

[REDDIT](#)

I'm Elijah Meeks, author of D3.js in Action and Semiotic. I do data visualization at Netflix and used to do it at Stanford in digital humanities. Ask me anything quick before data visualization dies.

ELIJAHMEEKS [R/SCIENCE](#)

Hi Reddit, I'm Elijah Meeks. I wrote [D3.js in Action](#) and I just open sourced [Semiotic, a data visualization framework focused on information modeling](#). I used to do data visualization in the digital humanities, including projects like [ORBIS](#), [Kindred Britain](#) and the [Digital Gazetteer of the Song Dynasty](#). Now I work at Netflix visualizing user behavior, algorithm performance and just big data more generally. Lately I've been pushing for the community to [take a critical look at professional data visualization](#) how we design roles, how data visualization is seen by leadership and how we evaluate data visualization products.

[Proof of Life](#)

[Follow me on Twitter](#)

[Read my pieces on Medium](#)

Some examples of my work:

[My Blocks](#)

[A visualization of Archer](#)

[ORBIS - Geographic and Transportation Data Visualization of the Roman Empire](#)

[A timeline of US Wars](#)

EDIT: Okay I came back and responded to a few more things and it was totally worth it.

[READ REVIEWS](#)

[WRITE A REVIEW](#)

CORRESPONDENCE:

DATE RECEIVED:

September 26, 2017

DOI:

10.15200/winn.150634.43856

ARCHIVED:

September 25, 2017

CITATION:

elijahmeeks, r/Science, I'm Elijah Meeks, author of D3.js in Action and Semiotic. I do data visualization at Netflix and used to do it at Stanford in digital humanities. Ask me anything quick before data visualization dies., *The Winnower* 4:e150634.43856, 2017, DOI: [10.15200/winn.150634.43856](#)

Why would data visualization die?

[Fxlrye](#)

When I look at the vibrancy of work being done and the boundaries being pushed in data visualization five years ago and compare it to today, it just doesn't seem to compare. Major consulting firms have reduced or entirely eliminated their data visualization units, here in the Valley some of the leading tech companies have scrapped internal data visualization teams. More abstractly, I feel like I hear from more and more dissatisfied people in the field, either in their personal career trajectories, or their lack of enthusiasm for the conferences and conference topics.

Data visualization can't really die, because even spreadsheets are data visualization, but the kind of exciting, ground-breaking work that tries to push forward the ability to communicate about complex phenomena can be overtaken by a more simplistic, more conservative structure that leaves us making bar charts and spreadsheets and scatterplots. When I wrote about professional dissatisfaction, there were a few people--prominent theorists like Stephen Few but also good friends of mine--who argued that data visualization isn't even a profession, just a skill, and I suppose my concern is that were it reduced to an ancillary skill, we'd see just that ossification of the practice.

© et al. This article is distributed under the terms of the [Creative Commons Attribution 4.0 International License](#), which permits unrestricted use, distribution, and redistribution in any medium, provided that the original author and source are credited.



Are Tufte's books still relevant and good places to get started? What resources do you recommend to people starting out with data visualization targeted at internet-based audiences?

[srm561](#)

Yes I think Tufte is still relevant but remember he had in mind a particular rhetorical moment: The summary communication with the busy decision maker. That moment was much more prominent back when Tufte started writing his books but is less so today. As far as resources, I'd look at the work of Nicky Case if you want to see how to really communicate with visuals, the books of Alberto Cairo which are quite good in spite of their moralizing titles (Data Visualization is no more "truthful" or "honest" than any other rhetorical form) because of their focus on journalism and then just more generally Andy Kirk's work for its accessible typology of chart forms that goes beyond the usual gestalt and bar charts dance most books do.

Regarding your "why people are leaving dataviz jobs" post, you seem to be in a good spot at Netflix. Getting time to build out things such as Semiotic (or Susie Lu being able to open source d3-annotation) and I hear that data visualization is also more and more appreciated and applied within Netflix. Do you then feel that you just "got lucky" in terms of your data visualization job, not wanting to leave, or does Netflix have a model on how to make proper use of their visualization focused employees that other companies don't have?

[nbremer](#)

I think in some regard I did just get lucky. Netflix has a great culture with a lot of latitude (as was evidenced by a couple unauthorized tell all AMAs) that allows for more innovation when there's not necessarily the structure or process in place to support it. I was also lucky in their hiring Susie, since any design practice--and I firmly believe data visualization is a design problem not an engineering problem--is improved by having a someone to collaborate with.

Ultimately I feel like I was successful because I have a forceful personality and a willingness to push for innovation. What worries me is that without the explicit structural buy in from leadership that only people with strong personalities will have a chance to succeed, and being loud and forceful does not necessarily correlate with being a good coder, designer or creator in any field.

Can you remember a time where the use of statistics dramatically changed your opinion on something? A scenario where the stats disproved many of your preconceived notions about a topic?

[zonination](#)

Sure, but first I'd like to focus on the times when that was scary and not positive. See, I've sat in a number of presentations where I watched a statistical technique or data visualization of model results that showed a different pattern than people were expecting, and watched experts start to theorize about how that different pattern made sense, only to find out it was an error in the modeling or analysis. Human beings are very good at ex post facto justification, especially when the explanation is wrapped in complex language or imagery. So we should be careful.

Hi, I'm an Economics undergraduate student. I'd love to pursue a career in Data Science. What do you think it's the best "career path" to follow?

Edit: Spelling

[pierpa17](#)

I hear some folks don't think there will be data scientists by the time you have the credentials to get a job at a Netflix--but I'm sure there's an [r/datascience](#) that can better address that. From a data visualization perspective I think the most impactful data scientists are those who are skilled enough in the use of the existing tools, like ggplot, to produce charts that don't just let them explore the data but also collaborate with other scientists and stakeholders. I've seen too many times when a data scientist, in love with their particular chart, just cannot seem to recognize that the chart is completely arcane to her audience. I was a philosophy major as an undergraduate, with little experience with statistical methods, but good scientific communicators can cut through that and enable me to contribute to making their research and products better. It's challenging, and it's one of those "soft" skills that are maddeningly difficult to describe or achieve, but it's critical in the modern collaborative environment.

So for data science, become a solid statistician but don't avoid all those opportunities to learn how to actually communicate to others.

You call for industries to allow space for more complex visualizations including "charts that at first look like art" and "scrollytelling" in your blog post. I love these too (I'm in this subreddit after all) but in practicality, working in a large company, most of my data needs to be understandable by the widest range of people in the shortest feasible amount of time.

Have you found any methods or success stories for raising the general data literacy of whatever group usually consumes your visualisations?

[heyheyhedgehog](#)

No.

Okay maybe one: Find out what kind of less complex charts the final complex chart is related to and build in an evolutionary tree leading to your chart.

Also I'd like to challenge your claim. There are all kinds of data at companies, some of it needs to be understood quickly by a lot of people, some of it needs to be understood slowly by a few people, and in that spectrum are suitable chances for complex charts. I make a lot of bar charts at Netflix and that's okay.

1. Opinions of Edward Tufte, for good or ill?
2. What's the biggest mistake(s) amateur/part-time data visualizers make?
3. What concrete principles should these folks be following to improve their visualizations?

[greginni](#)

Tufte is good, he knows what he's talking about and established a set of solid no-nonsense rules about unnecessary decoration that should be the default (though decoration is not the bugbear we've come to treat it as). The problem with Tufte is that many of his core examples are flawed:

- The Challenger critique is based entirely on hindsight, as an academic paper addressed a while back
- Minard's Map is another in a long line of anti-Russian chauvinism that pretends the Russian army never defeats its enemies, only winter

After writing [your post](#) regarding people leaving data visualization for other professions like front end

development or data science, have you felt any change within the community? Do more people stay/leave, do they speak more open about data visualization in general?

[ostedog](#)

No, I haven't, it's been very disheartening and I figured my energy would be better spent on productive activities. That's one of the reasons I open-sourced Semiotic, I thought it would provide a nice example of what I thought was a good way forward that didn't have to do with thought leadering.

How much of the problem is managers who don't want poor data quality revealed by data visualizations? I know these managers exist; my last supervisor didn't want any kind of data visualization. Once I figured this out, to avoid antagonizing him, I never made any more statistical graphics. Obviously, my role was not data visualization.

[QueueLinx](#)

While there are terrible situations to be in, I like to pretend that everything is perspective. If it's just an unethical situation, there's nothing you can do, but there are ways that data visualization practitioners can be dogmatic and antagonistic and not recognize their role in crafting an analytical view that allows for meaningful work to be done. If the data quality is not an issue or is so bad that you cannot fix it, then there are other patterns that may be available.

I feel like I'm doing one of those Tony Robbins things. DATA VISUALIZATION FOR GOOD TO FIX BROKEN HEARTS

Hi Elijah,

What are your thoughts around 3D data visualizations? Perhaps in VR or AR?

[ostedog](#)

This is where I'm definitely an old man who doesn't understand kids and their pokemon. When I look at VR and AR viz I think it's a goofy waste of time. I hope I'm wrong. The one place where I could see value is the intersection between AR and data visualization in video games, which I think is underexplored theoretically.

What is your favorite example of a good data visualization?

What about your favorite example of a bad data visualization?

[zonination](#)

I love the bump area chart from NYT that shows movement of peoples from different states and to different states.

I think most data visualization is very bad.

are there any underutilized ways to visualize time that you like?

[MettaWorldSteveBlake](#)

Tom Shanley just made a Sankey diagram that allows for cycles. I think everyone should try it out.

System visualization like that is typically only done by experts for experts, hence the terrible tensor flow visualization that gets lauded as wonderful, and yet they're so recognizable by large audiences. We should all be doing more of them.

What is the best way for someone interested in data visualization to pursue a career in the field?

[The RagingCaucasian](#)

I think it's still building your portfolio. It's so hard to evaluate whether someone is good at data visualization because it's at the crossroads of coding and design, and to make up for that we look through the work people have done.

Regarding Semiotic, I think that one of the most opinionated decision you made is to split visualisations in layers, with amongst other ones, one layer for the graphs and one for the interactions zones.

This can lead to adding quite a few elements only for interaction. I was for example suprised to see that you create rectangles on top of line graphs, that are the sensitive zones determining which point of the graph is highlighted on hover.

What drove you to that design? Is this something that you do a lot at Netflix? Is this something that you did to have a general solution for interactivity?

[madewulf](#)

Creating interaction regions on line charts is an old idea. Mike Bostock showed off voronois for line charts, like... five years ago? So that's an expectation at Netflix and pretty much in any modern data viz environment. You see built in support for that in libraries like Victory and elsewhere all the time.

The challenge with semiotic is when you want actual interactivity for actual graphical shapes, which you can do but you have to do instead of using the built-in functionality, rather than in tandem with it.

Hi Elijah, I am a business analytics student without a big math background and have two questions.

1.) Can you tell me what specific areas of math you encounter most when creating visualizations? I've heard probability theory, bayesian, & multivariate calculus are big. Is the calculus absolutely necessary? Just another impatient guy looking to build a career before the industry I want to move into dissolves...

2.) Alluding to this ^ do you think technology is moving faster than the rate at which one can build a career out of something before it evolves into something else? I know im being generic but I think about software replacing even the lowest level of tasks that used to require years of study... Thanks!

[anonadado](#)

I typically see trigonometry and geometry because what I'm most interested in is showing shapes on-screen. My stakeholders are varied in the math that goes into what they want to show on-screen, but since we don't do much visualization of the actual models and rather visualize the effect they have (so rather than visualizing the recommendation algorithm, we visualize the recommendations it makes) there's more emphasis on traditional statistics than on higher level math or Bayesian statistics. But I'm just wrapping up an LDA project and there's another one where we're using TSNE and another TSNE-like dimensional reduction, so it's definitely there.

What is your porcelain chicken's favorite graph?

[ostedog](#)

It's obviously a goose based on its beak and plumage. Come on. Given that geese are notoriously aggressive and small-brained, it would probably be the kind of graph that insider traders and tech bros love, so some kind of line chart with range bars that has a horrible vomit-inspired color scheme and is festooned with logos and jargon.

He is otherwise a wonderful goose.

How do you feel about moving from the public to the private sector? I mean it in a general way - whatever strikes you as worth sharing - but specifically would be curious about lifestyle, job satisfaction, and... Life satisfaction, I guess? In terms of the impact you feel like you're making on the world. Thanks! :)

[finalfronteer](#)

Academia is kind of messed up these days so I find that people in industry are happier and less political. The stuff that I do every day doesn't feel as meaningful as when I was working in the digital humanities. The pay is roughly 8000x what I made when I was working at Stanford but I feel like I'm making less of an impact. Fortunately, I can always bloviate on Medium and Twitter and now on Reddit to make myself feel better, and because of my business card, more people will pay attention.

What role do you see the digital humanities having in our society?

[annieyfly](#)

If you'd asked me five years ago I thought it was going to set the standard for dynamic documents that integrated text and data visualization to represent complex systems. Nowadays, I feel like it provides good post-modern critiques of technological utopianism and nice skills-building in GIS/SNA/traditional stats for disciplines where it hasn't been emphasized. I think the next big movement in that area will happen when the online education companies start to see the need to integrate research publication with traditional pedagogical material so that we can do what we thought we were going to do with ORBIS and build an application that delivers research findings to peer scholars but is also suitable for public audiences and as teaching material for undergraduates and high schoolers.

What are some resources you'd recommend to learn data viz?

[chronicpenguins](#)

Most of the academic stuff starts with visual cognition and that can be pretty eye-glazing (pardon the pun) so I'd look more at the coffee table books, especially the ones that look at histories of data visualization, like Manuel Lima's Book of Trees. That opens up the possibility space, the how you get there depends on what career you're looking at. There's this amazing book "D3.js in Action" that teaches you D3 and a lot of practical advice and theory, too, but that's only useful if you want to learn D3. I have no idea how you get started with, say, ggplot.

Which media/news site do you think does the most effective and creative data visualizations?

[eggn00dles](#)

Anywhere that Adam Pearce is working.

Netflix is (in)famous for the culture of operating like a pro sports team. When you're not of strategic value, you get cut from the team. "We're not a family," is the quote, I believe. What would cause you to be cut from the team? What's it like working under that kind of system? What are the pros and cons, and does it make you better or worse as an employee and human? Thanks :-)

[Ruckdive](#)

I kind of figured all jobs were like that. If you aren't doing well, you find a new place to work, and that decision either comes from the employee or the company. It gets a lot more attention from outside than internally (people do get fired but it doesn't cause people to operate under fear) because I guess all these other companies are old-fashioned government jobs or 1970s zaibatsus?

I suppose they'd get rid of me if it seemed like the work I was doing wasn't having an impact, which is stressful from a certain perspective because we've done such a bad job of creating evaluation metrics for data visualization.

Hi! Thanks for doing this.

I'm a blind date scientist working in the technology industry. Do you see visualizations ever being able to convey information to folk with visual impairments? IE: auditory or tactile methods of simply conveying complex information well?

[TalesOfT](#)

There has been some work done on data sonification but I think it's pretty weak. I think we should seriously invest in a natural language "reader" of data visualization both for visually impaired users and also for analysis and repackaging of content. The problem is most evaluation work done in data visualization is academic and so it doesn't produce much that's reusable or operationizable.

Are there any data science books you'd recommend reading to start getting into it?

Also what is your opinion on your job? What are the good and bad things?

[xwdaniel2803](#)

I'm not a data scientist, and so my exposure to ML and other data science is through papers and implementations in code. I like my job a lot, it's incredibly fulfilling but I worry that not everyone is having the same experience. One thing that's a double-edged sword at Netflix is its Freedom & Responsibility culture, which means I don't have the dictatorial authority I might want to make people use certain techniques.

2 Questions:

How do you combat user expectations around poor visualizations in a professional setting for simple metrics? I redesigned many of the Bloomberg visualizations years ago around financials which are exceedingly complex, but for very simple things, like dashboards, it's quite difficult to persuade

otherwise. Pushing people away from unhelpful trendlines is a beast.

I'm a principal architect and in a powerful position to augment or change directions around dataviz, but I find that I might not have read everything I need to or be aware of modern theories that I can distill for my audience (deveopers/pm/exec). What books/papers would you suggest for me to polish my end-game? (besides Tufte, Few; or maybe I'm not using these two authors appropriately?) Any feedback at all is welcome. Thanks.

[RonUSMC](#)

Typically in a dashboard environment the advice I give is to give your stakeholders what they ask for and then also give them another view into the data that is a natural extrapolation of their initial ask. You have to provide the bar chart because if you don't you're breaking their trust, but next to the bar chart you could provide a slope graph or a dot plot, which is only one step away from a bar chart. That also happens at a longer-term scale where you build up credibility with stakeholders and then you spend that credibility on some new view into the data. We introduced a connected scatterplot on one of our views here at Netflix and even though the stakeholders were suspicious, they okayed it because we had such a good track record. It ended up receiving a very positive response from some influential folks and established that as a chart we could use later.

As far as books I always recommend Andy Kirk and Alberto Cairo's books for industry types because they do a good job of speaking to design and storytelling while operating in the same sphere/language as Few and Tufte.

Hi Elijah, I've been thinking a lot about how to visualize uncertainty. Often, estimates come with an uncertainty interval but this is not well captured in a standard visualization. Have you seen any visualizations that do this well?

One example: If we see a map of countries colored by their population size, we might also be able to imagine the colors fluctuating depending on the uncertainty. If we had that, we would be able to see that, for Nigeria and North Korea for example, there is a lot of uncertainty.

[neburoc3](#)

There's a paper in my sketchy article that talks about using non-photorealistic ("sketchy") rendering to show uncertainty. People like error bars. I like to think about uncertainty and significance as two parts of the same coin, so take a look at any techniques used to show significance and think of the other side as uncertain. Hope that helps.

Do you think that Netflix moved to the "thumbs up and down" rating system specifically to mask the fact that the vast majority of its library is low quality? It seems that since changing the system, more terrible moves get high match percentages. (I say this as a happy Netflix subscriber who just gets frustrated with the completely inaccurate recommendations and ratings under the new system)

[Youknowimtheman](#)

You want one of those salacious tell-all anonymous Netflix AMAs, this one is about data visualization.

I'm an undergrad that uses D3.js in my research lab to visualize algae genomes and construct algae phylogenies-- a different use of D3 but still just as helpful and interesting. What motivated you to begin writing data visualization/use D3?

[cheese-queen](#)

I got started in GIS in grad school and slowly transitioned from making maps into more generic data visualization. One of the reasons I have such high standards for data visualization is that cartography is so much better established and introspective than data visualization is. I wish someone would do more exploratory work with cladograms in D3.

1. Do you consider data visualisation more artistic or technical?
2. If you had to tell someone with no knowledge of the field why it was necessary, what would you say?
3. What is the most interesting data set you've worked with?

[J_tt](#)

1. It's design field, so it's technical but not in an engineering sense. I think art is in there but there's less artistic expression in data visualization than some other design sub-fields.
2. We are constrained in our ability to communicate and understand our world based on the raw material that we use as language. If we can only use words and bar charts and spreadsheets, then we'll have a more limited understanding of our world than if we can also use network diagrams and other complex data visualization forms. Likewise, we can more efficiently communicate if we are all highly literate and don't have to rely on scribes to write for us, so if we were all better at reading and making data visualization, we'd be more productive.
3. I worked on the IUCN Red List back at Stanford, it really opened my eyes to how we're causing a 6th Great Extinction.

Are you going to answer anything?

[ihrtrox](#)

Hey come on, it's an "ask" me anything, not an "answer" anything.

I (PhD student in earth sciences) use a lot of R and python for data visualization. What's the benefit of js over something like [shiny app](#) in R?

[KyrgyzManas](#)

I know so little about the capabilities of shiny that I couldn't say. It seems like the difference between deployed data science web solutions and custom ones built with web technologies is that the deployed stuff doesn't give you much control over UI/UX/HCI kinds of things, which are really critical for serious analytical applications and not such a big deal if you're just sharing an interactive view among a small team.

How do you feel about Netflix changing the 5 star rating system to the thumbs up / down system?
There's a lot less you can extrapolate with just a positive/negative rating

[Nalopotato](#)

With a hundred million users all over the world doing all sorts of interesting things, reducing the dimension complexity of that one action isn't dramatically reducing the amount of data we have to work with. As far as whether or not it was a good idea, I don't know.

Do you test your visualizations beforehand to ensure people are understanding or getting the point you're trying to make?

[NotQuiteTooTall](#)

We build our data visualization in a design process, so we're constantly engaged in a dialogue with our stakeholders and listening for when they misinterpret it. That's because I'm building data visualization for experts to communicate and explore, I'm not trying to present my own research or analysis, which is one of the reasons why a dedicated data visualization role can provide a company with value.

[removed]

[\[deleted\]](#)

I found the opportunities and challenges of digital humanities work to be extremely rewarding, both personally and professionally. But it's also a particular social structure and can be off-putting to some folks in the academy.

I think we had the Challenger-type disaster already in the form of climate change, which was not properly communicated via data visualization and as a result we're heading toward 2+ degree temperature change and the resultant suffering. If environmental scientists had done a better job communicating that, especially if they'd used better charts, then maybe we wouldn't be in this mess.

Why does the landing page for your book as well as your post completely omit the name of Mike Bostock, a person without whom (principle creator of d3.js) none of this would be possible?

[mozennymoproblems](#)

I'm probably trying to stab him in the back out of a sense of selfish pride and desperate insecurity.

Interested in how you regard digital humanities as a field. How do companies see value in it as opposed to academia?

Taking critical looks at data visualization as a profession seems interesting, but also critical looks at all kinds of tech/development fields: machine learning, ai, iot, and so on. I'm interested in how you bring critique to these industries. When you bring up these topics to data viz pros, companies, and academics, are they receptive? Defensive? How do we have these conversations with engineers and investors?

[ViennettaLurker](#)

I think being able to integrate critical discourse and socio-historical themes into my practice and dialogue makes me a better designer, because I know that everything is contingent and open to interpretation and so I don't get offended when people don't understand me, or don't understand something I make or want to challenge it.

Engineers, on the other hand, are hopelessly locked in this aggressive belief that one can "win" an argument and find the perfect solution. It's a shame, really, that they're allowed to operate unsupervised.

Hi there! What are some of the best and worst ways you've seen social research visualized. For

example, if I was doing research for a Phd. in social work or sociology, what are ways I could really screw this up vs. what are ways I could use this to kick ass?

[andlaughlast](#)

Networks.

Network visualization is simultaneously the best and worst form of data visualization. Networks are so important in all of our social and cultural activities and yet we're really terrible at showing them except sometimes when we're like geniuses.

Do they types of shows watched change dramatically when a Lange even happens? Such as a natural disaster or terrorist attack or something of that magnitude.

[lcodyl](#)

You'll have to save these questions for those anonymous Netflix AMAs where they claim to tell you all the secret inner workings of the great red N.

Hey Elijah. What is your recommendation for breaking into Data Science jobs? I have the projects, I just need the experience in the form of the perfect internship as a stepping stone into Data science. What does that perfect internship look like?

[Blytheway](#)

I don't know we don't even do interns at Netflix. I found working for a library at a R1 university gave me a chance to work on really amazing projects.

I've quickly explored your Semiotic library for React dataviz, could you explain what is it for? Doesn't seem very different from existing React dataviz libraries, in my humble opinion...

Thanks for any clarification

[minchialepaste](#)

It allows you to transition easily between a lot of different chart types within a similar information model and it integrates annotations better than existing libraries. Ultimately, though, it is not "the best" and is just something we've found useful for deploying analytical applications here at Netflix.

What are some good visual representations of the relationships between large multinational corporations and the companies they own?

[77Zaxxonsynergy77](#)

I prefer the great big octopus because it implies a certain sinisterness that is accurate. Except now I've learned that octopuses aren't sinister at all, so I guess there's have to be another, more evil multi-armed creature.

As far as procedural forms, some kind of nested hierarchical diagram that combines the enclosure signal of treemaps or circle packs with the connection signal of dendrograms would probably be best.

What are some good visual representations of the relationships between large multinational corporations and the companies they own?

[77Zaxxonsynergy77](#)

I prefer the great big octopus because it implies a certain sinisterness that is accurate. Except now I've learned that octopuses aren't sinister at all, so I guess there's have to be another, more evil multi-armed creature.

As far as procedural forms, some kind of nested hierarchical diagram that combines the enclosure signal of treemaps or circle packs with the connection signal of dendrograms would probably be best.

Is DH rolling out easily or is there a lot of resistance? Do you or other DH scholars regularly sit down with the C.S. scholars to coordinate curriculums? If not, what prevents that dialogue?

I'm super interested in developing a STEM-Humanities (I am Humainities) curriculum centered on rhetoric for a high school or undergrad audience. Any suggestions for good models and/or resources

[binx85](#)

It's been a while since I was in DH. The going theory was that there was no such thing, that everything would involve "digital" but in my experience there was pretty decent resistance to quantitative and computational approaches. I can't think of any specific models but would look to particular resources like Voyant for introducing non-scientists to established techniques like NLP.

Who else beyond Edward Tufte is required reading these days? Who inspires you to do better visualization?

[technofiend](#)

I have a whole stack of books and the usual suspects are all there. Few, Tufte, Cairo, Andy Kirk, Evergreen, Cole Nussbaumer, and then because I'm an academic all the academics.

But you mean if you could only read one book? Tamara Munzner's book and then all her papers. And her tweets. Basically anything she writes is correct.

My inspiration has always been Ben Fry as far as practitioners go.

How is semiotic different from airbnbs superset? Main advantage & disadvantage?

[SlipperySteve71](#)

I only played with superset a little back when it was called caravel. We use Druid a lot at Netflix, so it made sense to explore it. It seemed more designed for quick views into data for exploration. While Semiotic has a lot in common with exploratory data analysis, it's really designed for building analytical applications and not for enabling an individual to have a quick look into their data.

What do you think of online courses like Udacity's Data Foundations and Data Analyst "nanodegree"?

[youarewrongstfu](#)

I don't know I haven't taken them or courses like them.

Hi there. I was wondering how much data plays a role in what tv shows Netflix will add and how intricately this data is used I.e. certain age groups like this style of film so we should be paying this director x number of dollars to create a specific film using this actor to appeal to this audience the most.

Do you think this will affect originality in film and tv and do you think tv shows will ultimately be written and based on algorithms.

[melbdude1234](#)

There was some amazing research out of Spain a decade ago that showed all songs really were starting to sound the same so I think it doesn't take algorithms, human beings are already really good at making boring cultural content.

Hi! Thanks for taking the time to do this. I have been trying to think up a good way to represent data in a scholarly article, where GIFs and videos are not especially welcome. The context is basically how the weather changes over time in specific areas. Right now I'm using weather map snapshots from different seasons but it can be a bit cumbersome when I want to show 5 or 6 areas over the course of 50 years. Any ideas for a better alternative?

[PurpleRhymer](#)

Show the deltas rather than the absolute, so that then you're only dealing with the change in this region or that. Look at the work of May Yuan, she's amazing, as Shirley Wu might say.

Hey thanks for doing this AMA! I'm also doing a lot of datavis professionally, some questions:

- Are you also using React for rendering and d3 for everything math related? If so how did you solve pitfalls that react based rendering contains for this sort of thing like no clear enter/update/exit loop and how to handle specific subcomponent updates. For example, in almost all d3-react charts, I notice the slow tooltips/guidelines because the whole component tree is checked if it needs to update.
- There are like a dozen great charting libraries currently: plotable, nvd3, c3js..., do you think there is a big use case missing that all of these libraries don't really solve? I always see realtime visualisations as a use case that is not really handled, but I'm also not sure if people would want a library if I would offer one.
- Do you think data centered languages like Clojure/Clojurescript or Elm offer a big advantage when working in this field?
- Any general suggestions what to look at for the future?

[T_N1ck](#)

- I try to separate the interactive bits in Semiotic into different layers from the data bits. This helps performance. React 16 will make this even better with async rendering.
- Most charting libraries that I saw were designed to solve the problems of "How do I make a bunch of charts easily" and "How do I make charts in React" which really weren't problems I was having. I needed a library for designing analytical applications in tandem with stakeholders and fellow designer/developers.

- I don't know I really only know how to code in JavaScript.
- I think all this reactive stuff is going to change things a lot. So learn rxjs even if it means you're forced to see Ben Lesh's guns on Twitter all the time.

I work for a large government agency and the program I work on collects large volumes of data but due to lack of initiative we stop at the "collecting" phase. I imagine this data would be very valuable as it directly relates to the work we do and how we serve the people. What do you recommend we do to operationalize this data?

[tiroc12](#)

Try to find someone with a concrete use case where this data can help and then focus on empowering them. My best work has always been in tandem with a researcher or a team of stakeholders who I try to enable to achieve things they didn't even imagine when we first start the project. This goes against the competitive Western "single genius" approach but I find it's particularly well-suited for big data visualization projects, where someone can either be too invested in one view of the data or too overwhelmed by the possibilities to succeed. From a more practical perspective, the best way to describe a large dataset is to provide summary statistics and then case studies where the data doesn't follow those summary statistics. That can be done through exploratory data analysis or anomaly detection. I can't get much more specific than that because datasets are all very different. If you have geodata put it on a map, they're the one form of complex data visualization that everyone understands.

How long have you worked with netflix? Also whats the experience in your career?

[LoveLunaafi](#)

I've worked at Netflix for almost three years. Before that I worked in the Stanford Library doing software development for researchers for five years. Before that, I was a graduate student studying the effects of climate change on state formation.

Why does everyone hate JavaScript? Is it dying? As a Computer Science undergrad I feel out of the loop.

[FrendlyNbrhdCanadian](#)

Java engineers are scared they have to learn a new language, also it's harder to act like some kind of arcane wizard if you're not programming a LISP machine and making your own compiler in vi with EMACS.

Wait, what? Why would it die? That's what I'm learning to do and I love it and, wait, what?

[jaysprout](#)

Well then you better work to save it. Basically, all the old people doing data visualization are angry, unimaginative, bitter, dried up sourpusses and they want to turn data visualization into a minivan. Stop them.

When making data visually appealing how do you ensure that the data does not get skewed or

changed when building the final product?

[Succotash88](#)

You work constantly with the people who are going to use it, and ask them over and over again "Hey we're not messing this up, right?"

How defensively do you or should others be designing their data visualizations? How much attention and focus is spent on ensuring information isn't misconstrued or warped down 3 links of the social media chain?

[boss1000](#)

There's only so much you can do. It's definitely present in the design, this attempt to ensure that it's not misinterpreted, but applications are living things and can be adjusted. Oftentimes, the misinterpretation comes from someone using a dashboard or viewing a data visualization and from that developing a more sophisticated view of the system being visualized so that they have the ability to misinterpret it. So misinterpretation can be a good sign.

What are some good resources for someone who doesn't know anything about data visualization to start learning about it?

[Lankshire](#)

Dig into the communication pieces by folks like Moritz Stefaner, Nadieh Bremer and Giorgia Lupi.

Do you feel like the shorter attention spans are effecting the more detailed forms of data visualization?

Personally I have zero experience in this field and at the start of your Archer link I was very amazed. Shortly after it became pretty overwhelming and required a lot of reading to understand what was being displayed.

[khaos_kyle](#)

I think most data visualization is optimized for Busy Professionals Who Need To Make a Decision Now. That's definitely a prominent audience for data visualization but the field suffers when it pretends there are not different levels and kinds of engagement with our work.

I've made some basic data visualizations with D3 but I almost always end up making the graph or chart in a 3rd party application like Excel and then simply save and upload it as an image. It's quicker and less clunky. What kind of use cases do you see D3, or SVG graphics in general, being superior to simply exporting and uploading an image from a data visualization application?

[DataGuru314](#)

When you're trying to build a different visual metaphor from the ground up or you're doing combinatorial work. Otherwise, the only reason most people make everything in D3 is convenience or to maintain their skills.

Your favorite data visualization by far? What are some things you see in some data visualizations that

you think people tend to forget to include or can be improved in general (excluding the fact that it sounds like data visualization is dying :(thats sad to hear but maybe we can change that)

[okBroThatsAwkward](#)

I hope people can change it, I like data visualization. My favorite data visualization is Ben Fry's visualization of the changes over different editions of The Origin of Species. I think it's elegant, important and thoughtful, three things 99% of data visualization being produced today is not.

D3 is absolutely fantastic. I've baked it into several control panels and visualization tools I've built over the years. Thank you for the incredibly performant visualization platform.

Now to my question. As an automations and software engineer I've found that I'm building platforms to make the decisions in place of a human being. Once the patterns are identified and the business needs are properly accounted for the actual visualizations aren't necessary or are relegated to a secondary role where they're only used to retarget the core algorithms.

What is your opinion and perspective on the higher emphasis of machines and automated systems analyzing the data compared to a human being visually processing it?

[Ganondorf_Is_God](#)

One of the major initiatives we're exploring now is in anomaly visualization. I think real innovation in data visualization will be done in integrating reports and schematics and anomalies into more sophisticated diagrams of systems that are run and optimized by algorithms. Once the computers can read and make those, though, I plan to retire to a goat farm.

Elijah, I feel that I understand how to study things like algorithms, data structures, languages, math and so forth, but what does it take to get a deeper understanding of how to create a visualization that is richer than just your average 2d plot?

[KDallas_Multipass](#)

I really started to understand the challenge of data visualization when I finally understood gestalt principles. Shapes and signals via shapes (and color and animation) are not easy but they're also not just artsy unstructured stuff, either.

Hi, thanks for this valuable AMA. You said your background is Philosophy, which of course gives you wider perspective on some things and mechanisms. But was it hard to learn JavaScript and use/learn mathematics/statistics or you had this knowledge before?

[proce55or](#)

Yes, and I'm sure I'm not a great programmer as a result, but we've reached an inflection point with programming where being a great programmer is less important than being good at design and understanding the needs of stakeholders.

The D3 code base has had almost no outside contribution since 2015 giving it a pretty small bus factor. How important do you think that is to the long term health of the project?

[gulup](#)

I think as the information visualization kernel for js projects, nothing is going to touch the core aspects of D3. I assume the other bits about DOM manipulation, event management, formatting, et cetera will all eventually recede in importance.

Can you show us any of the statistics/data you are looking at from Netflix? Would be cool to see :)

[gust1609](#)

No, of course not. If I'm any good at my job at all then any data visualization products I build would be just the kind of thing Netflix wouldn't want me showing of on Reddit, right?

Your favorite data visualization?

[reddit_lonely](#)

Ben Fry's visualization of Darwin's Origin of Species. It's the best data visualization ever made.

Having no DV skills myself how can I use today's technology to build images. Is it better for me to seek the right person or the right software or a combination of the two?

[mydaddyisadrunkass](#)

I think finding people is always better than finding technology. Real knowledge is made socially, not by individual geniuses, that's just a myth so someone can be on top.

I'm not sure if your edit was in jest or not, but I hope Sense8 gets renewed! Was there any data visualization in terms of viewers compared to other Netflix Originals that went into the initial decision to cancel Sense8?

[cox-in-sox](#)

It was entirely in jest I don't have any authority on content decisions.

As an undergrad graduating with a Data Science degree, what types of jobs should I apply to?

[midnightbanana35](#)

At Netflix data scientists typically have advanced degrees, whereas analysts don't always. I can't speak to the rest of industry, though.

Sorry Noob question, but can this be extended to 3D objects/framework? Thanks.

[tcdoe](#)

Check out DeckGL for awesome 3D data visualization.*

- I mean it's still basically 3D pie charts, right?

What do you think is most important in conveying information clearly?

[dovahart](#)

Annotation. And not just axes or legends, but little arrows and circles and inline explanations.

Why can't i use english subtitles in europe? Can you visualise that data for me? :(

[finalflash42](#)

I don't know I'm in America!

I'm Mr. Meeks, LOOK AT MEE

..sorry I bet you get that a lot now

[Mysteroo](#)

It's a good look.