Beyond Usability: Evaluation Aspects of Visual Analytic Environments

Jean Scholtz

Pacific Northwest National Laboratory

ABSTRACT

A new field of research, visual analytics, has recently been introduced. This has been defined as "the science of analytical reasoning facilitated by interactive visual interfaces" [20]. Visual analytic environments, therefore, support analytical reasoning using visual representations and interactions, with data representations and transformation capabilities, to support production, presentation, and dissemination. As researchers begin to develop visual analytic environments, it will be advantageous to develop metrics and methodologies to help researchers measure the progress of their work and understand the impact their work will have on the users who will work in such environments. This paper presents five areas or aspects of visual analytic environments that should be considered as metrics and methodologies for evaluation are developed. Evaluation aspects need to include usability, but it is necessary to go beyond basic usability. The areas of situation awareness, collaboration, interaction, creativity, and utility are proposed as the five evaluation areas for initial consideration. The steps that need to be undertaken to develop systematic evaluation methodologies and metrics for visual analytic environments are outlined.

CR Categories and Subject Descriptors: H.5.1 [Information Interfaces and Presentation] : Evaluation/ methodology; H.5.3 [Information Interfaces and Presentation]: Group and Organization Interfaces

Additional Keywords: visualization; analytic environments; metrics

1 INTRODUCTION

Visualizations are a means of taking advantage of some of the unique capabilities of humans. The human visual channel has a high bandwidth, and we are able to take in a large quantity of visual information. However, we can typically attend to only a small portion of this information. The challenge, then, is to produce visualizations that help users focus on the most relevant or most interesting aspect of the data being presented. Our society today is faced with a rapidly increasing amount of data. Information users in all domains have more information than they can deal with.

Herbert Simon observed that, "What information consumes is rather obvious: it consumes the attention of its recipients. Hence a wealth of information creates a poverty of attention, and a need to allocate that attention efficiently among the overabundance of information sources that might consume it [18]."

_____

e-mail: jean.scholtz@pnl.gov

The goal of visualization researchers is to produce visual representations of complex relationships that help users efficiently allocate their attention. While this appears straightforward, developing metrics to measure the success of visualizations is a complex task involving a number of factors. In this paper, we focus on metrics and methodologies for visual analytic environments for information analysts. We use the term visual analytic environments to encompass software and hardware, computing power as well as display technology. A visual environment for analytic work implies an interaction between the users in the environment and the data with which they are working. This interaction uses techniques supported by the software and hardware.

Standard metrics exist today in many areas of information technology. These metrics are used to measure performance, either of the system alone or of the combination of the user and the system for interactive technologies. For example, speech recognition research uses word error rate to evaluate progress in the field. Information retrieval has long used precision and recall [24]. Usability evaluations use efficiency, effectiveness, and user satisfaction to measure human-computer performance. The metrics for speech recognition and information retrieval are independent of the user but rely on having "ground truth" of the dataset being used. Usability evaluations depend on empirical studies of users interacting with the software. To account for individual differences in users, several methods are used. Either participants in the evaluations are selected to be representative of actual users, or large numbers of participants must be used to mitigate individual differences.

The evaluation of visual analytic environments will require going beyond performance evaluations and usability evaluations. This paper discusses metrics and methodologies for evaluation including the assessment of situation awareness, collaboration, interaction, creativity, and utility. Each of these e areas is discussed along with ways to develop metrics for each area. Recommendations for an approach to metrics development are proposed.

2 USABILITY OF VISUALIZATIONS

Freitas et al. [7] suggest that usability evaluations of visualizations involve three issues: presentation of the data, interaction with the data, and the usability of the data itself. The usability of the data is based on three principles: reliability of the data; access to original data – that is, avoid changing the data in a way that the user cannot access it as it was originally; and assessing the usability of the data in the user's decision-making process. Freitas et al. also suggest that distinctions need to be made about the type of visualizations. In their work, they use only two categories: visualizations that display data characteristics and values and visualizations that display data structure and relationships. They define three classes of interaction: help and orientation; browsing and searching; and data reduction. To help define evaluation criteria, they propose five properties of visual representations that should be

considered: limitations, cognitive complexity; spatial organization; information coding; and transitions between states. For each of these properties, they define metrics. These are shown in Table 1. Interactions with visualizations also need to be evaluated. Table 2 presents metrics proposed by Freitas et al. for evaluating different types of interactions. In addition, the limitations of the visualizations should be evaluated in terms of scalability and flexibility.

Table 1. Metrics for properties of visual representations suggested by Freitas et al.

Properties of visual representations	Suggested metrics
Cognitive complexity	Data density; data dimension; display of relevant information
Spatial organization	Logical order; occlusion; display of details; reference context
Information coding	Information mapping; realistic techniques
State transitions	Image generation time; visual spatial orientation

Table 2: Metrics proposed by Freitas et al. for interactions with visualizations.

Type of Interaction	Suggested Metrics
Orientation and help	Control of additional detail; undo; representation of additional information
Navigation and querying	Selection of objects; viewpoint manipulation; geometric manipulation; growing; searching and querying
Data set reduction	Filtering; clustering; planning

Data visualization metrics have been proposed by Brath [2,3], who looked at such aspects as the number of data points, the density of data points, the number of simultaneous dimensions, appropriate representations, and the percentage of data that are occluded.

Tufte [21,22,23] also provides valuable guidelines and examples for visualization. Tufte provides examples of simple visualizations for complex and multiple types of data, including the famous Napoleon's March visualization. He uses backgrounds to plot data, thereby giving data a context for rapid and increased comprehension.

These metrics and evaluations at this level are invaluable. However, they are only the starting point for evaluations when we consider visual analytic environments.

One limitation of the work reviewed above is that it seems to have neglected the problem of change over time as a relationship with a major emphasis on static visualizations. In the context of visual analytic work an important aspect of what an analyst may want to see is some sort of change over time. Indeed, the final solution to an analytic problem may emerge only after a period of the visual analog of "onion peeling" or "pearl growing" which have been studied in the context of change over time in bibliographic database searching [8].

Amar and Stasko [1] propose a framework for the design and evaluation of visualizations based on "analytic gaps." They define these gaps as "obstacles faced by visualizations in facilitating higher level analytic tasks". In particular they define tasks for two types of gaps: rationale gaps and world view gaps. A rationale gap is the gap between seeing a relationship and being able to explain the nature of the relationship. A world view gap is the difference between what is shown in the visualization and the visualization that would be needed in order to be directly used in a decision-making process. Amar and Stasko propose three subtasks to evaluate if these gaps have been eliminated. For tasks to test rationale gaps, they propose the subtasks of exposing uncertainty, concretizing relationships, and formulating cause and effect. For evaluating world view gaps, they propose the subtasks of determining domain parameters, constructing multivariate explanations, and confirming hypotheses. This framework represents a shift to user-centered metrics. The proposed subtasks go beyond the perceptual tasks of viewing the visualization and reporting information about the individual pieces of data. Amar and Stasko are focused on the utility that the visualization provides to the analyst in a decision-making task.

While this represents an excellent start, analytic tasks performed by information analysts are at a higher level than those suggested by Amar and Stasko. Therefore it would be appropriate for visualizations to provide support at a higher level. This support can be classified into areas or aspects and that metrics can eventually be identified to assess support in these areas. These areas are discussed in the following section.

3 AREAS FOR EVALUATION

Performance metrics are certainly needed. If a software algorithm is developed to calculate and display the highest ranked documents in a document collection given a particular ranking scheme, then the software must be evaluated to ensure that it performs correctly. If the software algorithm is developed to do this calculation on large amounts of data in a short time, then performance evaluations are needed to verify this.

Usability evaluations are also needed as visualizations are effective only if users can interpret the information provided and use the interactions provided by the software to manipulate the data and produce different views. The software must support intuitive interactions so that analysts can focus on the information they are manipulating, not the manipulation technique itself. Therefore, in addition to the typical usability measures of effectiveness, efficiency, and user satisfaction, we add a measure of cognitive workload. Visual analytic environments should strive to reduce the overall cognitive workload of analysts today. If analysts have to think less about the user interactions, they will have more time to think about analysis.

However, visual analytic environments will be used by multiple analysts from multiple disciplines for multi-source analysis. We propose that the areas of situation awareness, collaboration, interaction, creativity, and utility need to be considered as well.

3.1 Situation Awareness

Information analysts seek information used in sense making. Sense making is an understanding of a given situation at a given period of time. Thus, one way to evaluate visualizations is to assess the user's situation awareness as gleaned from the visualization. Endsley [4] defines three levels of situation awareness (SA): perception, comprehension, and projection. Perception is the basic level of situation awareness (SA level 1).

This level of awareness is achieved if operators are able to perceive in the user interface the information that is needed to do their job. The next level is comprehension (SA level 2). Not only must the information be perceived, it must be combined with other information and interpreted correctly. The third level (SA level 3) is projection or the ability to predict what will happen next based on the current situation. As situations are dynamic, time is critical to situation awareness as well. User interfaces need to be designed to facilitate the continuous acquisition of SA.

Operator interfaces for control of semi-autonomous systems, such as aircraft, power plants, and manufacturing systems, have been assessed for situation awareness. Performance-based methods, knowledge-based methods, subjective ratings, and direct assessment methodologies have been used to evaluate operators' SA.

Performance or outcome methods look at the outcome [15]. Was the correct decision made? One problem with using outcome methods is that even people with perfect situation awareness can make bad decisions; hence, the outcomes will be less than ideal. The converse is also true. People with less than ideal situation awareness can make decisions that turn out to be correct.

Knowledge-based methods are used in experimental conditions to assess situation awareness. These methods isolate particular components and assess them individually. Knowledge-based methods are better at uncovering declarative information than procedural information. Verbalization methods, such as think-aloud and talk-aloud protocols [6] are also used to discover what information users are relying on when making their decisions.

Subjective measures [11] ask users to assign a numerical value to their situation awareness at any given time. While this gives us an indication of the level of awareness, it fails to help us understand what information is missing. This method, however, can be used in an operational environment.

Scholtz et al. [16,17] have developed a methodology for evaluating human-robot interfaces modeled after the Situational Awareness Global Assessment Technique (SAGAT) developed by Endsley [5]. This evaluation methodology uses expert knowledge to develop questions that assess the users' awareness of a particular situation. A simulation is used, and the user is stopped during the simulation and given a quick series of questions to answer. These questions assess the three levels of situational awareness. After answering these questions, the users are returned to the simulation. The SAGAT methodology allows comparison of interface designs with respect to how well each facilitates the acquisition of situation awareness by users. The questions used in the assessment are based on the domain and must be developed in conjunction with domain experts.

In addition to awareness of the situation being researched, users also need to be aware of where they are in their own process, as well as what others are doing if the environment is being used collaboratively.

3.2 Collaboration

The usual questions in collaboration environments are who, what, where, and when. If multiple people are working in a synchronous environment, it is important to be able to determine who is doing what and where in the collaboration environment people are working. In an asynchronous environment, we are interested in knowing who did what, when it was done, and where we can find the work. There has been extensive research on providing this type of awareness in CSCW (computer supported cooperative work) environments.

Information analysts follow an iterative process of information seeking and sense-making. It is helpful for analysts to know what sources they have already visited, what queries they have already done, and what information they have already incorporated into their sense-making. It may also be useful for analysts working together to have an understanding of this same information for their collaborators. We keep track of this type of information in many ways: to-do lists explicitly written down, lists in our heads, formal representations in project management software, and sticky notes, to mention a few. If this type of information is provided in the user interface, an SA measure could be developed to evaluate it.

Collaborators may have different domain expertise. That is, one analyst might be a signals intelligence expert and might want to view data in an entirely different way than a political analyst. The ability to share and discuss data at a data level while using different views is a necessary feature of visual analytic environments that are collaborative.

As systems become more intelligent and act more like human collaborators, analysts will want to know what the system is doing and why the system is making recommendations. Metrics developed for evaluating collaboration should include the typical who, what, when, and where. As analysts need to justify their recommendations, the why also becomes important.

3.3 Interaction

Visual analytic environments will not rely solely on static displays. Users will be able to interact with the displays. While usability is certainly one aspect of interaction, it will be necessary to go beyond usability to evaluate the capabilities provided by the interaction. Such capabilities should include:

- ability to view occluded information
- ability to move up and down in the level of abstraction of the views
- high-level "undos" (see discussion on creativity)
- data sharing, not view sharing (see collaboration).

This is only a rudimentary beginning. This list of capabilities will certainly be increased through contextual work with analysts using their current tools.

Wehrend and Lewis [25] identified 11 tasks that one might do with a visualization:

- Locate
- Identify
- Distinguish
- Categorize
- Cluster
- Distribute
- Rank
- Compare between entities
- Compare within relations
- Associate
- Correlate.

Zhou and Feiner [26] added additional tasks to this list: identify; encode; emphasize; reveal; generalize and switch. Is it possible or even preferable to provide interactions at this level? The actual set of interactions need to be identified by working with information analysts to determine what it is they actually want to achieve using visualizations and how much effort is currently needed to carry out such tasks currently. For each of the interactions identified, it will be necessary to ensure the usability of each as well as to evaluate the utility of the interaction. This means that the various interactions should provide some insight into the information. This may not be true for all types of data or for all types of analysts. As we all differ

in the ways we can best comprehend information, a number of different views should be present for the analyst to use. In addition, analysts should be able to easily configure their workspaces to use the views they use most often.

ISO 9241 – Part 10 provides seven principles of dialog (human-computer interaction):

- Suitability for the task
- Self–descriptiveness of the action
- Controllability
- Conformity with user expectations or consistency
- Error tolerance
- Suitability for individualization or customization
- Suitability for learning.

These principles are certainly appropriate for evaluating a given set of interactions provided within a visual analytic environment. The challenge is in the first principle—suitability for the task. We must consider at what level(s) we want users to be able to interact. Should this be at the level of comparing entities, comparing categories, or comparing different combinations of categories?

3. 4 Creativity

Creativity is not a term usually associated with analysis. However, as Hewett [9] has pointed out, there are reasons to believe that creative work takes place in a large variety of contexts and domains. Furthermore, Hewett [9] has argued that design considerations for supporting creative work must take into account the design and development of the entire computer-based support environment.

The National Science Foundation recently sponsored a workshop on user interfaces to support creativity [14]. Based upon an analysis of work reported in The Handbook of Creativity [19], this report proposes that support tools for creative or innovative work must recognize that creativity is a multi-dimensional construct that involves a problem space with at least seven dimensions. These problem space dimensions are

- Creativity is a property of people, products, and a set of cognitive processes.
- Creativity can be thought of as a personal, social, societal or cultural phenomenon.
- Creativity can be common and frequent, or rare.
- Creativity may be domain-specific or domain-independent.
- Creativity may be qualitative or quantitative.
- Creativity can occur in an individual or in a group.

The multiple dimensions of creativity imply that support tools for creativity must take into account the environment in which work will take place. For example, designers of group brainstorming tools have made contributions anonymous as this encourages people to speak more freely when in groups.

One way to approach metrics for creativity is to consider methods for studying creativity [13]. These include:

- the use of psychometric tests
- experimental methodologies
- biological methodologies
- computational modelling
- contextual studies.

If we want tools to support creativity, those tools should enhance the personal experience of the user(s), improve the products or outcomes, and improve the processes used in creating the product or producing the outcome.

The National Science Foundation (NSF) study group suggested the following metrics to be used in evaluating creativity support tools:

- Quality of solutions
- Number of unique alternatives considered
- Degree of radicalism/conservatism of alternatives considered
- Serendipitous solutions
- Time to come up with solutions
- Satisfaction with solutions
- Cost (person-time) to come up with solutions
- Cost of the solution versus the utility of the solution
- Ease of use of the support tool
- Buy-in to the use of the support tool.

3.5 Utility

One of the most important measures of visual analytic environments is the utility of the environment from the user perspective. This includes metrics already suggested for evaluating creativity. The focus here is on the improvement of the process and an increased quality of product. Process time should be decreased, or if not, the individual should be able to produce a better product in the same time as a baseline condition. It is also preferable if new tools and environments can reduce the cognitive workload on users. In essence, the environment should allow the user to spend more time on task and less time on the tool or environment being used.

Quantitative measures for evaluation should include:

- number of citations in the analytic product
- analyst's confidence in product recommendations
- number of hypotheses investigated in product.

We have had senior analysts rate a number of analytic products (on the same task) and provide us with their criteria. We are in the process of compiling and analyzing this data to determine if there is a correlation between more highly rated products and our quantitative measures.

3.6 Other Concerns

Obviously, there are also logistics and interoperability issues involved. Data must be quickly imported in the environment and easily exported as well. Interoperable knowledge representations underlie all of this work. While necessary for effective and efficient visual analytic environments, this area of research is out of the scope of visual analytic environments. In addition, the area of trust should be considered for intelligent software algorithms. End users will take advantage of intelligent software only if they trust the information returned, hypotheses suggested, and measures of uncertainty computed.

4 PROCESS FOR METRICS DEVELOPMENT

In the previous section, five areas for evaluation of visual analytic environments were presented. For each of these areas, some possible measures were presented. Future studies are planned to determine the methodologies to use to accurately capture the needed data. Moreover, these metrics will need to be validated. That is, to determine if these metrics discriminate "good" systems from "not so good" systems.

In addition, metrics should also be defined to determine if the rationale for the development of visual analytic environments (VAE) are supported. Table 3 provides a starting set of hypotheses. For each of these hypotheses, it is possible to construct measures for that hypothesis and methodologies for capturing the measurements. These hypotheses and measures should be considered only as discussion points and will need to be refined and later validated in experiments.

The third column of Table 3 maps the hypothesis to a possible contributing area(s) of evaluation. In upcoming evaluations the

quantitative measures listed in column two will be collected along with measures for the different areas. The resulting data will be analyzed to determine which hypotheses are supported and to ascertain which evaluation areas contribute to supporting the various hypotheses. A similar methodology was used in producing evaluation methodologies for question-answering systems [12].

The validation procedure proposed is to work with senior analysts to develop criteria for ranking analytic products. These

criteria will be applied to products produced in experiments. The various measures for the hypotheses and areas of evaluation will be analyzed to determine if these measures produce results that are in line with the product rankings.

5 NEXT STEPS

Having formulated some starting hypotheses, the next steps are to start discussions with researchers in visual analytic environments, users of visual analytic environments, and managers of visual analytic environment programs—in essence, the stakeholders. The metrics need to reflect the needs of the researchers. The metrics also need to be based on the needs of other stakeholders. Once that is done, appropriate methods for collecting the necessary data will need to be determined and eventually the metrics will need to be validated. It will be necessary to identify current environments and produce baseline metrics that can be used for subsequent comparisons. Researchers should be encouraged to try out subsets of these metrics in their ongoing work to determine which metrics are useful in providing feedback on their work.

Table 3. Hypotheses and measures.

Hypothesis	Possible measures	Contributing Evaluation Area(s)
VAE should increase the amount of information/ data that analysts can incorporate into their analysis	-Number of documents looked at -Number of entities considered (some notion of relationships, attributes) - Amount of evidence extracted and used	Interaction
Visualizations (created by the analyst) should increase the comprehension of analytic products	-Accuracy of customer comprehension of analytic products - Accuracy of comprehension of analytic products by other analysts	Interaction Creativity
VAE should increase the efficiency and effectiveness of analytic collaboration	-Time for each party in collaboration to comprehend a situation -Accuracy in shared understanding	Collaboration Situation Awareness
VAE will increase the number of alternative paths that analysts are able to explore	-Increased analyst confidence in products -Paths explored	Interaction Utility
VAE should allow analysts to view information at different levels of abstraction efficiently and effectively	-Time to answer questions about different abstraction levels -Accuracy in answering questions	Interaction Creativity
VAE should allow analysts to efficiently and effectively view data from multiple intelligence sources	-Time/accuracy in answering questions about different intelligence sources	Interaction Situation Awareness
VAE should improve the comprehension that analysts have of the situation	-Situation awareness measures for levels 1, 2, and 3	Situation Awareness
VAE should improve the process of the analysts	-Time spent in each step of the process -Time spent in overhead (tool use) -Productivity of information-seeking tasks	Interaction Usability Utility
VAE should improve the analytic products	-Quality of products -Understanding of customer	Situation Awareness Utility
VAE should allow the analysts to be more creative in their work	-Quality of products -Radical nature of alternative solutions considered	Creativity Utility

As a case study, the results of a visualization contest for the 2006 Workshop of the IEEE Symposium on Visual Analytics Science and Technology (VAST) will be analyzed. A dataset and a set of questions to be answered by contest participants using their visualizations has been developed by the contest committee. These questions have been developed in line with Endsley's three levels of situation awareness. Participants in the contest are required to submit both the answers to the questions and the process they used to arrive at the answers. Robust visual analytic environments may apply for the opportunity to participate in a "live" version of the contest, where analysts will work with the systems to answer similar questions. This will allow the contest judges to view the process and the products produced by these analysts within a given time frame.

The metrics proposed here are specifically to support information analysts using visual analytic environments. However, visualizations and environments should support analytic work in numerous domains. Once the metrics and methodologies for the intelligence domain have been refined, the next step will be to apply them in other domains such as medicine, health care, weather research, and education. The National Institutes of Health (NIH)/NSF report on Visualization Research Challenges [10] outlines the following challenges for visualizations (among others):

- Systematically explore the design space of visualizations.
- Evaluate the effect of visualization techniques and approaches.
- Collaborate with domain experts and use applied problems to drive and evaluate visualizations.

The metrics presented in this paper should provide ways to assess the progress being made in addressing these challenges.

REFERERNCES

[1] Robert A. Amar, and John T. Stasko: BEST PAPER: A Knowledge Task-Based Framework for Design and Evaluation of Information Visualizations. INFOVIS 2004: 143-150.

[2] Richard Brath, Interactive Information visualization Guidelines. *Proceedings of HCI International '97.*

[3] Richard Brath, "Metrics for Effective Information Visualization". In *Proceedings of IEEE Symposium on Information Visualization '97,* 108-111. IEEE, 1997.

[4] Mica Endsley, "Theoretical Underpinning of Situation Awareness: Critical Review" in Mica R. Endsley and Daniel J. Garland (Eds.) *Situation Awareness Analysis and Measurement.* Lawrence Erlbaum Associates, Mahwah, New Jersey, 2000. 3-32.

[5] Mica Endsley, "Design and Evaluation for situation awareness enhancement". In *Proceedings of the Human Factors Society 32nd Annual Meeting,.* Santa Monica, CA: Human Factors Society. 1988. vol. 1, 97-101.

[6] K.Anders Ericsson and Herbert Simon, *Protocol Analysis: Verbal Reports as Data.* Cambridge, MA. MIT Press. 1993.

[7] Carla Freitas, Paulo Luzzardi, Richard Cava, Marco Windler, Marcelo Pimenta, and Luciana Nedel, "Evaluating Usability of Information Visualization Techniques." In *CHI 2002 – 5th Workshop on Human Factors in Computer Systems*. Brazilian Computer Society Press. 2002.

[8] Thomas Hewett and Sari Scott, "The use of thinking-out-loud and protocol analysis in development of a process model of interactive database searching." In H.-J. Bullinger & B. Shackel (Eds.), *Human computer interaction--INTERACT '87* (pp. 51-56). Amsterdam: North-Holland. 1987.

[9] Thomas Hewett, "Informing the design of computer-based environments to support creativity." *International Journal of Human-Computer Studies,* 63, 383-409, 2005.

[10] Chris Johnson, Robert Moorhead, Tamara Munzner, Hanspeter Pfister, Penny Rheingans, and Terry Yoo. NIH/NSF Visualization Research Challenges, Jan. 2006.

[11] Debra Jones, "Subjective Measures of Situation Awareness" in Mica R. Endsley and Daniel J. Garland (Eds.) *Situation Awareness Analysis and Measurement.* Lawrence Erlbaum Associates, Mahwah, New Jersey, 2000. 113-128.

[12] Diane Kelly, Paul Kantor, Emile Morse, Jean Scholtz, Ying Sun. "User-centered Evaluation of Interactive Question-answering Systems, Proceedings of the Workshop on Interactive Question Answering at the Human Language Technology Conference (HLT-NAACL '06), New York, NY, June. 2006.

[13] Richard Mayer, "Fifty Years of Creativity Research" in Robert Sternberg (Ed.) *Handbook of Creativity*. Cambridge University Press, Cambridge, UK, 1999.

[14] NSF Workshop Report Creativity Support Tools, **http://www.cs.umd.edu/hcil/CST**, accessed March 1, 2006.

[15] Amy Pritchett and R. John Hansman, Use of Testable Responses for Performance-Based Measurement of Situation Awareness" in Mica R. Endsley and Daniel J. Garland (Eds.) *Situation Awareness Analysis and Measurement.* Lawrence Erlbaum Associates, Mahwah, New Jersey, 2000. 189-209.

[16] Jean Scholtz, Brian Antonishek, and Jeff Young, "Evaluating human-robot interfaces: development of a situational awareness assessment methodology". Hawaii International Conference on System Science 37 (HICSS 37). Hawaii, Jan.2004.

[17] Jean Scholtz, Brian Antonishek, and Jeff Young, "Implementation of a Situation Awareness Assessment Tool for Evaluation of Human-Robot Interfaces." *System Man and Cybernetics Journal, Part A.* 2004.

[18] Herbert Simon, "Designing organizations for an information-rich world". In Martin Greenberg (Ed.) *Computers, Communications, and the Public Interest*, pp. 40 – 41. Johns Hopkins Press, 1971.

[19] Robert Sternberg, *Handbook of Creativity.* Cambridge University Press, Cambridge, UK, 1999.

[20] James J. Thomas and Kristin A. Cook (Eds.), *Illuminating the Path: the Research and Development Agenda for Visual Analytics.* 2005. IEEE Press

[21] Edward Tufte, The Visual Display of Quantitative Information. Graphics Press, 1983.

[22] Edward Tufte, *Envisioning Information*, Graphics Press, 1990.

[23] Edward Tufte, *Visual Explanations*, Graphics Press, 1996.

[24] Ellen Voorhees and Lori Buckland (Eds), *The Fourteenth Text REtrieval Conference Proceedings (TREC 2005)* Gaithersburg, MD. November 15-18, 2005.

[25] Stephen Wehrend and Clayton Lewis, "A problem –oriented classification of visualization techniques." *Proceedings IEEE Visualization '90.* 139-143, 1990.

[26] Michelle Zhou and Steve Feiner, "Visual task characterization for automated visual discourse synthesis." *Proceedings of CHI '98.* 392-399,1988.