Big Data

Session Organizers: Alberto-Giovanni Busetto, University of California, Santa Barbara, and Takeaki Uno, National Institute of Informatics, Japan

Processes and exchanges around us generate data at an escalating volume, velocity, and variety. In engineering, science, business, and medicine, insights from big data already enable better decision making and greater operational efficiency. To meaningfully extract value from data, however, big data technology will need to infuse analytics into systems, sensors, and devices while protecting security and privacy. In social analytics and the Internet of Things a dynamic tension is already visible between what individual data producers are willing to share with others and what data consumers might want to get access to. There is an urgent need for the design of mechanisms to balance competing values and priorities. These challenges need to be taken into account from the conception of each infrastructure for big data. Database technologies, for instance, will need to take these factors into account while rapidly incorporating data from heterogeneous sources such as sensors and mobile devices. There is significant pressure to develop efficient, simple, and secure indexing and storage systems. At present machine learning plays a central role in transforming data into information, but it will also need to help us transform information into useful knowledge. New challenges are already emerging not only in storage and analytics but also in communication, visualization, and governance. We are at the beginning of a revolution that will touch every machine and every life on earth.

The selection of our speakers reflects the breadth of this rapidly expanding field. Starting from the analytics, Cynthia Rudin (Massachusetts Institute of Technology) will talk about big data analysis to assist in practical decision making. While Hiromi Arai (University of Tokyo) will connect data mining to privacy, Vitaly Feldman (IBM) will discuss how the concept of differential privacy can be useful for preserving validity in adaptive data analysis. Takahiro Hara (University of Osaka) will talk about the reliable incorporation of data into database technology for the Internet of Things.