Narrative Visualization

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

Ph.D. candidate @ UW computer science
Advised by Jeff Heer

Background in data journalism
FiveThirtyEight, CNN, New Yorker,
New York Times (Elections, Covid-19)

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

@mathisonian
mathisonian.com
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...
**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
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.
While 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 communities. 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.
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: Herd Immunity
   Affordances and challenges of the format

Research highlight
   Authoring Tools
STORYTELLING across various media

PEOPLE TELL STORIES
WORDS TELL STORIES
IMAGES TELL STORIES
COMICS TELL STORIES
MOVIES TELL STORIES
STORYTELLING
across various media

PEOPLE TELL STORIES
WORDS TELL STORIES
IMAGES TELL STORIES
COMICS TELL STORIES
MOVIES TELL STORIES
DATA TELLS STORIES?
Lives of others

Jennifer Stolp, Canada's privacy commissioner, worked with Facebook in August last year to secure changes to its privacy policies. Since then, the company appears to be moving in an unfortunate direction, which involves increasing their data collection and sharing data with third parties.

Facebook is not the only company guilty of this, but Google is also under pressure. This week, it announced a change in its privacy policy, which will allow companies to collect and share data from users.

Privacy laws are changing, and businesses are adapting to keep up. The challenge is to do so without compromising user privacy. Facebook and Google face a delicate balance between growing their businesses and ensuring user privacy.

The changes are likely to have a significant impact on the social media landscape. Companies will need to re-evaluate their privacy policies and ensure they are compliant with the new laws.

For more information on privacy laws and regulations, visit <https://www.dataprotectionoffice.gc.ca>.
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)
SHENANIGANS

Change in teen shenanigans from 1993 to 2013.
Showing % difference from 2003 levels.

- Violent Arrest
- Pregnancy
- Alcohol
- Non-Violent Arrest
- Drugs
- Dropout
- Abortion
- Marijuana

1993  2013

Conlen 2014
Obesity Map Vadim Ogievetsky
Of the 7,666 times that police officers killed people in the U.S. between 2013 and 2019... 25 (0.3%) resulted in a conviction.

74 (1.0%) resulted in a charge but no conviction.

And...

Total disbelief about the system doing the bare minimum. George Floyd deserved so much more than this, he deserved to live.

Some detail about those 25 sentences:
- Unknown sentence = 4 police killings
- Just probation = 3
- 3 months in jail = 1
- 1 year in jail, 3 years suspended = 1
- 1 year in prison = 1
- 18 months in prison = 1
- 2.5 years in prison = 1
- 4 years in prison = 1
- 5 years in prison = 1
- 6 years in prison = 1
- 16 years in prison = 1
- 20 years in prison = 1
- 30 years in prison = 2
Narrative Devices

What devices communicates best?
Highly dependent on: audience, context, format

Format

Interactive Article
Animation
Lecture
Still Image
Video
Narrative Visualization
Design Space
CASE STUDIES

70% Journalism
20% Business
10% Research

Segel & Heer, 2010
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 took lead
- Home runs after 16 seasons: 600
- Bonds: 567
- Aaron: 554
- Ruth: 516

20 seasons
- Bonds was injured last season. He played 14 games and hit 5 homers.

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 home runs slightly outpace projected home runs for five seasons.

**Ruth**
- Averaged 46.4 home runs a season from age 30 to 34.
- Averaged 42.5 for next four seasons.

**Bonds**
- From age 35 to 39, he averaged 14 more home runs 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**
  - 15 times hit 30 or more (M.L. most).
  - Hit only 20 over first five seasons.

- **Babe Ruth**
  - Averaged 40 home runs a season from age 31 to 36.

- **Barry Bonds**
  - Hit only 20 over first five seasons.

- **Willie Mays**
  - Averaged 20 home runs a season from age 31 to 36.

- **Sandy Koufax**
  - Averaged 24 from 1960 to 1964.

- **Frank Robinson**
  - Hit 36 home runs in a season.

- **Mark McGwire**
  - Hit 50 home runs in a season.

- **Ken Griffey Jr.**
  - Hit 40 home runs in a season.

- **Alex Rodriguez**
  - Hit 40 home runs in a season.

- **Albert Pujols**
  - Hit 40 home runs in a season.
Observed Narrative Devices
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

---

**Genres**
- Interactive slideshow
- Drill-down story

**Glass**
- Martini glass
Visual Design

MAGAZINE STYLE
ANNOTATED CHART
SCIENCE FAIR POSTER
FLOWCHART
COMICSTRIP
SLIDESHOW
MOVIE

Duo-Specific Annotations
Captions Attached Article
Summaries Interpret Headlines

Tacit Tutorial Selection
Navigation Filtering Highlighting
Timelines Details on Demand

Messaging Interactivity

MORE
Interactive Articles
When Could the United States Reach Herd Immunity? It’s Complicated.

By Matthew Conlen and Charlie Smart  Feb. 20, 2021

With the vaccine rollout underway and coronavirus cases declining after a dark winter surge, it may seem as though the end of the pandemic is in sight. In reality, how soon could we get there?

One answer lies in herd immunity, the point when enough people are immune to the virus that it can no longer spread through the population. Getting there, however, depends not just on how quickly we can vaccinate but on other factors, too, like how many people have already been infected and how easily the virus spreads.

An estimate for the path to herd immunity

100% of population immune

80%

Herd immunity range
Motivation

- More engaged audience
  Greussing & Boomgaarden, *Digital Journalism* 2019

- Better learning outcomes
  Mayer’s *Multimedia Principles*, 2005

- Promote active reading
  Bret Victor’s “Explorable Explanations”, 2011

- Transparency in journalism
  “A sense of the tentativeness of truth.”
  - Philip Meyer, *Precision Journalism*, 2002
Challenges

- Integration between text and media
- Data-driven
- Computation
- Interactivity
- Multiple output targets (desktop, mobile, homepage, print, social)

- Many people working together
  - two authors, multiple editors, designers, researchers, interviewees
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

Tied together research from HCI, multimedia learning, infovis, digital journalism

Hohman, Conlen, Heer, Chau, Distill, 2020
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

Areas of Application
Research Dissemination
Journalism
Education
Policy & Decision Making
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

**Five affordances of the format**

1. **Connecting People and Data.**
   Make data pleasant to work with. Happy readers are engaged readers.

2. **Making Systems Playful.**
   Run interactive simulations directly in the browser. No setup required.

3. **Prompting Self-Reflection.**
   Help readers learn by asking them to reflect in a low pressure environment.

4. **Personalizing Reading.**
   Let readers choose the content that is relevant to their own experience.

5. **Reducing Cognitive Load.**
   Use effective representations to make complex topics more intuitive.
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.
Connecting people and data

Making systems playful
Promoting self reflection
Personalizing reading
Reducing cognitive load

5 AFFORDANCES OF THE FORMAT

Making Systems Playful
Run interactive simulations directly in the browser. No setup required.

There's a lot you can learn just from playing around with this editor. For example, can you figure out the order the pixels are stored in?

Something strange in the example above is that changing some numbers doesn't seem to impact the image at all, while setting the 17 on line one to 0 completely ruins the image! Other actions, like setting the 7 on line 1988 to 254 change the color, but only for subsequent pixels.
Connecting people and data
Making systems playful
**Promoting self reflection**
Personalizing reading
Reducing cognitive load

---

**5 AFFORDANCES OF THE FORMAT**

You Draw It: What Got Better or Worse During Obama’s Presidency

By Larry Buchanan, Hartoun 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**…

![Graph showing unemployment rate during Bush and Obama years]

Draw the line for the Obama years.

Show me how I did.
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

Explorable explanation

From Wikipedia, the free encyclopedia.

An explorable explanation (often shortened to explorable) is a form of informative media where an interactive computer simulation of a given concept is presented, along with some form of guidance on how to use the simulation. Explorable explanations encourage users to discover potential emergentbehaviour against its actual behaviour, promoting a more active learning experience.

Definition [edit]

The term 'explorable explanation' was first used in passing by [citation needed], where it included an explorable explanation of a digital model with a deliberate manipulation of the attention of their audience towards particular aspects of the model. Explorable Explanations is my umbrella project for ideas that enable and encourage truly active reading. The goal is to change people's relationship with text. People currently think of text as information to be consumed. I want text to be used as an environment to think in.

Some of the ideas Victor espoused in the essay occurred to him while working with Al Gore on the app version of the 2009 book Our Choice. He had proposed that the book should contain interactive models, but this idea was rejected on the basis that all numerical values proposed regarding climate change would need to have a citation, and the interactive models would generate unclad numbers.

The term has since also been characterized as being about learning through play. The related term “enthusiasm” was used by Alan Kay to refer to text-based explorable explanations.
Connecting people and data
Making systems playful
Promoting self reflection
Personalizing reading

Reducing cognitive load

5 AFFORDANCES OF THE FORMAT

Internal Combustion Engine

The invention of the internal combustion engine in the 19th century has revolutionized transportation over land, water, and air. Despite their omnipresence in modern day, the operation of an engine may be cryptic. Over the course of this article I’d like to explain the functionality of all the basic engine parts shown in the demonstration below. You can drag it around to see it from other angles:
Connecting people and data
Making systems playful
Promoting self reflection
Personalizing reading

Reducing cognitive load
**Challenges**

- Turned a writing task into a web development task
- Breadth of skills required
- Multiple output targets
- Accessibility
- Misaligned Incentives
- Hard to evaluate success
**CHALLENGES**

- Turned a writing task into a web development task
- Breadth of skills required
- Multiple output targets
- Accessibility
- Misaligned Incentives
- Hard to evaluate success

**OPPORTUNITIES**

<table>
<thead>
<tr>
<th></th>
<th>Research</th>
<th>Practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authoring</td>
<td>Next generation tooling</td>
<td>Evaluate in production setting, identify bugs</td>
</tr>
<tr>
<td>Designing</td>
<td>Developing theory, conducting laboratory studies</td>
<td>Evaluate specific design decisions in the wild, understand constraints</td>
</tr>
<tr>
<td>Publishing</td>
<td>Tools, guidelines, and best practices</td>
<td>Concrete examples for others to follow, available source code, accessible archives, DOI, branding</td>
</tr>
</tbody>
</table>
Narrative Implementation

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

Third-Party Support
Helper Libraries
Interactive Document Formats: Idyll
A toolkit for creating data-driven stories and explorable explanations.
Text + Parameterized Visuals

Write articles in *markdown*.

Connect *reactive variables* to user input widgets to add interactivity.
Text + Parameterized Visuals

Write articles in **markdown**.

Connect *reactive variables* to user input widgets to add interactivity.

---

**INPUT (EDITABLE)**

```markdown
# Hello World
[var name:"x" value:5 ]

The value of x is [Display value:x format:"d "].
[Range value:x min:0 max:10 ]
```

---

**Hello World**

The value of x is 5.
A Model of Interactive Articles

Idyll Markup

[Header
title: "Hi Idyll!" ]
This is the text of my article.

[Graphic
value: x ]
And that was a custom graphic.

AST

Component
type: Header
properties:
title: "Hello Idyll"

Text
content: "This is the text of my article."

Component
type: Graphic
properties: value: x

Text
content: "And that was a custom graphic."

Live Web Page

Hello Idyll
April 22, 2021

This is the text of my article

And that was a custom graphic
28 Built-in Components

Input

Presentation

Layout

Helpers
Bring your own graphics

A JavaScript component API is exposed to end users.

Use with libraries like D3, Vega-Lite, Mapbox, Processing (P5), React, ...
Le long de la route BR-163

UN AXE DE LA DÉFORESTATION EN AMAZONIE BRÉSILIENNE

Thomas Ansart, Antoine Rio (Sciences Po - Atelier de cartographie)
23 décembre 2019
There’s a lot you can learn just from playing around with this editor. For example, can you figure out the order the pixels are stored in?

Something strange in the example above is that changing some numbers doesn’t seem to impact the image at all, while setting the $17$ on line one to $0$ completely ruins the image! Other actions, like setting the $7$ on line $1988$ to $254$ change the color, but only for subsequent pixels.
Examples

idyll-lang.org/gallery
Usage

Open-source release
Pilot Studies
Usage in the wild

UK Infrastructure Transitions Research Consortium
Minneapolis Star Tribune
Vis courses in universities (UMD, UIUC, UW, CMU, …)
Teaching materials (CU Denver, UBC)
IEEE VisXAI
Texas2036
Recurse Center
Parametric Press - featured in FastCo Design
…and more!
Idvll...

Seems great but...

...still requires learning and writing markup syntax

...requires use of general-purpose programming tools, e.g. to make a new post, run development server
Idyll Studio

Built a **structured editor** for editing and creating Idyll programs

Reduce use of general-purpose programming tools.

Eliminate syntax errors.

Reify Idyll model.

Conlen, Vo, Tan, Heer, *UIST 2021*
Idyll Studio Demo

Introducing UIST to the structured editor.

Editing text

Text can be $y = mx + b$.

Text is written in a markup language based on something called Markdown. Markdown is a syntax that is meant to make it easy to add common styles to your text.

For example, you can use this syntax to create `**bold**` (\*\*\*bold\*\*) and `*italic*` (\*italic\*) styles. Or use three backticks to create code blocks:

```markdown
* lists
  * of
  * items,
```
Thanks!

Feel free to reach out:

mconlen@uw.edu

**FIGURE 1:** "Global Temperature." NASA. NASA Global Climate Change, 2020.