Selenium ve Scrapy ile Web Scraping

Why Selenium? Isn’t Beautiful Soup enough?

Web scraping with Python often requires no more than the use of the Beautiful Soup to reach the goal. Beautiful Soup is a very powerful library that makes web scraping by traversing the DOM (document object model) easier to implement. But it does only static scraping. Static scraping ignores JavaScript. It fetches web pages from the server without the help of a browser. You get exactly what you see in “view page source”, and then you slice and dice it. If the data you are looking for is available in “view page source” only, you don’t need to go any further. But if you need data that are present in components which get rendered on clicking JavaScript links, dynamic scraping comes to the rescue. The combination of Beautiful Soup and Selenium will do the job of dynamic scraping. Selenium automates web browser interaction from python. Hence the data rendered by JavaScript links can be made available by automating the button clicks with Selenium and then can be extracted by Beautiful Soup.

References

1- <https://medium.com/ymedialabs-innovation/web-scraping-using-beautiful-soup-and-selenium-for-dynamic-page-2f8ad15efe25>

2 - <https://www.youtube.com/watch?v=rONhdonaWUo>