26. The Web as Medium for Data Visualization

Elliot Bentley

Abstract

Exploring the types of graphics made possible by the web, including interactive dataviz, games and virtual reality (VR).

Keywords: interactive graphics, data visualization, web development, JavaScript, infographics, newsgames

Not all media are created equal. A 20-episode television series is able to tell a story differently than a two-hour film, for example. In the same way, the humble web page can provide its own possibilities for data visualization.

The web was originally designed for simple, hyperlinked documents consisting of mostly text and static images. The addition of JavaScript and a slow drip of new features and tools has expanded the palette available to work with.¹

Although traditional data visualization theory and techniques (e.g., Edward Tufte, Jacques Bertin) are still mostly applicable to graphics on the web, the unique features of the web provide vast potential for new forms of data journalism. These works are often referred to as “interactives,” an awkward word that obscures some of the web’s unique strengths.

Below is a list illustrating some of the ways in which graphics on the web can take advantage of their host medium.

¹ Other “multimedia” platforms, especially Flash, provided a wealth of options long before the open web did. However, for better or worse these have since been phased out. And while all of these features—more in fact—are available in native applications, the web is far easier to work with and distribution is practically free.

Huge, Explorable Data Sets

A classic use of interactivity is to present the reader with a huge data set and allow them to “dive in” and explore in as much depth as they like. Sometimes this takes the shape of a giant table; other times, a big interactive map.

This format is often looked down upon nowadays, since it expects the reader to find the interesting bits themselves; but it can still be valuable if the data is juicy enough. I find that the most successful versions accept the fact they are simply tools (as opposed to being articles), such as the extremely popular *Wall Street Journal* College Rankings or *ProPublica*’s public-service news apps.²

Guide the Reader Through Complex Charts

A now-common format begins with a single chart and then proceeds to manipulate it—zooming in and out, travelling through time, switching out data—in order to fully explore the data set. This pairs exceptionally well with scrollytelling and is especially valuable on mobile, where there may not be enough space to show all elements of a chart at once.³

In the now-classic piece “A Visual Introduction to Machine Learning” (Figure 26.1), the same data points transition between multiple chart formats, helping readers keep track of how the machine learning algorithms are sorting them.⁴ Another good example is “100 years of tax brackets, in one chart,” a *Vox* piece that zooms in and out of a data set that might be overwhelming if presented otherwise.⁵

Up-to-the-Second Live Data

Why settle for a static data set when you can use the latest numbers of whatever you’re charting? Elections, sport coverage, weather events and financial data are obvious sources of live data interesting enough to display in real time. Even more cool is providing context for these live figures in

2 <https://www.wsj.com/articles/explore-the-full-wsj-the-college-rankings-11567638555>, <https://www.propublica.org/newsapps/>

3 <https://pudding.cool/process/how-to-implement-scrollytelling/>

4 <http://www.r2d3.us/visual-intro-to-machine-learning-part-1/>

5 <https://www.vox.com/2015/10/26/9469793/tax-brackets>

Figure 26.1. A visual introduction to machine learning. Source: R2D3. <http://www.r2d3.us/visual-intro-to-machine-learning-part-1/>

interesting ways—for example, showing which countries benefit from the current price of oil (Figure 26.2).

Ampp3d, a short-lived experimental pop-data journalism outlet, used live counters to bring numbers to life in interesting ways, such as the number of immigrants entering the United Kingdom, and footballer Wayne Rooney's earnings.⁶ Sadly, these have since been taken offline.

Placing the Reader Within a Data Set

Another twist on the “huge data sets” idea—and one that I’ve found to be incredibly compelling to readers—is to show the reader where they fall in a data set, usually by asking for a couple of personal details. *The New York*

6 <https://onlinejournalismblog.com/2015/05/13/the-legacy-of-ampp3d-usvsth3m-and-row-zed/>

Figure 26.2. Countries that benefit from the current price of oil. Source: *Wall Street Journal*. <http://graphics.wsj.com/oil-producers-break-even-prices/>

Times' 2013 dialect quiz map (Figure 26.3) famously became the publication's most popular article of the year—despite only being published on December 20th.

The BBC seem to do these pretty frequently, often as a public service tool, with things like “UK fat scale calculator.”⁷ I like this *Quartz* piece on how people in different cultures draw circles, which opens by asking the reader to draw a circle, a compelling introduction to an otherwise (potentially) dull feature.⁸

Collecting Original Data Sets

A step beyond the previous category are projects that not only use readers' submitted data to give an immediate response, but also to compile a new data set for further analysis.

The Australian Broadcasting Corporation collaborated with political scientists on a “Vote Compass” to help readers understand their place

7 <https://www.bbc.com/news/health-43697948>

8 <https://qz.com/994486/the-way-you-draw-circles-says-a-lot-about-you/>

How Y'all, Youse and You Guys Talk

By JOSH KATZ and WILSON ANDREWS DEC. 21, 2013

What does the way you speak say about where you're from? Answer all the questions below to see your personal dialect map.

QUESTION 1 OF 25

How would you address a group of two or more people?

- ☐ you all
- ☐ yous / youse
- ☐ you lot
- ☐ you guys
- ☐ you 'uns
- ☐ yinz
- ☐ you
- ☐ other
- ☐ y'all

Next >

Figure 26.3: *The New York Times*' 2013 dialect quiz map. Source: *The New York Times*. <https://www.nytimes.com/interactive/2014/upshot/dialect-quiz-map.html>

in the political landscape—and then wrote a series of articles based on the data.⁹

More recently, *The New York Times* used the same idea on a softer subject, asking readers to rate *Game of Thrones* characters and plotting the results on live charts (Figure 26.4).

The Infinite Canvas

The web is infinite in its scope and capacity, but more specifically web pages can be as wide or tall as they like—an “infinite canvas” on which to work. I borrowed this term from artist Scott McCloud, who argues that

9 <https://www.abc.net.au/news/nsw-election-2015/vote-compass/>, <https://www.abc.net.au/news/nsw-election-2015/vote-compass/results/>

Your Estimates, and Everyone Else's

Tyrion Lannister

Introduced as "The Imp," Tyrion was a hard-drinking, whoring, black sheep of the royal family; now, he's a respected strategist.

Daenerys Targaryen

A stunning beauty who feels for the oppressed. But she has no problem sending people to their deaths as she conquers kingdoms.

Figure 26.4: A plot chart rating *Game of Thrones* characters. Source: *The Upshot*. <https://www.nytimes.com/interactive/2017/08/09/upshot/game-of-thrones-chart.html>

there is “no reason that longform comics have to be split into pages when moving online.”¹⁰ And indeed, why should our graphics be constrained to the limits of paper either?

In *The Washington Post*’s “The Depth of the Problem,” a 16K-pixel-tall graphic is used to show the depth of the ocean area being searched for missing flight MH370 (Figure 26.5).¹¹ Sure, this information could have been squeezed into a single screen, but it would have lacked the level of detail and emotional impact of this extremely tall graphic.

In *The Guardian*’s “How the List Tallies Europe’s Migrant Bodycount,” tens of thousands of migrant deaths are powerfully rendered as individual dots that appear one by one as the reader scrolls down the page.¹²

¹⁰ <http://scottmccloud.com/4-inventions/canvas/index.html>

¹¹ <https://www.washingtonpost.com/apps/g/page/world/the-depth-of-the-problem/931/>

¹² <https://www.theguardian.com/world/2018/jun/20/the-list-europe-migrant-bodycount>

Figure 26.5. Graphic that shows the depth of the ocean area being searched for the missing flight MH370. Source: *The Washington Post*. <http://apps.washingtonpost.com/g/page/world/the-depth-of-the-problem/931/>

Data-Driven Games

“Newsgames,” interactive experiences that borrow mechanics from video games to explore news subjects, have existed for a while, with varying levels of success.

The Upshot’s “You Draw It” series (Figure 26.6) challenges readers’ assumptions by asking them to fill in a blank chart, before revealing the answer and exploring the subject in greater depth.

Some games are more involved, perhaps asking the reader to solve a simplified version of a real-world problem—such as how to fund the BBC—to prove just how difficult it is.¹³

¹³ <https://ig.ft.com/sites/2015/bbc/>

Figure 26.6. A chart from *The Upshot*'s "You Draw It" series.
Source: *The Upshot*. <https://www.nytimes.com/interactive/2015/05/28/upshot/you-draw-it-how-family-income-affects-childrens-college-chances.html>

These could be considered toys that only present the reader with surface-level information, but done right they can provide a fresh perspective on played-out subjects. *FiveThirtyEight*'s "How to Win a Trade War," in which the reader chooses a trading strategy and competes against a previous visitor to the page, brings to life the otherwise potentially dry economic theory.¹⁴

Live, Randomized Experiments

A related format is to allow the reader to run a live simulation in their browser. More than just an animated explainer, this introduces a degree of randomness that leads to a unique result each time and is a great way to bring abstract statistical probabilities to life.

The *Guardian* piece in Figure 26.7 simulates a measles outbreak across ten populations with varying rates of vaccination. The web graphics make the results starkly clear in a way that percentages alone could not convey. In Nathan Yau's "Years You Have Left to Live, Probably," a simple line chart

¹⁴ <https://fivethirtyeight.com/features/how-to-win-a-trade-war/>

Figure 26.7: A simulation of a measles outbreak across ten populations with varying rates of vaccination. Source: *The Guardian*, <https://www.theguardian.com/society/ng-interactive/2015/feb/05/sp-watch-how-measles-outbreak-spreads-when-kids-get-vaccinated>

(“probability of living to next year”) is made more poignant with “lives” that die at random and then pile up.¹⁵

These simulations don’t have to use imaginary data. “The Birthday Paradox” tests the probability of shared birthdays using data from previous visitors to the page.¹⁶

3D, VR and AR

3D graphics and virtual reality are difficult to harness in service of data journalism, outside of maps of terrain.

Two notable experiments, both from 2015 and on the subject of financial data (“Is the Nasdaq in Another Bubble?” and “A 3-D View of a Chart That Predicts the Economic Future: The Yield Curve”), are clever novelties but failed to spark an explosion of three-dimension charts.¹⁷ Perhaps for the best.

¹⁵ <https://flowingdata.com/2015/09/23/years-you-have-left-to-live-probably/>

¹⁶ <https://pudding.cool/2018/04/birthday-paradox/>

¹⁷ <http://graphics.wsj.com/3d-nasdaq/>, <https://www.nytimes.com/interactive/2015/03/19/upshot/3d-yield-curve-economic-growth.html>

The potential of augmented reality, in which a camera feed of the real world is overlaid with graphics, has yet to be proven.

Conclusion: How New Formats Arise

Some of the web graphics listed above are new formats that have only emerged over the past few years; some have stuck around, such as the guide through a complex chart (typically using a scrollytelling interaction pattern). Others, like three-dimensional charts, were mere flashes in the pan.

Yet it's not just taste that determines which types of graphics are in vogue on the web: Available technology and readers' consumption habits shape trends, too.

Take, for example, the widely used interactive map. In addition to being a visually attractive and easily grasped form, the proliferation and familiarity of this format was doubtless helped by tools that make them easy to create and manipulate—Google Maps and Leaflet being two of the most common.

Without any hard data to hand, it at least feels as though fewer interactive maps are being published nowadays. While it would be easy to attribute this trend to a growing realization among journalists that such interactivity (or even the map itself) can often be superfluous, new technologies likely also contributed to this drop.

A high proportion of readers now access the web using mobile phones, and interactive maps are a particularly poor experience on small touchscreens. In addition, there is a new technological solution that in many ways is superior: *aizhtml*, a script open-sourced by *The New York Times* that generates a responsive html snippet from Adobe Illustrator files.¹⁸ Maps built with *aizhtml* can leverage a traditional cartographer's skill set and still have sharp, machine-readable text. The lack of interactivity in such maps is often a blessing, even if it is in many ways limiting.

This is just one example of how data journalists should be thoughtful in their use of the web's unique features. With so many possibilities to hand, it's important to carefully evaluate those and use them only when truly necessary.

¹⁸ <https://github.com/newsdev/aizhtml>

About the Author

Elliot Bentley has worked in *The Wall Street Journal's* graphics department since 2014 and is also the creator of open-source transcription app oTranscribe.