Michael Legemah

516-273-1611 | michaellegemah@gmail.com | LinkedIn | Portfolio | GitHub | New York, NY

SUMMARY

Principal / Staff AI Engineer with 13+ years of experience designing, architecting, and delivering production software platforms, with deep expertise spanning enterprise AI, agentic systems, distributed architectures, cloud platforms, and full-stack product engineering. Currently architecting and delivering production AI solutions at Amazon Web Services, building LLM-powered applications, agentic workflows, Retrieval-Augmented Generation (RAG) systems, AI evaluation frameworks, and cloud-native AI platforms using Python, TypeScript, AWS Bedrock, SageMaker, OpenSearch, Lambda, AppSync, DynamoDB, and distributed services. Specialized in translating ambiguous business problems into reliable AI systems through agent orchestration, tool calling, RAG, model evaluation, LLM-as-a-Judge, guardrails, confidence scoring, validation pipelines, observability, human-in-the-loop controls, and inference optimization. Strong background in distributed systems, API architecture, event-driven systems, microservices, and platform engineering. Technical leader experienced in setting architectural direction, establishing engineering standards, mentoring engineers, leading architecture and design reviews, and partnering with product, engineering, applied science, enterprise customers, and executive stakeholders. Brings an uncommon combination of AI engineering depth, software architecture, cloud infrastructure, and product-development experience.

CORE EXPERTISE

AI & Machine Learning
Agentic AI • LLM Applications • RAG • Multi-Agent Systems • Agent Orchestration • Tool Calling • LLM Inference • Prompt Engineering • Model Evaluation • LLM-as-a-Judge • Evaluation-Driven Development • Guardrails • Responsible AI • AI Observability • Confidence Scoring • Human-in-the-Loop Systems • Inference Optimization • Semantic Search • Vector Search • Embeddings • Enterprise AI

AI Frameworks & Platforms
AWS Bedrock • Amazon SageMaker • LangGraph • LangChain • Hugging Face • DeepEval • MCP • OpenAI APIs • Google Gemini APIs • Claude • Amazon Q

Cloud & Infrastructure
AWS • Azure • Docker • Kubernetes • Terraform • CI/CD • MLOps • IAM • KMS • CloudWatch • ECS • EC2 • S3 • RDS • Lambda • AppSync • API Gateway • DynamoDB • OpenSearch • EventBridge

Distributed Systems & Backend
Python • FastAPI • Node.js • Fastify • TypeScript • REST • GraphQL • gRPC • Microservices • Event-Driven Architecture • Kafka • RabbitMQ • WebSockets • Distributed APIs • Workflow Automation • Data Pipelines

Data & Storage
PostgreSQL • DynamoDB • MongoDB • Redis • OpenSearch • Vector Databases • SQL

Frontend & Product Engineering
React • Next.js • React Query • Redux • Storybook • TypeScript • JavaScript • HTML5 • CSS3 • Design Systems • React Native

Engineering Leadership
Technical Strategy • System Design • Architecture • Platform Engineering • Architecture Reviews • Engineering Standards • Technical Mentorship • Cross-Team Leadership • Developer Productivity • Code Reviews • Technical Roadmaps • Agile/Scrum

Testing & Reliability
LLM Evaluation • DeepEval • Jest • Playwright • Cypress • React Testing Library • TDD • CI/CD • Observability • Production Debugging • Incident Response

WORK EXPERIENCE

Software Development Engineer IV (Contract) | **Amazon Web Services** | New York, NY 10/25 – Present

- Architect and deliver production-grade AI platforms and applications using Python, TypeScript, React, AWS Bedrock, SageMaker, Lambda, AppSync, DynamoDB, and OpenSearch for enterprise workflows and intelligent customer experiences.

- Designed agentic AI architectures that combine deterministic business workflows with LLM-powered agents through confidence scoring, validation pipelines, automated fallback mechanisms, guardrails, and human-in-the-loop controls to improve reliability in production environments.
- Built Retrieval-Augmented Generation (RAG) services integrating enterprise knowledge repositories, vector search, REST APIs, and LLM-powered retrieval pipelines to support intelligent document synthesis and decision-support workflows.
- Developed AI evaluation and validation frameworks using LLM-as-a-Judge, regression testing, confidence scoring, automated validation, and human feedback loops to systematically evaluate model behavior and improve production reliability.
- Designed scalable APIs, event-driven microservices, and distributed cloud architectures integrating AI agents with enterprise applications and AWS services.
- Optimized LLM applications through prompt engineering, retrieval tuning, embedding strategies, vector search, caching, and model selection, balancing latency, throughput, scalability, and operational cost.
- Implemented CI/CD, observability, monitoring, authentication, production debugging, and incident-response practices to improve operational reliability and deployment confidence for cloud-native AI applications.
- Partnered with software engineers, data scientists, applied science stakeholders, solution architects, product managers, and enterprise customers to translate ambiguous requirements into scalable AI capabilities and production-ready systems.
- Drove architecture decisions across AI application and platform components while balancing product requirements, security, maintainability, reliability, governance, and long-term scalability.
- Leverage AI-native engineering workflows using Claude Code, Amazon Q, and agent-assisted development across prototyping, implementation, testing, debugging, and documentation while maintaining software engineering quality standards.

Staff Software Engineer (Contract) | **About Objects** | Gaithersburg, MD (Remote) 1/25 – 10/25

- Architected and deployed production RAG pipelines and LangGraph-based multi-agent systems that automated complex enterprise document workflows through durable AI orchestration and scalable cloud services.
- Established AI engineering practices spanning prompt engineering, evaluation datasets, automated validation, guardrails, observability, and evaluation-driven development to improve reliability of customer-facing AI capabilities.
- Built scalable Python microservices, FastAPI services, distributed APIs, and AI inference services emphasizing performance, security, observability, and maintainable architecture.
- Designed AI-powered capabilities from initial concept through production implementation, establishing reusable patterns for agent orchestration, retrieval, evaluation, validation, and deployment.
- Integrated AI capabilities into enterprise applications while maintaining maintainability, reliability, security, and production operational requirements.
- Led architecture and technical design across frontend, backend, cloud, and AI components, establishing reusable engineering patterns across product teams.
- Mentored engineers on AI architecture, testing strategies, evaluation-driven development, deployment automation, and software design.
- Led architecture reviews, code reviews, and engineering discussions while promoting scalable software architecture and high engineering standards.

Lead Front-End Engineer (Contract) | **Apex Systems** | Pittsburgh, PA (Remote) 4/24 – 1/25

- Led development of enterprise-scale React and TypeScript applications for the U.S. Army's AIDE platform used by Army officers for mission and operational workflows.
- Modernized distributed service communication by transitioning legacy REST integrations to GraphQL and gRPC, improving interoperability, type safety, maintainability, and frontend data access.
- Delivered end-to-end capabilities using React, TypeScript, Node.js, and Python while influencing architecture, engineering standards, API governance, and long-term platform evolution.
- Served as a technical bridge across frontend, backend, platform, and product engineering teams, aligning API and system-design decisions across squads.
- Established reusable engineering patterns and standards to improve maintainability and consistency across distributed development teams.
- Participated in architecture discussions, code reviews, technical design, and cross-functional delivery for mission-critical government software.

Software Engineer IV | **Zivaro** | Denver, CO (Remote) 4/22 – 4/24

- Built mission-focused React and TypeScript dashboards for U.S. Space Force and Air Force officers to analyze constellation and satellite data across COMMAND and Hyperion applications.
- Architected distributed backend services using Python, FastAPI, Docker, Kubernetes, RabbitMQ, and cloud infrastructure supporting mission-critical government applications.
- Designed event-driven microservice patterns that improved service isolation, deployment independence, fault tolerance, and scalability.
- Integrated backend services using Python and Lua with RabbitMQ, Keycloak, Docker, Kubernetes, and Azure Cloud for secure service-to-service communication.
- Tested and validated mission applications on-base within SIPRNet air-gapped environments, gaining hands-on experience operating software within constrained security, networking, and deployment environments.
- Established reusable frontend standards, shared component patterns, and engineering practices across defense-focused engineering teams.
- Collaborated with government stakeholders, engineers, and cross-functional teams to translate mission requirements into scalable software while balancing usability, reliability, security, and operational constraints.

Senior Front-End Engineer | **Reflexions Data** | New York, NY 10/20 – 10/21

- Led a team of six engineers delivering high-availability transaction platforms using React, TypeScript, GraphQL, Node.js, and CircleCI for MBTA, OMNY, and BART.
- Drove technical direction, architecture decisions, mentoring, and end-to-end ownership across multiple concurrent product initiatives.
- Established engineering practices around performance, reliability, testing, scalability, and maintainable distributed-system design.
- Influenced engineering strategy through reusable architecture patterns and technical standards while improving developer productivity and long-term maintainability.
- Contributed to an 80% revenue increase through technical leadership and end-to-end product ownership.

Senior Front-End Engineer (Contract) | **Mondo** | New York, NY 2/19 – 10/20

- Designed and implemented a Storybook-driven design system and component library using React, TypeScript, and Vue, enabling consistent user experiences and accelerating frontend development across a high-traffic healthcare platform.
- Built reusable design-system patterns that improved consistency and maintainability across product teams.
- Modernized legacy frontend systems into modular, cloud-native AWS services using EC2, CloudFront, RDS, and Route53.
- Improved platform reliability and reduced operational downtime through modernization of legacy architecture and deployment patterns

Senior / Mid-Level Front-End & Web Engineering Roles | **Various Companies** | New York, NY 1/13 – 11/18

- Delivered full-stack e-commerce platforms using React, Node.js, MySQL, and secure payment integrations, contributing to more than $3M in online revenue across high-traffic consumer properties.
- Built and optimized customer-facing web applications with a focus on performance, scalability, conversion, responsive design, and user experience.
- Revamped AKC.org using modern UX, responsive design, and performance optimization practices, improving customer engagement and conversion.
- Developed full-stack applications spanning frontend interfaces, backend services, databases, APIs, and third-party integrations.

SKILLS

Production AI Architecture: Agentic AI systems, Multi-agent orchestration, LLM-powered applications, RAG architectures, Enterprise knowledge retrieval, Tool-use and function calling, Human-in-the-loop systems, AI workflow automation

AI Reliability & Evaluation: LLM-as-a-Judge, Automated evaluation, Regression testing, Evaluation datasets, Confidence scoring, Guardrails, Validation pipelines, AI observability, Responsible AI practices, Production monitoring

AI Performance & Optimization: Prompt optimization, Retrieval tuning, Embedding strategies, Vector search optimization, Semantic search. Caching, Model selection, Latency optimization, Throughput optimization, Cost optimization

AI Platform Engineering: Python AI services, FastAPI, AWS Bedrock, SageMaker, OpenSearch, Lambda, AppSync, API Gateway, DynamoDB, Event-driven architectures, Containerized workloads, CI/CD and MLOps

AI / ML: AWS Bedrock, SageMaker, LangGraph, LangChain, Hugging Face, DeepEval, RAG, LLM Evaluation, LLM-as-a-Judge, Agentic AI, Multi-Agent Systems, Prompt Engineering, Tool Calling, Guardrails, AI Observability, Responsible AI, Vector Search, Embeddings, MCP

Technical Leadership: Architecture and system design, Platform strategy, Technical roadmap development, Cross-team architecture alignment, Architecture and design reviews, Engineering standards, API governance, Mentorship, Code reviews, Developer productivity, Technical decision-making, Customer and stakeholder engagement, Translating ambiguous requirements into technical solutions, Balancing innovation with reliability, security, and operational excellence

Frontend: React, Next.js, TypeScript, JavaScript, React Query, Redux, Storybook, React Native, Tailwind, Material UI, Expo, NativeWind, ShadCNUI

Backend: Python, FastAPI, Node.js, Fastify, REST, GraphQL, gRPC, WebSockets, Microservices, Distributed Systems, Event-Driven Architecture

Data: PostgreSQL, DynamoDB, MongoDB, Redis, MySQL, OpenSearch, Vector Databases

Messaging: Kafka, RabbitMQ

AWS: Bedrock, SageMaker, Lambda, AppSync, API Gateway, DynamoDB, OpenSearch, S3, RDS, ECS, EC2, IAM, KMS, CloudWatch, EventBridge, Cognito, CloudFront, CodePipeline, CodeDeploy, Neprtune, SQS, SNS, Athena, Glue, Q, Connect

Cloud / Infrastructure: AWS, Azure, Docker, Kubernetes, Terraform, MLOps, CI/CD

Engineering: GitHub Actions, CircleCI, Git, npm, Yarn, Bun, Agile/Scrum

Testing: DeepEval, Jest, Vitest, Playwright, Cypress, React Testing Library, TDD

Languages: Python, TypeScript, JavaScript, Go, SQL, GraphQL, YAML, HTML5, CSS3

Tooling & Process: Figma, JIRA, git, yarn, npm, Bun, Stripe, Scrum/Agile

Certifications: US Department of Defense Secret Security Clearance