Introduction to ParseHub

ParseHub is a powerful and user-friendly web scraping tool that allows users to extract data from websites without the need for any external coding. It provides a visual interface that enables users to interact with web pages, select the data they want to scrap, and organize it into structured formats like CSV, Excel, or JSON.

ParseHub is particularly useful for those who have limited programming skills but need to gather data from the web for various purposes, including research, business intelligence, or analysis.

Here are the steps to follow for web scraping using ParseHub:

Installation and Sign-Up

- Visit the ParseHub website (https://www.parsehub.com/).

- Sign up for a free account if you don't already have one.

- Download and install the ParseHub desktop application on your computer.

Create a New Project

- Launch ParseHub and click on "New Project."

- Enter the URL of the website you want to scrape data from and click "Start Project."

Set Up the Template

- ParseHub will load the webpage in a visual interface.

- Use the "Select" tool to click on the data elements you want to extract. ParseHub will automatically detect similar elements on the page.

- Customize your selections using the "Select" tool and refine them as needed.

- Create new templates for different types of data on the same page.

Start the Run

- Click on the "Get Data" button.

- ParseHub will scrape the data based on your selections.

- You can preview the data before proceeding.

Export the Data

- After confirming the scraped data, click on the "Run" button.

- ParseHub will extract and organize the data.

- Once the run is complete, click on "Download Data" and choose the desired format (CSV, Excel, JSON, etc.).

- Save the file to your computer.

Here's an example of what the data might look like in a CSV file:

A screenshot of a computer

Description automatically generated

I tried to extract the small data because I’m unable to load parser hub application properly even after multiple installations.

Snippet of my issue.

A screenshot of a computer

Description automatically generated