Session: Reproducibility in an era of confidential data

Julia Lane (New York University, University of Strasbourg, University of Melbourne)

Stefan Bender (Bundesbank)

How Amazon and Tripadvisor can help central banks

This paper describes how the application of metadata documentation techniques from Amazon and Trip Advisor have improved search, discovery and reproducibility of confidential microdata.

The inability to search and discover data and code is a major challenge for reproducible research. Similar challenges have been addressed in the private sector, particular by companies like Amazon and Tripadvisor. This paper describes the development of new approaches to metadata documentation and curation which draw on the techniques used by such companies and apply them to financial microdata held by central banks.

The case is particularly relevant since banks collect very detailed microdata about – for example - banks, non-financial institutes and/or securities, which are needed to analyze causal relationships. Although there is often not a great deal of metadata documentation, researcher access is critical to enable high quality analysis. We describe the development of new software platforms, inspired by Amazon and Tripadvisor, that engages all users in the systematic documentation of all relevant steps in a data generating process.

keywords: reproducibility, curation, confidential data, financial data, big data, data access, software solution.