PAVEL HOLYAVKIN

Forward Deployed / Applied AI Engineer
San Diego, CA (remote; travel OK) | hire@holyavkin.com | linkedin.com/in/holyavkin | holyavkin.com
github.com/holyavkin | **Book a call:** cal.com/holyavkin/meeting (live availability)

SUMMARY

Senior engineer who puts AI systems into production at real businesses and proves they work with evals, not demos. 17+ years shipping software. Formerly Head of Client Engineering at MileIQ (1M+ users, acquired by Bending Spoons). Since October 2025 I run App Incubator full-time: fractional CTO and AI engineering for six clients in consumer apps, legal, education, construction, media, and insurance, covering agents, retrieval, workflow automation, and the reliability work that keeps them running. Upwork Expert-Vetted (top 1%), 100% Job Success. Available for a full-time role; client engagements are hand-off ready.

EXPERIENCE

App Incubator (independent studio) | Fractional CTO & AI Engineer | Oct 2025 - Present (studio since 2020) | San Diego, remote

- **Okee (family screen-time app), CTO:** built and run the full stack: SwiftUI iOS app with four Screen Time extensions, an Express/TypeScript/MongoDB API on AWS, a Next.js admin console, and the marketing site; 800+ commits. Blue/green API deploys with automatic verification and rollback, isolated staging, a beta health dashboard, PostHog analytics. Engineering runs through an AI-agent harness: Claude Code, Codex, and Hermes agents building in parallel under hooks that enforce review and test gates, with 850+ PRs merged across the iOS app, backend API, web admin, and infrastructure. Private beta on TestFlight.
- **Real-estate rehab company (construction estimating):** unattended pipeline that turns a walkthrough video into a priced scope of work and estimate. Designed a 3-tier eval (price delta, line-item coverage, scope fidelity); reached 72% scope fidelity, and used deterministic extractors to fix the model's silent omissions.
- **Law firm:** repaired production email capture on Power Automate. Reconciled mailbox, run history, and logs; rehearsed a controlled failure path; audited 742 attachment candidates with zero gaps; verified rollback and removed access at handoff. Integrated a Clio MCP connector with custom field and notification fixes.
- **Education SaaS company, fractional CTO:** production readiness for a live CRM and test platform. Built an append-only storage backup (21K objects, 7.8 GB) with a byte-verified restore drill, set up isolated staging on Vercel + Neon, and fixed a Postgres pool failure that was terminating requests.
- **Documentary studio:** media archive pipeline with an atomic job queue, crash recovery, tiered subtitle-to-Whisper transcription, and chunked cloud upload on systemd.
- **Insurance sales and regulatory research:** lead-scoring and message-drafting pipeline (Apify + Claude). Citation-aware FDA/PubMed/ClinicalTrials.gov research workspace (FastAPI, pgvector, Next.js).

MileIQ (acquired by Bending Spoons, 2025) | Apr 2022 - Sep 2025 | Remote

iOS Engineer -> Guild Lead -> Head of Mobile -> Head of Client Engineering; Chair, AI Committee

- Led the 50+ engineer client engineering org for a 1M+ user app through a 2M-line rewrite from React Native to native iOS and Android, keeping 99.9%+ crash-free.
- Increased retention 9% through A/B-tested features with Product and Design.
- As AI Committee chair, took a 60+ person company from 1-2 to 6+ hours/week saved per person with AI Fridays and cross-functional office hours.

App Incubator, 2020 - 2025 (alongside MileIQ)

- Shipped 50+ apps and automations for clients from startups to Fortune 500, including ARKit/3D work and Private Notes (on-device LLM and voice, idea to App Store in 21 days). 29 Upwork contracts.

Earlier

- AI chatbot platform (2013 - 2020): built and ran 10 production bots on Messenger and Telegram with custom backends and analytics.
- EasyWakeup (2008 - 2012): built the original sleep-cycle alarm for iOS; #1 Top Grossing in 5 countries.

SKILLS

AI: Claude/OpenAI/Gemini APIs, agents and tool calling, MCP, RAG (pgvector), eval harnesses and holdouts, prompt/context engineering, on-device LLMs, Claude Code and Codex agentic workflows with review and test gates **Engineering:** Python, TypeScript/Node, Swift, SQL, FastAPI, Express, Next.js, Postgres/Supabase/Neon, MongoDB, Vercel, AWS, Power Automate, systemd/launchd, CI/CD, blue/green deploys **Delivery:** customer discovery, scoping, stakeholder demos, runbooks and handoff, recovery drills

EDUCATION

B.S. Microelectronics & Semiconductor Devices, Far Eastern National University; Machine Learning (Stanford/Coursera)