



# New data, new questions, new problems? Online behavioral data in social science research

Ruben L. Bach

(University of Mannheim)

14.06.2023, 16.00 (c.t.)

Online via Zoom (Link)

(Meeting-ID: 913-2473-4411; Password: StatsCol22)

Records of individuals' online activities obtained from devices like personal computers and smartphones have received a lot of interest in the social sciences in recent years. Many have praised such data for allowing fine-grained observations of individuals' online activities which would be impossible with more traditional data sources such as surveys. Recent work, however, warns that many data quality aspects of these novel data are so far poorly understood. Yet, as the number of observations in such data can quickly reach several millions, researchers sometimes seem tempted to treat such data as gold standard and ignore what their data may be missing, and which other systematic biases may be present. In this talk, I present both applied and methodological work using online behavioral data in a typical social science setting. First, using within-between random effects models, I show how online behavioral data combined with panel survey data allows us to understand the effects of news media consumption from populist alternative news platforms on individuals' political attitudes. Second, I show that online behavioral data, although containing detailed records of individuals' social media use, are far from being complete. Using hidden Markov models, combined online behavioral data, survey records, and donated social media data, I show that the online behavioral data seem to completely fail in capturing social media use for about one third of the sample.