

Max M. Awad 1.0

San Francisco Bay Area | 747-777-3918 | maxmawad@gmail.com | LinkedIn | maxawad.com

AI ENGINEER | STAFF-LEVEL SYSTEMS ARCHITECT | EX-META, APPLE, GOOGLE

Fractional CTO and Staff-Level Systems Engineer specializing in Generative AI, automation, and scalable infrastructure. Experience spans Meta, Apple, Google, and high-growth companies, building and operating production-critical systems used by hundreds of engineers and millions of users.

Focused on designing and delivering end-to-end AI systems—from LLM integration and RAG pipelines to agentic workflows using MCP and APA—built to operate reliably in real-world environments. Combines deep expertise in distributed systems, full-stack development, and cloud infrastructure to create systems that are scalable, secure, and production-ready.

Operates as an embedded technical leader, owning architecture and execution across AI-driven products and platforms. Known for turning complex ideas into working systems that don't just prototype, but execute, automate, and scale.

TECHNICAL SKILLS

Languages: JavaScript (ES6+), TypeScript, Python, Go, Swift, Java, C++, Rust, PHP, Solidity, Ruby, Bash/Shell, SQL, HTML/CSS, HCL (Terraform), Jupyter Notebooks

AI / ML Systems: LLM Integration, Prompt Engineering, AI Agents & Automation, NLP Pipelines, Experimentation Pipelines, Model Observability, Search Infrastructure, GPT-4, Claude, Gemini, LLaMA, Model Context Protocol (MCP), TensorFlow, PyTorch, ONNX Runtime, Pinecone (Vector DB), OpenAI SDK, Anthropic SDK, Vercel AI SDK, Perplexity AI SDK, Speaker Verification (Resemblyzer), WebRTC VAD, OpenWakeWord

Architecture & Design Patterns: MVC, MVVM, MVP, Microservices Architecture, Distributed Systems Design, Event-Driven Architecture, API Architecture, GraphQL Architecture (Apollo Client / Codegen), Relay-style Patterns, Monorepo Architecture, Object-Oriented Design, Strategy Pattern

Frontend: React, React Native, Next.js, SvelteKit / Svelte 5, Expo, Vite, Turbopack, TailwindCSS, Radix UI, shadcn/ui, Material UI, NextUI, Headless UI, Flowbite, React Bootstrap, Styled Components, Framer Motion, Storybook, Recharts, Chart.js, VisX, ProseMirror, React-Beautiful-DND, i18next, React PDF

Backend: Node.js, Express, Fastify, FastAPI, Flask, Django, Ruby on Rails, Spring Boot, Uvicorn

Mobile & Native: React Native, SwiftUI, Xcode, Swift Package Manager, CocoaPods, RealityKit, visionOS, Metal (GPU Graphics), Core Audio, AVFoundation, ScreenCaptureKit, macOS Accessibility APIs

Databases: PostgreSQL, MySQL, MongoDB, Redis, Supabase, Neon (Serverless Postgres), Vercel Postgres, SQLite, Firestore, DynamoDB, AWS Timestream, Google BigQuery, Google Cloud SQL, Google Cloud Spanner, NoSQL

ORMs & Data Access: TypeORM, Mongoose, SQLAlchemy, pg-promise, Supabase Migrations

Cloud & DevOps: AWS (EC2, Lambda, S3, ALB/NLB, DynamoDB, KMS, SageMaker), Google Cloud (Cloud Run, GKE, Cloud Functions, Cloud Storage, BigQuery, Cloud SQL, Cloud Monitoring), DigitalOcean, Vercel, Firebase, Supabase, Railway, Docker, Docker Compose, Kubernetes, Helm, Terraform, Skaffold, Kustomize, GitHub Actions CI/CD, Nginx, systemd, WireGuard VPN, Bash Automation

Auth & Security: NextAuth, Auth0, Firebase Auth, JWT, WebAuthn / Passkeys, LDAP, bcrypt, Helmet, Identity Verification (SumSub)

Payments & Financial APIs: Stripe, Plaid, Teller, ethers.js, web3.js, bitcoinjs-lib, WalletConnect, Coinbase Wallet SDK

Browser Automation: Puppeteer, Playwright, BeautifulSoup

Messaging: Twilio, SendGrid, Resend, Slack API, WhatsApp Web.js, WebSockets, Socket.io, Web Push Notifications

Testing & Code Quality: Vitest, Jest, Playwright Test, Mocha, Chai, pytest, Supertest, ESLint, Ruff, mypy

Observability & Analytics: PostHog, Sentry, Google Analytics, Vercel Analytics, Intercom, Winston Logging

Tools & Platforms: Git, GitHub Actions, Prettier, Husky, lint-staged, Swagger/OpenAPI, FFmpeg, node-cron, MJML, Sanity CMS, ButterCMS, SuiteCRM, Twitter API v2, Modal (Serverless Python), GitPython, Jira, Trello, Asana

Web3 / Blockchain: Solidity, Truffle, ethers.js, web3.js, bitcoinjs-lib, WalletConnect, Coinbase Wallet SDK, Polygon

PROFESSIONAL EXPERIENCE

META

Machine Learning Infrastructure Engineer

September 2025 – Present

Architect ML experimentation infrastructure supporting Meta Avatars and Reality Labs research platforms used by hundreds of engineers.

- Architected large-scale C++ experimentation infrastructure enabling experiment tracking, debugging, and results aggregation across machine learning workflows used by ~300 engineers.
- Built distributed validation systems including end-to-end, regression, and integration testing frameworks preventing regressions across complex ML research pipelines.
- Developed centralized observability platform aggregating logs, debugging signals, and model outputs using Plog and internal infrastructure, enabling real-time monitoring by research teams and Reality Labs leadership.
- Integrated compute infrastructure including ALA servers, internal cloud environments, Skills tooling, and MTP developer platforms supporting large-scale ML experimentation.
- Automated distributed ML pipelines using Bash, Python, and Jupyter tooling across GPU-intensive compute environments.
- Engineered high-performance compute orchestration for systems utilizing 150+ GPU cores and ~2TB RAM, enabling avatar generation pipelines requiring ~55 minutes per render.

AI ENGINEER / TECHNICAL CONSULTANT — Freelance

January 2024 – September 2025

Partnered with startups and product teams to design and deploy AI-driven applications, automation systems, and scalable software platforms.

- Built LLM-powered sales automation platform integrating GPT-4 with Apollo APIs to automate prospecting, messaging workflows, and engagement validation.
- Developed AI content intelligence platform analyzing social media signals to identify viral content trends and strategic insights for businesses.
- Architected generative AI platform enabling personalized image generation and AI-driven creative tools.
- Developed automated video response system integrating generative AI tools including GPT-4 and HeyGen.
- Built Zoom productivity plugin integrating AI teleprompter functionality and automated FAQ generation for sales calls.
- Maintain and contribute to open-source C++ networking library resolving connectivity issues for Call of Duty: Black Ops 3 Mac OS release.

APPLE

September 2019 – June 2020 | December 2020 – December 2023

Senior Software Engineer

Delivered distributed backend systems and full-stack applications powering Apple Store platforms used globally.

- Architected scalable backend services and APIs supporting Apple Store applications using distributed systems and microservices architecture.
- Developed React-based user interfaces using Apple internal UI libraries aligned with Apple Design Guidelines.
- Implemented GraphQL data services and optimized backend data pipelines improving query efficiency and performance.
- Managed CI/CD deployment pipelines using Ansible and AWS infrastructure.
- Led cross-team engineering collaboration and Agile processes improving feedback loop efficiency by 32%.
- Mentored engineers and developed internal documentation improving onboarding speed and team productivity.

INSTAGRAM (META)

July 2020 – November 2020

Senior Software Engineer – Privacy & Compliance

Developed privacy infrastructure used across Instagram, Facebook, and WhatsApp platforms.

- Built Account Center privacy controls enabling proactive ad-category blocking aligned with EU privacy regulations.
- Designed and implemented GraphQL APIs and backend services using Python Django and Relay.
- Developed cross-platform UI rendering infrastructure supporting multiple Meta applications.
- Collaborated with privacy, compliance, and ads teams to ensure regulatory alignment and platform integrity.

GOOGLE

August 2018 – September 2019

Software Engineer – Search Infrastructure (Travel)

Built large-scale search infrastructure powering travel discovery across Google Search, Maps, and Assistant.

- Developed backend infrastructure supporting travel search features serving 55M+ monthly search result views.
- Implemented C++ grammar rules and NLP pipelines improving query intent recognition and search accuracy.
- Built internal developer tools used by 150+ engineers, reducing manual experimentation workflows by ~1 hour per day.
- Conducted large-scale A/B testing improving click-through rates while maintaining <0.01% global search performance impact.
- Supported emergency response search infrastructure enabling localized SOS alerts during global crises including wildfires and COVID-19.
- Developed distributed backend services in Go supporting travel data ingestion, indexing, and retrieval systems.

CHOPSTIX

June 2013 – July 2018

Software Engineer

Built and scaled full-stack food delivery platform supporting online ordering and logistics systems.

- Architected backend systems using Django and Stripe for secure payments and order processing.
- Developed responsive web and mobile applications using React, React Native, and Material UI.
- Designed scalable APIs and distributed systems supporting real-time order tracking.
- Implemented AWS and DigitalOcean deployments with Nginx load balancing and CI/CD automation.
- Applied test-driven development to improve reliability and maintainability across application workflows.

EDUCATION

HARVARD UNIVERSITY — Intro to Computer Science

NEW JERSEY INSTITUTE OF TECHNOLOGY — Bachelor of Science – Computer Science & Finance

UDACITY — Machine Learning Engineer Nanodegree · iOS Developer Nanodegree · React Native Developer Nanodegree

CISCO CERTIFICATIONS — CCNA Routing and Switching · Introduction to Networks · Routing and Switching Essentials

SALARY EXPECTATIONS

California State law requires the salary range to be provided on job listings and provided during interviews. Following industry standards; equity over compensation per market value.

EXAMPLES OF PROJECTS & COMPLETED WORKS

Personal website: maxawad.com

Google Search — when googling for travel, worked on the maps, reviews, and links that come up at the top of a search for this topic.