



WEB Scrapping

Done by Kaveh Karimadini

TABLE OF CONTENTS

Web crawling

01



Web scraping

02



crawling VS
scraping

03



Hands On

04





Web Crawling

The ART of Data Collection



Web Crawling

What is web crawling?

- Google use crawlers to discover and index web pages.
- Crawling process:
 - ◆ following links, discovering new pages, and storing information.
- Benefits:
 - ◆ Helps users find relevant information on the web



Web Scraping

What is web scraping?

- web scraping extracts specific data points from web pages.
- Techniques:
 - ◆ Selenium with python
- Applications:
 - ◆ Data Collection





Web Crawling VS Scraping

General to Specific



Key Differences

1


Crawling: Discovers and indexes web pages. It's like exploring the entire library to find all the books.

Scraping: Typically targets specific websites or sections of a website, like a product category page.

2

Crawling: Often stores basic information about web pages, like titles and URLs, for indexing purposes.

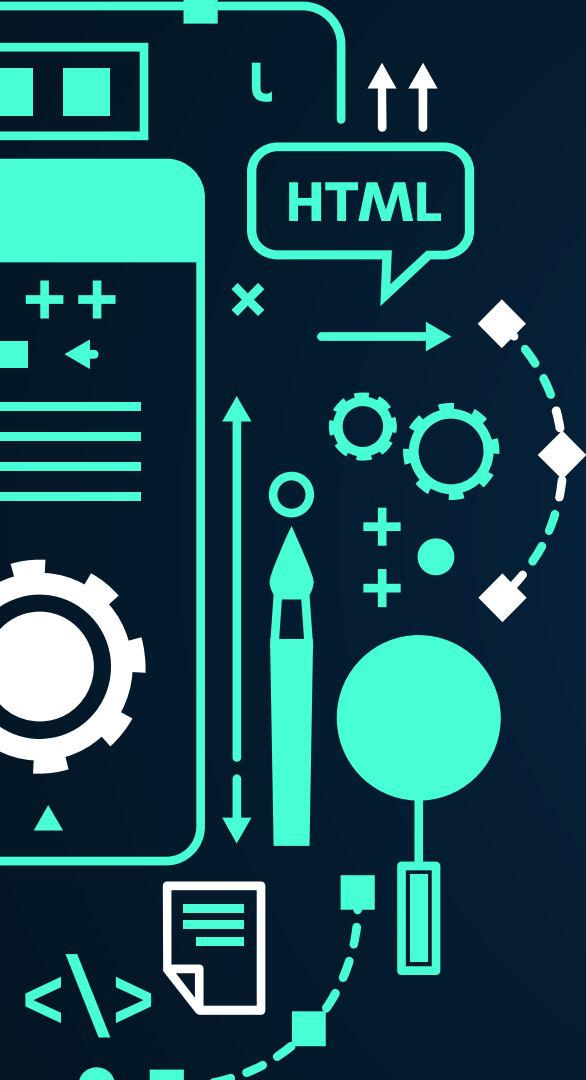
Scraping: Needs to parse and clean the extracted data to make it usable (e.g., removing HTML tags, converting formats).



"Practice makes perfect"

— **Kaveh Karimadini**

Hands On Project: `LinkedIn_scraper.ipynb`



THANKS!

Does anyone have any question?

kavehkarimoodini@gmail.com
+98 09058605654

