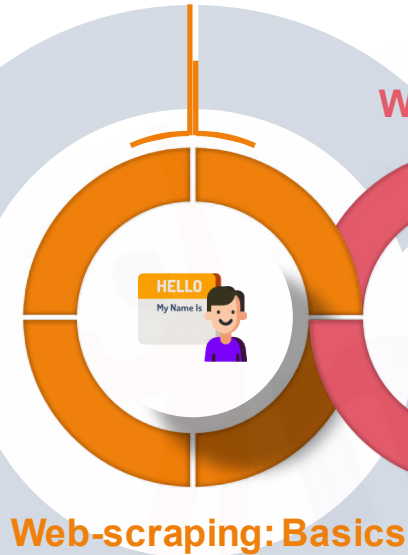


A stylized illustration of a globe with blue oceans and green landmasses. A black spider is positioned on the globe, with a thin line extending upwards from its body to a spider web at the top center of the frame.

Python Web-scraping

Agenda for Today's Session

01 Get a brief understanding of what exactly is Web Scrapping



Web-scraping: Usecase

02 Learn about how web-scraping can be used for different business scenarios



03 Learn about various Python Libraries which are used for web scraping

Python Libraries



HTML Basics

04 Quick overview of basic HTML Tags



05 Scrape data from a website using python libraries beautiful soup and urllib2

Demo: Scrapping a Website

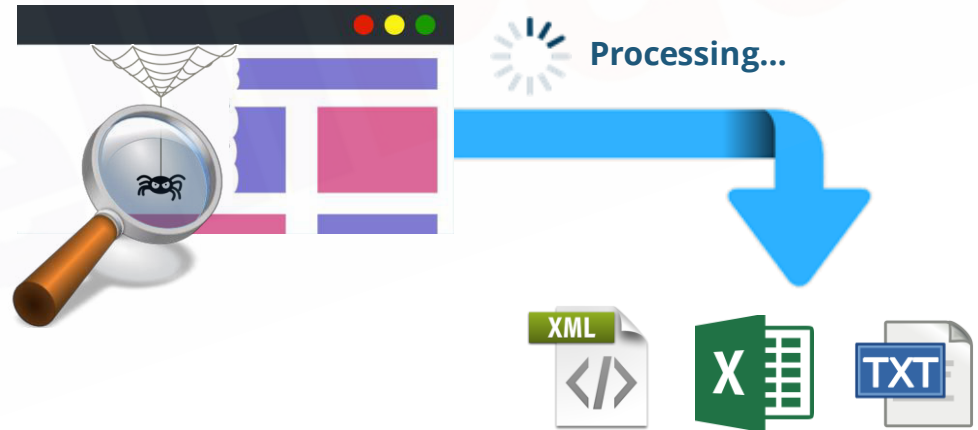


Web-scraping Basics



What is Web-scraping?

Extracting unstructured data from the websites and transforming it to a structured data...



Web-scraping Basics

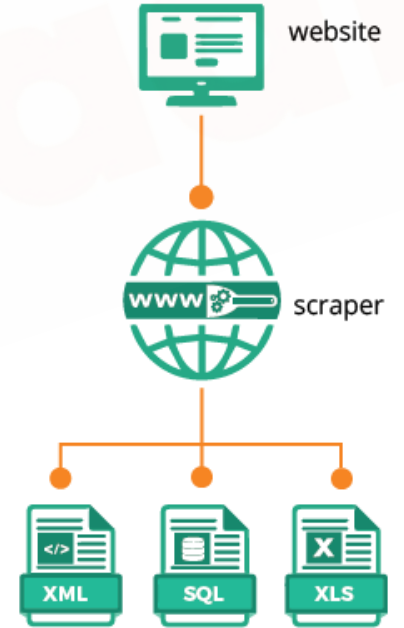


Are they same?

WEB CRAWLER



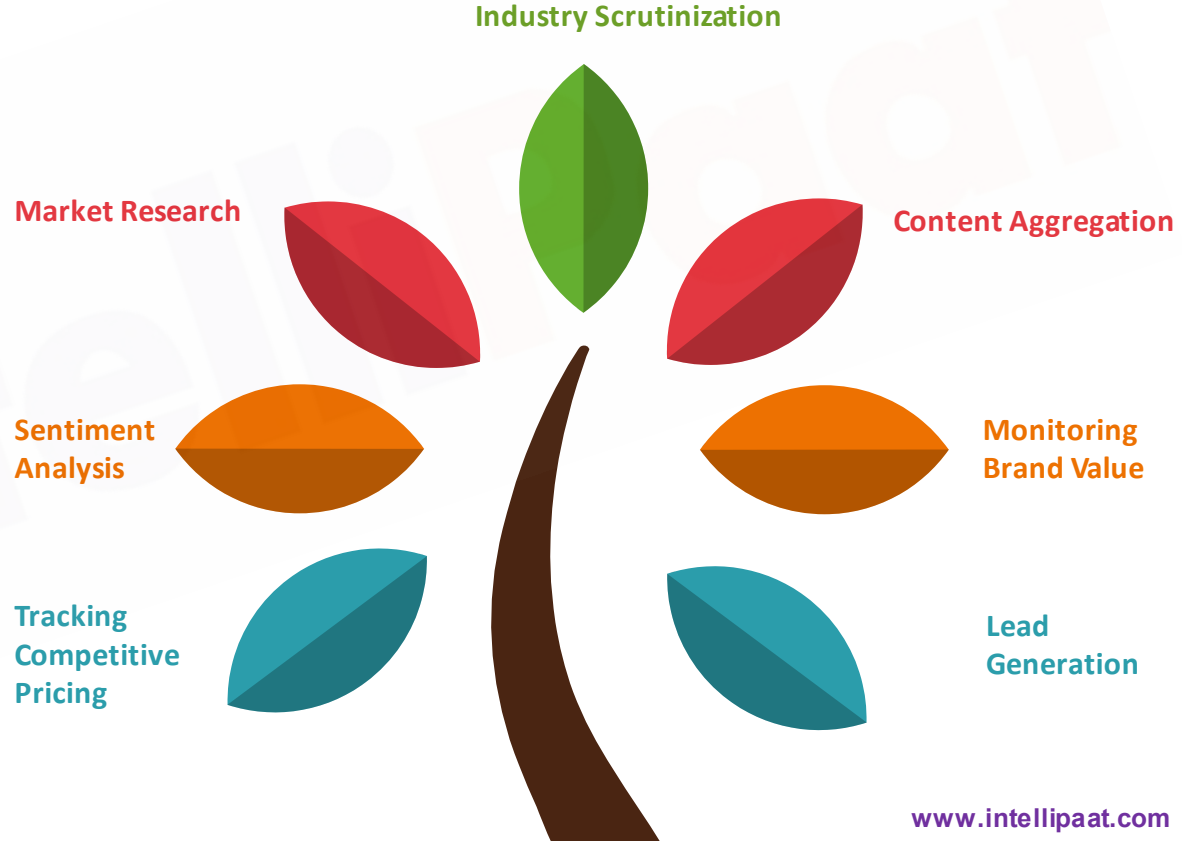
WEB SCRAPING



Need of Web-scraping



Business Application of Web-scraping



Is Web-scraping Legal?



Is Web-scraping / Web-Crawling Legal?

Maybe..



Is Web-scraping Legal?



Is Web-scraping / Web-Crawling Legal?

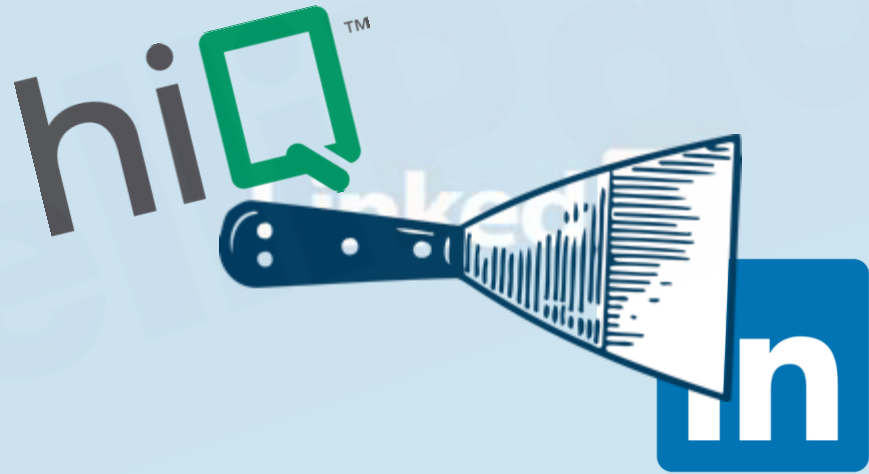
Know The Rules...

- Robots.txt
- Public Content
- Terms of Use
- Authentication Based Websites
- Crawl Delay



Is Web-scraping / Web-Crawling Legal?

Is Web-scraping Legal?



Python Libraries Required



Python Libraries required for Web-Scraping

Urllib2

Used for fetching URLs

<http://>

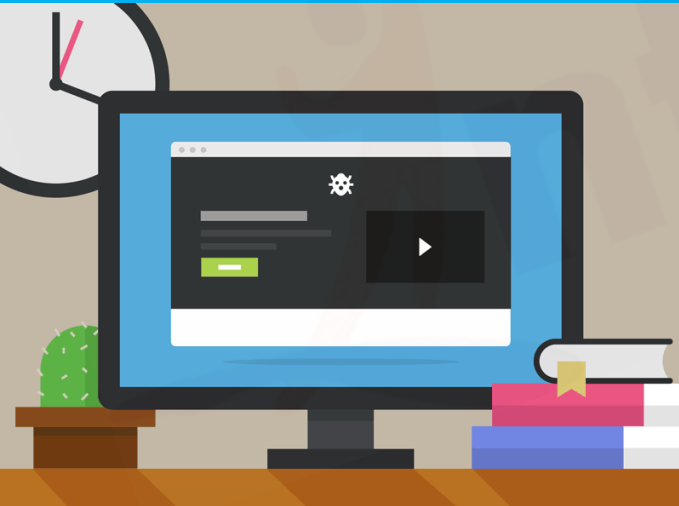


Beautiful Soup

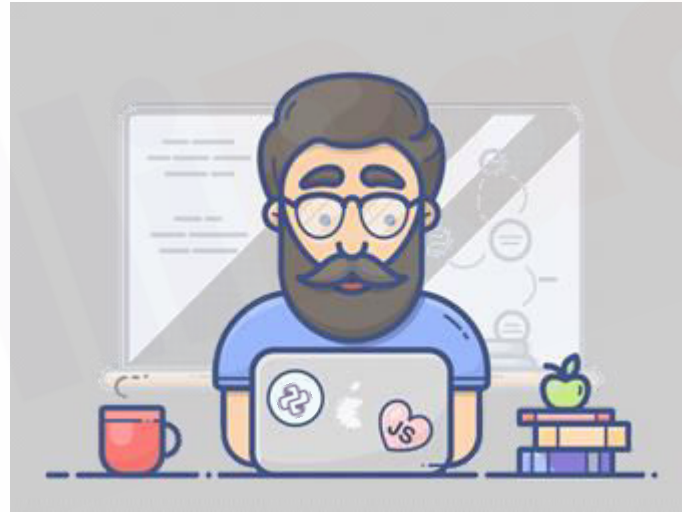
Used for pulling out information from the webpage



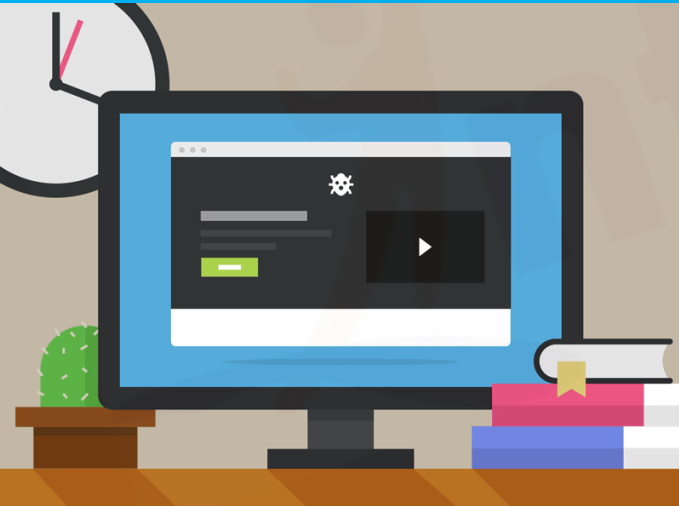
HTML Basics



Get familiar with basic HTML Tags



HTML Basics



Get familiar with basic HTML Tags

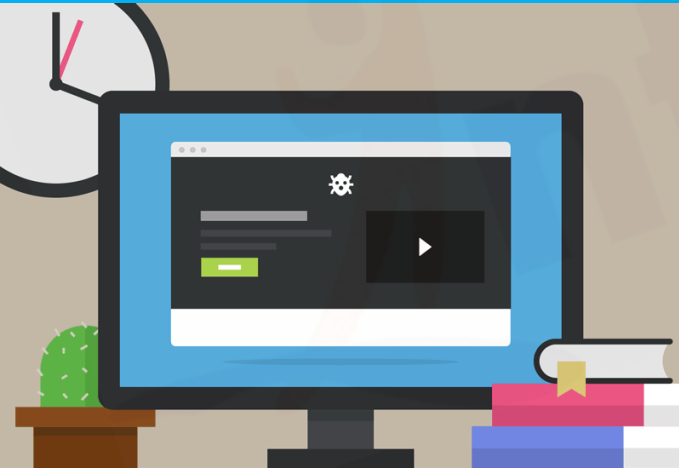
```
<!DOCTYPE html>
<html>
<body>

<h1> My first Header/Headline </h1>
<p> My first Paragraph </p>

</body>
</html>
```

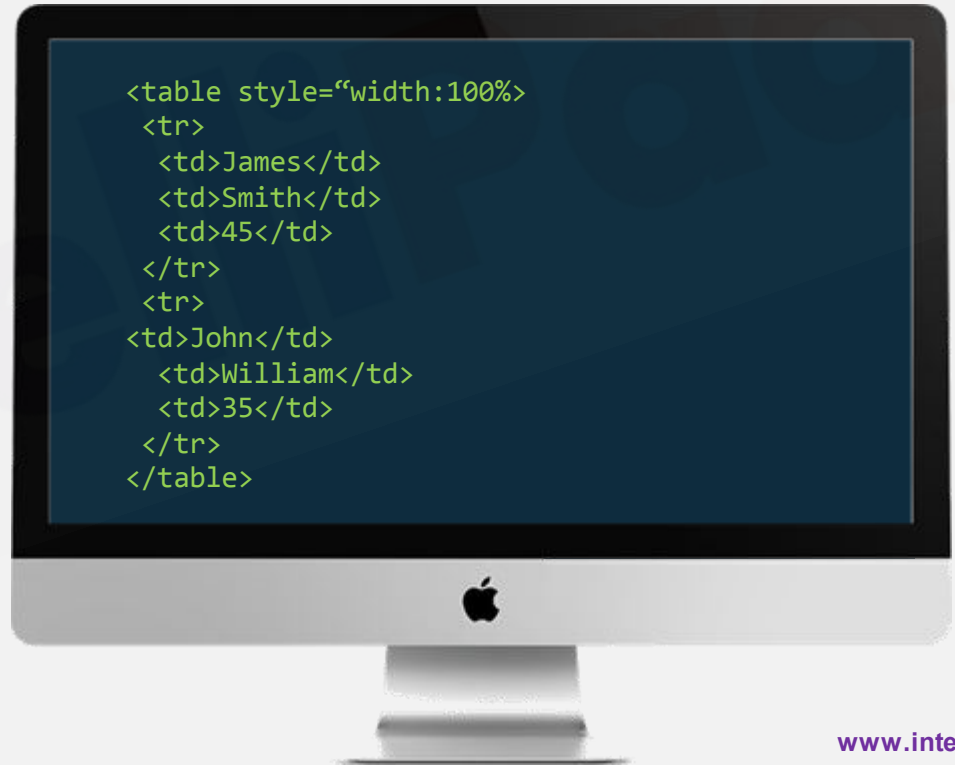


HTML Basics

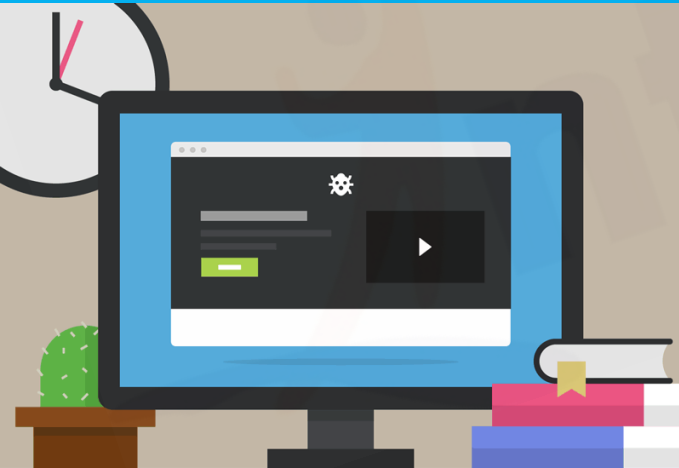


Get familiar with basic HTML Tags

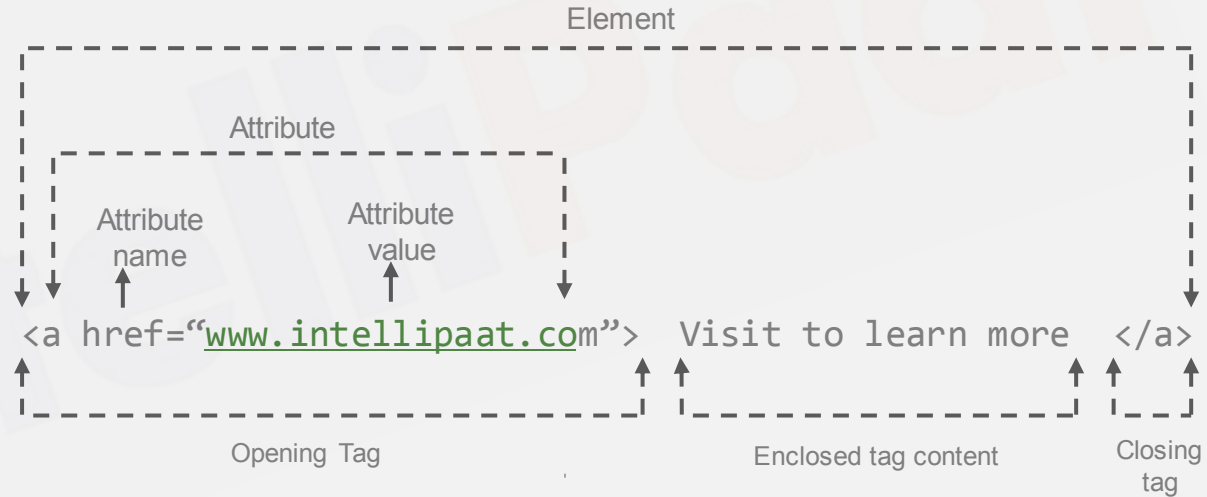
```
<table style="width:100%>
  <tr>
    <td>James</td>
    <td>Smith</td>
    <td>45</td>
  </tr>
  <tr>
    <td>John</td>
    <td>William</td>
    <td>35</td>
  </tr>
</table>
```



HTML Basics



Get familiar with basic HTML Tags



Scraping a Webpage



Thank
You



www.intellipaate.com



CALL US NOW

India : +91-7847955955

US : 1-800-216-8930 (TOLL FREE)

sales@intellipaate.com