Department of Computer Engineering

Senior Design Project
High Level Design Report

Supervisor:

Asst. Prof. Mustafa Özdal

Jury Members:

Asst. Prof. Ercüment Çiçek

Asst. Prof. Hamdi Dibeklioğlu

Innovation Expert:

Ahmet Kocamaz

Presented by:

Erkam Berker Şenol

Mert Aytöre

Gökhan Şimşek

Dias Alymbekov

Figali Taho

This report is submitted to the Department of Computer Engineering of Bilkent University in partial fulfillment of the requirements of the Senior Design Project course CS491/2.

Contents

1. Introduction

With the rise of social media platforms mainly in the last decade, the face of journalism and reporting has experienced a dramatic change from the traditional media sources and traditional reporting methods, to a more digital landscape, with the proliferation of Twitter and Facebook usage. According to a research by Pew Research Center, about four in ten Americans get their news online, and this number is expected to rise with the younger audience shifting to web sources [1]. This digitalization has not only made it more possible for media companies to find and reach a wider audience, but has also enabled non-professional civilians to partake in the news creation cycle. This has led to the rise of concepts such as citizen journalism and collaborative journalism.

Collaborative journalism is where multiple sources act together in order to create a news story, and not only professionals but also citizens can collaborate in this form of journalism. What we consume as news is usually a result of collaborative journalism, and the part citizen journalism plays in this collaboration is rising.

While there are numerous news outlets that provide collaborative journalism, and social media platforms that showcase these news pieces, there does not exist a popular platform focused exclusively on news that displays information on events not only via professional news outlets but also from the citizens' point of view. Furthermore, it often happens that we have to search extensively to be able to see beyond what algorithms choose as "appropriate/personalized" material, or to be able to escape filter bubbles we often find ourselves in nowadays on the web [2]. With Diafano, we propose to fill this void and create a platform that combines citizen stories with professional news, and provides multi-perspective, transparent journalism.

1.0.1   Purpose of the system

With Diafano, we aim to provide our users a unique experience of getting information. Diafano will be a platform that combines the already available data scattered through various platforms and applications that represent news outlets, as well as data the people feed to the system, making it a platform for receiving hands on information that is unbiased, and independent of user. The users will be able to view information pertaining to a specific location in the world, a certain event going on currently, via searching in a map, or in a search box. Additionally, we want to users to have easy access to local and global news, so we aim to provide this functionality too.

1.0.2   Design Goals

1.0.2.1   Usability

Being that Diafano should help people reach information and news of all formats in a unified place, it should be easy to use. We aim to maximize thus the usability of the web and Android application. Usability will further be promoted with the filtering of content via filtering algorithms for images, text, and user defined filtering via expression of abusiveness or inappropriateness of a content.

1.0.2.2   Reliability

Diafano's reliability is a serious matter. We aim to have 100% trustworthy content that is sourced from users as well as trustable news outlets. As already mentioned, users themselves will need to interact with the application in order for it to become more reliable. On the back end, our algorithms will make sure content is discarded - deleted - flagged in case it encompasses any sign of abusiveness, violence, profanity.

1.0.2.3   Efficiency

The system should be fast, and the response time should be very fast. Since there will be many interactions with the system, it is crucial to make the data retrieval and server-client

interaction processes as efficient as possible.

1.0.2.4   Security

Our application will be accessible only when the user logs in to his account with his password. Data will not be corrupted, and it will be stored safely in our servers. Anything the user opts out of sharing, will no longer be shared whenever they please.

1.0.2.5   Portability

The system should be able to be used in different software and hardware platforms.

1.0.3   Definitions, Keywords, Acronyms and Abbreviations

Explore local: Browse section for the local news.
Explore global: Browse section for the global news.
Stream: Streaming button for video streaming live. Upon user's request, the video will be stored/posted by user on the system to be later accessed by anyone whom the user has permitted via settings.
UI: User Interface
API: Application Programming Interface
HTTP: Hypertext Transfer Protocol
TCP: Transmission Control Protocol
Server: The part of the system responsible from logical operations, scheduling, and data management
Client: The part of the system the users interact with

2. Current System

In this chapter, existing apps and systems that might be considered similar to the proposed system are presented. Towards the end of the chapter, the qualities that set the proposed system apart from these existing apps will be explained.

2.1 Aggregated news apps and platforms

2.1.1 Flipboard

Flipboard is one of the most popular and successful news aggregation apps in the field.

- Available on mobile, for iPhone, Android, and Windows Phones. Flipboard also has a web platform.

- Personalized news feed, known as "Smart Magazines". Users can select interests and create magazines that are updated frequently and improved by learning users' preferences.

- "Cover Stories" feature lets users find an overview of news in topics that they are interested in.

- Users can opt out of seeing content from certain news outlets. This feature grants users the chance to limit the sources that they see.

2.1.2 Google News

Google News is a computer-generated news site that aggregates headlines from news sources worldwide, groups similar stories together and displays them according to each reader's personalized interests. [3]

- Google News is offered both as mobile apps and as a web platform.

- Google News collects news from over 50000 sources and utilizes a ranking algorithm to show the users only the relevant news. [4]

- Users can see news on the topic that they choose, and Google learns from these choices to make personalized suggestions in its "For You" tab. Google also utilizes information about users that they have collected previously, such as their search history. This results in a more personalized experience compared to other products.

- Users can view local news from the "Local" tab. Users can also change locations to view news from different parts of the world, and add new locations to be shown on their tab.

- Users can see news via default sections that focus on broad fields, such as business, technology, and politics. Furthermore, they can create personal sections by defining search terms.

- Users can search for news articles by keyword and date.

2.1.3 Apple News

Apple's official News app, formally known as Newsstand before iOS9. Current features and limitations:

- This app is available only in the US, UK, and Australia.[5] Also, it is not offered as a web platform.

- Personalized news stories. Users can follow channels and topics to get more relevant stories on their dashboard.

- Spotlight. This feature is Apple's version of story of the day.

- A trending topic feature that is similar to that of Twitter.

- Users can save stories for consuming them at a later time.

- Users can see their story history.

2.2   Broadcasting apps and platforms

2.2.1   Periscope

Periscope is a popular live video streaming app acquired and owned by Twitter.

- Periscope is available for iOS and Android, and has a web platform as well. Streaming on Periscope can only be done on mobile devices such as smart phones and tablets.

- Users can sign up with their Twitter, Facebook or Google accounts.

- Users can watch live videos from any person that is broadcasting.

- Users can search channels by hashtag, and see how many people are broadcasting about the said topic or hashtag.

- Users can comment and like broadcasts as they are happening.

- Broadcaster can disable comments, and has the option to make their video public or private.

- Broadcasters can save their videos.

- Users can broadcast whatever they want, as long as it conforms with the policy of Periscope. Thus, Periscope is not focused on news only.

2.2.2 Facebook Live

Facebook Live was introduced as a feature of Facebook gradually over the years of 2016 and 2017. Its features include:

- Facebook Live is available on mobile apps of Facebook and on PCs.

- Users can broadcast live streams, with features such as filters, drawing and doodling on the video.

- Users can interact, react to and comment on live streams.

- Users can get notified when a person they follow starts streaming, granted they have set their notification settings accordingly.

- Maximum time for a stream is four hours.

- Facebook Live Map allows users to see a world map with people currently live streaming represented as dots. Thus, people can find live streams from all over the world.

- Users can search keywords to find live streams.

- Broadcasters can share their screens while streaming.

2.2.3 Instagram Live

Popular photo sharing platform Instagram, owned by Facebook, rolled out its live-video functionality in November 2016. Its features include:

- Users can stream using their mobile devices on Instagram app.

- Viewers can comment and react on a live video.

- Users are notified, if their settings allow it, when a person they follow is live streaming.

- Broadcasters have the option to preserve their live stream as an Instagram story for 24 hours.

- With a current functionality, rolled out in October 2017, two users can broadcast using the same stream.

- "Explore" tab shows personalized live stream suggestions to users.

2.2.4   stre.am

Stream - Live Video Community is an application offered for iphone and Android smartphones.

- Users can view streams even if they're not logged into the application.

- Users can register with their Google+, Facebook or email accounts.

- Users can follow other users and can't thus miss notifications from the live streams of those they follow.

- While users are following a stream, they can share, tweet or re-stream this content. They can also react to the stream, and also comment.

2.3   Citizen journalism apps and platforms

2.3.1   Twitter

While not aimed only at citizen journalism, Twitter has become a milestone for numerous protests and movements across the globe, influencing events such as Gezi Protests and the Arab Spring. [6] [7] With its concise style and broad user base, Twitter has become the main platform for collaborative journalism. Some of Twitter's features include:

- Twitter is available on mobile devices as well as PCs.

- Users can share Tweets with at most 140 characters, although this cap is planned to be 280 in the future. Users can also share media such as photos and videos with the tweets.

- Users can follow other users, and send direct messages if the people they follow allow direct messaging.

- Users can see trending topics by specifying a region, e.g. their local city, such as Ankara, or a country, or worldwide.

- Users can search for content using hashtags and other well developed search options.

2.3.2 Reddit

Reddit is one of the most popular discussion websites in the world. Every post on Reddit is created by a user, and Reddit allows users to write their own posts. This, along with commenting on news articles in a forum style, has made Reddit a powerful news source with both official and citizen journalism. Some of Reddit's features include:

- Reddit is available on mobile as well as a web platform.

- Users can subscribe to different subreddits depending on their interests.

- Users can navigate through their personalized Front Page with posts from their subscribed subreddits.

- Users can post links, or write posts of their own.

- Users can upvote and downvote posts and comments, altering the visibility and popularity of the said post or comment if enough upvotes or downvotes are accumulated.

- During a major event happening worldwide, Reddit provides live coverage with links to tweets and articles posted by users. This allows for a smooth consuming of information and opinions from different perspectives during crises.

2.3.3 News stations - e.g. The Guardian, CNN

Multiple official news outlets have started allowing users to submit information, in various forms of text and media. Some examples include CNN iReport [8], BBC's Have Your Say [9] and ABC's World News Story Contribution [10]. The biggest downside of these contributions is that the submission undergo an election process by the people working at the news outlets. A submitted story might not be published on numerous grounds, including censorship. This

option provides an indirect platform for users, and is less powerful compared to social media platforms.

2.4   Diafano

The app that is proposed aims to gather together the best parts of the services that are mentioned above. The main difference of Diafano compared to other social media platforms is its sole focus on news. Diafano aims to offer its users a variety of views on the news that are happening around the world. It aims to make away with long chain of comments, and the tiresome job of scrolling through sub-content for information, by introducing cleaner and more concise ways of sharing information.

3.   Subsystem Decomposition

3.1   Overview

In the subsystem decomposition, the subsystem structure of our system is described in detail. First of all, the partitioning of the system is shown with diagrams and the classes within each layer are presented. Afterwards, hardware/software mapping of the system is provided which shows the allocation of resources and how various parts of our system will work in different hardware components. The persistent data management is also depicted; our database system and persistent objects are described. Access control and security defines the access boundaries of the users and how the security is managed in our system. In global software control, we illustrate how the software is controlled and how the general flow behaves in our system. Finally, in the boundary conditions we specify the initialization, termination, and failure conditions of the system.

3.2   Subsystem Decomposition

Diafano will be built based on the Client-Server architecture. The server side will be responsible for handling the broadcasting of the video to the clients. We do not want to put the load of the broadcasting to the edge devices, therefore the content of an event is transfered to the server and distributed from there.

The mobile application is the client part that presents the data a user and provides means to a user to post some news. The server controls manages the network of users, controls and aggregates the data produced by the users. The Client/Server architecture ensures that the scope of the project preserved. The server manages the data, applies filters and presents data

to users. In this way, the client side can focus solely on responding to user inputs. Within the Server/Client architecture, we adopt the 4-layer concept, in which the system is divided into the following layers: Presentation, Control, Logic and Data. The lowest data layer manages the persistency of the data we store. The stored data is passed to the logic layer (tier), in which the major logical operations are handled. These include filtering and aggregating data. In the end the processed data is passed to the presentation layer which communicates to the server using control layer. The data taken from a user is going through the same pathway backwards.
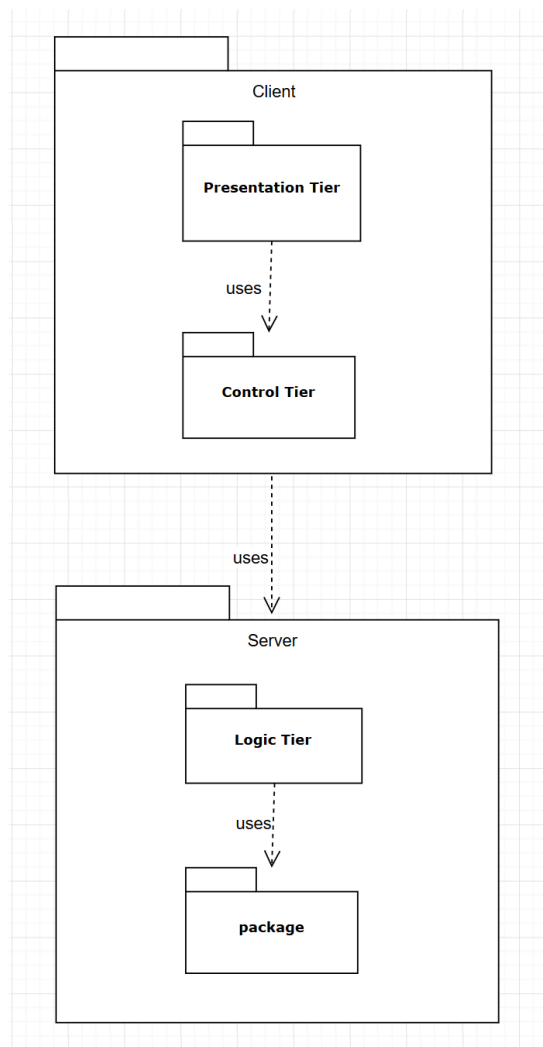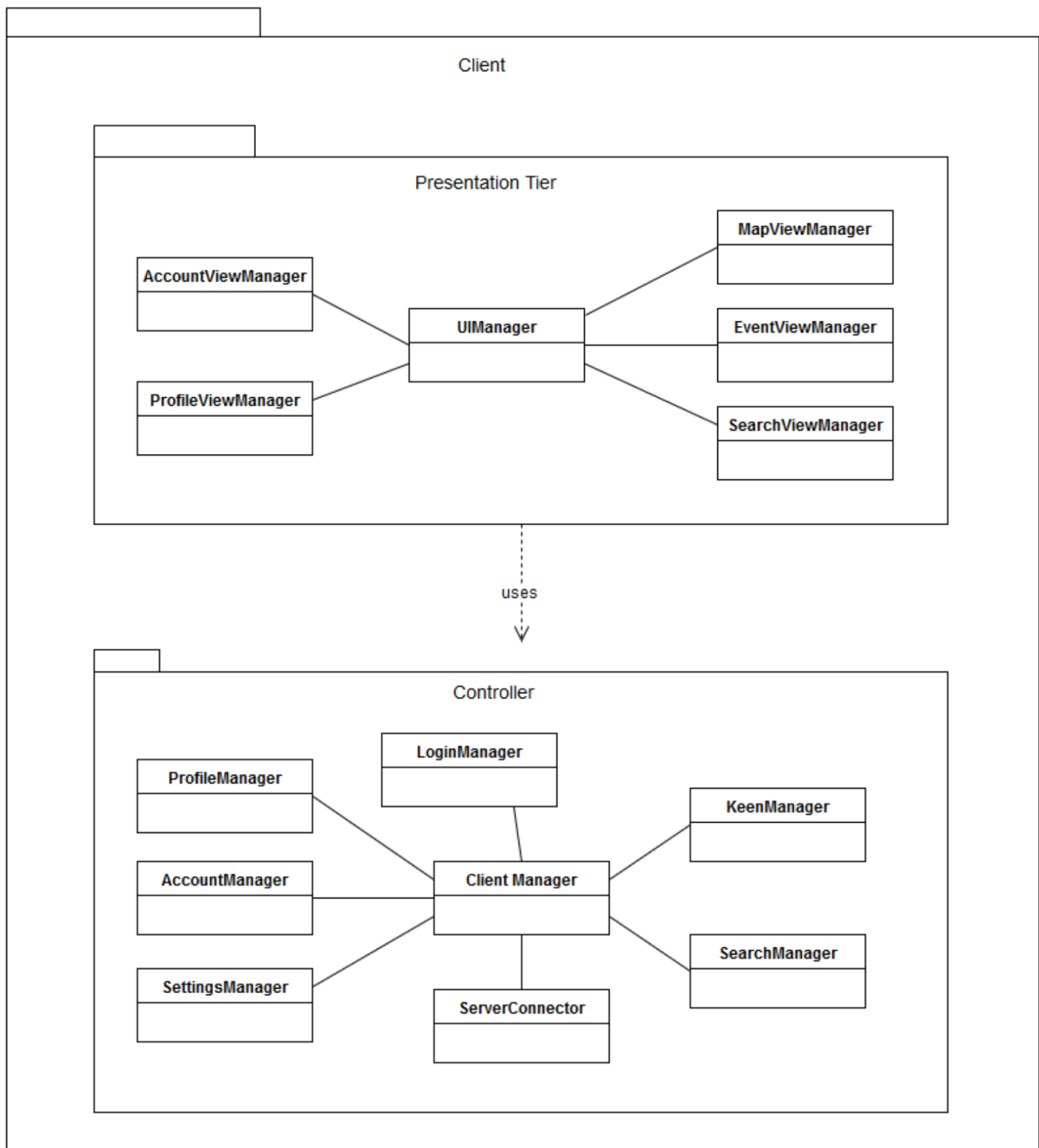

Figure 3.1: Subsystem Decomposition

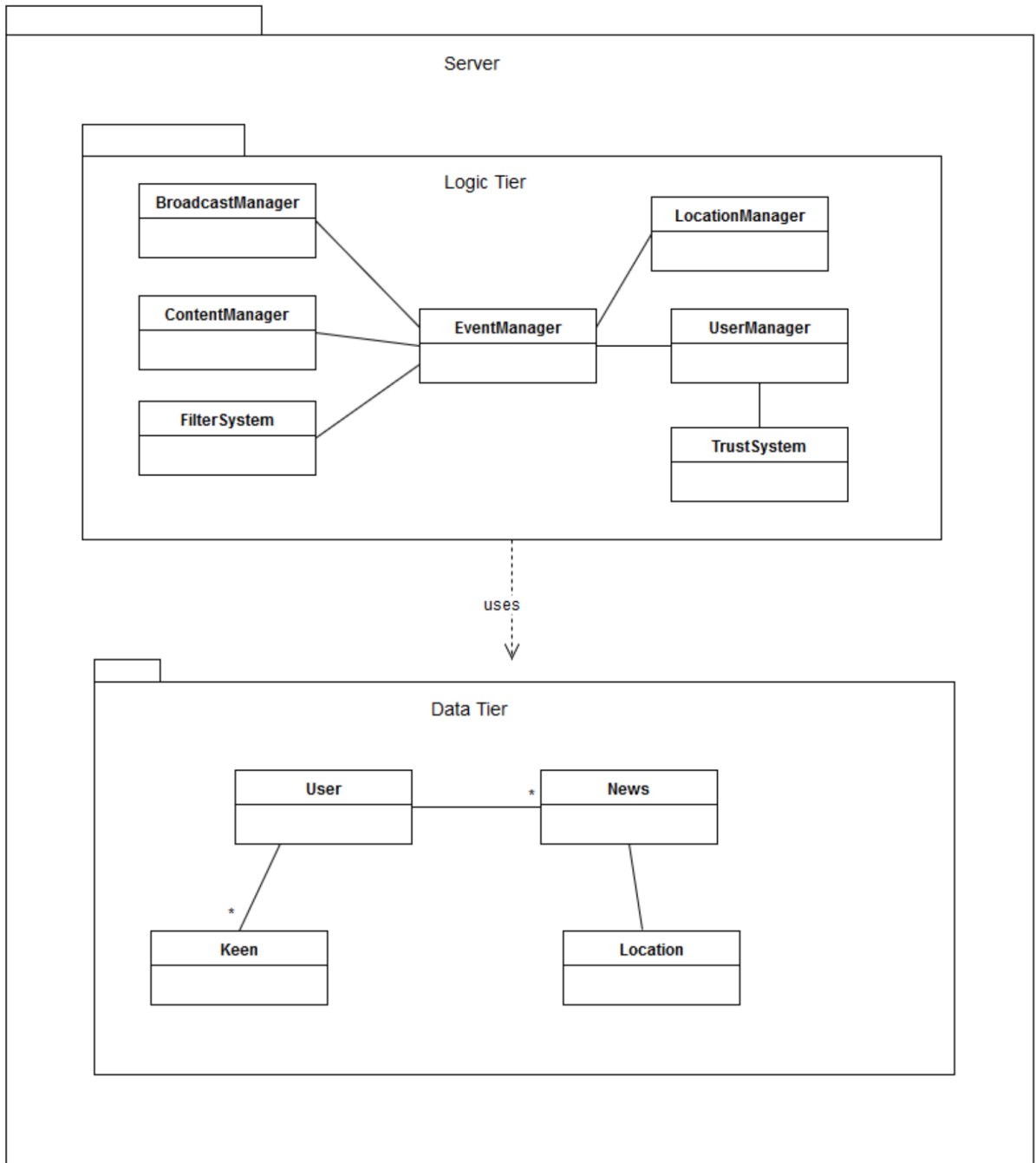Figure 3.2: Detailed View of Subsystem Decomposition: Client

Figure 3.3: Detailed View of Subsystem Decomposition: Server
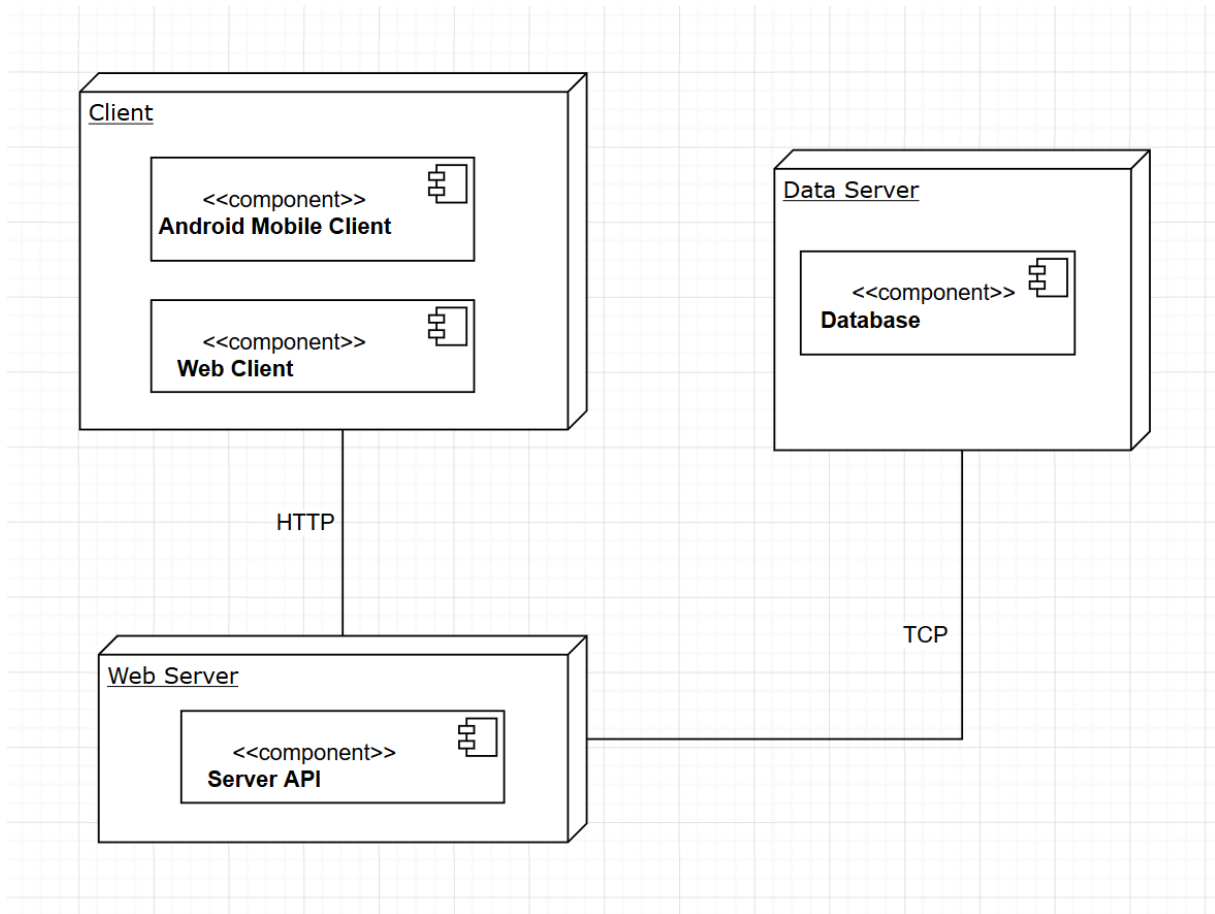
3.3 Hardware/Software Mapping

Figure 3.4: Hardware/Software Mapping

The client application can either be an Android phone, or a web browser. Due to nature of our application, not all of the functionalities will be available in the web client. Diafano will use the location services of each respective client application. The data related to users and news will be kept in the Data Server. As the application requires an active Internet connection, the data will be queried from the Data Server, so the user's device will not have to allocate much memory resource to accommodate the application.

The Web Server will handle requests from the users via HTTP. Similarly, based on the queries of the users and the inputs, Web Server will handle how these data are sent to the database. The Web Server is also responsible for filtering content.

3.4 Persistent Data Management

Considering specifications provided above, we need to store objects such as User, Map, Event and Content. To store this entities we consider three dimensions of data: velocity, variety and volume. In our scenario, velocity and volume are more important than variety. Since the scope of the project is focused on specific types of data, there is a little necessity to think about management of unstructured data.

As a part of the scope of our project we will need to store multimedia data, that requires special treatment. Therefore we will need to explore possibilities to benefit from video compression techniques. In this way we will be able to manage volume of the data. (Remark: Velocity of the data in the real time will be handled in the application layer using UDP protocol.)

In order to achieve ACID properties, we could use relational database. Usage of relational database will allow us to perform swift operations with previously developed queries. Since the scope does not enclose heavy analytical operations, we are not required to use different types of databases such as graph databases. Using relational database we plan to cover the minimum viable scope of the project.

3.5 Access Control and Security

Our system recognizes only one type of user, however, we do allow authentication as a verified source. The users are allowed to makes changes in their account, edit their profiles, view their past activity, and see all the streams or content they have shared in the application. Users can choose to share location information, or rather while permission is granted location of user/content can be shared, which will in turn be used to identify a specific type of content based on location where it was shared. In order to ensure security, our system requires a username/email and a password from all users to login to the system. The content of the profile of the user is shown to user only if the credentials are approved. A third-party password manager will be used to ensure the security of the passwords. Diafano does not allow any user who has not signed up to use the application. It is not possible to share streams or content

without user authentication. Therefore, Diafano allows access to users that are recognized in the database. Our system also ensures the personal data security; the information about the user is never shared unless the user opts in to.

3.6   Global Software Control

Our system has an event-driven software control. The user logins to his account by entering his credentials. When user requests to login, the server checks the database to make sure that the credentials are correct and gives permission. When the user wants to make changes in his account such as changing his/her username or password, or settings regarding the location sharing, the system updates the account of the user.

The user can view past shared content. In this case, the system presents them, along with details. Furthermore, if the user marks some other content as spam/abusive etc. , the system stored this information. This is to ensure the data remains intact, and some user account doesn't just blindly delete/flag other content just to have it removed from the system. There will be therefore, some implicit ranking performed to avoid such spam situations.

3.7   Boundary Conditions

3.7.1   Initialization

The user needs to have the application downloaded in a phone running Android operating system. To run the web application, the user needs to be using a browser like Google Chrome, Mozilla Firefox or Safari. In order to be able to interact with the application the user needs to create an account/register and log in with a username or password. A failure in the login process will lead to the login screen being displayed again. Additionally, the application retrieves real-time data to perform its functionalities thus it requires an internet connection.

3.7.2 Termination

The user can logout of the application by clicking a Logout button. In the same time if he/she clears the cache or the application data he/she will be automatically logged out. Thus, unless the user specifically decides to log out he/she will be signed in at all times even when the application is closed. If the application is not closed, it will run on background.

3.7.3 Failure

Failure of the application can occur when there is no Internet connection. There might be cached data in such cases, especially if the application was not closed on Android, so the information displayed may be out of date. Internet connection is necessary for updating the content.

4.   Subsystem Services

4.1   Client

The client corresponds to the mobile application mainly, because all of the functionalities will be implemented for mobile application. The client is the presentation layer of our system. Users can login to the system with the client and a request is made to server from the client for login. The user can specify the settings for their application from the client. Client is responsible for managing users operations on the system, presenting the data from the server to the user and also notifying the user when necessary. Client subsystem includes Presentation Tier and Control Tier. Presentation Tier is responsible for all of the user interface interactions and it uses Control Tier in order to communicate with the server.
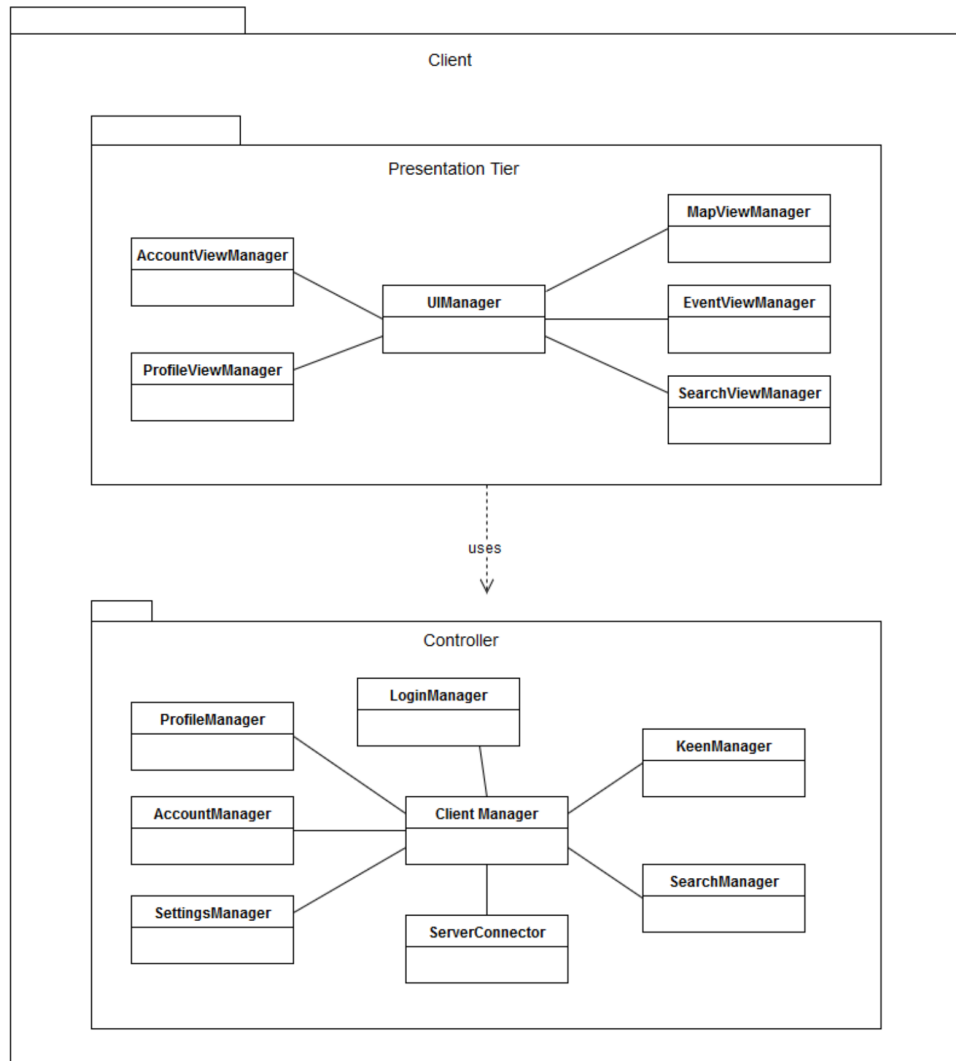
Figure 4.1: Detailed View of Client Subsystem
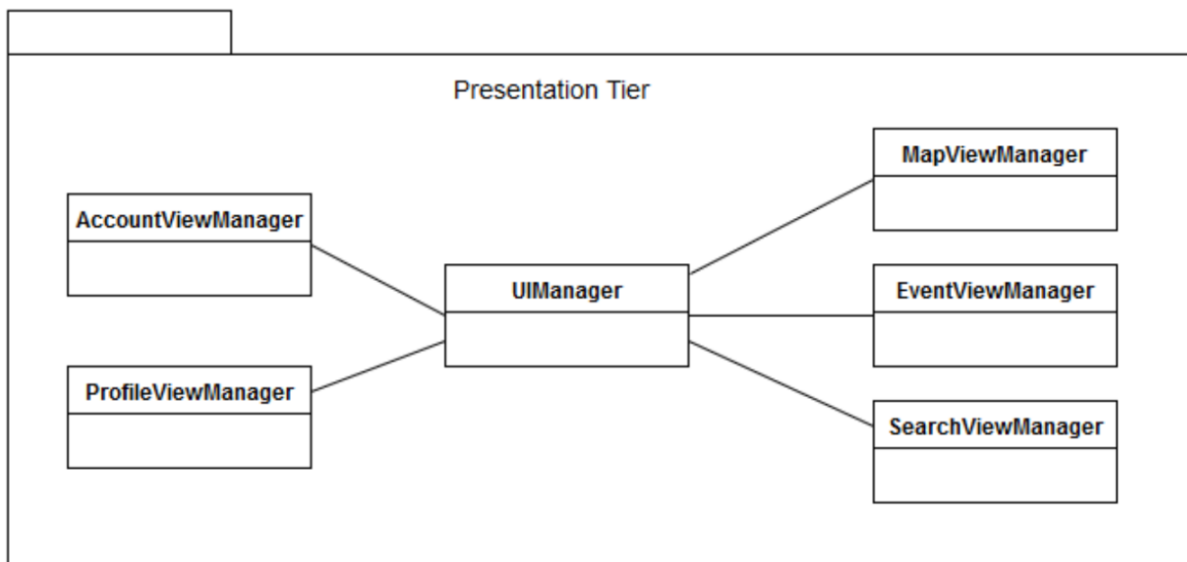
4.1.1 View Subsystem

Figure 4.2: View Subsystem in Server

View subsystem includes operations related to the user interface (UI).

- **MapViewManager:** This class handles all relations related to map presentation.

- **SearchViewManager:** This class handles search UI, including searchbar and results.

- **TextContentViewManager:** This class handles the UI for textual content based news.

- **MultimediaContentViewManager:** This class handles the UI for all news part that include multimedia, such as videos, saved broadcasts, and images.

- **AccountViewManager:** This class handles all operations related to UI for a user's account and settings.
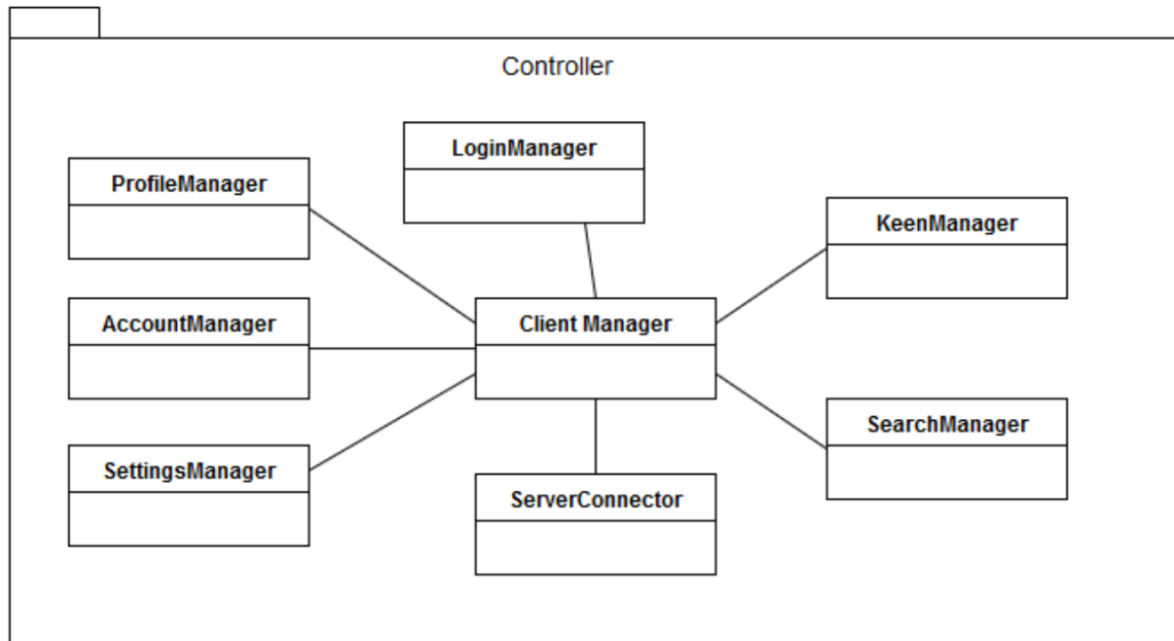
4.1.2 Controller Subsystem

Figure 4.3: Controller Subsystem in Server

Controller subsystem is responsible for managing the client operations and the communication between client and server.

- **LoginManager:** Class that handles login operations.

- **ServerConnector:** Class that handles the communication between the client and the server.

- **ProfileManager:** Class that updates the server with the changes made by the user.

- **AccountManager:** Class that updates the server about user's activity.

- **SettingsManager:** Class that updates the settings of the user according to their changes.

- **KeenManager:** Class that keeps the changes on the user's keens and notifies updates the server.

- **SearchManager:** Class that keeps the previous searches and directs the search to the server.

- **ClientManager:** Main class, global manager of operations of other managers.

4.2 Server

Server is an integral part of our system. The live video is recorded on the client side and delivered to the server. The server receives the data and binds it to event. The users requesting the video from an event will receive live video from the server. The server is responsible for consistent representation of the data and proper binding.

Furthermore server handles filtering operations. The filtering will allow user choose if they want to be exposed to a sensitive information. All the controllers are located on the server. It controls the persistency of the data and performs operation on top of it.

The server accomodates logic and data layers. In this way entities presented within this layer are allocated on the server. Logic tier is the backend application part responsible for handling requests and the data layer is the database responsible for permanent storage of entities in the applciation.
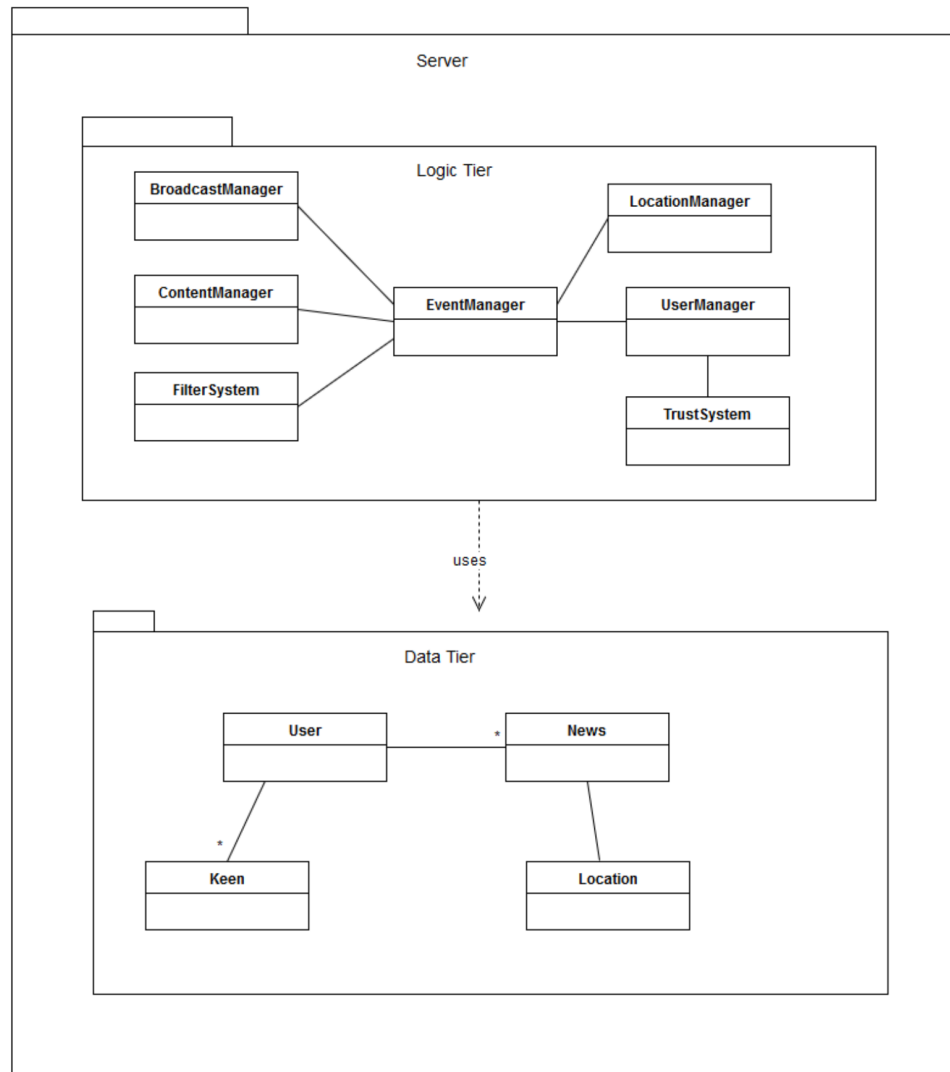
Figure 4.4: Detailed View of Server Subsystem

4.2.1 Logic Tier

The Logic tier is application layer that controls the flow of information between presentation layer and data layer. This tier accomodates all the heavy operations the application needs to handle.
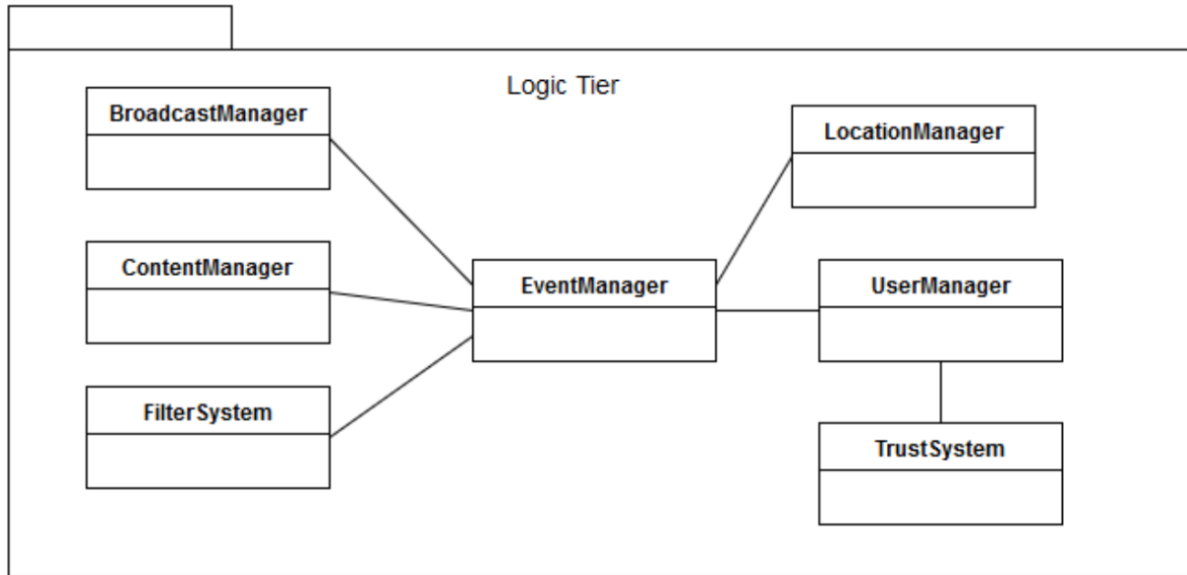
Figure 4.5: Logic Subsystem in Server

- **EventManager:** Fundamental class that will be responsible for events. It will gather all the information about the event.

- **BroadcastManager:** Class that is responsible for broadcasting. It will manage the interaction between the client that is broadcasting with the server as well as clients that will watch the broadcast through the server.

- **ContentManager:** This class is responsible for the news post.

- **FilterSystem:** This class is used to filter the events and their content post by the users.

- **LocationManager:** Class that is responsible for dealing with news and users' location data.

- **UserManager:** This class handles all the preferences, notifications about the user.

- **TrustSystem:** Class that is responsible for assigning a trust score to each user using the data of their news posts, votes, reports and more. This class is
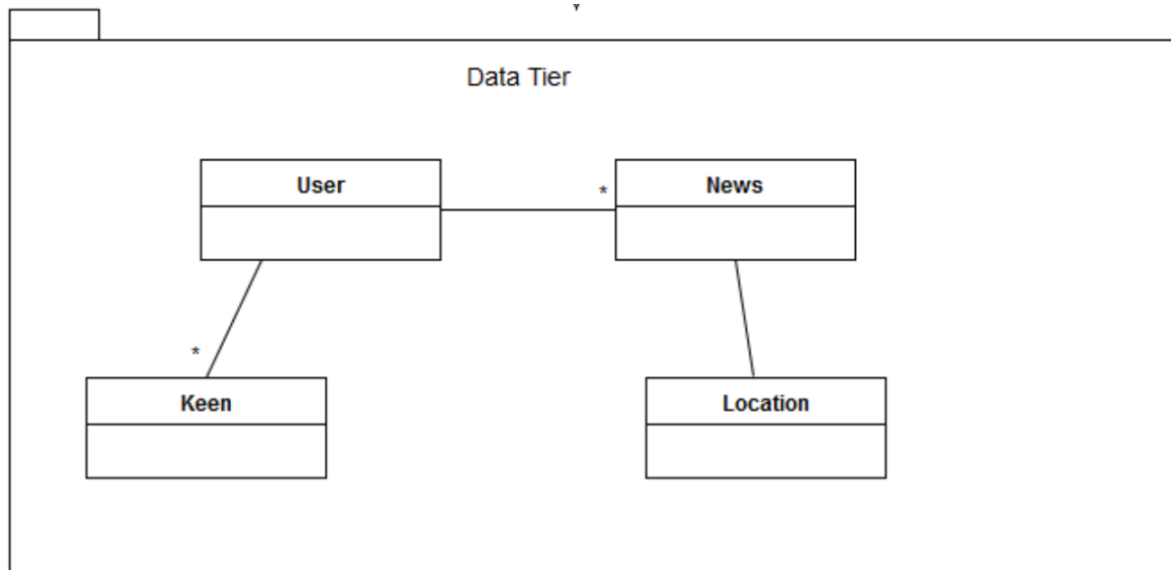
4.2.2 Data Tier

Figure 4.6: Data Subsystem in Server

Data Tier is the lowest layer that provides permanent storage of the data. Basic retrieval mechanisms such as search are implemented in this layer. This layer stores the data in a structured way and provides that data to the logic tier on demand.

- **User:** Data class representing users of Diafano. User class includes data such as user email, nickname, phone number, trustrank, and other information.

- **Keen:** Data class that keeps information about users following certain pieces of news for notifications.

- **Location:** Data class that keeps information on locations, such as their name, people close to that location, latest events that have happened there.

- **News:** Data class that keeps information about pieces of news, including date, source, location, content.

Bibliography

[1] Pew Research Center: How Americans Get Their News, 2016. `http://www.journalism.org/2016/07/07/pathways-to-news/`. Accessed: 2017-10-8.

[2] Beware Online Filter Bubbles, ted.com, 2011. `https://www.ted.com/talks/eli_pariser_beware_online_filter_bubbles`. Accessed: 2017-10-8.

[3] About Google News, google.com, 2017. `https://www.google.com/intl/en_us/about_google_news.html`. Accessed: 2017-10-31.

[4] Google News: the secret sauce — Technology, the guardian, 2017. `https://www.theguardian.com/technology/2013/feb/25/1`. Accessed: 2017-10-31.

[5] iOS - Feature Availability - Apple, apple.com, 2017. `https://www.apple.com/ios/feature-availability/#news`. Accessed: 2017-10-31.

[6] These Charts Show How Crucial Twitter Is for the Turkey Protesters , the atlantic, 2013. `https://www.theatlantic.com/international/archive/2013/06/these-charts-show-how-crucial-twitter-is-for-the-turkey-protester 276798/`. Accessed: 2017-10-31.

[7] New study quantifies use of social media in Arab Spring , the atlantic, 2011. `http://www.washington.edu/news/2011/09/12/new-study-quantifies-use-of-social-media-in-arab-spring/`. Accessed: 2017-10-31.

[8] CNN iReport , cnn. `http://edition.cnn.com/specials/opinions/cnnireport`. Accessed: 2017-11-04.

[9] Have Your Say , bbc news. `https://bbcnewsupload.streamuk.com/`. Accessed: 2017-11-04.

[10] Send 'World News' Your Story Idea , abc news. `http://abcnews.go.com/WN/mailform?id=10691207`. Accessed: 2017-11-04.