Viktor Tiulpin

Amsterdam, the Netherlands viktor@tiulp.in tiulp.in tiulpin tiulpin tiulpin

Profile

Senior infrastructure engineer, 7+ years across ML, developer tools and AI platforms. I run the Kubernetes, CI/CD and vLLM services behind JetBrains AI and build developer tools with 15M+ downloads for engineers and AI agents.

Experience

Senior Software / Infrastructure Engineer, AI Platform, JetBrains Mar 2024 – present

platform behind JetBrains AI: edge, service mesh, CI/CD, LLM serving and developer experience

- running the multi-region, multi-cloud Kubernetes platform (10K+ rps, 7 clusters): built and own its edge gateways, CDN and WAF, Istio service mesh, autoscaling, and Terraform for networking, databases and identity
- designed the gateway architecture other teams now build on: direct LLM routes, gateway-level CORS and weighted splits for GPU serving
- created TeamCity CLI for engineers and AI agents as a side project; it became an official product in a month, ships with the IDE plugin, migrates pipelines from other CI systems and passed 1M+ downloads
- cut p50 LLM request latency by 22% by routing inference straight to the model router; set up LLM platform latency measurement
- launched vLLM serving for self-hosted open-weight models on 80 H200 GPUs: 200B+ tokens and 8M+ requests a week
- unified CI/CD for 50+ services with shared, keyless pipelines that run 2M+ builds a year and halved merge-queue time
- on-call for the production platform; led incident response and postmortems for customer-facing outages

Software Engineer, Qodana, JetBrains July 2020 – Mar 2024

early engineer on Qodana, a then-new static code analysis and code quality platform

- built Qodana CI integrations for GitHub Actions, CircleCI and Azure Pipelines, adopted by 16K+ open-source projects
- shipped Qodana CLI (Go) and reproducible Docker images: 5M+ CLI downloads and 8M+ Docker pulls
- led a 3-person team shipping security and compliance features, including license audit and clone detection
- built a data pipeline that downloaded and analyzed 2.8M public GitHub repositories

Machine Learning Engineer, Center of Reactive Programming Feb 2019 – June 2020

built computer vision and natural language processing systems for clients

- trained object detection and text classification neural networks for client projects
- raised object detection precision by ~30% with data augmentation and transfer learning

Skills

Programming & Algorithms: Go · Python · Kotlin · TypeScript · Java · Data Structures & Algorithms · System Design
DevOps and Infrastructure: Docker · Kubernetes · Helm · Istio · Envoy · Terraform · OPA · Bazel · AWS · GCP
Observability: OpenTelemetry · Prometheus · Grafana · Loki
CI/CD and Version Control: git · GitHub Actions · Azure Pipelines · GitLab CI · CircleCI · Jenkins · TeamCity · Skaffold
AI Tooling & Developer Experience: agentic workflows · MCP · agent skills · LLM evals · CLI design · usage telemetry
Machine Learning: PyTorch · vLLM · TensorFlow · Scikit-learn · OpenCV · Hugging Face · Data Science
Databases and Message Brokers: PostgreSQL · MySQL · Redis · Kafka
Spoken Languages: English · Ukrainian · Russian · Dutch (A2)

Education

Computer Science Center, competitive admission bootcamp in Software Engineering Sept 2020 – Jan 2022

- mlcourse.ai: 8th/2360 participants · stepik: algorithms · c++ · java · python · databases · coursera: Stanford ML
- kaggle expert (ex. top 500): PANDA (13th/1010) · Bengali.AI (46th/2059) · TReNDS (31st/1047) · RSNA (48th/1345)

Saint Petersburg Polytechnic University (SPbPU), B. Sc. in Computer Science Sept 2016 – July 2020

Other Activities

volunteer @ zeilenvanvrijheid.nl · conference organizer & tutor @ log.tiulp.in · open-source contributor