AKASH SHANMUGANATHAN

AI Engineer | Solutions Engineer | Product Builder | AWS Certified | Agentic AI | Product Management

Building Production AI Systems That Deliver Measurable Business Outcomes

EXECUTIVE SUMMARY

AWS Certified AI Engineer and Full-Stack Product Builder with 4+ years of experience designing, building, and deploying production AI systems that transform enterprise operations. Combining deep technical expertise in multi-agent architectures, large language models, and cloud infrastructure with a product-first mindset that translates complex AI capabilities into measurable business value.

Currently leading AI product development at an AWS Advanced Tier Partner, owning the complete product lifecycle from competitive analysis and stakeholder alignment through architecture, implementation, and enterprise client delivery. Proven track record of building systems that achieve 10x operational efficiency gains, $180K+ in cost savings, and landing enterprise contracts through compelling technical demonstrations.

Specializing in two high-impact domains: (1) Healthcare AI featuring clinical NLP with Google MedGemma and Smart on FHIR interoperability for Revenue Cycle Management, and (2) Conversational AI with speech-to-speech interfaces using AWS Bedrock Multi-Agent Collaboration and Claude 4.5 for enterprise analytics. Passionate about responsible AI deployment with human-in-the-loop workflows and comprehensive observability.

PROVEN IMPACT & KEY ACHIEVEMENTS

$180K+	10x	85%	75%
Annual Cost Savings	**Efficiency Gains**	**AI Accuracy**	**Reduced Hallucinations**
Delivered through cloud migration and optimization	Production AI systems transforming workflows	Multi-agent autonomous decision-making	Enterprise RAG with semantic grounding
<800ms	**200+**	**40%**	**15+**
Voice Latency	**Workloads Migrated**	**Faster Development**	**Enterprise Projects**
Real-time speech-to-speech processing	On premises to AWS cloud infrastructure	AI-assisted development acceleration	Client-facing delivery and presentations

CORE COMPETENCIES & EXPERTISE

- **AI & LLM Engineering** - AWS Bedrock (Agents, Knowledge Bases, Guardrails), Claude 4.5 (Haiku, Sonnet, Opus), Multi-Agent Orchestration, LangChain, LangGraph, Model Context Protocol (MCP), AgentCore Framework, RAG Architecture, Prompt Engineering, Chain-of-Thought Reasoning, Few-Shot Learning, Conversational AI, Human-in-the-Loop Validation
- **Healthcare AI & Interoperability -** Google MedGemma, Smart on FHIR, HL7 FHIR R4, RCM (Revenue Cycle Management), HIPAA Compliance, EHR/EMR Integration, Claims Processing AI, Denial Prediction, API-First Healthcare Architecture
- **Vector Databases & Semantic Search** - Pinecone, ChromaDB, Amazon OpenSearch, Bedrock Titan Embeddings, OpenAI Embeddings, S3 Vector Storage, Semantic Retrieval, Similarity Search, Embedding Optimization, Hybrid Search Strategies
- **Cloud Architecture (AWS)** - Lambda, Step Functions, S3, DynamoDB, RDS (PostgreSQL, MySQL), Glue, Redshift, Athena, CloudWatch, CloudTrail, IAM, SNS/SQS, Serverless Architecture, Infrastructure as Code, Cost Optimization
- **Programming & Development** - Python, TypeScript, JavaScript, Node.js, SQL (Advanced), PySpark, React, FastAPI, Express.js, REST APIs, WebSockets, SocketIO, GraphQL, Git, Unit Testing, Integration Testing
- **Data Engineering** - ETL/ELT Pipeline Design, Medallion Architecture (Bronze/Silver/Gold), Data Lake Design, Data Warehousing, Apache Spark, Data Quality Frameworks, Schema Design, Data Partitioning, Query Optimization
- **Product & Business** - Full Product Lifecycle Ownership, 0-to-1 Product Development, Product Strategy, Roadmapping, Stakeholder Alignment, Technical Demos, POC Development, Client Presentations, Enterprise Sales Support, Agile/Scrum

PROFESSIONAL EXPERIENCE

AI Engineer — Product Development & Strategy **Onedata Software Solutions** *AWS Advanced Tier Partner*	**September 2025 – Present** **Fort Mill, SC, USA**

Leading AI product development and enterprise client delivery at an AWS Advanced Tier Partner, owning full product lifecycle from ideation to deployment. Driving AI transformation strategy across Healthcare RCM and MarTech verticals, combining hands-on engineering with technical consulting and C-suite stakeholder presentations.

Healthcare RCM AI Platform

0-to-1 Product Development | MedGemma | Smart on FHIR | HIPAA Compliant

End-to-end Healthcare Revenue Cycle Management AI platform designed to compete with industry leaders (Epic, Waystar, Cerner) through differentiated AI capabilities. Full ownership from competitive analysis through architecture, and development.

- o Leading full product lifecycle including market research, competitive analysis against Epic, Waystar, and Cerner, feature prioritization, and building compelling business cases for executive stakeholder buy-in and investment approval
- o Building AI chatbot powered by Google MedGemma for advanced clinical NLP capabilities including medical entity extraction, clinical reasoning for claims assistance, and conversational AI that differentiates from rule-based competitor solutions
- o Architecting Smart on FHIR (R4) implementation enabling seamless interoperability with any EHR/EMR system, designing API-first architecture for bidirectional healthcare data exchange across provider networks, payers, and clearinghouses
- o Developing predictive models for claims denial prevention, automated coding suggestions, and real-time eligibility verification to reduce denial rates and accelerate revenue collection cycles
- o Championing AI-assisted development tools (Claude Code, Replit, Amazon Kiro) across the engineering team, establishing best practices that accelerated time-to-production by 40% while maintaining code quality

Impact: Smart on FHIR Interoperability | 40% Development Acceleration | HIPAA Compliant Architecture | Competitive with Market Leaders

Speech-to-Speech Conversational Analytics Platform

Enterprise MarTech | AWS Bedrock Multi-Agent | Claude 4.5 | Nova Sonic | Real-Time Voice AI

Enterprise-grade conversational analytics platform enabling hands-free natural language interaction with marketing data through bi-directional voice AI. Built for MarTech clients requiring real-time insights without dashboard navigation.

- o Architected sophisticated 4-agent orchestration system using AWS Bedrock Multi-Agent Collaboration with Claude Haiku 4.5: Query Planner, Data Retriever (SQL generation, data fetching), Insight Generator (analysis), and Visualizer (chart selection)
- o Engineered bi-directional speech-to-speech interface with Amazon Nova Sonic, implementing real-time audio streaming via WebSocket/SocketIO architecture achieving sub-800ms end-to-end latency through iterative optimization with client teams
- o Built custom MCP server with 5 specialized tools for autonomous schema discovery, natural language to SQL generation, semantic data retrieval, context management, and response formatting.
- o Implemented production RAG system with Bedrock Titan Embeddings and S3-based vector storage, reducing hallucinations by 75% through semantic grounding, citation tracking, and configurable guardrails ensuring factual accuracy with marketing data
- o Established comprehensive monitoring with custom CloudWatch tracking token consumption, inter-agent communication latency, guardrail trigger rates, safety metrics, and cost attribution, enabling proactive performance optimization

Impact: 85% Autonomous Accuracy | <800ms Voice Latency | 75% Hallucination Reduction | Enterprise Client Deployment

Technical Consulting & Enterprise Client Delivery

Client Engagement | POC Development | Technical Demos | ROI Modeling | Sales Support

- o Lead technical consultant for enterprise engagements on AWS QuickSuite and agentic AI solutions, conducting discovery sessions to understand client pain points, current infrastructure, and strategic objectives
- o Rapidly developing proof-of-concept implementations that demonstrate AI capabilities with client-specific data, enabling stakeholders to visualize production potential and make informed investment decisions
- o Delivering compelling technical demonstrations to C-suite stakeholders, translating complex multi-agent architectures into quantifiable business value with clear ROI projections and implementation roadmaps
- o Driving Healthcare RCM AI transformation strategy with focus on responsible AI deployment, conducting technical assessments, architecting intelligent automation with human-in-the-loop workflows, and ensuring compliance

Software Engineer — Data Engineering & Client Delivery **Hexaware Technologies** *Global IT Services & Consulting*	**January 2021 – July 2023** Chennai, India

Data Engineer delivering cloud migrations, ETL pipelines, ML data infrastructure, and analytics solutions for Fortune 500 clients. Combined hands-on engineering with client-facing responsibilities across 15+ enterprise projects, developing strong foundations in cloud architecture and stakeholder communication.

- Delivered $180K annual cost savings by architecting and executing migration of 200+ SQL workloads from on-premises MySQL to AWS (RDS, Glue, Redshift), achieving 30% infrastructure cost reduction and 45% query performance improvement through optimized data partitioning strategies and query rewriting
- Built production ML training data pipelines with automated feature engineering, data validation, and preprocessing workflows supporting supervised learning algorithms for financial forecasting applications—handling millions of records with data quality
- Accelerated financial reporting cycles by 60% through systematic SQL query optimization, ETL architecture redesign, and implementing incremental processing patterns—reducing end-of-quarter processing from 5 hours to 2 hours and enabling faster business decisions
- Created automated Tableau dashboards serving 21 QA teams, tracking defect lifecycle metrics, sprint velocity, and quality KPIs across 15+ concurrent projects—improving visibility and enabling data-driven process improvements
- Conducted 15+ client-facing technical presentations demonstrating migration progress, architecture decisions, and optimization results, developing stakeholder communication skills that proved essential for later consulting roles

Impact: $180K Annual Savings | 45% Performance Improvement | 60% Faster Reporting | 21 Teams Served | 15+ Projects Delivered

CERTIFICATIONS & CREDENTIALS

AWS Certified Data Engineer – Associate

DEA-C01 • **Active**

Validates expertise in designing, building, and maintaining data processing systems on AWS including data pipelines, data lakes, and analytics.

Skills Validated: S3, Glue, Redshift, Athena, Kinesis, Lambda, Step Functions, Lake Formation, Data Quality

AWS Certified Gen AI Developer – Professional

In Progress • In **Progress**

Advanced certification validating skills in developing generative AI applications using AWS services including Bedrock, foundation models, and RAG architectures.

Skills Validated: AWS Bedrock, Foundation Models, RAG, Prompt Engineering, Fine-Tuning, Guardrails

AWS Certified Solutions Architect – Associate

SAA-C03 • In **Progress**

Demonstrates knowledge of designing distributed systems on AWS with focus on availability, cost optimization, and security.

Skills Validated: VPC, EC2, ELB, Auto Scaling, RDS, S3, CloudFront, Route 53, IAM, Well-Architected

EDUCATION

Master of Science in Engineering Management

Northeastern University • Boston, Massachusetts, USA

September 2023 – May 2025 • **GPA: 3.75 / 4.0**

Focus Areas: AI Product Management, Technology Strategy, Data-Driven Decision Making, Innovation Management, Operations Research

Bachelor of Engineering in Electrical & Electronics Engineering

Anna University • Chennai, Tamil Nadu, India

June 2017 – May 2021

Foundation: Signal Processing, Control Systems, Embedded Systems, Programming Fundamentals, Mathematics, Electronics

FEATURED PROJECTS — TECHNICAL DEEP DIVE

Healthcare RCM AI Platform with MedGemma

0-to-1 Product Development | Clinical NLP | FHIR Interoperability

Comprehensive Healthcare Revenue Cycle Management platform leveraging Google's MedGemma for clinical natural language processing, designed to automate claims processing, reduce denials, and accelerate revenue collection for healthcare providers.

Technologies: Google MedGemma, Smart on FHIR (R4), HL7 FHIR, Python, FastAPI, AWS Lambda, DynamoDB, S3, Claude Code, React

Architecture Highlights:

- Microservices architecture with API Gateway orchestrating clinical NLP, claims processing, and EHR integration services
- MedGemma-powered chatbot handling medical entity extraction, clinical reasoning, and conversational claims assistance
- Smart on FHIR implementation enabling bidirectional data exchange with Epic, Cerner, and other major EHR systems
- Event-driven processing pipeline using AWS EventBridge for real-time claims status updates and denial alerts

Key Features:

- Medical Entity Extraction: Automated identification of diagnoses, procedures, and billing codes from clinical notes
- Clinical Reasoning Engine: AI-powered suggestions for coding optimization and documentation improvement
- Real-Time Eligibility: Instant patient coverage verification integrated with major payers
- Audit Trail: Comprehensive logging for HIPAA compliance and claims audit support

Impact: FHIR Interoperability with Any EHR | 40% Faster Development | HIPAA Compliant | Enterprise-Ready Architecture

Speech-to-Speech Conversational Analytics

Multi-Agent AI | Real-Time Voice | Enterprise MarTech

Enterprise conversational analytics platform enabling natural language interaction with marketing data through bi-directional voice AI, eliminating dashboard navigation and enabling hands-free insights for marketing executives and analysts.

Technologies: AWS Bedrock, Claude Haiku 4.5, Amazon Nova Sonic, LangChain, MCP, SocketIO, Python, TypeScript, React, S3, CloudWatch

Architecture Highlights:

- 4-agent orchestration using AWS Bedrock Multi-Agent Collaboration with specialized agents for query planning, data retrieval, insight generation, and visualization
- Real-time bidirectional audio streaming via WebSocket/SocketIO achieving sub-800ms end-to-end latency
- MCP server providing 5 tools: schema discovery, SQL generation, semantic retrieval, context management, response formatting
- Enterprise RAG with Bedrock Titan Embeddings, S3 vector storage, and configurable guardrails for data governance

Key Features:

- Natural Language Queries: 'What were our top performing campaigns last quarter?' processed through multi-agent pipeline
- Voice-First Interface: Hands-free interaction with Nova Sonic speech synthesis and recognition
- Autonomous SQL Generation: Schema-aware query generation with validation and optimization
- Guardrails & Safety: Content filtering, PII protection, and query scope limitations for enterprise compliance

Impact: 85% Autonomous Accuracy | <800ms Voice Latency | 75% Fewer Hallucinations | Production Enterprise Deployment

Enterprise Cloud Migration & Data Platform

AWS Migration | Data Lake | Performance Optimization

Large-scale migration of legacy on-premises data infrastructure to AWS, establishing modern data lake architecture with optimized query performance and significant cost reduction for a Fortune 500 financial services client.

Technologies: AWS RDS, Glue, Redshift, S3, Athena, MySQL, PySpark, Python, SQL, Tableau, CloudWatch

Architecture Highlights:

- Medallion architecture (Bronze/Silver/Gold) data lake on S3 with Glue ETL pipelines and Redshift analytics layer
- Automated data quality framework with validation rules, anomaly detection, and data lineage tracking
- Comprehensive monitoring with CloudWatch dashboards tracking pipeline health, data freshness, and cost metrics

Key Features:

- 200+ Workload Migration: Systematic migration with zero data loss and minimal business disruption
- Query Optimization: Rewritten queries with proper indexing, partitioning, and materialized views
- Automated Reporting: Self-service Tableau dashboards replacing manual Excel-based reporting
- Cost Governance: Resource tagging, usage monitoring, and automated scaling for cost optimization

Impact: $180K Annual Savings | 45% Query Performance Improvement | 60% Faster Reporting | 30% Infrastructure Cost Reduction

WHY WORK WITH ME

End-to-End Ownership

I don't just build features—I own entire product lifecycles from competitive analysis and stakeholder alignment through architecture, implementation, deployment, and client success. This full-stack ownership mentality means I can step into any role the team needs.

Bridge Between Technical & Business

With experience delivering technical demos to C-suite executives and building business cases for AI investments, I translate complex architectures into quantifiable ROI. I speak both engineer and executive fluently.

Production AI Experience

Every system I've built runs in production with real users and real data. I understand the difference between demo-ready and enterprise-ready: observability, guardrails, compliance, scalability, and maintainability.

Healthcare AI Specialization

Deep domain expertise in healthcare technology including HIPAA compliance, HL7 FHIR interoperability, clinical NLP, and revenue cycle management—a rapidly growing market with complex regulatory requirements.

Responsible AI Advocate

I prioritize human-in-the-loop workflows, comprehensive guardrails, and transparent AI systems. Building trust through responsible deployment isn't optional—it's how enterprise AI succeeds.

Continuous Learner

From earning AWS certifications to championing AI-assisted development tools across teams, I stay at the cutting edge of AI engineering while helping others level up their skills.

Let's Build Something Extraordinary Together

I'm actively exploring opportunities in AI Engineering, Solutions Architecture, and Product Building roles.
Let's discuss how I can help transform your organization with production-ready AI systems.

Email	Phone
akashs101199@gmail.com	**(857) 693-8403**
LinkedIn	GitHub
linkedin.com/in/akash101199	**github.com/akashs101199**