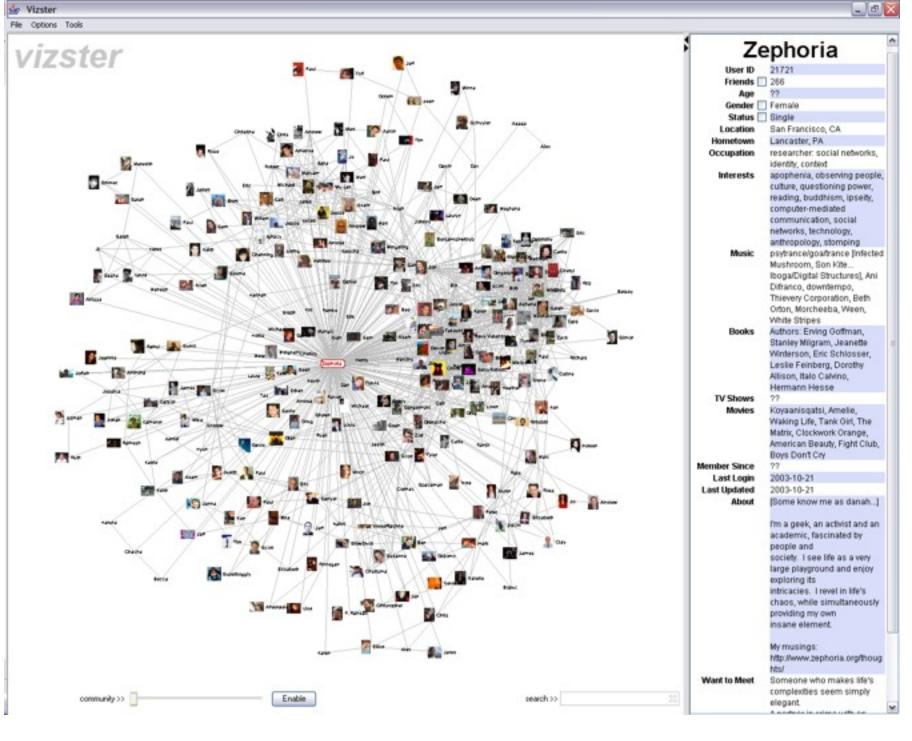
CSE512 :: 11 Mar 2014

Collaborative Analysis



Jeffrey Heer University of Washington

A Tale of Two Visualizations

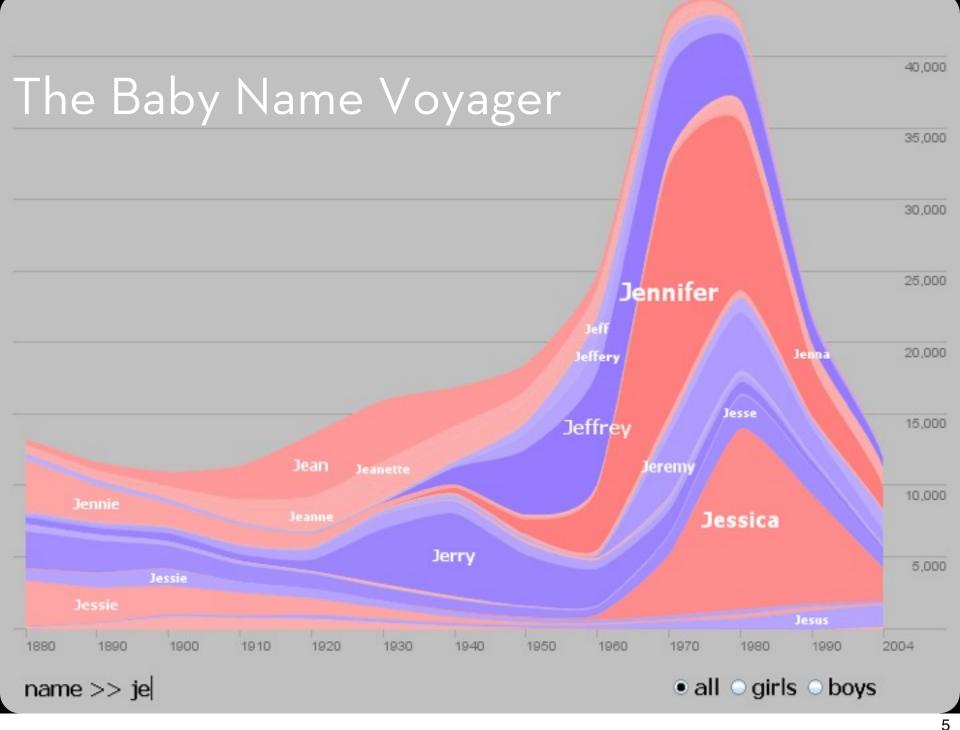


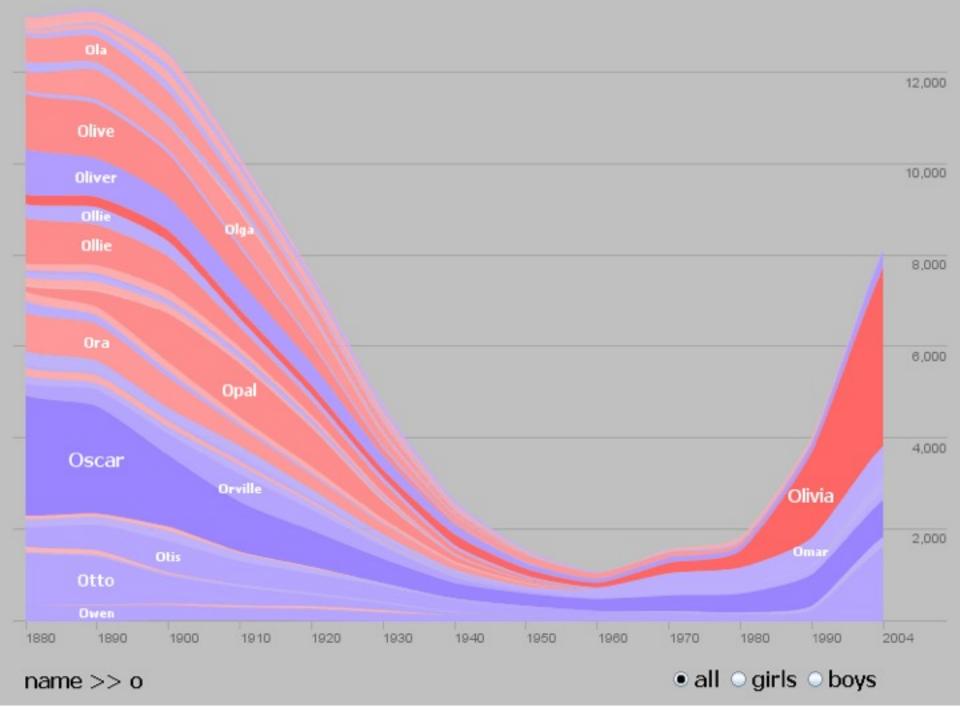
Observations

Groups spent more time in front of the visualization than individuals.

Friends encouraged each other to unearth relationships, probe community boundaries, and challenge reported information.

Social play resulted in informal analysis, often driven by story-telling of group histories.





Social Data Analysis

Visual sensemaking is a **social** process as well as a cognitive process.

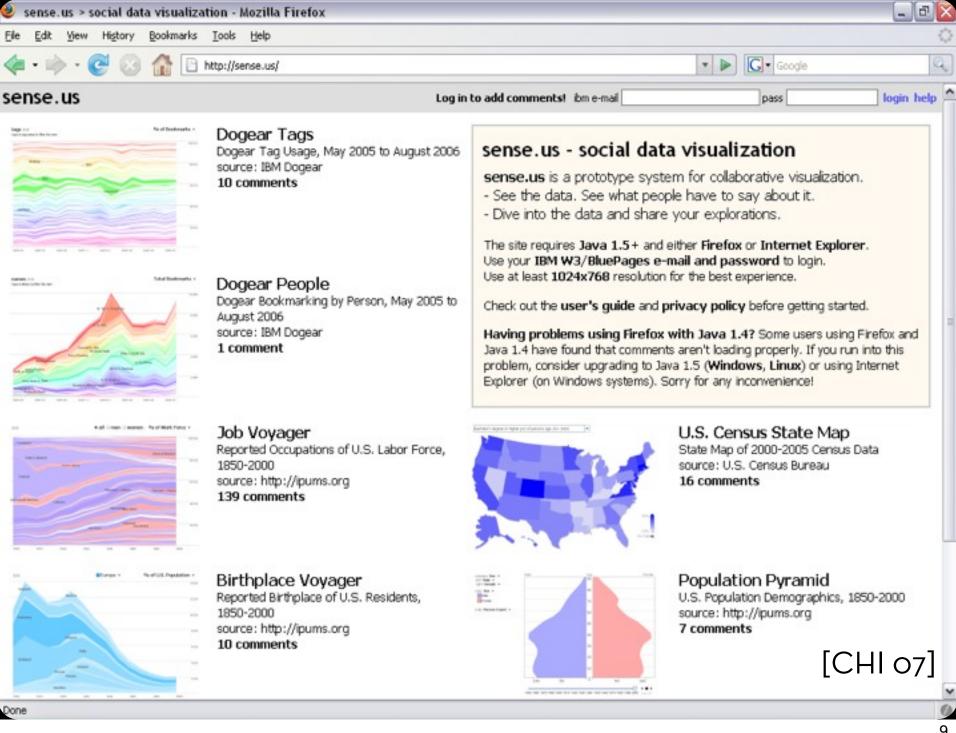
Analysis of data coupled with social interpretation and deliberation.

How can user interfaces catalyze and support collaborative visual analysis?

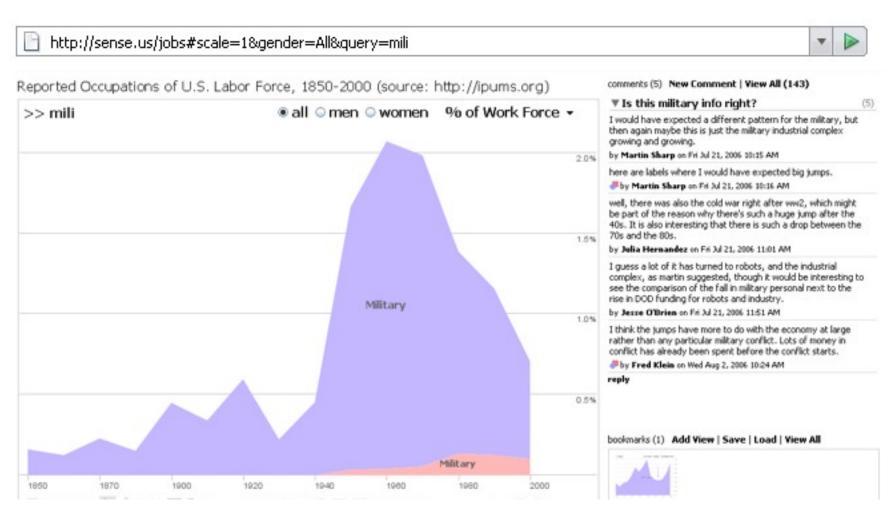
sense.us

A Web Application for Collaborative Visualization of Demographic Data

with Fernanda Viégas and Martin Wattenberg

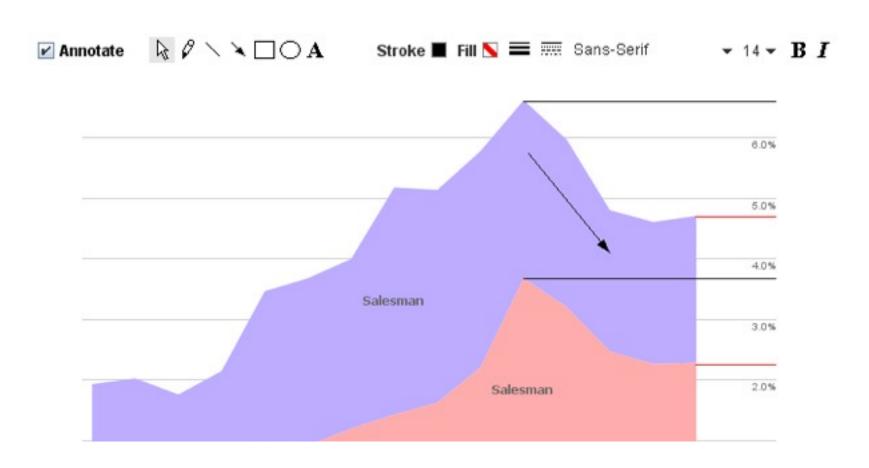


Sharing within visualizations and across the web



Sharing within visualizations and across the web

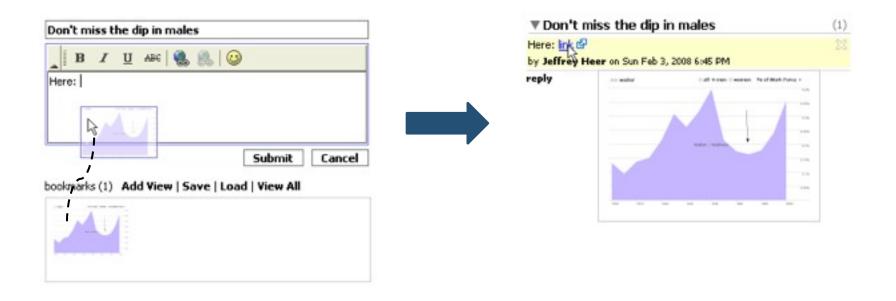
Pointing at interesting trends, outliers



Sharing within visualizations and across the web

Pointing at interesting trends, outliers

Collecting and linking related views

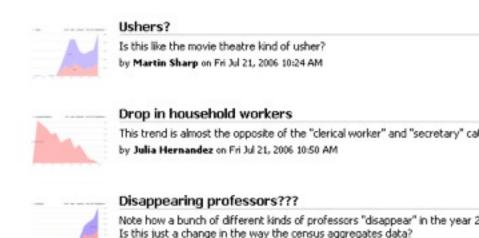


Sharing within visualizations and across the web

Pointing at interesting trends, outliers

Collecting and linking related views

Awareness of social activity



by Julia Hernandez on Fri Jul 21, 2006 11:07 AM

Sharing within visualizations and across the web

Pointing at interesting trends, outliers

Collecting and linking related views

Awareness of social activity

Embedding in external media (blogs, wikis, webpages)

Sharing within visualizations and across the web

Pointing at interesting trends, outliers

Collecting and linking related views

Awareness of social activity

Embedding in external media (blogs, wikis, webpages)

Don't disrupt individual exploration

User Study Design

30 participant laboratory study

25 minute, unstructured sessions with job voyager

3-week live deployment on IBM intranet

Employees logged in using intranet accounts

Data analyzed

- 12.5 hours of qualitative observation
- **258** comments (41 pilot, 85 ibm, 60 ucb, 72 live)

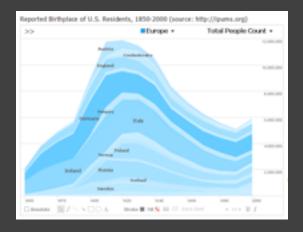
Usage logs of user sessions

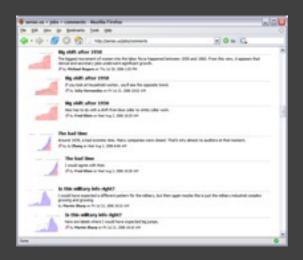
Voyagers and Voyeurs

Complementary faces of analysis

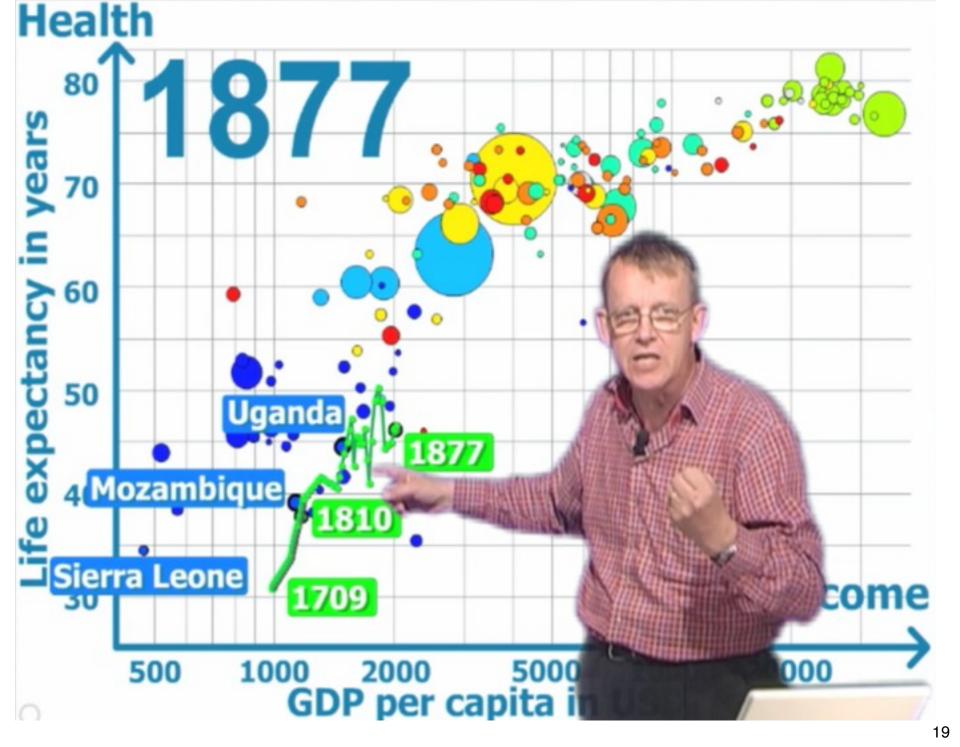
Voyager – focus on visualized data Active engagement with the data Serendipitous comment discovery

Voyeur – focus on comment listings Investigate others' explorations Find people and topics of interest Catalyze new explorations

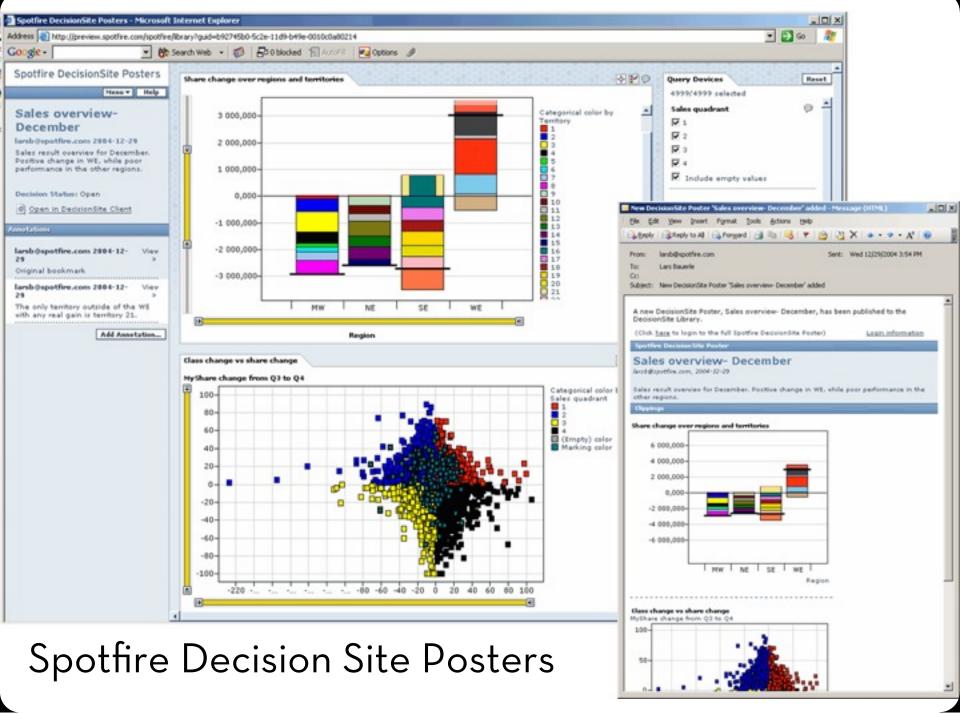




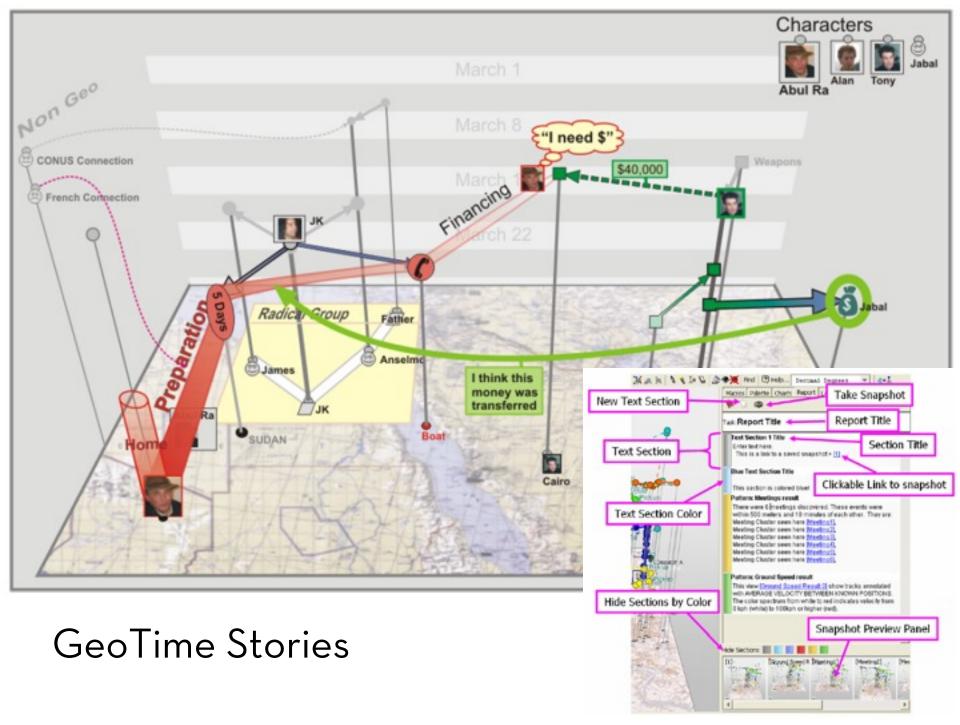
Social Data Analysis

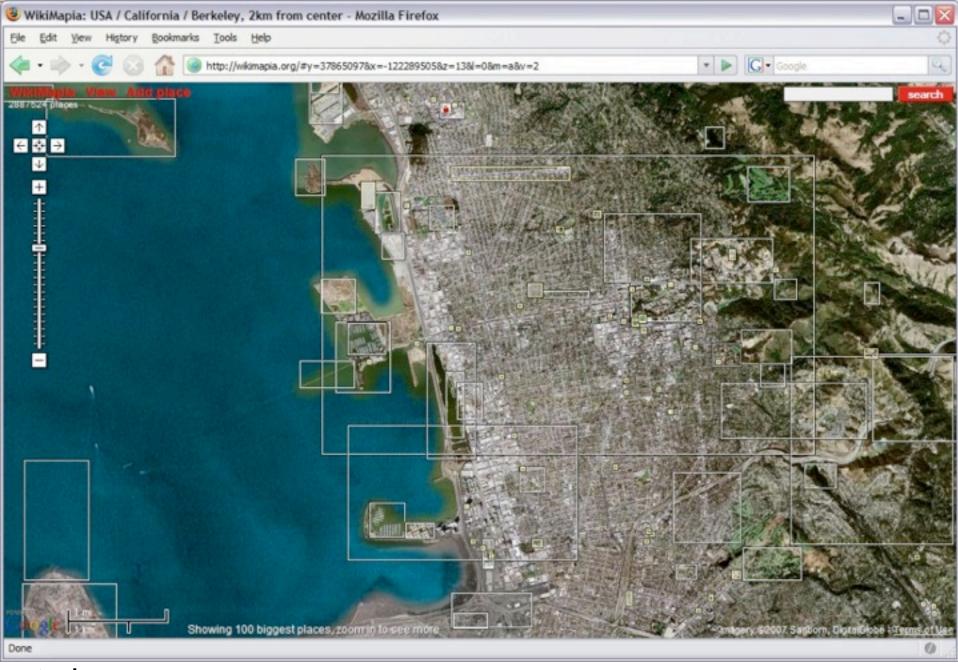




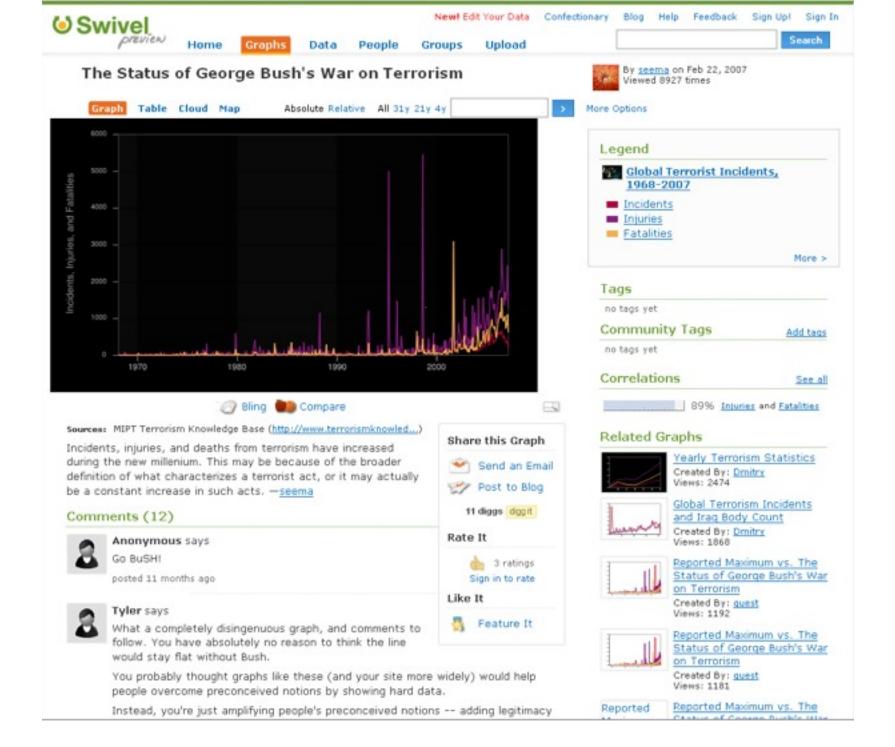


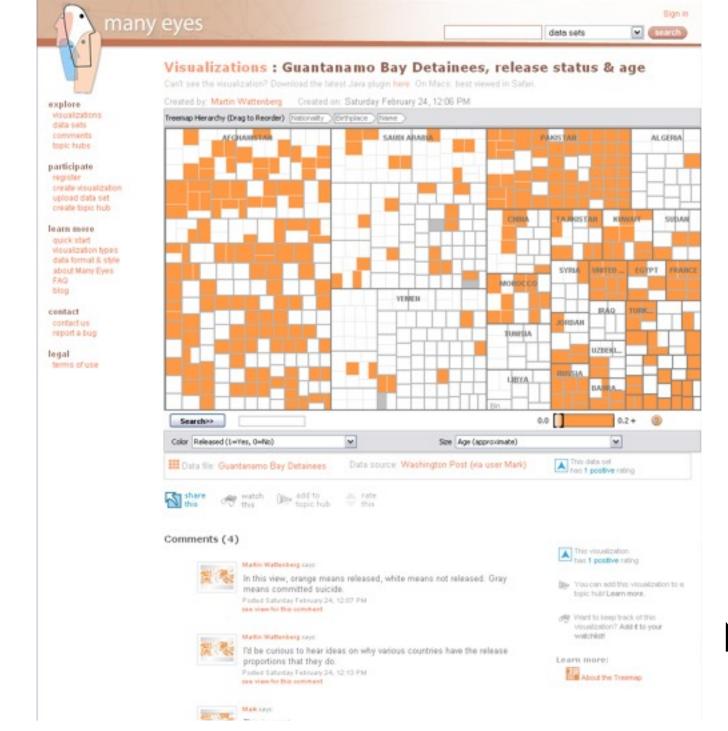






Wikimapia.org

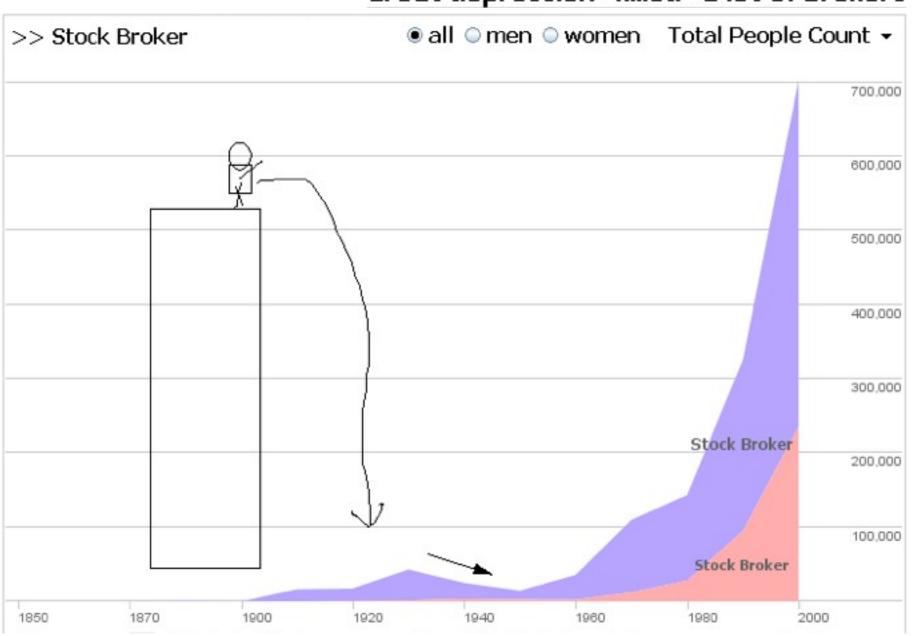


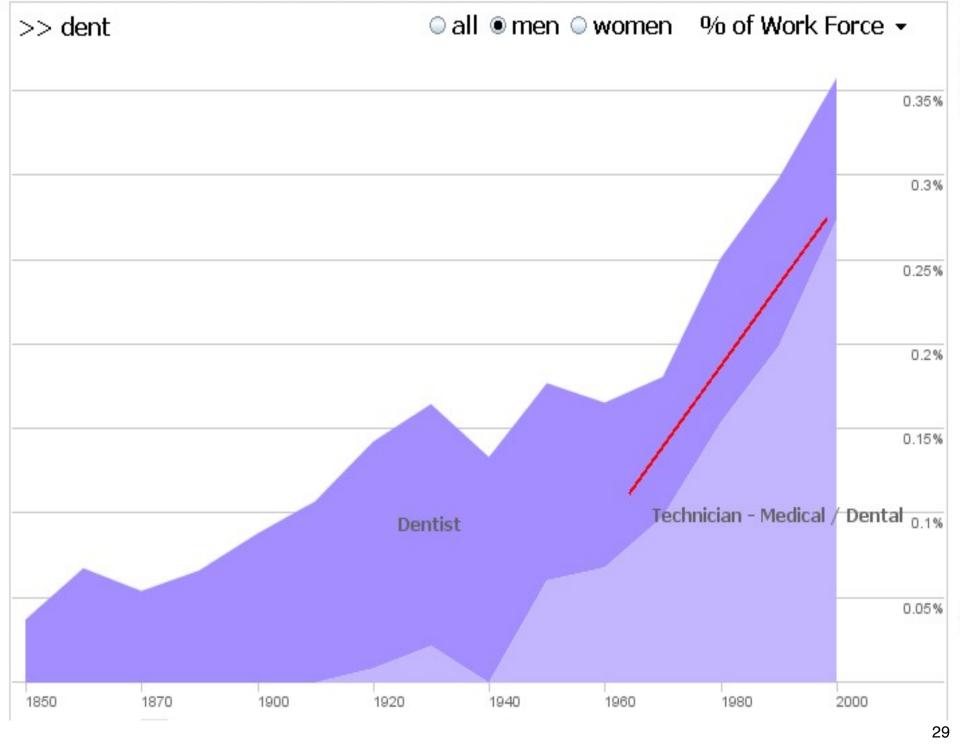


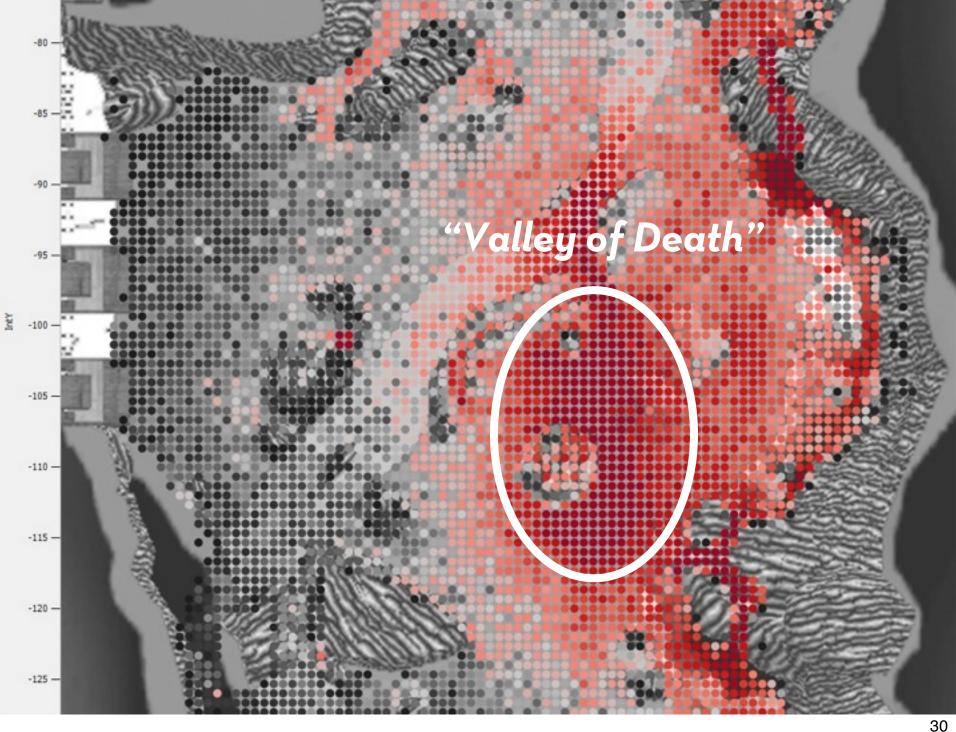
Many-Eyes

Discussion

Great depression "killed" a lot of brokers







explore comments topic hubs

participate: uproved data sed. create topic holi regreter

Searn more march start viscalization types data format & store about Many Eyes blog.

contact 0s. report a bug

hegal : terms of use

2007 2008 see blo tooks Census crim education election energy to health magazin people politic

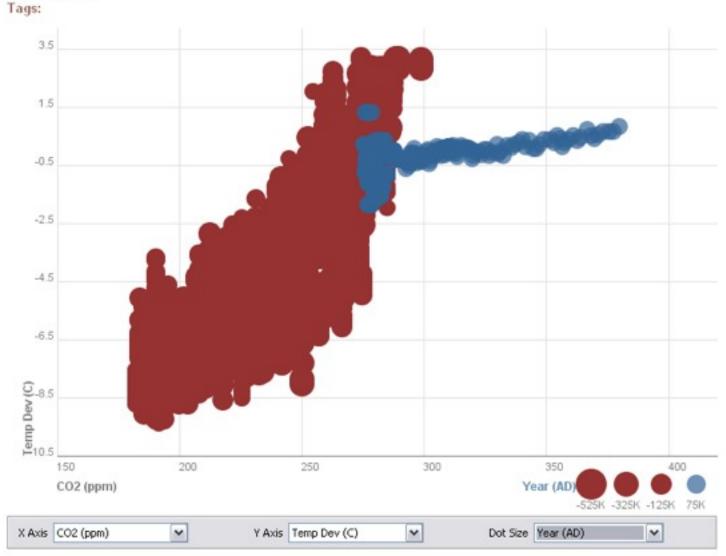
Popular Dataset Tags

population president prox map

speech state and money text US min WORLD

Visualizations: Flawed Data - Temp Over Time and CO2 Levels: This view shows the correlation between...

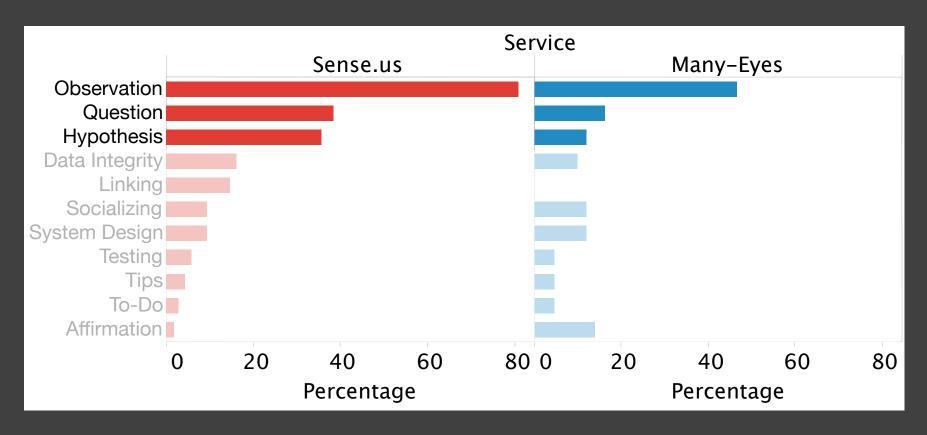
Creator: Bruce



years cluster. What does this mean though that such high CO2 levels are preceeding significant changes in temperature? Something, much worse is yet to come? Pulled Montay January 29 2007, 08:32 AM. day year to the comment.

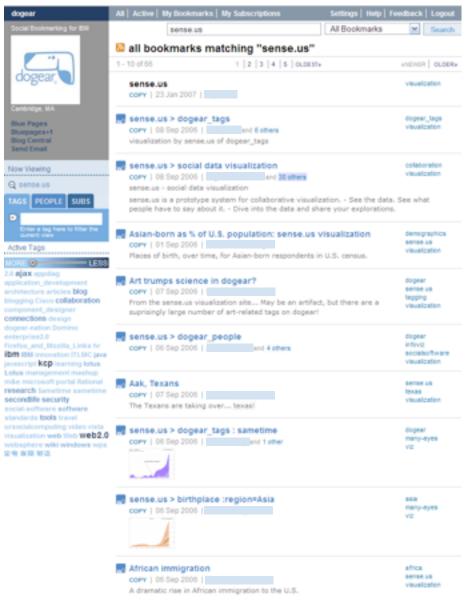
experienced numerous times in the 4+ billion year history of the planet. Pushed Friday February 09.2007, 03.46 PM the view to this section of

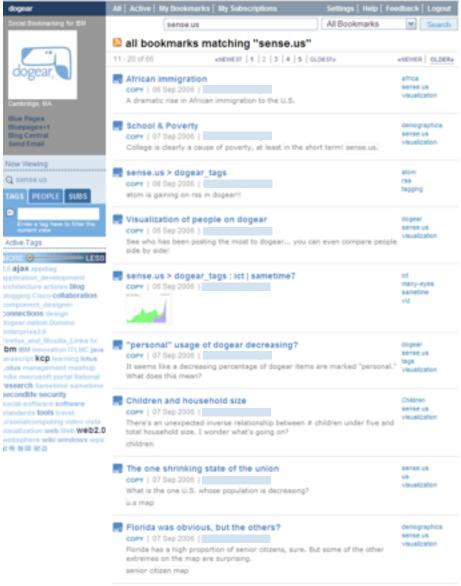
Content Analysis of Comments



Feature prevalence from content analysis (min Cohen's K = .74) High co-occurrence of Observation, Question, and Hypothesis

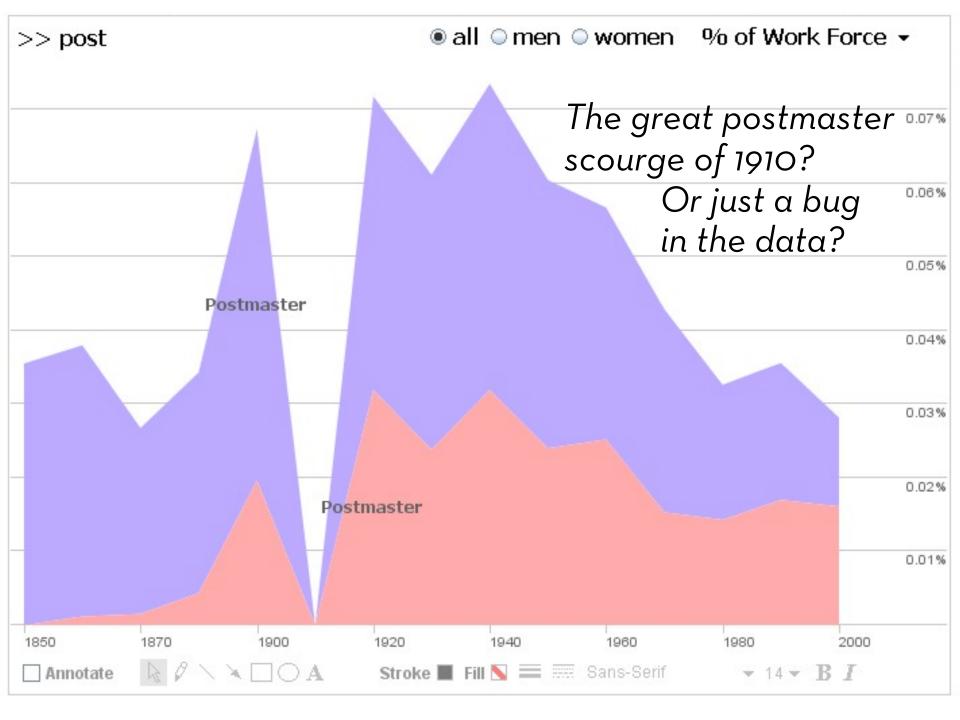
Sharing in External Media

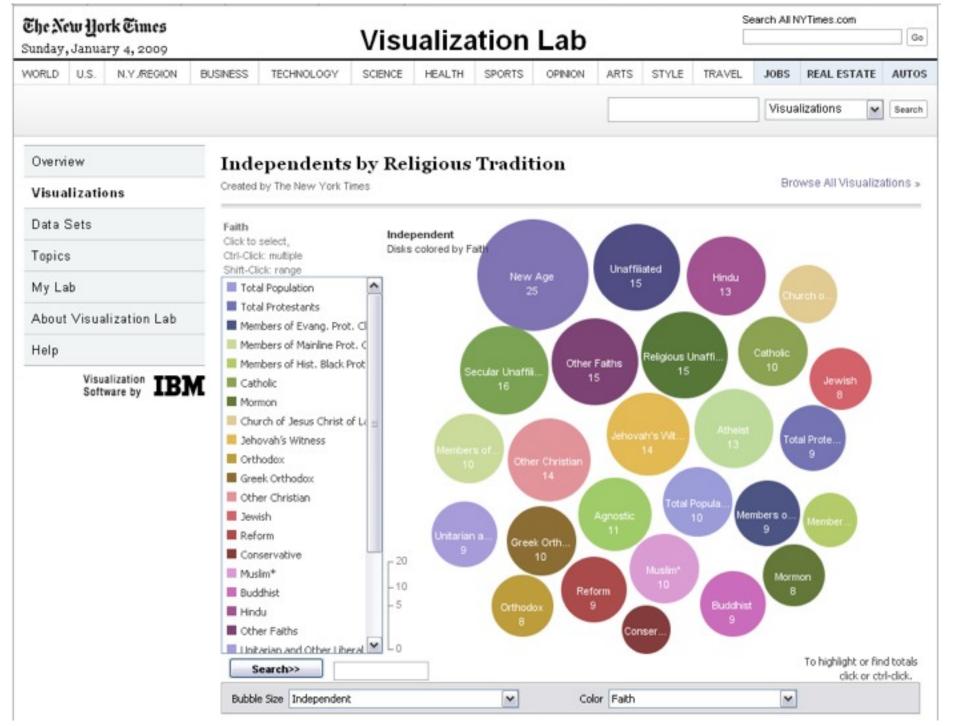


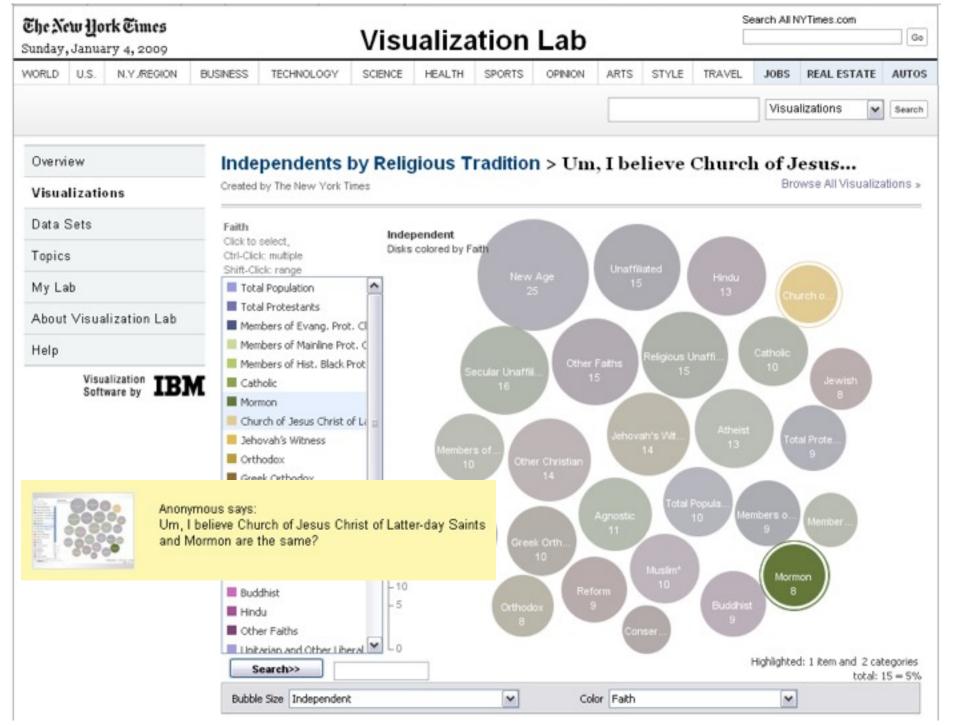


Data Quality

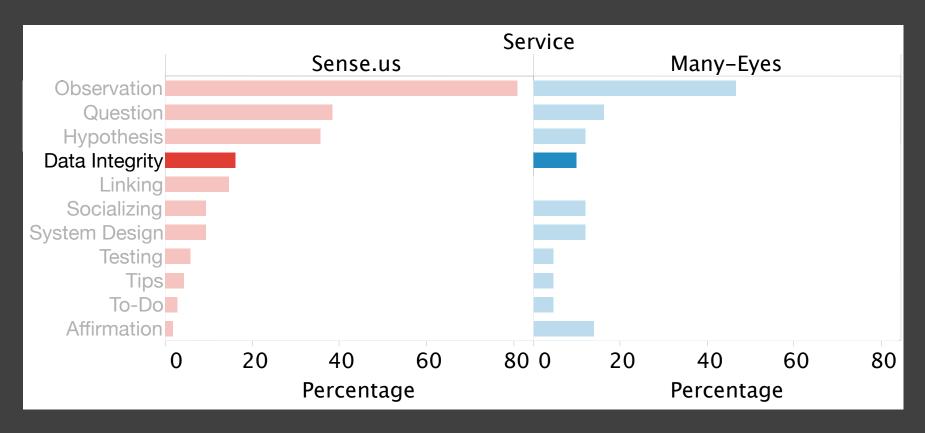








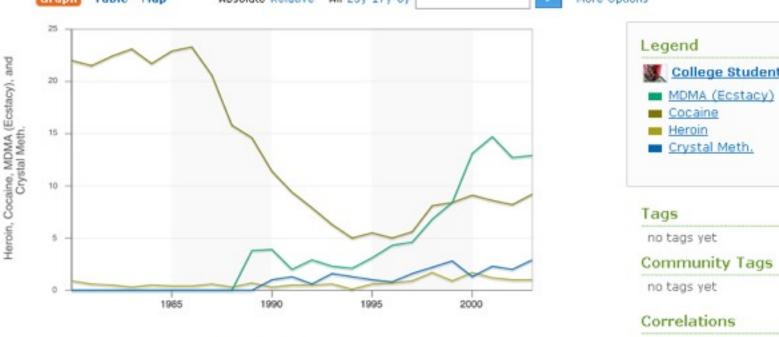
Content Analysis of Comments



16% of sense.us comments and **10%** of Many-Eyes comments reference data integrity issues.

Data Integration in Context





Sources: University of Michigan

Community Tags

no tags yet

Correlations

See all

Share this Graph

Related Graphs

Related Graphs

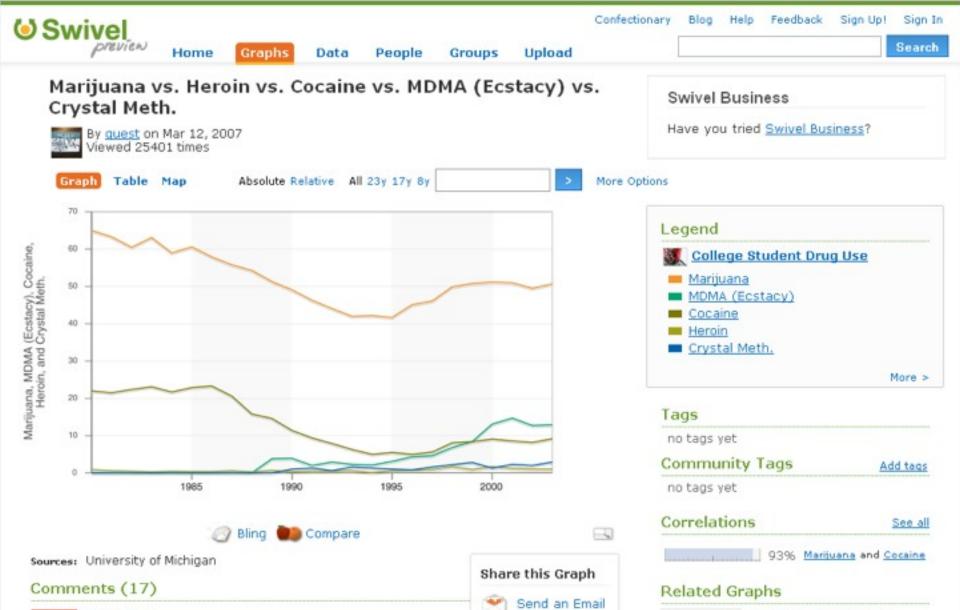
Send an Email

3 ratings

Where's the pot?

Heroin vs. Cocaine vs.
MDMA (Ecstacy) vs. Crystal
Meth. vs. Heroin vs.
Cocaine vs. MDMA
(Ecstacy) vs. Crystal Meth.
vs. Cocaine, NASDAQ
Adjusted Close, and Dow
Jones Industrial Average
Adjusted Close

More >



Post to Blog

Digg submit

O ratings

Rate It

visnu says

MARTY savs

pot pot pot pot

weed weed weed weed

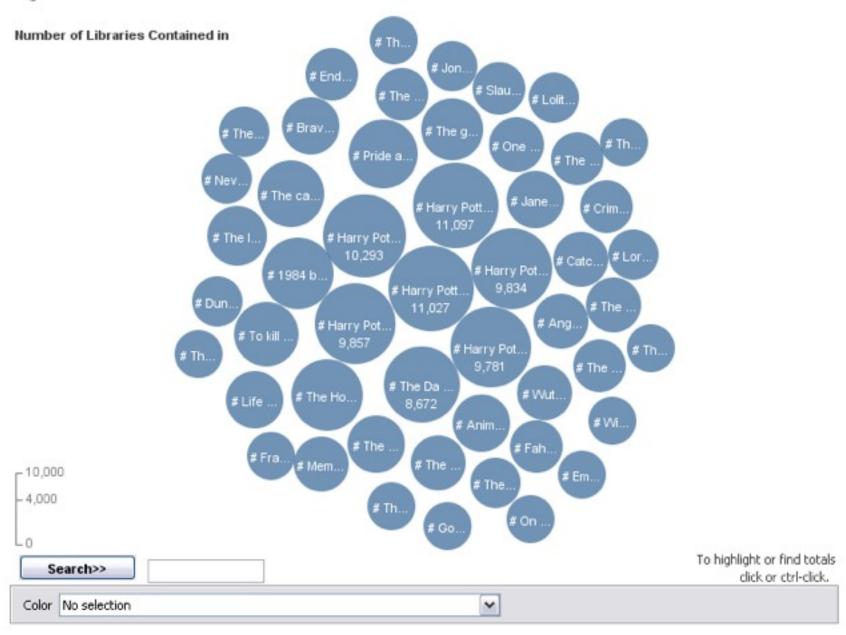
posted about 1 year ago

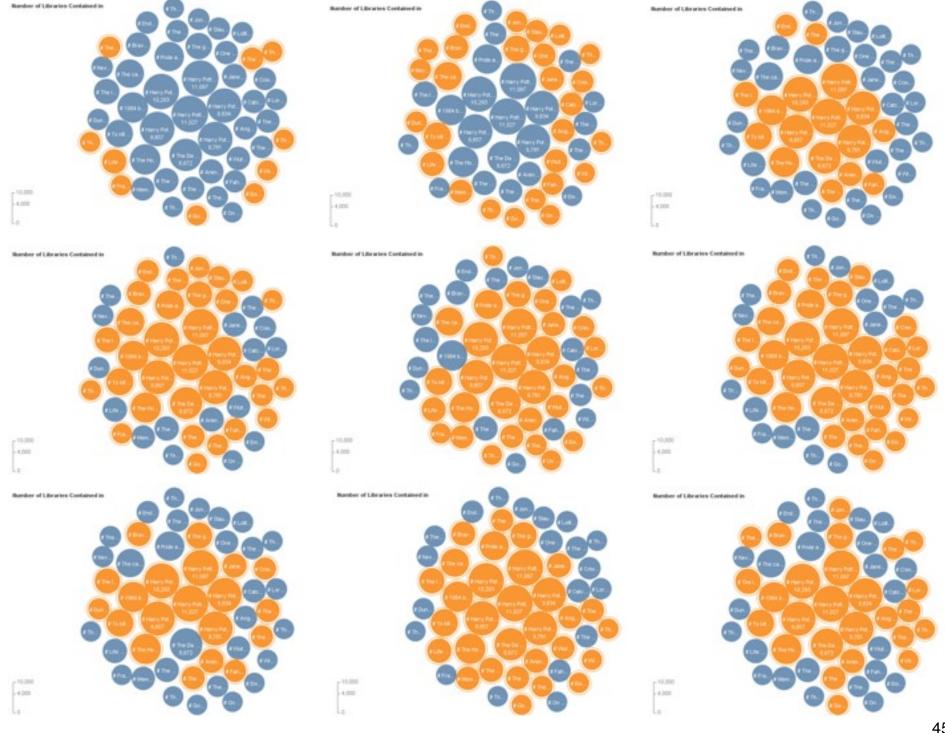
Heroin vs. Cocaine vs.
MDMA (Ecstacy) vs. Crystal
Meth. vs. Heroin vs.
Cocaine vs. MDMA
(Ecstacy) vs. Crystal Meth.
vs. Cocaine, NASDAQ
Adjusted Close, and Dow

Visualizations: Harry Potter is Freaking Popular

Creator: Alison

Tags:





What factors enable viable collaborations?

How might we design systems to facilitate social data analysis?

Administrivia

Final Project

Poster Presentations

Session is **Thu Mar 13 5-8pm** in CSE Atrium Bring **Poster + Laptop/Device** for demos Arrive early to setup!

Post Webpage on GitHub Pages

List team members, title, abstract, link to paper Include summary image for project!

Final Project Reports

Due **Thu Mar 20**, by **7am**, posted to GitHub 4-6 pages in ACM or IEEE TVCG format

Design Considerations for Collaborative Analysis

Modules of Contribution

Data Management

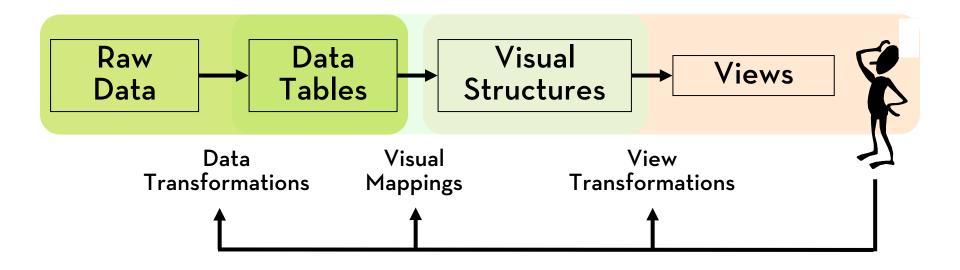
Contribute Data Clean Data Categorize Data Moderate Data Create Metadata

Visualization

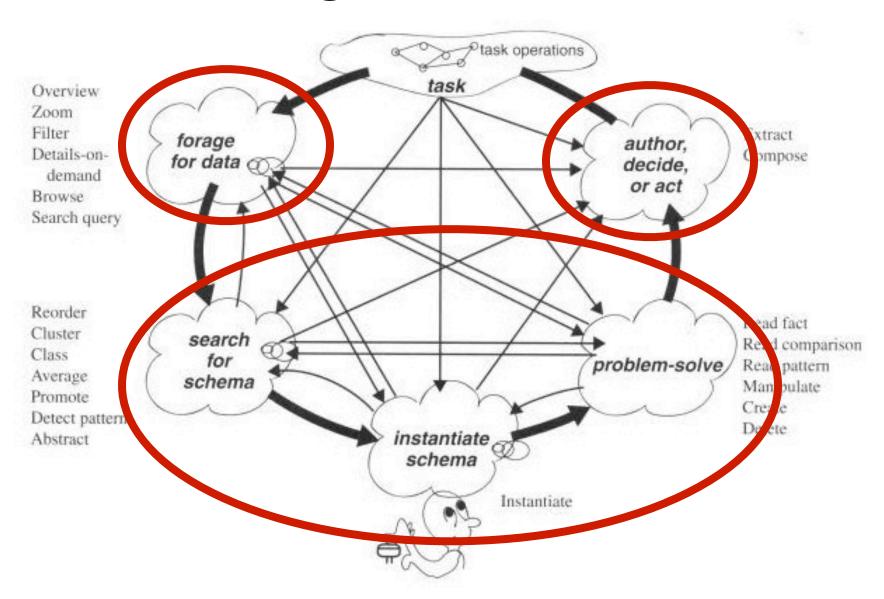
Select Data Sources
Apply Visual Encoding
Author Software

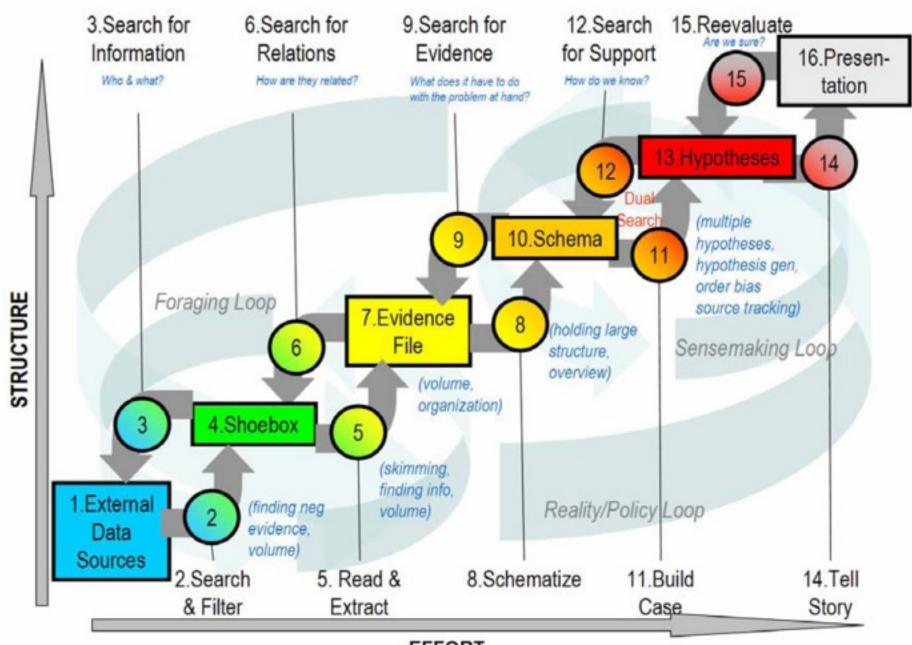
Visual Analytics

Observations
Hypotheses
Evidence (+/-)
Summarize
Report / Presentation



Sensemaking





EFFORT

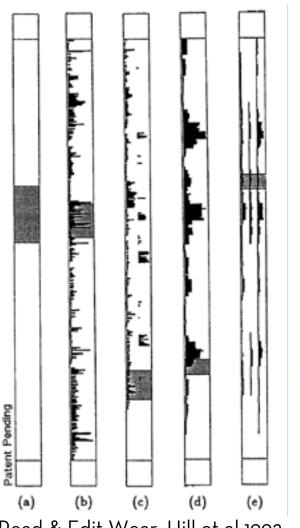
Design Considerations [VAST 07, IVS 08]

Division, allocation, and integration of work Common ground and awareness Reference and deixis (pointing) Identity, trust, and reputation Group formation and management Incentives and engagement Presentation and decision-making

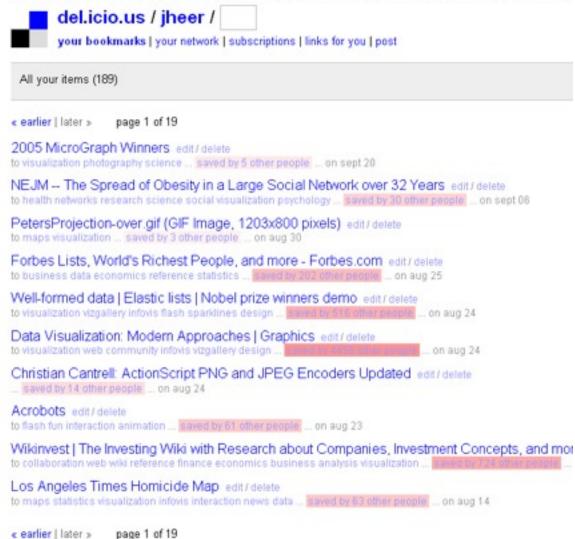
Social Data Analysis

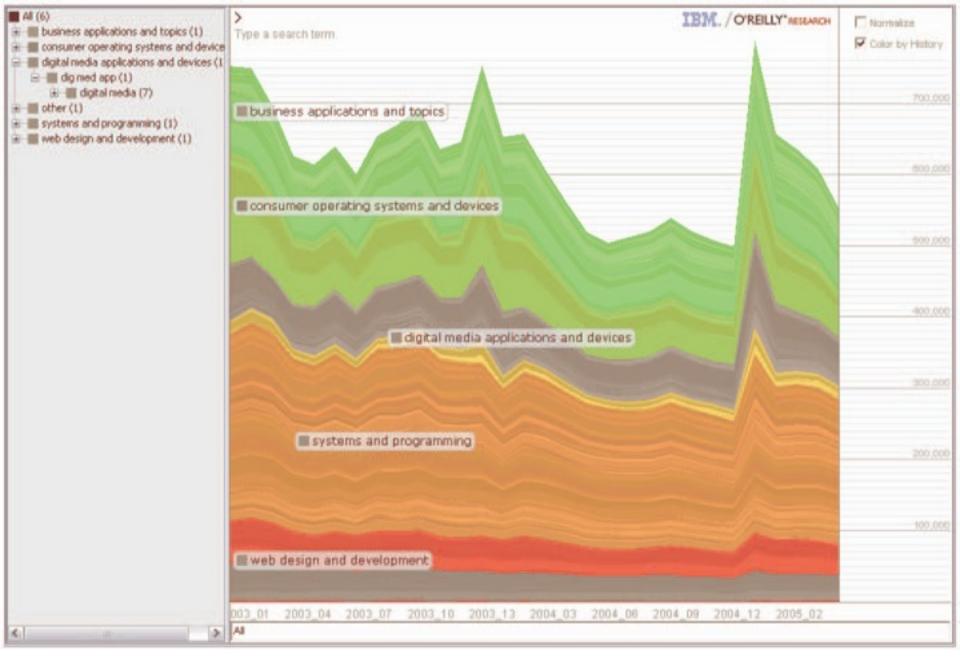
How can users' **activity traces** be used to improve awareness in collaborative analysis?

Social Navigation



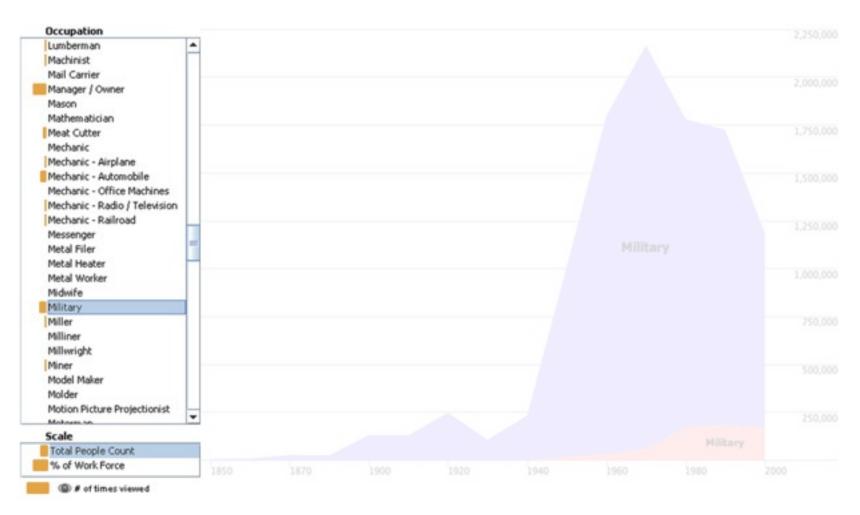
Read & Edit Wear. Hill et al 1992





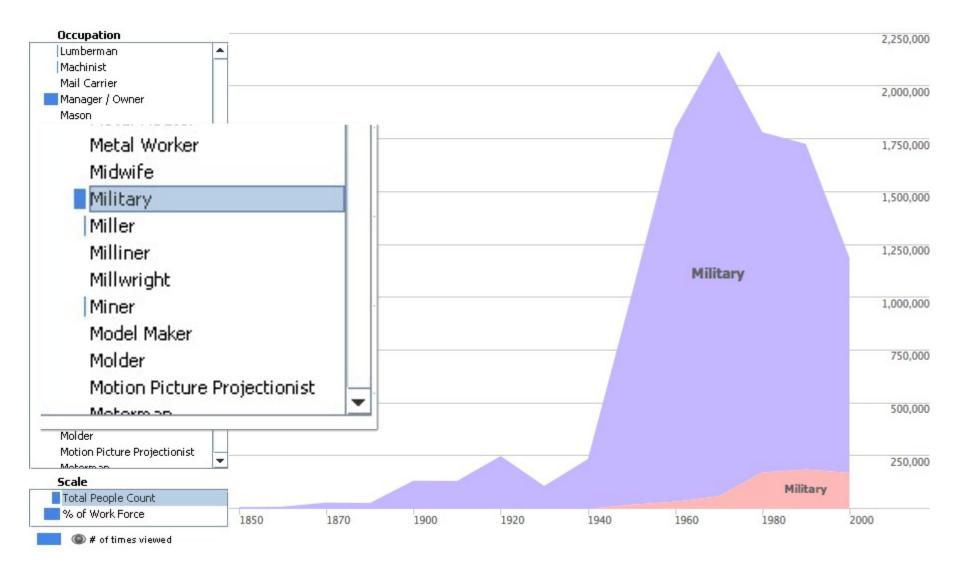
Wattenberg & Kriss - Color by history: gray regions have already been visited

Scented Widgets [InfoVis 07]

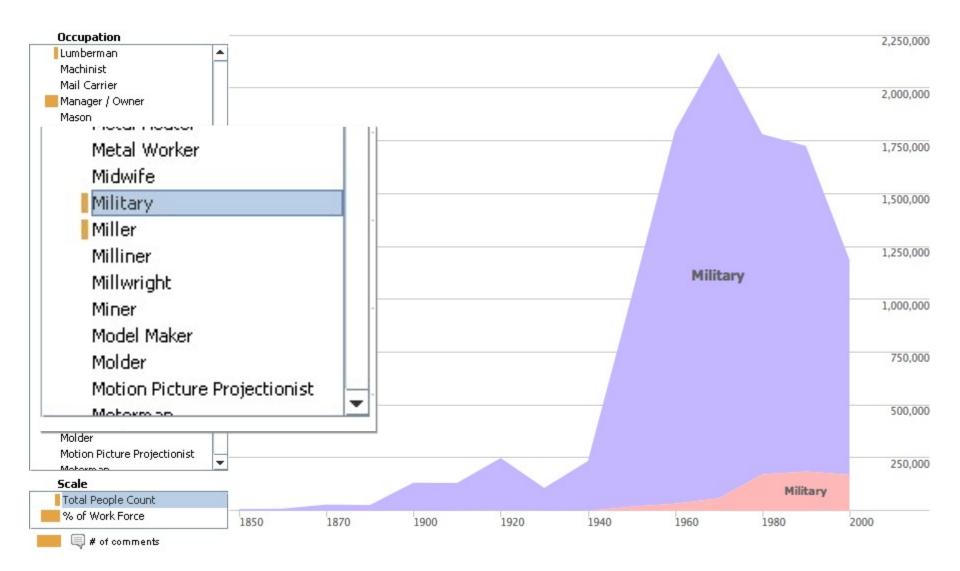


Visual navigation cues embedded in interface widgets

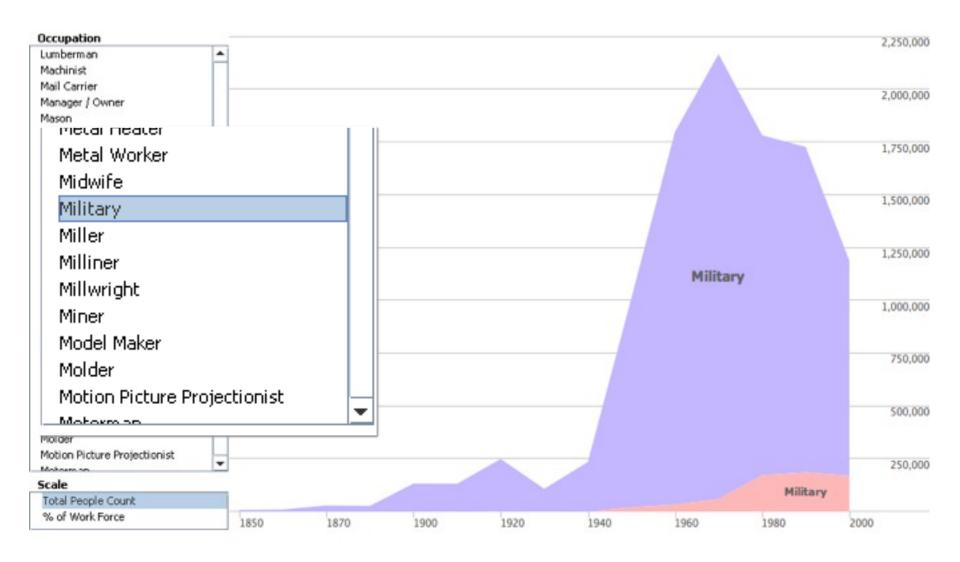
Visitation counts



Comment counts



No scent (baseline)



Do activity cues affect usage?

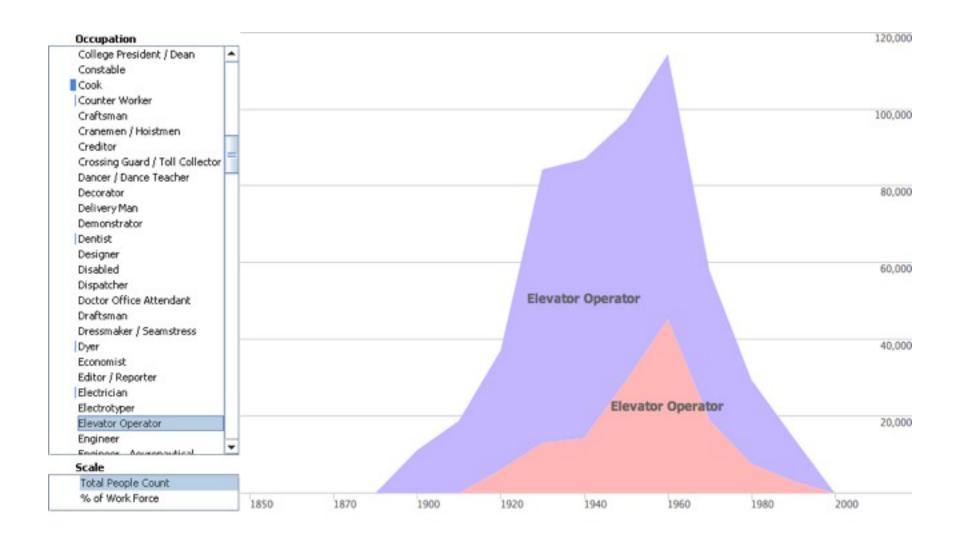
Hypotheses: With activity cues, subjects will

- 1. Exhibit more revisitation of popular views
- 2. Make more unique observations

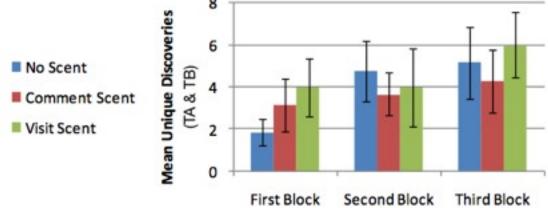
Controlled experiment with 28 subjects

Collect evidence for and against an assertion Varied scent cues (3) and foraging task (3) Activity metrics collected from sense.us study

"Technology is costing jobs by making occupations obsolete."



Results



Unique Discoveries

Visit scent had sig. higher rate of discoveries in first block. Less reliance on scent when subjects were familiar with data and visualization.

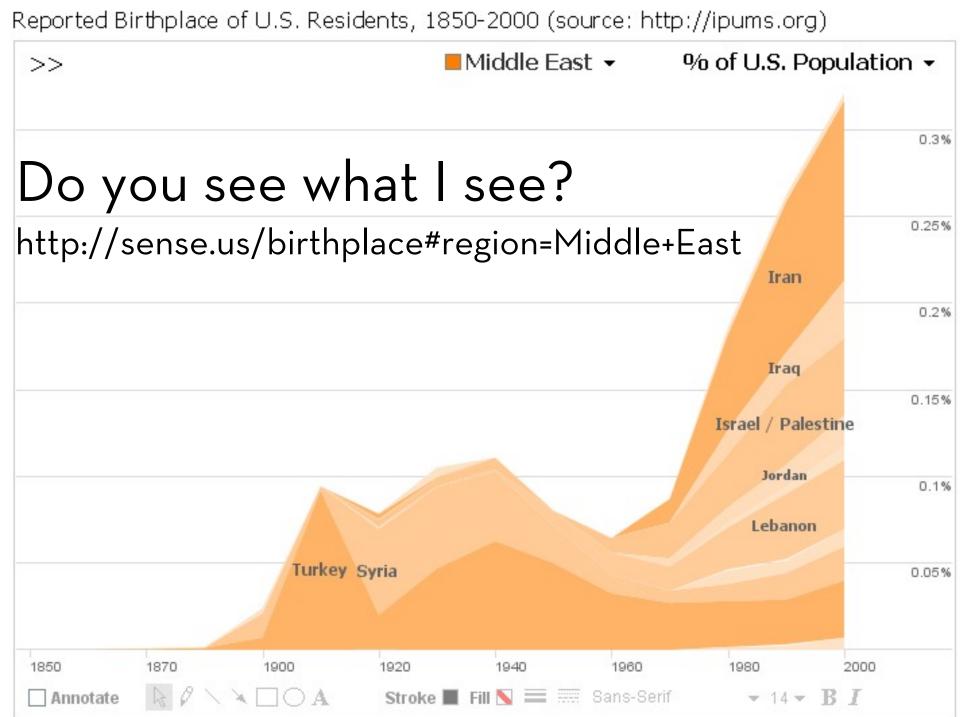
Revisitation

Visit and comment scent conditions correlate more highly with seed usage than no scent.

Social Data Analysis

How can users' **activity traces** be used to improve collaborative analysis?

How should **annotation techniques** be designed to provide nuanced pointing behaviors?



Common Ground

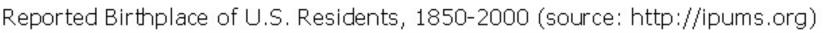
Common Ground: the shared understanding enabling conversation and collaborative action [Clark & Brennan '91]

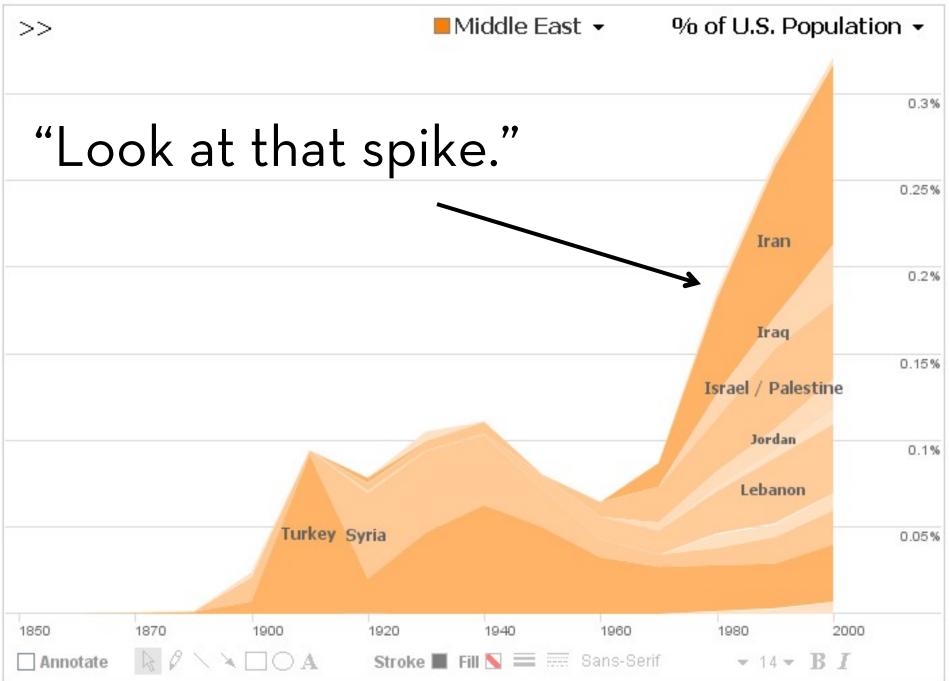
Do you see what I see? \rightarrow View sharing (URLs)

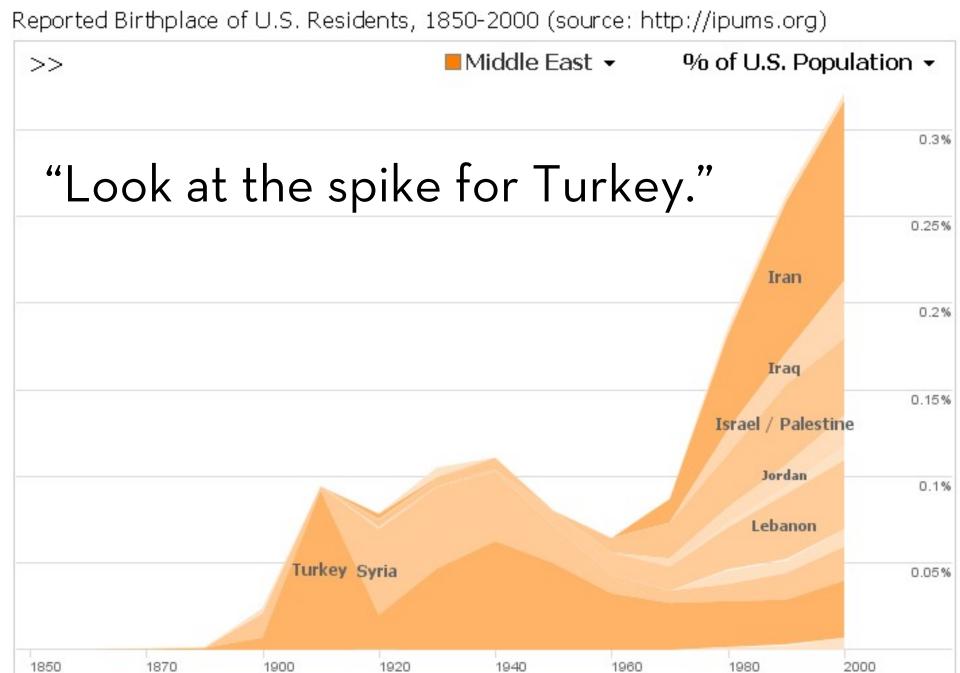
How do collaboration models affect grounding? Linked discussions vs. embedded comments vs. ...

Principle of Least Collaborative Effort: participants exert just enough effort to successfully communicate.

[Clark & Wilkes-Gibbs '86]

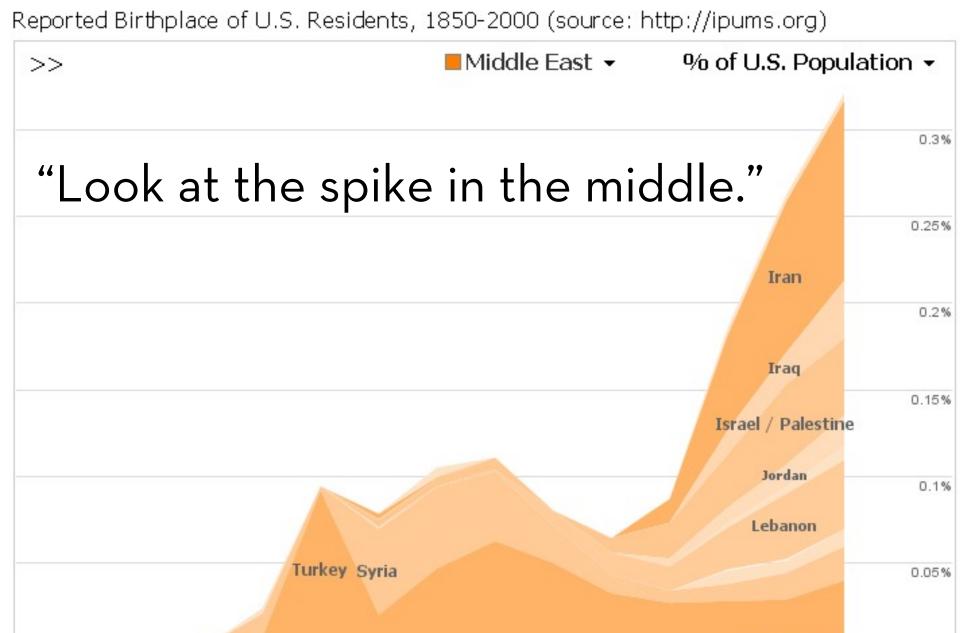






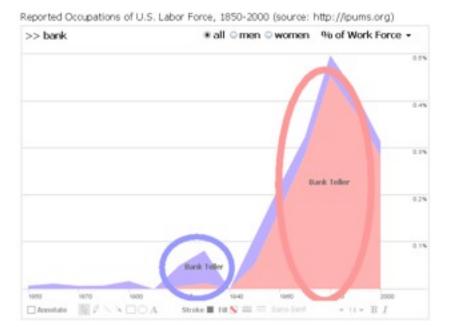
Annotate

Sans-Serif • 14 • B I



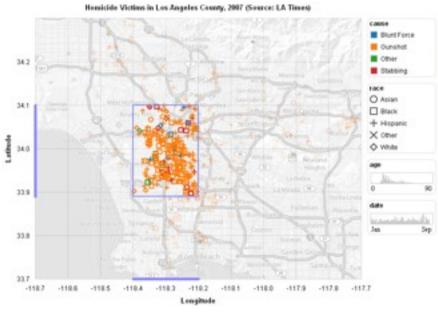
Sans-Serif • 14 • B I

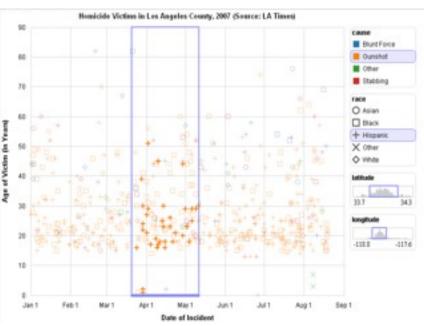
Annotate











Data-aware

Use of Annotations

Arrows	λ	25.1%	
Text	\mathbf{A}	24.6%	
Ovals	\circ	17.9%	
Pencil	Ø	16.2%	
Lines	\	14.5%	
Rectangles		1.7%	

39.0% of comments included annotations

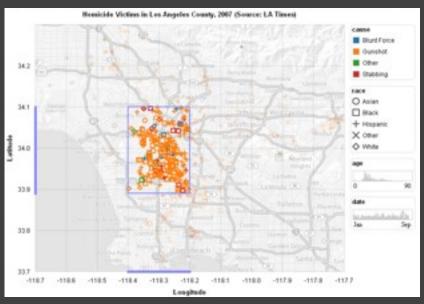
Pointing to specific points, trends, or regions (88.6%) Drawing to socialize or tell jokes (11.4%)

Variety of subject responses

'Not always necessary', but 'surprisingly satisfying' Some concern about professional look

Visual Queries

Model selections as **declarative queries** over interface elements or underlying data





 $(-118.371 \le lon AND lon \le -118.164) AND (33.915 \le lat AND lat \le 34.089)$

Visual Queries

Model selections as **declarative queries** over interface elements or underlying data

Applicable to dynamic, time-varying data

Retarget selection across visual encodings

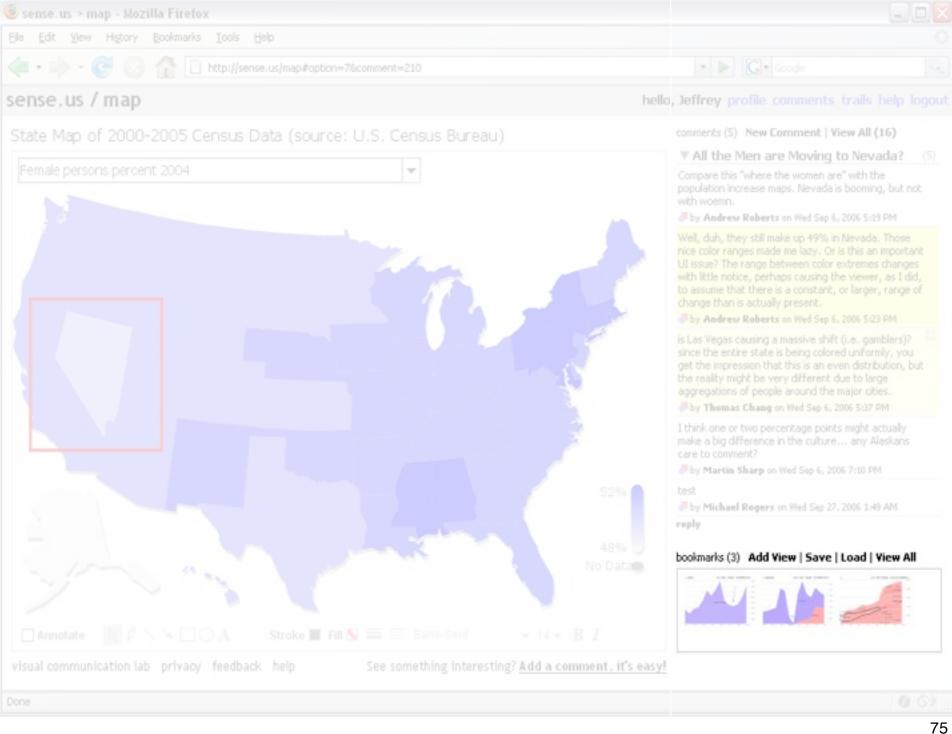
Support social navigation and data mining

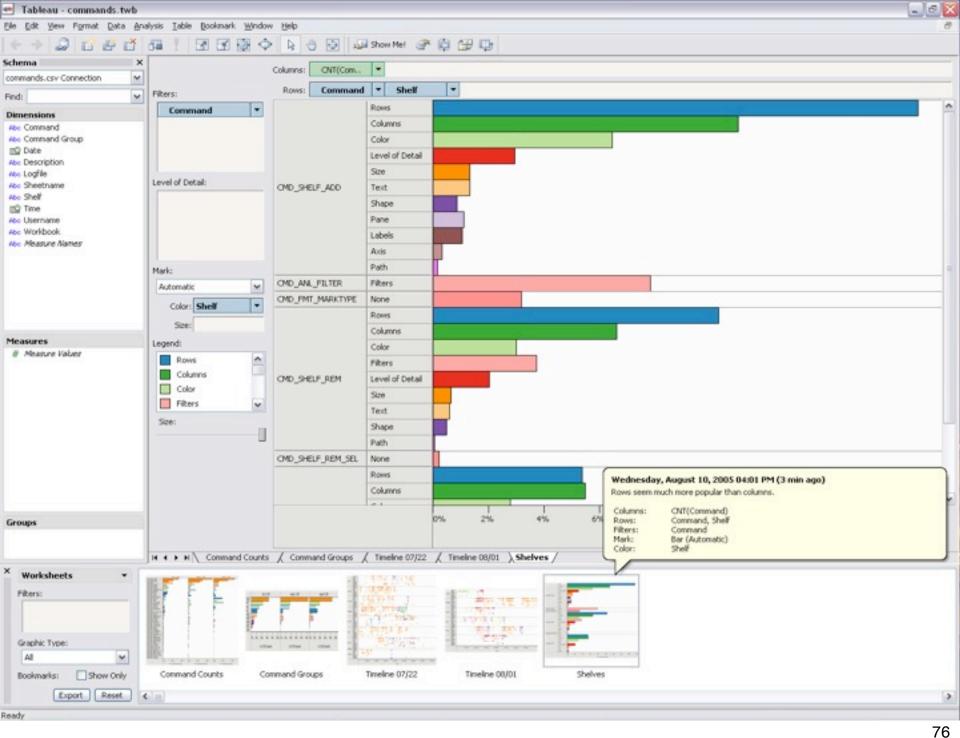
Social Data Analysis

How can users' **activity traces** be used to improve collaborative analysis?

How should **annotation techniques** be designed to provide nuanced pointing behaviors?

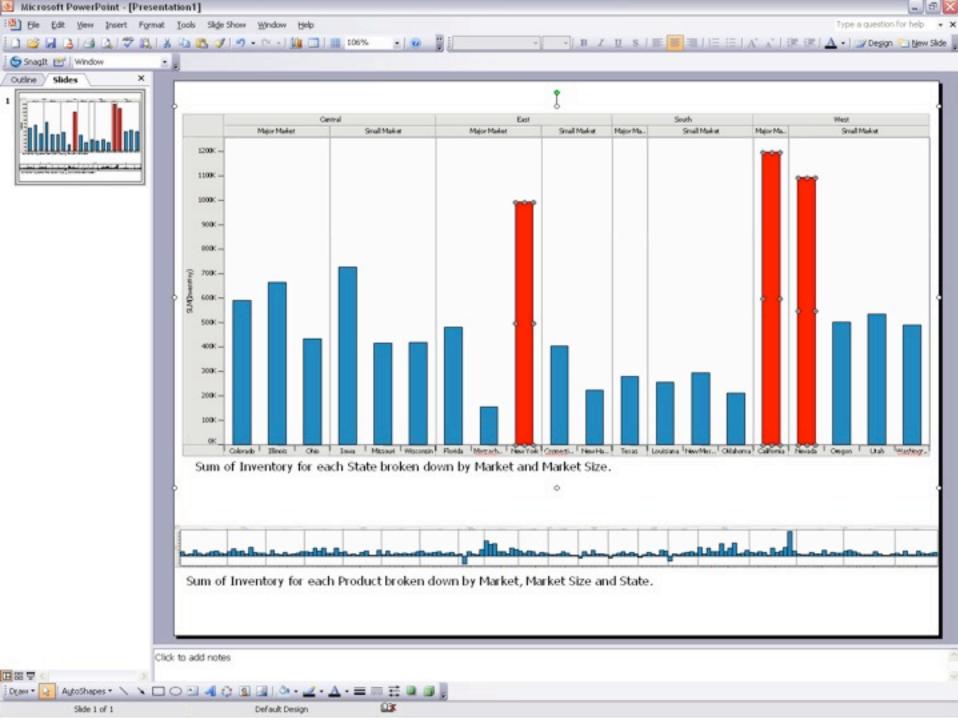
How can interface design better support communication of analytic findings?





Graphical Analysis Histories





Social Data Analysis

How can users' **activity traces** be used to improve collaborative analysis?

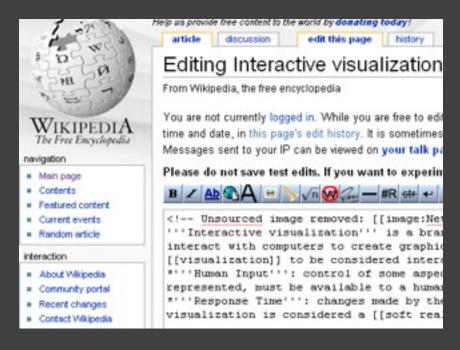
How should **annotation techniques** be designed to provide nuanced pointing behaviors?

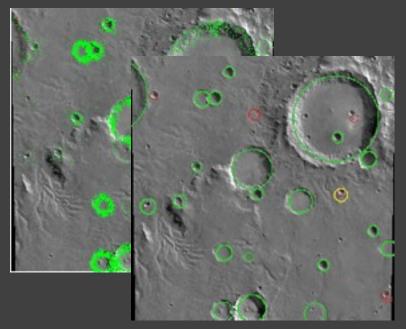
How can interface design better support presentation of analytic findings?

How can contributions be **better integrated**?

Structured Conversation

Reduce the cost of synthesizing contributions

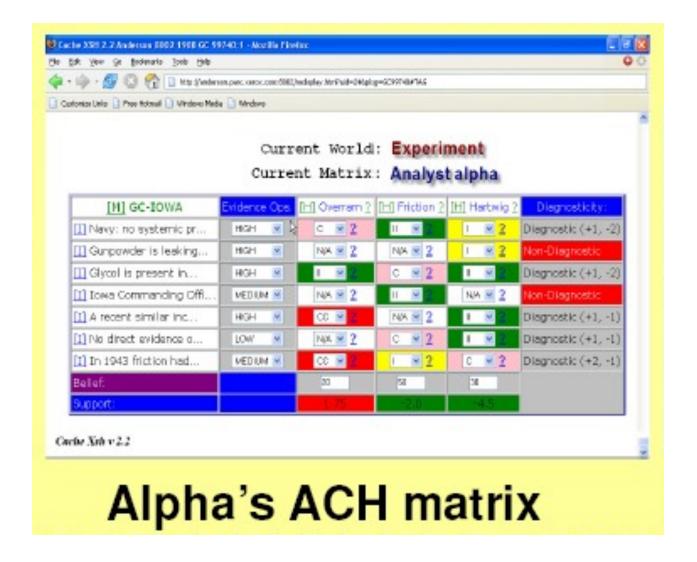




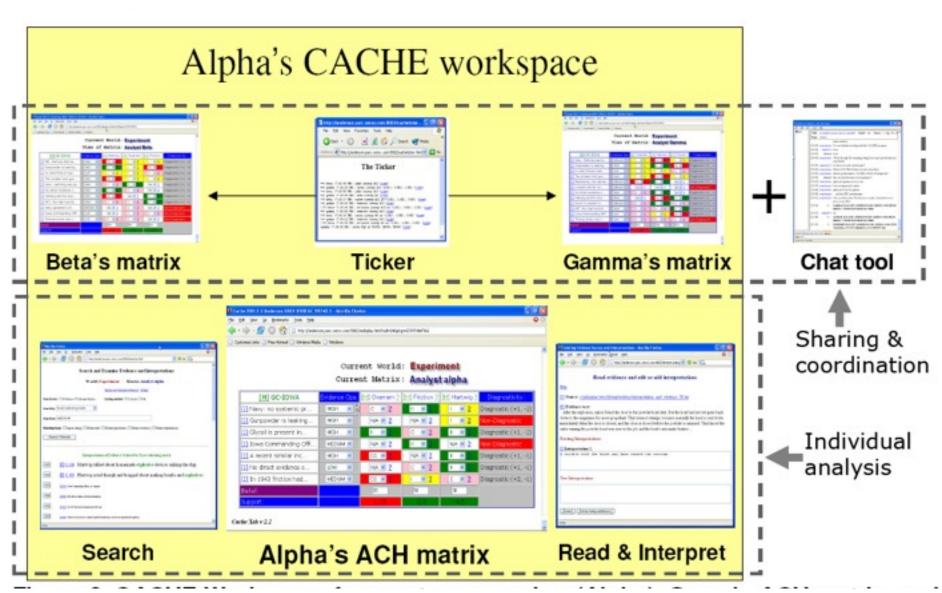
Wikipedia: Shared Revisions

NASA ClickWorkers: Statistics

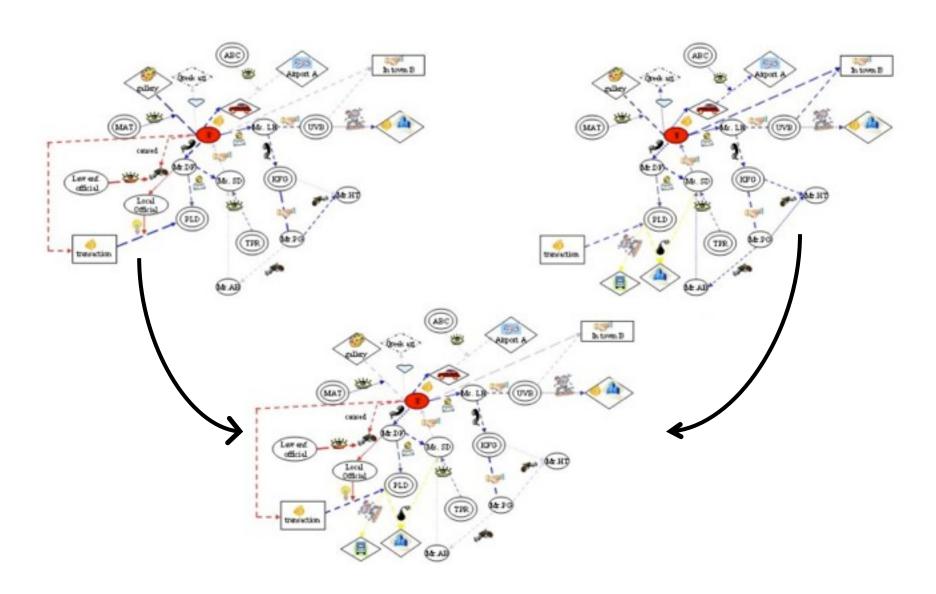
Integration: Evidence Matrices (Billman '06)



Integration: Evidence Matrices (Billman '06)



Merging Analysis Structures (Brennan et al '06)



Design Considerations [VAST 07, IVS 08]

Division, allocation, and integration of work Common ground and awareness Reference and deixis (pointing) Identity, trust, and reputation Group formation and management Incentives and engagement Presentation and decision-making

Ongoing Work

How to better structure analysis tasks?

Observe trends / patterns of interest Generate hypotheses Marshall evidence for/against a claim Structured tasks improve outcomes [Willett 2011]

How to better encourage participation?

Narrative: storytelling to spur exploration Financial incentives / crowdsourcing [Willett 2012]

Social Data Analysis

Visual sensemaking is a **social** process as well as a cognitive process.

Analysis of data coupled with social interpretation and deliberation.

How can user interfaces catalyze and support collaborative visual analysis?