

Web Scraping Tools Guide

Web Scraping Tools Guide

Tools to Extract Data from Websites:

Python Libraries

- BeautifulSoup - Parse and extract data from HTML/XML.
- Scrapy - A powerful framework for large-scale web scraping.

No-Code Tools

- Octoparse
- ParseHub
 - > Ideal for non-programmers to automate data extraction visually.

Key Tips:

- Understand HTML elements like <div>, <table>, and classes/IDs.
- Handle pagination, authentication, and JavaScript-rendered content.
- Always respect robots.txt and website TOS.

Output

- Export datasets as CSV, JSON, or insert into databases.
- Build custom datasets for analysis and visualization.