

Why Data Availability is Such a Hard Problem

Alan F. Karr

National Institute of Statistical Sciences, Research Triangle Park, NC, USA

karr@niss.org

Data availability is an issue about which it is very easy to be unconstructively dogmatic. In fact, the realities are extremely complex, with multiple stakeholders, each subject to multiple, and sometimes conflicting, incentives. The problem has scientific, behavioral and economic components, some of which are only incompletely recognized. “Solutions” that expect that people will behave in ways that are counter to their own self-interest do not advance the conversation, and often lead to a hardening of positions. In this talk, I will attempt to structure the data availability problem in a way that both accommodates its complexity and provides a basis for continuing progress.

Key Words: Data availability, data access, replicability, reproducibility