Collaborative Analysis

Jeffrey Heer  University of Washington
A Tale of Two Visualizations
Zephoria

User ID: 21721
Friends: 296
Age: ??
Gender: Female
Status: Single
Location: San Francisco, CA
Hometown: Lancaster, PA
Occupation: researcher, social networks, identity, context
Interests: apophenia, observing people, culture, questioning power, reading, Buddhism, ipsify, computer-mediated communication, social networks, technology, anthropology, stomping
Music: psytrance/goa trance [Infected Mushroom, Son Kite, Iboga/Digital Structures], Ani DiFranco, downtempo, Thievery Corporation, Beth Orton, Morcheeba, Ween, White Stripes
Books: Authors: Enving Goffman, Stanley Milgram, Jeanette Winterson, Eric Schlosser, Leslie Feinberg, Dorothy Allison, Balo Calhino, Hermann Hesse
TV Shows: ??
Movies: Koyaanisqatsi, Amelie, Waking Life, Tank Girl, The Matrix, Clockwork Orange, American Beauty, Fight Club, Boys Don't Cry
Member Since: ??
Last Login: 2003-10-21
Last Updated: 2003-10-21
About: [Some know me as danah...]
I'm a geek, an activist and an academic, fascinated by people and society. I see life as a very large playground and enjoy exploring its intricacies. I revel in life's chaos, while simultaneously providing my own insane element.
My musings: http://www.zephoria.org/thoughts/
Want to Meet: Someone who makes life's complexities seem simply elegant.
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
sense.us is a prototype system for collaborative visualization.

- See the data. See what people have to say about it.
- Dive into the data and share your explorations.

The site requiresJava 1.5+ and either Firefox or Internet Explorer.
Use yourIBM W3/BluePages e-mail and password to login.
Use at least 1024x768 resolution for the best experience.

Check out theuser's guide and privacy policybefore getting started.

Having problems using Firefox with Java 1.4? Some users using Firefox andJava 1.4 have found that comments aren't loading properly. If you run into this problem, consider upgrading to Java 1.5 (Windows, Linux) or using Internet Explorer (on Windows systems). Sorry for any inconvenience!
Supporting Collaboration

Sharing within visualizations and across the web

http://sense.us/jobs#scale=1&gender=All&query=mili


Is this military info right?
I would have expected a different pattern for the military, but then again maybe this is just the military industrial complex growing and growing.

by Martin Sharp on Fri Jul 21, 2006 10:15 AM

here are labels where I would have expected big jumps.

by Martin Sharp on Fri Jul 21, 2006 10:16 AM

well, there was also the cold war right after WW2, which might be part of the reason why there's such a huge jump after the 40s. It is also interesting that there is such a drop between the 70s and the 80s.

by Julia Hernandez on Fri Jul 21, 2006 11:01 AM

I guess a lot of it has turned to robots, and the industrial complex, as martin suggested, though it would be interesting to see the comparison of the fall in military personal next to the rise in DOD funding for robots and industry.

by Jesse O'Brien on Fri Jul 21, 2006 11:51 AM

I think the jumps have more to do with the economy at large rather than any particular military conflict. Lots of money in conflict has already been spent before the conflict starts.

by Fred Klein on Wed Aug 2, 2006 10:24 AM
Supporting Collaboration

Sharing within visualizations and across the web

Pointing at interesting trends, outliers
Supporting Collaboration

Sharing within visualizations and across the web

Pointing at interesting trends, outliers

Collecting and linking related views
Supporting Collaboration

Sharing within visualizations and across the web

Pointing at interesting trends, outliers

Collecting and linking related views

Awareness of social activity
Supporting Collaboration

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)
Supporting Collaboration

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
The chart illustrates the relationship between GDP per capita (y-axis) and life expectancy in years (x-axis) across different countries. It shows that countries with higher GDP per capita tend to have higher life expectancy. Notable points include Uganda, Mozambique, and Sierra Leone, which are highlighted on the graph. The year 1877 is prominently displayed, indicating a significant point in the historical trend.
Spotfire Decision Site Posters
GeoTime Stories
The Status of George Bush's War on Terrorism

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
Many-Eyes
Discussion
Great depression "killed" a lot of brokers

>> Stock Broker

*all  men  women  Total People Count*

Graph showing the decline in the number of stock brokers during the Great Depression and the subsequent increase in the 1980s and 1990s.
Graph depicting the trend of Dentist and Technician - Medical/Dental in the percentage of the workforce from 1850 to 2000. The graph shows an increase in the percentage of women in the workforce, with a significant rise post-1940.
“Valley of Death”
Visualizations: Flawed Data - Temp Over Time and CO2 Levels

This view shows the correlation between...

Creator: Bruce
Tags:
Content Analysis of Comments

Feature prevalence from content analysis (min Cohen’s $\kappa = .74$)
High co-occurrence of Observation, Question, and Hypothesis
Sharing in External Media
Data Quality
No cooks in 1910? ... There may have been cooks then. But maybe not.
The great postmaster scourge of 1910?
Or just a bug in the data?
Anonymous says:
Um, I believe Church of Jesus Christ of Latter-day Saints and Mormon are the same?
16% of sense.us comments and 10% of Many-Eyes comments reference data integrity issues.
Data Integration in Context
Heroin vs. Cocaine vs. MDMA (Ecstasy) vs. Crystal Meth.

By guest on Dec 19, 2006
Viewed 20481 times

Graphs

Legend
- College Student Drug Use
- MDMA (Ecstasy)
- Cocaine
- Heroin
- Crystal Meth.

Sources: University of Michigan

Comments (1 - 20 of 32)

<script src='http:' />jonny12</script> says
posted about 1 year ago

Anonymous says
Where's the pot?
Marijuana vs. Heroin vs. Cocaine vs. MDMA (Ecstasy) vs. Crystal Meth.

By guest on Mar 12, 2007
Viewed 25401 times

Legend
- Marijuana
- MDMA (Ecstasy)
- Cocaine
- Heroin
- Crystal Meth.

Tags
- no tags yet
Community Tags
- no tags yet

Correlations
- See all

Related Graphs
- Heroin vs. Cocaine vs. MDMA (Ecstasy) vs. Crystal Meth vs. Heroin vs. Cocaine vs. MDMA (Ecstasy) vs. Crystal Meth

Comments (17)
- visnu says
  pot pot pot pot
  weed weed weed weed
  posted about 1 year ago

MARTY says
Visualizations: Harry Potter is Freaking Popular

Creator: Alison
Tags:

Number of Libraries Contained in

To highlight or find totals click or ctrl-click.
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

---

**Raw Data** → **Data Tables** → **Visual Structures** → **Views**

- **Data Transformations**
- **Visual Mappings**
- **View Transformations**
Sensemaking
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

Occupation
- Lumberman
- Machinist
- Mail Carrier
- Manager / Owner
- Mason
- Total People Count
- % of Work Force
- # of comments

Scale

Military

1850 1870 1900 1920 1940 1960 1980 2000

2,250,000
2,000,000
1,750,000
1,500,000
1,250,000
1,000,000
750,000
500,000
250,000
No scent (baseline)
Do activity cues affect usage?

**Hypotheses:** With activity cues, subjects will
1. Exhibit more revisititation 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.

**Revisititation**
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?
Do you see what I see?
http://sense.us/birthplace#region=Middle+East
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]
“Look at that spike.”
“Look at the spike for Turkey.”
“Look at the spike in the middle.”
Use of Annotations

<table>
<thead>
<tr>
<th>Tool</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arrows</td>
<td>25.1%</td>
</tr>
<tr>
<td>Text</td>
<td>24.6%</td>
</tr>
<tr>
<td>Ovals</td>
<td>17.9%</td>
</tr>
<tr>
<td>Pencil</td>
<td>16.2%</td>
</tr>
<tr>
<td>Lines</td>
<td>14.5%</td>
</tr>
<tr>
<td>Rectangles</td>
<td>1.7%</td>
</tr>
</tbody>
</table>

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

\((-118.371 \leq \text{lon} \leq -118.164) \land (33.915 \leq \text{lat} \leq 34.089)\)
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?
State Map of 2000-2005 Census Data (source: U.S. Census Bureau)

Female persons percent 2004
Wednesday, August 10, 2005 04:01 PM (3 min ago)
Rows seem much more popular than columns.

Columns: CNT(Command)
Filters: Command, Shelf
Mark: Bar (Automatic)
Color: Shelf
Graphical Analysis Histories
Sum of Inventory for each State broken down by Market and Market Size.

Sum of Inventory for each Product broken down by Market, Market Size and State.
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?