

MUHAMMAD USMAN

Python Backend & AI Engineer

Islamabad, Pakistan | muawan2001@gmail.com | [LinkedIn](#) | [GitHub](#) | [Website](#)

SUMMARY

Python and AI engineer with three years of experience building backend services, production data pipelines, RAG systems, AI agents, and MCP integrations. Owns an automated regulation ingestion pipeline that processes over 250k regulations per month on Kubernetes with KEDA-driven autoscaling. Developed LangGraph-based agents for complex workflows and an MCP gateway enabling external AI agents to invoke product features directly. Specializes in backend AI product work, including messy data ingestion, accurate retrieval, and reliable, scalable systems.

EXPERIENCE

AAI Innovations GmbH

Software Engineer

Feb 2025 - Present

Remote, Germany

- Build Python backend services, data pipelines, and AI agent workflows for CORA, a regulatory intelligence platform, owning features from technical design through production deployment.
- Engineered the automated document ingestion pipeline (document parsing, semantic chunking, LLM-based image captioning, vector embeddings) that processes 250k regulations per month, running on Kubernetes with KEDA-driven autoscaling.
- Built MCP gateway hosting 5 MCP servers over Streamable HTTP, letting customers' AI agents (Claude, Cursor) call individual product features, secured with OAuth 2.1 (CIMD) with per-user credit metering.

Maanz AI

Software Engineer

Aug 2024 - Feb 2025

Islamabad, Pakistan

- Built and maintained Apache Airflow DAGs and Python services for high-volume daily data ingestion, transformation, and analytics, including vehicle homologation and safety-critical automotive data.
- Improved throughput and reliability across these workflows, cutting manual fixes and keeping results correct and traceable under strict quality requirements.

Condoroo

Software Engineer (Part-time)

Aug 2023 - Aug 2024

Remote, Portugal

- Built the company's first customer-facing AI chatbot on WhatsApp (Botpress plus custom code). Customers verify who they are, then get live account details from Airtable or file complaints with no staff involved. Thousands still use it.
- Built Airtable automations and Node.js backend workflows, including a document parsing pipeline, which streamlined document handling and supported more efficient operations

HIGHLIGHTED PROJECTS

CORA: AI Regulatory Compliance Assistant

AAI Innovations

- AI agent platform for regulatory compliance, covering 250k+ regulations from the EU, US, UN, ISO and others.
- Built agentic workflows with LangGraph and Vertex AI, using LangSmith for tracing, debugging and evaluation.
- Designed RAG pipelines using MongoDB Atlas Search for hybrid (vector + keyword) retrieval, delivering accurate, low-latency results across the full regulatory corpus.

MCP Gateway

AAI Innovations

- Built the MCP layer that lets customers plug specific product features into their own AI agents without buying the full product.
- One gateway hosts 5 MCP servers over the full Streamable HTTP protocol, working with clients like Claude and Cursor.
- Implemented OAuth 2.1 with Client ID Metadata Documents (CIMD), which replaces Dynamic Client Registration, plus per-request authorization.
- Tracks credit usage per user and routes each request to the existing product services, returning the results to the agent.

TECHNICAL SKILLS

Languages: Python, JavaScript

AI & LLMs: LangGraph, Google ADK, MCP (Streamable HTTP), LangSmith, RAG, OpenAI and Anthropic APIs, Vertex AI

Backend: FastAPI, REST APIs, OAuth 2.1

Data & Storage: MongoDB (Atlas Search), PostgreSQL, Apache Airflow, ETL, web scraping

Infrastructure: Kubernetes, KEDA, Docker, GitHub Actions, CI/CD, Git

EDUCATION

B.E., Mechatronics Engineering

National University of Sciences & Technology (NUST), College of E&ME

Nov 2020 - Jun 2024

Rawalpindi, Pakistan