

Markus Rothe

Dr.-Ing.

Meerbusch, Germany

+49 179 65 91 875

markus.rothe@rite.cc

Data Scientist and Software Engineer. Speaks German and English. Born in 1986.

Projects

Jan 2023 – Sep 2023 <i>Consumer Goods</i>	Performance optimizations and maintenance of a medium sized Hadoop cluster and its applications, both batch and streaming. <i>Hadoop, Spark, HPE Ezmeral Data Fabric (MapR), Ansible, Python, Scala, Java</i>
Jun 2023 – Sep 2023 <i>Consumer Goods (ibid.)</i>	Maintenance of containerized applications and deployment to an on-premises Kubernetes cluster. <i>Kubernetes, Python, Scala, Java</i>
Apr 2023 – Mai 2023 <i>Consumer Goods (ibid.)</i>	Support the migration of a data warehouse from Teradata to Snowflake by identifying and translating SQL statements between the language dialects. In addition, support the automation of the Snowflake configuration using Terraform. <i>Snowflake SQL, Teradata SQL, Terraform</i>
Sep 2022 – Dec 2022 <i>Automotive</i>	Set up of Kubernetes and Airflow on Azure. Development and deployment of ETL pipelines to track and statistically assess the remote update process of electronic control units (ECUs). <i>Azure Kubernetes Service (AKS), Python, Apache Airflow, Trino, OpenAPI/Swagger, Elasticsearch</i>
Mar 2021 – Dec 2022 <i>Public Sector</i>	Automated deployment of Hadoop clusters, based on Cloudera. <i>Cloudera (CDP Private Cloud), Ansible, Red Hat Identity Management / FreeIPA</i>
Dec 2021 – Aug 2022 <i>Public Sector (ibid.)</i>	Deployment and integration of identity and access management (IAM) based on LDAP, Kerberos, and SAML. <i>LDAP (389ds, OpenLDAP, LDIF), Kerberos, Keycloak (SSO, SAML), Ansible, Gitlab CI/CD</i>
Jul 2022 <i>Public Sector (ibid.)</i>	Data ingestion and cleansing of multiple state-run databases to create a unifying view on the data. <i>Spark, HBase, Hive, Impala, Kafka, Solr</i>
Nov 2021 – Mar 2022 <i>Construction</i>	Analyzing the content of email messages to predict the positive or negative outcome of projects. <i>Python, Hugging Face Transformers, Streamlit</i>

May 2021 – Sep 2021 <i>Metalworking</i>	Optimizing the dynamic operating point of a machine that shapes the form of blocks of steel. <i>SAP Data Intelligence, Python, xgboost</i>
Mar 2021 – May 2021 <i>Construction</i>	Data ingestion of email messages, followed by statistical evaluation of their meta data. <i>Python, Streamlit, pandas, scikit-learn</i>
Sep 2020 <i>Consumer Goods</i>	Model design for predictive maintenance of a manufacturing line that packages consumer products. <i>Databricks on Azure, Python, Spark, MLflow, pandas</i>
Aug 2020 <i>Consumer Goods (ibid.)</i>	Tuning parameters for a manufacturing process with black box optimization. <i>Black box optimization with botorch, Python</i>
May 2020 – Jul 2020 <i>Internet Services</i>	Development and deployment of an email proxy service. <i>C++, Python, Typescript, AWS: Simple Email Service (SES), Simple Notification Service (SNS), S3, DynamoDB, CodeBuild, CodeDeploy, Cloud Development Kit (CDK)</i>
Jan 2015 – Dec 2016 <i>Health Services</i>	Full stack web development of a web service for Tailor-Made Notched Music Training (TMNMT). <i>Go, SQL, JavaScript</i>
Mar 2015 – Sep 2015 <i>E-Commerce</i>	Development and deployment of an event registration website. <i>Go, SQL, JavaScript</i>

Certification

May 2021	Databricks Certified Professional Data Scientist
----------	--

Education

Nov 2011 – Mar 2019 <i>RWTH Aachen University</i>	Doctoral candidate and scientific staff member. Dissertation on Optimizing Large Scale Systems in Engineering.
Oct 2005 – Sep 2011 <i>RWTH Aachen University</i>	Diploma in Electrical Engineering and Information Technology. Diploma Thesis on Compressed Sensing for Multihop Relay Networks with Correlated Sources.