Patrick Selamy

New York, NY | (210) 704-7370 | pselamy@gmail.com | selamy.dev | linkedin.com/in/patrickselamy | github.com/pselamy

Forward Deployed Engineer | Generative AI, Production Systems, and Business Alignment

Summary

Google Cloud GenAI FDE with prior AI, product, and platform engineering experience at Meta and Google.

Experience

Google Cloud, Forward Deployed Engineer, Generative AI — July 2026 to Present

- Returned to Google Cloud in July 2026; completed the FDE onboarding capstone and began the first customer engagement specializing in generative AI.

Meta, Software Engineer, Facebook Dating and Marketplace — Oct 2025 to May 2026

- Built a Facebook Dating preview feature that let users try the experience before committing, with daily active use and engagement as the product goals.
- Expanded Facebook Marketplace partner APIs and seller-facing listing workflows.
- Built an ML-assisted Marketplace listing experience that predicted shipping weight from images, removed manual weight entry, and enabled shipping-label generation.

Google, Software Engineer — Apr 2018 to Sep 2025

- In Google Cloud Office of the CTO, built prototypes and ran rapid experiments to test executive hypotheses about market opportunities, de-risk technical directions, and hand promising work to product teams.
- Built a Gemini-based software-engineering agent, SWE-bench evaluation harnesses, planning-vs-execution metrics, and vector-search benchmarks using ScaNN, FAISS, and pgvector.
- In Google Play, helped build Play Games Leagues, a shared leaderboard across participating titles, and built large-scale smoke-testing infrastructure for third-party APIs.
- Built test infrastructure and reusable engineering frameworks in Google Shopping; developed insider-risk systems to help prevent intellectual-property exfiltration.

Earlier: Backend, platform, cloud, and healthcare engineering at Citi, Lumeris, and ClearDATA.

Forward-Deployment Approach

Start by aligning with stakeholders on the business need and success criteria. Combine implementation with architecture and deployment judgment; make cost and performance tradeoffs explicit. Help customers integrate the technology into their workflows, and bring their feedback back to product engineering.

Independent AI Systems Portfolio

- Built an autonomous-agent fleet with task decomposition, recoverable orchestration, MCP-style tool interfaces, reusable skills, and credential boundaries for document workflows, repository changes, and CI repair.
- Instrumented agent workflows with OpenTelemetry traces, metrics, and logs across tool calls, latency, token use, and recovery behavior; built CI-backed verification and evaluation loops.
- Modeled agent identity, runtime, skills, schedules, and secrets as infrastructure modules; operated the fleet with OpenTofu, Kubernetes, Argo CD, and per-agent isolation.

Selected Public Work

- Public agent-skills: reusable agent procedures and tool-authoring standards.
- Operating Agent Fleets: a field guide to durable agent systems, verification, and bounded authority.

Technical Skills

Languages:	Python, Java, Kotlin, JavaScript/TypeScript, Hack, PHP, Bash, SQL
AI Systems:	Generative AI, AI agents, MCP-style tools, tool calling, RAG, evaluation harnesses, context engineering, vector search
Platforms:	API design, distributed systems, workflow orchestration, internal IAM, secure tool execution, PostgreSQL, Redis, Kafka
Delivery:	Google Cloud, AWS, Kubernetes, OpenTofu/Terraform, Argo CD, Docker, GitHub Actions, CI/CD, OpenTelemetry