SHRIRAM BHAT

AI AND DATA ENGINEER

Bangalore | shrirambhat10@gmail.com | www.linkedin.com/in/shrirambhat10 | +91 9380677947

Summary

AI & Data Engineer skilled in Python, Azure, and Microsoft Fabric, with expertise in scalable data pipelines, ETL, and LLM/RAG solutions. Experienced in secure API development, data cleaning, and automation to improve accuracy, efficiency, and analytics.

Professional Skills:

- Scalable data pipeline & ETL development
- Data cleaning, normalization & deduplication
- Secure API design
- AI/LLM & RAG pipeline integration
- Workflow automation & optimization
- Problem-solving, debugging & performance tuning
- Agile collaboration & cross-functional teamwork

Technical Skills:

- Python, pandas
- Azure (Document Intelligence, AI Search, Function App, App Services, Container Registry, Container Instances, Data Factory), Microsoft Fabric (Notebooks, Data Pipeline, Data Lake)
- LLM APIs, Embeddings, LangChain, Langraph
- Docling, Tesseract OCR
- FastAPI, Flask, JWT Authentication
- PostgreSQL, SQL, Data Lake, Vespa, FAISS (vector search)
- Docker, Git
- Bitrix24

Projects:

Business Card OCR & Excel Integration – 6 months

- Developed a Python-based ETL process in Microsoft Fabric Notebooks that ingests both Excel master data and OCR-extracted contact information from scanned business cards (via Tesseract OCR). Applied data cleaning, including normalization, deduplication, and fuzzy matching, to produce unified customer contact records. The cleaned, curated datasets were stored in the Fabric Lakehouse, with the entire workflow orchestrated

through Fabric Data Pipelines, automating ingestion and integration for CRM and analytics teams.
- Automated ingestion workflows, reducing manual work by 80% and improving contact accuracy by 25% for CRM & analytics teams.

RAG Pipeline for Complex PDFs – 5 months

- Developed a Python + Azure Document Intelligence + Multi Modal LLM APIs pipeline to extract text/tables/images from PDFs, summarize, and index in Azure Cognitive Search with semantic filters.
- Orchestrated with Azure Functions, optimized indexing to cut latency by 30% while improving result relevance for enterprise searches. Built the Flask app for the user query and created the image using Docker. Pushed to Azure Container Registry and deployed in Azure App Service

RAG Agent (Langraph + FastAPI + Vespa) - 6 months

- Created an end-to-end multi-tenant RAG system using Python, LangChain, Langraph, Vespa, PostgreSQL, OpenAI Embeddings, JWT auth, and Docker with hybrid ranking for precise retrieval. The Docker image is pushed to the Azure Container Registry and deployed in the Azure App Service.
- Achieved a +15% relevance improvement via optimized chunking/embedding, with secure per-user data isolation and logging.

Conversational Assessment Agent – POC (2 months)

- Built a Python, LangChain, FastAPI, and PostgreSQL-based AI interview agent with LLM-assisted & rule-based scoring, dynamic questions, and timed assessments.
- Enabled reusable templates for multiple hiring workflows, supporting scalable and unbiased candidate evaluation.

LLM-driven Data Cleaning – POC (2 months)

- Designed Python (pandas) + LLM APIs + FastAPI engine converting natural language to deterministic cleaning transformations with schema validation, rollback, and audit logs.
- Automated deduplication, normalization, and formatting, reducing data prep time by 70%.

Experience:

AI & Data Engineer – *Crossroad ELF DSS Pvt. Ltd.*

Aug 2023 – Aug 2025 | Bangalore, India

- Designed and implemented scalable data pipelines and ETL workflows using Python, pandas, and Microsoft Fabric Notebooks to extract, clean, and unify data from business

cards extracted using OCR, Excel files,

- Developed advanced Retrieval-Augmented Generation (RAG) solutions with LLM APIs, OpenAI Embeddings, LangChain, and Langraph, integrating Azure AI Search and Vespa vector search for semantic indexing and retrieval.
- Automated data normalization, deduplication, and transformation workflows using a natural language-driven LLM-powered engine, improving data accuracy and reducing manual effort by up to 80%.
- Built secure, multi-tenant APIs with FastAPI and Flask, implementing JWT authentication and a PostgreSQL backend for persistent data management and audit logging, supporting enterprise compliance.
- Orchestrated cloud-native workflows via Azure Functions and containerized deployments using Docker and maintained source control with Git to deliver robust, scalable AI and data engineering solutions.
-

Data Analyst Intern—*Rubixe*
 Feb 2023 – Jul 2023 | Bangalore, India

- Analyzed business data using Power BI and MySQL to create interactive dashboards and generate actionable insights for decision-making.
- Learned and utilized Python with pandas and NumPy for data manipulation, cleaning, and basic statistical analysis to support reporting and data-driven projects.

Mechanical Design Engineer—*Devault Innovative Solutions Pvt. Ltd.*
Sep 2021 – Nov 2022 | Hubli, India

- Designed mechanical parts and assemblies using Solid Edge and AutoCAD, creating precise 2D and 3D models to support manufacturing and product development.
- Collaborated with engineering teams to refine designs, ensure manufacturability, and meet technical specifications, improving product quality and reducing design cycle time.

Education:

Bachelor of Engineering in Mechanical Engineering—Visvesvaraya Technological University
CGPA: 8.7 | Graduation Year: 2021 | Bangalore, India

Certifications / Additional Learning

- Certification ("Data Analytics – Datamites")
- I learned Python, pandas, NumPy, Power BI, and MySQL through internships and projects.