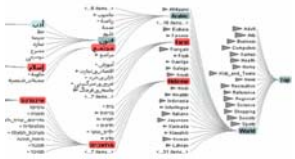
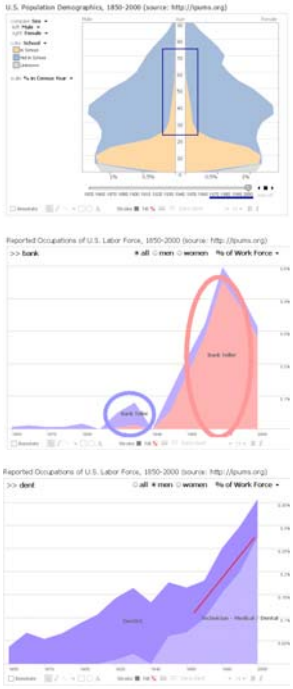
PERSONAL STATEMENT

Jeffrey Michael Heer

Visualization of word relations in the WordNet database, built using our open-source tools.



Degree-of-Interest Tree visualization of dmoz.org.



Annotated views from social data analysis in sense.us.

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

New technologies for sensing, simulation, and communication are helping people to both collect and produce data at exponential rates. Yet, acquiring and storing this data is, by itself, of little value. We must make sense of the data in order to produce real value from it. Such sensemaking—turning data sets into knowledge—is a fundamental challenge in human-computer interaction. It requires integrating data analysis algorithms with human judgments about the meaning and significance of patterns in the data. The goal of my research is to enhance our collective ability to analyze and communicate data through the design of interactive visualization tools.

My research in *visualization techniques* uses insights from studies of human perception and cognition to design visual representations and interaction techniques for data. For example, our "degree-of-interest" trees build a model of user's interest across a data set using input such as search queries and items clicked. This model then determines what information is shown and how it is displayed, smoothly updating the visualization in response to a user's changing focus of attention. We have used this technique to explore data sets on the order of a million items, including an analysis of time-varying hierarchies such as the political and military organization charts of nations.

Improved visualization techniques are of little use if they never make it into the hands of designers and developers. I also research *software architectures for visualization* that allow individual techniques to be combined to form customized visualization designs. With these systems, developers can craft visualizations by composing fine-grained "building blocks" for visual encoding and interaction. This research has led to open-source visualization toolkits that have been downloaded over 100,000 times and are actively used across academia and industry. Both myself and others have used these systems to build novel visual analysis applications for domains such as cellular protein interaction pathways, census data, online social networks, and the Enron e-mail corpus.

In deploying such systems, we observed that sensemaking is often a social process. The magnitude of available data and the diversity of expertise needed to fully analyze it require that our information interfaces enable us to work together to effectively forage, analyze, point, argue, and disseminate. My research on *collaborative visual analysis* explores how interfaces can catalyze social interpretation and deliberation. One such system is *sense.us*, a web site for social exploration of 150 years of U.S. census data. Users attach annotations to visualization views, engage in discussion, and share analysis trails; novel bookmarking and indexing features facilitate view sharing and reduce cross-talk. Our studies of system usage found that social features helped mobilize users in the process of identifying trends and generating hypotheses, and that exposing social activity regularly catalyzed new explorations by collaborators. This work has helped inspire a subsequent flowering of data sharing and visualization sites on the web, including IBM's Many-Eyes.com.

My group's ongoing research seeks to further advance these themes, designing interfaces for collecting, visualizing, and making sense of an increasing diversity of available data.