



INFORMATICS COLLOQUIUM

Speaker:

Prof. Felix Naumann, Chair for Information Systems, Hasso Plattner Institute, University of Potsdam, Germany

Data Profiling - A look back and a look forward

Abstract:

Data profiling is the act of extracting many different types of metadata from a given dataset. This research area has recently thrived, due to (i) its simple problem statements, such as "discover all key candidates", (ii) the high computational complexity of the problems, which are often exponential in the number of columns, (iii) the manifold opportunities for optimizations, such as apriori-inspired pruning or data sampling, and (iv) the various application areas for data profiling results, such as query optimization and data cleaning. In the talk, we will first cover the traditional problem statements and use cases and then highlight algorithmic intuitions of some solutions. Finally, we discuss open problems and promising research directions, for both data profiling and its use cases.

Bio:

Prof. Felix Naumann studied mathematics, economy, and computer sciences at the University of Technology in Berlin. After receiving his diploma (MA) in 1997 he completed his PhD thesis in the area of data quality at Humboldt University of Berlin in 2000. In 2001 and 2002 he worked at the IBM Almaden Research Center on data integration topics. From 2003 - 2006 he was assistant professor for information integration, again at the Humboldt-University of Berlin. Since 2006 he holds the chair for information systems at the Hasso Plattner Institute (HPI) at the University of Potsdam in Germany. He has been visiting researcher at QCRI, AT&T Research, IBM Research, and SAP. His research interests include data profiling, data cleansing, and data integration with over 200 scientific publications. Next to numerous PC memberships for international conferences, he has organized several conferences in various roles, including VLDB 2021 as PC co-chair, and he is trustee of the VLDB Endowment. More details are at <https://hpi.de/naumann/people/felix-naumann.html>.

Date and time: Friday July, 14th, 2023, 10.00 am
Location: Pérolles 21, room B130, Bd de Pérolles 90, Fribourg
Contact person: Dr. Alberto Lerner

The colloquium is free and open to the public.