

CSCI 15: AN INTRODUCTION TO THE MODERN INTERNET

Lecture 5: Privacy

PLAN FOR TODAY

- Talk about privacy
- Mostly slides with discussion, some activities

- No governments today



WHAT IS PRIVACY?

WHEN DO YOU WANT PRIVACY?

- Sometimes we give up our privacy
- Sometimes it's taken
- Sometimes it's not so clear

- What do we want out of privacy?

K-ANONYMITY

- One way to quantitatively measure privacy
- Definition: given some information, how many people share that information?



K-ANONYMITY

- Classic example: “anonymous” storage using gender, ZIP, date of birth
- <https://aboutmyinfo.org/identity>
- Wait a minute—what about privacy?

INTERSECTION ATTACK

- Without k-anonymity, two anonymous databases can be used to deanonymize



IMPLEMENTING K-ANONYMITY

- Background: what do websites know about your passwords?
 - They SHOULD know nothing (store an encoded version)
 - Sometimes doesn't happen
 - Bad implementation + hack = hackers know your passwords
 - Cheap, widely available tools that list common passwords

IMPLEMENTING K-ANONYMITY

- <https://www.troyhunt.com/ive-just-launched-pwned-passwords-version-2/#cloudflareprivacyandkanonymity>
- Database of passwords, tells you if yours has been cracked in a major hack
- Do we want to give them our info!? How can we avoid this?

WHAT'S CHANGED SINCE 1997?

- MUCH more data storage!
- Much more easily accessible
- Much more of life involves providing information

But also:

- More awareness!
 - (Less caring?)

WEAKNESSES OF K-ANONYMITY?

Let's say I get a $k > 1$. How can this still be attacked?

- Homogeneity: what if we share key information?

	Age	Gender	Test Score
1	12	M	79
2	12	M	98
3	13	F	99
4	13	F	99
5	14	M	82
6	14	M	90

What do I know about the test score of a female in the class?

WEAKNESSES OF K-ANONYMITY?

Let's say I get a $k > 1$. How can this still be attacked?

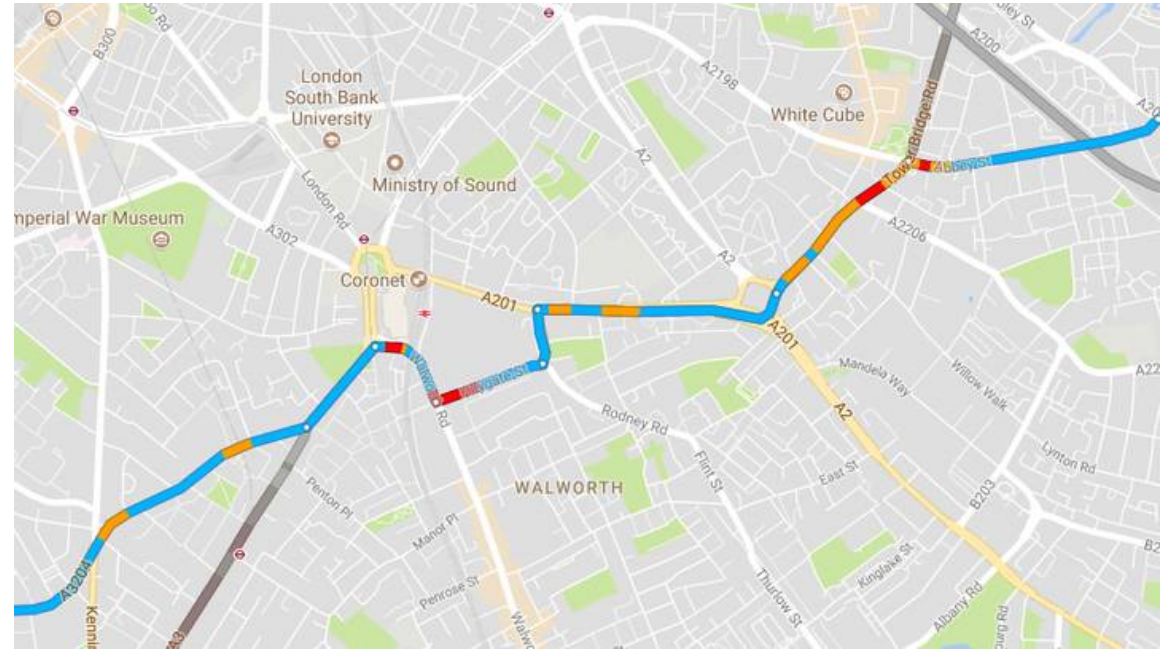
- Background knowledge

	Age	Gender	Test Score
1	12	M	79
2	12	M	98
3	13	F	99
4	13	F	99
5	14	M	82
6	14	M	90

I find out that Barry's (age 12) final average was a 96. Do I know anything about Andrew, who is also 12?

ANONYMITY IN THE AGE OF BIG DATA

- How does Google know this?
- What information does it leak?



DIFFERENTIAL PRIVACY

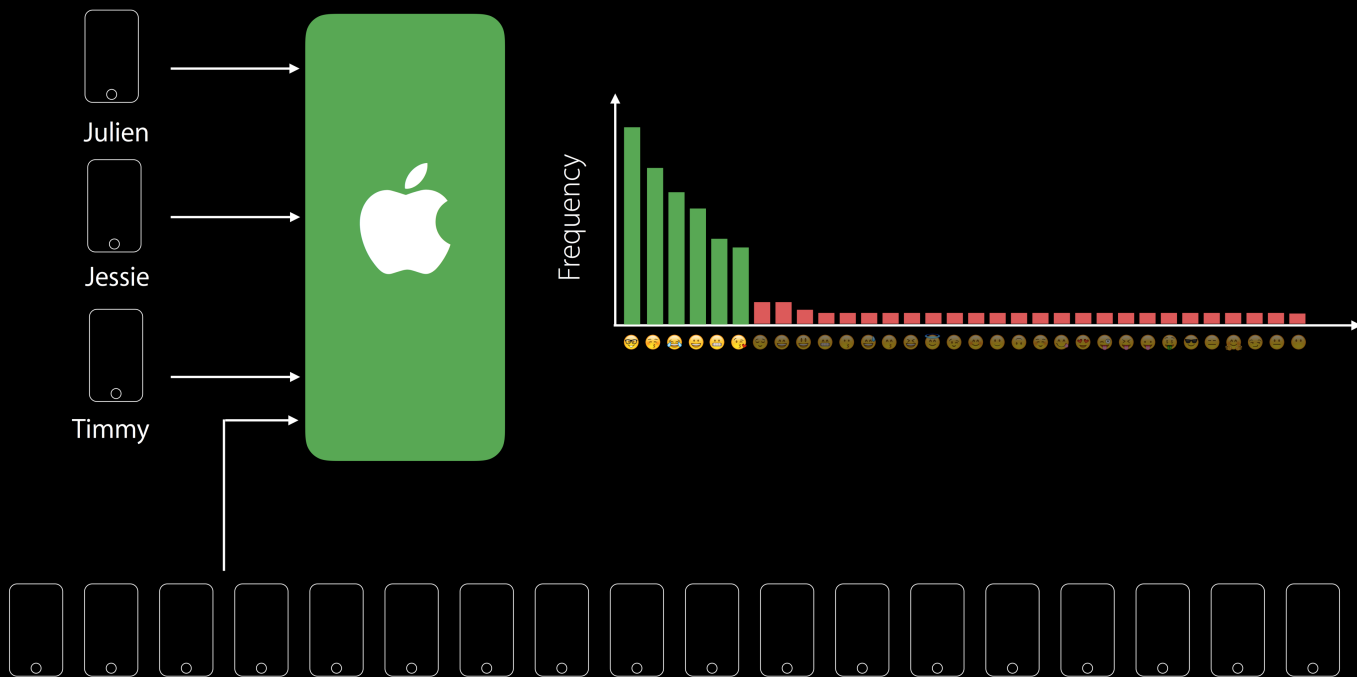
- Goal/definition: publish a large dataset of human behavior, without giving up information about any individuals
- Specifically: removing an individual from the database does not change the answer to any query
- How???
 - Add noise
 - Entire area of algorithmic and statistical study



WHY DIFFERENTIAL PRIVACY?

- Statistical algorithms don't need to be clever
- Any leaked information can be learned and used

Learning Popular Emojis with Privacy



- Apple is a major proponent (why?)
- Used for maps, words, emoji suggestions

DIFFERENTIAL PRIVACY USES

COST OF DIFFERENTIAL PRIVACY

Privacy in Pharmacogenetics: An End-to-End Case Study of Personalized Warfarin Dosing

Matthew Fredrikson*, Eric Lantz*, Somesh Jha*, Simon Lin†, David Page*, Thomas Ristenpart*
University of Wisconsin, Marshfield Clinic Research Foundation†*

Abstract

We initiate the study of privacy in pharmacogenetics, wherein machine learning models are used to guide medical treatments based on a patient's genotype and background. Performing an in-depth case study on privacy in personalized warfarin dosing, we show that suggested models carry privacy risks, in particular because attackers can perform what we call *model inversion*: an attacker, given the model and some demographic information about a patient, can predict the patient's genetic markers.

As differential privacy (DP) is an oft-proposed solution for medical settings such as this, we evaluate its ef-

machine learning over large patient databases containing clinical and genomic data. Prior works [36, 37] in non-medical settings have shown that leaking datasets can enable de-anonymization of users and other privacy risks. In the pharmacogenetic setting, datasets themselves are often only disclosed to researchers, yet the models learned from them are made public (e.g., published in a paper). *Our focus is therefore on determining to what extent the models themselves leak private information, even in the absence of the original dataset.*

To do so, we perform a case study of warfarin dosing, a popular target for pharmacogenetic modeling. Warfarin is an anticoagulant widely used to help prevent strokes in

HOW MUCH PRIVACY DO WE HAVE?

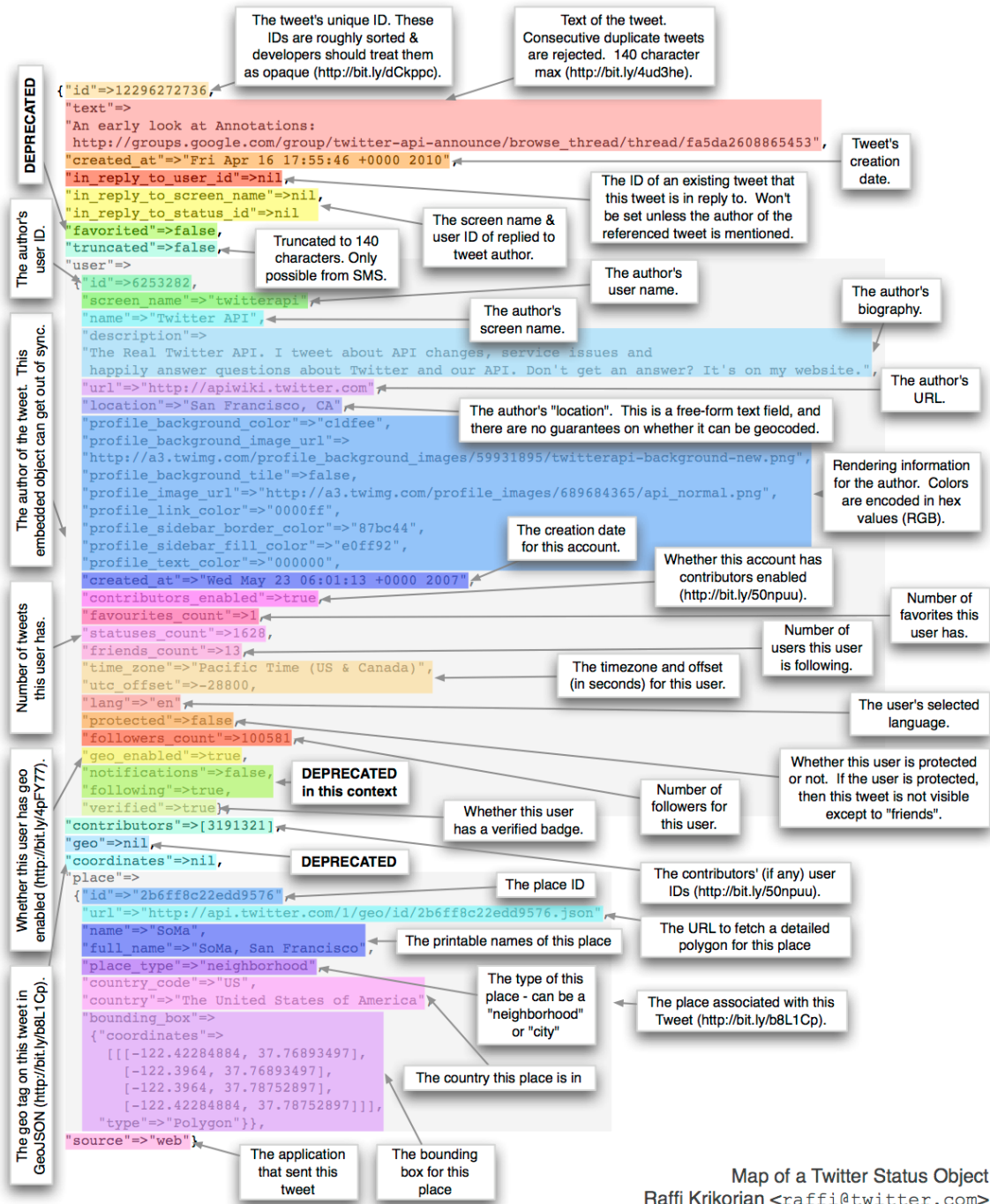
- Reminder: information transmitted over the internet
- All devices traversed (and anyone privy to traffic) can see:
 - Source
 - Destination
 - Content if not encrypted (if not HTTPS)

WHAT PRIVACY HAVE WE GIVEN UP?

- When do we knowingly give away privacy?
- What do we gain?

EXAMPLE: FACEBOOK

- What does Facebook know about us?
- Facebook -> settings -> ads -> interests
- What does Facebook not know about us?



EXAMPLE: CREDIT CARDS

- What does your credit card company know about you?
- How do they use it?

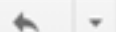
Capital One Fraud Protection Alert

Inbox x



Capital One <capitalone@email.capitalone.com>
to jim

12/16/11 ☆



Capital One

[Add us to your address book](#)

[Help prevent fraud](#)

[Log in to your account](#)

RE: Account ending in [REDACTED]

We recently noticed **potentially suspicious activity** on your VISA SIGNATURE account ending in 0336, and need to confirm that you authorized all of these charges. Please call **Customer Fraud Protection** as soon as possible at [800-427-9428](tel:800-427-9428).

If you are **calling from outside the United States**, the best way to reach us is by an operator-assisted collect call and Capital One® will accept the charges.

- Connect with a live international operator
- Tell the operator you are calling the U.S. collect
- Provide our number (00+1+804-934-2001) to the operator
- Give your name as "Capital One Customer" when prompted

If you have access to the Internet, you can look up how to call internationally from your specific location.

We would like to address this suspicious activity as soon as possible to **protect** your account. Thank you for your prompt attention to this matter.

Capital One
Fraud Protection Division

P.S. If you have already contacted us, thank you. There is no need to call again.



Hello, SEBASTIAN

Account Ending:



View Account

Make a Payment

Manage Alerts Preferences

Based on recent transactions, it appears that you will be traveling soon. We've made a notation on your account(s), so you don't need to contact us to advise us of your travel plans.

We're pleased to offer free and easy to use solutions that help make sure any fraud concerns can be resolved quickly. Before your trip, we recommend that our traveling Card Members do the following:

- **Download the Amex Mobile app** from the [Apple App Store](#) or [Google Play Store](#) and enable Push notifications. This will allow us to alert you of potential fraudulent activity on your account and, in many cases, allow you to resolve concerns instantly.
- **Confirm your current mobile phone number** is on file by logging into your account online. This will allow us to reach you quickly, should any concerns arise.

Thank you for your Card Membership.

American Express Customer Care

Yes, Your Credit Card Company Is Selling Your Purchase Data To Online Advertisers

Jim Edwards Apr 16, 2013, 9:02 AM



Depending on how you feel about online shopping, this is either old news or a huge betrayal of consumer trust: Mastercard and [American Express](#) are selling your data to online advertisers who then use it to target you with ads.



By amazing coincidence, she's start seeing shoe ads online soon. Flickr / Orin Zebest

As Ad Age notes, they're not shouting very loud about it for fear of a backlash.

Here's [what we learned today from reporter Kate Kaye](#):

1. Mastercard began doing this two and a half years ago.

This data is anonymized---only gives (say) number of purchases in a given ZIP code

EXAMPLE: CREDIT CARDS

- What does your credit card company know about you?
- How do they use it?

PRIVACY MIDDLE GROUND

- Sometimes we don't realize we're giving up privacy, but we may not be that surprised
- Google location timeline

PRIVACY MIDDLE GROUND

- What does a department store (Target, Walmart, Stop and Shop) know about me?
- How can they track me, my purchases, and my interests?
- What can't they (or won't they) track?

RETAIL IN 2019


- <https://www.businessinsider.com/walmart-kroger-walgreens-increasing-camera-sensor-data-collection-2019-4>

PRIVACY MIDDLE GROUND (?)

HOLIDAY MONEY

Malls track shoppers' cell phones on Black Friday

By Annalyn Censky @CNNMoneyTech November 22, 2011: 11:48 AM ET



COURTESY: FOREST CITY

Through this signage at Promenade Temecula, the mall is notifying shoppers that their phones may be tracked as they move throughout the premises.

NEW YORK (CNNMoney) -- Attention holiday shoppers: your cell phone may be tracked this year.

Starting on Black Friday and running through New Year's Day, two U.S. malls -- Promenade Temecula in southern California and Short Pump Town Center in Richmond, Va. -- will track guests' movements by monitoring the signals from their cell phones.

Mos

What v

Milleni

7 traits

Big Da

Your c

Tec

APPLE

Did J

TECH 1

ENTER

How

Tec

job ti

Grapt

Softw

SEE A

COOKIES

- First party: information stored on your computer by the website
- Third party: stored by another company (ads are most common)
- When are these good?
- When are these bad? (Who can see them?)

COOKIES

- Automatically login
- Website history
- Track across websites (press button to login)
- Ad targeting

CREDIT AGENCIES

THE VERGE  TWITTER  FACEBOOK



143 million compromised Social Security numbers: everything you need to know about the Equifax hack

Contributors: Verge Staff

   SHARE

It has been marked as the worst data breach in US history. Attackers stole half the US population's Social Security numbers from Equifax this spring, but the company only notified people in September. The fallout has been swift, with government agencies looking into the incident, class action lawsuits being filed, and consumers demanding free credit freezes.

Follow along with all of the updates as this story develops.

18 TOTAL UPDATES SINCE
SEP 7, 2017, 5:11PM EDT

 FOLLOW

July 22, 2019

21 comments

Equifax agrees to settlement of up to \$700 million over 2017 data breach

By Jon Porter | @JonPorty

It will have to pay as much as \$20,000 per person

June 29, 2019

23 comments

Former Equifax executive sentenced to prison for

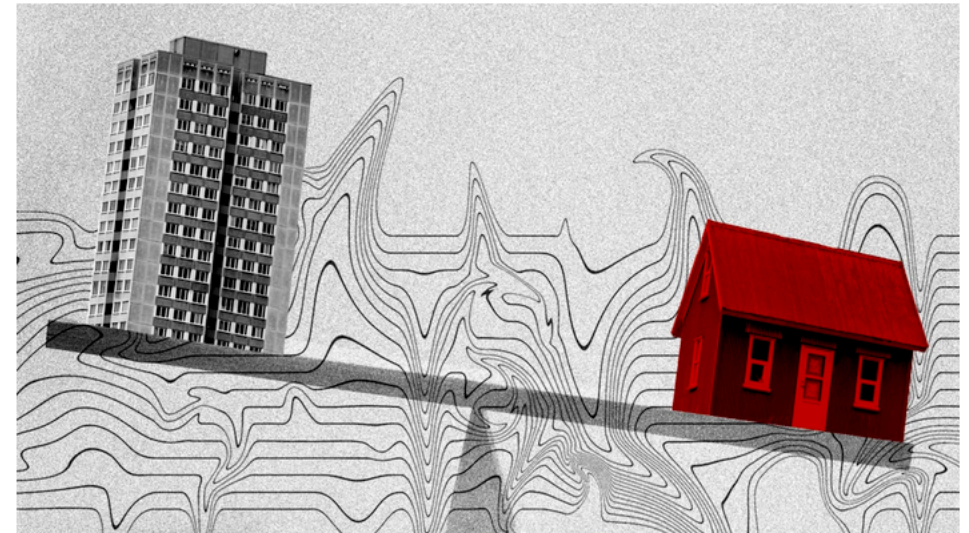
FASTCOMPANY

CO.DESIGN TECH WORK LIFE CREATIVITY IMPACT PODCASTS VIDEO NEWS RECOMMEND

04-06-19

Now wanted by big credit bureaus like Equifax: Your alternative data

Lenders and credit bureaus say a bold new data push can expand credit to more consumers, but some worry the shift could sting the people it's meant to help.



[Photo: [Simone Hutsch/Unsplash](#); [Luke Stackpoole/Unsplash](#); [LightFieldStudios/iStock](#)]



BY STEVEN MELENDEZ LONG READ

This story is part of *The Privacy Divide*, a series that explores the fault lines, disparities, and paradoxes that have developed around data privacy and its broader impacts on society. [Read the series here.](#)

PRIVACY WE DON'T KNOW WE'RE LOSING

3,113,177 views | Feb 10, 2012, 11:02am

How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did



Kashmir Hill Former Staff
Tech

Welcome to *The Not-So Private Parts* where technology & privacy collide

This article is more than 2 years old.

f

Every time you go shopping, you share intimate details about your consumption patterns with retailers.

t

And many of those retailers are studying those details to figure out what you like, what you need, and which coupons are most likely to make you happy. [Target](#) TGT +0%, for example, has figured out how to data-

mine its way into your womb, to figure out whether you have a baby on the way long before you need to start buying diapers.

Charles Duhigg outlines in the [New York Times](#) how Target tries to hook parents-to-be at that crucial moment before they turn into rampant -- and loyal -- buyers of all things pastel, plastic, and



Target has got you in its aim

“

[Pole] ran test after test, analyzing the data, and before long some useful patterns emerged. Lotions, for example. Lots of people buy lotion, but one of Pole's colleagues noticed that women on the baby registry were buying larger quantities of unscented lotion around the beginning of their second trimester. Another analyst noted that sometime in the first 20 weeks, pregnant women loaded up on supplements like calcium, magnesium and zinc. Many shoppers purchase soap and cotton balls, but when someone suddenly starts buying lots of scent-free soap and extra-big bags of cotton balls, in addition to hand sanitizers and washcloths, it signals they could be getting close to their delivery date.

Or have a rather nasty infection...

“

As Pole's computers crawled through the data, he was able to identify about 25 products that, when analyzed together, allowed him to assign each shopper a “pregnancy prediction” score. More important, he could also estimate her due date to within a small window, so Target could send coupons timed to very specific stages of her pregnancy.



- Transmits information over very short ranges
- Used in credit cards and passports, retail
- Is this an issue?
 - Not as much in 2019

RFID

RFID COSTS AND USAGES

- In 2011, tag cost 9 cents
- Many products you buy use it
- EZ-Pass
- Library books
- Races

WEBSITE FINGERPRINTING

- You have a fingerprint on the internet!
- What information must your browser give out?
- <https://panopticlick.eff.org>

PRIVACY WE DON'T KNOW WE'RE LOSING

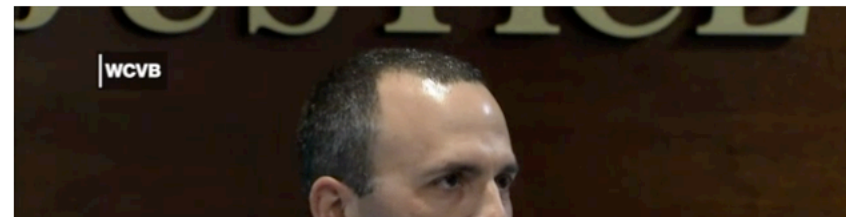
- DNA websites
- Pharmacies (anonymized)

Man arrested in 1993 cold case rape after he's identified by DNA: Police

The suspect was 28 at the time and not known to the victim.

By [Emily Shapiro](#)

October 15, 2019, 3:19 PM • 5 min read



EMERGING TECH

OPINION

PRIVACY & SECURITY

Do iRobot's cloud-stored maps of my home represent a major privacy risk?



Brian Jackson @brianjjackson

Published: December 4th, 2018

I guess when I tweeted about using a cloud-connected robot that maps the inside of my home in conjunction with an always-on microphone that persistently uploads data over the Internet, I should have expected privacy advocates might be triggered.

Rightly so, as it's only been in the last couple of years that my home has suddenly

SOLUTION

Jake Ragusa, Technology Director for the Ascension Parish School Board, credits the Fujitsu fi-Series scanners with eliminating the need for a \$2 million dollar expansion planned to deal with space issues.

[Read More](#)

Related Content





- Cell phones?
 - Conversations resulting in ads?
- No (probably)
 - They say they don't
 - Too many resources
- Why not?
 - They don't need to

WHO IS LISTENING TO US NOW?

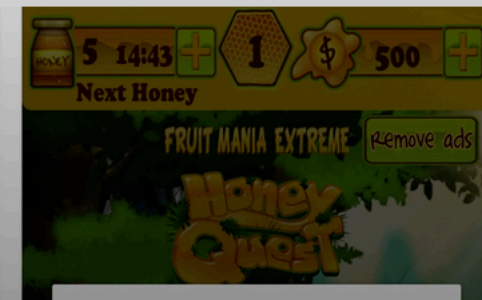
OK, ALSO YES (BUT NOT LIKE YOU'D THINK)

- Evidence that some small apps retain “fingerprints”
- Can keep track of (say) what you watch on TV, or what ads you view

www.nytimes.com/2017/12/28/business/media/alphonso-app-tracking.html?_r=1

The New York Times

*That Game on Your Phone
May Be Tracking What
You're Watching on TV*





- Yes (have to)
- Also workers listening to anonymized requests (2019)

WHAT ABOUT SMART SPEAKERS?

SMART SPEAKERS AND CRIME



Join Extra Crunch

Login

Search Q

Startups

Apps

Gadgets

Videos

Audio

Newsletters

Extra Crunch

Advertise

Events

—

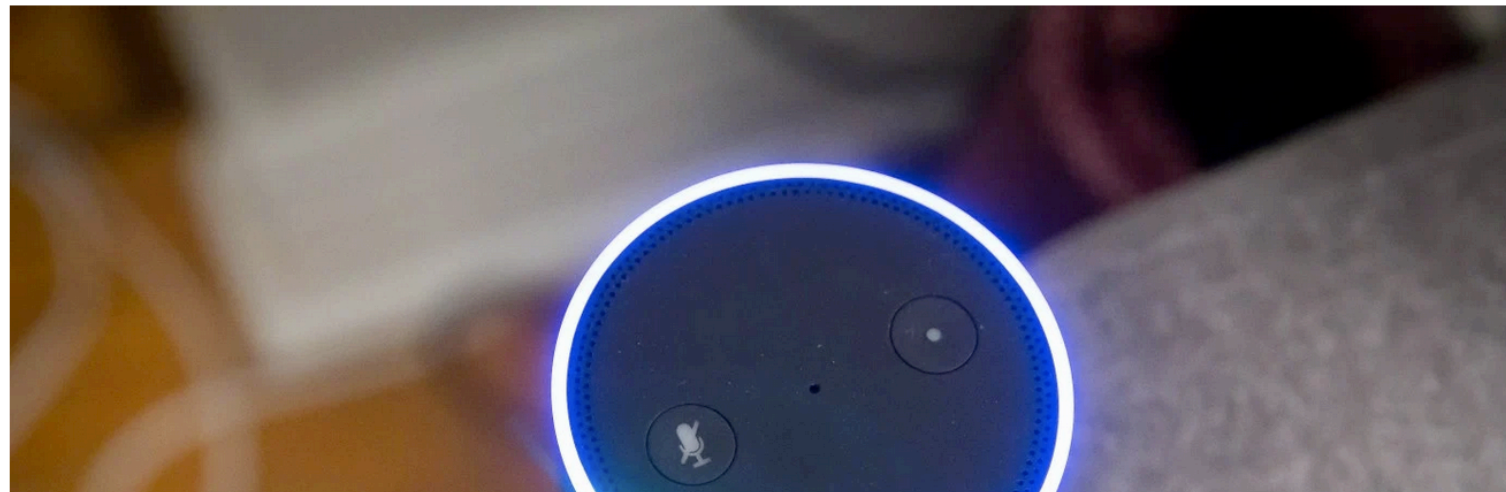
More

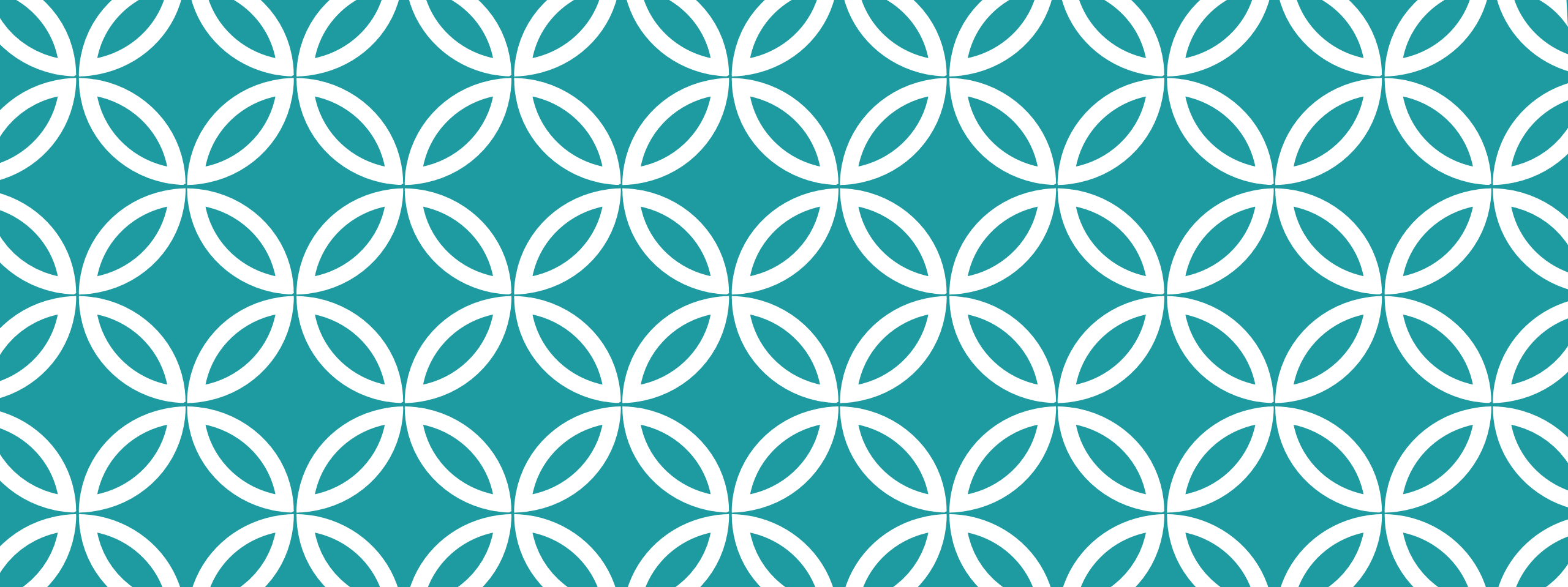
Judge orders Amazon to turn over Echo recordings in double murder case



Zack Whittaker @zackwhittaker / 11:46 am EST • November 14, 2018

Comment





WHAT CAN WE DO?



GDPR

- European law (2016)
- Websites must notify people of cookies
- “Right to be forgotten”
 - Request to be removed from search results
- Is this effective?
- Downsides?

VPN

- What is it?
 - Middleman: forwards your data to websites
 - Encrypted
- What protection does it provide?
- What does it not provide?
- Who runs these?