

# Data Engineer

Python or Java



## PROFILE

---

We are looking for a hands-on eager to take the challenge and unleash the power of technology and data. We constitute the core of Data Engineering Team which is responsible for development of next generation cloud-native data hub with built-in Artificial Intelligence Capabilities. We collect data, consolidate, process, visualize and apply machine learning models to generate business insights and value.

- A passionate data engineer who loves software development and distributed data processing and is **familiar with data processing and analysis in Python and AWS**
- Familiar with principles and common design patterns regarding distributed systems based on microservices and/ or serverless architecture running on AWS Lambda, AWS ECS or AWS EKS
- Familiar with AWS data processing stack as AWS Kinesis, Athena, S3, Lambdas, Firehose
- Experienced in deploying production-grade software with fully automated CICD pipeline and configuring fully automated logging and monitoring of a distributed system
- **Fluency in English**
- Willing to travel to Munich and Zurich from time to time

## WORKING ENVIRONMENT

---

- Develop API and software for our entire data pipeline to process massive amounts of data in real-time in the public cloud (AWS)
- Implement our main use case: analyze data of online customers across brands and make user behavior and content insights comprehensible for product and content recommendations
- Be comfortable with deploying and running production-grade microservices
- Collaborate within a cross-functional international team to plan, design, implement and run data processing solutions, including running DNN models in production
- Develop technical designs for advancing the complex data processing and machine learning pipelines across multiple AWS products and services
- Ensure customer experience for platform API users is as intuitive as possible

## MUST HAVE SKILLS

---

- Experience in using **Java or Python**
- Working knowledge of DevOps and IaC tooling
- Openness to new technologies and languages
- Understanding the problems of distributed systems design
- Knowledge of distributed data storage systems/formats using parallel processes and/or columnar data stores such as AWS Redshift, Elastic Search
- SOLID, Clean Code



# Data Engineer

Python or Java



## NICE TO HAVE

---

- Experience with online and offline data processing tools and technologies (Storm, Kafka, Spark, Flink)
- Knowledge of Cloud Computing on multiple Amazon Web Services or any other public cloud offering
- Exposure to Continuous Integration/Continuous Deployment Environment
- CICD for serverless infrastructure using Github Actions
- Background in Data Collection + Ingestion, Data Visualization, ETL, Data Processing Pipeline and Data Modelling
- Knowledge of machine learning and data science
- Ability to deep dive into log files and application monitoring alerts

## OUR OFFER

---

- Professional development in the fields of Big Data, Data Science and Machine Learning
- Learn on the job from our experienced team of international software engineers
- Work on bleeding edge public cloud architecture and data engineering framework
- Bring in your own ideas and be part of the decision making

