

MANOJ KARKI

Senior Web Scraping Engineer • Python Automation Specialist • Data Extraction Architect

Email: karkimanoj2020@gmail.com | **WhatsApp:** +977 9847059875 | **Location:** Kathmandu, Nepal (Remote)
Website: manojkarki2020.com.np | **LinkedIn:** linkedin.com/in/karkimanoj2020 | **Apify:** apify.com/devil_port369-owner

PROFESSIONAL SUMMARY

Senior Web Scraping Engineer with 5+ years of experience building large-scale data extraction systems, automation pipelines, and production-grade scraping infrastructure for international clients. Specialized in Python-based scraping, browser automation, anti-bot bypassing, distributed scraping architecture, and real-time data pipelines. Proven track record operating 250+ concurrent scrapers on AWS infrastructure while processing millions of records daily with high reliability and uptime. Strong expertise in Scrapy, Playwright, Selenium, APIs, MongoDB, cloud deployment, and structured data engineering for AI and analytics workflows.

TECHNICAL SKILLS

Programming	Python, SQL, JavaScript
Web Scraping	Scrapy, Playwright, Selenium, BeautifulSoup4, Requests
Cloud & Infrastructure	AWS EC2, Lambda, Docker, Apify, n8n
Databases	MongoDB, PostgreSQL, NoSQL, Pandas
Protocols & APIs	REST APIs, JSON/XML Parsing, HTTP/S, WebSockets
Anti-Bot Expertise	Fingerprint Evasion, CAPTCHA Handling, Rotating Proxies
AI & Automation	OpenAI API Integrations, Lead Scoring Pipelines

PROFESSIONAL EXPERIENCE

Lead Technical Consultant | Cosmofeed | Remote [Contract] Sep 2022 – Feb 2026

- Led technical strategy for large-scale automation systems and data-driven product initiatives on a contract basis.
- Architected modular scraping and workflow systems using Python, Apify, AWS, and n8n.
- Collaborated with cross-functional teams to define data requirements and optimize extraction pipelines.
- Built scalable and reusable scraping frameworks focused on reliability and maintainability.

Senior Web Scraping Engineer | EntTelligence | USA / Remote Aug 2023 – Nov 2025

- Designed and maintained a distributed scraping infrastructure extracting real-time cinema pricing and seating intelligence globally.
- Managed 250+ concurrent AWS-hosted scrapers with 99.9% uptime and millions of records processed daily.
- Implemented advanced anti-bot bypass systems including browser fingerprinting, CAPTCHA handling, and dynamic rendering.
- Improved pipeline stability, scalability, and throughput through modular scraper architecture and automation.
- Delivered validated and structured datasets for analytics, reporting, and business intelligence teams.

Data Scraper & Automation Engineer | Freelance Consultancy | Remote 2021 – 2023

- Built production-grade Apify Actors and automation systems for international clients.
- Developed AI-powered lead generation and community intelligence pipelines.
- Integrated OpenAI APIs into automated prospect scoring and data enrichment workflows.
- Delivered clean JSON/XML structured outputs for downstream analytics systems and dashboards.

KEY PROJECTS

- **Global Cinema Intelligence Engine:** Built a real-time monitoring system tracking movie ticket pricing and seat availability across thousands of global cinema locations using distributed AWS-hosted scrapers.
- **FreshForum AI Pro:** Designed a Reddit/forum intelligence pipeline extracting structured market trend and AI community data for competitive analysis.
- **InstaLeads AI Pro:** Developed a large-scale Instagram lead extraction and AI-powered prospect qualification system.
- **Apify Actor Portfolio:** Created and maintained multiple production-grade public and private Apify Actors for enterprise clients.

EDUCATION

BSc (Hons) Computing and Software Development

The British College, Kathmandu — Affiliated with Leeds Beckett University, UK
2018 – 2021

CORE STRENGTHS

- Large-scale static and dynamic web scraping
- Production-grade automation pipeline development
- Distributed scraper architecture and cloud deployment
- Anti-bot protection bypassing and browser automation
- Data cleaning, validation, and structured transformation
- API integrations and real-time data extraction
- Remote collaboration with international teams and clients