

Yousuf Mohammad



I take healthcare GenAI from idea to production, **owning strategy, product, architecture, engineering and delivery** end to end.

Richmond, Texas | 20+ Years | Previously Fannie Mae & AWS | Doctoral Researcher, UTHealth Houston
Claude Production AI Specialist · AWS Golden Jacket · CNCF Kubestronaut

AT A GLANCE

FORWARD-DEPLOYED GENAI

3 weeks

from idea to MVP for a Medicare Part D GenAI solution, embedded with business teams

AGENTIC AI

2×

regulated cloud adoption with CLARA, a multi-agent AI system I designed and built

ENTERPRISE GENAI

\$3M

saved per year with LLM-powered document intelligence at Fannie Mae

AI-READY PLATFORMS

\$67M+

annual cloud spend made AI-ready across 160+ healthcare migrations

FOCUS Forward-Deployed AI · Generative & Agentic AI · Healthcare & Regulated Industries · AWS & Azure

Perspective

“ Most AI never makes it past the demo. It stalls in the gaps between strategy, product, architecture, engineering and operations, where every handoff loses context and no one owns the outcome end to end. ”

Selected Work

UNITEDHEALTH GROUP

GenAI for Medicare Part D

CHALLENGE

Identify the highest-impact drugs to inform the Medicare Part D bidding process.

APPROACH

Embedded with business teams, owning strategy, product, architecture and hands-on engineering of an evidence-based GenAI solution.

3 weeks

from idea to MVP; now being taken to production

UNITEDHEALTH GROUP

CLARA: Agentic Cloud Governance

CHALLENGE

Architecture governance was slowing regulated healthcare workloads on their way to the cloud.

APPROACH

Designed and built a multi-agent AI system that automates architecture reviews and risk scoring for regulated workloads.

STACK Agentic AI · Multi-Agent Systems · Automated Risk Scoring

2×

regulated cloud adoption



GenAI Document Intelligence

CHALLENGE

Make a large estate of lending documents faster to find and review.

APPROACH

Re-architected document management as a plugin-based platform with metadata tagging, search and analytics, and LLM-powered summarization.

STACK OpenAI API · Amazon S3 Tables · Amazon Athena · Plugin-Based CMS

\$3M

saved per year, with 60% faster retrieval

Enterprise Kubernetes Platform

CHALLENGE

Standardize and speed up how enterprise teams build and deploy containerized workloads.

APPROACH

Built a global, enterprise Kubernetes framework, backed by reference architectures and proofs of concept.

STACK Kubernetes · AWS · Reference Architectures

30%

faster deployment cycles

Experience

UnitedHealth Group

Sep 2025 – Present
Minnetonka, MN

Forward Deployed AI Lead · Jul 2026 – Present

Mandate expanded to AI strategy, product and delivery

- Led the team that took a Medicare Part D GenAI solution from **idea to MVP in 3 weeks**; now driving it to production (see [Selected Work](#)).
- Embedded with business teams, owning strategy, product, architecture, hands-on engineering and delivery end to end.
- Claude Code Ambassador Program: certified Claude Production AI Specialist; helped 20+ colleagues put AI to work.

Principal Architect, Medicare & Retirement · Sep 2025 – Jul 2026

- Pioneered CLARA, a multi-agent AI system for architecture governance and risk scoring that helped double regulated cloud adoption from **25% to 50%**.
- Led the transformation of legacy healthcare platforms into AI-ready, cloud-native architectures: **\$67M+** annual spend, **160+ migrations**.
- Set technical direction for **11+ architects and 80+ engineers** across multiple vendors; standardized Cloud & GenAI accelerators and enterprise Kubernetes platforms.

STACK GenAI · Agentic AI · Azure · AWS · Kubernetes · HIPAA-Aligned Architecture

Fannie Mae

Aug 2023 – Jul 2025
Reston, VA

Principal, Enterprise Architecture & Technology · Director Level

- Brought GenAI into document management with OpenAI-powered summarization and metadata extraction (see [Selected Work](#)).
- Designed a domain-driven, API-first microservices architecture for the lending platform, reducing processing fees by 0.3%.
- Replaced manual, PDF-based lender risk assessments with an event-driven, configuration-driven React forms engine that achieved **99%** approval ratings.
- Built a self-service analytics platform on Redshift APIs with Lambda-based pipelines, cutting development cycles by **40%**.

STACK OpenAI API · AWS Lambda · Redshift · Athena · S3 Tables · React · DDD · Event-Driven Architecture

Amazon Web Services

Apr 2021 – May 2023
Houston, TX

Solutions Architect

- Drove **\$40M ARR** (65% YoY) across North America, the UK and Germany, aligning Fortune 500 C-suites on cloud strategy through executive briefings.
- Delivered **30%** cost savings and faster disaster recovery through Well-Architected Framework reviews.
- Built an enterprise Kubernetes platform and a React micro-apps framework that cut duplicate development by **40%**.
- Inducted into the AWS Technical Field Community; trained 300+ engineers through workshops and game days.

STACK AWS · Kubernetes · Well-Architected · Micro Frontends · FinOps

Earlier Career

2002 – 2023

Principal Cloud Architect at Kessel Run (US Air Force); Solutions Architect at Corteva Agriscience; and Software Engineer to Technical Architect at Schlumberger, Shell, JPMorgan Chase, Wells Fargo and others, building secure, polyglot systems (Java, C#, Python, JavaScript) for regulated industries.

Doctoral Research

DOCTOR OF HEALTH
INFORMATICS
UTHEALTH HOUSTON

I've spent years on the systems behind care. Now I'm taking that rigor to the point of care.

NephroIQ: Nephrology Guidelines in Everyday Practice

NephroIQ is a platform for AI agents that support nephrologists and their patients. Its first agent, **Filtra**, prepares a one-page, evidence-backed summary before each chronic kidney disease (CKD) visit, built on KDIGO guidelines. It shows the evidence and its grade behind every point and leaves every decision with the nephrologist. It is being developed with a community nephrology practice in Beaumont, Texas.

RESEARCH QUESTION

Does a pre-visit evidence summary, compared with usual chart review, increase guideline-recommended treatment for adults with CKD over 12 weeks?

HOW FILTRA WORKS



Filtra prepares; the clinician decides. It never orders, prescribes or alerts.

Writing

Sep 2026 · 5 min read

FDE Junkies: Addicted to the Last Mile

AI made demos nearly free, which is exactly why the last mile is now the whole job. On forward-deployed engineering and building for real use, not for the reveal.

Mar 2026 · 4 min read

From Code to Connection

My playbook for AI-powered customer obsession: customers don't buy technology, they buy better versions of themselves.

Mar 2026 · 2 min read

From Static PDFs to Dynamic Web Apps

How a multi-agent AI pipeline turned 70+ medical PDF forms into validated, accessible web forms in minutes instead of weeks.

Mar 2026 · 4 min read

Healthcare Document Intelligence

From chaos to clinical clarity: moving medical documents from passive storage to intelligent understanding, so clinicians find answers instead of searching for them.

Technical Depth

Generative AI

- Agentic & Multi-Agent Systems · RAG ·
- LLM Integration (OpenAI, Claude) ·
- Evaluation & Guardrails · NLP & ML Pipelines ·
- Explainable AI

Architecture

- Domain-Driven Design ·
- Event-Driven, CQRS & Event Sourcing ·
- API-First Microservices · Micro Frontends ·
- Legacy Modernization (Strangler Pattern) ·
- HIPAA-Aligned Compliance

Cloud Platforms

- AWS & Azure Strategy ·
- Kubernetes at Enterprise Scale · Serverless ·
- CI/CD & IaC · Observability ·
- FinOps & Cost Optimization

Engineering

- Python · Node.js · React · Blazor ·
- Java / Spring Boot · .NET Core · Kafka ·
- Amazon Redshift · Athena · Azure Data Lake

Credentials & Recognition

CERTIFICATIONS

Amazon Web Services AWS GOLDEN JACKET

All 8 active AWS certifications

Solutions Architect Professional	DevOps Engineer Professional	Solutions Architect Associate	Data Engineer Associate	Developer Associate	SysOps (now CloudOps) Associate	AI Practitioner	Cloud Practitioner


Kubestronaut


CKA Administrator


CKAD Application Developer


CKS Security Specialist


KCNA Cloud Native Associate


KCSA Security Associate


Solutions Architect Expert


DevOps Engineer Expert


Administrator Associate


Azure Fundamentals


AI Dojo Generative AI


Claude Production AI Specialist


TOGAF® 9 Certified


Project Management Professional

EDUCATION


Doctor of Health Informatics
UTHealth Houston, McWilliams School of Biomedical Informatics. In progress; translational project: [NephroIQ](#)


Master of Science
Computer Science, University of Houston–Clear Lake


Bachelor of Science
Computer Science, Osmania University

RECOGNITION

6× BEAT Award
Fannie Mae

CIO Award
Team Contribution

Author
Four articles on forward-deployed AI and healthcare AI (see [Writing](#))

Solutions Architect of the Quarter
AWS

Speaker
AWS Energy Conference, London

Volunteer
Microsoft TEALS, teaching computer science in high schools

Open to conversations about taking healthcare AI from strategy to production.
yousuf.mohammad@gmail.com · Richmond, Texas