Yaroslav Salimov — AI Engineer / Forward Deployed Engineer CV

Oct 20, 2026 · @Yaroslav

Yaroslav Salimov

AI Engineer / Forward Deployed Engineer — Node.js, TypeScript, LLM features, integrations with customer systems

Kraków, Poland (UTC+1/+2) · Permanent residence, no sponsorship needed · Remote · Available immediately

htcsalimyaro@gmail.com · Telegram @salimov_yr · linkedin.com/in/y-salimov
github.com/salimov-y · salimov.dev

Summary

Software engineer with nearly 5 years of commercial experience who connects products to customers' real systems and ships LLM features to production. At RepMove I build two-way integrations with customers' ERP and CRM systems (NetSuite, Epicor P21, Texada SRM, Point of Rental) and migrate enterprise data of up to 175k records per customer. At Thinkable I shipped OpenAI and Gemini features to clinicians and patients: a clinical safety screen, recommendations and a 13-language translation pipeline, and cut LLM latency from 57 s to about 5 s.

I built an MCP server for a public reporting API and work AI-native every day (Claude Code, OpenAI Codex) with project context, subagents and review guardrails, so a full two-way CRM integration takes me about a week.

Core skills

- **LLM & AI engineering:** OpenAI API (Responses, Assistants), Google Gemini, structured JSON output with tolerant parsing, prompt design, safety screening and crisis alerts, model tiering, latency and cost tuning, AI usage accounting, Model Context Protocol (MCP) servers
- **AI-native development:** Claude Code (subagents, skills, persistent memory, CLAUDE.md), OpenAI Codex (custom skills), agent-generated test suites, AI code review turned into coding guidelines

- **Customer integrations & data:** NetSuite (SuiteQL), Epicor Prophet 21, Texada SRM, Point of Rental, QuickBooks, Dodge, ConstructConnect, Zapier, Integration.app, Stripe, Revolut; Salesforce and Microsoft Dynamics migrations; idempotent dry-run imports
- **Backend:** Node.js, TypeScript, Express, REST, OpenAPI (Scalar), Zod, WebSockets, JWT
- **Cloud & data:** Google Cloud (Cloud Functions, Cloud Run, Cloud Tasks, Pub/Sub, Scheduler, Logging, Monitoring/PromQL), Firebase, AWS (EC2, S3, SES, SNS), MongoDB Atlas, Firestore, PostgreSQL, MySQL
- **Frontend & mobile:** React 18/19, Redux Toolkit, RTK Query, Ant Design, React Flow; Flutter releases; Swift and SwiftUI
- **Quality & delivery:** Jest, Playwright, k6 load testing, esbuild, Docker, GitHub Actions, nginx, pm2

Experience

Full Stack Developer (integrations & AI) — RepMove (field-sales CRM SaaS) · Remote · Feb 2026 – Present

Multi-tenant CRM for equipment dealers: about 200 Firebase Cloud Functions on Firestore and MongoDB Atlas. 180+ merged pull requests; in Jun–Aug 2026 alone 323 commits and 110 merged PRs with zero reverts, most of the time on the hardest (level 4–5) tickets.

- **Customer-system integrations:** built two-way syncs with customers' ERPs on Cloud Scheduler and Cloud Tasks — a multi-tenant Texada SRM integration from scratch (customers, contacts, quotes, reservations and contracts mapped to deals, outbound push), Epicor P21 (designed for a 154k-customer backfill), Integrated Rental, Dodge and ConstructConnect; extended NetSuite, Point of Rental, QuickBooks and Integration.app syncs.
- **Enterprise onboarding & migrations:** moved new customers in through idempotent, dry-run-first import endpoints: 47k accounts / 72k contacts / 70k opportunities from Microsoft Dynamics, 175k accounts / 181k contacts from Salesforce and Oracle (type-scoped matching keys prevented 14k wrong merges), 57k clients for another dealer, and an 8-into-1 tenant consolidation.
- **Data integrity:** found silent data loss in production syncs — an external-ID collision that had merged ~39k contacts into ~420 records, and pagination that advanced the cursor after partial fetches; traced 5,925 missing deals to a cascade delete and planned the resync that restored them.
- **LLM integration:** built an MCP server for the public reporting API (stateless Streamable HTTP on Cloud Functions, 10 tools reusing the REST endpoints' Zod schemas and OpenAPI docs, per-company gating) and used it with Claude to turn 66k

accounts into a 12.4k-row address-quality report for a customer.

- **Performance & reliability:** measured ~2,000 cold starts a day at ~10 s each and prototyped esbuild bundling that loads 2.6× faster (now the team's bundling spec); designed k6 load tests and Atlas connection limits for 197 functions; found the root cause of a production outage with Cloud Logging and PromQL.
- **AI-native delivery:** wrote the project context the agents run on (CLAUDE.md/AGENTS.md, 10 per-integration LLM docs of ~21k words, an integration blueprint), turned ~340 AI code-review findings into coding guidelines and a pre-PR checklist, and built a Playwright API test framework in 6 days whose Claude subagents read backend routes and Zod schemas, plan tests and generate them.

Full Stack Developer — GGtude (Thinkable mental-health platform) · Remote · Jun 2024 – Oct 2026

Main developer of the clinical dashboard and the backend for three iOS/Android apps: 1,300+ commits, 495 releases. Part-time; full-time Oct 2025 – Jan 2026.

- **Clinic pilot in 3 weeks:** wrote the partner API for a TMS-clinic pilot from scratch (screening intake with a payer eligibility rules engine, patient stages, clinic handoff emails, links into the clinical dashboard) and the patient-facing track player; the first clinic went live in September 2026.
- **LLM safety in production:** added a safety screen on patients' free text (keyword pass plus Gemini, risk threshold tuned after the model over-flagged ordinary text), a re-check for unscreened text and crisis detection that sends the clinic a "call today" alert; covered by ~118 automated tests.
- **LLM cost & latency:** moved the AI service to one model-tier config, parallel batches and the OpenAI priority tier, and added per-route AI usage accounting: personalization 57 s → ~5 s, thought-catcher 32 s → 8 s, about 10× cheaper per call. Migrated off retiring models to the OpenAI Responses API.
- **LLM features:** OpenAI-based module recommendations from clients' free text, AI progress summaries, an LLM risk screen on therapist notes, and a translation pipeline (Gemini translation plus an OpenAI verification pass) that grew the content from 7 to 13 languages.
- **Security & mobile:** moved every AI call out of the Flutter app behind the API and removed the OpenAI key from the app binary; upgraded the app to Flutter 3.47, iOS 27 and Android API 36 and shipped releases to TestFlight and Google Play.
- **Platform:** built the multi-tenant backend and React admin dashboard from scratch (~255 documented API routes, invite onboarding via deep links, PHQ/GAD clinical reporting, English/Hebrew RTL), per-project feature flags for a hospital pilot, and CI/CD from GitHub Actions to AWS EC2.

Full Stack MERN Developer — ArtiD (client: TaxZap, tax refunds in Ireland and the UK) · Kyiv / Remote · Mar 2023 – Sep 2025

2,700+ commits across the TaxZap backend, admin and client apps.

- **Backend from the first commit:** Node.js/Express/TypeScript on MongoDB with role- and region-based access and ~50 REST endpoints for clients, cases, documents and questionnaires; CSV imports and data migrations.
- **No-code questionnaire builder:** a visual flow editor on React Flow (~12 question types, AND/OR conditions, live preview) plus the client-side engine that renders it with validation, autosave and progress; it powers the Ireland and UK tax-return intake.
- **Tax products:** the Irish client web app end to end (sign-up with e-signature, a case tracker with ~15 statuses, PWA), ~15 UK Self Assessment sections including Open Banking transaction categorisation by tax area, HMRC 64-8 and P87 PDF forms, and a Making Tax Digital checker.
- **Payments & automation:** a Revolut Checkout paywall, invoice and agreement PDFs (PrinceXML, pdf-lib), an income tax / USC / credits calculation engine covered by Jest tests, ~22 email templates and Agenda reminder jobs; GitHub Actions and nginx deploys, Node 20 upgrade.

Full Stack Developer — DevPuzzle · Kyiv · Nov 2021 – Mar 2023

- Delivered web apps for early-stage startups, from REST APIs on MongoDB and PostgreSQL to React front ends, including a scheduling and booking SaaS (calendar, Google Calendar and Zoom, real-time notifications, a scheduling engine).
- Implemented authentication, access control, WebSocket real-time features and timezone-aware scheduling.

Own products

- **Modbook — Car Build Tracker (iOS, App Store, Jul 2026):** built solo with Claude Code; v1.0 went to App Store review 9 hours after the first commit, now 1.5 with iPad support. Swift 6, SwiftUI, SwiftData mirrored to CloudKit, ~120 tests, CI on GitHub Actions; an in-app LLM prompt turns free-form notes into a clean CSV import.
- **TickDown (iOS, App Store, Sep 2025):** countdown app with Home/Lock Screen widgets, a Live Activity, Siri shortcuts and iCloud sync; rewrote it as 2.0 in one day while keeping existing users' widgets and data.

Education & languages

- Bachelor's degree, International Business — National Aviation University, Kyiv

- Courses: Full Stack Web Development (2020–2021), IT Project Management (2021)
- English: Upper-Intermediate (B2) · Ukrainian, Russian: native