Narrative Visualization



Matthew Conlen

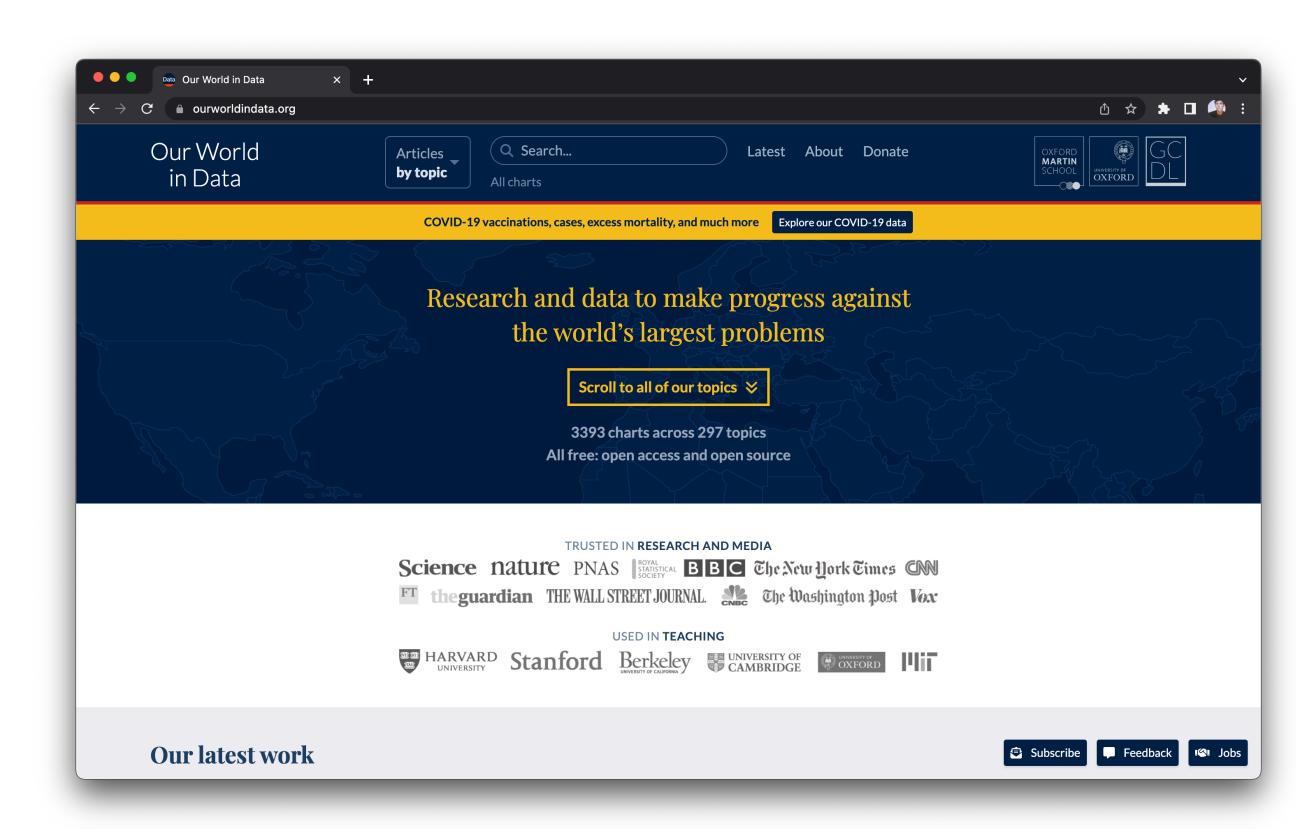
(with material from Jeff Heer, Edward Segel)

About Me

I am an HCI researcher and data journalist

- Ph.D. from UW in 2021
- Work at NYTimes, FiveThirtyEight, CNN, New Yorker

Currently Head of Product & Design at Our World in Data, contributor to The Upshot at NYT



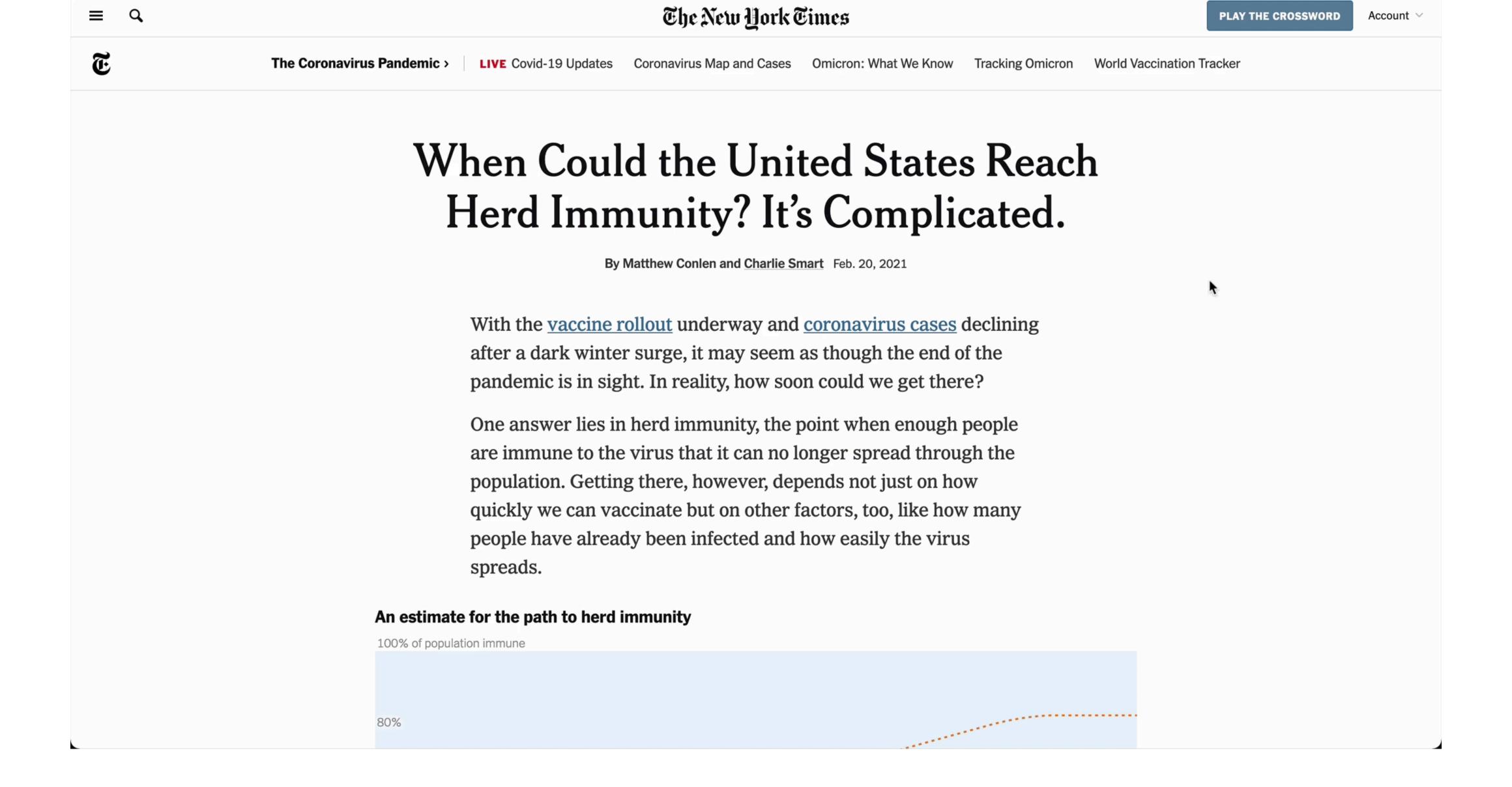
@mathisonian mathisonian.com

About Me

My CS research is inspired by working as a journalist

- Tools for interactive and data-driven storytelling
- Interactive research papers and educational materials
- Better decision making tools for policymakers
- Transparent and data-driven journalism on a deadline

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



When Could the United States Reach Herd Immunity? It's Complicated. Conlen & Smart, New York Times 2021

Motivation

Transparency in journalism

"A sense of the tentativeness of truth."

- Philip Meyer, Precision Journalism, 2002

More engaged audience

Greussing & Boomgaarden, Digital Journalism 2019

Better learning outcomes

Mayer's Multimedia Principles, 2005

Promote active reading

Victor's "Explorable Explanations", 2011

...also many challenges!



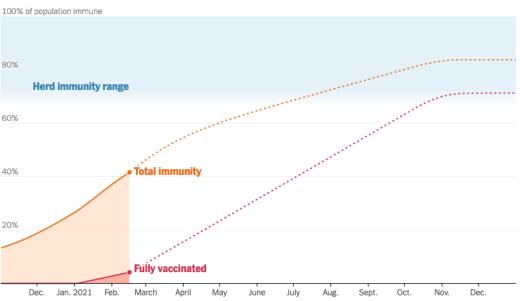
When Could the United States Reach Herd Immunity? It's Complicated.

By Matthew Conlen and Charlie Smart Feb. 20, 2021

With the <u>vaccine rollout</u> underway and <u>coronavirus cases</u> 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



This chart shows the current path to herd immunity in the United States, based on a model developed by PHICOR, a public health research group. It looks at the number of people who have been fully vaccinated and combines that with an estimate of the number of people who have been infected and have recovered to measure

When the orange line crosses into the blue area, that means we have entered the **herd immunity range**. The exact threshold for herd immunity for the coronavirus is unknown, but recent estimates range from 70 percent to 90 percent.

At first, this looks like pretty good news — under these assumptions, we could reach herd immunity as early as July. But a lot could happen between now and then. The speed and uptake of vaccination, and how long immunity lasts are big factors. The rise of new virus variants and how we respond to them will also affect the path to herd immunity.

In most scenarios, millions more people will become infected and tens or hundreds of thousands more will die before herd immunity is reached.

What if we speed up vaccinations?

More than 15 million people have been fully vaccinated, and the U.S. is currently administering about 1.7 million shots per day. Some experts say we could nearly <u>double that pace</u> by April as new vaccines are approved. (Because the current vaccines require two doses spaced weeks apart, the number of people fully vaccinated each day is smaller.)

The more people we vaccinate, the faster we could reach the threshold for herd immunity.

Topics

Storytelling

- How different mediums can be used to tell stories

Design of narrative visualizations & interactive articles

- Segel & Heer, 2010
- Hohman, et al 2020

Research on Authoring Tools

- Idyll, *UIST 2018*
- Idyll Studio, UIST 2021

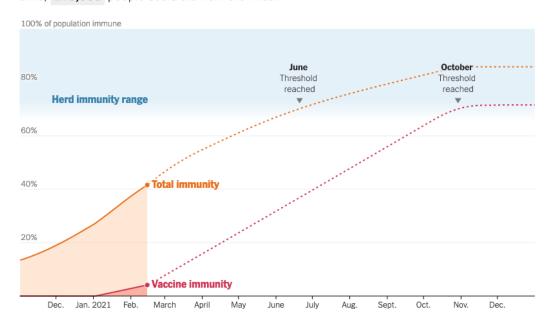
Design exercise

While this means we would reach herd immunity sooner because more people would get sick and develop antibodies, that speed would come at a grim cost.

CHOOSE A SCENARIO S

Keep measures in place	Ease up in the spring	End precautions now
Continue social distancing until most people are vaccinated.	Lift restrictions when 15 percent of people are vaccinated.	Lift restrictions when few people are vaccinated.

If we end restrictions in April, we could reach the herd immunity threshold by **June**. But in that time, **170,000** people could die from the virus.



Precautions remain especially important as new variants of the virus emerge. If social distancing measures aren't followed, a stronger virus could rapidly infect and kill hundreds of thousands of people before they can be vaccinated.

What if a more contagious variant spreads?

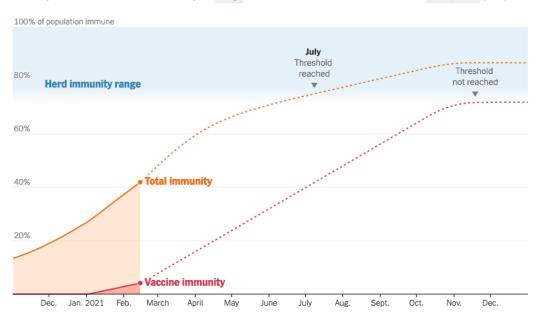
A more contagious variant of the virus, like the one first identified in Britain that is now <u>spreading</u> throughout the United States, could further complicate the path to herd immunity.

If the virus becomes more contagious, the threshold for herd immunity will go up. It may be hard for vaccines to keep pace, and precautions will be even more necessary to stem the spread.

CHOOSE A SCENARIO

Current variant with precautions Transmissibility and precautions remain the same.	More contagious variant with precautions Transmissibility increases but preventative measures continue.	More contagious variant with no precautions Preventative measures end in March and transmissibility increases.
---	---	---

A more contagious variant will spread faster and raise the herd immunity threshold. We could still expect to reach herd immunity in **July**, but new deaths could double, to **200,000** people.

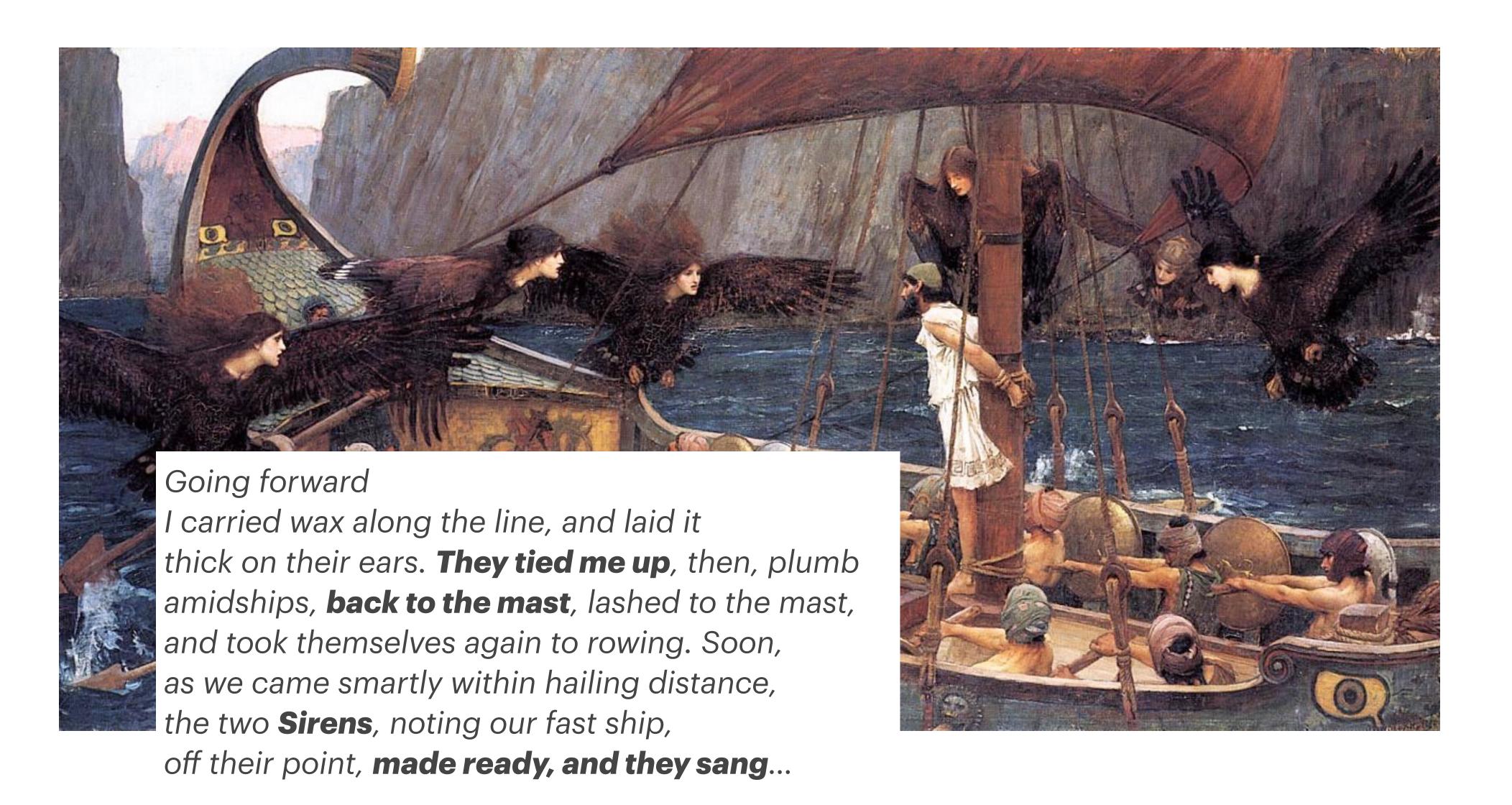


The spread of new virus variants makes it impossible to put a firm date on when we'll reach herd immunity or when the pandemic will end. There's a chance a mutation could lead to a version of the virus that doesn't respond to existing immunity, leading us to start the journey to herd immunity all over again.

Dr. Lee said that coronaviruses have relatively high mutation rates and that it is likely that new variants of the virus will continue to emerge. "The question will be how different might these variants be," he said. If a variant of the virus stops responding to the

STORYTELLING

as ancient as mankind



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











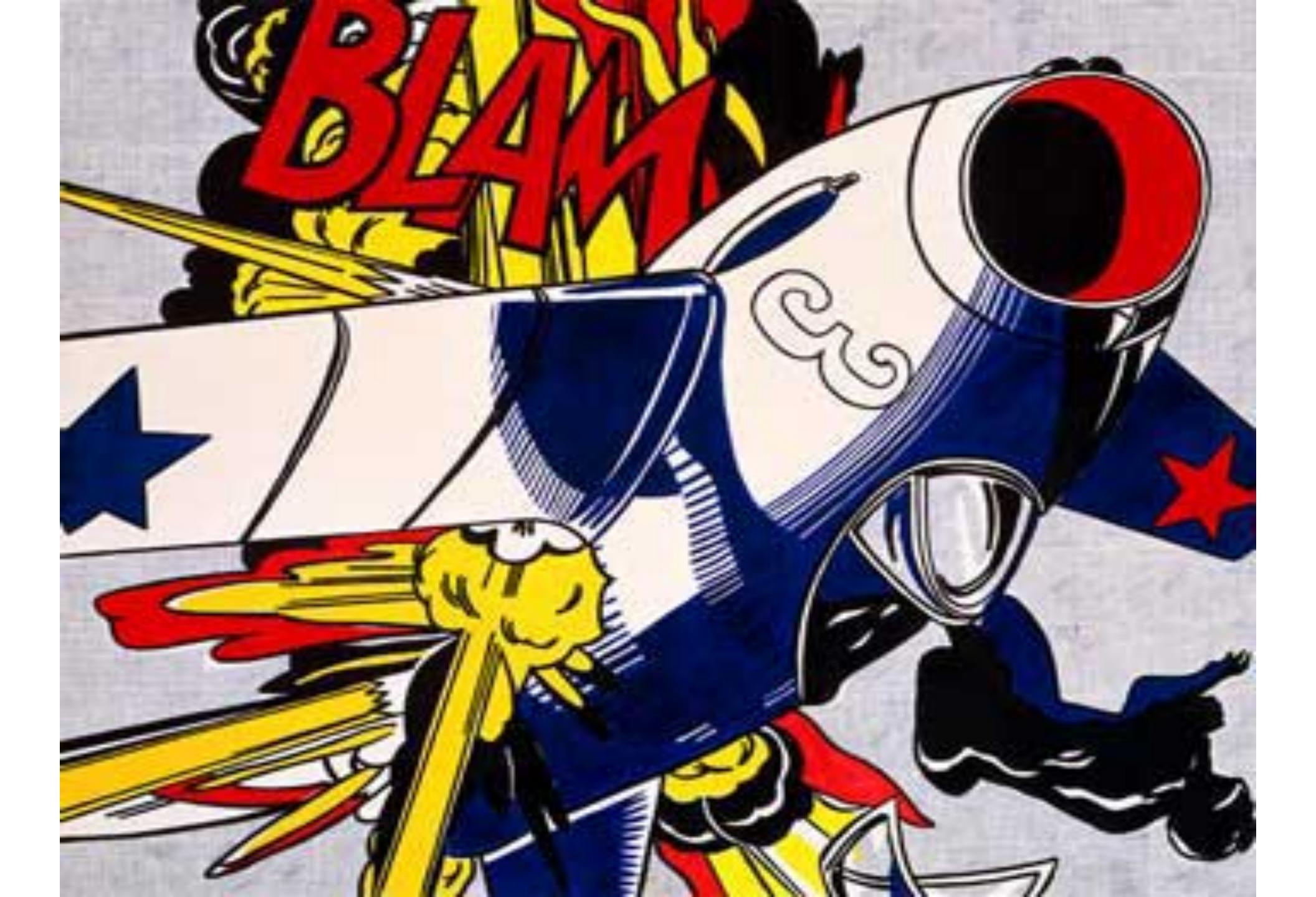












Justice Stephen G. Breyer to Retire From Supreme Court



Justice Stephen G. Breyer was appointed to the Supreme Court in 1994 by President

WASHINGTON — Justice Stephen G. Breyer, the senior member of the Supreme Court's three-member liberal wing, will retire, two people familiar with the decision said, providing President Biden a chance to make good on his campaign pledge to name a Black woman to the court.

Mr. Biden is expected to formally announce the retirement at the White House on Thursday, according to one person familiar with the planning for the event.

Justice Breyer, 83, the oldest member of the court, was appointed in 1994 by President Bill Clinton. After the death of Justice Ruth Bader Ginsburg in 2020 and the appointment of Justice Amy Coney Barrett by President Donald J. Trump, he became the subject of an energetic campaign by liberals who wanted him to step down to ensure that Mr. Biden could name his successor while Democrats control the Senate.

Justice Stephen G. Breyer to Retire From Supreme Court



Justice Stephen G. Breyer was appointed to the Supreme Court in 1994 by President

WASHINGTON — Justice Stephen G. Breyer, the senior member of the Supreme Court's three-member liberal wing, will retire, two people familiar with the decision said, providing President Biden a chance to make good on his campaign pledge to name a Black woman to the court.

Mr. Biden is expected to formally announce the retirement at the White House on Thursday, according to one person familiar with the planning for the event.

Justice Breyer, 83, the oldest member of the court, was appointed in 1994 by President Bill Clinton. After the death of Justice Ruth

Bader Ginsburg in 2020 and teaching arrange for the Barrett by President Donald J. Trump, he became the subject of an energetic campaign by liberals who wanted him to step down to ensure that Mr. Biden could name his successor while Democrats control the Senate.

Justice Stephen G. Breyer to Retire From Supreme Court



Justice Stephen G. Breyer was appointed to the Supreme Court in 1994 by President

nutshell paragraphol the Senate.

WASHINGTON — Justice Stephen G. Breyer, the senior member of the Supreme Court's three-member liberal wing, will retire, two people familiar with the decision said, providing President Biden a chance to make good on his campaign pledge to name a Black woman to the court.

Mr. Biden is expected to formally announce the retirement at the White House on Thursday, according to one person familiar with the planning for the event.

Justice Breyer, 83, the oldest member of the court, was appointed in 1994 by President Bill Clinton. After the death of Justice Ruth

Bader Ginsburg in 2020 and the appointment of Justice Amy Coney

Barrett by President Donald J. Trump, he became the subject of an energetic campaign by liberals who wanted him to step down to ensure that Mr. Biden could name his successor while Democrats

STORYTELLING

across various media

PEOPLE TELL STORIES

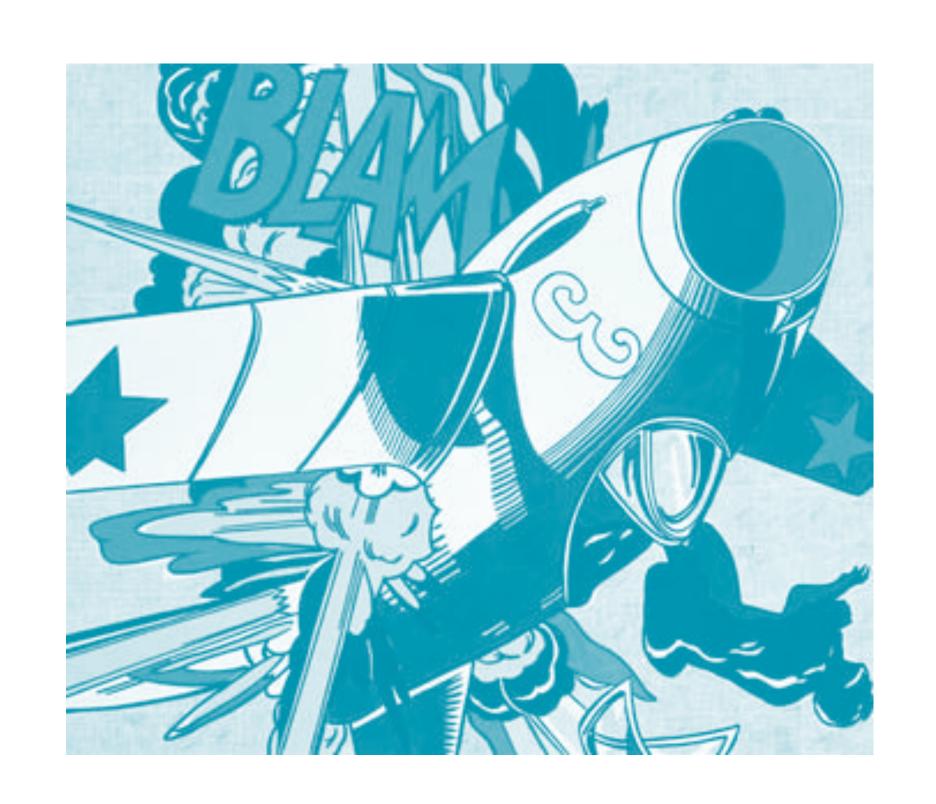
WORDS TELL STORIES

IMAGES TELL STORIES

COMICS TELL STORIES

MOVIES TELL STORIES

DATA TELLS STORIES?





Principled Thy interest.

Lives of others

COURTS OF A MEDICAL PROPERTY.

Throshook and Google face a hocklash, from a were and regulators with a respective year. they have hundled sensitive date

TENNINGS STOCODART, Canada's privacy comment our whoten with facilities. in Abgust acres the secral appreciang the strack a deal, agreeing to charge its politics within a year in comply with the country's privacy line is not used the Seedfact, the correctory appears to be lened up int or import of portel for dual, which ineralized graving several their and every-to-imphrantiplicity consults the triplian price are daru with third parties, "It doesn't stem to our that translands is going in the right fiection on this uses." The says, burning that, without a thorage of course, the farm could soon become the value of all mother foorwise corporer by her organization.

Hacebrook is may the early in turnet giant to encycle the fee of their sustefulous floor. all and and wifiting arthurn this week. following revove that 9 Intil reported stems personal community blance call over unsedependent investigation late Google's butoreiner, claiming that it had reimply discharged courtail railes in the development and asage of vertice in."

The tases highlight rising tention between grantiens of privacy and internet Bress, And Buy reflect toncers arrangweb uses about how process that are

typ in with as Cony Doctorous a session for the rather, and has lapped, a productor. have at independ friethook. Siles such as Qualify or look Disposors are unging offered to proper remarks May you be a made from book "wants".

This is unlikely to dop the meret runise. of Suchook, which is police to chain full. w hillion report on and which drows even more insiture as a whole to its old foreathant. For nerves, have been ranked at the comparts basisquarters in silicon yelley. robust books an orolling may how to respond Several unior tolk are now hinting that Prochook will your soll out simpler. princip commit to make it expect to keep energy data frightn, MySpopu, a gival, is already making to controls along for teraminifort to was distillested terchooken to its HELTY NO.

A myoli over taechook's handling at pared With their natworks in bores and - privacy has hear herwing for sains time. offices he some no constrain. On May aris - to Demember the ownel national changed Pater School, Germany's furtional common. The district settings on transferors common. dense for data protection, called for an or - so that coditionals' personal information. would be shared with "averyone" rather than selected friends Tsasbook argued this reflected a shift in socially streams. months comment and rained that times model will adjust privacy services back agric. But femented privacy activities lobbiedfort to be owned.

The switch should not have come to a mode public Served processest internet surprise tarly on meny regal networks.

Also in this section

- 68 Suring conspanies
- #9 The recovery of beneval Motors

The Connectat No. word with \$7.

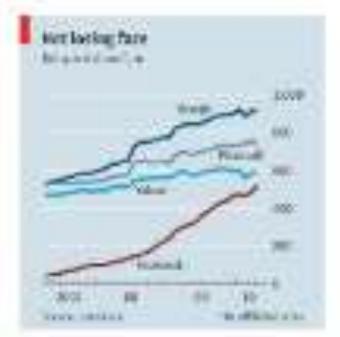
- 86 Business crime to Urine
- 76 Indian retail.
- 70. Japan's drug firms on the copys
- 74 Brazilies telecome
- 71 Schumpeter: Deinstrebched

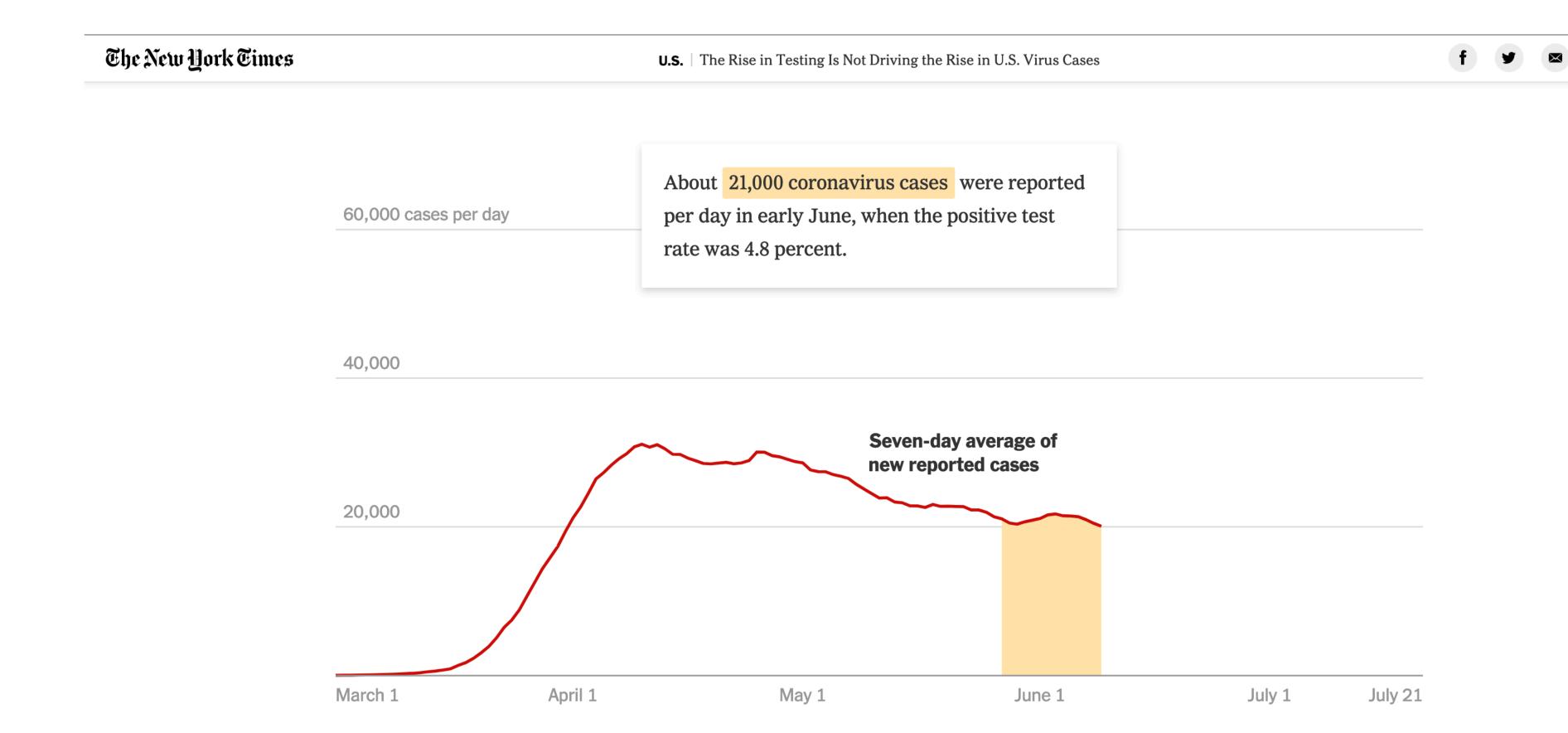
he dela anaboto end dirboto colombiamo and Consent M. com, St. clinical Section

Legions built to ally authory policies in an tier to attach and increase tests. But an more join, connot are gradually icosened. tis encourage came during his people share from Pacifical can increase the trafthe against which it selfs advertising. And the more it from about users from and eighbor, the better recontanget edution geneurobsenhads of millions of dollars.

Protests great lauren still following a caveleges, porcurence last mostly at teletics Mark Zaglerburg, Ficelantics been. announced yet crueker series of policy diarroges. Once that totaled introduces were as: "Yesteri pseuirallector" fustore that letnorsety third-party technics across Pacybook data when people visit. Cities say that Saabuok bas mada i sticky to deable this feature, which raity explain why his Sinddam cidifines it so much.

Suropean officials are grantfoling about racebook test. This morth a group of dataprotection experts who advise the handpican Commission whose to the social new words, codling its discovered to become the disfact ortings 'unacceptable', and in the in-





CASE STUDY COVID-19 TESTING Conlen, 2020

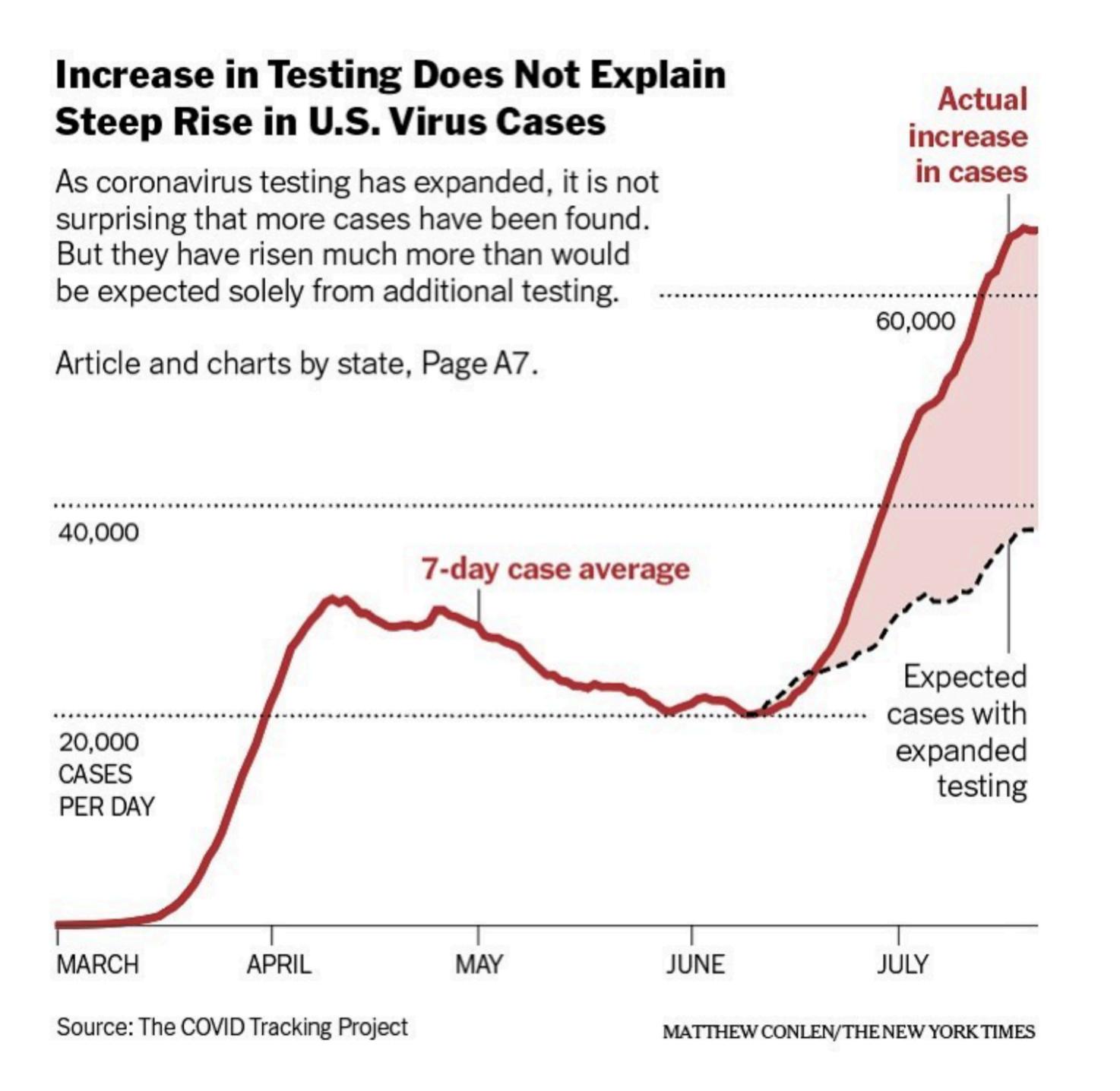
CASE STUDY COVID-19 TESTING

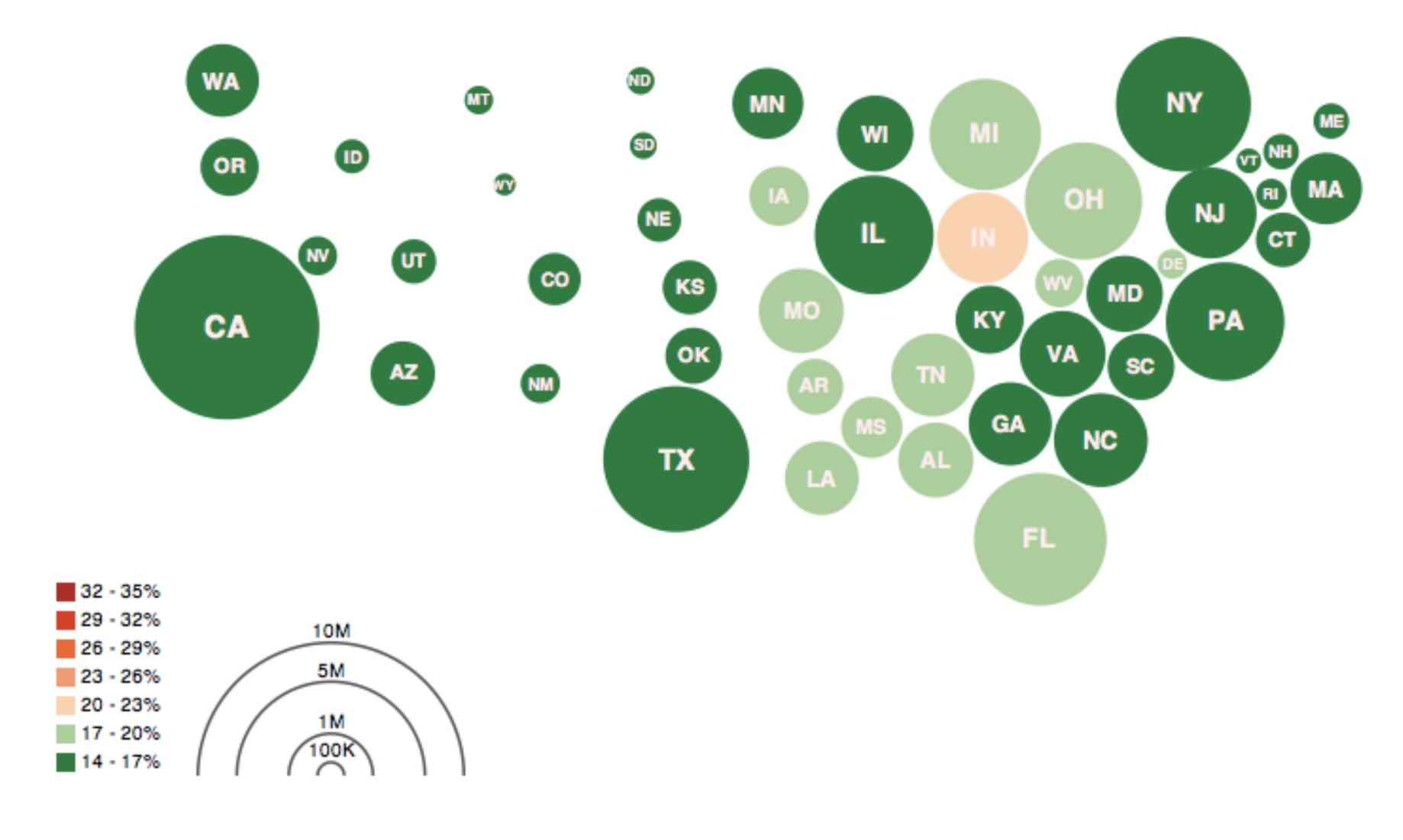
Input Data

Daily case and test counts.

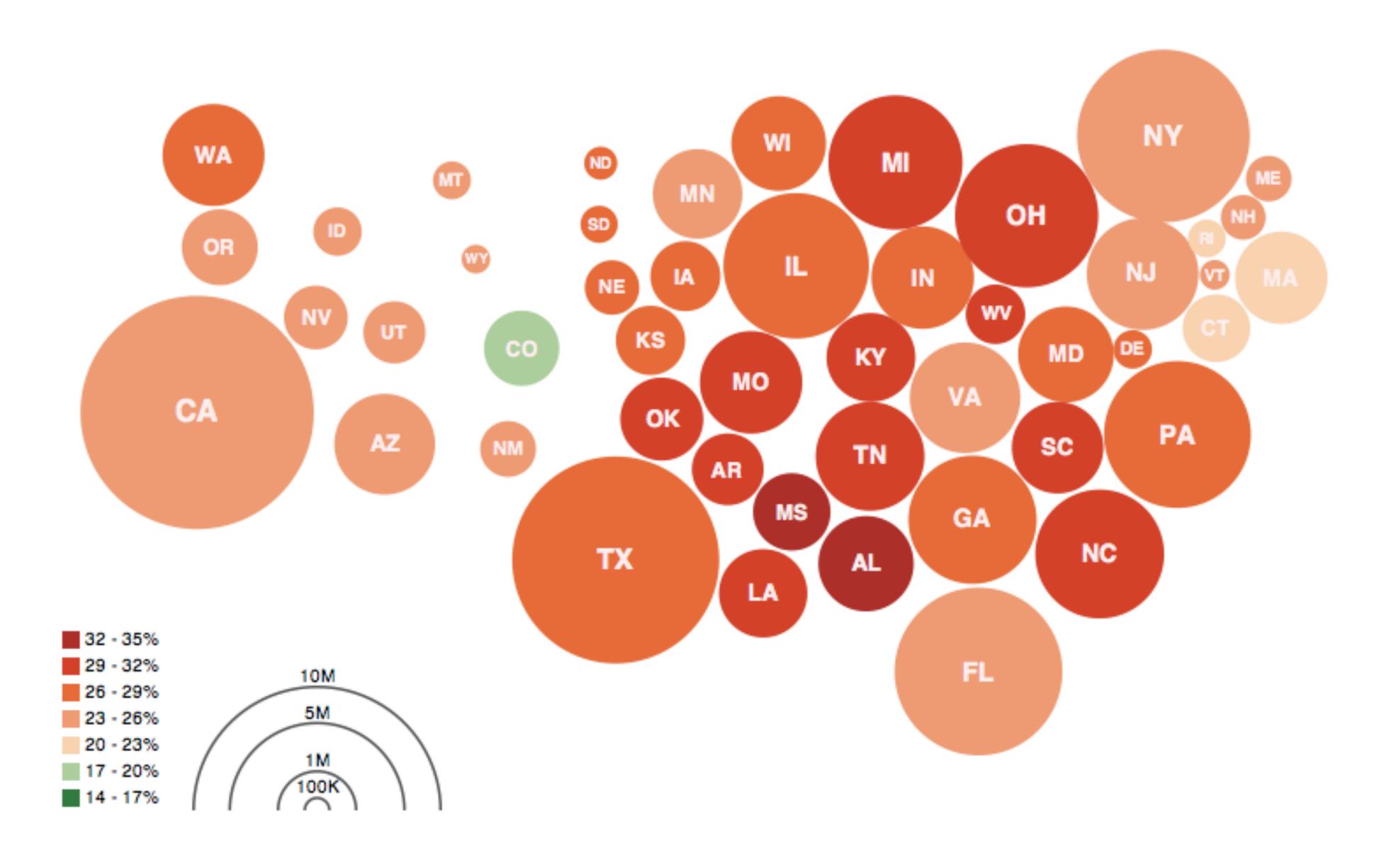
Points of Focus

- 1. Baseline (smoothed average of daily cases)
- Counterfactual case curve (baseline positivity rate * daily test counts),
- 3. Actual case curve (daily cases)





Obesity Map Vadim Ogievetsky 1995



Obesity Map Vadim Ogievetsky 2008

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





monachalabi 🌣 • Follow

•••

United States



monachalabi 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





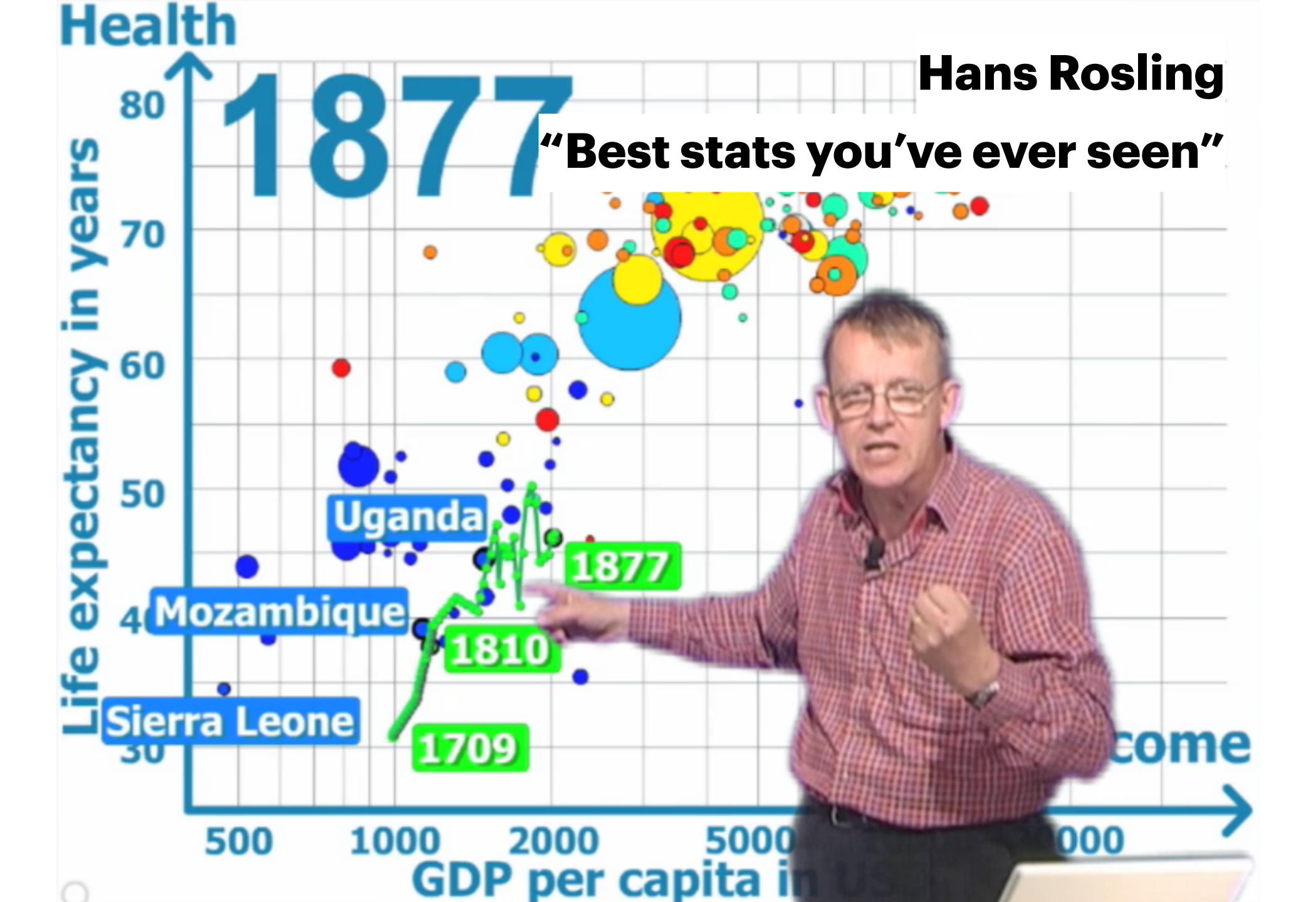




22,976 likes

3 HOURS AGO

Log in to like or comment.



Narrative Devices

What devices communicates best?

Highly dependent on: audience, context, format

Format

Interactive Article

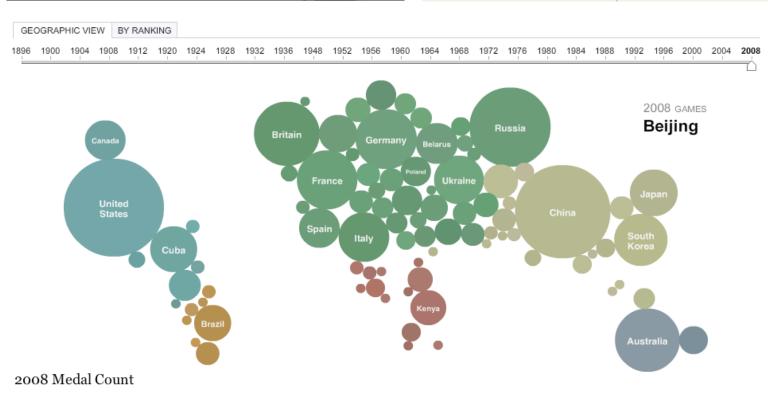
Animation

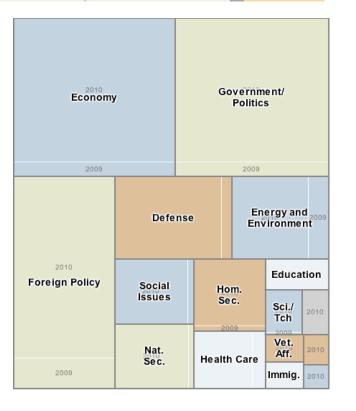
Lecture

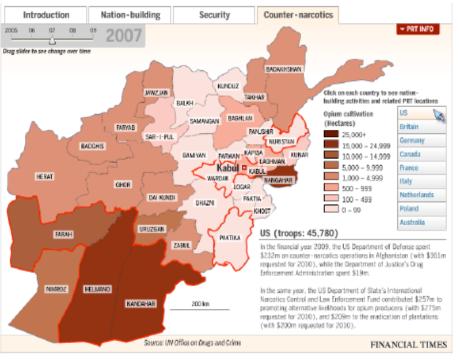
Still Image

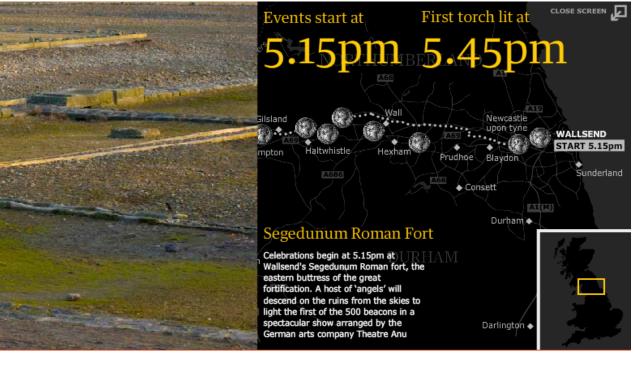
Video

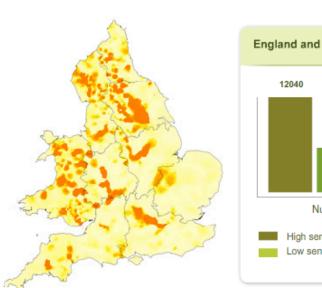
rettes | Coling | Co

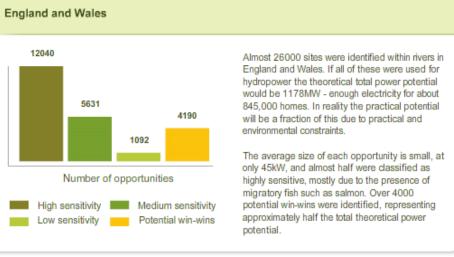










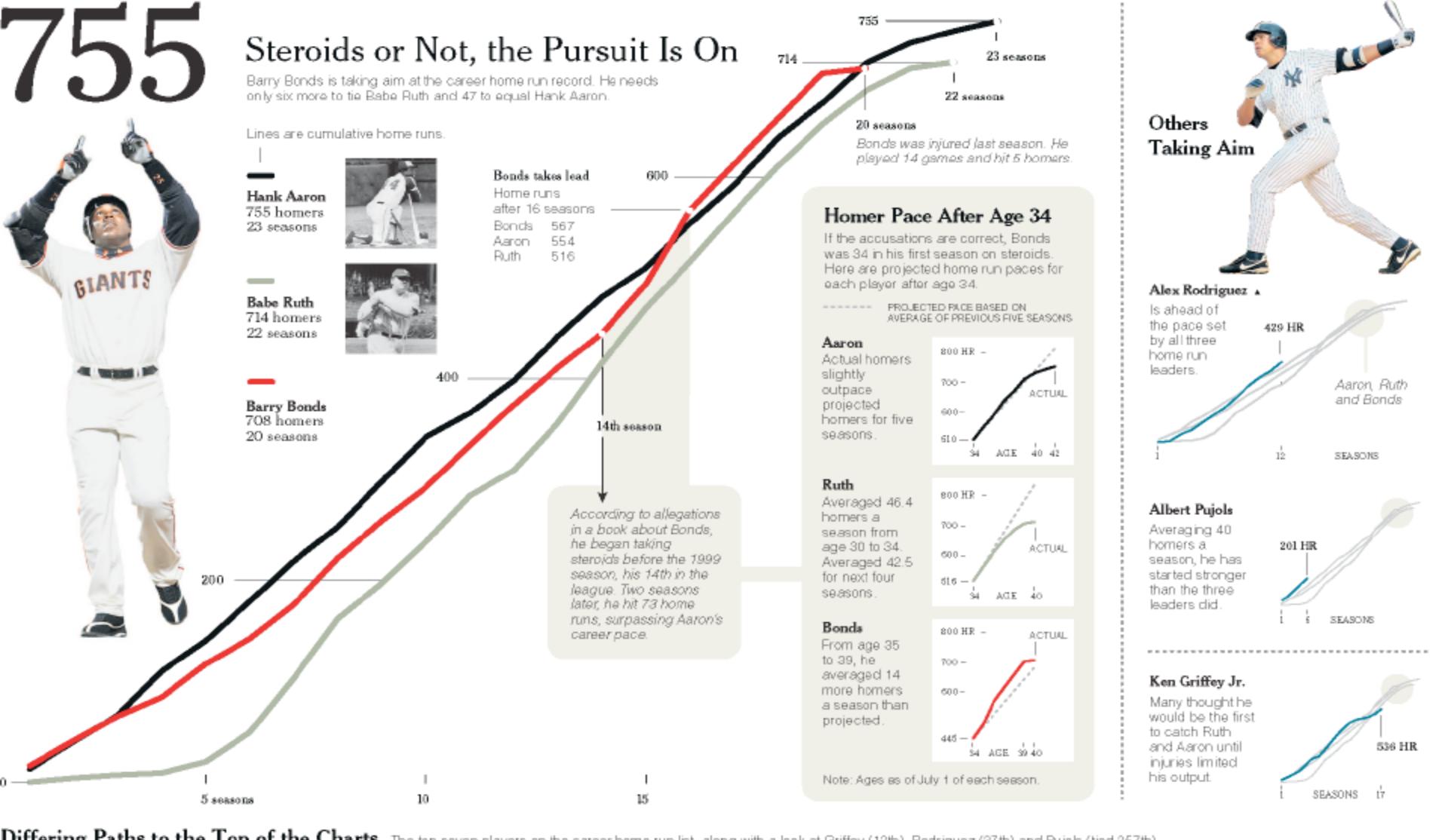




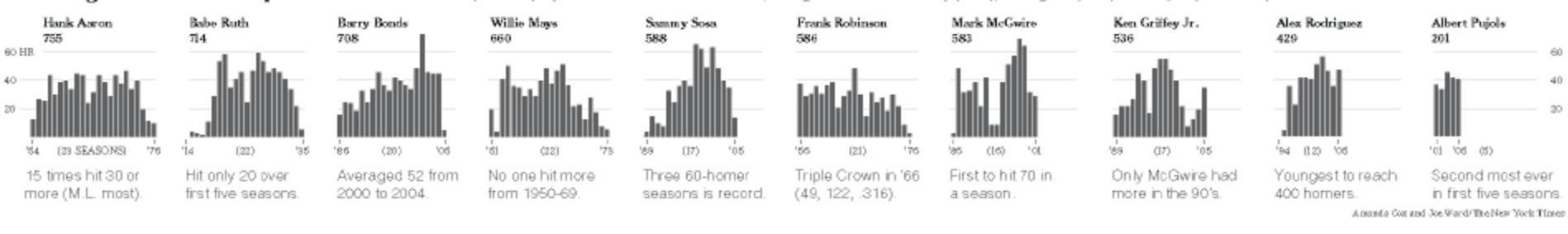
58 CASE STUDIES

70% Journalism20% Business10% Research

Segel & Heer, 2010



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).





(49, 122, .316).

a season.

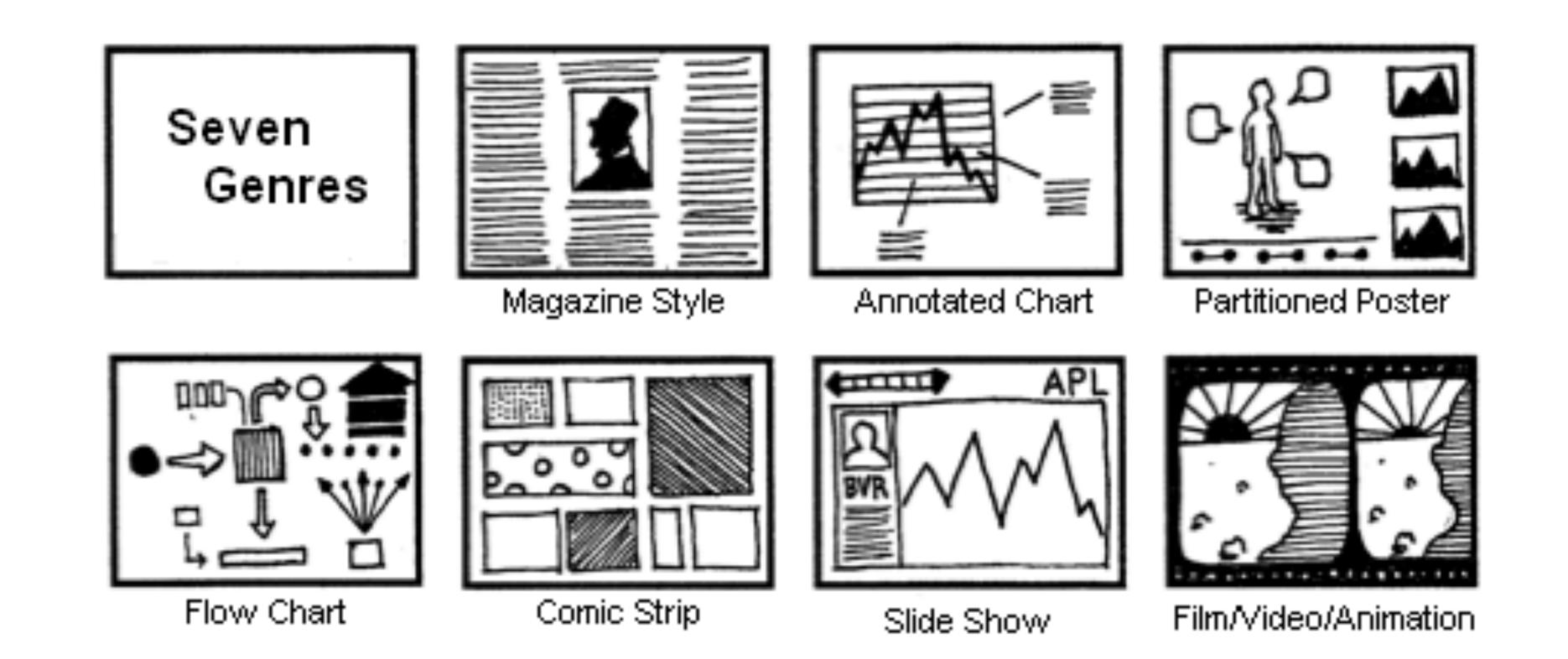
2000 to 2004.

more (M.L. most).

in first five season

400 homers.

more in the 90's.



Genres for Narrative Visualization (2010)

Genres + Interactivity + Messaging =

DESIGN SPACE

STORYTELLING

CLARITY

SPEED

Author Driven

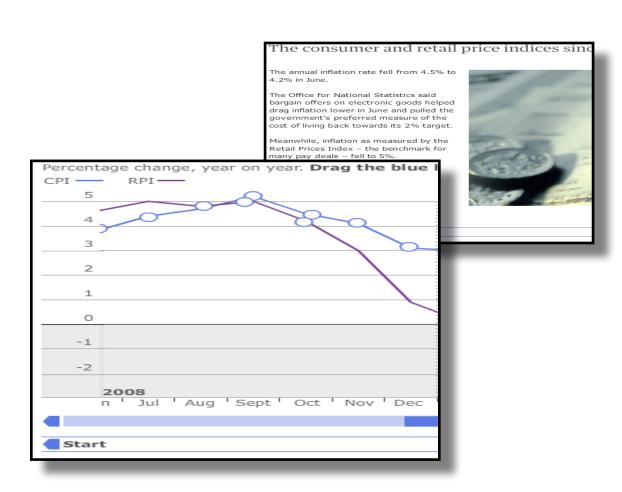
strong ordering heavy messaging limited interactivity **Reader Driven**

weak ordering light messaging free interactivity

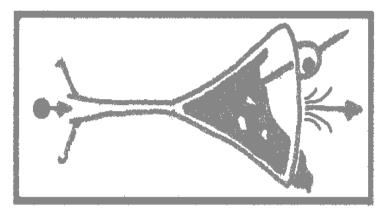
ASK QUESTIONS

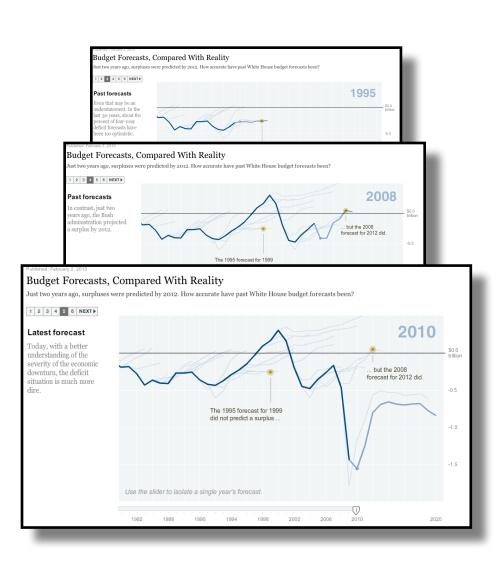
EXPLORE

FIND

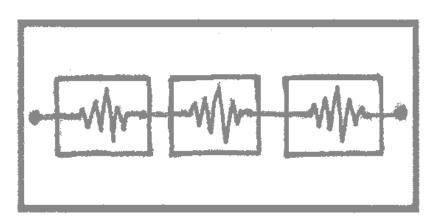


martini glass



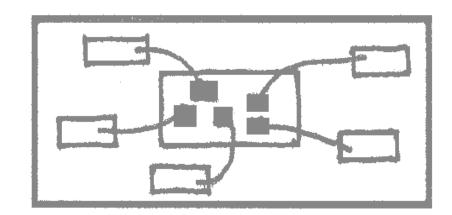


interactive slideshow





drill-down story

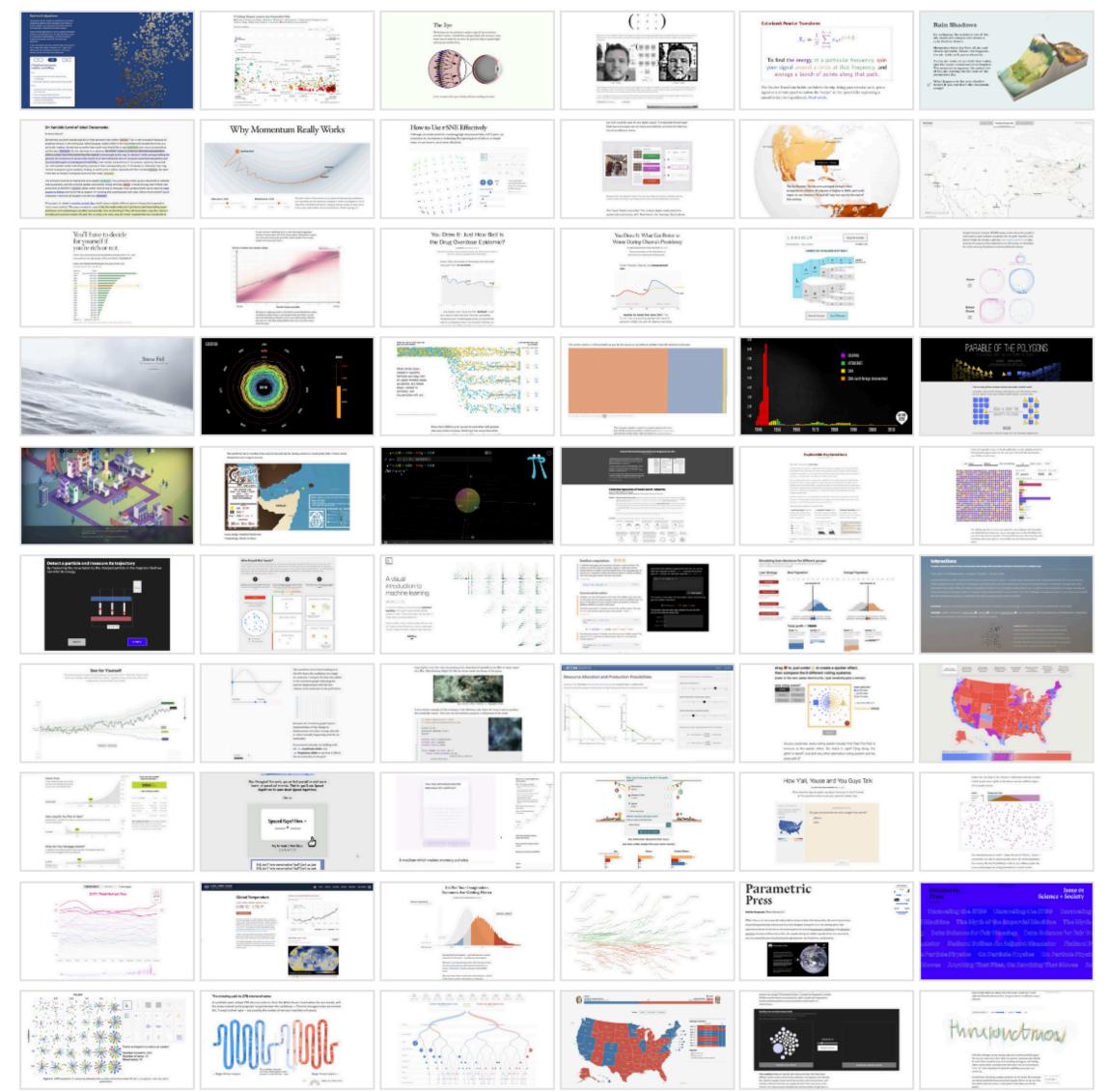


Communicating with Interactive Articles

60 Interactive Articles

NYTimes, WaPo, Distill, VisXAI, ...

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



Hohman, Conlen, Heer, Chau, Distill, 2020

Five affordances of the format

Connecting People and Data.

Make data
pleasant to work
with. Happy
readers are
engaged readers.

Making Systems Playful.

Run interactive simulations directly in the browser. No setup required.

Reducing Cognitive Load.

Use effective representations to make complex topics more intuitive.

Personalizing Reading.

Let readers choose the content that is relevant to their own experience.

Prompting Self-Reflection.

Help readers learn by asking them to reflect in a low pressure environment.

Five affordances of the format

Connecting People and Data.

Make data pleasant to work with. Happy readers are engaged readers.

Making Systems Playful.

Run interactive simulations directly in the browser. No setup required.

Reducing Cognitive Load.

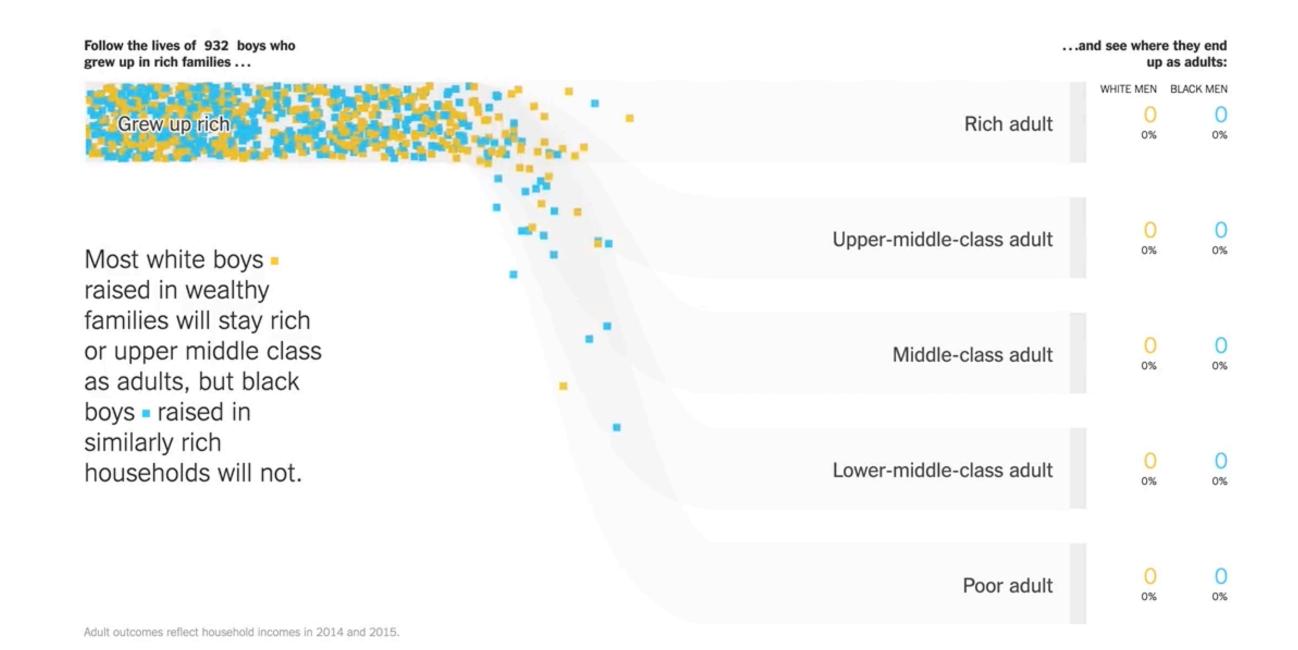
Use effective representations to make complex topics more intuitive.

Personalizing Reading.

Let readers choose the content that is relevant to their own experience.

Prompting Self-Reflection.

Help readers learn by asking them to reflect in a low pressure environment.



Even when children grow up next to each other with parents who earn similar incomes, black boys fare worse than white

Five affordances of the format

Connecting People and Data.

Make data pleasant to work with. Happy readers are engaged readers.

Making Systems Playful.

Run interactive simulations directly in the browser. No setup required.

Reducing Cognitive Load.

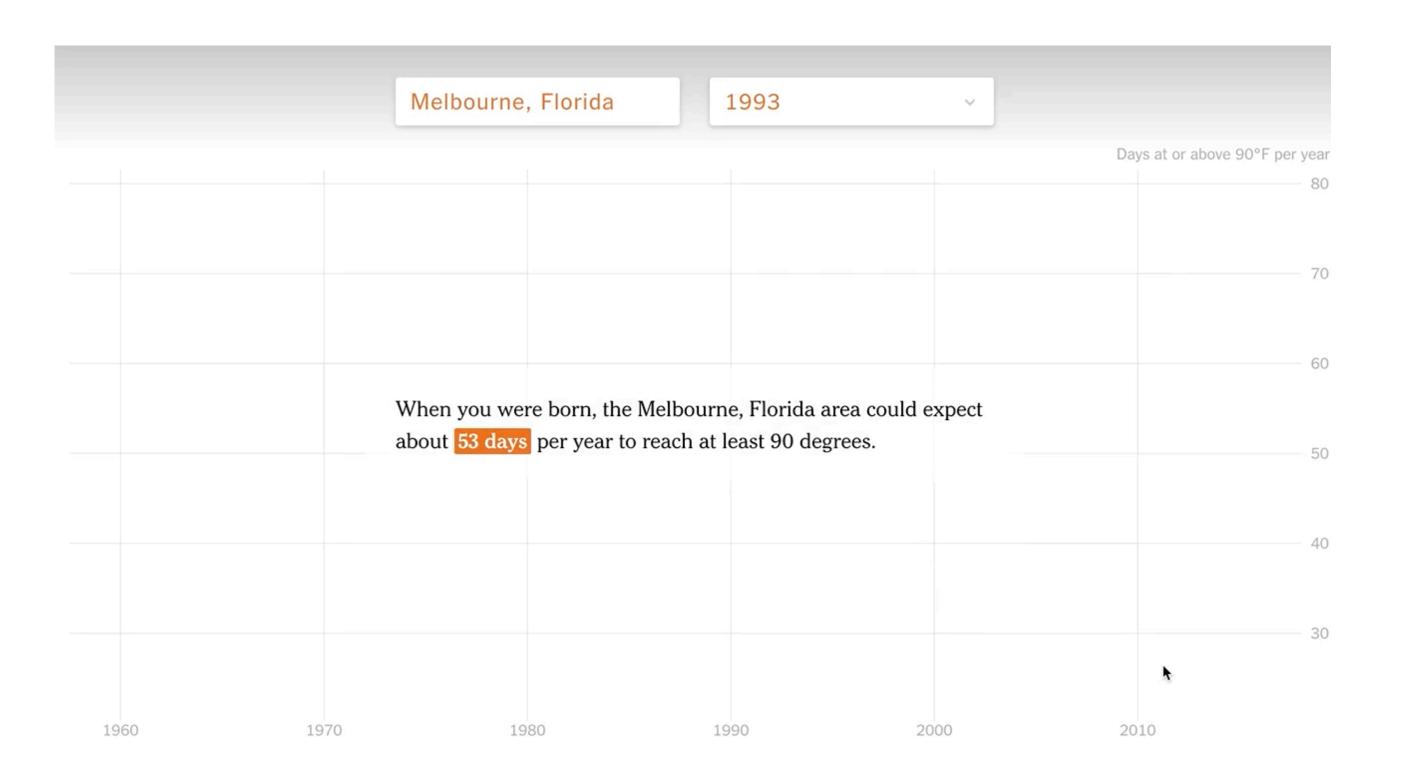
Use effective representations to make complex topics more intuitive.

Personalizing Reading.

Let readers choose the content that is relevant to their own experience.

Prompting Self-Reflection.

Help readers learn by asking them to reflect in a low pressure environment.



Popovich, Migliozzi, Taylor, Williams, Watkins *NYTimes*, 2018

Five affordances of the format

Connecting People and Data.

Make data pleasant to work with. Happy readers are engaged readers.

Making Systems Playful.

Run interactive simulations directly in the browser. No setup required.

Reducing Cognitive Load.

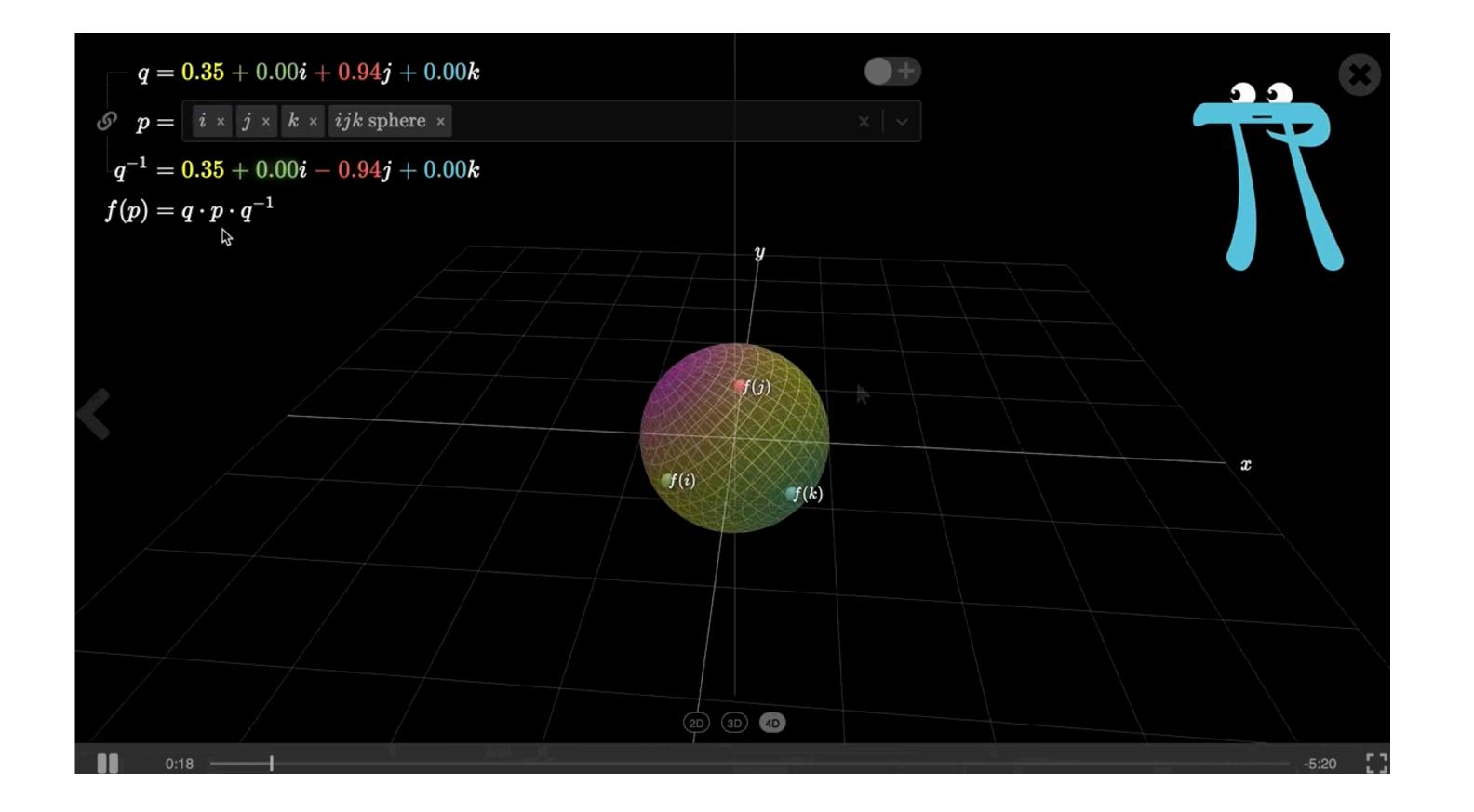
Use effective representations to make complex topics more intuitive.

Personalizing Reading.

Let readers choose the content that is relevant to their own experience.

Prompting Self-Reflection.

Help readers learn by asking them to reflect in a low pressure environment.



Five affordances of the format

Connecting People and Data.

Make data pleasant to work with. Happy readers are engaged readers.

Making Systems Playful.

Run interactive simulations directly in the browser. No setup required.

Reducing Cognitive Load.

Use effective representations to make complex topics more intuitive.

Personalizing Reading.

Let readers choose the content that is relevant to their own experience.

Prompting Self-Reflection.

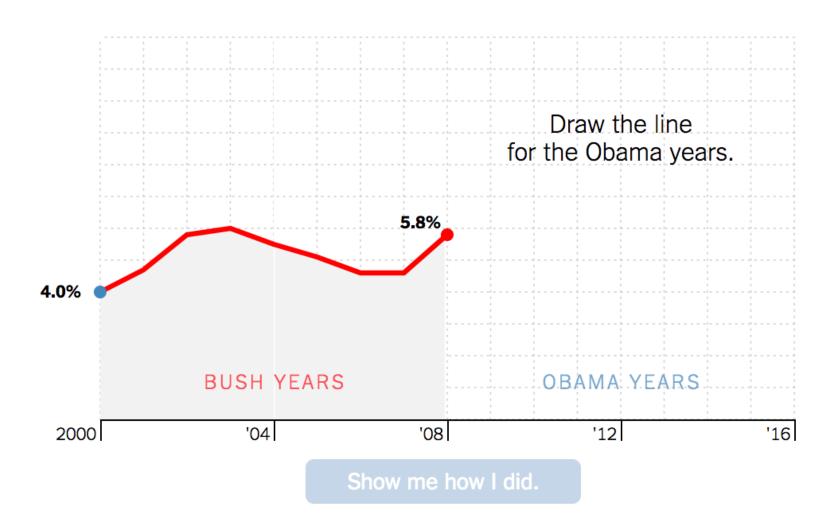
Help readers learn by asking them to reflect in a low pressure environment.

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 ...



Buchanan, Park, Pearce, NYTimes, 2017

Five affordances of the format

Connecting People and Data.

Make data pleasant to work with. Happy readers are engaged readers.

Making Systems Playful.

Run interactive simulations directly in the browser. No setup required.

Reducing Cognitive Load.

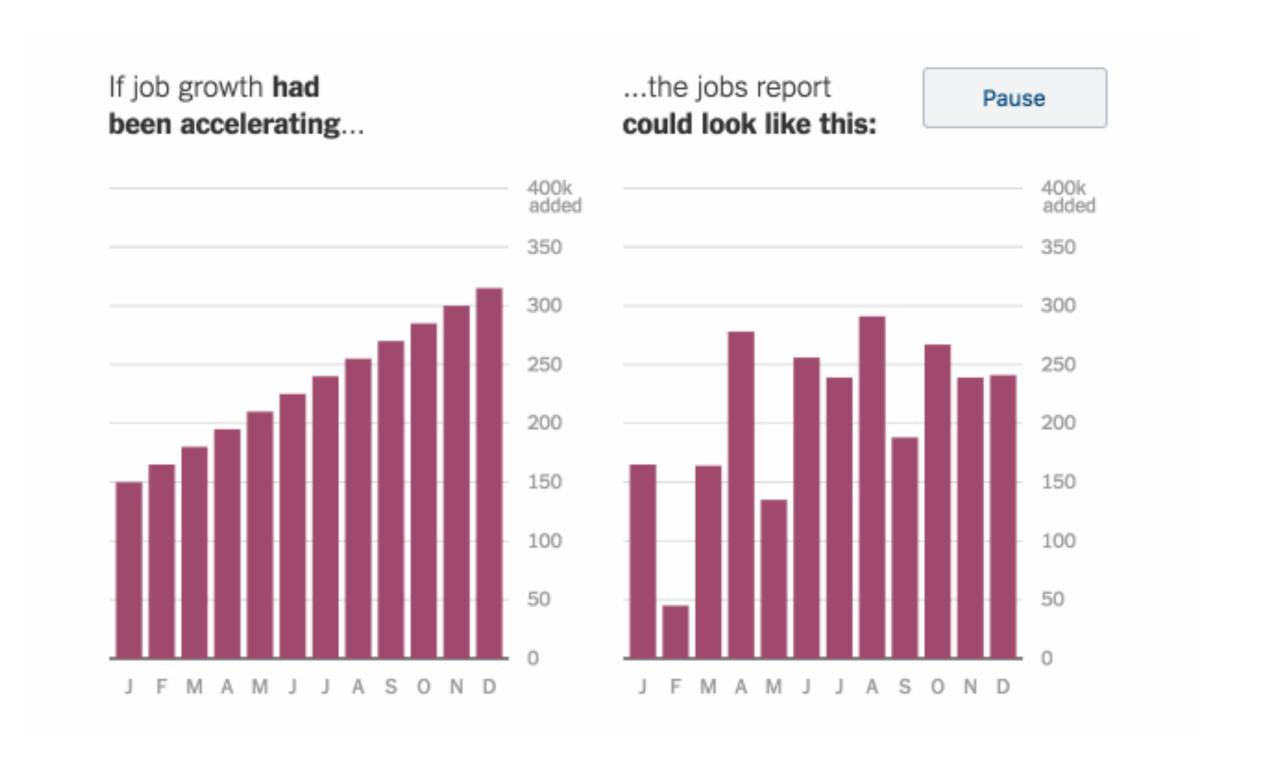
Use effective representations to make complex topics more intuitive.

Personalizing Reading.

Let readers choose the content that is relevant to their own experience.

Prompting Self-Reflection.

Help readers learn by asking them to reflect in a low pressure environment.



Challenges

Few authoring tools available

Those that do exist are too low level or lack expressivity Not suitable for production usage

Requires general purpose programming

In practice need to be a web developer

Difficult to evaluate

Limited empirical evidence; hard to analyze usage data; lack of metrics

Lack of design guidance

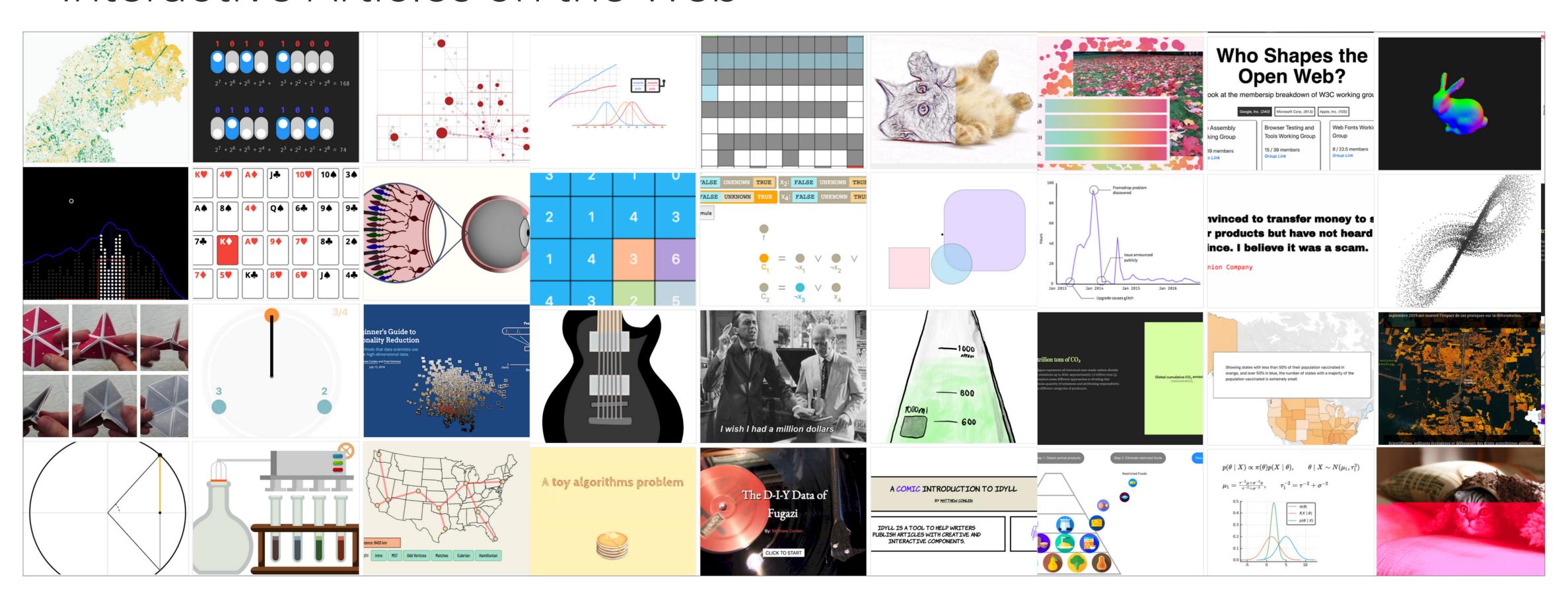
Integration between text and media

Multiple output targets

How to make content accessible?

Idyll

A Markup Language for Authoring and Publishing Interactive Articles on the Web



DSL for Interactive Articles

Idyll Markup

```
[Header
  title:"Hi Idyll!" /]

This is the text
  of my article.

[Graphic
    value:x /]

And that was a
  custom graphic.
```

Input:

extended markdown domain-specific graphics

DSL for Interactive Articles

Idyll Markup

```
[Header
  title:"Hi Idyll!" /]

This is the text
  of my article.

[Graphic
    value:x /]

And that was a
  custom graphic.
```

Input:

extended markdown domain-specific graphics

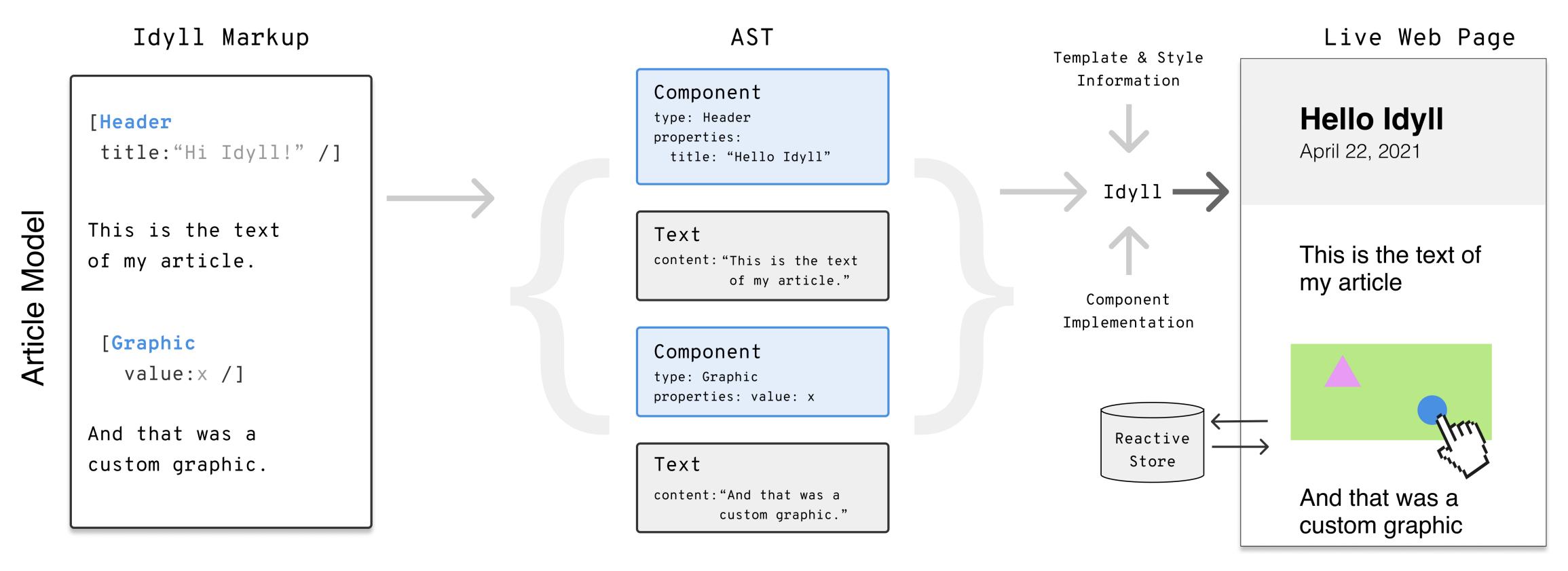
Live Web Page



Output:

javascript, html, css

DSL for Interactive Articles



Input:

extended markdown

domain-specific graphics

interactive document schema

(Interactions, linked text+graphics, events)

Output:

javascript, html, css

Text + Parameterized Graphics

Write articles in markdown.

Embed parameterized graphics.

Connect reactive variables to user input widgets to add interactivity.

Text + Parameterized Graphics

Write articles in markdown.

Embed parameterized graphics.

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 /]
```

Text + Parameterized Graphics

Write articles in markdown.

Embed parameterized graphics.

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" /].

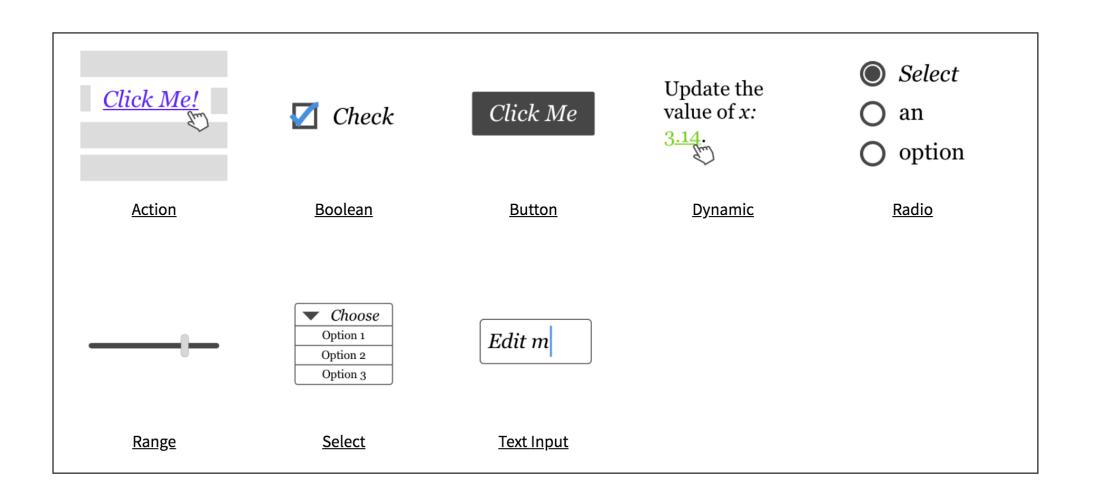
[Range value:x min:0 max:10 /]
```

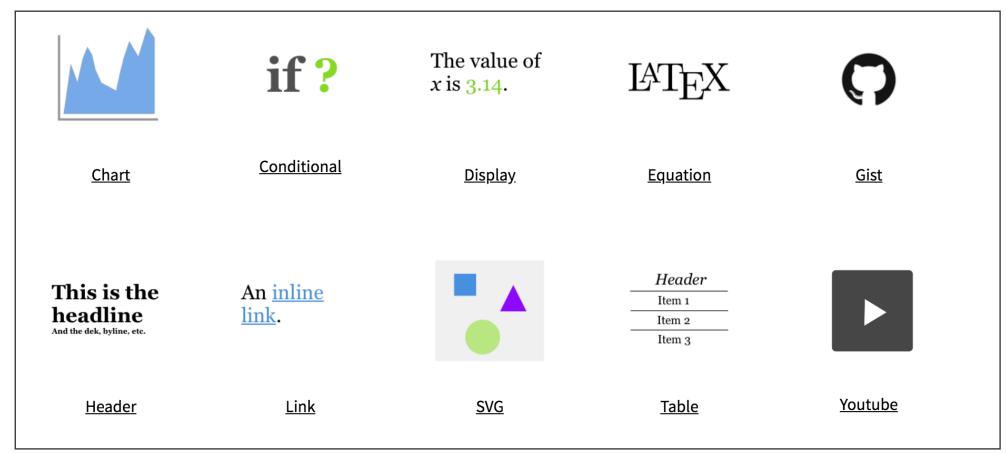
Hello World

The value of x is 5.



Component Library

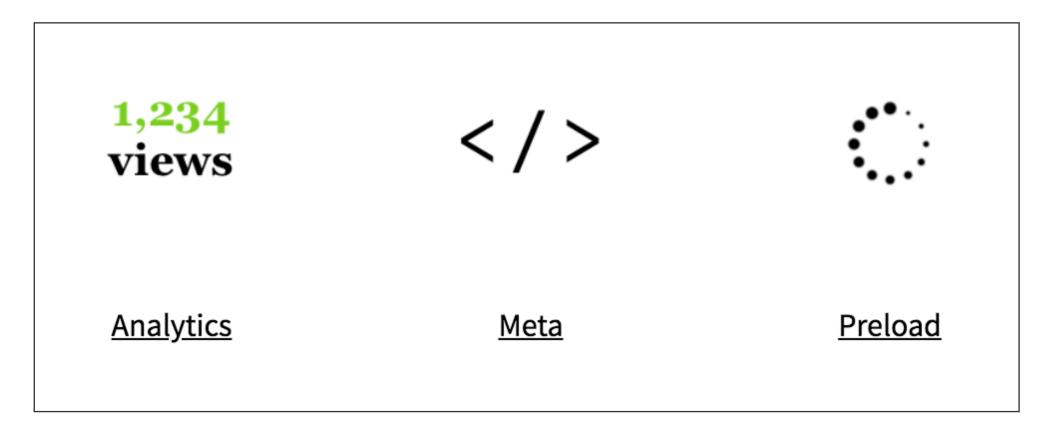




Input



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, ...

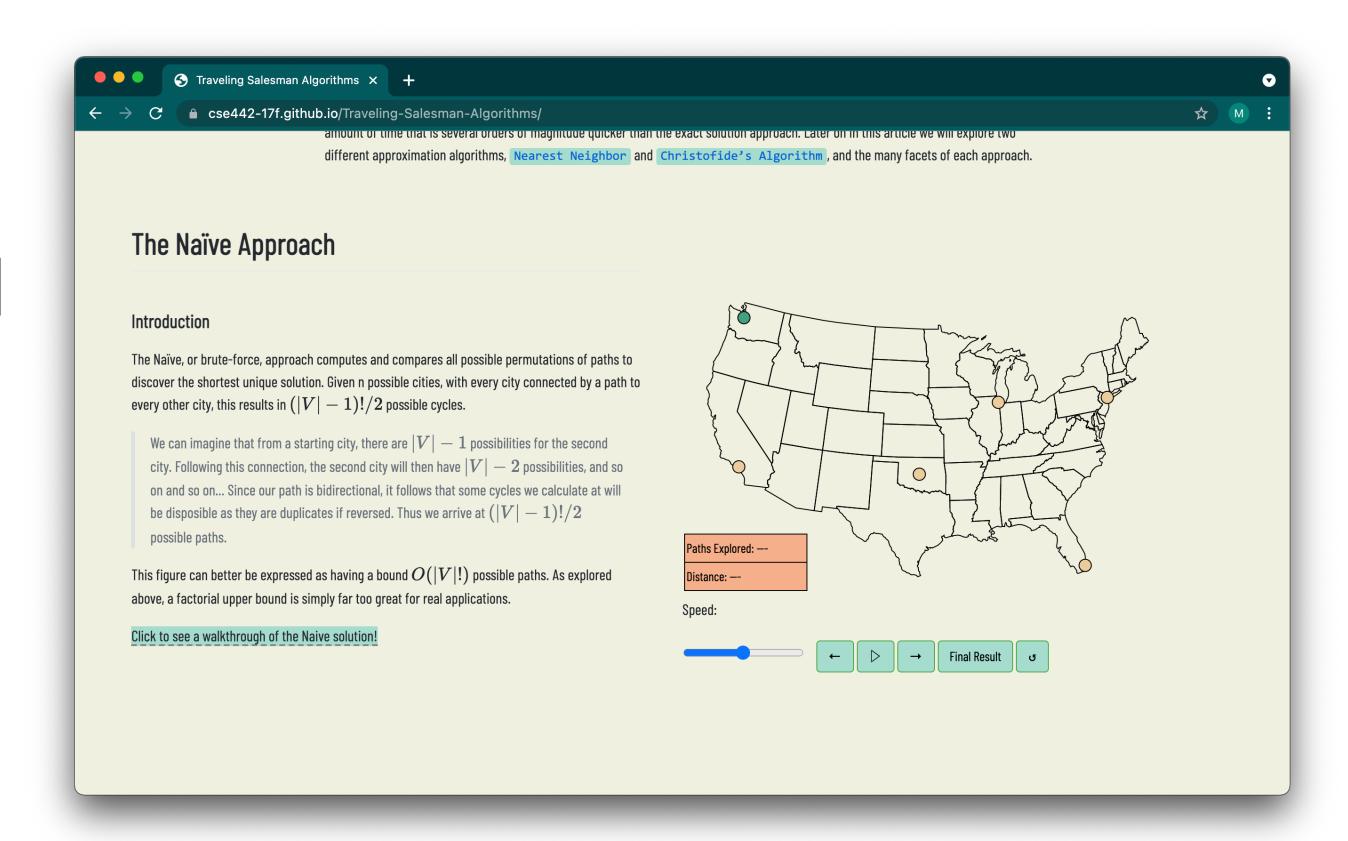


Usage

Refined design through expert interviews.

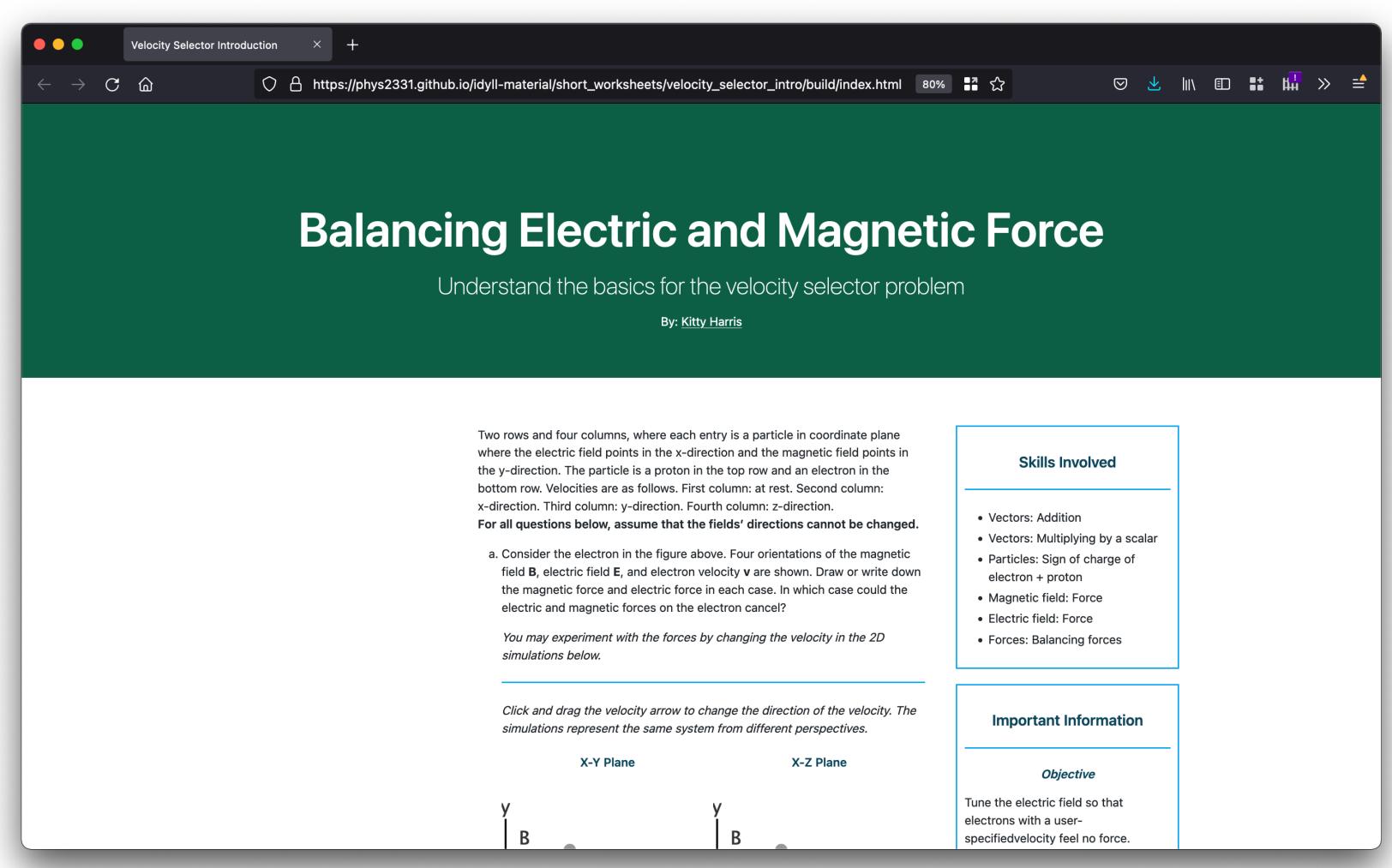
Deployed with undergrad CS students at UW.

Students successfully in creating highly customized explorable explanations of algorithms.



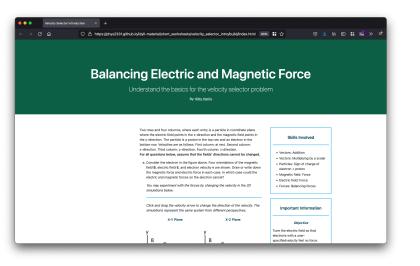
Mathew, Cherukupalli, Pusich, Zhao, Traveling Salesman Algorithms, CS442 '17

Idyll in the wild

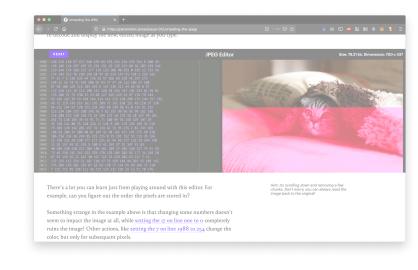


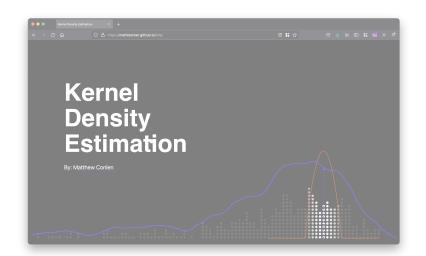
UC Denver

Undergraduate physics material (interactive worksheets)









IdyII...

Seems great but...

...still requires learning and writing markup syntax

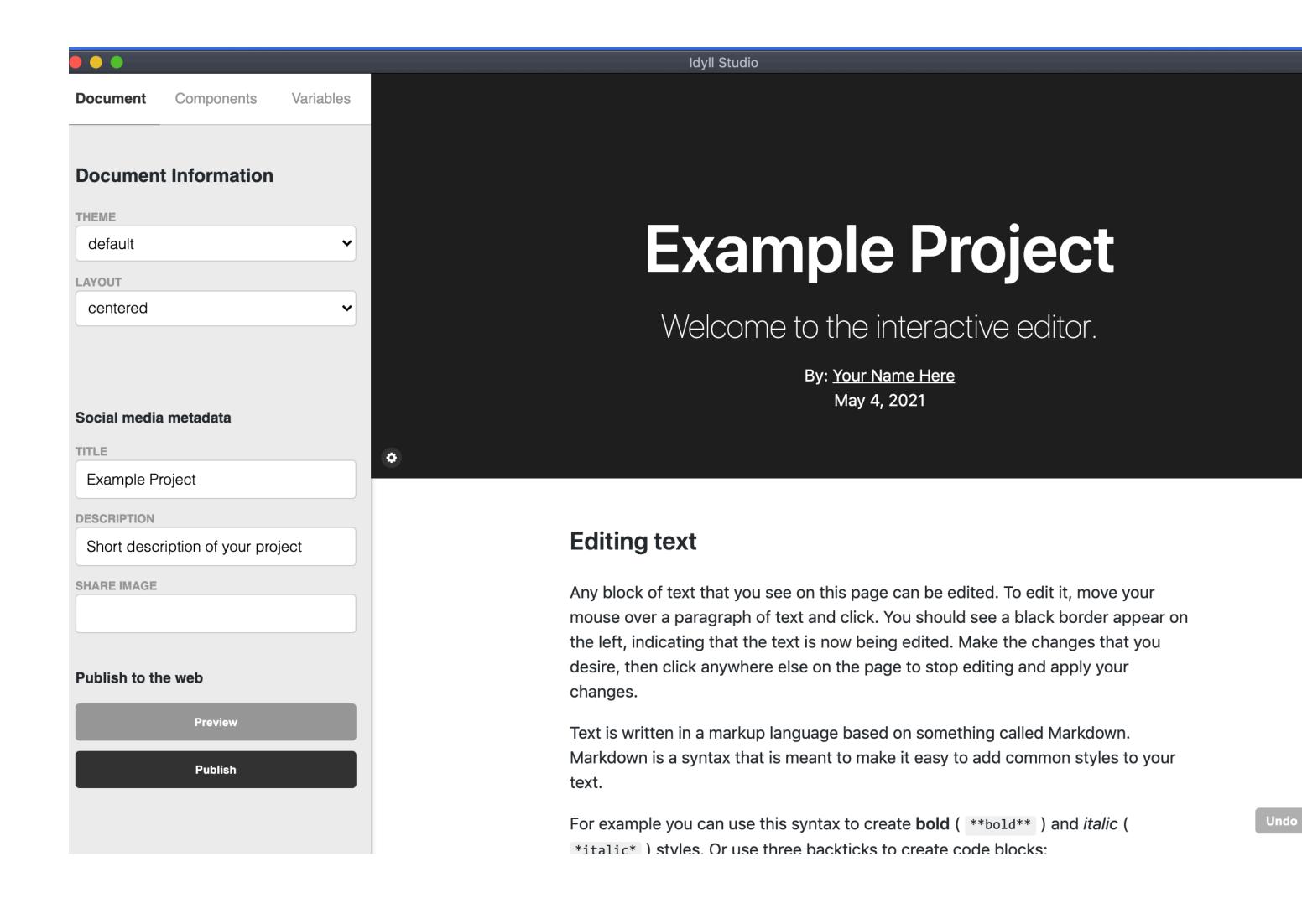
...requires use of general-purpose programming tools, e.g. to make a new post, run development server

Idyll Studio

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

Reduce use of generalpurpose programming tools.

Eliminate a class of syntax errors.



Reify Idyll model.

Conlen, Vo, Tan, Heer, UIST 2021

Publish

Idyll Studio

Introducing UIST to the structured editor.

By: Matthew Conlen, Megan Vo, Alan Tan and Jeffrey Heer Aug 11, 2021

Editing text

•

Idyll Studio

Any block of text that you see on this page can be edited. To edit it, move your mouse over a paragraph of text and click. You should see a black border appear on the left, indicating that the text is now being edited. Make the changes that you desire, then click anywhere else on the page to stop editing and apply your changes.

Text is written in a markup language based on something called Markdown Undo

Redo

Markdown is a syntax that is meant to make it easy to add common styles to your

Idyll Studio Findings

- 1. The interface enabled users to **rapidly create designs with less stress**.
- 2. But some users struggled to synthesize multiple concepts. Others **needed time to understand the reactive model**.
- 3. The interface **should provide more visual feedback**, promote experimentation, provide documentation on demand, and have **guardrails to prevent likely mistakes**.
- 4. The interface empowered users and provided better support to less technical users compared to existing tools.

	Role	Tech.	I1	D1	I2	D2	I3	D3
P1	comp. biologist	4.2	У	У	У	У	у	y
P2	programmer	4.2	У	у	У	у	у	y
P3	hci researcher	3.8	У	у	у	у	y	y
P4	vis. practitioner	3.7	у	y	у	у	y	y
P5	hci researcher	3.5	У	у	У	У	у	y
P6	journalist	3.4	У	у	У	у	у	y
P7	hci researcher	3.4	у	у	у	у	y	y
P8	lab manager	3.4	у	y	у	у	y	y
P9	programmer	3.2	У	у	У	У	у	y
P10	vis. practitioner	2.9	у	y	у	у	y	y
P11	editor	2.9	у	y	y	у	y	y
P12	cs/journ. student	2.5	У	у	У	у	у	y
P13	data scientist	2.5	У	у	у	у	n	n
P14	analyst	2.3	у	y	у	у	n	n
P15	designer	2.3	у	y	у	у	n	n
P16	writer	1.8	У	y	У	n	n	n
P17	ux researcher	1.8	У	у	У	У	У	У
P18	designer	1.3	У	У	У	У	n	n

Design Prompt

You want to keep the public as informed as possible of the status of the 2024 presidential election as it happens.

May want to convey:

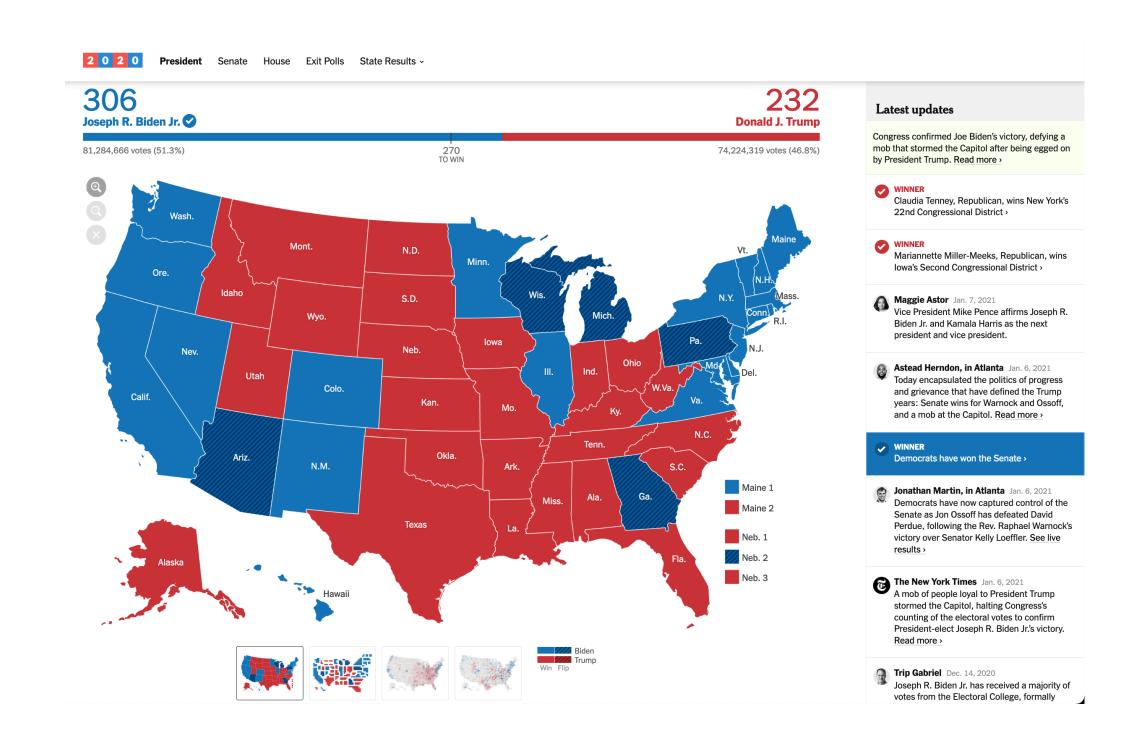
Current snapshot

Likely outcomes

Awareness of uncertainty

Historical context

Any unusual circumstances (e.g. delayed returns)



Input Data

Vote counts candidate, state, method, timestamp, # votes

Race calls candidate, state, timestamp

Constrains

Near real time

Never seen the data before

Data may contain errors

Should not erode public trust

Input Data

Vote counts candidate, state, method, timestamp, # votes

Race calls candidate, state, timestamp

Constrains

Near real time

Never seen the data before

Data may contain errors

Should not erode public trust

How would you visualize this?

What data would you use?

Input Data

Vote counts candidate, state, method, timestamp, # votes

Race calls candidate, state, timestamp

Constrains

Near real time

Never seen the data before

Data may contain errors

Should not erode public trust

How would you visualize this?

What data would you use?

You want to keep the public as informed as possible of the status of the 2024 presidential election as it happens.

May want to convey:

Current snapshot

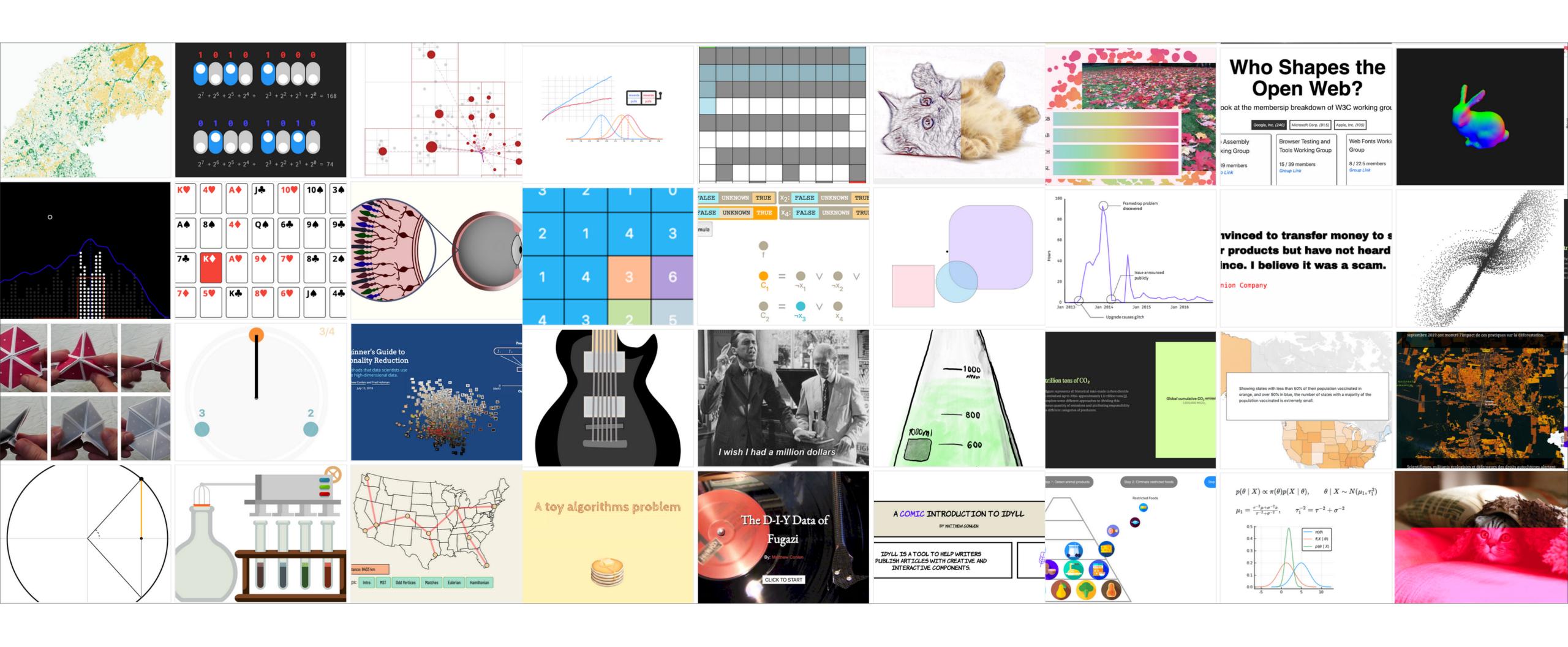
Likely outcomes

Awareness of uncertainty

Historical context

Any unusual circumstances

Thank you!



Contact:

mpconlen@gmail.com @mathisonian