Subhayu Kumar Bala

✉ balasubhayu99@gmail.com 📞 +91 93828 77751 🔗 subhayu99.github.io in subhayu-kumar-bala

 subhayu99

Intro

A Data & Infrastructure Engineer with 3+ years of experience building scalable, high-performance, intelligent systems that drive measurable business outcomes. I specialize in turning complex data problems into efficient, scalable software.

I leverage deep expertise in Python, SQL, and multi-cloud architecture (AWS/Azure/GCP) to deliver transformative results, from slashing a 27-hour data process to just 5 seconds to engineering production-grade agentic LLMs.

My work has directly contributed to critical business milestones for clients, including enabling a startup to secure a $6M Series A funding round and ensuring enterprise-grade data compliance for regulated industries.

Technologies

Core Languages & Practices: Python, SQL, BASH, JavaScript, Git, Agile (Scrum/Kanban), Code Reviews, Technical Consultation & Solution Design

Data Engineering & Orchestration: ETL/ELT Pipelines, Data Modeling, PySpark, Pandas, NumPy, DuckDB, DBT, Airflow, GCP Workflows, Kafka, RabbitMQ, Microsoft Fabric, ADF, Treasure Data

AI & LLM Engineering: OpenAI APIs, Gemini, LangChain, Agentic Architecture (MCP, A2A), RAG, Axolotl, LLM Fine-tuning (SFT, DPO), Prompt Engineering, Semantic Deduplication, TensorFlow, Keras, SerpAPI

Cloud & DevOps: AWS, Azure, GCP, S3, Docker, Kubernetes, Terraform, Pulumi, ARM Templates, CI/CD Pipelines, Jenkins, Azure DevOps, NGINX, Linux, Cloud Cost Optimization

Application & API Development: FastAPI, Django, DRF, Flask, REST APIs, Typer CLI

Databases & Vector Stores: PostgreSQL, MongoDB, BigQuery, MS SQL Server, MySQL, Neo4j, Chroma, FAISS

Security & Quality Assurance: Zero-Trust Architecture, Cryptography (E2EE, AES-GCM, RSA), IAM, OAuth2/JWT, SSO/SAML, RBAC, Secret Management, Pytest, Unittest, Pydantic, Pandera, Great Expectations

Specialized & Visualization: Geospatial Analysis (GeoPandas, PostGIS), CDPs, SFMC, Web Scraping, Looker Studio, PowerBI, Streamlit, Plotly, Matplotlib, Quantum Computing Simulation

Experience

FiftyFive Technologies, Data Engineer — Gurugram, India (Remote) — Jun 2022 – present

- **Project: QxLab** Engineered a production-grade agentic LLM system from scratch using fine-tuned Mistral/Llama models to achieve >95% accuracy in real-time API data processing. Orchestrated the GPU-accelerated deployment on Docker, creating a scalable foundation for AI model serving.
- **Project: LoopKitchen** Architected the foundational data platform from scratch on GCP for a high-growth startup. This work, including building multi-platform ingestion pipelines and real-time analytics, was instrumental in the company securing a $6M Series A funding round.
- **Project: CV Advisors** Delivered a high-impact PoC that re-architected a 27.5-hour legacy SQL procedure into a 5-second Python/DuckDB process. This >99.99% performance gain unlocked significant operational efficiency for the client.
- **Project: Johnson & Johnson** Implemented robust data pipelines for a major healthcare client, engineering complex consent logic and ensuring enterprise-grade data quality to meet strict data governance and privacy standards in regulated markets.
- **Project: Wade Insight** Enhanced core orchestration features for an Azure SaaS platform and led the migration of enterprise data pipelines from Azure Data Factory to Microsoft Fabric, managing ARM template adaptation and deployment.

- **Project: SlideNinja** (PoC for McKinsey Partner) Delivered a cutting-edge RAG platform for a McKinsey partner during the initial GenAI boom, leveraging the then-new GPT-3.5 API and GPT Researcher to demonstrate rapid, high-value AI adoption.

FiftyFive Technologies, Software Engineer Intern — Gurugram, India (Remote) / Jan 2022 – May 2022
- Contributed across the full development lifecycle, from building a geospatial backend platform that earned direct client commendation to establishing CI/CD pipelines on Azure DevOps for services supporting 50k+ daily users.

Personal Projects

DatasetPipeline — May 2025
- Developed a production-ready CLI tool for transforming messy datasets into ML-ready formats; supports SFT, DPO, semantic deduplication, and quality analysis with plugin architecture.
- Published on PyPI ↗ and open-sourced on GitHub ↗ with extensible architecture for custom loaders, formatters, and analyzers.

Smart Commit — May 2025
- Built an AI-powered CLI tool that generates context-aware git commits via OpenAI or Anthropic models, enhancing developer productivity and project history quality.
- Published on PyPI ↗ and GitHub ↗ with Model Context Protocol (MCP) server for direct AI assistant integration.

DocumentAccessPOC — Jan 2025
- Designed and built a zero-trust secure document system featuring end-to-end encryption (AES-GCM) and secure key exchange (RSA) to enforce data-level permissions.
- Built a FastAPI interface for secure document sharing and revocation, with the full project documented and open-sourced on GitHub ↗.

Publication

QuDiet: A Classical Simulation Platform for Qubit-Qudit Hybrid Quantum Systems — Mar 2023

Subhayu Kumar Bala, Turbasu Chatterjee, Arnav Das

10.1049/qtc2.12058 ↗ (IET Quantum Communication)

Education

CIEM, (B.Tech - Information Technology) — Kolkata / Jul 2018 – Jun 2022
- CGPA: **8.57**/10 (Certificate ↗)