# Sanidhya Singh

**AI/ML Engineer | GenAI, RAG & Multi-Agent Systems**

Indore, India | +91 92430 63155 | sanidhya018@gmail.com | linkedin.com/in/sanidhya-aiml | github.com/sanidhya-ai-ml

## Professional Summary

AI/ML Engineer with 2+ years of experience shipping production GenAI systems: RAG pipelines, multi-agent orchestration (LangGraph) and LLM-driven automation. Currently building an Azure OpenAI + Milvus ticket-resolution agent for a global enterprise client. Previously built a clinical-trial verification platform end to end, plus MLOps anomaly detection and edge computer vision systems. M.Tech in Data Science (2026). Open to remote and on-site roles.

## Technical Skills

**GenAI & LLMs:** RAG, LangChain, LangGraph, Multi-Agent Systems, MCP Servers, Azure OpenAI, OpenAI API, Prompt Engineering, Hybrid Search, Fine-tuning (LoRA/PEFT), RAGAS, Hugging Face

**Vector Databases:** Milvus, ChromaDB, Pinecone, FAISS, Embeddings

**ML & Deep Learning:** PyTorch, TensorFlow, Scikit-learn, Transformers, NLP, OCR, NER, Pandas, NumPy

**Computer Vision:** YOLOv3, DeepSORT, MTCNN, OpenCV, TensorRT

**Backend & APIs:** Python (Async, Pydantic), FastAPI, REST APIs, Microservices, Celery, Redis, PostgreSQL, SQL, NoSQL

**MLOps & Cloud:** Azure AI, GCP, AWS S3, Docker, Kubernetes, MLflow, Apache Beam, Dataiku, Git

## Professional Experience

**AI/ML Engineer** | *BestPeers* Jun 2026 - Present

- Automated first-line IT and access support for a global enterprise client across multiple business-application use cases: an AI agent triages each ticket, auto-resolves what it can and escalates the rest to a human agent.
- Engineered **hybrid search over Milvus** with embeddings to match each ticket to the right resolution template; when nothing similar exists, it falls back to a prompted **Azure OpenAI GPT** response.
- Designed a three-tier fallback (template, then LLM, then human agent) in **Python** services, so no ticket is left without a resolution path.

**Data Scientist Intern / Jr. Data Scientist** | *MoreYeahs Technologies, Indore, India* Jun 2022 - Apr 2024

- Shipped **Sentry-B**, a real-time edge surveillance system running at **25+ FPS** on a Jetson Nano: person detection (YOLOv3), tracking (DeepSORT), face extraction (MTCNN) and identity verification from live CCTV, optimised with TensorRT.
- Moved **Dish/Sling TV** operations from reactive to proactive monitoring: Dataiku ETL over logs from **10 critical machines** into an AWS S3 data lake, an **Isolation Forest** model that predicts failures, and MLflow tracking with automated alerting dashboards.
- Delivered OCR-based financial-document pipelines and LLM-powered NLP applications as **FastAPI** microservices on **GCP**, tuning extraction accuracy directly with clients.

## Projects

**GenAI Clinical Trial Source Data Verification Platform** | *LangGraph, RAG, MCP Servers, OCR, NER, FastAPI, Celery, PostgreSQL*

- Replaced a fully manual, human-driven clinical-data verification process for a life sciences client with a **LangGraph multi-agent** pipeline that maps source documents to **EDC records in CDISC ODM XML**, eliminating manual cross-checking.
- Built a **RAG + OCR + NER** document-intelligence layer that extracts clinical facts from PDFs and links each one to structured database records with **full evidence traceability**, enabling auditable, source-backed verification at scale.
- Developed a **mismatch-detection engine** that surfaces discrepancies between extracted clinical facts and database entries, and integrated **MCP servers** to expose tools to agents as modular, composable components.
- Owned the project solo from requirements to production: chose LangGraph for orchestration, designed the RAG retrieval and chunking strategy, and delivered async **FastAPI** microservices with **Celery** task queues for concurrent processing of large document batches.

**AI-Powered Resume Optimization Engine** | *LangChain, ChromaDB, FastAPI, Pydantic, OpenAI API, Embeddings*

- Built a **RAG-based resume evaluation tool** (LangChain, ChromaDB, embeddings) that semantically compares resumes against job descriptions, surfacing genuine skill gaps rather than surface-level keyword mismatches.
- Designed a custom **ATS scoring engine** that jointly evaluates keyword alignment, skill relevance and experience context to rate how well a resume fits a role.

## Education

**M.Tech in Data Science** | DAVV University, School of Data Science & Forecasting, Indore 2024 - 2026

**B.Tech in Information Technology** | Medi-Caps University, Indore 2019 - 2023

## Certifications

**The MCP Course: Unit 1 - Fundamentals of MCP** | Hugging Face Apr 2026

**Microsoft Azure AI Essentials: Workloads & ML on Azure** | LinkedIn Learning Apr 2026