

Muhammad Umer Sohail

mailtoutersohail@gmail.com | linkedin.com/in/callmeumer | umersohail.dev

SUMMARY

Full Stack Engineer with 4+ years of experience building complex systems across healthcare, fintech, and AI. Currently building a healthcare case management platform at siParadigm handling AI triage, lab bidding, insurance automation, and compliance workflows. Previously built an autonomous AI browser agent from concept to beta, a real-time PII redaction system, and a financial trading platform with real-time market data streaming. Most comfortable working in small teams on hard problems where I have to own the system end-to-end.

EXPERIENCE

Backend Developer 2 at siParadigm (Jan 2026 – Present)

- Building a healthcare case management platform from the ground up with a small team
- Hooked up the AI triage system using message queues and async task processing so clinical diagnosis models feed directly into the case workflow
- Built the full lab ecosystem: registration, licensing checks, fee scheduling, per-payer insurance catalogs, and a bidding engine that scores labs on geolocation, network status, and other criteria then auto-awards the top pick
- Developed ICD-10 to CPT mapping (AI-assisted) that filters labs by the procedures they actually perform and scores them into the bidding pipeline
- Built the document and attachment layer handling test results, images, and medical records across cases and claims
- Integrated FedEx API for medical kit shipping: label generation, tracking, and delivery updates

Backend Developer at siParadigm (Jul 2024 – Dec 2025)

- Reduced microservice orchestration time from 2 seconds to 110ms by replacing chained HTTP calls in the SAGA architecture with gRPC
- Built a real-time PII redaction system that runs live during screen recording without any LLM, using OCR, pattern matching, and intelligent frame analysis
- Developed a workflow recording system that captures screen, audio, and user I/O, builds an action graph from multiple sessions, and replays tasks from natural language input
- Built the recorder's tool suite: replay engine, file management, system control, and I/O capture, enabling the agent to interact with the full desktop environment
- Built field-level AES-256 encryption by overriding Django's ORM field classes. Decryption is controlled dynamically through RBAC, with deterministic encryption on searchable columns
- Developed a comprehensive AI evaluation pipeline tracking true/false positives, false negatives, token usage, and accuracy breakdowns across multiple dimensions
- Implemented microservice authentication with RBAC through Traefik with dynamic routing. Requests are authenticated before reaching internal services

Full Stack Engineer at Agileloop Ltd. (Feb 2024 – Jun 2024)

- Led development of an AI browser agent that executes multi-step browser tasks from natural language, no user intervention needed
- Built the core execution engine that reads multiple input signals and acts on them in real time
- Set up async background processing so the next action is ready before the current one finishes
- Added a verification loop where each action is validated before the agent moves to the next step
- Made the agent undetectable: randomized timing, natural cursor movement, stealth browser sessions
- Took the product from concept to beta; owned the core agent logic
- Co-developed the Chrome extension frontend
- Maintained backend services: auth, billing, usage metering, and AI model integration

Backend Engineer at Agileloop Ltd. *(Sep 2023 – Feb 2024)*

- Built the backend from scratch for an AI-powered computer-use agent that executes tasks on behalf of users through natural language
- Designed a microservices architecture with Node.js/Express handling auth, billing, and account management, and Django handling AI model integration and agent logic
- Implemented OAuth-based multi-account linkage (Google, Zoho, and others), enabling the agent to map and call third-party APIs on the user's behalf
- Designed complex data models across PostgreSQL and MongoDB to support user sessions, billing records, token usage, and agent activity logs
- Integrated Stripe for billing, token-based usage metering, token auditing, and recurring payment management
- Integrated OpenAI models with prompt tuning, context management, response caching, and token streaming
- Built conversation history and session state persistence with end-to-end agent activity logging and monitoring
- Implemented rate limiting and request throttling to protect services under load
- Set up the CI/CD pipeline with Dockerized containers and automated deployments
- Wrote unit and end-to-end test suites across services

Full Stack Developer (Freelance) at Biads Consulting Ltd. *(Nov 2021 – Jan 2023)*

- Built the backend for a real-time portfolio management platform serving ~700 users
- Integrated Finnhub WebSocket streams for live financial data ingestion
- Developed portfolio tracking and management APIs for user holdings, performance, and alerts
- Implemented prediction features using technical indicators and in-house vector-based models
- Worked with a microservices architecture using Django, PostgreSQL, Redis, and Kafka
- Wrote unit test suites across services to maintain reliability
- Deployed and managed production workloads on Google Cloud Platform

EDUCATION

Bachelor of Computer Systems Engineering

University of Engineering & Technology, 2023

SKILLS

Python, JavaScript, TypeScript, Django, FastAPI, Node.js, Express, React, PostgreSQL, MongoDB, Redis, Kafka, RabbitMQ, Celery, gRPC, Docker, GCP, Traefik, Stripe, OpenAI APIs, Git