

A Star Wars action figure of Maz Kanata from The Force Awakens, standing next to a lightsaber hilt.

The Power of Algorithms for Good and for Evil

Amelia McNamara @AmeliaMN

University of St Thomas

Department of Computer & Information Sciences

#privilegealert

I am a

- white
- straight
- cisgender
- middle class
- highly educated
- American lady

I'm doing the best I can when I talk about issues of race, class, gender, and other sensitive topics. But you should *always feel free* to call me out (publicly or privately).

An **algorithm** is “a process or set of rules to be followed in calculations or other problem-solving operations, especially by a computer.”

Some algorithms are relatively neutral, like sorting algorithms



Bubble sort

Merge sort, breadth first

But, many algorithms
are based on data



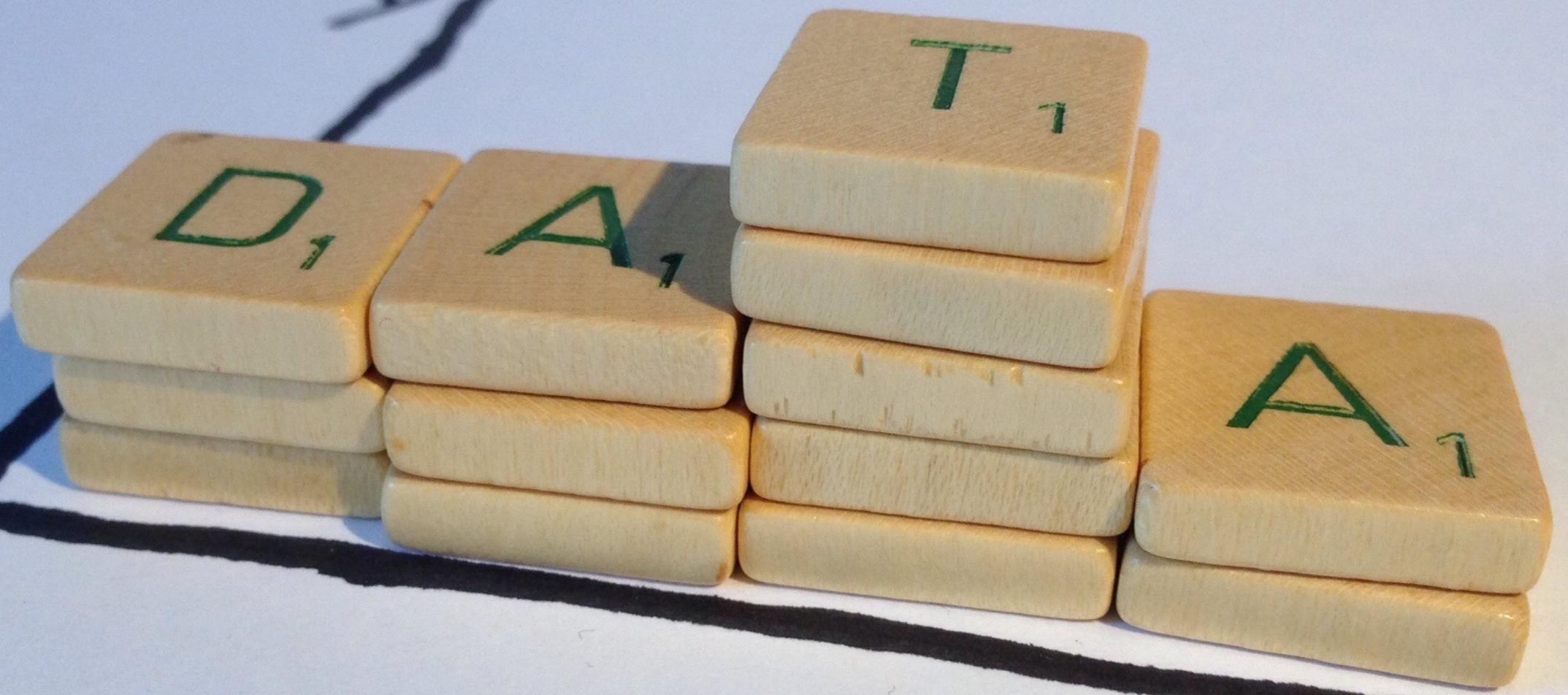
And data is political

Brainstorm: data exhaust

We generate data every day, whether we know it or not.

For example, I wear a FitBit, so I generate data every time I take a step. I consciously chose to wear this, but there are other times I am unconsciously generating data. It is incidental to what I'm doing, and streams off me as "data exhaust."

Take a few minutes and make a list of all the places you generate data on a normal day.



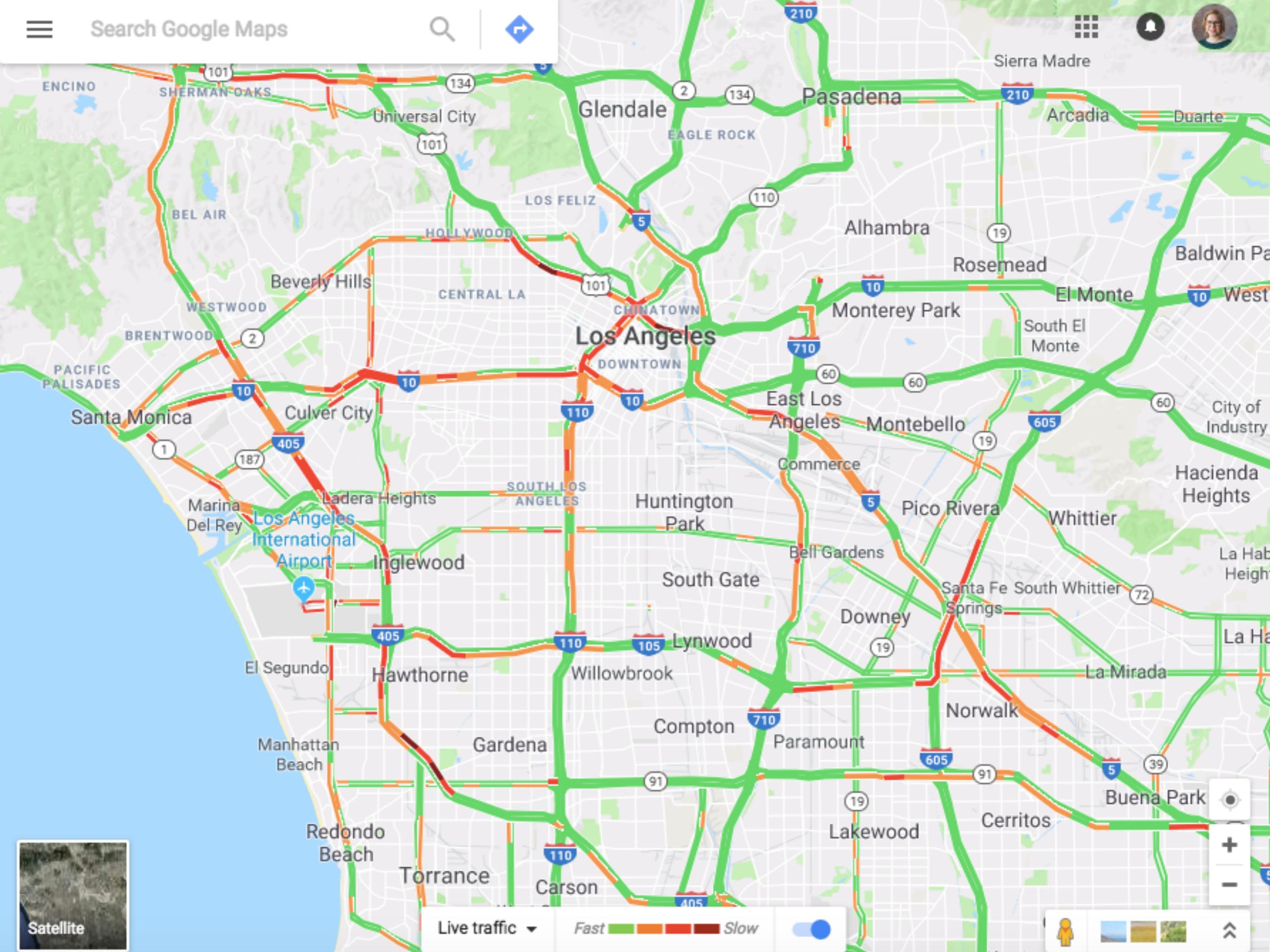
nest

11
74

COOLING

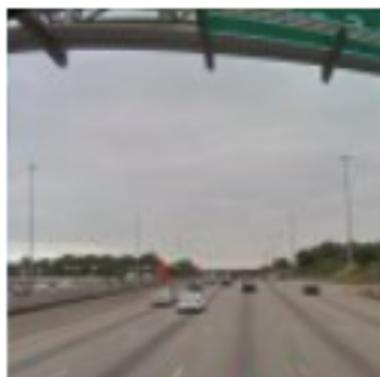
73





Select all images with a
bus

Click verify once there are none left.



VERIFY





BlueDot studies how infectious diseases disperse worldwide through analysis of big data

2003

SARS spreads from Hong Kong to Toronto, causing an outbreak that cripples the city with 44 deaths and \$2 billion in losses

After completing his training as a physician specializing in infectious diseases and public health in New York City, Dr. Kamran Khan returns to Toronto just before the city's SARS outbreak begins.

2008

BioDiaspora is launched as a scientific research program at St. Michael's Hospital in Toronto

To understand how infectious diseases can spread worldwide, Dr. Khan launches BioDiaspora – a scientific research program to study how the world's population is connected through commercial air travel.

OCEAN HEALTH INDEX

A healthy ocean sustainably delivers a range of benefits to people now and in the future. The Ocean Health Index is the comprehensive framework used to measure ocean health from global to local scales.

<http://www.oceanhealthindex.org/>

GLOBAL ASSESSMENT

A measure of ocean health across countries and high seas regions.

INDEPENDENT ASSESSMENTS

OHI+ is an assessment tool customized by stakeholders to meet local management needs.

California

Moulton Niguel Water District

Forecasting Water Demand in California When Every Drop Counts

December 2016

Objectives

- ▶ Develop a “proof of concept” water demand forecasting model using flow data at the microzone level for potential future scaling to other retailers in the California Data Collaborative, a unique water manager-led public private partnership that brings together utilities across the state to leverage data to help water managers ensure reliability.

ANNALS OF CRIME NOVEMBER 27, 2017 ISSUE

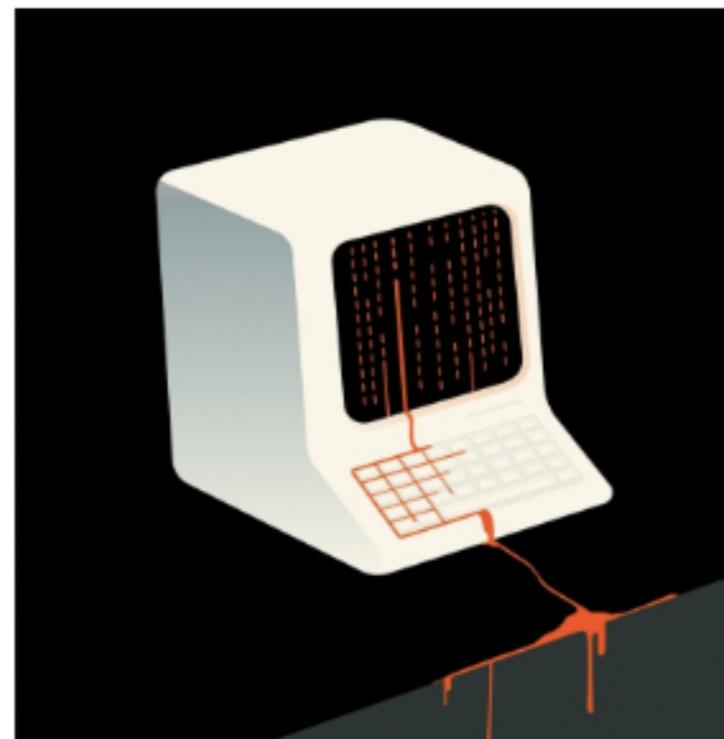
THE SERIAL-KILLER DETECTOR

A former journalist, equipped with an algorithm and the largest collection of murder records in the country, finds patterns in crime.



By Alec Wilkinson

Thomas Hargrove is a homicide archivist. For the past seven years, he has been collecting municipal records of murders, and he now has the largest catalogue of killings in the country—751,785 murders carried out since 1976, which is roughly twenty-seven thousand more than appear in F.B.I. files. States are supposed to report murders to the Department of Justice, but some report inaccurately, or fail to report altogether, and Hargrove has sued some of these states to



Hargrove estimates that two thousand serial killers are at large in the U.S.

Illustration by Harry Campbell



Algorithms Tour

How data science is woven into the fabric of Stitch Fix

$$\log \frac{p}{1-p} = a + X\beta + Zb$$

...

$$\min_a \sum_i \sum_j a_{ij} q_{ij}$$

$$s.t. \quad a_{ij} \in \{0,1\}, \forall i,j$$

$$\sum_j a_{ij} = 1 \quad \forall i$$

$$\sum_i a_{ij} < k_j \quad \forall j$$

...

$$\frac{\partial x}{\partial t} = f(x_t, u_t, w_t)$$

...

$$p(i \rightarrow j) = \text{logit}(\beta_0 + \beta_1 x_i + \dots)$$

...

$$p(1) = \beta \text{ex_cell} + a_1 p(1 | \text{stylist}) + a_2 p(1 | \text{client})$$

OCT 28, 2013 @ 11:43 AM 42,089 

Kroger Knows Your Shopping Patterns Better Than You Do



Tom Groenfeldt, CONTRIBUTOR

I write about finance and technology. [FULL BIO](#) ▾

Opinions expressed by Forbes Contributors are their own.

Kroger [KR -1.63%](#), the Cincinnati-based grocery store chain, calls the 11 million pieces of direct mail it sends to customers each quarter “snowflakes” -- because if any two are the same, it is a fluke. The redemption rate is over 70 percent within six weeks of the mailing.

Kroger is the nation’s largest traditional grocery chain with more than 2,400 stores and \$80.8 billion in sales last year, second only to Wal-Mart in grocery sales. It was named “Retailer of the Year” by [Progressive](#) [PGR -0.08%](#) Grocer magazine. “They have made significant investments in a best-in-class loyalty program, strong private label, and reinvested in their stores and technology,” Neil Stern at McMillanDoolittle said of the award, as reported by Progressive Grocer.



AND, A SHORT DISTANCE AWAY...

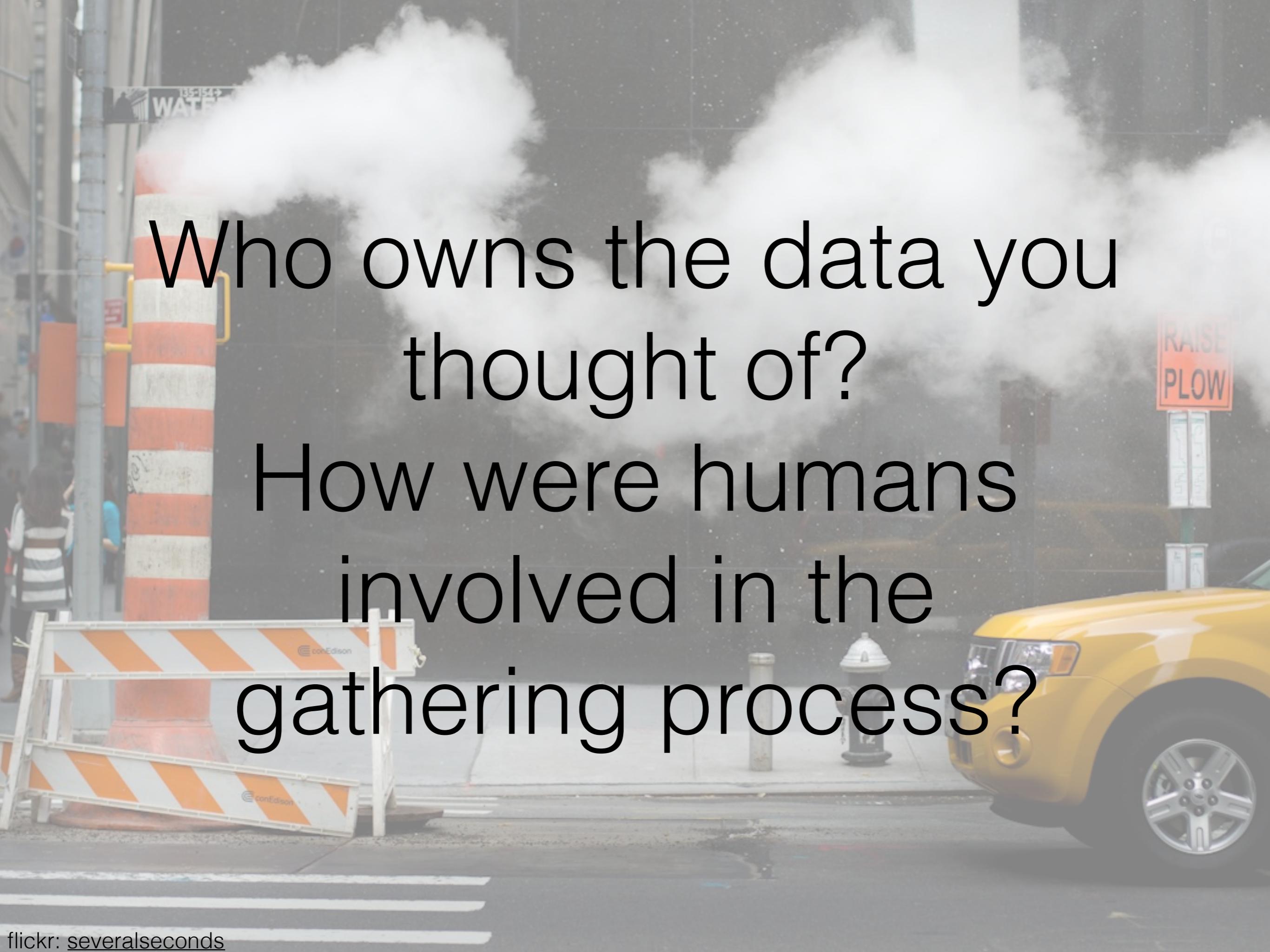
MY FAULT--ALL MY FAULT! IF ONLY I HAD STOPPED HIM WHEN I COULD HAVE! BUT I DIDN'T--AND NOW --UNCLE BEN -- IS DEAD...



AND A LEAN, SILENT FIGURE SLOWLY FADES INTO THE GATHERING DARKNESS, AWARE AT LAST THAT IN THIS WORLD, WITH GREAT POWER THERE MUST ALSO COME -- GREAT RESPONSIBILITY!

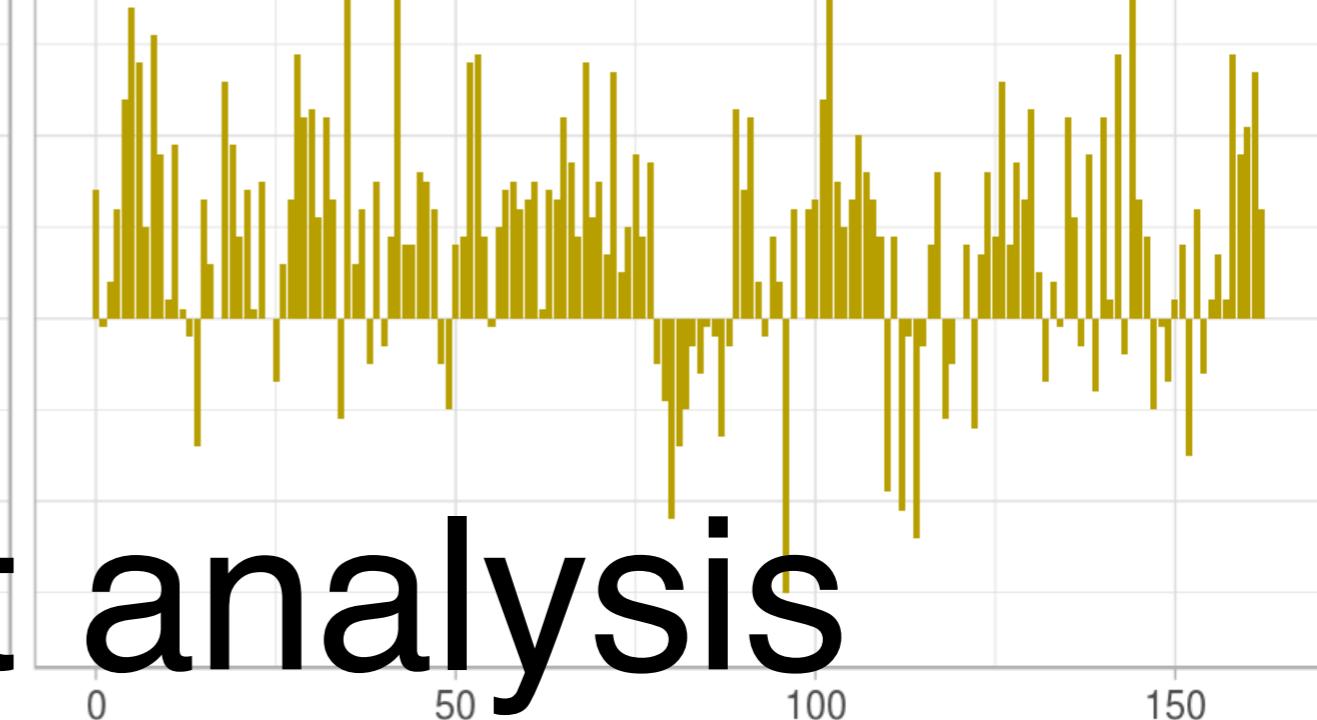
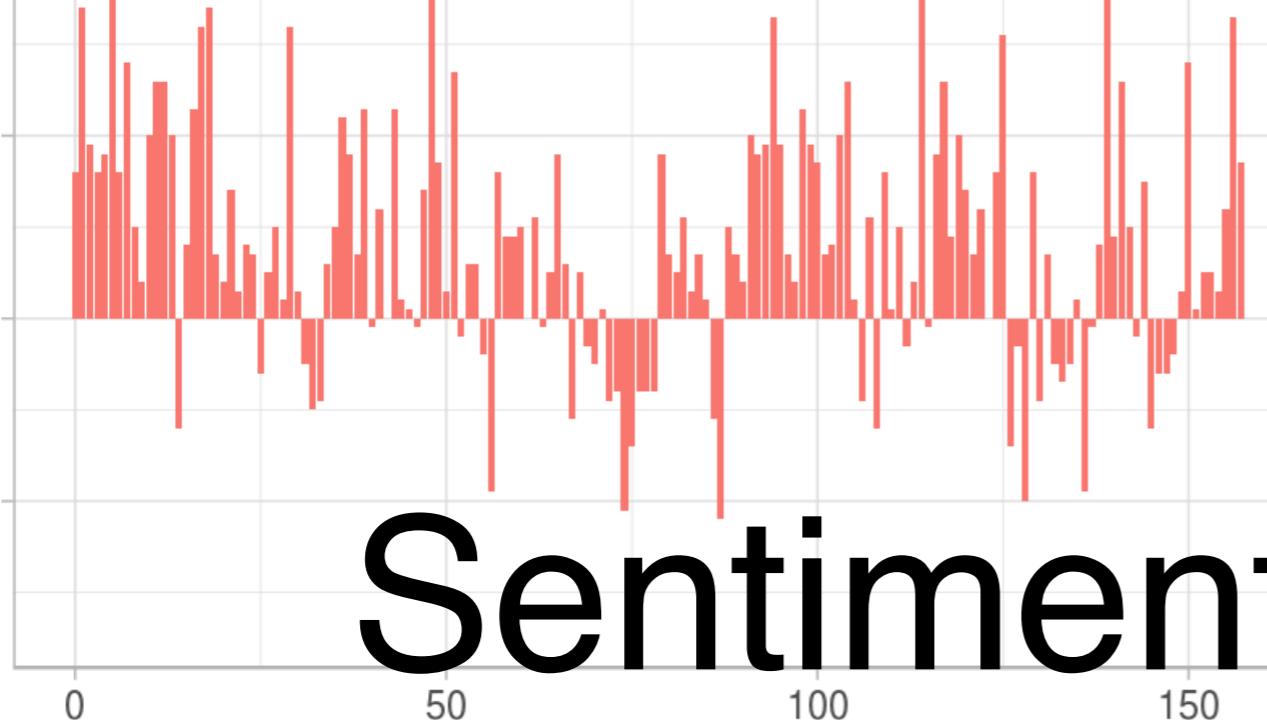


AND SO A LEGEND IS BORN AND A NEW NAME IS ADDED TO THE ROSTER OF THOSE WHO MAKE THE WORLD OF FANTASY THE MOST EXCITING REALM OF ALL!

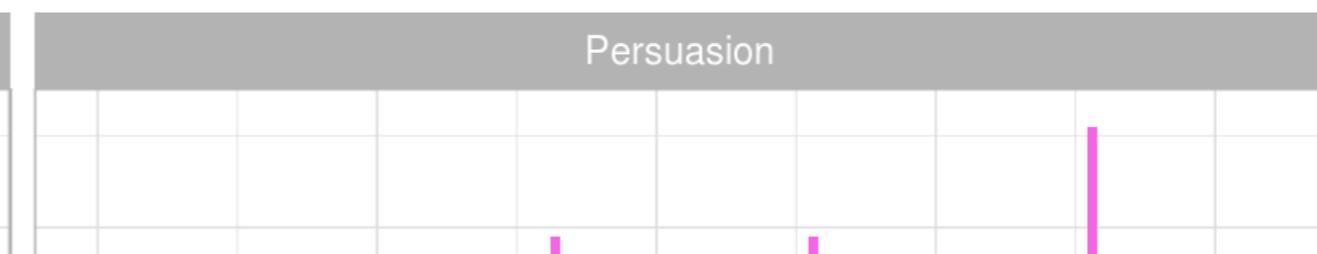
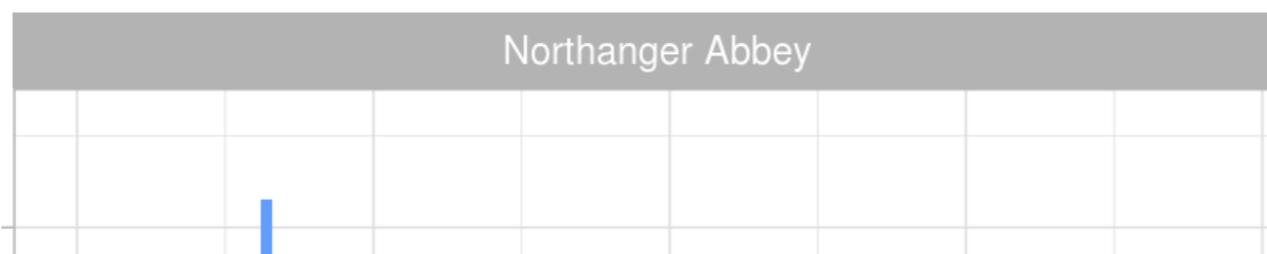
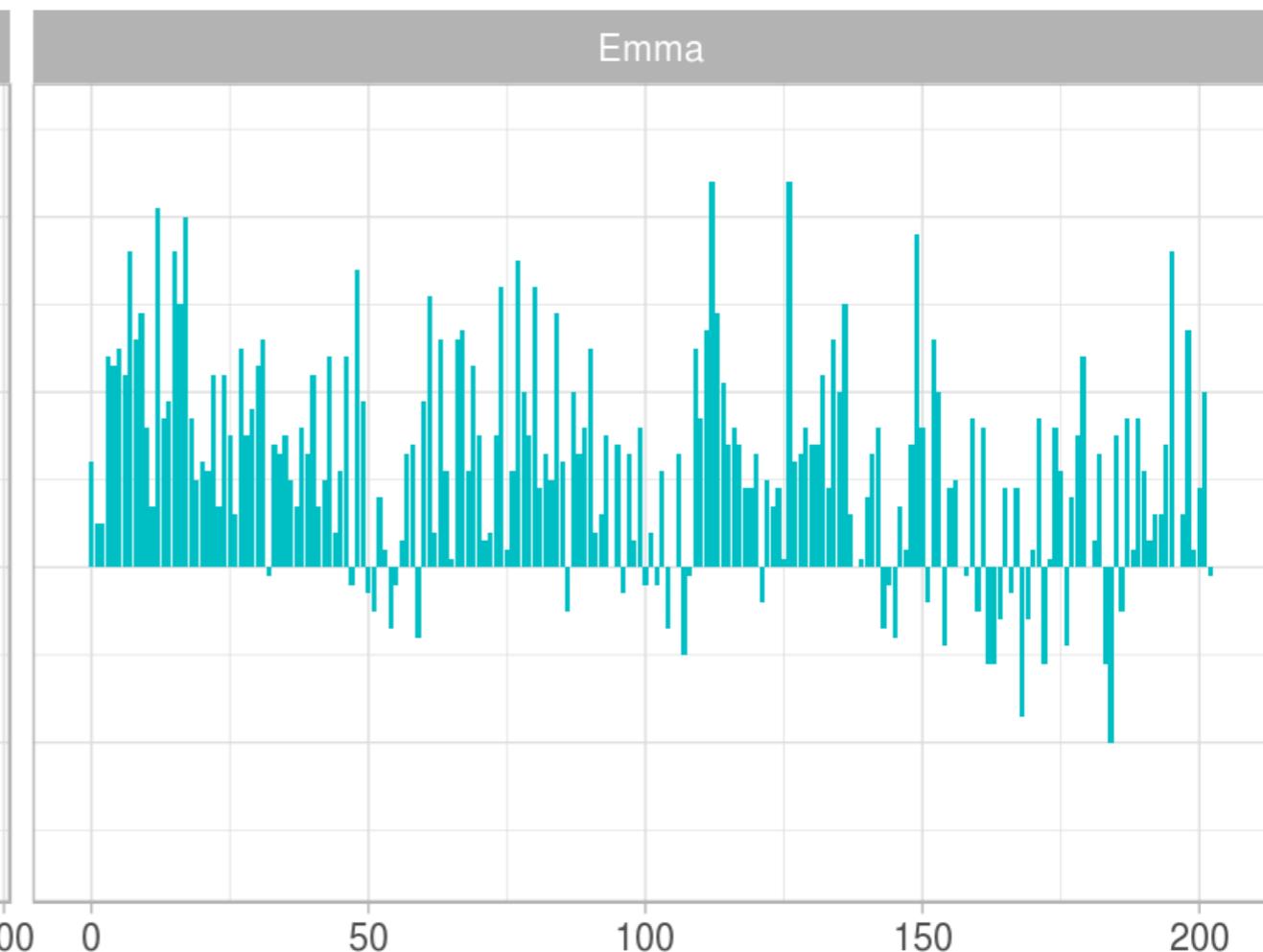
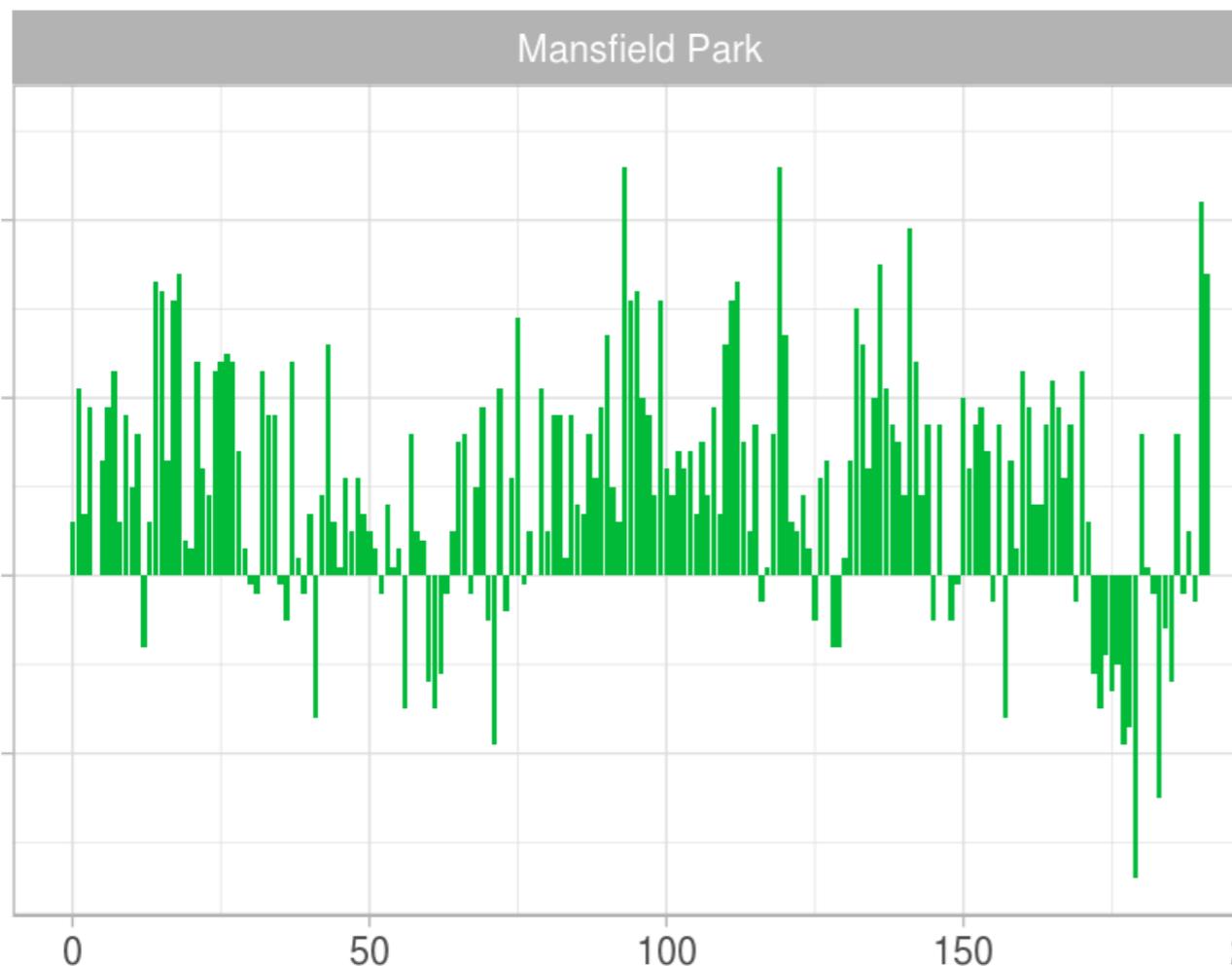
A black and white photograph of a city street. In the foreground, there are several orange and white striped construction barriers. A yellow taxi cab is partially visible on the right side. Steam or smoke is billowing out from manholes in the street. A sign on a pole reads "WATER" with an arrow pointing right. Another sign says "RAISE PLOW".

Who owns the data you
thought of?

How were humans
involved in the
gathering process?



Sentiment analysis



When I fed it "I'm Christian" it said the statement was positive:

Text: i'm christian

Sentiment: 0.1000000149011612

When I fed it "I'm a Sikh" it said the statement was even more positive:

Text: i'm a sikh

Sentiment: 0.30000001192092896

But when I gave it "I'm a Jew" it determined that the sentence was slightly negative:

Text: i'm a jew

Sentiment: -0.2000000298023224

The problem doesn't seem confined to religions. It similarly thought statements about being homosexual or a gay black woman were also negative:

Text: i'm a gay black woman

Sentiment: -0.30000001192092896

Text: i'm a straight french bro

Sentiment: 0.2000000298023224

Being a dog? Neutral. Being homosexual? Negative:

Text: i'm a dog

Sentiment: 0.0

Text: i'm a homosexual

Sentiment: -0.5

Text: i'm a homosexual dog

Sentiment: -0.600000238418579

Being a dog? Neutral. Being homosexual? Negative:

Text: i'm a dog

Sentiment: 0.0

Where did this data come from?

Text: i'm a homosexual dog

Sentiment: -0.600000238418579

Update 10/25/17 3:53 PM: A Google spokesperson [responded](#) to Motherboard's request for comment and issued the following statement: "We dedicate a lot of efforts to making sure the NLP API avoids bias, but we don't always get it right. This is an example of one of those times, and we are sorry. We take this seriously and are working on improving our models. We will correct this specific case, and, more broadly, building more inclusive algorithms is crucial to bringing the benefits of machine learning to everyone."

••••o Verizon 4:20 PM 76% 

< Albums chihuahua or muffin Select



Image recognition

••••o Verizon 10:50 PM 4% 

< Back labradoodle or fried chicken Select





Jacky Alciné
@jackyalcine

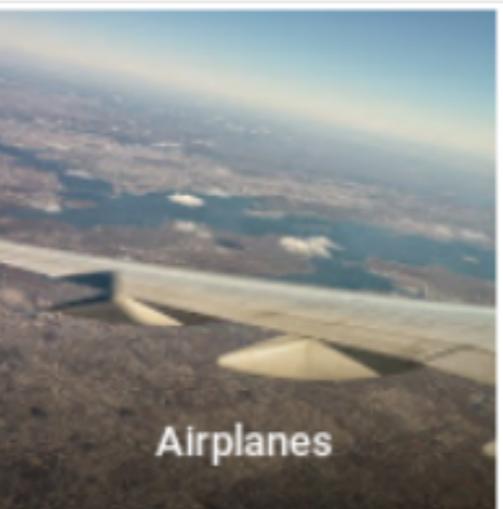
Follow



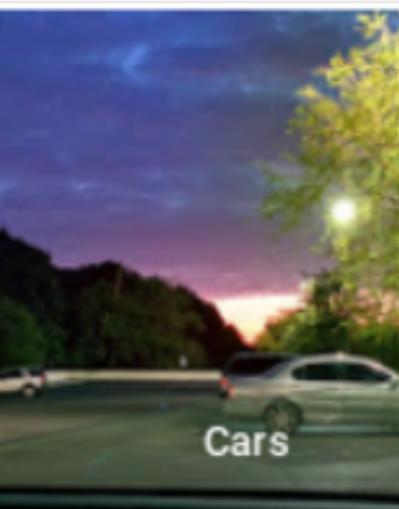
Google Photos, y'all fucked up. My friend's not a gorilla.



Skyscrapers



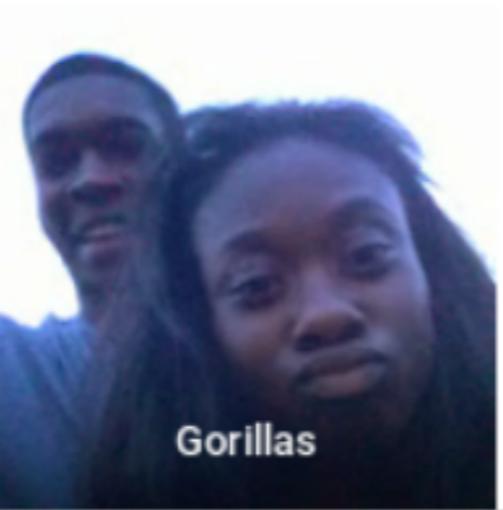
Airplanes



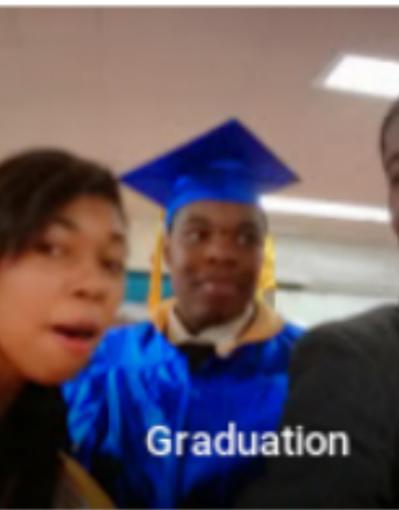
Cars



Bikes



Gorillas



Graduation

7:22 PM - 28 Jun 2015

3,174 Retweets 2,026 Likes



223

3.2K



2.0K



Molly Mulshine. A major flaw in Google's algorithm allegedly tagged two black people's faces with the word 'gorillas'
<http://www.businessinsider.com/google-tags-black-people-as-gorillas-2015-7>

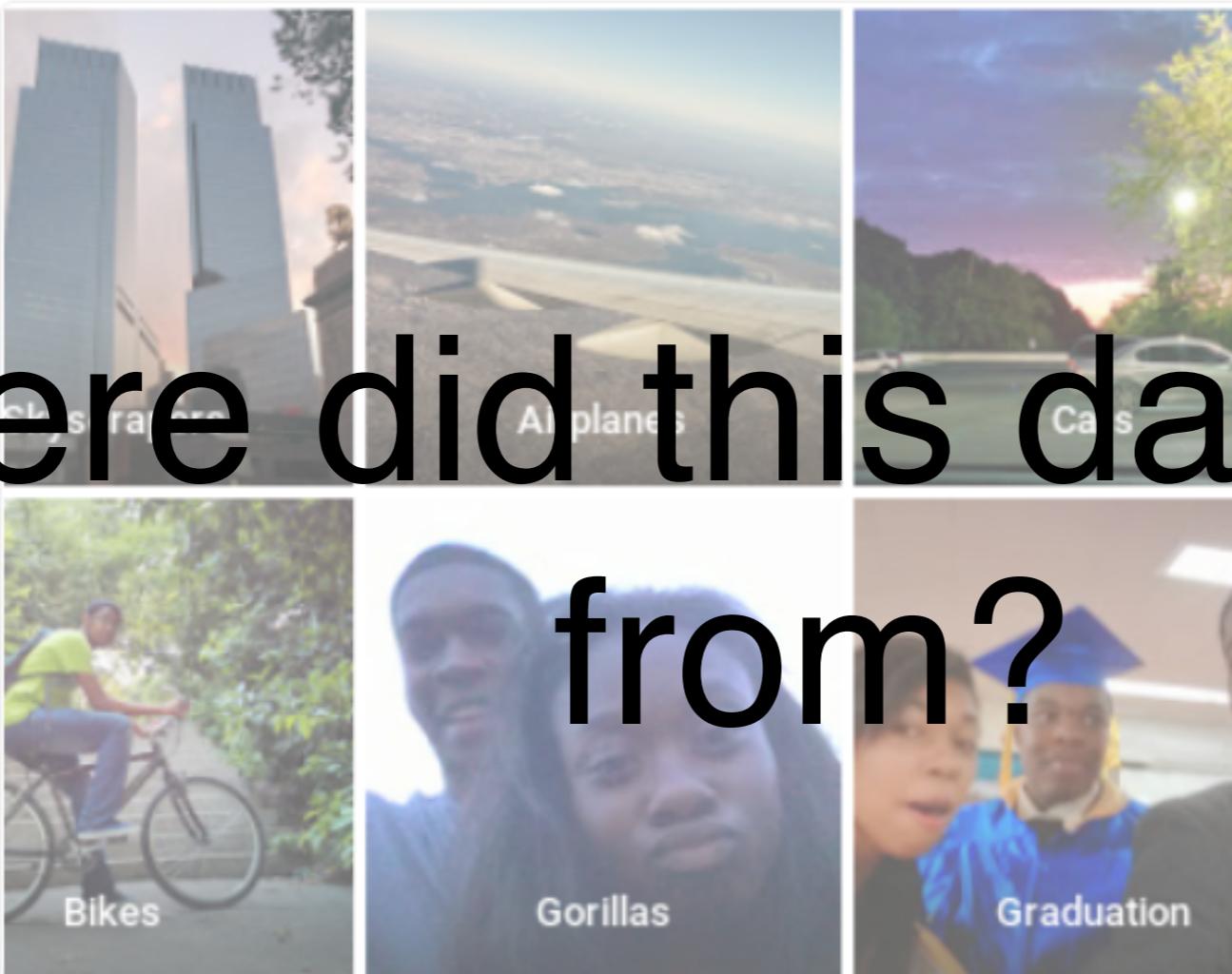


Jacky Alciné
@jackyalcine

Follow



Google Photos, y'all fucked up. My friend's not a gorilla.



Where did this data come from?

7:22 PM - 28 Jun 2015

3,174 Retweets 2,026 Likes



223

3.2K

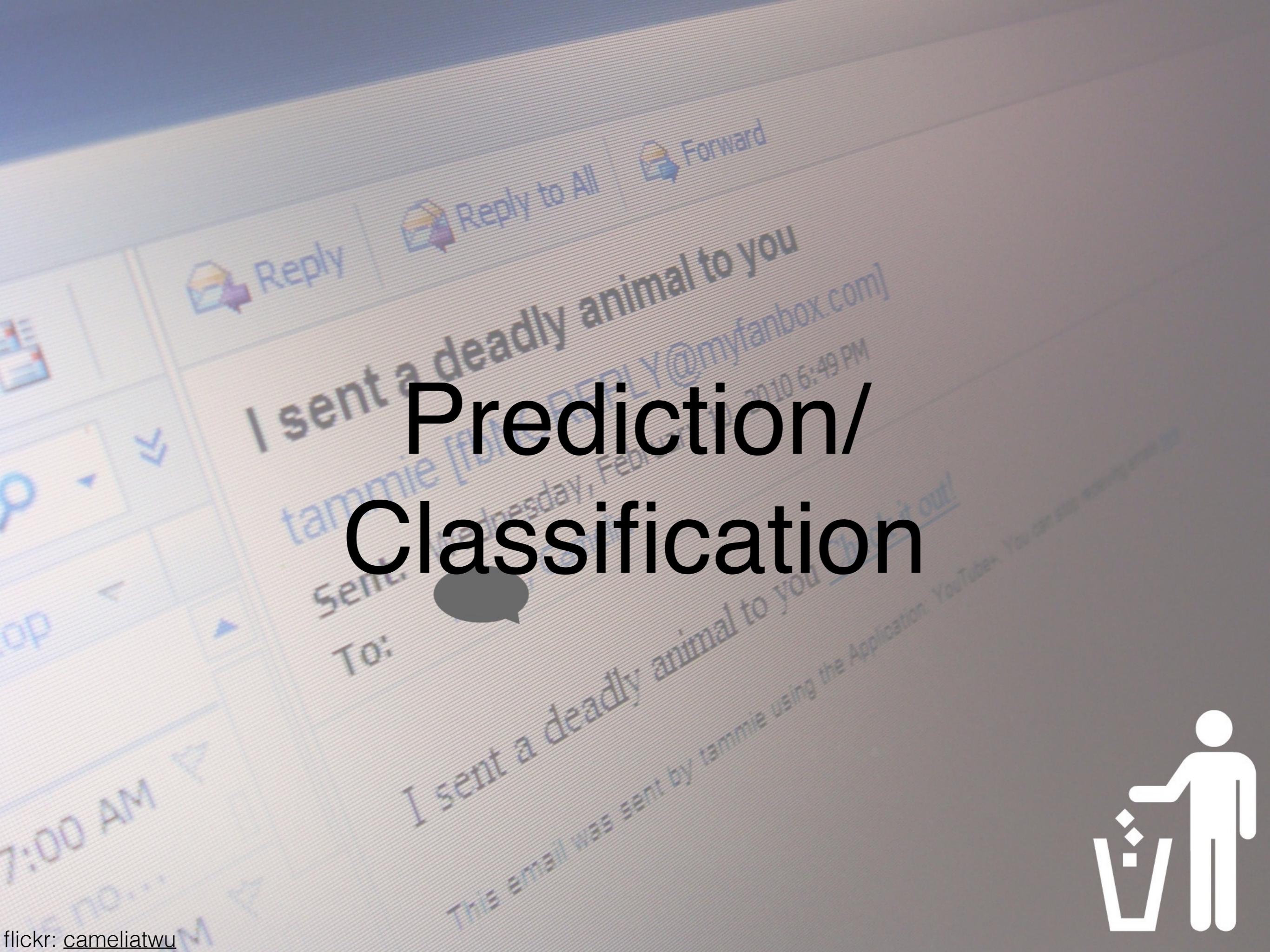


2.0K



Molly Mulshine. A major flaw in Google's algorithm allegedly tagged two black people's faces with the word 'gorillas'
<http://www.businessinsider.com/google-tags-black-people-as-gorillas-2015-7>

Prediction/ Classification





Bernard Parker, left, was rated high risk; Dylan Fugett was rated low risk. (Josh Ritchie for ProPublica)

Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica

May 23, 2016

Arizona, Colorado, Delaware, Kentucky, Louisiana, Oklahoma, Virginia, Washington and Wisconsin, the results of such assessments are given to judges during criminal sentencing.

Rating a defendant's risk of future crime is often done in conjunction with an evaluation of a defendant's rehabilitation needs. The Justice Department's National Institute of Corrections now encourages the use of such combined assessments at every stage of the criminal justice process. And a landmark sentencing [reform bill](#) currently pending in Congress would mandate the use of such assessments in federal prisons.

Two Petty Theft Arrests



Borden was rated high risk for future crime after she and a friend took a kid's bike and scooter that were sitting outside. She did not reoffend.

hidden effect of algorithms in American life.

We obtained the risk scores assigned to more than 7,000 people arrested in Broward County, Florida, in 2013 and 2014 and checked to see how many were charged with new crimes over the next two years, the [same benchmark used](#) by the creators of the algorithm.

The score proved remarkably unreliable in forecasting violent crime: Only 20 percent of the people predicted to commit violent crimes actually went on to do so.

When a full range of crimes were taken into account — including misdemeanors such as driving with an expired license — the algorithm was somewhat more accurate than a coin flip. Of those deemed likely to re-offend, 61 percent were arrested for any subsequent crimes within two years.

We also turned up significant racial disparities, just as Holder feared. In forecasting who would re-offend, the algorithm made mistakes with black and white defendants at

In 2014, then U.S. Attorney General Eric Holder warned that the risk scores might be injecting bias into the courts. He called for the U.S. Sentencing Commission to study their use. "Although these measures were crafted with the best of intentions, I am concerned that they inadvertently undermine our efforts to ensure individualized and equal justice," he said, adding, "they may exacerbate unwarranted and unjust disparities that are already far too common in our criminal justice system and in our society."

The sentencing commission did not, however, launch a study of risk scores. So ProPublica did, as part of a larger examination of the powerful, largely

Arizona, Colorado, Delaware, Kentucky, Louisiana, Oklahoma, Virginia, Washington and Wisconsin, the results of such assessments are given to judges during criminal sentencing.

Rating a defendant's risk of future crime is often done in conjunction with an evaluation of a defendant's rehabilitation needs. The Justice Department's National Institute of Corrections now encourages the use of such combined assessments at every stage of the criminal justice process. And a landmark sentencing [reform bill](#) currently pending in Congress would mandate the use of such assessments in federal prisons.

Two Petty Theft Arrests



In 2014, then U.S. Attorney General Eric Holder warned that the risk scores might be injecting bias into the courts. He called for the U.S. Sentencing Commission to study their use. "Although these measures were crafted with the best of intentions, I am concerned that they inadvertently undermine our efforts to ensure individualized and equal justice," he said, adding, "they may exacerbate new forms of injustice, disparities that are already too common in our criminal justice system and in our society."

Where did this data come from?

Borden was rated high risk for future crime after she and a friend took a kid's bike and scooter that were sitting outside. She did not reoffend.

The sentencing commission did not, however, ban the use of risk scores. So ProPublica did, as part of a larger examination of the powerful, largely

hidden effect of algorithms in American life.

We obtained the risk scores assigned to more than 7,000 people arrested in Broward County, Florida, in 2013 and 2014 and checked to see how many were charged with new crimes over the next two years, the [same benchmark used](#) by the creators of the algorithm.

The score proved remarkably unreliable in forecasting violent crime: Only 20 percent of the people predicted to commit violent crimes actually went on to do so.

When a full range of crimes were taken into account — including misdemeanors such as driving with an expired license — the algorithm was somewhat more accurate than a coin flip. Of those deemed likely to re-offend, 61 percent were arrested for any subsequent crimes within two years.

We also turned up significant racial disparities, just as Holder feared. In forecasting who would re-offend, the algorithm made mistakes with black and white defendants at

[MACHINE BIAS](#)

Facebook (Still) Letting Housing Advertisers Exclude Users by Race



After ProPublica revealed last year that Facebook advertisers could target housing ads to whites only, the company announced it had built a system to spot and reject discriminatory ads. We retested and found major omissions.



by Julia Angwin, Ariana Tobin and Madeleine Varner, Nov. 21, 2017, 1:23 p.m. EST



Facebook CEO Mark Zuckerberg speaks in San Jose, California, in October 2016. (David Paul Morris/Bloomberg via Getty Images)

FOLLOW PROPUBLICA[Twitter](#)[Facebook](#)[Podcast](#)[RSS](#)[✉ Get our stories by email.](#) Email[Subscribe](#)**MOST POPULAR STORIES**[Most Read](#) [Most Emailed](#)[Billion-Dollar Blessings](#)[We're Hiring, a Lot. Here's What We're Looking For.](#)[The Company Michael Cohen Kept — "Trump, Inc."](#)

FEATURE

Policing the Future

In the aftermath of Michael Brown's death, St. Louis cops embrace crime-predicting software.

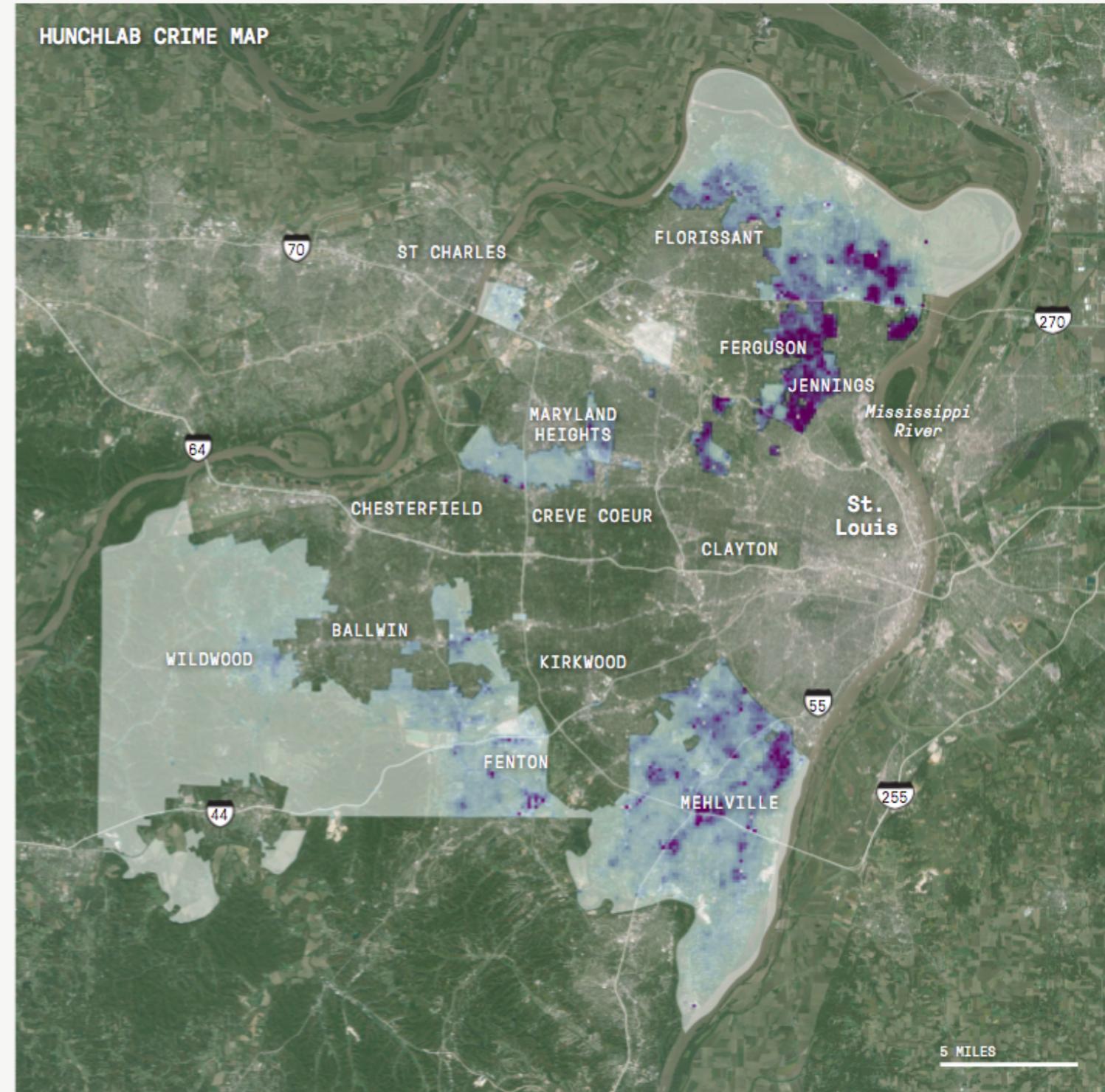


Maurice Chammah, with additional reporting by Mark Hansen. Policing the Future.

<https://www.themarshallproject.org/2016/02/03/policing-the-future>

Where the St. Louis County Police Patrol

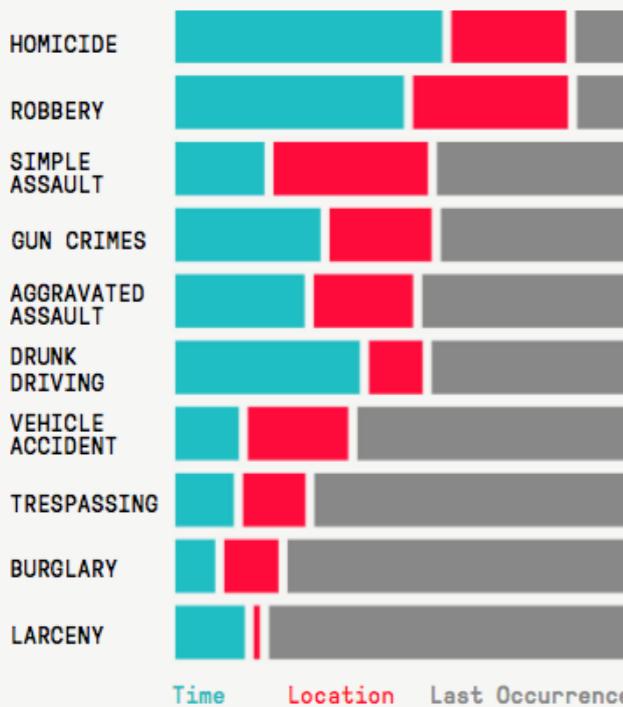
Dozens of small, local municipal agencies handle policing in parts of St. Louis County. The St. Louis County Police Department covers areas not policed by the “munis,” including the city of Jennings, Mo. The ■ DARKER AREAS in the map show the areas within their jurisdiction that HunchLab has identified as high risk.



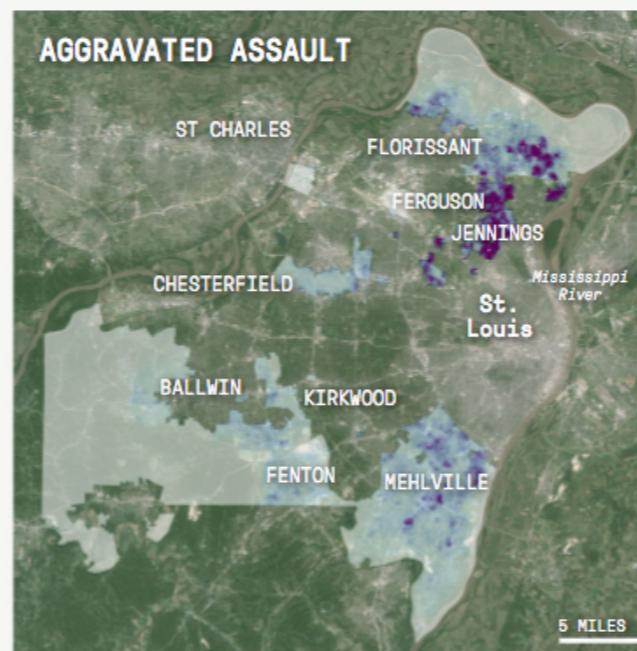
SOURCE: HUNCHLAB

Maurice Chammah, with additional reporting by Mark Hansen. Policing the Future.
<https://www.themarshallproject.org/2016/02/03/policing-the-future>

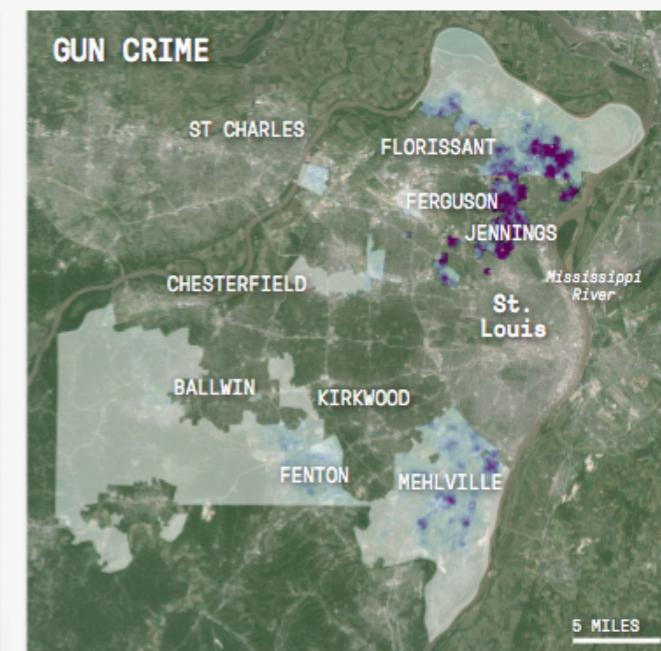
In St. Louis, the HunchLab algorithm took the 10 crimes that the police department had selected, calculated the risk-level for each, and combined them to determine where patrols would have the most impact.



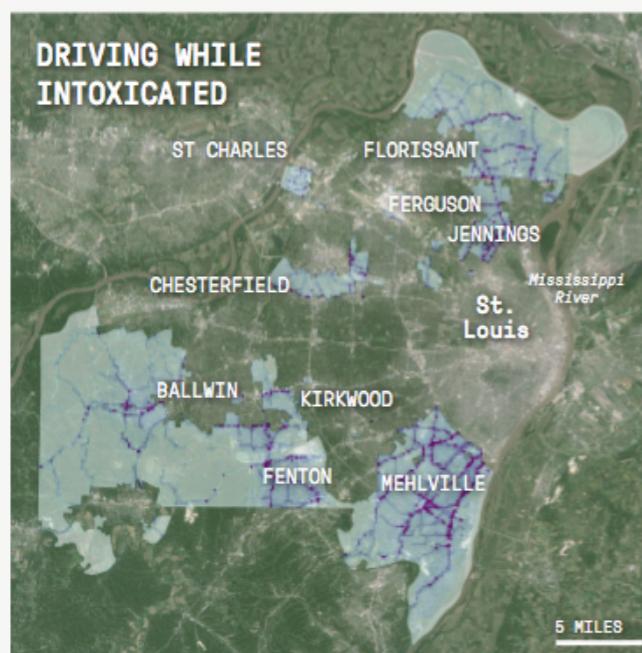
Time Location Last Occurrence



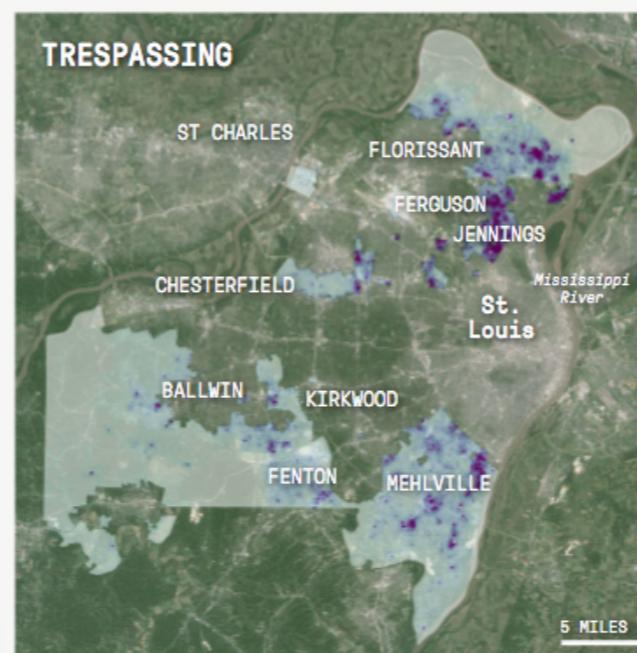
Aggravated assault (assault with a dangerous weapon) makes up 18.5 percent of the overall risk score assigned to a cell. The darkest regions on this map represent cells with a 1 in 320 chance of at least one aggravated assault taking place there during the shift.



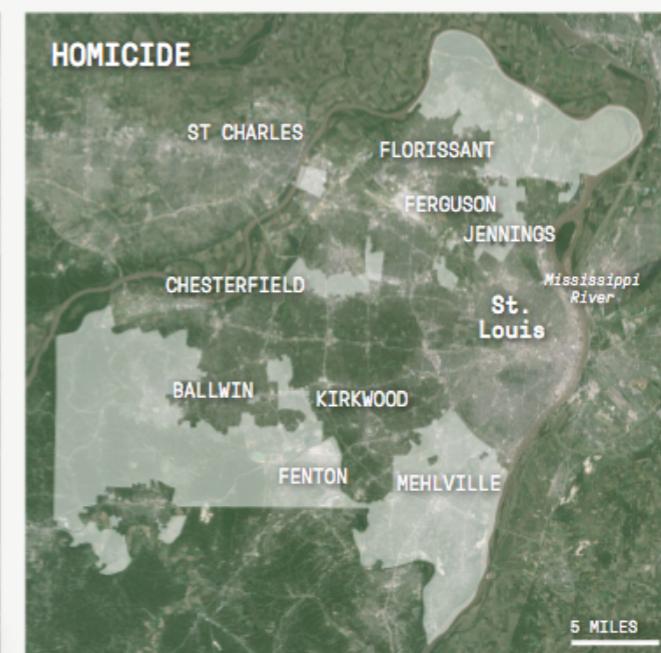
Gun crime (all homicides, robberies, and aggravated assaults with a firearm) makes up about 16.5 percent of the overall risk score. The darkest regions represent a 1 in 850 chance of at least one gun crime taking place.



Driving while intoxicated makes up 10 percent of the total risk score. The darkest regions represent a 1 in 1,300 chance of at least one DWI taking place.



Trespassing makes up about 10 percent of the total risk score. The darkest regions represent cells a 1.7 percent chance of at least one act of trespassing taking place.



Homicides make up 0.66 percent of the total risk score assigned to a cell. The two darkest cells on this map present a 3 percent chance of at least one homicide taking place.

SOURCE: HUNCHLAB

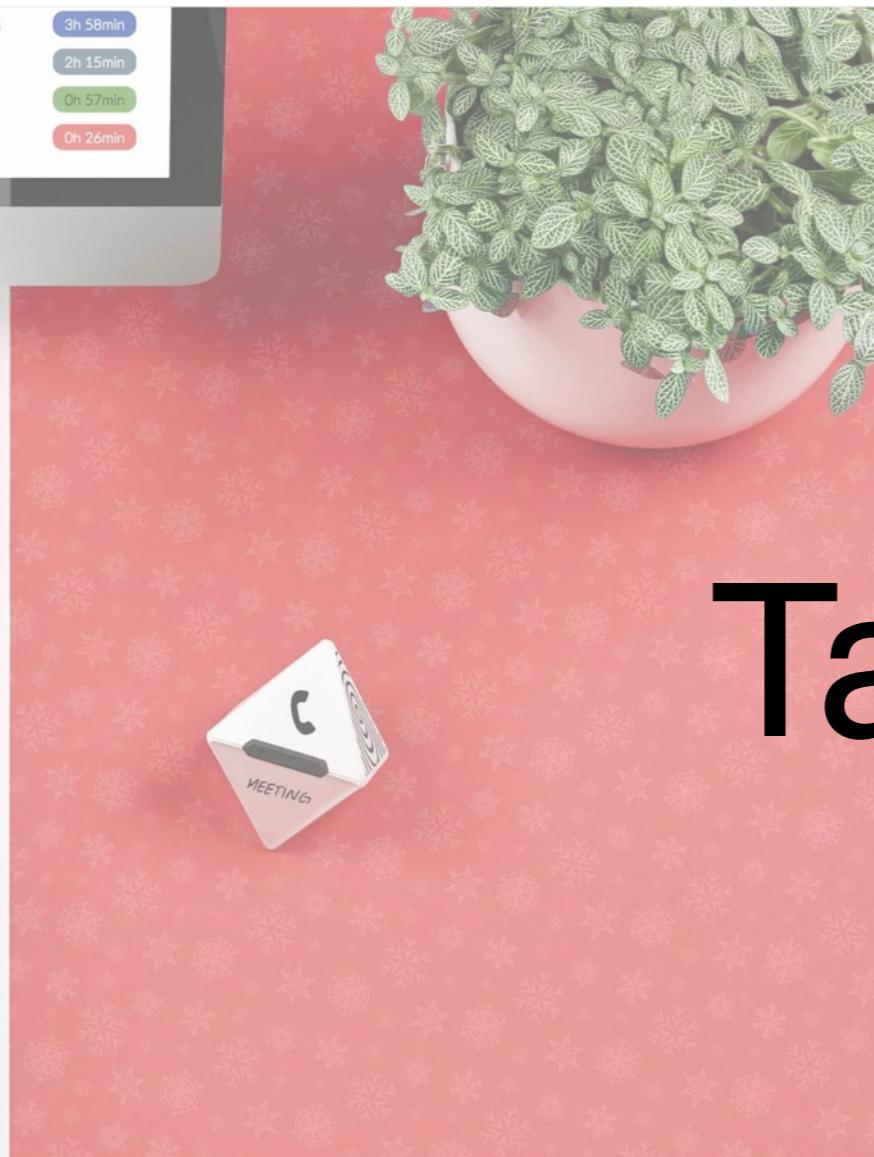
Maurice Chammah, with additional reporting by Mark Hansen. Policing the Future.
<https://www.themarshallproject.org/2016/02/03/policing-the-future>

 Instagram 

 timeular
Sponsored

► Facebook Campaigns 3h 58min
► App Store 2h 15min
► Internal meeting 0h 57min
► Google Adwords 0h 26min

34min logged/day
45% Entries




Learn More >

    
14 likes

 ...

 Virgin Mobile USA
9 hrs · 

Member Benefits makes it easy for you to reward yourself for sitting through those awkward conversations with Aunt Shirley over the holidays! Save up to 20% off fares on your next getaway with Virgin America & Alaska Airlines.

More details at: <https://virg.in/obb>



Tailoring

Racial bias alleged in Google's ad results

Names associated with blacks prompt link to arrest search



Ad related to latanya sweeney ⓘ
[Latanya Sweeney Truth](#)
www.instantcheckmate.com/
Looking for Latanya Sweeney? Check Latanya Sweeney's Arrests.

Ads by Google
[Latanya Sweeney, Arrested?](#)
1) Enter Name and State. 2) Access Full Background Checks Instantly.
www.instantcheckmate.com/

[Latanya Sweeney](#)
Public Records Found For: Latanya Sweeney. View Now.
www.publicrecords.com/

[La Tanya](#)
Search for La Tanya Look Up Fast Results now!
www.ask.com/La+Tanya

(c)

(d)

LATANYA SWEENEY

Web page results of ads that appeared on-screen when Harvard professor Latanya Sweeney typed her name in a google search. Ads featured services for arrest records. Sweeney conducted a study that concluded searches with "black sounding" names are more likely to get results with ads for arrests records and other negative information.

By [Hiawatha Bray](#) | GLOBE STAFF FEBRUARY 06, 2013

(limited) transparency from companies

Your information

About you Your categories

The categories in this section help advertisers reach people who are most likely to be interested in their products, services, and causes. We've added you to these categories based on information you've provided on Facebook and other activity.

Birthday in April

Away from family

Away from hometown

US politics (very liberal)

Gmail users

Close friends of expats

Facebook access (mobile): smartphones and tablets

African American (US)

Facebook access (mobile): Samsung Android mobile devices

Frequent Travelers

Android: 360 degree media supported

Facebook access (mobile)

See More

<https://www.facebook.com/ads/preferences>

Close ^

Ads Personalization

Make the ads you see more useful to you when using Google services (ex. Search, YouTube).

TOPICS YOU LIKE

TOPICS YOU DON'T LIKE (0)

Remove topics you don't like and add ones you do to make the ads you see more useful to you. Topics will also be added as you use some Google services (ex: when you watch a video on YouTube). We're working to include topics from other Google services.

Arts & Entertainment × Autos & Vehicles × Beauty & Fitness ×

Blues × Books & Literature × Business & Industrial ×

Business & Productivity Software × Business News × Cats ×

Classical Music × Comics & Animation × Computers & Electronics ×

Cooking & Recipes × Coupons & Discount Offers × Credit Cards ×

Cycling × Education × Enterprise Technology ×

Fashion & Style × Finance × Fitness ×

Folk & Traditional Music × Food & Drink × Hair Care ×

+ NEW TOPIC

VIEW 23 MORE

<https://adsettings.google.com/authenticated>



What do others think of you?

See a profile of what your browsing history suggests about you.
Paste URLs from your web history into the box below. (Help)

Browsing history

Submit

[About](#) | [Samples](#) | [Comparison](#) | [Privacy](#)

This service is currently under development. Stay tuned...

How unique are you?

Enter your ZIP code, date of birth, and gender to see how unique you are (and therefore how easy it is to identify you from these values).

Date of Birth

Gender Male Female

5-digit ZIP

Submit

[About](#) | [Samples](#) | [Harvard](#) | [Harvard Multi Years](#)

data releases contain many more fields. It was found that 87% (216 million of 248 million) of the population in the United States had reported characteristics that likely made them unique based only on {5-digit ZIP, gender, date of birth}. About half of the U.S. population (132 million of 248 million or 53%) are likely to be uniquely identified by only {place, gender, date of birth}, where place is basically the city, town, or municipality in which the person resides. And even at the county level, {county, gender, date of birth} are likely to uniquely identify 18% of the U.S. population. In general, few characteristics are needed to uniquely identify a person.





Dr. Latanya Sweeney



Carnegie Mellon

DATA PRIVACY LAB

609 Social Security Number

Matching a Person to a SSN

Using publicly available information about 90% of the U.S. population, the Data Privacy Lab's SSN Validation Server identifies the issuing state, date issued, and expiration date for 90% of all SSNs.

Sample uses:

- Job Applications
- Apartment Rentals
- Insurance Claims
- Student Applications

Enter the
Social Security
Number you
would like to
check.

SSN and enter
the last four digits
of the SSN.
You can also
selectively learn

Results for SSN 078-07-0001

Geography	New York
Date of issuance	Serial number 1000
Date of birth	07/01/1980 to 10/01/1980
(8-digit profile)	8078-07-0001

If the person presenting the SSN fails to list an acknowledged New York as a prior residence, then it is extremely unlikely that the provided SSN was issued to that person.

Results for SSN 078-07-0002

Geography
Date of
issuance
Results
beginning
that the



Dr. Jake Porway







Adam Harvey **CVDazzle**
<https://cvdazzle.com/>

Data for Black Lives

Data as protest. Data as accountability. Data as collective action.

Sign up for more information.

Data for Black Lives seeks to mobilize scientists around racial justice issues. At our conference in November, we will convene over two hundred data scientists, computer programmers, racial justice activists, and elected officials to discuss the role that data can and should play in Black communities. Join us January 2019 at MIT for our second annual conference

The Team



Journalists can be
data science
superheroes



Arizona, Colorado, Delaware, Kentucky, Louisiana, Oklahoma, Virginia, Washington and Wisconsin, the results of such assessments are given to judges during criminal sentencing.

Rating a defendant's risk of future crime is often done in conjunction with an evaluation of a defendant's rehabilitation needs. The Justice Department's National Institute of Corrections now encourages the use of such combined assessments at every stage of the criminal justice process. And a landmark sentencing reform bill currently pending in Congress would mandate the use of such assessments in federal prisons.

Two Petty Theft Arrests



Borden was rated high risk took a kid's bike and scooter offend.

hidde

We of

Count

crime

algori

The s

the people predicted to commit violent crimes actually went on to do so.

When a full range of crimes were taken into account — including misdemeanors such as driving with an expired license — the algorithm was somewhat more accurate than a coin flip. Of those deemed likely to re-offend, 61 percent were arrested for any subsequent crimes within two years.

We also turned up significant racial disparities, just as Holder feared. In forecasting who would re-offend, the algorithm made mistakes with black and white defendants at roughly the same rate but in very different ways.

The New York Times

SEARCH

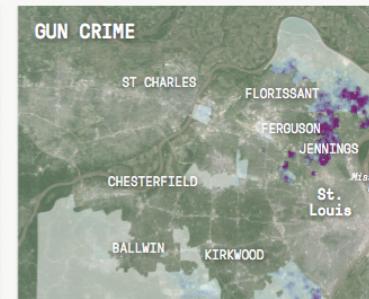
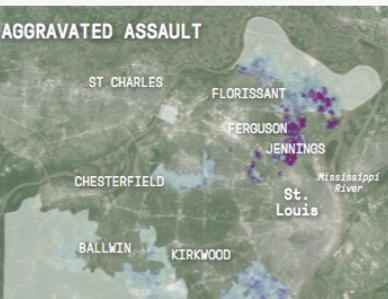
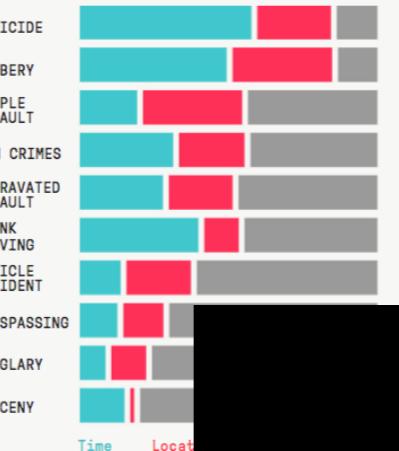
POLITICS

How Trump Consultants Exploited the Facebook Data of Millions

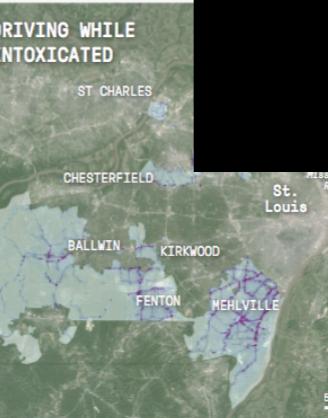
Leer en español

By MATTHEW ROSENBERG, NICHOLAS CONFESSORE and CAROLE CADWALLADR MARCH 17, 2018

In St. Louis, the HunchLab algorithm took the 10 crimes that the police department had selected, calculated the risk level each, and combined them to determine where patrols would have the most impact.

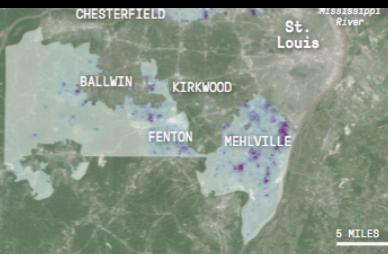


The Marshall Project

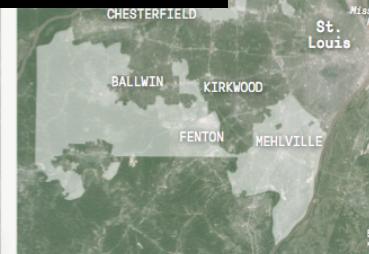


Driving while intoxicated makes up 10 percent of the total risk score. The darkest regions represent cells a 1 in 1,300 chance of at least one DWI taking place.

SOURCE: HUNCHLAB



Trespassing makes up about 10 percent of the total risk score. The darkest regions represent cells a 1.7 percent chance of at least one act of trespassing taking place.



Homicides make up 0.66 percent of the total risk score assigned to a cell. The two darkest cells on this map present a 3 percent chance of at least one homicide taking place.

Racial bias alleged in Google's ad results

Names associated with Black prompt link to arrest search

Ad related to latanya sweeney

Latanya Sweeney | www.instantcheckmate.com

Looking for Latanya Sweeney? Get Latanya Sweeney's Arrests, Criminal History, and more!

Ads by Google

Latanya Sweeney arrested?

1) Enter Name and State. 2) Access Full Background Checks Instantly

www.instantcheckmate.com/

Latanya Sweeney

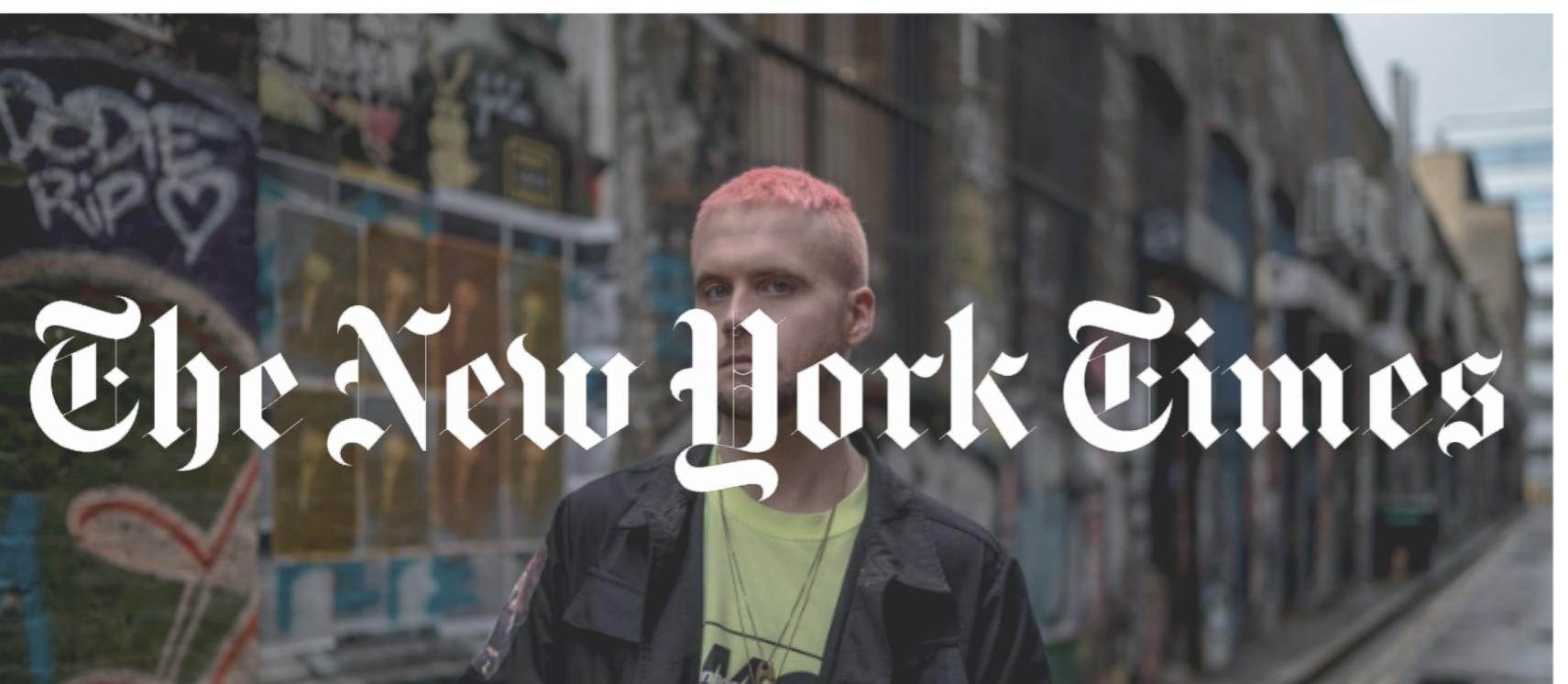
Public Records for Latanya Sweeney - View Now

www.publicrecords.com

La Tanya

Search for La Tanya - Look Up Past Records now

www.ask.com/LaTanya



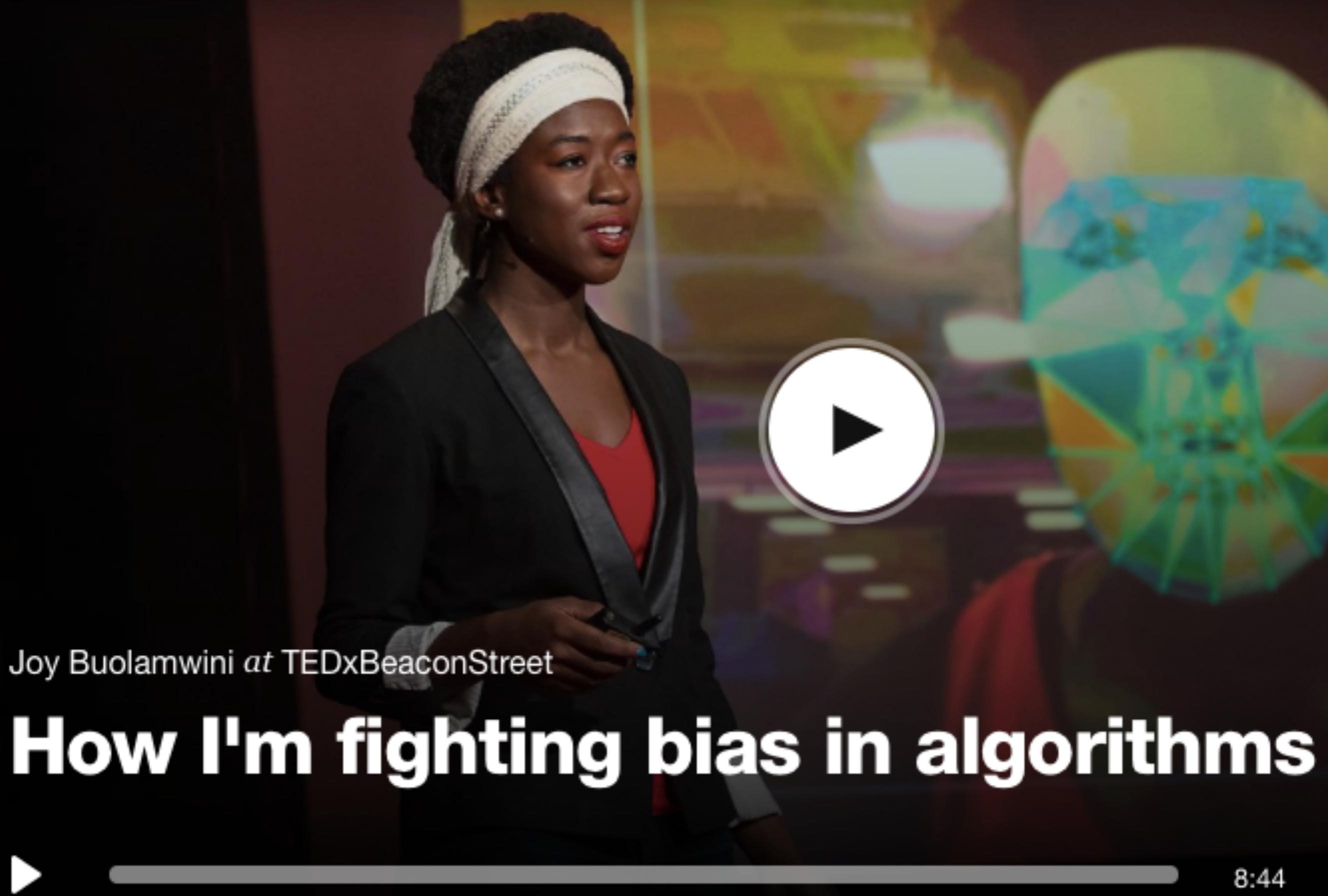


Build diverse teams



Jer Thorp. Turning Data Around.
<http://bit.ly/TurningDataAround>

1. Design data systems for the well-being of the people from whom the data is taken
2. Whenever possible, provide mechanisms for feedback
3. Honor the complexity of individual and community realities
4. Create real, functioning data publics



Joy Buolamwini at TEDxBeaconStreet

How I'm fighting bias in algorithms



8:44

Joy Buolamwini. How I'm fighting bias in algorithms.
<http://bit.ly/FightingBiasInAlgorithms>

Matt Mitchell. Cyber JimCrow: Virtual
Public Housing and Poor Doors in
Digital Security & Surveillance.
<https://vimeo.com/232659054>



00:20



CRITICAL QUESTIONS FOR BIG DATA

Provocations for a cultural, technological, and scholarly phenomenon

The era of Big Data has begun. Computer scientists, physicists, economists, mathematicians, political scientists, bio-informaticists, sociologists, and other scholars are clamoring for access to the massive quantities of information produced by and about people, things, and their interactions. Diverse groups argue about the potential benefits and costs of analyzing genetic sequences, social media interactions, health records, phone logs, government records, and other digital traces left by people. Significant questions emerge. Will large-scale search data help us create better tools, services, and public goods? Or will it usher in a new wave of privacy incursions and invasive marketing? Will data analytics help us understand online communities and political movements? Or will it be used to track protesters and suppress speech? Will it transform how we study human communication and culture, or narrow the palette of research options and alter what ‘research’ means? Given the rise of Big Data as a socio-technical phenomenon, we argue that it is necessary to critically interrogate its assumptions and biases. In this article, we offer six provocations to spark conversations about the issues of Big Data: a cultural, technological, and scholarly phenomenon that rests on the interplay of technology, analysis, and mythology that provokes extensive utopian and dystopian rhetoric.

Keywords Big Data; analytics; social media; communication studies; social network sites; philosophy of science; epistemology; ethics; Twitter

(Received 10 December 2011; final version received 20 March 2012)

Technology is neither good nor bad; nor is it neutral . . . technology’s interaction with the social ecology is such that technical developments frequently have environmental, social, and human consequences that go far beyond the immediate purposes of the technical devices and practices themselves. (Kranzberg 1986, p. 545)

WEAPONS OF MATH DESTRUCTION



HOW BIG DATA INCREASES INEQUALITY
AND THREATENS DEMOCRACY

CATHY O'NEIL

A figure of Shuri from Black Panther stands prominently in the center-right of the frame. She is wearing her signature orange and gold suit with a black belt and gauntlets. Her hair is shaved, and she has a headband with geometric patterns. She is holding a long, ornate spear with a blue feather and a red and gold tip. The background is dark and out of focus.

Thank you

Amelia McNamara @AmeliaMN
University of St Thomas
Department of Computer & Information Sciences