

# InfoVisExplorer

Jaroslav Tyman, Grant P. Gruetzmacher, John Stasko

College of Computing/GVU Center

Georgia Institute of Technology

Atlanta, GA 30332 USA

[dabel@cc.gatech.edu](mailto:dabel@cc.gatech.edu), [gte133q@mail.gatech.edu](mailto:gte133q@mail.gatech.edu), [stasko@cc.gatech.edu](mailto:stasko@cc.gatech.edu)

## ABSTRACT

In this paper we briefly describe 3 tools developed to visualize the history of information visualization papers. The visualization consists of a standard 3D scatterplot view enhanced with “bubbles,” lines, text, and colors aimed at making comparisons between authors and topics found in the papers. Three components were developed to translate and display raw XML data using OpenGL and Cocoa. We use the visualization tool to perform five tasks and discuss its weaknesses.

## Keywords

Information Visualization, scatterplot, bubbles, conference publications, publications, references, links, OpenGL, evaluation, dynamic query

## INTRODUCTION

There are many hundreds of articles published covering various topics related to information visualization. The systems described here are entries into the InfoVis 2004 competition. Systems are to support four tasks outlined on the project website. The system consists of two major sections: a set of programs to translate, view, and organize the massive amount of data, and the visualization program.

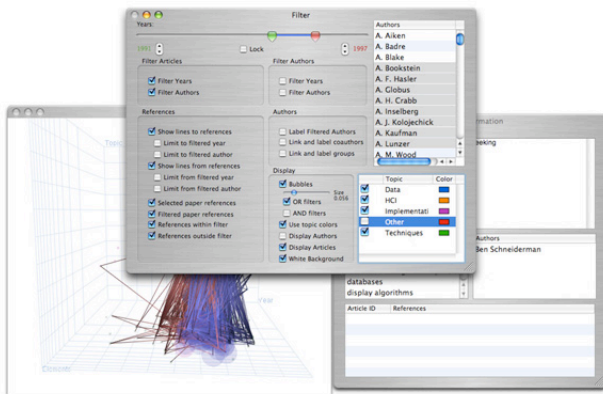


Figure 1: InfoVisExplorer interface.

## DATA

Our aim was to minimize modifications of the data, even though we were aware that the data set was incomplete and contained errors. We used our XMLImport tool to translate the XML into a Cocoa friendly property list. We then used InfoVisBrowser to browse the resulting data and group keywords into topics.

We decided to categorize the keywords into five categories based on topics presented in an Information Visualization class at Georgia Tech: data, HCI, Implementation, Other, and Techniques.

## Visualization

Once both the data file and the topic file are loaded, there are several ways to modify the visualization. The visualization begins merely as an enhanced scatterplot, influenced by Film Finder[2] and SpotFire, displaying articles according to year, topic category, and number of elements.

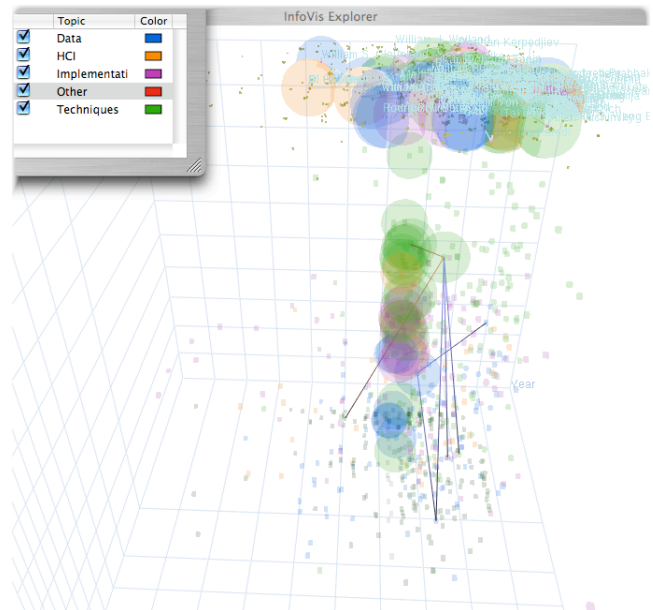


Figure 2: Topics and topic colors shown, along with bubbles (colors corresponding to the legend in the top left), reference links (red and blue), and a complete connected set of authors for the filter (top of the picture, text in light blue).

Papers are placed on the topic axis according to a weighted average of the topics that the keywords in the article fit in. The elements axis maps the sum of the number of keywords, authors, and references plotted on a logarithmic scale. The logarithmic scale was necessary in order to neutralize the effect of a single article with an element count of over 150, whereas most papers range below 30 elements.

Along the top plane of the 3D scatterplot is a 2D scatterplot representing authors. Authors are plotted according to the average of the year and number of elements

of their papers. We chose to eliminate the topic axis mapping for authors in order to facilitate easier viewing when both authors and articles are visible.

### Investigation Process

To investigate the five tasks, various visualization settings can be adjusted. The visualization lends itself to a top-to-bottom investigation of the data. Starting with all of the articles, users can dynamically query[1] for years and authors. The power of the visualization comes with the display options. Topic colors, the author projection, links between elements, and adjustable bubbles provide a set of powerful tools with which to investigate the data set.

Topic colors combined with bubbles allow an easy categorization of the filtered set. Adjusting the size of the bubbles is analogous to adjusting the focus on a camera. Large bubbles lend themselves to a broad view of the data, while small bubbles allow for a finer grain of evaluation, a technique similar to volume rendering for aggregate data [3].

The author projection provides a way of investigating the latter tasks of the InfoVis contest. The elements within this projection can also have bubbles displayed around them, as well as be linked to co-authors or the complete connected set of co-authors.

The article display provides a complex set of display options to maintain flexibility in investigating the data. Visible links can be limited to within the filtered set, outside the filtered set. Users can also select to display references from the selected set, or to the selected set. The combination of these features allowed us to find erroneous data in the dataset.

### Weaknesses

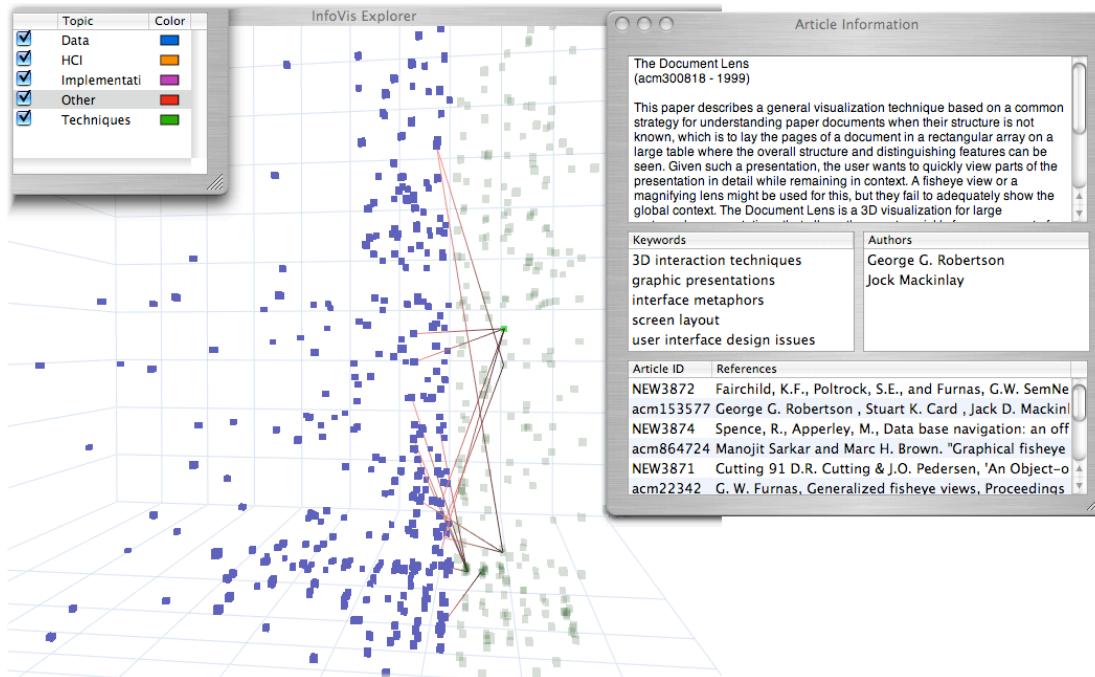


Figure 3: Filtered articles in blue reveal references (in red) to articles in the future. Selecting one of the referenced articles reveals an error in the dataset: The Document Lens is labeled as being published in 1999, however many of the references to it are for a 1993 paper.

Though potentially powerful, the application is far from complete. Walking through the data is not as simple as it should be; one way to ease this would be to auto-center on selected elements, walk the dataset using the Article Information window, and show a trace of references through paper s. We could also build a minimal spanning tree using the references and provide an easy navigational tool that way as well. The number of visualization options suggests some form of automated configuration for common tasks. Though much could be improved, current work is a solid foundation for a powerful information visualization tool.

### ACKNOWLEDGMENTS

We thank Prof. John Stasko and Georgia Tech for providing the opportunity to work on this project along with resources, an education, and advice.

We also thank the sponsors of the InfoVis 2004 contest, and the many individuals who contributed to the creation of the dataset.

### REFERENCES:

1. B. Shneiderman, "Dynamic Queries for Visual Information Seeking", IEEE Software, Vol. 11, No. 6, pp. 70-77, 1994.
2. C. Ahlberg, B. Shneiderman (1994), Visual Information Seeking with the FilmFinder, in Proceedings of CHI'94: Human Factors in Computing Systems, page 433.
3. Li Yang, n23tool: a tool for exploring large relational datasets through 3D dynamic projections, Proceedings of the ninth international conference on Information and knowledge management, p.322-327, November 06-11, 2000, McLean, Virginia, United States