

# PARSA KAMALI

M.Sc. Data Science Student | Working Student Data Engineering | Cloud Data Engineering  
Bochum, Germany | [parsa.kamali.data@gmail.com](mailto:parsa.kamali.data@gmail.com) | [www.linkedin.com/in/parsa-kamalii](https://www.linkedin.com/in/parsa-kamalii) | [github.com/Parsaka](https://github.com/Parsaka)

**Portfolio & full project documentation:** [parsa-kamali.de](https://parsa-kamali.de)

## PROFILE

M.Sc. Data Science student at TU Dortmund University with practical experience in data analysis, data quality evaluation, Python, SQL, and academic data projects involving large-scale datasets and machine learning workflows. Focused on building a career in Data Engineering and Cloud Data Engineering, with hands-on learning in AWS, Databricks, Spark, and ETL/ELT pipeline design. Seeking a Working Student / Werkstudent position in Data Engineering, Cloud Data Engineering, or Data Analytics in Germany.

## TECHNICAL SKILLS

**Programming & Querying:** Python, SQL, Bash basics

**Data Processing & Analysis:** Pandas, NumPy, Polars, Jupyter Notebook, data cleaning, data preprocessing, exploratory data analysis, feature engineering, data visualization

**Data Science & Machine Learning:** Scikit-learn, XGBoost, PyTorch, BERT-based text processing, supervised learning, classification, model evaluation, precision/recall/F1, confusion matrix

**Data Engineering:** ETL/ELT concepts, data pipelines, batch processing, data quality evaluation, data transformation, data modeling basics, workflow orchestration basics

**Cloud & AWS:** Amazon S3, AWS Glue, Amazon Athena, Amazon Redshift, AWS Lambda, AWS Step Functions, IAM, CloudWatch, KMS

**Databricks & Big Data:** Databricks, Apache Spark, PySpark, Delta Lake, Lakehouse architecture, Lakeflow Jobs

**Graph Analytics:** PyTorch Geometric, Graph Neural Networks, GAT/GATv2, reply graph modeling, edge classification

**Tools & Development:** Git, GitHub, Docker basics, Linux basics, VS Code, HTML, CSS, Bootstrap, PHP basics

## EDUCATION

**M.Sc. Data Science | TU Dortmund University | Dortmund, Germany | Present**

- Relevant areas: data analysis, machine learning, statistics, data mining, simulation-based inference, computational intelligence, and graph-based modeling.

**B.Sc. Software Engineering | Golestan University | Iran**

## WORK EXPERIENCE

**Data Analyst | TELUS Digital AI Data Solutions | Essen, Germany | May 2022 - Present**

- Work on data analysis and quality evaluation tasks in an international AI data solutions environment.
- Review, analyze, and evaluate datasets to identify patterns, inconsistencies, and data quality issues.
- Follow structured workflows and quality standards to support reliable data-driven processes.
- Developed strong attention to detail, consistency, and reliability when working with data-related tasks.

**Software Engineer | Poshtibanan Pardazesh Aseman | Gorgan, Iran | 2017 - Jun 2020**

- Worked on backend development for database-driven web and mobile application projects, mainly using PHP, SQL, and Microsoft SQL Server.
- Designed, maintained, and queried relational database structures to support business workflows, user data, and application features.
- Developed backend logic connecting PHP-based application components with SQL Server databases and structured data operations.
- Worked with SQL queries, data retrieval, data updates, and backend application logic for production-oriented software use cases.
- Built a strong backend and database foundation through practical experience in debugging, problem solving, and project delivery.

## SELECTED PROJECTS

**Telegram Message Deletion Analysis Pipeline | Independent Data & Graph Analytics Project | Python, Polars, PyTorch Geometric, XGBoost**

- Built a Python-based data processing and graph analytics pipeline for Telegram message deletion analysis.
- Cleaned and processed large-scale message datasets containing user IDs, timestamps, reply relationships, topics, and deletion labels.
- Created structured features from temporal patterns, user behavior, reply chains, message metadata, and graph-based interactions.
- Trained and evaluated deletion prediction models using XGBoost and graph neural network approaches.
- Produced analysis-ready outputs, evaluation metrics, and visual summaries for model comparison and interpretation.

### **AWS & Databricks Lakehouse Pipeline | Independent Portfolio Project | AWS, Databricks, Spark, Delta Lake, ETL/ELT**

- Designed an end-to-end cloud data engineering pipeline based on AWS and Databricks Lakehouse concepts.
- Planned Bronze, Silver, and Gold data layers for raw ingestion, data cleaning, transformation, and analytics-ready outputs.
- Practiced AWS services including Amazon S3, AWS Glue, Amazon Athena, Amazon Redshift, IAM, AWS Step Functions, CloudWatch, and KMS.
- Applied Databricks concepts including Apache Spark, PySpark, Delta Lake, Lakehouse architecture, Lakeflow Jobs, and workflow orchestration.
- Focused on scalable data processing, data quality checks, partitioning concepts, reproducible transformations, and job-ready cloud data engineering workflows.

### **Eye-Tracking Data Analysis Case Study | Academic Data Analysis Project | Python, Data Cleaning, EDA, Visualization, Statistical Reporting**

- Analyzed eye-tracking fixation data to study participant-level visual attention patterns.
- Worked with fixation-based measurements, participant information, and experimental data structures.
- Performed data cleaning, exploratory data analysis, and structured interpretation of behavioral patterns.
- Prepared analytical outputs, visual summaries, and written reporting based on experimental results.

### **Personal Portfolio Website | Independent Web Portfolio Project | PHP, HTML, CSS, Bootstrap, GitHub**

- Developed a personal portfolio website to present technical projects, CV, skills, contact information, and career profile.
- Built structured sections for profile, experience, projects, skills, and contact.
- Used PHP, Bootstrap, HTML, CSS, GitHub, and basic local development tools for implementation and presentation.

**Portfolio note: The projects above are selected summaries. Full project documentation, screenshots, additional projects, and supporting materials are available on my portfolio website: [parsa-kamali.de](https://parsa-kamali.de)**

## **CERTIFICATIONS & TRAINING**

---

- AWS Certified Data Engineer - Associate (DEA-C01)
- Databricks Certified Data Engineer Associate - exam preparation.
- Databricks Academy training focused on Lakehouse architecture, Spark, Delta Lake, and data engineering workflows.

## **LANGUAGES & ADDITIONAL**

---

**Languages:** Persian: Native | English: C1 | German: B1, currently improving toward B2

**Driving Licence:** Class B