CSE 412
Narrative Visualization

Matthew Conlen
(with material from Jeff Heer, Edward Segel, and Jessica Hullman)
About Me

Ph.D. student @ UW
    Advised by Jeff Heer

Background in data journalism
    FiveThirtyEight, CNN, New Yorker, New York Times (current)

Lecture will be biased toward data journalism but the material is more generally relevant (education, scientific publishing, policymaking).
Narrative Storytelling

**Story:**
The events or facts you want to convey

**Narrative:**
How you convey those facts
Narrative Storytelling

**narrative (n):** An account of a series of events, facts, etc., given in order and with the establishing of connections between them.

Effective storytelling “require[s] skills like those familiar to movie directors, beyond a technical expert’s knowledge of computer engineering and science.” - Gershon & Page ‘01
Going forward
I carried wax along the line, and laid it thick on their ears. They tied me up, then, plumb amidships, back to the mast, lashed to the mast, and took themselves again to rowing. Soon, as we came smartly within hailing distance, the two Sirens, noting our fast ship, off their point, made ready, and they sang...
STORYTELLING
across various media

PEOPLE TELL STORIES
WORDS TELL STORIES
IMAGES TELL STORIES
COMICS TELL STORIES
MOVIES TELL STORIES
By Denise Grady and Carl Zimmer

April 14, 2021, 9:34 p.m. ET

The pause in the use of Johnson & Johnson’s Covid-19 vaccine may continue for a week to 10 days, after expert advisers to the Centers for Disease Control and Prevention determined on Wednesday that they needed more time to assess a possible link to a rare but serious blood-clotting disorder.

The decision not to reinstate the vaccine has painful consequences, nationally and globally. It may further erode public confidence in vaccination in general and slow the rollout of desperately needed shots to rural and underserved areas and homebound people. The vaccine is considered ideal for hard-to-reach people and places because it requires only one shot and is more easily stored and shipped than the vaccines made by Moderna and Pfizer-BioNTech, which must be kept at very low temperatures.
By Denise Grady and Carl Zimmer

April 14, 2021, 9:34 p.m. ET

The pause in the use of Johnson & Johnson’s Covid-19 vaccine may continue for a week to 10 days, after expert advisers to the Centers for Disease Control and Prevention determined on Wednesday that they needed more time to assess a possible link to a rare but serious blood-clotting disorder.

The decision not to reinstate the vaccine has painful consequences, nationally and globally. It may further erode public confidence in vaccination in general and slow the rollout of desperately needed shots to rural and underserved areas of the country and people. The vaccine is considered ideal for hard-to-reach people and places because it requires only one shot and is more easily stored and shipped than the vaccines made by Moderna and Pfizer-BioNTech, which must be kept at very low temperatures.
By Denise Grady and Carl Zimmer

April 14, 2021, 9:34 p.m. ET

The pause in the use of Johnson & Johnson’s Covid-19 vaccine may continue for a week to 10 days, after expert advisers to the Centers for Disease Control and Prevention determined on Wednesday that they needed more time to assess a possible link to a rare but serious blood-clotting disorder.

The decision not to reinstate the vaccine has painful consequences, nationally and globally. It may further erode public confidence in vaccination in general and slow the rollout of desperately needed shots to rural and underserved areas and homebound people. The vaccine is considered ideal for hard-to-reach people and places because it requires only one shot and is more easily stored and shipped than the vaccines made by Moderna and Pfizer-BioNTech, which must be kept at very low temperatures.
Topics

Narrative visualization
  Case Study: Coronavirus Testing
  Design space

Interactive articles
  Case Study: Carbon Footprint
  Authoring tools
Storytelling... with data?
Privacy and the internet

Lives of others

Facebook and Google face a backlash from users and regulators alike over the way they have handled sensitive data.

Jennifer Scott-Clark, Canada’s privacy commissioner, recently issued a warning to Facebook in August that the social networking giant should overhaul its privacy practices. In a letter to Facebook, the commissioner expressed concern about the company’s handling of user data, particularly in light of recent controversies involving data breaches.

There is a growing awareness among users of the risks associated with sharing personal information online. Many are calling for greater transparency and accountability from social media companies.

The Facebook data scandal, which first surfaced in 2018, highlighted how the company can collect and share user data with third parties, raising concerns about user privacy. In response, Facebook implemented changes to its privacy settings and launched a new feature allowing users to download their data.

Experts believe that companies need to be more transparent about how they collect and use user data. They argue that users should have more control over their personal information and that companies should be held accountable for their data practices.

The European Union’s General Data Protection Regulation (GDPR) has set new standards for data protection in Europe, but many other countries have yet to implement similar regulations.

The Facebook data scandal has also sparked a broader conversation about trust in social media. Users are demanding more control over their data and are increasingly skeptical of social media companies.

In response, Facebook has pledged to make changes to its privacy practices and to be more transparent with users. The company has also launched a new feature allowing users to download their data and to control who can see it.

The Facebook data scandal has raised important questions about the role of social media in our lives and how we should balance the benefits of sharing information with the risks of privacy.

In the coming years, it is likely that social media companies will continue to face pressure to improve their privacy practices and to build trust with users. The Facebook data scandal has set a new standard for data protection and has shown that companies must be held accountable for their actions.
About 21,000 coronavirus cases were reported per day in early June, when the positive test rate was 4.8 percent.
CASE STUDY COVID-19 TESTING

Input Data
Daily case and test counts.

Points of Focus
1. Baseline (smoothed average of daily cases)
2. Counterfactual case curve (baseline positivity rate * daily test counts),
3. Actual case curve (daily cases)
Narrative Devices
What devices communicates best?
Highly dependent on: audience, context, format

Format
Interactive Article
Animation
Lecture
Still Image
Video
Change in teen shenanigans from 1993 to 2013. Showing % difference from 2003 levels.
Obesity Map Vadim Ogievetsky
Obesity Map Vadim Ogievetsky
Narrative Visualization
Design Space
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.

Lines are cumulative home runs.

Hank Aaron
755 homers
23 seasons

Babe Ruth
714 homers
22 seasons

Barry Bonds
708 homers
20 seasons

Bonds takes lead
Home runs
755
after 16 seasons
567
Aaron
554
Ruth
516

Bonds was injured last season. He played 14 games and hit 5 home runs.

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.

PROJECTED BASED ON AVERAGE OF PREVIOUS FIVE SEASONS

Aaron
Actual home run rate
slightly outpace projected home run rates for five seasons.

Ruth
Averaged 46.4 home run per season from age 30 to 34. Averaged 42.5 for next four seasons.

Bonds
From age 35 to 39, he averaged 14 more homers a season than projected.

Note: Ages as of July 1 of each season.

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 (37th) and Pujols (tied 257th).

Hank Aaron
755
(23 SEASONS)
755
1974
1976
15 times hit 30 or more (M.L. most).
Hit only 20 over first five seasons.
Averaged 52 from 2000 to 2004.
No one hit more from 1950-69.
Three 60-home run seasons is record.
Triple Crown in '66 (49, 122, 316).
First to hit 70 in a season.
Only McGwire had more in the 90's.
Youngest to reach 400 home runs.
Second most ever in first five seasons.

Babe Ruth
714
(22 SEASONS)
714
1965
1973
Hit only 20 over first five seasons.
Averaged 52 from 1950-69.
Three 60-home run seasons is record.
Triple Crown in '66 (49, 122, 316).
First to hit 70 in a season.
Only McGwire had more in the 90's.
Youngest to reach 400 home runs.
Second most ever in first five seasons.

Barry Bonds
708
(20 SEASONS)
708
1960
1966
Averaged 52 from 1950-69.
Three 60-home run seasons is record.
Triple Crown in '66 (49, 122, 316).
First to hit 70 in a season.
Only McGwire had more in the 90's.
Youngest to reach 400 home runs.
Second most ever in first five seasons.

Willie Mays
660
1965
1966
Averaged 52 from 1950-69.
Three 60-home run seasons is record.
Triple Crown in '66 (49, 122, 316).
First to hit 70 in a season.
Only McGwire had more in the 90's.
Youngest to reach 400 home runs.
Second most ever in first five seasons.

Sonny Sosa
588
1965
1966
Averaged 52 from 1950-69.
Three 60-home run seasons is record.
Triple Crown in '66 (49, 122, 316).
First to hit 70 in a season.
Only McGwire had more in the 90's.
Youngest to reach 400 home runs.
Second most ever in first five seasons.

Frank Robinson
586
1965
1966
Averaged 52 from 1950-69.
Three 60-home run seasons is record.
Triple Crown in '66 (49, 122, 316).
First to hit 70 in a season.
Only McGwire had more in the 90's.
Youngest to reach 400 home runs.
Second most ever in first five seasons.

Mark McGwire
583
1965
1966
Averaged 52 from 1950-69.
Three 60-home run seasons is record.
Triple Crown in '66 (49, 122, 316).
First to hit 70 in a season.
Only McGwire had more in the 90's.
Youngest to reach 400 home runs.
Second most ever in first five seasons.

Ken Griffey Jr.
536
1965
1966
Averaged 52 from 1950-69.
Three 60-home run seasons is record.
Triple Crown in '66 (49, 122, 316).
First to hit 70 in a season.
Only McGwire had more in the 90's.
Youngest to reach 400 home runs.
Second most ever in first five seasons.

Alex Rodriguez
429
1965
1966
Averaged 52 from 1950-69.
Three 60-home run seasons is record.
Triple Crown in '66 (49, 122, 316).
First to hit 70 in a season.
Only McGwire had more in the 90's.
Youngest to reach 400 home runs.
Second most ever in first five seasons.

Albert Pujols
261
1965
1966
Averaged 52 from 1950-69.
Three 60-home run seasons is record.
Triple Crown in '66 (49, 122, 316).
First to hit 70 in a season.
Only McGwire had more in the 90's.
Youngest to reach 400 home runs.
Second most ever in first five seasons.
Steroids or Not, the Pursuit Is On

Bonds is trying to break the career home run record held by Hank Aaron. He needs 66 home runs in the next 6 seasons to match the record. The graph shows Bonds' progress towards breaking the record, with projections for future seasons.

Hank Aaron
755 home runs
23 seasons

Babe Ruth
714 home runs
22 seasons

Barry Bonds
708 home runs
20 seasons

According to allegations in a book about Bonds, he began taking steroids before the 1989 season. He hit 146 home runs in 1990, the season after the book was released. Two seasons later, he hit 73 home runs, surpassing Aaron's career pace.

Bonds was injured last season. He played 14 games and hit 5 home runs.

Homer Pace After Age 34

If the allegations are correct, Bonds was 34 in his first season on steroids. Here are projections for his home run pace for each player after age 34:

- **Aaron**
  - Actual home run pace
  - Averaged 46.4 home runs a season from age 30 to 34. Averaged 42.5 for next four seasons.
- **Ruth**
  - Averaged 46.4 home runs a season from age 30 to 34.
- **Bonds**
  - From age 35 to 39, he averaged 14 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 (37th) and Pujols (tied 257th):

- **Hank Aaron**
  - Hit only 20 over first five seasons.
  - Averaged 52 home runs from 1955 to 1966.
  - Three 60-home run seasons in a row.
- **Babe Ruth**
  - Hit 60 or more in 12 seasons.
  - Hit 60 or more in 12 seasons.
  - Hit 60 or more in 12 seasons.
- **Barry Bonds**
  - Hit 60 or more in 12 seasons.
  - Hit 60 or more in 12 seasons.
  - Hit 60 or more in 12 seasons.
- **Willie Mays**
  - Hit 60 or more in 12 seasons.
  - Hit 60 or more in 12 seasons.
  - Hit 60 or more in 12 seasons.
- **Sandy Koufax**
  - Hit 60 or more in 12 seasons.
  - Hit 60 or more in 12 seasons.
  - Hit 60 or more in 12 seasons.
- **Frank Robinson**
  - Hit 60 or more in 12 seasons.
  - Hit 60 or more in 12 seasons.
  - Hit 60 or more in 12 seasons.
- **Mark McGwire**
  - Hit 60 or more in 12 seasons.
  - Hit 60 or more in 12 seasons.
  - Hit 60 or more in 12 seasons.
- **Ken Griffey Jr.**
  - Hit 60 or more in 12 seasons.
  - Hit 60 or more in 12 seasons.
  - Hit 60 or more in 12 seasons.
- **Alex Rodriguez**
  - Hit 60 or more in 12 seasons.
  - Hit 60 or more in 12 seasons.
  - Hit 60 or more in 12 seasons.
- **Albert Pujols**
  - Hit 60 or more in 12 seasons.
  - Hit 60 or more in 12 seasons.
  - Hit 60 or more in 12 seasons.
58
CASE STUDIES

70% Journalism
20% Business
10% Research

Segel & Heer, 2010
Observed Narrative Devices

Visualization Description

Case Studies

Architecture and Justice (Brooklyn Crime Blocks)

Who: Did Your County’s Jobs Disappear?

Acid Politics House Price Index

Band Earnings: How Compensation Relates to Performance

Dead offensive: Taliban attacks in Pakistan

CDO Moves by Sector

UK Economic Data

Budget 2010: Reaction from around the UK

Ford One 2010: driver’s rankings

Lighting up Hadrian’s wall

Mapping hydropower hotspots across the UK

Meteor’s metric bombs: interactive map

The grid economy turns the corner

The 100 of Olympic Medals

All 12 (inflation’s Little Parts)

Bar: Bonds visualization (by perspective)

The history and flow of multinational’s cost of war, 1966—2008

The bliss Rate for People Living in the U.S., 1970-1999

Advertisement: Bus

Advertisement: Helicopter

Analysing Obama’s schedule

Oct 2, 2010: News you can use: recession

The consumer and the consumer

UK: Unemployment

Bar: Bonds Visualization (base image)

Comparison of Bear Markets

How Americans Spend Their Day

Middle Obama’s Family Tree

Net Rentals

Van Gogh’s Olympic Venue

On the Map: Five Major North Korean Prison Camps

A Visual Guide to the Financial Crisis

Economic Meltdown of 2008–2009

Did the Trades Go the Way of the Dodo?

Life cycle of a beetle through a year

McCullin’s “Making Comics”

In the Government Dealt with Past Recession

Afghanistan: Behind the front line

Troy: A Timeline: a company history

Global Human Development

Earthquakes: why they happen

Iraqi nuclear programme

Shit: White’s Double McTwist

Troy’s stick accelerator problem

Afghani: Skiing, From Technical Turns to Tucks and Speed

Budget Forecasts vs. Reality

Mentor: Presentation Video

Defibrillating (Adobe Illustrator)

The Story of Stuff

Virgin America Airplane Safety Video
Genres for Narrative Visualization (2010)
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
Interactive Articles
In comparison, only 12% of YouTube’s traffic went through cellular networks, but they were by far the most expensive part of YouTube’s content delivery pipeline, accounting for approximately \textbf{8,500 Gigawatt-hours of electricity usage—enough to power over 750,000 U.S. homes} \footnote{[3]}. At over 10 times the electricity usage per unit of traffic, the relative inefficiency of cellular transmission is clear.
Communicating with Interactive Articles

Examining the design of interactive articles by synthesizing theory from disciplines such as education, journalism, and visualization.

60+ Interactive Articles
NYTimes, WaPost
Distill, VisXAI
+more

Hohman, Conlen, Heer, Chau
2020
Communicating with Interactive Articles

Examining the design of interactive articles by synthesizing theory from disciplines such as education, journalism, and visualization.

5 affordances of the format
Connecting people and data
Making systems playful
Promoting self reflection
Personalizing reading
Reducing cognitive load

Hohman, Conlen, Heer, Chau
2020
5 AFFORDANCES OF THE FORMAT

Connecting people and data
Making systems playful
Promoting self reflection
Personalizing reading
Reducing cognitive load

Follow the lives of 932 boys who grew up in rich families...

Grew up rich

Most white boys raised in wealthy families will stay rich or upper middle class as adults, but black boys raised in similarly rich households will not.

...and see where they end up as adults:

<table>
<thead>
<tr>
<th>Class</th>
<th>White Men</th>
<th>Black Men</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rich adult</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Upper-middle-class adult</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Middle-class adult</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Lower-middle-class adult</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Poor adult</td>
<td>0%</td>
<td>0%</td>
</tr>
</tbody>
</table>

Adult outcomes reflect household incomes in 2014 and 2015.

Even when children grow up next to each other with parents who earn similar incomes, black boys fare worse than white boys in terms of income as adults regardless of demographic factors.
Connecting people and data

Making systems playful
Promoting self reflection
Personalizing reading
Reducing cognitive load

5 AFFORDANCES OF THE FORMAT

Teachable Machine

Train a computer to recognize your own images, sounds, & poses.

A fast, easy way to create machine learning models for your sites, apps, and more – no expertise or coding required.

Get Started

What is Teachable Machine?
You Draw It: What Got Better or Worse During Obama's Presidency

By LARRY BUCHANAN, HAEYOUN PARK and ADAM PEARCE  JAN. 15, 2017

Draw your guesses on the charts below to see if you're as smart as you think you are.

Under President Obama, the unemployment rate ...

Draw the line for the Obama years.
Connecting people and data
Making systems playful
Promoting self reflection

Personalizing reading
Reducing cognitive load

5 AFFORDANCES OF THE FORMAT

When you were born, the Melbourne, Florida area could expect about 53 days per year to reach at least 90 degrees.
Connecting people and data
Making systems playful
Promoting self reflection
Personalizing reading
Reducing cognitive load

5 AFFORDANCES OF THE FORMAT

If job growth **had been accelerating**...  
...the jobs report **could look like this:**
Narrative Implementation

Do-It-Yourself
Monitor scroll position & element positions
Guidance: M. Bostock, J. Vallandingham

Third-Party Support
Helper Libraries
Interactive Document Formats: Idyll
idyll

A toolkit for creating data-driven stories and explorable explanations.
Text + Parameterized Graphics

Write articles in markdown.

Embed parameterized data visualizations.

Connect reactive variables to user input widgets to add interactivity.
## Hello World

[var name:"x" value:5 ]

The value of x is [Display value:x format:"d" ].

[Range value:x min:0 max:10 ]
28 Built-in Components

Input
- Action
- Boolean
- Button
- Dynamic
- Select
- Option
- Range
- Select
- Text Input

Presentation
- Chart
- Conditional
- Display
- Equation
- Gist
- Latex
- If
- Inline Link
- SVG
- Table
- Youtube

Layout
- Aside
- Full Width
- Fixed
- Float
- Inline

Helpers
- Analytics
- Meta
- Preload
Bring your own graphics

A JavaScript component API is exposed to end users.

Use with libraries like D3, Mapbox, Processing (P5), React, Vega-Lite
# A Sad Guitar.

Take a second and strum the guitar. It doesn't sound so good, does it?

We've just taken it out of storage and *it's all out of tune...*

# Electric Tuner to the Rescue.

Tune the guitar using the tuner knobs on the right to tighten.  
// Need a reward state to let

[Tuner selectedString: targetString]

[conditional if:isInTune]  
Great work, scroll on.  
[/conditional]

[/**/section]
In comparison, only 12% of YouTube’s traffic went through cellular networks, but they were by far the most expensive part of YouTube’s content delivery pipeline, accounting for approximately **8,500 Gigawatt-hours of electricity usage**—enough to power over 750,000 U.S. homes [3]. At over 10 times the electricity usage per unit of traffic, the relative inefficiency of cellular transmission is clear.
**Texas 2036 COVID Reopening**

The average number of cases per day in Texas has been **decreasing** by 0.64 cases per day over the last 14 days. Based on the data available, the true rate of change may be anywhere from -135.56 to 134.29 new cases per day. There is **not enough** evidence to show a downward trend in new cases per day.

---

**OR a Downward trajectory of positive tests as a percent of total tests within a 14-day period (flat or increasing volume of tests)**
Examples

See full gallery online.
Thanks!

Feel free to reach out with questions: mconlen@cs.washington.edu


@mathisonian
mathisonian.com