

# **STATISTICS COLLOQUIUM**

THURSDAY, NOVEMBER 12, 2009

TALK: 4:00 PM — SCIENCE CENTER, ROOM 705

RECEPTION: 5:15 PM — SCIENCE CENTER, 7TH FLOOR

## **“Data Wars – The Agency Strikes Back”**

**Jörg Drechsler**  
**Institute for Employment Research (IAB)**  
**Germany**

### ABSTRACT

Disseminating microdata to the public that provide a high level of data utility while at the same time guaranteeing the confidentiality of the survey respondent is a difficult task. Generating multiply imputed synthetic datasets is an innovative statistical disclosure limitation technique with the potential of enabling the data disseminating agency to achieve this twofold goal. So far, the approach was successfully implemented only for a limited number of datasets in the U.S. In this talk I present the first successful implementation outside the U.S.: The generation of partially synthetic datasets for a German establishment survey, the IAB Establishment Panel. I will describe the evolution of the project from the imputation of the missing values in the original survey and the discussions which variables to synthesize to the final synthesis. I will also present some disclosure risk evaluations and provide first results on the data utility of the generated datasets.