

FEATURED PROJECTS

iMigrator Platform Evolution (v6.4 → v6.6/v7)

Python 3.12 • Polars • DuckDB • Snowflake Pushdown • Multi-DB Connectors (17+ Platforms) • AWS Bedrock • Claude • Jenkins

- High-throughput data validation platform adopted across 60+ enterprise projects; expanded core engine from 1,000 to 2,560+ LOC.
- Engineered Snowflake pushdown compute modules generating temporary views to execute reconciliation directly in-warehouse, eliminating multi-terabyte network data transfer.
- Built automated MD5 hash keyless fallback reconciliation, extraction adapters for MongoDB/DynamoDB document schemas, and automated DML execution blockers to prevent accidental destructive SQL operations.

DataCraft — Autonomous SQL & Data Engineering Assistant

Python • Google Agent Development Kit (ADK) • Gemini API • Firestore • Snowflake • BigQuery • PostgreSQL • Vertex AI

- Architected an autonomous multi-modal AI agent using the **Google Agent Development Kit (ADK)** during Google's 4-Hour "Build with Gemini" hackathon, earning the official Google Developer Badge & Credly Certification.
- Engineered a dynamic Firestore schema catalog (**list_tables**, **get_table_details**, **add_table**) registering database schemas, primary keys, and data types across multiple database platforms.
- Built automated SQL anti-pattern detection using **sqlparse** (flagging full scans, missing filters, and uncapped sorting) with dialect-specific optimizations for Snowflake, BigQuery, and PostgreSQL.
- Implemented multi-modal generation tools producing visual ER diagrams via **gemini-3.1-flash-lite-image** and animated Kafka event-streaming architecture videos via **gemini-omni-flash-preview** on Vertex AI with GCS storage.
- Integrated Vertex AI Memory Bank (**PreLoadMemoryTool**) for session-level dialect persistence, paired with **AgentEngineSandboxCodeExecutor** and a responsive FastAPI/A2UI card interface.

Enterprise Packaging & Multi-Tier Self-Updater

Python • PyInstaller (onedir/onefile) • Windows API • SQLite • UNC Network Shares

- Architected 3-tier zero-downtime hot-swap updater (**Launcher → Updater → Real App**) with build repair, rollback protection, and CLI maintenance utilities (**IMIG_Updater.cmd**).

Healthcare Payer Data Integration Pipeline

Informatica PowerCenter • Oracle SQL • XML / XSD • Flat File Processing

- Built multi-source ETL pipeline processing fixed-width and delimited patient insurance data, cleansing and routing errors to exception tables, and generating personalized XML welcome letters.

ARCHITECTURE & PRACTICES

Engine Architecture

AST parsing, multiprocessing, pushdown compute, dynamic RAM allocation.

Data Governance

DDL schema drift detection (strict/normalized/loose), STTM & Data Vault audits.

Resilient Deployment

Zero-downtime hot-swap binaries, rollback protection, UNC share sync.

Pipeline Concurrency

Collision-free multi-script runs via Jenkins lockable resources & dynamic envs.

EDUCATION

B.Tech in Computer Science & Engineering (AI)

Noida Institute of Engineering & Technology (NIET), Greater Noida
2021 – 2025 | CGPA: 7.18

Intermediate & Matriculation (CBSE)

Jesus & Mary Academy, Darbhanga, Bihar
2018 – 2021