

Oleg Kossoy

Software Architect — AI/ML & agentic systems · almost 40 years in software, 15 at SAP

Netanya, Israel · +972-52-477-6277 · oleg@kossoy.com · linkedin.com/in/kossoy · github.com/kossoy · cv.oleg.kossoy.com

Software Architect (Principal Engineer at SAP) shipping software since 1987 — from OS/2 system programming to autonomous AI-agent fleets. Today: production ML for expense-fraud detection at SAP Concur (Python, AWS EKS/SQS/DynamoDB, ONNX) plus applied LLM/agentic engineering — I design and operate a fleet of autonomous AI coding agents shipping 20–30 hands-off releases/day across 9 products, and teach AI-assisted development from prompt engineering to agent fleets. 3 SAP patents. Available for fractional/contract architect-level AI/ML roles.

AI EXPERTISE

- Agentic engineering: Maestro — Go orchestrator: issue → git worktree → PR → review gate → auto-merge; 6-agent personal fleet w/ semantic memory (RAG)
- AI-assisted development: Claude Code, Cursor, Codex, OpenCode, Antigravity CLI — skills, MCP servers, subagents, context management, agent-failure debugging
- LLM platform economics: multi-provider routing w/ cost-aware fallbacks (Anthropic, OpenAI, Google, Fireworks, Ollama); self-hosted gateway
- Classical ML in production: fraud/anomaly detection, document & line-item classification (ONNX), OCR post-processing at multi-region scale
- Prompt & context engineering: production prompts for code agents & LLM pipelines; structured output, tool-use design, eval discipline
- Teaching & enablement: author of AI-101 (ai-101.pages.dev) — EN/RU/HE course: non-programmers → automating real workflows with Claude Code + MCP

EXPERIENCE

- SAP Labs Israel — Software Architect / Principal Engineer (SAP Concur) Mar 2019 - present
—Architect and operator of production ML audit services for global Concur customers: ml-audit, fake-receipt detection, tax extraction, policy audit, line-item classification — Python, AWS, ONNX, FluxCD
—Incident ownership & RCA, multi-region production (US/EU/APJ/US Public Sector); on-call for ML-pipeline alerts
- SAP Labs — Software Architect (Innovation Center Network Israel) Jan 2016 - Mar 2019
—Owned Web & App architecture end-to-end: took the product's UI from zero to production and led its full-stack team hands-on
- SAP Labs — Senior Developer (SAP Portal) Dec 2010 - Jan 2016
—Full-stack on SAP Portal; co-invented patented real-time messaging framework
- Teleweb Services — Senior Software Engineer (e-commerce) 2006 - 2010
- USBPhone — Senior Software Engineer (VoIP) 2003 - 2006
- CreoSoft.net — PM / Sr. Engineer / Team Leader · Inopsys — Sr. Engineer, Team Leader 1999 - 2002
- Admiral Makarov National University of Shipbuilding — Systems Programmer (from half-time student job): OS/2 system programming/admin, CGI-generated HTML for internal hyperweb, AutoCAD/AutoLISP 1987 - 1998

INDEPENDENT ENGINEERING (PUBLIC, 2024-PRESENT · [GITHUB.COM/BEFEAST](https://github.com/befeast))

Maestro (Go) — AI-agent orchestrator, 9-project fleet hands-off, ~30 releases/36h · RehearseKit (rehearsekit.uk) — stem separation → DAW-ready projects · karaoke — vocals/instrumental split + synced lyrics (GPU stems, Whisper), built hands-off by the Maestro fleet · scribe-service — YouTube→transcript→summary (Whisper + LLM) · ok-gobot — Telegram AI agent, single 18 MB Go binary · ok-player (C#/Rust) · audio-DSP plugins (C++/JUICE) · AI-101 — multilingual AI course

SKILLS · LANGUAGES · EDUCATION

Stack: Python, Go, TypeScript, C++/JUICE; AWS (EKS/SQS/DynamoDB/S3), Kubernetes, Docker, FluxCD, PostgreSQL/Redis; React/Next.js

Languages: Russian (native), Ukrainian, English, Hebrew (professional) · Patents: 3 (SAP) · Education: M.Sc. Mechanics, Admiral Makarov National University of Shipbuilding, 1992

Full CV: cv.oleg.kossoy.com