

Health Care Privacy in the Context of Global Privacy Policy

The Twenty-Eighth National HIPAA Summit

John Verdi

March 4, 2019

Who is FPF

The Members

130+
Companies

25+
Leading Academics

10+
Advocates

The Mission

Bridging the policymaker-industry-academic gap in privacy policy

Developing privacy protections, ethical norms, and workable business practices

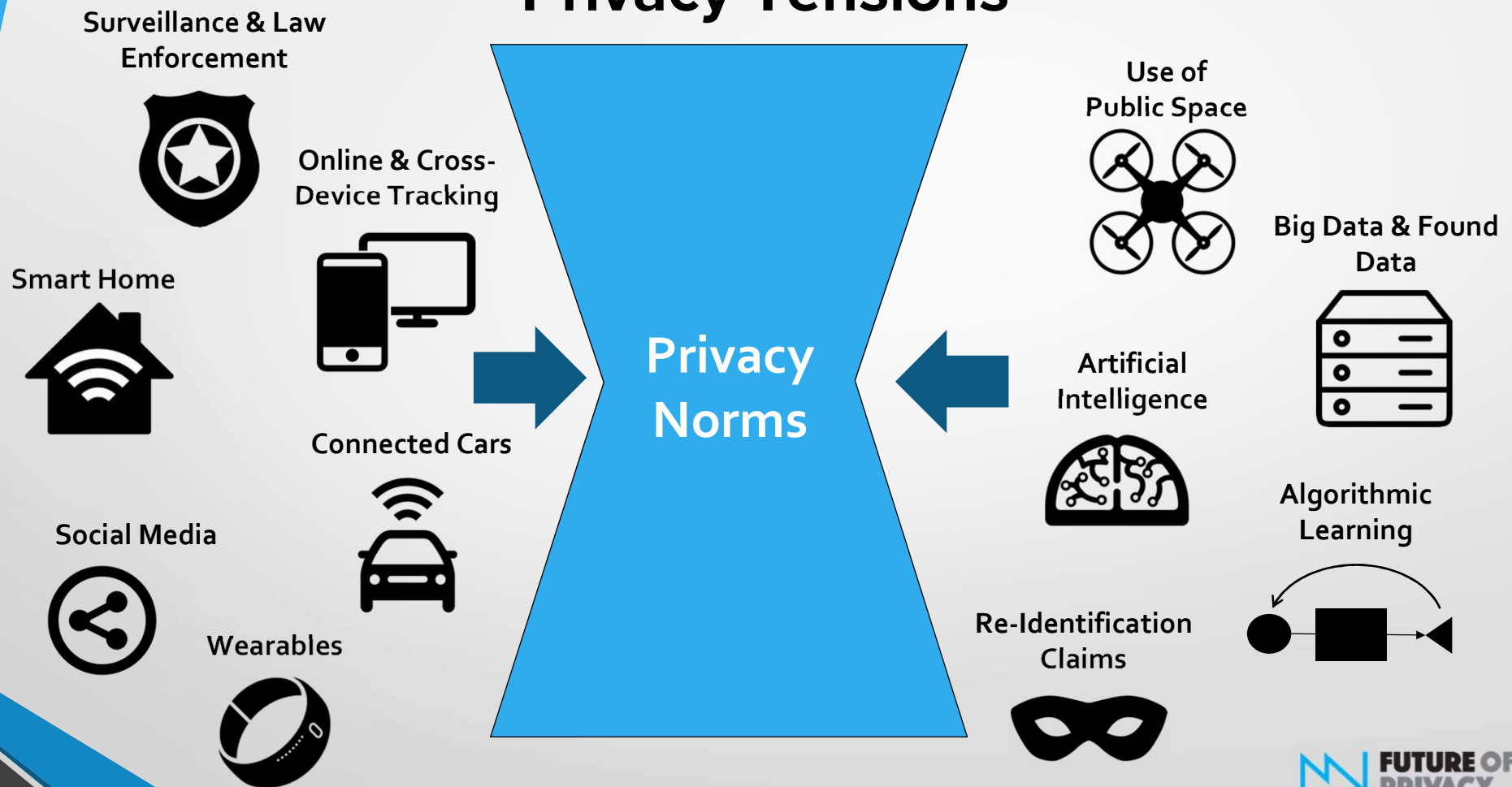
The Workstreams

Connected Cars
Student Data

Location & Ad Tech
Internet of Things
Health & Genetics

Ethics & De-identification
Smart Cities

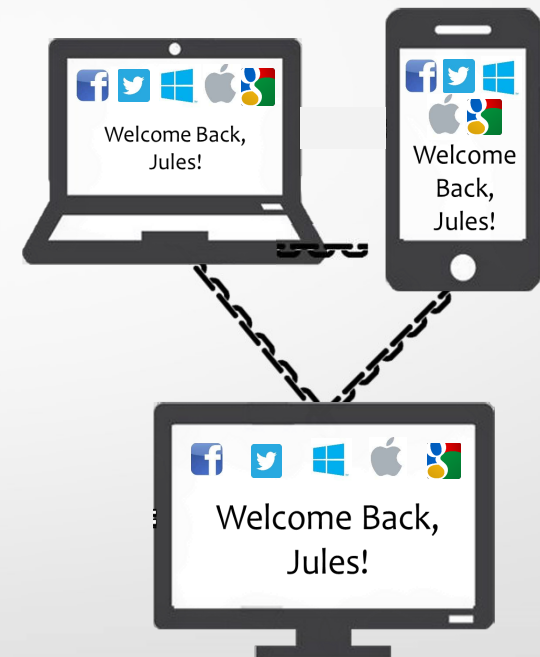
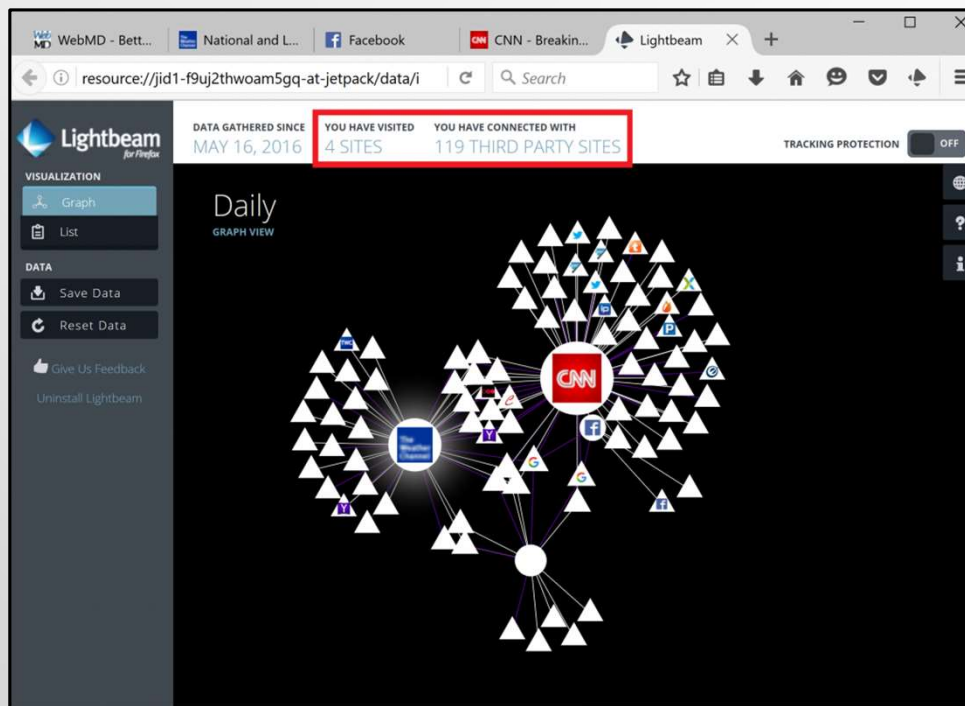
New Technologies Contribute to Privacy Tensions



Surveillance & Law Enforcement



Online and Cross-Device Tracking



Smart Home

Toys...



Appliances...



Home Assistants...



Energy Management...

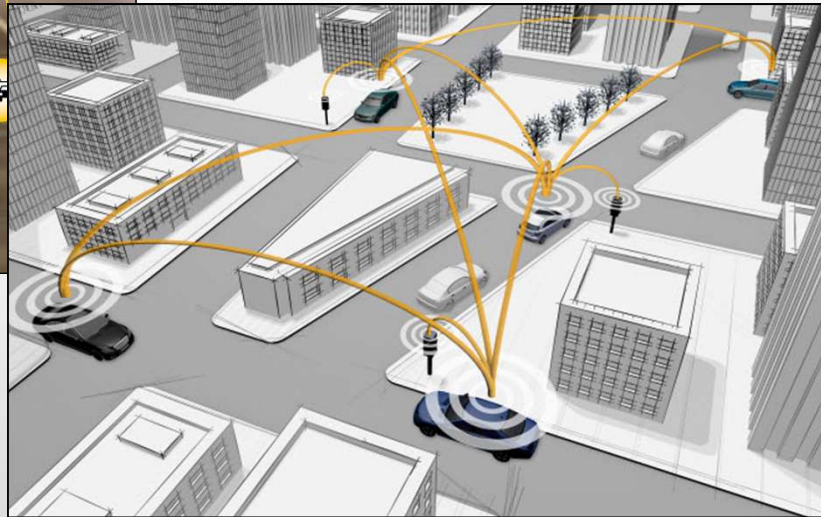


Connected Cars



**“Smart”
Car**

V2V and V2I Communication



Social Media

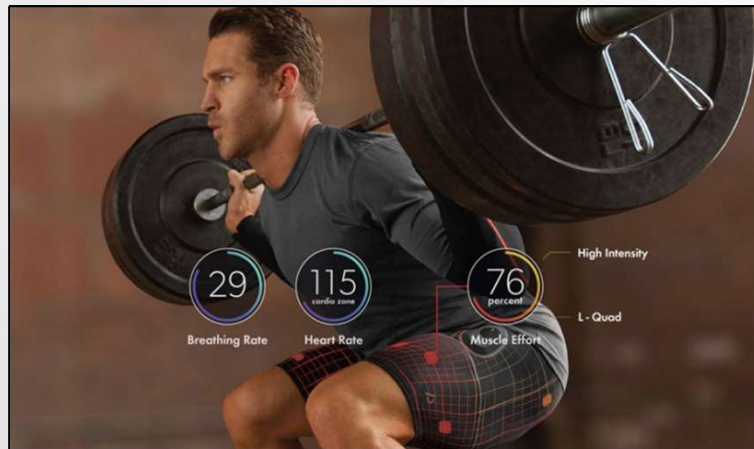


Controversial...

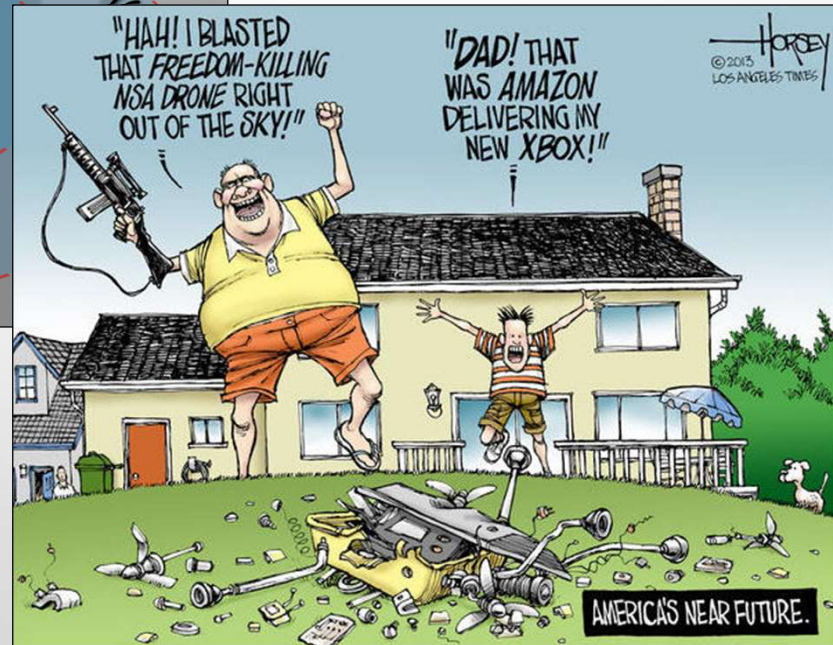
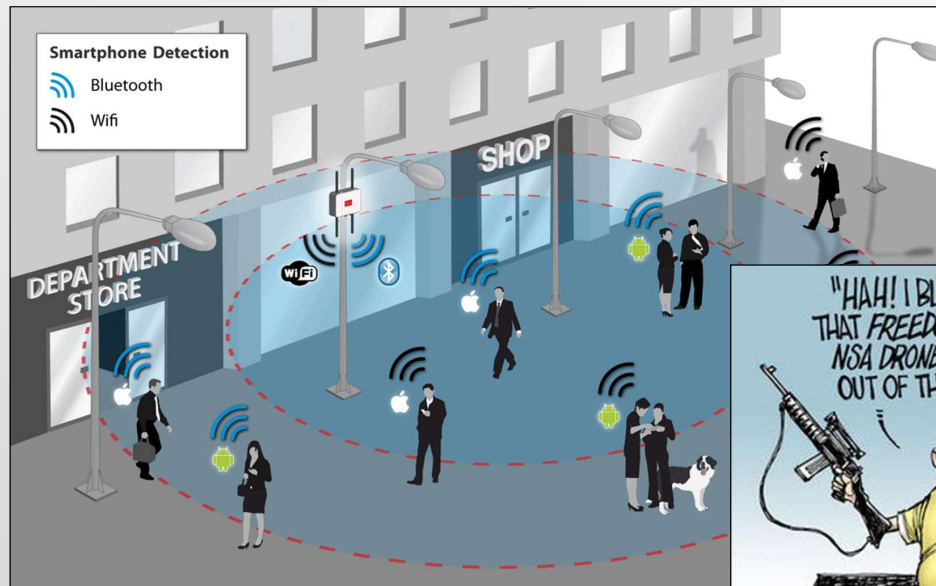


Hip!

Wearables



Use of Public Spaces



Big Data & Found Data

**Notice
Choice**

Data Quality & Integrity

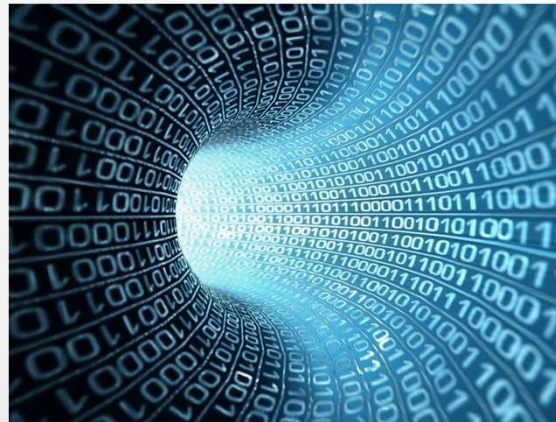
Purpose Specification

Use Limitation

Data Minimization

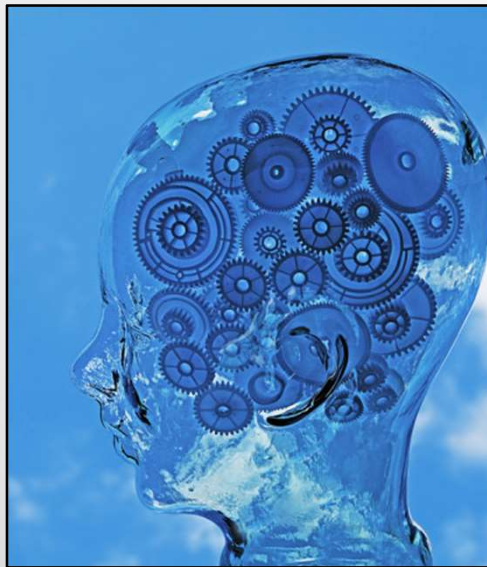
Security

Accountability



New data sets and corporate research challenge Fair Information Practice Principles (FIPPS) and ethical research principles.

Artificial Intelligence



INSIDER Sign In | Register

OPINION

Artificial intelligence needs your data, all of it

Today's concerns about giving up privacy will seem quaint in the coming years. A.I. will need everything, and we'll happily give it.





By Mike Elgan
Contributing Columnist, Computerworld | FEB 22, 2016 3:15 AM PT

THE WALL STREET JOURNAL.

This copy is for your personal, non-commercial use only. To order presentation-ready copies for distribution to your colleagues, clients or customers visit <http://www.djreprints.com>.

<http://blogs.wsj.com/cio/2017/01/27/building-public-policy-to-address-artificial-intelligences-impact/>

CIO JOURNAL.

Building Public Policy To Address Artificial Intelligence's Impact

By IRVING WLADAWSKY-BERGER

Jan 27, 2017 1:04 pm ET

Algorithmic Learning

When Discrimination Is Baked Into Algorithms

As more companies and services use data to target individuals, those analytics could inadvertently amplify bias.

LAUREN KIRCHNER | SEP 6, 2015 | BUSINESS

[Share](#) [Tweet](#) [...](#)

We want to hear from you! Help shape our future by taking the 2017 Attitudinal Survey. [Click here](#) to get started.

A recent [ProPublica analysis](#) of The Princeton Review's prices for college tutoring shows that customers in areas with a high density of Asian students often charged more. When presented with this finding, The Princeton Review called it an "incidental" result of its geographic pricing scheme. The company is now reportedly reviewing its pricing strategy.

 **International Business Times** 

Technology

Fighting crime with computers: Is predictive policing the future of law enforcement?

■ Predictive policing is causing crime rates to fall – but it comes with privacy concerns.

 **By Jason Murdock**
June 23, 2016 17:23 BST

[f](#) [t](#) [g+](#) [d](#) [in](#)

Transparency? Accountability?



Privacy Concerns

Sections 

The Washington Post
Democracy Dies in Darkness

Technology

Data of 143 million Americans exposed in hack of credit reporting agency Equifax




Danaher agrees to...
Buffett: I was close to making a 'very large' acquisition in the...
Prosecutors charge Patriots Robert Kraft in prostitution...

CYBERSECURITY

Adultery site Ashley Madison hacked, user data leaked

PUBLISHED MON, JUL 20 2015 • 2:42 PM EDT UPDATED MON, JUL 20 2015 • 2:42 PM EDT

Arjun Kharpal
[@ARJUNKHARPAL](#)

THE VERGE TECH SCIENCE CULTURE CARS REVIEWS LONGFORM MORE   

WEB CYBERSECURITY

Hack leaks hundreds of nude celebrity photos

Jennifer Lawrence among stars whose pictures were stolen

By Rich McCormick | Sep 1, 2014, 2:29am EDT

One Solution: De-identification

A VISUAL GUIDE TO PRACTICAL DATA DE-IDENTIFICATION

Produced by
**FUTURE OF
PRIVACY
FORUM**
FPP.ORG

In collaboration with
EY

What do scientists, regulators and lawyers mean when they talk about de-identification? How does anonymous data differ from pseudonymous or de-identified information? Data identifiability is not binary. Data lies on a spectrum with multiple shades of identifiability.



This is a primer on how to distinguish different categories of data.

DEGREES OF IDENTIFIABILITY

Information containing direct and indirect identifiers.

PSEUDONYMOUS DATA

Information from which direct identifiers have been eliminated or transformed, but indirect identifiers remain intact.



DE-IDENTIFIED DATA

Direct and known indirect identifiers have been removed or manipulated to break the linkage to real world identities.



ANONYMOUS DATA

Direct and indirect identifiers have been removed or manipulated together with mathematical and technical guarantees to prevent re-identification.



DIRECT IDENTIFIERS
Data that identifies a person without additional information or by linking to information in the public domain (e.g., name, SSN)



INDIRECT IDENTIFIERS
Data that identifies an individual indirectly. Helps connect pieces of information until an individual can be singled out (e.g., DOB, gender)



SAFEGUARDS and CONTROLS
Technical, organizational and legal controls preventing employees, researchers or other third parties from re-identifying individuals

SELECTED EXAMPLES

	EXPLICITLY PERSONAL	POTENTIALLY IDENTIFIABLE	NOT READILY IDENTIFIABLE	KEY CODED	PSEUDONYMOUS	PROTECTED PSEUDONYMOUS	DE-IDENTIFIED	PROTECTED DE-IDENTIFIED	ANONYMOUS	AGGREGATED ANONYMOUS
DIRECT IDENTIFIERS	INTACT	PARTIALLY MASKED	PARTIALLY MASKED	ELIMINATED or TRANSFORMED	ELIMINATED or TRANSFORMED	ELIMINATED or TRANSFORMED	ELIMINATED or TRANSFORMED	ELIMINATED or TRANSFORMED	ELIMINATED or TRANSFORMED	ELIMINATED or TRANSFORMED
INDIRECT IDENTIFIERS	INTACT	INTACT	INTACT	INTACT	INTACT	INTACT	ELIMINATED or TRANSFORMED	ELIMINATED or TRANSFORMED	ELIMINATED or TRANSFORMED	ELIMINATED or TRANSFORMED
SAFEGUARDS and CONTROLS	NOT RELEVANT due to nature of data	LIMITED or NONE IN PLACE	CONTROLS IN PLACE	CONTROLS IN PLACE	LIMITED or NONE IN PLACE	CONTROLS IN PLACE	LIMITED or NONE IN PLACE	CONTROLS IN PLACE	NOT RELEVANT due to nature of data	NOT RELEVANT due to high degree of data aggregation
SELECTED EXAMPLES	Name, address, phone number, SSN, government-issued ID (e.g., Jane Smith, 123 Main Street, 555-555-5555)	Unique device ID, license plate, medical record number, cookie, IP address (e.g., MAC address 68:A6:6D:35:65:03)	Same as Potentially identifiable except data are also protected by safeguards and controls (e.g., hashed MAC addresses & legal representations)	Clinical or research datasets where only cursor retains key (e.g., Jane Smith, diabetes, Hgb 15.1 g/dL = Crk123)	Unique, artificial pseudonyms replace direct identifiers (e.g., HIPAA Limited Datasets, John Doe = SLT LMS192) (unique sequence not used anywhere else)	Same as Pseudonymous, except data are also protected by safeguards and controls	Data are suppressed, generalized, perturbed, swapped, etc. (e.g., GPA: 3.2 = 3.0-3.5, gender: female = gender: male)	Same as De-identified, except data are also protected by safeguards and controls	For example, noise is calibrated to a data set to hide whether an individual is present or not (differential privacy)	Very highly aggregated data (e.g., statistical data, census data, or population data that 52.6% of Washington, DC residents are women)

Re-Identification



DATA PRIVACY LAB

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

LaTanya Sweeney & Gov.
William Weld
Netflix
AOL Searcher No. 4417749
Paul Ohm's "Database of Ruin"

The New York Times

With a Few Bits of Data, Researchers Identify 'Anonymous' People

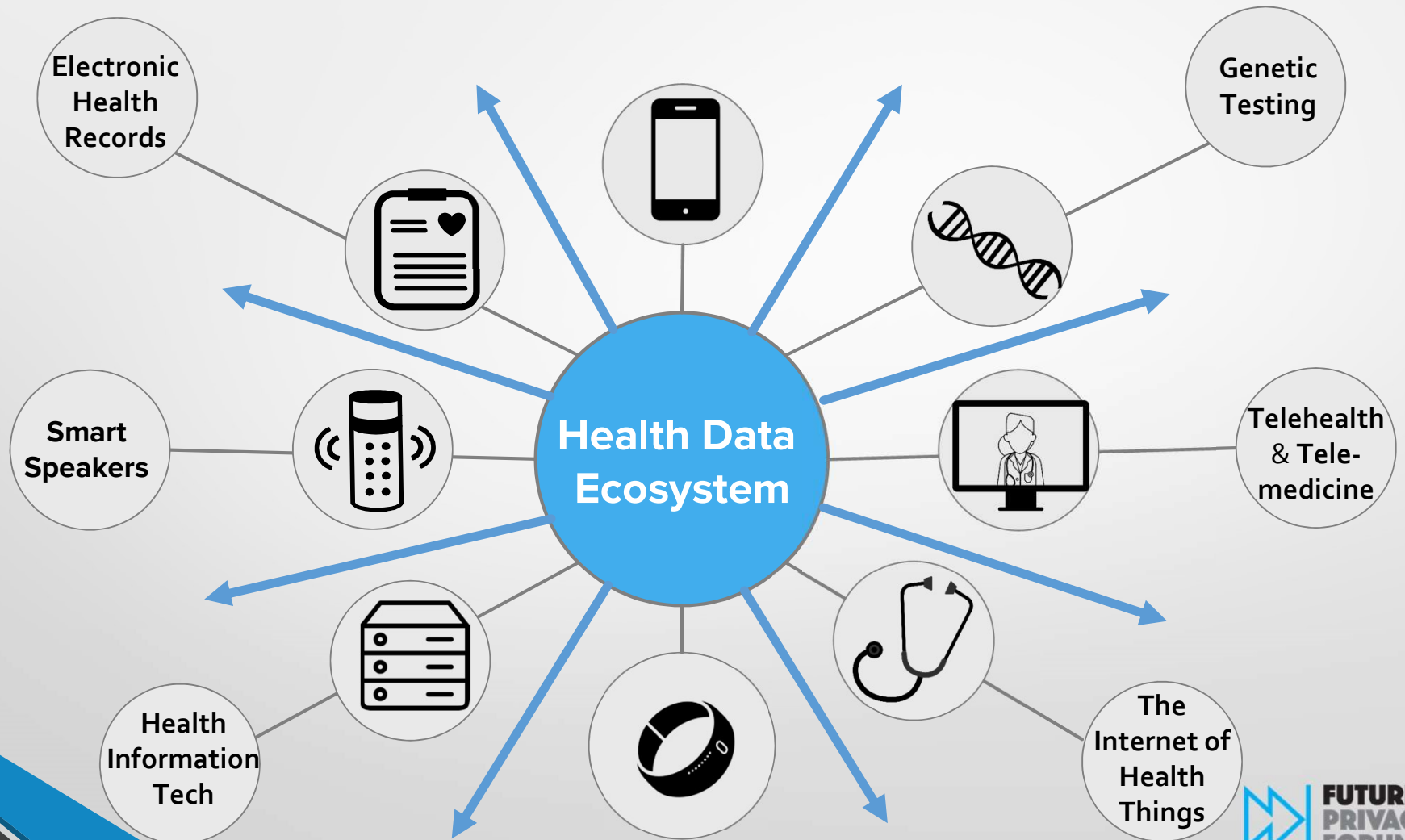
By **Natasha Singer** January 29, 2015 2:01 pm

Even when real names and other personal information are stripped from big data sets, it is often possible to use just a few pieces of the information to identify a specific person, according to a study to be published Friday in the journal *Science*.

In the study, titled "Unique in the Shopping Mall: On the Reidentifiability of Credit Card Metadata," a group of data scientists analyzed credit card transactions made by 1.1 million people in 10,000 stores over a three-month period. The data set

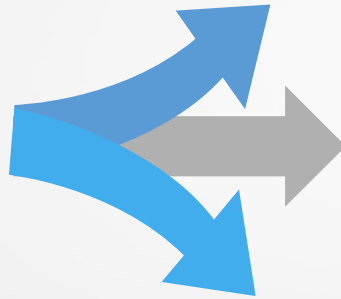
Demonstrations of re-identification cast doubt on anonymization.

Health Data Ecosystem is Expanding



Effects of Expansion

Technological
Development
Expanding the
Health Data
Ecosystem



New Data Sources & Types



Innovative Data Uses

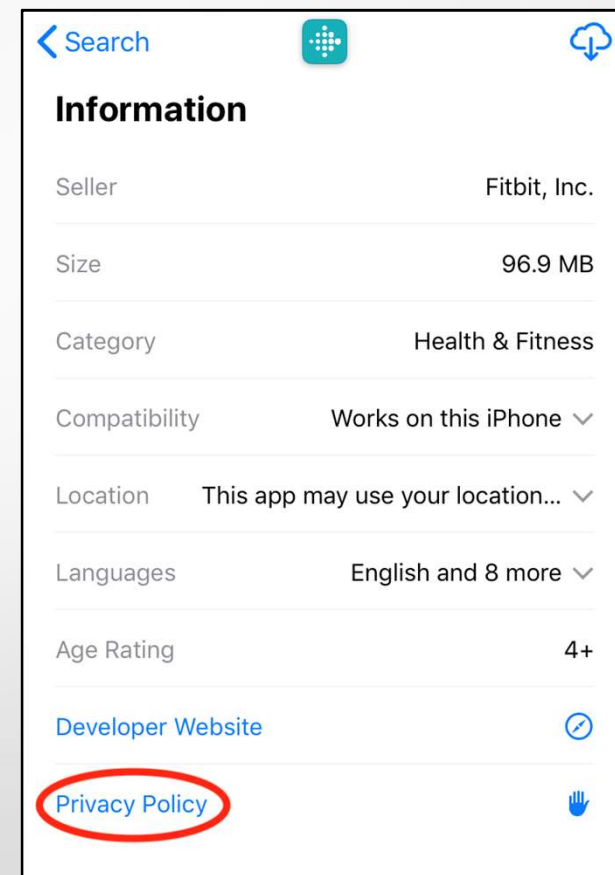


New Stakeholders

Also bringing new privacy questions about the responsible collection, use, and sharing of data and a need for standardized language...

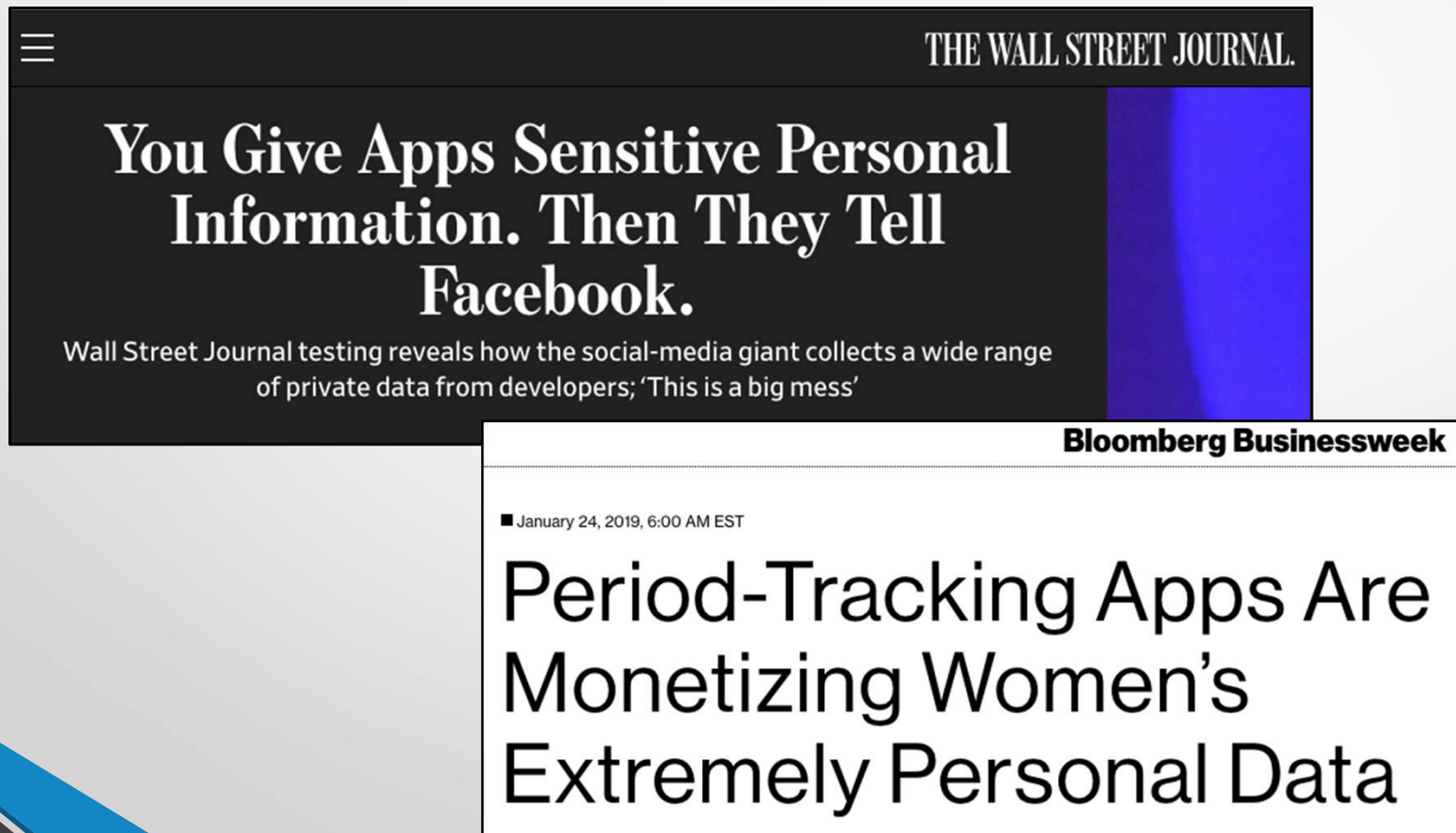
FPF Mobile App Survey

- Question: Do the most popular health apps on the most used platforms have a linked privacy policy?
- Commissioned by California's Attorney General
- Looked at: 100 Health and Fitness Apps, 91 Sleep Aid Apps, 41 Fertility-Tracking App
- Outcomes:
 - For Health and Fitness apps, free apps were more likely to have linked privacy policies than paid apps
 - Fertility-Tracking apps had the highest rates of having a linked privacy policy than any other group



Apps were not tested for compliance with stated privacy policies

App Privacy Still a Challenge



The image shows a screenshot of a news article. The top part is from The Wall Street Journal, with a black background and white text. The headline reads: "You Give Apps Sensitive Personal Information. Then They Tell Facebook." Below the headline, a sub-headline states: "Wall Street Journal testing reveals how the social-media giant collects a wide range of private data from developers; 'This is a big mess'". To the right of the text is a blue abstract graphic. The bottom part of the screenshot is from Bloomberg Businessweek, with a white background and black text. It includes a timestamp: "January 24, 2019, 6:00 AM EST" and a headline: "Period-Tracking Apps Are Monetizing Women's Extremely Personal Data".

THE WALL STREET JOURNAL.

You Give Apps Sensitive Personal Information. Then They Tell Facebook.

Wall Street Journal testing reveals how the social-media giant collects a wide range of private data from developers; 'This is a big mess'

Bloomberg Businessweek

■ January 24, 2019, 6:00 AM EST

Period-Tracking Apps Are Monetizing Women's Extremely Personal Data

Best Practices for Consumer Wearables and Wellness Apps and Devices

- Released August 17, 2016
- Provides a detailed set of guidelines the responsible companies can follow to protect consumer-generated health and wellness data
- Supported by the Robert Wood Johnson Foundation

Best Practices for
Consumer Wearables &
Wellness Apps & Devices

August 17, 2016



The Best Practices for Consumer Wearables & Wellness Apps & Devices was produced with support from the Robert Wood Johnson Foundation.

Highlights

- Requires opt-in consent for sharing with third parties
- Bans sharing with data brokers, information resellers, and ad networks
- Requires opt-out options for tailored first-party advertisements
- Provides access, correction, and deletion rights
- Supports interoperability with global privacy frameworks and leading app platform standards



Privacy Best Practices for Consumer Genetic Testing Services

- Released July 31, 2018
- Provides a policy framework for the collection, retention, sharing, and use of genetic data generated by consumer genetic and personal genomic testing companies
- Supported by Ancestry, 23andMe, Helix, MyHeritage, African Ancestry, and Living DNA

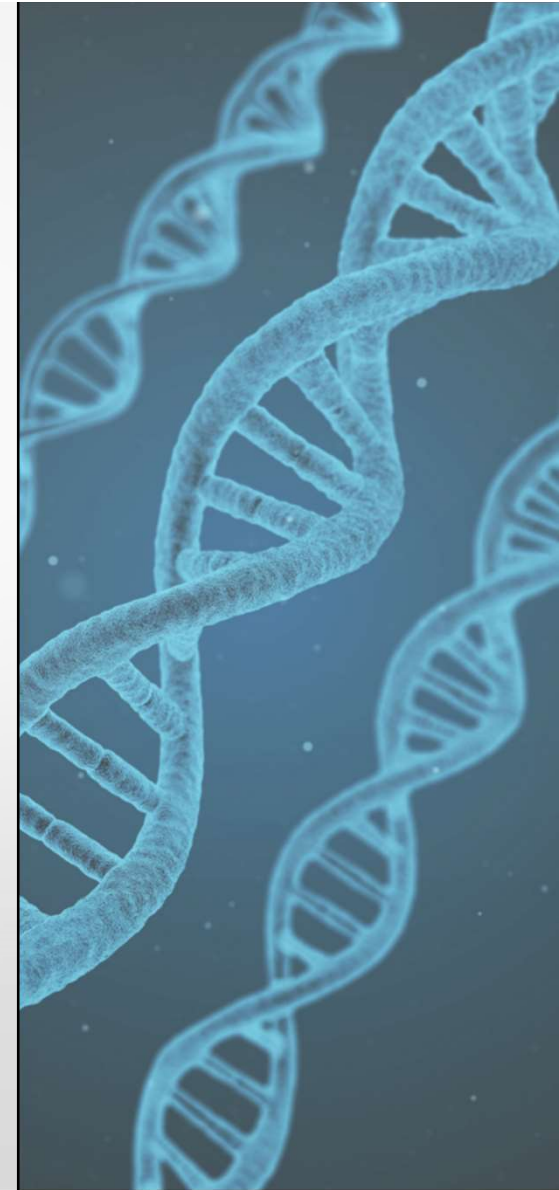
Privacy Best Practices for
Consumer Genetic Testing
Services

July 31, 2018

 **FUTURE OF
PRIVACY
FORUM**
1400 Eye Street, NW, Suite 450
Washington, DC 20005
fpf.org

Highlights

- Requires detailed transparency about collection, use, sharing, and retention of genetic data
- Provides access and deletion rights
- Requires valid legal process for the disclosure of genetic data to law enforcement and transparency reporting on at least an annual basis
- Restricts marketing based on genetic data
- Requires strong data security protections and privacy by design



Emerging Health Privacy Policy Questions

General Data
Protection
Regulation
(GDPR)



California
Consumer
Privacy Act
(CCPA) + other
state laws



US Federal
Privacy
Legislation on
the Horizon

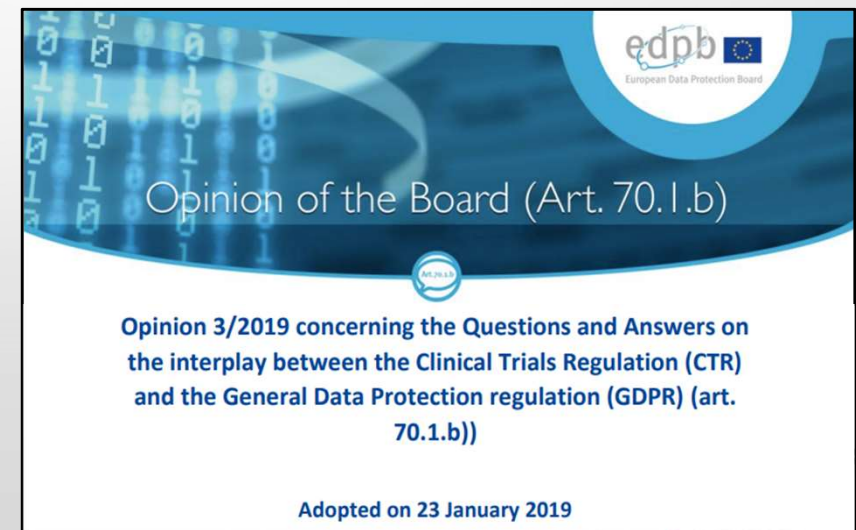
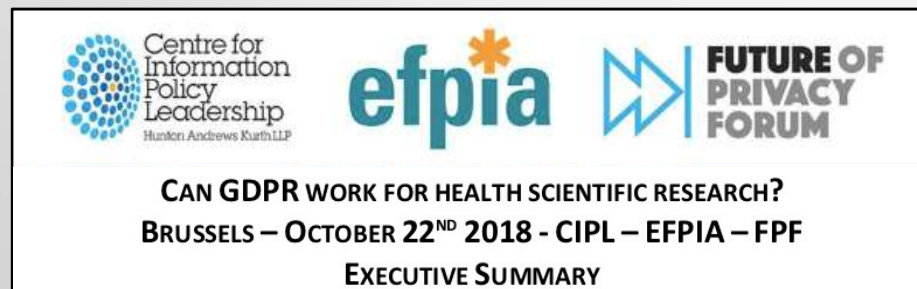
Pushing us to re-think
how we have
traditionally protected
the privacy of health
information in the
United States



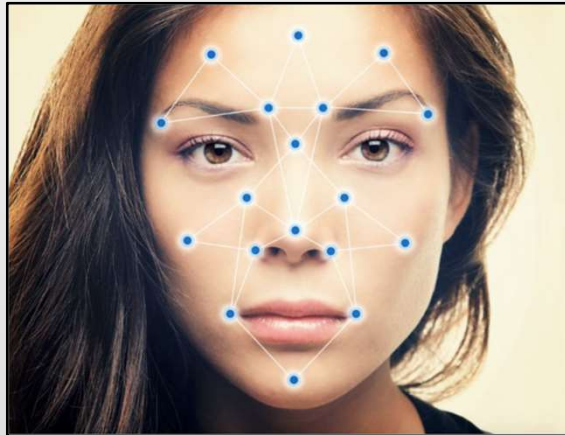
Will HIPAA be exempted from, amended by, or repealed by future privacy legislation?

Emerging Challenge: Health Research

- Uncertainty around the legal basis for processing data for clinical research under the GDPR
- European Data Protection Board (EDPB) issued a Q&A in January clarifying acceptable bases and the interplay between GDPR and the Clinical Trials Regulation (CTR)
- Continuing issue: harmonization across member states



Emerging Challenge: Facial Characterization



LETTERS | FOCUS
<https://doi.org/10.1038/s41591-018-0279-0>

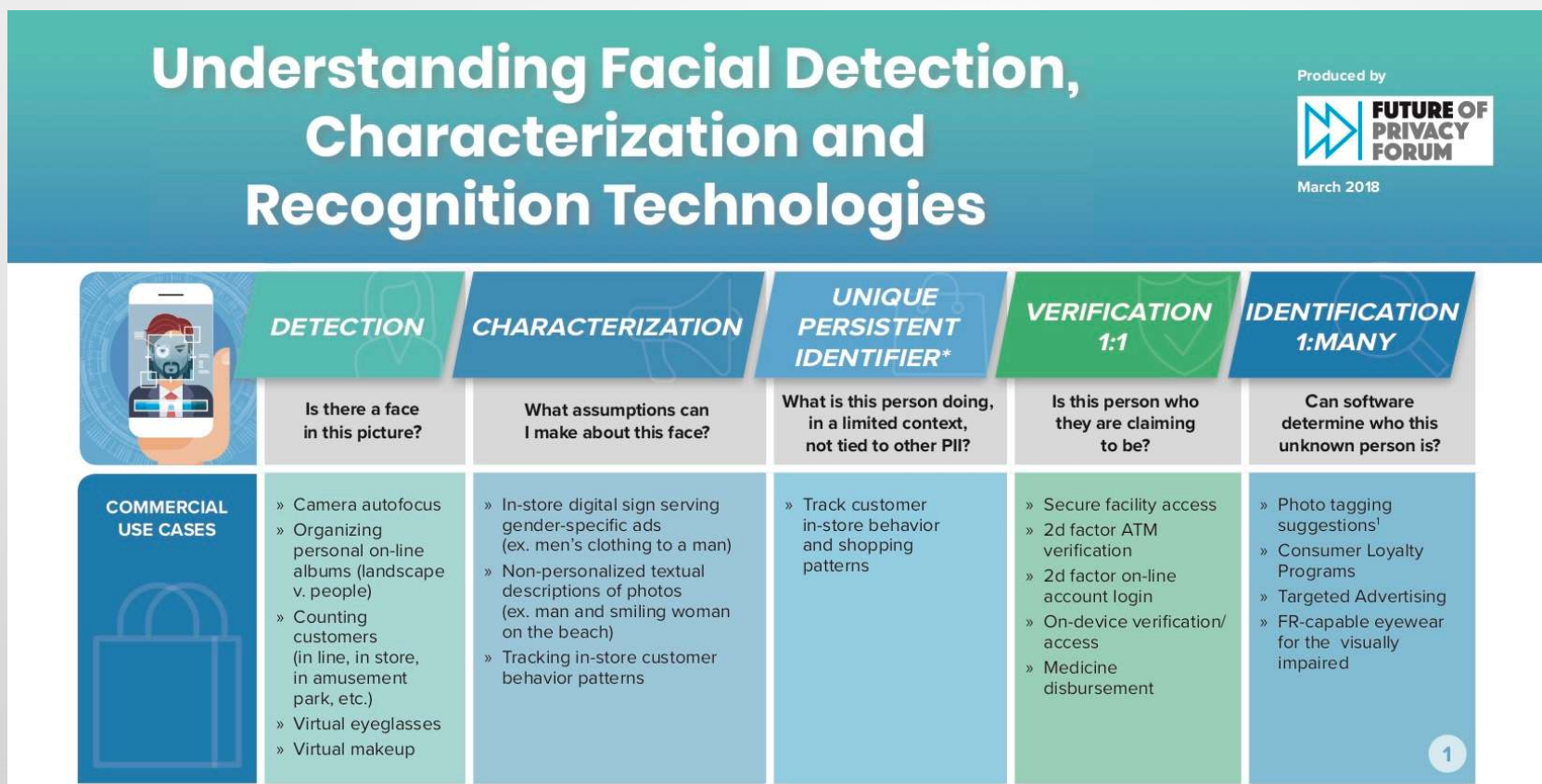
nature
medicine

Identifying facial phenotypes of genetic disorders using deep learning

Yaron Gurovich^{1*}, Yair Hanani¹, Omri Bar¹, Guy Nadav¹, Nicole Fleischer¹, Dekel Gelbman¹, Lina Basel-Salmon^{2,3}, Peter M. Krawitz⁴, Susanne B. Kamphausen⁵, Martin Zenker⁵, Lynne M. Bird^{6,7} and Karen W. Gripp⁸

- Facial characterization tools are increasingly able to reveal health information
- When does facial characterization analysis constitute a privacy invasion?
- Accuracy of analysis and the importance of context will need to be addressed

FPF Infographic on Facial Detection, Characterization, and Recognition Technologies

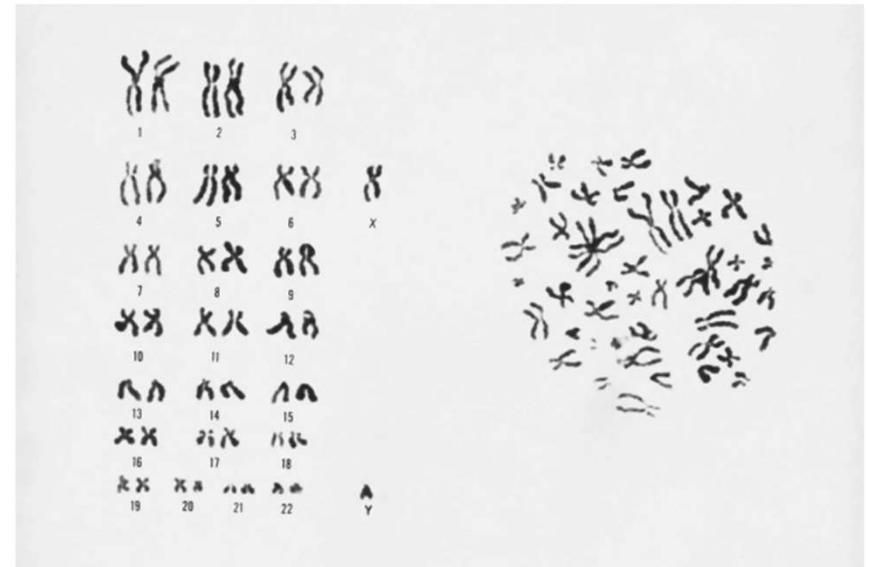


Emerging Challenge: Genetic Data

- What does it mean when your personal information implicates the privacy of others?
- De-identification challenges also exist for this sensitive information
- Additional protections (contractual controls, access controls, and security protocols) may need to be employed

The New York Times

Most White Americans' DNA Can Be Identified Through Genealogy Databases



Only two percent of the population needs to have done a DNA test to identify nearly everyone else, researchers found. Leonard Lessin/Science Source

Thank You

John Verdi

Vice President of Policy,
Future of Privacy Forum

jverdi@fpf.org

@JohnVerdi

Follow us!



- www.fpf.org
- facebook.com/futureofprivacy
- [@futureofprivacy](https://twitter.com/futureofprivacy)