



Scrapy



O que é o Scrapy

- Framework de web scraping em Python
- Fácil de usar
- Extensível



Por que usar o Scrapy

- É um framework, com soluções para problemas comuns em qualquer parte do processo de web scraping: no acesso ao site, na extração dos dados, no armazenamento, etc.
- Rápido e eficiente por usar o Twisted para fazer requisições de forma assíncrona, mas o programador não precisa lidar com essa complexidade quando usa o Scrapy



Diferenças entre o Scrapy e o requests

- O requests apenas lida com HTTP, o Scrapy tem ferramentas para extrair e armazenar os dados
- O requests faz requisições de forma síncrona/bloqueante



Diferenças entre o Scrapy e o BeautifulSoup

- O BeautifulSoup apenas extrai dados do HTML
- O Scrapy tem ferramentas para extrair dados do HTML através de seletores XPath ou CSS
- É possível usar o BeautifulSoup com o Scrapy
 - <https://doc.scrapy.org/en/latest/faq.html#can-i-use-scrapy-with-beautifulsoup>