Beniamin Levin

Hands-on Engineering Team Lead / Tech Lead | Java / Spring Boot | Kafka | Kubernetes | AI-Assisted Engineering

benlevin0213@gmail.com | Israeli citizen | linkedin.com/in/beniaminlevin | github.com/beniamin-levin

Hands-on engineering team lead: 9 years as Lead Software Engineer at Netcracker, managing up to **11 direct reports** in Europe and India while coding in **Java** and **Spring Boot**. Owned technical design and delivery of the team's 13 microservices for telecom OSS/BSS order entry, CPQ, and product catalog (~3,500 events a second, **Kafka** on **Kubernetes** in AWS, GCP, Azure) for operators such as Vodafone, Telenet, and Rakuten. Owned the team's roadmap and ran sprint and release planning, performance reviews, and mentoring; used Claude Code and Cursor daily and rolled out Cursor for AI-assisted development in the team. Independent projects: ikite.fyi, a production Go service, and FixedYet, a Python AI agent for GitHub bug reports.

Delivery record: hired **9 engineers**, 5 later promoted; incoming defects down **~75%**; triage down from **~2h to ~30min** with an AI agent adopted by 40 teams (an estimated ~1,500 engineer-hours saved a month); response times 30–88% lower in an earlier role.

CORE COMPETENCIES & TECHNICAL SKILLS

Leadership, Architecture & Delivery: Engineering Leadership, **People Management**, Hiring & Interviewing, Performance Reviews, Team Mentoring, Code Review; System Design, **Distributed Systems**, Microservices Architecture, Event-Driven Architecture, OOP, Design Patterns, SOLID Principles; L3 Production Support, Performance Tuning, Release Management, Agile/Scrum, Kanban, Sprint Planning, Jira, Story Mapping, Root-Cause Analysis

Backend: Java, Java EE, **Spring Boot 3**, Spring Framework 6, Spring MVC, Spring Data JPA, Spring Security, Spring Cloud (Gateway, Config), Hibernate, jOOQ, Tomcat, JUnit, Mockito, Maven, Gradle, Quarkus; Multithreading & Concurrency; **Microservices**; Go (Golang), Node.js, Python

APIs & Security: REST, OpenAPI/Swagger, gRPC, GraphQL; OAuth, OpenID Connect, JWT, Keycloak, RBAC

Messaging & Data: Kafka, RabbitMQ; Redis, Elasticsearch, NoSQL, Cassandra; SQL, PL/SQL, Stored Procedures, PostgreSQL, Oracle, MySQL, MongoDB

Cloud & DevOps: AWS, GCP, Azure, Oracle Cloud; **Kubernetes**, Docker, Helm, OpenShift, Terraform; **CI/CD**, Git, Jenkins, GitLab, Argo CD; Observability, Prometheus, Grafana, JMX, JFR; Linux, Bash, systemd, Nginx

AI-Assisted Engineering: Claude Code, Cursor; LLM Prompt Engineering, Agent Skill Design, RAG-Grounded Agents, MCP Agent Tooling; Repository-Aware Code Analysis, AI-Assisted Defect Triage; LLMs (Claude, GPT, Gemini, Kimi)

Frontend: JavaScript, TypeScript; React, Angular, RxJS; HTML5, CSS3

Earlier: WebLogic, GWT; C#, ASP.NET; PHP, Laravel, WordPress, WooCommerce; SOAP, XML; jQuery, Bootstrap; Android

PROFESSIONAL EXPERIENCE

Netcracker Technology, Ra'anana | **Lead Software Engineer** — **Jul 2017 – Sep 2026**

- Managed up to **11 direct reports** in Europe and India while coding ~80% of the time in Java/Spring. Owned the team's roadmap, turned new product requirements into technical designs, and ran sprint and release planning, performance reviews, mentoring, code reviews, and incident resolution. Delivered releases to **5–6 operators a year**, including end-to-end order entry for Telenet's BSS transformation (mobile, TV, fixed; residential and business) and the Rakuten and Vodafone Hungary launches.
- Hired **9 engineers** in India, including a 5-engineer CPQ pod in Pune (70+ interviews in three months); onboarded them on site during two multi-month handovers. Four were later promoted to Technical Lead and one to Senior Software Engineer.
- Owned technical design and delivery of the team's **13 microservices** for telecom OSS/BSS (order entry, CPQ, product catalog, customer information management) on Java/Spring, Quarkus, Node.js, Kafka/RabbitMQ, PostgreSQL/Oracle, Cassandra, Redis, and Elasticsearch on Kubernetes (AWS, GCP, Azure), handling **~3,500 events a second** under a 99.99% availability SLA.
- Coded REST APIs for sales orders, product characteristics, and customers, and Angular order-entry UI features and fixes.
- Diagnosed and resolved production JVM issues: a classloading deadlock in TELUS's CPQ quotation engine and full-GC pauses on Change Ownership screens; also backported the Kubernetes 1.26 HPA autoscaling configuration to CPQ 22.3.
- Decreased incoming Order Capture defects **~75% over two years** (from ~40 to ~10 a month) by fixing recurring defects at their root cause and tightening code reviews and CI/CD validation; maintained the team's Jenkins and GitLab CI pipelines.
- Shortened average defect-triage time from **~2 hours to ~30 minutes** by writing the agent skill behind Netcracker's Order Capture triage agent on Claude, Kimi, and GPT, grounded through MCP tools (Jira, GitLab, Jenkins) and RAG over docs and past tickets. Adopted by 40 teams for 1,000+ tickets a month (an estimated ~1,500 engineer-hours saved a month).
- Resolved **~2,000 L3 tickets** for **30+ operators** (Telefónica/O2 UK, TELUS, Nuuday, Charter) typically in 1–3 days, prioritizing escalations with product managers and using Python scripts to parse service logs and query Oracle and PostgreSQL data.
- Lowered unit-test creation time **~40%** by rolling out Cursor for AI-assisted development in the team's Java/Spring and Angular/React code, with repository-aware workflows for coding, refactoring, code analysis, and test generation.

Netcracker Technology, Nizhny Novgorod | **Senior Software Engineer** — **Jul 2015 – Jul 2017**

- Reduced order-entry response times **30–88%** across critical workflows (key responses from **~4s to ~0.5s**) by profiling with JFR, removing duplicate service calls, adding caching, and tuning Oracle SQL in the Java/Spring backend.

Symphony Teleca (now Harman), Nizhny Novgorod | **Principal Software Engineer** — **Feb 2011 – Jul 2015**

- Engineered the Java backend and Android client of SiriusXM's connected-car infotainment platform, shown live at CES.
- Optimized core messaging infrastructure for flagship Samsung Android devices on site at Samsung HQ in Suwon, South Korea, and built Android UI and widgets for Nokia and Savelli.

Onlyoffice (then TeamLab), Nizhny Novgorod | **Software Engineer** **Nov 2009 – Jan 2011**

- Built C# and ASP.NET modules for an enterprise collaboration and document-management platform.

MERA NN, Nizhny Novgorod | **Software Engineer** **Nov 2006 – Nov 2009**

- Built high-availability Java web applications for remote administration and monitoring of Nortel communication systems.

PERSONAL PROJECTS

FixedYet (Nebius x NVIDIA Global AI Hackathon entry) **Sep 2026 – Present**

- Built an autonomous AI agent that validates GitHub bug reports: it reproduces the failure, compares the reported release with current code, verifies fixes and backports, and opens draft PRs; its fixes and backports led to four upstream PRs, including Black (42k stars), Pydantic, and RDFLib (merged).
- Results: reproduced 11–13 issues of a 15-issue SWE-bench Verified subset per run at ~$0.03 of model cost each and gave correct verdicts on 8 of 10 recent live GitHub issues.
- How it works: NVIDIA Nemotron on Nebius Token Factory writes pytest reproductions, Docker isolates each run, and deterministic rules, not the model, decide the verdict: fixed, open, or not reproduced.
- Stack: Python 3.12, asyncio, FastAPI, Pydantic, Docker, MCP server; 1,700+ pytest tests.
- Code at gitlab.com/beniamin-levin/fixedyet; live at fixedyet.benlevin.fyi.

ikite.fyi **Feb 2023 – Present**

- Designed, built, and operate a production wind-monitoring app on Oracle Cloud (OCI) with live station readings, historical charts, beach cameras, and wind, rain, and storm alerts; rewrote it from PHP (2023) to Go in 2026.
- Built Go collectors that ingest Windguru and Israeli station data into MySQL every minute and send wind and storm alerts through Telegram and browser notifications; run them on systemd timers, with a Docker-based dev setup.
- Added live rain and storm alerts from radar and satellite data: it tracks rain-cell motion across consecutive radar images, projects it forward, and alerts through Telegram before rain reaches the beach; replayed over a month against a rain gauge ~2 km away, it warned of 7 of 10 rain spells a median of ~40 minutes ahead, and it self-tunes after each storm.

EDUCATION

Master's in Applied Informatics (Computer Science) with Honors **Sep 2003 – Jul 2008**

N. I. Lobachevsky State University of Nizhny Novgorod, Faculty of Computational Mathematics and Cybernetics

LANGUAGES

English: fluent | Hebrew: professional working proficiency | Russian: native