

Alonso Gonzalo Valdivia Quispe

Senior Web Scraping Engineer | Anti-Bot Systems | Distributed Infrastructure

alonso@algova.dev | +51 950326992 | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

Based in Europe | Open to global relocation

PROFESSIONAL SUMMARY

Web scraping engineer and infrastructure architect with 7+ years building high-load data extraction platforms at scale. Expert in anti-bot bypass (Cloudflare, Akamai), browser fingerprinting (CDP, TLS/JA3), CAPTCHA solving, and proxy management. Architected Estela, an open-source distributed scraper orchestration platform deployed on Kubernetes across AWS, GCP, and DigitalOcean. Built 35+ production scrapers for e-commerce, airlines, and government sites. Currently operating as independent consultant delivering web scraping, data extraction, and Data as a Service to international clients, and co-founding Deeger.io for AI agent web infrastructure.

WORK EXPERIENCE

Independent Consultant | algova.dev

Remote — Europe | Apr 2025 – Present

- Providing web scraping, data extraction, automation, and Data as a Service to international clients.
- Green Badger (USA): Extracted ~50,000 PDF documents from 2 certification websites, delivering structured data for product compliance tracking.
- Transsa (Chile): Data as a Service for real estate portals — Portal Inmobiliario and Assetplan.
- CityTech: Estela platform installation, spider development, and infrastructure maintenance.
- Built and manage scraping infrastructure independently using Claude Code and Cursor as primary development tools.

Co-founder & Engineer | Deeger.io

Ireland (Remote) | 2024 – Present

- Co-founded Deeger.io with Breno Colom — building open-source infrastructure for AI agent interactions on the web.
- Designed WAIS (Web Agent Interaction Standard): specification for autonomous agent discovery, authentication (SD-JWT, DPoP), and structured interactions within the MCP ecosystem.
- Built SerpHub (serphub.algova.dev): Google Search results API as a WAIS-compatible reference implementation.
- Developed Google Maps and Shein data extraction services on the Deeger platform.

Software Engineer → Tech Lead → Solution Architect → Project Manager | Bitmaker

Arequipa, Peru | Jan 2018 – Apr 2025 (40% equity co-founder)

- Built 35+ production scrapers for e-commerce (price monitoring, product catalogs), airline package tracking, and government data extraction using Scrapy, Selenium, and Playwright.
- Anti-bot bypass: Cloudflare (curl_cffi with TLS fingerprint rotation), Akamai, JA3/TLS fingerprinting, HTTP/2, and browser-level detection evasion.
- CAPTCHA solving: Integrated Capsolver pipelines for reCAPTCHA, hCaptcha, and image-based challenges at scale.
- Proxy management: Designed rotation strategies with Rayobyte and Geonode — geo-targeted residential proxies, session persistence, error-rate monitoring.
- Architected Estela (github.com/bitmakerla/estela): open-source distributed scraper orchestration platform on Kubernetes with Kafka, Elasticsearch, Redis, deployed across AWS, GCP, DigitalOcean via Terraform.
- Built Spider Factory Agent: AI-powered tool orchestrating LLM calls (Claude, GPT-4) to auto-generate Scrapy spiders from network traffic analysis.
- Automated CI/CD with GitHub Actions, GitLab CI/CD; containerized all services with Docker/Kubernetes/Helm.
- Reduced infrastructure costs by 50% migrating Estela from AWS to DigitalOcean.
- Led Professional Services unit turnaround: deficit to +\$60,000 profitability managing 15 simultaneous projects.
- Mentored team of 5 engineers, conducted code reviews, established quality standards.

Cloud Infrastructure & DevOps Manager | LinkBusters

Amsterdam, Netherlands (Remote) | Jan 2022 – Jan 2023

- Managed AWS infrastructure: 60 EC2 servers, 18 ECS clusters, achieving 20% cost reduction through optimization.
- Administered PostgreSQL databases (128GB RAM, 15TB storage), executing critical migrations with minimal downtime.
- Built inter-service communication using SQS and RabbitMQ; integrated CI/CD with CircleCI, AWS CodeDeploy.
- Implemented monitoring, alerting, and observability across distributed infrastructure.

Full-Stack Developer | June

San Francisco, USA (Remote) | Oct 2024 – Mar 2025

- Built the MVP for June (tryjune.ai), integrating AI models via Baseten for real-time inference in a React/Next.js application with Python backend.
- Managed deployment on Vercel, shipped continuously in a fast-paced startup environment.

KEY PROJECTS

Estela (Open Source) — Distributed scraper orchestration platform. Kubernetes scheduling, Kafka event processing, Elasticsearch indexing, multi-cloud deployment (AWS/GCP/DO). github.com/bitmakerla/estela

Spider Factory Agent — AI-powered tool orchestrating LLM calls (Claude, GPT-4) to auto-generate Scrapy spiders by analyzing network traffic with Playwright. Full agent lifecycle: traffic analysis, code generation, validation, deployment.

WAIS — Web Agent Interaction Standard — Specification for AI agent discovery, authentication (SD-JWT, DPoP), and structured interactions. deeger.io/wais

SerpHub — Google Search results API built on Deeger platform. Real-time SERP data extraction. serphub.algova.dev

TECHNICAL SKILLS

Scraping: Scrapy, Selenium, Playwright, curl_cffi, Capsolver, Rayobyte, Geonode, headless Chromium

Anti-Bot: Cloudflare bypass, TLS/JA3 fingerprinting, CDP protocol, HTTP/2, browser stealth plugins

Languages: Python (async, 7+ years), Go (familiar), TypeScript, JavaScript, Bash

Backend: FastAPI, Django, Flask, Node.js, Celery

Infrastructure: Kubernetes, Docker, Helm, Terraform, Minikube

CI/CD: GitHub Actions, GitLab CI/CD, CircleCI, AWS CodeDeploy/CodeBuild

Messaging: Kafka, RabbitMQ, AWS SQS, Redis

Databases: PostgreSQL (15TB scale), MongoDB, MySQL, Redis, Elasticsearch

Cloud (AWS): EC2, ECS, Lambda, S3, SQS, RDS, Route53, CloudFront, IAM, OpenSearch

Cloud (GCP): GKE, Cloud Storage, BigQuery, Artifact Registry, Cloud Build

AI/ML: LLM API integration (Claude, GPT-4), PyTorch, Computer Vision, CNNs, GANs

Monitoring: Prometheus, Grafana, CloudWatch, observability at scale

EDUCATION

M.S. Convolutional & Adversarial Neural Networks | 2023

Focus: ResNet architectures, GANs, Computer Vision

B.S. Computer Science | San Pablo Catholic University, Arequipa | 2018

Graduated first in class (Valedictorian), GPA 4.0/4.0

ADDITIONAL

Languages: English (Fluent), Spanish (Native)

Co-founder: Deeger.io — Open-source infrastructure for the agentic web (Ireland)