https://github.com/jawadtahir
https://linkedin.com/in/jawadtahir

Dr. Jawad Tahir

Researcher/Architect/Big Data/Machine Learning

ranajawadtahir@gmail.com
+49 1525 2959791

Education

Ph.D. (Computer Science)
Technical University of Munich
09/2020 – 10/2025

M.S. (Computer Science)
Technical University of Munich
10/2017 – 02/2020

M.S. (Computer Science)
Lahore University of Management Sciences
09/2013 – 06/2016

B.Sc. (Computer Engineering)
University of Engineering and Technology, Lahore
09/2009 – 08/2013

Certifications

Oracle Certified Associate: Java SE 8 Programmer

Microsoft Specialist: Programming in HTML5 with JavaScript & CSS3

Skills

Languages: Java, Python, Kotlin, Scala, JavaScript
Databases: MongoDB, PostgreSQL, MySQL, IBM DB2
Distributed processing tools: Apache Kafka, Kafka Streams, Apache Storm, Apache Spark, Apache Flink, Apache Hadoop
***Containerization*:** Docker, Podman, Docker Swarm, Kubernetes
Machine Learning: TensorFlow, Hugging Face, LLMs, LangChain
Vector Databases: ClickHouse, LanceDB, Faiss
Build tools: Maven, Gradle, SBT, pip, make, CI/CD pipelines
***Web frameworks*:** Spring, Django, Vert.x, Scrappy
Frontend tools: jQuery, Node.js
Infrastructure tools: Terraform, Ansible
Monitoring: Prometheus, Grafana, WandB
Cloud providers: AWS, Azure, GCP, OpenStack, OpenNebula

Professional Experience

Technical University of Munich | Munich, Germany
09/2020 – 09/2024 | Scientific Researcher

- Designed and developed an open-source tool, PGVal, to benchmark fault tolerance of Kafka Streams, Apache Storm, and Apache Flink for reliability and performance. Infrastructure faults were injected to learn the fault-tolerance behaviors of these systems.
  - Research findings were published in VLDB 25.
- Taught *Cloud Databases* lab course for 5 semesters.
- Went on a research exchange to the *University of Toronto*, where I worked as a Teaching Assistant for the Distributed Systems course.
  - Developed an auto-grader, pTA, to reduce the workload of instructors. It further resulted in increased interest for the students.

IBM T.J. Watson Research Centre | NY, USA
03/2024 – 06/2024 | Visiting Researcher

- Benchmarked various schedulers and dispatchers of Kubernetes to optimize distributed ML workloads by gang-scheduling. Studied the effects of varying workload characteristics, such as the number of pods, resource requirements, and compute resources. Repo

Intel | Munich, Germany
01/2018 – 12/2019 | Software Engineer

- Designed and developed *Task List Processor: Transpiler and Patcher*. The transpiler enabled developers to write firmware for 5G modems in Python, rather than using hexadecimal codes. The patcher employed a clever use of jump instructions to enable over-the-air firmware updates.
  - Was awarded *Intel Invention Discovery* for this work.

NETSOL Technologies | Lahore, Pakistan
06/2015 – 03/2017 | Software Engineer
06/2013 – 08/2014 | Software Engineer

- Worked on the development of an enterprise insurance management system using Java, J2EE, Hibernate, and Spring framework. Provided round-the-clock support to clients. Worked in the SCRUM framework to optimize deliverables.

Voluntary Scientific Work

- Chaired the Grand Challenge track of the *ACM DEBS conference* in 2021, 2022, 2024, and 2025.
  - Developed an open-source benchmarking tool for the submitted solutions using a virtualized and containerized evaluation infrastructure on a multi-tenant Kubernetes cluster. The tool used REST to transfer data. Paper

Publications

- Published 8+ papers at top-tier publishing venues, such as *VLDB, Middleware, DataEd,* and *DEBS*. Google Scholar profile