Frank Windmeijer, Economics, Social Sciences and Law

Causal inference, Instrumental Variables Estimation

With Epidemiology, Mendelian Randomization, Genetic Markers

Not really using "Big Data" for analysis, but machine learning techniques

Lasso for identifying and removing invalid instruments (markers), Adaptive Lasso with initial consistent median estimator

Hard thresholding, again for identifying valid instruments (markers)

Currently applying these latter techniques to summary data, i.e. only (partial) estimates available from various large data sets, to combine efficiently and correctly (again valid/invalid genetic markers).