# SCUFFY WEB SCRAPER

**A PROJECT REPORT**

***Submitted by***


**PARIKH HETVI SUNILBHAI**

**(190130107074)**


***In fulfillment for the award of the degree of***

**BACHELOR OF ENGINEERING**

***in***

**Computer Engineering**


**Government Engineering College**

**Gandhinagar**




**Gujarat Technological University, Ahmedabad**

**April 2023**



# GOVERNMENT ENGINEERING COLLEGE

Gandhinagar - Vijapur Rd, Sector 28 GIDC, Sector 28, Gandhinagar, Gujarat

# CERTIFICATE

This is to certify that the project report submitted along with the project entitled **SCUFFY WEB SCRAPER** has been carried out by **PARIKH HETVI SUNILBHAI** under my guidance in fulfillment for the degree of Bachelor of Engineering in COMPUTER ENGINEERING,8th Semester of Gujarat Technological University, Ahmedabad during the academic year 2022-23.

Prof. Hemani Shah

Internal Guide

D. A. Parikh

Head of the Department

# Company Certificate

Achyut labs
Unshakable Values

717 - Times Square Arcade,
Thaltej - Shilaj Rd, opp. Rambaug,
nr. Ravija Plaza, Thaltej,
Ahmedabad, Gujarat 380059

**Date:** 21.04.2023

**TO WHOM IT MAY CONCERN**

This is to certify that Parikh Hetvi Sunilbhai, a student of Government Engineering College Gandhinagar, has successfully completed her internship in the field of Web Development(Vue Js) as a frontend intern from 1st February 2022 to 28th April 2022 (Total number of Weeks: 12) under the guidance of Achyut labs Team.

Her internship activities include Web Development Activities.

We wish her every success in her life and career.

Yours Faithfully,
**For, Achyut Labs Private Limited**

**Sandipkumar Patel**
**Director**



## GOVERNMENT ENGINEERING COLLEGE

**Gandhinagar - Vijapur Rd, Sector 28 GIDC, Sector 28, Gandhinagar, Gujarat**

# DECLARATION

We hereby declare that the Internship report submitted along with the Internship entitled **SCUFFY WEB SCRAPER** submitted in fulfillment for the degree of Bachelor of Engineering in Computer Engineering to Gujarat Technological University, Ahmedabad, is a bonafide record of original project work carried out by me at Achyut Labs Pvt. Ltd. under the supervision of Prof.Hemani Shah and that no part of this report has been directly copied from any students' reports or taken from any other source, without providing due reference.

| | Name of the Student | Sign of Student |
|---|---|---|
| 1 | Parikh Hetvi Sunilbhai | _____ |

# Acknowledgement

I would like to extend my deepest gratitude to ACHYUT LABS PVT. LTD. Consultancy and Mr.Sandip patel, who gave me the golden opportunity to do this wonderful internship in Web development.

I express my deep & sincere gratitude to my mentor Mr. Jayesh Patel, Tech Lead and Business Analyst, Guidance, supervision, expert suggestion & encouragement which helped me to tide over the hardship encountered during study.

Last but not the least my sincere gratitude to all people who knowingly or unknowingly supported me, for my morals to make this project a reality.

I extend warm thanks to Dr. Dhaval Parikh, H.O.D., Dept. of Computer Engineering. GEC, Gandhinagar for his constant encouragement, motivation and guidance.

I would like to thank my internal guide **Prof. Hemani Shah** for giving me constant guidance during internship and helping me a lot in finalizing this internship within the limited time frame.

**Place**: Ahmedabad **Parikh Hetvi**

**Date**: 190130107074

# Abstract

Scuffy Web Scraper is a website that provides metadata for websites to help them rank higher in search engine results. The website is built using Vue.js and Python Django, and allows users to enter a URL and obtain key metadata such as the website title, description, and keywords. This information can be used to optimize a website for search engine optimization (SEO) purposes, and ultimately increase its visibility and ranking on popular search engines such as Google. The Scuffy Web Scraper is designed to be user-friendly and efficient, allowing users to quickly obtain the information they need without any technical knowledge. The project is an innovative solution for website owners and digital marketers looking to improve their SEO efforts, and can be easily integrated into existing workflows. Overall, the Scuffy Web Scraper is an effective tool for optimizing websites and improving their search engine visibility.

# List of Figures

# List of Tables

# Table of Contents

# Chapter 1
# Overview of the Company

## 1.1 History

Achyut Labs is a private limited company based in Ahmedabad, Gujarat, India. It was founded in 2015 and provides software development and IT consulting services to businesses across various industries. The company specializes in web and mobile application development, cloud computing, blockchain, and artificial intelligence.

Achyut Labs aims to create innovative and high-quality software solutions that help their clients achieve their business objectives. The company's team consists of experienced developers, designers, and project managers who work collaboratively to deliver customized solutions that meet their clients' needs.

## 1.2 Scope of Work

- **Web and Mobile App Development**: Achyut Labs specializes in building custom web and mobile applications for businesses across various industries. This includes developing software from scratch or improving existing applications.
- **Cloud Computing:** The company also provides cloud computing services, such as AWS and Google Cloud Platform, to help businesses reduce their IT infrastructure costs and improve scalability.
- **BlockChain Development**: Achyut Labs has expertise in developing blockchain-based solutions for businesses, such as creating smart contracts, decentralized applications, and digital wallets.
- **IT Consulting:** Achyut Labs provides IT consulting services to help businesses make strategic decisions about their IT infrastructure,processes and applications.

- **Front End Development**: We build designs that support smooth navigation and are aesthetically beautiful using vue js.
- **API Development**: Make smoother and hassle-free shifts in your
- **Graphic Design**: Let the enticing visuals create a deep impact and speak for your brand.
- **Product Support & Maintenance:** With our product support and maintenance service, foster innovation, client satisfaction, and cut down later stage maintenance cost.

# Chapter 2

# Overview of

# Different Unit

## 2.1 Work Being Carried out in each department

**Design:** The design department is responsible for creating the user interface (UI) and user experience (UX) design of the web scraper. This involves creating wireframes, designing the layout and look and feel of the application, and ensuring that the design is consistent with the brand identity of the company.

**Backend Development:** The backend development team is responsible for creating the server-side of the web scraper using backend technologies such as Python Django. This involves implementing the business logic, handling requests and responses, and storing and retrieving data from databases.

**Frontend Development:**The frontend development team is responsible for implementing the UI and UX design in the web scraper using frontend technologies such as Vue.js. This involves creating responsive web pages, integrating with APIs, and ensuring that the frontend is user-friendly and easy to navigate.

**Testing:** The testing and quality assurance team is responsible for ensuring that the web scraper works as intended and is free of bugs and errors. This involves performing functional testing, user acceptance testing, and performance testing, and ensuring that the application meets the quality standards of the company and the client.

## 2.2 List of used Equipment

Table 2.1 List of Used Equipment

| | |
|---|---|
| System | Windows 11 |
| IDE | Visual Studio |
| Device | NA |
| System Specification | 16 GB RAM, 256 GB SSD , Intel(R) |

## 2.3 Schematic layout

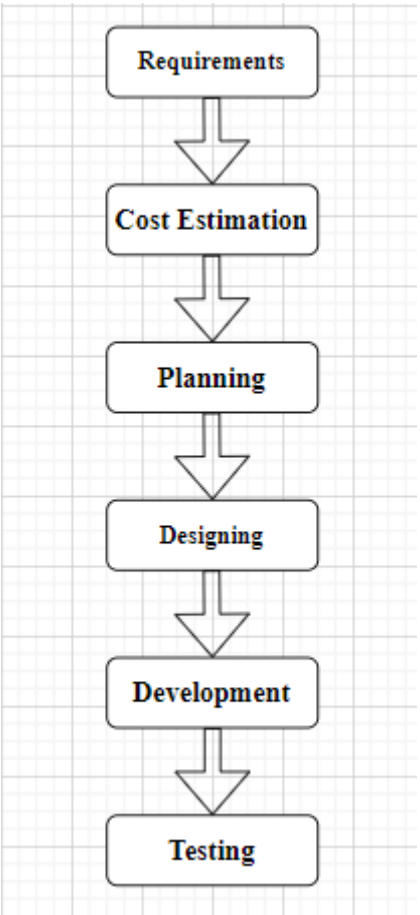



Fig 2.1 Schematic Layout

## 2.4 Explanation of each production stage

**Requirements:** In this stage first understand the requirements of the project. Discuss the workflow of the project. What the user will experience during its work. Mostly during this stage, we need to understand exactly how different types of users will use the application.

**Cost of Estimation:** This is one of the important stages of the project structure. Cost of estimation will work on how much time required during the project. It requires looking at the tasks, duration, and resources required to forecast a project's total cost to deliver

**Planning:** In this stage the team creates technical designs, task lists, resources plan, budget etc. how much cost will require , how much time will be required to complete the project, how much resources required to complete the project all the things are discussed during this stage.

**Designing:** The design department is responsible for creating the user interface (UI) and user experience (UX) design of the web scraper. This involves creating wireframes, designing the layout and look and feel of the application, and ensuring that the design is consistent with the brand identity of the company.

**Development:** The frontend development team is responsible for implementing the UI and UX design in the web scraper using frontend technologies such as Vue.js. This involves creating responsive web pages, integrating with APIs, and ensuring that the frontend is user-friendly and easy to navigate.

The backend development team is responsible for creating the server-side of the web scraper using backend technologies such as Python Django. This involves implementing the business logic, handling requests and responses, and storing and retrieving data from databases.

**Testing:** Testing process is one of the most important processes of the production stage. Testing is useful for finding bugs and errors in any project. There is manual testing and auto testing. Testing is important for good user experience. And it measures the capability of the project.

# Chapter 3
# Introduction of
# Project

## 3.1 Project Summary

Scuffy Web Scraper is a web scraping tool that extracts data from various websites and presents it in an easy-to-use format. The project was developed using Vue.js and Python Django technologies.

Scuffy Web Scraper has several benefits. It allows users to easily collect and analyze data from different websites without manual intervention, saving time and effort. It also provides an efficient and accurate method for collecting data, reducing errors and inconsistencies. Additionally, the project is highly customizable, allowing users to tailor the tool to their specific data extraction needs.

## 3.2 Project Purpose

The project purpose is to make a website that allows users to input a URL and retrieve the meta-data of the webpage, including keywords, to understand how it ranks highly in search engine optimization (SEO). The website will use web scraping techniques to extract the meta-data, and will display the relevant information in a user-friendly format. The tool will be useful for web developers and marketers to optimize their own websites for better search engine ranking.

## 3.3 Project Objective

The main objective of Scuffy Web Scraper is to automate the process of data extraction from websites. It does this by scraping the HTML code of the targeted websites and then filtering and organizing the extracted data based on the user's needs.

The project has multiple modules, each with its own functionality. The Front-end module provides a user-friendly interface for entering the website URLs and data fields to be extracted. The Backend module handles the web scraping process by utilizing various Python libraries and packages. The Development module is responsible for developing

new features and updating the existing ones. The Testing module ensures that the web scraper works properly and meets the user requirements.

## 3.4 Project Scope

The scope of this project involves building a website that allows users to input a URL and retrieve the meta-data of the webpage, with a focus on identifying the keywords that are used to rank highly in search engine optimization (SEO). The website is built using Django as a backend framework, and web scraping techniques with Beautiful Soup are used to extract the meta-data from the URL. The website utilizes Django's built-in views and templates to display the extracted meta-data in a user-friendly format. The website is designed to be responsive and easy to navigate, with a simple and intuitive interface for entering URLs and viewing the extracted data. The project also includes testing and validating the website to ensure accuracy and functionality, providing documentation and support for users and developers, and contributing to the field of web development and digital marketing by providing a valuable tool for SEO optimization.

## 3.5 Technology Overview

Table 3.1 Technology Used

| Category | Technology Used |
|---|---|
| Programming Language | Vue Js |
| IDE | Visual Studio |
| Dependencies | Axios,Nuxt,Vueityfy,webpack,jspdf,core-js |
| Database | MySQL |

### 3.5.1 HTML/CSS/JAVASCRIPT

HTML stands for HyperText mark-up Language.HTML describes the structure of web pages using mark-up. HTML elements are the building blocks of HTML pages. HTML elements are represented by tags.

CSS stands for Cascading Style Sheets. CSS describes how HTML elements are to be displayed on screen, paper, or in other media. CSS saves a lot of work. It can control the layout of multiple web pages all at once. External style sheets are stored in CSS files.

Introduction JavaScript is a dynamic computer programming language. It is lightweight and most commonly used as a part of web pages, whose implementations allow client-side scripts to interact with the user and make dynamic pages. It is an interpreted programming language with object-oriented capabilities.

### 3.5.2 Git

Git is a free and open source distributed version control system designed to handle everything from small to very large projects with speed and efficiency.Git is easy to learn and has a tiny footprint with lightning fast performance. It outclasses SCM tools like Subversion, CVS, Perforce, and ClearCase with features like cheap local branching, convenient staging areas, and multiple workflows.

### 3.5.3 GitHub

GitHub is a provider of Internet hosting for software development and version control using Git. It offers the distributed version control and source code management (SCM) functionality of Git, plus its own features. It provides access control and several collaboration features such as bug tracking, feature requests, task management, continuous integration and wikis for every project.

### 3.5.4 VS-Code

Visual Studio Code, also commonly referred to as VS Code, is a source-code editor made by Microsoft for Windows, Linux and macOS. Features include support for debugging, syntax highlighting, intelligent code completion, snippets, code refactoring, and embedded Git. Users can change the theme, keyboard shortcuts, preferences, and install extensions that add additional functionality.

### 3.5.5 Vue Js

Vue.js is an open-source JavaScript framework used for building user interfaces and single-page applications. It was created by Evan You in 2014 and has gained popularity due to its simplicity, flexibility, and ease of integration with other libraries and existing projects.

Vue.js uses a reactive data-binding system, allowing developers to create dynamic and interactive web applications. It offers a component-based architecture, allowing developers to build complex applications by breaking them down into smaller, reusable components. Vue.js also provides a set of tools for state management, routing, and form validation.

### 3.5.6 Nuxt

Nuxt.js is a framework for building server-side rendered (SSR) Vue.js applications. It builds upon the Vue.js framework by adding features such as server-side rendering, automatic code-splitting, and static file generation.

### 3.5.7 MYSQL

MySQL is a popular choice for web developers due to its ease of use, scalability, and robustness. It uses a client-server architecture, where the database server processes requests from multiple clients, making it suitable for high-traffic web applications.

## 3.6 Project Internship Planning

In Managing any project the whole plan of the project is made before its actual implementation.The plan of the project helps th team t work as per the schedule and helps to successfully complete the project.To plan a project the main requirements that are calculated are cost, duration, effort,scheduling, manpower, resource allocation, risk management etc.We plan on taking a time-based approach to deliver this product on time. To begin with, we create a set of deadlines to complete certain tasks. This will give us a general idea on how to meet the final deadline on time.

### 3.6.1 Project / Internship Development Approach and Justification

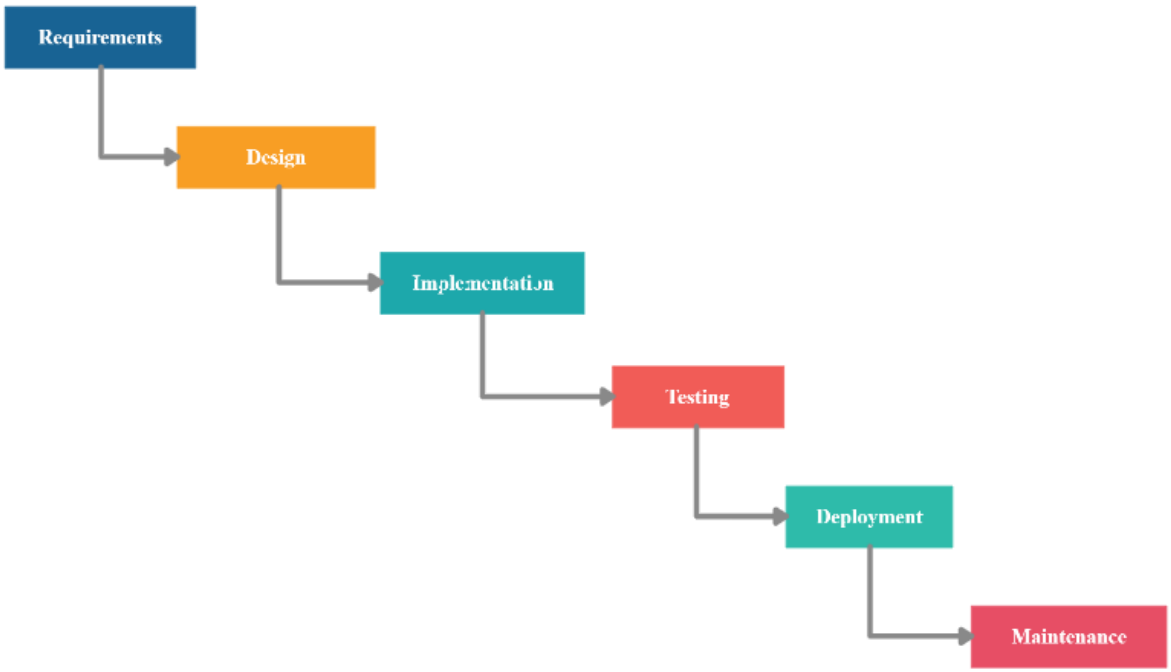



**Fig. 3.1 Iterative Waterfall Model**

### 3.6.2 Project / Internship Effort and Time, Cost Estimation

The entire internship was 3 months long. The internship was milestone based outlined mention in Table 3.2

Table 3.2 Internship timeline & progress

| Months | Internship stage |
|---|---|
| February | Leaning All Basics & Advanced concepts of frontend technology |
| March | Complete all assignment given by company for ability testing |
| April | Work on product issues |

Table 3.3 Roles and Responsibilities Table

| NAME | Role | | | | |
|---|---|---|---|---|---|
| | **Analysis** | **Designing** | **Coding** | **Testing** | **Documentation** |
| **Parikh Hetvi** | ✔ | ✔ | ✔ | ✔ | ✔ |

## 3.7 PROJECT / INTERNSHIP SCHEDULING (Gantt Chart / PERT /Network Chart)

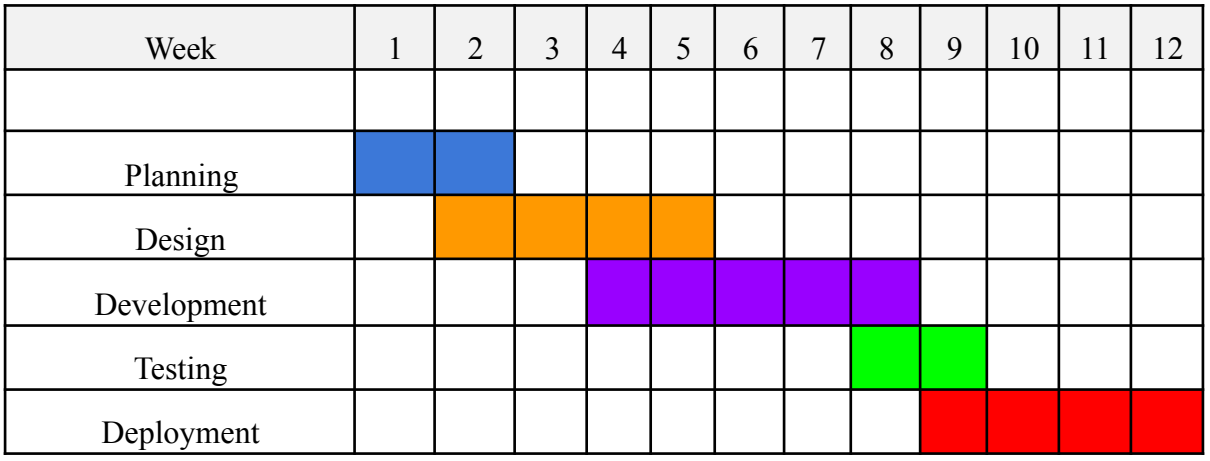



**Fig 3.2 Gantt Chart**

# Chapter 4

# System

# Analysis

## 4.1 Study of Current System

Web scraping is a popular technique used to extract data from websites automatically. It involves using specialized tools and software to extract data quickly and efficiently. Web scraping is used for a variety of purposes, such as data analysis, market research, and price monitoring. The data extracted can provide insights into customer behavior, trends in the market, and inform business decisions.

## 4.2 Problem of Current System

Web scraping can also result in inaccurate or incomplete data, as the data may be outdated, incomplete, or inconsistent. This can lead to errors and incorrect conclusions in data analysis web scrapers require maintenance and upkeep, as websites can change their design and data format over time. This can make it difficult to maintain and update web scrapers, leading to errors and inconsistencies in data extraction.

## 4.3 Requirements of New System

Scuffy provides robust data management features, allowing users to clean, transform, and store data extracted from websites. It also supports a range of data storage options, including databases, CSV files, and JSON files. This allows users to manage their data effectively and integrate it with external analysis tools for further insights.In addition, Scuffy has a user-friendly interface that makes it easy to use for non-technical users. Its user interface is designed to be intuitive, with clear documentation and customer support available to help users troubleshoot any issues. This allows users to focus on their core tasks of data extraction and analysis, without having to worry about the technical details of the scraping process.

## 4.4 System feasibility

- **Technical Feasibility**: Scrapy is a widely used and popular open-source web scraping framework that is supported by a large community of developers. It is built using Python, a widely used and supported programming language. Scrapy provides a range of features and capabilities that enable efficient and effective web scraping, including support for XPath and CSS selectors, built-in crawling and scraping algorithms, and support for distributed crawling. These features make Scrapy a technically feasible option for web scraping tasks.
- **Economic Feasibility**: Scrapy is a cost-effective option for web scraping tasks, as it is an open-source framework that can be downloaded and used for free. This makes it an affordable option for organizations of all sizes. In addition, Scrapy provides a range of features that enable efficient and effective web scraping, which can help reduce the time and resources required for data extraction and analysis. This makes Scrapy an economically feasible option for web scraping tasks.
- **Operational Feasibility:** Scrapy has a user-friendly interface that makes it easy to use for non-technical users. Its user interface is designed to be intuitive, with clear documentation and customer support available to help users troubleshoot any issues. This makes it an operationally feasible option for web scraping tasks, as it can be easily integrated into an organization's existing workflows and processes.

## 4.5 Activity in New System

The new system for web scraping should incorporate a range of activities to ensure efficient and effective data extraction and analysis. One of the key activities that should be included in the new system is data collection. This involves identifying and selecting the websites or sources from which data needs to be extracted. The new system should provide a user-friendly interface that allows users to input the relevant information, such as the URL of the target website, the data to be extracted, and any relevant parameters or filters. The system should also allow users to preview the data before extraction to ensure accuracy and completeness.

## 4.6 Features of New System

A variety of features that make it simple and effective to gather and analyze data from websites should be included in Scuffy Web Scraper. The new system ought to include an intuitive user interface that enables users to enter the necessary data for data collecting. The URL of the intended website, the information to be extracted, and any pertinent parameters or filters are all included. The system should also offer a variety of data collecting methods, including built-in crawling and scraping algorithms, support for XPath and CSS selectors, and support for distributed crawling. Scuffy Web Scraper should also have tools for data cleansing and transformation, such as the capacity to eliminate HTML tags and format data into CSV, JSON, or other common forms. This guarantees that the data is precise, comprehensive.

## 4.7 List of Modules

Scuffy Web Scraper may have a range of modules that provide specific functionalities for data collection, cleaning, transformation, management, and analysis. Some potential modules include a data collection module that provides a user-friendly interface for selecting websites and specifying data to be extracted. Another potential module is a scraping module that includes a range of techniques for extracting data from different types of websites. The system could also have a data cleaning and transformation module that includes features for removing HTML tags, filtering out irrelevant data, and formatting data into standardized formats. A data management module could also be included that provides options for storing and organizing the collected data, such as support for databases and file formats.

## 4.8 List of Software

Table 4.1 List of Software

| Category | Software |
|---|---|
| Programming Language(s) | Vue JS |
| Integrated Development Environment | Visual Studio |
| Dependencies | |
| Databases | MySQL |
| Operating System | Windows 11 |

# Chapter-5
# System
# Design

## 5.1 System Design and Methodology

The design of the Scuffy Web Scraper system should be based on sound software engineering principles to ensure that it is robust, efficient, and maintainable. This includes using a modular architecture that allows for flexibility and scalability as the system grows and evolves over time. The system should also be designed with the user in mind, ensuring that the interface is intuitive and easy to use. The methodology used to develop the system should be based on an agile approach that emphasizes iterative development, frequent testing, and continuous feedback from users. This ensures that the system is constantly being refined and improved to meet the evolving needs of users. Additionally, the system should be designed with security and privacy in mind, ensuring that any data collected is protected and that user data is not at risk.

## 5.2 Input/Output and Interface Design

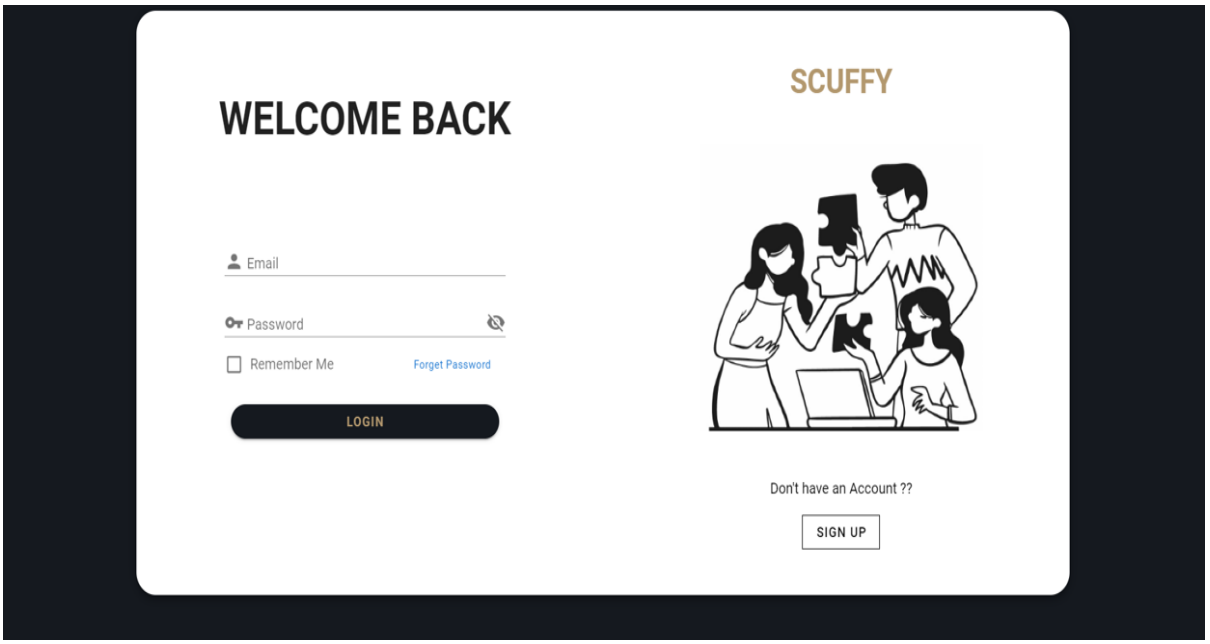



Fig 5.1 Login-Registration Screen

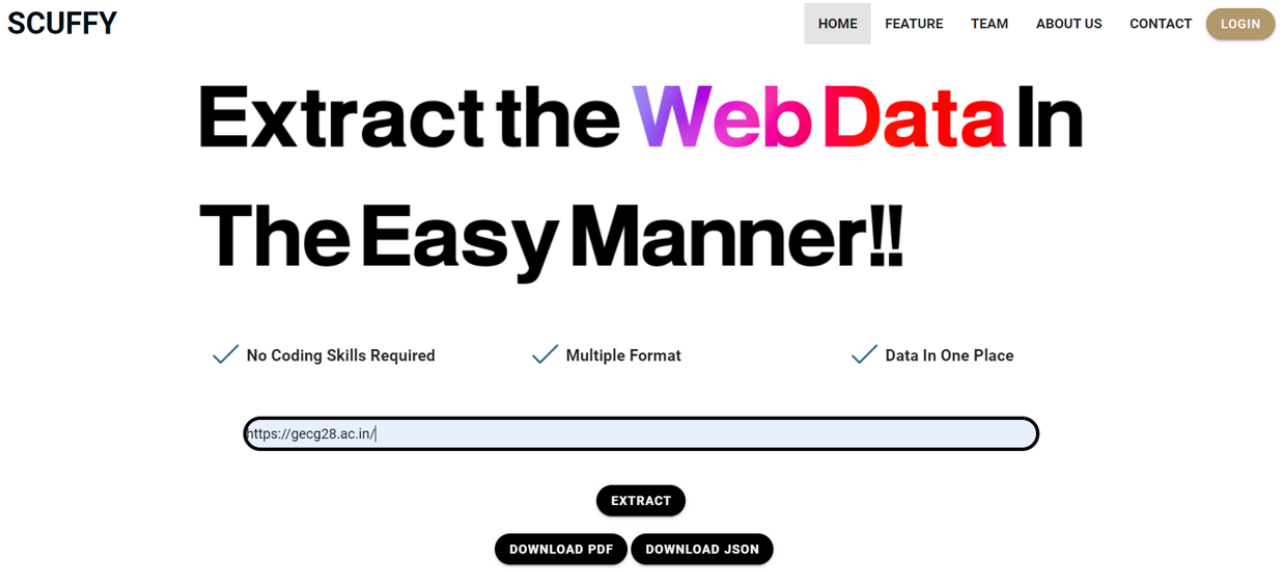



Fig 5.2 Home Page

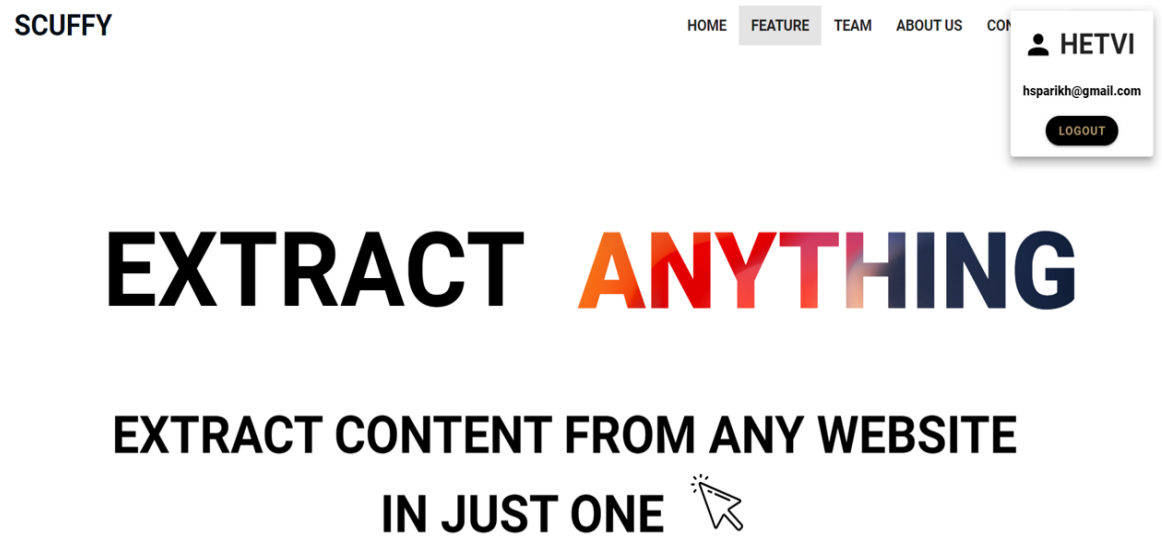



Fig 5.3 Feature page after Login

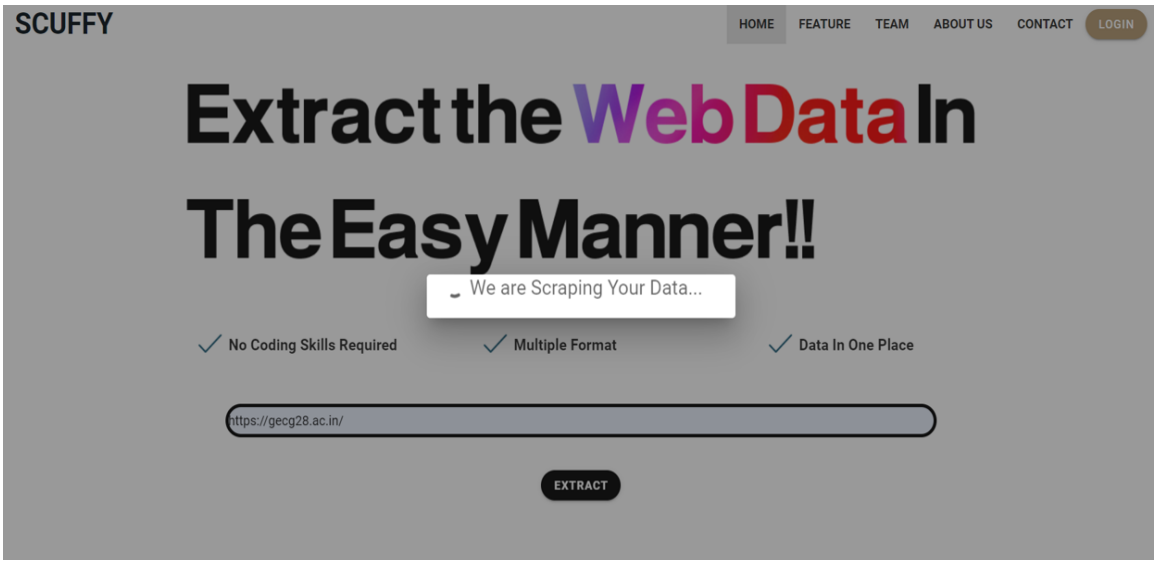
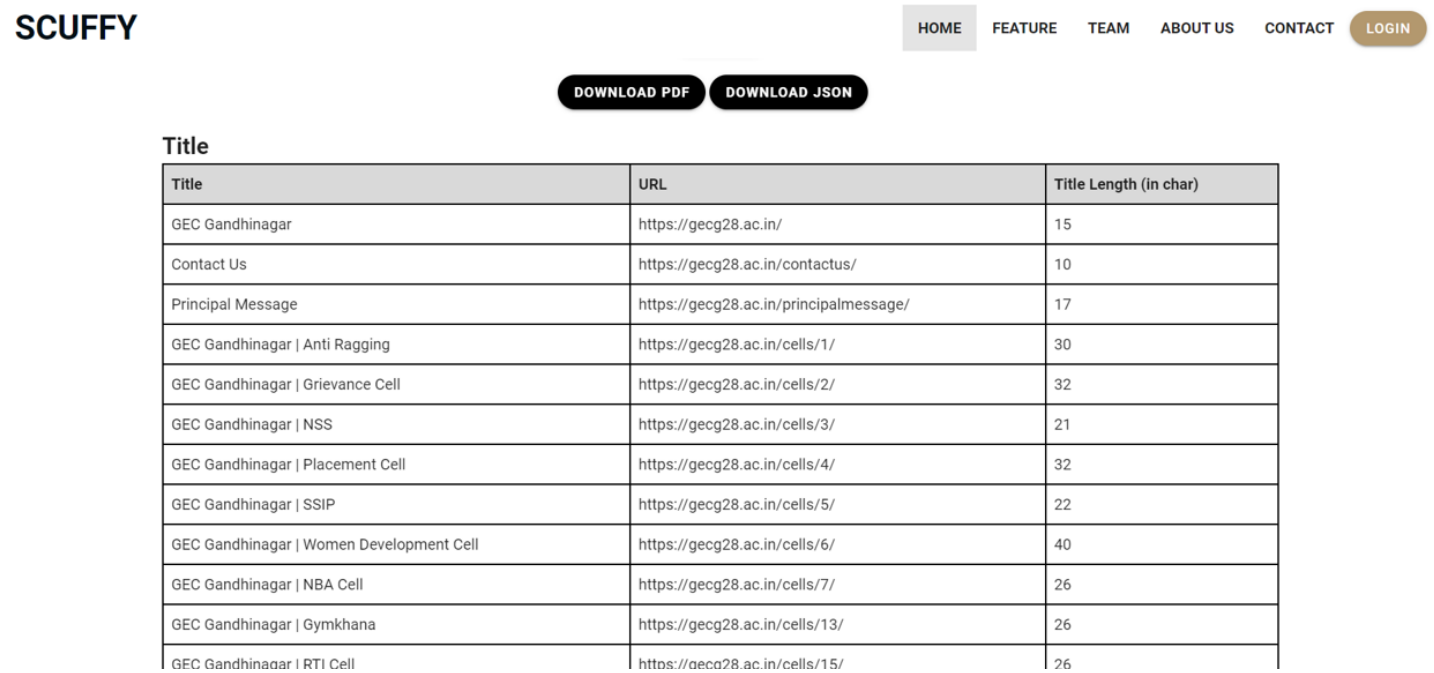



Fig 5.4 After Pasting the URL




| Title | URL | Title Length (in char) |
|---|---|---|
| GEC Gandhinagar | https://gecg28.ac.in/ | 15 |
| Contact Us | https://gecg28.ac.in/contactus/ | 10 |
| Principal Message | https://gecg28.ac.in/principalmessage/ | 17 |
| GEC Gandhinagar \| Anti Ragging | https://gecg28.ac.in/cells/1/ | 30 |
| GEC Gandhinagar \| Grievance Cell | https://gecg28.ac.in/cells/2/ | 32 |
| GEC Gandhinagar \| NSS | https://gecg28.ac.in/cells/3/ | 21 |
| GEC Gandhinagar \| Placement Cell | https://gecg28.ac.in/cells/4/ | 32 |
| GEC Gandhinagar \| SSIP | https://gecg28.ac.in/cells/5/ | 22 |
| GEC Gandhinagar \| Women Development Cell | https://gecg28.ac.in/cells/6/ | 40 |
| GEC Gandhinagar \| NBA Cell | https://gecg28.ac.in/cells/7/ | 26 |
| GEC Gandhinagar \| Gymkhana | https://gecg28.ac.in/cells/13/ | 26 |
| GEC Gandhinagar \| RTI Cell | https://gecg28.ac.in/cells/15/ | 26 |

INin

Fig 5.5 Fetched Metadata

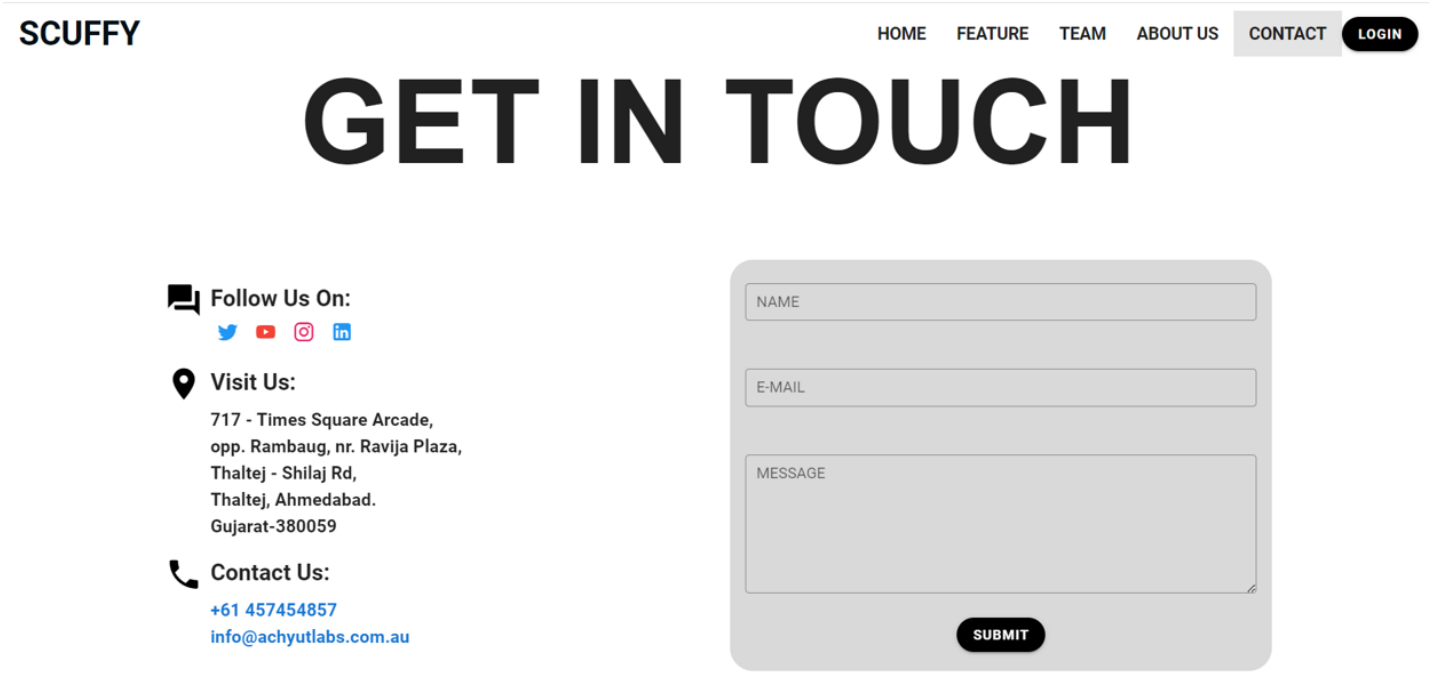



Fig 5.6 Contact Us Page

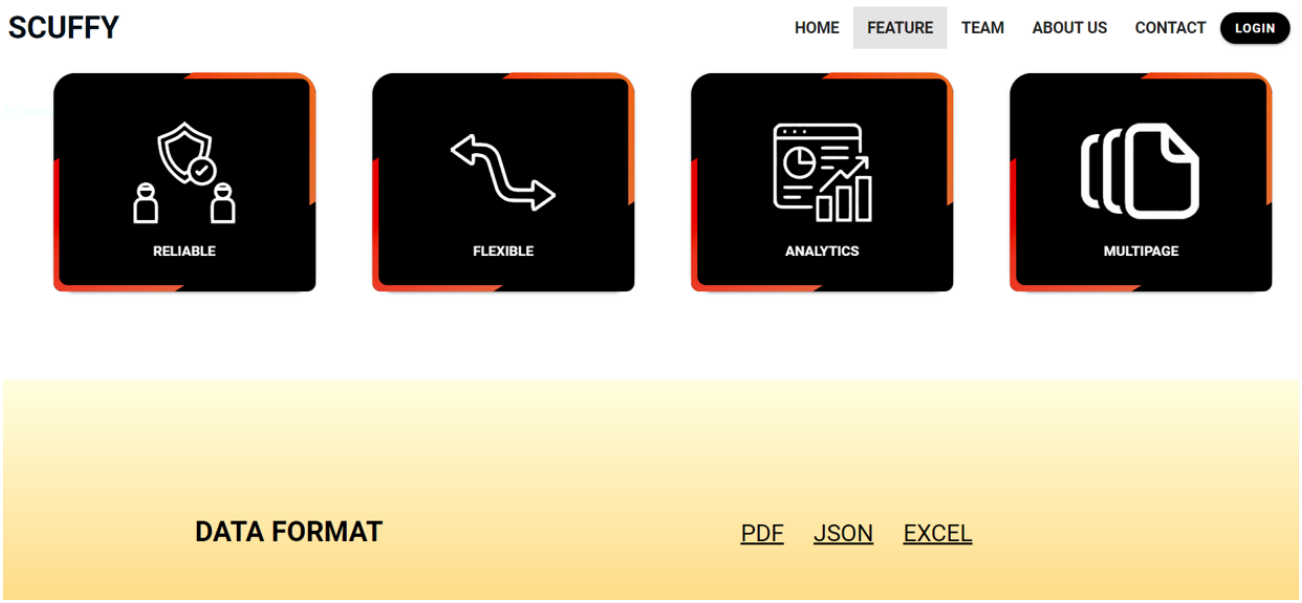



Fig 5.7 Feature Page

# Chapter6
# Implementation

## 6.1 Implementation Platform

Scuffy Web Scraper was implemented using a web browser as the runtime environment for the frontend code, and a localhost server running on port 3000 as the backend environment. The code was developed using the popular integrated development environment (IDE) called Visual Studio Code, which provided a rich set of tools and features to support the development process. The combination of Vue.js and Django frameworks allowed for efficient and effective development of the frontend and backend components respectively. The MySQL database was used to store the data scraped from websites, and testing was carried out using a variety of tools and techniques to ensure the scraper was functioning correctly. Overall, the implementation platform used for Scuffy Web Scraper was robust, scalable, and highly effective at delivering a reliable web scraping tool.

## 6.2 Module Specification

**Home Page:** A homepage in Vue.js is built by disassembling it into smaller, reusable components. A header component, navigation component, hero component, feature component, and footer component are a few examples of possible components. Each component has a unique function, and you may build a complicated web page by mixing them.

The logo, links to the website's navigation, and any other pertinent data are often included in the header component. Links to different parts of the website may be included in the navigation component, which aids in page navigation. The homepage's first prominently displayed portion, known as the hero component, may include text, graphics, or videos that highlight the site's principal message.

**Feature Page:**Both the large text that reads "ANYTHING" and the smaller text that

reads "ANY PAGE" in the first section have various opacity and position settings that are determined by code-defined variables. These factors regulate the texts' visibility and behavior, which appears to be animated or brought about by the user scrolling down the page.

The second section has a call-to-action text and button graphic that encourages users to click on in order to extract content from any website.

The third portion contains four cards with the headings "MULTIPAGE", "SCRAP TITLE", "RELIABLE", and "FLEXIBLE", each of which highlights a different advantage of the technology. Each card has a picture as well as text that describes the characteristic. The playing cards.

**Team Page:**The team page is made up of multiple parts, each of which is imported under its own name. A navigation bar component that is defined outside of this template opens the page.
The Vuetify framework includes the v-row and v-col classes, which enable responsive component layout. The img tag is used to display the banner image, and the src property is used to specify the image source.
The scoped style portion of the template defines the classes.banner-image and.contain. The.banner-image class is used to set the height of the banner image to 50vh (viewport height) and to center it both horizontally and vertically. Use the.contain class to make sure the image doesn't

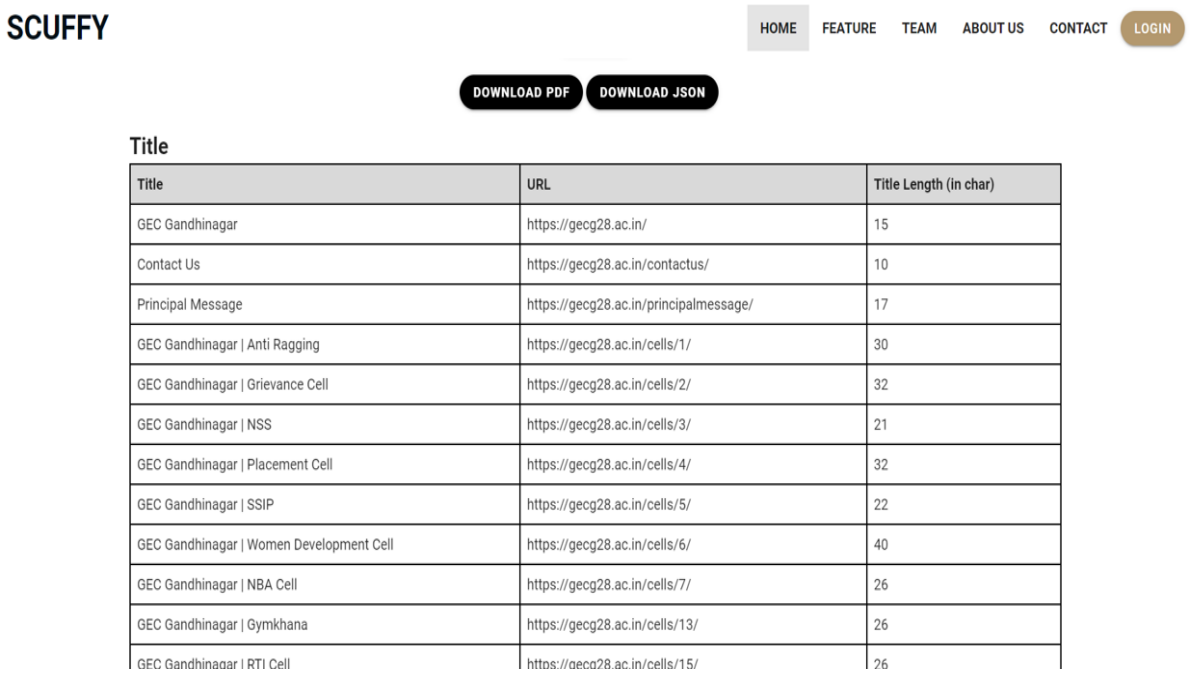
**Contact Us** :The contact page is styled with Vuetify, a UI framework for Vue.js. The component's first element is a Newnavbar navigation bar component. A centered container for the content is created by the v-container element. The v-row and v-col elements combine to generate a two-column layout, with the form appearing in the right column and the social network links, physical address, and phone number appearing in the left column. The href attributes on the v-btn elements used for the social media links point to various URLs. H3 elements with tags that also have href attributes are used to display the physical location and phone number.
A v-form element with three v-text-field elements for the name, email, and message, as

well as a v-textarea element, are found in the right column.

## 6.3 Results / Outcomes

Your life will be a lot easier if you use Fragments when you try to create applications for both phones and tablets. Since the fragments are so tied in with the Honeycomb APIs, you will want to use them on phones as well to reuse code. That's where the compatibility library comes in handy.

The Main advantage of using fragments is when you have a large application, which has more than 100+ screens. In that case if we use 100+ activities then it takes much. Time to call and load new Activity as well as this increases the size of application.




| Title | URL | Title Length (in char) |
|---|---|---|
| GEC Gandhinagar | https://gecg28.ac.in/ | 15 |
| Contact Us | https://gecg28.ac.in/contactus/ | 10 |
| Principal Message | https://gecg28.ac.in/principalmessage/ | 17 |
| GEC Gandhinagar \| Anti Ragging | https://gecg28.ac.in/cells/1/ | 30 |
| GEC Gandhinagar \| Grievance Cell | https://gecg28.ac.in/cells/2/ | 32 |
| GEC Gandhinagar \| NSS | https://gecg28.ac.in/cells/3/ | 21 |
| GEC Gandhinagar \| Placement Cell | https://gecg28.ac.in/cells/4/ | 32 |
| GEC Gandhinagar \| SSIP | https://gecg28.ac.in/cells/5/ | 22 |
| GEC Gandhinagar \| Women Development Cell | https://gecg28.ac.in/cells/6/ | 40 |
| GEC Gandhinagar \| NBA Cell | https://gecg28.ac.in/cells/7/ | 26 |
| GEC Gandhinagar \| Gymkhana | https://gecg28.ac.in/cells/13/ | 26 |
| GEC Gandhinagar \| RTI Cell | https://gecg28.ac.in/cells/15/ | 26 |

Fig. 6.1 Image of Result

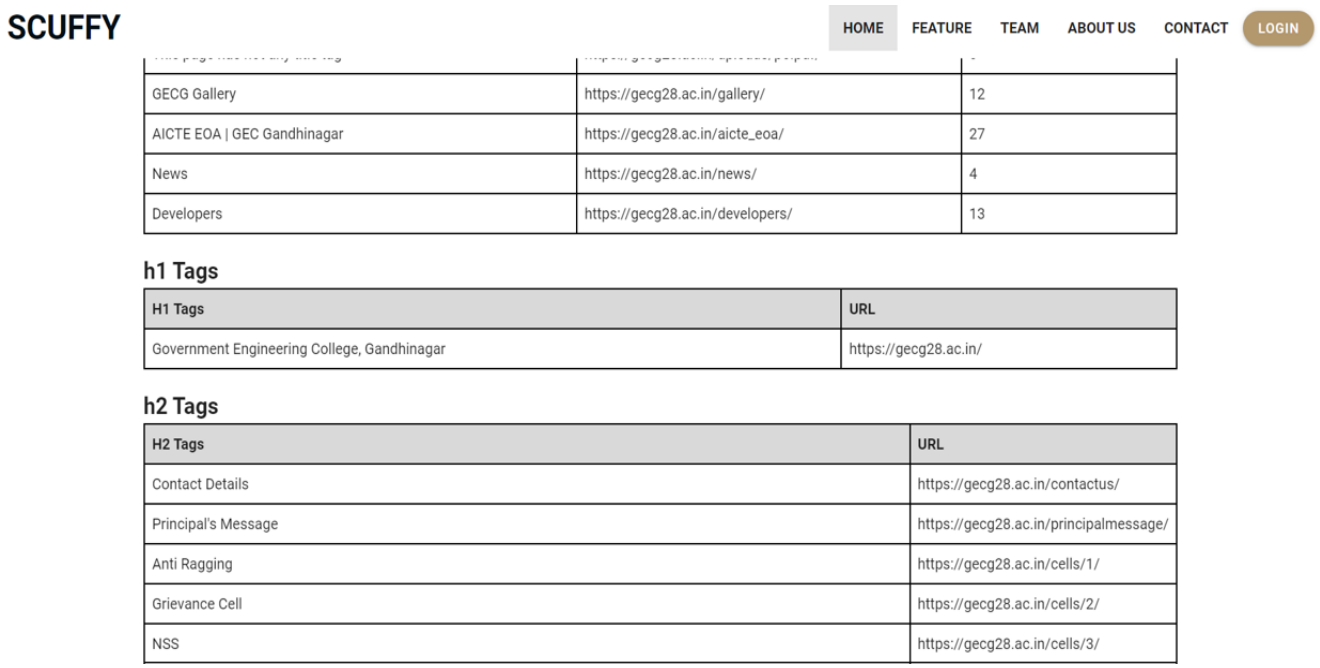



Fig 6.2 Image of the Result

# Chapter 7

# Testing

## 7.1 Testing plan / Strategy

In this project testing is done in different parts. First part is to match the design , css , colors and all drawables should be the same as per given figma. Second part is to check all validations of buttons , text fields , documents etc. third part is to check all data should be saved correctly. The fourth part is to check if all functionalities should work properly as per requirements.

## 7.2 Test Results and Analysis

### 7.2.1 Test Case (test ID, test condition, expected output, actual output,remark)

Table 7.1 Test Cases

| Test Cases ID | Purpose | Condition | Expected Output | Actual Output | Remarks |
|---|---|---|---|---|---|
| 1 | Test with incorrect email | Email: random@gmail.com, Password: something | Credentials do not match our records | Credentials do not match our records | pass |
| 2 | Test with incorrect email | Email:random,Password: something | The email must be a valid email address | The email must be a valid email address | pass |
| 3 | Submit Empty Register Form | name = "" email="" phone = "" password="" | The name field is required, email must be valid, password confirmation do not match, phone number field is | The name field is required, email must be valid, password confirmation do not match, phone number | pass |

| | | | | | |
|---|---|---|---|---|---|
| 4 | Register with already existing email | Email: hsparikh02@gmail.com | The email has already been taken | The email has already been taken | pass |
| 5 | Register user with valid new user | Email:hsparikh02@gmail.com, Name: admin, password: something | User Registered successfully and verification link sent to email id | User Registered successfully and verification link sent to email id | pass |
| 6 | Valid Login | Email: hsparikh02@gmail.com, password: something | Successful login | Successfu l login | pass |
| 7 | Manage Account validation | name = "" phone = "" | please enter valid name please enter valid phone number | please enter valid name please enter valid phone number | pass |
| 8 | Area Range should work properly | range = 15 | 15 miles distance jobs should be displayed | 15 miles distance jobs are displaying | pass |
| 9 | Change Passwod Validation | old password= password 123 entered password= password | "please enter correct old password" | "please enter correct old password" | pass |
| 10 | Change Password Validation | new password = password 123 confirm password = password | "confirm password should be same as new password" | "confirm password should be same as new password " | pass |
| 11 | No internet Connectio n dialog | No internet connection dialog in every page when internet is not available | when internet is not available No Internet Connection Dialog should be displayed | when internet is not available No Internet Connection Dialog should be displayed | pass |

# Chapter 8

# Conclusion and Discussion

## 8.1 Overall Analysis of Internship / Project Viabilities

For those wishing to gain expertise in web scraping, an internship with Scuffy Web Scraper could be a beneficial opportunity. Working with a state-of-the-art web scraping tool could teach interns useful skills including data collecting, analysis, and visualization. An internship with Scuffy Web Scraper could provide you the confidence and knowledge with the newest tools and technology in the field that you need. Interns may work on practical projects that result in observable outcomes, allowing them to demonstrate their abilities and build their professional networks. Interns might maximize their time using Scuffy Web Scraper and get ready for a successful career by putting an emphasis on problem-solving, critical thinking, and learning from feedback.

## 8.2 Limitation

One potential limitation is the sources of data that it can scrape. Some websites may be difficult to scrape due to security measures or other obstacles. Additionally, legal and ethical considerations such as data privacy and intellectual property must be taken into account when scraping data. Technical limitations, such as the processing power and memory required for large data sets, could also be a challenge for some systems. Maintenance and support are also important considerations, as the system will require ongoing updates and monitoring to function effectively. Finally, user error is another potential limitation. Proper training and guidance must be provided to ensure that users are using the system effectively and inputting correct data. Despite these limitations, Scuffy Web Scraper can still provide a valuable tool for data collection and analysis.

## 8.2 Future Enhancement

As technology continues to evolve, there are many potential future enhancements for Scuffy Web Scraper. One area for improvement could be in the integration of artificial intelligence and machine learning. By leveraging these technologies, Scuffy Web Scraper could become even more accurate and efficient in its data collection and analysis. Another potential enhancement could be in the development of new algorithms or techniques for scraping data from difficult sources. This could expand the scope of the system and allow it to collect data from a wider range of sources. Additionally, enhancements to the user interface and user experience could make the system more intuitive and user-friendly. Finally, the addition of new features or modules could make Scuffy Web Scraper even more versatile and useful for a variety of applications. These enhancements could help to further establish Scuffy Web Scraper as a valuable tool for data collection and analysis in a wide range of industries

# Appendix

**Weekly Report 1**

# Weekly Report 2

# Weekly Report 3

212
394

Weekly
Report 3

# Weekly Report 4

212
394

Weekly
Report 4

# Weekly Report 5

212
394

Weekly
Report S

# Weekly Report 6

# Weekly Report 7

212
394

Weekly
Report 7

# Weekly Report 8

212
394

Weekly
Report 8

# Weekly Report 9

212
394

Weekly
Report 9

# Weekly Report 10

212
394

Weekly Report
10

# Weekly Report 11

212
394

Weekly Report
11

# Weekly Report 12

212
394

Weekly Report
12

# Feedback