

PremNews Functional Specification

Table of contents

1. Introduction	1
1.1 Overview	1
1.2 Glossary	2
2. General Description	2
2.1 Product / System Functions	2
2.2 User Characteristics and Objectives	2
2.3 Operational Scenarios	2
2.4 Constraints	3
3. Functional Requirements	3
3.1 Scraping information and scores from internet sources	3
3.2 Inserting the scraped information into a database	3
3.3 Searching for news/scores	4
4. System Architecture	5
5. High-Level Design	5
6. Preliminary Schedule	6
7. Appendices	6

1. Introduction

1.1 Overview

The product to be developed is an android application that provides news and information about football matches in the English Premier league. The app will display basic information about live or past matches from the last number of days alongside news articles from a number of major news websites. The need for this application is that the majority of sport statistics apps use their own journalists or news sources, our app will give a variety of insights and opinions by providing the user with information from multiple news sources.

1.2 Glossary

Web Scraping: The process of collecting structured web data in an automated fashion. It's also called web data extraction [1]

BeautifulSoup: Library that aids in pulling particular content from a webpage, removing the HTML markup, and saving the information. It is a tool for web scraping that helps clean up and parse the documents that have been pulled down from the web. [2]

2. General Description

2.1 Product / System Functions

The end product of this system will be a mobile application that runs on the Android operating system that can display news and match information from various sources about football teams in the English Premier League. The application will get all this information from a database stored in the cloud. This database will be populated by a number of programs gathering the necessary information from the internet.

2.2 User Characteristics and Objectives

The user community for this application are members of the general population that use a device on the Android operating system as their primary mobile device and have an interest in the English Premier League. The application should be easy to use with a simple UI, which means that the end user must not have high levels of expertise to be able to use it. From the users perspective the app should allow for easy browsing and search of information. The user would also require the information to be regularly updated so that the user is never behind on the latest news. To avoid unnecessary presses by the user should be presented with the most recent information right after they start up the application.

2.3 Operational Scenarios

When the user first opens the app with no specific information that they wish to look for the user will be shown a screen with a list of headlines from various news articles, the news articles will be sorted in chronological order with the most recent news being displayed at the top. Each headline will be a hyperlink to a webpage containing the article. The user should be able to scroll through the list and access the articles with ease.

If a user would like to find specific news articles, the app will have an easily accessible search bar. The user will then be able to search for news articles related to a specific team or search by a keyword that is contained in the headline. The application will then filter the news articles and display the found articles in the same style that the home screen displays all the information.

The user will also have the option of viewing match scores and other basic information about live and past matches without having to view news articles. This will be accessed by the user from an options menu, live matches will be displayed at the top of the screen and past

matches will be displayed in chronological order similar to the news articles. For each match the home score, away score and the date/time of the match will be displayed. The user will be able to click on each match and be brought to a new screen displaying more detailed information and also as the main focus of this application is news, articles related to the match and both teams playing will also be displayed on this screen.

2.4 Constraints

- For the application to work the users device must be running a modern version of the Android operating system
- The users device must also be connected to the internet so that it can access the online database
- The back end program should be able to search the web efficiently for information to reduce the time it takes to update the information in the application

3. Functional Requirements

3.1 Scraping information and scores from internet sources

Description

The main function of the program is to scrape scores and information from sources on the internet and insert them into a cloud based database, to be pulled and posted into the app for users to access

Criticality

Implementing the scrapers, as well as processing the information properly is the most critical part of this application as without it there is no app

Technical Issues

A possible technical issue would be making sure the scrapers are constantly working and making sure that it is still taking the right information

Dependencies

As the main feature of the app is giving users an all in one place for news about football teams in the premier league, this requirement is essential to every other requirement

3.2 Inserting the scraped information into a database

Description

After the scrapers have found the correct information and scores from games and teams, the information is inserted into a cloud based database, for the app to use.

Criticality

This part is essential to the app's functionality as without implementing some kind of database it would be impossible for users to look at past games / news articles for teams.

Technical Issues

A potential technical issue that could arise is ensuring the database is storing the information correctly on the cloud and posting it properly to the app.

Dependencies

This feature is essential to how the app will display and post data to the app, and most of the app's functionality depends on the database being implemented properly

3.3 Searching for news/scores

Description

When a user opens the application, they will have the option to search for a specific team's scores and related news stories.

Criticality

This feature is important as there will be many stories and a lot of information present on the application and without a feature for users to search for and filter out irrelevant information the application will be less appealing

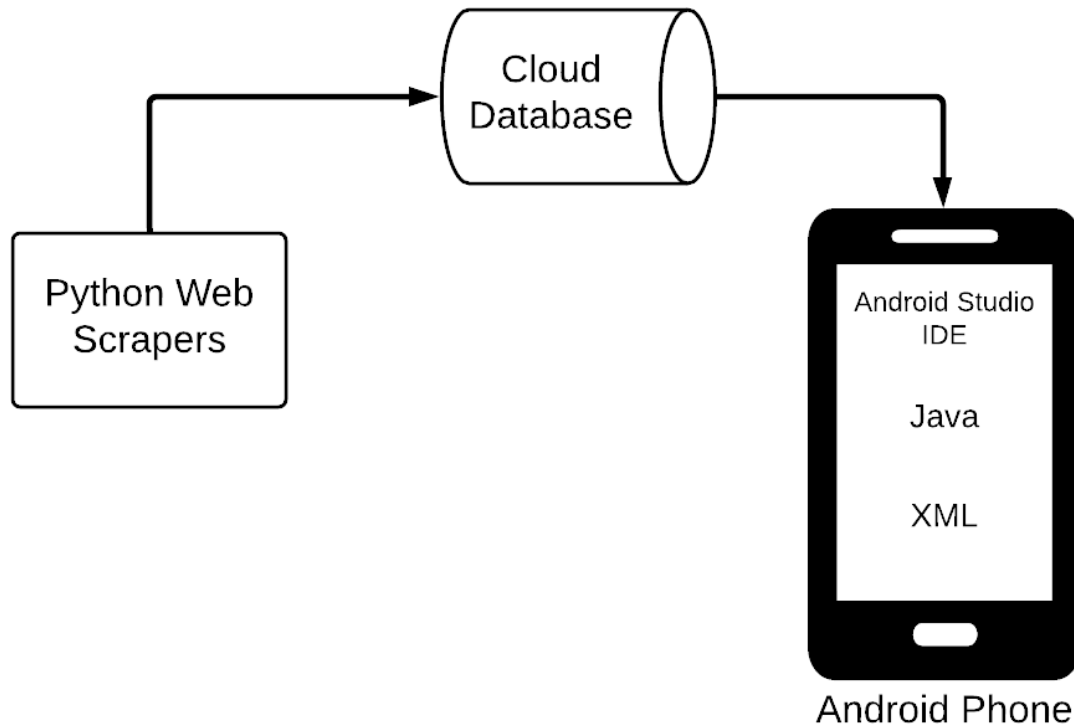
Dependencies

The search function is dependent on the database working correctly and the information being scraped and stored properly for users to search for stories and information about a particular team/ article

Technical Issues

A potential technical issue that could arise when developing this part of the application is that the searches may be too strict and not return the correct information or all information, so we would like to implement a feature that will bring up information of similar spelling to the user's input

4. System Architecture



The information that is to be displayed in the application will be gathered by a number of Python programs which are known as scrapers. The Python scrapers will search the HTML of specific web pages using the BeautifulSoup and Selenium libraries, filtering the necessary information. The gathered information will then be imported into a SQL database which for the moment will be stored and accessed in GCP (Google Cloud Platform).

The Android Application will be built using the Android Studio IDE and will be written in Java and XML. The application will access the database in the cloud and display the necessary data.

5. High-Level Design

Show information and results from database

When a user opens the application, they will be presented with some of the main stories of the day/week. This information will have been scraped from the internet, put in an online database and printed to the application

Search for results/information

If a user wants to find out about a particular game, they have the option to search for a particular keyword/s. The program will return information based on the search criteria

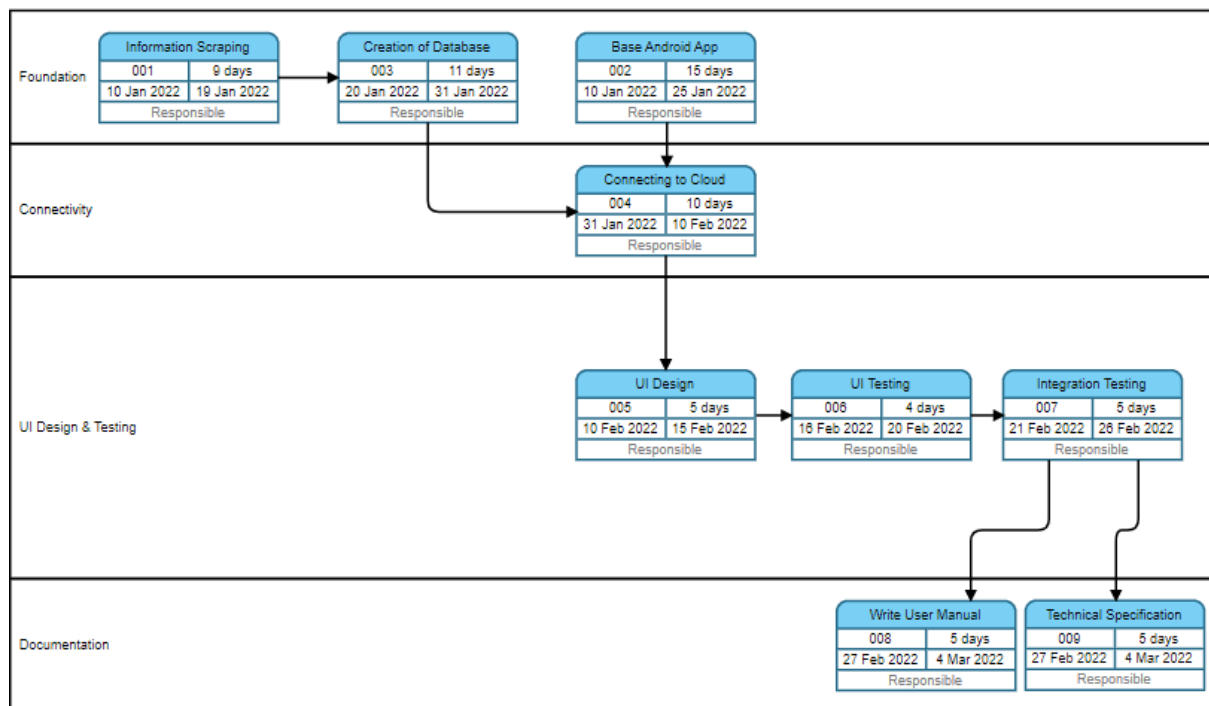
Open and read articles

When the user finds an article they would like to learn more about, they can then enter and read the full article in detail

Look at league table and see information about the league

The user can keep updated on the standings in the league as well as being kept up to date with the personal achievement leaders eg top goal scorer etc

6. Preliminary Schedule



7. Appendices

[1]C. Kenny, “What is web scraping and how does it work?,” *Zyte*, 03-Nov-2021.
[Online]. Available: <https://www.zyte.com/learn/what-is-web-scraping/>.

[2]J. Wieringa, “Intro to Beautiful Soup,” *Programming Historian*, 30-Dec-2012.
[Online]. Available:
<https://programminghistorian.org/en/lessons/retired/intro-to-beautiful-soup>.

