**Web scraping**

Web scraping is a data collection from website for personal or industrial use, wait what is this…. ok

First ask some questions with yourself:

* What is web scraping?
* Why we use web scraping?
* How its work?
* It is legal or not?

**What is web scraping?**

Web scraping is an automated method used to extract large amounts of data from websites. The data on the websites are unstructured. Web scraping helps collect these unstructured data and store it in a structured form. There are different ways to scrape websites such as online Services, APIs or writing your own code. Or in simple words, web scraping used to gain data from any website without visiting that website in easy and quickly way (extracting data from a website).

[**What is web scraping used for?**](https://www.octoparse.com/blog/web-scraping-for-businesses) **(Why we use web scraping?)**

**Example:** Web scraping is used to collect data from Social Media websites such as Twitter to find out what’s trending.

[**Is web scraping legal?**](https://www.octoparse.com/blog/is-web-scraping-legal-in-some-countries)

Web scraping is legal or not, some websites allow web scraping and some don’t. To know whether a website allows us web scraping or not, you can look at the website by a command.

**Which’s the best web scraping tool?**

**What’s the best programming language for web scraping?**

**Can I extract data from the entire web?**

**Is web scraping** [**data mining**](https://www.octoparse.com/blog/7-web-mining-tools-around-the-web)**?**

**How to avoid being blocked when scraping a website?**

**What is the difference between web scraping and** [**web crawling**](https://www.octoparse.com/DataCrawler)**?**

**What is a robots.txt file?**

**“robots.txt”** is a way to find out that web site allow web scraping on the site or not. You can find out the legal website for web scraping and illegal website this can be done by appending “/robots.txt” file to the URL that you want to scrape.

**Example:** www.abcd.com/robots.txt

**Can I scrape data behind a login page?**

**How do I extract the content from dynamic web pages?**

**Can you crawl Facebook/LinkedIn/Twitter?**

[**What are some strategies for keeping up with the latest literature in a field?**](https://www.researchgate.net/post/What_are_some_strategies_for_keeping_up_with_the_latest_literature_in_a_field)