

Web Scrapping

import scrapy

Data Scraping

Data scraping (raspagem de dados) é uma técnica computacional na qual um programa extrai dados de saída legível somente para humanos, proveniente de um serviço ou aplicativo. Os dados extraídos geralmente são minerados e estruturados em um formato padrão como CSV, XML ou JSON.

Metodologias de extração de dados

- Human copy-and-paste
- Text pattern matching
 - UNIX **grep** command or **regular expression-matching** facilities of programming languages (for instance Perl or Python)
- HTTP programming
- HTML parsing
- Semantic annotation recognizing
- Computer vision web-page analysis
 - **machine learning** and **computer vision**

Web crawling frameworks



Indicação

<https://www.udemy.com/python-web-scraping/>

<https://pythonhelp.wordpress.com/2014/08/05/web-scraping-com-scrapy-primeiros-passos/>