# Data Prospecting - the low hanging fruit in “big data”

Rahul Ramachandran, John Rushing, Amy Lin, Sara Graves
Information Technology and Systems Center
University of Alabama in Huntsville
Huntsville, Alabama
rramachandran@itsc.uah.edu

A new approach for exploiting "big data" is now possible with the availability of high performance computing and the advent of new techniques for efficient distributed file access. This new approach coined as “data prospecting” combines methods from both data analytics and mining to provide interactive online data exploration capabilities. Just as prospecting focuses on locating the site within the vast land and determining the type of deposit that is located at that site. Data prospecting focuses on providing interactive “first look” analytics for finding the right subset of data amongst all the data files and determining the value of the information contained within the subset. An initial prototype was developed to explore the viability of an interactively exploring large Earth science data. The Special Sensor Microwave Imager (SSM/I) gridded products available publicly from Remote Sensing Systems (http://www.ssmi.com) were chosen for this study. This presentation will describe the “data prospecting” prototype architecture, interactive data exploration capabilities and demonstrate some science discoveries that have been made using this tool.