Khaleel Ahmad

📞 +49 15563611714 | ✉ khaleel.eu@gmail.com | 📍 Berlin, Germany

⊙ github.com/khaleel-git | in linkedin.com/in/khaleel-ahmad | 🔗 khaleel.eu

Platform Engineer with 3+ years of experience designing scalable data models, ETL/ELT pipelines, and data platform solutions. Expert in SQL, data modeling, Apache Airflow, and cross-functional collaboration. Proven track record implementing data quality standards, schema design, and automation that improved processing efficiency by 90%. Currently pursuing M.Sc. in Artificial Intelligence at BTU Cottbus-Senftenberg.

Education

BTU Cottbus-Senftenberg, Germany   Apr 2025 – Present

Master of Science in Artificial Intelligence
Relevant Coursework: Network and System Security, Distributed Computing in ML, Logic in Databases

Information Technology University, Pakistan   Sep 2016 – Feb 2021
Bachelor of Science in Computer Science

Skills

Languages: Python, Java, C, C++, Bash, Zsh, HCL (terraform)

Data Engineering & Databases: SQL, PostgreSQL, MySQL, MSSQL, MongoDB, SQLite, Data Modeling, Schema Design, pandas, NumPy, Apache Spark, Apache Airflow (ETL/ELT Pipelines), Data Quality Assurance

Data Formats & Processing: JSON, CSV, Parquet, XML, Data Validation, Data Enrichment

DevOPS: Linux (Debian, CentOS), Kubernetes, Docker, Git Versioning, Ansible, Terraform, CI/CD (GitHub Actions)

Frameworks & Monitoring: Flask, Django, Node.js, ReactJS, Prometheus, Grafana

Cloud: VMware, AWS, GCP, Azure, IBM, Digital Ocean, Alibaba

Experience

Ericsson   Aug 2022 – Feb 2025
Data (Platform) & AI Engineer
- Designed and maintained scalable ETL/ELT pipelines using Apache Airflow and Python, processing 30,000+ XML files daily with schema validation and version control.
- Improved distributed data processing performance by 90% through custom framework, reducing data pipeline execution time from 12 hours to 1 hour.
- Automated MSSQL workflows with Python, pandas, and pymssql, implementing data contracts and validation rules that reduced processing time by 30%.
- Built analytics-ready and ML-ready datasets through data validation, enrichment, and transformation workflows using pandas and SQL.
- Optimized heavy-load T-SQL queries and managed database schemas for PostgreSQL and MSSQL environments.
- Evaluated distributed processing frameworks including Spark and DuckDB concepts.
- Managed VMware/vCenter environments, RAID storage, and ILO server infrastructure.

Hello App   Jan 2020 – July 2022
Linux System Administrator & Python Developer
- Managed multi-cloud infrastructure for 50+ applications across AWS, GCP, Azure, DigitalOcean, Vultr & Linode, ensuring 99.9% SLA uptime and high availability.
- Deployed and administered 50+ websites using WHM/cPanel, Vesta, CloudPanel, and ISPConfig, streamlining web hosting operations.
- Deployed scalable WordPress and Python applications, enhancing cloud infrastructure efficiency and performance.

Projects

Tech Migration from Ericsson GSC China in   Mar 2023 – Jun 2024

Tech Stack: Python, Apache Airflow (Workflow Orchestration / CI/CD), AWS, MSSQL, Linux, Bash, On-prem VPS
- Migrated critical infrastructure and automated workflows from Ericsson GSC China to AWS and internal VPS infrastructure, ensuring zero-downtime system operations.
- Re-engineered ETL/ELT data pipelines using Apache Airflow, Python, and pandas with robust schema design and versioning, improving data reliability and reproducibility while reducing manual intervention by 40%.

Diversity Audit | Ericsson/Jazz in   Mar 2023 – Jun 2024

Tech Stack: Python, Custom Distributed Processing (VMware-based VPS Cluster), MSSQL, Pandas, NumPy
- Built resilient distributed data processing framework achieving 12x performance improvement (12 hours to 1 hour) by processing 30,000 XML files across 4 servers with data validation and enrichment workflows.
- Designed analytics-ready datasets with strong schema management, data contracts, and version control for diversity reporting.
- Implemented automated data quality checks and schema versioning using pandas and SQL, ensuring reproducible analytics workflows and maintaining data contracts across multiple downstream reporting systems.

Delta Scanner | Ericsson/Jazz in   Mar 2023 – Jun 2024

Tech Stack: Python, pymssql, Custom Distributed Processing, Multiprocessing, XQuery

- Developed real-time ETL pipeline using multiprocessing and distributed computing to validate and analyze 30,000 XML files for parameter changes with schema enforcement.
- Generated automated, reproducible delta reports for network infrastructure enabling proactive configuration management with data contracts ensuring consistency.

Certifications

Python Data Structures

Certified Kubernetes Administrator (CKA)

Containerized Applications on AWS

Mastering Ansible Automation

Languages

English (C1)

Deutsch (A2)