

Sahil

Backend & AI Engineer

📍 Haryana, India | 📞 +91 9588132980 | ✉️ sahil95881@gmail.com
🌐 linkedin.com/in/sawhill | 🐙 github.com/Sahil-958

Professional Summary

Backend & AI Engineer specializing in Python, high-throughput microservices, and Agentic AI/RAG architectures. Proven track record architecting enterprise procurement platforms, building resilient scraping engines with Playwright, and integrating LLMs with PostgreSQL (pgvector) for semantic search and caching. Adept at designing scalable systems with FastAPI, robust authentication (RBAC/OAuth2), and containerized deployments.

Professional Experience

AI/ML Trainer & Backend Engineer

Fusionpact Technologies

June 2025 – June 2026

Noida, UP

- **3-Tier Microservices Architecture:** Engineered and integrated 3 dedicated backend services (*TenderAxis Core*, *Data Scraper API*, and *Agent API*) using FastAPI and Node.js to orchestrate real-time ingestion, processing, and evaluation of enterprise procurement data.
- **Agentic AI & Document Intelligence:** Developed autonomous multi-agent workflows using LLMs (OpenAI, Gemini) and Agno framework to parse complex IFB/RFx/BOQ documents, converting unstructured PDFs and spreadsheets into structured JSON schemas.
- **Automated Portal Ingestion:** Architected resilient Playwright scraping engines to extract real-time bids from government e-procurement portals (NEA, GPL, CPPP, GeM), implementing automated CAPTCHA and OTP handling for continuous data ingestion.
- **Semantic Caching & Vector Search:** Implemented vector similarity search and semantic caching with PostgreSQL (*pgvector*), slashing API latency by 30% and lowering recurring LLM inference costs.
- **Security & Enterprise Integration:** Implemented Role-Based Access Control (RBAC), OAuth2/JWT authentication, and integrated automated evaluation pipelines with internal Document Management Systems (DMS) and financial scoring tools.

Featured Projects

Agentic AI Framework (agentic-api) | *FastAPI, Python, pgvector, Docker, Agno, LLMs*

- Architected a containerized multi-agent backend providing asynchronous RAG analysis for multi-format enterprise documents (PDF, DOCX, XLSX).
- Implemented semantic caching and vector indexing using PostgreSQL with pgvector, optimizing response times and reducing compute overhead.
- Engineered Docker Compose configurations and automated deployment workflows for seamless containerized production orchestration.

IMS - AI-Powered Platform | *Node.js, React.js, Transformers.js, Hugging Face, Bash*

- Built a full-stack image intelligence platform featuring automated captioning, object detection, and OCR via local on-device neural inference.
- Engineered a parallelized Bash data processing pipeline integrated with Node.js for high-throughput batch media processing and data extraction.

Study Companion / Telemetry Agent | *Python, Flutter, RAG, Asynchronous APIs, Linux APIs, SQLite*

- Developed a modular system service and mobile client with real-time background telemetry, process monitoring, and device-level API hooks.
- Integrated an asynchronous RAG-driven AI assistant providing context-aware analytical insights and automated task management.

Technical Skills

Languages: Python, TypeScript, JavaScript, SQL, Bash, C++, Dart

AI & Data: RAG, Agentic AI, LLMs (OpenAI, Gemini), Agno Framework, pgvector, Hugging Face, NLP, Transformers.js

Backend Frameworks: FastAPI, Node.js, Next.js, Express.js, SQLAlchemy, Alembic, Redis, RESTful APIs, Microservices

DevOps & Databases: PostgreSQL, Docker, CI/CD (GitHub Actions), Linux (Arch/Ubuntu), Git, Nginx, RBAC, OAuth2/JWT

Core Competencies: Distributed Systems Architecture, High Concurrency, Asynchronous Programming, Semantic Search, System Design

Education

Master of Computer Applications (MCA) | 8.67 CGPA

Chaudhary Ranbir Singh University

2023 – 2025

Jind, Haryana

Bachelor of Science (Hons) in Information Technology | 7.9 CGPA

Kurukshetra University

2020 – 2023

Kurukshetra, Haryana