

Deploying Crawlers Using Scrapy Cloud



Janani Ravi

CO-FOUNDER, LOONYCORN

www.loonycorn.com

Overview

Scrapinghub Ltd. is the developer and maintainer of the Scrapy library

Offer a wide variety of products and scraping services

Developer tools on the cloud which allow you to scrape at scale

Deploy a Scrapy spider to Scrapy Cloud

Build a spider using Portia, the UI-based tool

Scraping on the Cloud

Scrapinghub

Scrapy Cloud

Deploy your locally built spiders to the cloud

Crawlera

A smart proxy that never gets banned and doesn't need IP rotation

Splash

A browser behind an API - can scrape dynamic content as well

Portia

Point and click UI tool to build spiders on the cloud

Scrapinghub

Scrapy Cloud

Deploy your locally built spiders to the cloud

Crawlera

A smart proxy that never gets banned and doesn't need IP rotation

Splash

A browser behind an API - can scrape dynamic content as well

Portia

Point and click UI tool to build spiders on the cloud

Demo

**Deploying a locally built spider to the
Scrapy Cloud**

Demo

**Building a spider on the cloud using
Portia**

Summary

Developer tools on the cloud which allow you to scrape at scale

Deploy a pre-built Scrapy spider to Scrapy Cloud

Build a spider using Portia, the UI-based tool



Websites

Scrapy documentation:

<https://scrapy.org/>

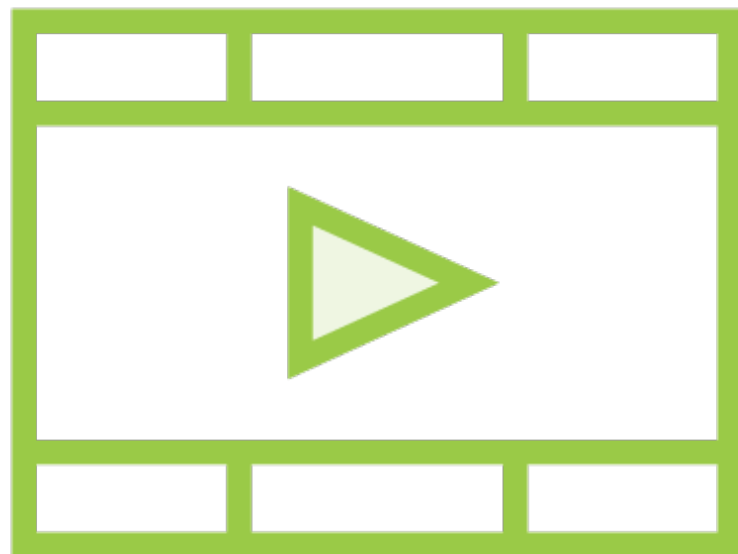
Scrapinghub website:

<https://scrapinghub.com/>

BeautifulSoup:

<https://www.crummy.com/software/BeautifulSoup/bs4/doc/>

Related Courses



Getting Started with Spark 2

**Getting Started with Stream Processing
with Spark Streaming**