

JOHN CULKIN

AI Solutions Architect & Production Engineer

San Francisco Bay Area | john@culkintech.com | linkedin.com/in/johnculkin

SUMMARY

AI Solutions Architect with 15+ years building production-grade systems. Expert in deploying scalable AI applications across AWS, Azure, and GCP using serverless architectures, Infrastructure as Code, and modern DevOps practices. Specialized in taking AI from prototype to production with enterprise-grade reliability.

EXPERIENCE

Senior AI Solutions Architect | Independent Consultant | 2020 - Present

- Architected and deployed production AI chatbots using AWS Lambda, Bedrock, API Gateway, and CloudFront serving 100K+ monthly users
- Built multi-agent financial planning systems with Aurora Serverless, coordinating specialized AI agents for complex financial analysis
- Implemented full Infrastructure as Code pipelines with Terraform, reducing deployment time by 80%
- Created automated CI/CD workflows with GitHub Actions for zero-downtime deployments
- Designed cost-optimized multi-cloud strategies across AWS, Azure, and GCP

Lead Platform Engineer | Tech Startup | 2015 - 2020

- Led platform engineering team building SaaS products with Next.js, FastAPI, and modern cloud infrastructure
- Implemented observability stack (CloudWatch, LangFuse) for monitoring AI system performance and cost

- Reduced infrastructure costs by 60% through serverless architecture and intelligent caching strategies
- Built RAG systems with vector databases and custom embedding models for enterprise search

Software Engineer | Fortune 500 Tech Company | 2009 - 2015

- Developed large-scale distributed systems handling millions of daily transactions
- Contributed to cloud migration initiatives moving legacy systems to AWS
- Mentored junior engineers on cloud architecture and best practices

TECHNICAL SKILLS

Cloud Platforms: AWS (Lambda, Bedrock, SageMaker, S3, CloudFront, Aurora), Azure, GCP

AI/ML: LLMs, RAG, Multi-agent systems, Embeddings, Prompt engineering

Infrastructure: Terraform, Docker, Kubernetes, CI/CD, GitHub Actions

Languages: Python, TypeScript, JavaScript, SQL

Frameworks: FastAPI, Next.js, React, Node.js

Databases: PostgreSQL, Aurora Serverless, DynamoDB, Vector DBs

CERTIFICATIONS

- AWS Solutions Architect Professional
- AI in Production: Gen AI and Agentic AI at Scale (Ligency/Udemy)

EDUCATION

Bachelor of Science in Computer Science & Engineering

Stanford University | 2009

RECENT PROJECTS

- Digital Twin AI: Production chatbot with conversation memory, deployed serverless on AWS
- Multi-Agent Financial Planner: Coordinated AI agents for portfolio analysis and recommendations
- Healthcare AI SaaS: Full-stack application with authentication, billing, and AI-powered insights
- Cybersecurity Analyst: RAG-based system with MCP integration for security research

EOF