



Workshop Discussion Notes: Inferences & Connections

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 Inferences & Connections 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/InferencesConnectionsPrimer.pdf>

Overview

The workshop on Inferences and Connections generated a great deal of discussion on two major issues: who or what is reading inferences or making connections; and when it is appropriate or inappropriate to make those connections. The underlying theme of the discussion is that drawing inferences and connections based on “big data” is less about the “big data” and more about power dynamics. That is, who is holding sway over whomever else by reading inferences into the data? Which communities have a say in how data is collected, used, or repurposed? Where is the human intermediary who draws a particular set of conclusions at the data-analyst level? Similarly, what are the harms or benefits to an individual or user who is reading an inference into the products presented to them as a result of data-driven associations? One of the points that evolved thematically in the discussion is that “big data”, as a phenomenon, can obscure or make difficult to interrogate more problematic heuristics.

The other major focal point of the discussion centered on how data-driven inferences are more risky on some domains than in others. Can we identify areas where data-driven inferences are more useful than in others? The group discussed how to weigh the risks and benefits of data-driven inferences in different domains, like medical research, education, and policing. There was an echo of ‘we need a comprehensive framework for thinking about this’ that characterized many of the discussions, while some participants also acknowledged that context and specificity is fundamentally important for understanding risks and benefits. Some participants brought into question how truthful any inference can really be based on a stray or composite of stray pieces of

data. This led to a discussion beyond specific domains, and centered on how the data environment influences the data that is produced or made visible about people, and how correlating it with other patterns in human behavior is challenging.

Themes and Discussion Points

Human in the Loop

In the initial discussion, some members of the group observed that the need to have a human in the decision-making process is evident, but that humans are also a huge source of bias. Similarly, human and machines have complementary capabilities. To what extent can a human actor decide their own course of action with a certain set of values and norms embedded in their decision-making process, and in which ways are they compelled to act? A human “data-thermostat” might present a useful point of accountability, and some participants sought out precedents, alternatives, or elaborations on this metaphor. In law, there are tools to determine when there is a liability attached to an inference, such as evidentiary standards. Several participants acknowledged the difficulty with applying a similar expectation to determine liability with regard to an inference is that it is difficult to know what causes the inference to be made, particularly by a machine that may be using proxies for, rather than direct and taboo variables, like race, gender, sexuality, etc. To exemplify the point, some participants addressed the Grindr case, where a man using that app was shown a link to a sex offender search app, as though they were naturally related or associative interests. Perhaps that link appeared because of some statistical correlation, but we don’t know if the correlate evolved from language, content, downloads patterns, or some other variable.

It’s clearly challenging to untangle algorithmic judgments transparently. Some members of the group questioned if we could create an audit structure for this. To the extent that inferences are applied to make decisions about individuals that could harm them or others, some members of the group suggested that there could be more scientific validation or rigorous examinations around more sensitive sectors, like policing, education, or healthcare. These inferences happen at the interpretive layer as well as the technical layer. For example, with regard to 23andMe, it’s clear that there can be transparency at the data level, but lingering problems with interpretation when it hits the public.

In the medium term, there is a place for people to act as mediators. Some participants offered that the public, and people generally, are comfortable with a person making an inference based on data. So, if you can continue to have an accountability

point where there's a person who's in the system, and you can apply auditing or oversight there, that might be a more acceptable way to proceed with "big data" developments. The fear around computational inferences has to do with programs like Watson, IBM's stand-alone group for developing big data solutions to a range of solutions by using what it calls cognitive technology. The cognitive element has to do with its claim that its algorithms are on a constantly improving learning curve that improves when they are fed more and more data. The group discussed how the unease around Watson has to do with the probable fact that no-one on the outside knows what is going on inside the Watson group. This example brought the group discussion back to the desire for a human in the loop has to be acting as a human with subjectivity, not an algorithm. We need a person with degrees of freedom to depart from data and organizational algorithms. Some participants pointed out that, historically, we haven't seen many examples of that going well. For instance, if you ask a person to be transparent in how they're making their decisions, it's very hard to get a straight answer, and we'll have the same problem with human-like intelligence machines.

Some participants pointed out that we couldn't assume that there would be a static 'human.' Other participants discussed how marginalized communities might be invisible or uninvolved with the mainstreaming of data analytics. Who will be in the room with data analysts, algorithm makers, or other stakeholders to ensure that the datasets are comprehensive reflections of diverse communities? At a certain point, this is a question of design. The group discussed how there's not a clear beginning or end to data collection, which makes it very difficult to put a human in the loop beyond a quality control person. Data collection is always evolving and always generating more data. However, it is always appropriate for someone to periodically audit or validate that an algorithm is working the way it is intended or supposed to.

The group also discussed how you could have explicit inferences embedded in algorithms to bring about a desired level of transparency, but as some participants pointed out, this might be a complementary way of eliciting accountability rather than a replacement for the human in the loop. It applies readily to computations, but if you ask a person if they considered gender, for example, in making a decision, they can just lie. With an algorithm, you can know if it's directly discriminating, and then you can look at disparate impact and proxies, and you can attack the problem statistically. If the decision-making tree is fully inside a person's head or a smoke-filled room, it's very hard to know what is going on. It's not easy to fix an algorithm, but there are ways forward. If it's a machine learning algorithm, it's potentially harder to apply this sort of explicit inference approach, but there are actually more tools than you might imagine, according to some participants. There can also be algorithmic ways of validating the algorithms.

Some members of the group also discussed the concept of cyborg epistemology. (For some research on what this entails, see [here](#)). Namely, what is it good for? When does it need to be filtered or tested by human epistemology? Also, the group observed that simple global problems could potentially be solved using these approaches. For example, instead of having “data shields” against eating unhealthy food (apps that tell us what to eat), maybe we should be paying attention to regulating the food industry. Thus, a discussion about the social opportunities of “big data” offer could also carry an implicit opportunity to de-politicize certain domains, and re-approach problems in those areas using technical innovation.

Methodological issues: feedback loops and sampling

Some members of the group discussed how the datasets we have are very uneven, which affects the kinds of inferences that can reasonably be drawn from them. We tend to think big = good without thinking about how the dataset is sampled or who it leaves out. With regard to the ethical, social, or cultural obligations or responsibilities of individuals when they are faced with the choice to opt in or out of data collection, some participants offered that when you opt out of data collection, you opt out of being in the sample. Thus, they reasoned, there’s almost an ethical imperative to participate because opting out makes you invisible for decision-making and analytics. In previous approaches, datasets were designed with more specificity, and they gave you more ability to think about where the data was coming from. Now it’s coming from who knows where? You can’t do as much up front thinking about issues of sampling and interpretation. One question that arose in the discussion asked, is there a place for an algorithm in the loop with regard to sampling methods? For example, can an algorithm incentivize the collection of data along particular lines?

When the discussion turned to ideas about feedback and feedback loops, the human element re-entered the group discussion. Inferences feed back into reality; people are consciously or unconsciously sensitized to inferences, and thus they often act as prophetic, in that people will select from the options inferred to apply to them. For example, one of the fire-starter’s in the morning session referenced how her daughter was using a reading app on an iPad, and the book selections were all labeled with pink tags. She wondered if the inferences imputed to her daughter resulted in gender-silos in consumerist settings. This complicates the idea of calculating the accuracy of inferences. The idea of high versus low accuracy is troublesome because it implies that there’s a system for evaluating this, when in fact there are feedback loops. For instance, we don’t know what else girls could like in their reading collections besides pink things. The accuracy of the prediction is not independent from the work that the prediction does.

And then, what are the social costs? It's very hard to calculate probabilities for one-off events. There may not be an independent standard for checking that.

In some contexts, having specified inferences can be helpful, such as with drug treatment plans, and less helpful in others. Several members of the group were concerned with thinking about what is motivating the nudge or prompts in the feedback cycle.

Risks and Benefits

The discussion re-oriented towards thinking about the risks and benefits of using algorithms to make inferences and connections about people based on large amounts of data gathered on them, or that relates to them, from a multiplicity of sources. Some participants articulated the need to compare those within different domains, and how challenging it would be to have an overarching framework for doing. The more specific way to address this calculation is by adding context into that mix. This applies both in data collections, initial applications, further uses, etc. The group often returned to the idea that this balances shakes out differently in different contexts (e.g. an Amazon recommendation versus a criminal justice matter).

The group discussed the domain of education as an example that kind of complexity - the data collected on students can be used productively to improve their education, but there is lots of potential for exacerbating existing inequities in student learning experiences. To make matters more problematic, the risks and benefits within the educational population are different. Student data is all considered equally risky in terms of disclosure; once it's aggregated, however, it's a different story. How are individual student records protected when there are tools for de-anonymizing metadata? What constitutes an educational record, exactly? Though, as some participants observed, what constitutes an "educational record" is also changing (directory information? metadata used by educational app makers?).

Some members of the group also considered how this plays out in the healthcare and medical research domain. For example, 3% of the genome budget is set aside for looking at bioethics issues. There is an attempt to map out the conflicting values and trade-offs in genomics, or in other words, to sequence the ethical, legal, and social implications around privacy, and the potential for discrimination. For instance, there is a federal prohibition on excluding people from health insurance plans by using genetic information, but this is an incomplete resolution. In another example of domain-specific concerns, some members of the group discussed how several banks, like J.P. Morgan and Chase, are analyzing their financial data to uncover human trafficking operations. But what if people are mis-identified as traffickers? Or what if this results in a

crackdown on undocumented workers instead? Some risks result from low accuracy inferences, and some from high accuracy inferences. The group discussed how we map out the connections between advantages and risks systematically – how do they correlate? Some participants discussed a general desire for a comprehensive framework for thinking about this.

One of the recurrent threads that emerged from that discussion is the idea that people are not rational deliberators, a topic that elicited a wide variety of questions. You can give people information, but they don't necessarily make rational inferences. This brings into question both the decision-makers, and the ability of anyone to maximize the knowledge they gain from data access. What's the value of having empirical data at all? If we had unlimited empirical data, would we be able to solve all problems? How trustworthy is any given instance of 'data'? For example, what people say on Facebook may not correlate well to what they would confide in their physician, or to a confidante. Some people like to have all of their own data, and others only like to have what is of immediate use to them. How does the notion of social acceptability and propriety figure into the bites of data that is collectable on people? What kinds of truths can be inferred from mere visibility of some information on specific platforms? In that discussion, there is an implicit concern with how people trust in particular institutions, organizations, or platforms.

The context in which people feel like they're engaging in a data environment may be different than it really is. Does a tainted platform impair the integrity of the data collected from it? How comprehensive can "big data" really be? Is part of the buzz about it more about having a new silver bullet to continue the march of progress that scientism promises? For example, just because Amazon knows you're more likely to click on a book that it highlights in association with a purchase you already made, does it know how you would respond in a different data environment? People tend to strongly care about who is giving them information because they understand that what they are capable of knowing is limited. They trust in physicians to evaluate their blood pressure, for example, but also expect doctors to know up-to-date scientific research that is applicable to them.

On a less abstract scale, there are rational reasons to doubt the data, too. A good data scientist can manipulate data to say a particular thing by setting the levers correctly. It's hard to figure out when someone is doing this, and this happens in scientific research, too. In the same vein, it's also important to recognize the technology shift. Some participants offered that in Computer Science, you used to have to structure the data in a certain way, and now you don't have to do it that way anymore. The technology is so promising that people say let's throw it all in there, which reflects a certain fallible notion that we can know everything if we have enough data.

Further Exploration

A lot of the discussion focused on the ongoing debates about the more negative possibilities for “big data” to cleave wider gaps into existing inequalities, either by leaving out certain groups from data sampling, or from not having appropriate auditing mechanisms in place for evaluating the impact of inferences and connections being wrought by algorithmic associations. However, part of the discussion addressed one overarching concern that isn’t always explicitly included in the debates on inferences and connections, which has to do with distrust in data. That is, many people don’t even buy into the scientific method or use of data to drive decision-making. Is it rational to make decisions on the basis of data that is collected in large quantities from partially nebulous sources? How much do we want an expert intermediary to make interpretations for us? Many stakeholders don’t buy into the utility of data as having any policy force in our social contract, as some participants voiced. Other participants offered that this is mainly a power issue that can be explored further. Who gets a say over how data are used? Who are the communities? Who gets to decide for the individual what the risk profile is? Does the organization or the law determines what you can or can’t consent to? What does big data replace? What are its alternatives?