

TIRTH SHAH

Data & AI Engineer | Agentic AI Engineer | Data Engineer

College Station, Texas | F-1 OPT EAD | Available Immediately | Open to Relocate | (979) 635-2045 | tirthcshah1999@gmail.com

[linkedin.com/in/tirth-chirayu-shah](https://www.linkedin.com/in/tirth-chirayu-shah) | github.com/Tirth-1999 | tirthcshah.com

SUMMARY

Data and AI engineer with 4+ years building **Python/SQL ETL**, **Data Warehousing**, and guarded **Agentic AI** across financial services, higher education, energy, and telecom. Delivered 10M+ daily-record systems, cut manual reporting about 95%, and shipped from orchestration through evaluation, deployment, and handoff

TECHNICAL SKILLS

Data Engineering: Python, SQL, PySpark, ETL, ELT, Airflow, dbt, Medallion Architecture, Data Quality, Data Warehousing, Star Schema, GCP, AWS, Databricks, Delta Lake, Unity Catalog, SQL Server, PostgreSQL, pandas

AI Engineering: Agentic AI, LangChain, LangGraph, Multi-Agent Orchestration, RAG, Deep Agents, ADLC, Tool Calling, Prompt Engineering, NL-to-SQL, Gemini, Claude, ChromaDB, Human-in-the-Loop, Model Evaluation

Applications & Delivery: FastAPI, Streamlit, REST APIs, Docker, GitHub Actions, Google OAuth 2.0, CI/CD, RBAC, Webhooks, Log Monitoring, Streamlit Cloud

Certifications: Azure Data Engineer Associate; Azure Administrator Associate; Azure AI Fundamentals; Azure Data Fundamentals; Azure Fundamentals; PSM I

AI Accreditations: Introduction to LangChain; Databricks AI Agent Fundamentals; Databricks Generative AI Fundamentals; Databricks Fundamentals

PROFESSIONAL EXPERIENCE

HCLTech | Global Engagement Management Intern

Feb 2026 - May 2026

Internship | Client: Verizon | [GitHub](#) | [Demo](#)

Remote, Texas

- Designed an **Agentic AI** network-operations architecture for the **Verizon** MNS account with **Multi-Agent Orchestration**, **NWDAF** signals, **RAG/runbook** context, **Human-in-the-Loop** approval, rollback, and audit controls
- Co-built a React/TypeScript 5G intelligence simulation that translated multi-domain diagnosis and remediation into stakeholder-visible workflows, KPI recovery views, and governed execution choices for leadership demos
- Authored **NWDAF-to-Agentic-AI ART** strategy materials mapping Observer, Diagnostic, Strategic, and Operator agents to service assurance, **LLM** gateway guardrails, cost-risk scoring, and controlled execution

Mays Business School (TAMU) | Data & AI Graduate Student Worker

Nov 2025 - Feb 2026

Part-time | [GitHub](#) | [Demo](#)

College Station, Texas

- Built the Mays Flex Online analytics platform with **Python ETL**, a 17-table SQLite Data Warehouse, **Data Quality** checks, and **Streamlit** delivery across 7 graduate programs and 3 cohorts, cutting manual reporting about 95%
- Shipped a Gemini, **ChromaDB**, and **LangChain Deep Agents** assistant with schema **RAG**, **NL-to-SQL**, **Prompt Engineering**, **ADLC Tool Calling**, and SELECT-oriented safeguards for non-technical directors
- Implemented **ARIMA**, **Prophet**, and **scikit-learn** forecasting plus marketing ROI analysis across about \$803K in spend, while owning requirements, OAuth access, releases, testing, runbooks, and handoff
- Engineered state-snapshot admissions and marketing **ETL** with dynamic sheet discovery, 40+ program-alias normalization, idempotent upserts, fiscal mapping, **Data Quality** validation, and spend-to-outcome joins

Texas A&M UES | Data Engineer Student Worker

Apr 2025 - Nov 2025

Part-time

College Station, Texas

- Orchestrated **Python/SQL Server ETL** across 13 projects and 15+ sources ingesting 50K+ daily energy, weather, ERCOT pricing, billing, and sensor records with **Data Quality** gates and retry controls
- Delivered an 8-stage CHP operations **ETL** pipeline that reduced a 3-hour daily reporting process to about 5 minutes with freshness checks, auto-retry, PDF capture, alerts, and distribution
- Consolidated operational silos into a quality-checked **Star Schema** Data Warehouse and **Power BI**-ready models, cutting typical query paths from 15 minutes to under 2 seconds and replacing 15+ Excel reports
- Designed free-tier **Databricks + SQL Server** architecture exposing queryable feeds for **Airflow**-style DAG orchestration of **Python ETL**, cutting failed-job recovery from 30+ minutes to under 2 minutes

Black Tie Concierge, Inc. | Founding Engineer Intern

May 2025 - Aug 2025

Internship | [Demo](#)

College Station, Texas

- Shipped BlackTieCars as founding engineer in 12 weeks using Next.js, TypeScript, Supabase Postgres, Stripe, Google Maps, and Vercel, taking the booking platform from zero to paid operations
- Designed the users, bookings, and payments data layer with constraints, triggers, RBAC, webhooks, and lightweight **ETL** of Stripe and form events into reportable ops records supporting \$10K+ revenue and about 3x early growth

Tata Consultancy Services | Data Engineer

Aug 2021 - Aug 2024

Full-time | Client: Equifax

Gandhinagar, Gujarat

- Migrated **Equifax** credit-data **ETL** pipelines from on-premises systems to **GCP** and built a normalized data fabric across 7+ partner categories handling 10M+ records per day with **Airflow** DAG orchestration
- Improved **Entity Matching** hit ratio about 45% and reduced lookup latency about 93% through canonical schemas, address normalization, match-key deduplication, indexed storage, and **Data Quality** validation
- Led a 5-person proof of concept modernizing 60+ SAS workflows into **Python** with **pandas**, **NumPy**, multiprocessing, shared configuration, and audit reconciliation, cutting runtime about 80% and compute cost about 50%
- Automated row-count, field-mismatch, outlier, and output-parity **Data Quality** checks with PII masking, cutting validation cycles from about 3 days to about 2 hours and protecting bank-facing reporting from silent drift

PROJECTS

AI-Powered Regulatory Document Classifier | [GitHub](#) | [Demo](#)

Won TAMU Datathon 2025; Claude/Gemini with OCR, PII detection, dual-model verification, **Human-in-the-Loop**, and **Model Evaluation** at 98% accuracy

Retail Analytics Lakehouse on Databricks | [GitHub](#) | [Case Study](#)

PySpark/SQL Medallion Architecture lakehouse with **Delta Lake**, **Unity Catalog**, 25+ **Data Quality** rules, **Star Schema** models, and governed BI KPI exploration

Sous - Personal Chef AI | [GitHub](#) | [Demo](#)

LangChain/LangGraph Agentic Workflows with **Tool Calling**, **Prompt Engineering**, Tavily search, Gemini/Groq vision, memory, and **Streamlit** delivery

EDUCATION

Texas A&M University | M.S. Management Information Systems

Aug 2024 - May 2026

GPA 3.90/4.00 | Focus: data engineering, BI, analytics, AI systems, project management

College Station, Texas

Sardar Vallabhbhai Patel Institute of Technology | B.E. Computer Engineering

Jun 2017 - Jul 2021

CGPA 9.50/10.00 | Coursework: OOP, DBMS, Data Mining, AI, Python, Java, Software Engineering

Vasad, Gujarat