

Tech for Good: Data for Social Empowerment
Google DC
Washington D.C.
September 10, 2015
Keynote Remarks of Commissioner Terrell McSweeney

Thank you for having me here today as part of Google’s DC Tech Talk series. I want to thank Susan Molinari for those kind opening remarks. I also want to extend my thanks to all the panelists who are joining the discussion today. I’m looking forward to what will certainly be an interesting conversation.

I’ll start with the usual disclaimer: the views expressed in this speech are my own and are not those of the FTC.

Today’s talk is entitled “Tech for Good: Data for Social Empowerment” and I think the title of this session says a lot about where we are as a society, as consumers, and as policymakers. In the last few years, the words “data” and “technology” have become two sides of the same coin. Technology is increasingly about how much, and what types, of data can be collected, stored, sorted, and fused from disparate sources to create new uses and applications. Data is the fuel that powers much of our technological progress, and provides innovators with the raw materials they need to make better apps, better services, and better products.

It was fifty years ago this year that Gordon Moore published his seminal work that predicted an ever increasing capacity to, as he said, “Cram more and more components onto integrated circuits.”¹ He further quantified this “cramming” into what we now know as Moore’s Law – the regular doubling of the components in an integrated circuit. Since that paper was published, we’ve essentially seen the processing capacity of components double nearly every two years.

This has meant greater data processing at less cost than could have ever been imagined before. We’re on the cusp of an explosion of data from sensors and Internet connected devices that will facilitate the digitization of almost every part of our lives. In 2013 researchers estimated there were about four zettabytes of data generated worldwide. Cisco recently released a report that the Internet of Things will generate more than 400 zettabytes of data a year by 2018.²

How much is that, you might ask? A zettabyte is one trillion gigabytes. One gigabyte is approximately the amount of data in 4,500 books. So very roughly, we will be generating more data in a year from all our devices than all the data created from the dawn of the written word to the creation of the Internet.

¹ Gordon Moore, *Cramming More Components onto Integrated Circuits*, ELECTRONICS 114 (April 19, 1965).

² White Paper, Cisco, *Cisco Global Cloud Index: Forecast and Methodology, 2013-2018* (2014) at 17, available at http://www.cisco.com/c/en/us/solutions/collateral/service-provider/global-cloud-index-gci/Cloud_Index_White_Paper.pdf.

How this vast amount of data – our data – is handled, from its collection and use to its storage and protection, raises some of the thorniest issues policymakers face during this new Information Age. These are complex and important issues that have real and potentially harmful consequences for consumers – but they are not the focus of my talk this afternoon.

Today we are discussing some of the many benefits and insights we can gain from data – and how transformative it can be for communities. The growing abundance of data and our increasing capacity to use it, presents a wonderful opportunity for data-driven governing and policymaking.

For example, we all know the very real problems that the drought in California is creating in many communities. Just a decade ago, cities would have been limited to issuing citations for watering lawns or maybe hiring extra inspectors to review water bills and water meters. That is no longer the case. Now some cities are using new flow meters, water sensors, and even soil monitors to do more than just enforce watering restrictions. They are not only targeting water use, but also dispatching maintenance crews to leaky underground pipes, engaging citizens on social media, preventing wild fires, monitoring soil moisture for agriculture, and saving tens of thousands of hours and hundreds of thousands of dollars by efficiently managing their response.³

In another example, last winter the city of Chicago employed a “Plow Tracker” so that residents could, in real time, track and request the city’s snowplows in order to more efficiently monitor and dispatch street clearing services. The online portal also organized Snow Corps, groups of neighbors who would shovel walks, driveways, and stairs for elderly or infirm Chicagoans.⁴

But that’s not the end of what happened with this data. The results were further crunched by the University of Chicago to create new maps of the city. Clusters of underserved blocks were identified and the city had real-time evidence of where to dispatch other services beyond snow removal, including pest control, road and sidewalk repair, and weatherization assistance.

Good things are also happening when governments open up their data for access by the public. Like public health authorities in Baltimore, New York, and Chicago that have shared their data with the private sector to better target free medical services and information campaigns – in some cases substantially increasing access to preventive care in underserved communities.⁵

³ Molly Peterson, NAT’L PUB. RADIO, “In Long Beach, Calif., Smart Meters Spot Wasteful Water Users” (Apr. 6, 2015), available at <http://www.npr.org/2015/04/06/397774754/in-long-beach-calif-smart-meters-spot-wasteful-water-users>.

⁴ See Building Resilient Regions, Institute of Governmental Studies, UC Berkeley, “Big Data and Urban Planning” (Jan. 22, 2013), available at <http://brr.berkeley.edu/2013/01/big-data-and-urban-planning/>; City of Chicago, “Snow Corps,” available at http://www.cityofchicago.org/city/en/depts/mayor/snowportal/snow_corps.html; Monica Davey, N.Y. TIMES, “Snow Site Lets Chicago See if Plows Are Really in a Rut” (Jan. 3, 2012), available at <http://www.nytimes.com/2012/01/03/us/chicagoshovels-web-site-gives-lowdown-on-snow.html>.

⁵ See Shefali Luthra, Kaiser Health News, “Big Data Offer New Strategy for Public Health Campaigns” (Dec. 3, 2014), available at <http://khn.org/news/big-data-offer-new-strategy-for-public-health-campaigns/>.

On the federal level, President Obama has been on the forefront of freeing up data for innovation for the public good. He not only appointed the first Chief Technology Officer and first US Chief Data Scientist, but also issued the Open Government Directive that was later codified as an executive order. Under President Obama, more than 130,000 data sets have been made available to the public for innovation and entrepreneurship on data.gov – with the necessary safeguards to prevent the release of sensitive and personally identifiable information.⁶

This has provided new tools empowering people in a variety of ways – creating methods for better weather forecasts, providing searchable stats about colleges for students and parents, improving our response to natural disasters, helping consumers lower their utility bills, enabling a new era of medicine by making treatment and prevention plans more personalized and effective, and much more.

And through Challenge.gov, agencies throughout the government are empowered to enlist the help of citizen innovators to solve public policy dilemmas. Data sets and other tools are released to the public, often a prize is created, then innovators and entrepreneurs create solutions.

For example, at the FTC we've used this platform to get help fighting robocalls, which thwart the Do Not Call list. Robocalls present particular enforcement challenges – defeating them is more or less like playing a game of whack-a-mole. A couple of years ago we held our first "FTC Robocall Challenge."⁷ We made available the data collected on robo-calls such as the offending phone numbers, the times of day calls were most frequent, and the subject of the calls, then challenged developers to create their own solutions. The winner was a programmer named Aaron Foss. The service he founded, Nomorobo, is available to all consumers, and has so far stopped more than 36 million robocalls.

This year, our "Robocalls: Humanity Strikes Back" contest challenged entrants to develop technologies that automatically block and forward the unwanted calls to a crowd-sourced honey pot.⁸ The winner, RoboKiller, uses audio fingerprint technology to identify calls and enables users to personalize their filtering settings.

These are terrific examples of how we in government can take our data, partner with the private sector, and develop innovative solutions to difficult problems.

Many of the other challenges focus on empowering Americans with information so they can make good decisions. We've seen challenges that have generated new apps addressing everything from tracking air quality to campus rape prevention to calculating the full cost of

⁶ DJ Patil, The White House Blog, "Two Years of Transformative Open Data for Public Good" (May 11, 2015), available at <https://www.whitehouse.gov/blog/2015/05/11/two-years-transformative-open-data-public-good>.

⁷ See Press Release, Fed. Trade Comm'n, "FTC Announces Robocall Challenge Winners" (April 2, 2013), available at <https://www.ftc.gov/news-events/press-releases/2013/04/ftc-announces-robocall-challenge-winners>.

⁸ See Press Release, Fed. Trade Comm'n, "FTC Awards \$25,000 Top Cash Prize for Contest-Winning Mobile App That Blocks Illegal Robocalls" (Aug. 17, 2015), available at <https://www.ftc.gov/news-events/press-releases/2015/08/ftc-awards-25000-top-cash-prize-contest-winning-mobile-app-blocks>.

college tuition, to connecting veterans to job opportunities, even to providing help with where to eat and shop while on a military base.

For all the innovation and empowerment that can be unleashed by our data revolution, I think my colleague Commissioner Ohlhausen put it best by saying, "...big data isn't knowledge or wisdom."⁹ It is a tool, and like any other tool, it can be used for good or ill.

That is where policy makers, data scientists, regulators, and enforcers come in, and it is where we at the FTC will have to take our consumer protection mission. To begin with, data can only positively affect people's lives if the data itself is good. Too often in this country, minorities, the poor, the disabled, the elderly, and new immigrants are not adequately counted.

For example, two years ago, the city of Boston unveiled an app called Street Bump. It allowed drivers to report potholes and other road hazards directly to the public works department from their smartphones. After a few weeks, the city began noticing that there were far more hazards and potholes in wealthy neighborhoods than in poorer areas of the city.¹⁰ Clearly, the data collected was skewed. Wealthy people downloaded the app and drove private cars, while poorer people didn't have smartphones often took the bus. It became obvious to the city that an innovative and cost-effective way of tackling a nagging problem was exacerbating inequality.

In fact, many of the most intractable problems we face as a nation have their roots in how we gathered and used data as policy makers and businesses. It wasn't done with computers and algorithms, but for most of the twentieth century, banks and government agencies produced data sets on where, and for whom, loans could be made and houses could be sold. They created maps that redlined communities and enforced a rigid segregation based on race and class. This was big data used for discriminatory purposes, and it has left a legacy that we are still struggling to correct.

With that tragic history as a backdrop, the Leadership Conference on Civil Rights issued a set of principles to ensure that our new data era is one of empowerment and opportunity rather than one of data driven prejudice.¹¹ Among their principles is the recommendation that audit mechanisms be used to limit risks of profiling and discrimination. They also ask that computerized decision making – especially in the areas of employment, credit, education, and health – be subjected to human oversight.

I believe these are necessary safeguards for government and industry. Last year, the FTC held a workshop entitled "Big Data: A Tool for Inclusion or Exclusion?" While we are still

⁹ Maureen K. Ohlhausen, Commissioner, Fed. Trade Comm'n, Remarks at the Center for Data Innovation: The Social Impact of Open Data (July 23, 2014) at 4, *available at* https://www.ftc.gov/system/files/documents/public_statements/571281/140723socialimpactofopendata.pdf.

¹⁰ See Nick Heath, TECHREPUBLIC, "Big data: Neither snake oil or silver bullet" (May 9, 2014), *available at* <http://www.techrepublic.com/blog/tech-decision-maker/big-data-neither-snake-oil-or-silver-bullet/>; Elizabeth Good Christopherson, RITA ALLEN FOUND., "Confronting the Data Dilemma" (July 25, 2013), *available at* <http://www.ritaallenfoundation.org/blogs/confronting-data-dilemma.php>.

¹¹ See The Leadership Conference on Civil and Human Rights, "Civil Rights Principles for the Era of Big Data" (2014), *available at* <http://www.civilrights.org/press/2014/civil-rights-principles-big-data.html>.

working on our report on that workshop, it became clear to me that without transparency and oversight, consumers could face real harms.

Every day research reveals the subtle and not so subtle ways fusion of data can lead to situations in which different groups are served different kinds of information. For example, a study by Northeastern University revealed that consumers could pay different prices for the same product depending on the type of device a consumer used or the location of their device. Studies have also shown that homestay rentals offered by minorities, even in the same or better neighborhoods, have lower returns than those offered by white renters. I could go on.

Pinpointing why a result like this occurs poses a challenge. Algorithms that are neutral on their face can uncover and perpetuate our own biases. But outcomes like these have the potential to cause real harm to consumers. In this environment it is more important than ever to test results of algorithms to detect potentially discriminatory consequences of them.

At the FTC we recently established a new Office of Technology Research and Investigations. Part of its mission – which I hope will expand over time – will be to uncover some of these possible harms. For years, the FTC has promoted “Privacy by Design” and “Security by Design” – a process-based approach to improving protections for and consideration of consumers’ privacy and security from the outset of product design.¹² We believe that responsible companies employing these strategies can potentially address problems before products get to market.

In an era when algorithms and big data sets can hire and fire, inform health care decisions, extend financial opportunities and much more, it is time to engage in responsibility by design – in other words, to incorporate a process-based approach that values nondiscrimination and fairness throughout the lifecycle of products and to take meaningful steps to assess and mitigate harmful or discriminatory consequences if they are revealed.

Getting the balance right will require an openness in the uses and methods we attach to data, testing and engagement with stakeholders and communities that are typically not engaged in product development, and most importantly, empathy when problems come to light. An approach to fairness and nondiscrimination will ensure that Big Data can truly be a tool of social empowerment and that the promise of technology is kept for everyone.

Thank you again for having me here. I look forward to what will be a very valuable discussion.

¹² See generally Report, Fed. Trade Comm’n, *Protecting Consumer Privacy in an Era of Rapid Change: Recommendations for Businesses and Policymakers* (Mar. 2012), available at <http://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-report-protecting-consumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf>; Report, Fed. Trade Comm’n, *Careful Connections: Building Security in the Internet of Things* (Jan. 2015), available at <http://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-report-protecting-consumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf>.