

Sept. 17, 2008

# Visualization for Information Analysis and Exploration

John Stasko

Information Interfaces Research Group  
School of Interactive Computing  
Georgia Institute of Technology



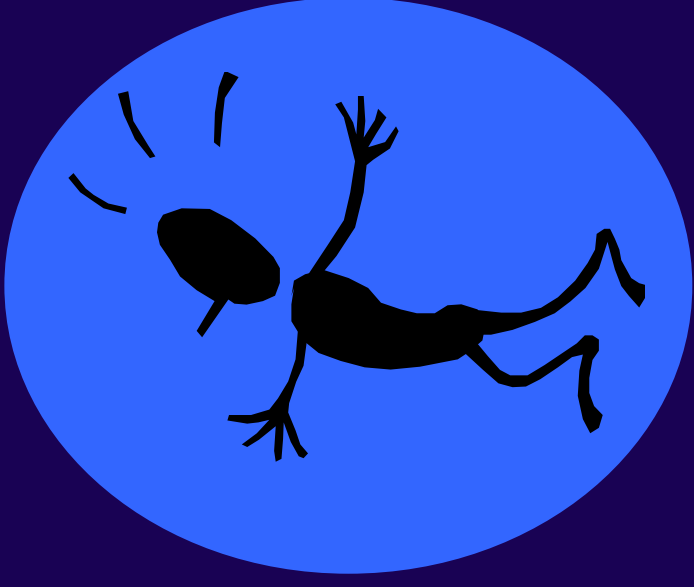
# Exercise

- Get out pencil & paper



# Data Explosion

- Society is more complex
  - There simply is more “stuff”
- Computers, internet and web give people access to an incredible amount of data
  - news, sports, financial, purchases, etc...



# Data Overload

- Confound: How to make use of the data
  - How do we make sense of the data?
  - How do we harness this data in decision-making processes?
  - How do we avoid being overwhelmed?

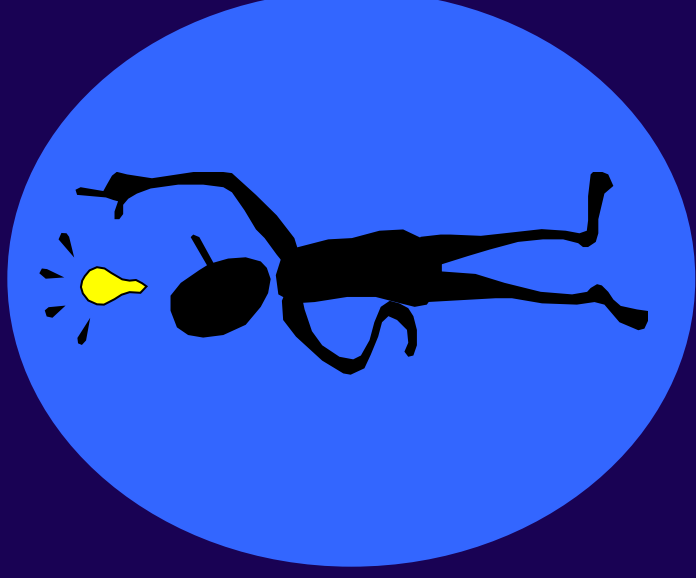


© 2004 E-TRAIL SCOTTADAMS@aol.com



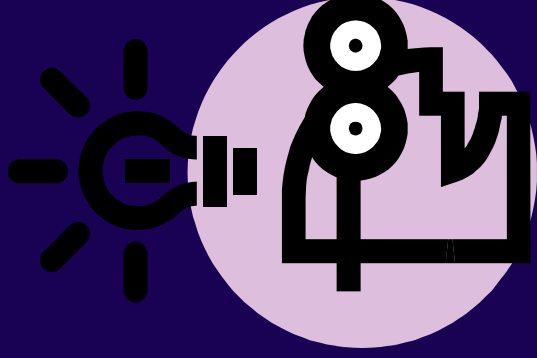
# The Challenge

- Transform the *data* into *information* (understanding, insight) thus making it useful to people



## Premise of my Work

- Visualization of data helps people understand it better



# Human Vision

- Highest bandwidth sense
  - ~100 MB/s
  - Parallel
  - Strong pattern recognition
  - Much done preattentively, ie, without thought



# Visualization

- Definition
  - “The use of computer-supported, interactive visual representations of data to amplify cognition.”
  - From [Card, Mackinlay Shneiderman '98]





# Visualization

- Often thought of as process of creating a graphic or an image
- Really is a cognitive process
  - Form a mental image of something
  - Internalize an understanding
- “The purpose of visualization is insight, not pictures”
  - Insight: discovery, decision making, explanation, analysis, exploration, learning



## Main Idea

- Visuals help us think
  - Provide a frame of reference, a temporary storage area
- Cognition → Perception
- Pattern matching
- External cognition aid
  - Role of external world in thinking and reason

Larkin & Simon '87  
Card, Mackinlay, Shneiderman '98



## When to Apply?

- Many other techniques for data analysis
  - Data mining, DB queries, machine learning...
- Visualization most useful in **exploratory data analysis**
  - Don't know what you're looking for
  - Don't have a priori questions
  - Want to know what questions to ask



## Part of our Culture

- "I see what you're saying"
- "Seeing is believing"
- "A picture is worth a thousand words"



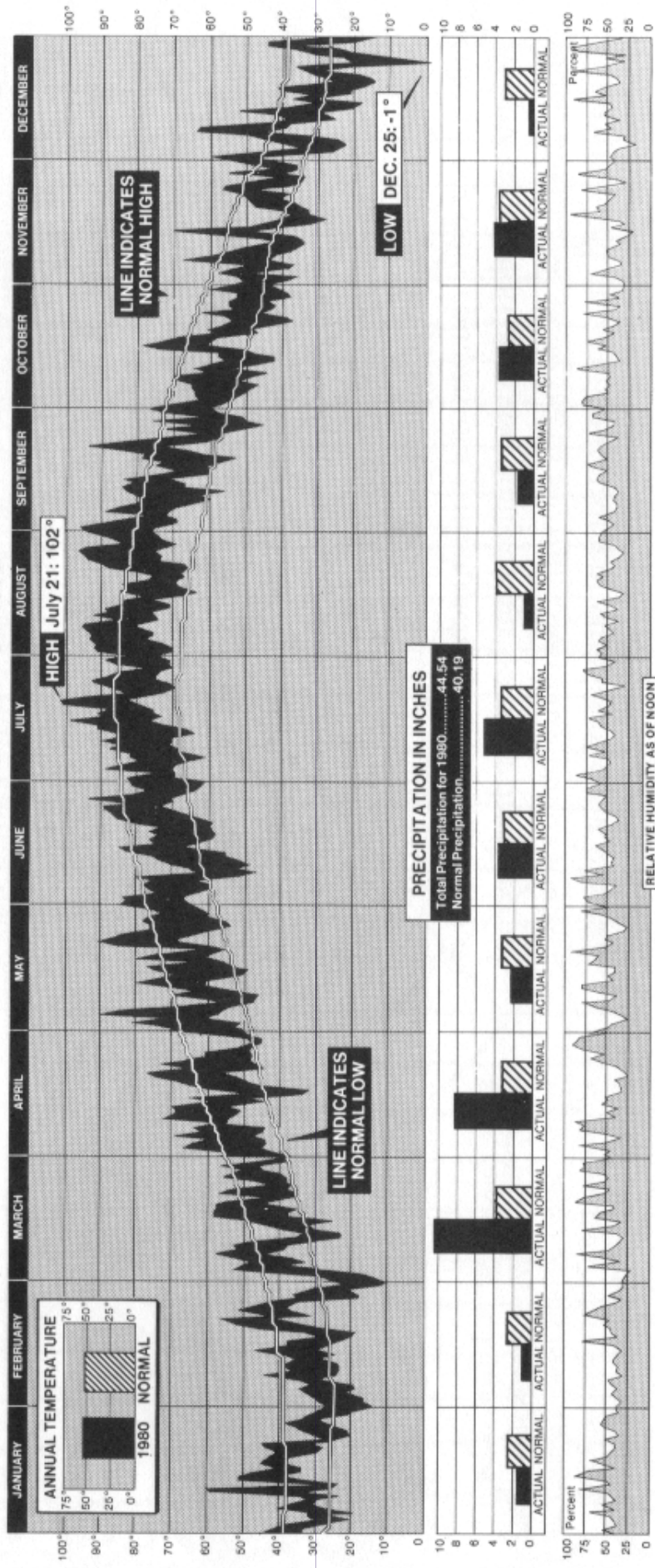
# Some quick (static) examples...



# NYC Weather

2220 numbers

## NEW YORK CITY'S WEATHER FOR 1980



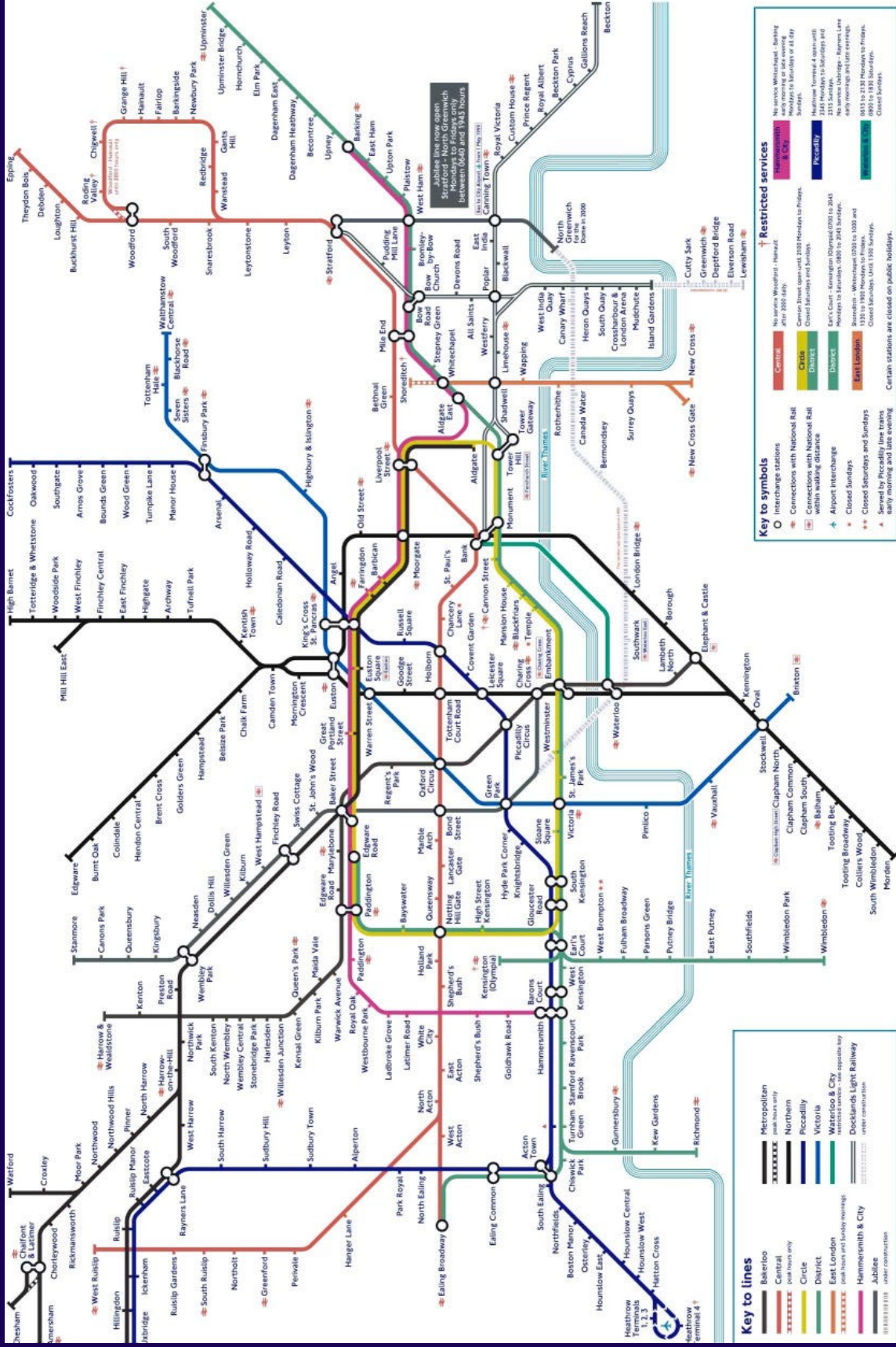
New York Times, January 11, 1981, p. 32.

E. Tufte, *Visual Display of Quant Info*

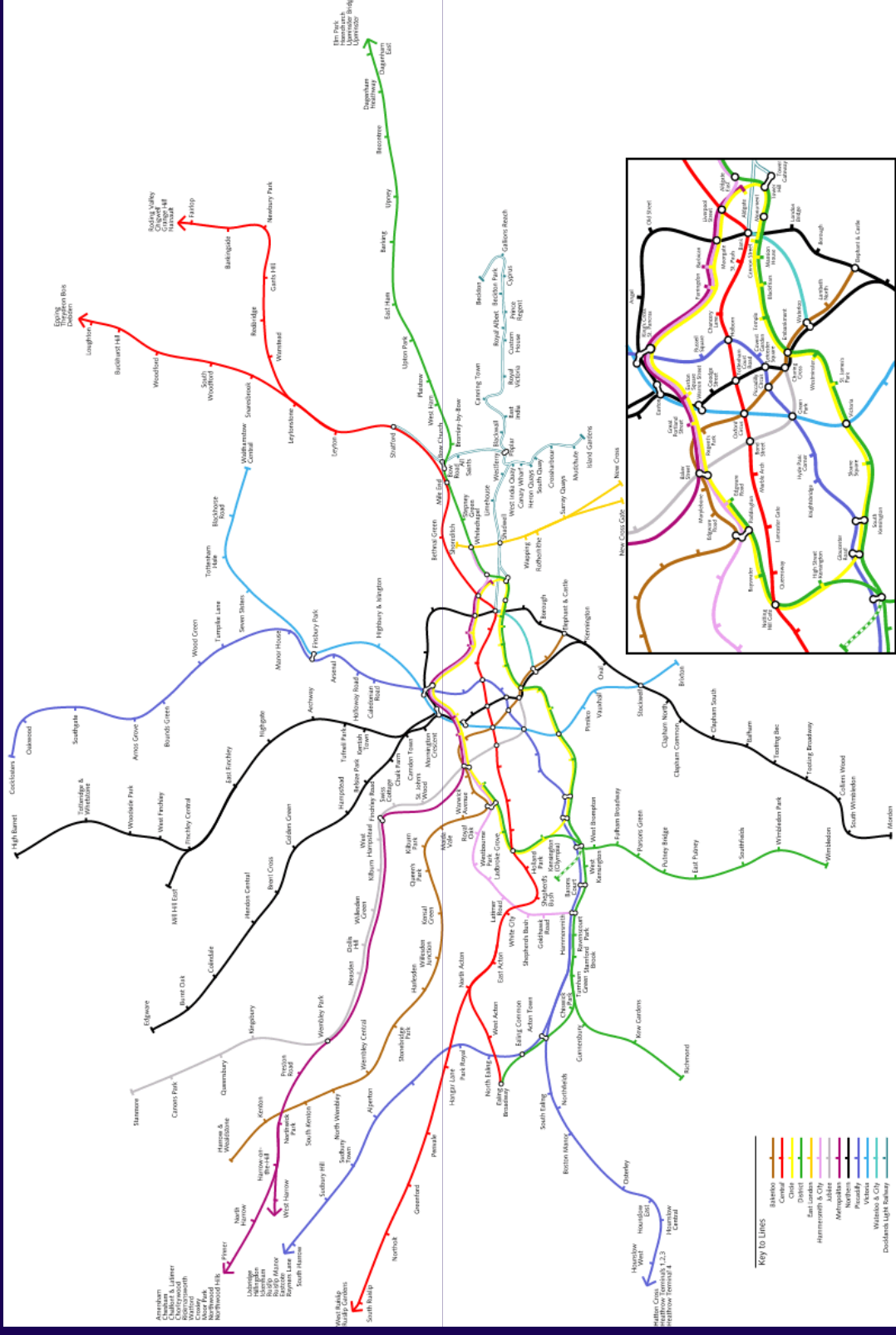


# London Subway

www.thetube.com



# True Geography

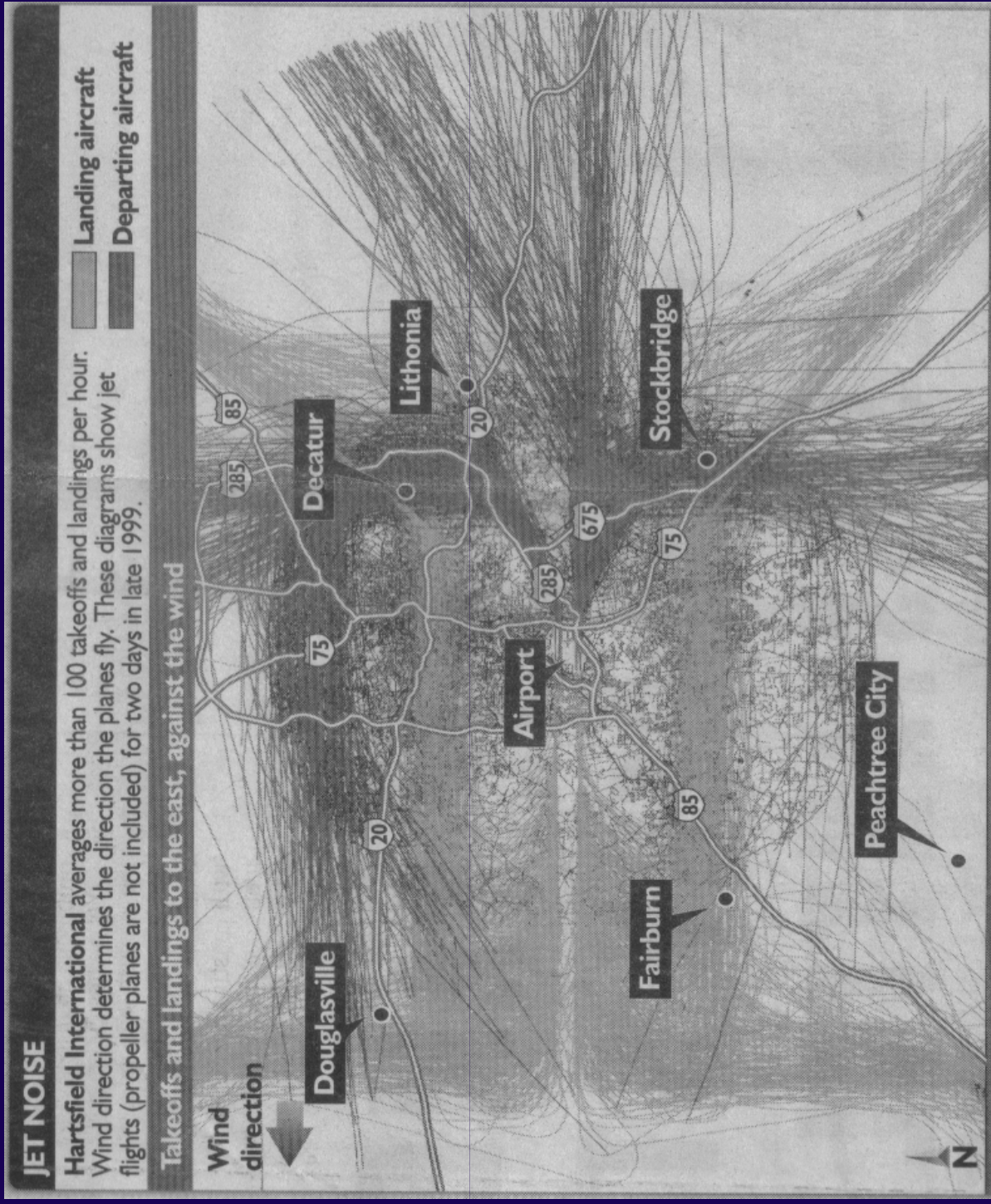




# Easy Walking Lines Added



# Atlanta Flight Traffic



Atlanta Journal  
April 30, 2000



## Casual Information Visualization: Depictions of Data in Everyday Life

Zachary Pousman, John T. Stasko, *Member, IEEE*, and Michael Mateas

**Abstract**—Information visualization has often focused on providing deep insight for expert user populations and on techniques for amplifying cognition through complicated interactive visual models. This paper proposes a new subdomain for infovis research that complements the focus on analytic tasks and expert use. Instead of work-related and analytically driven infovis, we propose Casual Information Visualization (or Casual Infovis) as a complement to more traditional infovis domains. Traditional infovis systems, techniques, and methods do not easily lend themselves to the broad range of user populations, from expert to novices, or from work tasks to more everyday situations. We propose definitions, perspectives, and research directions for further investigations of this emerging subfield. These perspectives build from ambient information visualization [32], social visualization, and also from artistic work that visualizes information [41]. We seek to provide a perspective on infovis that integrates these research agendas under a coherent vocabulary and framework for design. We enumerate the following contributions. First, we demonstrate how blurry the boundary of infovis is by examining systems that exhibit many of the putative properties of infovis systems, but perhaps would not be considered so. Second, we explore the notion of insight and how, instead of a monolithic definition of insight, there may be multiple types, each with particular characteristics. Third, we discuss design challenges for systems intended for casual audiences. Finally we conclude with challenges for system evaluation in this emerging subfield.

**Index Terms**—Casual information visualization, ambient infovis, social infovis, editorial, design, evaluation.

### 1 INTRODUCTION

Much of the work in information visualization assumes a population of expert users who have knowledge and experience in analyzing problems in specific domains. Workers in widely varying domains from finance to government to journalism use information visualization tools to explore data, generate, refine and test hypotheses, and ultimately to produce insight. This user population of information workers and

Are these types of tools really infovis systems? The question arises, where are the limits of infovis with respect to the everyday uses of computational artifacts.

Card, Mackinlay, and Shneiderman define information visualization broadly: infovis is the use of computers to interactively amplify cognition, using visual representations [10]. Therefore if we take this as our definition, systems must be computer-based, interactive, pro-



# Reinforce my point with two examples

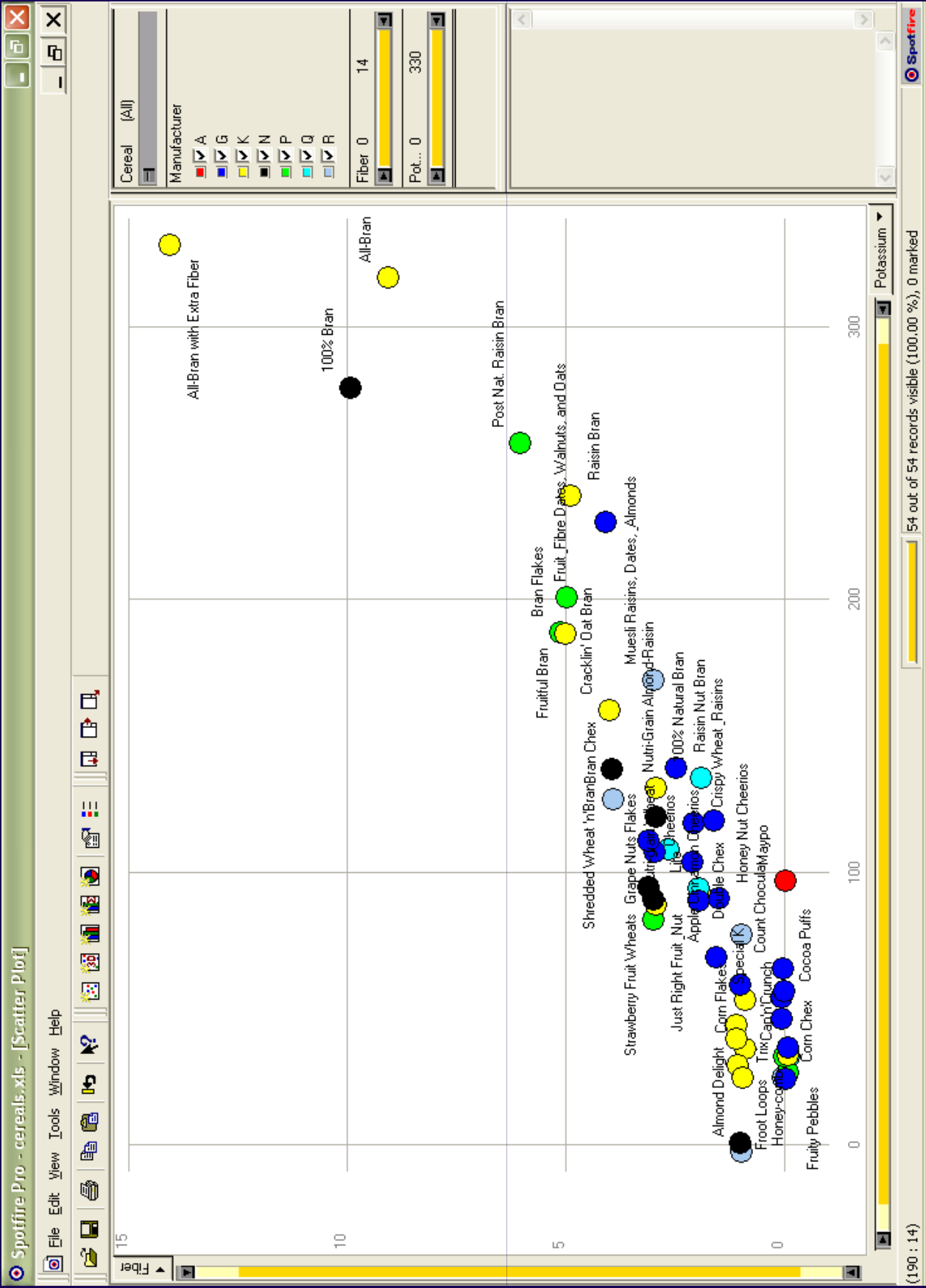


## Questions:

- Which cereal has the most/least potassium?
- Is there a relationship between potassium and fiber?
- If so, are there any outliers?
- Which manufacturer makes the healthiest cereals?

A	B	C	D
Cereal	Manufacturer	Fiber	Potassium
1	Cereal		
2	100% Bran	N	10
3	100% Natural Bran	Q	2
4	All-Bran	K	9
5	All-Bran with Extra Fiber	K	14
6	Almond Delight	R	1
7	Apple Cinnamon Cheerios	G	1.5
8	Bran Chex	R	4
9	Bran Flakes	P	5
10	Cap'n Crunch	Q	0
11	Cheerios	G	2
12	Cocoa Puffs	G	0
13	Corn Chex	R	0
14	Corn Flakes	K	1
15	Count Chocula	G	0
16	Cracklin' Oat Bran	K	4
17	Cream of Wheat (Quick)	N	1
18	Crispy Wheat & Raisins	G	2
19	Double Chex	R	1
20	Froot Loops	K	1
21	Frosted Flakes	K	1
22	Fruit & Fibre Dates, Wal	P	5
23	Fruitful Bran	K	5
24	Fruity Pebbles	P	0
25	Golden Grahams	G	0
26	Grape Nuts Flakes	P	3
27	Honey Nut Cheerios	G	1.5
28	Honey-comb	P	0
29	Just Right Fruit & Nut	K	2
30	Life	Q	2
31	Lucky Charms	G	0
32	Maypo	A	0
33	Muesli Raisins, Dates, &	R	3
34	Multi-Grain Cheerios	G	2
35	Multi-Grain Almond-Rais	K	3
36	Nutri-grain Wheat	K	3
37	Oatmeal Raisin Crisp	G	1.5
38	Post Nat. Raisin Bran	P	6
39	Product 19	K	1
40	Quaker Oatmeal	Q	2.7
41	Raisin Bran	K	5
42	Raisin Nut Bran	G	2.5
43	Rice Krispies	K	0
44	Shredded Wheat	N	3
45	Shredded Wheat 'n Bran	N	4
46	Shredded Wheat spoon	N	3
47	Smacks	K	1
48	Special K	K	1
49	Strawberry Fruit Wheats	N	3
50	Total Corn Flakes	G	0
51	Total Raisin Bran	G	4
52	Total Whole Grain	G	3
53	Trix	G	0
54	Wheaties	G	3
55	Wheaties Honey Gold	G	1





Fiber



# Potassium

## Even Tougher?

- What if you could only see one cereal's data at a time? (e.g. some websites)
- What if I read the data to you?



# Four Data Sets

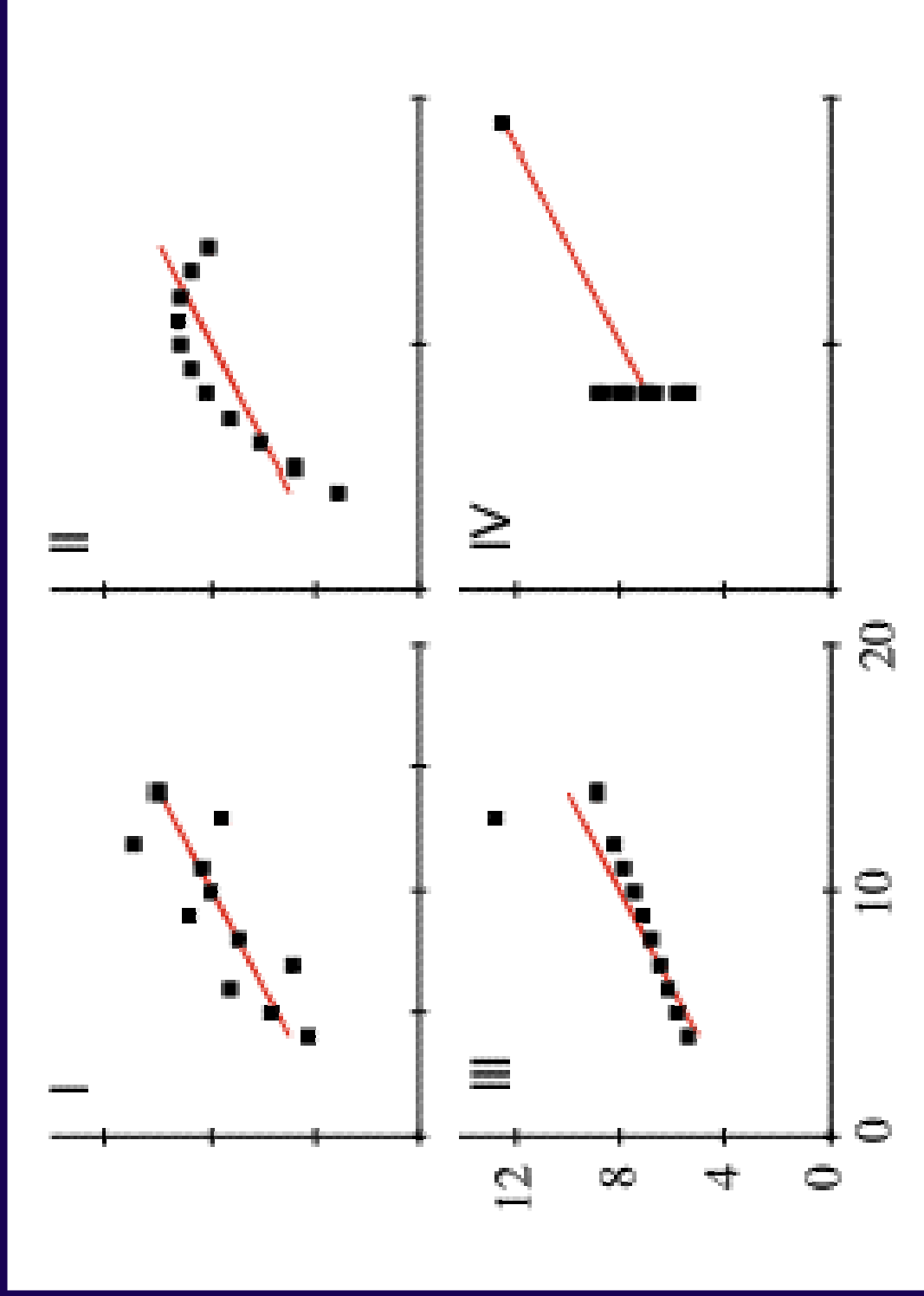
- Mean of the x values = 9.0
- Mean of the y values = 7.5
- Equation of the least-squared regression line is:  $y = 3 + 0.5x$
- Sums of squared errors (about the mean) = 110.0
- Regression sums of squared errors (variance accounted for by x) = 27.5
- Residual sums of squared errors (about the regression line) = 13.75
- Correlation coefficient = 0.82
- Coefficient of determination = 0.67

<http://astro.swarthmore.edu/astro121/anscombe.html>





# The Data Sets



# The Values

1	2	3	4
10.0, 8.04	10.0,9.14	10.0, 7.46	8.0, 6.58
8.0, 6.95	8.0,8.14	8.0, 6.77	8.0, 5.76
13.0, 7.58	13.0,8.74	13.0,12.74	8.0, 7.71
9.0, 8.81	9.0,8.77	9.0, 7.11	8.0, 8.84
11.0, 8.33	11.0,9.26	11.0, 7.81	8.0, 8.47
14.0, 9.96	14.0,8.10	14.0, 8.84	8.0, 7.04
6.0, 7.24	6.0,6.13	6.0, 6.08	8.0, 5.25
4.0, 4.26	4.0,3.10	4.0, 5.39	19.0,12.50
12.0,10.84	12.0,9.13	12.0, 8.15	8.0, 5.56
7.0, 4.82	7.0,7.26	7.0, 6.42	8.0, 7.91
5.0, 5.68	5.0,4.74	5.0, 5.73	8.0, 6.89



# Revisit Starting Exercise

- What did you put on paper?



# Two Related Disciplines

- Information Visualization
- Visual Analytics



# Information Visualization

- Using interactive computer visualizations to represent and communicate abstract data
  - Statistics, databases, software, ...
- Area emerged approximately 1990



# Information Visualization

- Recent research trends
  - InfoVis for the Masses
  - Challenges of evaluation
  - Interaction is crucial



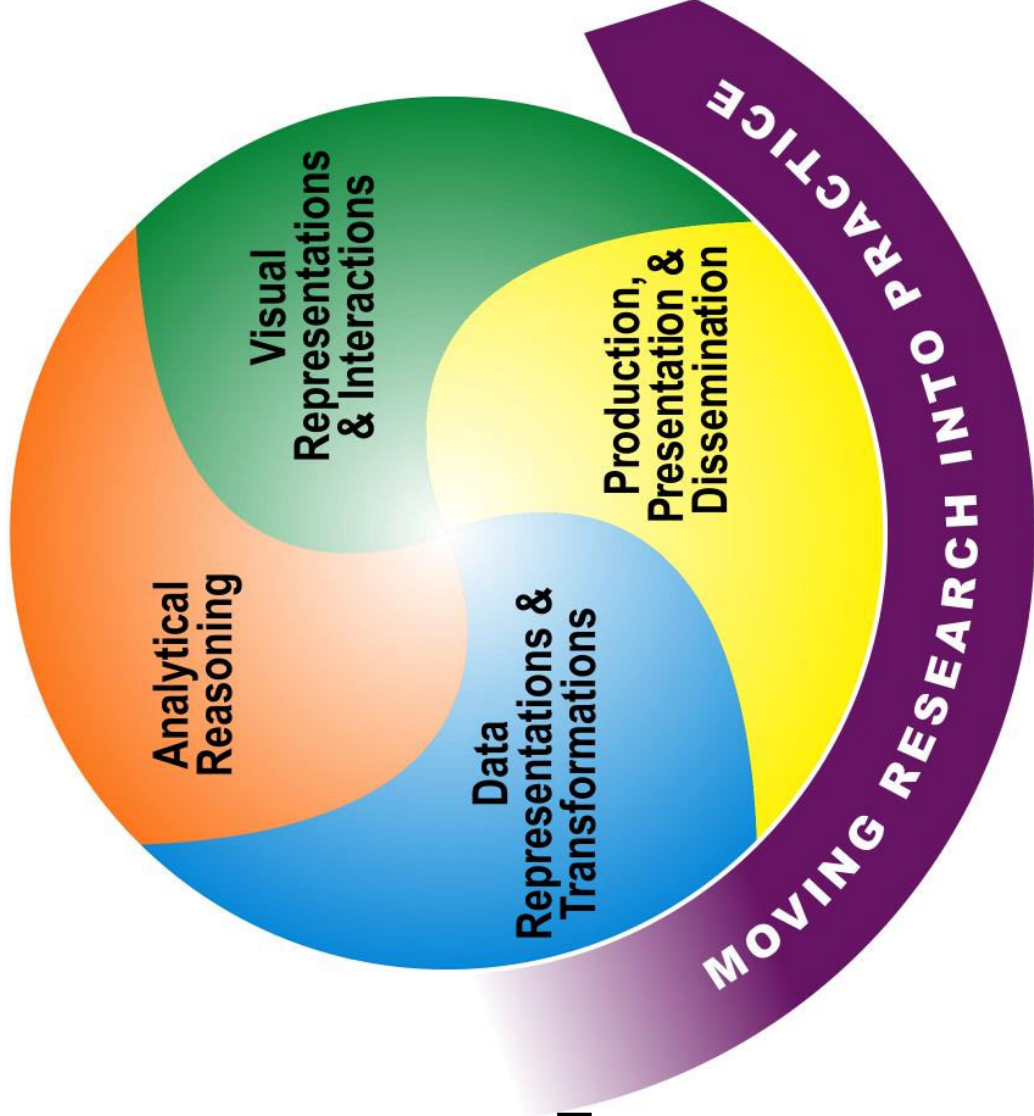
# Visual Analytics

- Informal: Using visual representations to help make decisions
- Formal: The science of analytical reasoning facilitated by interactive visual interfaces
- InfoVis++
- Area emerged approximately 2005



# Overview of the R&D Agenda

- **Challenges**
- **Science of Analytical Reasoning**
- **Science of Visual Representations and Interactions**
- **Data Representations and Transformations**
- **Production, Presentation, and Dissemination**
- **Moving Research Into Practice**
- **Positioning for an Enduring Success**





# Visual Analytics: Beyond InfoVis

- Statistics, data representation and statistical graphics
- Geospatial and Temporal Sciences
- Applied Mathematics
- Knowledge representation, management and discovery
- Ontology, semantics, NLP, extraction, synthesis, ...
- Cognitive and Perceptual Sciences
- Communications: Capture, Illustrate and present a message
- Decision sciences



# Academic Context

Visual  
Analytics  
~2005

Information  
Visualization  
~1990



# IEEE InfoVis

IEEE VisWeek 2008 - Windows Internet Explorer

http://vis.computer.org/VisWeek2008/InfoVis/

1331 Susquehanna 101.9 KUJ 98.9 KWJZ WGLT CC Webmail Indian Hills

IEEE VisWeek 2008

## InfoVis08

VIS • INFOVIS • VAST

- > Welcome
- >> Vis • InfoVis • VAST
- >> Week-at-a-Glance
- >> VisWeek Sessions
- >> Exhibition
- >> Registration
- >> Student Volunteers
- >> Call for Participation
- >> Committees
- > Presenter Information
- > Site Map
- > Contact Us

Computer-based information visualization centers around helping people explore or explain data through interactive software that exploits the capabilities of the human perceptual system. A key challenge in information visualization is designing a cognitively useful spatial mapping of a dataset that is not inherently spatial and accompanying the mapping by interaction techniques that allow people to intuitively explore the dataset. Information visualization draws on the intellectual history of several traditions, including computer graphics, human-computer interaction, cognitive psychology, semantics, graphic design, statistical graphics, cartography, and art. The synthesis of relevant ideas from these fields with new methods and techniques made possible by interactive computation are critical for the development of information visualization with the torrents of data confronting them.

For questions, please email [infovis@vis.computer.org](mailto:infovis@vis.computer.org).

### General Chair

Jarke van Wijk, Eindhoven University of Technology

>> **Daners Chairs**

## Welcome » IEEE VAST08

IEEE Symposium on Visual Analytics Science and Technology (IEEE VAST) founded in 2005, is the international symposium dedicated to advances in Visual Analytics Science and Technology. This year, the IEEE VAST08 will be held in conjunction with the IEEE Visualization Conference (IEEE Vis) and the IEEE Information Visualization Conference (IEEE InfoVis). IEEE VAST08 includes both fundamental research contributions within visual analytics as well as applications of visual analytics, including research in science, security and investigative analysis, engineering, medicine, health, media, business, and social interaction. We invite you to participate in IEEE VAST 2008 by joining us in Columbus, Ohio.

For questions, please email [vast@vis.computer.org](mailto:vast@vis.computer.org).

### VAST Symposium Cochairs

Thomas Ertl, University of Stuttgart  
David Ebert, Purdue University

IEEE VAST 2008 is part of **VisWeek 2008**, which also includes:  
IEEE InfoVis 2008    IEEE Information Visualization Conference  
IEEE Visualization 2008    IEEE Visualization Conference

>> **Important Dates**

IEEE VisWeek 2008 - Windows Internet Explorer

http://vis.computer.org/VisWeek2008/vast/index.html

1331 Susquehanna 101.9 KUJ 98.9 KWJZ WGLT CC Webmail Indian Hills Maty G T-Square

IEEE VisWeek 2008

## VAST08

VIS • INFOVIS • VAST

- > Welcome
- >> Vis • InfoVis • VAST
- >> Week-at-a-Glance
- >> VisWeek Sessions
- >> Exhibition
- >> Registration
- >> Student Volunteers
- >> Call for Participation
- >> Committees
- > Presenter Information
- > Site Map
- > Contact Us

Computer-based information visualization centers around helping people explore or explain data through interactive software that exploits the capabilities of the human perceptual system. A key challenge in information visualization is designing a cognitively useful spatial mapping of a dataset that is not inherently spatial and accompanying the mapping by interaction techniques that allow people to intuitively explore the dataset. Information visualization draws on the intellectual history of several traditions, including computer graphics, human-computer interaction, cognitive psychology, semantics, graphic design, statistical graphics, cartography, and art. The synthesis of relevant ideas from these fields with new methods and techniques made possible by interactive computation are critical for the development of information visualization with the torrents of data confronting them.

For questions, please email [infovis@vis.computer.org](mailto:infovis@vis.computer.org).

# IEEE VAST



# Sensemaking

“A motivated , continuous effort to understand connections (which can be among people, places, and events) in order to anticipate their trajectories and act effectively.”

– Klein, Moon and Hoffman



# Jigsaw

- Visualization for Investigative Analysis across Document Collections



# The Jigsaw Team

Carsten Görg  
Zhicheng Liu  
Vasili Pantazopoulos  
+ 4 new students



Gennadiy Stepanov  
Sarah Williams  
Neel Parekh  
Kanupriyah Singhal



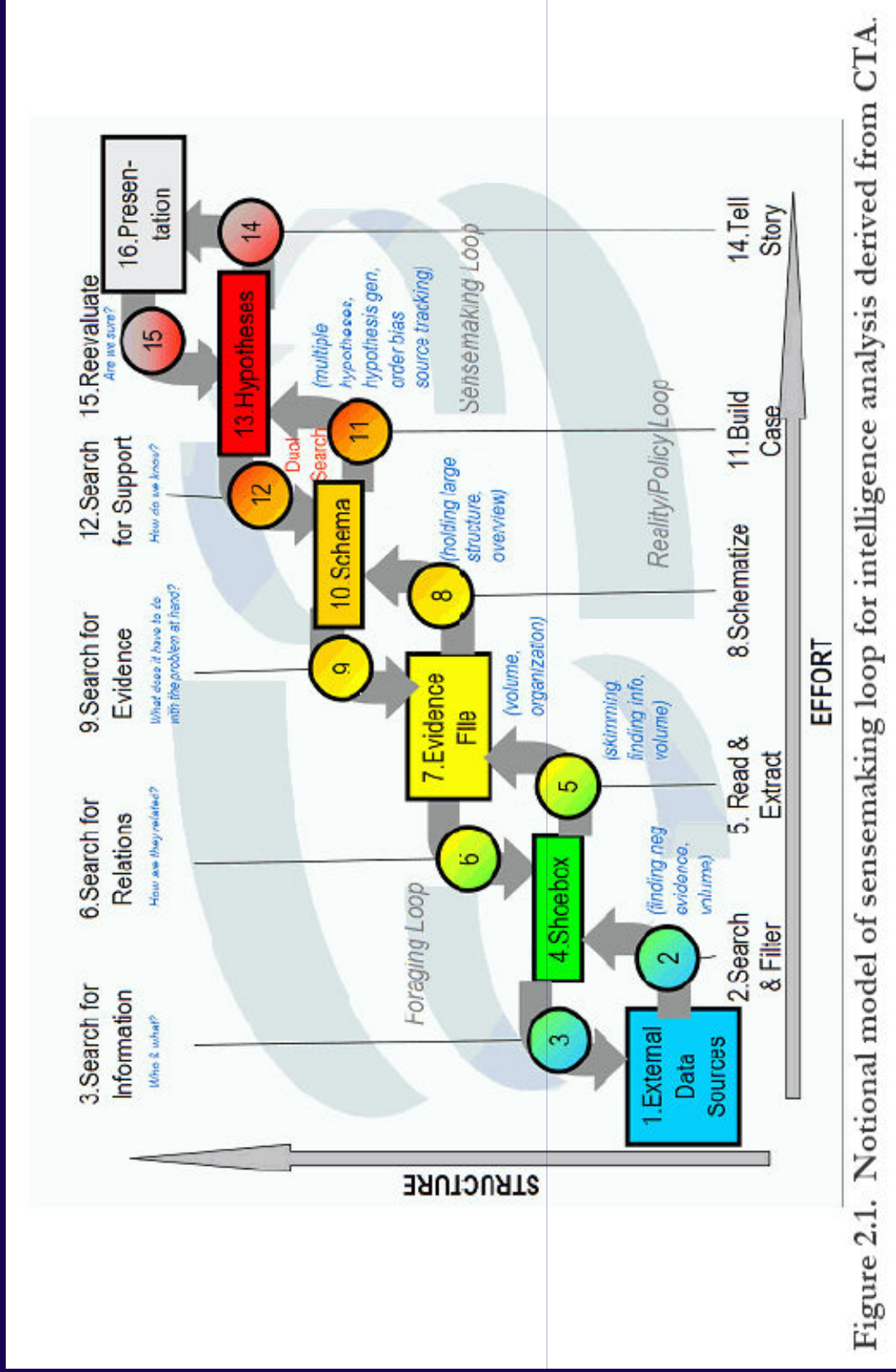


Figure 2.1. Notional model of sensemaking loop for intelligence analysis derived from CTA.

# Pirolli & Card, ICIA '05



## Pain Points

- Cost structure of scanning and selecting items for further attention
- Analysts' span of attention for evidence and hypotheses

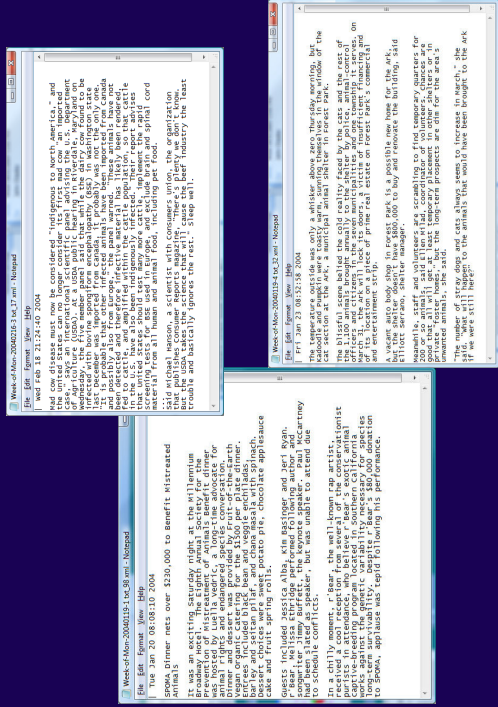




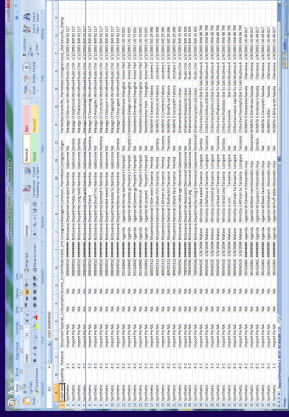
# Problem Addressed

# •Analogy

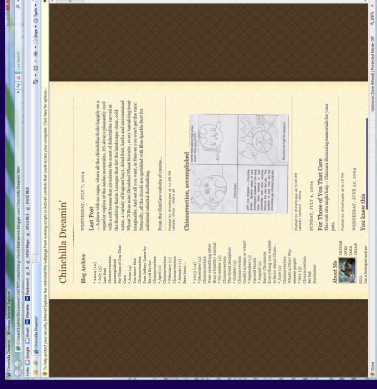
- Help investigative analysts discover plans, plots and threats embedded across the individual documents in large document collections



Documents/  
case reports



DBS



Blogs



# Example Document

Report: 20040510-4\_16  
May 14 2004

VANCOUVER, British Columbia - A Canadian immigration panel is considering whether accused environmental saboteur Tre Arrow can apply for refugee status in Canada.

Arrow, 30, who is wanted for fire bombing logging and cement trucks in Oregon, asked the Canadian authorities to remain in Canada as a political refugee at a hearing in Vancouver on Tuesday.

A key issue will be whether Arrow is affiliated with a terrorist group, which would immediately disqualify him from receiving refugee status in Canada, authorities said.

The Immigration and Refugee Board is scheduled to decide by May 31 whether Arrow is affiliated with the Earth Liberation Front, a group the FBI considers a terrorist organization responsible for scores of attacks on property over the past dozen years.



## Our Focus

- Entities within the documents
  - Person, place, organization, phone number, date, license plate, etc.
- Thesis: A plot/threat within the documents will involve a set of entities in coordination



# Entity Identification

- Must identify and extract entities from plain text documents
  - Crucial for our work
- Not our main research focus – Collaborate with or use tools from others



# Entities Identified

**Source:**

**Date:** May 14, 2004

**VANCOUVER, British Columbia** - A Canadian immigration panel is considering whether accused environmental saboteur **Tre Arrow** can apply for refugee status in **Canada**.

**Arrow**, 30, who is wanted for fire bombing logging and cement trucks in **Oregon**, asked the Canadian authorities to remain in **Canada** as a political refugee at a hearing in **Vancouver** on **Tuesday**.

A key issue will be whether **Arrow** is affiliated with a terrorist group, which would immediately disqualify him from receiving refugee status in **Canada**, authorities said.

The **Immigration and Refugee Board** is scheduled to decide by **May 31** whether **Arrow** is affiliated with the **Earth Liberation Front**, a group the **FBI** considers a terrorist organization responsible for scores of attacks on property over the past dozen years.



# Connections

- Entities relate/connect to each other to make a larger “story”
- Connection definition:
  - Two entities are connected if they appear in a document together
  - The more documents they appear in together, the stronger the connection



# Jigsaw

“Putting the pieces together”

- Multiple visualizations (views) of documents, entities, & their connections
- Views are highly interactive and coordinated
- User actions generate events that are transmitted to and (possibly) reflected in other views



# System Views

against agents **arrow** arson authorities being **canada** canadian charges construction **Elf** enforcement **environmental fbi** federal government group logs

Documents

- 1 20031027-8,51
- 0 2004015-3,65
- 0 2004019-2,75
- 0 20040510-4,54
- 0 20040614\_27

Source:

Date: May 14, 2004

**Vancouver, British Columbia** - A Canadian immigration panel is considering whether accused environmental saboteur **Travis Arrow** can apply for refugee status in **Canada**.

**Arrow**, 30, who is wanted for the bombing, logging and cement trucks attacks, saved the Canadian authorities to remain in **Canada** as a poll refugee at a hearing in **Vancouver** on **Tuesday**.

A key issue will be whether **Arrow** is admitted with a terrorist group which would immediately disqualify him from receiving refugee status in **Canada**, authorities said.

The immigration and **Business Express** is scheduled to decide by **May 21**.

**Arrow** is affiliated with the **Earth Liberation Front**, a group the **FBI** considers a terrorist organization responsible for scores of attacks on property over the past dozen years.

Filters:

- US: 35
- North America: 15
- South America: 3
- Chile: 1
- Australia: 15
- Asia: 21
- South Africa: 3
- Chile: 1
- Europe: 21
- Other: 21
- South Africa: 3
- Chile: 1
- Australia: 15
- Asia: 21
- South Africa: 3
- Chile: 1
- Europe: 21
- Other: 21

JGOSAW

Blue Iguanodon

Color Legend:

- place
- organization
- category
- date
- money
- time

Search:

Substituting search:

Test search:

Network graph showing connections between nodes.

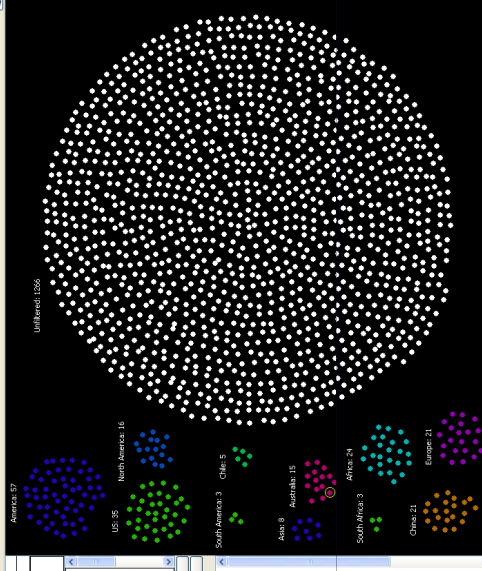
Nodes include:

- Animal Liberation Front
- California Institute of Technology
- Department of Fish and Wildlife
- Earth First!
- Liberation front
- IEL Web
- FEDA
- Fourth World Way Redwave Board
- Joan Terrence Task Force
- Pearland State University
- Wildlife Conservation Society
- Schuppert Logging Company
- U.S. Department of Agriculture
- U.S. Forest Service
- Wildlife Society
- WFO TALLEY JANE

Network graph showing connections between nodes.

Nodes include:

- Alaska
- Arizona
- California
- Colorado
- Connecticut
- Delaware
- Florida
- Georgia
- Idaho
- Illinois
- Indiana
- Iowa
- Kansas
- Kentucky
- Louisiana
- Maine
- Massachusetts
- Michigan
- Minnesota
- Mississippi
- Montana
- Nebraska
- Nevada
- New Hampshire
- New Jersey
- New Mexico
- New York
- North Carolina
- North Dakota
- Ohio
- Oklahoma
- Oregon
- Pennsylvania
- Rhode Island
- South Carolina
- South Dakota
- Tennessee
- Texas
- Utah
- Vermont
- Virginia
- Washington
- West Virginia
- Wisconsin
- Wyoming



Network graph showing connections between nodes.

Nodes include:

- Atlanta
- Boston
- Chicago
- Denver
- Houston
- Los Angeles
- London
- Madrid
- Mumbai
- New York
- Paris
- Rome
- San Francisco
- Seattle
- Singapore
- Tokyo
- Washington DC

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Atlanta	0	0	0	0	0	0	0	0	0	0	0	0
Boston	0	0	0	0	0	0	0	0	0	0	0	0
Chicago	0	0	0	0	0	0	0	0	0	0	0	0
Denver	0	0	0	0	0	0	0	0	0	0	0	0
Houston	0	0	0	0	0	0	0	0	0	0	0	0
Los Angeles	0	0	0	0	0	0	0	0	0	0	0	0
London	0	0	0	0	0	0	0	0	0	0	0	0
Mumbai	0	0	0	0	0	0	0	0	0	0	0	0
New York	0	0	0	0	0	0	0	0	0	0	0	0
Paris	0	0	0	0	0	0	0	0	0	0	0	0
Rome	0	0	0	0	0	0	0	0	0	0	0	0
San Francisco	0	0	0	0	0	0	0	0	0	0	0	0
Seattle	0	0	0	0	0	0	0	0	0	0	0	0
Singapore	0	0	0	0	0	0	0	0	0	0	0	0
Tokyo	0	0	0	0	0	0	0	0	0	0	0	0
Washington DC	0	0	0	0	0	0	0	0	0	0	0	0

Network graph showing connections between nodes.

Nodes include:

- Atlanta
- Boston
- Chicago
- Denver
- Houston
- Los Angeles
- London
- Madrid
- Mumbai
- New York
- Paris
- Rome
- San Francisco
- Seattle
- Singapore
- Tokyo
- Washington DC





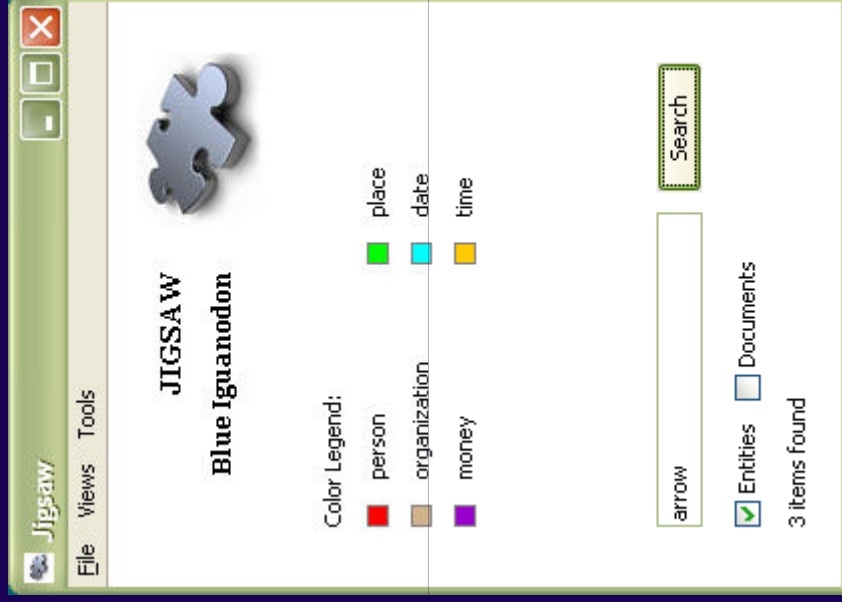
# The Need for Pixels



# Demo



# Console



# Document View

The screenshot shows a 'Document View' window with a microphone icon in the top right. The main text area contains a news article with several words highlighted in colored boxes: 'arrow' (pink), 'canada' (green), 'environmental fbi' (orange), 'arson authorities' (pink), 'self' (orange), and 'group logging' (orange). Below the text is a list of documents with columns for a document ID, a count, and a document ID. The document '20040510-4\_16' is highlighted in yellow. To the right of the list is a 'Source' section with a date and a paragraph of text containing more highlighted words like 'VANCOUVER', 'British Columbia', 'saboteur', 'Arrow', 'Canada', 'Oregon', 'Tuesday', 'Arrow', 'May 31', 'FBI', and 'Refugee Board'.

Microphone

Document View

Edit View Bookmarks

against agents **arrow** arson **authorities** being  
**canada** canadian charges construction **self** enforcement  
**environmental fbi** federal government **group logging**

Only Entities

Documents		
1	20031027-8_51	
0	20040315-3_65	
0	20040419-2_25	
1	20040510-4_16	
0	20040510-4_54	
0	20040614_27	

**Source:**  
**Date:** May 14, 2004  
**VANCOUVER, British Columbia** - A Canadian immigration panel is considering whether accused environmental **saboteur Tre Arrow** can apply for refugee status in **Canada**.  
**Arrow**, 30, who is wanted for fire bombing logging and cement trucks in **Oregon**, asked the Canadian authorities to remain in **Canada** as a political refugee at a hearing in **Vancouver** on **Tuesday**.  
A key issue will be whether **Arrow** is affiliated with a terrorist group, which would immediately disqualify him from receiving refugee status in **Canada**, authorities said.  
The **Immigration and Refugee Board** is scheduled to decide by **May 31** whether **Arrow** is affiliated with the **Earth Liberation Front**, a group the **FBI** considers a terrorist organization responsible for scores of attacks on property over the past dozen years.



# List View

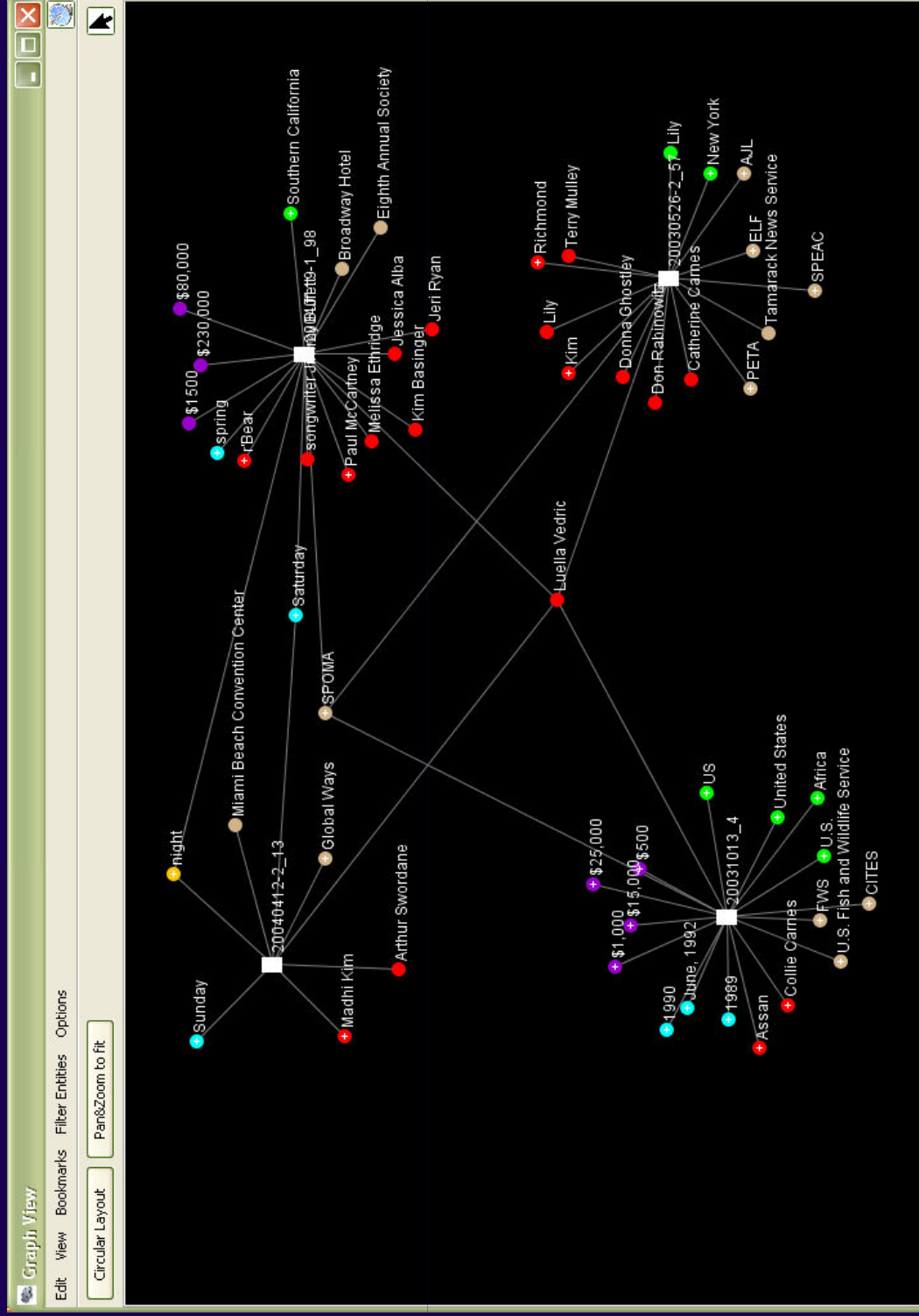
The screenshot displays a web application interface titled "List View" with a menu bar containing "Edit", "View", "Bookmarks", "Lists", and "Options". The main content area is divided into three vertical panels, each with a search bar and a "Show all connections" button.

- place panel:** Lists geographical locations such as Abbottsford, Alberta, America, Blaine, British Columbia, British Columbia, Buckner, Calgary, Canada, central Alberta, Chiapas, Colo., Colorado, Columbia, Dawson City, Eagle Creek, Edmonton, Fort Collins, Fort Nelson, Great Slave Lake, Hollywood, Iraq, Long Island, LOS ANGELES, Los Angeles, Mexico, Montana, N.M., N.Y., northern British Columbia, Northwest Territories, Oregon, Portland, and Puget Sound.
- person panel:** Lists individuals including Arnold Schwarzenegger, Arrow, Connor Cash, Connor Cash, 22, Cottrell, 23, Craig Rosebraugh, David Barbarash, Duane Christopher Bradley, Gary R. Perlstein, George Bush, Jack, Jeff Guidry, Jeff Koerings, Jim Wichmier, Keith, Keith Jespersen, Michael Scarpitti, News Headlines, Pumpkin, Robert Blecksmith, Ron Arnold, Rosebraugh, 31, saboteur Tre Arrow, Scott McInnis, Tre Arrow, Wallace Stegner, Wesley Smith, WILDLIFE EXPERT CAUGHT SMUGGLING REPTILES, William Jensen Cottrell, and Wolf.
- organization panel:** Lists various organizations and groups such as ALF, Animal Liberation Front, California Institute of Technology, CWD, Department of Fish and Wildlife, Earth First, Earth Liberation Front, ELF, ELF Web, FBI, FDA, Fourth World War, Immigration and Refugee Board, Joint Terrorism Task Force, Portland State University, San Bernardino Central Detention Cen., Sarvey Wildlife Center, Schoppert Logging Company, Supreme Court, SUV, U.S. Department of Agriculture, U.S. District Court, US Forrest Service, Washington Fish and Wildlife Commission, and WHO KILLED JANE.

Lines connect items across the panels, showing relationships. For example, "Tre Arrow" in the person panel is connected to "Arrow" in the person panel and "Tre Arrow" in the organization panel. "WILDLIFE EXPERT CAUGHT SMUGGLING REPTILES" in the person panel is connected to "WILDLIFE EXPERT CAUGHT SMUGGLING REPTILES" in the organization panel.



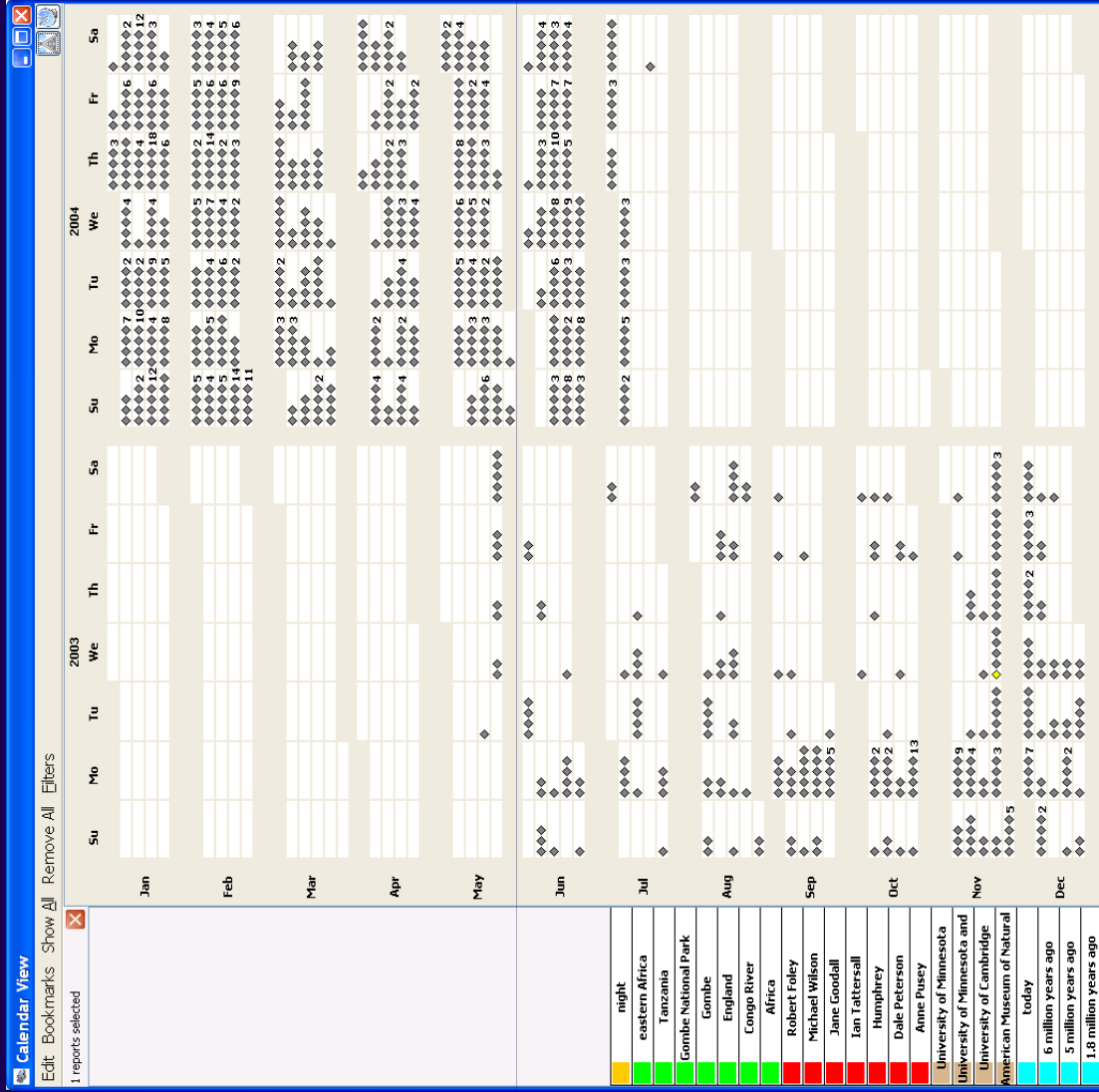
# Graph View



# Scatterplot View

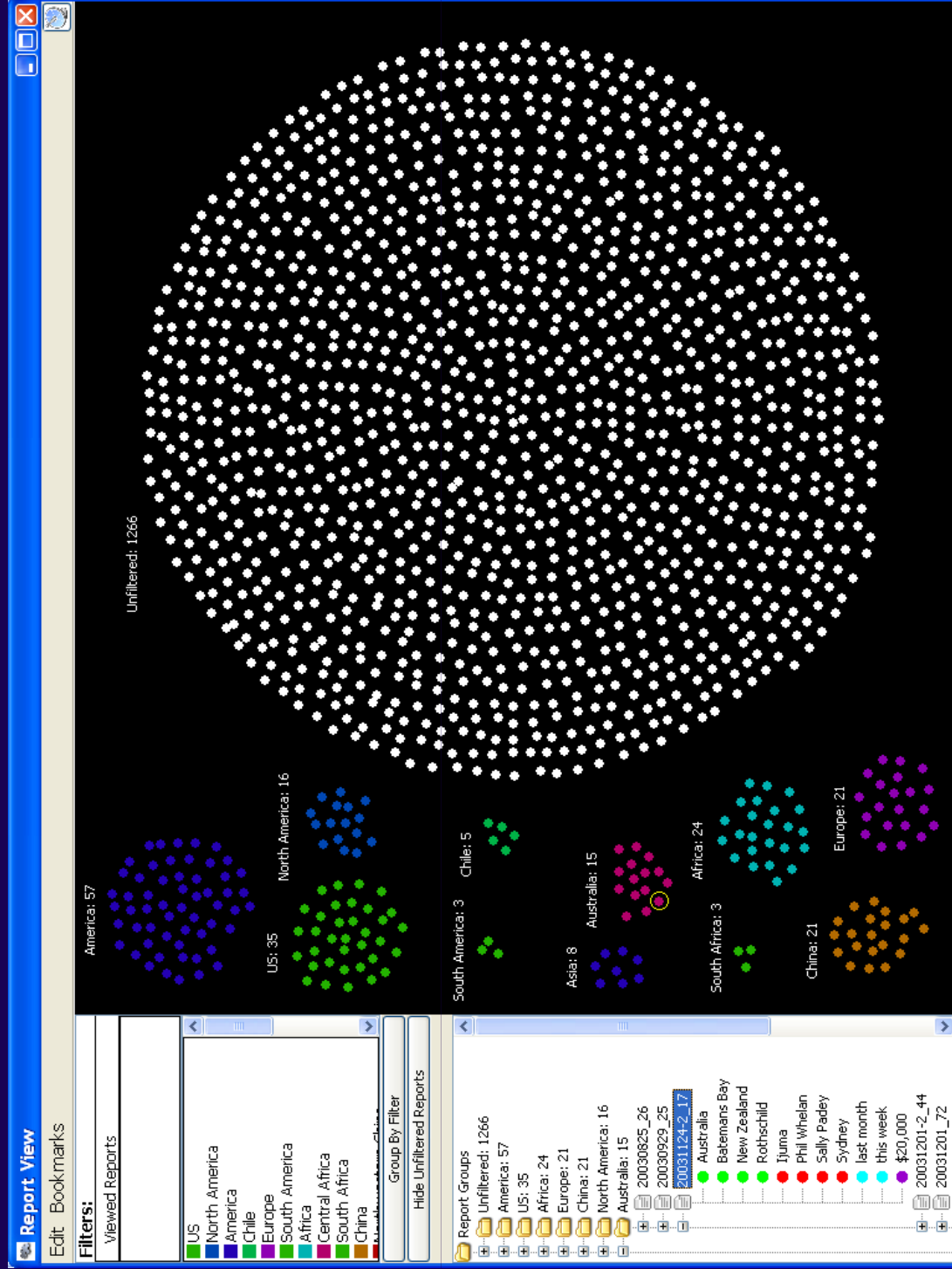


# Calendar View

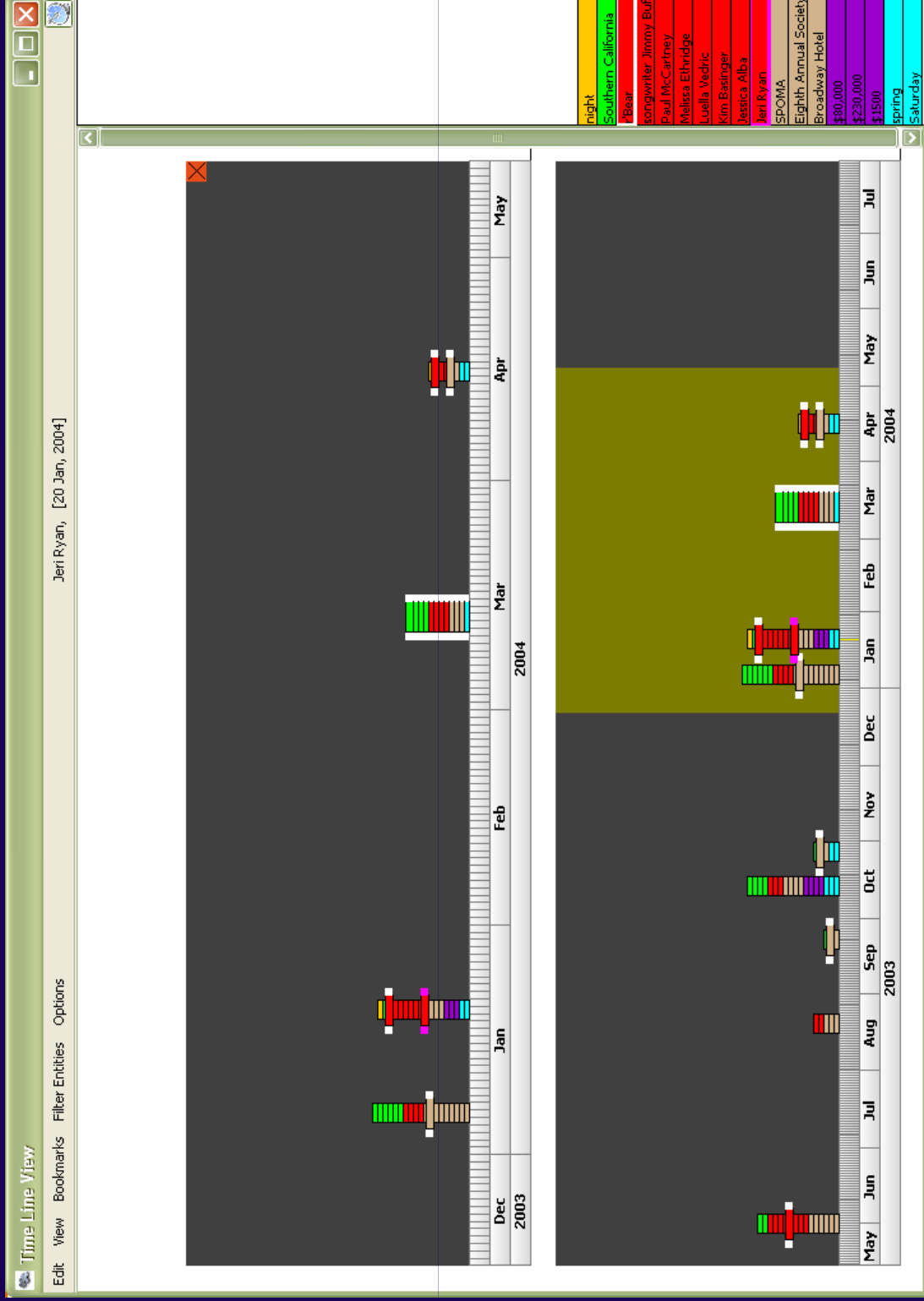




# Report Cluster View



# Timeline View



# Shoebox

The screenshot shows the 'Shoebox' application window. At the top, there is a navigation bar with 'UP' and 'DOWN' buttons. Below this is a list of nodes, each with a colored dot and a label. The nodes are: Lily (red), Madhi Kim (red), South America (green), Florida Department (brown), FWS (brown), Global Ways (brown), Sanchez (brown), Aquatics (red), Arthur Swordane (red), Kim (red), Leslie (red), Luella Vedic (red), r'Bear (red), Singer (red), Alabama (green), America (green), California (green), New York (green), Texas (green), United States (green), Zimbabwe (green), 1989 (cyan), 1990 (cyan), \$15,000 (purple), and \$25,000 (purple). A search bar at the top right contains 'chinchilla Flocke ???' with a red dot next to it. A filter bar below the search bar has 'Kim visited' (red dot), 'New York last' (green dot), and 'spring' (cyan dot). Two pop-up windows are open. The first, titled 'Hypothesis: Luella is involved in bad activities', lists nodes: Assan (red), Colle Carnes (red), r'Bear (red), Global Ways (brown), 20030526-2\_57 (brown), 20031013\_4 (white), and Luella Vedic - Luella's Connections (yellow). The second, titled 'Chinchilla smuggling', lists nodes: Gil Breeders (red), Los Angeles (green), Cesar Gil (red), Chile (green), Argentina (green), Andes (green), and Peru (green). A third pop-up window titled 'Global Ways' lists nodes: CDC (brown), Fish and Wildlife Service (brown), US Department of Agriculture (brown), Florida (green), South America (green), Ed Parker (red), and Madhi Kim (red). A white arrow points from the 'Luella Vedic - Luella's Connections' node in the first pop-up to the 'Chinchilla smuggling' node in the main graph.



## Trial Use

- Transitioning system to real clients



# Future Work

- Entity Identification
- Evaluation
- Collaborative version
- Themes/concepts
- Enhanced evidence marshalling
- Present/browse investigation history
- Scalability issues
- Wikipedia & Intellipedia
- Geospatial View
- Connectivity search
- Reliability/uncertainty
- Other types of data
- Web search & situational awareness
- Display wall?
- Deployment



## Take Away Point

- Design your visualization systems and tools to facilitate analysis and exploration
  - Not to just illustrate and reconfirm existing knowledge
- Including flexible, useful interaction is one of the best ways to do this



# To Learn More

- <http://www.gvu.gatech.edu/ii>

The screenshot shows the website for the Information Interfaces Research Group. The browser's address bar displays the URL <http://gvu.cc.gatech.edu/ii/>. The website features a navigation menu with links for About, People, Projects, Publications, Resources, Talks, and Videos. A central banner reads: "The Information Interfaces Group, an HCI research group in the Gvu Center at Georgia Tech, develops computing technologies that help people take advantage of information to enrich their lives. More about the lab approach". Below this, a "Projects" section lists several initiatives: "Jigsaw" (using visualization and visual analytics for intelligence analysis), "Information Visualization" (helping people understand and analyze data through interactive techniques), "The Buzz" (supporting end-user mashup creation and content aggregation), and "InfoCanvas" (developing information art). A "Hot News" section on the right contains several updates, including a keynote at the 2008 VL/HCC and SoftVis Symposia, three papers accepted at the 2008 IEEE InfoVis Conference, a visit to the CEPS Center and DSTO in Australia, a contest win at VAST 2007, and the availability of 7450 InfoVis lectures. A "People" section at the bottom left shows headshots of group members: John Skasko, Carsten Görg, James Eagen, Youn ah Kang, Zhicheng Liu, and Chris Plaue. The website also includes a "Downloads" section with "Active" and "Downloads" buttons for various papers.



# Acknowledgment

- Some slides in this presentation borrowed from overviews of visual analytics by Jim Thomas, NVAC Director





# Acknowledgments

- Work conducted as part of the Southeastern Regional Visualization and Analytics Center, supported by DHS and NVAC



- Supported by NSF IIS-0414667



This material is supported in part by:  
**National Science Foundation**  
Grant #0414667



**End**

- Thanks for your attention!
- Questions?

