# **Software Requirements**

# **Specification**

# **for**

# **Web scraper for latest news**

**Version 1.0 approved**

**Prepared By :**

**Aastha Patel, Roll no: CE-001**

**Priyavardhan Panchal, Roll no: CE-086**

**Mauneel Sorathia , Roll no : CE-133**

# **Department of Computer Engineering**

# **B.Tech. CE Semester – VI**

**Guided by : Prof. Jigar M Pandya**

# **Table Of Content**

|  |  |  |
| --- | --- | --- |
| **Sr No.** | **Content** | **Page No.** |
| 1 | Introduction | 3 |
| 2 | Overall Description | 4 |
| 3 | External Interface Requirements | 5 |
| 4 | System Features | 6 |

# **Introduction**

## **1.1 Purpose**

A web scraper is a specialized tool designed to accurately and quickly extract data from a web page. Web scrapers vary widely in design and complexity, depending on the project. An important part of every scraper is the data locators (or selectors) that are used to find the data that you want to extract from the HTML file - usually xpath, css selectors, regex or a combination of them is applied. Our project is one of the applications of web scraping. We are using web scraping to extract the latest news articles from different websites and show them on our web application.

## **1.2 Intended Audience and Reading Suggestions**

This document is intended to both the stakeholders as well as the developers of the system.

## **1.3 Product Scope**

This system is intended to:

* Allow users to view news articles in chosen language by user
* Allow user to search news by keyword
* Show weather report
* Update news feed every 10 minutes
* Helps user to go through the news without any interrupts of advertisements

## **1.4 References**

The following material was used in creating this document:

IEEE Std 830-1998, IEEE Recommended Practice for Software Requirements Specifications.

# **Overall Description**

## **2.1 Product Perspective**

This web application is an efficient way to keep the users updated with the latest news from various news websites at one destination without any disturbances.

## **2.2 Product Functions**

* User can choose language of the news (By default, news are displayed in english)
* Users can sort the articles categorically
* User can filter the news according to the keywords
* User can go through the highlights of past 6/12 hours
* The news feed will be updated every 10 minutes
* System also shows weather report
* System will filter out the advertisements
* The web application is accessible to all

## **2.3 Design and Implementation Constraints**

* System is wirelessly networked with an encryption..
* Database is password protected.
* System should use less RAM and processing power.
* The system is accessible to all.

## **2.4 Assumptions and Dependencies**

* Internet connection is mandatory
* Browser in user’s system must support the frameworks

# **External Interface Requirements**

**3.1 User Interfaces**

* Users can filter the news based on keywords.
* Homepage to display the headlines and a short description for each
* A section for weather reports.
* Users can choose the language of the news.
* Users can choose the category of the news.

**3.2 Software Interfaces**

Any latest web browser which supports modern technologies such as HTML5,CSS3.

**3.3 Hardware Interfaces**

* Any device on which we can run the web browser.
* The device should have enough RAM(4 GB), capable processor, and Internet connection to run the website smoothly.

**3.4 Communication Interfaces**

* Data sender and receiver should be connected via LAN or WAN in order to communicate.
* Website server and User can be connected through WAN.

1. **System Features**

**Functional requirements**

**4.1 Manage News**

**4.1.1 Display News**

**Description** : The latest 10 news in english (by default) are displayed to the

user on the home page.

**Input** : No input

**Output** : Headlines with short descriptions and source of the news are

displayed.

**4.1.2 Update News**

**Description :** The news will be updated after every 10 min.

**Input :** User has to refresh the page.

**Output :** Latest news are added on the top.

**4.1.3 Choose category**

**Description :** The user can choose the category of news to be displayed.

**Input :** User Selection

**Output :** The news of that category is displayed on the home page.

**4.1.4 Show weather report**

**Description :** The weather report is displayed in a section of the home page.

**Input :** No input

**Output :** Maximum, Minimum temperature along with type of weather of the day are displayed.

**4.1.5 Change language**

**Description :** The user can change the language of the news to be displayed.

**Input :** User selection

**Output :** News in that particular language is displayed.

**4.1.6 Filter News based on time**

**Description :** The user can filter out the news according to their convenience

based on time.

**Input :** User can choose either 6 hours or 12hours

**Output :** News of past that particular time is displayed.

**4.1.7 Search News**

**Description :** The user can filter news based on a keyword.

**Input :** Keyword

**Output :** News related to that keyword is displayed.

**4.1.8 Filter out advertisements**

**Description :** The news is shown without any advertisements.

**Input :** No input.

**Output :** News is displayed free from advertisements.