

# Heni Ben Arab

Senior Data Engineer

65933 Frankfurt am Main  
Deutschland

📞 \*\*\*\*\*

✉ [ben.arab.heni@gmail.com](mailto:ben.arab.heni@gmail.com)

🌐 [hbenarab.com](http://hbenarab.com)



## Professional Experience

- since Oct. 2021 **Senior Data Engineer**, *DB Systel GmbH*, Frankfurt, Germany.  
Design and development of Big-Data Architectures on AWS.  
**Tech-Stack:** Java, Scala, Python, Spark, Kafka, Docker, Airflow AWS
- Apr. 2017 **IT Consultant**, *ITinera projects & experts*, Eschborn, Germany.
- Sep. 2021 Consulting clients in banking, media and transportation industries.  
**Tech-Stack:** Java, SpringBoot, Spark, Kafka, Docker, Airflow, AWS
- Mar. 2015 **Working Student**, *University of Passau*, Passau, Germany.
- Dec. 2015 Research and implementation of Neural Networks models for NLP usecases.  
**Tech-Stack:** Python, Machine Learning/Deep Learning, NLP, Knowledge Graph
- Jun. 2014 **Trainee**, *Prologic Group*, Tunis, Tunisia.
- Aug. 2014 Implementation and Configuration of the Network&Security Module *vShield* on a VMware Cloud environment.  
**Tech-Stack:** Cloud Computing, Virtualization, VMware Cloud Suite (ESX, vSphere, vShield, vCloud Director)
- Jul. 2013 **Summer Trainee**, *Vermeg*, Tunis, Tunisia.
- Aug. 2013 Implementation and Configuration of CI/CD-Pipelines.  
**Tech-Stack:** Java, Maven, Hudson

## Education

- 2014–2017 **M.Sc. in Computer Science**, *University of Passau*, Passau, Germany.  
Specialization: Information and Communication Systems.  
Master thesis: Using Recurrent Neural Network for Knowledge Graph Extraction.
- 2012–2015 **Telecommunication Engineering Degree**, *Higher School of Communication of Tunis (Sup'Com)*, Tunis, Tunisia.  
Double degree agreement with the University of Passau.
- 2010–2012 **Classes Préparatoires**, *Tunis Preparatory Engineering Institute*, Tunis, Tunisia.  
Specialization: Mathematics-Physics  
Success in the competitive national examination for access to engineering studies. (ranked 89th out of 4000 students)
- 2010 **Baccalaureate Degree**, *L'Aouina High School*, Tunis, Tunisia.  
Mathematics, with distinction (16.28/20)

## Technical skills

<b>Programming</b>	Java, Scala, Python
<b>Java Frameworks &amp; Standards</b>	Spring Framework (Spring Boot, Spring-Data, Spring-Security, Spring-Batch), JPA, Hibernate, JUnit, Maven
<b>Database</b>	MySQL, PostgreSQL, MongoDB, InfluxDB
<b>Big Data</b>	Hadoop-Ecosystem (HDFS, YARN, Spark, Hive), Kafka, Nifi, Hudi
<b>Cloud</b>	<b>Amazon Web Services:</b> EC2, S3, RDS, EMR, ECS, Lambda, CloudWatch, Glue, CloudFormation, Comprehend. <b>Google Cloud Platform:</b> ComputeEngine, CloudStorage, CloudSQL, BigQuery, PubSub, CloudDataproc
<b>Other Frameworks &amp; Tools</b>	Git, Jenkins, Gitlab CI/CD, Docker, Airflow, Grafana, Camel

## Certifications

Jan. 2023	AWS Certified Solutions Architect - Associate
Aug. 2020	AWS Certified Cloud Practitioner
Jan. 2020	Google Cloud Certified Professional Data Engineer
Aug. 2018	Certified Scrum Developer
Sep. 2017	Certified Scrum Master
Mar. 2016	Deutsche Sprachprüfung für den Hochschulzugang (DSH 2: 78%)
Jan. 2014	Test of English for International Communication (TOEIC: 925/990)

## Languages

German	Fluent
English	Fluent
French	Bilingual
Arabic	Native

## Activities and interests

since Apr. 2023	Member of TS-Griesheim Club (Tennis & Table tennis)
since Feb. 2013	Co-founder of "Rainbow Events" (Organization of trips and excursions)
2013-2014	Chair of IEEE Sup'Com Student Branch
Hobbies	Reading, DJing, Tennis

## Projects

- since Nov. 2021 **MOVAS Data-Platform**, *Extension and further development of Vendo Datalake.*
- ◇ *Industry: Transportation*
  - ◇ *Role: Software-Engineer / Data-Engineer / DevOps-Engineer*
  - ◇ *Tasks:*
    - *Design and implementation of a containerized data management tool with Java and SpringBoot.*
    - *Design and implementation of Lakehouse usecases with Hudi.*
    - *Design and implementation of real-time streaming usecases with Spark Streaming.*
    - *Connecting external systems to Datalake with Apache Camel.*
    - *Interim takeover of Product Owner tasks.*
    - *Migration of CI/CD-Pipelines from Jenkins to GitlabCI/CD.*
  - ◇ *Tech-Stack: Java, Scala, Python, Spark, Hudi, Kafka, Camel, Docker, AWS, Airflow, GitlabCI/CD.*
- Jun. 2021 **Data driven Ads Platform**, *Extension of a data driven Ads Platform on GCP.*
- Sep. 2021
- ◇ *Industry: Media*
  - ◇ *Role: Data-Engineer*
  - ◇ *Tasks:*
    - *Extension of existing datalake architecture on GCP.*
    - *Implementation of Data Quality Control workflows with Great Expectations.*
    - *Orchestration and Scheduling of Workflows with Airflow.*
  - ◇ *Tech-Stack: GCP, Python, Airflow, Great Expectations.*
- Jul. 2019 **Vendo-Datalake**, *Design of a Datalake for the sales platform "Vendo".*
- Mai. 2021
- ◇ *Industry: Transportation*
  - ◇ *Role: Software-Engineer / Data-Engineer*
  - ◇ *Tasks:*
    - *Design of the Datalake architecture.*
      - *Stage 1: Based on Hortonworks Suites HDF and HDP.*
      - *Stage 2: Migration to AWS.*
    - *Development of ETL-Applications with Apache Spark.*
    - *Orchestration and Scheduling of ETL-Workflows with Airflow.*
    - *Development of Data Pipelines with NiFi.*
    - *Design and implementation of a KafkaStreams-based Application to anonymize personal data.*
    - *Monitoring of workflows, services and infrastructure with Grafana.*
  - ◇ *Tech-Stack: Java, Scala, Spark, Kafka, Kafka-Streams, NiFi, Hadoop/HDFS, YARN, Hive, Docker, AWS (EC2, EMR, S3, Lambda), Python, Airflow, Grafana, Jenkins, MongoDB, InfluxDB.*
- Mai 2018 **ReACT**, *Building a BigData Platform to prevent against money laundering and terrorism financing.*
- Jul 2019
- ◇ *Industry: Banking*
  - ◇ *Role: Software-Engineer / Data-Engineer*
  - ◇ *Tasks:*
    - *Design and development of an ETL-Framework with Java und Apache Spark.*
    - *Development of a DSL to help with data modeling and automatic generation of test scripts.*
  - ◇ *Tech-Stack: Java, Spark, Hive, Xtext.*

Sep. 2017 **Digital Strategy - Cross Channel Banking**, *Implementation of new Backend-*  
Mai. 2018 *Features for the Online Banking System.*

◇ *Industry: Banking*

◇ *Role: Java-Developer*

◇ *Tasks:*

- *Implementation of a SpringBatch-based batch architecture for importing foreign bank accounts and transactions.*

- *Design of data model on Oracle.*

◇ *Tech-Stack: Java, Spring, Spring-Batch, Spring-Data, Hibernate, OracleDB, Jenkins.*

Apr. 2017 **Timecard**, *Development of an internal management tool.*

Aug. 2017 ◇ *Industry: -*

◇ *Role: Java-Developer*

◇ *Tasks:*

- *Development of a WebApplication with Spring and PrimeFaces.*

- *Design of data model on MySQL.*

◇ *Tech-Stack: Java, Spring, Spring-Data, Spring-Security, PrimeFaces, Hibernate, MySQL, Jenkins.*