Shawn Seiki

Plano, TX, US

(510) 920–9911

shawnseiki46@gmail.com

Summary

AI/ML Full–Stack Engineer with deep expertise in building intelligent, scalable, and high–performance applications. Skilled in architecting modern platforms using React, Next.js, Node.js, Python/FastAPI, and cloud–native ecosystems. Experienced in integrating LLMs, embeddings, RAG pipelines, vector databases, and real–time data processing into production systems. Adept at designing AI–powered product experiences, implementing event–driven microservices, optimizing full–stack performance, and leading engineering initiatives. Passionate about delivering clean architecture, strong testing practices, and robust, secure AI–driven applications.

Experience

Paramount, *Midtown Manhattan, NYC* **— Senior AI/ML Full-Stack Engineer**

Sep 2022 - Oct 2025

- Architected an AI-enabled analytics platform using React, Next.js, TypeScript, and Python/FastAPI, delivering real-time personalized insights that scaled across millions of user interactions.

- Integrated LLM-driven semantic search, multi-document summarization, and embeddings retrieval using OpenAI APIs and vector databases, significantly elevating platform intelligence.

- Designed scalable RAG pipelines for metadata enrichment, improving content discovery, ranking accuracy, and high-volume personalization workflows.

- Built distributed microservices with Node.js and Python to automate content tagging, transcription, and classification across large ingestion pipelines.

- Created GPU-optimized inference endpoints for recommendation scoring and high-throughput AI workloads, ensuring minimal latency at peak usage.

- Developed a large-scale design system using Storybook, Tailwind, and atomic React patterns to support consistent UI/UX across multiple engineering teams.

- Reduced TTI and overall load time by over 40% via advanced code-splitting, caching layers, and deep Vite-based build optimizations.

- Refactored GraphQL schemas and Apollo resolvers to support AI contextual queries, enabling richer filtering, ranking, and data orchestration.

- Implemented real-time observability with CloudWatch, Prometheus, and custom logging dashboards to monitor inference health and model performance.

- Built internal tools for analyzing embedding similarity, tracking model drift, and profiling inference latency in production environments.

- Led end-to-end CI/CD pipelines for AI microservices using GitHub Actions and AWS CodePipeline, improving deployment reliability and frequency.

- Introduced AI-assisted testing workflows across Jest/Vitest and Playwright, streamlining test coverage for evolving ML-powered features.

- Mentored engineers on AI integration patterns, prompt engineering fundamentals, scalable architecture, and end-to-end observability best practices.

Wells Fargo, *San Francisco, CA* — Senior Frontend Engineer (AI-Enhanced Systems)

Mar 2021 - Sep 2022

- Modernized enterprise dashboard systems with AI-powered anomaly detection using Python and Node.js, enabling teams to surface insights more quickly across large financial data flows.

- Built advanced React/TypeScript visualization layers driven by ML-based classification and predictive scoring models to support real-time decision making.

- Developed natural-language insights panels using LLM summarization APIs, allowing business users to interpret complex operational metrics more efficiently.

- Implemented AI-assisted form validation and automated classification workflows, reducing manual review time across multiple high-volume financial operations.

- Enhanced accessibility and Lighthouse performance through semantic refactoring, improved layout structures, and fully WCAG-compliant UI patterns.

- Created reusable Storybook and Material UI component libraries that increased design consistency and development speed across multiple banking product teams.

- Designed scalable frontend state management architectures using Redux and Context API to handle ML-enriched datasets and dynamic real-time updates.

- Implemented real-time data syncing using WebSockets and event-driven streaming APIs, ensuring highly responsive and continuously updated dashboards.

- Built CI/CD pipelines for ML-adjacent UI features using GitHub Actions and AWS CodePipeline, improving deployment reliability and reducing time to release.

- Collaborated closely with data engineering and ML teams to integrate telemetry, model signals, and unified UX patterns, ensuring consistency across AI-powered interfaces.

Target, *Minneapolis, MN* — Full Stack Engineer

Mar 2019 - Mar 2021

- Developed a customer-facing ecommerce platform using Next.js, React, Node.js, and Python services, improving page performance, SEO rankings, and user engagement.

- Built reusable frontend component systems using React Hooks and Tailwind CSS while coordinating backend data flows to support dynamic, ML-informed product displays.

- Implemented secure authentication and authorization flows with JWT, OAuth2, and role-based access controls across both frontend and backend systems.

- Designed and maintained Node.js and Express microservices integrated with MongoDB, handling API orchestration, business logic, and multi-service data pipelines.

- Improved overall site performance by reducing bundle size, implementing code splitting, optimizing WebP/AVIF images, and refining server-rendered pages.

- Architected a scalable micro-frontend ecosystem enabling independent development, deployment, and runtime integration of multiple full-stack modules.

- Built automated end-to-end tests using Cypress and React Testing Library, validating UI flows, API integrations, and backend data consistency.

- Integrated GraphQL APIs with Apollo Client and server-side resolvers, improving data-fetch efficiency and reducing over-fetching across the app.

- Implemented responsive UI layouts using Tailwind, and Figma-driven design specs while ensuring backend query performance supported dynamic content.

- Contributed to architectural discussions involving caching layers, API schema design, ML-assisted personalization strategies, and long-term application scalability.

Apple, *Cupertino, CA* — **Full Stack Developer**

May 2017 - Feb 2019

- Developed internal dashboards using React, Redux, TypeScript, and Node.js services, supporting high-volume workflows across multiple engineering and operational teams.

- Built interactive visualization components with D3.js and SVG, paired with backend aggregation endpoints to deliver optimized, real-time data to the UI.

- Improved backend performance by optimizing API response times, refining data models, and reducing latency across critical internal systems.

- Integrated GraphQL APIs and contributed to resolver design, streamlining data retrieval across various applications and backend services.

- Worked on frontend and backend performance optimization efforts, enhancing rendering efficiency, caching behavior, and Lighthouse scores.

- Implemented Jest and Cypress test suites to validate UI behavior and backend interactions, ensuring consistent reliability across deployments.

- Migrated legacy JavaScript modules to TypeScript and improved security through CSP enforcement, strict validation, and hardened API communication.

Skills

AI & ML:

- LLM integration (OpenAI, Anthropic, Gemini)
- RAG Systems, Vector Databases (Pinecone, Chroma, FAISS)
- Embeddings Pipelines, Semantic Search, NLP
- Python (FastAPI, NumPy, Pandas)
- Model Monitoring, Drift Detection, AI Observability

Frontend Development:

- TypeScript, React, Next.js, Vue, Angular, Redux, Context API, Hooks, Nuxt, Vuex, Pinia, RxJS, NgRx, Zustand, Storybook, D3.js, HighChart
- Tailwind CSS, Material UI, SASS, LESS
- HTML5, CSS3, jQuery, JavaScript

Backend Development:

- Node.js, TypeScript, Python/FastAPI, Java, Express
- RESTful APIs, Microservices, GraphQL, Apollo
- SQL, NoSQL, PostgreSQL, MongoDB, DynamoDB, Redis, MySQL, CosmosDB, Firebase
- Event-Driven Architecture

Cloud & DevOps:

- AWS (Lambda, EC2, S3, CloudWatch), Docker, Kubernetes, Azure
- CI/CD (GitLab, Jenkins, AWS CodePipeline, Github Actions)
- Serverless Architecture, Infrastructure as Code (IaC)
- Prometheus, Grafana for monitoring and alerting

Software Engineering Practices:

- Agile Development, Test-Driven Development (TDD)
- Scalable System Design & Architecture
- Code Reviews, Mentorship, Team Leadership

Others:

- npm, yarn, Webpack, Vite, Parcel
- Git, Github, Gitlab, Bitbucket
- Jest, React Testing Library, Cypress, Playwright, Vitest

Education

University of Phoenix, Phoenix, AZ — Bachelor's *Degree*

2013 – 2017