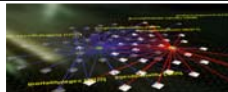
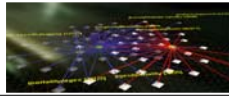


L701, Guest Lecture
Data Mining and Information Visualization

Katy Börner
School of Library and Information Science
INDIANA UNIVERSITY
BLOOMINGTON
katy@indiana.edu



Research Projects



Information Visualization

“Information Visualization is a process of transforming data and information that are not inherently spatial, into a visual form allowing the user to observe and understand the information.”

(Source: Gershon and Eick, *First Symposium on Information Visualization*)

- Rooted in geography
- Not even 15 years old
- Far reaching (IR, etc.)
- Tremendous potential

Humans can detect a single dark pixel in a 500 x 500 array of white pixels in less than a second. This screen can be replaced every second by another, enabling a search of 15 million pixels in a minute (Ware, 2000).

Also, people have a truly remarkable ability to recall pictorial images. In one study, Standing, Conezio, & Haber (1970) showed S's 2560 pictures, each for 10 seconds over 7 hours, in a 4-day period. Afterwards, S's were asked to classify pictures presented at a rate of 16 pictures/min and they achieved better than 90% accuracy.

Katy Börner: SLIS Orientation Session, Fall 2004.



Knowledge Domain Analysis, Visualization & Modeling

- Examining the Evolution and Distribution of Patent Classifications.** Daniel O. Kutz
- Visualization of Weblog Topic Spaces.** Elijah Wright
- Visual Interface to SRS.** Ketan K. Mane
- Visualizing the Work of the Supreme Court: 59 Years of Data.** Peter A. Hook
- Mapping Medline Papers, Genes, and Proteins Related to Melanoma Research.** Kevin W. Boyack, Ketan K. Mane and Katy Börner
- Visualizing the Blogosphere.** Ning Yu, Susan Herring, Inna Kouper, John Paolillo, Lois Ann Scheidt, Mike Tyworth, and Elijah Wright
- PNAS Mapping Knowledge Domains**
- Unstructured Peer-to-Peer Networks: Topological Properties and Search Performance.** George Fletcher, Hardik Sheth and Katy Börner
- Analysis and Visualization of the IV 2004 Contest Dataset.** Weimao Ke, Katy Börner and Lalitha Viswanath

Analysis and Visualization of Diffusion Patterns

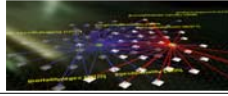
- The ActiveWorlds Toolkit.** Shashikant Penumarthy
- The Spatial-Semantic Impact of a Collaborative Information Virtual Environment on Group Dynamics.** Chaomei Chen and Katy Börner
- Mapping Virtual Worlds and Their Inhabitants**

InfoVis Cyberinfrastructure

- InfoVis Cyber-Infrastructure Database.** Todd Holloway et al.
- InfoVis Learning Modules.** Katy Börner
- InfoVis Cyberinfrastructure Software Framework.** Shashikant Penumarthy, Bruce Herr and Katy Börner

<http://ella.slis.indiana.edu/~katy/gallery/04-openhouse/>

Katy Börner: SLIS Orientation Session, Fall 2004.



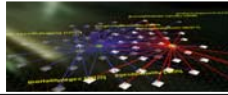
Knowledge Domain Analysis and Visualization

To answer questions such as:

- What are the major research areas, experts, institutions, regions, nations, grants, publications, journals in xx research?
- Which areas are most insular?
- What are the main connections for each area?
- What is the relative speed of areas?
- Which areas are the most dynamic/static?
- What new research areas are evolving?
- Impact of xx research on other fields?
- How does funding influence the number and quality of publications?

Answers are needed by funding agencies, companies, and researchers.

Katy Börner: SLIS Orientation Session, Fall 2004.



Process of Mapping Knowledge Domains

DATA EXTRACTION	UNIT OF ANALYSIS	MEASURES	LAYOUT (often one code does both similarity and ordination steps)		DISPLAY
			SIMILARITY	ORDINATION	
SEARCHES ISI INSPEC Eng Index Medline ResearchIndex Patents etc.	COMMON CHOICES Journal Document Author Term	COUNTS/FREQUENCIES Attributes (e.g. terms) Author citations Co-citations By year THRESHOLDS By counts	SCALAR (unit by unit matrix) Direct citation Co-citation Combined linkage Co-word / co-term Co-classification VECTOR (unit by attribute matrix) Vector space model (words/terms) Latent Semantic Analysis (words/terms) incl. Singular Value Decomp (SVD) CORRELATION (if desired) Pearson's R on any of above	DIMENSIONALITY REDUCTION Eigenvector/ Eigenvalue solutions Factor Analysis (FA) and Principal Components Analysis (PCA) Multi-dimensional scaling (MDS) LSA, Topics Pathfinder networks (PFNet) Self-organizing maps (SOM) includes SOM, ET-maps, etc. CLUSTER ANALYSIS SCALAR Triangulation Force-directed placement (FDP)	INTERACTION Browse Pan Zoom Filter Query Detail on demand ANALYSIS
BROADENING By citation By terms					

Börner, Katy, Chen, Chaomei, and Boyack, Kevin. (2003) Visualizing Knowledge Domains. In Blaise Cronin (Ed.), *Annual Review of Information Science & Technology, Volume 37*, Medford, NJ: Information Today, Inc./American Society for Information Science and Technology, chapter 5, pp. 179-255.

Katy Börner: SLIS Orientation Session, Fall 2004.

<http://www.indiana.edu/aag05/>

Mapping Humanity's Knowledge and Expertise in the Digital Domain

At the **101st Annual Meeting** of the **Association of American Geographers** | Denver, CO: April 5-9, 2005.

Session Organizers

Katy Börner, **Indiana University**
André Skupin, **University of New Orleans**

Sponsors

Cartography and GIS specialty groups

Description

This session will bring together leading researchers and practitioners that aim to develop techniques, tools, and infrastructures to map humanity's knowledge and expertise for the improvement of science and education.

Knowledge and expertise is typically extracted from digitally available literature, news, computer mediated communication data as well as from information about the producers and consumers of those data sets. Advanced data analysis techniques in combination with spatial metaphors, geographic principles, and cartographic methods are applied to organize, visualize, and communicate the semantic relationships inherent in the data.

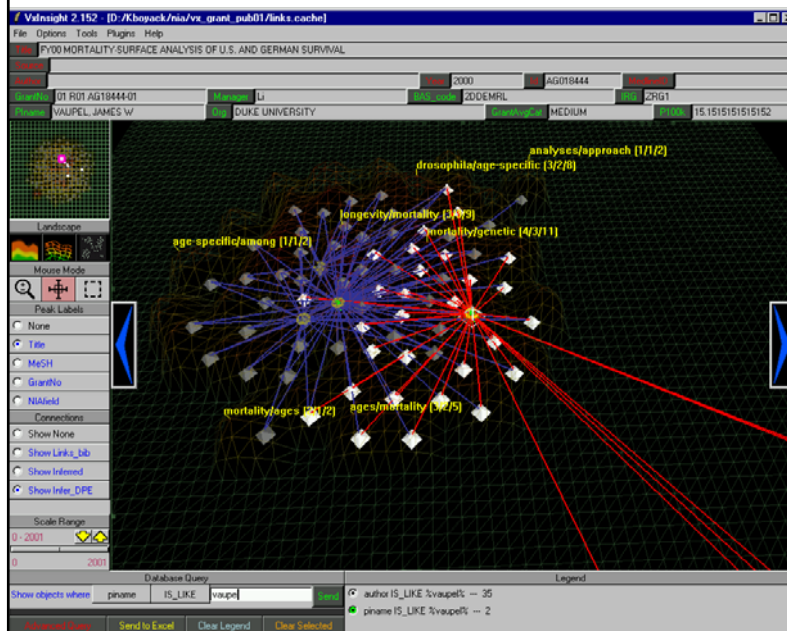
The ultimate goal of this work might be an interactive cartographic map of all of science, with continents representing the major research areas such as, e.g., biology or physics, dots denoting major authors, PIs, papers or news, dynamically evolving research frontiers, blinking 'hot' papers and topics, etc. This map could be used to teach and understand the evolving structure of all of science, to identify major experts, to find and read the most relevant papers and news, to see the effects of resource allocation decisions, to study social networks, etc. Last but not least, it would provide a unique bird's eye view of major experts in specific areas and mankind's knowledge in general.

Some of the leading-edge research on this topic is found where geography intersects with information/library science, computer science, and cognitive science. We invite papers on the broad foundations, computational methods, software systems, and evaluation of such data analyses and visualizations, as they have emerged in this interdisciplinary endeavor.



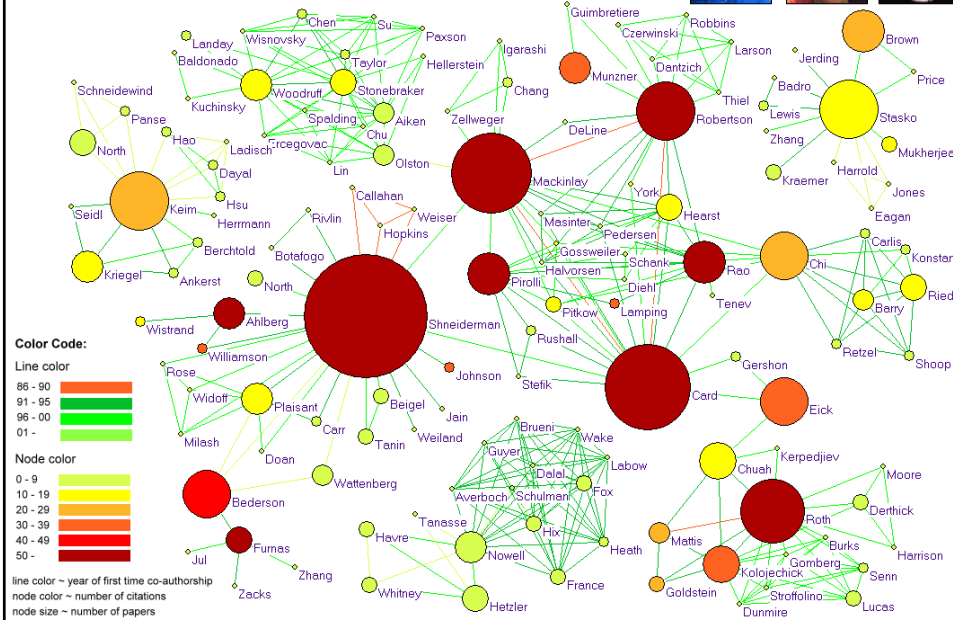
Indicator-Assisted Evaluation and Funding of Research

Visualizing the influence of grants on the number and citation counts of research papers (Boyack & Börner, 2003)



Mapping InfoVis Co-Authorships ([Interactive Map](#))

IV Contest Submission (Ke. Viswanath & Börner, 2004)



Information Diffusion Patterns

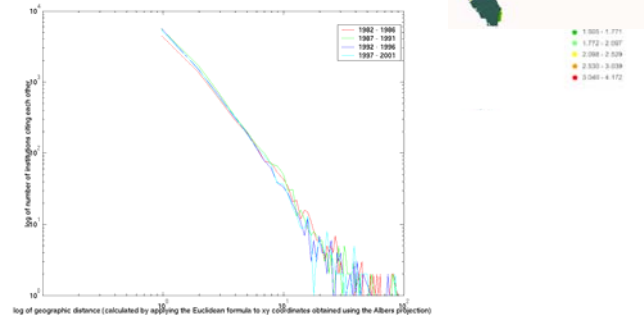
(Börner & Penumathy, 2004)



Top 500 most highly cited U.S. institutions.

Each institution is assumed to produce and consume information.

Does Internet lead to more global citation patterns, i.e., more citation links between papers produced at geographically distant research institutions?



Katy Börner: SLIS Orientation Session, Fall 2004.

Examining the Evolution and Distribution of Patent Classifications

1 Patents Granted Over the Last 20 Years

Year	Number of Patents
1983	1000
1984	1100
1985	1200
1986	1300
1987	1400
1988	1500
1989	1600
1990	1700
1991	1800
1992	1900
1993	2000
1994	2100
1995	2200
1996	2300
1997	2400
1998	2500
1999	2600
2000	2700
2001	2800
2002	2900

Top Classes 1976 - 1982

Class	Number of Patents
2600	1000
2601	800
2602	700
2603	600
2604	500
2605	400
2606	300
2607	200
2608	100
2609	100
2610	100
2611	100
2612	100
2613	100
2614	100
2615	100
2616	100
2617	100
2618	100
2619	100
2620	100
2621	100
2622	100
2623	100
2624	100
2625	100
2626	100
2627	100
2628	100
2629	100
2630	100
2631	100
2632	100
2633	100
2634	100
2635	100
2636	100
2637	100
2638	100
2639	100
2640	100
2641	100
2642	100
2643	100
2644	100
2645	100
2646	100
2647	100
2648	100
2649	100
2650	100
2651	100
2652	100
2653	100
2654	100
2655	100
2656	100
2657	100
2658	100
2659	100
2660	100
2661	100
2662	100
2663	100
2664	100
2665	100
2666	100
2667	100
2668	100
2669	100
2670	100
2671	100
2672	100
2673	100
2674	100
2675	100
2676	100
2677	100
2678	100
2679	100
2680	100
2681	100
2682	100
2683	100
2684	100
2685	100
2686	100
2687	100
2688	100
2689	100
2690	100
2691	100
2692	100
2693	100
2694	100
2695	100
2696	100
2697	100
2698	100
2699	100
2700	100

In the United States, each patent gets assigned to one out of more than 450 classes covering broad application domains. An examination of the size and growth of patent classes provides insight about patenting trends.

Tree maps, a space filling method developed in the HCI Lab at the University of Maryland, are used to communicate major results. Tree maps represent a tree structure as nested rectangles with each rectangle representing a node. A rectangular area is first allocated to hold the representation of the tree, and this area is then subdivided into a set of rectangles that represent the next level of the tree. This process continues recursively on the resulting rectangles to represent each lower level of the tree. The parent-child relationship is indicated by enclosing the child rectangle by its parent rectangle. Typically the size of each rectangle corresponds to the size of the node. Additional information about a node, e.g., its age or value, can be represented by the color of the innermost rectangle.

2 Fast Growth Domains 1983 - 1987 / 1998 - 2002

Slow Growth Domains 1983 - 1987 / 1998 - 2002

Shown is a comparison of the patent class space for 1983 to 1987 and 1998 to 2002. There is a preponderance of growth in the 1998 to 2002 patent space, which correlates to the increase in patent grants during this period. By comparing the growth in categories, one can distinguish between domains that have been recouping a larger amount of patent grants.

3 Apple Computers

Depicted above is how Apple Computers' portfolio has changed in yearly increments from 1980 to 2002.

Lemelson's patent holdings below show a more even distribution over multiple classes. No class dominates over a majority of the years for granted patents, instead they are distributed more broadly over the intellectual space.

Jerome Lemelson

Katz, Daniel O. Examining the Evolution and Distribution of Patent Classifications. Accepted for the Information Visualization Conference, London, UK, July 2004.

The material is based upon work supported by the National Science Foundation under Grant No. IIS-0238261.

For more information, contact Katy Börner at katz@indiana.edu.

Mapping Virtual Worlds and Their Inhabitants

Addressed User Tasks

The developed visualization tools are intended to support social navigation in three dimensional virtual worlds, to help evaluate and optimize the design of virtual worlds, and to provide a means to study the communities evolving in virtual worlds.

Design Concept

The figures show the layout and utilization of diverse virtual worlds. Information on the position, size and rotation of all three-dimensional objects as well as on interaction possibilities are used to generate a map of a world. Overlaid are user interaction data such as movements, web click, or chat activity recorded during virtual events in a particular world.

Design Implementation

All virtual objects are rendered in transparent green to preserve the visibility of layered objects. A reference grid indicates the size of the virtual world. To show the evolution of a world, darker colors are used for older objects and lighter colors for younger ones. Web links and teleports are indicated by green square and purple plus signs respectively. Color-coding is used to denote the chronological sequence of user interactions.

Map of the 1000x1000 in large (published in *AvatarWorld: Information Systems*, October 2004) of a virtual world. The map shows the layout and utilization of diverse virtual worlds. Information on the position, size and rotation of all three-dimensional objects as well as on interaction possibilities are used to generate a map of a world. Overlaid are user interaction data such as movements, web click, or chat activity recorded during virtual events in a particular world.

A few snapshots of user tracks recorded during the virtual conference on learning in Three Dimensions (L3D) in November 2003.

This visualization tool has been applied to map and visualize numerous virtual worlds, including virtual worlds, virtual worlds, virtual worlds, and virtual worlds.

The spatial movement of a large group can be visualized by displaying this a position of the world of all group members for each time step and identifying the movement path of the group centroid. The global transparency - reserved by the member - allows an overview of the movement of the group.

Katy Börner and Shashikant Penumarty (2003) Social Diffusion Patterns in Three-Dimensional Virtual Worlds. *Information Visualization Journal*, 2(3):182-198.

This material is based upon work supported by the National Science Foundation under Grant No. IRI 0411846.

For more information, contact Katy Börner at katz@indiana.edu.

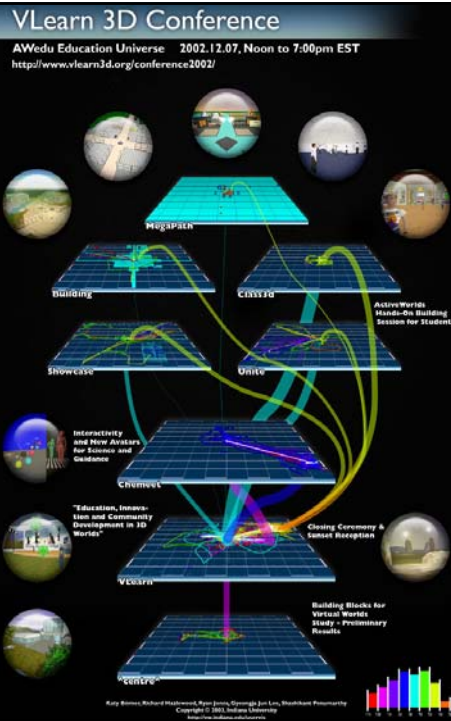
VLearn 3D Vis

(Börner, Hazelwood, Jones, Lee & Penumarthy, 2003)

Temporal-spatial distribution of Conference attendees

- Conference worlds are represented by square, perspective maps, each labeled by its name.
- Worlds accessed at the beginning of the conference are placed at the bottom, worlds accessed later toward the top.
- Next to each world is a circular snapshot of the virtual venue. Short descriptions of the main sessions are added as text.
- Major jumps between worlds are visualized by transparent lines. The thickness of each line corresponds to the number of traveling users. Color coding was used to denote the chronological paths of the conference sessions.

Katy Börner: SLIS Orientation Session, Fall 2004.



Information Visualization CyberInfrastructure

The InfoVis CyberInfrastructure provides access to data, software code and learning modules as well as computing resources in support of the analysis, modeling and visualization of diverse data sets.

DATABASES

An Oracle database provides access to publications, patents, grants and grant opportunities. The database is continuously and automatically updated. (<http://ivislab.indiana.edu/>)

COMPUTING RESOURCES

The InfoVis CyberInfrastructure is hosted at Indiana University's Research Database Complex, comprising of two Sun V2200 servers with 12 900MHz processors and 96 GB of memory each. 6 TB fiber channel disks are attached to both servers. A Sun V880 system with 4 CPUs and 4GB memory serves as the web front-end for the database servers. (<http://ivislab.indiana.edu/>)

SOFTWARE

An open source IVC framework was designed to facilitate the integration of diverse data analysis, modeling and visualization algorithms. New algorithms, data persistence methods, look and feels for the interface and even entire toolkits can be easily "plugged in" or "unplugged". (<http://ivislab.indiana.edu/>)

LEARNING MODULES

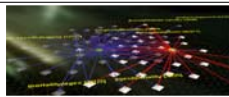
A set of associated learning modules aims to equip learners with a practical skill set by providing code and advice to quickly modify and run different algorithms, test diverse interaction techniques and design features, and to quickly generate and compare information visualizations. (<http://ivislab.indiana.edu/>)



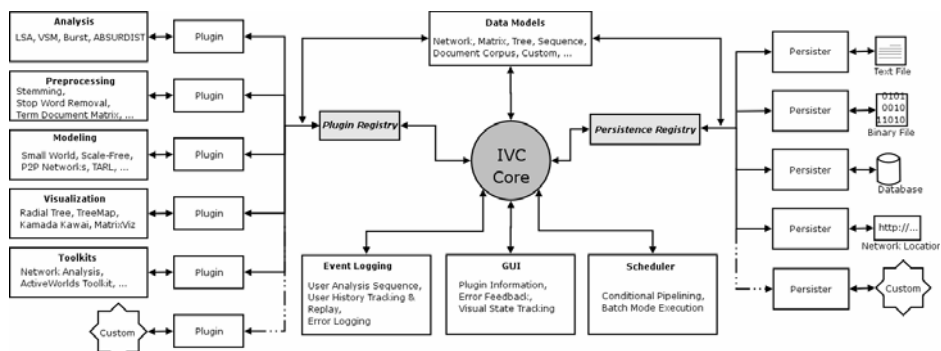
InfoVis Lab, School of Library and Information Science, Indiana University (2004).
 For more information, contact Katy Börner at kborner@indiana.edu

This material is based upon work supported by the National Science Foundation under Grants No. IRI-0206261 and DUE-0339623.

Katy Börner: SLIS Orientation Session, Fall 2004.



IVC Software Framework



See SLIS Colloquium talk on the *InfoVis Cyberinfrastructure* on Nov 19, 2004.

Katy Börner: SLIS Orientation Session, Fall 2004.

Workshop on Information Visualization Software Infrastructures

Sat. Oct. 9th, 2004, 10am-6pm
Room Hill Country A-B

Right before the [IEEE Symposium on Information Visualization \(InfoVis\) 2004](#).

Workshop Chairs

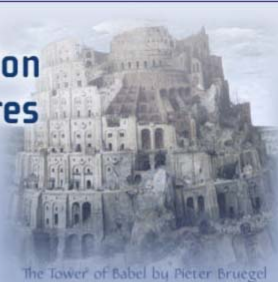
- **Jean-Daniel Fekete**, INRIA Futurs, France (Author of [The InfoVis Toolkit](#)),
Jean-Daniel.Fekete@inria.fr, Home page: <http://www.lri.fr/~fekete/index.en.html>
- **Katy Börner**, Indiana University, USA (Co-Author of the [InfoVis Cyberinfrastructure](#))
katy@indiana.edu, Home page: <http://ella.slis.indiana.edu/~katy/>

Description

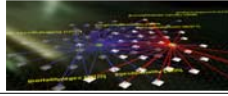
Information visualization systems and toolkits are becoming available for a large range of visualization and interaction techniques and are used in diverse application domains. This workshop is aimed at gathering experts involved in building such infrastructures to share their views, understand the issues involved and trying to find ways to avoid fragmentation and improve collaborations.

To participate in the workshop, you should submit a semi-structured position paper explaining your view of what an infrastructure should provide, describe what you consider as the main challenges for such infrastructures and describe the capabilities of toolkits of systems you have already built, following a form available [here](#) by Sept 30th, 2004.

<http://vw.indiana.edu/ivsi2004/>



The tower of Babel by Pieter Bruegel



InfoVis Learning Modules



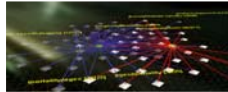
Learning Modules

Most information visualizations are highly interactive. While a number of excellent textbooks exist, the two-dimensional printouts on paper often cannot convey their true visual appearance and interactive performance. Several textbooks come with accompanying web sites that contain snapshots of user interfaces as well as animations and movies. However, none of them facilitates the exploration, application, evaluation, and comparison of algorithms.

This web page will provide access to a number of learning modules. Each learning module comes with an:

- ◆ Introduction
- ◆ Discussion of Existing Algorithms
- ◆ Learning Task
 - ◊ A challenging scenario to use the code in the [XML Toolkit](#) to visualize a data set to support a specific user group.
- ◆ Programming Exercise
 - ◊ An programming exercise plus an explanation of possible solutions.
- ◆ Opportunities & Challenges, and
- ◆ References to research papers, online demos, (commercial) applications)

Katy Börner: SLIS Orientation Session, Fall 2004.



Visualizing Tree Data

<http://iv.slis.indiana.edu/lm/lm-trees.html>

[Learning Modules](#) > Visualizing Tree Data

[Description](#) | [Usage Hints](#) | [Learning Task](#) | [Discussion](#) | [References](#) | [Acknowledgments](#)

Description

Many data sets come in tree format. There are family trees, organizational charts, classification hierarchies, and directory structures. The figure below shows an inheritance tree by Ernst Haeckel ('Stammbaum' in German). Read also [To Draw a Tree](#) by Pat Hanrahan.

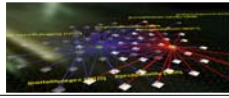


[Click image for larger version](#)

A tree graph is a set of straight line segments (edges) connected at their ends containing no closed loops (cycles). You can also call it a simple, undirected, connected, acyclic graph (or, equivalently, a connected forest). A tree with n nodes has $n-1$ graph edges. All trees are bipartite graphs.

Many trees have a root node and are called rooted trees. Trees without a root node are called free trees. Subsequently, we will only consider rooted trees. In rooted trees, all nodes except the root node have only one parent node. Nodes which have no children are called leaf nodes. All other nodes are referred to as intermediate nodes.

Katy Börner: SLIS Orientation Session, Fall 2004.



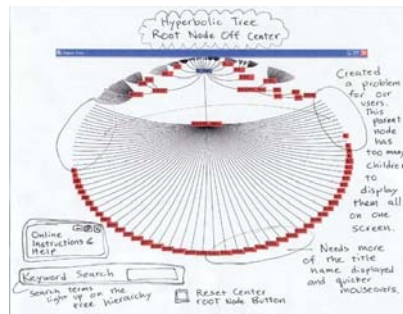
Student's Project Results

User & Task Analysis for Visualizing Tree Data

- Visualizing the structure of IU's Decision Support System
- Visualizing the co-occurrences of keywords in DLib Magazine articles.
- Visualization of the Java API
- Visualizing the the Library of Congress Classification System to retrieve legal materials in a library.

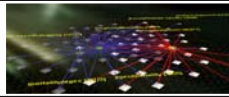
See Handin pages at

<http://ella.slis.indiana.edu/~katy/handin/L579-S04/cgi/handinlogin.cgi>



Katy Börner: SLIS Orientation Session, Fall 2004.

Image by Peter Hook and Rongke Gao



Time Series Analysis & Visualization

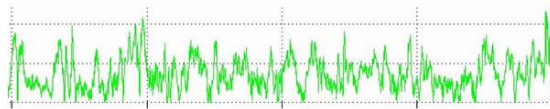
<http://iv.slis.indiana.edu/lm/lm-time-series.html>

[Learning Modules](#) > Visualizing Time Series Data

[Description](#) | [Usage Hints](#) | [Learning Task](#) | [Discussion](#) | [References](#) | [Acknowledgments](#)

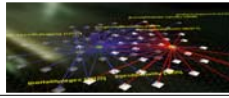
Description

A time series is a sequence of events/observations which are ordered in one dimension, e.g., time. Frequently, successive observations depend on each other and it makes sense to display them in a (time) sorted fashion, e.g., as a scatter plot. Alternatively, one could be interested to know how many observations of a certain value have been made. Here one would sort the observations by value, count the number of observations for each value and derive a histogram. Time series data can be continuous, i.e., there is an observation at every instant of time see figure below, or discrete, i.e., observations exist for regularly or irregularly spaced intervals.



Time series are recorded, analyzed and used in diverse domains of science. Check out the [Time Series Data Library](#) maintained by Rob Hyndman and Muhammad Akram for numerous data sets from Agriculture, Chemistry, Crime, Demography, Ecology, Finance, Health, Hydrology, Industry, Labour market, Macro-Economics, Meteorology, Micro-Economics, Physics, Production, Sales, Simulated series, Sport, Transport & Tourism or Utilities.

Katy Börner: SLIS Orientation Session, Fall 2004.



Student's Project Results

Time Series Analysis & Visualization

- Using Timesearcher and the Burst Detection Algorithm to Analyze the Stock Market from 1925 to 1945
- Applying Burst and TimeSearcher to Chat Data
- Lab Access Trends
- Quest Atlantis Chat Log Data

See Handin pages at

<http://ella.slis.indiana.edu/~katy/handin/L579-S04/cgi/handinlogin.cgi>

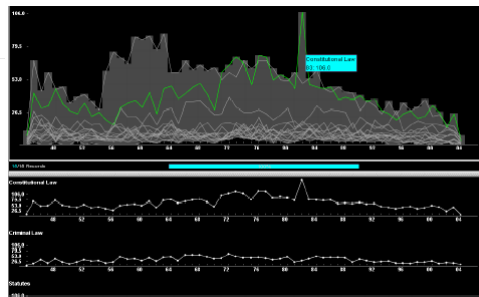
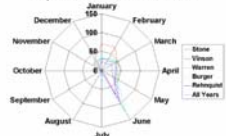
Katy Börner: SLIS Orientation Session, Fall 2004.

Visualizing the Work of the United States Supreme Court Based on Time Data and Top Level West Topics

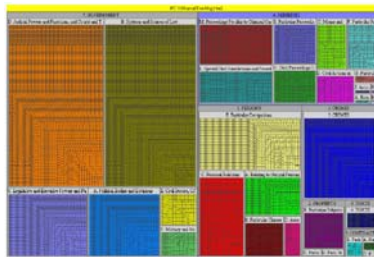
by Peter A. Hook & Rongke Gao



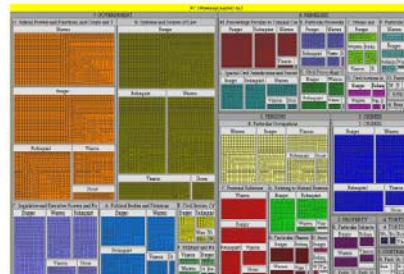
Topics Per Month Per Court



Top fifteen most occurring topics from 1944 to 2004 in Timesearcher



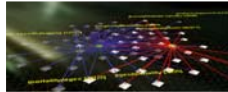
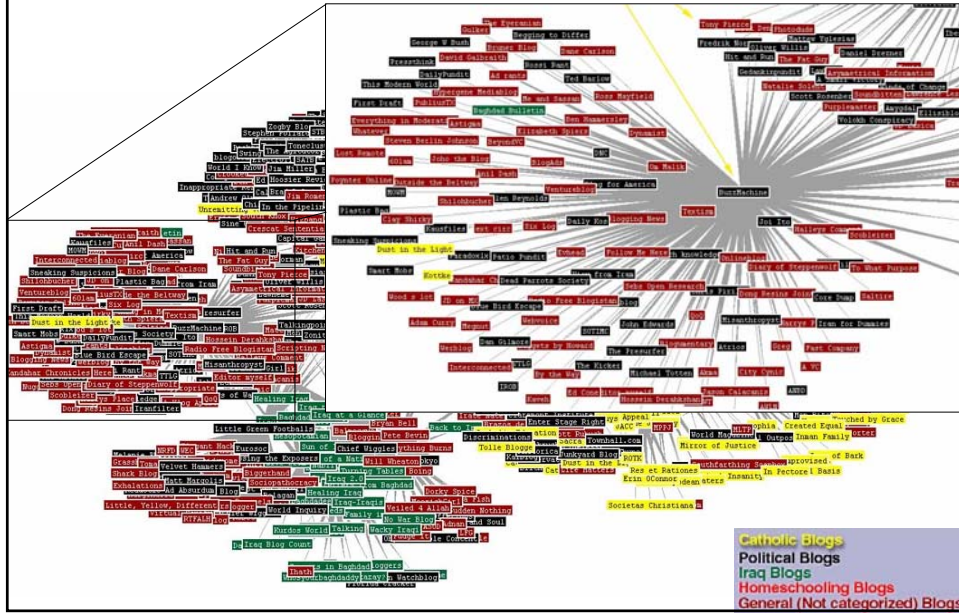
All topics grouped by West Category and Sub-Category grouped over the entire lengths of the data set



All topics by West Category and Sub-Category grouped corresponding to the five chief justices

Visualizing Niches of the Blog Universe

BY Mike Tyworth and Elijah Wright



IVC DB Data Sets (<http://iv.slis.indiana.edu/db>)

Papers and Patents



Medline
 Number of Entries: 11,693,477
 Years covered: 1963-2002
 Size: 135 MB (gunzipped)



Proceedings of the National Academy of Science (PNAS)
 Number of Entries: 16,169
 Years covered: 1987-2002
 Size: 583 MB



United States Patent and Trademark Office (Patents)
 Number of Entries: 2,582,647
 Years covered: 1976-2003
 Size: 350 MB

Grant Awards



National Science Foundation (NSF)
 Number of Entries: 181,132
 Years covered: 1985-2002
 Size: 400 MB



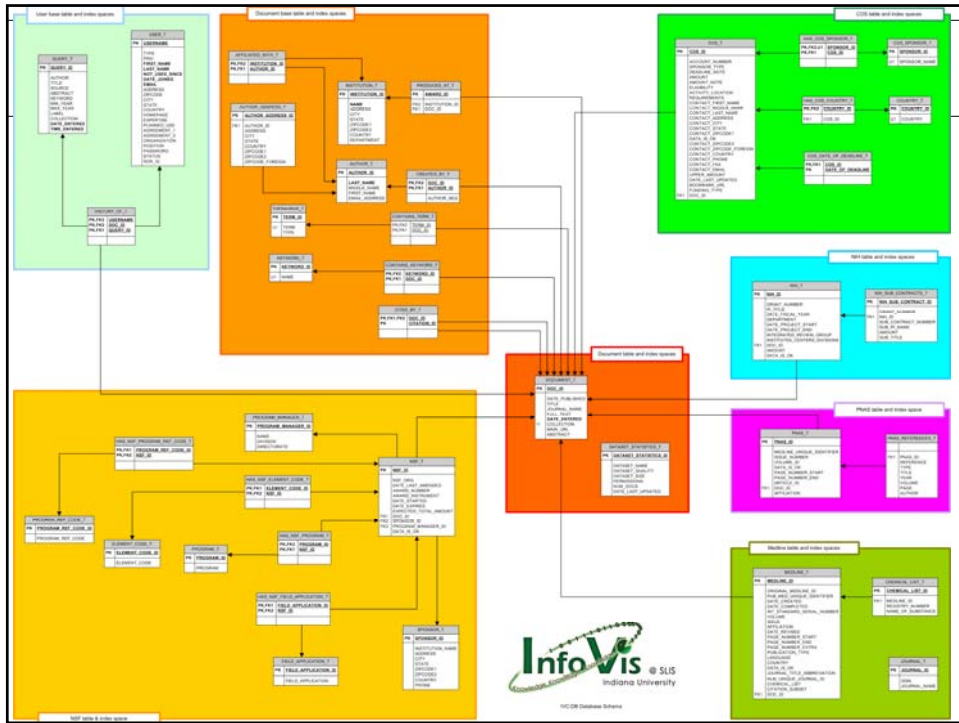
National Institute of Health (NIH)
 Number of Entries: 1,003,521
 Years covered: 1972-1992 and 1994-2002
 Size: 2.3 GB

Funding Opportunities



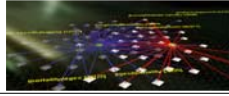
Community of Science (COS)
 Number of Entries: 38,154 (5,000 new entries per month)
 Years covered: 2001-present
 Size: 60 MB

Katy Börner: SLIS Orientation Session, F:



Questions?

Katy Börner: SLIS Orientation Session, Fall 2004.

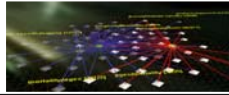


Courses

- ❖ **L597** Structural Data Mining and Modeling
Fall 2004 (<http://ella.slis.indiana.edu/~katy/L597>)
- ❖ **L579** Information Visualization (formerly L697)
Spring 2005 (<http://ella.slis.indiana.edu/~katy/L579>)

Both courses have final projects in which students work with outside collaborators.

Katy Börner: SLIS Orientation Session, Fall 2004.

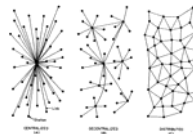


L597 Structural Data Mining and Modeling

This course

- Introduces students to major methods, theories, and applications of structural data mining and modeling.
- Covers elementary graph theory and matrix algebra, data collection, structural data mining, data modeling, and applications.

Upon taking this course students will be able to analyze and describe real networks (power grids, WWW, social networks, etc.) as well as relevant phenomena such as disease propagation, search, organizational performance, social power, and the diffusion of innovations.

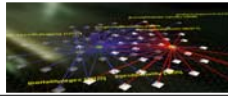


Format: Lectures and 4-5 labs.

Class Webpage: <http://ella.slis.indiana.edu/~katy/L597>

Fall 2004 Talk Series on Networks and Complex System

<http://vw.indiana.edu/talks-fall04/>



L579 Information Visualization



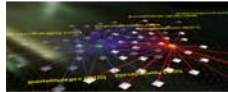
This course covers

- Perceptual basis of information visualization.
- Data mining algorithms that enable extraction of relationships in data.
- Visualization and interaction techniques.
- Discussions of systems that drive research and development, and
- Future trends and remaining fundamental problems in the field.

Students do weekly readings, provide a presentation on specific readings, do Java projects, and participate in class & online discussion.

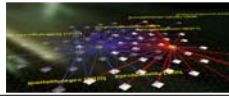
Class Webpage: <http://ella.slis.indiana.edu/~katy/L579>

Katy Börner: SLIS Orientation Session, Fall 2004.



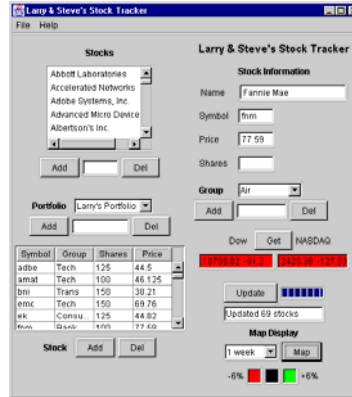
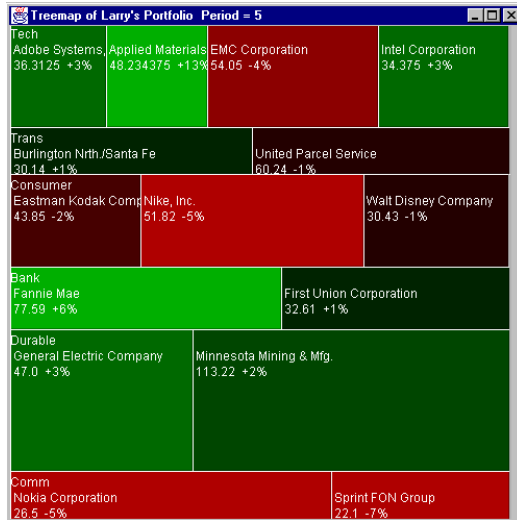
Student Projects

Katy Börner: SLIS Orientation Session, Fall 2004.

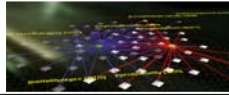


Stock Tracker

Larry Mongin & Steve Rice

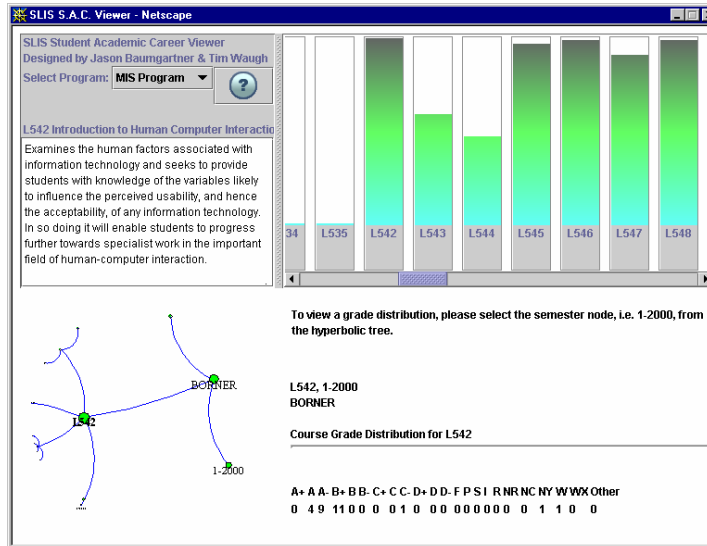


Katy Börner: SLIS Orientation Session, Fall 2004.

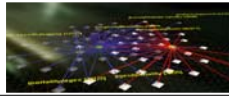


SLIS Student Academic Career Viewer

Jason Baumgartner & Tim Waugh

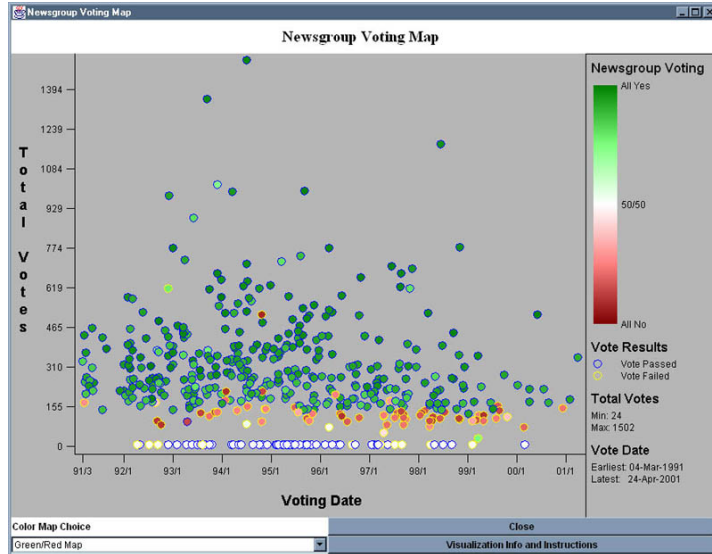


Katy Börner: SLIS Orientation Session, Fall 2004.

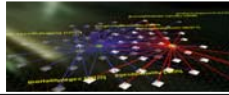


Newsgroup Votes Visualization

David Heald (Collaborator John C. Paolillo)

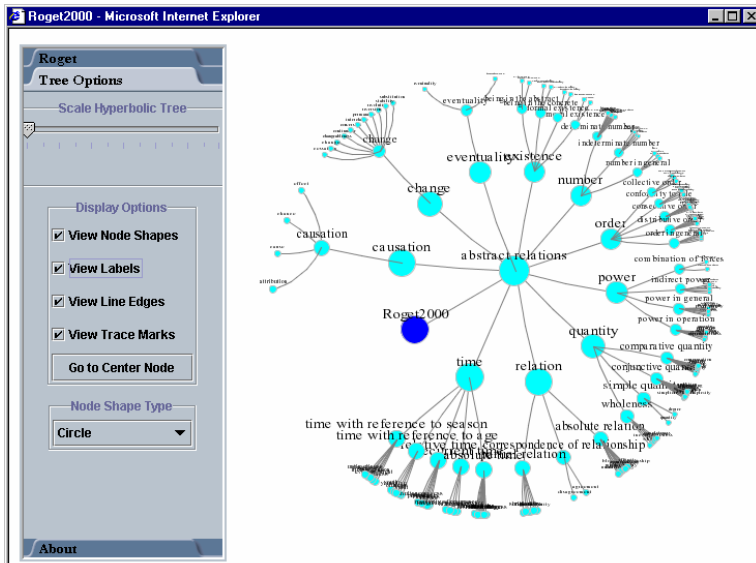


Katy Borner: SLIS Orientation Session, Fall 2004.

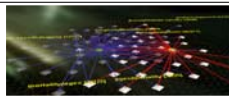


Hyperbolic Tree Visualization of Roget's Thesaurus

Jason Baumgartner & Tim Waugh (Collaborator John Old)



Katy Borner



Oncosifter – Personalized Cancer Information



- ✦ Filters cancer related news and medical information from Medlineplus and Cancer.gov.
- ✦ Provides hierarchical search & browsing interface.

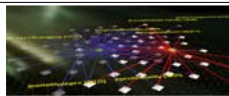
Faculty Client: Javed Mostafa,
SLIS, IUB

To be presented at the Information Visualization Interfaces for Retrieval and Analysis Workshop @ JCDL, Houston, TX, May 31, 2003.

The screenshot displays the Oncosifter web application with several key components:

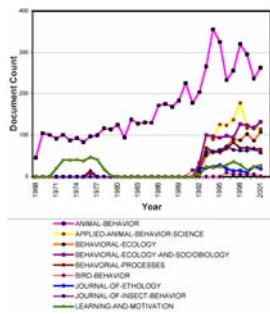
- Directory Interface:** A navigation menu with categories like Endocrine, Eye, Genitourinary, Germ Cell, and Gynecologic.
- Visual Interface:** A circular radial chart showing hierarchical data.
- Text Based Interface:** A search results page for the keyword "BRAN" showing 9 types of "Brain" cancer, including Adult Brain Tumors.
- Profile Based Interface:** A table of search results for various cancer types with columns for document counts (C1, C2, C3, C4).

Katy Börner: SLIS Orientation Session, Fall 2004.

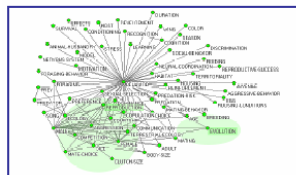
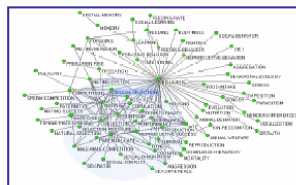
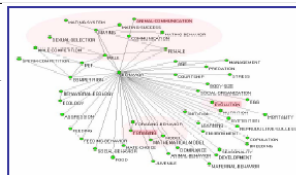


Visualizing the Animal Behavior Domain

Faculty Client:
Emilia Martins & Terry Ord, Biology, IU



Sidharth Thakur, Ketan Mane, Katy Börner, Emilia Martins & Terry Ord. Content Coverage of Animal Behavior Data. *Submitted to IEEE InfoVis, 2003.*



Katy Börner: SLIS Orientation Session, Fall 2004.