

Rob McElvenny

Founder, Product Design, Engineer

2026-Present

Aviron Software

Senior Software Engineer - AI-Native Product Engineering, Contract Remote

- Congregation of the Mission: built a production C# and .NET Azure OpenAI integration for natural-language report generation, constraining model output to validated, executable report specifications.
- accountingNode: helped build an AI-native accounting platform from the ground up, combining RAG, ambient business context, global copilot workflows, and secure custom scripts, triggers, jobs, and pages.
- Sheds for Sale: used AI-native engineering to modernize an established Web 2.0 ecommerce and administration platform across React, Next.js, TypeScript, AWS serverless services, Cognito, Stripe, and data workflows.

C#, .NET, Azure OpenAI, RAG, TypeScript, React, Next.js, Bun, PostgreSQL, AWS, serverless

2025-Present

Hyper

Founder & AI Product Engineer - Miami, FL

- Architected Core, a local-first LLM and RAG agent runtime with durable memory, retrieval, planning, tool execution, and a typed HyperBus event fabric across 240 production modules.
- Built an evidence-first mission harness that plans, executes, validates, repairs, retries, and resumes long-running work while streaming observable tool activity to Studio and voice clients.
- Optimized Qwen3.8-27B inference on one RTX 4090 from 38.62 to 103.98 tok/s in a reproducible experimental profile using CUDA offload, quantization, Flash Attention, KV-cache tuning, and MTP speculative decoding.
- Integrated PostgreSQL, Redis, Qdrant, WebRTC, speech services, model evaluation, prompt compaction, and isolated application runtimes without loading generated code into Core PID1.

LLM, RAG, Qwen, CUDA, llama.cpp, TypeScript, Bun, PostgreSQL, Redis, Qdrant, WebRTC

2025-Present

Independent Client Work

Fractional CTO, Product & Software Development - AI Engineering

- Led product discovery through production delivery, translating informal business processes into focused workflows and AWS-hosted TypeScript, React, Bun, Python, PostgreSQL, and REST API systems.
- Designed and built Skewuo as a structured intake and analysis platform with guided capture, validation, image-quality checks, adaptive analysis, preserved evidence, and human review.
- Designed The Cobbler Guy's luxury repair experience around clearer customer intake, internal review, service decisions, quote preparation, and follow-up communication.

AI, computer vision, TypeScript, React, Bun, Python, PostgreSQL, APIs

2023-25

Defense Unicorns

Platform Engineer - Defense and Government Delivery Contexts

- Built and supported RKE2 Kubernetes platforms across 5 active development, staging, and production-style environments.
- Automated provisioning and configuration with Terraform, Ansible, Python, and Bash, reducing drift and dependence on undocumented manual procedures.
- Standardized secure delivery with GitLab CI/CD, GitOps, containerized builds, validation, and rollback; diagnosed scheduling, networking, storage, secrets, and application-health failures.

RKE2, Kubernetes, GitLab CI/CD, GitOps, Terraform, Ansible, Linux, Bash, Python

2021-23

OPSWAT

Site Reliability Engineer - Cybersecurity Environment

- Operated cybersecurity applications against SLI, SLO, and SLA commitments including a 99.95% availability guarantee while supporting up to 5 production deployments per day.
- Shortened incident diagnosis through logs, metrics, alerts, dashboards, runbooks, deployment health, and customer-impact signals.
- Turned recurring production failures into measurable reliability work through incident timelines, SLO trends, health checks, and release validation.

Linux, cloud infrastructure, CI/CD, observability, networking, incident response, SLI/SLO/SLA

2018-21

Motional

Lead DevOps Engineer - Autonomous Vehicle Engineering

- Led DevOps delivery for 20+ engineers building C and C++ systems with Bazel, Jenkins, Git, CI/CD, CMake, and Linux.
- Automated compilation, testing, packaging, artifact publication, and releases while improving caching, parallelism, traceability, and reproducibility across workstations and CI.
- Designed autonomous-vehicle data-ingestion and storage-optimization infrastructure that saved approximately \$2.3M.

Bazel, Jenkins, Git, CI/CD, CMake, Linux, Bash, Python, C, C++, build systems

Contact
Projects
AI Systems

robert.mcelvenny@gmail.com · [Website](#) · [LinkedIn](#) · [GitHub](#)

[Hyper](#) · [Skewuo](#) · [The Cobbler Guy](#)

artificial intelligence, machine learning, LLM, RAG, local inference, autonomous agents, tool execution, retrieval, vector databases, model evaluation, computer vision, voice AI, TypeScript, React, Bun, Python, Bash, C, C++, PostgreSQL, REST APIs, AWS, Kubernetes, RKE2, Docker, CI/CD, GitOps, Terraform, Ansible, Linux, observability, incident response

Credentials

AWS Certified Solutions Architect - Associate · Credential 9ZLJD4R1E2Q11GCL

Patent

Named inventor · U.S. Provisional Patent Application 62/473,720 · filed February 1, 2017

Background

English and Spanish · Self-directed study in distributed systems, artificial intelligence, cloud infrastructure, and software engineering