Sergei Solovev

Python Backend Tech Lead | AI Engineering Platforms | Performance Engineering
Limassol, Cyprus | hello@takentui.ru | +357 95 170784 | takentui.ru | linkedin.com/in/takentui | github.com/takentui

Professional Summary

Python Backend Tech Lead with 10+ years building high-load backend and production AI systems. Specializes in AI engineering platforms, performance optimization, and hands-on technical leadership. Multiple-time winner of AI and computer vision hackathons and industry competitions; works across international teams in English.

Professional Experience

Python Backend Tech Lead, Exness – Limassol, Cyprus Mar 2022 - present

Owns technical direction and production reliability for KYC backend platforms, document intelligence, and AI-assisted software delivery.

- Built an internal **AI engineering platform and coding-agent harness spanning 50+ repositories**, with multi-repository context and AI code review; adopted by 2+ teams to accelerate delivery and reduce review effort across **100+ Jira tasks and 400+ merge requests in six months**.
- Set technical direction and held final architecture authority for a **16-person cross-functional team** (8 engineers, 4 QA, 3 analysts, 1 DevOps); led design reviews and advised 3+ engineering teams.
- Replaced a FastAPI gateway with Traefik, increasing critical API throughput from **200 to 2,500 RPS**. Added alerting and developer on-call response, reducing production incidents by **67%**.
- Designed a verification platform processing **about 300K verifications per month** and **about 100 events per second** across 10+ FastAPI services; introduced Temporal with bounded sub-workflows, idempotency, an outbox, and safe rollout and fallback.
- Owned reliability for on-prem Kubernetes services using Helm, Terraform, and ArgoCD, including SLOs, capacity planning, observability, safe rollouts and rollbacks, and postmortems.
- Built an LLM-powered document intelligence pipeline for OCR, English translation, classification, summarization, and support validation; cut processing time from **5 hours to 86 minutes** and staffing needs, saving an estimated **$945K annually**.

Team Lead, Computer Vision, ORI Team – Russia Aug 2019 - Mar 2022

Led a four-person startup team from customer discovery and architecture through computer vision models and edge deployment.

- Translated customer requirements into architecture, domain schemas, delivery plans, and estimates while coordinating a four-person engineering team.
- Took multiple industrial computer vision pilots from prototype to production, including complete installation of edge hardware and software within customer environments.
- Designed an asynchronous Python/asyncio computer vision framework achieving **100+ FPS** across **10 cameras** on one NVIDIA Jetson device.

Full-Stack Developer, Tensor – Russia Jan 2016 - Aug 2019

Backend and web development for billing and analytics systems using Django and PostgreSQL.

- Optimized complex PostgreSQL queries for high-load services handling more than **1,000 requests per second**.
- Implemented analytics capabilities for billing workflows using Django, PostgreSQL, Celery, RabbitMQ, Docker, and CI/CD.

Awards

- **Hackathon winner, Exness Brainstorm 2023:** AI-assisted customer support chat.
- **Industrial computer vision:** winning solutions for Norilsk Nickel's flotation process and a strawberry farm.
- **OMK Accelerator 2021-2022:** selected as one of the winners for industrial computer vision solutions.

Technical Skills

AI engineering: AI coding agents, coding-agent harnesses, AI code review, agent orchestration, context engineering, human-in-the-loop approval gates, AI guardrails

Technical leadership: Technical direction, system architecture, design reviews, mentoring, hiring, cross-team enablement

Backend: Python, FastAPI, Django, SQLAlchemy, asyncio, microservices, API design, automated testing

Distributed systems: Temporal, Kafka, Saga pattern, event-driven architecture, transactional outbox, idempotency

Data and messaging: PostgreSQL, SQL performance tuning, Redis, RabbitMQ

Infrastructure: On-prem Kubernetes, Helm, Terraform, ArgoCD, Docker, Traefik, CI/CD, SLO/SLI, Sentry, Grafana

AI and computer vision: LLM-powered applications, LangChain, OCR, Google APIs, NVIDIA Jetson

Languages: Russian (native), English (professional working proficiency)

Education

P. G. Demidov Yaroslavl State University, Specialist in Computer Security 2012 - 2018