



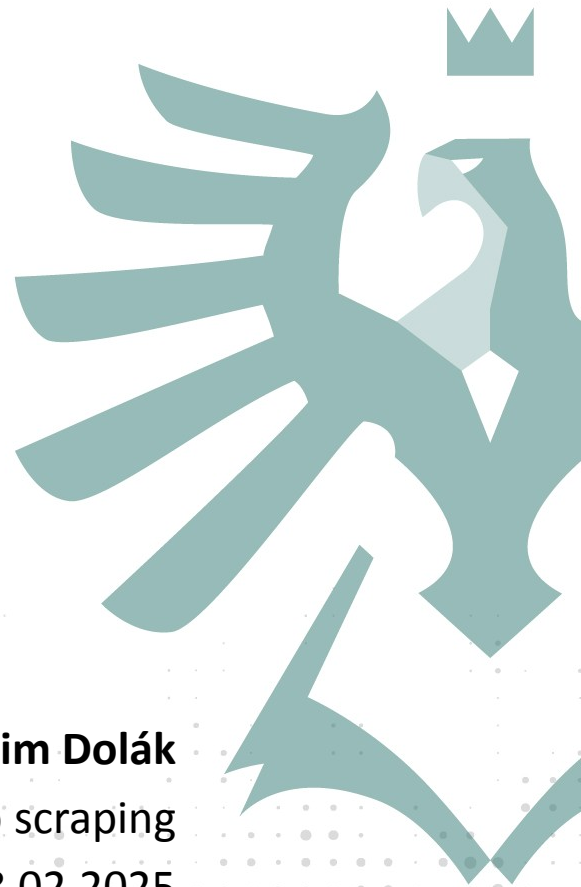
**SLEZSKÁ
UNIVERZITA**
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVINĚ

WEB SCRAPING

Efektivní extrakce dat z webu

www.slu.cz

Radim Dolák
Workshop na téma Web scraping
18.02.2025



WEB SCRAPING – DEFINICE POJMU

- ❖ **Web scraping (čti jako „web skrejping“) je technika automatického získávání, třídění a využívání dat z webových stránek pomocí specializovaných nástrojů.**
- ❖ **Tato data se následně ukládají do souborů ve formátech jako XLS, CSV nebo JSON.**
- ❖ **Cílem web scrapingu je především úspora času a rychlé získávání informací.**

WEB SCRAPING – VÝHODY

- ❖ **Rychlost a efektivita** – sběr velkého množství dat z různých webů vyžaduje trpělivost a spoustu práce, ale díky automatizaci se tento proces mnohonásobně zjednodušil.
- ❖ **Všestrannost** – web scraping lze využít v různých odvětvích a pro různé účely. Například e-shopy jej využívají k porovnávání cen a sledování dostupnosti produktů. Marketingové firmy pak mohou snadno shromažďovat a studovat data o chování uživatelů.
- ❖ **Web scraping je legální** – musí se však provádět s respektem k soukromí uživatelů i autorským právům. Lidé pracující na web scrapingu by tak měli ideálně dodržovat veškeré i etické zásady.

WEB SCRAPING – NÁSTROJE

- ❖ **APIFY** (apify.com)
- ❖ **BROWSE.AI** (<https://www.browse.ai>)
- ❖ **OUTSCRAPER** (<https://outscraper.com>)

Apify Store

Search for Actors

Social media

AI

Lead generation

E-commerce

SEO tools

Jobs

News

Real estate

Developer tools

Travel

Videos

Automation

Integrations

Open source

Other

All Actors

Explore 3,000+ pre-built Actors for your web scraping and automation projects.

View all →



Website Content Crawler

apify/website-content-crawler

Crawl websites and extract text content to feed AI models, LLM applications, vector databases, or RAG pipelines. The Actor supports rich formatting using Markdown, cleans t...



Apify

35.8k ☆ 994



Instagram Scraper

apify/instagram-scraper

Scrape and download Instagram posts, profiles, places, hashtags, photos, and comments. Get data from Instagram using one or more Instagram URLs or search queries...



Apify

72.8k ☆ 972



Google Maps Email Extractor

lukaskrivka/google-maps-with-contact-details

Extract Google Maps contact details. Scrape websites of Google Maps places for contact details and get email addresses, website, location, address, zipcode, phone...



Lukáš Krivka

14.9k ☆ 469



Facebook Posts Scraper

apify/facebook-posts-scraper

Extract data from hundreds of Facebook posts from one or multiple Facebook pages and profiles. Get post URL, post text, page or profile URL, timestamp, number of likes,...



Apify

18k ☆ 278



Contact Details Scraper

vdrmta/contact-info-scraper

Free email extractor and lead scraper to extract and download emails, phone numbers, Facebook, Twitter, LinkedIn, and Instagram profiles from any website. Extrac...



Vojta Drmota

27.8k ☆ 311



TikTok Data Extractor

clockworks/free-tiktok-scraper

Extract data about videos, users, and channels based on hashtags or scrape full user profiles including posts, total likes, name, nickname, numbers of comments, shares,...



Clockworks

20.9k ☆ 355

Category

Added Recently ✨

- Automations
- Competitive Intelligence
- Digital Publishing
- E-Commerce monitoring
- Education
- Entertainment
- Intellectual Property
- Jobs/Hiring
- News
- Real Estate
- Recruiting
- Search Engine
- SEO
- Social Media Marketing
- Technology



Extract Stories and More from a Medium Topic

Gather the latest stories, key influencers, and top recommendations from any topic page on Medium. Ideal for content aggregation, market research, and personalized recommendations across diverse subjects.



Monitor Company Info on LinkedIn

Monitor a company's details on LinkedIn and get an alert when their details or employees list change.



Extract Products from ProductHunt

Scrape the list of products on Product Hunt for a specific keyword and/or category. Download it as a spreadsheet or monitor it for changes over time.



Extract Jobs Listings from Indeed US

Extract a list of all jobs posted on Indeed in the US for a specific role and city or state.

Site

- Airbnb
- Amazon
- Apartments.com
- AppSumo
- Audible
- Bark
- BBB
- Best Buy
- Binance
- Bing



Extract job listings information from LinkedIn

With this Prebuilt Robot, you can extract job listings from LinkedIn for a job position & location



Extract Apps list from Google Workspace Marketplace

Extract Apps list from Google Workspace Marketplace



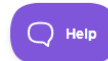
Extract Apps list from Zapier

Scrape the list of Apps from Zapier



Extract App Details from Zapier

With this prebuilt robot you can extract an App details from Zapier





Booking Search Scaper


[Pricing](#) | [API Docs](#) | [Product Page](#)

Returns search results


* Search URLs

or select a CSV/XLSX/TXT/Parquet file 



Limit per one query 

> Other parameters (result format, task tags)

Once you select the parameters, you can start the task by clicking this button. 

• •

WEB SCRAPING – POUŽITÉ ZDROJE

<https://coderslab.cz/cz/blog/co-je-webscraping>



**SLEZSKÁ
UNIVERZITA**
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVINĚ



DĚKUJI ZA POZORNOST

www.slu.cz