


# Mohammed Tahir Siddique

[mohammedsiddique729@gmail.com](mailto:mohammedsiddique729@gmail.com) | 9141500407

 [TahirSiddique092](#)

 [mohammed-tahir-siddique](#)

## Education

Bachelor of Technology in AIML | **PES University**, Bangalore CGPA: 8.32 | 2024-28  
Class XII (PUC) | **Sri Sapthagiri Pre-University College** 96% | 2023-24  
Class X | **The Presidency Public School** 91% | 2021-22

## Experience

**Founding Engineer | GoSquad** - UK Based EdTech Startup July 2025 - Present

- Core backend engineer in a 5-member founding team, working directly with founders on product architecture and AI-driven feature design.
- Designed and led the core video generation engine, a multi-stage RAG pipeline intersecting learning beats and story structure blueprints, feeding Writer and Director LLM agents for scene-by-scene script generation, assembled via Remotion Lambda, achieving under 4 minutes latency at \$0.25 per video.
- Designed and led Ben, an adaptive AI tutor on Gemini with cognitive state estimation and dynamic system prompt adaptation, backed by a Redis job queue with 5 parallel workers and global API rate limiting, deployed on a VPS using Docker.

## Projects

### Vesper - AI Agent Infrastructure Framework (Active Development)

- Open source infrastructure layer for AI agent fleets that handles lifecycle management, persistent memory, cost tracking, and multi-agent coordination through a single YAML spec.
- Designed a 7-layer architecture: Agent Registry, Lifecycle Manager, Memory Engine, Communication Bus, Eval & Quality Gate, FinOps Engine, and Security & Guardrails.
- Built in Python with SQLite, gRPC communication bus, and a CLI using Click. Works alongside LangGraph, CrewAI, and the Anthropic SDK without replacing agent logic.

### Forge - One Command Deployment CLI for Hackathons

[\[website\]](#)

- Built a CLI tool that deploys a full stack project in one command, automating backend deployment to Railway and frontend deployment to Vercel without touching any dashboard.
- Architected a FastAPI backend with an async Redis-based job queue coordinating Dockerized workers to execute build and test commands in isolated environments.
- Shipped a React frontend dashboard for monitoring deployment status, build history, and live logs across projects.

### Nano PESU - RAG Chatbot for PES University.

[\[website\]](#)

- Building a RAG-based question answering system to answer student academic queries using college-provided PDF materials with contextual citations.
- Implemented the entire backend ingestion pipeline in FastAPI, including PDF upload with metadata, text extraction, cleaning, and sentence-level chunking.
- Generated embeddings using Sentence Transformers, structured vector storage in ChromaDB, and collaborating on Ollama integration for context-aware responses.

## **Skills** ---

**Languages:** Python, JavaScript, C/C++

**Backend & Frameworks:** FastAPI, Node.js, Express.js

**AI/ML:** RAG, Sentence Transformers, ChromaDB, RASA

**DevOps & Tools:** Git, GitHub, Docker (Basic), Redis, CI/CD, n8n