

Nicholas Marler

Senior Software Engineer

📍 Austin, TX ✉ nicholasmarler0236@outlook.com ☎ (765) 800-9091 🔗 [linkedin.com/in/nicholas-david-marler-91a770431](https://www.linkedin.com/in/nicholas-david-marler-91a770431)

SUMMARY

Senior Software Engineer specializing in Generative AI, Agentic AI, cloud-native architecture, and microservices development. Experienced in designing and delivering AI-powered applications using AWS Bedrock, Claude, MCP, LangChain, LangGraph, CrewAI, LlamaIndex, Pinecone, Databricks, and Snowflake. Proven expertise in building scalable platforms with Python, FastAPI, Node.js, event-driven architectures, Kong Gateway, and Kubernetes, while implementing AI governance, security, observability, and multi-cloud solutions across AWS and GCP. Strong track record of delivering intelligent healthcare and wellness solutions with AI-driven recommendations, voice assistants, real-time analytics, and personalized user experiences.

TECHNICAL SKILLS

Generative AI & LLM Engineering: Agentic AI, Generative AI, Large Language Models (LLMs), Retrieval-Augmented Generation (RAG), Prompt Engineering, AI Workflow Automation, Multi-Agent Systems, Tool Calling, Function Calling, Knowledge Retrieval

LLM Evaluation, AI Safety & Governance: LLM Evaluation, Answer Relevance, Context Relevance, Faithfulness, Hallucination Reduction, Citation Accuracy, Content Moderation, Grounding, Prompt Safety, PII Detection & Masking, Blocked-Topic Controls, Escalation Logic, Responsible AI, AI Guardrails

AI Platforms & Frameworks: Azure OpenAI, AWS Bedrock, Snowflake Cortex AI, Azure AI Search, LangChain, LangGraph, Amazon SageMaker, OpenAI API

Data & Analytics: Snowflake, Vector Embeddings, Vector Search, Semantic Search, Data Warehousing, Data Pipelines, Real-Time Analytics, Structured & Semi-Structured Data Processing

Backend Development: Python, FastAPI, Java, Spring Boot, Node.js, REST APIs, GraphQL, Microservices, WebSockets, Apache Kafka, Amazon SQS, Amazon SNS, AWS Lambda, Event-Driven Architecture

Frontend Development: TypeScript, JavaScript, React.js, Next.js, AngularJS, Tailwind CSS

API Management & Observability: Kong Gateway, Kong AI Gateway, API Security, Traffic Management, Rate Limiting, New Relic, Application Performance Monitoring (APM), Distributed Tracing, Logging, Telemetry

Databases: PostgreSQL, MySQL, MongoDB, SQL, NoSQL

Cloud, DevOps & Infrastructure: Microsoft Azure, AWS, Docker, Kubernetes, Azure DevOps, CI/CD, GitHub Actions, Jenkins, Bamboo, Infrastructure as Code (IaC), Cloud-Native Architecture

PROFESSIONAL EXPERIENCE

Senior Software Engineer | MyNetDiary

June 2022 – July 2026, Austin, TX

- Architected healthcare/wellness platform using AWS Bedrock, Claude, Agentic AI, MCP, LangChain, LangGraph, CrewAI, LlamaIndex, Pinecone, PydanticAI for personalized health-coaching, meal-recommendation, exercise-planning, AI-voice-interaction, notification-service, and digital-health-twin capabilities.
- Designed Nutrition, Exercise, Health Risk, and AI Coach multi-agent workflows using retrieval-augmented generation (RAG), semantic search, memory management, vector search, tool calling, Bedrock Agents, reducing manual health analysis over 80% and improving recommendation quality.
- Built enterprise-scale AI knowledge and retrieval systems using LlamaIndex, LangChain, Pinecone, vector embeddings, and Bedrock Knowledge Bases, enabling intelligent analysis of nutrition data, wearable device metrics, health records, and behavioral patterns.
- Developed scalable microservices-based backend solutions using Python, FastAPI, Node.js, PostgreSQL, Redis, OpenSearch, Kafka/MSK, and Kong Gateway, supporting food tracking, health analytics, wearable integrations, AI assistants, notifications, and high-volume real-time workloads.
- Implemented event-driven architecture and asynchronous processing pipelines using Kafka/MSK, improving system scalability, resiliency, and real-time data processing across healthcare, fitness, and AI services.
- Engineered modern data and analytics platforms using Databricks, Snowflake, Cortex AI, feature engineering pipelines, vector processing, and real-time analytics, enabling advanced health insights, personalized recommendation engines, and digital body modeling.
- Architected and deployed secure multi-cloud infrastructure across AWS and GCP utilizing Kubernetes (EKS), Terraform, GitHub Actions, and CI/CD pipelines, reducing deployment time by approximately 70% and improving platform reliability and scalability.
- Implemented enterprise-grade security and AI governance controls using AWS Bedrock Guardrails, PII detection and masking, prompt injection protection, encryption, audit logging, and Responsible AI practices, ensuring secure handling of sensitive healthcare and personal data.

- Leveraged Claude, GitHub Copilot, and Cursor AI for architecture design, code generation, refactoring, automated testing, documentation, and troubleshooting, accelerating development velocity by approximately 40-50% while maintaining high code quality and engineering standards.
- Translated business requirements from product managers, healthcare specialists, user-experience/user-interface (UX/UI) teams, operations teams, executive stakeholders into scalable AI solutions. Instrumented New Relic observability/distributed tracing/telemetry/AI monitoring/operational dashboards, reducing issue-resolution time 50% and improving production visibility.

Software Engineer | Charles Schwab

April 2020 – June 2022, Westlake, TX

- Engineered enterprise backend applications with Python, Java, and Spring Boot, reducing average request latency by 15% while increasing transaction-processing capacity by approximately 2x.
- Implemented RESTful and GraphQL APIs and microservices using established API integration and security practices, reducing internal integration time by 30% and improving service reliability.
- Optimized PostgreSQL data models, queries, and database access patterns, improving query performance by 25% and supporting reliable downstream service integration.
- Deployed cloud-based applications using AWS EC2, RDS, S3, and Lambda, increasing system availability to 99.9% and reducing infrastructure provisioning time by 40%.
- Refactored backend processing through concurrency optimization, application profiling, and architectural improvements, reducing CPU utilization by approximately 20% and improving batch-processing performance.
- Automated build, deployment, and release workflows using scripting and CI/CD tools, reducing manual deployment effort while contributing to API documentation, operational procedures, production troubleshooting, and engineering standards.

Software Developer | General Motors

March 2019 – April 2020, Austin, TX

- Assembled customer-facing and internal customer relationship management (CRM) workflows with React.js, Next.js, TypeScript, JavaScript, Tailwind CSS, Node.js, and Python/FastAPI, improving usability and reducing manual data-entry activities by 25%.
- Integrated React components with Node.js APIs and Azure services to deliver CRM pages and business workflows, reducing manual data-entry activities by 25%.
- Implemented scalable backend services and APIs using FastAPI and microservice architecture, providing reliable integration between CRM functionality and enterprise systems.
- Optimized PostgreSQL schemas and SQL queries supporting CRM functionality, reducing average query latency by 15% and improving application response times.
- Deployed and supported applications through Azure App Service and Azure Kubernetes Service (AKS) while implementing Apache Kafka event-driven communication, accelerating release cycles and improving system reliability.
- Partnered with product, QA, and engineering teams to test, debug, and optimize Node.js and FastAPI services, contributed Azure DevOps and Bamboo pipelines for automated deployments, and authored technical documentation, reducing production incidents and speeding release cycles.

Frontend Developer | L3Harris Technologies

January 2018 – February 2019, Spring, TX

- Created and enhanced scalable frontend applications with AngularJS and TypeScript, delivering responsive interfaces for internal and customer-facing platforms, reducing user-error reports by 40% and increasing adoption by 30%.
- Built Progressive Web Applications (PWA) with React, Next.js, and Tailwind CSS, improving application performance, responsiveness, and cross-device user experience.
- Integrated backend services using Swagger-documented REST APIs and implemented automated frontend testing to improve application reliability and maintainability.
- Managed source code using Git and collaborated with Agile engineering teams during sprint planning, code reviews, and iterative feature delivery, ensuring timely releases and higher code quality.

EDUCATION

Bachelor's degree in Computer Science | Texas Tech University

August 2011 – December 2017, Lubbock, TX

CERTIFICATIONS

Microsoft Specialist: Architecting Microsoft Azure Solutions | Microsoft

July 2016

Microsoft Certified Professional | Microsoft

July 2016