Contribution ID: **32**   Type: **not specified**

Hadoop Workshop

Usage of Apache Hadoop in large scale Data Analysis Projects are on the way to become mainstream. But what are the required skills and how do I start with an Apache Hadoop project? The workshop shows and compares several aspects which should be considered in the beginning of large projects. How do I start with a POC and how works this: "scale out"? What data is stored how and how do I access data in my Hadoop cluster? What programming skills are required and what are the processing paradigms I should know in the beginning? Such questions are discussed and possible solutions are presented during this interactive hands on session. The example use case is a data driven market study, which combines social media, time series data, and network analysis in one project.

Participants will receive a download link for the latest Workshop-VM and a preparation survey two weeks before the workshop.

Presenter: KÄMPF (CLOUDERA), Mirko

Track Classification: Big Data and Storage Systems