Evaluating Visualizations to Unearth Behavior and Insight

Jeffrey Heer
Computer Science Division and Berkeley Institute of Design
University of California, Berkeley
Berkeley, CA 94720-1776
jheer@cs.berkeley.edu

ABSTRACT

In this position paper, I describe three case studies of evaluations of visualization systems focused on social interaction. For the Vizster social network visualization, we conducted qualitative observation of a public installation of the visualization. For the sense.us collaborative visualization environment, we conducted both unstructured usage sessions and a live deployment on the IBM corporate intranet, followed by a formal content analysis of user activity. To evaluate the use of visual social navigation cues in social data analysis environments, we used a task-driven experimental design to measure the number of unique discoveries uncovered by subjects. Though quite different in terms of methods, each of these evaluations helped unearth how the visualization contributed to an analysis process (whether formal or informal) and revealed patterns of behavior and insight generation that have served to guide future designs.

1. Vizster

Before the advent of sites such as MySpace.com and Facebook.com, Friendster.com was the first social networking service to achieve widespread viral growth, amassing over a million members in 2003. Based on insights from a 9-month ethnographic study of Friendster users, danah boyd and I designed Vizster [2], a social network visualization to help end-users of online social networks better explore their online communities. Our primary goal was to augment users existing, playful practices on social network services. These practices included exploring social connectivity, hunting for friends, and playing games with online identity production. As such, neither task time nor errors were of particular relevance to our design goal. Instead, we were interested in users' subjective experiences and higher-level patterns of visual exploration.

After conducting standard usability sessions to correct basic usability issues and solicit additional design features, we deployed Vizster in a semi-public space: a 500 person all-night party in San Francisco. The attendees were highly representative of Friendster's early adopters: young adults affiliated with tech culture and/or the Burning Man arts festival. Our deployment consisted of both a kiosk with which users could interact and a large projected display allowing on-lookers to follow along. We then observed usage and took field notes throughout the night.

Our study netted two sets of insights. First, we observed a number of qualitative usage patterns. Usage was routinely coupled with social play—for example, games to see if one could hunt down specific people they knew. Onlookers tended to participate as well, engendering a shared experience around the visualization as people retold stories about people, their relationships, and shared experiences. Particularly salient were the stories told in response to our semi-automatic community analysis features. Comments included "Look, it found all the people from Harvard" and "It was smart enough to know not to group them together, they hate each other." In general, people imparted more sophistication to the algorithm than was actually implemented.

Second, we noticed a difference between usage by groups and by individuals. Group story-telling and inter-group challenges to find particular friends or communities led to longer usage sessions and deeper analyses (though of an informal nature) than exhibited by individual usage. This insight led us to consider how social activities that deepen analysis might be incorporated into more visualization systems.

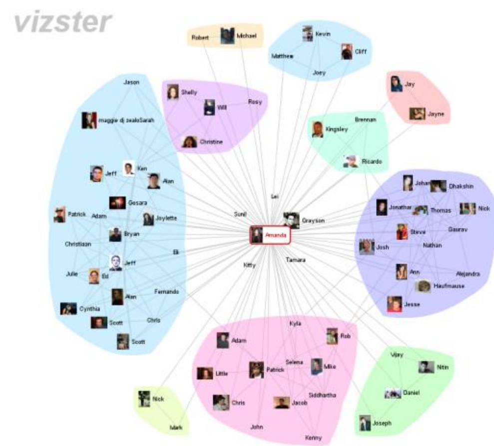

Figure 1. Vizster, a visualization of online social networks. Colored regions show community groupings inferred by the system.

2. sense.us

The insights gained from our evaluation of Vizster, coupled with observations of social data analysis on the web around the Name Voyager applet [4], inspired FernandaViégas, Martin Wattenberg, and I to design sense.us [3], a web site for collaborative analysis of 150 years of U.S. census data. The site provides interactive visualizations of data concerning immigration, labor, population, and other demographic variables. In addition, we developed features to support sharing both within and external to the site: application bookmarking, discussions doubly-linked to visualization views, graphical annotations, and bookmark trails enabling the creation of both tours and inter-view linking.

To better understand asynchronous collaboration practices around visualizations, we conducted a study of the sense.us system. Our study had two goals: first, to better understand emergent usage patterns in social data analysis; second, to learn how well the various features of the sense.us system supported this analysis. We ran the study in two parts: a pair of controlled lab studies and a 3-week deployment on the IBM corporate intranet. To analyze the data, we employed a mixed-methods analysis approach [1] combining qualitative and quantitative observations.

In our lab studies, we observed subjects in unstructured usage sessions—to better simulate web usage, no tasks were assigned. A think-aloud protocol was used so that we could follow subjects' thought processes. In our live deployment, we logged all user actions on the site, while also monitoring other aspects of IBM's intranet, including blogs and a social bookmarking service. We provide a full description of the study results elsewhere [3].

Here, we relate two findings of interest. First, we conducted a formal content analysis of comments posted to the site to analyze the forms of collaboration. This resulted in identifying cycles of collaborative data analysis involving observations, questions, and hypotheses around the data. In particular, the inclusion of questions and hypotheses in one's comments appears to increase the probability of responses by others. Furthermore, we found that many comments did not only concern the data itself, but also reflected on the process of data production, such as collection biases or missing data points. Users helped flag problematic data that might lead to false inferences if taken at face value.

Second, by running both lab studies and a live deployment, we gained insights not achievable by one approach alone. The think-aloud protocol in our lab studies enabled us to note issues that would be difficult to discern from usage logs, such as subject's reactions to comments (e.g., laughing or saying "I agree" out loud) and their motivations for particular actions. In the live deployment, we were first disappointed by a lower-than-expected number of comments being left on the site. We then found that users were using our bookmarking feature to instead post their findings to IBM's social bookmarking service, which was already widely adopted and highly visible within the organization.

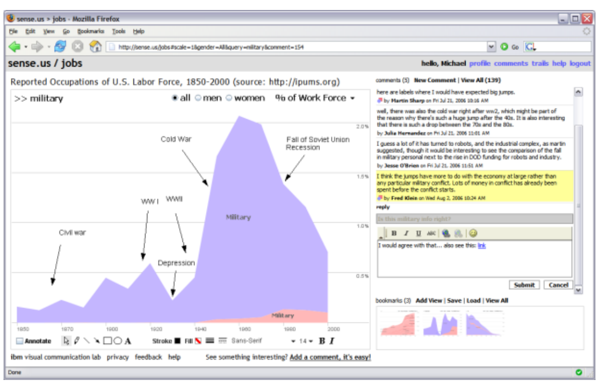

Figure 2. The sense.us collaborative visualization environment. The left side shows a visualization with graphical annotations. The panel on the right includes collaborative commentary on the current view.

3. Scented Widgets

Motivated by the ways in which sense.us users engaged in social navigation to find interesting topics for further analysis, Wesley Willett, Maneesh Agrawala, and I developed *Scented Widgets*, a general framework for embedding visualizations in navigation controls (e.g., dynamic query widgets) to enhance exploration [5]. We conducted a controlled experiment in which scented widgets were used to provide social navigation cues in a collaborative data analysis environment.

In our study, we asked subjects to explore a visualization of labor statistics and collect evidence concerning a task hypothesis, such as "Technology is costing jobs by making occupations obsolete." We asked subjects to make at least seven observations that provided evidence for or against the current task hypothesis. At least two of the observations had to be unique findings on views not yet commented upon. Subjects noted their observations by leaving new comments on corresponding views. Across tasks, we varied the dynamic query widgets to display visualizations of comment counts, visit counts, or neither for destination views.

Our goal was not to discover if embedded visualizations made tasks faster or less-error prone; rather, we wanted to discover if they would result in more unique discoveries by indicating under-explored regions of the data. To compute a metric of unique findings we collected all comments on visualization states that previously had none, followed by manual filtering to remove jokes and unrelated questions. The result was a count of unique discoveries made in each task trial. For the first task performed we found a significant advantage for social navigation cues, but this equalized over subsequent trials. In other words, scenting led to more unique discoveries when users were unfamiliar with the data, but these benefits equalized over time, suggesting a transfer from social to semantic navigation.

4. Conclusion

In the aforementioned studies, different evaluation approaches were employed to assess patterns of usage and discovery within interactive information visualizations. With Vizster, qualitative observation of a public installation revealed unexpected usage patterns and suggested new research directions. For sense.us, multiple studies employing a mixed-methods approach allowed us to triangulate patterns of both individual and social interest. We also found content analysis to be a useful method for reliably coding data across collaborators with differing opinions and inclinations. Finally, for "scented widget" navigation cues, we used a standard experiment design, but one targeted at measuring unique insights. In each case, our evaluations served to measure higher-level aspect of the visual analysis process, unearthing the behaviors constituent of the analysis process, contextual factors influencing those behaviors, and the resulting insights.

5. References

[1] Creswell, J. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Sage Publications, 2002.

[2] Heer, J. and boyd, d. Vizster: Visualizing Online Social Networks. *Proc. IEEE InfoVis '05*. Oct 2005.

[3] Heer, J., Viégas, F., and Wattenberg, M. Voyagers and Voyeurs: Supporting Asynchronous Collaborative Information Visualization. *Proc. ACM CHI '07*. Apr 2007.

[4] Wattenberg, M. Baby Names, Visualization, and Social Data Analysis. *Proc. IEEE InfoVis '05*. Oct 2005

[5] Willett, W., Heer, J., and Agrawala, M. Scented Widgets: Improving Navigation Cues with Embedded Visualizations. *Proc. IEEE InfoVis '07*. Oct 2007