RESEARCH STATEMENT

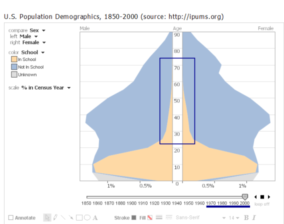
Jeffrey Michael Heer

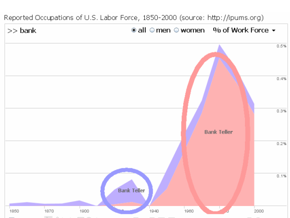
As society generates increasing amounts of digital information, new computational and perceptual aids are needed to analyze and communicate the escalating quantity of data. To this aim, visualization research leverages human visual processing and cognition to increase the scale of information with which humans can effectively work. My research in human-computer interaction focuses on the design, implementation, and evaluation of novel computational tools for the visual analysis and communication of information. I employ user-centered design methods grounded in analyses of real-world phenomena and evaluated with representative users. I adopt a holistic view of the visualization process, conducting research in three areas: collaborative visualization systems that leverage the insights of multiple analysts; software tools to improve and simplify visualization creation; and novel visualization techniques for more effectively exploring data.

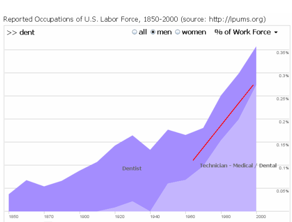
COLLABORATIVE VISUAL ANALYTICS

Most visualization research to date focuses on a single-user model, considering the loop between a user and an interactive display. However, this approach overlooks the social nature of visual media. Visualizations are used not only to explore and analyze, but to communicate findings. People may disagree on how to interpret data and contribute contextual knowledge that deepens understanding. Furthermore, some data sets are so large that thorough exploration by a single person is unlikely. Such scenarios arise regularly in scientific collaboration, business intelligence, and public data consumption. To address these issues, my dissertation research explores collaborative visual analytics, building both targeted techniques and holistic systems for supporting social sensemaking.

To explore the potential of collaborative visualization, we built *sense.us*, a web application for social data analysis of 150 years of United States census data. As described in our CHI 2007 paper, the site features visualizations of demographic data and novel features for collective analysis. Users can attach commentary and annotations to views, share collections of views, and engage in discussion. Novel bookmarking and indexing features facilitate view sharing and help reduce cross-talk between related visualization states. We studied usage of the system through a live deployment and a series of laboratory studies, and conducted a content analysis of recorded usage. Our results revealed patterns of social data analysis, as users combined their knowledge in cycles of observation and hypothesis to make sense of trends in the data. For example, an observed decline in the number of dentists netted multiple explanations, including the fluoridation of the water supply and an increasing stratification between dentists and hygienists over the last century.







Annotated views from social data analysis in sense.us. Users can annotate, discuss, and link visualization views.

1. The rise of adult education from the 1970s onward
2. Reversal of the dominant gender of bank tellers
3. Stratification of dentistry into dentists and hygienists

Scented widgets embed visualizations into user interface controls to aid navigation. For example, bars might indicate the number of previous visits to an adjacent interface state.

We have applied the insights gained from sense.us to develop mechanisms that further support social processes in visualization. Our INFOVIS 2007 paper on *scented widgets* introduces a framework for embedding visualizations in common user interface controls. We have used this system to provide social navigation cues in sense.us, visualizing visitation and comment counts for visualization views reachable from the current state. A controlled experiment found that such cues can simultaneously promote visits to popular or controversial views and, by revealing under-visited regions of the data, increase the number of unique discoveries made by users.

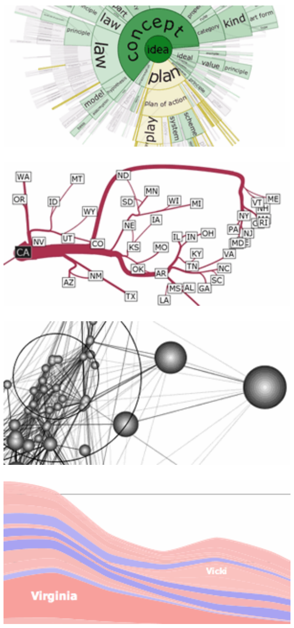
To aid analysis and communication, we also created data-aware selection techniques that tie annotations to underlying data sets by modeling selections as declarative queries. As described in our CHI 2008 paper, this representation allows reuse of annotations across varied visual encodings, allowing collaborative annotations to be verified under diverse viewing conditions. Our direct manipulation techniques couple declarative selection queries with a *query relaxation* engine with which users can generalize their selections along multiple dimensions. We conducted a study and found that users were significantly more accurate at creating selection queries when using our generalization techniques.

Future Research. Research in collaborative visual analytics is still in an early stage and many research challenges remain. Our current work focuses on new tools for weaving visualizations into the web. One phase of this research, the *flare* visualization toolkit, is described in the next section. We are combining this framework with work on selection techniques and analysis histories to create new tools for social sensemaking. We outline such challenges in our VAST 2007 paper and are writing grant proposals in this area.

More generally, collaborative visual analytics provides a lens onto the phenomenon of social computing now taking place on the web. How can the activity of various actors be channeled into useful results? For example, sites such as Wikipedia rely on human editing to integrate contributions, whereas structured approaches such as Luis von Ahn's "games with a purpose" afford statistical aggregation. I am excited to explore systems that lie on the spectrum between these examples, providing representations (*e.g.*, of hypotheses and evidence) that facilitate aggregation while also enabling unhindered communication. A related issue is how to combine the efforts of human and machine collaborators. How can human activity guide data mining processes, and conversely, how should automated techniques better suggest interesting observations or data regions in need of exploration?

TOOLKIT SUPPORT FOR INFORMATION VISUALIZATION

A consistent challenge in user interface development is finding software tools that enable novel interface designs while helping minimize development time. In the domain of



Visualizations built with the prefuse visualization toolkit. http://prefuse.org/gallery

information visualization, I noticed that although many effective techniques had been developed, wide-spread use had yet to take hold. One reason for this is that visualization applications are often difficult to build, requiring mathematical skills to implement complex layout algorithms and dynamic graphics. Another reason is that infovis applications do not lend themselves to "one size fits all" solutions; successful visualizations often reuse established techniques, but in a manner tailored to their application domain, requiring customization. While toolkits attempting to fill this gap have been created, most provide only a library of canned visualization "widgets" rather than a set of reusable techniques for building customized or novel visualization designs.

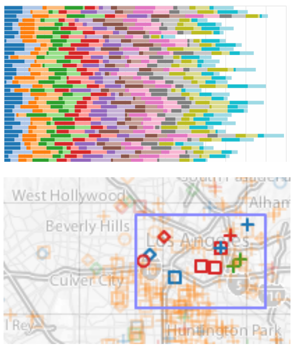
In response, I developed *prefuse*, a toolkit for authoring information visualization applications. Instead of providing infovis "widgets" that are applied much like buttons or checkboxes in traditional GUI tools, prefuse provides finer-grained building blocks for constructing tailored visualizations. As described in our CHI 2005 and INFOVIS 2006 papers, this approach simplifies the composition of established methods—such as layout and encoding algorithms, dynamic queries, and zooming—while providing an integrated structure in which to develop novel techniques and domain-specific designs.

To evaluate prefuse, we built a number of example applications to demonstrate both the expressiveness and economy of the toolkit. Additionally, we conducted usability studies of programmers using the toolkit to create new visualizations. The studies demonstrated the usability of the toolkit, but also provided insights for further design. In addition to inspiring improved abstractions, the study highlighted the importance of supplementary materials, particularly example code, in the process of learning third-party tools.

Prefuse is available as an open-source project (http://prefuse.org) and public use has provided another means of evaluation. Over three years, the toolkit has been downloaded over 40,000 times and is actively used by corporations, academic researchers, students, and hobbyists. Perhaps the most rewarding aspect has been the stream of creative visualization tools that others have built, a number of which have resulted in research papers published in top-tier conferences such as ACM CHI and IEEE INFOVIS.

More recently, we released *flare*, a toolkit for web-based visualization using Adobe Flash (http://flare.prefuse.org). Our goal is to make visualizations more accessible by weaving them into the web, providing tools for web developers and furthering our research agenda in collaborative visualization. In the first month of its release, flare has been downloaded over 2,000 times and bookmarked over 800 times on the del.icio.us bookmarking service.

Future Research. I am excited to continue developing novel tools for visualization and interface creation. We are extending our work on the flare toolkit to develop richer tools



Web-based visualizations built with the flare toolkit.
http://flare.prefuse.org

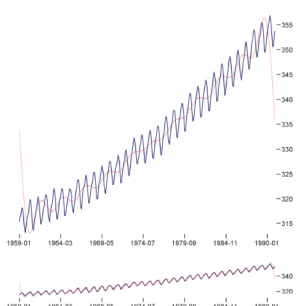
for online collaboration and communication around visualizations. I am also interested in tools for emerging display configurations, including mobile devices and large displays. For example, our CHI 2004 short paper introduced toolkit support for "seam-aware" interfaces that utilize multiple monitors in a more perceptually effective manner.

While most of the tools I have designed focus on developers, an important area of future research is the creation of end-user tools for authoring interactive visualizations. Modular component architectures such as prefuse lend themselves to declarative description, and I would like to explore both declarative languages for authoring visualizations and related graphical interfaces for creating novel visualizations. I am also interested in the further use of optimization techniques for specifying visualizations.

VISUALIZATION TECHNIQUES AND APPLICATIONS

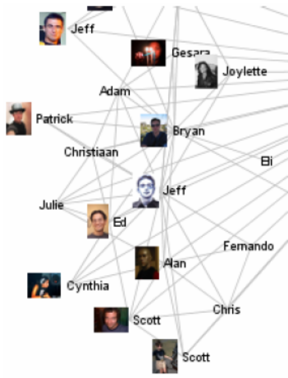
Creating perceptually effective visualizations remains a vital and exciting research area. I have developed a number of novel visualization techniques and applications and intend to continue working in this domain. My approach is to bridge real-world problems with computational techniques for effective perceptual displays of information.

Animated Transitions in Statistical Data Graphics. Though favored by many users for its engaging properties, animation is a controversial topic in information visualization, with skeptics arguing for carefully designed static depictions over animated graphics. To explore the issue and provide guidance for designers, we conducted a study of animated transitions between data graphics. Our INFOVIS 2007 paper presents guidelines and animation techniques for crafting effective transitions and details two controlled experiments demonstrating that carefully designed animated transitions can improve viewers' graphical perception at both syntactic and semantic levels of analysis.



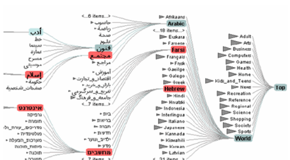
Atmospheric CO_2 levels plotted at aspect ratios determined by multi-scale banking.

Multi-Scale Banking. Subtle changes to a visualization can have significant perceptual impacts. For example, changing the aspect ratio of a line chart can change which trends become apparent: narrow aspect ratios favor low-frequency patterns, where as wider ratios better emphasize high-frequency detail. Our INFOVIS 2006 paper on *multi-scale banking* introduced techniques to automatically determine perceptually optimized aspect ratios for trends at multiple scales in the data. The technique uses spectral analysis to identify potential trends of interest, followed by perceptual optimization of the orientations of line segments in generated trend lines. Furthermore, our work proposed and compared a number of optimization approaches. The figure to the left shows the result of our technique on a plot of atmospheric CO_2 levels: the first aspect ratio emphasizes the accelerating rise of CO_2 levels, while the second aspect ratio allows the gradual onsets and steep declines of yearly fluctuations to be seen more clearly.



A segment of a social network visualized in the Vizster system.

Vizster and Enronic. Social networking services, typified today by sites such as Facebook and MySpace, first entered the mainstream consciousness with the rise of Friendster.com in late 2002. Fascinated by this social phenomenon, we built Vizster, a visualization system for end-users to explore their articulated social networks. As described in our INFOVIS 2005 paper, the design of Vizster was grounded in ethnographic study of these emerging social network services. The system visualizes network structures, user profile attributes, and community groupings. Groupings are automatically derived by clustering the link structure of the social graph. Evaluation of the system in both laboratory and semi-public settings resulted in a number of insights, including observations of end-user social data analysis that have influenced our subsequent research on collaborative visual analytics. Encouraged by the success of Vizster, we extended our software to create Enronic, a network analysis tool for investigating communication patterns in e-mail archives, particularly the infamous Enron e-mail corpus. Using the tool (and lacking prior knowledge of the actors involved), we identified the primary culprit behind the California energy crisis. Both the Vizster and Enronic systems received attention from the media, and have been featured in television, newspapers, and popular blogs.



Degree-of-Interest Tree visualization of the Open Directory Project (dmoz.org).

Degree of Interest Trees. Despite a large body of work in visualizing tree structures, we found that most existing systems suffered from scale issues at either the systems level, perceptual level, or both. Our AVI 2004 paper describes Degree-of-Interest Trees, a scalable focus+context visualization technique that uses a model of user interest to dynamically adjust the tree layout and guarantees that the tree fit in a bounded region of space. We have used the system to explore hierarchies on the order of a million nodes while maintaining real-time interactive performance. As described in our VAST 2006 paper, we have worked with intelligence analysts to extend this work to support time-varying hierarchies. The resulting visual analysis tool, called TimeTrees, has been used to investigate the movement of individuals within and between organizations over decades.

Future Research. I plan to continue developing visualization techniques and applications. For example, our recent work on animation is a first step in a larger research area; more research is needed to refine our design principles and create tools to automate the design of effective animations. As the number of relevant data sets continues to increase, so does the potential for new visualization solutions. One attractive aspect of a faculty position is the increased potential for forming partnerships with domain experts in other disciplines, enabling novel visualization techniques to be grounded in relevant problems and provide results that positively impact research in the sciences.

Electronic versions of publications (many with videos) are available at **http://jheer.org**.