

Yehor Brodskiy

Senior Developer

Kyiv, Ukraine | yehor@yeskiy.com | yeskiy.com | github.com/yeskiy | linkedin.com/in/yehor-brodskiy

Summary

I'm a Senior Developer from Kyiv, Ukraine. I take the IT problems that don't really fit a SaaS model: backends, infrastructure, streaming pipelines, embeddable forms, on-site builds - and handle them through from the initial interaction to the handover. I'm currently a Senior Developer at The Parker Avery Group, working on enterprise systems for US retail and CPG.

Experience

Senior Developer, The Parker Avery Group

Mar 2022 - Present

Enterprise systems integration for US retail and CPG: the Unified Demand Intelligence platform and Ask Udi.

- Designed and solely built an OpenAI-compatible Batch API gateway in Rust (Axum, Tokio, SQLite) in front of a self-hosted vLLM server: durable job state that survives restarts, a two-class scheduler that prioritizes synchronous traffic while never starving batch work, schema-validated responses with a bounded re-ask budget, and retry-with-backoff that honors Retry-After.
- Main author of the retrieval layer of a retrieval-augmented, natural-language-to-SQL assistant (LangGraph, LangChain, OpenAI-compatible models) for a large enterprise retail analytics platform. Built a hybrid retriever (pgvector cosine similarity plus PostgreSQL full-text search) that pulls the most relevant prior question and SQL examples into the prompt, in-database embedding generation with sentence-transformers, schema-aware resolution of free-text mentions against an embedded catalog of dimension values over a complicated multi-table snowflake schema, and confidence-driven routing to a cheaper or stronger model by retrieval score.
- Designed and built a GraphQL query gateway that stitches several internal services into one endpoint, with a separate real-time channel for live updates.
- Added first-class self-hosted inference to a multi-provider LLM pipeline (local vLLM alongside hosted OpenAI, Gemini, and Cerebras-served gpt-oss), behind one OpenAI-compatible interface with shared retry, timeout and structured-output handling.

Senior Technical Consultant, Mobilauto

2022 - 2024

I'd already moved to The Parker Avery Group when Mobilauto came back to life. I owned a lot of its parts, so I pretty much got pulled into a 2-year knowledge transfer.

Full Stack Engineer / DevOps (incl. SRE), Team Lead, Mobilauto

Jan 2020 - Mar 2022

Ran the taxi platform end to end: the Cabinet24 CRM, the Express Soft dispatch platform, Mobilauto Web, the gateway and the partner bridge. prpcow and rust-wasmpack-loader came out of this work as open source. It ended in spring 2022, with the full-scale invasion.

- Built a program set for taxi services on Kubernetes (k0s) for deployment orchestration. Node.js backend (including worker threads), Kafka for driver-point delivery, Redis for cache, history and pub/sub, PostgreSQL, React frontend, full Asterisk and A2Billing telephony integration.
- Built a CRM (client base, task board, calls, mailbox, a file cloud on MinIO and S3 on bare-metal, and an in-browser SIP phone) with full Asterisk and A2Billing integration.
- Designed low-cost infrastructure on Kubernetes (k0s) for high availability of every critical service.
- Led a team of six.

Full Stack Engineer / DevOps (incl. SRE), Qberry

Apr 2019 - Dec 2019

My first salaried job, at an outsourcing firm. Production reliability and product work.

- Customer projects (React and Vue frontends, Node.js including worker threads, MariaDB, MySQL and MongoDB), plus support for bare-metal servers including PowerDNS.
- Built gamification add-ons (Node and RabbitMQ on the backend, React on the frontend).
- Built many customer-requested environments: Docker-based, bare-metal, several cloud platforms, and Kubernetes clusters including air-gapped setups.
- Led 2 or more developers.

Full Stack Engineer / DevOps (incl. SRE), Tech Lead, Cropman.io

Jan 2019 - Apr 2019

A multi-image editor right in the browser. Built it with classmates from DAN.IT and sold it as a product.

- An online batch photo editor (layers, crop, filters, watermarks). Node and Sharp (libvips) with Firebase for auth and storage, MongoDB, and a React frontend.
- Used Firebase Cloud Functions as compute for user requests, with everything else on Docker Compose.
- Led a team of four.

Freelance, Self-taught builds

2015 - 2018

Skills

Languages

TypeScript / JavaScript, Python, Rust, Go, Java, C#

Frontend

Next.js, React, Tailwind CSS, Angular, Vue, Nuxt.js

Backend & messaging

Node, prpcow, REST, RPC, GraphQL, WebSockets, gRPC, Protobuf, Spring, Microservices

Runtimes	Bun, Node, WASM in the browser, Edge runtimes
Orchestration	Kubernetes, K8s on bare metal (Talos, k0s), ArgoCD, Devtron, Custom CSI / Operators, k3s, Docker / Podman, CI/CD automation, Terraform, Ansible
Multimedia	HLS, Streaming pipelines, nginx-vod-module
Auth & identity	OIDC / OAuth2, SSO / IdP federation, Keycloak, Zitadel
CMS	Strapi, Custom headless
Data & AI	OpenAI APIs, ClickHouse, pgvector, qdrant, Langfuse, LangFlow, LangGraph, RAG / embeddings, LangChain, Apache Spark, DuckDB
Queues & gateways	Kafka, Valkey / Redis, RabbitMQ, KubeMQ, Higress
Observability	OpenReplay, Microsoft Clarity, SigNoz, Datadog, Prometheus, Grafana
Cloud & hosting	Self-hosted infrastructure, Bare metal, Major clouds, Private, air-gapped and hybrid environments, EDGE computing
Realtime & telephony	Asterisk, A2Billing, WebRTC softphones, Custom TURN / STUN with strong auth
Databases	PostgreSQL, MySQL / MariaDB, SQL Server, Oracle DB, MongoDB, Couchbase, SQLite, Cassandra
Storage	S3 / MinIO, GCS, Azure Blob, B2, NFS / SMB, Ceph
LLM serving	OpenAI-compatible APIs (Chat and Batch), vLLM, Ollama, Quantization (w4a16, FP8), Multi-GPU serving over NVLink, Throughput and latency load-testing
Security	Network hardening, VPN-only and no-public-inbound topologies, Secrets management, Vulnerability scanning, Compliance controls

Education

Kyiv Aviation Institute (National Aviation University), Bachelor's, IoT	2021 - 2025
Faculty of Electronics, Robotics, Monitoring and IoT Technologies. Average mark 89/100. Qualification work: Network Protocol for Remote Function Calls (RPC-like) as the primary seamless connection between EDGE devices in IoT.	
Kyiv Aviation Institute (National Aviation University), Master's, Electronics, Robotics and IoT	2025 - 2026
Faculty of Air navigation, Electronics and Telecommunications.	
DAN.IT Education, Full stack course	2019 - 2021
Average mark 80/100.	

Languages

Ukrainian Native | **English** C1/C2 | **Russian** Native