Adrian Szczepański

ML / AI Engineer · Python · Apple Silicon & Edge AI

Warsaw area, PL | hi@aszc.dev | aszc.dev | github.com/aszc-dev | linkedin.com/in/aszc

Software engineer with almost 10 years of experience and a deep Python core, specializing in machine-learning acceleration on Apple Silicon and real-time edge AI. Maintainer of an open-source toolkit that runs Stable Diffusion through Core ML on the Apple Neural Engine. Comfortable across the stack: production AWS, backend APIs, LLM integration, and low-latency on-device inference. I ship measurable results and document what I learn.

Technical Skills

Languages	Python (primary, 8+ yrs production), ClojureScript / Clojure, SQL
ML / AI	Core ML, Apple Neural Engine, MPS, MLX, PyTorch, YOLO, LLM integration (LangChain, OpenAI / Anthropic APIs), ComfyUI
Edge / Real-time	Raspberry Pi, Hailo, NVIDIA Isaac Sim, low-latency inference pipelines
Backend	FastAPI, Flask, Django, REST APIs, SQLAlchemy, Pydantic
Cloud / Infra	AWS (Lambda, ECS, EC2, S3), Docker, Nginx, Coolify, RunPod
Distributed	OpenStack, RabbitMQ, Zookeeper, HA test environments
Tooling	Git, GitLab CI/CD, modern Python tooling (pytest, ruff, uv)

Experience

aszc.dev — Founder & Principal Engineer 2022 — Present

Independent Consulting *Otwock, PL · Remote*

- Independent engineering practice: I take on ML, AI, and production-cloud work as a contractor, and build and ship my own products under the aszc.dev brand.
- Author and maintainer of **ComfyUI-CoreMLSuite** (open source): custom nodes that run Stable Diffusion through Core ML on the Apple Neural Engine, delivering 1.5–2× faster image generation on an M2 Pro versus the standard backend.
- Built and shipped **edgevox.app**: a commercial product generating text-to-speech voice packs for EdgeTX RC firmware, with self-hosted inference and payment integration (Stripe / Paddle).
- Publish reproducible engineering benchmarks (ANE vs. GPU/MPS energy and fidelity studies for Core ML diffusion) at log.aszc.dev.

Python & Integration Lead — Autonomous Drones (Contract) Oct 2025 — Jan 2026

Real-time on-device inference *Remote*

- Python and integration lead on a multidisciplinary drone team: the engineer responsible for joining flight control, perception, and AI components into one coherent real-time software/hardware pipeline.
- Engineered a low-latency image-recognition pipeline on **Raspberry Pi + Hailo** accelerator, tuned to minimize lag, frame drops, and buffer overflow for closed-loop flight control.
- Designed training scenes in **NVIDIA Isaac Sim**, integrated multiple drone communication protocols in simulation, and wrote a minimal custom simulator to validate control logic before hardware deployment.
- Drove fine-tuning of YOLO-based detection models for precise feature localization on high-resolution imagery, after evaluating several classical numerical approaches.

Software Developer Nov 2024 — Sep 2025
Xfaang Sp. z o.o. *Warsaw, PL · Remote*

- Enhanced a Flask-based API and built supporting service clients to enable data exchange and automate business workflows.
- Deployed and hosted an AI model for **emotion recognition from video**, exposed through a web application.
- Built an **audio analysis** component for voice activity detection (VAD).
- Worked hands-on across AWS in production (Lambda, ECS, EC2): maintaining web applications, running Lambda data-aggregation jobs, and handling domain configuration.

Software Developer Dec 2022 — Feb 2023
Igoria Trade S.A. *Warsaw, PL*

- Built a Python automation service with strong resilience guarantees (race-condition safety, high availability); managed servers, databases, and deployments.

Software Engineer Dec 2017 — Apr 2021
CodiLime sp. z o.o. *Warsaw, PL*

- Developed an OpenContrail / Tungsten Fabric Neutron plugin; repeatedly stood up minimal high-availability OpenStack test environments (multi-instance Neutron wired to a central Contrail node running Zookeeper and RabbitMQ) to validate the plugin under HA conditions.
- Led migration of enterprise-grade networking systems from Python 2 to Python 3, ensuring backward compatibility and performance.
- Designed and maintained backend solutions; collaborated with global QA teams on E2E integration testing and stress-test analysis. Built in-house CI/CD automation tooling and mentored junior engineers.

Selected Projects

ComfyUI-CoreMLSuite — Open source · github.com/aszc-dev/ComfyUI-CoreMLSuite
One of the few tools running Stable Diffusion through Core ML on the Apple Neural Engine. 1.5–2× faster generation on M2 Pro.

CodeAid — Django + LangChain
An AI Python tutor that evaluates student code and returns Socratic guidance. LLM orchestration via LangChain (LLMChain + PromptTemplate, OpenAI) with a sandboxed external code-execution API (Judge0).

edgevox.app — Commercial product
Text-to-speech voice packs for EdgeTX RC firmware; self-hosted inference, Stripe / Paddle payments.

Education

Engineer (inż.) in Telecommunications 2012 — 2017
Warsaw University of Technology

Languages: Polish (native), English (B2/FCE, professional working proficiency).