HOW TO TELL STORIES WITH DATA (REALLY)
edward segel
why interactive?

ANALYSIS  PERSONALIZATION  SOCIAL  STORYTELLING
why interactive?


Businessweek
Top 100 M&A Deals

BGOV
Federal Spending

Analysis
Personalization
Social
Storytelling
why interactive?

- **ANALYSIS**
  - NYTimes: Jobless Rate for People Like You
  - Aaron Koblin: Arcade Fire’s Wilderness Downtown
  - NYTimes: Households like Yours

- **PERSONALIZATION**
  - Make relevant.
  - Increase engagement.
  - Make emotional.

- **SOCIAL**

- **STORYTELLING**
  - UK is offering £6.9bn
  - Offering £25.6bn
  - EFSF: Needed up to £85bn
  - Size of fund: £376.2bn
why interactive?

**PERSONALIZATION**

**ANALYSIS**

**SOCIAL**

**STORYTELLING**

Aaron Koblin  
*Sheep, Cash, Bicycle*

Stamen Design  
*MTV VMA Twitter*

NYTimes  
*Reactions to Osama*

Sharing.  
Collaborative.  
Real-Time sentiment.
why interactive?
interactive storytelling is...

- a new way of telling stories
- a guided tour through data
- interactive and engaging
- becoming wildly popular
STORYTELLING
as ancient as mankind
STORYTELLING changes with technology

PEOPLE TELL STORIES
WORDS TELL STORIES
IMAGES TELL STORIES
COMICS TELL STORIES
MOVIES TELL STORIES
HOW CAN YOU TELL STORIES WITH DATA?
Privacy and the return of Lives of others

Jennifer Mogford, Canada’s privacy commissioner, has written to Facebook in August 2013 to say the social networking site should agree to change its privacy policies within a year. In response, the company has already modified some of its policies. For example, it now allows users to opt out of certain types of data sharing. But as more people use the site, concerns are growing about how data is shared and who can access it. Facebook has been under pressure to change its policies after a series of data breaches, including one in 2012 in which the personal information of millions of users was exposed. The company has also been criticized for collecting too much data on users, including their location, search history, and browsing habits. Facebook has denied these allegations, saying that the data collected is used to improve the user experience and that it is necessary to provide services like advertising. But critics argue that the company is collecting too much data and that it is not transparent about how the information is used. Facebook has also faced criticism for its handling of user privacy, particularly in light of the Cambridge Analytica scandal in 2018. In that case, Facebook was accused of allowing the political consulting firm to access the personal information of millions of users without their consent. Facebook has said that it was not aware of the data sharing and has taken steps to prevent similar incidents from happening in the future. Despite these challenges, Facebook remains one of the most popular social media platforms in the world, with more than 2.5 billion active users as of 2020.
what does it mean?

exciting!

beautiful!

technical!
Budget Forecasts, Compared With Reality

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

Latest forecast

Today, with a better understanding of the severity of the economic downturn, the deficit situation is much more dire.

Use the slider to isolate a single year's forecast.
THE NARRATIVE FRAMEWORK
70% Journalism
20% Business
10% Research
Visual Design

Messaging
- Summaries
- Interpret
- Headlines
- Attached Article
- Captions
- Duo-Specific

Interactivity
- Timelines
- Filtering
- Highlighting
- Details on Demand
- Navigation
- Selection
- Tacit Tutorial

Content Types:
- Magazine Style
- Annotated Chart
- Science Fair Poster
- Flowchart
- Comic Strip
- Slideshow
- Movie
Genres + Interactivity + Messaging = DESIGN SPACE

**Author Driven**
- strong ordering
- heavy messaging
- limited interactivity

**Reader Driven**
- weak ordering
- light messaging
- free interactivity

**STORYTELLING**
- CLARITY
- SPEED

**ASK QUESTIONS**
- EXPLORE
- FIND

martini
glass

interactive slideshow
drill-down story
Common Patterns
- Exponential Tree
- Funnel
- Hub & Spoke
- Linear
- Diamond
- Hour Glass
- Network

Popular Variations
- Loop Backs with Removal
- Precedence (or state-based)
- Linear + Diamond
- Diamond + Diamond

Adapted from Simon Puleo and Carolyn Miller

(more patterns)
THE DESIGN CATALOGUE
INTERACTIVITY
engage

MESSAGING
tell

VISUAL DESIGN
show

INTERACTIVITY
depend

MESSAGING
tell

VISUAL DESIGN
design
Choose your visualization type carefully. Know your options. Even obscure ones.

“Cool” and “readability” are at odds. Recognize the trade-off and choose your audience.
Avoid “chart junk” (Tufte). Extra marks distract from the data.

But...
chart junk may reflect cool design choices

Businessweek...
does great with labeling only as much as needed
Make it clear where to start.
Don’t let readers defect.

World Population: 6,853,328,460
Migrants in the world: 215,738,321

Almost 216 million people, or 3.15% of the world population, live outside their countries.

Click on a country box to know more about migration flow to/from that country.
The more linear, the more like a story. Stories have a beginning, middle, and end.
Consistent visual frameworks. Keep things tidy.

The Social Network

Many people have the Golden Globes winner down as the film to ruin the King's Speech. A slick script and well-paced performances impressed critics and picking a film often is not the easiest task. The Social Network is a film about the initial days of Facebook. The Social Network is a film about the initial days of Facebook. The Social Network is a film about the initial days of Facebook.

8 nominations. 3 wins. The Social Network is a film about the initial days of Facebook. The Social Network is a film about the initial days of Facebook. The Social Network is a film about the initial days of Facebook.

When Dodd-Frank was signed into law a year ago, the lobbying reform and regulatory agencies Now with the ISEF, the financial firms are attempting to craft rules that will determine how Dodd-Frank will change the system.

Dodd-Frank One Year Later: The Key Players

Source: Bloomberg
The jump

Five judges award points to competitors based on factors such as variety of tricks, the height snowboarders reach and the difficulty of tricks and rotations. Points are deducted for bad landings and a lack of speed. White's jump was almost perfect.
Use staging and animation for complicated transitions.
Stage big transitions to avoid confusing readers.
Use establishing shots.
Situate the viewer before diving in.
Use establishing shots. Situate the viewer before diving in.

The course is more gentle at the top and begins to build up speed about midway.

Kumarashvili lost control of his sled as he rounded the final turn on the track.

Speed:
43 m.p.h.

START
Elevation
3,077 feet

Course length:
4,508 feet

Kumarashvili can be seen riding high on the first half of curve 16.

SPHERES OF INFLUENCE: The Bush Campaign Pioneers

George W. Bush reached into each of the spheres surrounding his life to find fundraisers for his first presidential bid. These Pioneers, who raised a minimum of $100,000 in individual contributions of $1,000 or less in 2000, have changed the face of national campaign finance.

ROLL OVER SPHERES TO LEARN MORE

SPHERES OF INFLUENCE: The Bush Campaign Pioneers
Highlighting Techniques

Character Direction
Feature Distinction
Close-Ups
Zooming
Framing
Motion
Audio

Make it clear what to look at and when. Guide readers through the story or they’ll get lost.
Steroids or Not, the Pursuit Is On

Barry Bonds is taking aim at the career home run record. He needs only six more to tie Babe Ruth and 47 to equal Hank Aaron.

Hank Aaron
- 755 home runs
- 23 seasons

Babe Ruth
- 714 home runs
- 22 seasons

Barry Bonds
- 708 home runs
- 20 seasons

Bonds takes lead
- Home runs after 16 seasons
  - Bonds 567
  - Aaron 554
  - Ruth 516

According to allegations in a book about Bonds, he began taking steroids before the 1999 season, his 14th in the league. Two seasons later, he hit 73 home runs, surpassing Aaron's career pace.

Homer Pace After Age 34

If the accusations are correct, Bonds was 34 in his first season on steroids. Here are projected home run paces for each player after age 34.

Aaron
- Actual homers: 46.4
- Projected homers: 42.5

Ruth
- Averaged 64.4 homers a season from age 30 to 34.
  - Averaged 42.5 for next four seasons.

Bonds
- From age 35 to 38, he averaged 14 more homers a season than projected.

Others Taking Aim

Alex Rodriguez
- 429 HR
- Ahead of the pace set by all three home run leaders.

Albert Pujols
- 201 HR
- Average 40 home runs a season, he has started stronger than the three leaders did.

Ken Griffey Jr.
- 538 HR
- Youngest to reach 400 home runs.

Differing Paths to the Top of the Charts

The top seven players on the career home run list, along with a look at Griffey (12th), Rodriguez (9th) and Pujols (tied 5th).
Steroids or Not, the Pursuit Is On

Barry Bonds is taking aim at the career home run record. He needs only six more to tie Babe Ruth and 47 to equal Hank Aaron.

According to allegations in a book about Bonds, he began taking steroids before the 1999 season, his 14th in the league. Two seasons later, he hit 73 home runs, surpassing Aaron's career pace.

Homer Pace After Age 34
If the accusations are correct, Bonds was 24 in his first season on steroids. Here are projected home run rates for each player after age 34.

- Aaron: Actual homers slightly outpace projected homers for five seasons.
- Ruth: Averaged 4.4 homers a season from age 30 to 34. Averaged 42.5 for next four seasons.
- Bonds: From age 35 to 36, he averaged 1.4 more home runs a season than projected.

Differing Paths to the Top of the Charts
The top seven players on the career home run list, along with a look at Griffey (12th), Rodriguez (9th) and Pujols (tied 25th).
Steroids or Not, the Pursuit Is On

Barry Bonds is taking aim at the career home run record. He needs only six more to tie Babe Ruth and 47 to equal Hank Aaron.

According to allegations in a book about Bonds, he began taking steroids before the 1999 season, his 14th in the league. Two seasons later, he hit 73 home runs, surpassing Aaron’s career pace.

Homer Pace After Age 34

If the accusations are correct, Bonds was 34 in his first season on steroids. Here are projected home run rates for each player after age 34.

**Hank Aaron**
- Actual homers: 68
- Outpace projected homers for five seasons.

**Babe Ruth**
- Actual homers: 464
- Average 4.8 homers a season from age 30 to 34.

**Barry Bonds**
- From age 35 to 36, he averaged 1.4 more home runs a season than projected.

Differing Paths to the Top of the Charts

The top seven players on the career home run list, along with a look at Griffey (12th), Rodriguez (97th) and Pujols (tied 257th).

- Hank Aaron: 735
- Babe Ruth: 714
- Barry Bonds: 708
- Willie Mays: 660
- Sammy Sosa: 588
- Frank Robinson: 568
- Mark McGuire: 583
- Ken Griffey Jr.: 531
- Alex Rodriguez: 429
- Albert Pujols: 201
Toyota accelerator
Roll over the circles

The problem
Grooves on the pedal head mesh with grooves on the shoe creating friction. Occasionally this has been so great that the pedal has jammed against the shoe.
Paths to the Top of the Home Run Charts

The players with the most home runs are shown below; click on a player’s name or on a line below to highlight a player’s trajectory.

Hank Aaron
755 home runs
One home run every 16.4 at-bats. 1954 – 1975
The model of power consistency: he had 15 seasons of more than 30 homers, but no seasons of more than 47.

Ted Williams
A Navy pilot in World War II.

In 1974, Aaron broke the home run record that Ruth set in 1935.

Lines above Aaron’s show players who had hit more home runs at that age.
Lines below show players who had hit few home runs.

Base jumping
Metal prices, $ terms, January 1st 2008=100

Source: London Metal Exchange
Weave text into the graphic—not just at the beginning.
Text and graphics work better together than apart.

Connect the text to the relevant graphics.
See Fig. 5

**1.5 million**

Women sued Wal-Mart for discriminatory pay & promotion practices, using mostly statistics to make their case.

Women filled 70% of hourly jobs...
...but only 33% of management.

But Wal-Mart's numbers showed their women fared better than elsewhere in the country.

Women were paid less than men in every region, even when they had higher ratings and seniority.
Weave text into the graphic—not just at the beginning. Text and graphics work better together than apart.

**Word-Specific**
Pictures illustrate the words

**Picture-Specific**
Words accentuate aspects of the scene

**Duo-Specific**
Words and pictures send the same message

**Intersecting**
Words and pictures contribute information independently

**Interdependent**
Word and pictures combine to convey an idea neither conveys alone

**Parallel**
Words and pictures seem to be independent.

**Montage**
Words and pictures combine pictorially.
Start with an editorially interesting view.
Default views can be boring.
Curate the experience from the beginning.

Make data relatable. Put numbers and facts in context.
250 thousand square miles means nothing. It's the size of Texas!
Mind your precision.
Significant digits, tickmarks, and labels suggest what deserves attention.

Good: Lance Armstrong, Popularity issue

### Medicare and Medicaid Spending Show No Signs of Slowing Down

Centers for Medicare and Medicaid Services projects increases in spending compared to gross domestic product and national health expenditures.

**Projected spending**
- Medicare
- Medicaid

**Projected spending as a percent of NHE**
- $922.0 billion
- $561.1 billion

**Projected spending as a percent of GDP**
- Total 2011-2020: $11.1 trillion

---

**Table: Fiscal 2010 contract spending**

<table>
<thead>
<tr>
<th>Representative</th>
<th>District</th>
<th>Fiscal 2010 contract spending</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mike Coffman</td>
<td>Colo. 6</td>
<td>$3.26 billion</td>
</tr>
<tr>
<td>Doug Lamborn</td>
<td>Colo. 5</td>
<td>$2.98 billion</td>
</tr>
<tr>
<td>Rob Bishop</td>
<td>Utah 1</td>
<td>$2.76 billion</td>
</tr>
<tr>
<td>Kenny Marchant</td>
<td>Texas 24</td>
<td>$2.56 billion</td>
</tr>
<tr>
<td>Jeff Duncan</td>
<td>S.C. 3</td>
<td>$2.40 billion</td>
</tr>
<tr>
<td>Sandy Adams</td>
<td>Fla. 24</td>
<td>$2.36 billion</td>
</tr>
<tr>
<td>Lamar Smith</td>
<td>Texas 21</td>
<td>$1.55 billion</td>
</tr>
<tr>
<td>Steve Scalise</td>
<td>La. 1</td>
<td>$1.49 billion</td>
</tr>
<tr>
<td>Jeff Landry</td>
<td>La. 3</td>
<td>$1.47 billion</td>
</tr>
<tr>
<td>Roscoe Bartlett</td>
<td>Md. 6</td>
<td>$1.44 billion</td>
</tr>
<tr>
<td>John Carter</td>
<td>Texas 31</td>
<td>$1.42 billion</td>
</tr>
</tbody>
</table>

---

**Thoughtful**

---

**Too precise**

---

**Fine**

---

**The New York Times** reports a second former teammate has backed allegations of cheating by Armstrong.
Interactive features should scream interactivity. Avoid a click-and-seek experience.

Be explicit.

Click here.

suggested actions

We first looked at whether policy was counter-cyclical in any given quarter.

CONTINUE

visual cues

familiar widgets
Interactive features should react to the user. Depress buttons, highlight items, animate widgets.

Don’t obscure data. Avoid letting pop-ups obscure data.
Make the visualization “look alive”. Things move even without the user!

Demonstrate interactivity. Animate interactive widgets for tacit tutorials.
Restrict interactivity to key dimensions. The more interactivity, the less story.

Include a progress bar and “back” and “reset” buttons. Linear navigation is important—especially with lots of interactivity.
Sound & Video

Mobile

Linking

Fluid

THE FUTURE
DATA JOURNALISM IS NOT GRAPHICS AND VISUALIZATIONS. IT’S ABOUT TELLING THE STORY IN THE BEST WAY POSSIBLE. SOMETIMES THAT WILL BE A VISUALIZATION... BUT SOMETIMES IT’S A NEWS STORY. SOMETIMES, JUST PUBLISHING THE NUMBER IS ENOUGH.

SIMON ROGERS, THE GUARDIAN // 2011
PEOPLE HAVE BEGUN TO FORGET HOW POWERFUL HUMAN STORIES ARE, EXCHANGING THEIR SENSE OF EMPATHY FOR A FETISHISTIC FASCINATION WITH DATA... THE HUMAN STUFF IS THE MAIN STUFF, AND THE DATA SHOULD ENRICH IT.

JONATHAN HARRIS // 2008
(1) Choose your visualization type carefully. Know your options. Even obscure ones.

(2) The more linear, the more like a story. Stories have a beginning, middle, and end.

(3) Use establishing shots. Situate the viewer before diving in.

(4) Make it clear what to look at and when—especially where to start. Guide readers through the story or they'll get lost.

(5) Limit complexity at first. Reveal as needed. Don't confuse the reader with extraneous information.

(6) Consistent visual frameworks. Keep things tidy.

(7) Use staging and animation for complicated transitions. Stage big transitions to avoid confusing readers.

(8) Design Matters. Duh.

(9) Avoid “chart junk” (Tufte). Extra marks distract from the data.

(10) “Cool” and “readability” are at odds. Recognize the trade-off and choose your audience.
(1) The more text, the more storytelling. Graphs are abstract. Words give concrete interpretations.

(2) Say the point you’re making with the graphic. Don’t make the reader figure it out on his own.

(3) Use headlines, captions, & annotations. Quickly draw attention to what’s important.

(4) Weave text into the graphic—not just the beginning. Text and graphics work better together than apart.

(5) Start with an editorially interesting view. Default views can be boring. Curate the experience from the beginning.

(6) Make data relatable. Put numbers and facts in context. 250 thousand square miles means nothing. It’s the size of Texas!

(7) Connect the text to the relevant graphics. See Fig. 5

(8) Don’t ignore summaries & conclusions. Answer the “so what?” or the reader leaves empty handed.

(9) Mind your precision. Significant digits, tickmarks, and labels suggest what deserves attention.
(1) Demonstrate interactivity. 
Animate interactive widgets for tacit tutorials.

(2) Interactive features should scream interactivity. 
Avoid a click-and-seek experience.

(3) Interactive features should react to the user. 
Depress buttons, highlight items, animate widgets.

(4) Restrict interactivity to key dimensions. 
The more interactivity, the less story.

(5) Make the visualization “look alive”. 
Things move even without the user!

(6) Don’t obscure data. 
Avoid letting pop-ups obscure data.

(7) Include a progress bar and “back” and “reset” buttons 
Linear navigation is important—especially with lots of interactivity.

New York Times. How many households are like yours?

New York Times. The Death of a Terrorist: a Turning Point?

http://www.nytimes.com/ref/sports/20070731_BONDS_GRAPHIC.html


New York Times. Aerial Photographs of Joplin Before and After the Tornado

New York Times. How different groups spend their day.


New York Times. Alpine Skiing, From Technical Turns to Tucks and Speed


The Guardian

Guardian. Toyota's sticky accelerator problem.
http://www.guardian.co.uk/business/interactive/2010/feb/04/toyota-automotive-industry

Guardian. Shaun White's Double McTwist.
http://www.guardian.co.uk/sport/interactive/2010/feb/19/winterolympics2010-vancouver

Guardian. Earthquakes: Why they happen.
http://www.guardian.co.uk/world/interactive/2008/jan/23/earthquakes

Guardian. Oscars 2011: an interactive guide to this year’s best picture nominees

Guardian. Inflation in the UK.

Guardian. The World’s Economy Turns the Corner.
http://www.guardian.co.uk/business/interactive/2010/Jan/26/recession-gdp

Miscellaneous

Aaron Koblin. Personal website with work.
http://www.aaronkoblin.com/work.html

Jeff Heer. Seneus.
http://vis.stanford.edu/papers/senseus

Stamen. MTV Video Music Awards.
http://stamen.com/clients/mtv

Simon Rogers. Data journalism at the Guardian: what is it and how do we do it?

GapMinder
http://www.gapminder.org


People Movin
http://peoplemov.in/

Visual Complexity. Football Drawings.
http://www.visualcomplexity.com/vc/project_details.cfm?id=452&index=452&domain

Dan Archer. Nisoor square shootings of 2007
http://www.cartoonmovement.com/icomic/11
BGOV

BGOV. Federal Programs on Debt-Limit Bubble
http://about.bgov.com/2011/07/12/august-invoices-show-u-s-treasury%E2%80%99s-limited-choices/

BGOV. Hospitals Wary of Data Breaches Seek Cyber Insurance From AIG
http://www.bgov.com/news_item/4D53U6teO2rko9nUq8XDoQ

BGOV. Medicare-Medicaid Costs $1.5 Trillion by 2020: Chart of the Day.
http://www.bgov.com/news_item/XwqM2luxD58Dh0n6lvp8WA

BGOV. Kentucky Bourbon’s Seoul Appeal May Rise With Trade Agreement.
http://www.bgov.com/news_item/PtRA9_pVGxAHlVVsiZev_q

BGOV. Preparing Americans for Death Lets Hospices Neglect End of Life.
http://www.bgov.com/news_item/On7zzinP-22zll6mQUeiqw

BGOV. Tea Party Debt Deal Opponents Gain From U.S. Awards.
http://www.bgov.com/news_item/JcPzhvW8bn8Luuho8KNA

BGOV. August Invoices show Treasury’s Limited Choices.
http://about.bgov.com/2011/07/12/august-invoices-show-u-s-treasury%E2%80%99s-limited-choices/

BGOV. Dodd-Frank One Year Later: The Key Players
http://www.bgov.com/news_item/WMr8I6JoxhifG-KDDpytg

Businessweek

Businessweek. The Credit Boom’s Best and Worst Deals.
http://images.businessweek.com/mz/10/34/1034mergers47.pdf?chan=magazine+channel_news++markets+%2B+finance

Businessweek. Amid doping furer, Lance Armstrong’s reputation heads downhill.
http://images.businessweek.com/mz/10/34/1034endorsments22.pdf?chan=magazine+channel_news++companies+%2B+industries

Bloomberg Interactive Insights. America tied up by debt record.

Businessweek. How to save Greece.
http://www.businessweek.com/magazine/content/11_27/b4235000567802.htm

Businessweek. The Big Apple is Recovering its Shine.
http://images.businessweek.com/mz/10/34/1034therecovery11.pdf?chan=magazine+channel_news++global+economics

Businessweek. Problem? What Problem?
http://www.businessweek.com/magazine/content/11_26/b4234013596135.htm

Businessweek. Oil on Troubled Waters.
http://images.businessweek.com/mz/10/34/1034oilbp53.pdf
The Washington Post
Washington Post. Obama’s Calendar.
http://projects.washingtonpost.com/potus-tracker/
