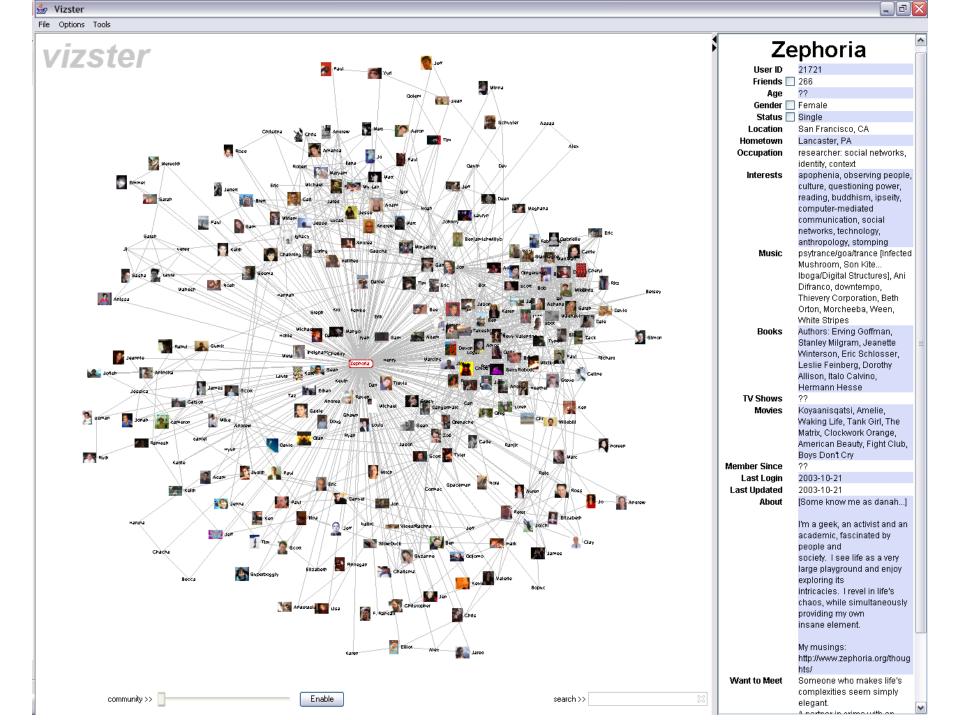
CSE 512 - Data Visualization

Collaborative Analysis



Jeffrey Heer University of Washington

A Tale of Two Visualizations

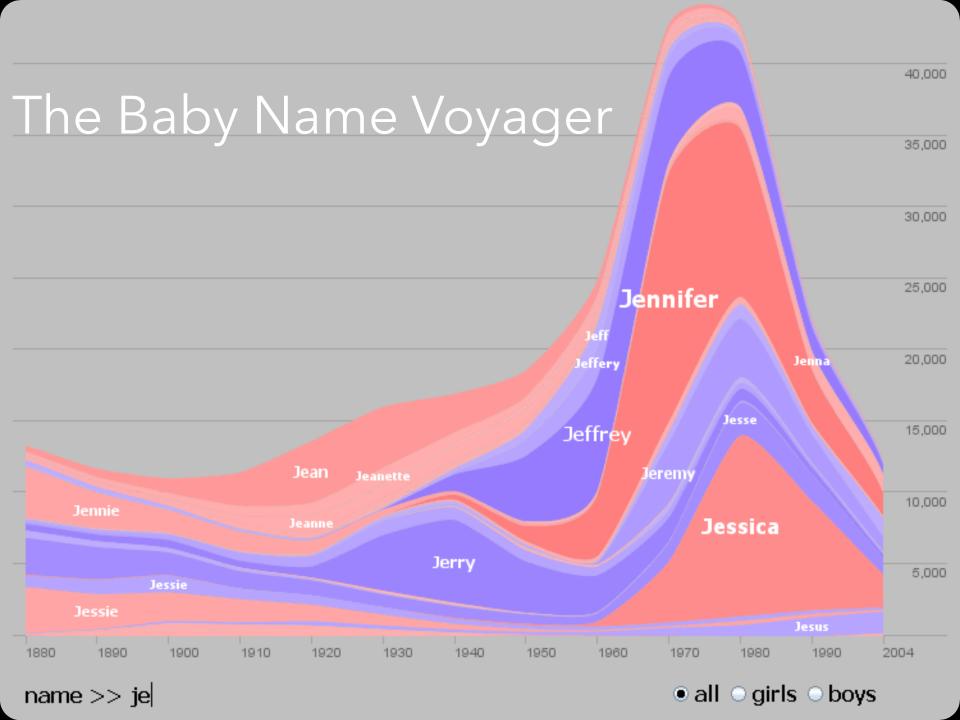


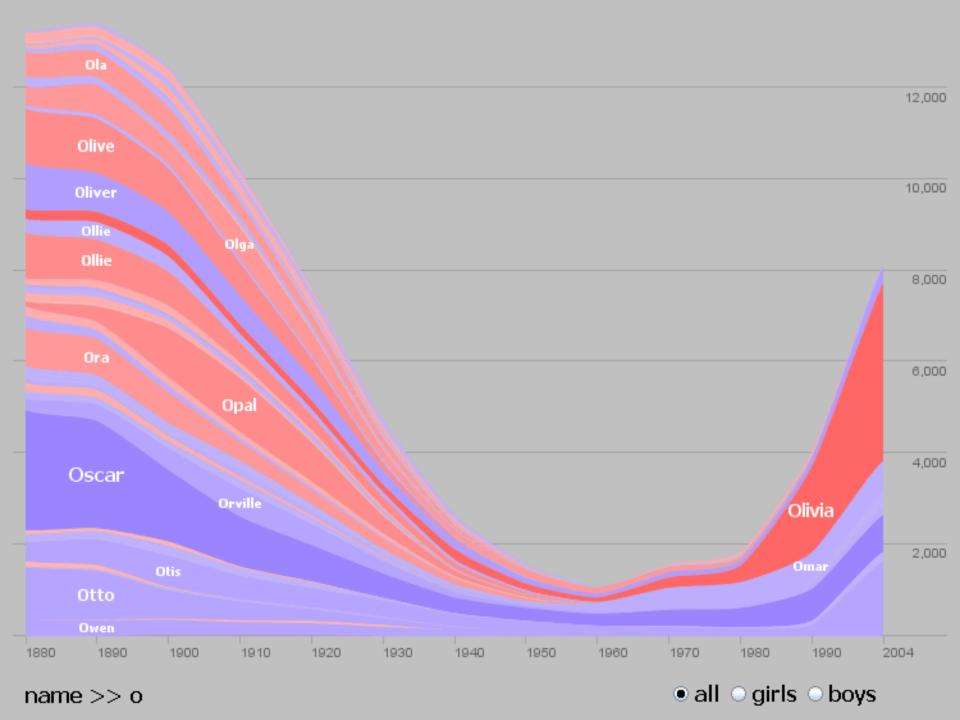
Observations

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.

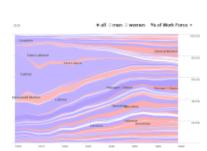
Analysis of data coupled with **social interpretation and deliberation**.

How can user interfaces catalyze and support collaborative visual analysis?

sense.us

A Web Application for Collaborative Visualization of Demographic Data





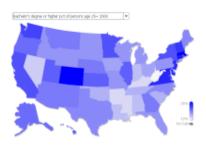
Job Vovager

Reported Occupations of U.S. Labor Force,

1850-2000

source: http://ipums.org

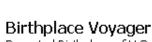
139 comments



U.S. Census State Map

State Map of 2000-2005 Census Data source: U.S. Census Bureau

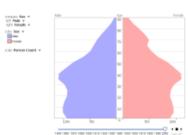
16 comments



Reported Birthplace of U.S. Residents, 1850-2000

source: http://ipums.org

10 comments



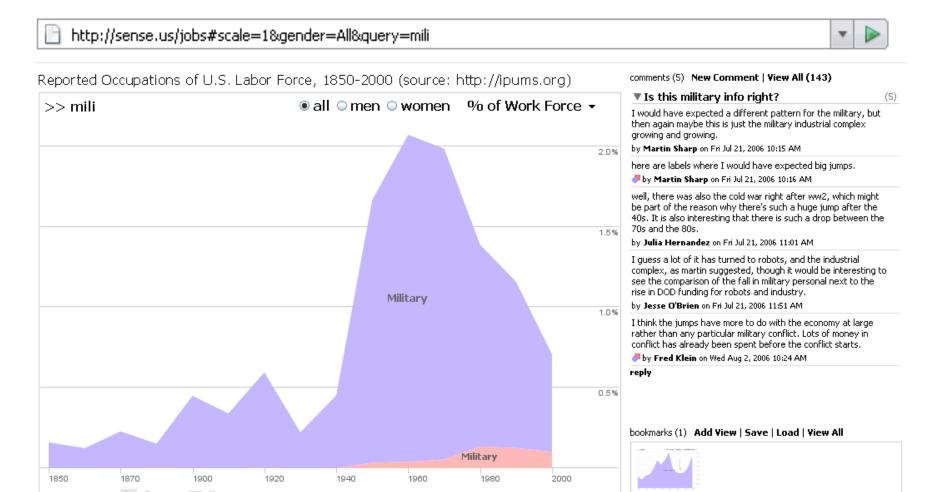
Population Pyramid

U.S. Population Demographics, 1850-2000 source: http://ipums.org

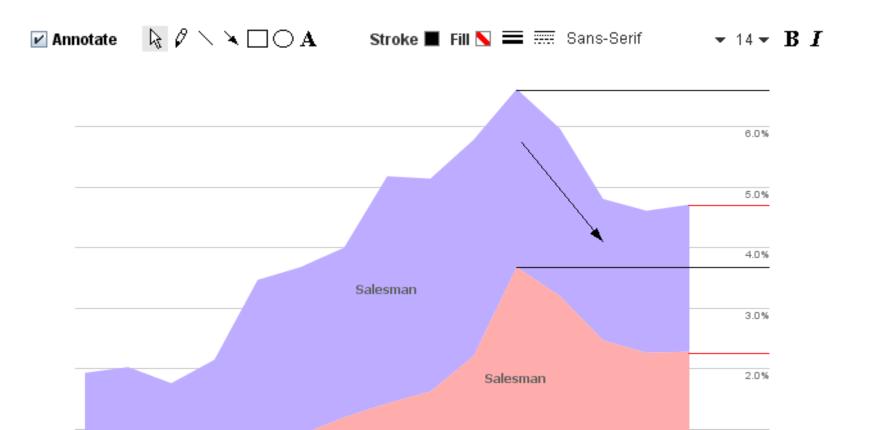
7 comments

[CHI 07]

Sharing within visualizations and across the web

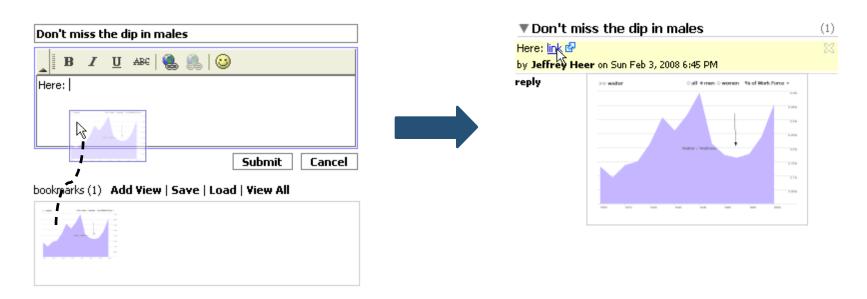


Sharing within visualizations and across the web **Pointing** at interesting trends, outliers



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



Ushers?

Is this like the movie theatre kind of usher? by Martin Sharp on Fri Jul 21, 2006 10:24 AM



Drop in household workers

This trend is almost the opposite of the "clerical worker" and "secretary" call by **Julia Hernandez** on Fri Jul 21, 2006 10:50 AM



Disappearing professors???

Note how a bunch of different kinds of professors "disappear" in the year 2 Is this just a change in the way the census aggregates data?

🏴 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, ...)

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, ...)

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

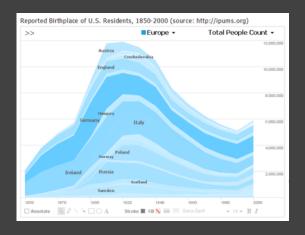
- **12.5** hours of qualitative observation
- **258** comments (41 pilot, 85 ibm, 60 ucb, 72 live)

Usage logs of user sessions

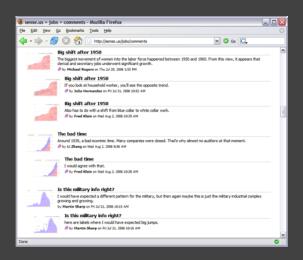
Voyagers & Voyeurs

Complementary faces of analysis

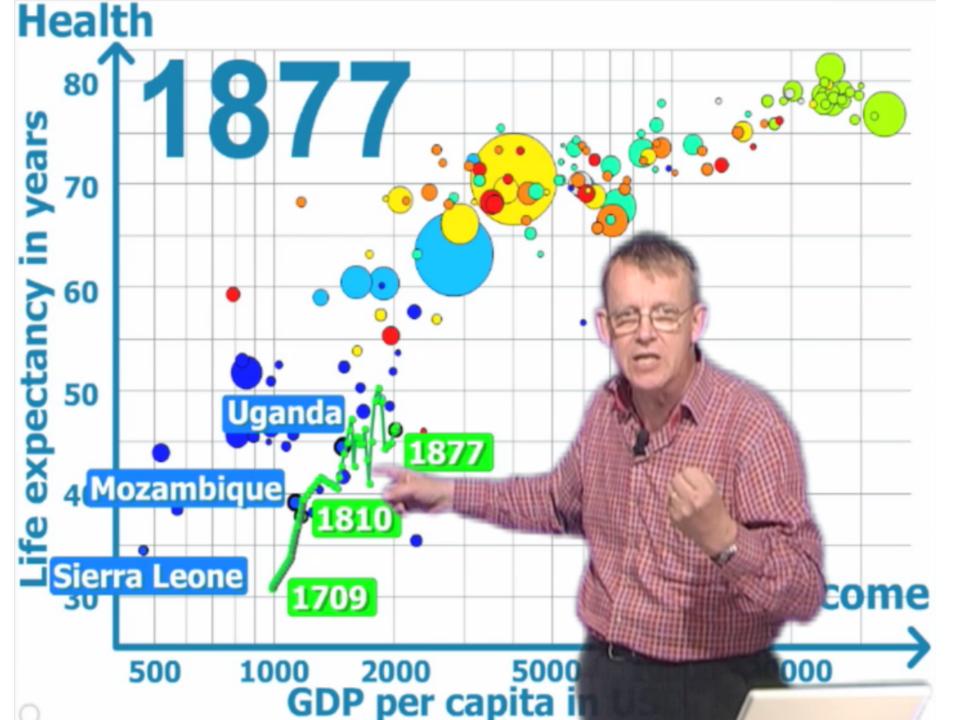
Voyager - focus on visualized data Active engagement with the data Serendipitous comment discovery



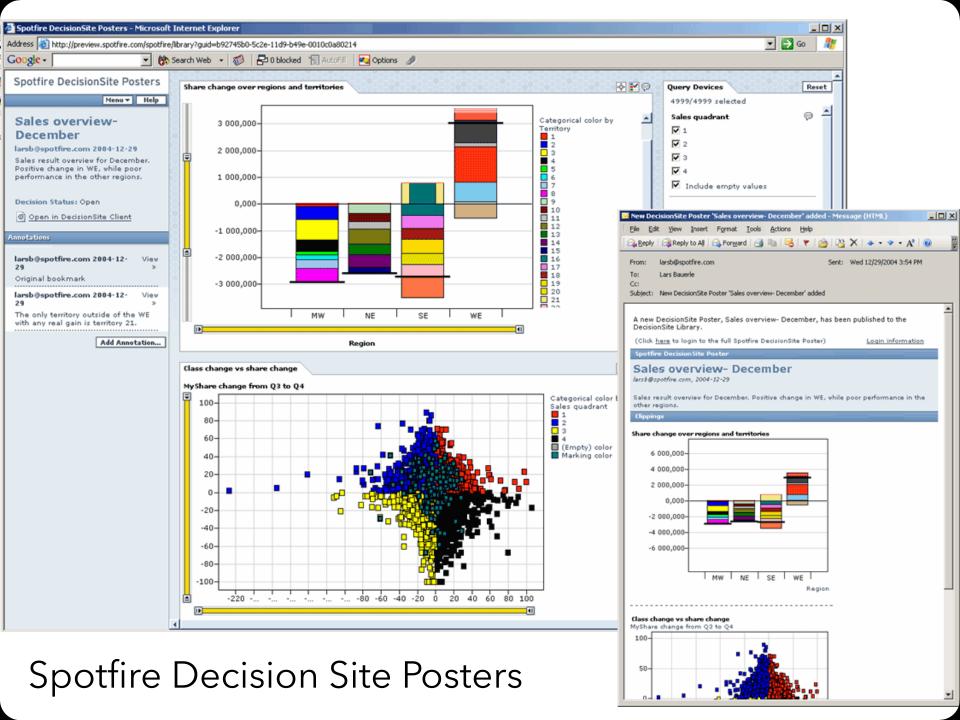
Voyeur - focus on comment listings Investigate others' explorations Find people and topics of interest Catalyze new explorations

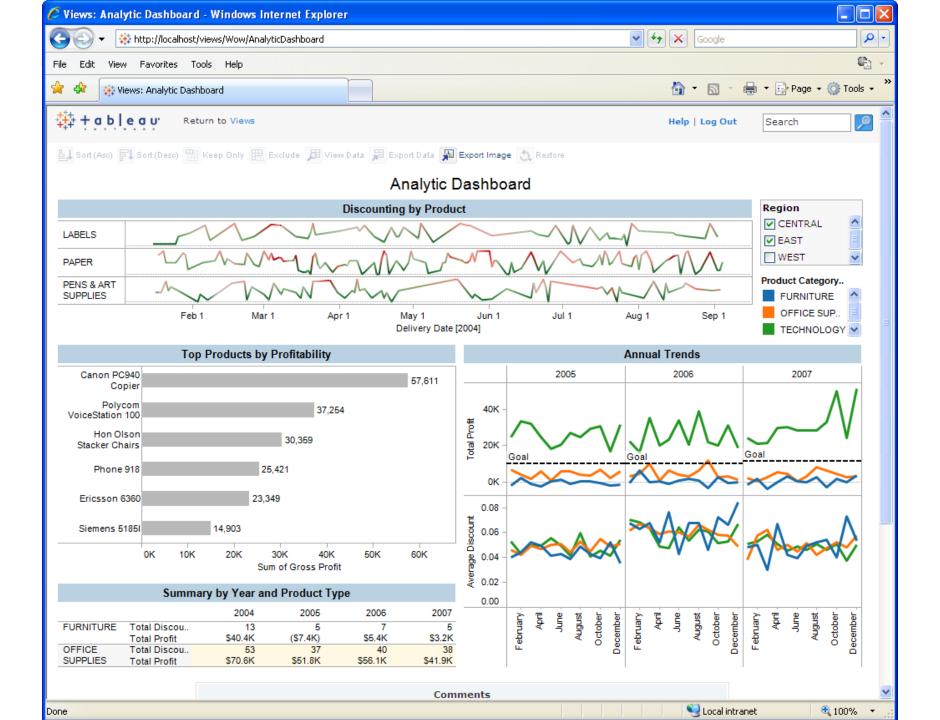


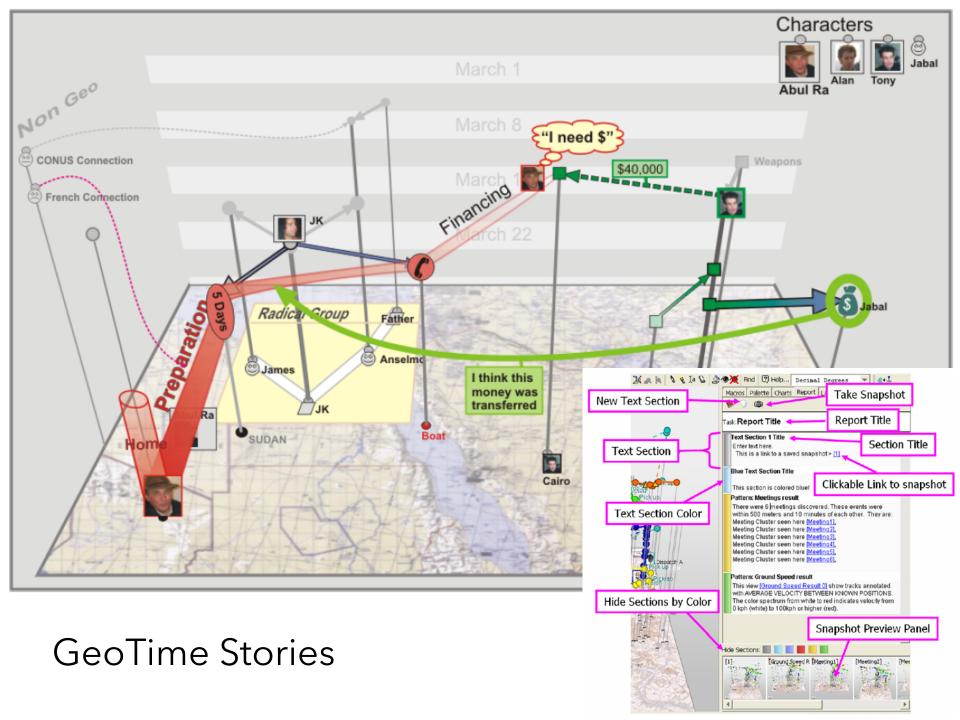
Social Data Analysis

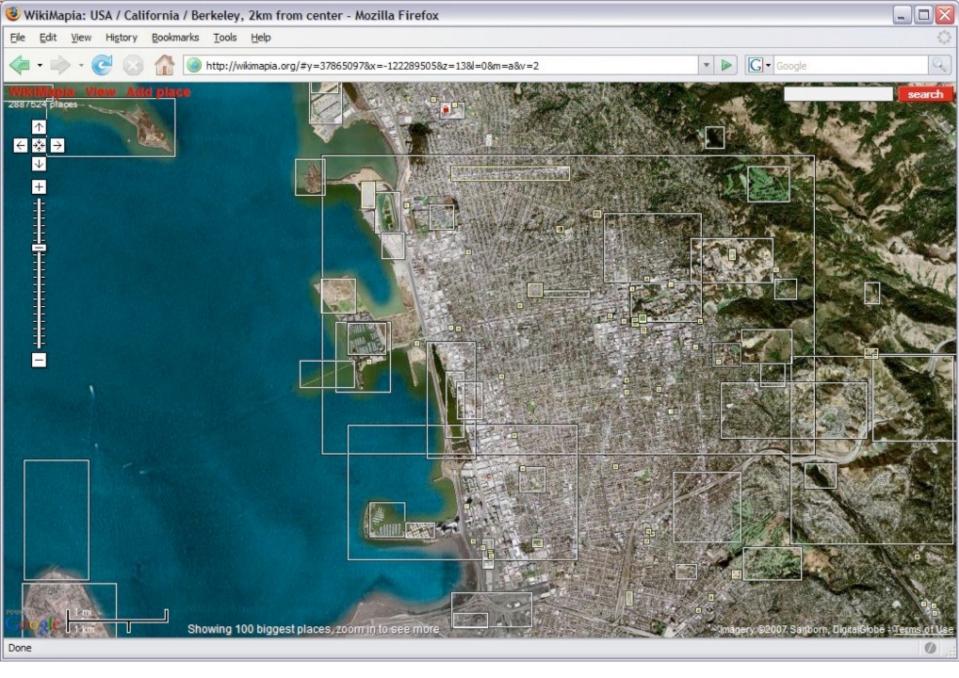












Wikimapia.org

Sign Up! Sign In Search

Swivel

Home



Data People Groups

Upload

-

Share this Graph

🏋 Post to Blog

11 diggs diggit

Rate It

Like It

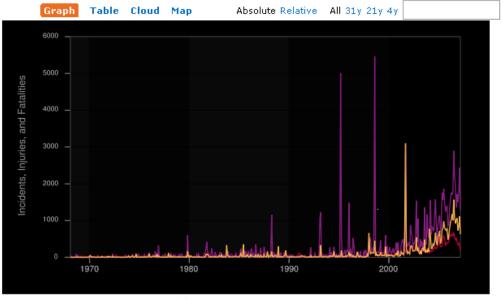
Send an Email

3 ratings

Sign in to rate

Feature It

The Status of George Bush's War on Terrorism





🦳 Bling 👅 Compare

Sources: MIPT Terrorism Knowledge Base (http://www.terrorismknowled...) Incidents, injuries, and deaths from terrorism have increased during the new millenium. This may be because of the broader definition of what characterizes a terrorist act, or it may actually be a constant increase in such acts. -seema

Comments (12)



Anonymous says

Go BuSH!

posted 11 months ago



Tyler says

What a completely disingenuous graph, and comments to follow. You have absolutely no reason to think the line would stay flat without Bush.

You probably thought graphs like these (and your site more widely) would help people overcome preconceived notions by showing hard data.

Instead, you're just amplifying people's preconceived notions -- adding legitimacy



More Options

By <u>seema</u> on Feb 22, 2007 Viewed 8927 times



Tags

no tags yet

Community Tags

Add tags

no tags yet

Correlations

<u>See all</u>

89% <u>Injuries</u> and <u>Fatalities</u>

Related Graphs



Yearly Terrorism Statistics

Created By: Dmitry Views: 2474



Global Terrorism Incidents and Iraq Body Count

Created By: Dmitry Views: 1868

Reported Maximum vs. The Status of George Bush's War on Terrorism

Created By: quest Views: 1192



Reported Maximum vs. The Status of George Bush's War on Terrorism

Created By: quest Views: 1181

Reported

Reported Maximum vs. The Status of Coorgo Buchle IAI







Sign in

explore visualizations data sets comments topic hubs

participate

register create visualization upload data set create topic hub

learn more

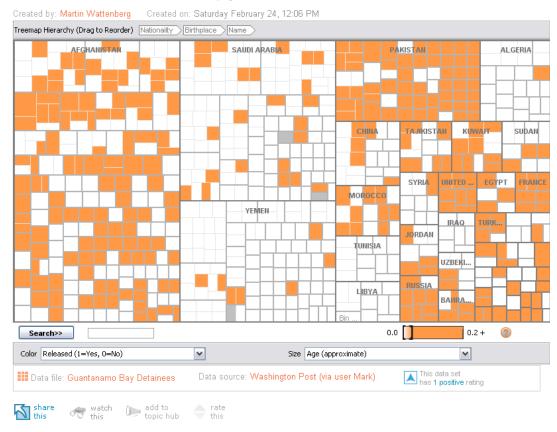
quick start visualization types data format & style about Many Eyes FAQ blog

contact contact us report a bug

legal terms of use

Visualizations: Guantanamo Bay Detainees, release status & age

Can't see the visualization? Download the latest Java plugin here. On Macs: best viewed in Safari.



Comments (4)



Martin Wattenberg says:

In this view, orange means released, white means not released. Gray means committed suicide.

Posted Saturday February 24, 12:07 PM see view for this comment



Martin Wattenberg says:

I'd be curious to hear ideas on why various countries have the release proportions that they do.

Posted Saturday February 24, 12:13 PM see view for this comment



Want to keep track of this visualization? Add it to your watchlist!

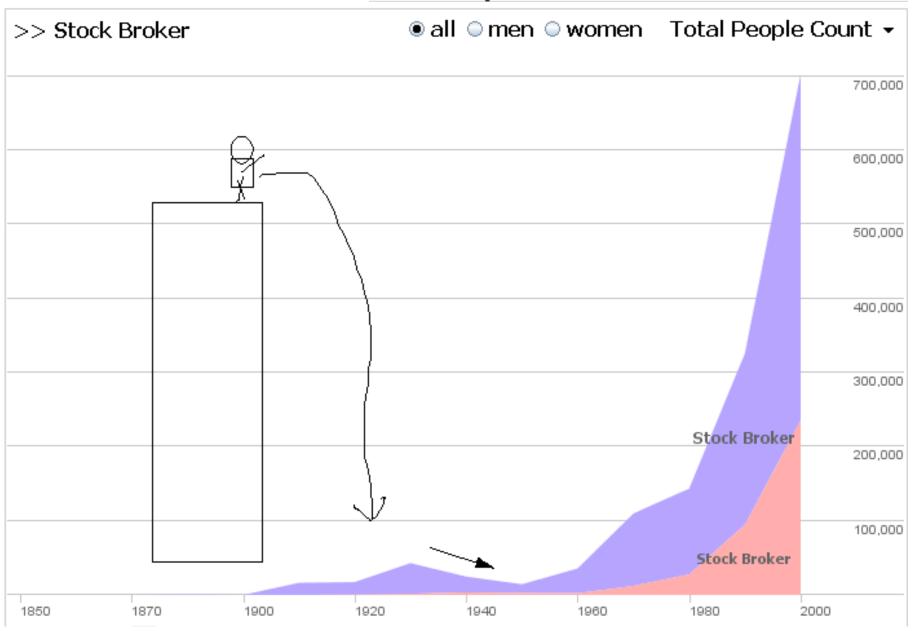
Learn more:

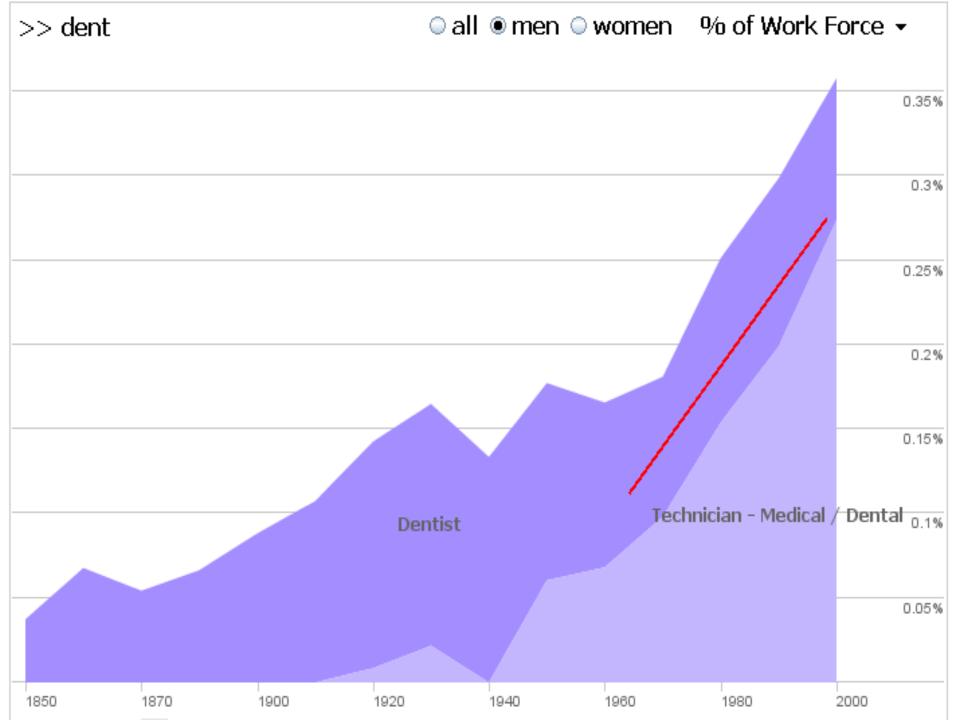


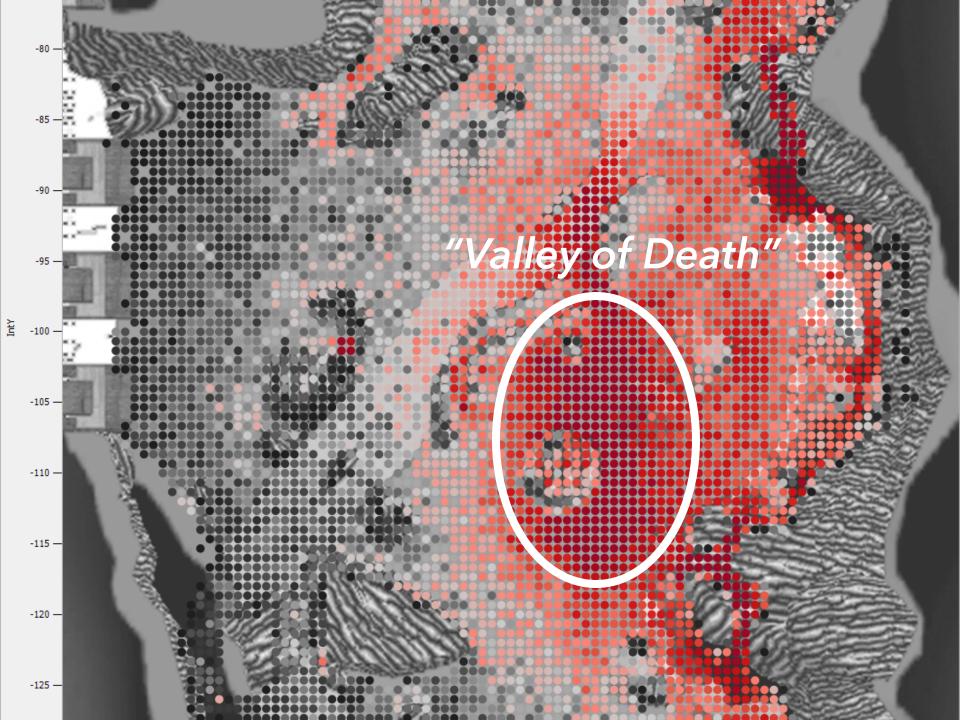


Discussion

Great depression "killed" a lot of brokers











visualizations data sets comments topic hubs

participate create visualization upload data set create topic hub register

learn more quick start visualization types data format & style about Many Eyes FAQ blog

contact Us contact report a bug

legal terms of use

Popular Dataset Tags
2007 2008 bible blo
books CENSUS crim
education eharmor
election energy for
health inauguratic
internet ireland literatu
lyrics media mus
network oban

people politic population president prices religion

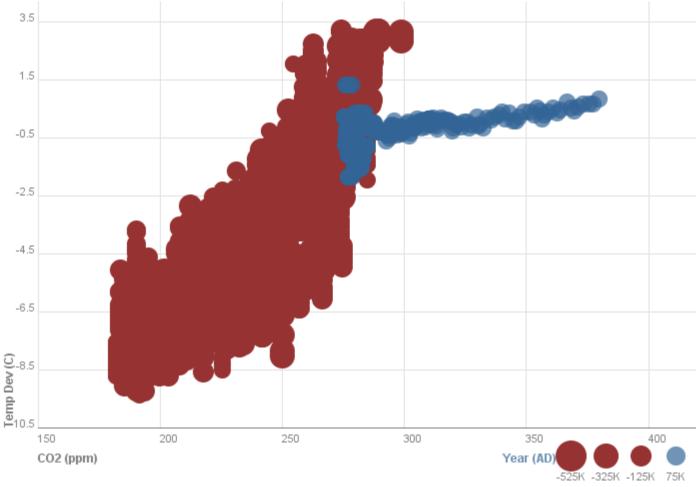
speech state state statistics supply-cha survey text US web World

See Mr

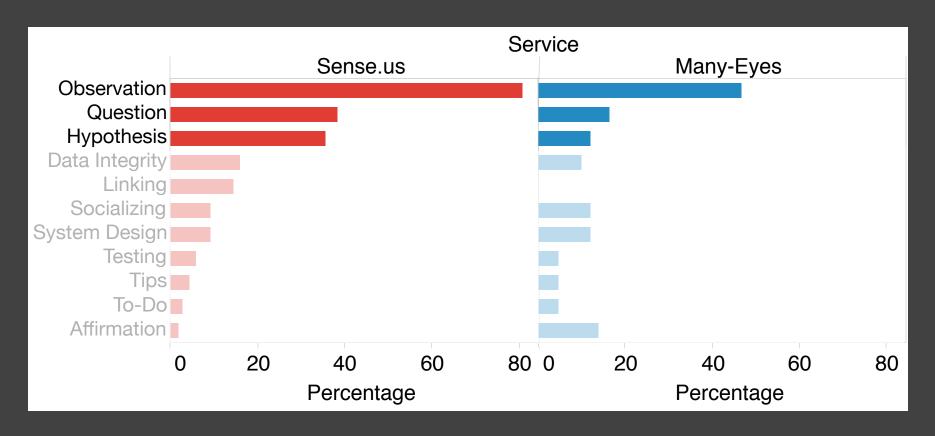
Visualizations: Flawed Data - Temp Over Time and CO2 Levels: This view shows the correlation between...

Creator: Bruce



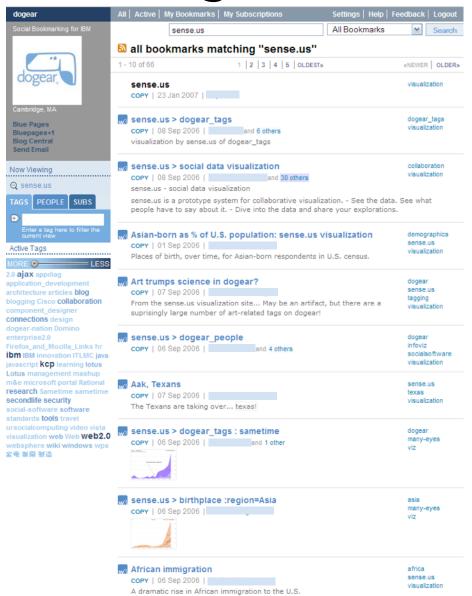


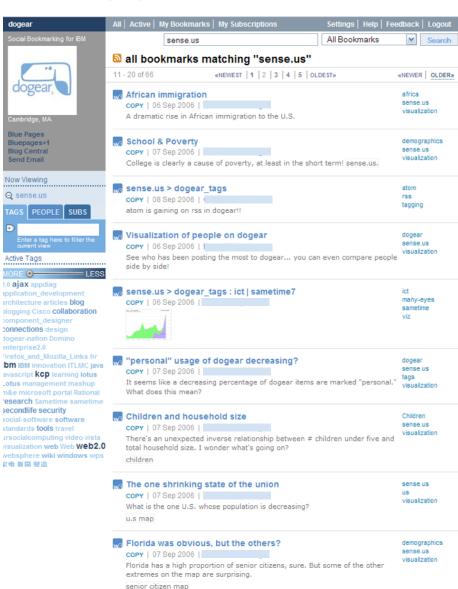
Content Analysis of Comments



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

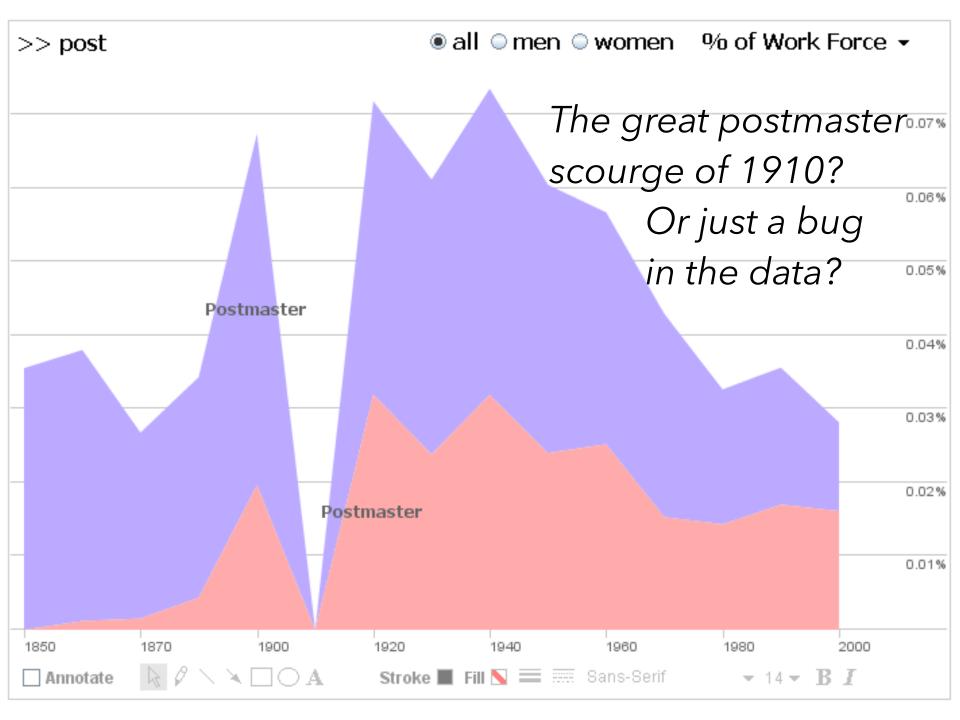
Sharing in External Media

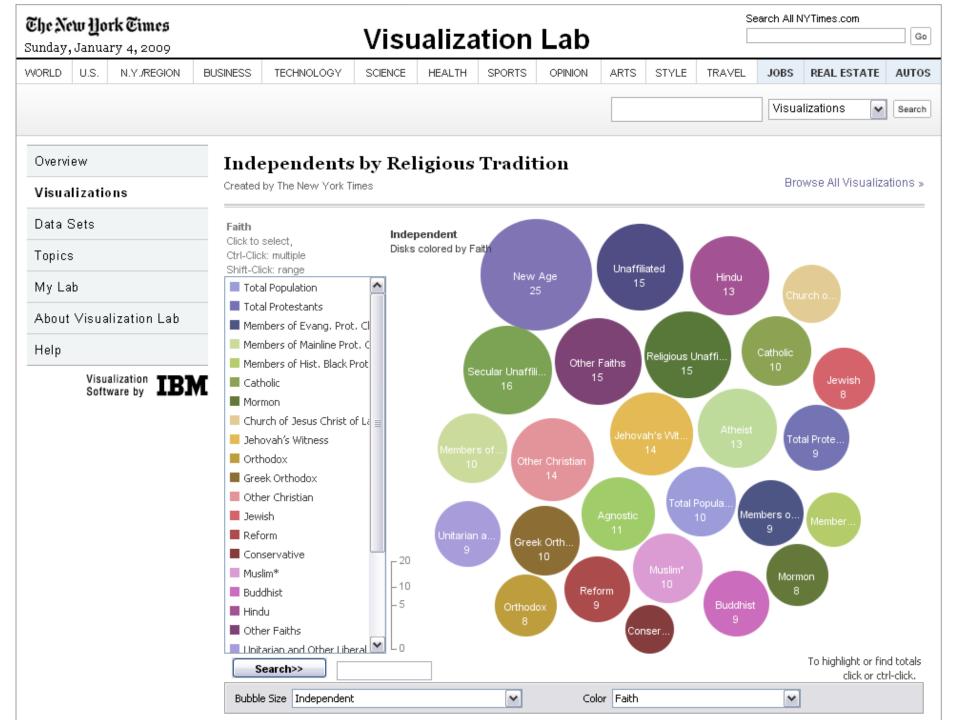




Data Quality

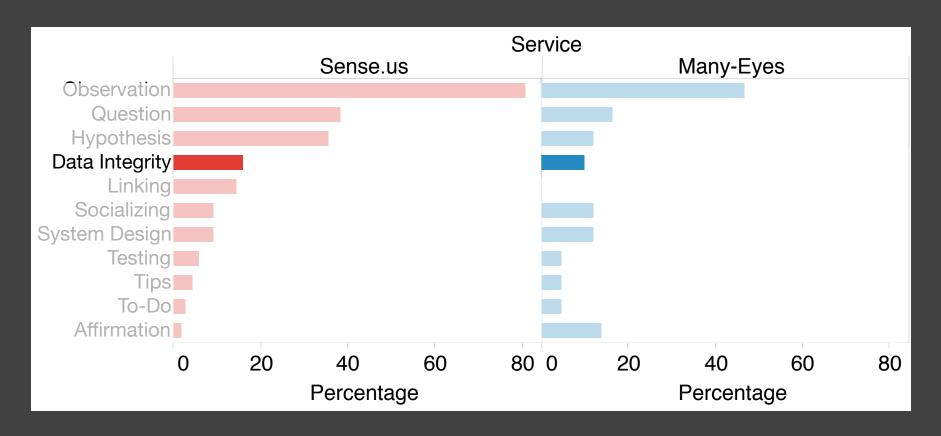






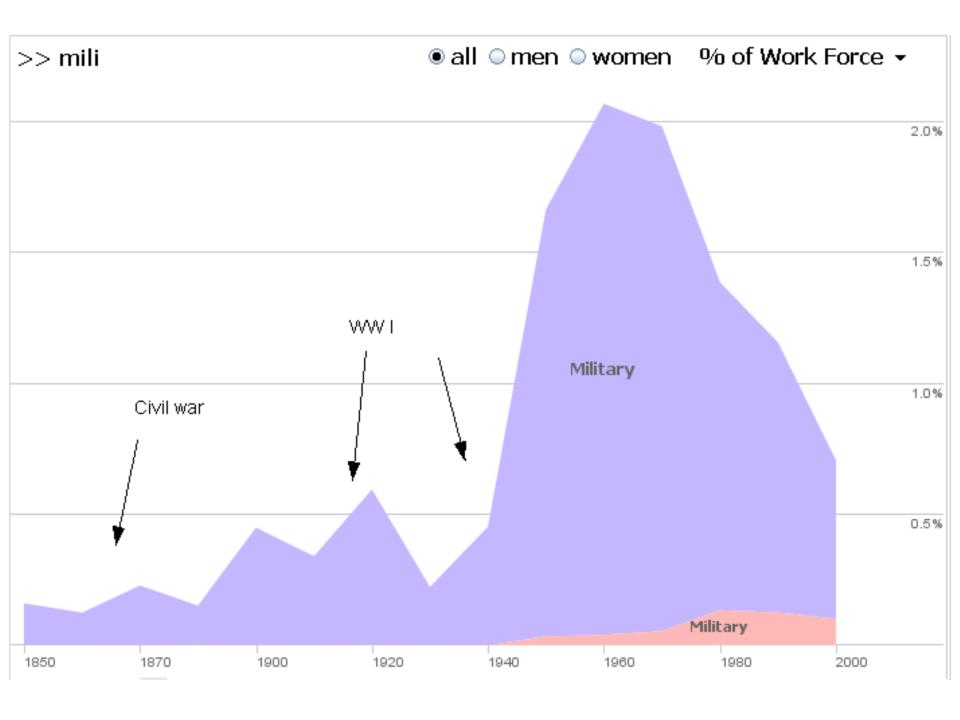
The N	ern Ho	rk Times										Search All NYTimes.com		
The New York Times Sunday, January 4, 2009				Visualization Lab										Go
Sunday	, Janua	ry 4, 2009	I						_					
WORLD	U.S.	N.Y./REGION	BUSINESS	TECHNOLOGY	SCIENCE	HEALTH	SPORTS	OPINION	ARTS	STYLE	TRAVEL	JOBS	REAL ESTATE	AUTOS
												Visua	ilizations 🕒	Search
			.											
Overv	iew		Inde	pendents	by Relig	gious T	raditioi	1 > Um	, I be	lieve	Churc			
Visua	alizatio	ns	Created	by The New York T	imes							Bro	wse All Visuali:	zations »
Data	Sets		Faith		Indo	pendent								
- ·			Click to			pendent colored by F	aith							
Topics			Ctri-Click: multiple											
My L	ah			al Population	^		New 2:		15		Hindu			
, 2				al Protestants			2				13	Chu		
Abou	t Visua	lization Lab		nbers of Evang, Pro										
Help Visualization IBM Software by			nbers of Mainline Pro											
			Members of Hist, Black Prot Secular Unaffili Other Faiths 15 Catholic 10 Lewish											
	Soft	ware by	Mor				16						8	
				mon irch of Jesus Christ (et i –									
				ovah's Witness	, LG =				Jehova	ah's Wit	Atheist	Tot	al Prote	
			Ort			Member				14	13		9	
				ek Orthodov		10	Othe	er Christian 14						
-														
1000	000	Anon	ymous says:						Agnostic		Popula	mbers o		
The second secon	00 0			eve Church of Jesus Christ of Latter-day Saints								9	Member	
The state of the s	0000	and f	Mormon are	the same?			Gree	k Orth						
(Marci)		was congress						10		Muslim*				
			■ Bud	ldbict	-10			Ref	orm			Morn 8	non /	
			Hin		-5		Orthod	_			Buddhist			
				er Faiths			8		Carr	sor	9			
				er Faiths :arian and Other Lihr	wal 🕶 Lo				Con	ser				
				earch>>	ial — I							Highlighte	d: 1 item and 2 c	ategories
				car cn>>									total	15 = 5%
			Bubbl	e Size Independent			~	Col	or Faith			~		

Content Analysis of Comments



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

Data Integration In Situ



Adjusted Clase

Home

Graphs

People Data

Groups

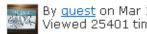
Upload

Blog Help Feedback Sign Up!

Search

Sign In

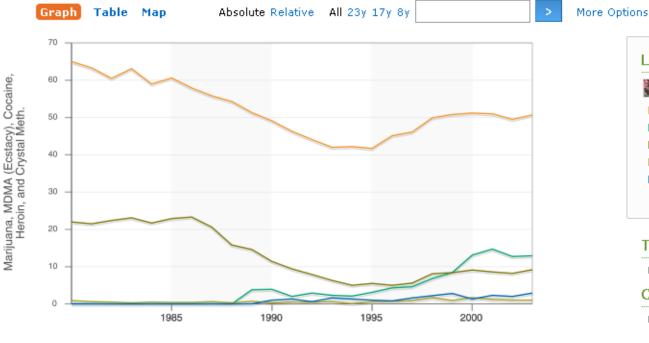
Marijuana vs. Heroin vs. Cocaine vs. MDMA (Ecstacy) vs. Crystal Meth.



By guest on Mar 12, 2007 Viewed 25401 times



Have you tried Swivel Business?



Compare

Blina (

Legend

Confectionary

College Student Drug Use

💻 Marijuana

MDMA (Ecstacy)

Cocaine

Heroin

Crystal Meth.

More >

Tags

no tags yet

Community Tags

Add tags

no tags yet

Correlations

See all

93% Marijuana and Cocaine

Related Graphs



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

Rate It

A O ratings

Digg submit

Share this Graph

Send an Email

Post to Blog

pot pot pot pot weed weed weed posted about 1 year ago

visnu says

Sources: University of Michigan

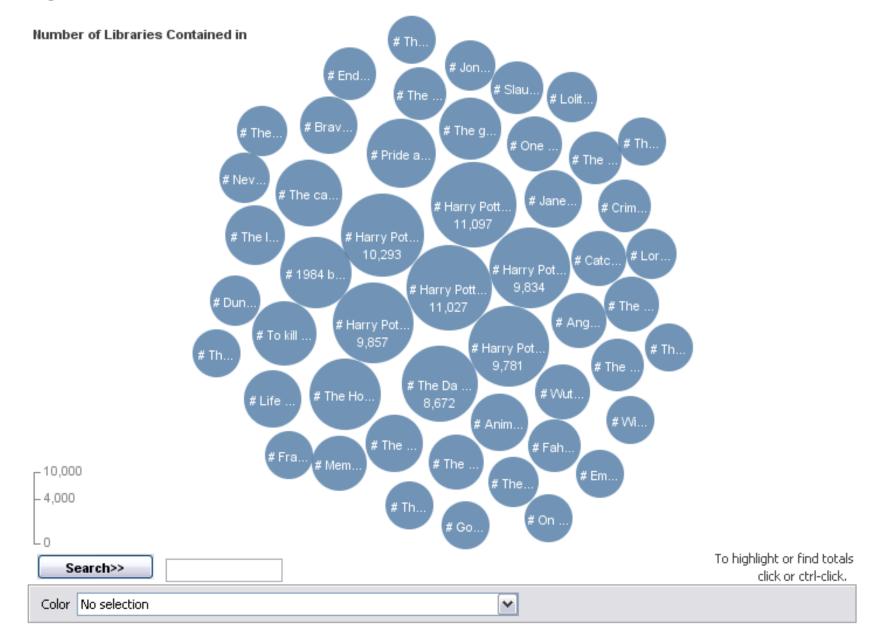
Comments (17)

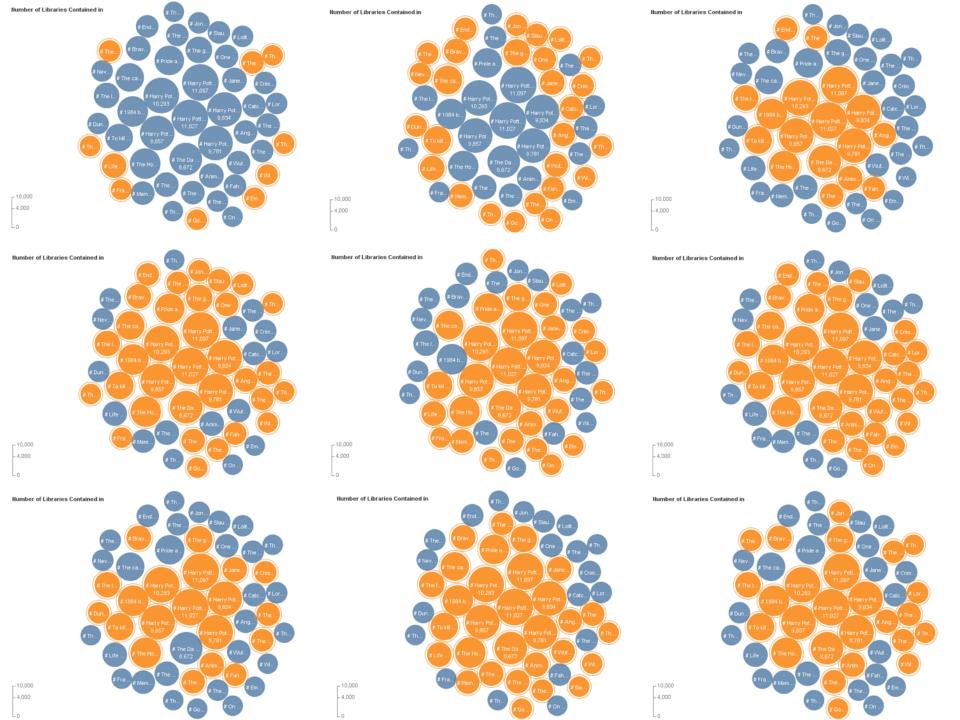
MARTY save

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 **Mon Jun 8 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 Jun 11**, by 8**am**, posted to GitHub 4-6 pages in ACM or IEEE TVCG format

Design Considerations for Collaborative Analysis

Modules of Contributions

Data Management

Contribute Data

Clean Data

Categorize Data

Moderate Data

Create Metadata

Visualization

Select Data Sources

Apply Visual Encoding

Author Software

Visual Analytics

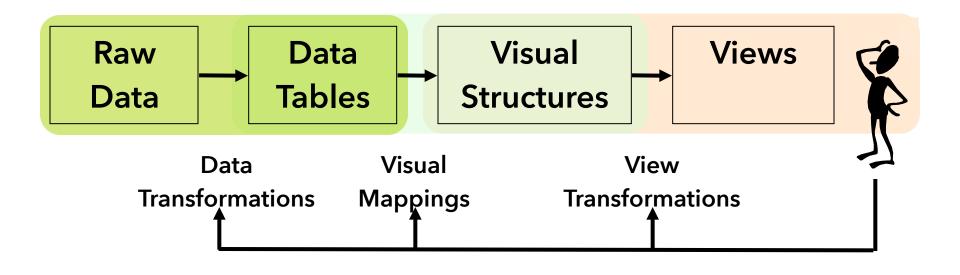
Observations

Hypotheses

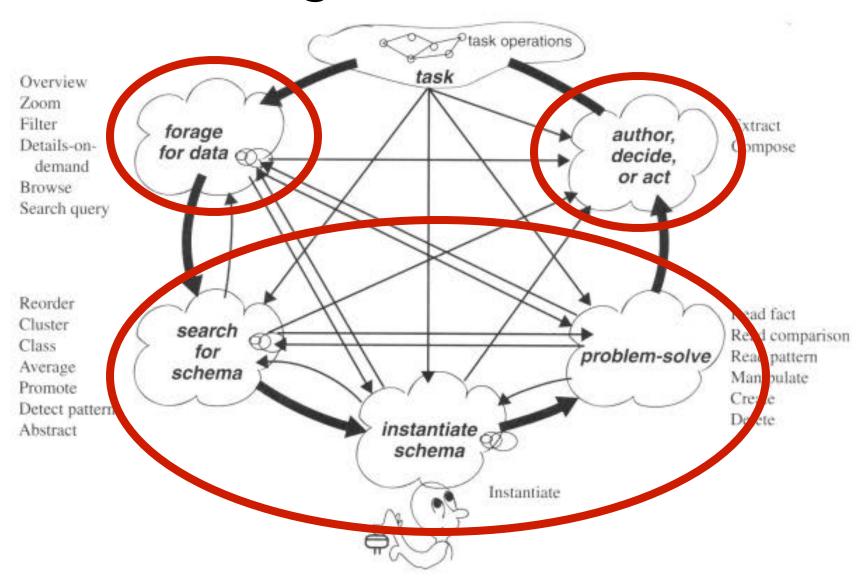
Evidence (+/-)

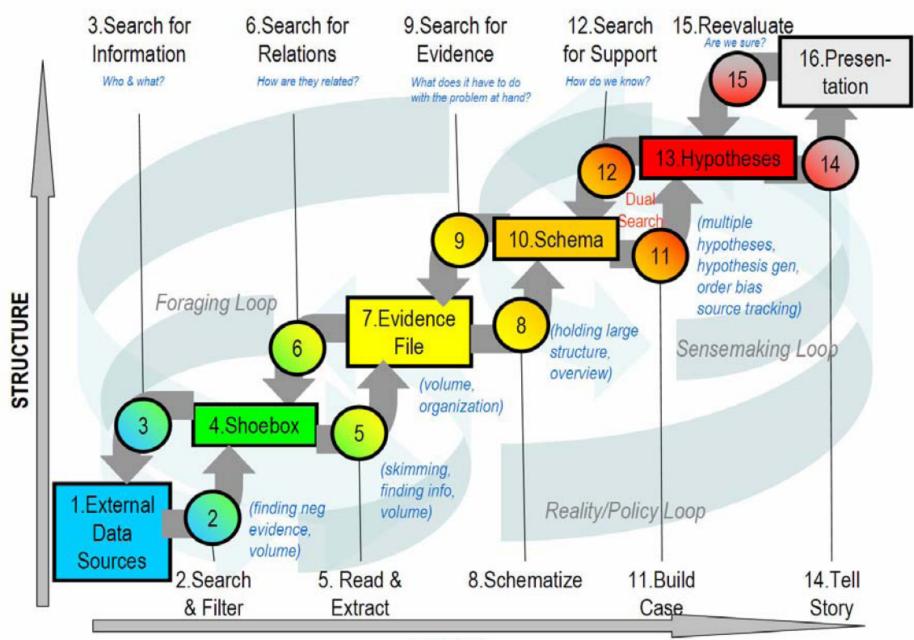
Summarize

Report / Presentation



Sensemaking





EFFORT

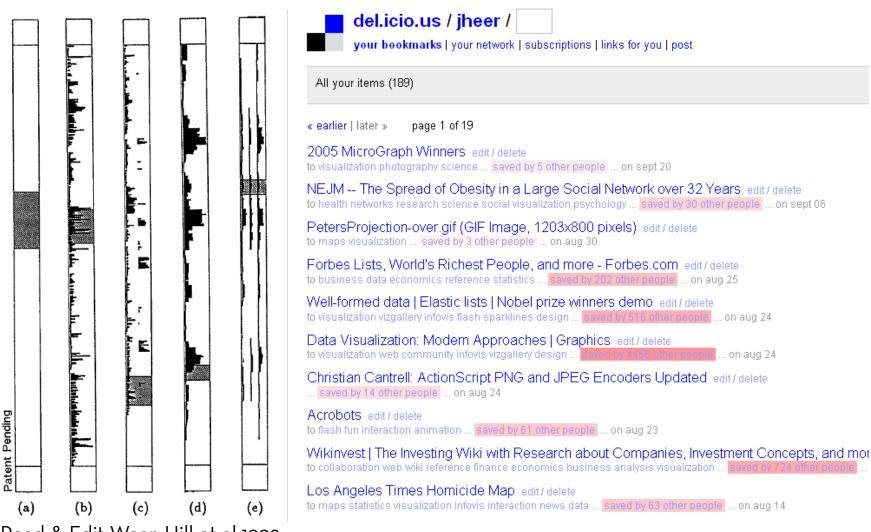
Design Considerations

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

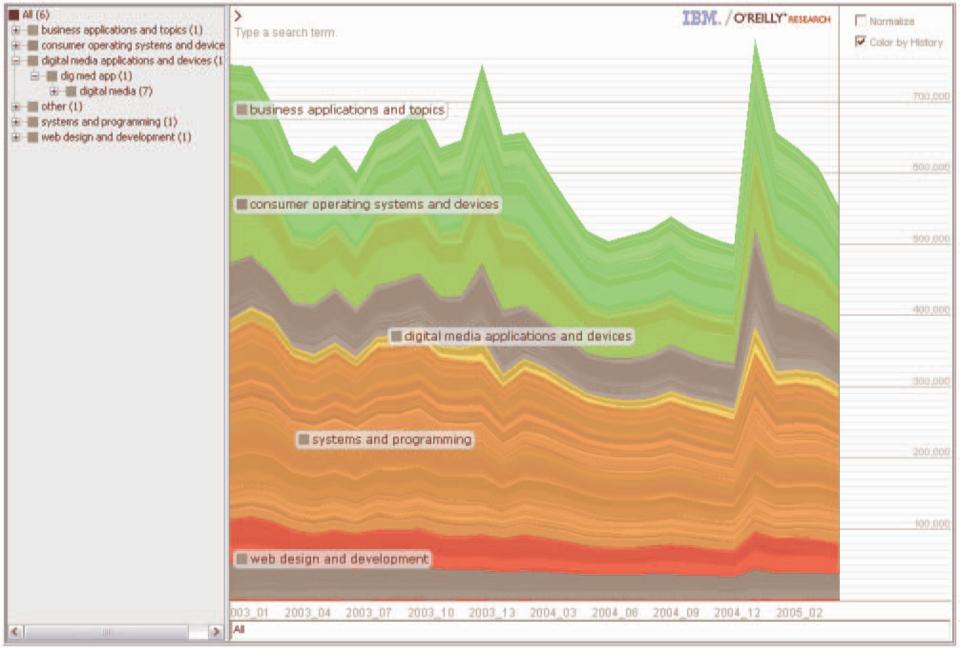
Social Data Analysis

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

Social Navigation

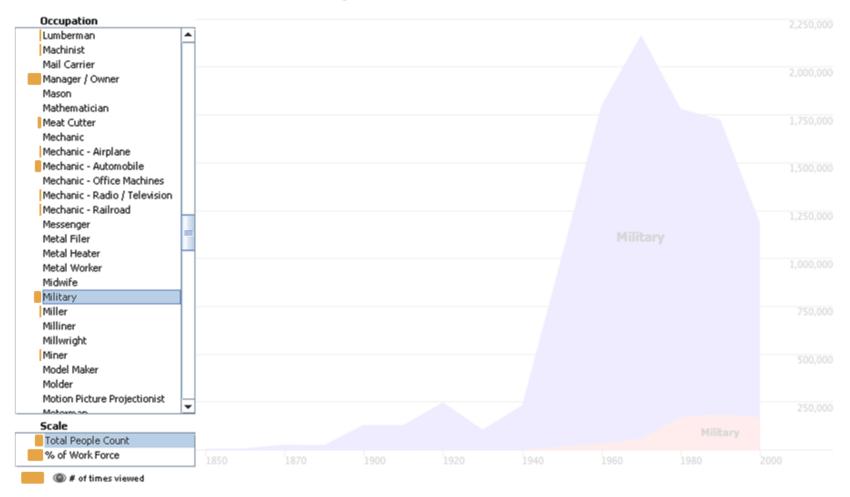


Read & Edit Wear, Hill et al 1992 « earlier | later »



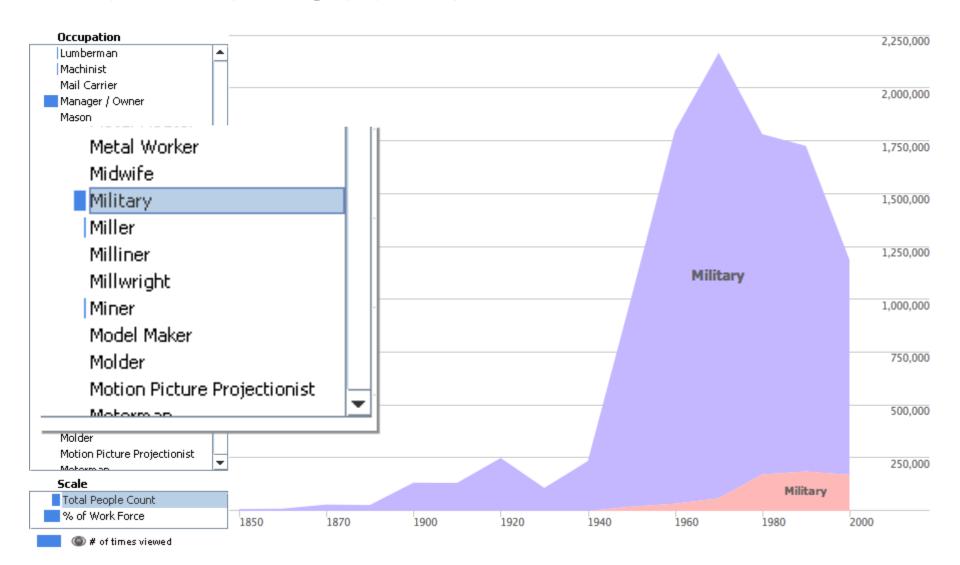
Wattenberg & Kriss - Color by history: gray regions have already been visited

Scented Widgets [InfoVis'07]

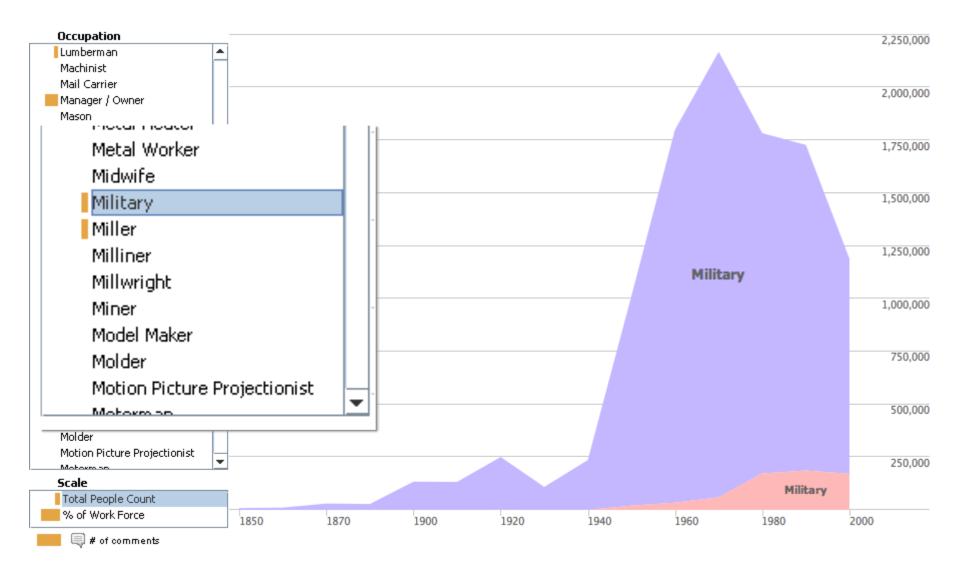


Visual navigation cues embedded in interface widgets

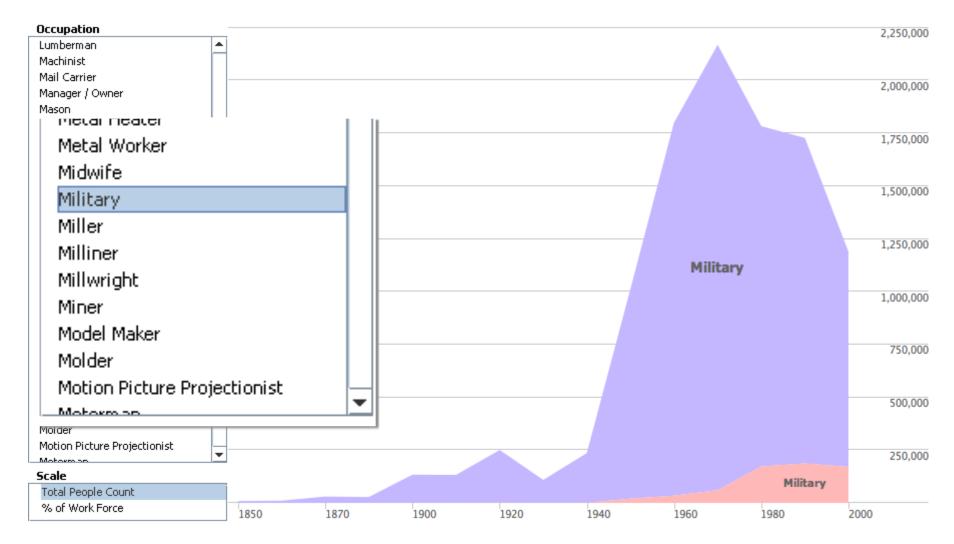
Visitation Counts



Comment Counts



No Scent (Baseline)



Do Activity Cues Affect Usage?

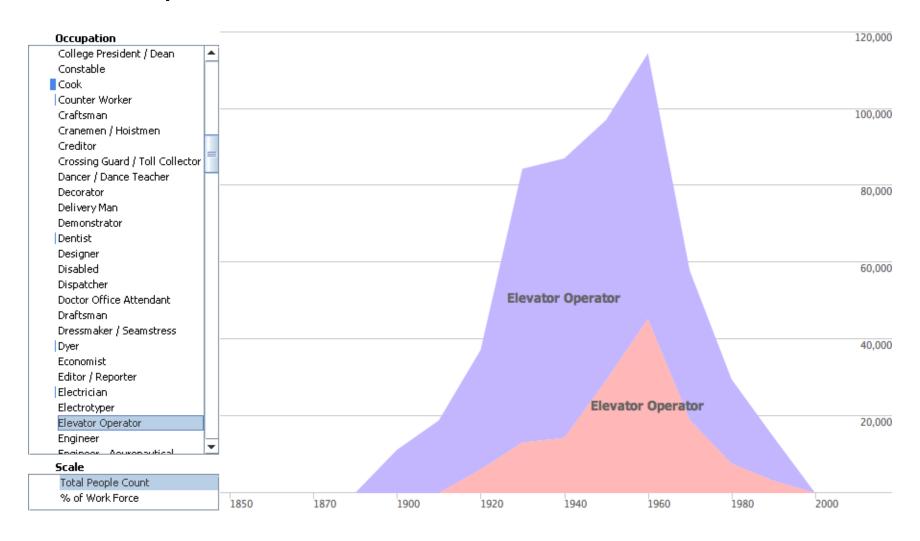
Hypotheses: With activity cues, subjects will

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

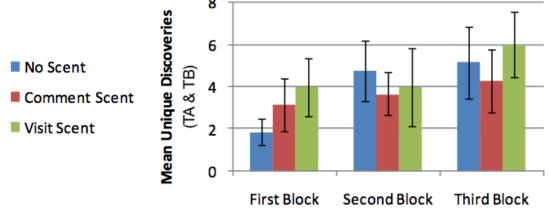
Controlled experiment with 28 subjects

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.

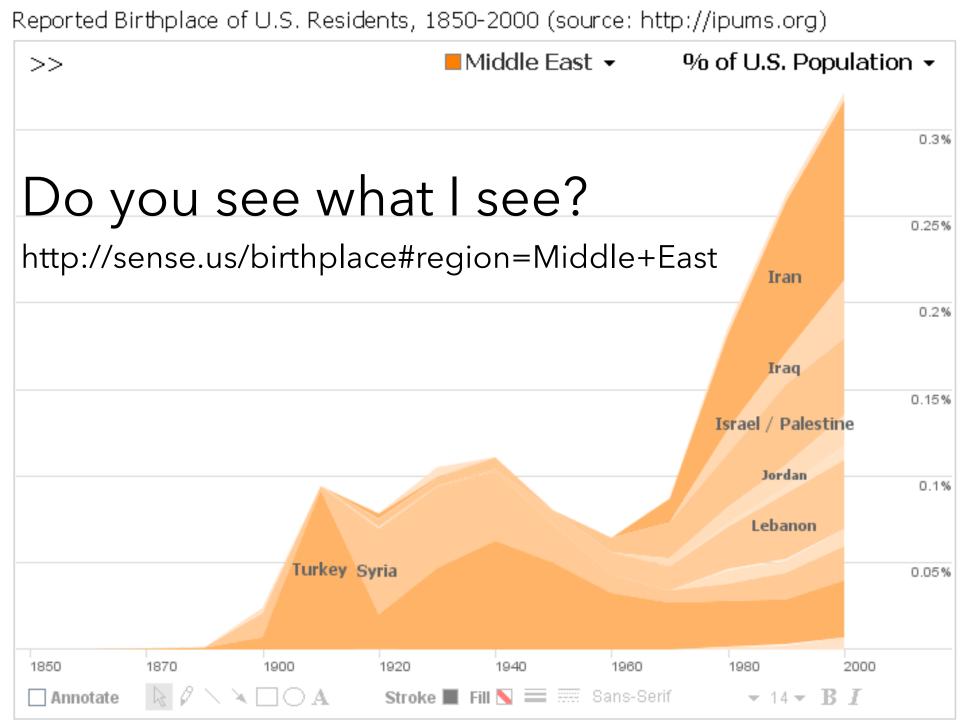
Revisitation

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?



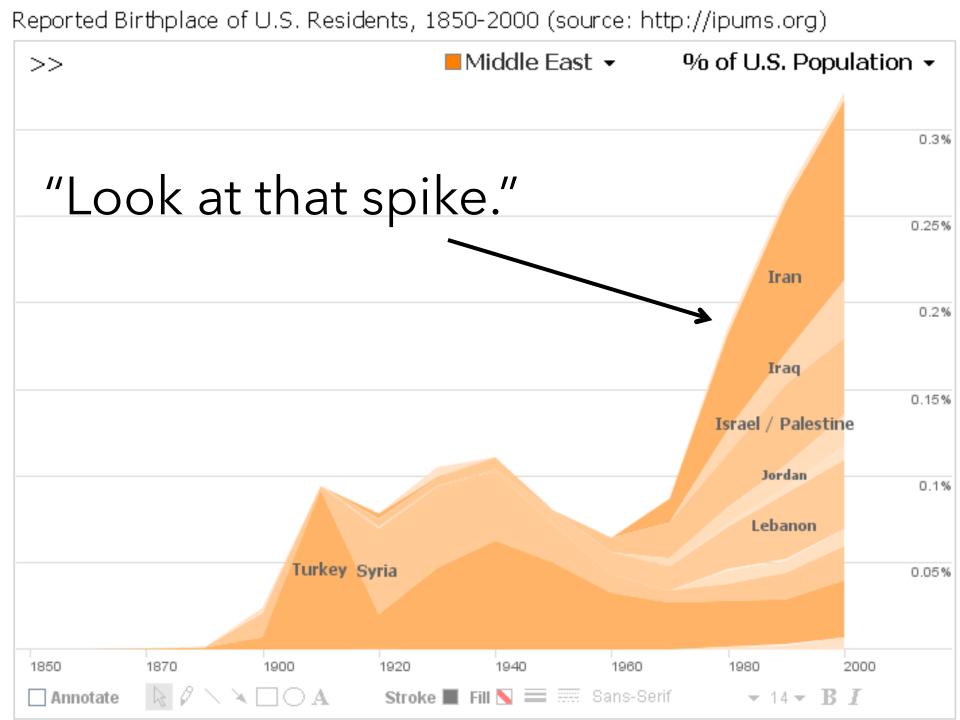
Common Ground

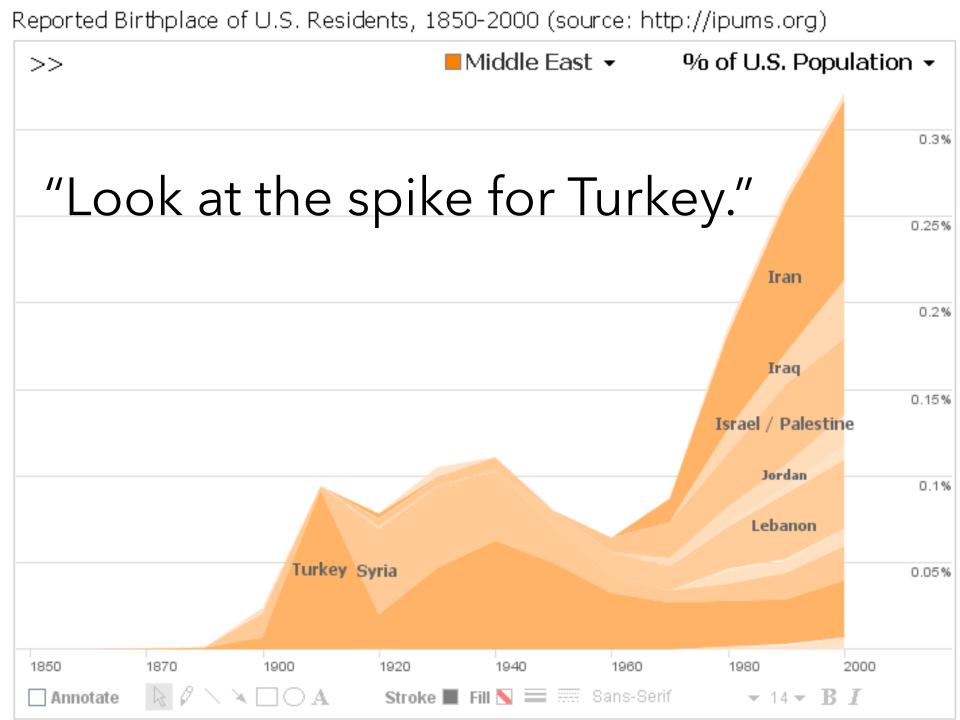
Common Ground: the shared understanding enabling conversation and collaborative action [Clark & Brennan '91]

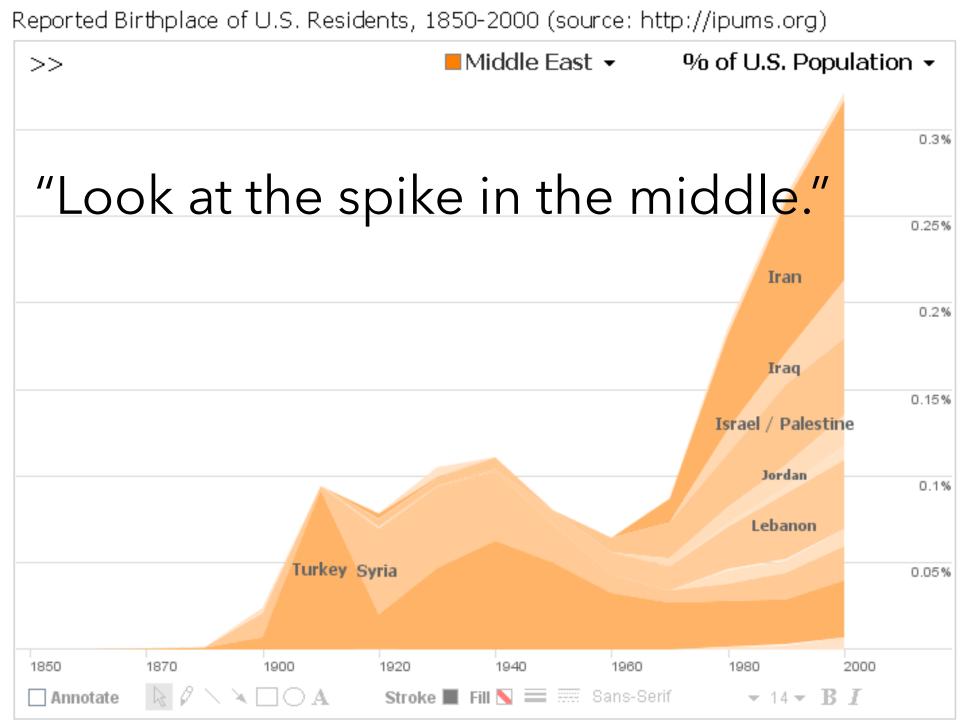
Do you see what I see? -> View sharing (URLs)

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]







Reported Occupations of U.S. Labor Force, 1850-2000 (source: http://ipums.org)

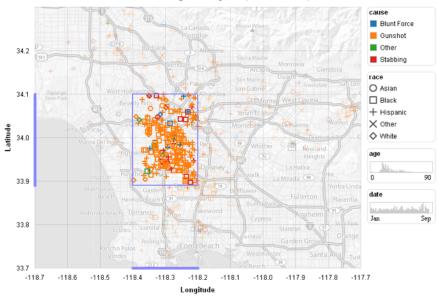


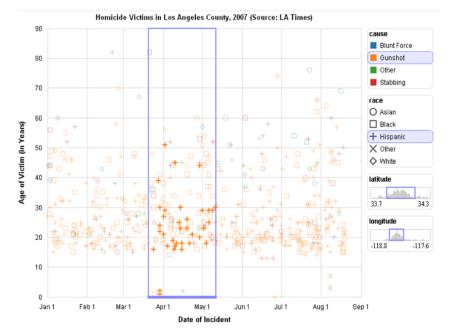




Free-form







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

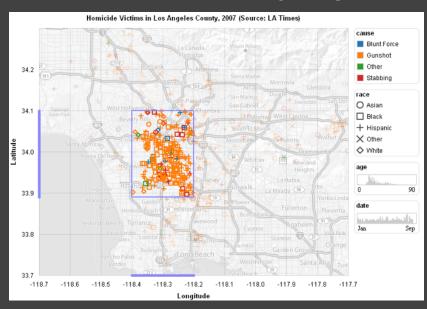
Pointing to specific points, trends, or regions (88.6%) Drawing to socialize or tell jokes (11.4%)

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





Visual Queries

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

Applicable to dynamic, time-varying data

Retarget selection across visual encodings

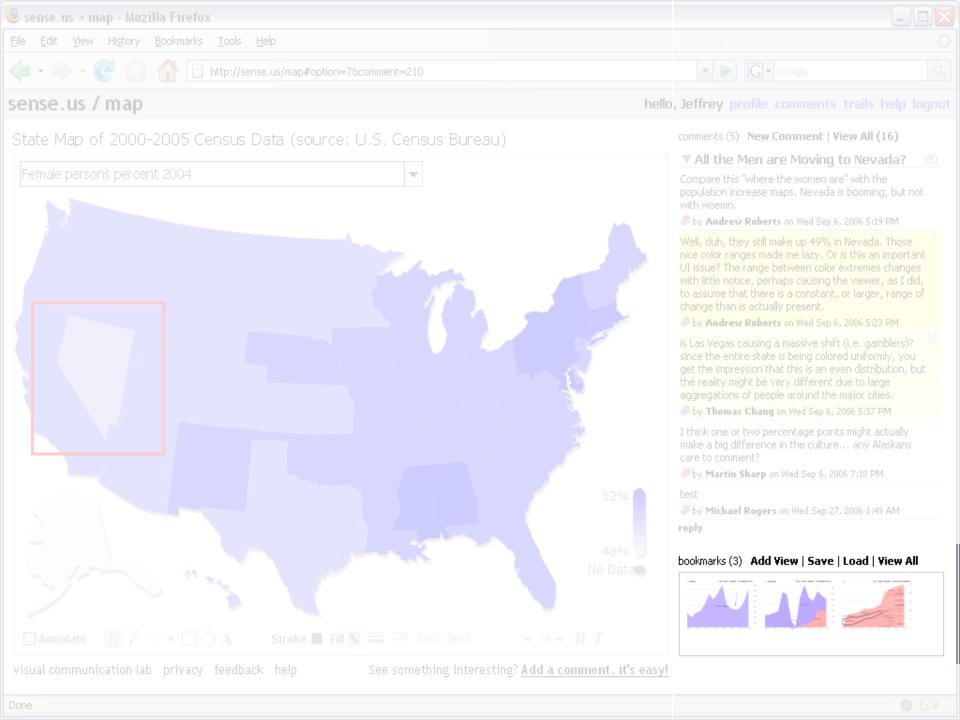
Support social navigation and data mining

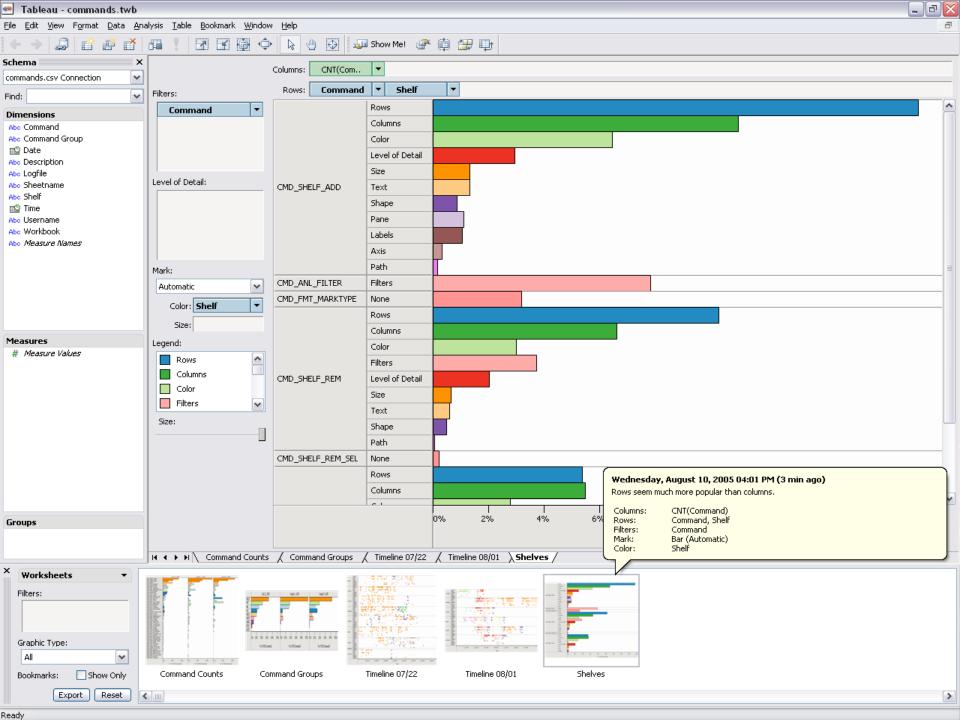
Social Data Analysis

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

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

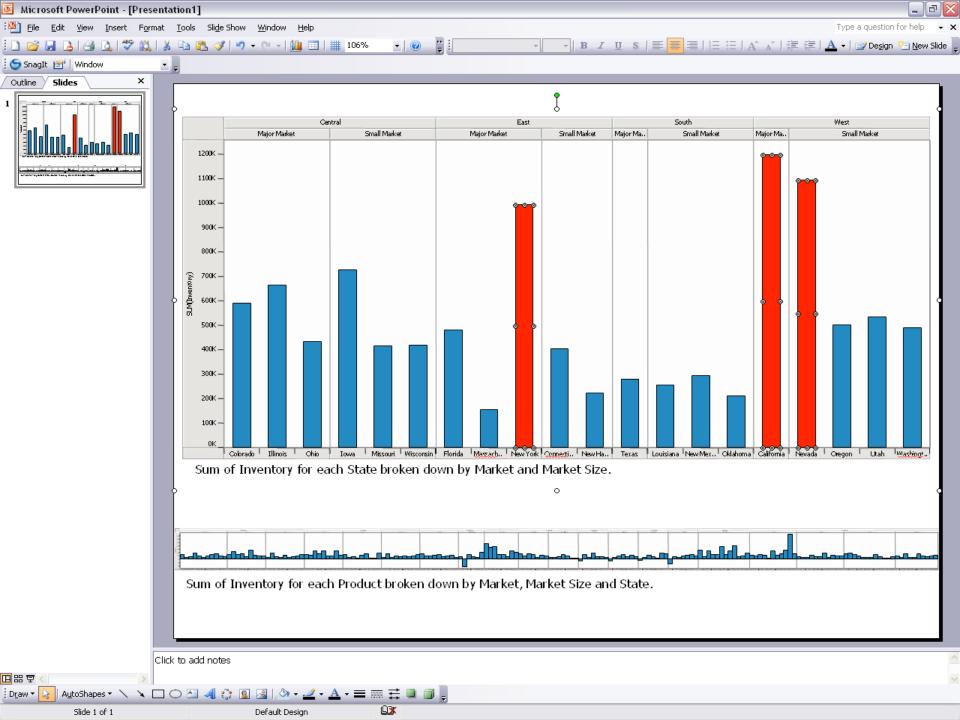
How can interface design better support communication of analytic findings?





Graphical Analysis Histories





Social Data Analysis

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

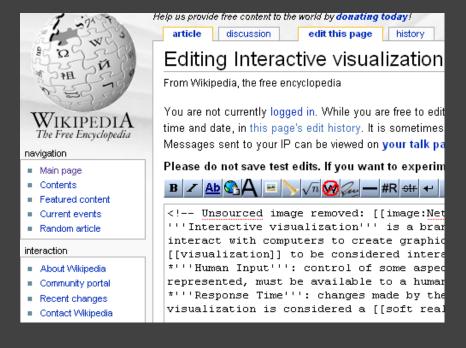
How should **annotation techniques** be designed to provide nuanced pointing behaviors?

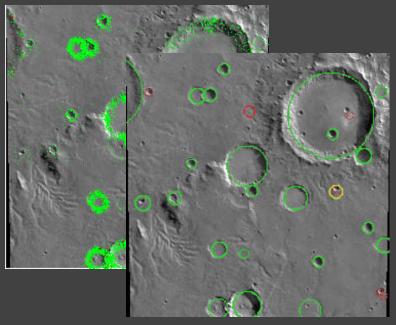
How can interface design better support presentation of analytic findings?

How can contributions be better integrated?

Structured Conversation

Reduce the cost of synthesizing contributions

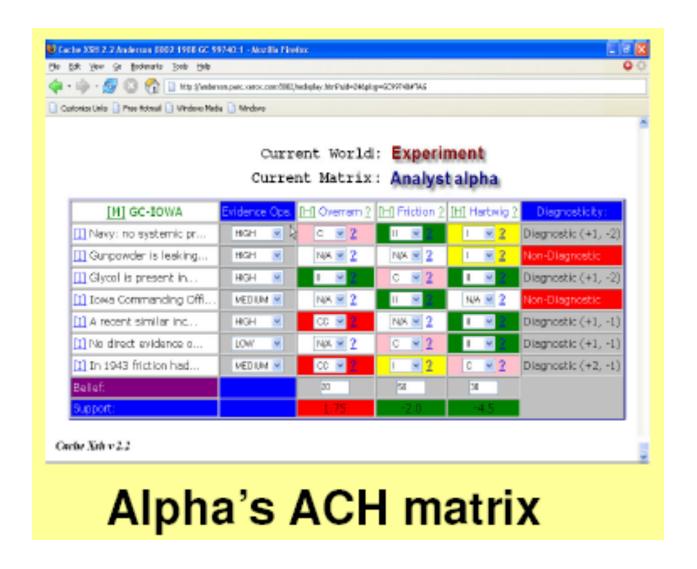




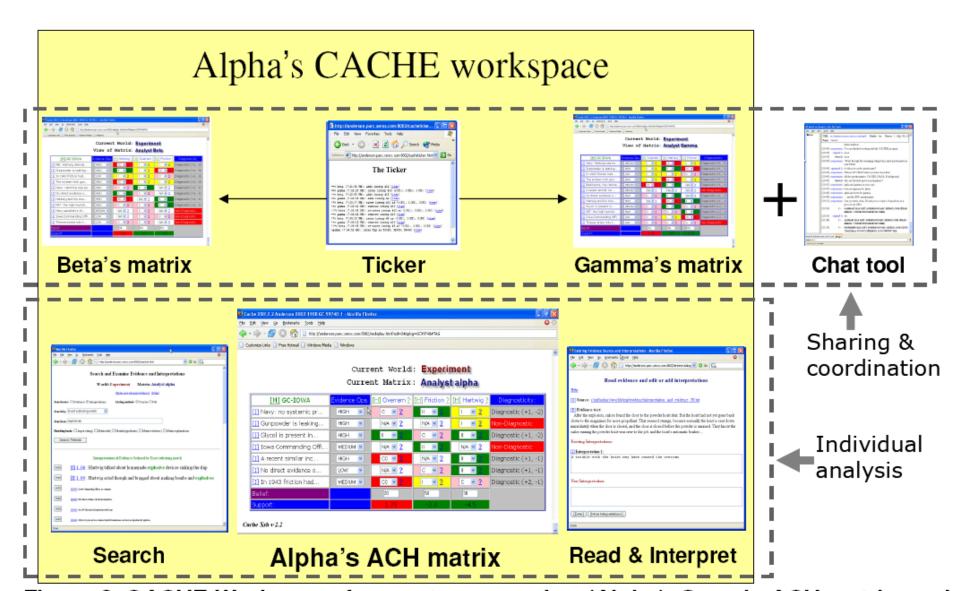
Wikipedia: Shared Revisions

NASA ClickWorkers: Statistics

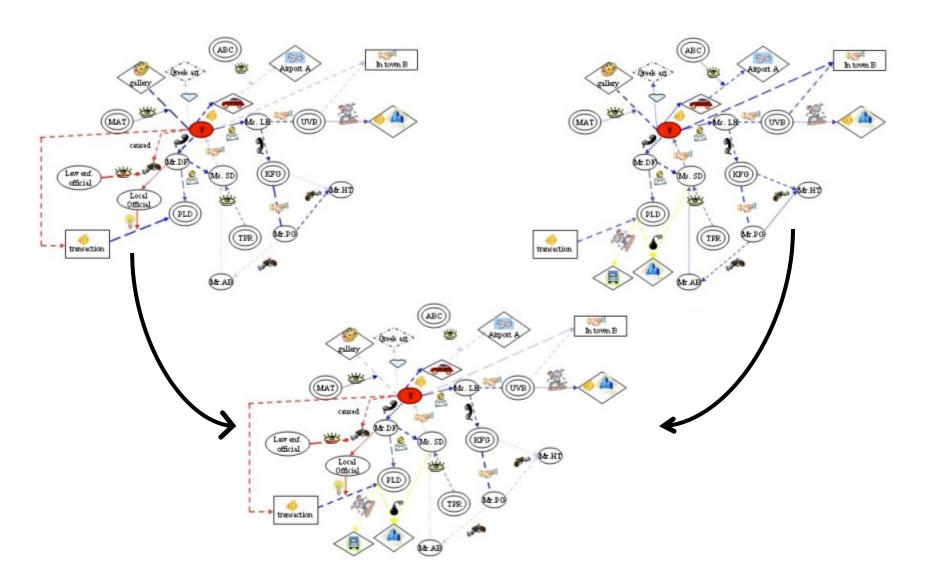
Integration: Evidence Matrices



Integration: Evidence Matrices



Merging Analysis Structures [Brennan '06]



Design Considerations

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?