



Privacy, Confidentiality, and Big Data

Stephen E. Fienberg
Carnegie Mellon University, USA – fienberg@stat.cmu.edu

As the size and complexity of databases grow, so do concerns regarding the protection of data from inappropriate use by others. On the one hand we want big data to be shared and useful for a variety of purposes, and on the other hand we know that increased access to more and more data poses increased risk to individual privacy. Is it reasonable to have as our goal maximizing the benefits of big data while minimizing the risk to individual privacy? I review some recent advances in methodology for confidentiality and privacy protection and consider the extent to which they scale to be of use in the sharing of privacy-protected big data sets.

Keywords: complexity of databases; risk to privacy; privacy protection;