

# Data Mining as a Key Enabler of Computational Social Science

Jaideep Srivastava

Department of Computer Science & Engineering  
University of Minnesota, Twin Cities, USA  
srivasta@cs.umn.edu

**Abstract.** Observation and analysis of a phenomenon at unprecedented levels of granularity not only furthers our understanding of it, but also transforms the way it is studied. For instance, invention of gene-sequencing and computational analysis transformed the life sciences, creating fields of inquiry such as genomics, proteomics, etc.; and the Hubble space telescope has furthered the ability of humanity to look much farther beyond what we could otherwise. With the mass adoption of the Internet in our daily lives, and the ability to capture high resolution data on its use, we are at the threshold of a fundamental shift not only in our understanding of the social and behavioral sciences (i.e. psychology, sociology, and economics), but also the ways in which we study them. Massively Multi-player Online Games (MMOGs) and Virtual Worlds (VWs) have become increasingly popular and have communities comprising tens of millions. They serve as unprecedented tools to theorize and empirically model the social and behavioral dynamics of individuals, groups, and networks within large communities. The preceding observation has led to a number of multi-disciplinary projects, involving teams of behavioral scientists and computational scientists, working together to develop novel methods and tools to explore the current limits of behavioral sciences.

This talk consists of four parts. First, we describe findings from the Virtual World Exploratorium; a multi-institutional, multi-disciplinary project which uses data from commercial MMOGs and VWs to study many fields of social science, including sociology, social psychology, organization theory, group dynamics, macro-economics, etc. Results from investigations into various behavioral sciences will be presented. Second, we provide a survey of new approaches for behavioral informatics that are being developed by multi-disciplinary teams, and their successes. We will also introduce novel tools and techniques that are being used and/or developed as part of this research. Third, we will discuss some novel applications that are not yet there, but are just around the corner, and their associated research issues. Finally, we present commercial applications of Game Analytics research, based on our experiences with a startup company that we've created.