John Stasko, Professor
College of Computing & GVU Center, Georgia Institute of Technology
Position Statement for joint NSF-NIH Workshop on Visualization

The most important "grand challenge" problems in computer science are those that go beyond the technology and address real problems of real people. These are the challenges whose solution will affect millions of people in crucial ways. A long-term challenge for the visualization community, I feel, is to develop techniques and systems that assist people in complex decision-making and analysis tasks, particularly those involving large amounts of data.

As examples consider

- A doctor analyzing a myriad of test and conditions data from a patient and seeking to determine an ailment and the best treatment
- A financial analyst who wants to recommend investing strategies for clients by examining a multitude of financial, news, and historic data
- An intelligence analyst who is combing over a wide assortment of data from the field in order to assess potential threats

In the past, acquiring data has often been a difficult roadblock. Today, with advances in science, technology, and networked information infrastructures, we can imagine scenarios close to "knowing all the data." What would it be like to have all the relevant, and for that matter irrelevant, facts available? Clearly, this is an overwhelming proposition, and such a situation would make focusing and filtering out the irrelevant items be primary operations. Therefore, our challenge shifts from data acquisition to data analysis and understanding in order to make decisions and recommend courses of action.

Visualization holds tremendous promise as an external cognition aid and a tool to help people organize and analyze data to transform it into knowledge. Essentially, this is the fundamental premise of *information visualization*, to help people give meaning to data that is formerly just numbers and text.

The primary challenge of information visualization within the broader context of all visualization is that no underlying representation model exists. That is, we do not know what information looks like. Thus, the researcher's challenge is not only "How do I make it look just so?" but "What should it look like?"

The visualization research community must develop new techniques and systems that allow people who work with complex data sets to experiment with "what if" scenarios, to examine the data under multiple perspectives and assumptions, to seek connections between any number of attributes of the data, and to understand the reliability of any conclusions reached by analyzing the data.

I suspect it will never be the case that we provide sufficient information representations and tools for all the different analysis domains that may arise. Thus, another challenge of our discipline is to enable more people who are not visualization experts to build analytic representations of their own data. We need to create a kind of "information visualization Photoshop," a system that, while clearly not comprehensive and all-powerful, does help to enable non-experts to perform tasks otherwise beyond their capabilities in any reasonable timeframe. In addition to pushing the upper boundaries of visualization, we must expand the area's reach and benefits in all other directions.