

ANNE-SOPHIE MLOPS- STADELMAIER

SENIOR MLOPS ENGINEER

CONTACT

✉ annesophie.stadelmaier@example.de
☎ +49 30 4485 6620
🏠 Berlin, Deutschland
🌐 annesophie-stadelmaier.io
👤 linkedin.com/in/annesophie-stadelmaier
📁 github.com/asstadelmaier

PROFIL

MLOps Engineer mit 7 Jahren Plattform-Erfahrung bei Zalando SE und Continental AG. Schwerpunkt Kubeflow, Argo Workflows und Terraform mit messbarer Wirkung auf Modell-Deploy-Frequenz, MTTR und Trainings-Compute-Kosten. AWS ML Specialty + CKA-zertifiziert und Maintainer einer Open-Source-Pipeline-Bibliothek mit 1.080 GitHub-Stars.

AUSBILDUNG

M.Sc. Informatik
10/2014 - 09/2017
TU Berlin Berlin, Deutschland
Informatik (Verteilte Systeme + ML)
1,3

B.Sc. Informatik
10/2011 - 09/2014
TU Berlin Berlin, Deutschland
Informatik GPA: 1,5

BERUFSERFAHRUNG

Senior MLOps Engineer 07/2021 - heute
Zalando SE Berlin, Deutschland

Senior MLOps-Engineer im Data-Science-Plattform-Team

- Aufbau einer Kubeflow-Pipeline-Plattform fuer 24 ML-Teams und 380 produktive Modelle in 14 Produkt-Bereichen
- Migration von Modell-Serving auf Triton Inference Server, p95-Latenz von 240 ms auf 38 ms reduziert
- Senkung der Trainings-Compute-Kosten um 38 Prozent (480.000 EUR p.a.) durch Spot-Instances und Auto-Scaling
- Aufbau einer Drift-Erkennungs-Pipeline mit Evidently und Prometheus, 14 stille Modell-Degradierungen frueh erkannt

MLOps Engineer 10/2017 - 06/2021
Continental AG ADAS+AD Hannover, Deutschland

MLOps-Engineer im ADAS-Trainings-Plattform-Team

- Aufbau einer Multi-Cluster-Kubernetes-Plattform fuer 48 ADAS-ML-Engineers ueber 4 Standorte
- Integration von DVC und MLflow in den Standard-Trainings-Workflow, 1.840 Experimente pro Quartal nachverfolgbar
- Konfiguration von Argo Workflows fuer 240 nightly Trainings-Jobs auf einem 480-GPU-Cluster
- Terraform-Module fuer reproduzierbare Cluster-Konfigurationen ueber AWS, Azure und On-Prem geliefert

FÄHIGKEITEN

- MLflow & Kubeflow
- Pipelines
- Argo Workflows & Airflow
- Kubernetes, Helm &
- ArgoCD
- Terraform & Pulumi
- Python (FastAPI, gRPC,
- OpenTelemetry)
- Triton Inference Server &
- BentoML
- AWS SageMaker & Vertex
- AI
- Prometheus, Grafana &
- Loki

ZERTIFIKATE

AWS Certified Machine Learning Specialty
02/2025

Certified Kubernetes Administrator (CKA)
09/2024

HashiCorp Certified Terraform Associate
06/2023

Google Cloud Professional Machine Learning Engineer
03/2022

SPRACHEN

Deutsch Muttersprache
Englisch C1

STÄRKEN

Plattform-Reife
Sehe ML-Plattform als Produkt fuer interne Engineers mit Service-Level-Objectives und Roadmap

PROJEKTE

kubeflow-helm-charts (Open Source) 06/2024 - heute

Production-ready Kubeflow Helm-Charts fuer DACH-Konzern-Kontexte, 1.080 GitHub-Stars und 14 Konzern-Adoptionen

Multi-Tenant-Modell-Serving bei Zalando 01/2024 - 05/2024

Aufbau einer Triton-Inferenz-Plattform fuer 24 ML-Teams, p95-Latenz 38 ms, 480.000 EUR p.a. Kosten gesenkt

Observability-Fokus

Bestehe auf End-to-End-Tracing mit OpenTelemetry, Daten- und Modell-Drift-Alarmen vor Rollouts

Pragmatismus

Bevorzuge stabile, gut dokumentierte Standards vor neuesten Frameworks und liefere reproduzierbare Deployments