The Future of Information Discovery:

Designing to Support Social Media and Exploratory Search

Ben Shneiderman ben@cs.umd.edu

Founding Director (1983-2000), Human-Computer Interaction Lab Professor, Department of Computer Science Member, Institute for Advanced Computer Studies





Information Retrieval/Search Technologies Seminar 2008 IBM Research Haifa December 16, 2008

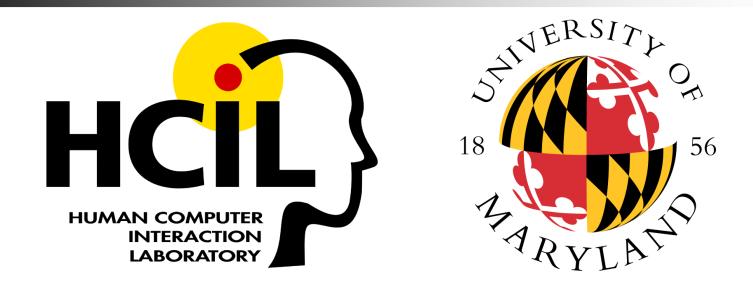
The Future of Information Discovery

Ben Shneiderman ben@cs.umd.edu

Founding Director (1983-2000), Human-Computer Interaction Lab Professor, Department of Computer Science Member, Institute for Advanced Computer Studies







Interdisciplinary research community

- Computer Science & Info Studies
- Psych, Socio, Poli Sci & MITH

(www.cs.umd.edu/hcil)



U.S. Library of Congress

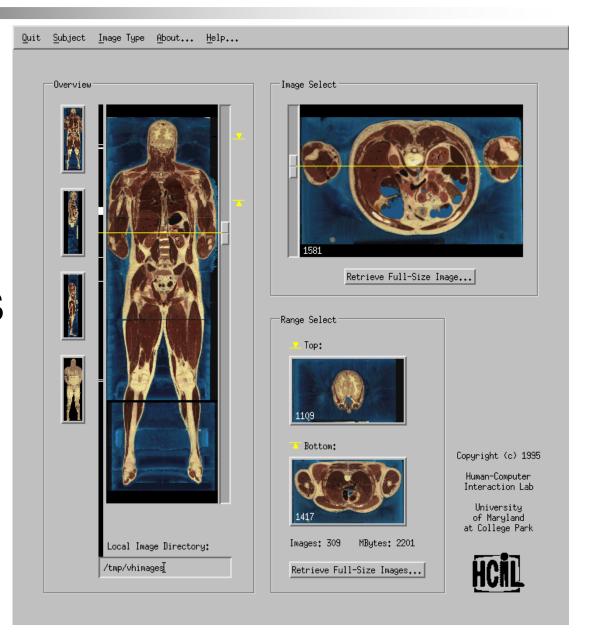


- Scholars, Journalists, Citizens
- Teachers, Students



Visible Human Explorer (NLM)

- Doctors
- Surgeons
- Researchers
- Students



NASA Environmental Data

- Scientists
- Farmers
- Land planners
- Students



Bureau of the Census



- Economists, Policy makers, Journalists
- Teachers, Students

American FactFinder



NSF Digital Government Initiative



Census, NCHS, BLS, EIA, NASS, SSA

- Find what you need
- Understand what you Find





www.ils.unc.edu/govstat/

FEDSTATS



International Children's Digital Library



Books

Se

in English

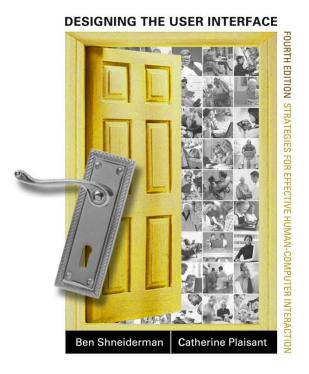
Keywords

www.childrenslibrary.org

University of Maryland, Jeanette has three daughters.

Design Issues

- Input devices & strategies
 - Keyboards, pointing devices, voice
 - Direct manipulation
 - Menus, forms, commands
- Output devices & formats
 - Screens, windows, color, sound
 - Text, tables, graphics
 - Instructions, messages, help
- Collaboration
- Help, tutorials, training

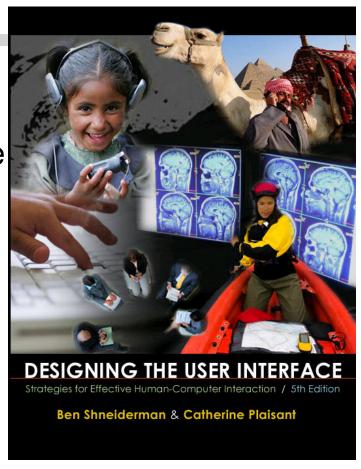


www.awl.com/DTUI



Design Issues

- Input devices & strategies
 - Keyboards, pointing devices, voice
 - Direct manipulation
 - Menus, forms, commands
- Output devices & formats
 - Screens, windows, color, sound
 - Text, tables, graphics
 - Instructions, messages, help
- Collaboration & Social Media
- Help, tutorials, training
- Search & Visualization

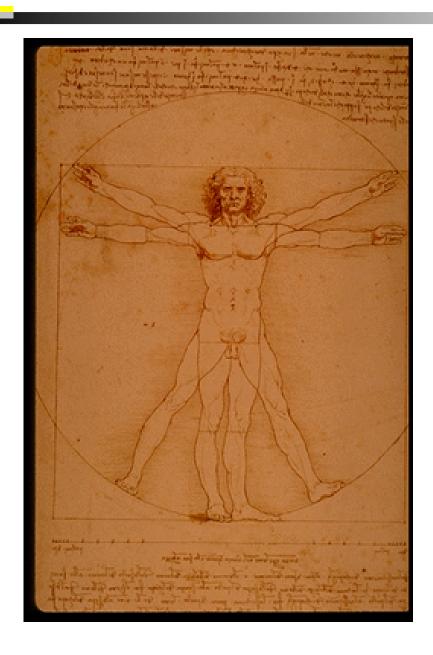


www.awl.com/DTUI

Fifth Edition: March 2009



Information Visualization



The eye...
the window of the soul,
is the principal means
by which the central sense
can most completely and
abundantly appreciate
the infinite works of nature.

Leonardo da Vinci (1452 - 1519)



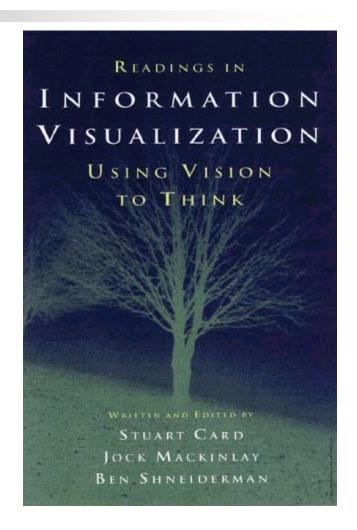
Using Vision to Think

Visual bandwidth is enormous

- Human perceptual skills are remarkable
 - Trend, cluster, gap, outlier...
 - Color, size, shape, proximity...
- Human image storage is fast and vast

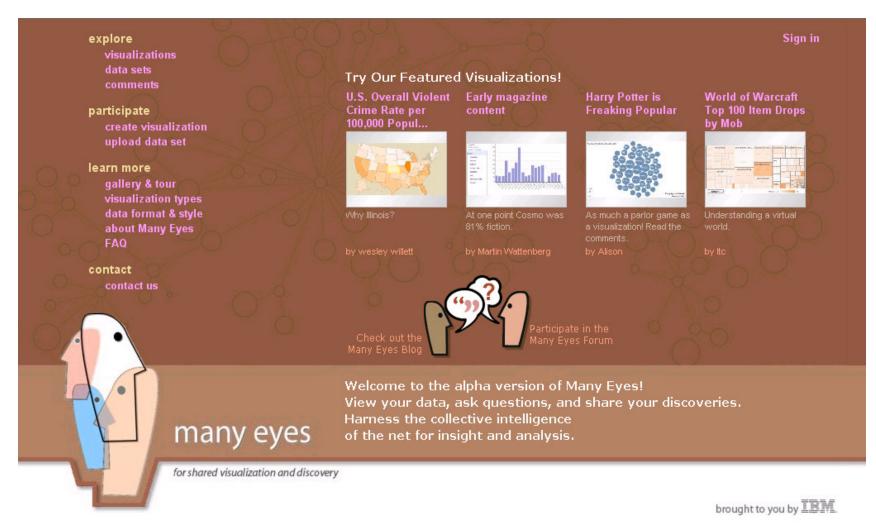
Opportunities

- Spatial layouts & coordination
- Information visualization
- Scientific visualization & simulation
- Telepresence & augmented reality
- Virtual environments





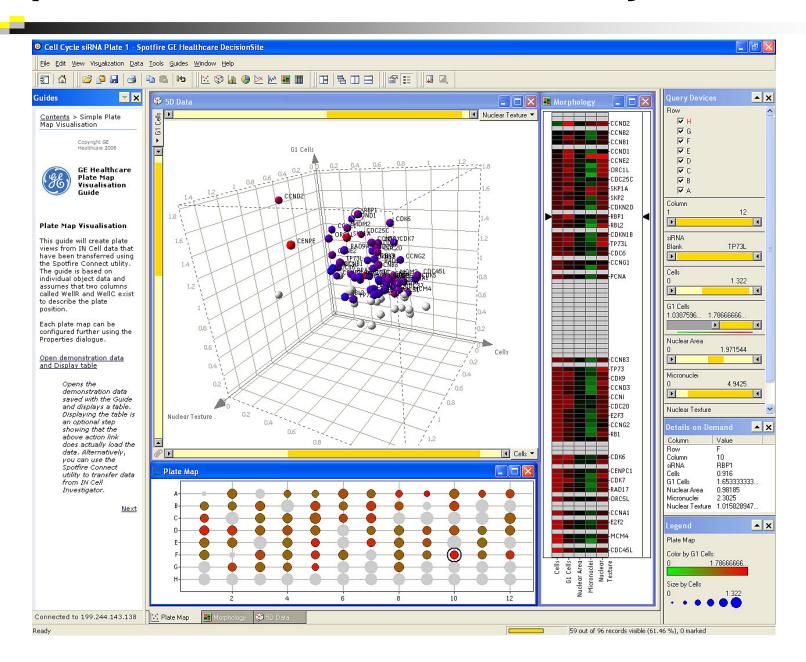
ManyEyes: A web sharing platform



http://services.alphaworks.ibm.com/manyeyes/app

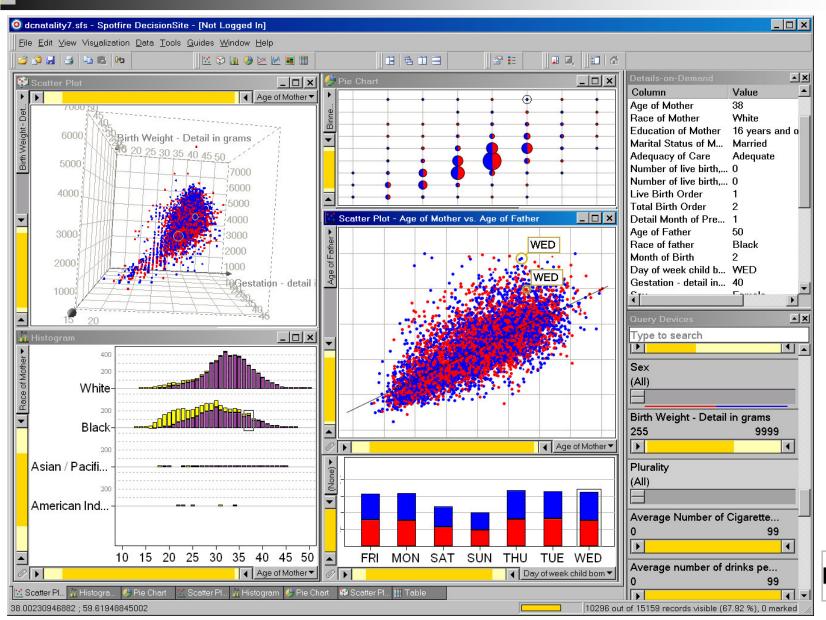


Spotfire: Retinol's role in embryos & vision





Spotfire: DC natality data



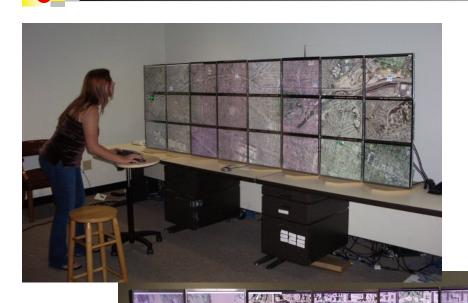








100M-pixels & more

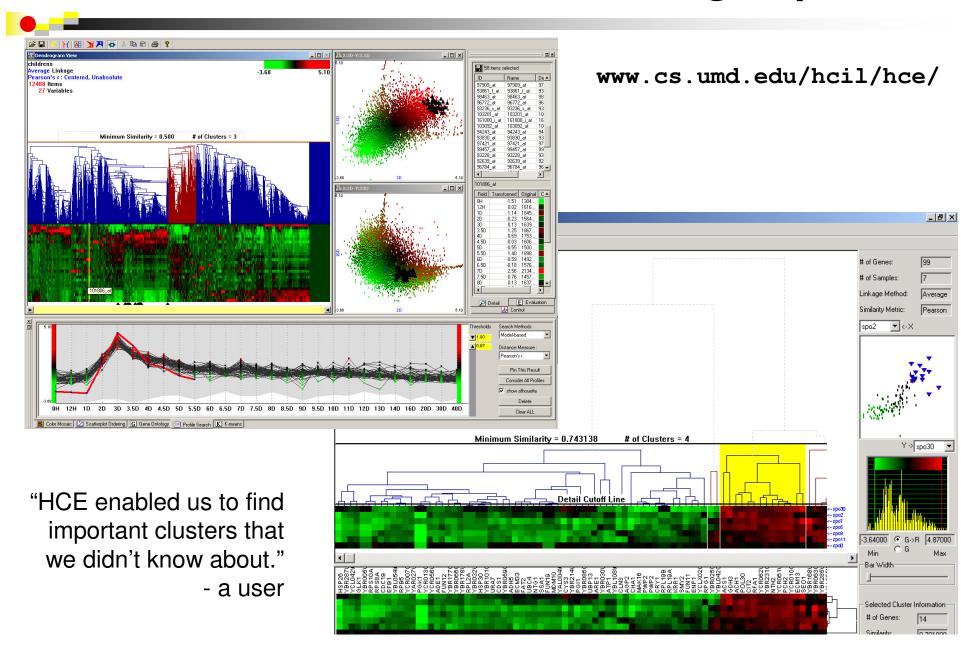


Large displays for single users



infovis.cs.vt.edu/gigapixel

Multi-Var: Hierarchical Clustering Explorer



Information Visualization: Mantra

- Overview, zoom & filter, details-on-demand



Information Visualization: Data Types

SciViz

1-D Linear

Document Lens, SeeSoft, Info Mural

2-D Map

GIS, ArcView, PageMaker, Medical imagery

3-D World

CAD, Medical, Molecules, Architecture

InfoViz

Multi-Var

Spotfire, Tableau, XGobi, TableLens, ParCoords

Temporal

LifeLines, TimeSearcher, DataSpiral, DataMontage

Tree

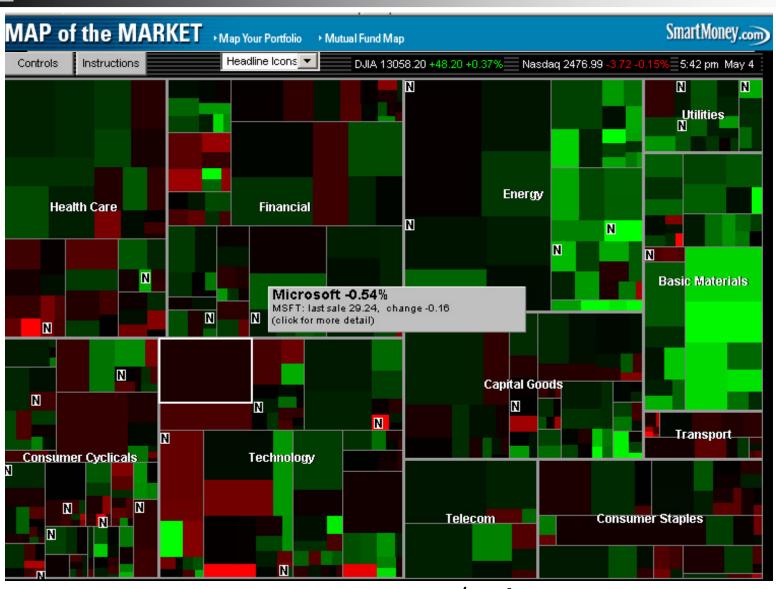
Cone/Cam/Hyperbolic, SpaceTree, Treemap

Network

Pajek, JUNG, SocialAction, .netmap



Treemap: Stock market, clustered by industry





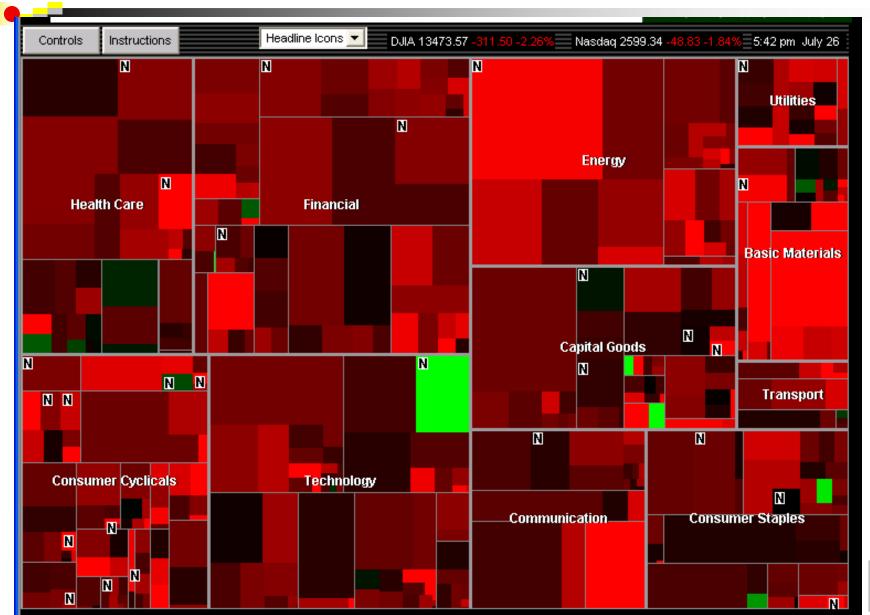
www.smartmoney.com/marketmap

Market falls steeply Feb 27, 2007, with one exception





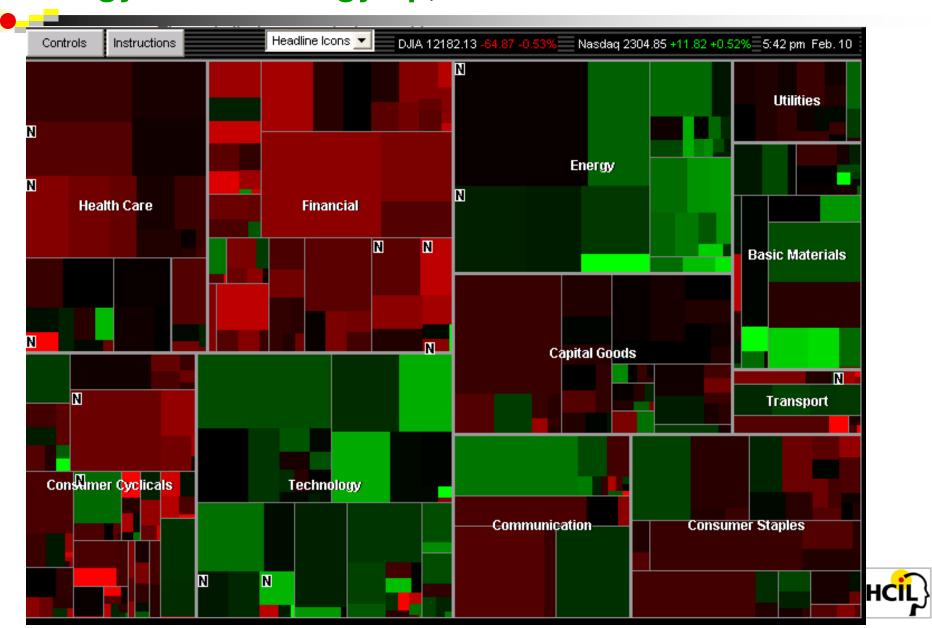
Market falls 311 points July 26, 2007, with a few exceptions



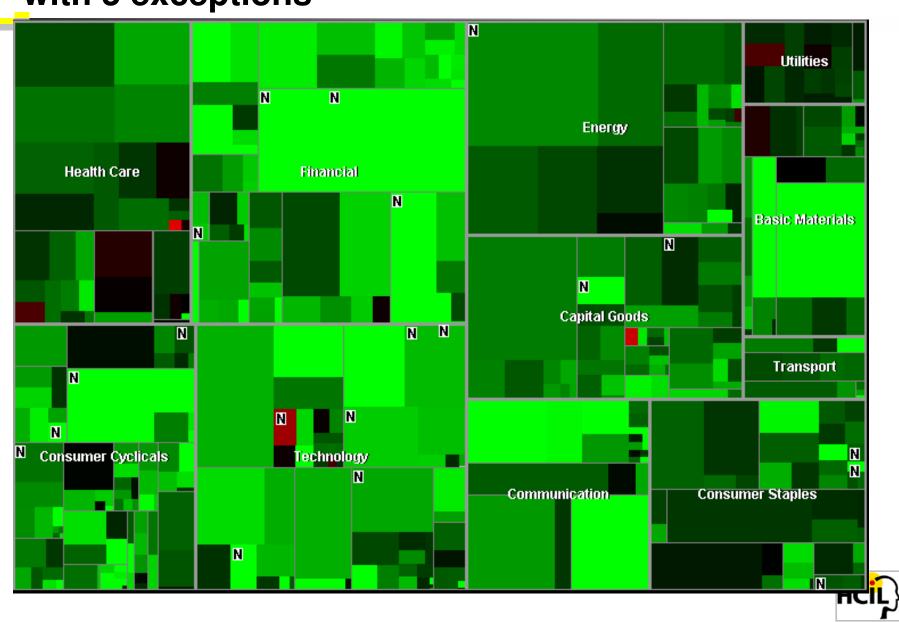


Market mixed, February 8, 2008

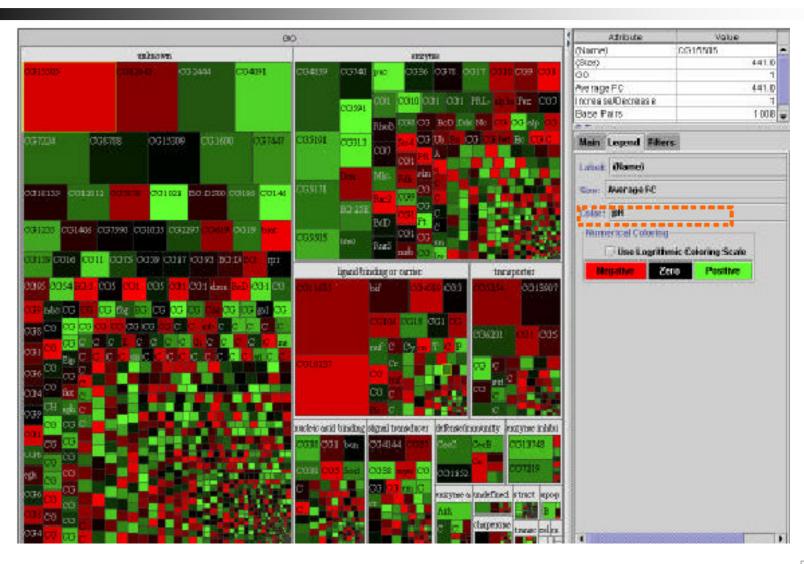
Energy & Technology up, Financial & Health Care down



Market rises 319 points, November 13, 2007, with 5 exceptions



Treemap: Gene Ontology

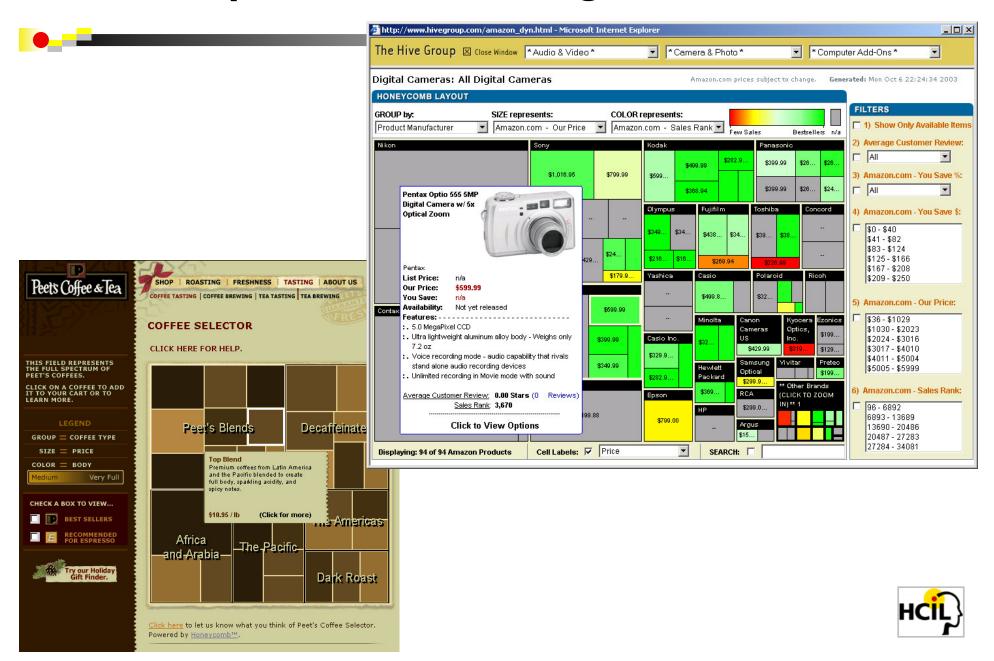






Treemap: Product catalogs

www.hivegroup.com

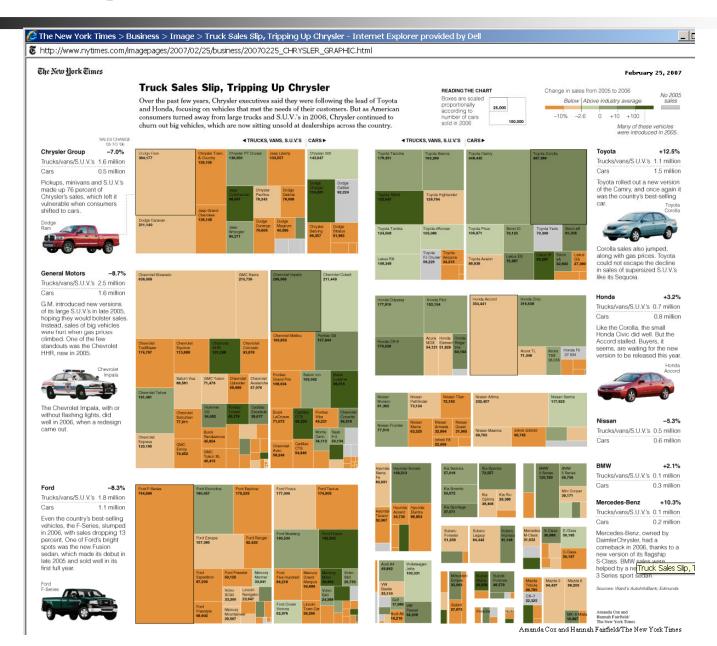




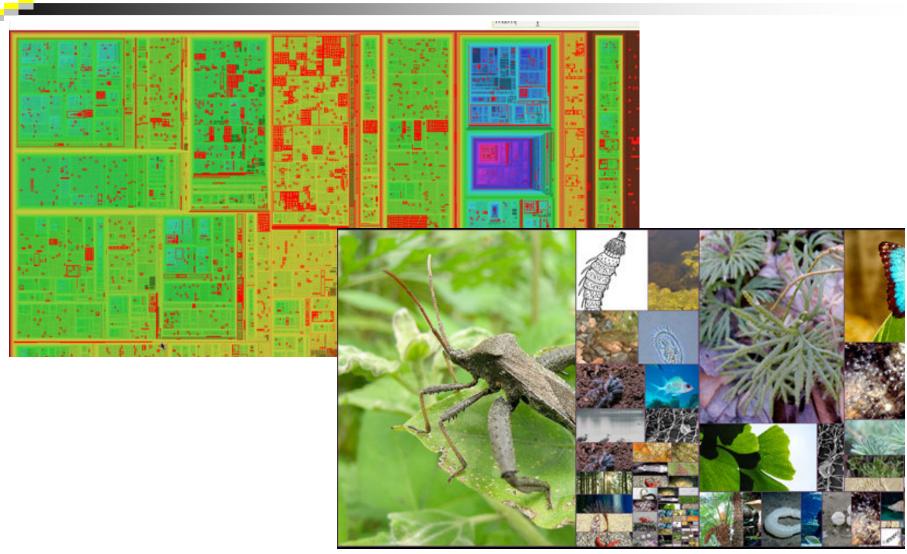
MARINE CORPS EQUIPMENT READINESS INFORMATION TOOL tome I Readiness by Commodity I Readiness by Functional Area I Feedback Logous GROUP SIZE COLOR FILTER BY PROBLEMS MR rating - Current Functional Assa W S rating - Current This This This Month Querter Year MR 72.50 FA 46 - Infantry Weapons stelligence Serveill \$8.28 - Engineers Support Г View Deadlined Bens for this TAMCN Generate Control Chart for this TAMON 800. 825. FA 10, Padios Generate Historical View for this TAMCN FILTER BY ITEMS Show Tabular data for this TAMCN Selected items will be show 800. 00 - Urknown A08... A20. A19 Track this TAMON 10 - Rados Section 1 Zoom in on this group 11 - Communications Support Equipme DOUGOSALISTO D 13 - Air Command/Control Equipment Zoom Out to view all 14 - Air Support Rada/IFF Equipment Close Menu 16 - Electronic Equipment BOY_ B2460 17 - Ground Support Radar 18 - Tactical Remote Sensor Equipmen Tractor, Full-Tracked, W/AngleBlade, T-5 19 - Intelligence/Surveillance Equipme 20 - Generators Last This : This : This 21 - Environmental Control Equipment TA 25 - Towers 011 Week: Month: Quarter: Year Motor Transport 23 - Earthmoving Equipment 101% 101% 100% 94% 26 - Materials Handling Equipment Equipment 112% 110% 29 - Engineer Support 111% : 113% \$11. 100 30 - Trucks 91% : 89% 89% 85% 35 - Towed Motor Transport Equipmen 40 - Tanks Click mouse for reports and options. 41 - Assault Amphibious Vehicles FA 23 -42 - Light Armored Vehicles FA 13 - Air Command Earthmoreus 43 - Artillery 11 Westinery 18:55 OOF guiget Fried ACCESS TO 45 - Infantry Weapons Equipment 48 - Anti-Amor Weapon Systems and C E100. ADDUSCHA! 43 - Missile Systems 50 - High Density/Low Deadline Currently displaying: 186 of 186 items 97 - Communications Security Equipme 98 - Assault Crafts Search for TAMON # Search by: TAMON Title Filter by Label by Show Only 99 - Nuclear Biological Chemical Equip Pacing Items thickes. ENDOSE FELTER BY SEARCH CRITERIA EXCESS FILTER & LABEL FILTER BY PACING ITEM OCHs meeting search criteria will be shown. Pacing Items will be shown Inniate Hamp Wencess equipment Hanage MyHERIT Portfolio 🛴 Filter by MyMERIT Portfolio 📗 Label by MyHERIT Portfolio (Interhente colling Corple)



Treemap: NY Times – Car&Truck Sales

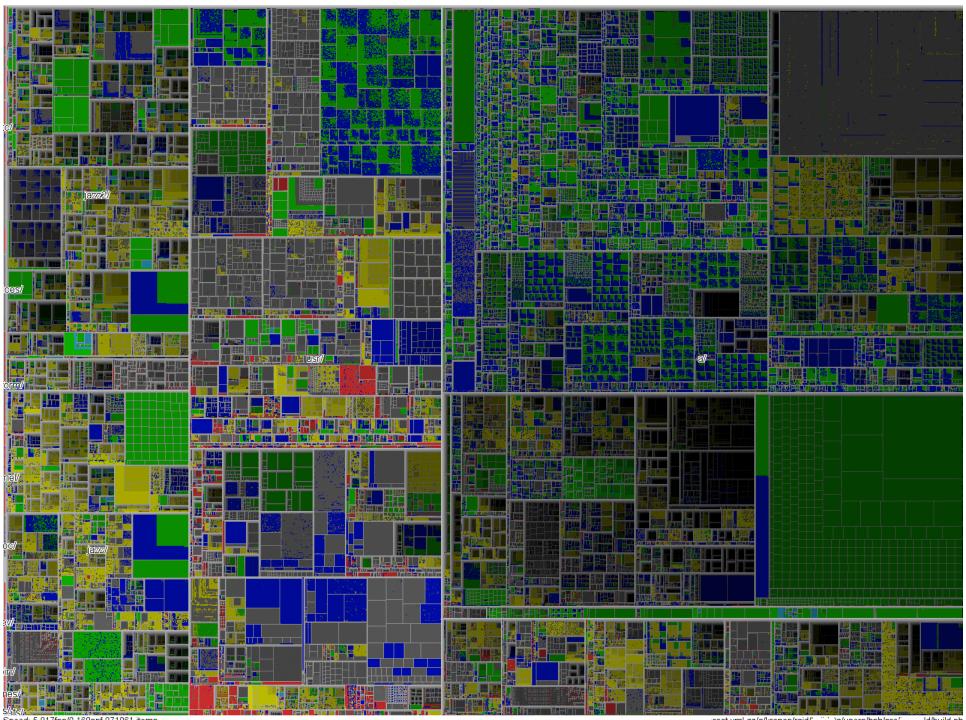


Treemap: Encyclopedia of Life (1.8M species)



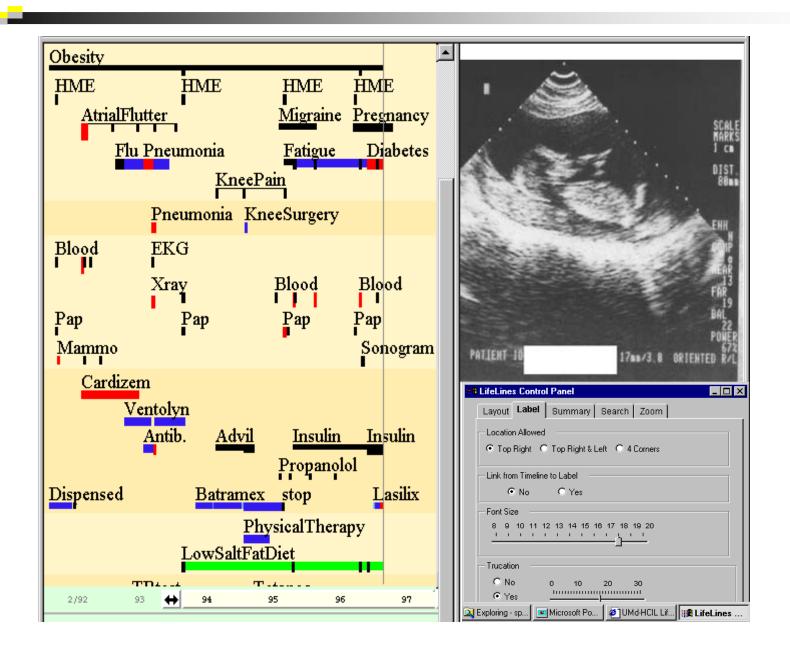
http://eol.org





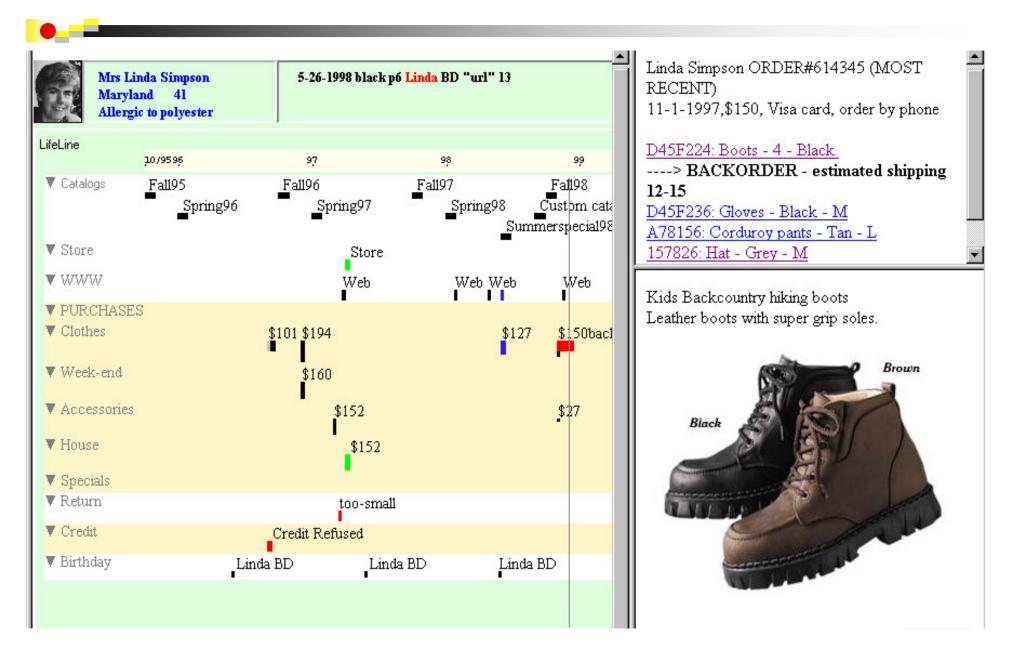
Speed: 5.917fps/0.169spf 971061 items :root.xml.gz/a/kronos/raid/ :a/users/hsh/src/ ild/build.sh

LifeLines: Patient Histories



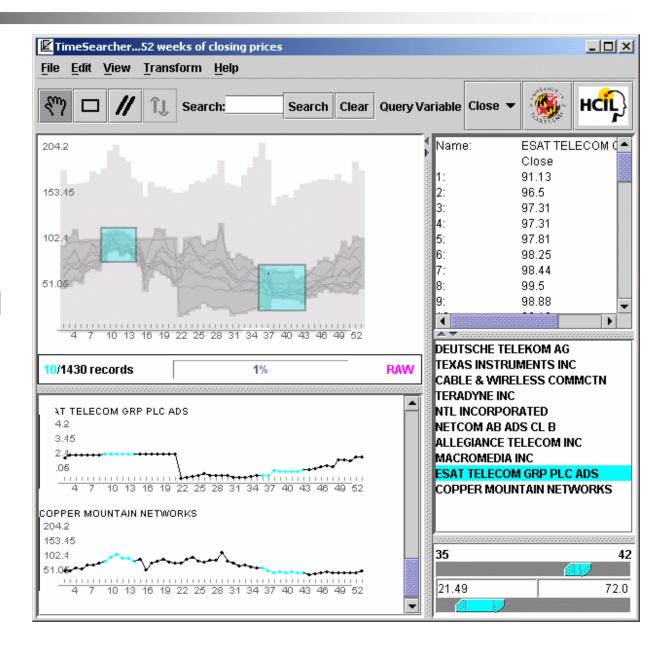


LifeLines: Customer Histories



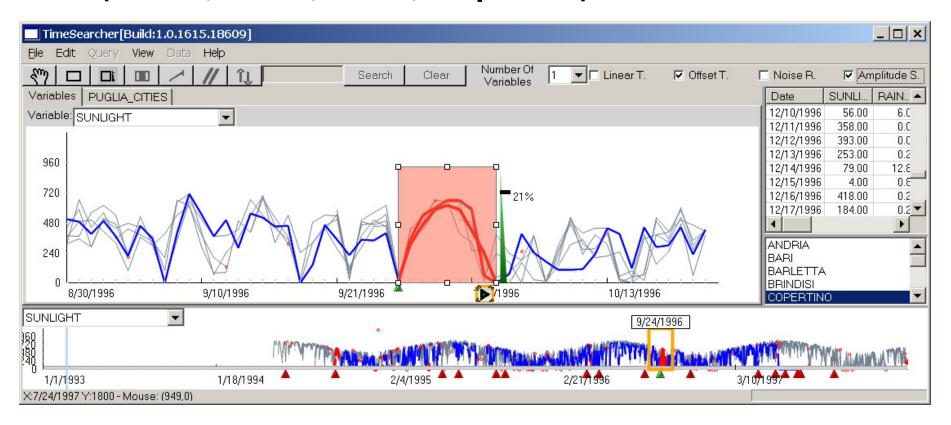
Temporal Data: TimeSearcher 1.3

- Time series
 - Stocks
 - Weather
 - Genes
- User-specified patterns
- Rapid search

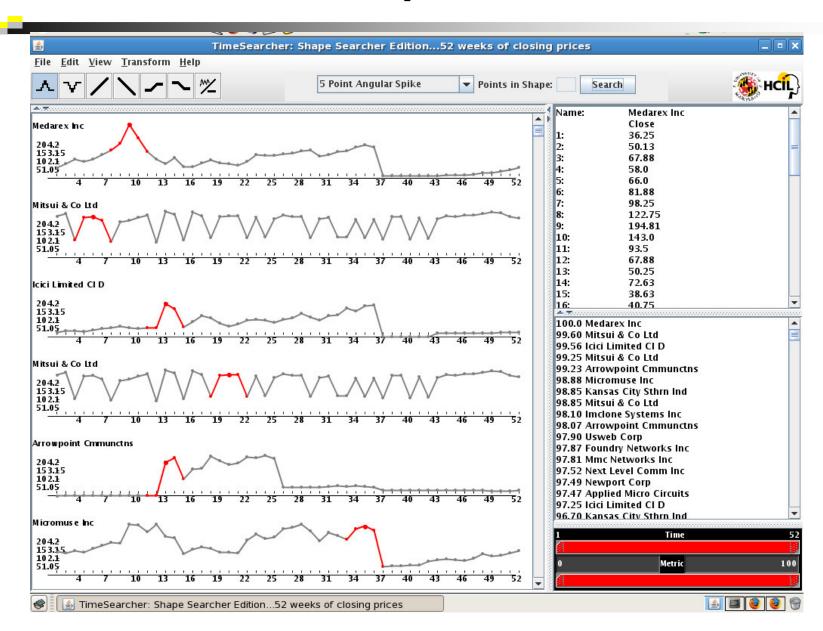


Temporal Data: TimeSearcher 2.0

- Long Time series (>10,000 time points)
- Multiple variables
- Controlled precision in match (Linear, offset, noise, amplitude)

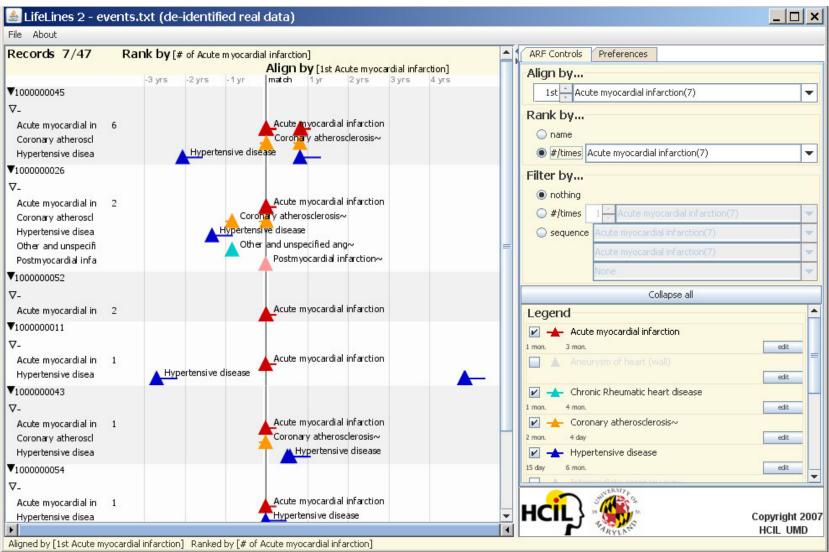


TimeSearcher: Shape Searcher Edition



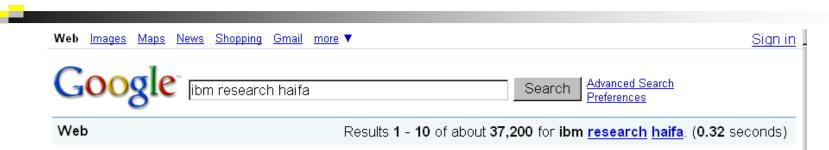


LifeLines2: Align-Rank-Filter





Search: Task Analysis



IBM Research Lab in Haifa

Perform **research** and development such as VLSI design, verification technology, storage subsystems, computer systems, programming languages and environments, ... www.**haifa**.il.**ibm**.com/ - 12k - Cached - Similar pages

Careers at IBM Haifa LabsResearch CareersHaifa Research LabRehovot SiteHaifa SiteIBM homeLeadership SeminarsMy Account

More results from ibm.com »

IBM Research | IBM Haifa Research Lab | Formal Verification and ...

IBM Haifa Research Lab | Formal Verification and Testing Technologies. ... IBM presents RuleBase PE at the annual DAC conferences and other industry events ... www.haifa.il.ibm.com/projects/verification/Formal_Methods-Home/index.html - 16k - Cached - Similar pages

Haifa Verification Conference 2007

The event will take place at the **IBM Research** Lab in **Haifa**, located on the University of **Haifa** campus, Mount Carmel, **Haifa**, Israel. ... www.**haifa.ibm**.com/Workshops/verification2007/ - 19k - Cached - Similar pages

IBM R&D Labs in Israel | IBM Haifa Research Lab

The **IBM Research** Lab in **Haifa** (HRL) has conducted decades of **research** that have been vital to **IBM's** success. R&D projects are being executed today in areas ... www.haifa.ibm.com/research.html - 11k - <u>Cached</u> - <u>Similar pages</u>
More results from www.haifa.ibm.com »



Search: Task Analysis



Sign in



on what day was Barack Obama born?

Search

Advanced Search Preferences

Web Video

Results 1 - 10 of about 4,400,000 for on what day was Barack Obama born?. (0.24 seconds)

Barack Obama - Wikipedia, the free encyclopedia

Obama was born on August 4, 1961 in Honolulu, Hawaii to Barack Hussein Obama,
[103] On October 2, 2002, the day Bush and Congress agreed on the joint ...
en.wikipedia.org/wiki/Barack Obama - 467k - Cached - Similar pages

The Biography Channel - Born on this Day

SIGN UP FOR EMAIL: Start your day with Born On This Day! ... Election 2008: Barack Obama » Election 2008: Hillary Clinton » Election 2008: Rudy Guiliani ... www.biography.com/search/botd.do - 33k - Cached - Similar pages

Barack Obama Biography - Biography.com

Barack Hussein Obama was born Aug. 4, 1961, in Honolulu, Hawaii. His father, Barack Obama, Sr., was born of Luo ethnicity in Nyanza Province, Kenya.... www.biography.com/search/article.do?id=12782369 - 34k - Cached - Similar pages



YouTube - Barack Obama: South Carolina Victory Speech

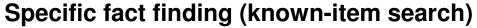
On election day we need foot canvassers. This means sign up ...

Watch video - 17 min - ****
www.youtube.com/watch?v=i8mG5qfDXL4

Posted Stories: BARACK OBAMA WAS BORN ON THIS DAY

BARACK OBAMA WAS BORN ON THIS DAY. 4 August 1961. Barack Hussein Obama is a United States Senator from Illinois and a leading candidate for the Democratic ... postedstories.blogspot.com/2008/02/barack-obama-was-born-on-this-day.html - 99k - Cached - Similar pages





On what day was Barack Obama born?

<Google succeeds>



Specific fact finding (known-item search)

On what day was Barack Obama born?

<Google succeeds>

Extended fact finding (vague query)

What cities did Barack Obama live in since he became a Senator?

Exploration of availability (vague result request)

What genealogical information on Barack Obama is at the National Archives?



Specific fact finding (known-item search)

On what day was Barack Obama born?

<Google succeeds>

Extended fact finding (vague query)

What cities did Barack Obama live in since he became a Senator?

Exploration of availability (vague result request)

What genealogical information on Barack Obama is at the National Archives?

Open-ended browsing and problem analysis (hidden assumptions)

How has Barack Obama's position on the environment changed since 2001?

Mismatch with metadata (requires exhaustive search)

How has Barack Obama's choice of clothing changed during his campaign?



known-item search

vague query

vague result request

hidden assumptions

requires exhaustive search



1-minute

known-item search

vague query

Weeks

&

Months

vague result request

hidden assumptions

requires exhaustive search



Discovery Process: Design Challenges

Enrich query formulation

Expand result management

Enable long-term effort

Enhance collaboration

Deal with special cases:

- Comprehensive: Legal, patent, medical
- Negation: Patent, pharma
- Exceptions: Outlier documents
- Bridges: Connecting



Discovery Process: Design Challenges

Enrich query formulation

Expand result management

Enable long-term effort

Enhance collaboration



Enrich query formulation

Expand result management

Enable long-term effort

Enhance collaboration







Expand result management

Enable long-term effort

Enhance collaboration

Previous & Similar Structured input Spell check Query previews

Finding Aids

Limit:

Time

Geography

Language

Sources

Media

- - -



Enrich query formulation

Expand result management

Enable long-term effort

Enhance collaboration

Ranking (Social, ...)
Snippets + Handles
Highlighting

Clustering:
Dynamic generation
Meaningful & Stable

Visualization: By attributes

Summarization



Enrich query formulation

Expand result management

Enable long-term effort



Enhance collaboration

Save & Replay
Edit & Grab
Mark & Annotate
Compare

Notebooks

History-keeping

Systematic Yet Flexible Strategies



Enrich query formulation

Expand result management

Enable long-term effort

Enhance collaboration



Send email
Announce, publish...
Export: web, slides...
Record video

Comment, blog... Wiki Share screen

Tagging & voting Feedback to search



Discovery Process: Design Challenges

Enrich query formulation

Expand result management

Enable long-term effort

Enhance collaboration

Deal with special cases:

- Comprehensive: Legal, patent, medical

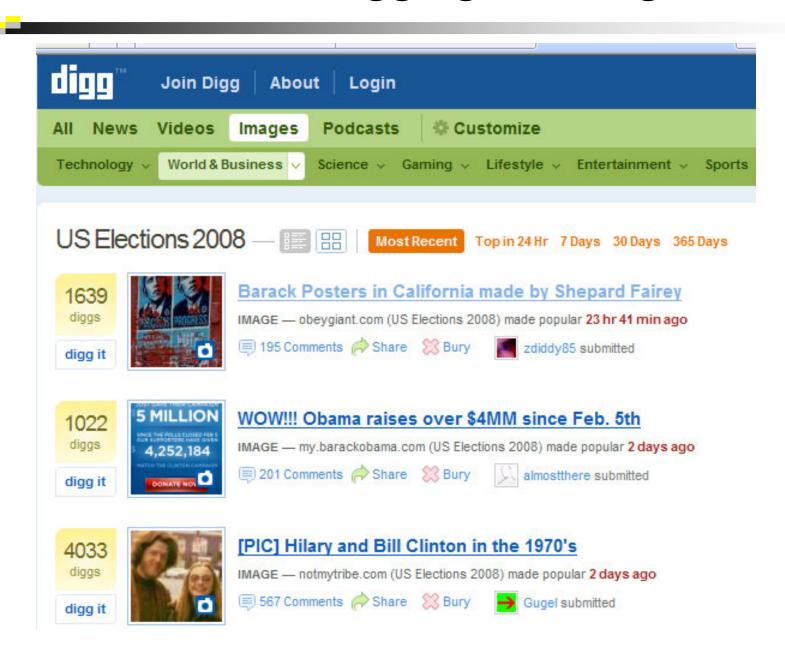
- Negation: Patent, pharma

- Exceptions: Outlier documents

- Bridges: Connecting



Collaboration: Tagging & voting





Treemap: Digg Overview

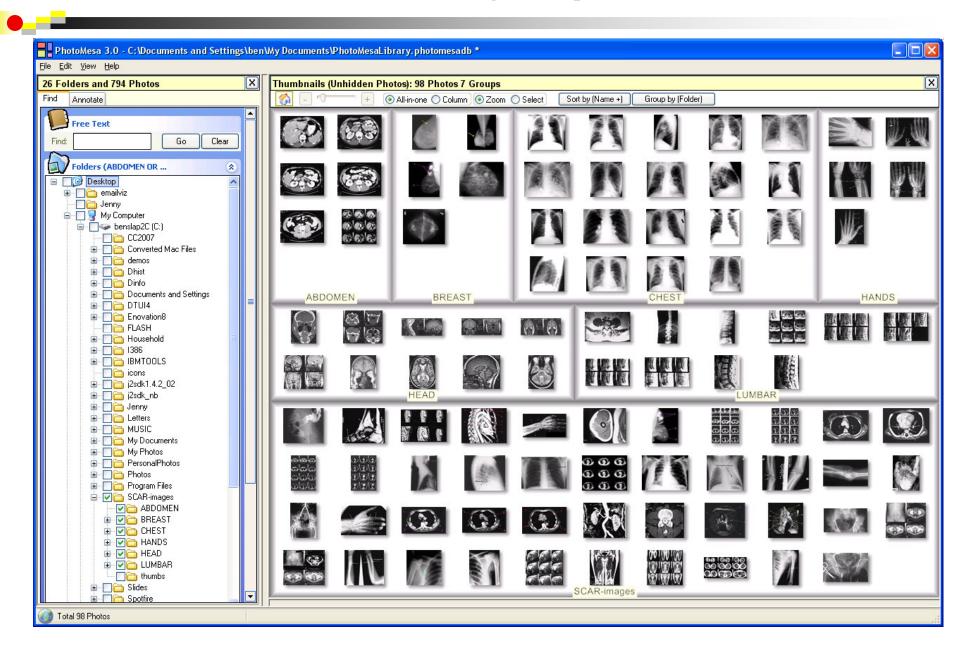


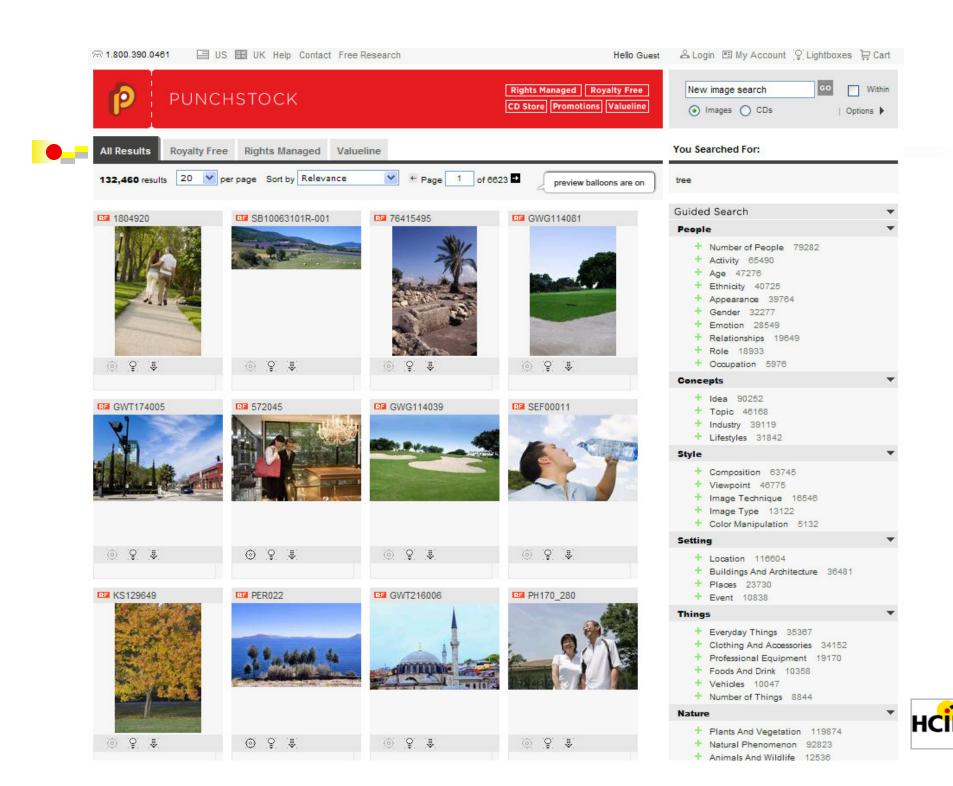
Treemap: Newsmap



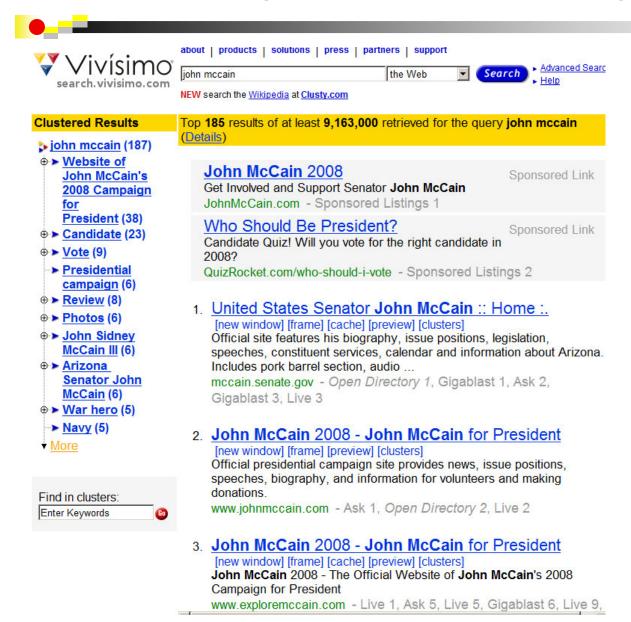


PhotoMesa: Zooming Exploration



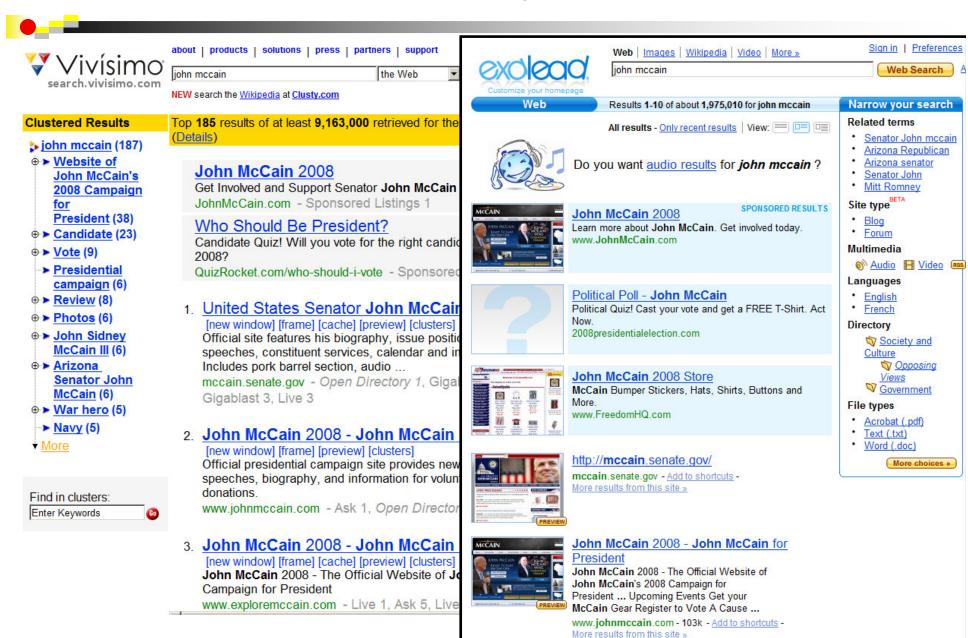


Discovery Process: Strategies





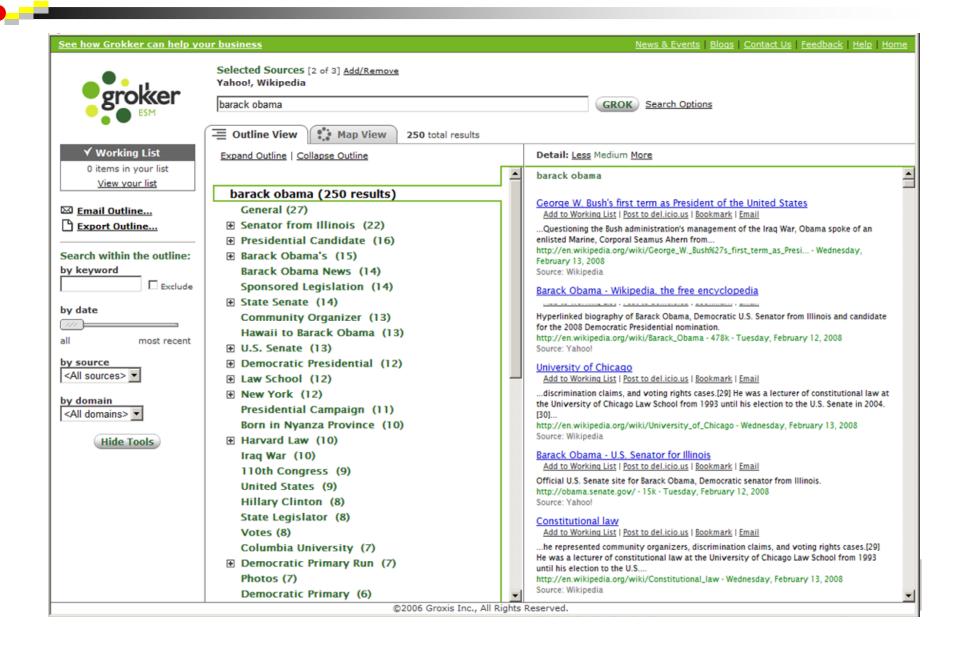
Vivisimo & Exalead: Dynamic overviews



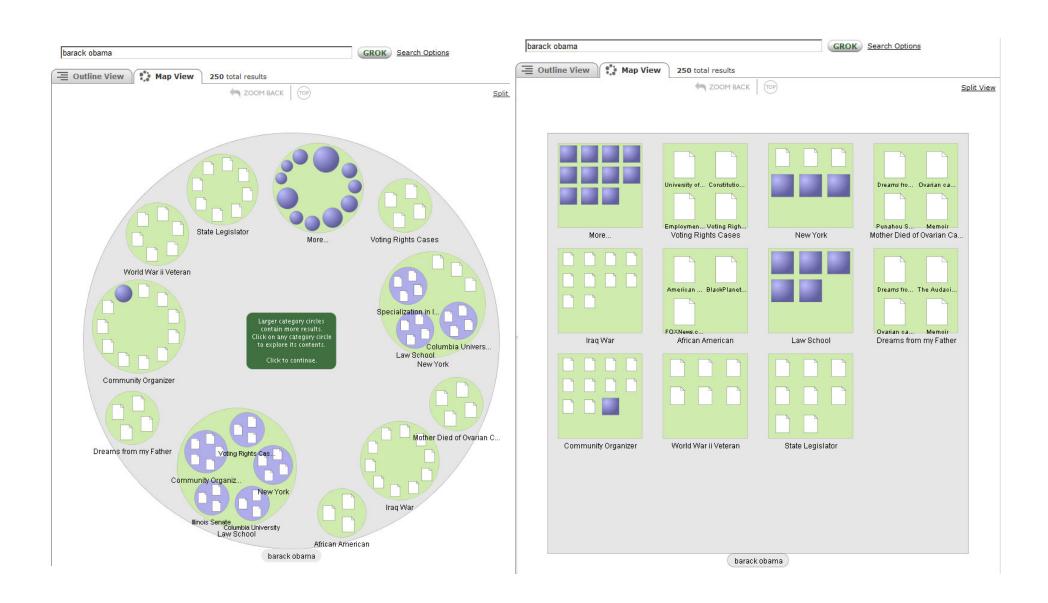
SERVICE: Categorized Overviews



Grokker: Overviews & Maps



Grokker: Overviews & Maps

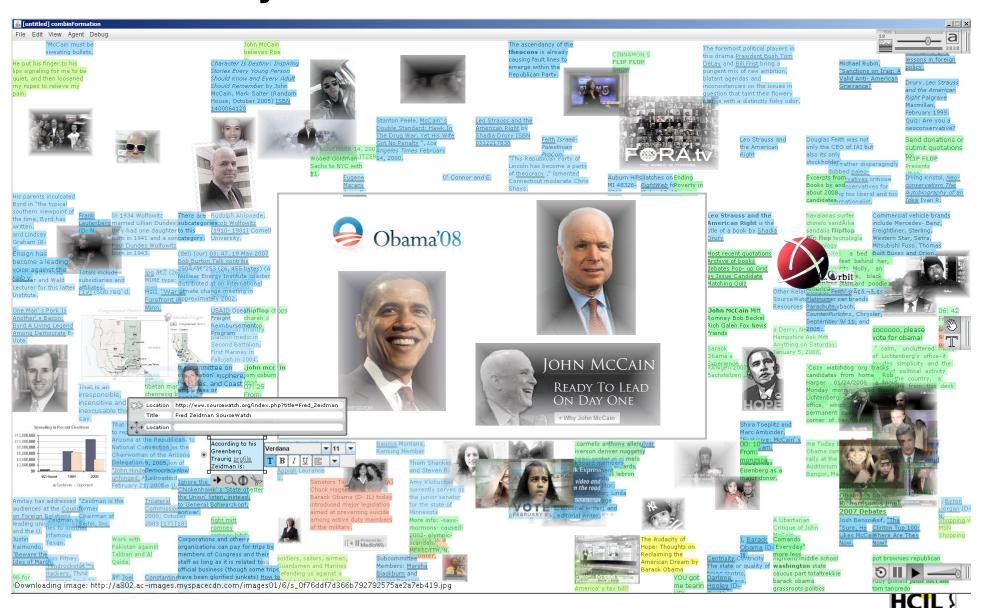


Discovery Process: Combinformation





Discovery Process: Combinformation



Discovery Process: Design Challenges

Enrich query formulation

Expand result management

Enable long-term effort

Enhance collaboration

Deal with special cases:

- Comprehensive: Legal, patent, medical

- Negation: Patent, pharma

- Exceptions: Outlier documents

- Bridges: Connecting



Research Methods

Controlled experiments

Repeated interventions

Logging usage patterns

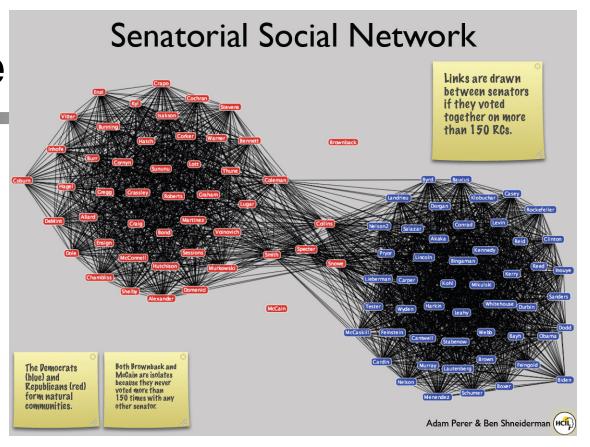
Multi-dimensional In-Depth Long-Term Case Studies (MILCs)

Domain experts doing their own work for weeks & months



MILC example

Evaluate SocialAction



- Focused on integrating statistics & visualization
- 4 case studies, 4-8 weeks (journalist, bibliometrician, terrorist analyst, organizational analyst)
- Identified desired features, gave strong positive feedback about benefits of integration

(Perer & Shneiderman, CHI2008)

Science 1.0

- Reductionist
- Controlled Experiments
- Laboratory
- Natural World



Science 1.0 + Science 2.0

- Natural World → Made World



Science 1.0 + Science 2.0

- Natural World → Made World

- Hypothesis Testing → Hypothesis Testing
- Predictive Theories -> Predictive Theories
- Replications → Replications







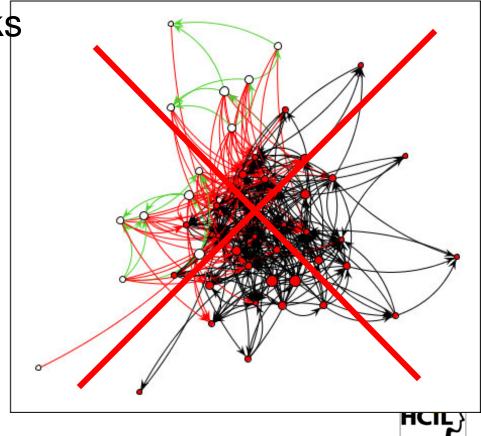
26th Annual Symposium May 28-29, 2009

www.cs.umd.edu/hcil



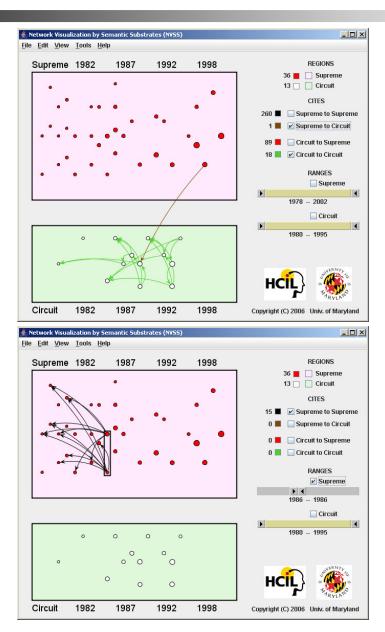
Network Data

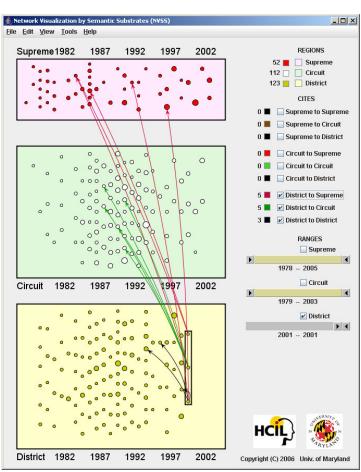
- Nodes & Links
 - Relationships & communication
 - Scientific/legal citations
- Difficult to complete tasks
 - Occlusion
 - Complexity



Network Visualization with Semantic Substrates

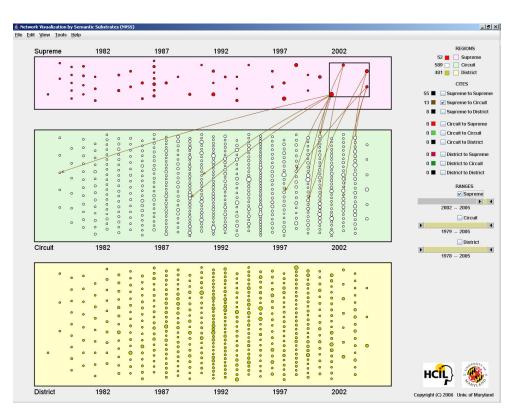
- Meaningful layout of nodes
- User controlled visibility of links

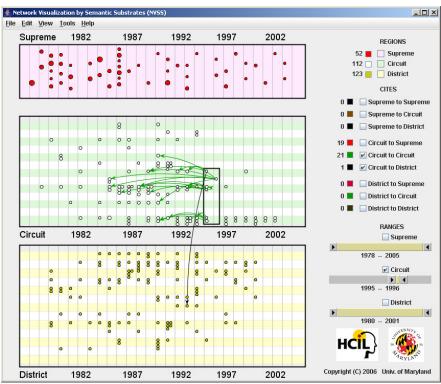






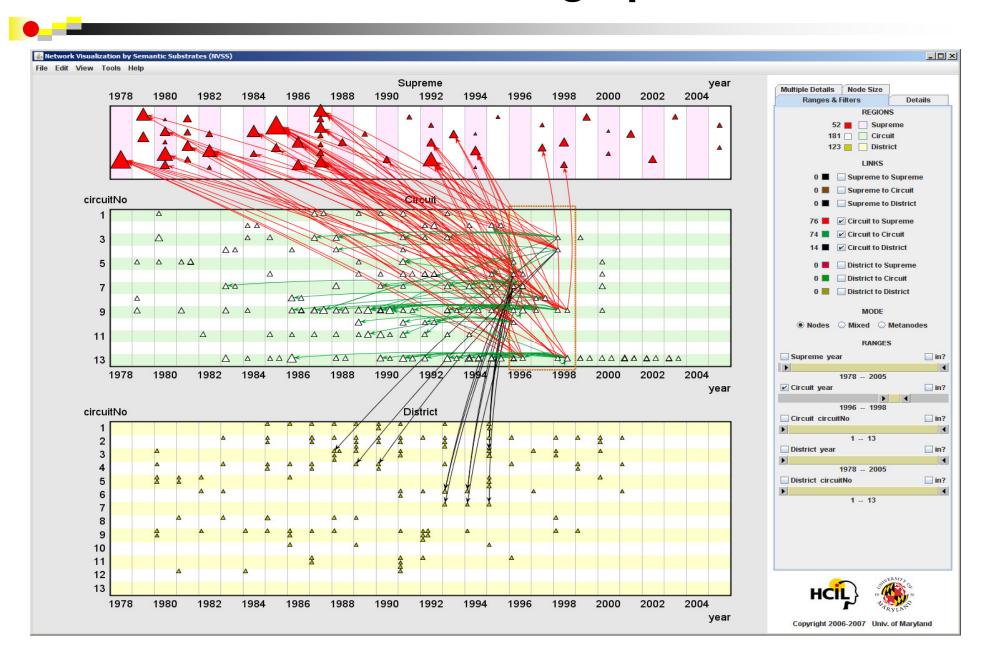
Network Visualization with Semantic Substrates





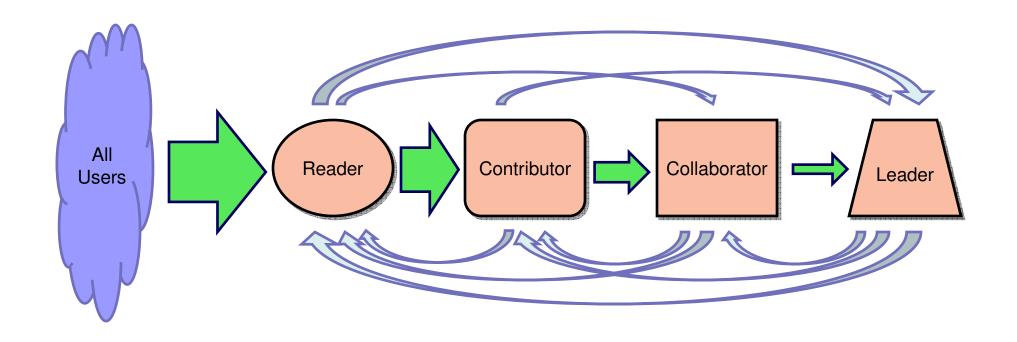


NVSS: Citation Historiographs



The Reader-to-Leader Framework:

Motivating Technology-Mediated Social Participation





Research Methods

Controlled experiments

Repeated interventions

Logging usage patterns

Multi-dimensional In-Depth Long-Term Case Studies (MILCs)

Domain experts doing their own work for weeks & months



Case Study Methodology

- 1) Interview (1 hr)
- 2) Training (2 hr)
- 3) Early Use (2-4 weeks)
- 4) Mature Use (2-4 weeks)
- 5) Outcome (1 hr)







26th Annual Symposium May 28-29, 2009

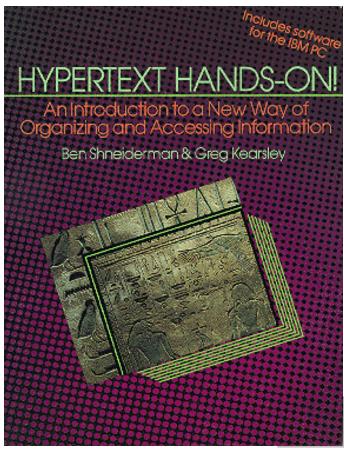
www.cs.umd.edu/hcil



World-Wide Web links

- Developed & tested embedded menus (1983) that became "hot spots" (Berners-Lee, 1989)
- First electronic book (1989):
 Hypertext-Hands On





Hypertext on Hypertext – CACM July'88



Special Section of the 1919, 1988 Communications of the AGM

Prepared by the Human-Computer Interaction Laboratory, University of Maryland, College Park, Maryland.

© Association for Computing Machinery, Inc., 1988. ACM order number 217880

Hyperties[™] © Cognetics Corporation and the University of Maryland. 55 Princeton-Hightstown Road, Princeton Jct., NJ 08550

EClick mouse button to start (or press '?' for instructions)]



