

# Introduction to Web Scraping Tools

Join us to learn about What is web scraping? How do web scrapers work? Types of web scraping tools...in the article below

As of January 2020, there were close to 1.3 billion websites online. These invaluable repositories of data hold useful business data such as product details, prices, stock prices, company contacts, and sports statistics, to name but a few.

Picture 1 of Introduction to Web Scraping Tools

The benefit of Big Data analytics is not lost to businesses that depend on these insights for decision-making. Access to the actionable insights held within these pages is not humanly possible through a manual copy and paste method.

Technology has come up with automated web scraping tools that speed up the data collection process eliminating any human error.

## What is web scraping?

Web scraping is the process of extracting data from online sources using a bot or spider. The data scraped from web pages is downloaded and saved in usable file formats such as spreadsheets. An API can also direct the data collected to a business's business management software.

These automated tools are preferred to manual methods due to the obvious advantages of speed and affordability. Web scraping tools have advanced over time, that any person can use them to mine data. They have a low learning curve meaning that programming skills are not a necessity for business web scraping needs.

## How do web scrapers work?

There are many different programming languages out there used to build websites. Web pages, therefore, differ in form, build, and content. For this reason, web scrapers vary in features and functionality too.

Nevertheless, the basic functioning of a web scraper begins with the input of a URL that directs the tool towards the focus website. Some sophisticated web scrapers will scrape the whole website, including content such as Javascript and CSS features. Some only render HTML elements.

It is possible to limit the scraping function of a tool to ensure that it only scrapes the necessary information. You can, for instance, gather price data only from an Amazon page and ignore the descriptions or reviews if you do not need his data for your business insights.

After data collection, your tool of choice will output the information collected into usable formats, such as an Excel sheet. Tools with API support will save the data in JSON and other formats.

## **Types of web scraping tools**

As mentioned, there are different types of websites and information databases out there. Consequently, there are various web scraping tools designed in different programming languages and optimized for different websites and database structures. Some of the different types of web scraping-out there include;

### **Self-built web-scraping tools**

These tools are built by tech-savvy developers with a keen knowledge of programming languages such as Python. Self-built web scrapers can collect data from minimal online sources. The build of robust data scrapers requires extensive programming knowledge and support.

You can also download simple pre-built web scrapers that you can put into action immediately after download. They, of course, do not have as many advanced or support features as premium tools but may have useful but straightforward features such as JSON exports or scrape schedulers.

### **Browser extensions**

These are the most common web scraping tools. They are software-like programs that are installed on browser programs such as Firefox or Google Chrome. These tools are easy to run and manage since they integrate easily with a browser.

Their main disadvantage is that the browser also limits their function. They can only use features that browsers can support. Browsers were not designed with web scraping in mind, so some useful web scraping functions such as IP rotations cannot operate from browser extensions.

### **Scraper software**

Web scraping software is downloaded and installed directly into a computer. These web-scraping tools have better features than browser extensions. They, however, are not as easy to get working as browser extensions are.

## **Other varying features in web scraping tools**

### **User interfaces**

Simple web scrapers have very minimalistic user interfaces. They might also just have a command line, a feature that can be a bit challenging for users that do not have any web scraping experience. There are sophisticated web scrapers with impressive user interfaces that make them much easier to use.

Such scraping solutions can render whole websites at a simple click. They also have premium features such as helpful suggestions and tips and customer support services.

### **Location of web scraper**

A locally run web scraping tool will feed off the resources of your computer. The strength of your internet connection will also influence its working. If your computer's RAM or CPU resources are inadequate, then the locally installed scraper will not function as it is designed to.

On the flip side, should your scraper encounter a website with tons of URLs, the scraping operation will slow down the functioning of even the robustness of computers. This form of scraping could also affect your internet service provider's rates or bandwidth cap.

Cloud-based data mining tools run from off-sites servers. Their servers are provided, managed, and maintained by the tool's providers. These web scraping tools free up computer resources allowing you to work and perform web scraping at the same time.

Businesses are turning to web scraping tools for access to business intelligence and Big Data analytics. Is your business ready for the digital age?

You finished reading the article "**Introduction to Web Scraping Tools**" edited by the [TipsMake](#) team. We hope this article has provided you with many useful tech tips and tricks. You can search for similar articles on tips and guides. Thank you for reading and for following us regularly.