

# Infovis 2004 Contest



Jean-Daniel Fekete, INRIA Futurs/LRI

Georges Grinstein, U Mass Lowell

Catherine Plaisant, University of Maryland

# Contest Goals

- Encourage the development of **benchmarks**
  - dataset, list of tasks, and non-trivial discoveries
- Push forefront of Infovis **research**
  - making available difficult problems
- Establish a forum to promote **evaluation**
- Interesting and **exciting event** at the conference

# 2004 Topic

## the history of Infovis

- IEEE Infovis is 10 years old today!
- Dataset: 614 Infovis papers and their 8507 references  
(from IEEE and ACM digital libraries)
- Tasks
  - Characterize the **research areas** and their evolution.
  - Where does a particular **author fit within the research areas**
  - What are the **relationships** between two or more or all authors?
- 5 months to submit (February to June)

# Preparing the data set

- 1000+ hours
- Several passes at cleaning
  - Often a hand process
- From imperfect to less imperfect
  - References unprocessed by digital library, or wrong
  - Multiple names for same authors (1161 > 1036)
  - Missing data (e.g. abstracts, keywords, all of 2003)
- Lots of help

# Thank you!

- ACM and IEEE and in particular Mark Mandelbaum & Bernard Rous
- Caroline Appert (*University of Paris South*)
- Urska Cvek, Alexander Gee, Howie Goodell, Vivek Gupta, Christine Lawrence, Hongli Li, Mary Beth Smrtic, Min Yu and Jianping Zhou (*University of Massachusetts Lowell*)
- Kevin Stamper, Tzu-Wei Hsu, Dave McColgin, Chris Plaue, Jason Day, Bob Amar, Justin Godfrey & Lee Inman Farabaugh (*Georgia Tech*)
- Niklas Elmqvist (*Chalmers, Sweden*)
- Jeff Klingner (*Stanford*)
- Jung-Rung Han, Chia-Ning Chiang and Tamara Munzner (*UBC*)
- Maylis Delest (*Université de Bordeaux*)
- Paolo Buono (*University of Bari, Italy*)
- Shabnam Tafreshi (*U. of Maryland*) for website
- and many others who made small corrections...

# Submission materials

- Two-page summary (see compendium p.47)
- Structured web form:
  - Process
  - Screen shots
  - Insights identified ← most important
- Video
- Materials now available in [Infovis Benchmark Repository](http://www.cs.umd.edu/hcil/InfovisRepository)  
[www.cs.umd.edu/hcil/InfovisRepository](http://www.cs.umd.edu/hcil/InfovisRepository)

[\(example\)](#)

Contents

- Home
- Benchmarks
- Contributing
- Links

## What is the Information Visualization Benchmarks Repository?

The Information Visualization Repository contains resources to improve the evaluation of information visualization techniques and systems.

[Benchmark datasets and tasks](#) are being made available, as well as results submitted by teams demonstrating their visualization tools at work using the benchmark data. The first benchmark datasets and tasks were created for the [InfoVis2003 Contest](#). The contest continued with the [InfoVis2004 contest](#) and we will continue populating the site with new benchmarks and their results.

You can help improving this site by suggesting interesting datasets to visualize or by trying your system on datasets and tasks contained in the repository and returning your results to [infovis-repository@cs.umd.edu](mailto:infovis-repository@cs.umd.edu)

### News

October 6, 2004:

[Results of Infovis2004 contest](#) added in repository

October 19, 2003:

1st announcement of the Repository during Infovis2003

### Credits

The Repository is maintained by Jean-Daniel Fekete, Catherine Plaisant and Shabnam Tafreshi.

### Sponsors



The Hive Group

[www.cs.umd.edu/hcil/InfovisRepository](http://www.cs.umd.edu/hcil/InfovisRepository)

or see page 47 of compendium

# Submissions

- 18 submissions
- 6 countries
- 7 student teams
  
- Comparing with 2003
  - More submissions (only 8 in 2003)
  - Most participants did provide insights about data
  - Still very hard to compare and select



# Awards

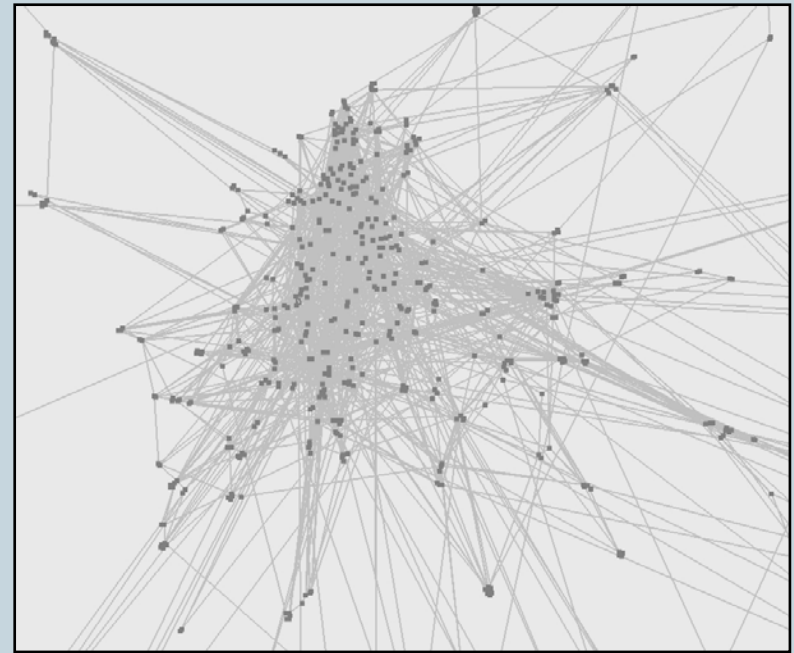
- Three “1<sup>st</sup> Place” + One “Student 1<sup>st</sup> Place”
  - nVidia card [Quadro FX (NV40)]
  - Kitware VTK books
- Eight “2<sup>nd</sup> Place”
  - Xbox



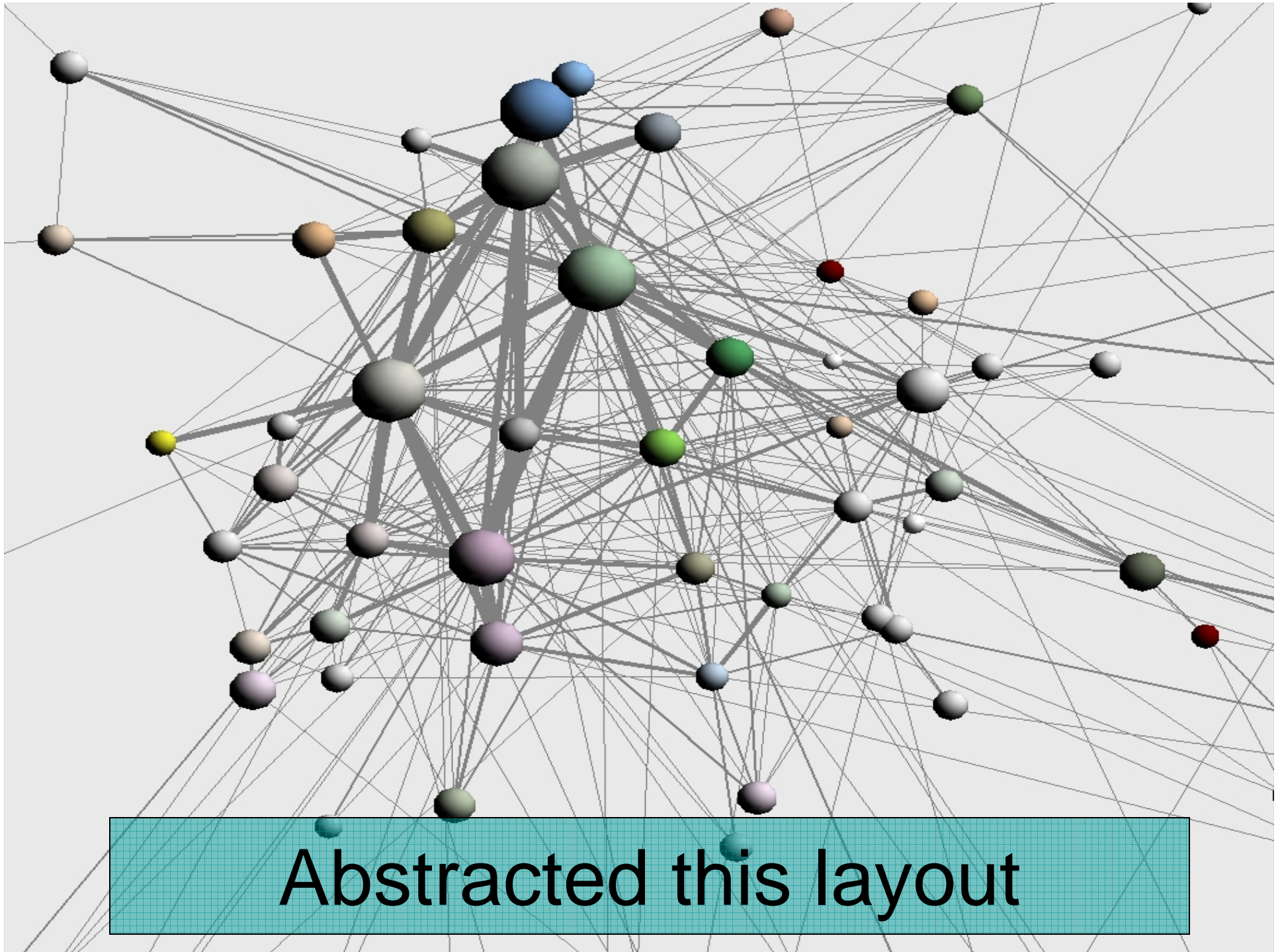
Here the 3 First Place had 10 minutes each to present ...  
(I don't have the slides)

And then the 8 Second Place had 3 minutes each  
(slides follow!)

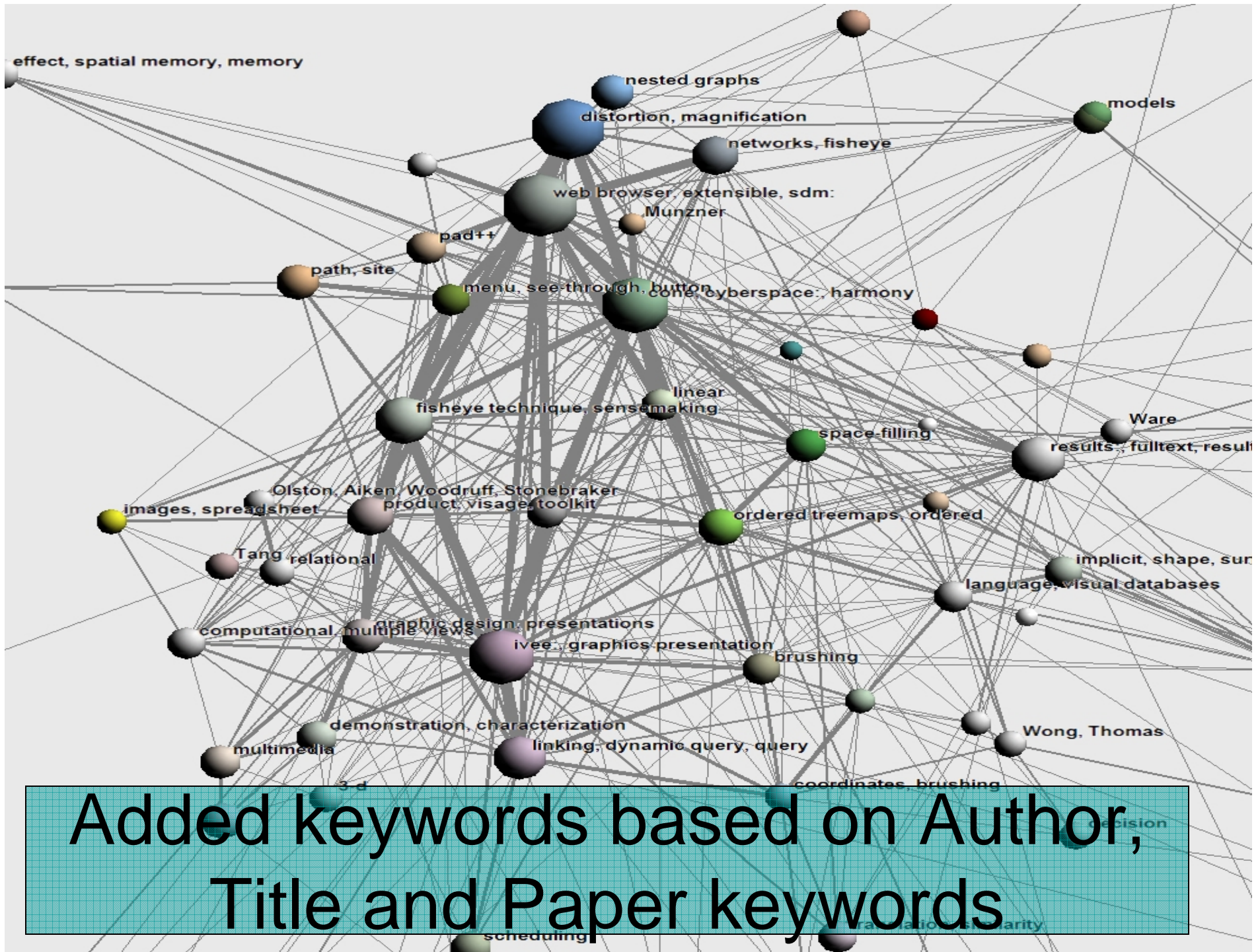
# Frank van Ham



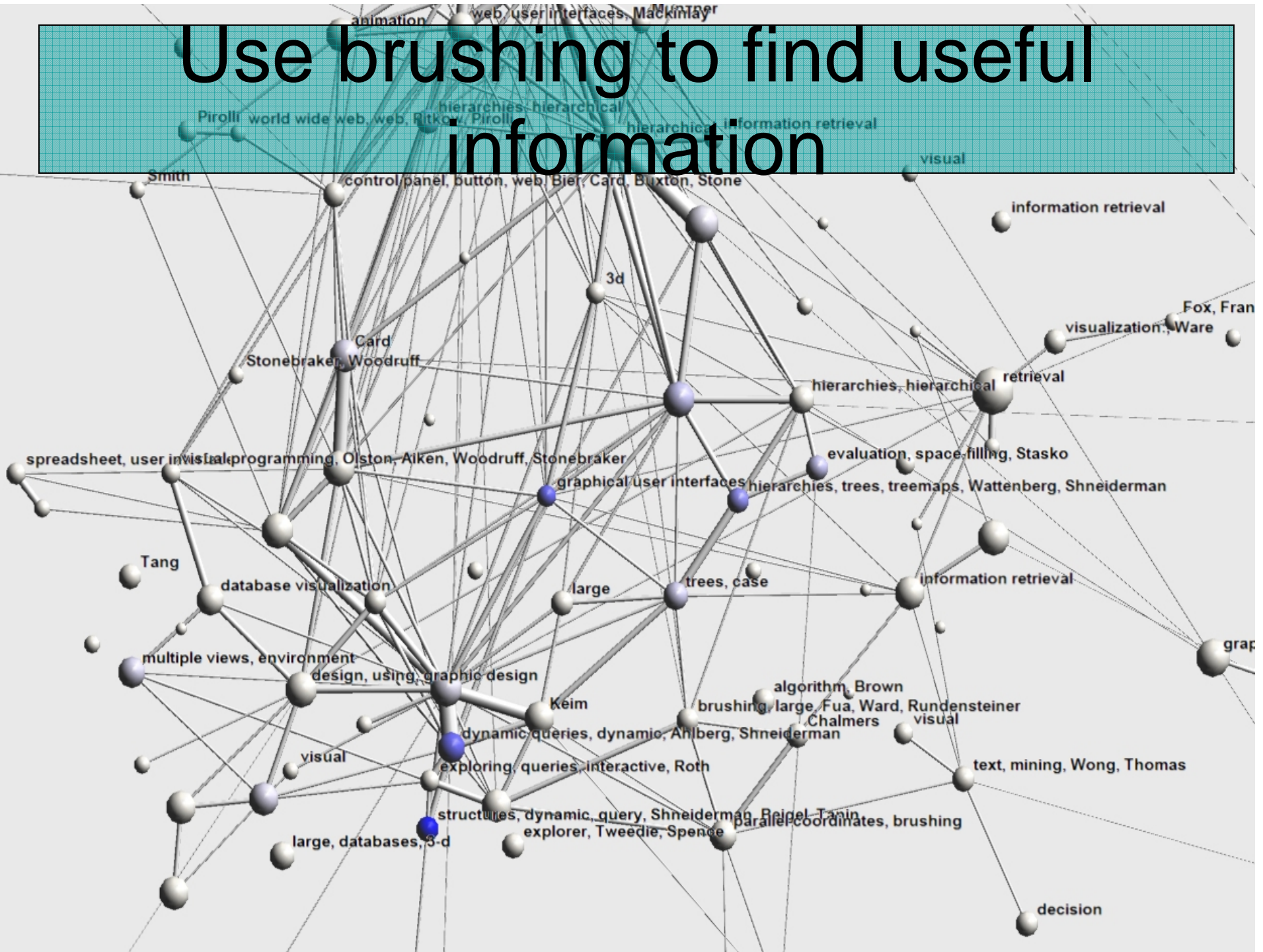
- Created FD layout of entire graph (messy)
- Pulled apart using modified force model



Abstracted this layout



# Use brushing to find useful information



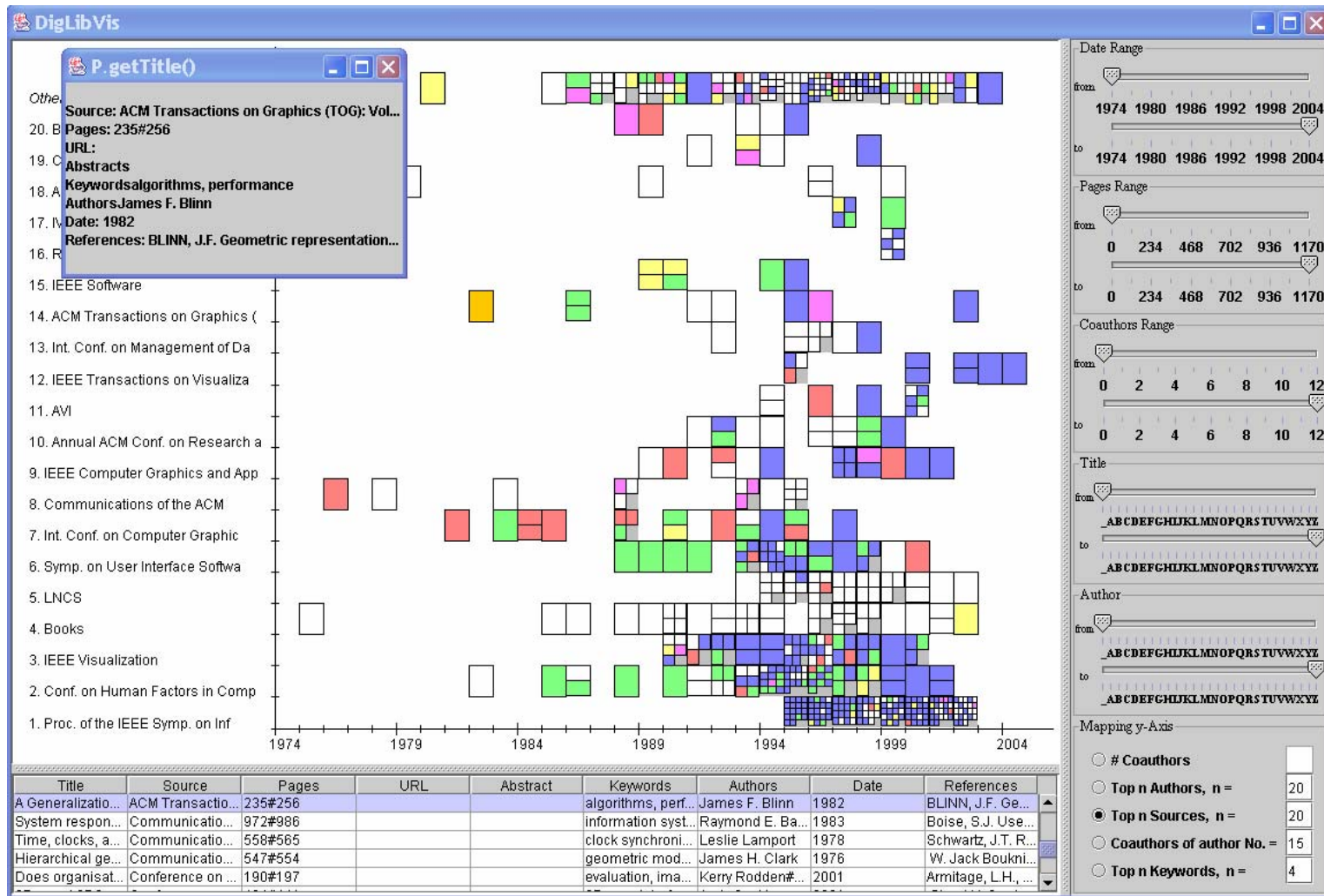
# InfoVis Contest: Exploring and Visualizing the History of InfoVis

**University of Konstanz**

*Daniel Keim, Christian Panse, Mike Sips,  
Joern Schneidewind, Helmut Barro*

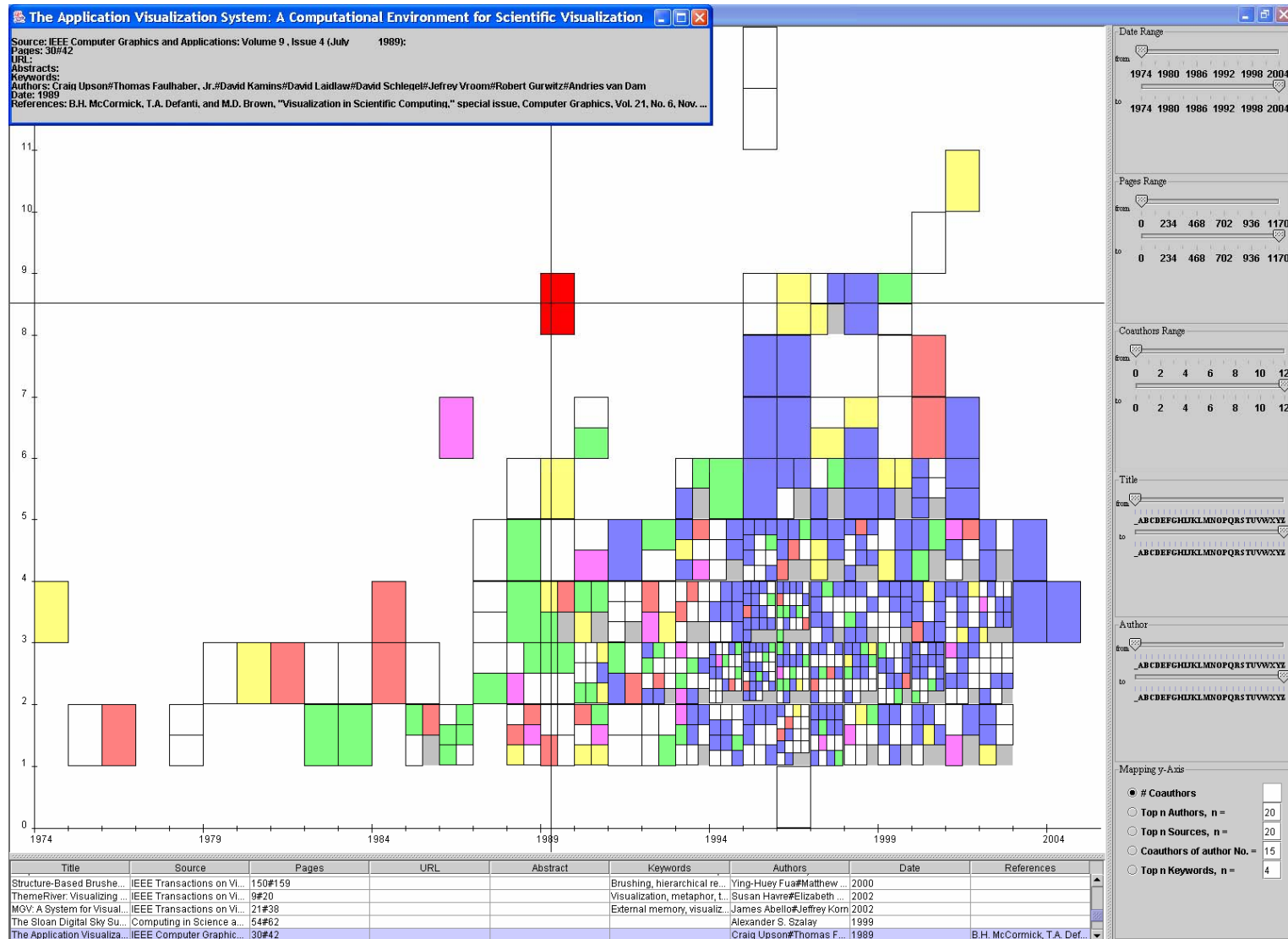
**Databases, Data Mining and Visualization Group**

# Overview first

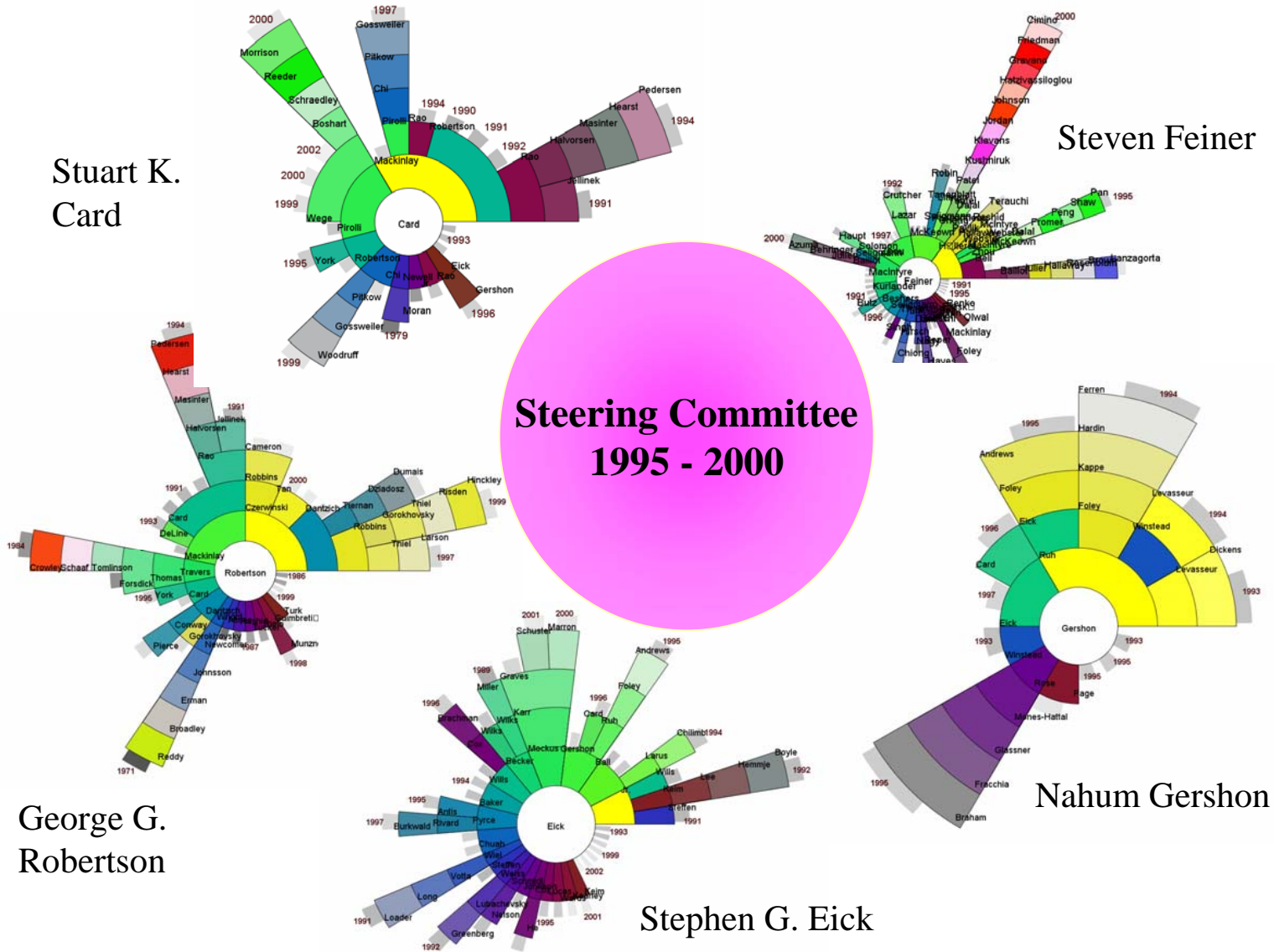




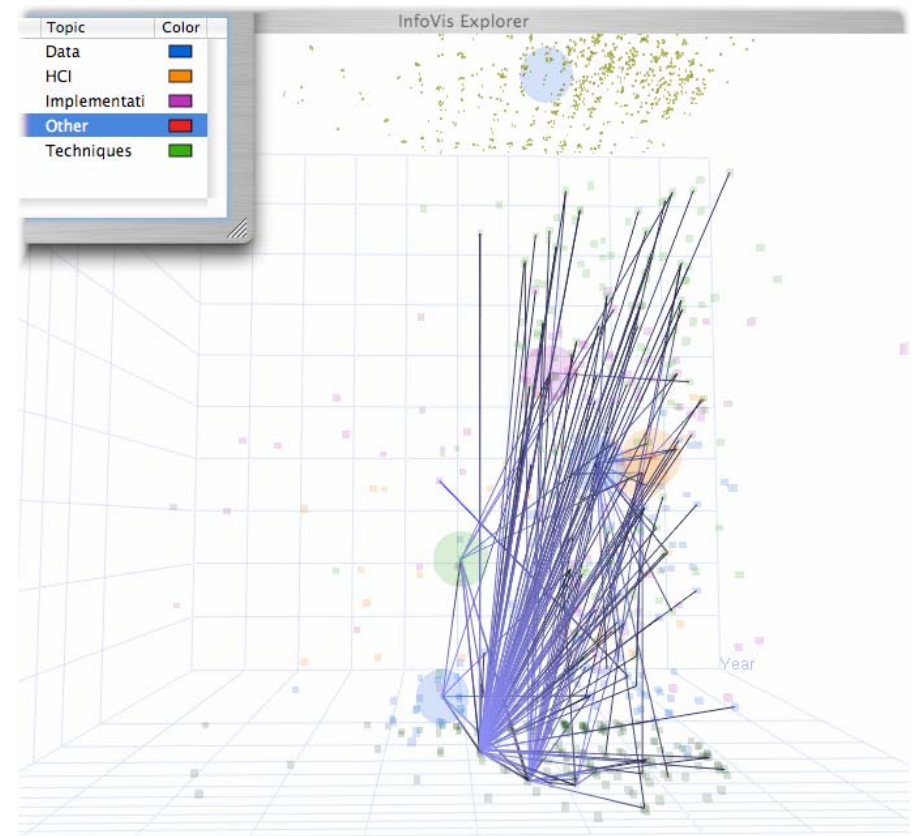
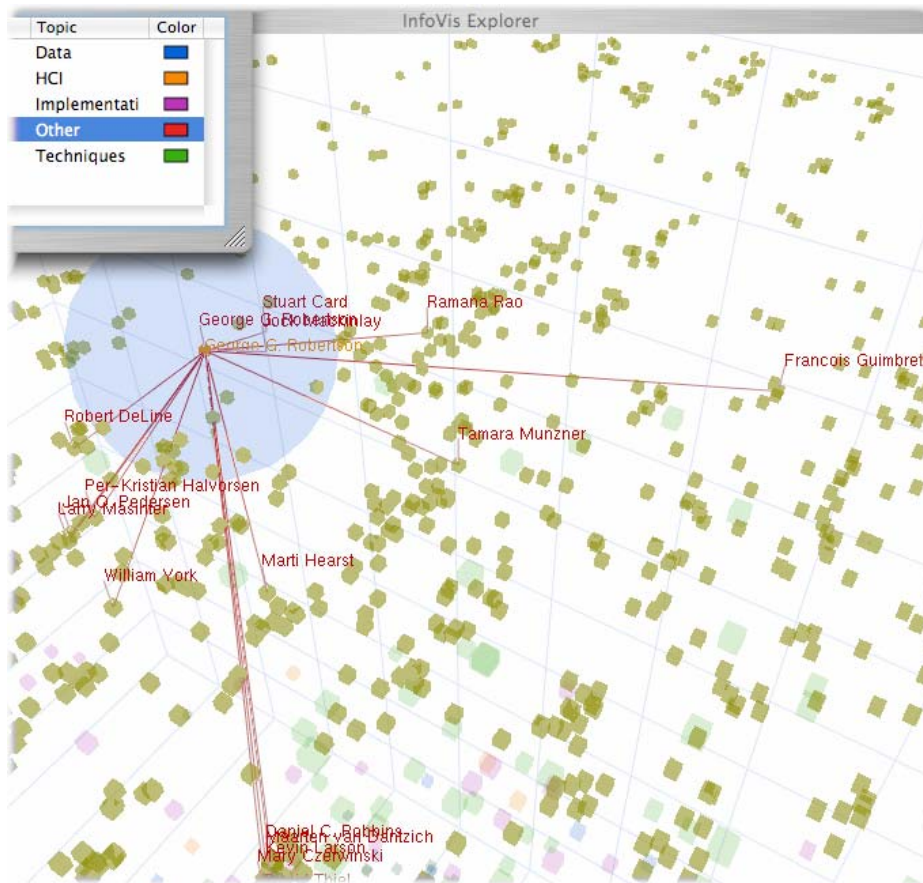
# Interactive Querying



# Details about InfoVis People



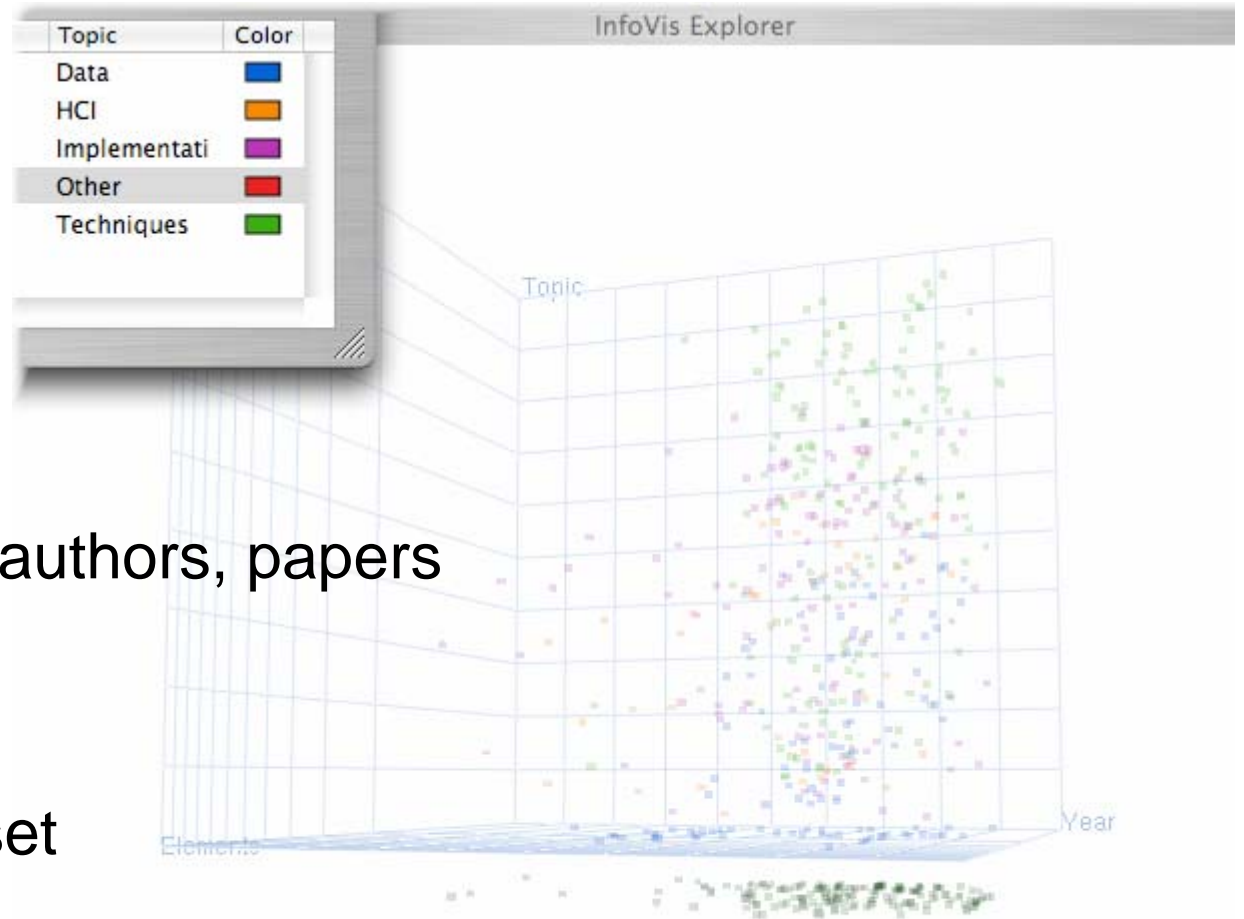
# InfoVis Explorer



Jaroslav Tyman, Grant P. Gruetzmacher, John Stasko  
Georgia Institute of Technology

# InfoVis Explorer

- 3 tools:
  - Import
  - Organize
  - Visualize
- Visualize:
  - Links between authors, papers
  - Groups
  - Topics
  - Errors in data set





# VisualLink: Associative Information Visualizer

Xia Lin

Jan Buzydlowski

Howard D. White

Drexel University

# Main Feature –1

## Balance of global mapping and localized mapping

- Global mapping for overviews
  - Emphasis on the power of abstraction and summarization
- Localized mapping for interaction
  - Emphasis on view selections and instant responses.

# Main Feature - 2

## Interactive Cross-Mapping

- Let the user select any mapping views
  - Search by author, map subjects ...
  - Search by subject, map documents ...

### VisualLink: *Associative Information Visualizer*

This visualizer generates instant maps that summarize various topics/authors/documents from the *The Proceedings of IEEE Symposium on Information Visualization* from 1995 to 2002.

Search:

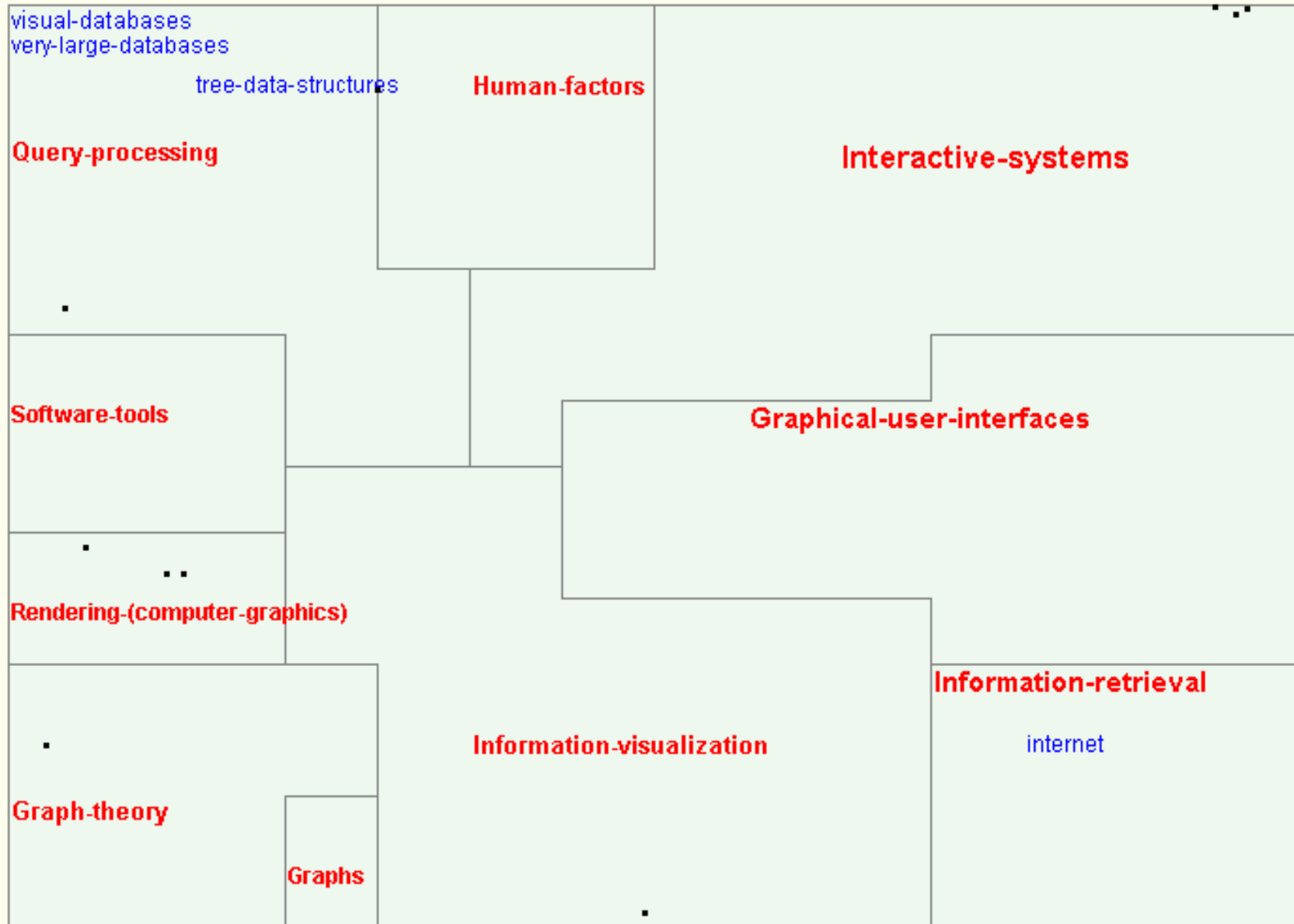
Instant Co-Occurrence Citation Map of **INFORM** **TRIEVAL**

# Main Feature -3

- Instant Understandability
  - Not just pretty pictures
  - Use Text labels
  - Utilize the power of graphics: distance, sizes, links, neighborhoods.



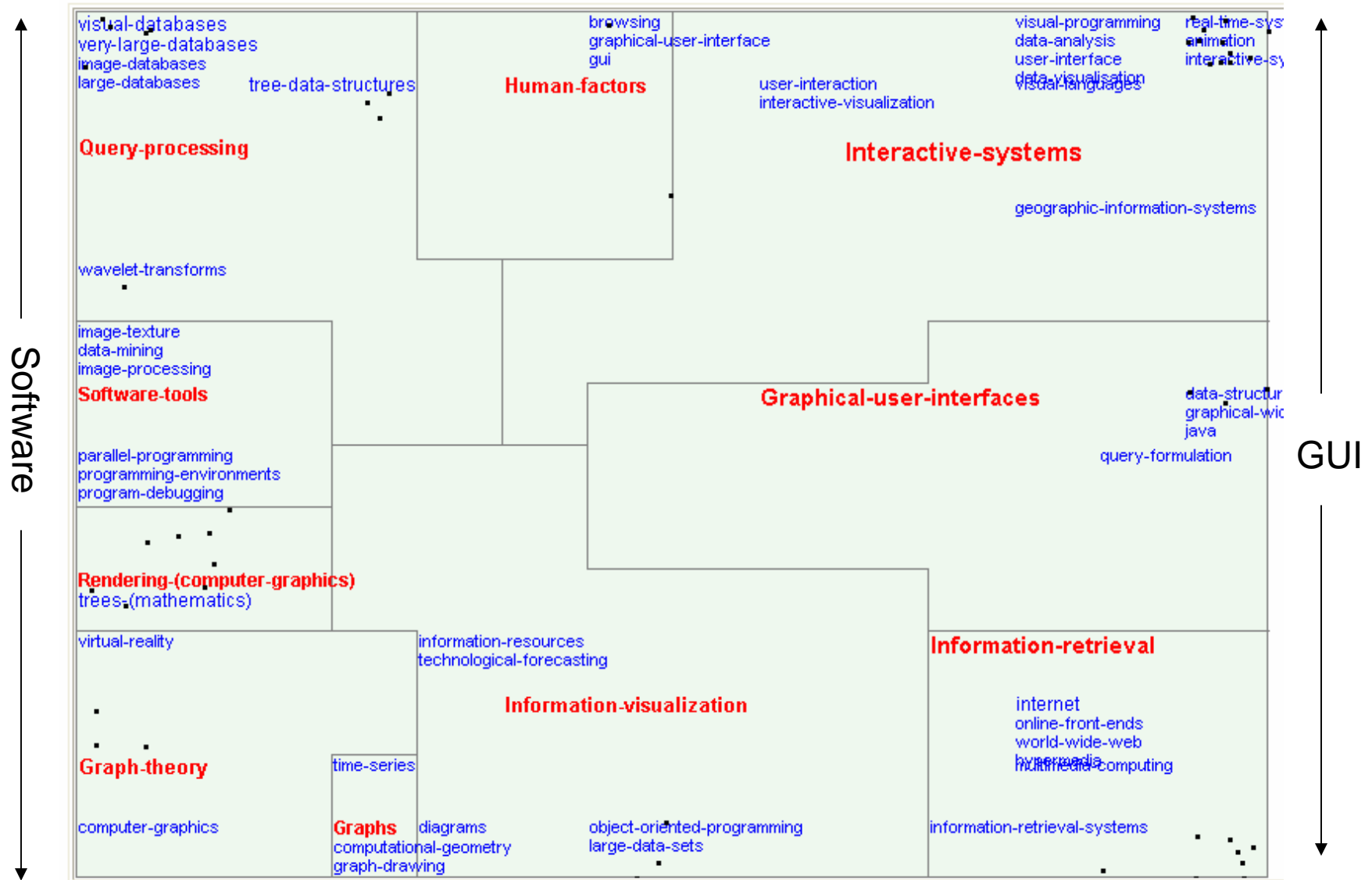
# Overview Map



TECHNOLOGY

HCI

INTERACTION

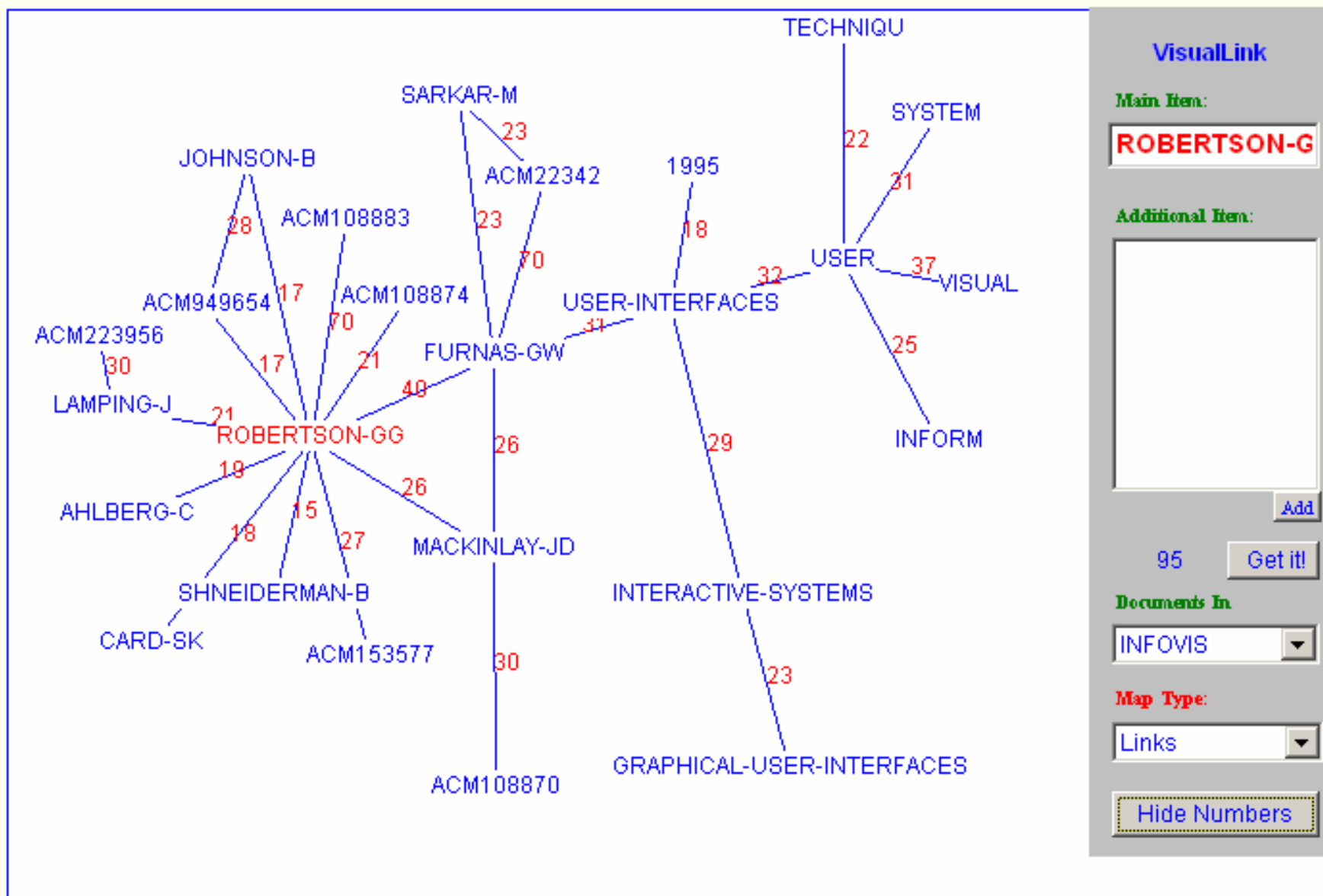


THEORY

Visualization

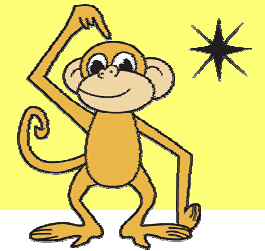
CONTENT

## Instant Co-Occurrence Map of ROBERTSON-GG



Monk\*Ellipse\*

Visualizing the History  
of Information Visualization



Lee Inman Farabaugh (lee@cc.gatech.edu)

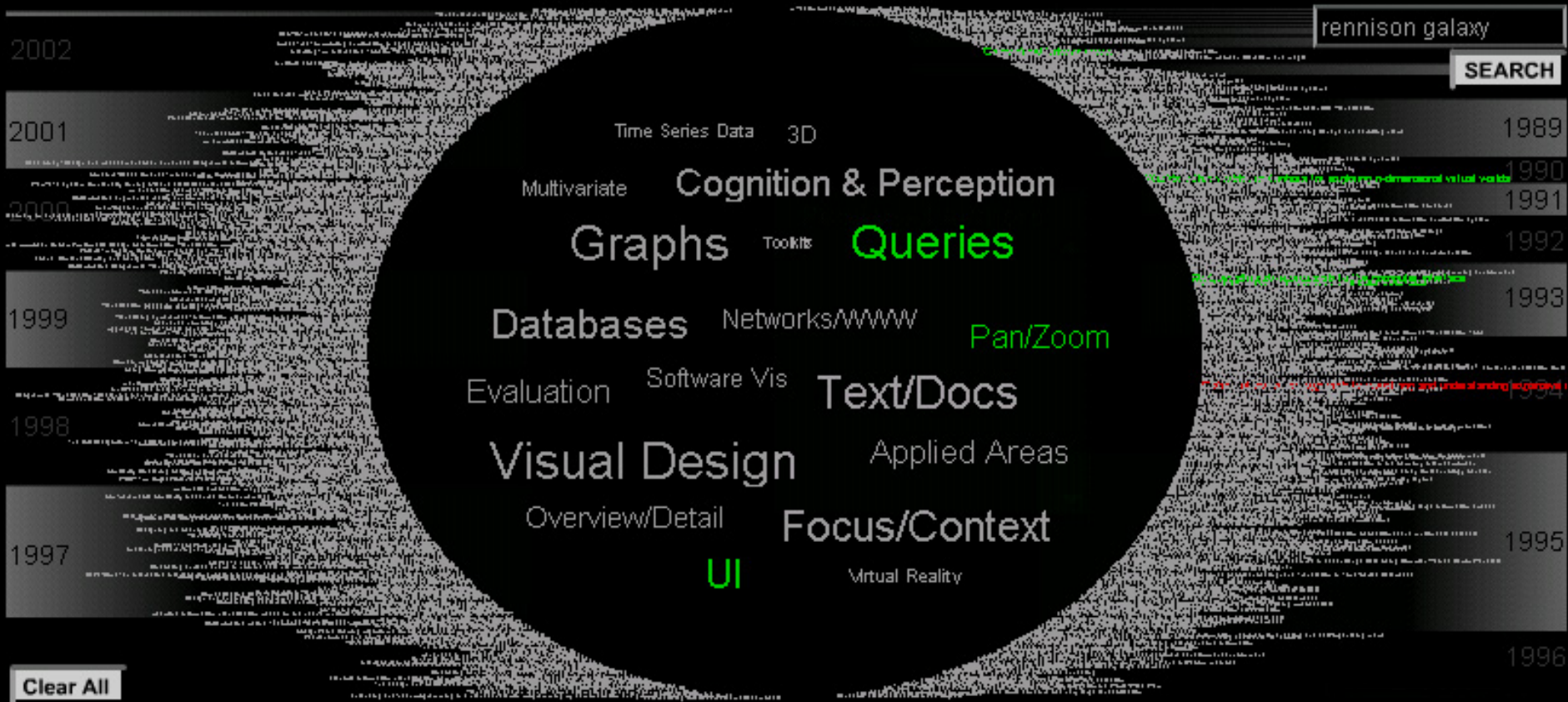
Tzu-Wei Hsu (thsu@cc)

Dave McColgin (davemail@cc)

Kevin Stamper (kstamper@cc)

M.S. in Human-Computer Interaction

Georgia\*Institute\*of\*Technology



Galaxy of news: an approach to visualizing and understanding expansive news landscapes

**Title:** Galaxy of news: an approach to visualizing and understanding expansive news landscapes

**Authors:** Earl Rennison

**Date:** 1994

**ACM Link:** [acm192429](https://doi.org/10.1145/192429)

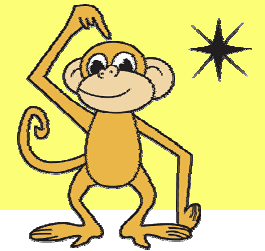
**Abstract:**

The Galaxy of News system embodies an approach to visualizing large quantities of independently authored pieces of information, in this case news stories. At the heart of this system is a powerful relationship construction engine that constructs an associative relation network to automatically build implicit links between related articles. To visualize these relationships, and hence the news information space, the Galaxy of News uses pyramidal structuring and visual presentation, semantic zooming and panning, animated visual cues that are dynamically constructed to illustrate relationships between articles, and fluid interaction in a three dimensional information space to browse and search through large databases of news articles. The result is a tool that allows people to quickly



# Monk\*Ellipse\*

Visualizing the History  
of Information Visualization



## Find Out More!

What:

Talk With Us

Watch Our Demo

See Our Poster

Play With MonkEllipse

Where:

Symposium Reception & Poster Session

Monday, October 11, 2004

7:00 – 9:00 PM

Georgia\* Institute \*of \*Technology

# Exploring InfoVis Publication History with Tulip

Maylis Delest, LaBRI Bordeaux

Tamara Munzner, UBC

David Auber, LaBRI Bordeaux

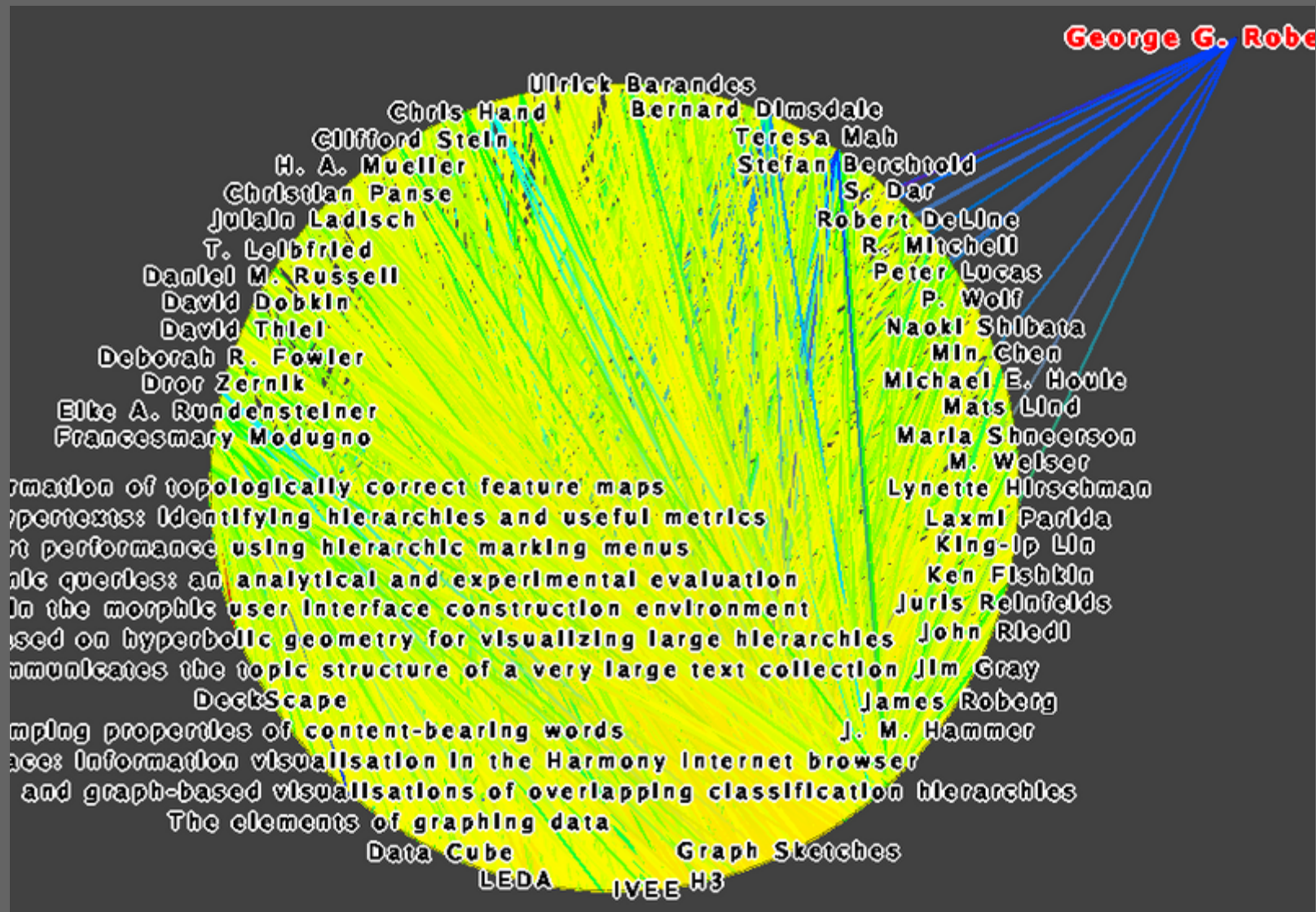
Jean-Philippe Domenger, LaBRI Bordeaux

# Tulip

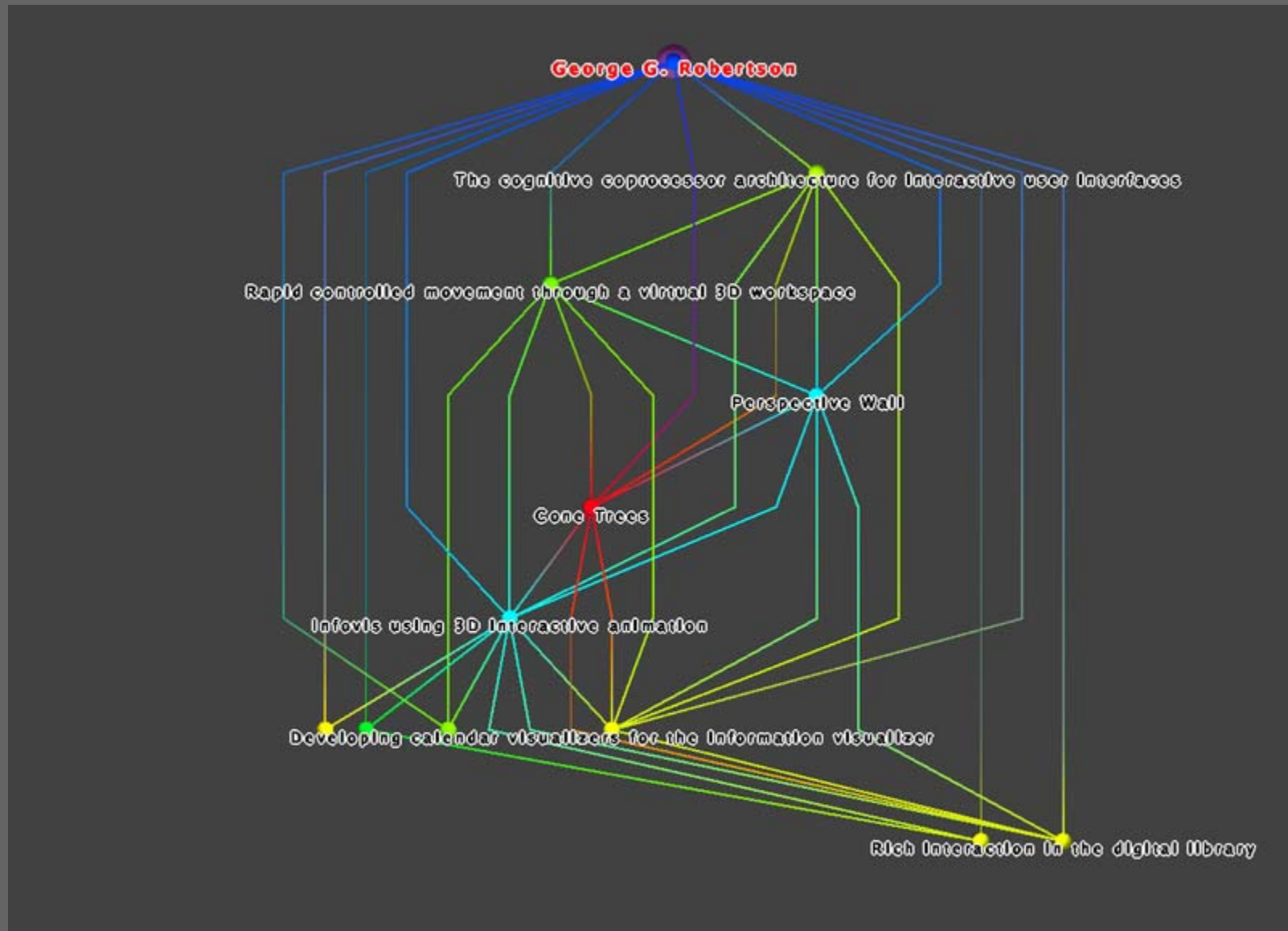
- graph drawing testbed
  - scalable
  - powerful
  - flexible
- functionality
  - clustering
  - layout
  - interaction
  - guaranteed frame rate rendering



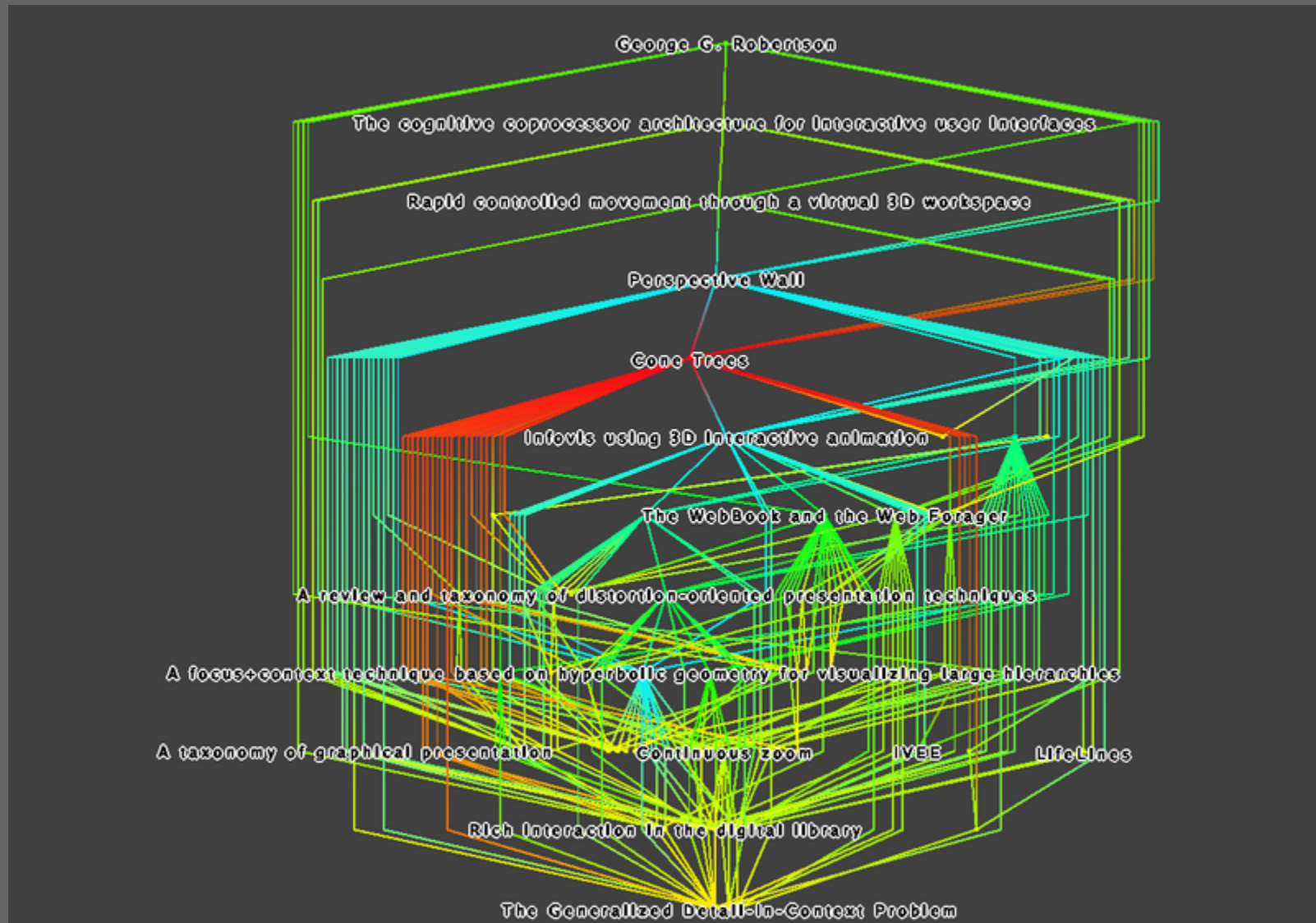
# Finding GGR: interactively move



# Reachable subgraph 1 hop: papers



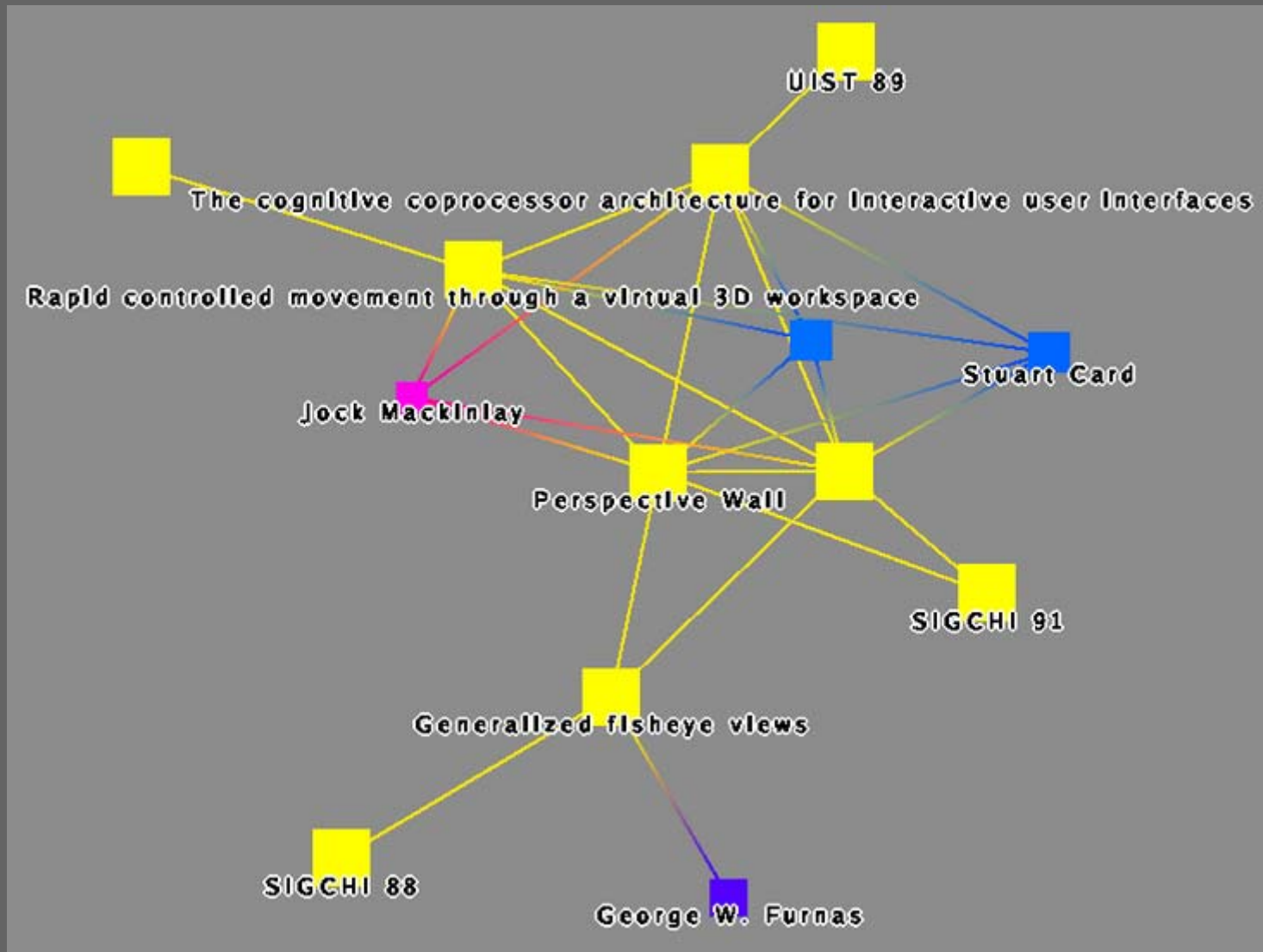
# Reachable subgraph 2 hops: citers



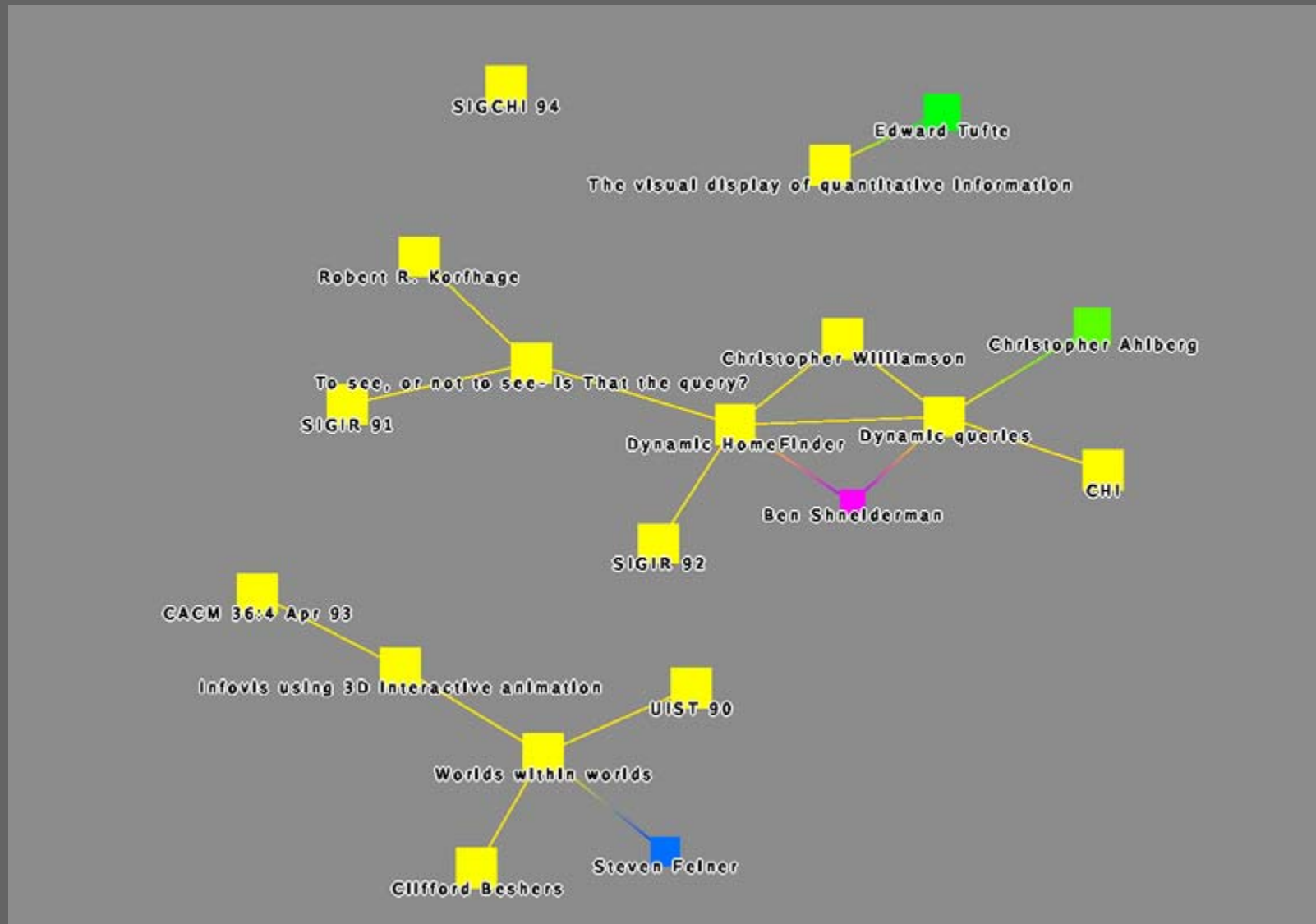
# Convolution Clustering

- visually determine best number of clusters
  - Strahler based graph clustering using convolution.  
David Auber, Maylis Delest, and Yves Chiricota.  
8th Int'l IEEE Conference on Information Visualisation,  
London, 2004
- clusters quite stable, show off core topics
- Strahler metric measures "centrality"

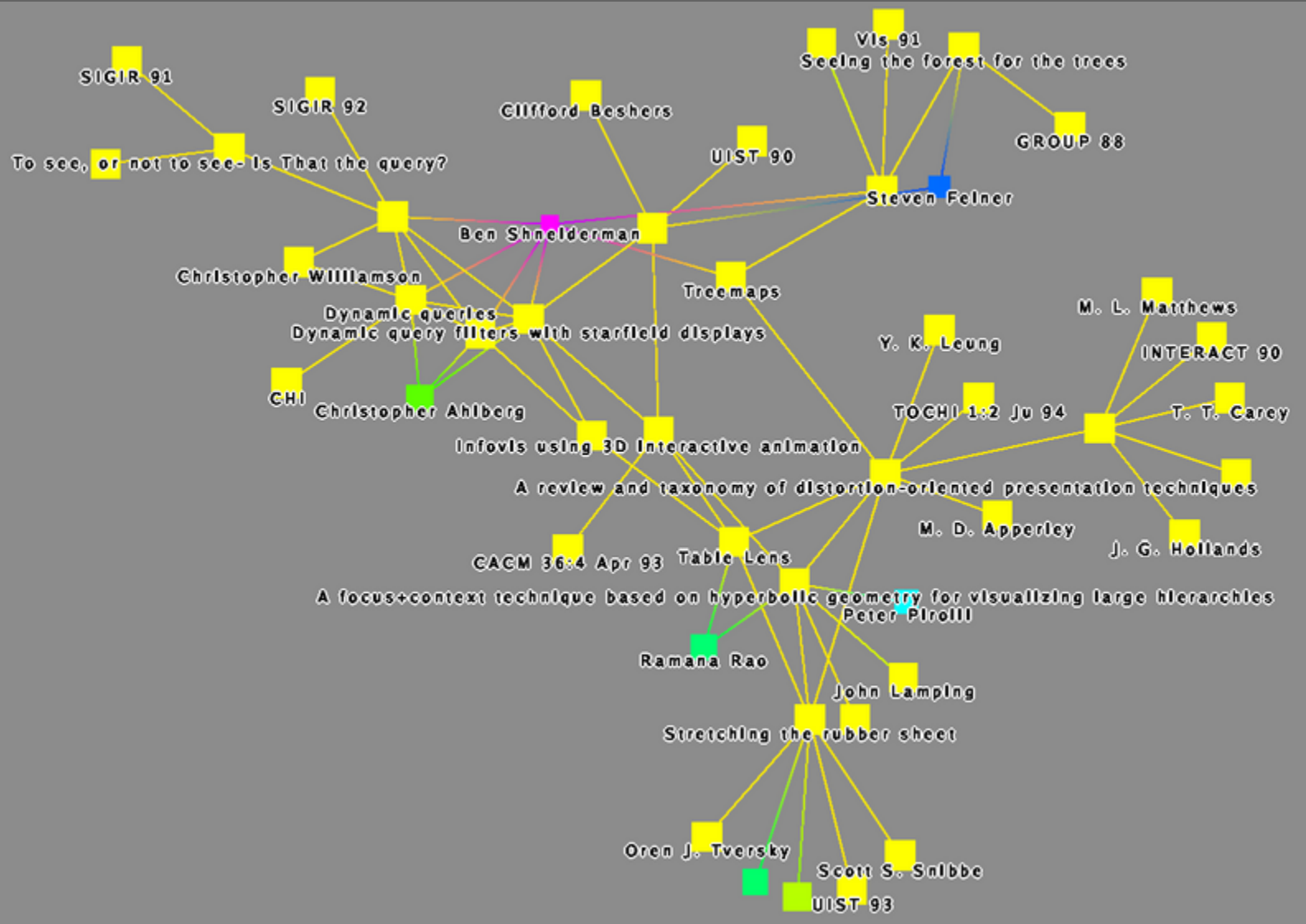
# Cluster 1: PARC/Furnas, F+C



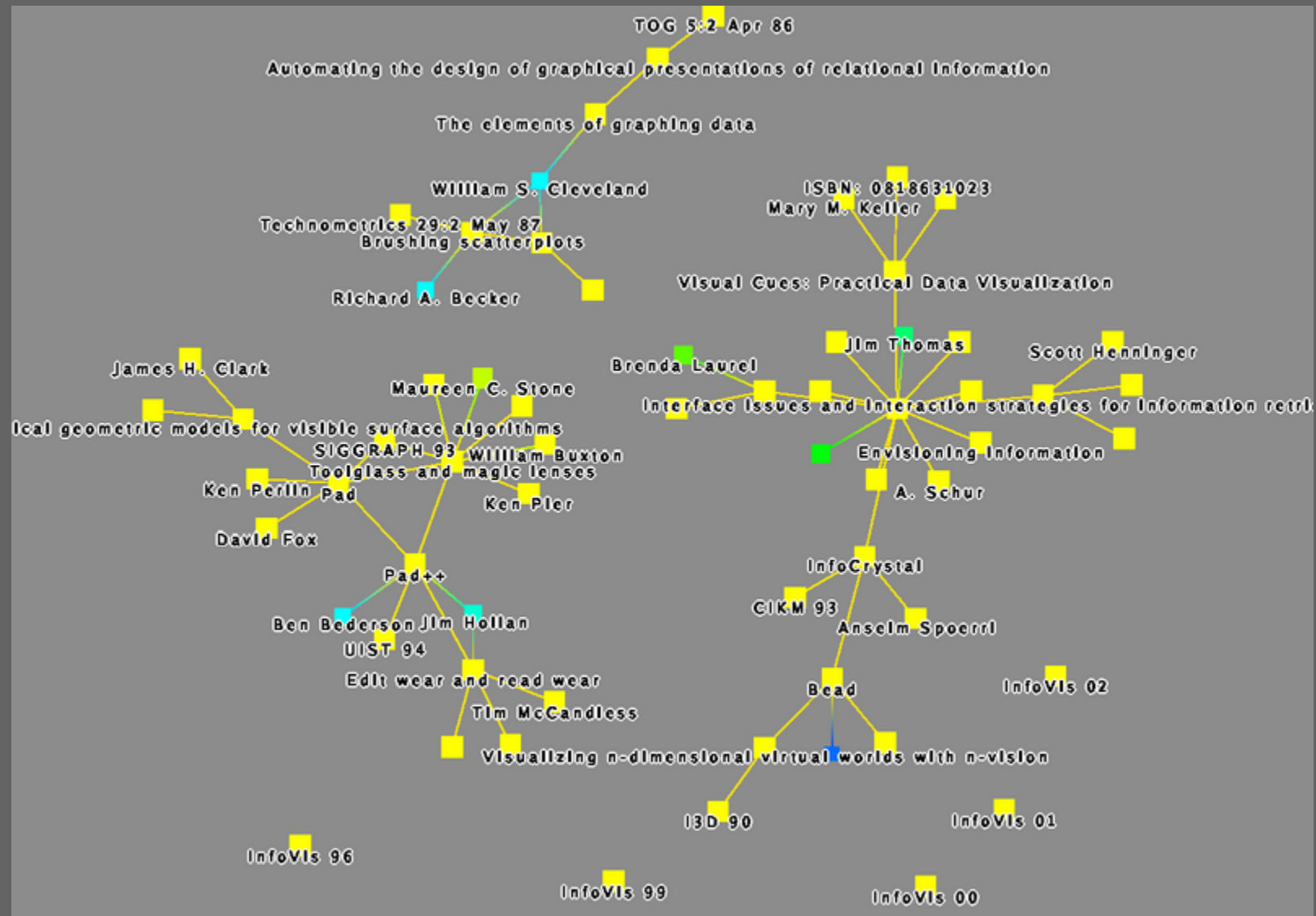
# Cluster 2: Dynamic Queries, Tufte



# Cluster 3: Focus+Context

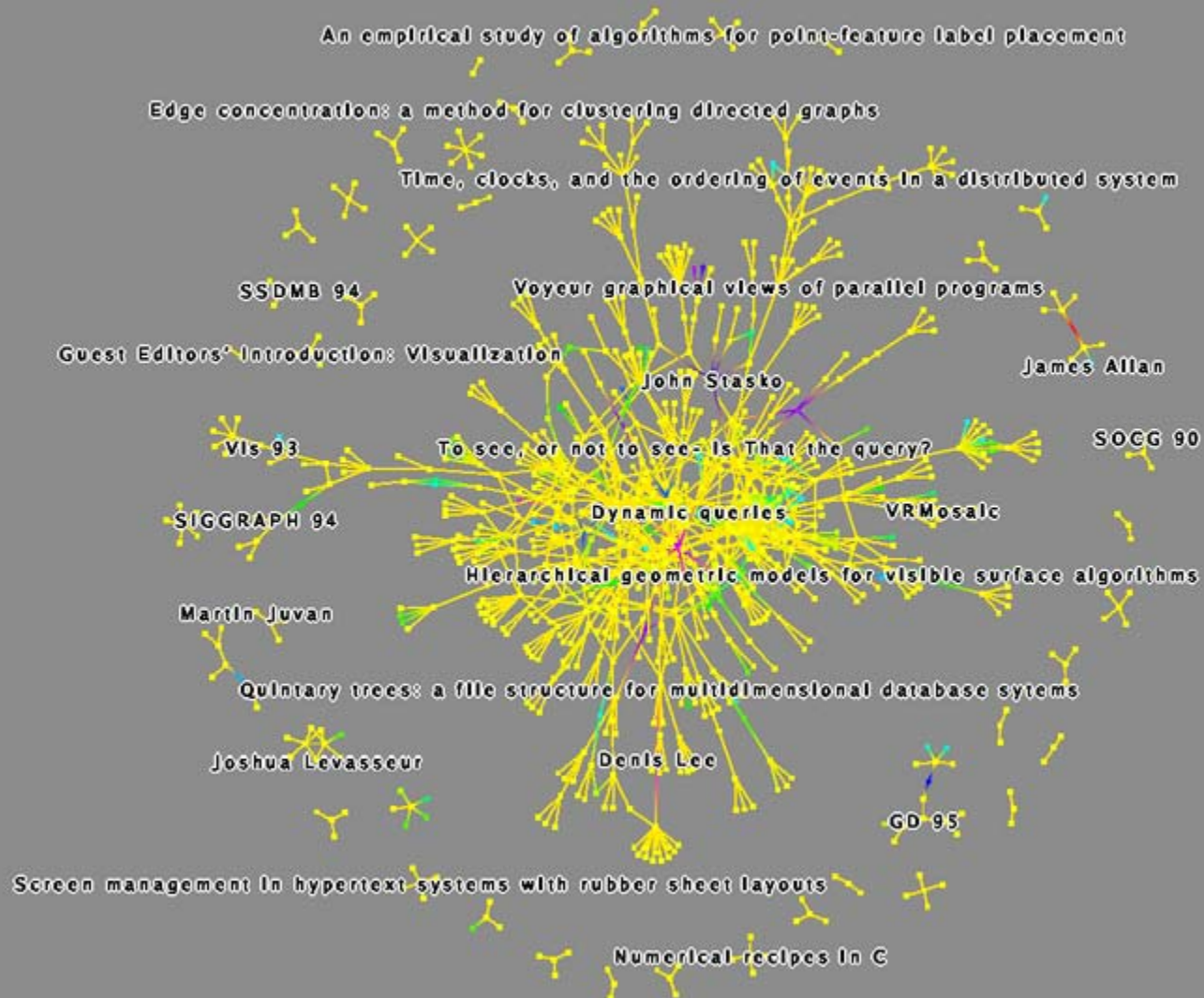


# Cluster 4: ZUIs, high dim, brushing

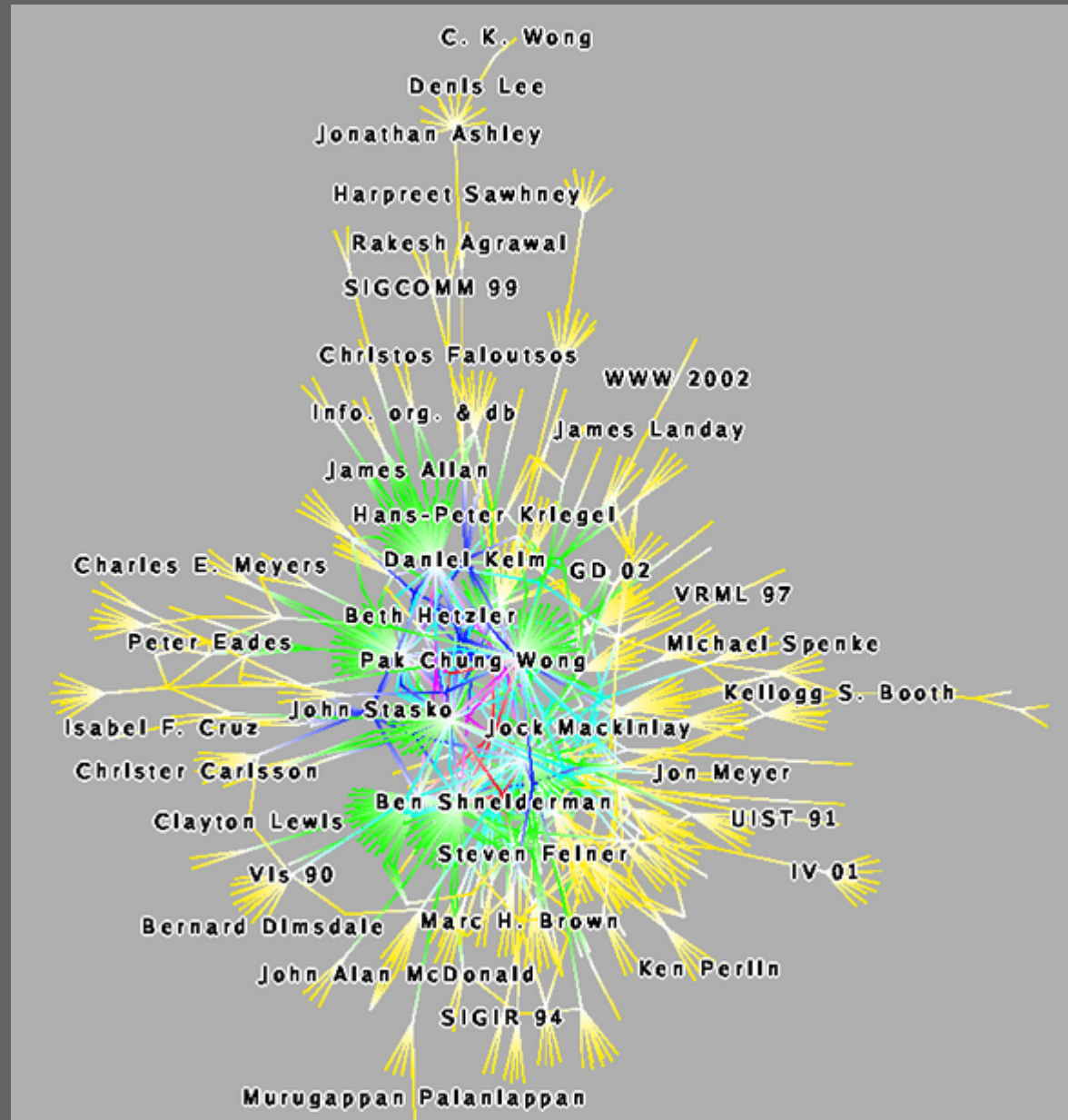




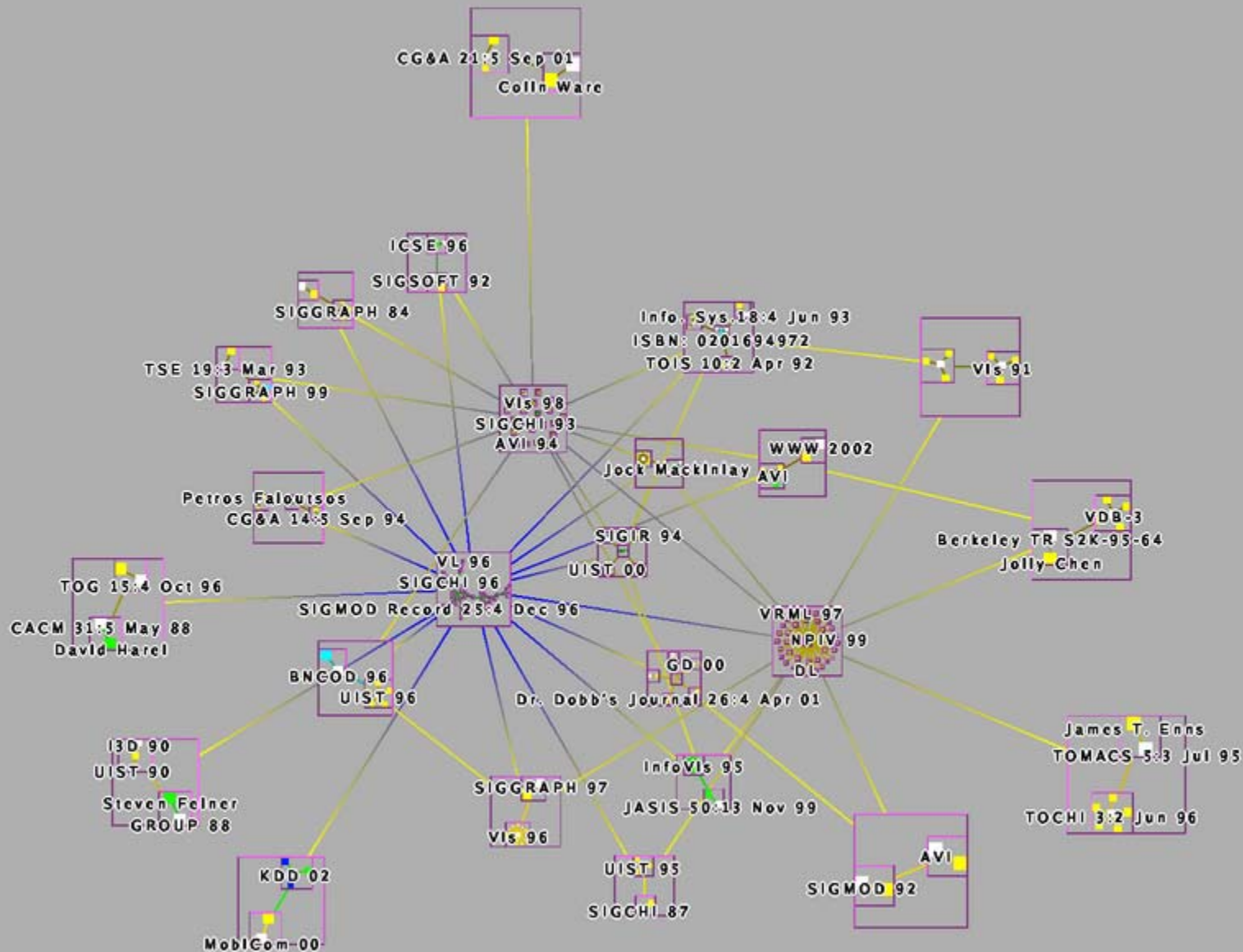
# Cluster 5: everything else



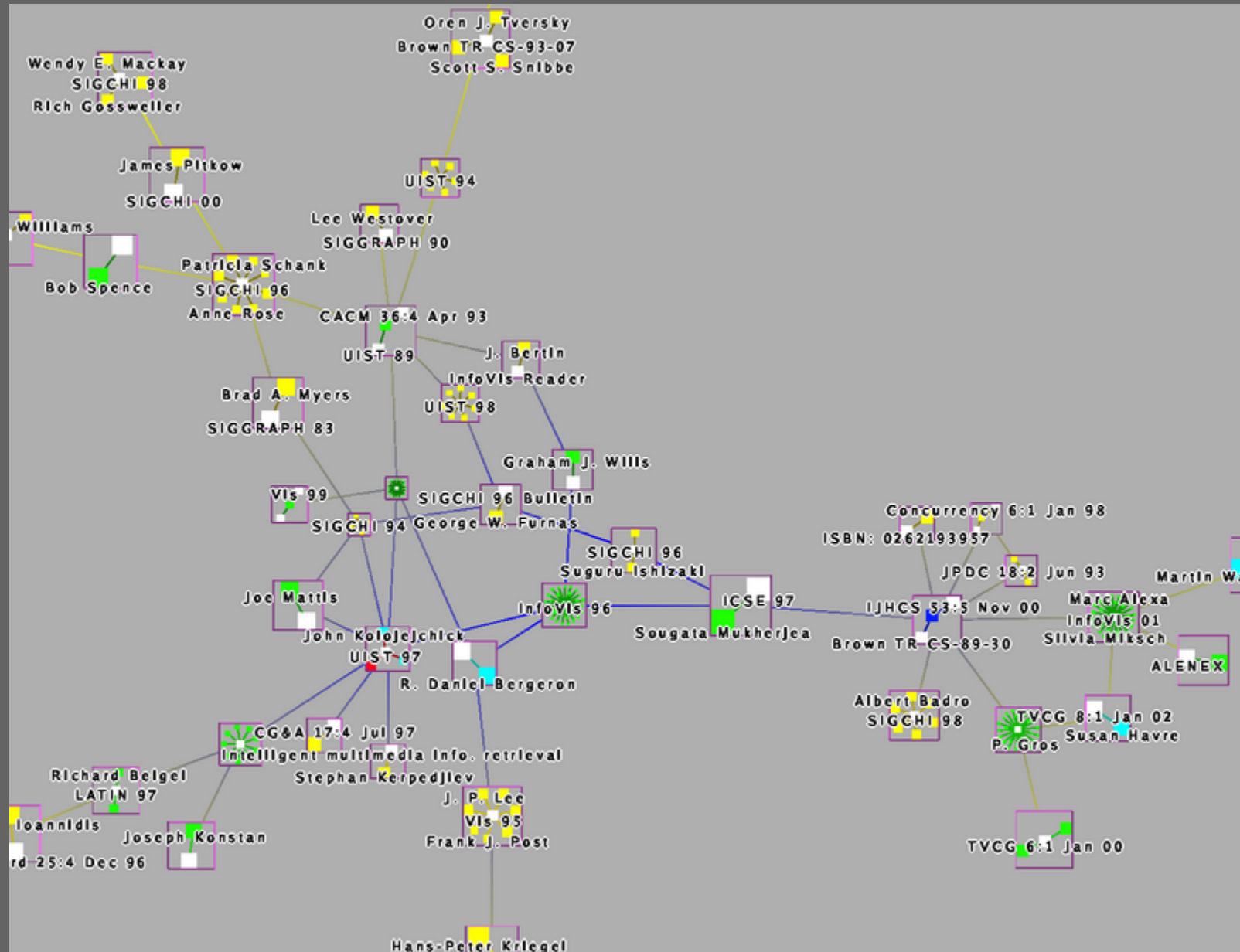
# Core Clutter with Auth-Conf



# Small-World Clustering



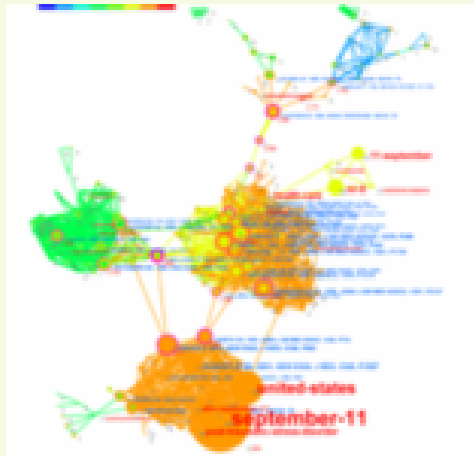
# Recursively Cluster



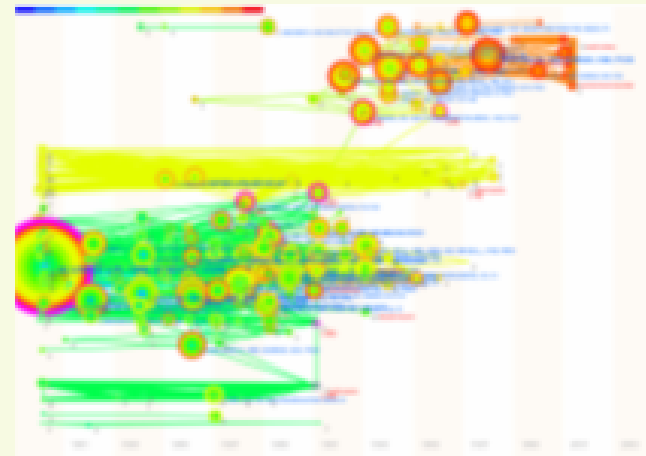
# Visualizing Trends and Turning Points in *CiteSpace*

Chaomei Chen  
Drexel University

Trends and Turning Points Visualized in CiteSpace [\*More\*](#)



Terrorism (1990-2003)  
300dpi (775K)



Mass extinction (1981-2003)  
300dpi (920K)

<http://www.pages.drexel.edu/~cc345/>

# CiteSpace

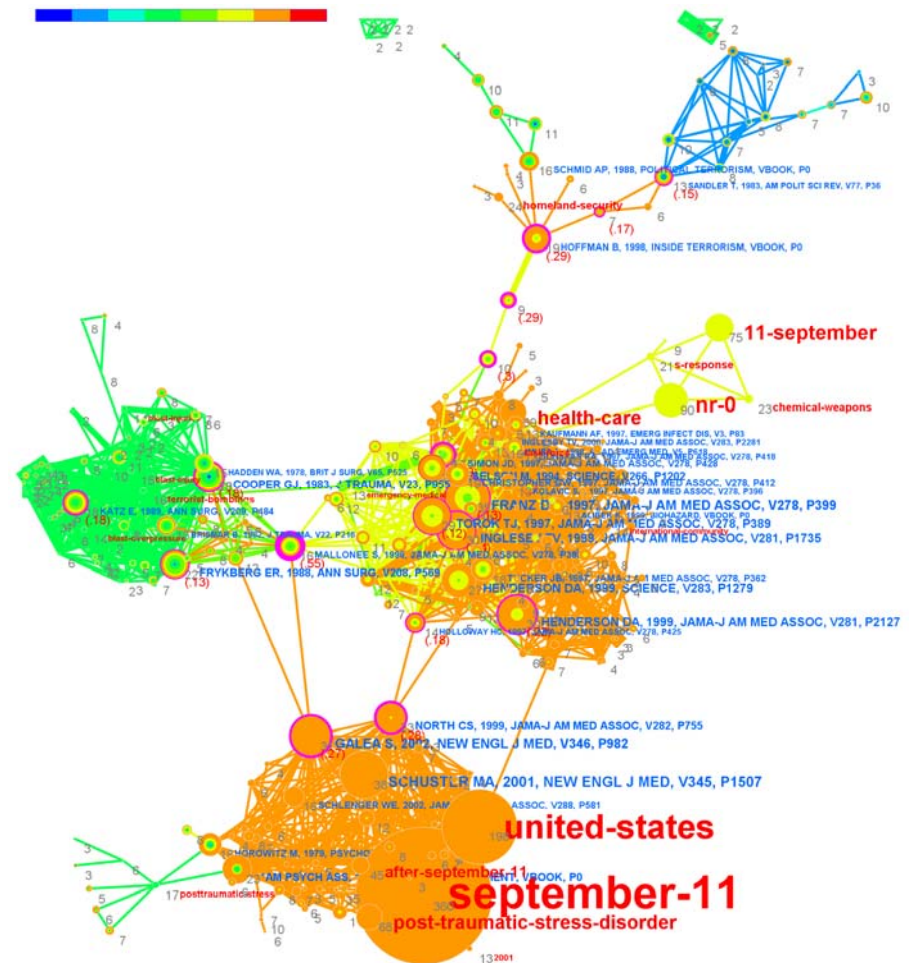
<http://cluster.cis.drexel.edu/~cchen/>

1. Written in Java
2. Freely available
3. Highly streamlined
4. Instant data fusion

Try it now!

CiteSpace 1.028

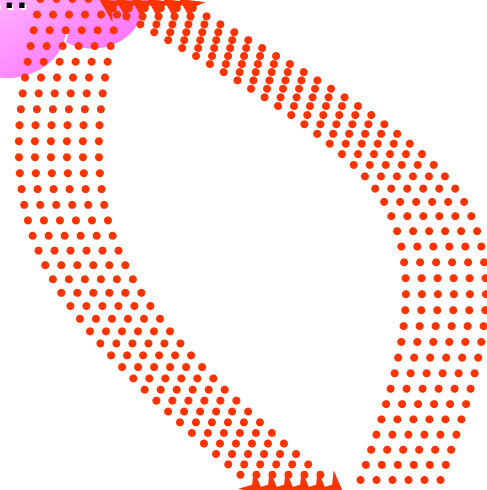
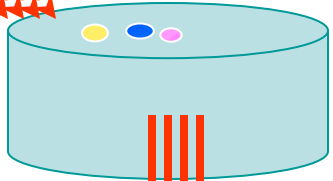
Chaomei Chen, Drexel University



Science Citation Index, Bibliographical data sources

Patents, Grant Abstracts, Digital Libraries

PubMed, Thesaurus, .....



CiteSpace v. 1.02B - (c) 2003-2004 Chaomei Chen

File Edit Data Analyze View Demo Help

PubMed Search: lung cancer [maj] 1990 2004 25 Search

WOS Directory: Unreads and Settings\Chaomei\Documents\WOS

WOS Demos: Mass Extraction, PubMed Demo, GDE, Reset

Year Range	Pubmed	WOS	Total	Used
1990-1990	3121015	120	0	0
1991-1991	3121015	3	0	0
1992-1992	3121015	133	0	0
1993-1993	3121015	486	1	0
1994-1994	3121015	291	1	0
1995-1995	3121015	672	2	1
1996-1996	3121015	720	9	7
1997-1997	3121015	287	1	0
1998-1998	3121015	154	0	0
1999-1999	3121015	586	0	0
2000-2000	3121015	452	4	0
2001-2001	3121015	222	0	0

Total JVM Mem (M): 7577 Used JVM Mem (%): 55

Time Slicing: Range begins with 1985, Range ends with 2003, Length of each slice: 1, 11, 21

Modeling:
 

- Normalization by cosine coefficients
- Normalization by Dice coefficients
- Normalization by Jaccard coefficients
- Tally within slices
- Tally across the whole range

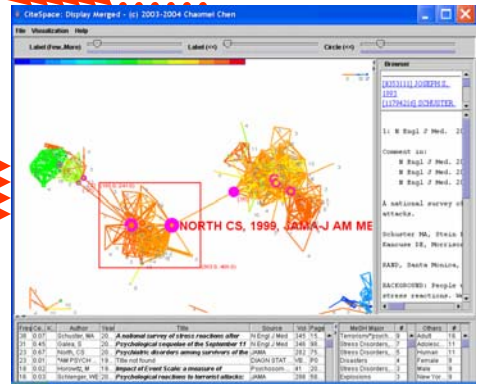
Thresholding (z, cc, ccv): z: 2, 2, 15, 3, 4, 15, 3, 2, 15; cc: 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0

Pruning:
 

- Pruning by Pathfinder
- Pruning by Minimum Spanning Tree
- Pruning by Weighted Mean Silhouette
- Pruning by Weighted Mean Silhouette and Modularity

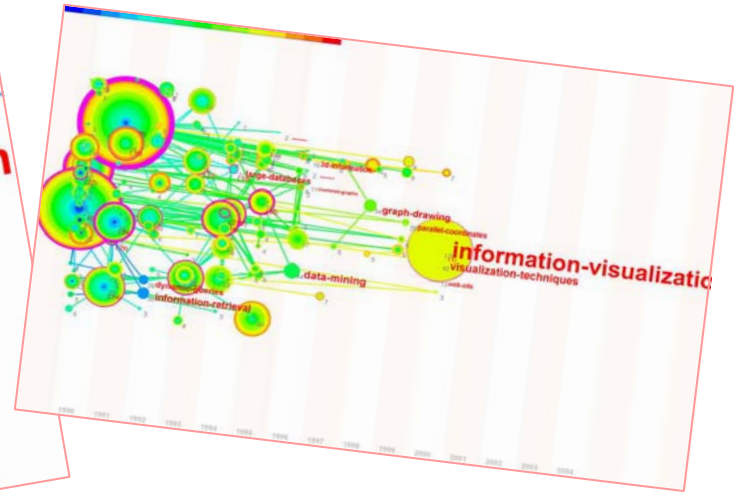
Layout: Time Zone, Label Key Terms Only, Layout Slice, Layout Merged

Reset scale: 1, 11, 21, 31, 41; Reset font size: 1, 11, 21, 31, 41

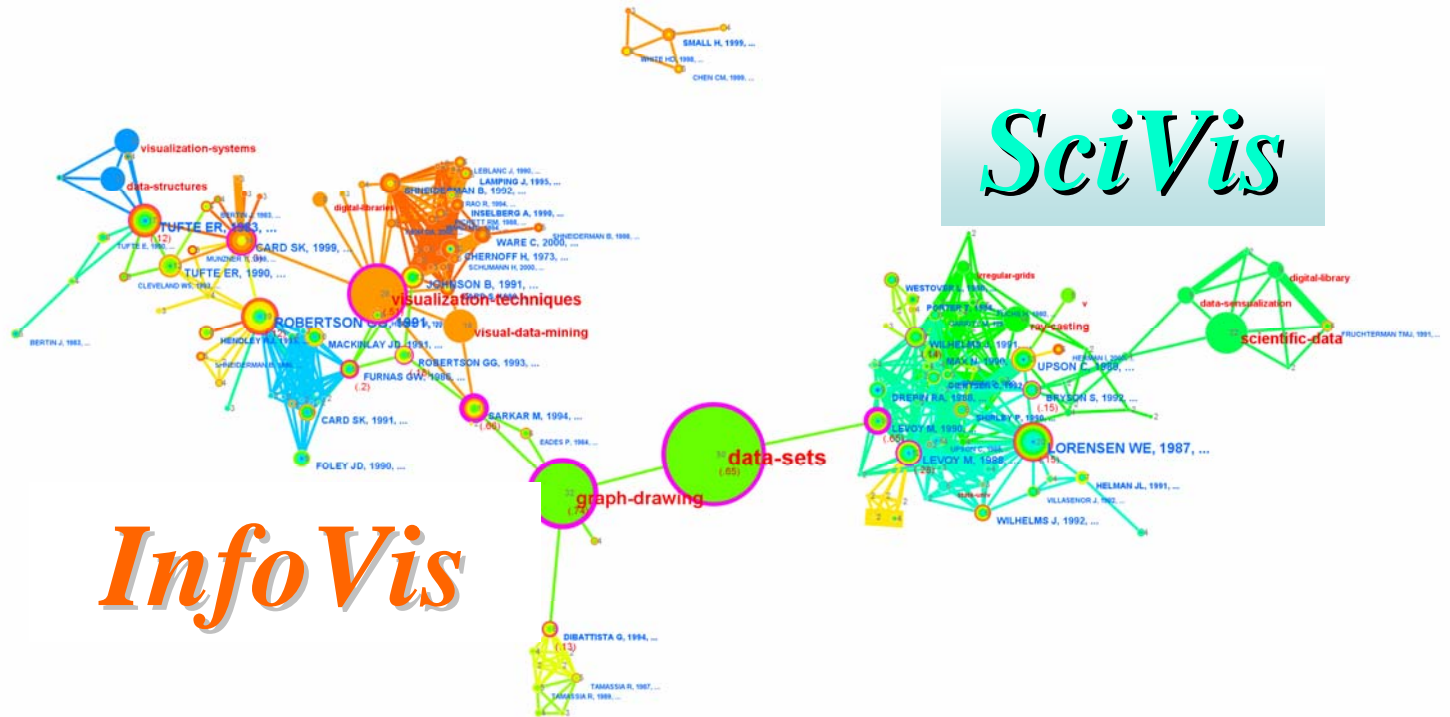


- Trend detection
- Turning point discovery
- Research front tracking

# InfoVis Contest 2004



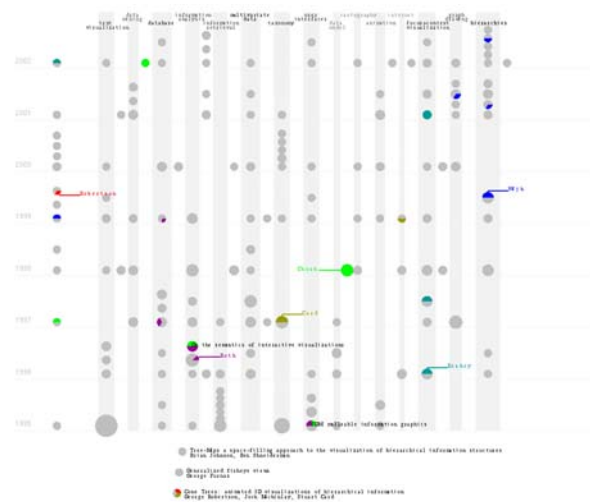
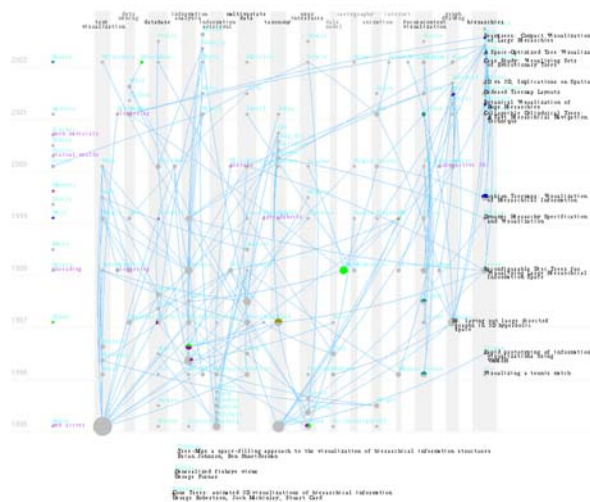
More ...



SciVis

InfoVis





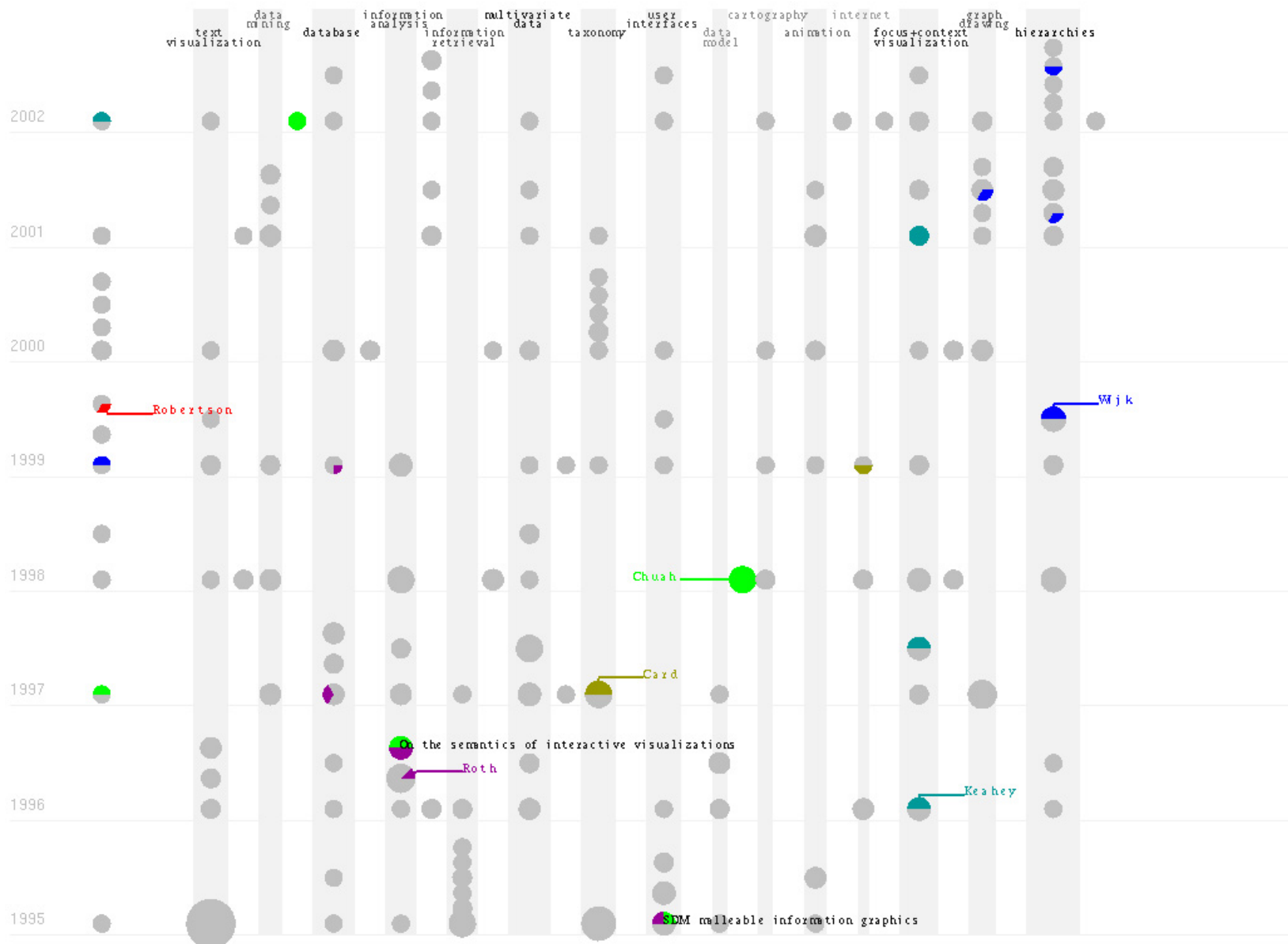
# ONE FOR ALL

Bibliography Visualization of the Information Visualization Symposia

Soon Tee Teoh and Kwan-Liu Ma

Department of Computer Science

University of California, Davis

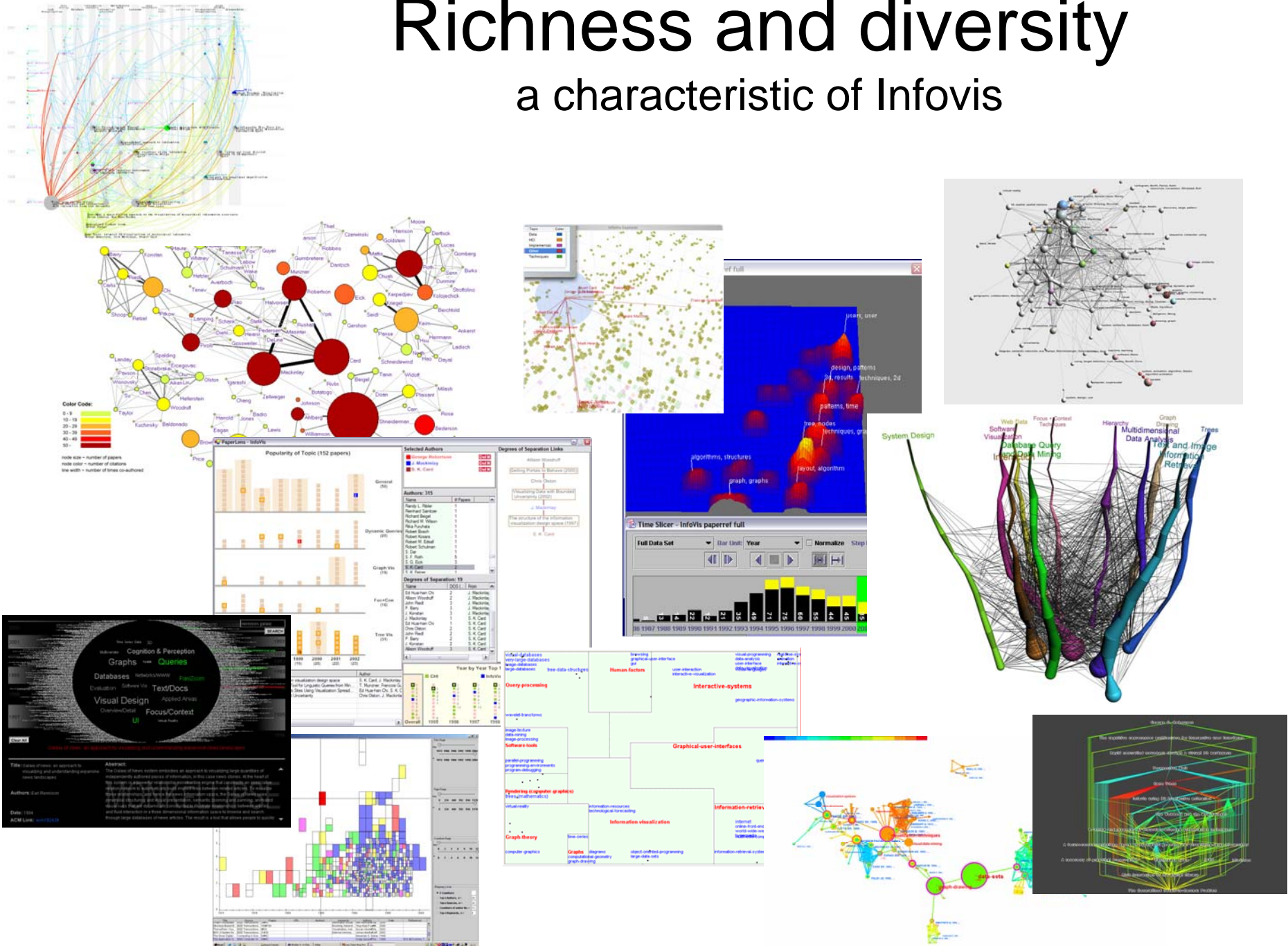


- Tree-Maps a space-filling approach to the visualization of hierarchical information structures  
Brian Johnson, Ben Shneiderman
- Generalized fisheye views  
George Furnas
- Cone Trees: animated 3D visualizations of hierarchical information  
George Robertson, Jock McKinlay, Stuart Card



# Lessons Learned

# Richness and diversity a characteristic of InfoVis

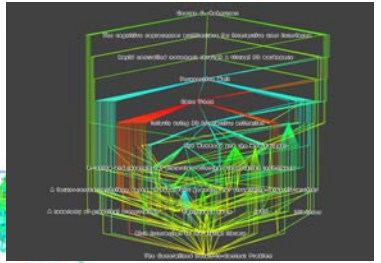
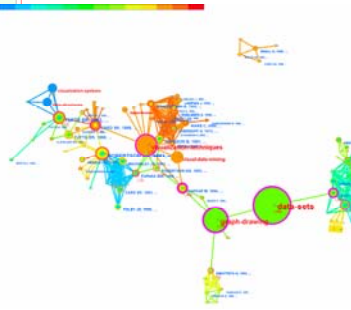
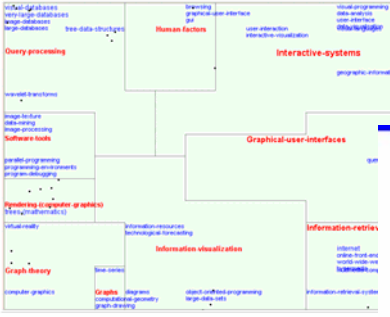
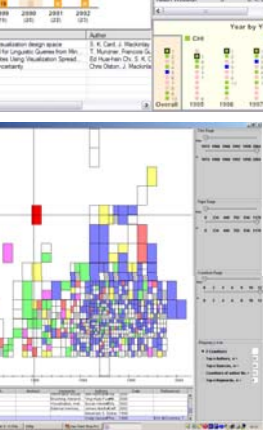
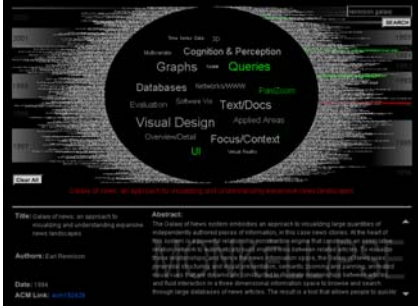
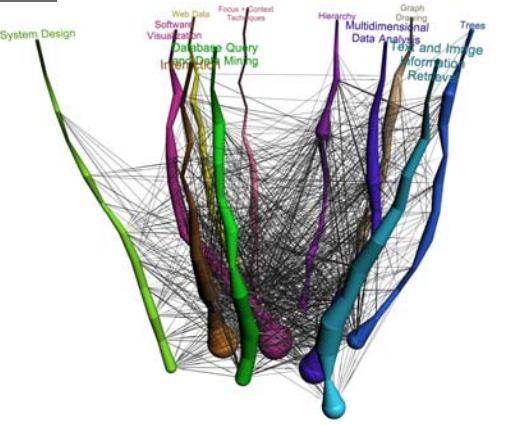
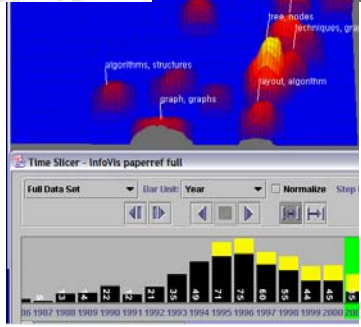
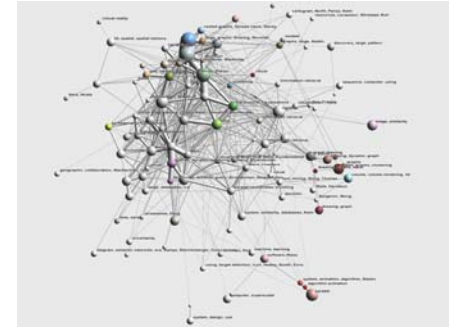
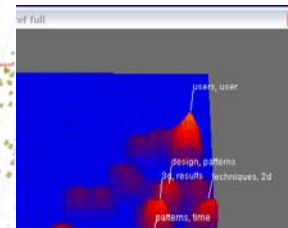
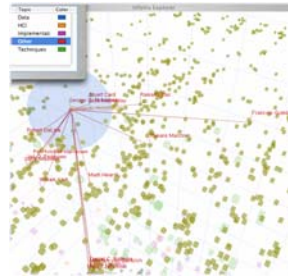
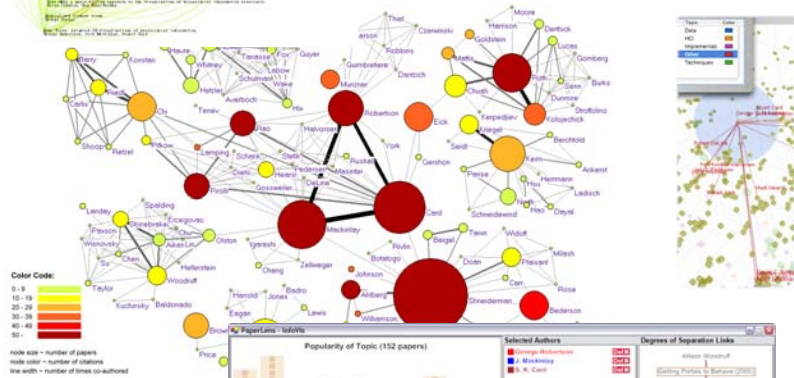


# Richness and diversity

a characteristic of InfoVis

Interaction methods

Underlying analysis methods

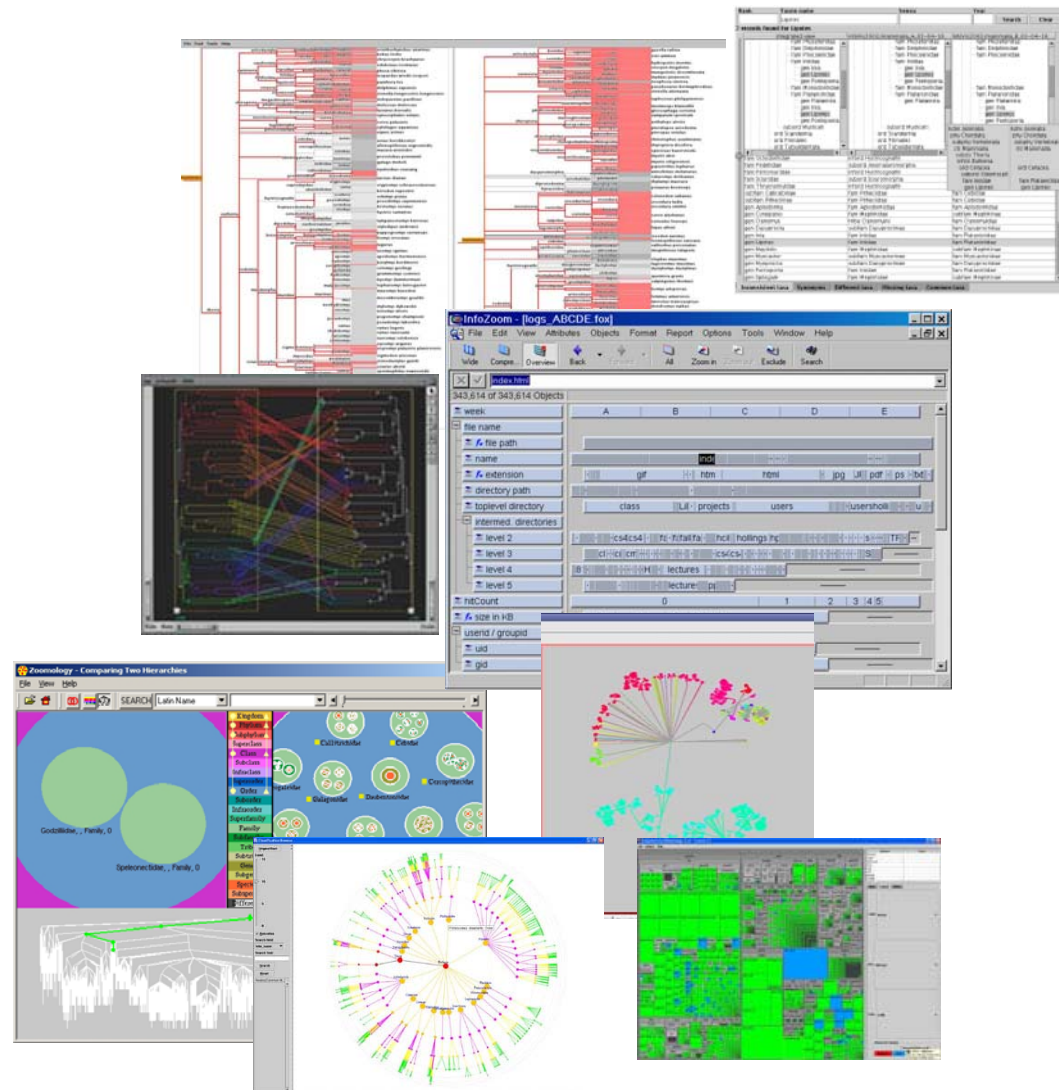


# Contest Outcomes

- Generate new techniques and case studies
- Enrich the Domain Knowledge Analysis research area
- Teach about benchmark design
- Demonstrate challenge of evaluation
  - “the purpose of visualization is insight”
    - Not a common exercise for system designers
      - Structured answer form was very helpful
    - User expertise important
  - Partial systems
  - Surprises

# Progress since 2003 contest

- 2003:  
Pairwise  
comparison  
of trees





# Progress since 2003 contest

- 2003:
  - Benchmark too complex (too realistic?)
    - 3 tree types, many tasks → only partial answers
  - Barely any insights reported (mostly descriptions of tools)
- 2004:
  - We simplified... (1 dataset – fewer tasks)
  - Authors described INSIGHTS better
- Better benchmark
  - Realistic ↔ Simple
  - Generic ↔ Specific
  - Interesting to community
- Still issue with tasks
  - Open ↔ simple

# Progress since 2003 contest

- 2003:
  - Benchmark too complex (too realistic?)
    - 3 tree types, many tasks → only partial answers

- B

- 2004

- V

- A

- Better

- R

- G

- In

- Still

- C

## **Insights :**

most cited papers are CHI papers

Two references dominate

## **See what was not there:**

No references to other visualisation conferences

Extent of missing data

## **Anomalies:**

References to yet-unpublished papers

# Progress since 2003 contest

- 2003:
  - Benchmark too complex (too realistic?)
    - 3 tree types, many tasks → only partial answers
  - Barely any insights (mostly descriptions of tools)
- 2004:
  - We simplified... (1 dataset – fewer tasks)
  - Authors described INSIGHTS better
- Better benchmark
  - Realistic ↔ Simple
  - Generic ↔ Specific
  - Interesting to community
- Still issue with choosing tasks
  - Open ↔ simple

# Progress since 2003 contest

- 2003
  - Example: **Example: relationship between authors?**
    - Co-authors**
      - who works with who?
      - Where are the « cliques » or « empires »
    - Co-citations**
      - Authors working on similar topics?
    - Number of co-authors**
      - Who works with big teams?
    - Authors with more publications**
    - Authors with more references**
    - etc.
- 2004
  - V
  - A
- Between 2003 and 2004
  - P
  - C
  - I
- Still to be done
  - C
    - First author only OR all authors**
    - Infovis papers only OR also include all references**
    - Mixing authors+keywords**

# Progress since 2003 contest

- 2003:
  - Benchmark too complex (too realistic?)
    - 3 tree types, many tasks → only partial answers
  - Barely any insights (mostly descriptions of tools)
- 2004:
  - We simplified... (1 dataset – fewer tasks)
  - Authors described INSIGHTS better
- Better benchmark
  - Realistic ↔ Simple
  - Generic ↔ Specific
  - Interesting to community
- Still issue with choosing tasks
  - Open ↔ simple

# Repository

- Materials (2003 and 2004) are online  
[www.cs.umd.edu/hcil/InfovisRepository](http://www.cs.umd.edu/hcil/InfovisRepository)
- You can:
  - Explore datasets and tasks with your tools
  - PUBLISH your materials and results in repository
    - e.g. we contacted Refvis (Thomson ResearchSoft)
    - Encourage your colleagues

# Why run more contests?

Benchmarks are difficult to:

create

promote

use

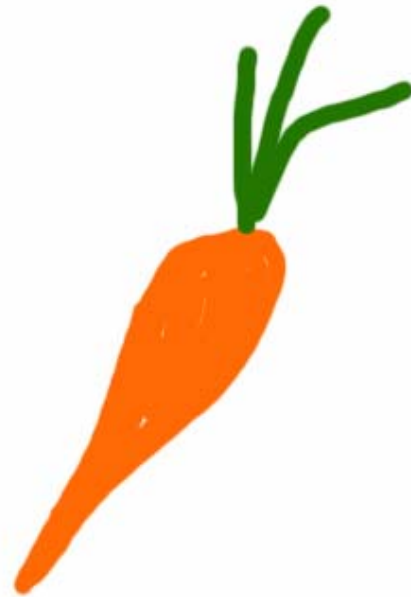
# Why run more contests?

Benchmarks are difficult to:

create

promote

use





# Infovis 2005 Contest

- Chairs:
  - Georges Grinstein
  - 3 others
- Detail announcement by end of year
  - Dataset
  - Tasks
  - Deadlines (similar to 2004)

# Richness and diversity a characteristic of InfoVis

