FTC Big Data: A Tool for Inclusion or Exclusion? Workshop
September 15, 2014
Segment 1
Transcript

EDITH RAMIREZ: Thank you, Tiffany. And welcome everyone to our new facility, those of you who haven't been here before. I want to thank everyone for joining us here today for our workshop, Big Data: A Tool for Inclusion or Exclusion. And I also want to take this opportunity to think that Tiffany George as well as all the other FTC staff members who've worked so hard to organize today's event, and also to thank the speakers for sharing their expertise with us.

We're at a pivotal stage in the information age. Thanks to smartphones and smart meters, wearable fitness devices, social media, connected cars, and retail loyalty cards, each of us is generating data at an unprecedented rate. In 2013, it was reported that an astonishing 90% of the world's data was generated in the two preceding years. Today, the output of data is doubling every two years. Advances in computational and statistical methods mean that this mass of information can be examined to identify correlations, make predictions, draw inferences, and glean new insights.

This is big data. It has the capacity to save lives, improve education, enhance government services, increase marketplace efficiency, and boost economic productivity. But the same analytic power that makes it easier to predict the outbreak of a virus, identify who is likely to suffer a heart attack, or improve the delivery of social services also has the capacity to reinforce disadvantages faced by low income and underserved communities.

As businesses segment consumers to determine what products are marketed to them, the prices they are charged, and the level of customer service they receive, the worry is that existing disparities will be exacerbated. Is this discrimination?

In one sense, yes. By its nature, that's what big data does in the commercial sphere. It analyzes vast amounts of information to differentiate among us at blinding speed through a complex and opaque process. But is it unfair, biased, or even illegal discrimination? And if so, can steps be taken to level the playing field? Those are the questions we'll be exploring today.

Big data, in its 21st century form, is at an early stage. We have the ability to shape its development and its outcomes. If we are alert to the risks presented by big data, we can take steps to guard against them. We can help ensure that big data can be a tool for economic inclusion, not exclusion. That's the weighty subject before us today.

But before we begin the discussion, I'd like to address three questions. First, how did we get here? Second, what's our aim with today's program? And finally, where do we go from here?

Let me start by tackling the first question, how did we get here, very literally. Whatever mode of transportation you used to get to this workshop, there were apps or connected devices available to assist your commute. Those of you who came here using public transportation may have availed yourselves of apps to tell you when the next bus or train would arrive. If you came by car, you may have benefited from GPS technologies that gave you directions, sent you real time traffic alerts, or allowed you to summon a taxi or driver by tapping on a smartphone app. And for the virtuous among us who biked or walked here, you may have used a wearable device to track the distance traveled and calories burned.

No matter your mode of transportation, once in the vicinity, an app or a website may have helped you to find a spot nearby to buy a cup of coffee before arriving at the workshop. These various devices and services that help many of us get here physically are also what brought us here figuratively. The popularity of smartphones and other mobile devices, the array of mobile apps we have at our fingertips, and the burgeoning internet of things phenomenon more generally means that countless individuals actively and passively generate information in an extensive ecosystem throughout the day.

The proliferation of connected devices, the plummeting cost of collecting, storing, and processing information, and the ability of data brokers and others to combine offline and online data means that companies can accumulate virtually unlimited amounts of consumer information and store it indefinitely. Using predictive analytics, they can learn a surprising amount about each of us from this data.

While powerful algorithms can unlock the value from immense data sets, their ability to draw correlations and make fine-grained distinctions also raises the prospect of differential treatment of low income and underserved populations. This is a risk suggested by the commission's recent report on the data broker industry.

The commission's study of a cross-section of nine data brokers revealed that data brokers aggregate online and offline data from disparate sources to make inferences about consumers' ethnicity, income, religion, age, and health conditions, among other characteristics. As the FTC and others have found, some brokers create segments or clusters of consumers with high concentrations of minorities or low income individuals.

There may be legitimate reasons why businesses would want to sort consumers in this fashion. But the practice also raises the possibility that these segments will be used for what I've called discrimination by algorithm, or what others have called digital redlining. We heard these concerns this past spring at the FTC seminar on predictive scoring. There are now products beyond traditional credit scores that purport to predict or score everything from the chances that a transaction will result in fraud to the efficacy of sending consumers catalogs and the best prices to offer consumers.

Some speakers lauded the benefits of such predictions, emphasizing that they enable the personalization many consumers want and help minimize the risk of fraud. But other speakers worried that certain predictive scoring products could fall outside the reach of the Fair Credit Reporting Act and the Equal Credit Opportunity Act despite having an impact on consumers' access to credit, housing, employment, and insurance.

For example, if a company lowers my credit limit based on a score that reflects my own credit history, I would be entitled to certain protections under the FCRA. If, however, the same company lowers my credit limit based on the scores of a group of which I am a member, the application of the FCRA may be less clear. Will these scores be used in ways that influence the opportunities of low income, minority, or other populations to get credit, jobs, housing, or insurance in ways that fall outside of the protections of at FCRA or ECOA?

Could the use of geographic information such as ZIP codes, for example, lead to Americans in low income rural neighborhoods being charged higher prices? And if so, is this a worrisome function of big data, or just a continuation of age old pricing practices and market forces?

These and other issues figured prominently also in the White House's wide ranging report on big data, which squarely raised the concern that large scale information analytics will be used for disparate or discriminatory outcomes for certain consumers, even absent discriminatory intent. It's these questions and concerns raised by these prior initiatives that bring us to today's program.

And to my second question, what is our goal today? We'll explore whether and how big data helps to include or exclude certain consumers from full opportunity in the marketplace. And to help shed light on this issue, we convened experts from industry, consumer, and civil rights groups, academia and government, all of whom are representing a wide variety of perspectives. Our panelists and speakers will provide us a framework for our conversation today, assess current big data practices in the private sector, discuss possible developments on the horizon, present pertinent research, and offer potential ways to ensure that big data is a force for economic inclusion.

It's my hope that our participants will discuss in depth the benefits and risks of big data to low income and underserved populations. On the benefits side, let me start the discussion with one example. New York City is developing a tool that combines addiction data with emergency shelter admission information and other data to predict when individuals or families are on the brink of homelessness.

Using this information, the city's able to deploy social workers to help these families and prevent them from ending up on the street. This is an example of positive government use rather than a business use. I hope our speakers will provide examples showing how companies can also use big data to benefit those in low income or underserved groups.

And as for real world examples of possible risks, let me cite a study conducted by Latanya Sweeney, who's here from Harvard serving as the Commission's chief technologist. Professor Sweeney found that web searches for distinctively black names were 25% more likely to produce an ad suggesting the person had an arrest record, regardless of whether that person had actually been arrested than web searches for distinctively white names.

This could have devastating consequences for job applicants and others, by creating the impression that the individual has been arrested. While the research did not establish why the algorithm yielded these racially disparate results, it does provide a concrete example of how an algorithm may have adverse repercussions for a particular population. I suspect we'll hear more illustrations today, including from Professor Sweeney, who will be presenting results of a more recent study.

After we conclude our workshop, the question naturally arises, where do we go from here? We may all have an array of apps to guide us home when we leave this afternoon, but there's no clear path for navigating the use of big data in a way that advances opportunities for all consumers while diminishing the risks of adverse differential impact on vulnerable populations. We may not yet know what the best course ought to be, but I believe we should have at least three objectives going forward.

First, we should identify areas where big data practices might violate existing law. Where they do, the FTC is committed to vigorous enforcement of the law, as demonstrated by cases such as our recent action against Instant Checkmate, a website that promoted some of its background checks as tools for screening tenants and employees. The FTC alleged that Instant Checkmate did so without regard for the FCRA, and we obtained a $525,000 fine and a permanent injunction against the company. In addition to helping the FTC and others to enforce existing laws, today's program should also help identify any gaps in current law and ways to fill them.

Second, we need to build awareness of the potential for big data practices to have a detrimental impact on low income and underserved populations. I'd like today's program to help foster a discussion about industry's ethical obligations as stewards of information detailing nearly every facet of consumers' lives.

Third, and relatedly, we should encourage businesses to guard against bias or disparate impact on low income and vulnerable populations when designing their analytics systems, algorithms, and predictive products. A good example is the Boston Street Bump app highlighted in the White House Big Data Report. Like any big city, Boston has its share of potholes and faces the ongoing challenge of staying on top of street repairs. To help address the issue, the city released a mobile app residents could use to identify potholes in need of repair.

But the city also recognized that because lower income individuals are less likely to carry smartphones, the data might skew road services to higher income neighborhoods. They addressed this problem by issuing the app to road inspectors who service all parts of the city equally and used the data gathered from the inspectors to supplement what they received from the public. This illustrates how consideration of risks before launching a product or service could help avoid them.

So big data can have big consequences. Those consequences can be either enormously beneficial to individuals in society or deeply detrimental. It will almost certainly be a mixture of the two. But it's the responsibility of the FTC and others to help ensure that we maximize the power of big data for its capacity for good while identifying and minimizing the risks it represents.

As we navigate the transformative terrain of big data, it's vital that we work to ensure that technological innovation benefits all consumers, whatever their backgrounds. I look forward to hearing the thoughts and ideas of our panelists on how to do just that. And I thank you all for your contributions to that endeavor. Thank you.

[APPLAUSE]

EDITH RAMIREZ: Let me hand it back to Tiffany.

TIFFANY: Thank you, chairwoman. We will now begin with our first presentation, Framing the Conversation, which will be led by Solon Barocas, a postdoctoral research associate at the Princeton University Center for Information Technology Policy.

SOLON BAROCAS: good. Morning. Let me begin by saying how thankful I am to be here. I really appreciate the opportunity to speak with you all, and I particularly want to think Catherine and Tiffany for putting together what I think will be an excellent day.

I am Solon Barocas. I'm a postdoctoral fellow at the Center for Information Technology Policy at Princeton. And I'll be presenting today what I hope will be a way of framing the conversation today and hopefully going forward as well. This draws on some of the work that I've been doing, and well, OK. And I encourage people who are interested in what I'm presenting to take a look at my website, where you can find this paper, if you want to follow along while I present, in more detail. Let me begin.

OK, so big data, we've come I think to know these three V's as a common definition, that the volume of data is exploding, that the velocity at which the data is accumulating is increasing, and the variety of formats of data is also likewise proliferating. This is a useful definition.

But I tend, I think, to focus instead on the traditional categories from the social sciences, observational data, what we might call self-reported or user generated data, and experimental data. And what I mean by this, then, is that there are actually three rather different things happening here, all of which have interesting consequences for consumer protection. One is that there are many more ways to actually observe consumers and consumer behavior, things like transactional data.

But of course, we can now think of things like mobile phones and various health devices. Self-reported user generated being the vast variety of social media that people use, and finally, the experimental, which I think has now become slightly more familiar to people in the wake of this Facebook experiment that got a fair amount of press.

What I mean by that is there are now platforms upon which to perform large scale experiments in the wild in ways that were basically impossible before. And I think these are the useful ways perhaps to think about it. For our purposes today, however, I'm going to focus on what I call data mining. This is the more traditional term from both industry and the academy, which is in some ways what we might call a subfield of machine learning, which is a field within computer science that is devoted to the automated computational analysis of large data sets.

And again, I focus on this in part because I think, for our purposes today, it is the analysis and use of the data that is interesting, perhaps less so the technical challenges that large data sets present to those who accumulate them. So the remainder of my talk will focus specifically on the analytic techniques and why those analytic techniques present some kinds of troubles for us when thinking about consumer protection.

So what I'll say then is, as a kind of starting place, that we can define data mining as the automated process of extracting useful patterns from large data sets, and in particular, patterns that can serve as the basis for subsequent decision making. What I'm saying here, in quotes, learning, meaning I learned from the previous examples that there is some general trend, some relationship in the data that I imagine will hold true in the future. And I can use that as a way to make future guesses and inferences as mentioned earlier already.

For terminology, I thought I'd also point out that within the field, this accumulated set of relationships within the data is commonly referred to as a model. So you might have heard the term predictive model. What that refers to, then, is all the various kinds of patterns that have been extracted from a large data set that then inform future decision making. And these models can be used in a variety of ways. To begin with, they can be used to classify entities. So the most common example of this would be spam. I think many people are familiar with this.

Your computer, often webmail in fact, will make guesses about whether or not your message is spam or not. And again, it arrives at a rule to determine what is spam and what is not spam based on the history of examples of spam. Likewise, it can estimate values of unobserved attributes, so it can guess your income, for instance, as also mentioned. And finally, it can also make predictions about what you're likely to do, so future consumer behavior of all sorts.

Now, you might say, as again was already mentioned, that of course data mining is discriminatory. The very intent and purpose of the activity is to be able to differentiate and draw distinctions. And what I would say, too, is that it is in some sense it's a statistical form of discrimination that is almost by necessity a rational form because it is being driven by apparent statistical relationships in the data. These are not arbitrary. This is not a case of caprice. This is in fact evidence suggesting that there are reliable patterns in the data.

And using that, you can confer to the individual those qualities which happen to be similar to those who appear statistically similar. So if I reside in one particular statistical category that has been revealed by the analysis, they can impute to me those same qualities.

So the remainder of the talk will focus on this five part taxonomy, which is me basically trying to explain how the process of actually mining data lends itself to a variety of issues that can raise concerns of discrimination and fairness. So let me jump right into it.

Again, a technical term is target variable. What this basically refers to is, when I set about trying to determine if there are useful patterns that correlate with some outcome, I need to be very specific about what I mean by that outcome. So when I'm looking for good customers, I actually need to arrive at a formal definition of what "good customer" means. Does good customer mean that it's the one from whom I can extract the most profit? Is it the one I can have a long term relationship with? Is it the one that if I provide some inducement will stay a customer?

And there's no way to actually avoid this formalization process. You must specify in some definable way what it is that you're looking for. And so the exercise of mining data always begins with actually having to establish some translation from a business problem into a problem that can be solved by predicting the value of this target variable. And in general, the art of data mining, the kind of creative work of data mining, involves this process of translation, finding a smart, clever way of actually translating some kind of business problem into one that can be solved by predicting the target variable, by inferring the value of the target variable.

And I think here, what's interesting is that the way that the business goes about defining the target variable can have serious consequences for whether or not the data mining process has a disparate impact. In my own work, I look at employment. And you might say that trying to predict whether or not someone is going to be particularly productive as compared to predicting whether or not are going to remain an employee for a certain period of time, trying to avoid turnover for instance, those differences and definitions will have very different consequences for how you rank potential applicants. And the same would likewise be true with consumers.

The second part of the taxonomy is what, again, data miners refer to as training data. Training data is the large set of information that you use to extract some kind of useful rule. It's the set of examples that you look at in order to decide if there are actually useful patterns to guide future behavior, future decision making. And I think, in this case, there are really two different although related problems with training data that again can have consequences for fairness.

One is that, as also mentioned, the set of examples can be skewed in some way. And the second, that the examples that you draw on could actually be in some way tainted by prior prejudice. So let me walk through this a bit.

When trying to derive some general rule from a set of particular examples, the only way that rule will actually generalize to future cases is if the particular set of examples happens to be representative of future cases. And as we know from the street bump case, we know this is not always the case.

And even more interestingly, I think, oftentimes companies are often seeking ways to try to change the composition of their customer base such that to suggest that you can draw general rules from a customer base that you're purposely changing again should put into doubt the idea that this is representative data. That in fact, you're dealing with a subset of all possible customers, and the particular subset you're dealing with changes over time.

We could also point out, I think, that the reason why data is unlikely to be particularly representative in certain cases is for reasons having to do with the following. So to begin with, it might well be that certain populations are less involved in the formal economy and the various mechanisms involved in producing these kinds of traces. They might have unequal access to, and less fluency in, the technology that's required to produce those kinds of digital traces. And finally, they simply might be less profitable or important constituents and therefore not the subject of ongoing observation.

And I think the more problem here is that oftentimes, the under or over representation of particular populations is not always evident. Sometimes, when the geographic distribution of the population is skewed in some obvious way, as in Street Bump, we might have intuition as to the fact that there's a problem. But many times, it will be much, much more difficult.

Finally, you could also say then that when you have this skewed example, it also suggests that companies should be devoting their attention to some populations and not others. And over time, this can have a compounding effect where certain populations are discounted further and further because you have less and less opportunity for those populations to disprove your sense that they are not in fact good customers. You're in fact limiting opportunities for those customers to buck the apparent trend. And this is a serious problem, pardon me, in credit scoring, where the industry has long worked on problems trying to deal with that.

Labeling examples, this is the process of actually trying to specify what is in fact a good customer and what is in fact a bad customer from examples. So I mentioned the example of spam. Let me actually jump to this example. So during the debates leading up to the Equal Credit Opportunity Act, Fair Isaac pointed out in those congressional debates that in fact any way of drawing some rules about how to extend credit to customers that looked to previous ways that consumers were evaluated as potential customers of credit would simply reproduce any prejudice involved in those past decisions.

Meaning, Fair Isaac could not simply draw on the history of credit decisions to automate the process. It actually had to find new ways to decide what in fact is a good target for credit. And what this reveals, then, is that any decision that uses past decisions as a basis for inferring rules must be sensitive to the fact that those decisions might be tainted by prejudice in some way.

Finally, in the same theme, along the same line, we can point out then that it's not only the case that data mining can inherit past prejudice, but it can continue to reflect the persistence of prejudice in the behavior taken as input to some kind of model. And this, I think, is a way of categorizing some of the work that Latanya Sweeney and others have done, showing then that if the input the algorithm receives is itself biased or prejudiced in some way, it will simply be reflected back in the recommendations of that system.

Feature selection, this is the process of deciding what variables, what criteria associated with each person would you actually fold into your analysis. And here, again, I think this is an interesting issue because you would imagine that big data presents opportunities to vastly increase the amount of features and variables you consider.

Of course, adding additional features to the analysis can often be costly. And it may well be that your analysis does very well when considering a certain set of features, but it doesn't do particularly well for some populations because it doesn't actually carve up the population in a particularly precise way. Redlining is a traditional example of this, using neighborhood alone as a way to decide who is worthy of credit is an extremely coarse way of making that determination.

And I think those same kinds of problems can actually translate to this new area because it is still possible that additional data would be useful in drawing distinctions for particularly marginalized populations. That simply might just be very costly. It might be very difficult to obtain that information. And the question therefore becomes, I think, does it justify subjecting these populations to less accurate determinations simply because it actually costs additional money or resources to gain that kind of information?

This fourth point of the taxonomies is what we call proxies. And what this refers to is the fact that oftentimes, many of the features that are legitimately relevant in making some kind of predictions about customers might also be highly correlated with their class membership. Meaning certain features, certain attributes, are both proxies for the thing you care about and proxies for the person's class membership.

And what's worse here, then, is that it may well be that this is actually simply reflecting the fact of inequality in society. And it's a particular form of inequality where members of historically marginalized and particular classes are disproportionately in a less favorable position.

And big data is in a position potentially to simply further exposed the exact extent of that inequality. I will, in the interest of time, jump over this. The final part of this taxonomy is masking, which refers to the idea that is possible to mask intentional discrimination by relying on any of the number of ways I've identified here of having discrimination happen unintentionally.

Decision makers additionally can rely on data mining to infer whether or not you belong to a protected class, and then to use that information in secret to discriminate against you. I want to emphasize, though, and this is, I think, one of the most important points I'll make today, is that unintentional discrimination of the sort identified in the first four parts of the taxonomy is far more likely to be occurring. And it has potentially far more sweeping consequences than the kinds of intentional discrimination that could be pursued through masking.

And I'll simply conclude by saying that I think there's a serious issue here about the unintentionality of the discrimination that might be occurring. And in my own research, I've looked at Title VII and employment decisions. And my sense actually is that this aspect of the problem, the unintentionality of the problem, will pose serious issues for trying to bring to bear legal remedies. It's unclear that we have the tools when looking at existing laws to actually address this form of unintentional discrimination.

Additionally, if the problem is that we are exacerbating inequality, it's also unclear whether or not using discrimination law as a way to deal with that issue is the correct mechanism. And finally, I think that for many of the kinds of problems identified earlier, there's no ready answer both at a technical and I think legal level, and we really require a conversation that involves both parts of this debate, the technical and the legal dimension. So thank you very much, and I hope people will speak with me if they have further questions. Thanks.

[APPLAUSE]