

ASHIK ROSHAN I
Data & AI Engineer – L2

Email :ashikroshan261@gmail.com

• [Github](#) • [Linkedin](#) • [Medium](#) • [Website](#)

TECHNICAL SKILLS

Languages: Python, SQL, JavaScript, HTML, CSS	Cloud – Azure: Blob Storage, Data Lake, SQL Database, Azure OpenAI, Databricks, ADF, Azure VM, Azure Document Intelligence, Function App
Cloud – AWS: S3, Lambda, Glue, Step Functions, EC2, CloudWatch, Textract, Bedrock	Data Engineering: PySpark, DBT, Informatica, Snowflake, Pandas, ADF
Databases: SSMS, pgAdmin, MySQL, Oracle, SQL Server, Snowflake	BI & Analytics: Power BI, ThoughtSpot, Plotly, Streamlit
AI & GenAI: Azure OpenAI, Amazon Bedrock, Gemini 2.5 Pro, LangChain, Neo4j, RAG, Prompt Engineering, Azure Document Intelligence	Automation & Web: Selenium, Web Scraping, Web Crawling, FastAPI, PyAutoGUI, Apify, Flask
DevOps & Tools: GitHub, Azure DevOps, CI/CD, PuTTY, ServiceNow, Rally, SharePoint	Libs & Frameworks: asyncio, PyVis, PyPDF2, pyodbc, Snowflake Connector, xmldict, smtplib, oracledb, pywin, pytesseract, dotenv

PROFESSIONAL EXPERIENCE

Data Engineer – L2 Optisol Business Solutions	April 2025 – Present
Data Engineer – L1 Optisol Business Solutions	August 2024 – March 2025
Data Engineer Intern – L0 Optisol Business Solutions	March 2024 – July 2024
Trainee – Software Engineer Blue Cloud	June 2023 – March 2024

EDUCATION

Bachelor of Engineering in Computer Science KLN College of Engineering Grade: A+	2019 – 2023
--	-------------

CERTIFICATIONS

- ❖ SnowPro Core Certification – Snowflake
- ❖ Microsoft Certified: Azure Data Fundamentals – Microsoft
- ❖ Databricks Lakehouse Fundamentals – Databricks
- ❖ dbt Learn Fundamentals – dbt Labs
- ❖ SQL (Basic) Certificate – HackerRank
- ❖ Academy Accreditation: Generative AI Fundamentals – Databricks
- ❖ 100 Days of Code: Python Pro Bootcamp – Udemy
- ❖ Snowflake Masterclass – Udemy

AWARDS & RECOGNITION

OptiSol Business Solutions | 2024 – Present

◆ Most Valuable Person (MVP) Award | 2024–2025

Highest organizational honor for consistent performance excellence, leadership in cross-functional collaboration, and contributions toward the company's long-term vision and project success.

◆ Spot Award – Project Excellence & Leadership | Jan 2026

Awarded by the CTO for mature professional handling of project success and fostering a culture of colleague recognition.

◆ Spot Award – RS ARP Project Go-Live | Nov 2025

Recognized for exceptional contribution to the "Beatty Go-Live" rollout and managing complex "Delta Load" activities under tight timelines.

◆ Spot Award – AI Tool Innovation (NotebookLLM) | May 2025

Took initiative to evaluate, demo, and drive team adoption of NotebookLLM to enhance project efficiency.

◆ Spot Award – Community Mentorship | Mar 2025

Delivered technical sessions for college students on interview preparation and emerging technologies.

◆ OKR Top Contributor (Q4) | Oct–Dec 2024

Honored for playing a pivotal role in attaining company-wide Objectives and Key Results (OKRs).

◆ Spot Award – Client Excellence (Ontology Mapping) | Dec 2024

Received high praise from client for a solution-oriented presentation on Ontology Mapping during an on-site visit.

◆ Spot Award – Gen AI & Automation | July 2024

Implemented a Generative AI-based solution to automate data inventory and served as SME resolving DBT blockers.

DATA ENGINEERING PROJECTS

1. Python-Based Data Migration: Google Sheets to Azure SQL Database

Client: Internal Business Team | **Company:** Blue Cloud

Built a Python-based ETL pipeline to migrate structured data from Google Sheets to Azure SQL Database. Implemented full and incremental loads, data transformation using Pandas, and optimized bulk inserts with pyodbc. Automated execution via Cron jobs on Azure VM and performed data validation and reconciliation.

Tech: Python, Pandas, gspread, pyodbc, SQL, Azure SQL, Azure VM, GitHub

2. On-Premises to Snowflake Data Warehouse Migration

Client: Republic Services (Waste Disposal Company) | **Company:** Optisol Business Solutions

Led Oracle/SQL Server to Snowflake migration, including schema conversion and CDC-based incremental loads. Developed ETL pipelines using Informatica and ELT transformations using dbt for ODS layer creation. Automated orchestration with AWS Step Functions and monitored via CloudWatch.

Tech: Snowflake, Informatica, dbt, Oracle, SQL Server, AWS (EC2, Step Functions, CloudWatch), GitHub

3. Enterprise Database Migration: Oracle to SQL Server (On-Premises)

Client: Republic Services (Waste Disposal Company) | **Company:** Optisol Business Solutions

Executed enterprise-scale Oracle to SQL Server migration using in-house ETL framework. Designed RAW → Staging → Mirror architecture, automated schema and data load scripts, managed constraints and identity columns, and performed comprehensive data validation. Supported UAT and production deployment through structured change management.

Tech: Oracle, SQL Server, Python, T-SQL, ETL Framework, ServiceNow, GitHub

4. API-Driven Data Migration: Podio to Azure SQL Database

Client: Jiffy (Cultural Exchange Agencies) | **Company:** Optisol Business Solutions

Developed API-driven data migration from Podio to Azure SQL Database using Python and REST APIs. Processed JSON data with Pandas, staged files in Azure Blob Storage, and built ADF pipelines across Staging → Production. Implemented high-performance bulk loading and ensured data reconciliation.

Tech: Python, REST API, Pandas, Azure Data Factory, Azure SQL, Azure Blob Storage, GitHub

AI-Driven Automation Initiatives

1. AI-Powered Automated Data Profiling Platform

Analyzes database schemas and datasets to generate AI-driven profiling insights and quality reports. Improves data governance by automating anomaly detection and intelligent data quality recommendations.

2. AI-Powered Pandas Code Generation & Self-Healing Data Transformation Agent

Generates dynamic Pandas transformation code from natural language using Azure OpenAI. Reduces manual scripting effort with automated error correction and self-healing execution.

3. Ontology Kit – Data Mapping Agent

Automatically maps source system data to destination schemas using AI-driven domain understanding.

Reduces manual data mapping effort by 40–50% and improves accuracy in system migrations.

4. AI-Powered Document Processing & Structured Data Extraction Tool

Converts scanned PDFs and images into structured JSON using Textract and Bedrock. Automates document data extraction and reduces manual processing time and errors.

5. Credit Risk Reporting & JSON Data Intelligence Platform

Processes and classifies large JSON credit risk data using async LLM orchestration.

Improves underwriting efficiency through automated mapping, enrichment, and risk block generation.

6. Knowledge Graph Builder (KGB) with RAG

Transforms unstructured data into interactive knowledge graphs with Neo4j and LLM integration. Enables structured querying and RAG-based insights from complex enterprise data sources.