

siam news

Volume 45/ Number 3
April 2012

Got Data: Now What?

Our ability to collect and store data of many different kinds has improved at an enormous pace over the last few decades. From personalized gene sequencing, to video-uploads, to the latest radar systems, real-time generation of terabytes of data is quickly becoming reality. This phenomenon arises independently in the biological, social, and engineering sciences and is common to commercial and academic enterprise. A large industry has evolved around the task of storing this data, creating large data warehouses in which the information is kept, so it is fair to say that the “warehousing problem” is being addressed quite effectively.

But the question remains, what is to be done with all this data? Specifically, how do we go from “the bits” to actual understanding that can inform decision-making processes and yield more fundamental scientific understanding? The process of going from data to description to decision has lagged far behind the development of methods for data collection or data storage. The problem is twofold: size and structure. While the sheer enormity of the data complicates processing, the complex internal structure of the data can confound conventional methodologies for analysis.

The analysis of large data sets to provide

understanding, and ultimately knowledge, is one of the fundamental intellectual challenges of our time. It falls to practitioners of the mathematical sciences (mathematics, statistics, and computer science) to devise new methods for carrying out analysis tasks, as well as to construct new models or paradigms for thinking about data. Recent history is witness to the successful application of increasingly sophisticated methodologies—mathematical, statistical, and computational—in the study of high-dimensional data sets. For this reason, we are in a period that is particularly rich in intellectual opportunities for mathematical scientists to develop novel methods based on their domain expertise, and to see these developments translate into value for society. One reason for the choice of “The Data Deluge” as the theme of Mathematics Awareness Month 2012—celebrated in April—is to

WHAT WOULD YOU DO WITH ALL THIS DATA?

Mathematics and statistics provide the tools to understand ever-increasing amounts of data. To learn more, visit the Mathematics Awareness Month website and enter for a chance to win an iTunes gift card at www.mathaware.org.

Mathematics, Statistics, and the Data Deluge
MATHEMATICS AWARENESS MONTH

Sponsored by the Joint Policy Board for Mathematics—American Mathematical Society, American Statistical Association, Mathematical Association of America, Society for Industrial and Applied Mathematics

make everyone with interests in the mathematical sciences aware of the opportunities for innovation.—Gunnar Carlsson and Robert Ghrist, 2012 Mathematics Awareness Month Committee.