS]->(f)
    CREATE (u)-[r:FOLLOWS]->(f)
    RETURN r

```

Listing 3.2: Create Follow Relationship

- Read: This operation will allow users to read nodes and relationships from the database. For example, users can read information about anime and manga and relationships between users and media content. A user can read the list of liked media contents:

```

    MATCH (u:User {id: $userId})-[:LIKE]->(a:Anime)
    RETURN a

```

Listing 3.3: Read Liked Media Contents

  A user can read the list of followers:

```
    MATCH (u:User {id: $userId})<-[:FOLLOWS]-(f:User)
    RETURN f

```

Listing 3.4: Read Followers

- Update: This operation will allow users to update nodes and relationships in the database. For example, users can update their likes for anime and manga and relationships between users.
- Delete: This operation will allow users to delete nodes and relationships from the database. For example, users can delete their likes for anime and manga and relationships between users.

  A user can unlike a media content:

```
    MATCH (u:User {id: $userId})-[r:LIKE]->(a:Anime {id: $animeId})
    DELETE r
    RETURN r

```

Listing 3.5: Delete Like Relationship

A user can unfollow another user:

```
MATCH (:User {id: $followerUserId})-[r:FOLLOWS]->(:User {id: $followingUserId})
DELETE r
RETURN r

```

Listing 3.6: Delete Follow Relationship

Availability and Partition Tolerance

MangaVerse, as a social network, gives priority to the AP configuration of the CAP theorem, ensuring Availability and Partition Tolerance. This allows users to access the application and interact with other users and media content, even if the data is not always consistent (Eventual Consistency).

Redundancy

The performance of the application is critical, so we need to ensure that the system is highly available and fault-tolerant. To achieve this, we gave priority to fast responses, rather than reducing memory consumption.

Latest reviews

In the anime and manga collections, there's a field containing the latest 5 reviews written for that specific media content, in this way it's fast to retrieve.

Average rating

In the anime and manga collections, there's a field containing the average rating of the media content, this field is updated every time a new review is written.

Number of likes

In the anime and manga collections, there's a field containing the number of likes, this field is updated every time a new like relationship is created or deleted.

Followers and Followings

In the user collection, there are fields containing the number of followers and followings, this field is updated every time a new follow relationship is created or deleted.

User field in Reviews

In the reviews collection, there's a field containing the user data, such as id, username, picture, and also location and birthday, which are used for suggestion porpouses.

Media Content field in Reviews

In the reviews collection, there's a field containg information about the anime or manga the reivew is about. This field has information about the media content id and title.

Review Ids

A list of review ids is stored in the anime, manga and users collections, this is used to quickly retrieve the reviews of a media content and of a user.

Replicas

A cluster of three nodes is available for this project, allowing deployment of replicas: however, replicas were only implemented in MongoDB, as Neo4j required the Enterprise version for it. We have 3 replicas for MongoDB and 1 for Neo4J. In MongoDB we have one primary and two secondary replicas, the primary is used for write operations and the secondaries are used for read operations. This will allow us to distribute the load and improve the performance of the application. In case of failure of the primary node, one of the secondary nodes will be promoted to primary, ensuring high availability of the system.

Sharding

Sharding is a method for distributing data across multiple machines to meet the demands of data growth. As the size of the data increases, a single machine may not be sufficient to store the data nor provide an acceptable read and write throughput. Sharding typically solves the problem with horizontal scaling. However, in our application, sharding is not possible due to the nature of the collections and their interdependencies.

Our database consists of four collections: anime, manga, users, and reviews and each collection stores different types of information. Sharding these collections would be complex due to their interrelated nature. The data in these collections is highly interconnected; for example, reviews reference both anime/-manga and users, and each user can have multiple reviews linked to both anime and manga. Distributing such interdependent data across multiple shards would lead to significant overhead in maintaining relationships between documents, potentially causing performance degradation instead of improvement.

Therefore, while the database design is ready for sharding in theory, the practical constraints of our application's structure and the interconnectedness of our collections make sharding infeasible. Implementing sharding would introduce complexity and overhead that could outweigh the benefits of horizontal scaling in this context.

Implementation

Development Environment

To ensure efficient and successful Implementation of MangaVerse web application, choosing the appropriate development environment is one of the most important points of the project.

Programming Languages

- **Backend:** Java is the main programming language used in the project's backend development.
- **Frontend:** HTML, CSS, JavaScript are utilized for building user interface in the project.
- **Data Preprocessing:** Python and java were used in the project to conduct data preprocessing task with the help of its powerful libraries and ease of use features.

Database

- **Document Database:** MongoDB is used in the project to store and manage document-based data with the help of its flexibility and scalability features.
- **Graph Database:** Neo4j is used in the project to manage and query graph data and handle complex relationships and connections between user entities and media contents in an efficient way.

Integrated Development Environment

Intellij IDEA was used as an primary IDE. It is powerful Java integrated development environment for developing software in an efficient way.

Version Control

Github was used to provide a collaborative development with its version control system.

Web Server

Apache Tomcat was used as a web server to provide reliable environment for deploying and running the java based web application.

Build Automation

Maven was used as a build automation tool. It is used to manage the project's build, reporting, and documentation from a central piece of information.

Testing

JUnit was used as a testing framework for Java code. It is used to write and run repeatable automated tests. This ensures the reliability and efficiency of the codebase throughout the development process.

Main Modules

- Configuration
- Controller
- DAO (Data Access Objects)

- DTO (Data Transfer Objects)
- Model
- Service
- Utils
- User Interface

Configuration

Configuration module contains a class named *AppServletContextListener* which is responsible for initializing and managing database connections for the web application. The configuration class implements ServletContextListener interface. *@WebListener* annotation is used to provide listening for application lifecycle events. This annotation contains two methods, which are *contextInitialized(ServletContextEvent sce)* and *contextDestroyed(ServletContextEvent sce)*. The first method is called when the web application is started and the second method is called when the web application is shut down.

Database Connection Management: Database connection is provided with *openConnection()* and *closeConnection()* methods. They are both initialized for managing connection for MongoDB and Neo4j databases. Connections are managed with corresponding DAO classes which are BaseMongoDBDAO and BaseNeo4jDAO.

With using the configuration module for database connection, web application ensures robustness and reliability in its data access layer.

Controller

The controller modules plays a role as intermediary between the user requests and backend of the MangaVerse wab application as servlet classes. They receives the user requests, process them and returns with the corresponding response. The controller classes are implemented using HttpServlet to handle user requests and responses. Within the scope of their intermediary role, the controller classes are responsible of being a bridge between the user interface and backend logic. When a user interacts with the web application, their actions are translated into a HTTP request and these requests are handled by the related servlet class in the controller module. To be able to do do request translation in an efficient way each controller class extend 'HttpServlet' and has various methods to handle HTTP requests like GET and POST. Each controller class utilized a switch-case structure to determine the action requested and invokes the appropriate handler method accordingly. This structure allows for clear and organized routing of request to their corresponding handler method. After processing the request, the servlet generates a requested response.

Example code snippet from MediaContentServlet:

```
{
    protected void processRequest(HttpServletRequest request, HttpServletResponse response
    ) throws ServletException, IOException {
        String action = request.getParameter("action");

        switch (action) {
            case "toggleLike" -> handleToggleLike(request, response);
            case "addReview" -> handleAddReview(request, response);
            case "deleteReview" -> handleDeleteReview(request, response);
            case "editReview" -> handleEditReview(request, response);
            case "getMediaContent" -> handleGetMediaContentById(request,response);
            case "getMediaContentByTitle" -> handleSearchMediaContentByTitle(request,
    response);
            case null, default -> handleLoadPage(request,response);
        }
    }
}

```

The controller module contains the following classes:

- Exception
  NotAuthorizedException: This exception is thrown when the user is not authorized to access the requested resource.
- AuthServlet
  The AuthServlet class handles the user authentication and authorization processes. It includes login, logout and sign up functions
- MainPageServlet
  The MainPageServlet class is responsible for handling the main page of the web application. It includes the main page of the web application and the search functionality. It provides request related to displaying main page and searching media contents.
- ManagerServlet
  The ManagerServlet class manages administrative requests in manager page. These request are primarily about manga, anime and user analytics such *averageRatingByMonth(), trendMediaContentByYear(), getBestCriteria()...*
- MediaContentServlet
  The MediaContentServlet class is responsible for managing request related with media contents. These requests include like, adding,deleting or editing reviews and retrieving media content details.
- ProfileServlet
  The ProfileServlet class is responsible for managing user profile related requests. These requests include updating user profile, following/unfollowing other users, getting user profile details such as liked anime and manga and user reviews.
- UserServlet
  The UserServlet class is responsible for managing user related requests and interactions. These requests include retrieving followers list, following list and user information.

DAO (Data Access Objects)

The DAO module includes the logic for accessing and managing data in the database and provides data retrieval, storage and manipulation. This module includes classes with CRUD (create, read, update, delete) operations and query executions. It provides a layer of abstraction between the database and the rest of the application and ensures the separation of concerns. The DAO module contains the following classes:

- Enums
  - DataRepositoryEnum
- Exceptions
- Interfaces
  - MediaContentDAO
  - ReviewDAO
  - UserDAO
- Mongo
  - AnimeDAOMongoImpl
  - BaseMongoDBDAO
  - MangaDAOMongoImpl
  - ReviewDAOMongoImpl
  - UserDAOMongoImpl
- Neo4j
  - AnimeDAONeo4jImpl
  - BaseNeo4jDAO
  - MangaDAONeo4jImpl
  - UserDAONeo4jImpl
- DAOLocator

Example code snippet from MangaDAOMongoImpl:

```
{
    //MongoDB queries
    //Best genres/themes/demographics/authors based on the average rating
    @Override
    public Map<String, Double> getBestCriteria (String criteria, boolean isArray, int page
    ) throws DAOException {
        try  {
            MongoCollection<Document> mangaCollection = getCollection(COLLECTION_NAME);
            int pageOffset = (page-1)*Constants.PAGE_SIZE;

            List<Bson> pipeline;
            if (isArray) {
                pipeline = List.of(
                        match(and(exists(criteria), ne("average_rating", null))),
                        unwind("$" + criteria),
                        group("$" + criteria, avg("criteria_average_rating", "
    $average_rating")),
                        sort(descending("criteria_average_rating")),
                        skip(pageOffset),
                        limit(25)
                );
            } else {
                pipeline = List.of(
                        match(Filters.exists(criteria)),
                        group("$" + criteria, avg("criteria_average_rating", "
    $average_rating")),
                        sort(new Document("criteria_average_rating", -1)),
                        skip(pageOffset),
                        limit(25)
                );
            }

            List <Document> document = mangaCollection.aggregate(pipeline).into(new
    ArrayList<>());
            Map<String, Double> bestCriteria = new LinkedHashMap<>();
            for (Document doc : document) {
                Double avgRating = doc.get("criteria_average_rating") instanceof Integer?
                        doc.getInteger("criteria_average_rating").doubleValue() :
                        doc.getDouble("criteria_average_rating");
                if (criteria.equals("authors")) {
                    bestCriteria.put(doc.get("_id", Document.class).getString("name"),
    avgRating);
                } else {
                    bestCriteria.put(doc.get("_id").toString(), avgRating);
                }
            }

            return bestCriteria;

        } catch (Exception e) {
            throw new DAOException(DAOExceptionType.GENERIC_ERROR, e.getMessage());
        }
    }
}

```

Example code snippet from UserDAONeo4jImpl:

```
{
    /**
    * Retrieves a list of users following a specific user from the Neo4j database.
    *
    * @param userId The ID of the user whose followers are to be retrieved.
    * @param loggedUserId The ID of the user requesting the list of followers.
    * @return A list of RegisteredUserDTO objects representing the followers of the
    specified user.
    * @throws DAOException If an error occurs while retrieving the followers list.
    */
   @Override
   public List<UserSummaryDTO> getFirstNFollowers(String userId, String loggedUserId)
    throws DAOException {
       try (Session session = getSession()) {
           StringBuilder queryBuilder = new StringBuilder("MATCH (follower:User)-[:FOLLOWS
    ]->(:User {id: $userId}) ");
           if (loggedUserId != null) {
               queryBuilder.append("WHERE follower.id <> $loggedUserId ");
           }
           queryBuilder.append("RETURN follower AS user ");
           queryBuilder.append("ORDER BY follower.username ");
           queryBuilder.append("LIMIT 10");
           String query = queryBuilder.toString();

           Map<String, Object> params = new HashMap<>();
           params.put("userId", userId);
           if (loggedUserId != null) {
               params.put("loggedUserId", loggedUserId);
           }

           List<Record> records = session.executeRead(
                   tx -> tx.run(query, params).list()
           );

           return records.isEmpty() ? null : records.stream()
                   .map(this::recordToUserSummaryDTO)
                   .toList();

       } catch (Neo4jException e) {
           throw new DAOException(DAOExceptionType.DATABASE_ERROR, e.getMessage());

       } catch (Exception e) {
           throw new DAOException(DAOExceptionType.GENERIC_ERROR, e.getMessage());
       }
   }
}

```

DTO (Data Transfer Objects)

The DTO modules are the intermediary class between presentation layer and the DAO module in the web application. They transfer data structures between different layers and components of the application in a more standardized way.

Model

- Enums
- Media Content
  - Anime
  - Manga
  - Manga Author
  - Media Content
- Registered User
  - Mangager
  - Registered User
  - User
- Review

Service

Service module has also important role in the web application. The classes in the service module are responsible for containing the business logic and maintaining interaction between the DAO classes and the presentation layer. It handles complex operations with guarantying that the application's core functionalities are executed correctly. Some of the services that are provided in the service module are: *UserService, MediaContentService, ReviewService, TaskManager, ExecuterTaskService.* The package structure of Service module is as follows:

- enums
  - ExecuterTaskService
- exceptions
  - enums
  —— BusinessExceptionType
  - BusinessException
- impl
  - asinc media tasks
  —— CreateMediaTask
  —— DeleteMediaTask
  —— UpdateAverageRatingTask
  —— UpdateMediaRedundancyTask
  —— UpdateMediaTask
  —— UpdateNumberofLikesTask
  - asinc review tasks
  —— RemoveDeletedMediaReviewsTask
  —— RemoveDeletedUserReviewsTask
  —— UpdateReviewRedundancyTask
  - asinc user tasks
  —— CreateUserTask
  —— DeleteUserTask
  —— UpdateNumberOfFollowedTask
  —— UpdateNumberOfFollowersTask
  —- UpdateUserTask
  - AperiodicExecutorTaskServiceImpl
  - ErrorTaskManager
  - MediaContentServiceImpl

- PeriodicExecutorTaskServiceImpl
- ReviewServiceImpl
- UserServiceImpl
- interfaces
—— ExecuterTaskService
—— MediaContentService
—— ReviewService
—— Task
—— TaskManager
—— UserService

- ServiceLocator

Adopted Patterns and Techniques

Patterns

Techniques

Task Manager:
Task Manager class which is located in the service module of the system provides asynchronous task execution with using *PriorityBlockingQueue*. It helps to order the tasks according to their prioritizes. After that prioritization, it ensures that higher priority tasks will be executed first and if two tasks have the same priority the one which is created before will be executed first. While Task Manager class is able to start and stop the tasks within the functions inside, it can also take tasks to the queue in a thread-safe way. By using taskComparator for ordering the tasks, the system provides also effective scheduling and execution.

Aperiodic Executor Task Service:
Executor Task Service class which is located inside the service module of the system is an important part for providing the eventual consistency. Executing tasks in asynchronous way with threads guarantees eventual consistency across different collections, mongoDB and neo4j and different replicas. With the help of the Executor Task Service, tasks that are needed to be executed in an asynchronous way are handled by ensuring that changes propagate correctly across different part of the system. While using multiple databases and data replicas for this web application, it is important for maintain data integrity and eventual consistency. Executing the tasks in an asynchronous way using threads allows to perform operations without blocking the main execution flow. Aperiodic executer task service class is implemented by using the interface of executor service.

Description of Main Classes

Controller

Class	Description
AuthServlet	Handles business logic for authentication
MainPageServlet	Handles business logic for main page
ManagerServlet	Handles business logic for manager
MediaContentServlet	Handles business logic for media content
ProfileServlet	Handles business logic for profile
UserServlet	Handles business logic for user

DAO

Class	Sub-package	Description
MediaContentDAO	interfaces	Collection of methods for media content database related entities on mongoDB
ReviewDAO B	interfaces	Collection of methods for review database related entities on mongoDB
UserDAO	interfaces	Collection of methods for user database related entities on mongoDB
AnimeDAOMongoImpl	mongo	Contains all the method implementation for the MongoDB database anime entities
BaseMongoDBDAO	mongo	Contains all the method implementations for the MongoDB database
MangaDAOMongoImpl	mongo	Contains all the method implementations for the MongoDB database manga entities
ReviewDAOMongoImpl	mongo	Contains all the method implementations for the MongoDB database review entities
UserDAOMongoImpl	mongo	Contains all the method implementations for the MongoDB database user entities
AnimeDAONeo4jImpl	neo4j	Contains all the method implementation for the Neo4j database anime entities
BaseNeo4jDAO	neo4j	Contains all the method implementations for the Neo4j database
MangaDAONeo4jImpl	neo4j	Contains all the method implementation for the Neo4j database manga entities
UserDAONeo4jImpl	neo4j	Contains all the method implementation for the Neo4j database user entities
DAOLocator		Implements the locator pattern for accessing DAOs based on the specified data repository

DTO

Class	Sub-package	Description
AnimeDTO	mediaContent	Represents data transfer object containing attributes for animes
MangaDTO	mediaContent	Represents data transfer object containing attributes for mangas
MediaContentDTO	interfaces	Defines common attributes for media content
DashboardDTO	statistics	Contains statistical data for the dashboard
MongoDBStats	statistics	Provides statistics specific to MongoDB
LoggedUserDTO		Holds information about a logged-in user.
PageDTO		Represents pagination details
ReviewDTO		Contains attributes for reviews

Class	Sub-package	Description
UserRegistrationDTO		Holds data for user registration
UserSummaryDTO		Provides a summary of user information

Model

Class	Sub-package	Description
Anime	mediaContent	Provides unique anime attributes by extending parent class MediaContent and related getter and setter methods.
Manga	mediaContent	Provides unique manga attributes by extending parent class MediaContent and related getter and setter methods.
MangaAuthor	mediaContent	Contains manga author attributes and related getter and setter methods.
MediaContent	mediaContent	Contains all the attributes used by types of media contents and their getter and setter methods.
Manager	registeredUser	Provides unique manager attributes by extending parent class RegisteredUser and related getter and setter methods.
RegisteredUSer	registeredUser	Contains all the attributes used by types of registered users and their getter and setter methods.
User	registeredUser	Provides unique user attributes by extending parent class RegisteredUser and related getter and setter methods.
Review		Contains review attributes and related getter and setter methods.

Service

Class	Sub-package	Description
CreateMediaTask	impl/ asinc_media_tasks	Implementation of methods for media task creation for MediaContentService
DeleteMediaTask B	impl/ asinc_media_tasks	Implementation of methods for media task deletion for MediaContentService
RefreshLatestReviewsTasks	impl/ asinc_media_tasks	Implementation of methods for refreshing latest reviews for MediaContentService
UpdateAverageRatingTask	impl/ asinc_media_tasks	Implementation of methods for updating average rating for MediaContentService
UpdateMediaRedundancyTask	impl/ asinc_media_tasks	Implementation of methods for updating media redundancy for MediaContentService
UpdateMediaTask	impl/ asinc_media_tasks	Implementation of methods for updating media for MediaContentService
UpdateNumberofLikesTask	impl/ asinc_media_tasks	Implementation of methods for updating numbers of likes for MediaContentService
RemoveDeletedMedia ReviewsTask	impl/ asinc_review_tasks	Implementation of methods for removing reviews of deleted media for ReviewService

Class	Sub-package	Description
RemoveDeletedUser ReviewsTask	impl/ asinc_review_tasks	Implementation of methods for removing reviews of deleted user for ReviewService
UpdateReviewRedundancyTask	impl/ asinc_review_tasks	Implementation of methods for updating review redundancy for ReviewService
CreateUserTask	impl/ asinc_user_tasks	Implementation of methods for user creation for UserService
DeleteUserTask	impl/ asinc_user_tasks	Implementation of methods for user deletion for UserService
UpdateNumberOfFollowedTask B	impl/ asinc_user_tasks	Implementation of methods for updating number of followed for UserService
UpdateNumberOfFollowersTask	impl/ asinc_user_tasks	Implementation of methods for updating number of followers for UserService
UpdateUserTask	impl/ asinc_user_tasks	Implementation of methods for updating user for MediaContentService
AperiodicExecutor TaskServiceImpl	impl	Implementation of aperiodic tasks for ExecutorTaskService
ErrorTaskManager	impl	Implementation of TaskManager interface to handle error
MediaContentServiceImpl	impl	Implementation of MediaContentService, providing media content operations
PeriodicExecutor TaskServiceImpl	impl	Implementation of periodic tasks for ExecutorTaskService
ReviewServiceImpl	impl	Implementation of ReviewService, providing review operations
UserServiceImpl	impl	Implementation of UserService, providing user operations
ExecutorTaskService	interfaces	Collection of methods for task management
MediaContentService	interfaces	Collection of methods for media content service
ReviewService	interfaces	Collection of methods for review service
Task	interfaces	Collection of methods for execution operations
TaskManager	interfaces	Collection of methods for managing task prioritization
UserService	interfaces	Collection of methods for user service
ServiceLocator		Implements locator pattern for services

MongoDB queries

Some of the most important MongoDB queries for analytic and suggestion purposes.

USERS:

- GetDistribution query to get the user's location, gender, birthday year that gave the highest rating to the application:

```
// Match stage to filter documents where 'criteriaOfSearch' exists
db.collection.aggregate([
    {
        $match: {
            [criteriaOfSearch]: { $exists: true }
        }
    },
    // Project stage to include 'criteriaOfSearch' and 'app_rating' fields
    {
        $project: {
            [criteriaOfSearch]: 1,
            app_rating: 1
        }
    },
    // Group stage to count occurrences of each 'criteriaOfSearch'
    {
        $group: {
            _id: "$" + criteriaOfSearch,
            count: { $sum: 1 }
        }
    },
    // Sort stage to sort documents by 'count' in descending order
    {
        $sort: {
            count: -1
        }
    }
]);

```

Listing 4.1: GetDistribution

ANIME/MANGA:

- GetBestCriteria query, the criteria can be genres, demographics, themes, authors and serialization for manga; tags, producers, studios for anime:

```
  db.collection.aggregate([
      // Match stage to filter documents where 'criteria' exists and 'average_rating' is
     not null
      {
          $match: {
              criteria: { $exists: true },
              average_rating: { $ne: null }
          }
      },
      // Unwind stage to deconstruct the 'criteria' array field
      {
          $unwind: "$" + criteria
      },
      // Group stage to calculate the average rating for each criteria
      {
          $group: {
              _id: "$" + criteria,
              criteria_average_rating: { $avg: "$average_rating" }
          }
      },
      // Sort stage to sort documents by 'criteria_average_rating' in descending order
      {
          $sort: {
              criteria_average_rating: -1
          }
      },
      // Skip stage to skip the first 'pageOffset' documents
```

```
        {
            $skip: pageOffset
        },
        // Limit stage to limit the results to 25 documents
        {
            $limit: 25
        }
    ]);

```

Listing 4.2: GetBestCriteria

REVIEWS:

- GetMediaContentRatingByYear query to get the average rating of media content by year:

```
// Match stage to filter documents based on specified conditions
db.collection.aggregate([
    {
        $match: {
            [`${nodeType}.id`]: new ObjectId(mediaContentId),
            rating: { $exists: true },
            date: { $gte: startDate, $lt: endDate }
        }
    },
    // Group stage to group documents by year and calculate the average rating
    {
        $group: {
            _id: { $year: "$date" },
            average_rating: { $avg: "$rating" }
        }
    },
    // Project stage to shape the output documents
    {
        $project: {
            _id: 0,
            year: "$_id",
            average_rating: 1
        }
    },
    // Sort stage to sort documents by year in ascending order
    {
        $sort: { year: 1 }
    }
]);

```

Listing 4.3: GetMediaContentRatingByYear

- SuggestMediaContent query to suggest media content based on common criteria, like birthday or location:

```
db.collection.aggregate([
{
  // Match documents based on a dynamic user criteria
  $match: {
    ["user." + criteriaType]: criteriaValue
  }
},
{
  // Group documents by node type ID and calculate the first title and average rating
  $group: {
    _id: "$" + nodeType + ".id", // Group by the node type's ID
    title: { $first: "$" + nodeType + ".title" }, // Get the first title in the group
    average_rating: { $avg: "$rating" } // Calculate the average rating for the group
  }
},
{
  // Sort the grouped documents by average rating in descending order
  $sort: { average_rating: -1 }
```

```
  },
  {
    // Limit the number of results to the page size constant
    $limit: Constants.PAGE_SIZE
  }
]);
```

Listing 4.4: SuggestMediaContent

GraphDB queries

Some of the most important Neo4j queries for analytic and suggestion porpouses.

USERS:

- Suggest other users followings in common:

```
MATCH (u:User {id: $userId})-[:FOLLOWS]->(following:User)<-[:FOLLOWS]-(suggested:User)
WHERE NOT (u)-[:FOLLOWS]->(suggested) AND u <> suggested
WITH suggested, COUNT(DISTINCT following) AS commonFollowers
WHERE commonFollowers > 5
RETURN suggested as user, commonFollowers
ORDER BY commonFollowers DESC
LIMIT $n
```

Listing 4.5: SuggestUsersByCommonFollowings

- Suggest other users likes in common:

```
MATCH (u:User {id: $userId})-[r:LIKE]->(media:Manga)<-[:LIKE]-(suggested:User)
WHERE u <> suggested AND r.date >= $date
WITH suggested, COUNT(DISTINCT media) AS commonLikes
WHERE commonLikes > $min
RETURN suggested AS user, commonLikes
ORDER BY commonLikes DESC
LIMIT $n
```

Listing 4.6: SuggestUsersByCommonLikes

ANIME/MANGA:

- Suggest anime and manga based on following in common:

```
MATCH (u:User {id: $userId})-[:FOLLOWS]->(f:User)-[r:LIKE]->(a:Anime)
WHERE NOT (u)-[:LIKE]->(a) AND r.date >= $startDate
WITH a, COUNT(DISTINCT f) AS num_likes
RETURN a AS anime
ORDER BY num_likes DESC
LIMIT $n
```

Listing 4.7: GetSuggestedByFollowings

- Suggest anime and manga based on likes in common between followers:

```
MATCH (u:User {id: $userId})-[r1:LIKE]->(a:Anime)<-[:LIKE]-(f:User)
WHERE r1.date >= $startDate
WITH u, f, COUNT(a) AS common_likes
ORDER BY common_likes DESC
LIMIT 20
MATCH (f)-[:LIKE]->(a2:Anime)
WHERE NOT (u)-[:LIKE]->(a2)
WITH a2, COUNT(DISTINCT f) AS num_likes
RETURN a2 AS anime
ORDER BY num_likes DESC
LIMIT $n
```

Listing 4.8: GetSuggestedByLikes

- Get the trend of media contents in a specific year based on the number of likes:

```
MATCH (a:Anime)<-[r:LIKE]-(u:User)
WHERE r.date >= $startDate AND r.date < $endDate
WITH a, count(r) AS numLikes
ORDER BY numLikes DESC
RETURN a AS anime, numLikes
LIMIT $n
```

Listing 4.9: GetTrendMediaContentByYear

- Get the general trend of media contents based on the number of likes:

```
MATCH (u:User)-[r:LIKE]->(a:Anime)
WHERE r.date >= $startDate
WITH a, COUNT(r) AS numLikes
ORDER BY numLikes DESC
RETURN a AS anime, numLikes
LIMIT $n
```

Listing 4.10: GetMediaContentTrendByLikes

Testing

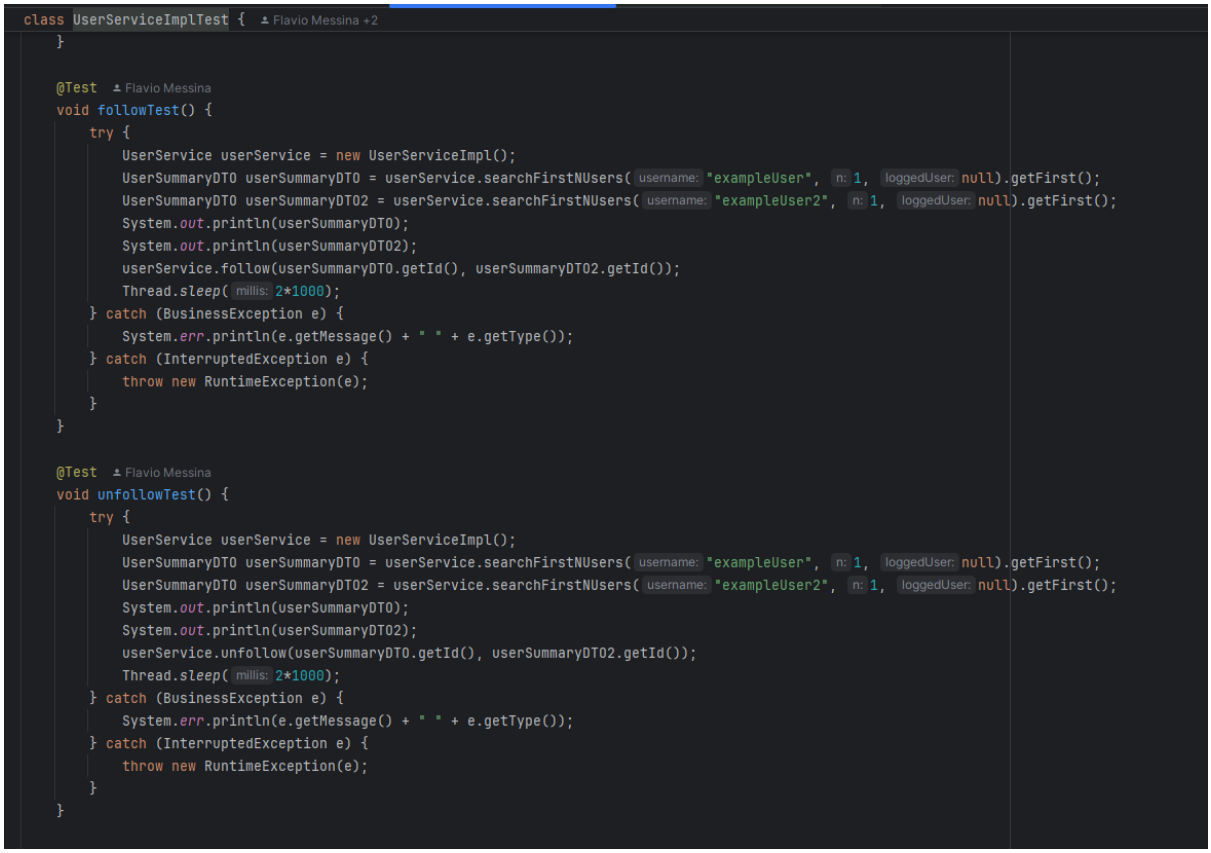
Testing is a substantial part of the MangaVerse web application project. Testing helps to ensure application's reliability, performance and correctness. To be able to conduct efficient testing process, two kind of tests are preformed. They are JUnit testing as a structural testing and functional testing.

Structural Testing

Structural testing also with other name white-box testing is based on testing the internal structure of the working application and it guarantees that the methods are working as expected.JUnit testing framework is used to conduct structural testing. JUnit testing is performed by testing different modules of the application such as DAOs and services. With that process each methods efficiency and correctness is guaranteed. Some examples of JUnit testing are shown below.

```
class MangaDAOMongoImplTest {  Flavio Messina +1 *
    // test 1 : update an existing manga (before that, I try to find the manga by title)
    // test 2 : update a non-existing manga
    @Test  Flavio Messina *
    void updateMediaContentTest() throws DAOException {
        MangaDAOMongoImpl mangaDAO = new MangaDAOMongoImpl();
        List<MediaContentDTO> mangaList = mangaDAO.search(List.of(Map.of( k1: "title",  v1: "Sample Manga")), Map.of( k1: "title",  v1: 1),  page: 1,  reducedInfo: false).getEntries();
        // test 1
        if (!mangaList.isEmpty()) {
            MediaContentDTO mangaToUpdate = mangaList.getFirst();
            System.out.println("Manga to update: " + mangaToUpdate);
            Manga manga = new Manga();
            manga.setId(mangaToUpdate.getId());
            manga.setTitle("Updated Manga");
            assertDoesNotThrow(() -> mangaDAO.updateMediaContent(manga));
            System.out.println("Manga updated");
        }
        // test 2
        Manga manga = createSampleManga();
        manga.setId("6635632b4276578429f29384");
        manga.setTitle("Non-existing Manga");
        assertThrows(DAOException.class, () -> mangaDAO.updateMediaContent(manga));
        System.out.println("Non-existing Manga not found");
    }
    // test 1 : delete an existing manga (before that, I try to find the manga by title)
    // test 2 : delete a non-existing manga
    @Test  Flavio Messina *
    void deleteMediaContentTest() throws DAOException {
        MangaDAOMongoImpl mangaDAO = new MangaDAOMongoImpl();
        List<MediaContentDTO> mangaList = mangaDAO.search(List.of(Map.of( k1: "title",  v1: "Sample Manga")), Map.of( k1: "title",  v1: 1),  page: 1,  reducedInfo: false).getEntries();
        // test 1
        if (!mangaList.isEmpty()) {
            MediaContentDTO mangaToDelete = mangaList.getFirst();
            System.out.println("Manga to delete: " + mangaToDelete);
            assertDoesNotThrow(() -> mangaDAO.deleteMediaContent(mangaToDelete.getId()));
            System.out.println("Manga deleted");
        }
        // test 2
        assertThrows(DAOException.class, () -> mangaDAO.deleteMediaContent( mangaId: "6635632b4276578429f29384"));
        System.out.println("Non-existent manga not deleted");
    }
```

Figure 5.1: MangaDAOMongoImpl Class Test Example



```
class UserServiceImplTest {  Flavio Messina +2
    }

    @Test  Flavio Messina
    void followTest() {
        try {
            UserService userService = new UserServiceImpl();
            UserSummaryDTO userSummaryDTO = userService.searchFirstNUsers( username: "exampleUser",  n: 1,  loggedUser: null).getFirst();
            UserSummaryDTO userSummaryDTO2 = userService.searchFirstNUsers( username: "exampleUser2",  n: 1,  loggedUser: null).getFirst();
            System.out.println(userSummaryDTO);
            System.out.println(userSummaryDTO2);
            userService.follow(userSummaryDTO.getId(), userSummaryDTO2.getId());
            Thread.sleep( millis: 2*1000);
        } catch (BusinessException e) {
            System.err.println(e.getMessage() + " " + e.getType());
        } catch (InterruptedException e) {
            throw new RuntimeException(e);
        }
    }

    @Test  Flavio Messina
    void unfollowTest() {
        try {
            UserService userService = new UserServiceImpl();
            UserSummaryDTO userSummaryDTO = userService.searchFirstNUsers( username: "exampleUser",  n: 1,  loggedUser: null).getFirst();
            UserSummaryDTO userSummaryDTO2 = userService.searchFirstNUsers( username: "exampleUser2",  n: 1,  loggedUser: null).getFirst();
            System.out.println(userSummaryDTO);
            System.out.println(userSummaryDTO2);
            userService.unfollow(userSummaryDTO.getId(), userSummaryDTO2.getId());
            Thread.sleep( millis: 2*1000);
        } catch (BusinessException e) {
            System.err.println(e.getMessage() + " " + e.getType());
        } catch (InterruptedException e) {
            throw new RuntimeException(e);
        }
    }
```

Figure 5.2: UserServiceImpl Class Test Example

Functional Testing

Functional testing also with other name black-box testing is based on testing the application's external functionalities. It checks the application from end-user's perspective. It ensures that specified requirements are provided efficiently by the web application and expected is outcome is created. With the help of the use cases and real world scenarios, functional testing is conducted. Some examples of functional testing are shown below.

Table 5.1: Admin Test case

Id	Description	Input	Expected Output	Output	Outcome
User_01	Login with correct information	email: nmiller@example.com, password: f6d6b3ffecb44a...	The user logs in successfully		
User_02	Login with wrong information	email: wrong@example.com, password: wrong	The user is not able to log in successfully		
User_03	Signup with all the mandatory info are filled				
User_04	Signup with missing info				
User_05	Update user information	description: manga lover	User profile is updated with new info.		
User_06	Follow another user	-	User is followed.		
User_07	Unfollow another user	-	User is unfollowed.		
User_08	Search manga by title	title: "Slam Dunk"	The list of manga which includes the words of "Slam Dunk" is shown.		
User_09	Search manga by detailed filtering				
User_10	Like anime	-	The anime is liked		
User_11	Add review to anime	review:"I like the anime"	The review is added to the anime and displayed in the anime page		
User_12	Update review	review: "I dont like this anime anymore"	The review is updaated with the new one.		
Admin_01	See users distribution analytics	-			
Admin_02	See manga analytics for get average rating by month	Year:2020	Average rating for each month in 2020 is displayed in the page		
Admin_03	See anime analytics for get trend media content by year				
Admin_04					

Conclusion

Conclusion

The MangaVerse is a web application project that provides a comprehensive web application for dynamic social platform for manga and anime enthusiasts. The web application allows users to explore, search media content and be in contact with other users by review system. Having a user-friendly interface, the application is designed to have a robust set of features. The applications offers functionalities for both unregistrered user and registered user including manager purposes such as browse media content, personalized recommendations, profile management and analytics checking for management purposes.

Beside the functional requirements, the application has also well-defined development process and architecture using different technologies and techniques. While java is used for main backend development programming language, as a database MongoDB and Neo4j are used.

Future Work

For the future: manager will be able to update add delete anime and manga and delete user.

Security

- Data Encryption: All user data, including passwords, should be securely encrypted during transmission and storage.
- Delete user accounts if necessary.user management