

## **New Tools for the Analysis of Massive Data**

Subhash Kuvelker,  
National Academy of Sciences

The presentation topic is “New Tools for the Analysis of Massive Data”, a study that is now being conducted by the Board on Mathematical Sciences and their Applications of the National Academy of Sciences, Washington, DC. The presenter will explain the scope of work of the study as briefly described below, the issues and challenges to be addressed, application to continuing CODATA efforts, and benefits of this study for the CODATA community.

The study will 1) assess the current state of data analysis for mining of massive sets and streams of data; 2) identify gaps in current practice and theory; and 3) propose a research agenda to fill those gaps. The study will consider analysis of all data, especially streaming data from national as well as international sources, and in all disciplines that produce massive amounts of data including, but not limited to, the physical sciences, economics & finance, and social networks.

**Keywords:** massive data, big data, data analytics, streaming data, scientific data, social networks, financial data, mathematical tools, machine learning, statistics