Extraindo dados da Web com Python

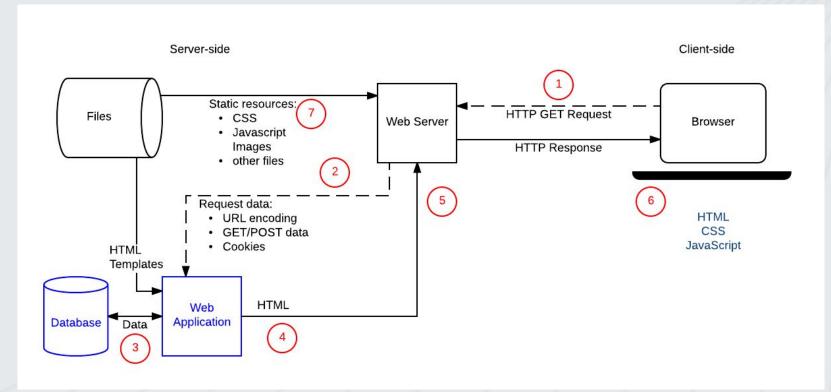


Rogers Prates de Pelle

Desenvolvedor Python na Zyte

Professor na UEMS

Requisição Cliente-Servidor



Web Crawler



Exemplos

01: Minerar dados sobre um conjunto de produtos

02: Extração de dados imobiliários

03: Extração de dados de hotéis

04: Extração de dados de mídia social

05: Obtenção de grandes massas de dados

06: Streaming de dados em tempo real

07: Obter dados de um site sem API oficial

scrapy



An open source and collaborative framework



In a fast, simple, yet extensible way



Maintained by

Zyte and many

other

contributors



github.com/rogersdepelle/scrapy-tutorial

toscrape.com

zyte.com/scrapy-cloud

01: JS (Selenium)

02: IP Banning (Proxy)

Desafios

03: reCAPTCHA (Zyte Splash)

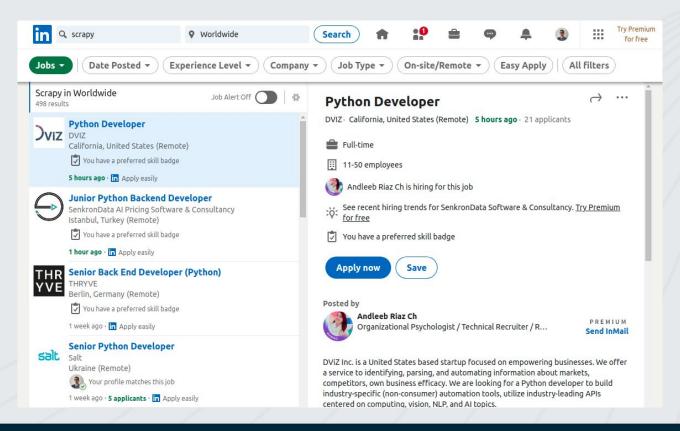
04: Antibot (User-Agent e outros Headers)

05: Site performance (DOWNLOAD_DELAY)

06: Legal Issues (robots.txt)



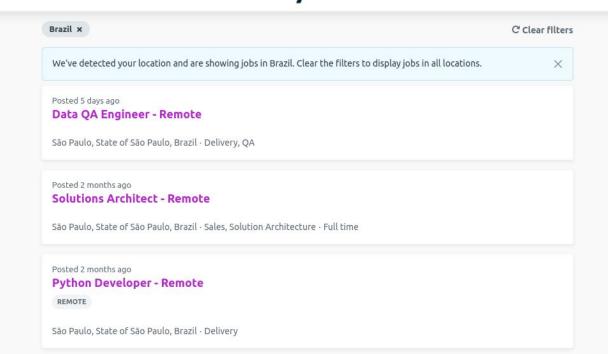
Mercado de Trabalho





Mercado de Trabalho

zyte



apply.workable.com/zyte

Thank You

github.com/rogersdepelle

linkedin.com/in/rogersdepelle