

ASIF IQBAL

+91 8789282509 ✉ asifiqbal7.iitkgp@gmail.com 🌐 github.com/maxasif

Education

Indian Institute of Technology, Kharagpur

Bachelor of Technology in Chemical Engineering

Aug. 2016 – April 2020

CGPA – 8.87

Summary

Software engineer with 6+ years designing scalable backend and cloud systems, with growing expertise in AI/LLM tooling. Track record of shipping reliable, production-grade platforms that cut costs and boost efficiency for enterprise clients

Technical Skills

Languages: Rust, Python, SQL, JavaScript, TypeScript

AI/ML Technologies: LLM gateway/proxy design (Anthropic, OpenAI, Gemini APIs), Model Context Protocol (MCP), agentic coding tools (Claude Code, Cursor, Copilot, Codex), ASR (NVIDIA Parakeet, OpenAI Whisper), real-time inference

Cloud & Infra: AWS (Nitro Enclaves, EC2, S3, EKS, ECS), GCP (Dataflow, BigQuery, Pub/Sub), Azure (Confidential VM, AKS), Docker, Kubernetes, Kata Containers, Grafana

Frameworks & Databases: Tokio, Effect-TS, Flask, Django, React, PostgreSQL, MongoDB, BigQuery, ClickHouse

Experience

Oblivious

May 2023 – Present

Software Engineer

Bangalore, Karnataka

- Architected a multi-provider LLM gateway/proxy (Anthropic, OpenAI, Gemini) fronting AI coding tools (Claude Code, Cursor, Copilot, Codex), serving 15K requests/day across 5 engineering teams
- Built prompt/tool-call tracing that links every Jira ticket, PR, commit, and diff to the underlying AI prompts, skills, and MCP tools used — giving leadership full auditability of AI-generated code
- Designed a unified request/response schema normalizing Anthropic, OpenAI, and Gemini APIs, with model routing, per-team rate limiting, and in-flight request intervention for policy enforcement
- Built PII/secrets redaction that detects and tokenizes sensitive data (API keys, passwords, credentials) in prompts before they reach external LLM providers, preventing data leakage from AI coding tools
- Shipped a usage analytics pipeline surfacing per-team, per-repo AI adoption metrics, used by engineering leadership to quantify AI ROI and prioritize tool rollout
- Created OBLV Deploy supporting AWS Nitro Enclaves, Azure Confidential VMs (AMD SEV-SNP attestation), and NVIDIA H100 confidential GPUs for multi-cloud confidential ML workloads
- Built a Layer 4 Rust proxy (TCP/UDP over vsock), improving enclave service throughput 30% and enabling secure external connectivity

Anakin Inc

January 2023 – May 2023

Software Engineer

Bangalore, Karnataka

- Engineered a resilient scraping proxy with automated rotation and fallback, increasing scrape success rate by 35%
- Used LLMs to automate product matching across 7+ e-commerce platforms, improving price-comparison accuracy for 50,000+ products

Amplus Solar by Petronas Group

July 2020 – August 2022

Senior Associate Software Engineer

Gurugram, Haryana

- Developed a custom CRM integrating 5+ internal services with advanced user management, doubling user adoption year-over-year
- Built a fault-tolerant task-queue service (Flask, Docker) on a serverless platform, sustaining 5K jobs/day
- Created an ETL pipeline on Google Dataflow ingesting 20 GB/day into BigQuery, powering downstream analytics

Projects

Rust Terminal Editor — Rust, crossterm | github.com/MaxAsif/rust-editor

- Built a terminal text editor from scratch with syntax highlighting, incremental search, and full Unicode grapheme support, using a modular component-based UI architecture

EODS Live Transcript Demo, Oblivious | *Python, ASR*

- Designed secure, real-time transcription using an attestation gateway for user/enclave authentication, processing audio with NVIDIA Parakeet and OpenAI Whisper
- Developed Python backend and Gradio UI supporting scalable, compliant audio capture and transcript visualization