

**MIEHLEKETO MATHEBULA**  
**CENTURION**  
**CELL: 0793653889**  
**EMAIL: [miehleketo93@gmail.com](mailto:miehleketo93@gmail.com)**

### **CAREER SUMMARY**

I am an accomplished data engineer with a strong track record of leading innovative data solutions that integrate AI to drive financial intelligence. My work bridges practical engineering and academic research, with a focus on AI engineering to enhance decision-making and governance in financial systems. I led the successful redesign of data processing workflows, achieving a 30% reduction in processing time and significantly improving data accuracy. Currently pursuing a PhD, my research investigates the intersection of artificial intelligence, engineering principles, and data-driven financial intelligence models to support smarter, more resilient financial systems. I excel at delivering scalable, intelligent data systems that turn complex data into actionable insights.

### **PROFESSIONAL EXPERIENCE**

#### **1. AVBOB: Lead Data Engineer (January 2025 – Present) (Full-time)**

- Lead the development and optimization of data pipelines for the enterprise Data Warehouse and internal systems.
- Manage and evolve the Data Engineering architecture to support growing data needs and strategic objectives.
- Implement robust data validation, reconciliation, and quality assurance processes to ensure trusted data assets.
- Design and develop scalable ETL processes that enable efficient data integration across multiple sources.
- Collaborate with cross-functional teams to analyse data and communicate actionable insights to stakeholders.
- Oversee data engineering projects from conception to deployment, ensuring alignment with organizational goals.
- Implement policies and practices for data quality, privacy, and compliance with legal standards.
- Provide leadership, mentorship, and foster a collaborative environment for skill development and knowledge sharing.
- Utilize Infrastructure as Code (IaC) to automate the deployment and management of data infrastructure.

#### **2. Absa Group: Data Engineer (January 2023 – December 2024) (Full-time)**

- Utilize Databricks for integrating data from diverse sources such as Hive and SQL Server, orchestrating workflows, and constructing feature stores for data sciences, effectively enhancing data accessibility and usability.
- Create and maintain a feature store in Databricks for data sciences, providing a centralized repository of curated features to accelerate model development and deployment.
- Create data tools for analytics and data science teams, facilitating the building and optimization of data sets to drive business insights and decision-making processes.
- Optimize data ingestion and processing pipelines using Delta Lake's optimization features, resulting in improved performance and scalability.
- Collaborate with cross-functional teams to design and implement robust data architectures leveraging Databricks Delta Lake for enhanced data reliability and efficiency.
- Implement and align to the Group Security standards and practices to ensure the undisputable separation, security, and quality of the organization's data.
- Assemble large, complex data sets that meet business requirements and manage the data pipeline.

- Build infrastructure to automate extremely high volumes of data delivery.

### **3. FNZ Group: Data Engineer (February 2022: - December 2022) (Full-time)**

- Analyse and understand requirements from stakeholders, and build up digital tools, workflows, and reports from scratch.
- ETL development into staging and data warehouse environments. Develop new data extracts and modify ETL processes as required to improve reporting datasets.
- Share collaboration best practices within the development team globally.
- Constantly improve and upgrade the collaborative tools.
- Transform FNZ's internal processes and systems into a digital business.
- Ingesting data from a variety of corporate data sources and performing required extract transformation and load to target databases.
- Data modelling capabilities, designing effective data models in line with the Agile methodology.
- Provide OLAP support and end-user training on the various cubes used for reporting downstream.

### **4. FNB South Africa: Data Engineer (April 2019 – January 2022) (Full-Time)**

- Design and development of ETL scripts using open-source software in a Linux environment.
- Perform Some operational support of batch jobs in production and responsible for EBucks monthly pay-outs.
- Build extensible data acquisition and integration solutions to meet the functional and non-functional requirements of the business.
- Implement processes and logic to extract, transform, and distribute data across one or more data stores from a wide variety of sources.
- Optimize data integration platform to provide optimal performance under increasing data volumes.
- Develop quality standards to ensure data quality and integrity across various database systems.
- Conduct a review of other integration developers' development efforts to ensure consistent methodologies are followed and to make recommendations where necessary.
- Work with Analysts and Business Users to translate functional specifications into technical designs for implementation and deployment.

### **5. SA Underwriters: SQL Developer (January 2019 – 2019 April) (Full-Time)**

- Design stable, reliable, and effective databases.
- Design, develop, and maintain appropriately scaled, effective solutions to support data cleansing, integration, and reconciliations.
- Deliver a robust, testable, repeatable framework of jobs and scripts for data management and producing reports to confirm success or failure.
- Build ETL packages for database maintenance and pay-out processes on Microsoft Server 2014 using SSIS.
- Creating Dynamic Reports using SSRS
- Ensure performance, security, and availability of databases.
- Ensure design and development consider all aspects of requirements - including data quality, customer experience, architecture, security, and operations.
- Ensure all development goes through appropriate approval, test, and deployment processes.

### **6. Momentum: SQL Developer (October 2016 – December 2018) (Full-Time)**

- Work with databases from a wide variety of vendors and database platforms.

- Dump findings into reports using SSRS.
- Troubleshoot SQL Queries
- Build ETL packages for database migration from DB2 Unix to Microsoft server 2008/2012 using SSIS.
- Creates Stored procedures and Views in Microsoft Data warehouse.
- Analysis of reporting requirements as defined by the business and documenting the desired solution and recommendation.
- Create Power BI reports based on business requirements with linked content, drill-downs, and nested parameters.
- Develop ad-hoc reports and automate them.

#### **EDUCATION**

- Doctor of Philosophy (Computer Sciences) University of Pretoria **(2025-Present)**
- Master of Science (Computer Sciences) University of Pretoria **(2023-2024) - Grade: Cum Laude**
- Bachelor of Science Honours (Computer Sciences) University of Pretoria **(2021) - Grade: B**
- Bachelor of Science (Computer Sciences) University of Limpopo **(2011 - 2015) - Grade: B**
- Nkwangulati Education Centre (High school) **(2006-2010) - Grade: B**

#### **ADDITIONAL SKILLS**

- **Databases:** SQL Server, Oracle, SAP Sybase, PostgreSQL, DB2, Hadoop Hive, **Vector Databases** (FAISS).
- **Data Integration & ETL:** StreamSets, SSIS.
- **Scripting & Programming:** Python, PySpark.
- **Reporting & Analytics:** SSAS, SSRS, Power BI.
- **Cloud Platforms:** Azure, Google Cloud Platform (GCP), Databricks.
- **AI & Machine Learning:** Scikit-learn, TensorFlow, Keras, Hugging Face Transformers, Large Language Models (LLMs), Prompt Engineering, Model Evaluation & Deployment.
- **Natural Language Processing (NLP):** Sentiment Analysis, Text Classification, Retrieval-Augmented Generation (RAG), Auto-Regressive Fine-Tuning (ARFT), ESG-specific NLP applications.
- **Proprietary Models & Research:** LFEAR (Language Feature Extraction and Adaptation for Reviews): A novel NLP model combining RAG and ARFT for enhanced sentiment analysis and contextual understanding.

#### **CERTIFICATES**

- Microsoft Certified: Azure Fundamentals

#### **TRAINING**

- Querying Microsoft SQL Server 2014
- Implementing a SQL Data Warehouse / Microsoft Azure
- Developing SQL Data Models
- Analysing Data with SQL Server Reporting Services

#### **PUBLICATIONS**

- *Mathebula, M., Modupe, A., Marivate, V. (2024). ChatGPT As a Text Annotation Tool to Evaluate Sentiment Analysis on South African Financial Institutions. [IEEE Access Journal](#)*
- *Mathebula, M., Modupe, A., Marivate, V. (2024). Fine-Tuning Retrieval-Augmented Generation with an Auto-regressive Language Model for Sentiment Analysis in Financial Reviews. [MDPI Journal](#)*

#### **RESEARCH SUPERVISION**

- *Bridging Explainability and Ethics in AI: Improving Trust and Reducing Misinformation*  
Honours Project (2025) – Student: Kumbirai Shonhiwa
- *SE-Spline Attention-Based Convolutional Kolmogorov-Arnold Network for Intrusion Detection*  
Honours Project (2025) – Student: Wadalisa Molokwe
- *Dynamic Context-Aware Multimodal Knowledge Graph for Real-Time Reasoning*  
Honours Project (2025) – Student: Zingisa Matwana

#### REFERENCES

- References available on request.
- Supporting Documents available on request