



# INFORMATICS COLLOQUIUM

**Speaker:**

**Dr. Fabio Porto, National Laboratory of Scientific Computing (LNCC), Brazil**

## Towards Data-Centered AI

**Abstract:**

The construction of machine-learning models is a complex task. One crucial element concerns the data used for training, which should be correct and comprehensively reflect the learning objective. In this talk, we will argue in favor of data-centered AI, which refers to focusing on improving the quality of the data as a major goal to achieve better quality models. We will discuss two main areas we are interested in: data cleaning and subdomain characterization. We will present data-centered algorithms for these two areas and discuss the challenges associated with their evaluation on large training datasets.

**Bio:**

Fabio Porto is a Senior Researcher at the National Laboratory of Scientific Computing (LNCC), in Brazil. He is the founder of the DEXL Laboratory, developing R&D activities in the context of scientific data analysis and management. He holds a PhD in Informatics from PUC-Rio, with sandwich at INRIA, in 2001, and a postdoc at Ecole Polytechnique Fédérale de Lausanne (EPFL). He has more than 80 research papers published in International Conferences and Scientific Journals, including VLDB, SIGMOD, SSDBM and ICDE. He was the General Chair of VLDB 2018 and SBBD 2015. Since 2018 he has been a member of the SBBD steering committee, and a member of SBC and ACM.

*Date and time:* Thursday, November 02<sup>nd</sup>, 2023, 10.00 am.  
*Location:* Pérolles II, room E230, Bd de Pérolles 90, Fribourg  
*Contact person:* Prof. Philippe Cudré-Mauroux

*The colloquium is free and open to the public.*