ZAKARIYA MANSURI

AI Engineer · Ahmedabad, India
mozakariyamansuri@gmail.com · +91 63558 34077 · Portfolio · LinkedIn · GitHub

SUMMARY

AI engineer who ships production LLM agents working directly with the client's business and IT teams, from requirements to launch.

EXPERIENCE

AI Engineer — Multi-agent AI assistant for mutual fund advisory Sept 2024 – Present

Estatic Infotech (Aug 2025 – Present) · The One Technologies (Sept 2024 – Jun 2025) · used by a national distributor's brokers

- Took ownership when speech-to-text lacked the multilingual support the client expected, two days before launch at their national conference; delivered multilingual voice input and output (Sarvam Saaras speech-to-text, Bulbul text-to-speech) to production in time for the launch.
- Worked directly with the client's business and IT teams to turn requirements into features, e.g. deterministic guardrails on text and image requests, and broker details placed onto AI-generated marketing banners in code so names and contact details are always exact.
- One of two AI engineers who designed the system together; built 12 of its 17 agents (40+ tools) on FastAPI with Socket.IO streaming, Redis session memory and ChromaDB retrieval.
- Built a RAG pipeline over 150,000 transcribed Q&A pairs from the client's voice-call data, filtering and structuring them with rule-based checks and LLM review before indexing.
- Built a document agent that extracts and analyses brokers' own files (PDF, images, Excel, CSV, DOCX), using metadata filters to keep each user's documents separate, and a market-summary agent updated monthly from fund house (AMC) and client documents, with developer review.
- Chose the production models (GPT-4.1 Mini, GPT-4o Mini, gpt-oss-20b) after comparing several models and providers; tested agent routing against prepared scenario sheets before each release, and later automated those tests.
- Wrote the system documentation: architecture, API reference, deployment and broker onboarding.

AI Engineer Intern — Chaitanya AI Jan 2024 – Apr 2024

- Built Docvision, invoice field extraction with YOLOv8 + OCR on 200+ invoices across 6 classes (mAP50 0.764, precision 0.763, recall 0.728), with a Flask upload-and-review interface.

PROJECTS

- **Multilingual video dubbing:** Whisper transcription → transformer translation → Bark speech → Wav2Lip lip-sync, across Indian and international languages.
- **AI Solution Planner:** six-agent system (planner, strategist, researcher, writer, editor, validator) that turns goals and constraints into solution blueprints (Phidata, Flask).

SKILLS

LLM & agents: multi-agent systems, tool calling, RAG (ChromaDB), deterministic guardrails, OpenAI, Groq, LangChain, CrewAI, Phidata

Backend & data: Python, FastAPI, Flask, Socket.IO / WebSockets, Redis, SQL Server

Speech & vision: Sarvam AI (speech-to-text, text-to-speech), Whisper, Hugging Face Transformers, YOLOv8, OpenCV, TensorFlow, scikit-learn

EDUCATION

MSc, Artificial Intelligence & Machine Learning — Gujarat University · CGPA 7.73/10 2022 – 2024

BSc, Artificial Intelligence & Machine Learning — Gujarat University · CGPA 7.71/10 2019 – 2022