**DEFENITION**

Web Scraping is the process of collecting structured web data in an automated fashion. To be honest, it became more of an art rather just a simple process. You can organize collected data in many unique ways. That’s why it is so cool.

**WHY WEB SCRAPING**

* In today’s world it became very hard to compete with big companies and start a business from a scratch. Web Scraping is an excellent way to fill your database with information and have a good starting point. (STEALING information. Sometimes author rules may apply! So be careful!!)
* Some companies use web scraping as a price regulator tool. Companies are able to monitor competitors’ prices on the same type of products and thus are able to adjust their pricing to compete for a marketplace. (EXAMPLE: WALMART collects information about different types of groceries from other markets and makes its prices more affordable)
* Another important point is Data Science (or Machine Learning). If you are familiar with Data Science at some point you know that programmers need to have a good database with information for their algorithms. (EXAMPLE: You are trying to guess the type of a tree by its dimensions. You are going to need measurements from many trees to base your algorithm on that collected data)

**TOOLS**

The tools we are going to use are open source and easy to find.

**THE DOM**

And now we need to understand with what we are going to work. The main Data Structure is going to be a tree. On the Web its called ‘THE DOM’. The Document Object Model represents a document (like HTML) with a logical tree. Each branch of the tree ends in a node and each node contains objects. We are going to use this structure.

**MAIN CONCEPTS**