Are you looking for a PHD in the high-impact fields of Data Management, Machine Learning Systems, and/or High-Performance Computing? Are you eager to learn, like to work independently, but also in a team? Would you like to make an important scientific contribution? We look forward to meeting you!

PhD in Data Management and ML Systems (38,5 h/w) in Graz

The data management for data science laboratory (DAMS Lab) – as a cross-organizational research group, uniting the data management group of Graz University of Technology and the research area data management of the co-located Know-Center – offers a fully funded PhD position for 4 years in the context of the EU project "DAPHNE".

DAPHNE Project: With an excellent consortium of 13 industrial and academic partners, the DAPHNE project (Integrated Data Analysis Pipelines for Large-Scale Data Management, HPC, and Machine Learning) aims at defining and building an open and extensible system infrastructure for integrated data analysis pipelines, including data management and processing, HPC, and machine learning (ML) training and scoring. Key observations are that (1) systems of these areas share many compilation and runtime techniques, (2) there is a trend towards complex data analysis pipelines that combine these systems, and (3) the used HW infrastructure converges as well. Yet, the programming paradigms, cluster resource management, as well as data formats and representations differ substantially. Therefore, this project aims – with a joint consortium of experts from the data management, ML systems, and HPC communities – at systematically investigating the necessary system infrastructure, language abstractions, compilation and runtime techniques, as well as systems and tools necessary to increase the productivity when building such data analysis pipelines, and eliminating unnecessary performance bottlenecks.

Open Position: In the context of the DAPHNE project, this PhD position at Know-Center is supported by several senior researchers from the Know-Center and other partners, and focuses primarily on a

Multi-Level, Optimizing Compiler Infrastructure for Integrated Data Analysis Pipelines,

which contributes to the overall system architecture, language abstractions for linear algebra programs, neural networks, HPC primitives, and relational algebra, and especially optimizing compiler and runtime techniques. This infrastructure will orchestrate and interface with specialized techniques – developed in the other work packages – for scheduling and resource sharing, computational storage, and heterogeneous hardware accelerators like GPUs and FPGAs. The overall system infrastructure will then be used in several industrial use cases, as well as related benchmarking efforts. A research assistant on this position is expected to pursue a PhD study at the Institute of Interactive Systems and Data Science at **Graz University of Technology**, under the supervision of Matthias Boehm (https://mboehm7.github.io/; for further questions, please send an email to mboehm@know-center.at).

Qualifications:
- University degree (MSc or equivalent) in computer science, software engineering, physics, or equivalent
- Programming experience in C/C++, and system-level programming
- Basic background in data management, machine learning, distributed systems, and compilers
- Motivation, team work, independent problem solving, and strong communication skills
- Very good English (written and spoken), German is advantageous but not a requirement

We offer:
- Close supervision of your PhD project, with international collaborators and networking opportunities
- High-impact, system-oriented research on data management, ML systems, and HPC
- Opportunities for professional and personal development
- Very good work-life balance and good working atmosphere in a research environment

The gross salary for this full-time-time job (38.5 h/week) is **2,750 EUR** per month (14 times per year).
We look forward to receive your application at career@know-center.at by the deadline **August 31, 2020**.