Workshop Discussion Notes: Interpretation Gone Wrong

The Social, Cultural & Ethical Dimensions of "Big Data"
March 17, 2014 - New York, NY
http://www.datasociety.net/initiatives/2014-0317/

This document was produced based on notes that were taken during the Interpretation Gone Wrong workshop as part of "The Social, Cultural, and Ethical Dimensions of 'Big Data'". Not all attendees were involved in every part of the conversation, nor does this document necessarily reflect the views and beliefs of individual attendees. All workshop participants were given workshop materials prior to the event to spark discussion. This primer can be found at: http://www.datasociety.net/pubs/2014-0317/InterpretationGoneWrongPrimer.pdf

Overview

The range of discussions in this workshop was broad and varied, including the use of statistics to misrepresent what the data tells us and other sources of anxiety in data interpreter communities; the difference between what data is perceived to represent, as opposed to other forms of representative realities, like a painting; the idea that numbers are better than nothing, such that even if data has a known bias, or is limited in its professed or implicit explanatory scope, the public will still consume it readily as proof that some notion has an underlying real value; the desire for there to be a point, or several points, of accountability such that when there are consequences to misinterpreted data, people cannot evade responsibility; the challenges associated with creating points of accountability at the levels of data collection, analysis, and interpretation; the question of who bears the burden of responsibility for data interpretation more generally; how we could think about this responsibility in terms of educational initiatives and tech literacy; and the possibilities for exploring and evaluating interpretation pathologies.

The tensions that emerged from discussions on the challenges of addressing interpretation gone wrong ranged from the universal – such as how people aren't necessarily swayed by data that challenges their beliefs – to specific situations or sectors where there is heightened concerns about interpretation gone wrong, namely in healthcare. Many of the discussion points raised the twin issues of truth and accountability, such that the desire for accountability underpins the desire for truth (or expertise). These discussion points evolved further dichotomies about how information can be repackaged to outwit current regulatory frameworks for protecting data disseminations, while also reflecting a desire for regulation, or guidelines and best practices, that can anticipate the challenges of data-driven products.

Themes and Discussion Points

Statistics and Skepticism

The initial theme of the discussion on how the interpretation of "big data" can go wrong centered on the fallibility of statistical rationales, insomuch as statistics can be used to mislead as readily as they can be used to elicit truths from piles of information. Statistics has garnered a tongue-in-cheek reputation for being the science of producing unreliable facts from reliable figures; similarly, when it comes to "big data", it can also be the science of using select datasets to create robust-looking figures. However, statistics was also credited with being the grammar of science, and a surefire way to establish patterns. The more challenging aspect is ensuring that those patterns are relational in meaningful ways, rather than spurious connections between unrelated items. This uneasy tension between truth and manipulation, or information and knowledge, led the group to discuss some of the anxiety in data-interpreter communities, and amongst groups of experts more generally. Some participants mentioned that many disciplines have access to data, and readily use it, without having a firm understanding of statistical methods or training in them. What's the problem with 'bigger' data? Some members of the group discussion posited that perhaps the issue with "big data" is that everyone wants a piece of it, without necessarily understanding what it is or what to do with it.

The group discussed how people tend to mistake data about "the thing" for "the thing." By contrast, nobody mistakes a book for "the thing," or a painting for "the thing." They understand that books and paintings are representations of reality. What's the difference? Data are measurements of something - is this about measurement, or is the relationship itself more direct? Some participants mused how this confusion may come from the appearance of data objectivity. One thread of the discussion offered that numbers appear more transparent and less mediated, and thus carry more weight, especially with the public. Some participants suggested that data is, in fact, less mediated than a painting. For example, the external temperature is directly correlated to mercury levels, and can be accurately interpreted. However, the idea that everything is quantifiable is leading to a new complexity of everyday objects. While it may be simple to trust that a thermometer just 'says' the temperature accurately, it's harder to assess the factors that go into a ranking of the best colleges when what it says is a function of a mysterious algorithm. In other words, not all data is made equal, either in its collection, analysis, or the intentions behind its actionable use. Some participants thought about how people make decisions based on data even if that data is disputable because, the logic goes, something is still better than nothing.

Tech Literacy and Education

Some of the major questions of the discussion centered around the question of, what is the role of tech literacy and education in bringing greater understanding to data interpretation? Peer reviews, accountability systems, credibility, and generally, a guarantee that the information being disseminated has some integrity to it is fundamental to the public trust in data and data-driven decision making processes. Several participants voiced the thought that if you offload data-interpretation to individuals, the organization or experts responsible for data use and interpretation evade a responsibility that should be theirs.

Some members of the group asserted that not everyone wants to be educated on data interpretation, and even if they did there are cognitive considerations and ingrained biases that would limit a sort of collective capacity to understand things, but most people do seek to hold experts or decision-makers accountable for misinformation. But, some participants asked, do the tools that allow for number-crunching to happen on a large scale (i.e. computing, algorithms, etc.) create a way for experts to evade responsibility for the information that is generated as a result of the numbers inputted into the system? One example raised in the discussion was about the 9/11 Museum. ¬¬Employees of the Museum were directed to say that the output footage and notes produced by the museum that ostensibly represented visitor viewpoints. If employees were asked what the criteria were for choosing the particulate footage or notes were instructed to say that the selection of representations were algorithmically created, thus diminishing liability over the selection process. However, the employees understood that the algorithms were themselves created typically by the staff. The tension here is still about having a real point of accountability, rather than questions about objectivity.

Evading Responsibility

One of the workshop themes centered on public concerns regarding the evasion of responsibility. What happens when badly interpreted data is networked into the public understanding of something? Is there some basic level of regulation that could create a stable system for data interpretation, like requiring private corporations to have expert oversight in some capacity when they develop algorithms? The key issue that emerged was what are the ways in which people who are experts can serve as educators, so that the good and bad sides of massively collected data is communicated, instead of just fears?

In terms of accountability, some members of the group discussed the possibility of publishing things like code to demonstrate how data is used. However, that still leaves just a few experts to grapple with something that few people will understand. With regard to regulating private companies and their processes of data interpretation, the discussion centered around whether or not this was even possible. Some participants offered the case of 23&Me as an example of a private company that could be regulated, to some extent, because it claimed to disseminate personalized health advice based on DNA swabs collected from its consumers. It did this by acting as an unregulated diagnostic test. Some members of the group further offered that the company didn't let outside experts in to examine its data, which brought its credibility into question.

Other participants focused on the claim that the company made to its consumers is problematic because of how health advice is packaged or labeled affects its interpretation, and non-physicians are not supposed to be offering personalized medical advise using unregulated diagnostic tests. Can the label, package, or framing of the issues that 23andMe's product tries to address allow it to disseminate quantified-self information in a way that doesn't violate FDA regulations on diagnostic testing? The issues discussed also broached the question of, who has a right to say that the consumers should not get the service about their own genome? The group discussion went on to debate the pros and cons of applying higher standards or requiring data interpretation training to some sectors, specifically healthcare, where people's physical lives hang in the balance of strong data interpretation skills. One of the questions that emerged from the discussion was, when and why do we tolerate error, and in which sectors is it more acceptable? How much does techno-positivism accept our tolerance for error in the marketplace of information?

Who Bears the Burden of Interpretation?

Another key theme of the workshop revolved around the question of who bears the burden of interpretation? Some members of the group thought about having a certification program whereby datasets are certified for clean collection methodology, although other participants highlighted how this would only apply to collection, not to analyses or further distribution. It was noted that the Open Data Institute has a set of four self-certification standards. Other participants mused on the pros and cons of journals that could publish negative results, or retracted results, so that data-driven solutions have a disclaimer naturally, and accessibly, attached to them. Some participants spoke about how that might be particularly relevant in instances where academic work that uses a small sample size, and has disclaimers that it shouldn't be used for policy, is still co-opted for policy because there's a number attached to it. The key issue was, what are the ways that we can see the missing data?

The issue of missing data fed into a secondary discussion on what best practices might look like in this area, although the main tension that became evident from this discussion was less about the guidelines that surround data collections, and more about its interpretation. In a larger sense, the group discussed what the pathologies of interpretation might be. By contrast, on a smaller scale, some participants expanded the discussion on education to think about teaching statistics to schoolchildren. This rounded the discussion out to its beginnings, whereby the bias inherent to any assessment of data is difficult to account for. Even transparency doesn't fix this sort of problem, as some group participants offered. People don't necessarily change their minds when given evidence of gaps between their preconceived beliefs and data, even high-quality data. This part of the discussion resonated back to the idea of exploring pathologies of interpretation and whether or not tech literacy would assist individuals with making accurate inferences from data. Does the accuracy or transparency of a data-driven result matter at the level of individual interpretation of a specific instance or event?

In the main takeaways of the workshop, the group discussed broader questions about the qualities of data and the people using it. What is the shelf life of data? How can data brokers be regulated? What are the ways in which people who are experts can serve as educators, so that the good and bad sides of massively collected data is communicated, instead of just fears?

Further Exploration

The workshop discussions reflected a number of significant tensions that often revolved around the notion that the numbers don't necessarily speak for themselves: the data are represented by numbers, and figures are considered more reliable than other items in asserting a fact or a truth. There is a great deal of tension about how numbers can misconstrue as well as illuminate points of probable fact. How do we ensure integrity, both in the selection of data sets and their analysis? This is especially important if a wide range of sectors, from government to media to healthcare is going to develop policies around the "big data" phenomenon. Some participants offered that much of the discomfort with data interpretation seems to come from either not having a singular person or organization responsible for the interpretative process of data analysis (i.e. a faceless, powerful entity), or from not having enough expertise available to explain or challenge data interpretation. Who should be monitoring or auditing processes of data interpretation? On a societal scale, how do we ensure that all groups of people are represented in data collections, especially those that drive policy decisions, so that we can account for people who might otherwise be rendered invisible? What would broader public education on data-interpretation look like? What are the pathologies of data interpretation, and how do we assess them?