KESHAV MAHESHWARI
DATA ENGINEER

CONTACT

S-300-H, Scheme 4-S,
New Loha Mandi, Macheda,
Jaipur, Rajasthan, India
302013
+916376065608
keshav.maheshwarics@gmail.com
Linked-In

EDUCATION

MCA • 2021 – 2023

BCA • 2018 - 2021

MAISM, JAIPUR

KEY SKILLS

| Microsoft Fabric | Python | SQL | ETL | Data Engineering | Hadoop | Apache Spark | PySpark | Azure Databricks | Azure Data Factory | Azure Data Lake | Generative AI | LLM | FastAPI | Data Modeling | Data Warehousing | API Development | Azure | KQL | AI & ML | Event Stream |

INTERESTS

Numismatics
Travel

CERTIFICATIONS

Microsoft Certified Fabric Analytics Engineer Associate: DP-600

Microsoft Certified: Azure AI Engineer Associate: AI-102

PROFILE

Results-oriented Data Engineer with 2+ years of experience in big data, machine learning, and cloud technologies. Skilled in designing, developing, and optimizing data pipelines and implementing ETL processes using tools like Microsoft Fabric and PySpark. Proven ability to apply AI and machine learning models to generate insights and enhance business performance. Adept at building scalable solutions for efficient data analysis, reporting, and management.

EXPERIENCE

MANDELBULB TECHNOLOGIES • OCT 2022 – PRESENT

Client/Project: Conventional SQL Warehouse Migration to Modern Cloud

• **Project Overview:** Led the migration of Fleet Energies' on-premises SQL data warehouse to **Azure**, improving processing speed and cost management. Converted SQL stored procedures into **PySpark scripts** and reconstructed the data model for BI reporting.

Responsibilities:

Designed the migration architecture from on-premises SQL to **Azure warehouse**.

- Migrated ETL processes to **Azure Pipelines**, extracting data from **local system & SQL Warehouse** to **Azure Data Lake Gen2**.
- Converted SQL stored procedures into **PySpark scripts** on **Azure Databricks** for optimized in-memory computation.
- Rebuilt the **data model** in **Azure** and integrated with **Power BI** for reporting.
- Reduced processing time from **2hr 50min** to **~5min** by optimizing pipelines with in-memory computation and **Delta Lake**.
- Developed **KPIs** from raw data to measure **IoT device efficiency** for business insights.

ACHIEVMENTS

• Led the implementation of a fiber splinter prediction project for a leading viscose staple fiber producer, which was recognized and published as a case study on Microsoft's official website.

Client/Project: Data Integration from Overseas Plant to Azure

Project Overview: Led the integration of data from the **Domsjo plant**'s local system to **Azure**, enabling data transfer to the **India unit's Azure environment**. Utilized **Azure SQL Warehouse** and **MS Fabric Lakehouse** for efficient ETL processing and reporting. Applied custom business logic using **PySpark** to enhance data transformation.

Responsibilities:

- Transferred data from the **Domsjo plant's local system** to an **Azure SQL Warehouse**, followed by migration to **MS Fabric Lakehouse** for ETL and reporting.
- Managed the seamless movement of data from the plant's local system to the **India unit's Azure environment** through **pipelines** and **gateways**.
- Designed and implemented robust **ETL processes**, transforming and loading data while applying custom business logic with **PySpark**.
- Developed a detailed **architecture diagram** to streamline data integration and ensure efficient workflows across systems.
- Created and delivered comprehensive **reports** for the **CO** and **CFO**, providing key insights to support informed business decisions.

Client/Project: BRSR Report

Project Overview: BRSR (Business Responsibility and Sustainability Report) for a world-leading manufacturer in **man-made fibre**. The project involved migrating data from the company's **SQL** database to **MS Fabric Lakehouse** and **SQL Warehouse** for ETL and reporting purposes.

Responsibilities:

- **ETL** migration of data from **SQL** to **MS Fabric Lakehouse** and **SQL Warehouse** for business reporting.
- Applied **PySpark** for custom business logic to transform and process the data.

Client/Project: Automated **fibre defect detection** using **deep learning** and **image segmentation** for improved **quality control**.

Responsibilities:

- Developed a **CNN model** on **Azure ML** for defect detection.
- Implemented **image segmentation** and stored results in **KQL** and **Fabric Lakehouse**.
- Built a **web app** for real-time defect detection, using **Microsoft Entra ID** for secure access.