

Krzysztof Kutniowski

AI Engineer (LLM, AI agents, RAG, MCP) · Full-stack Python / TypeScript · Process automation

Zabrze, Poland krzysztofkutniowski@gmail.com +48 572 042 069
github.com/GenorTG linkedin.com/in/krzysztofkutniowski
Portfolio: smart-samurai.pl Smart Knife: smart-knife.smart-samurai.pl
Finetune Studio (GitHub Pages): genortg.github.io/finetune-studio

PROFILE

AI engineer and full-stack developer with 8+ years of commercial software experience, the last 2 years building production AI systems. I design and ship LLM-based solutions: AI agents with a custom agent loop and tool calling, RAG pipelines, MCP integrations and n8n process automation — from prototype to a working deployment on my own infrastructure (Linux, Docker, PostgreSQL/PostGIS, Tailscale). As Lead IT Developer and IT Team Leader at Rentiers P.S.A. I build the company's systems and bring AI into other departments' work. I maintain public projects: Finetune Studio (RAG, LoRA/QLoRA fine-tuning, model evaluation, 2,725 automated tests) and Smart Knife (25 tools, an AI assistant that acts on each of them, an MCP server I built with 113 tools for external agents, a public instance online), and I run the open-source OpenClaw, configured by hand, on my own server — a hub that manages CLI agents, coding apps and automations. I code with AI agents (Claude Code, OpenAI, OpenCode) and own the quality of what ships: architecture, code review and tests.

EXPERIENCE

Owner · AI Systems Engineer, Krzysztof Kutniowski Smart Samurai — 11.2024 — present

Own B2B company

- Design and deploy local and cloud AI systems and agentic automation for business processes.
- Combine LLM inference, RAG, fine-tuning, and image, video and voice generation.

Lead IT Developer · IT Team Leader (previously Web Developer / IT Specialist), Rentiers P.S.A. — 05.2024 — present

formerly ASI Deweloperskie / Deweloperskie P.S.A.

- Designed a location-analysis AI agent: custom agent loop, tool calling (OpenAI, Anthropic), 20 tools, risk-flag scoring.
- Spatial data pipeline: 2M+ OpenStreetMap objects imported, PostgreSQL/PostGIS in Docker, 5/10/15-minute drive-time isochrones and Excel reports.

Senior Interactive Courses Developer · Team Lead, Weblink Software — 02.2018 — 05.2024

- Led a project team of more than 10 people, held frequent client meetings and translated requirements into technical solutions.
- Trained about 90% of the development team (30+ developers at its peak) and onboarded new hires.

SELECTED PROJECTS

Finetune Studio — genortg.github.io/finetune-studio

Public GitHub project and my main engineering lab: a self-hosted WebUI (Python, FastAPI) for RAG corpus creation, LoRA/QLoRA fine-tuning, local inference, model evaluation and benchmarks, and GGUF export. Project presentation: genortg.github.io/finetune-studio; code: github.com/GenorTG/finetune-studio. The app itself has no public live instance — it runs locally.

Python, FastAPI, Jinja2, LoRA, QLoRA, RAG, SentenceTransformers, Tesseract OCR, MCP, GGUF, CUDA

- Document RAG pipeline: multi-format ingestion with Tesseract OCR, configurable chunking, SentenceTransformers embeddings and live progress over SSE.
- Exports a corpus as a self-installing package serving search over MCP and HTTP API — fully offline, with bundled models.
- LoRA/QLoRA fine-tuning with live loss, gradient and learning-rate tracking, adapter merging and GGUF export.
- 981 commits in 7 weeks, CI on GitHub Actions and a suite of 2,725 automated tests.

Smart Knife

smart-knife.smart-samurai.pl

Smart Knife: my online service, a Polish-first toolbox with an AI assistant, live with no install at smart-knife.smart-samurai.pl; systemd releases with auto-update and rollback.

Next.js, TypeScript, React, Tailwind CSS, MCP, Python, Playwright, Docker

- AI assistant acts on every tool: several chats at once, durable tasks, vision, trust modes and confirmation of destructive steps.
- 25 tools, incl. CV studio (per-job versions, ATS, application tracker), Website studio (audits, GitHub PRs, Vercel), office, workflows, 13 Craft apps, SQL, PDF OCR.
- Built Smart Knife's MCP server: 113 tools through which external agents (Claude Code, Cursor, OpenClaw) work on a user's data, with a personal URL per account and read-only keys.
- Sandboxed Python, SSRF guards, sealed per-account files; 105 unit test files, 73 e2e suites (62 Playwright), a security probe.

OpenClaw

My own OpenClaw deployment: OpenClaw is an open-source (MIT) agent assistant by the OpenClaw Foundation, not my code — I configured it by hand for my whole stack and use it daily at work. The private layer on my always-on home server ties CLI agents and coding apps (Claude Code, OpenCode, Cursor), local and cloud models and automations into one Tailscale-linked network. One shared personal config — personas, prompts, skills and plugins — carries across projects, and the agents work in loops: collect data, monitor, deploy, test and repair, then report back.

AI agents / agentic workflows, Claude Code, OpenCode, Cursor, n8n, LLM inference, llama.cpp, Ollama, Linux, Docker, Tailscale, self-hosting

- One place to manage agents: CLI agents and coding apps (Claude Code, OpenCode, Cursor) share configuration, prompts, skills and plugins.
- Always-on home server plus several Tailscale-linked machines.
- Jobs go to local models (llama.cpp, Ollama, LM Studio) or cloud models (OpenRouter, OpenAI, Anthropic).
- Trigger-, schedule- and webhook-driven automation (including n8n): scraping, monitoring, deploys, tests and repair without manual step-by-step runs.
- Agents reach my self-hosted services, data and network devices inside the private Tailscale network; nothing is sent outside it.

SKILLS

Web and applications: TypeScript, JavaScript, React, Next.js, FastAPI, REST API, JSON, Server-Sent Events, OAuth (Google), D3 (d3-geo), HTML/CSS, Tailwind CSS, Playwright, pytest, mobile apps, Git, SEO

AI / ML: Python, LLM inference, AI agents / agentic workflows, tool calling, MCP (Model Context Protocol), OpenClaw (agent management hub), RAG, embeddings, OpenAI API, Anthropic API (Claude), LoRA / QLoRA fine-tuning, LLM evaluation & benchmarks, n8n, llama.cpp, Ollama, LM Studio, GGUF, ComfyUI, TTS, Claude Code, Cursor, OpenCode

Infrastructure and hardware: Linux, Docker / Docker Compose, PostgreSQL / PostGIS (Docker), SQLite, self-hosting, Tailscale, reverse proxy, DNS, tunneling, Nextcloud, Home Assistant, CI/CD (GitHub Actions), PM2, cloud deployment, custom PC/server builds, benchmarking

Other strengths: system architecture, automation, location analytics, report generation, mentoring, leadership, client communication

EDUCATION

Liceum Ogólnokształcące nr 4 w Tychach

General secondary education. Udemy course: Complete Blender Creator — Learn 3D Modelling for Beginners (11.2021). Continuously upskilling: Python, backend, LLMs and inference, model training (Boot.dev).

LANGUAGES

Polish — native English — fluent, near-native (C2)

HOW I WORK

I code with AI agents (daily Claude Code, OpenAI and OpenCode, previously Cursor) and I’m open about it — it’s my productivity multiplier. I own what the agent delivers: I understand the code and architecture, judge the trade-offs, delegate precisely, do code review and guard quality with tests and CI, so the result isn’t low-quality slop.

I consent to the processing of my personal data contained in this document for the purpose of the recruitment process, in accordance with Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 (GDPR).