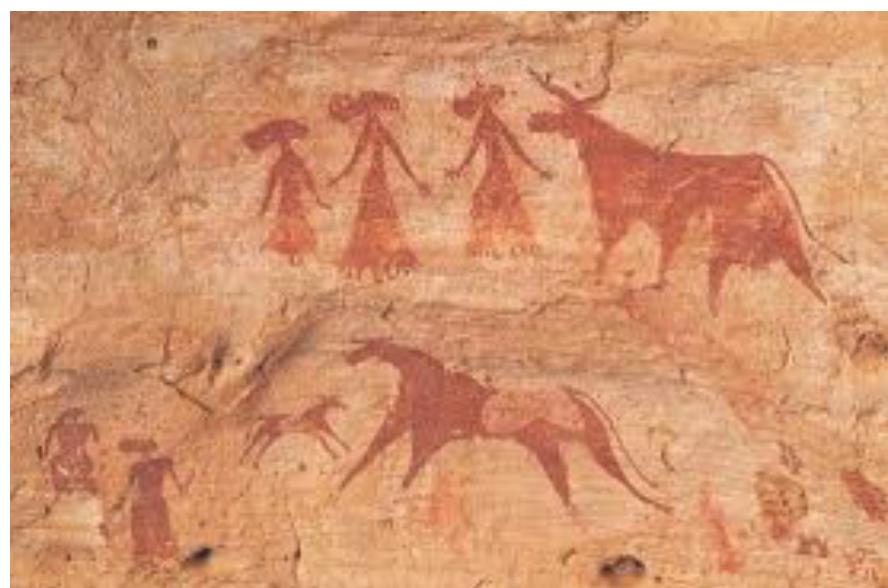


Narrative Visualization

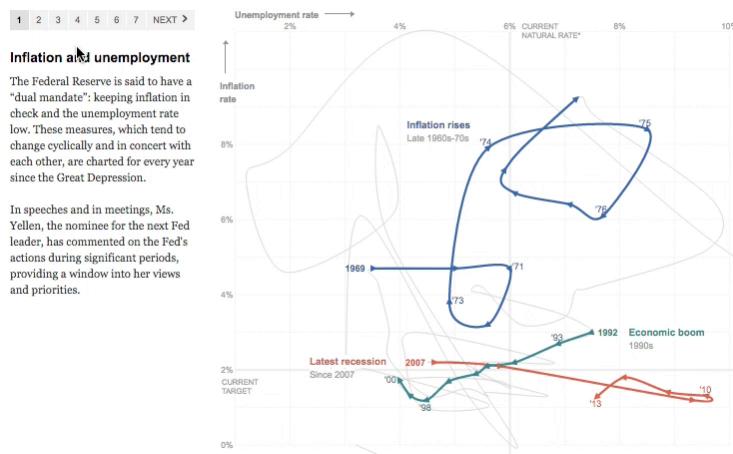
Jessica Hullman



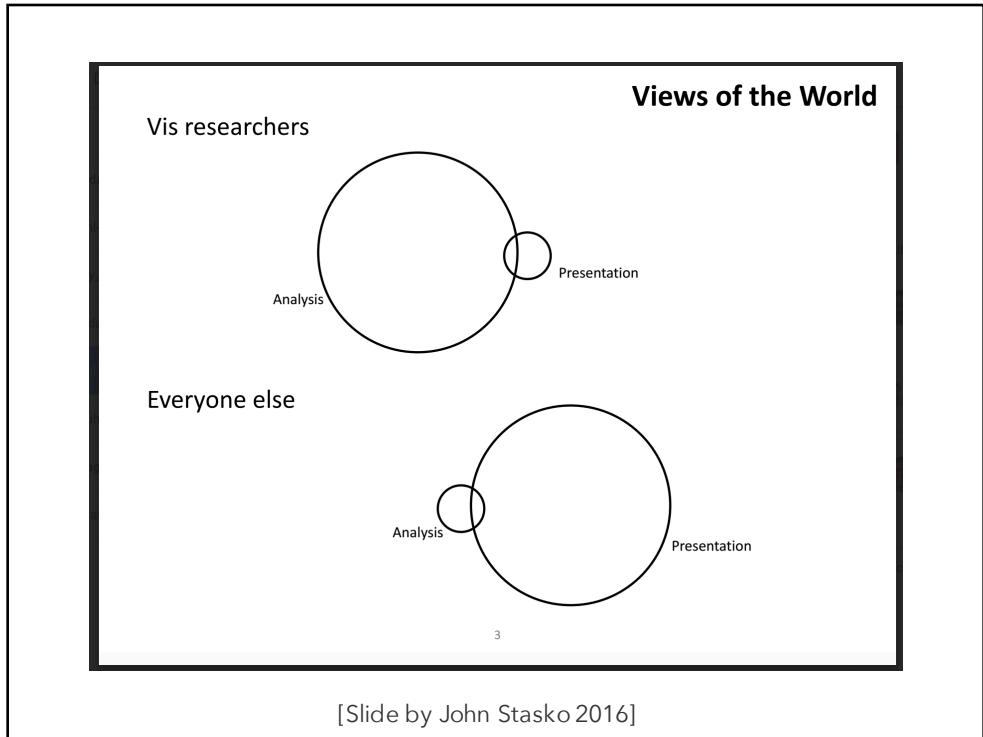
Cave painting [Ennedi Plateau, Chad]



Janet L. Yellen, on the Economy's Twists and Turns



[Giratikanon and Parlapiano 2013]



Topics

Design space

Theories of interpretation

Construction

Manual

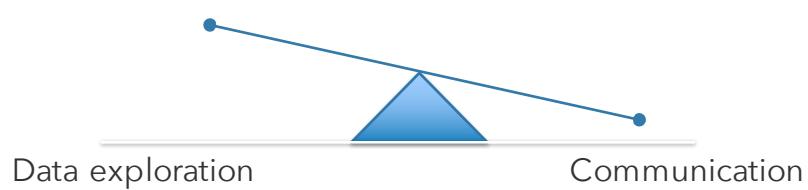
Automated

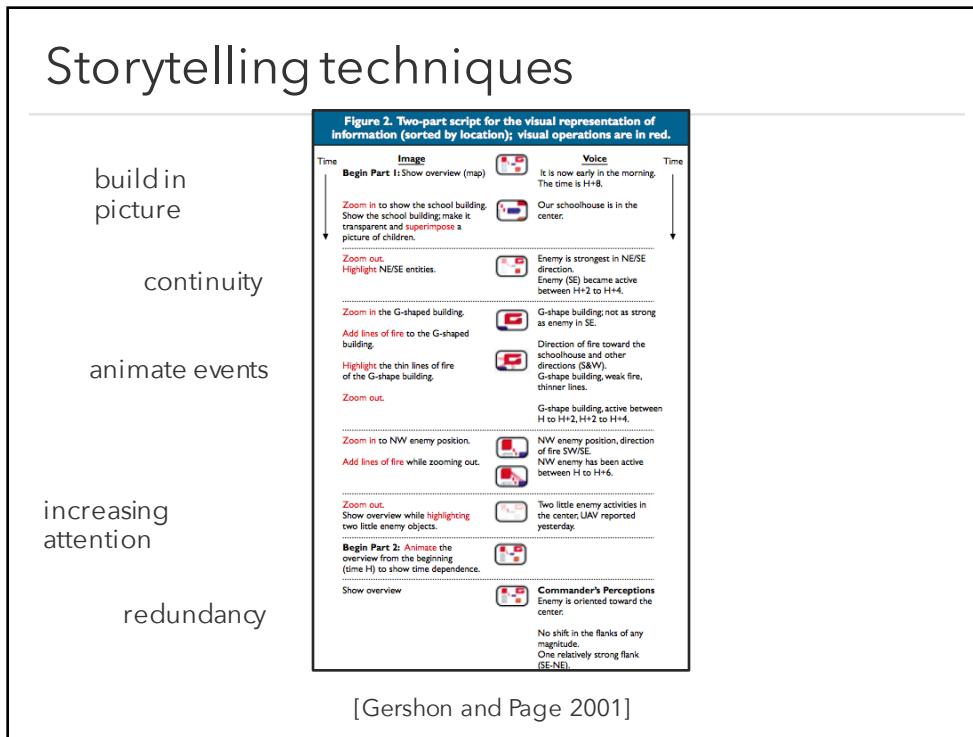
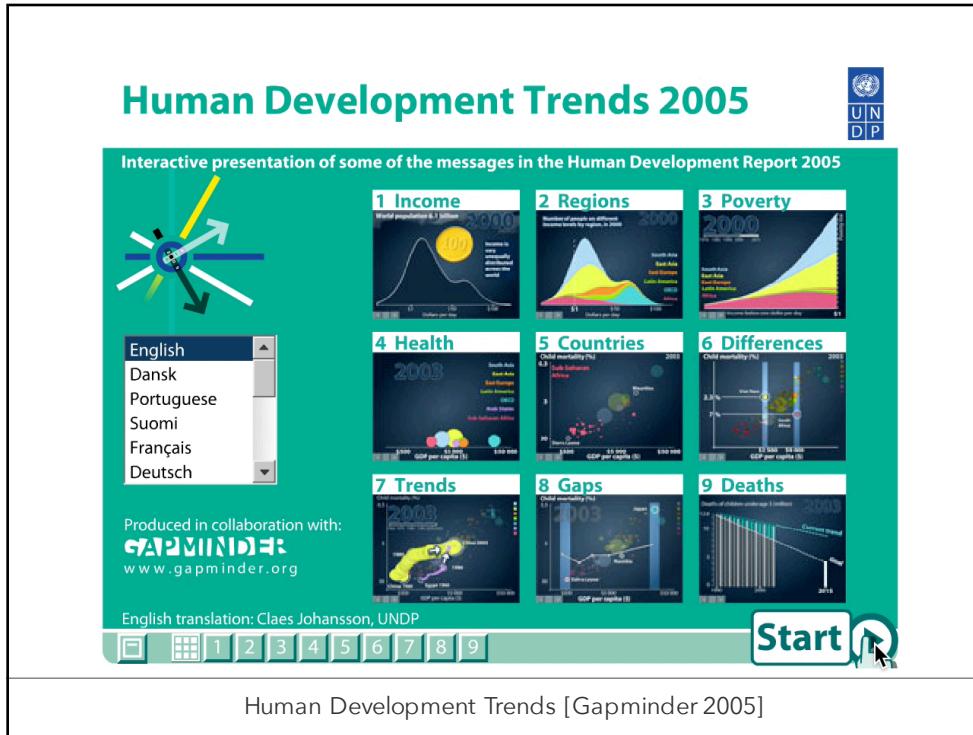
Design Space

Value of storytelling [Gershon and Page 2001]

A way of structuring information

- Easier to understand than lists
- Uncertainty, conflict, resolution
- Text and visuals can be complementary

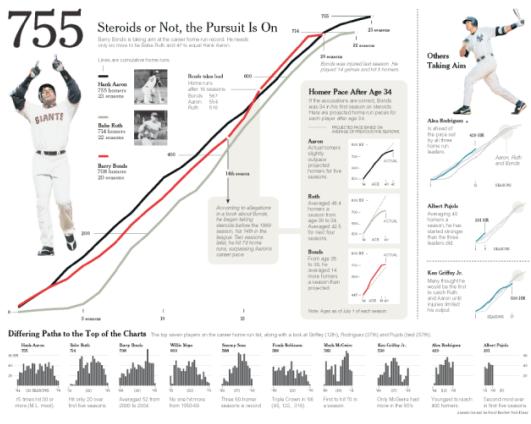




Narrative Visualization [Segel and Heer 2010]

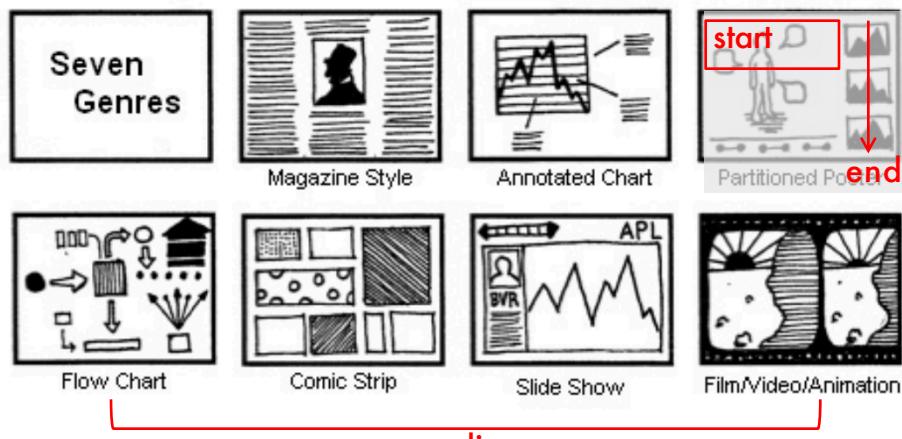
Studied 58 examples, characterized design space

- News media, blogs, instructional videos, research



[Cox and Ward 2006]

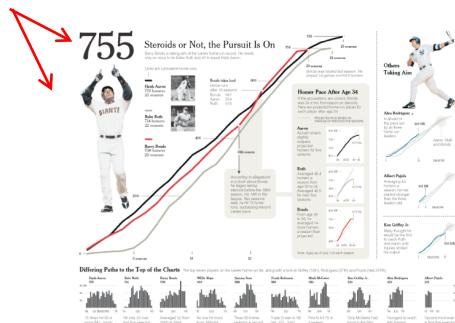
Design space dimensions: 1. Genre



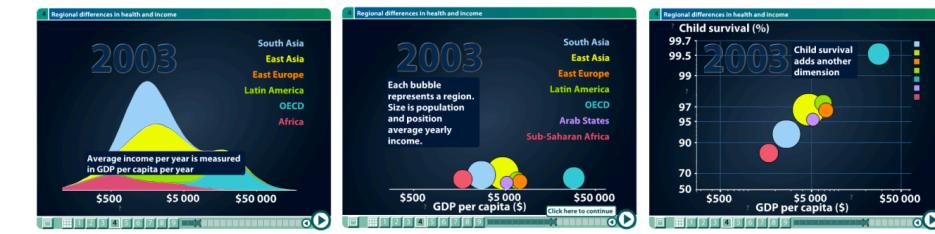
[Slide adapted from Segel 2010]

2. Visual narrative tactics

Highlighting



Transition guidance



3. Narrative structure tactics

Ordering Messaging Interactivity

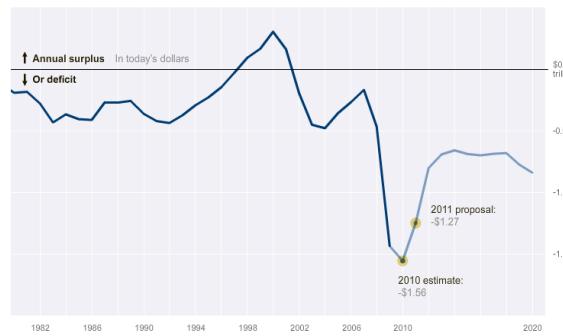
Budget Forecasts, Compared With Reality

Just two years ago, surpluses were predicted by 2012. How accurate have past White House budget forecasts been?

1 2 3 4 5 6 NEXT >

Falling short

President Obama's budget proposal estimates a deficit of \$1.6 trillion for the current fiscal year and \$1.3 trillion in 2011.



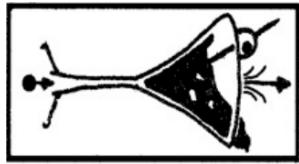
Author vs. reader driven stories

Author driven

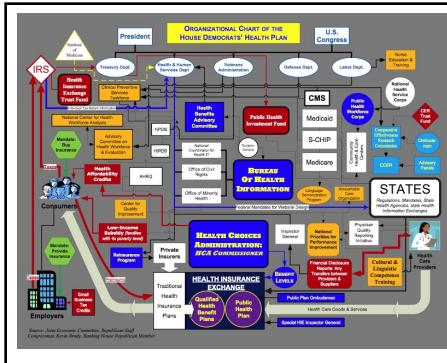
Prescribed ordering
Stronger messaging
Limited interactivity

Reader driven

Multiple orderings
Less messaging
More open interactivity



How do we make sense
of these examples?



An Organizational Chart of the House Democrat's Health Plan [GOP 2008, Robert Palmer 2008]

Visualization rhetoric [Hullman & Diakopoulos 2011]

Using data and visualization to persuade users to adopt certain interpretations (explicitly or implicitly)

Framing effects: small changes in presentation of an issue result in significant changes in opinion

Method

- 51 professional produced narrative visualizations
- NYT, BBC, Economist, local news, political outlets
- Iterative qualitative coding, seeded scheme with semiotics, persuasion concepts

Taxonomy of rhetorical strategies

Rhetorical Categories

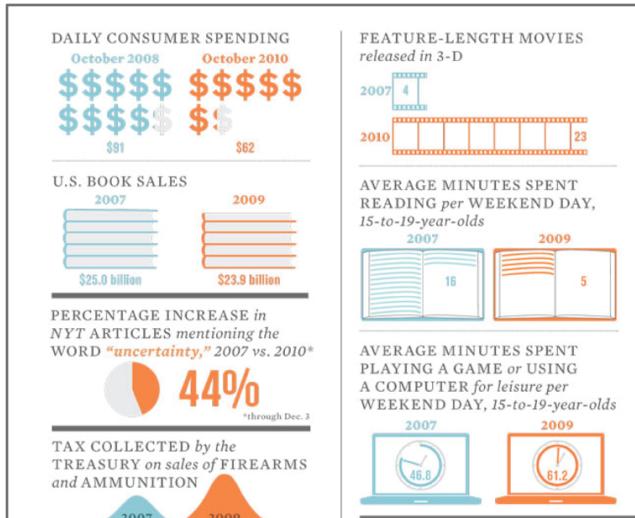
Information Access
Provenance
Mapping
Procedural
Linguistic

Editorial Layers

Data
Visual Representation
Annotation
Interactivity

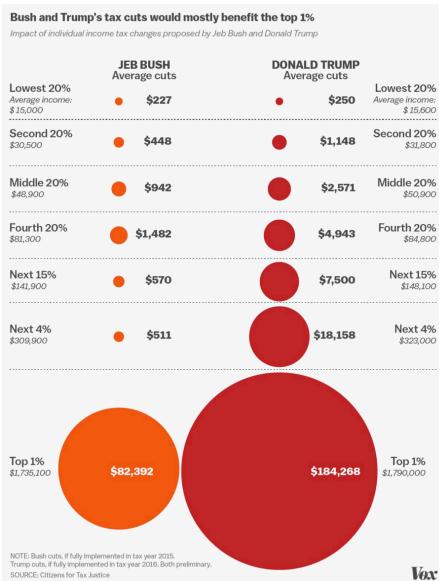
-
- ```
graph LR; A[Information Access] --> D[Data]; B[Provenance] --> D; C[Mapping] --> D; E[Procedural] --> V[Visual Representation]; F[Linguistic] --> An[Annotation]; G[Information Access, Provenance, Mapping, Procedural, Linguistic] --> I[Interactivity];
```
- Omission (variable selection)
  - Metonymy (aggregation, categorization)

## Enthymeme (information access)



How the Recession Changed Us [The Atlantic 2009]

## Metonymy (information access)



Donald Trump wants massive tax cuts for the rich [Vox 2015]

# Taxonomy of rhetorical strategies

## Rhetorical Categories

Information Access  
Provenance  
Mapping  
Procedural  
Linguistic

## Editorial Layers

Data  
Visual Representation  
Annotation  
Interactivity

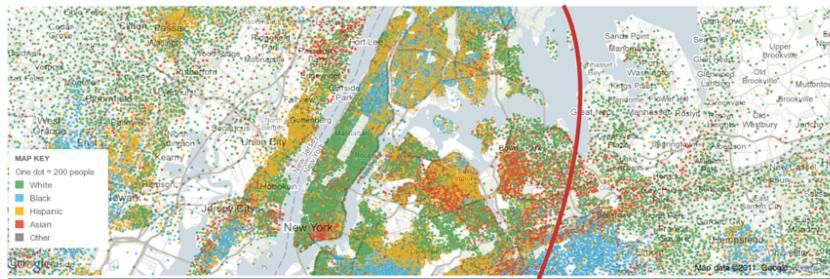
- Citing sources, methodology
- Exceptions, corrections
- Acknowledging uncertainty

## Citing sources /methods (provenance)

### Mapping America: Every City, Every Block

Browse local data from the Census Bureau's American Community Survey, based on samples from 2005 to 2009. Because these figures are based on samples, they are subject to a margin of error, particularly in places with a low population, and are best regarded as estimates.

#### Distribution of racial and ethnic groups



### Mapping America: Every City, Every Block

Browse local data from the Census Bureau's American Community Survey, based on samples from 2005 to 2009. Because these figures are based on samples, they are subject to a margin of error, particularly in places with a low population, and are best regarded as estimates.

#### Distribution of racial and ethnic groups



[Bloch, Carter, and McLean 2010]

# Taxonomy of rhetorical strategies

Rhetorical Categories

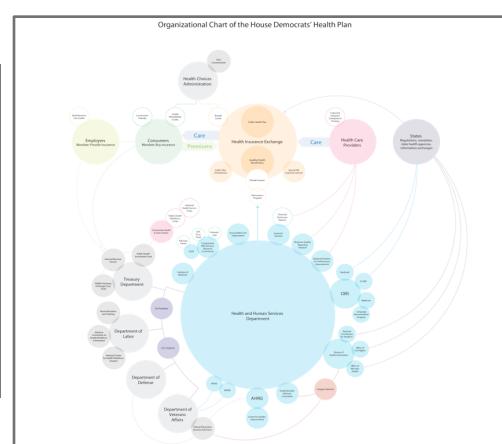
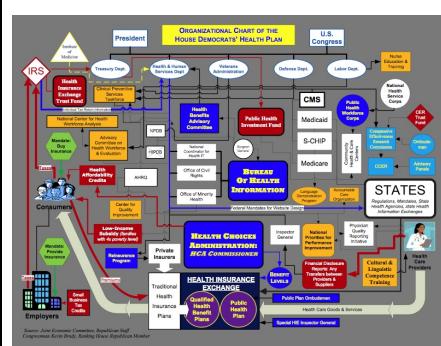
- Information Access
- Provenance
- Mapping
- Procedural
- Linguistic

Editorial Layers

- Data
- Visual Representation
- Annotation
- Interactivity

- Visual metaphor
- Semantic encodings (e.g., red:Rep, blue:Dem)

## Visual metaphor(mapping rhetoric)



# Taxonomy of rhetorical strategies

## Rhetorical Categories

Information Access  
Provenance  
Mapping  
Procedural  
Linguistic

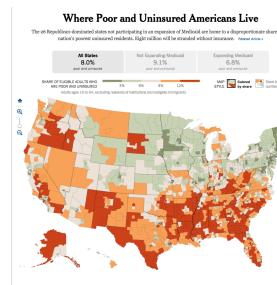
## Editorial Layers

Data  
Visual Representation  
Annotation  
Interactivity

- Rhetorical question, "scare quotes"
- Labeling choices
- Analogy, simile, double entendre

## Design process

An *editorial process* characterized by rhetorical decisions at various points.



define context / goal

filter, transform

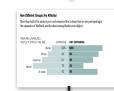
visualize



select



annotate



order, interactions



## Scope of narrative visualization

- Visualization genres that employ visual and narrative structures to guide attention
- Visualizations that use rhetorical devices to persuade

### **Discuss with 1 or 2 other people near you:**

In your own research or papers you've read/cited, is narrative visualization occurring? How do persuasion and rhetoric occur in these contexts?

## Theories of Interpretation

## Semiotics

"TREE"

Sign

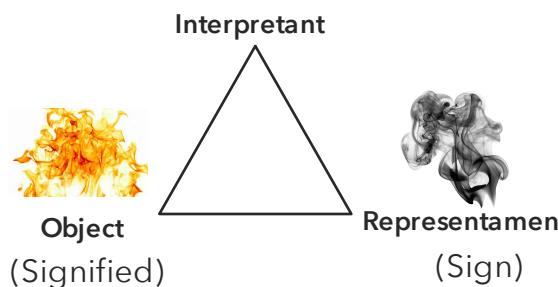


Signified

## Semiotics

Viewing codes: the perceptual mechanisms, cognitive mechanisms, conventions and other prior knowledge that influence how we interpret signs.

*There is a fire!*

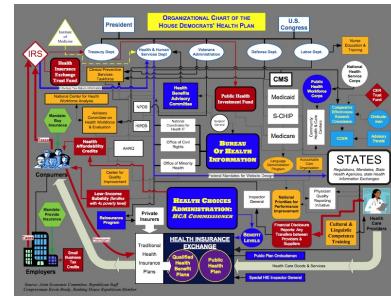


## Denotation vs. connotation

(Make explicit)

(Imply)

The Democrats proposed health-plan is disorganized and inefficient.



# Narrative theory

## Story grammars:

Models of narrative cognition based on systematic studies of what impacts peoples' ability to recall parts of a story

Reader mentally indexes events by time, space, protagonist, causality, intention [Zwaan 1995]

| SUMMARY OF REWRITE RULES FOR A SIMPLE STORY GRAMMAR <sup>6</sup> |                                                                                                                                  |
|------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------|
| FABLE                                                            | → STORY AND MORAL                                                                                                                |
| STORY                                                            | → SETTING AND EVENT STRUCTURE                                                                                                    |
| SETTING                                                          | → $\left\{ \begin{array}{l} \text{STATE}^* \text{ (AND EVENT}^*) \\ \text{EVENT}^* \end{array} \right\}$                         |
| STATE*                                                           | → STATE (( AND STATE)*)                                                                                                          |
| EVENT*                                                           | → EVENT (( $\left\{ \begin{array}{l} \text{AND} \\ \text{THEN} \\ \text{CAUSE} \end{array} \right\}$ ) EVENT) (( (AND STATE)*    |
| EVENT STRUCTURE                                                  | → EPISODE (( THEN EPISODE))                                                                                                      |
| EPISODE                                                          | → BEGINNING CAUSE DEVELOPMENT CAUSE ENDING                                                                                       |
| BEGINNING                                                        | → $\left\{ \begin{array}{l} \text{EVENT}^* \\ \text{EPISODE} \end{array} \right\}$                                               |
| DEVELOPMENT                                                      | → $\left\{ \begin{array}{l} \text{SIMPLE REACTION CAUSE ACTION} \\ \text{COMPLEX REACTION CAUSE GOAL PATH} \end{array} \right\}$ |
| SIMPLE REACTION                                                  | → INTERNAL EVENT (( CAUSE INTERNAL EVENT))                                                                                       |
| ACTION                                                           | → EVENT                                                                                                                          |
| COMPLEX REACTION                                                 | → SIMPLE REACTION CAUSE GOAL                                                                                                     |
| GOAL                                                             | → INTERNAL STATE                                                                                                                 |
| GOAL PATH                                                        | → $\left\{ \begin{array}{l} \text{ATTEMPT CAUSE OUTCOME} \\ \text{GOAL PATH (CAUSE GOAL PATH}^*) \end{array} \right\}$           |
| ATTEMPT                                                          | → EVENT*                                                                                                                         |
| OUTCOME                                                          | → $\left\{ \begin{array}{l} \text{EVENT}^* \\ \text{EPISODE} \end{array} \right\}$                                               |
| ENDING                                                           | → $\left\{ \begin{array}{l} \text{EVENT}^* \text{ (AND EMPHASIS)} \\ \text{EMPHASIS} \\ \text{EPISODE} \end{array} \right\}$     |
| EMPHASIS                                                         | → STATE                                                                                                                          |

## Remembrance of Things Parsed [Mandler and Johnson 1977]

# Narrative theory applied

European Alliances before World War I (1836-1914)

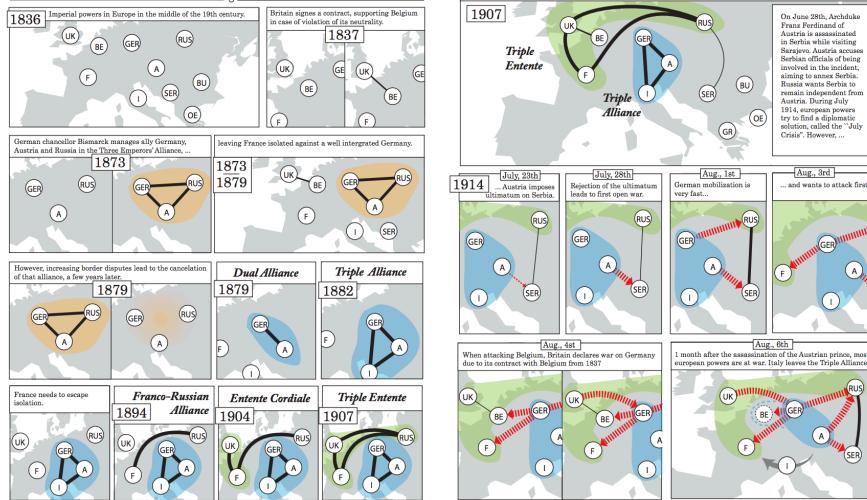


Figure 3. Graph comic example illustrating historical events preceding WWI, targeted to a general audience.

Graph Comics [Bach et al. 2016]

# Narrative theory applied

**C1. Element identity**—Characters in traditional comics are recognizable via their distinct visual features. In graph visualization, nodes and edges are usually represented by circles, sometimes using labels for identifications. Variations in shape and color reflect attributes in the data.

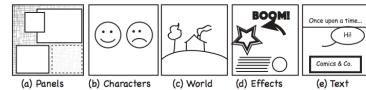
**C2. Depicting change**—In comics, readers can infer changes even if not shown explicitly. For example, a baby being shown implies it has been born. In graphs, a (new) node being shown does not mean it was not there before.

**C3. Order of events**—Many changes exhibit a natural order (e.g. humans grow old, objects fall down, day turns into night). Changes in graphs do not have a natural order nor duration; a cluster can grow or shrink, divide or merge.

**C4. Spatial context**—In comics, the spatial context of a scene is usually indicated by the panel background: the same background means the same place and often the same scene. Switching spatial context from one cluster to another may not be recognized by a reader (both consists of nodes and links).

**C5. Number of elements**—Comics usually show a small set of characters in every panel. Changes in graphs, may involve many more elements than it is possible to show in a panel.

## 1) Traditional Comics



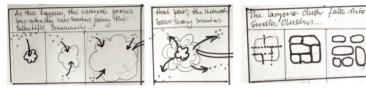
## 2) Designing for Graph Changes



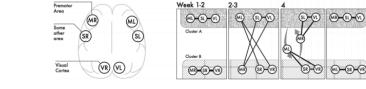
## 3) Design Principles



## 4) Creating Comics



## 5) Discussion with Domain Experts



## 6) Readability Study (see Section "Readability Study")

Graph Comics [Bach et al. 2016]

## Narrative theory applied

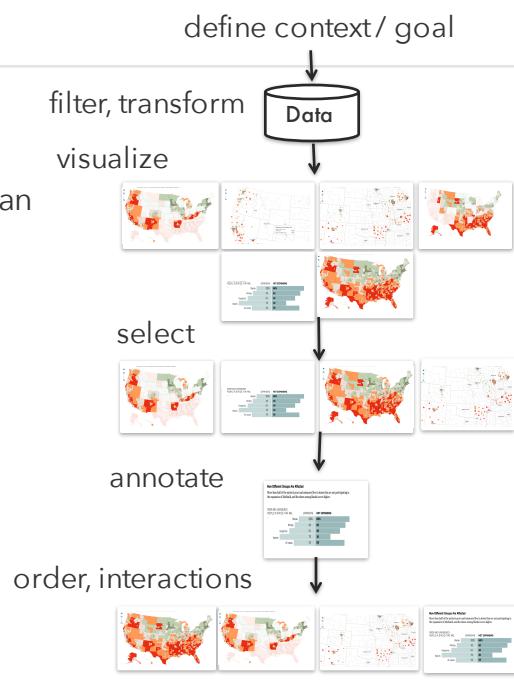
Presentation sequence → recall, interpretation  
[Black & Bower 1979, Pennington & Hastie 1982]



[Hullman et al. 2013]

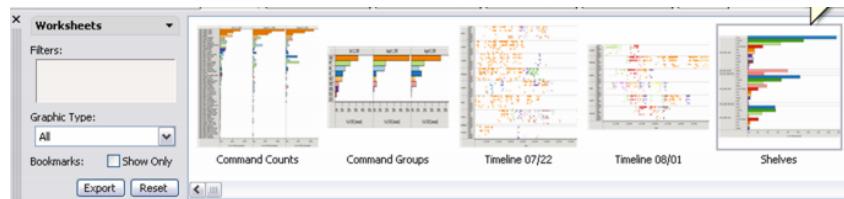
## Design tools

Can we automatically identify sequences to recommend to a human designer?



## Discovery order ≠ presentation order

Interactive dynamics in analysis: Filter, derive, etc.  
[Shneiderman 1996, Heer & Shneiderman 2012]



Provenance: How did I get to this result?

[Heer et al. 2008, Bavoli et al. 2005, Scheidegger et al. 2007]

## What makes an effective sequence?

What do designers and  
journalists do?



Examined 400+ transitions in  
explicitly ordered visualizations

- Interactive slideshows
- Scrollable visualizations
- Animated videos
- Analysis presentations

## Petterson's 7 types

- Dialogue (e.g., question/answer)
- Causal
- Temporal (e.g., chronology, future)
- Hierarchy (general-to-specific)
- Comparative Data (Dimension walk, Measure walk)
- Spatial (e.g., cardinal directions)

## Dimension walk - change I.V.

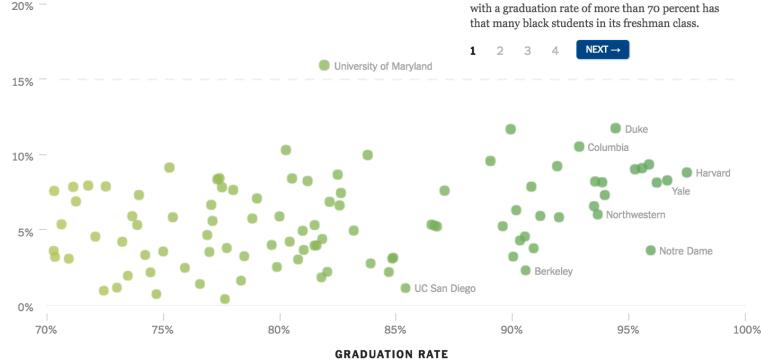
At Top Colleges, an Admissions Gap for Minorities

**BLACKS**  
as a percentage of  
2011–12 freshmen

### Few black students at elite colleges

Roughly 15 percent of public high school graduates are black. But despite the widespread use of affirmative action at elite colleges, only one college with a graduation rate of more than 70 percent has that many black students in its freshman class.

1 2 3 4 **NEXT →**



Note: For-profit colleges and colleges with fewer than 5,000 undergraduates are omitted from the chart.

By JOSH KELLER

Source: National Center for Education Statistics, 2011

## Some observations about transitions

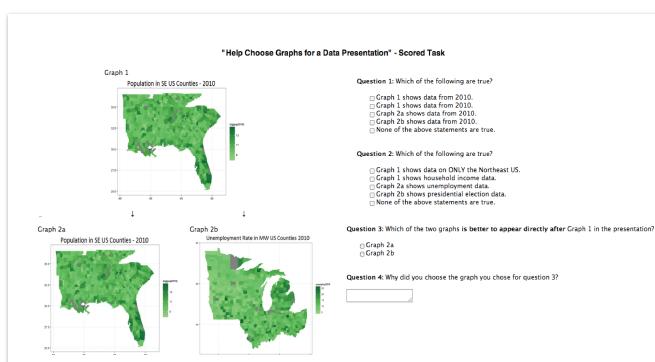
Some types (time, hierarchy, comparative) more common  
Defined by a single change to data schema

- Implies a goal of minimizing *cognitive cost*

## Examining the cost of transitions

Crowd workers decided between pairs of transitions  
and explained choices

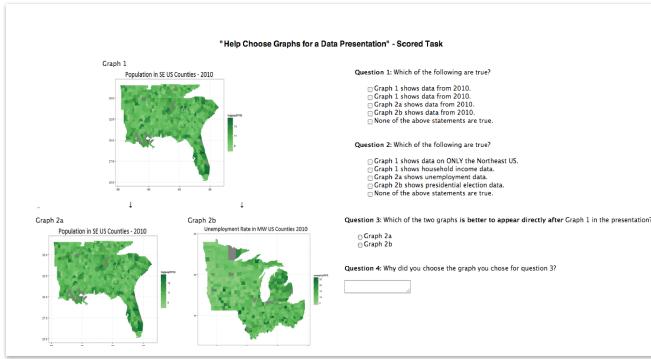
Lower cost transitions are preferred.



## Examining the cost of transitions

Type preferences among equal cost transitions preferred over others:

*Temporal > (Dimension | Measure ) > Hierarchy*



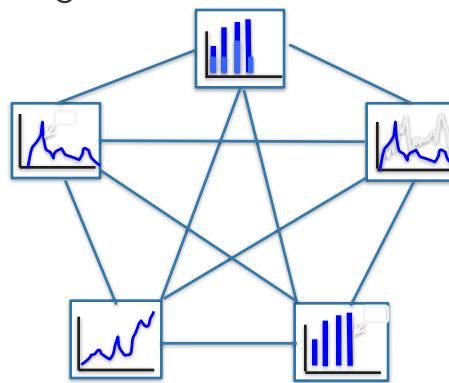
## Some observations about transitions

Graph-based approach

Minimize cost between adjacent visualizations

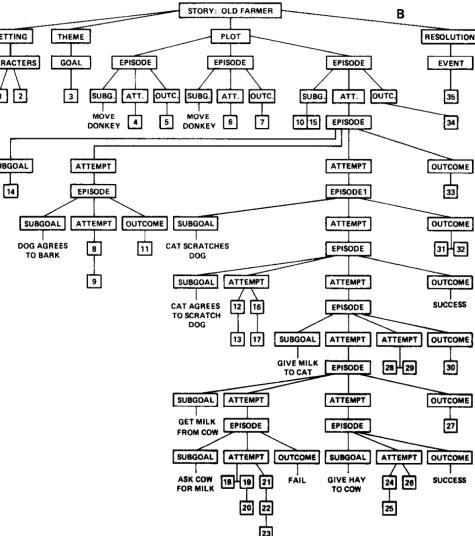
Cost model should account for preferences between types of transitions and changes to data schema

Goal:  
Automatic detection  
of effective sequences



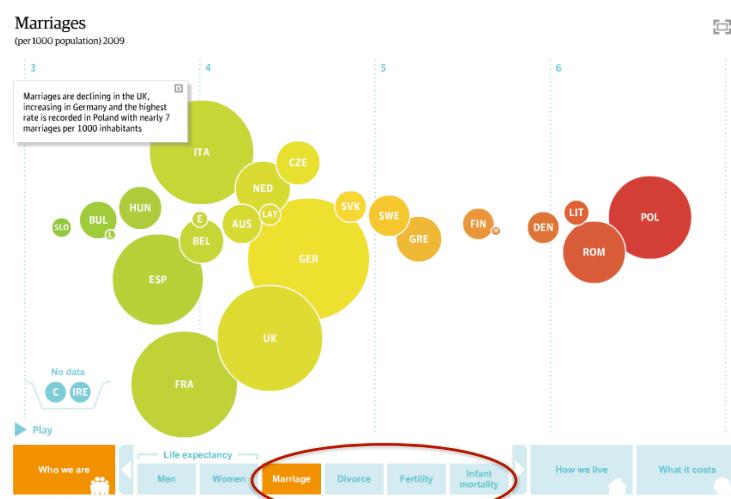
## Micro- vs. macrostructure

Fictional stories are described in terms of how *larger events* are organized (macro-), each of which is comprised of *sub-events* (micro-).

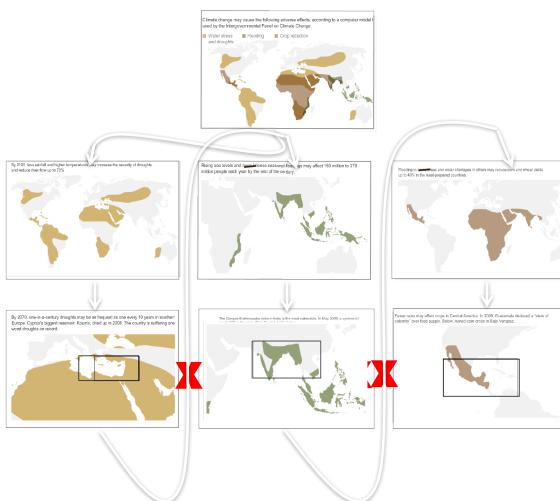


Remembrance of Things Parsed [Mandler and Johnson 1977]

## Multiple transition patterns

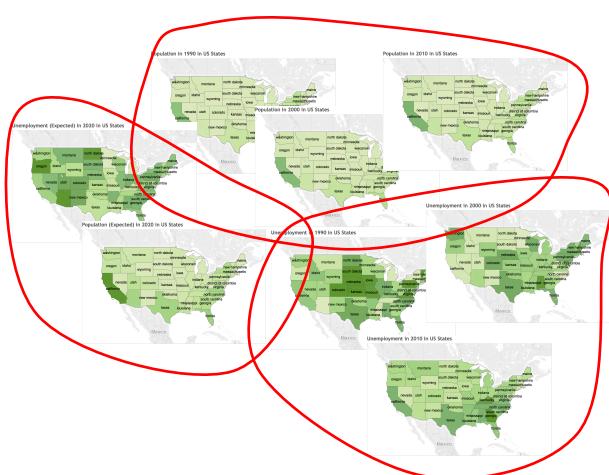


## Parallel structure



## Are groupings preferred despite cost?

How do authors balance high-level and low-level structuring strategies?



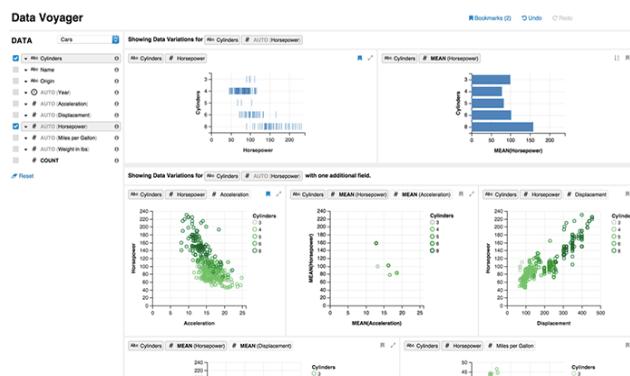
## Narrative structure in visualization: model

Top-down: parallel structure, semantic groupings



Bottom-up: cognitive cost of transitions

## Narrative patterns to support EDA

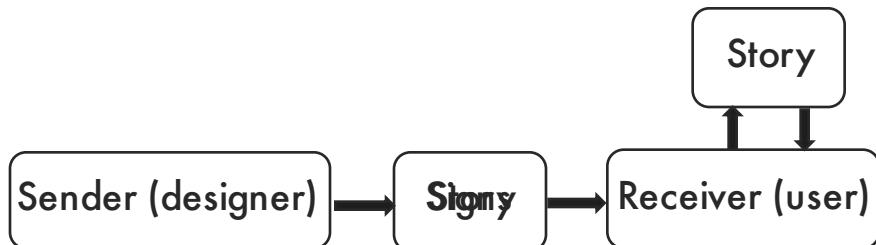


Voyager [Wongsuphasawat et al. 2015]

Suggest paths through recommended visualizations

## Theories: Narrative as constructive

Presentation sequence influences the stories we tell ourselves (Wilensky 1997)



## Theories: Narrative as constructive

### Theories of Data Analysis: From Magical Thinking Through Classical Statistics

Persi Diaconis  
Stanford University



Exploratory data analysis (EDA) seeks to reveal structure, or simple descriptions, in data. We look at numbers or graphs and try to find patterns. We pursue leads suggested by background information, imagination, patterns perceived, and experience with other data analyses.

## A fix for base-rate neglect?

~~1 out of 100~~

~~1%~~ of the women had breast cancer at the time of the screening.

~~80 out of 100~~

Of those with breast cancer, ~~80%~~ received a positive result on the mammogram.

~~15 out of 100~~

Of those without breast cancer, ~~15%~~ received a positive result on the mammogram.

All others received a negative result.

Suppose a woman gets a positive result during a routine mammogram screening. Without knowing any other symptoms, what are the chances she has breast cancer?

[Krynski and Tenenbaum 2007]

## A fix for base-rate neglect?

1% of the women had breast cancer at the time of the screening.

Of those with breast cancer, 80% received a positive result on the mammogram.

**30% of the women had a benign cyst at the time of the screening. Of those with a benign cyst, 50% received a positive result on the mammogram.**

All others received a negative result.

Suppose a woman gets a positive result during a routine mammogram screening. Without knowing any other symptoms, what are the chances she has breast cancer?

[Krynski and Tenenbaum 2007]

## Causal models matter

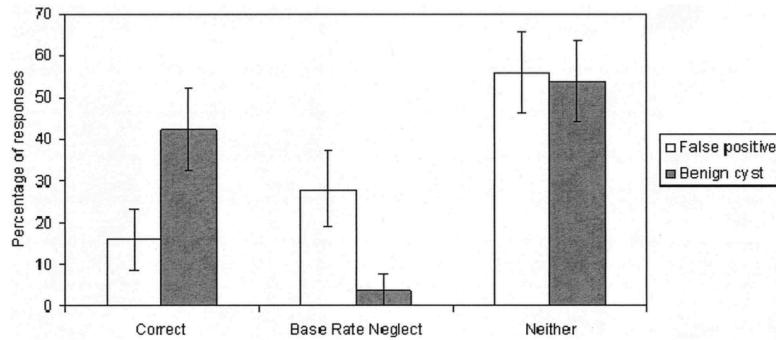


Figure 5. Histogram of responses to Experiment 2. The correct answer was 5.1%. Responses were classified as correct (5.1%), base-rate neglect ( $\geq 65\%$ ), and other. A significant difference was found between false positive and benign cyst scenarios (Fisher's exact test,  $p < .05$ ). Error bars represent the standard error of the normal approximation to the binomial distribution.

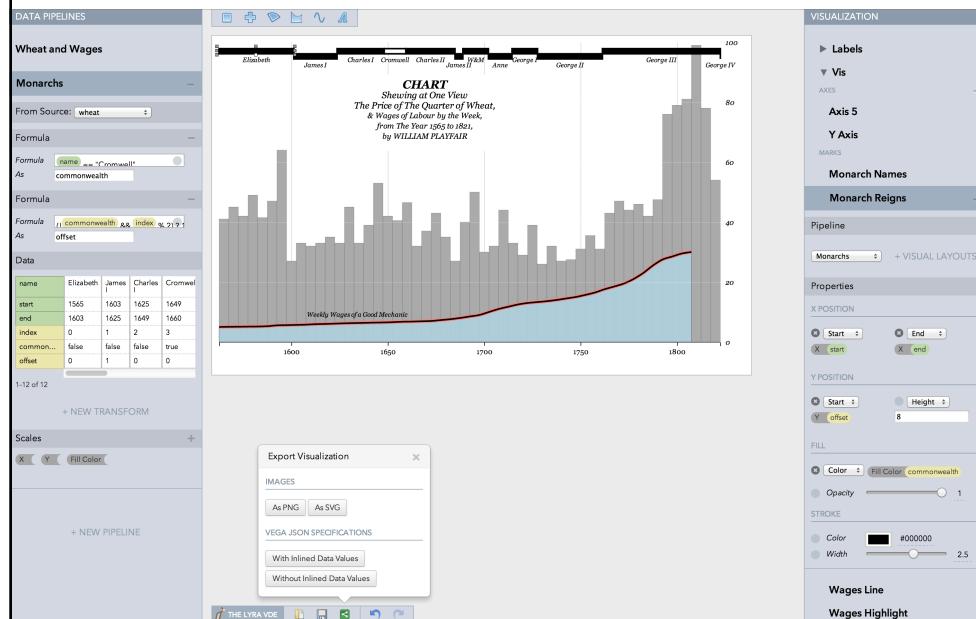
## Construction

# Manual design processes

"require skills like those familiar to movie directors, beyond a technical expert's knowledge of computer science and engineering...."

*Gershon and Page 2001*

## Lyra (Satyanarayan and Heer 2014)

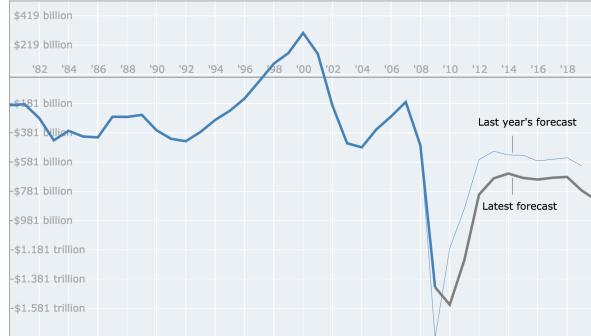


## Ellipsis (Satyanarayan and Heer 2014)

1 2 3 4 5 6

### Forecasts worsen

The forecast for the next decade is somewhat worse than it was a year ago, mostly because of a revised economic outlook. From 2011 to 2020, a total deficit of \$8.5 trillion is expected.



## Tableau Story Points

### Austin's Teacher Turnover Problem

Many Austin schools are still struggling to hold onto teachers, despite a multimillion-dollar effort, called Reach, to stem turnover in troubled schools since 2010.

AISD |  
Many Aust  
turnover in  
how that is

Use the si  
2013

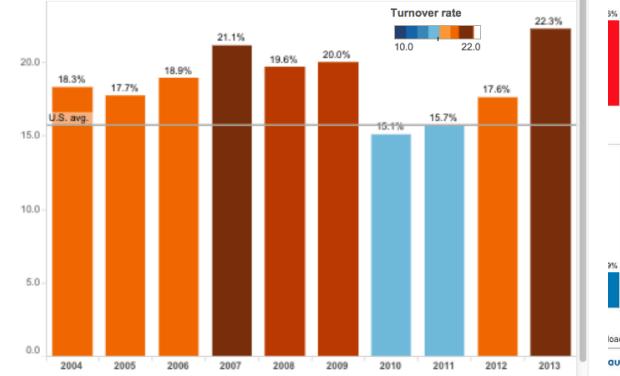
Austin's teacher  
turnover is above  
the national average.

Teacher turnover is a  
city-wide problem.

The problem is  
particularly acute at  
some schools.

And city-wide, the  
problem isn't getting  
better.

Austin's teacher turnover rate by year. Turnover was above the national average for most years.



## Automated Construction

### Requirements

---

A model of effective visual narrative that can be operationalized algorithmically

Examples:

- Automated support for chart reading operations
- Automated generation of annotated news visualizations
- Personalized text stories and graphics

## Brainstorming

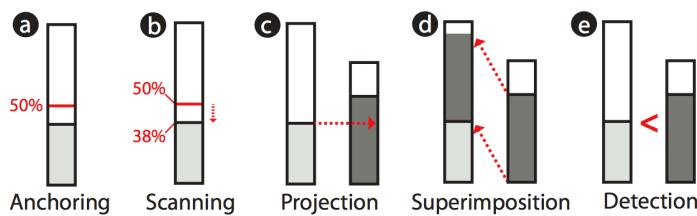
Think about annotations (text or visual) that are useful in presenting or analyzing visualizations.

Are any of these amenable to automated generation?

With 1 or 2 people near you, talk about what types of annotations you would support and how your algorithm would generate them.

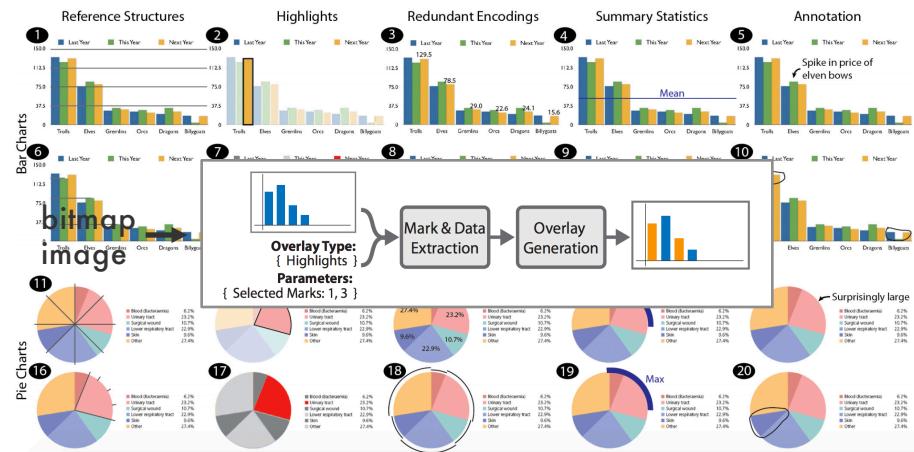
## Facilitating chart reading

Elementary processes in graphical perception



[Simkin and Hastie 1987]

# Facilitating chart reading

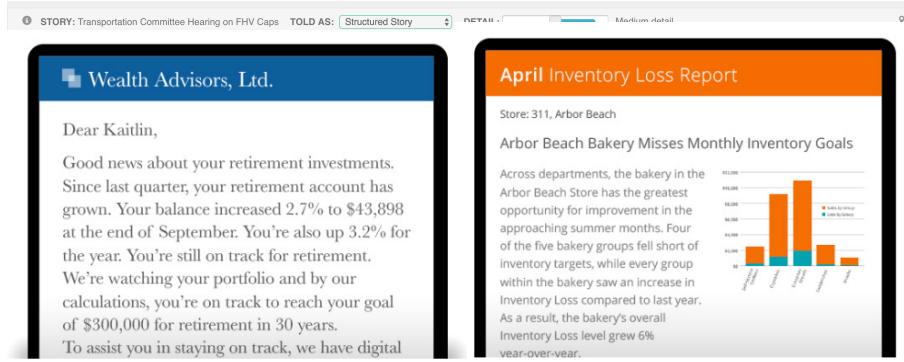


Graphical Overlays [Kong and Agrawala 2012]

# AI-generated stories

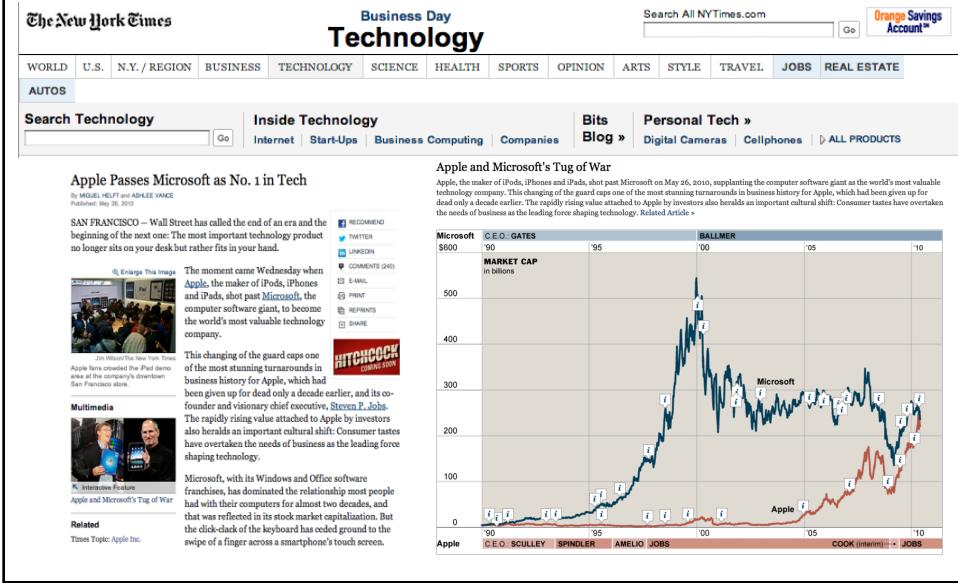
Artificial intelligence based on human narrative cognition

- Conceptual dependency theory (Schank 1969)
- Case based reasoning (Schank 1982)

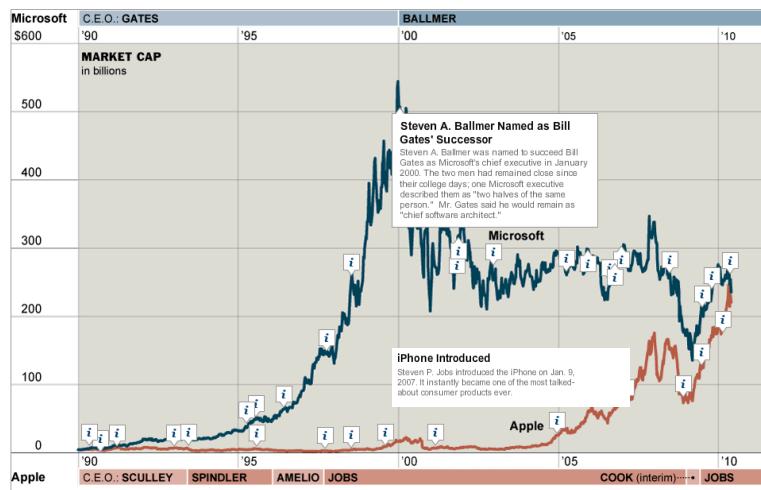


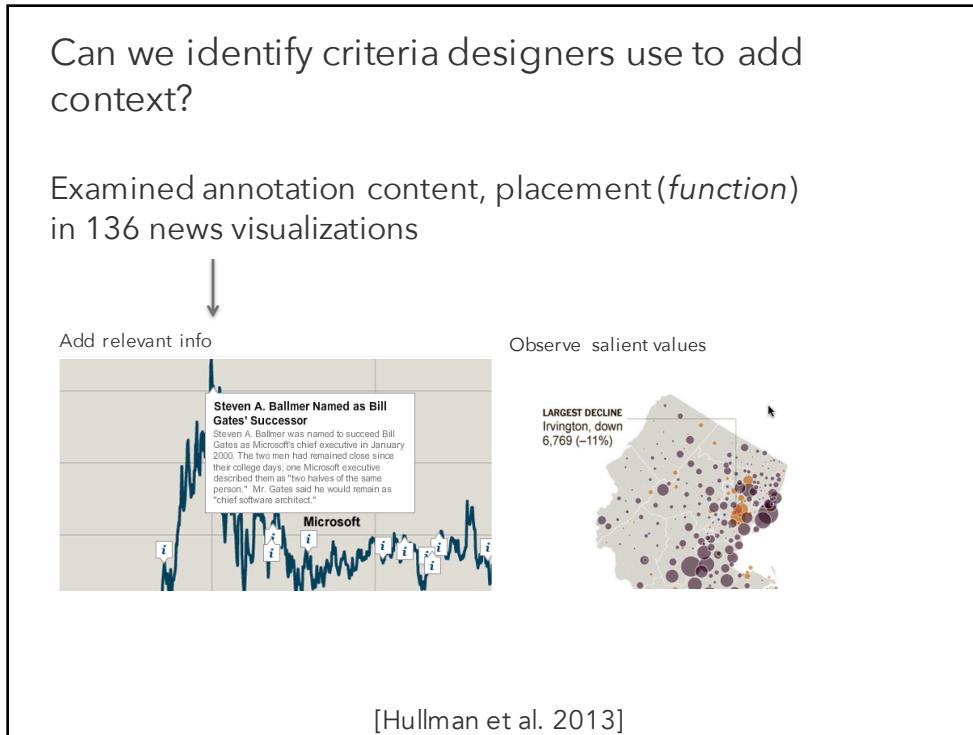
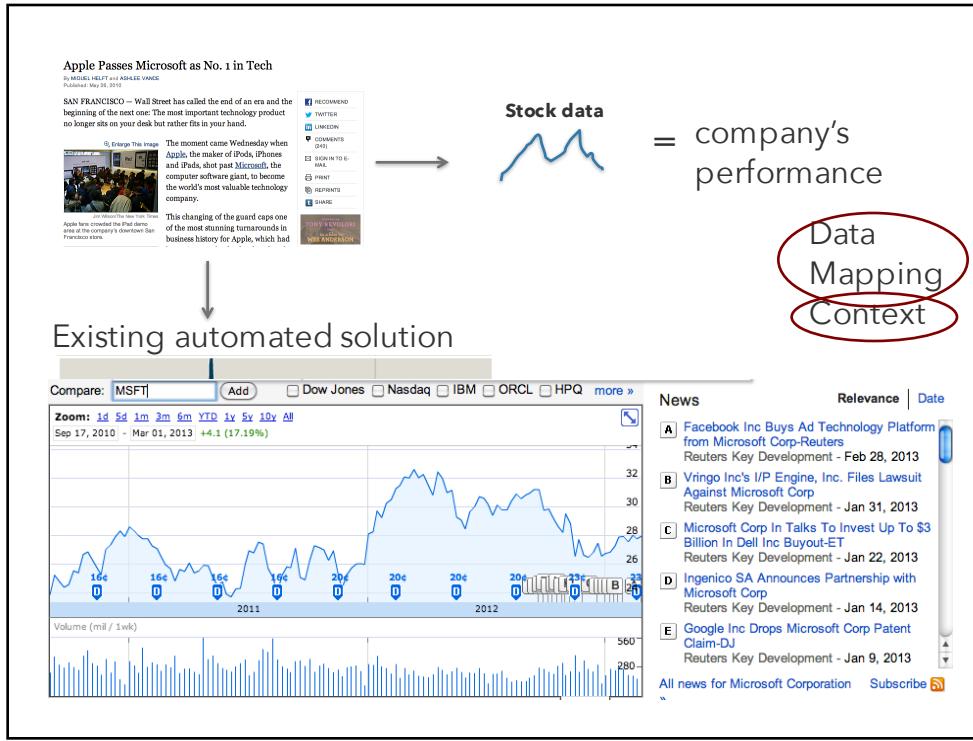
[Narrative Stories]

# Annotated visualizations with news



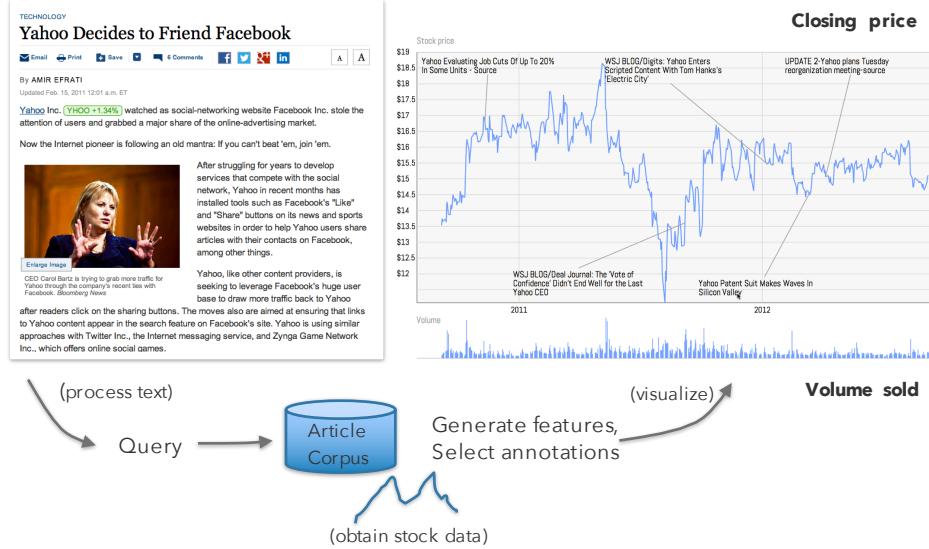
# Annotations tell the story





# Contextifier system

Input: news article about a company



## Relevancy feature

Technology  
Yahoo Decides to Friend Facebook

By AMIR EFRATI  
Updated Feb. 15, 2011 12:01 a.m. ET  
Yahoo Inc. (YHOO +1.3%) watched as social-networking website Facebook Inc. stole the attention of users and grabbed a major share of the online-advertising market.  
Now the Internet pioneer is following an old mantra: If you can't beat 'em, join 'em.  
 CEO Carol Bartz is trying to grab more traffic for Yahoo by teaming up with Facebook. In a recent letter to shareholders, Bartz said she wants to "leverage Facebook's huge user base to draw more traffic back to Yahoo after readers click on the sharing buttons. The moves also are aimed at ensuring that links to Yahoo content appear in the search feature on Facebook's site. Yahoo is using similar approaches with Twitter Inc., the Internet messaging service, and Zynga Game Network Inc., which offers online social games.

Technology Myspace Layoffs Slated for Tuesday By Nat Worden And Scott Morrison Myspace the social-

ALL THINGS DIGITAL: Worst But First:  
Yahoo Uses Words Of Facebook's Zuckerberg To Poke Him In Patent Lawsuit

Kullback-Leibler(KL)  
divergence

$$D_{KL}(P \parallel Q) = \sum_i P(i) \log \frac{P(i)}{Q(i)}$$

Article Corpus

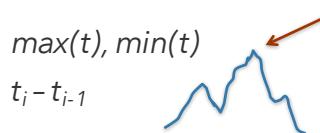
## Article volume



## Visual saliency



Visual saliency proxy via simple analyses on stock series.



**Animal Health Experts Warn Horse Owners About Early Danger Signs of Widespread Mosquito-Borne Diseases; Eastern Equine Encephalitis and West Nile Virus rising Experts Urging Vaccination**

NEW YORK Sept. 7 /PRNewswire/ -- Animal health experts warn that fatal cases of Eastern Equine Encephalitis and West Nile are being reported in numerous states even in areas where activity has been low for several years. This follows a rise in the early warning signs of mosquito-borne diseases that include Eastern Equine Encephalitis West Nile and Western Equine Encephalitis. This doesn't bode well for either horses or humans. To help prevent the spread of additional cases Pfizer Animal Health and other health organizations are strongly encouraging horse owners and veterinarians to follow the American Association of Equine Practitioners (AAEP) guidelines for vaccinations against mosquito-borne diseases. "If horses aren't vaccinated this situation could become much worse" says Kevin Hankins DVM MBA Equine Veterinary Services at Pfizer Animal Health. Though annual vaccinations should happen in early spring the AAEP also recommends boosters after five or six months. Pfizer Animal Health offers a trusted line of vaccines including WEST NILE-INNOVATOR to help protect against West Nile virus. All Pfizer Animal Health equine vaccines are backed by an Immunization Support Guarantee. Pfizer Animal Health will support reasonable diagnostic and treatment costs up to \$5,000 if a horse properly vaccinated by a veterinarian with one of its antigens contracts the corresponding equine disease (EEE WEE WNV Venezuelan Equine Encephalitis (VEE) Tetanus or Influenza).

Pfizer Profit Edges Up Sales Slip

Business Watson Pharmaceuticals Launches Generic Lipitor

UPDATE: Pfizer Files Plans for Animal Health Unit IPO

Pfizer Considers Options For Its Animal Health Nutrition Units

Business Briefs

2011 2012

Combine terms in an objective function

## Personalization

The Best and Worst Places to Grow Up: How Your Area Compares

Children who grow up in some places go on to earn much more than they would if they grew up elsewhere. [MAY 4, 2015](#) | [RELATED ARTICLE](#)

Washtenaw County is very bad for income mobility for children from poor families. It is better than only about 8 percent of counties.

Consider Washtenaw County, Mich., our best guess for where you might be reading this article. (Feel free to change to another place by selecting a new county on the map or using the search boxes throughout this page.)

It's among the worst counties in the U.S. in helping poor children up the income ladder. It ranks 20 out of 2,478 counties, better than only about 8 percent of counties. Compared with the rest of the country, it is also bad for rich boys and rich girls.

Here are the estimates for how much 20 years of childhood in Washtenaw County adds or takes away from a child's income (compared with an average county), along with the national percentile ranking for each.

| What a Childhood in Washtenaw County Does to Future Income |             |                         |          |             |          |
|------------------------------------------------------------|-------------|-------------------------|----------|-------------|----------|
| For poor kids                                              |             | For average-income kids |          |             |          |
| GROUP                                                      | INCOME CHG. | PER CENT                | GROUP    | INCOME CHG. | PER CENT |
| All Kids                                                   | \$2,688     | 8%                      | All Kids | \$3,769     | 4%       |
| Boys                                                       | \$2,790     | 6%                      | Boys     | \$5,190     | <1%      |
| Girls                                                      | \$2,350     | 13%                     | Girls    | \$1,399     | 8%       |

| For rich kids |             |             |          |             |          |
|---------------|-------------|-------------|----------|-------------|----------|
| GROUP         |             | INCOME CHG. |          |             |          |
| GROUP         | INCOME CHG. | PER CENT    | GROUP    | INCOME CHG. | PER CENT |
| All Kids      | \$4,800     | 4%          | All Kids | \$4,880     | <1%      |
| Boys          | \$7,200     | <1%         | Boys     | \$8,770     | <1%      |
| Girls         | \$1,440     | 12%         | Girls    | \$830       | 19%      |

**Location matters** – enormously. If you're poor and live in the Detroit area, it's better to be in Sanilac County than in Wayne County or Genesee County. Not only that, the younger you are when you move to Sanilac, the better you will do on average. Children who move at earlier ages are less likely to become single parents, more likely to go to college and more likely to earn more.

[Aisch, Bloch, Cox, and Quealy 2015]

## Summary

---

Narrative visualizations blend communication /  
exploratory techniques

Messaging, metaphor, sequencing, and other suggestive  
strategies have a powerful impact on interpretation

Semiotics, narrative theory, causal reasoning are critical

Automated systems possible by formalizing features