

Data Science in Large-Scale Data Infrastructure Management

Dr. Daniel Quoc Dung Cao

Research Data Scientist

Meta Platforms, Inc.

Abstract: From unprecedented growth of online users to a vast range of new products such as virtual reality or Instagram Reels, Meta's demand for storage, computing, I/O, and networking capacity started exceeding what its data infrastructure can supply. Therefore, Meta has to research into more efficient and reliable data infrastructure management solutions. This seminar will bring an overview of some challenges that Meta Infrastructure Foundation is facing and the initiatives to address them. One case study about data dependency analytics will be discussed in detail.

Bio: Dr. Daniel Quoc Dung Cao is a Research Data Scientist at Meta Platforms, Inc. He works on improving the efficiency and reliability of the data warehouses and machine learning training resources. Previously, he obtained a PhD in Industrial & Systems Engineering with Data Science & Machine Learning specialization at the University of Washington, Seattle.