Evaluation of Visual Analysis Environments

**Mission**

Our mission is to develop user-centered evaluations for visual analysis environments. While software performance is necessary for usability, it is not sufficient to guarantee utility for end users. User-centered evaluations measure the impact of software on the end users’ processes and products. To conduct evaluations we need data sets, user tasks, and metrics. The data sets and user tasks need to be realistic and resemble what users do in their real-world work environment. The metrics need to reflect the goals of the users and be diagnostic for the software researchers.

It is difficult to conduct evaluations in complex problem-solving domains as access to analysts is an issue. They are busy, they work in secure environments for the most part, their work is not easily accomplished in short time periods (two-hour time frames are typically used in evaluation work), and their data is often classified.

**Solutions**

We are approaching metrics development work from two perspectives. We are working with the Regional Visualization and Analytics Centers to understand their different research areas and consequently help them develop evaluation metrics, both to measure their progress and also to measure the utility to their end users.

We are working with potential end users to develop user profiles, including descriptions of tasks, constraints in the user work environment, and metrics. These profiles will be made available to the RVACs to help them understand the needs of different user types.
We have taken a more general approach in the IEEE Visual Analytics Science and Technology (VAST) symposium contests. We have run contests in 2006 and 2007 in conjunction with the VAST symposium. Participants use data sets developed by the NVAC Threat Stream Generator project and perform a specified task. To date we have developed accuracy metrics, product quality criteria, and rating criteria for the utility of the different software components and visualizations in analysis environments. Experts in usability and visualization along with senior analysts participate in reviewing the entries. Contest winners are invited to an interactive session during the symposium and spend several hours with an analyst using their software to investigate a similar problem. We observe the analysts interacting with the different software tools and use their observations to refine metrics for the next year.

**Impact**
Participants in the VAST 2008 Challenges will be invited to a workshop where the results and lessons learned will be discussed. The data sets produced for these contests are made publicly available along with the solutions. The data sets and metrics are being used by others, including university professors and corporations doing internal evaluation, and are used as a benchmark for other evaluations, such as link analysis.

The VAST 2007 contest is being revised and will become the VAST 2008 Challenge. In order to increase participation we are developing several “mini-challenges” in addition to the overall challenge.

Our National Science Foundation project titled “Establishing Scientific Evaluation Methods for Visual Analytics Science and Technology” (SEMVAST) was awarded in July 2007. This focuses on developing metrics and automated tools for evaluation at the system and component level and will seed an infrastructure for the coordination of long-term evaluation activities across the multiple core research and application domains.

**For more information, contact:**
Jean Scholtz
Phone: (503) 355-2792
jean.scholtz@pnl.gov

Mark A. Whiting
Phone: (509) 375-2237