Open, metadata enriched, non-proprietary dataformat for data dissemination

Claudia Saalbach*[1] and Xiaoyao Han*[1]

[1]DIW Berlin – Germany

Abstract

At the moment, social scientists are using different and sometimes proprietary software which processes metadata in different ways to analyze their data. Different data formats of statistical software packages, which are only partially compatible, represent an obstacle for replication studies. In particular, proprietary data formats endanger the requirement of interoperability anchored in the FAIR principles. Our project aimed to address this problem by introducing a metadata-enriched open data format that can be easily accessible, readable and interoperable in various statistical software. As a first milestone, we would like to present the conceptual model of an open data format, where DDI might come into play, and a minimal data example derived from survey data. Furthermore, we would like to discuss our preliminary work on importing the open-data format into (proprietary) statistical programs and converting data of a proprietary format into the open format.

*Speaker