

Review on Web Scrapping

Alen Geo Alex¹, Nakul S Kumar², Prithwiraj Prakash³

Student, Santhigiri College of Computer Sciences, Vazhithala, India^{1,2,3}

Abstract: *Web scraping (also known as Screen Scraping, Web Data Extraction, Web Harvesting, and so on) is a method of extracting publicly uploaded data, processing it, and constructing it into plain or processed data. The result may then be applied to a variety of domains such as Artificial Intelligence, Machine Learning, Market Evaluation, and so on.*

Keywords: Web, Wanderer, JSoup, Scraping, Browser, Data, Webpage, etc.

REFERENCES

- [1] Al Sweigart, Automate the Boring Stuff with Python, 2nd Edition: Practical Programming for Total Beginners 2nd Edition.
- [2] Michael Heydt, Python Web Scraping Cookbook: Over 90 proven recipes to get you scraping with Python, microservices, Docker and AWS.
- [3] Ryan Mitchell, Web Scraping with Python: Collecting More Data from the Modern Web 2nd Edition.
- [4] Bart Baesens, Practical Web Scraping for Data Science: Best Practices and Examples with Python 1st ed. Edition.
- [5] Jaime Buelta, Python Automation Cookbook: 75 Python automation ideas for web scraping, data wrangling, and processing Excel, reports, emails, and more, 2nd Edition.
- [6] Olgun Aydin, R Web Scraping Quick Start Guide: Techniques and tools to crawl and scrape data from websites.
- [7] SeppevandenBroucke, Practical Web Scraping for Data Science: Best Practices and Examples with Python 1st ed. Edition.
- [8] <https://webscraper.io/blog/brief-history-of-web-scraping>
- [9] <https://www.parsehub.com/blog/web-scraping-legal/>
- [10] <https://hevodata.com/learn/8-best-web-scraping-tools/#topWST>
- [11] https://en.wikipedia.org/wiki/Web_scraping