Technische Universität Berlin

Technische Universität Berlin offers an open position:

Research Assistant - salary grade E13 TV-L Berliner Hochschulen

Part time employment may be possible

The Berlin Institute for the Foundations of Learning and Data (BIFOLD) is one of six national AI centres in Germany and is funded by the State of Berlin and the Federal Ministry of Education and Research. BIFOLD currently consists of 12 research groups with over 150 employees, a graduate school and the BIFOLD office. Fellows from the major Berlin universities, Charité - Universitätsmedizin Berlin and various other national and international universities and non-university research institutions are also involved.

Faculty IV - The Berlin Institute for the Foundations of Learning and Data (BIFOLD)
Reference number: IV-177/25 (starting at the earliest possible / for 4 years / closing date for applications 20/06/25)

Working field:

The DEEM Lab (https://deem.berlin) is looking for a research associate in responsible data engineering. The research will be conducted in close collaboration with Prof. Julia Stoyanovich from New York University (https://airesponsibly.net/people/julia/).

Responsible data engineering is emerging as a new discipline at the intersection of data engineering and AI that treats ethics, legal compliance, and inclusivity as central design considerations. The holistic nature of this approach is based on the observation that the decisions we make during data collection and preparation profoundly impact AI systems we build and deploy.

The goal of this position is to create a new system which helps data engineers to design data preparation pipelines that optimize model performance along a rich set of responsibility objectives, including accuracy, robustness, fairness, and legal compliance. For that, the system will proactively guide data engineers through the selection and evaluation of a large set of data preprocessing, data augmentation and feature selection operations. A reliable, efficient and easy-to-use open source implementation of this system will be created as part of the research project.

This endeavor is technically challenging in multiple ways. First, data preparation and model selection need to be optimized for multiple objectives, in contrast to existing approaches, which focus on a single objective only such as overall prediction accuracy. Second, the system will have to create, rewrite and concurrently execute large numbers of different pipeline variants, which requires an efficient runtime and novel query optimization techniques. Third, the research needs to account for current dramatic changes in the development practices of AI applications, e.g., AI assisted programming, tabular foundation models and AI-based data science agents.

Requirements:

- Successfully completed university degree (Master, Diplom or equivalent) in Computer Science or Artificial Intelligence
- Very good programming skills in Python and at least one additional language (Java/Rust/C++)
- Knowledge in data processing with dataflow systems, relational databases and/or dataframe libraries (e.g., Apache Spark, DuckDB, pandas, etc.)
- Experience with increasing the efficiency, scalability and correctness of data-centric programs
- Basic knowledge of machine learning and common libraries (e.g., pandas, sklearn, pytorch, SparkML, etc.)
- Good knowledge of German and/or English required; willingness to acquire the respective missing language skills

Desirable:

- First-hand experience with real world data processing systems and/or ML deployments (e.g., from internships, jobs or entrepreneurial experience)
- Contributions to open source projects

Please send your application with the usual documents (at least a CV, a motivation letter and degree certificates) exclusively by e-mail to Prof. Dr. Sebastian Schelter at **schelter@tu-berlin.de**, quoting the **reference number**.

By submitting your application via email you consent to having your data electronically processed and saved. Please note that we do not provide a guaranty for the protection of your personal data when submitted as unprotected file. Please find our data protection notice acc. DSGVO (General Data Protection Regulation) at the TU staff department homepage: https://www.abt2-t.tu-berlin.de/menue/themen_a_z/datenschutzerklaerung/ .

To ensure equal opportunities between women and men, applications by women with the required qualifications are explicitly desired. Qualified individuals with disabilities will be favored. The TU Berlin values the diversity of its members and is committed to the goals of equal opportunities. Applications from people of all nationalities and with a migration background are very welcome.

The vacancy is also available on the internet at
https://www.personalabteilung.tu-berlin.de/menue/jobs/