Many Eyes
A Site for Collaborative Data Analysis

Frank van Ham
Visual Communication Lab
IBM Watson Research
Cambridge, MA USA
Traditional infovis: exploration oriented
But, also, communication oriented
private data

PostHistory (2003)
Fernanda B. Viégas
- email archives
- personal data
2 public data

NameVoyager (2005)
Martin Wattenberg
public discourse

2004 US presidential election results
3 public discourse
2004 US presidential election results

County-level election results
public discourse
2004 US presidential election results

Cartogram: counties rescaled according to population
Research Agenda: Massive public visualization

Traditionally visualization researchers look at scaling the size of the data. But what happens when the audience scales?

1. Massively collaborative
   - internet scale: huge potential community
   - tens of thousands of viewers, thousands of commentators
   - easy for end users

2. Not just for analysis, but for communication, conversation
   - discovery
   - personal expression
   - public debate

3. Visual
   - not simply collaboration around text
   - use human visual intelligence, not data mining

4. End user driven
   - allow users to look at their own data
   - allow users to select and customize a visualization
Demo

http://www.many-eyes.com

a site where people:
- view and discuss visualizations
- view and discuss data
- create visualizations from existing data
- upload their own data to visualize
Social data analysis: a growing area

WikiSky
Wikimapia
Google Earth community
Swivel
Data360
Gapminder
Communication outside Many Eyes

- Communications does not happen only on Many Eyes.
- The web is a whole ecosystem that consists of blogs, forums, wiki’s etc
- Many interactions observed between Many Eyes and this ecosystem.
1 User "crossway" uploads co-occurrence data for biblical figures to Many Eyes
1. User "crossway" uploads co-occurrence data for biblical figures to Many Eyes.

2. Crossway uses the network diagram tool to create a graph visualization.
1. User "crossway" uploads co-occurrence data for biblical figures to Many Eyes.

2. Crossway uses the network diagram tool to create a graph visualization.

3. Crossway writes about the visualization on ESV blog.

The long list of responses / trackbacks on the ESV blog entry: a discussion about the visualization and analysis.
1. User “crossway” uploads co-occurrence data for biblical figures to Many Eyes.

2. Crossway uses the network diagram tool to create a graph visualization.

3. Crossway writes about the visualization on ESV blog.

4. Many blogs (almost 100 by Google’s count) write about crossway’s blog entry.

The long list of responses / trackbacks on the ESV blog entry: a discussion about the visualization and analysis.
1. User “crossway” uploads co-occurrence data for biblical figures to Many Eyes.

2. Crossway uses the network diagram tool to create a graph visualization.

3. Crossway writes about the visualization on ESV blog.

4. Many blogs (almost 100 by Google’s count) write about crossway’s blog entry.

5. One of these bloggers posts new data to Many Eyes—and, of course, blogs about the results.

The long list of responses / trackbacks on the ESV blog entry: a discussion about the visualization and analysis.
Visualization use

Scientific discovery

Personal expression

Journalism and advocacy

Social interaction
Many, many data domains

- Science
  - Microarray data
  - Chemistry
  - single nucleotide polymorphisms
  - global CO2 vs. temperature vs. time

- Politics & War
  - Iraq
  - Political speech

- Literature
  - words in Swinburne’s poetry
  - Pride and Prejudice
  - Green Eggs and Ham

- Religion
  - Bible
  - Potential Converts

- Personal information
  - Family trees
  - Running, swimming, weight loss
  - Nick and Betty’s gift-giving network
  - Books read
  - Countries visited
  - Del.icio.us tags
Many Eyes as a living laboratory?

- User testing still problematic in information visualization.
- Small user groups (colleagues, CS students)
- Unclear what/how to measure (“insight?”)
- Many Eyes may help by providing a large user sample
- Already some interesting examples.
Closing remarks

- For the first 20 years of their existence telephones were marketed as “serious” communication instruments.
- Use for social purposes was actively **discouraged** by phone companies.
- But, current advertising campaigns focus **only** on the social aspects.
- Same holds for visualization
Some basic numbers

- Over 1500 registered users
- Users with data sets: 625  42%
- Users with visualizations: 425  29%
- Users with comments: 113  7.7%
- Data sets w/ source: 70%
- Data sets w/ source + URL: 41%
- Page retrievals that aren't bots: ~400,000
- Syndication feed hits: 174,036