Navigating Privacy in a Data Driven World

HoTSoS 2017

FUTURE OF PRIVACY FORUM

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

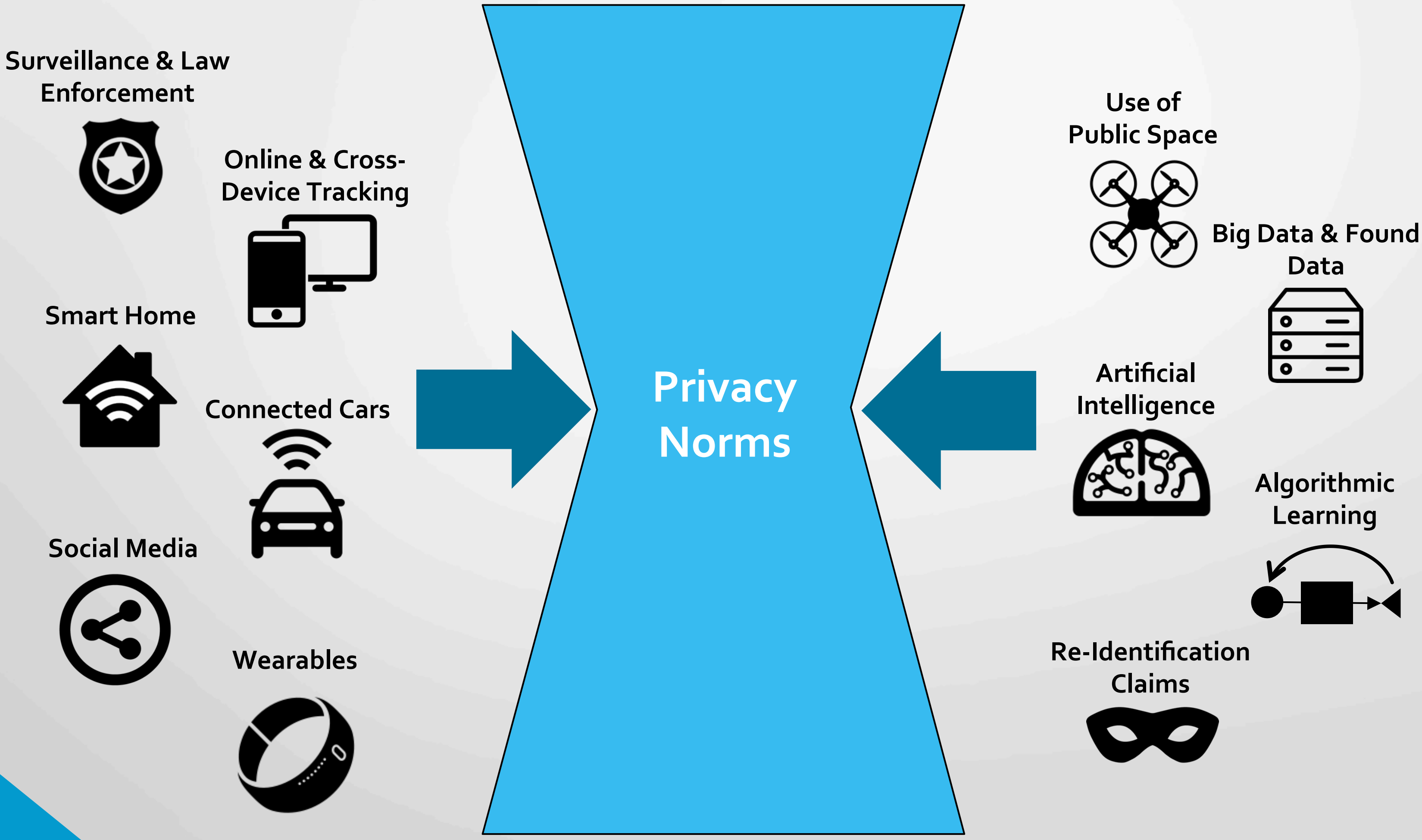
Connected Cars
Student Data

Location & Ad Tech
Internet of Things

Ethics & De-identification
Smart Cities

FUTURE OF PRIVACY FORUM

New technologies contribute to privacy tensions.



Surveillance & Law Enforcement

Online & Cross-Device Tracking

Smart Home

Connected Cars

Social Media

Wearables

Privacy Norms

Use of Public Space

Big Data & Found Data

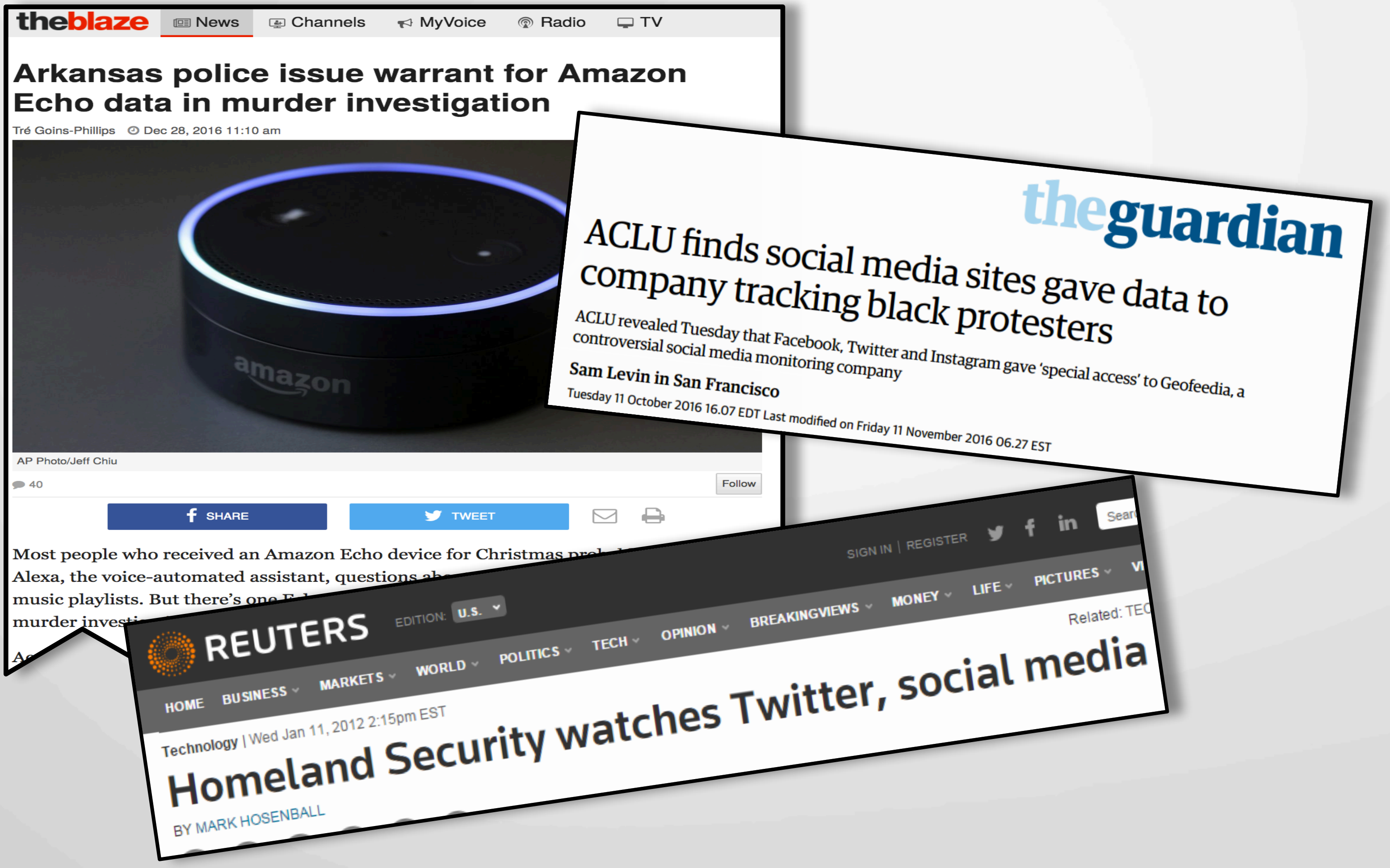
Artificial Intelligence

Algorithmic Learning

Re-Identification Claims

FUTURE OF PRIVACY FORUM

Surveillance & Law Enforcement



theblaze News Channels MyVoice Radio TV

Arkansas police issue warrant for Amazon Echo data in murder investigation

Tré Goins-Phillips Dec 28, 2016 11:10 am

AP Photo/Jeff Chiu

40 Follow

SHARE TWEET

Most people who received an Amazon Echo device for Christmas prob
Alexa, the voice-automated assistant, questions ab
music playlists. But there's one E
murder investi
A

theguardian

ACLU finds social media sites gave data to company tracking black protesters

ACLU revealed Tuesday that Facebook, Twitter and Instagram gave 'special access' to Geofeedia, a controversial social media monitoring company

Sam Levin in San Francisco

Tuesday 11 October 2016 16.07 EDT Last modified on Friday 11 November 2016 06.27 EST

SIGN IN | REGISTER Sear

REUTERS EDITION: U.S.

HOME BUSINESS MARKETS WORLD POLITICS TECH OPINION BREAKINGVIEWS MONEY LIFE PICTURES VI

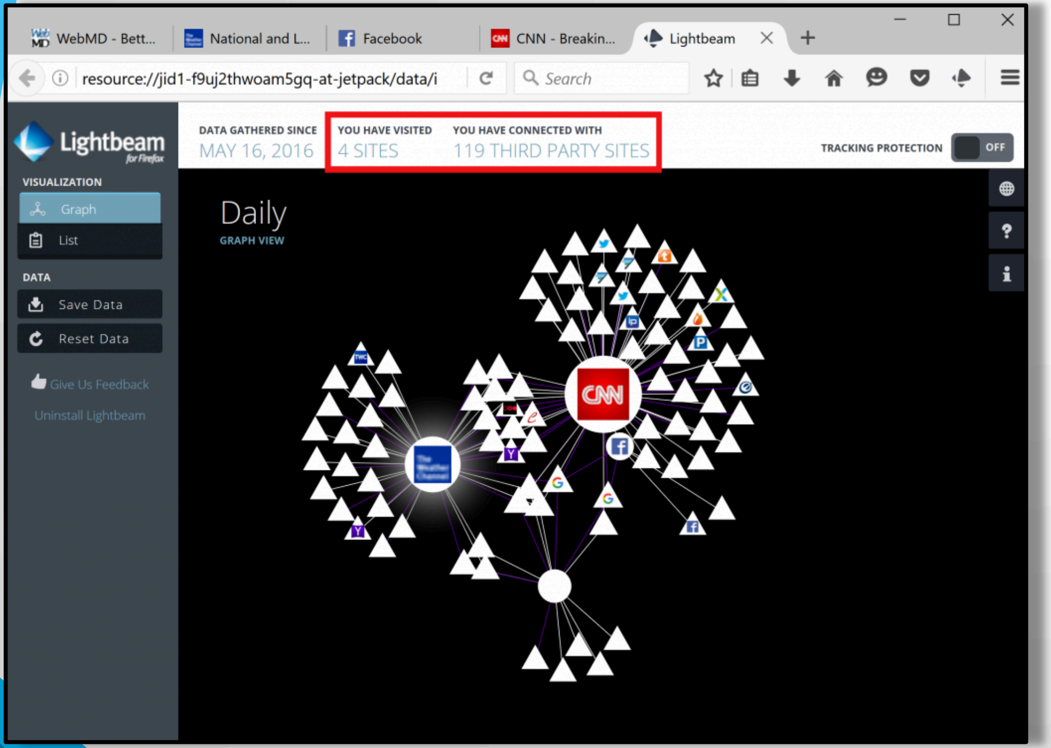
Related: TEC

Technology | Wed Jan 11, 2012 2:15pm EST

Homeland Security watches Twitter, social media

BY MARK HOSENBALL

Online and Cross-Device Tracking



resource://jid1-f9uj2thwoam5gq-at-jetpack/data/i

DATA GATHERED SINCE MAY 16, 2016

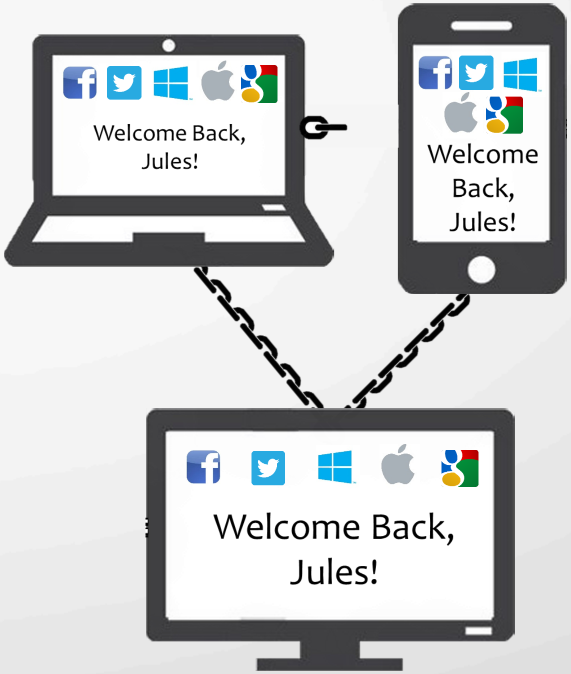
YOU HAVE VISITED 4 SITES

YOU HAVE CONNECTED WITH 119 THIRD PARTY SITES

TRACKING PROTECTION OFF

Daily

GRAPH VIEW



Smart Home

Toys…

Appliances…

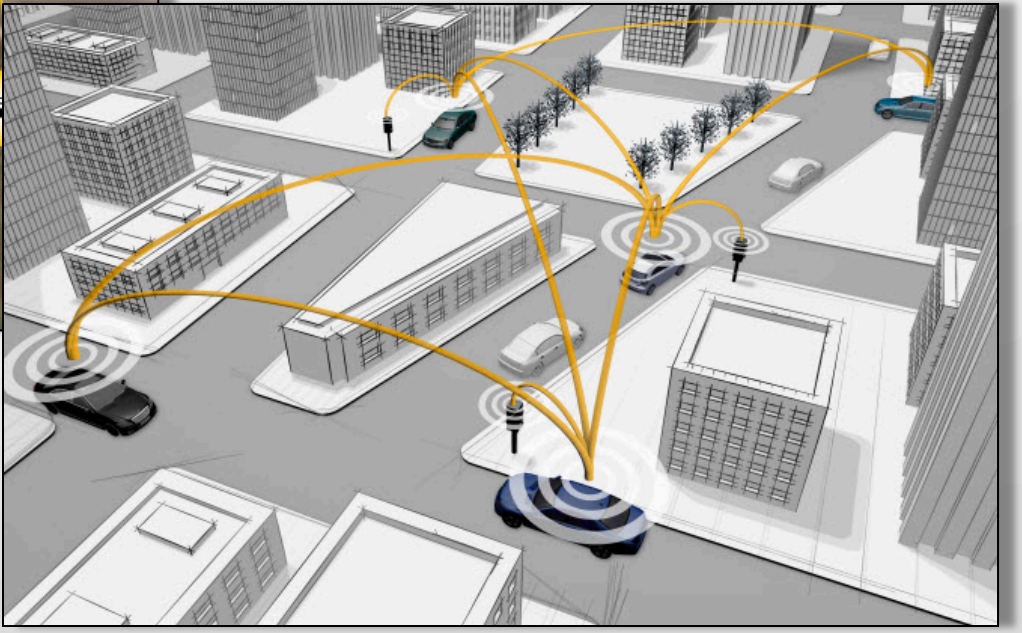
Home Assistants…

Energy Management…

Connected Cars

"Smart" Car

V2V and V2I Communication



Social Media

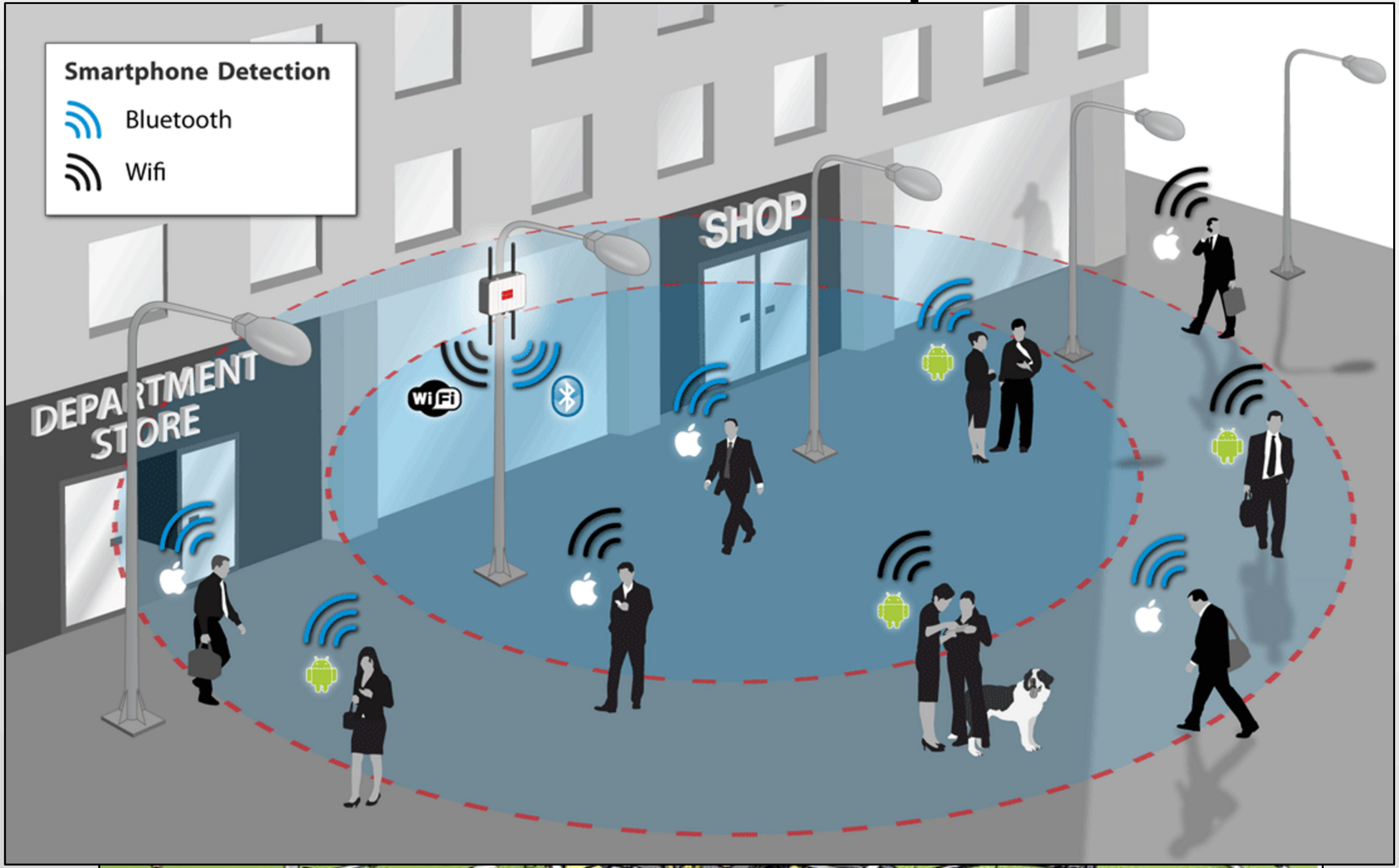
Controversial…

Hip!

Wearables

29
Breathing Rate
115
cardio zone
Heart Rate
76
percent
Muscle Effort
High Intensity
L - Quad
twelve
thirty
five

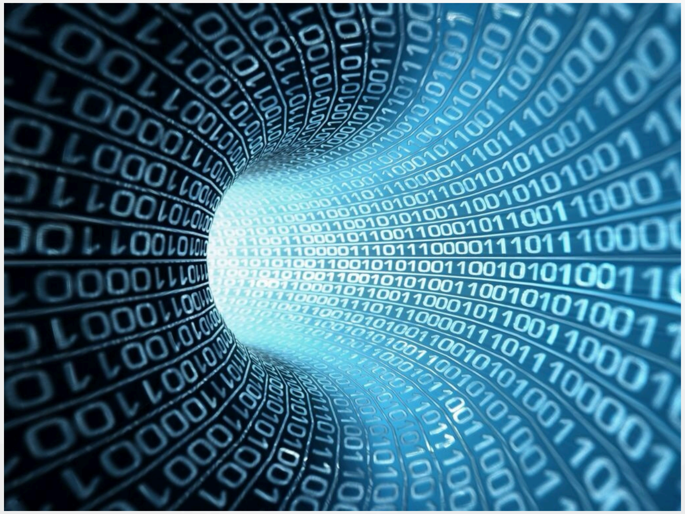
Use of Public Spaces



FUTURE OF PRIVACY FORUM

Big Data & Found Data

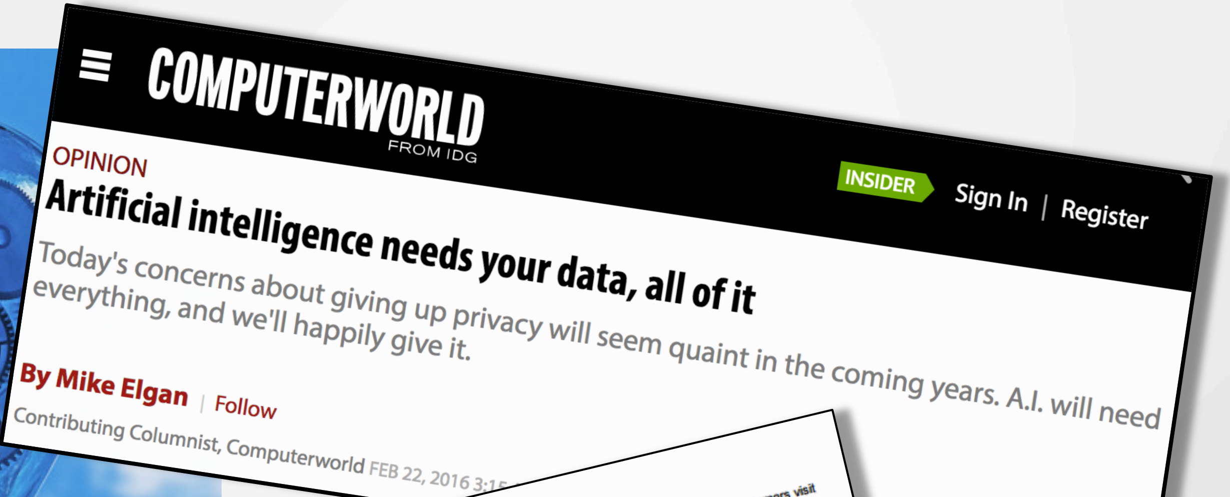
Notice
Choice
Data Quality & Integrity
Purpose Specification
Use Limitation
Data Minimization
Security
Accountability



COMMON
RULE

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

Artificial Intelligence



COMPUTERWORLD FROM IDG

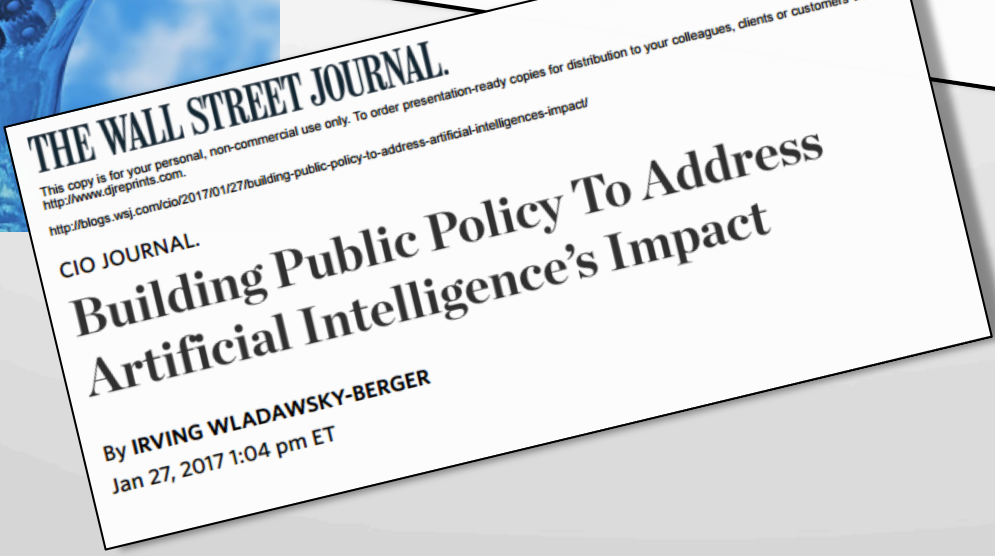
OPINION

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 | Follow

Contributing Columnist, Computerworld FEB 22, 2016 3:1



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/

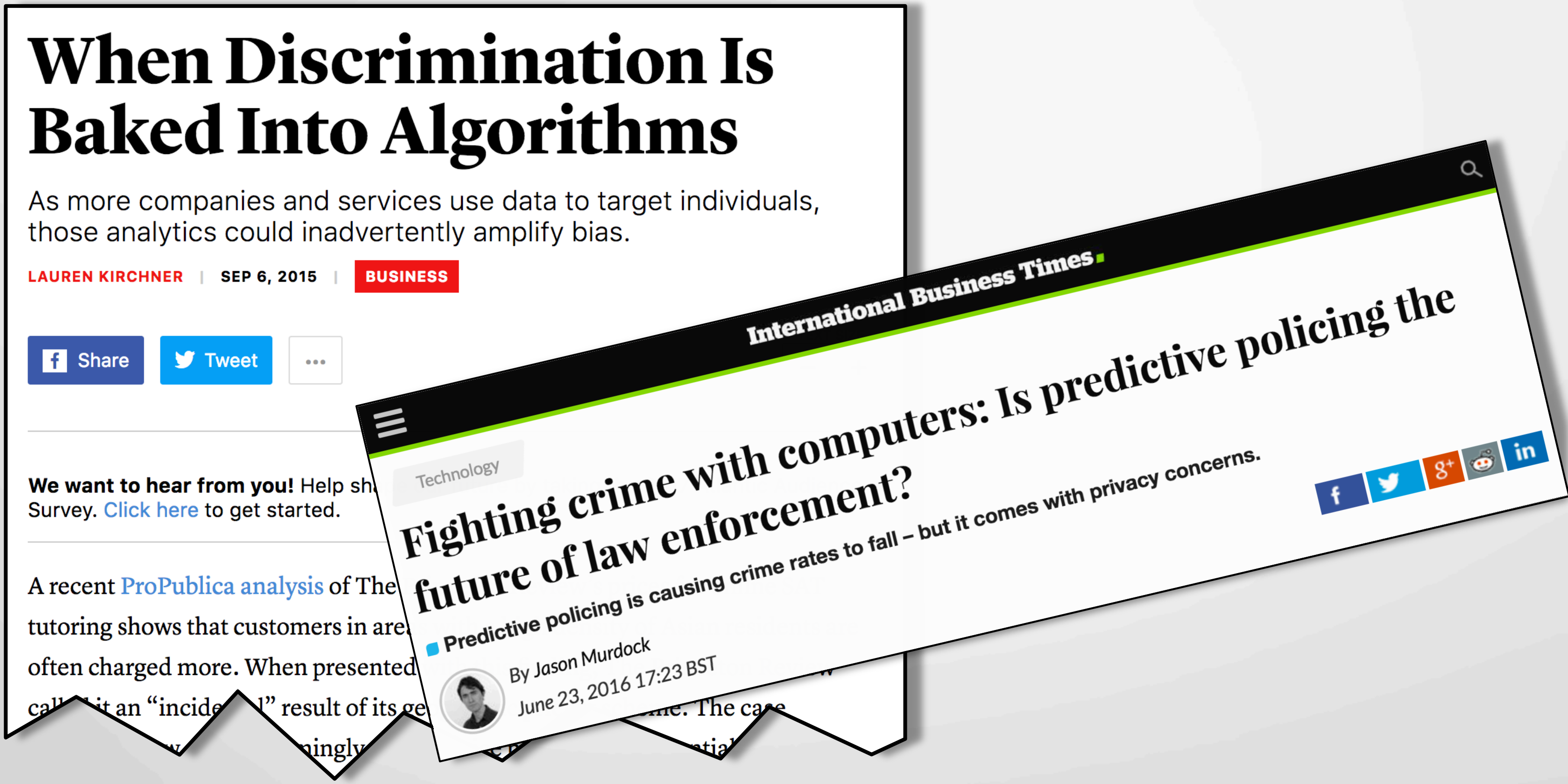
CIO JOURNAL.

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

We want to hear from you! Help sh
Survey. Click here to get started.

A recent ProPublica analysis of The
tutoring shows that customers in are
often charged more. When presented
call it an "incide l" result of its ge me. The case
ingly ntial

International Business Times

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

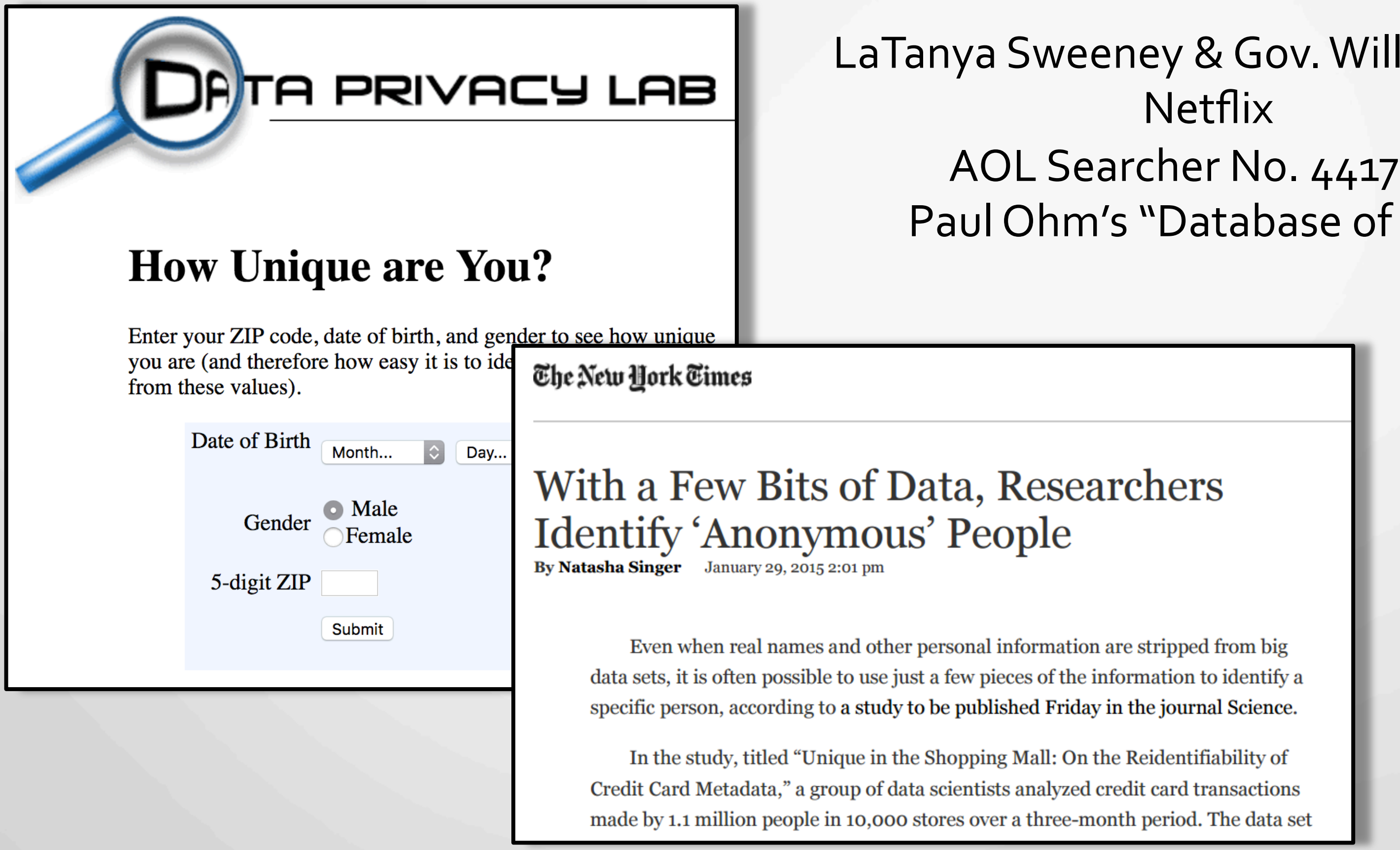
Predictive policing is causing crime rates to fall – but it comes with privacy concerns.

By Jason Murdock
June 23, 2016 17:23 BST

Transparency? Accountability?

FUTURE OF PRIVACY FORUM

Re-Identification



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 ide
from these values).

Date of Birth	Month... Day...
Gender	● Male ○ Female
5-digit ZIP	
	Submit

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

By Natasha Singer January 29, 2015 2:01 pm

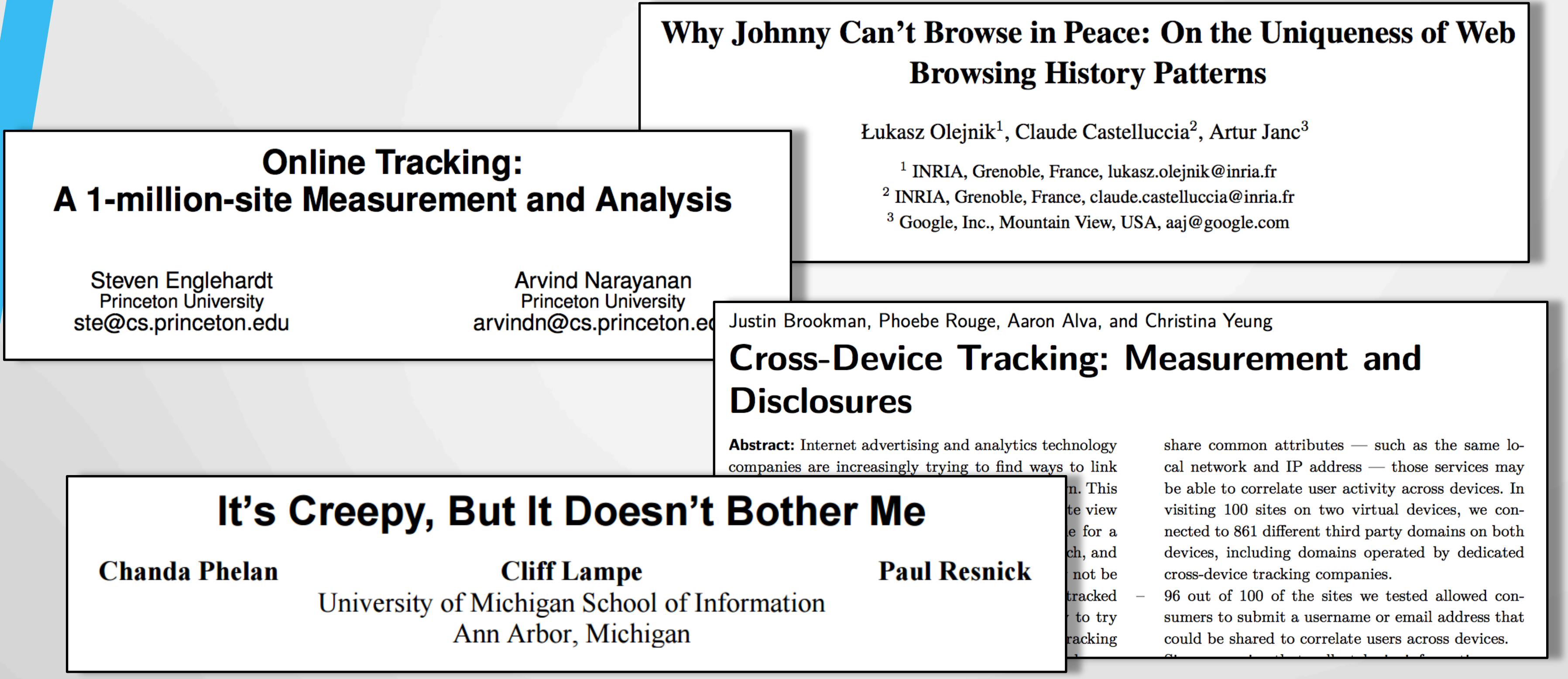
Even when real names and other personal information are stripped from big data sets, it is often possible to use just a few pieces of the information to identify a specific person, according to a study to be published Friday in the journal Science.

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

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

Demonstrations of re-identification cast doubt on anonymization.

Current State of Privacy Research



Why Johnny Can’t Browse in Peace: On the Uniqueness of Web Browsing History Patterns

Łukasz Olejnik[1], Claude Castelluccia[2], Artur Janc[3]

[1] INRIA, Grenoble, France, lukasz.olejnik@inria.fr
[2] INRIA, Grenoble, France, claude.castelluccia@inria.fr
[3] Google, Inc., Mountain View, USA, aaj@google.com

Online Tracking: A 1-million-site Measurement and Analysis

Steven Englehardt
Princeton University
ste@cs.princeton.edu

Arvind Narayanan
Princeton University
arvindn@cs.princeton.ed

Justin Brookman, Phoebe Rouge, Aaron Alva, and Christina Yeung

Cross-Device Tracking: Measurement and Disclosures

Abstract: Internet advertising and analytics technology companies are increasingly trying to find ways to link

n. This
te view
e for a
ch, and
not be
tracked
to try
racking

share common attributes — such as the same local network and IP address — those services may be able to correlate user activity across devices. In visiting 100 sites on two virtual devices, we connected to 861 different third party domains on both devices, including domains operated by dedicated cross-device tracking companies.

– 96 out of 100 of the sites we tested allowed consumers to submit a username or email address that could be shared to correlate users across devices.

It’s Creepy, But It Doesn’t Bother Me

Chanda Phelan **Cliff Lampe** **Paul Resnick**
University of Michigan School of Information
Ann Arbor, Michigan

Much privacy research is aimed at identifying and blocking privacy threats.

The Way Forward

National Privacy Research Priorities

3.1 Foster multidisciplinary approach to privacy research

3.2 Understand and measure privacy impacts and desires

3.3 Develop system design methods to incorporate privacy desires, requirements & controls

3.4 Increase transparency of data collection, sharing, use, and retention

3.5 Assure that information flows and use are consistent with privacy rules

3.6 Develop approaches for remediation and recovery

3.7 Reduce privacy risks of analytical algorithms

Privacy Research and Data Responsibility Research Coordination Network (RCN)

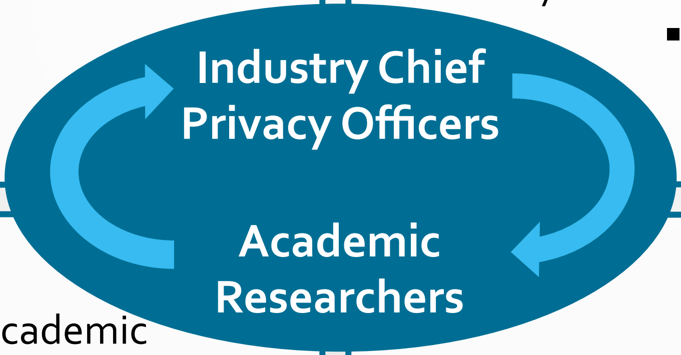
Challenge

How can industry and academia work together to advance the National Privacy Research Strategy?

Scientific Impact

Encourage multi-disciplinary research along a continuum of privacy challenges, e.g.

- Privacy risks of analytical algorithms
- Transparency of data collection and use
- De-identification



Industry Chief Privacy Officers

Academic Researchers

Ongoing Efforts

RCN fosters industry-academic collaboration by incentivizing and distributing privacy research

- Privacy Papers for Policymakers
- FPF-Capital Area Academic Network
- Privacy Scholarship Reporter
- Cross-sector workshops & symposia

New Efforts

- Privacy researcher clearinghouse
- Document industry data flows
- Methods to evaluate privacy controls

www.fpf.org/rcn

FUTURE OF PRIVACY FORUM

Research Issues

Ethical Review

IRB review is not well-suited to data-driven research.

- Found Data
- Corporate Research
- Informed Consent
- Common Rule Limitations

Access to Data

Analysis of large data sets—from the private and public sector—promises societal benefits and smarter policy-making. But researchers face significant hurdles:

- Privacy & Security Concerns
- Ethical Review Concerns
- Transaction Costs
- Re-Identification Risk
- IP & Trade Secrets

De-Identification

Powerful computing and ubiquitous data sets have cast doubt on traditional methods of de-identification

Solutions: Ethical Review

New structures for ethical review beyond the IRB can provide the processes required to authorize non-contextual data uses.

Ethical Review

Scope

- Data research & experimentation
- Non-contextual data use
- Disparate impact & algorithmic data use

Guiding Principles

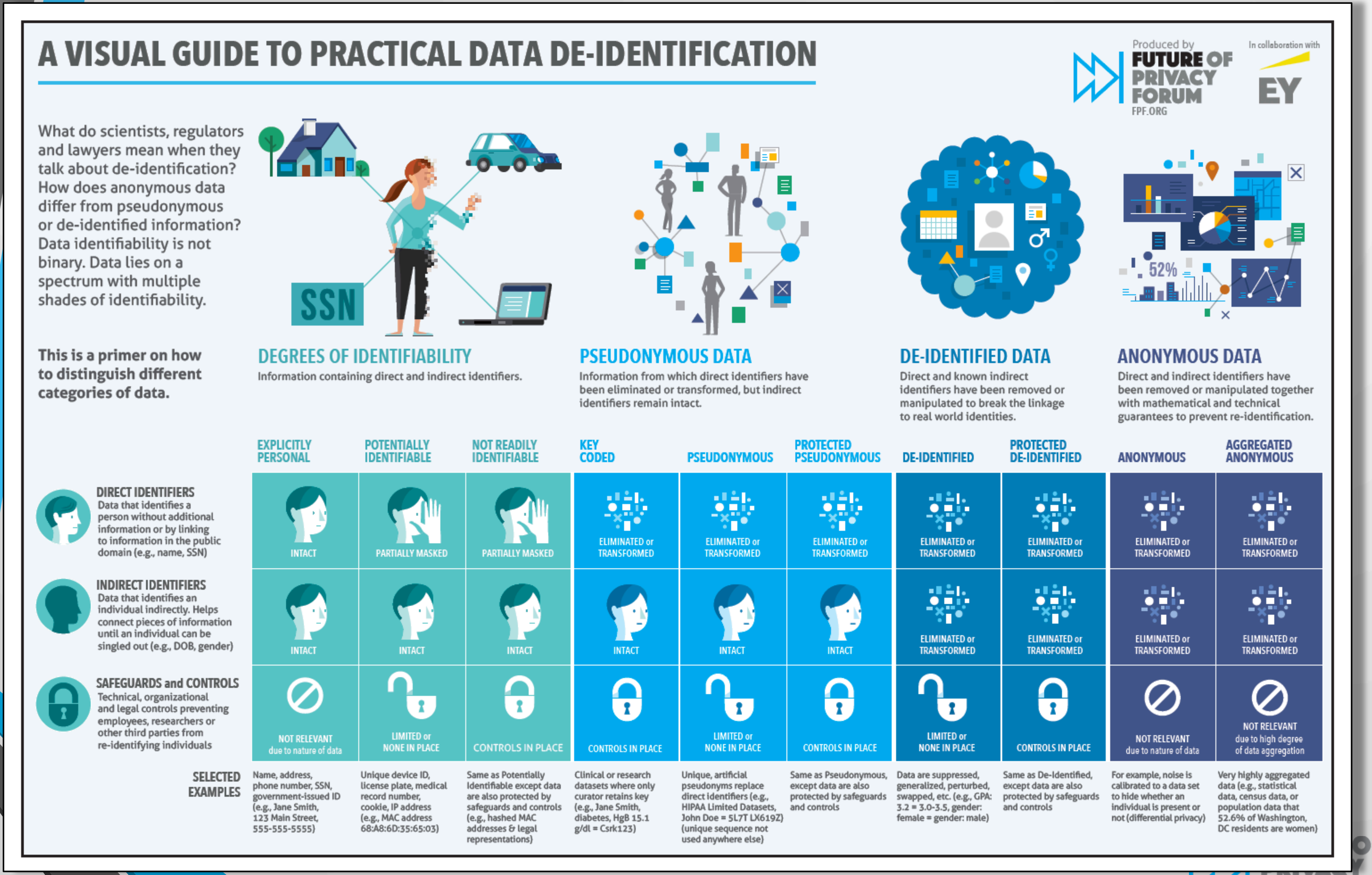
- Respect for persons
- Benefit-risk analysis
- Fairness & justice
- Due diligence
- Independent membership
- Process-oriented documentation

Governance

- Oversight
- Rapid response
- Confidentiality
- Transparency & Accountability

FUTURE OF PRIVACY FORUM

Solutions: Practical De-Identification



A VISUAL GUIDE TO PRACTICAL DATA DE-IDENTIFICATION

Produced by FUTURE OF PRIVACY FORUM FPF.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

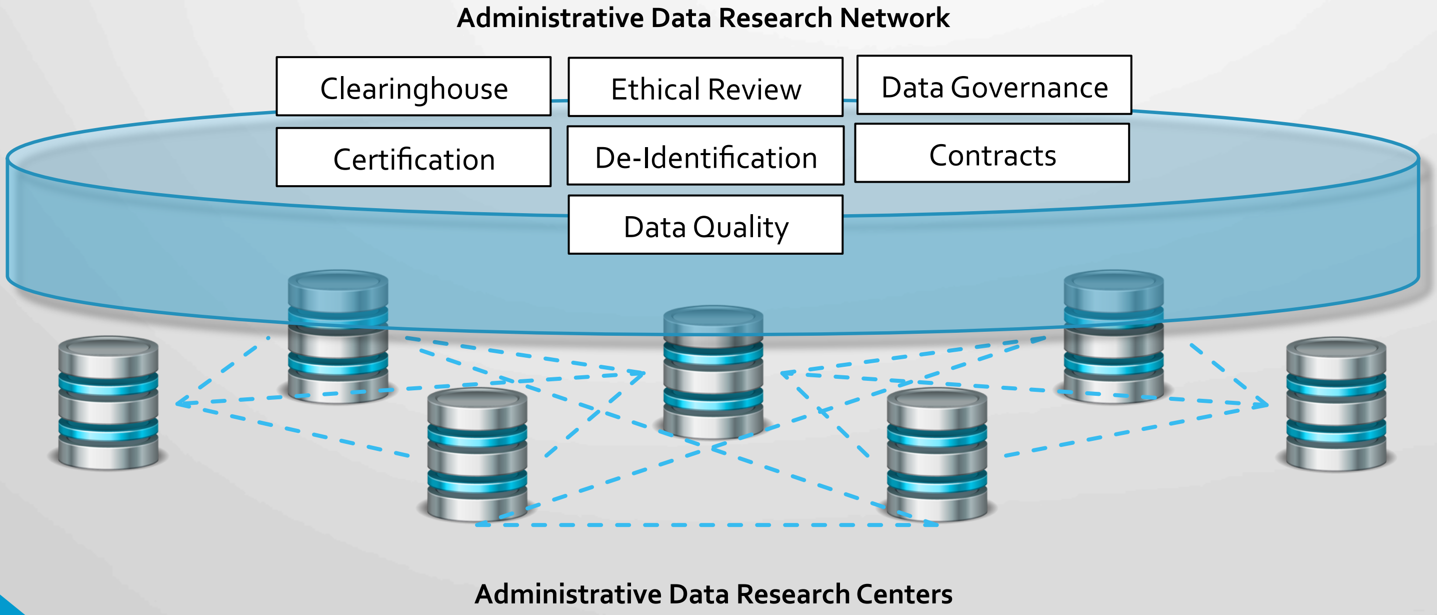
Direct and indirect identifiers have been removed or manipulated together with mathematical and technical guarantees to prevent re-identification.

	EXPLICITLY PERSONAL	POTENTIALLY IDENTIFIABLE	NOT READILY IDENTIFIABLE	KEY CODED	PSEUDONYMOUS	PROTECTED PSEUDONYMOUS	DE-IDENTIFIED	PROTECTED DE-IDENTIFIED	ANONYMOUS	AGGREGATED ANONYMOUS
DIRECT IDENTIFIERS Data that identifies a person without additional information or by linking to information in the public domain (e.g., name, SSN)	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 Data that identifies an individual indirectly. Helps connect pieces of information until an individual can be singled out (e.g., DOB, gender)	INTACT	INTACT	INTACT	INTACT	INTACT	INTACT	ELIMINATED or TRANSFORMED	ELIMINATED or TRANSFORMED	ELIMINATED or TRANSFORMED	ELIMINATED or TRANSFORMED
SAFEGUARDS and CONTROLS Technical, organizational and legal controls preventing employees, researchers or other third parties from re-identifying individuals	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:A8:6D: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 curator retains key (e.g., Jane Smith, diabetes, HgB 15.1 g/dl = Csrk123)	Unique, artificial pseudonyms replace direct identifiers (e.g., HIPAA Limited Datasets, John Doe = 5L7T LX619Z) (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)

Solutions: "ADRN/ADRC"

A **network** of **Administrative Data Research Centers (ADRCs)** can provide:

- Researcher & Data Clearinghouse
- Researcher Certification
- Ethical Review Structure
- De-Identification Expertise
- Data Governance
- Data Quality
- Standard Contracts



Administrative Data Research Network

Clearinghouse

Ethical Review

Data Governance

Certification

De-Identification

Contracts

Data Quality

Administrative Data Research Centers

Smart Cities

Issues

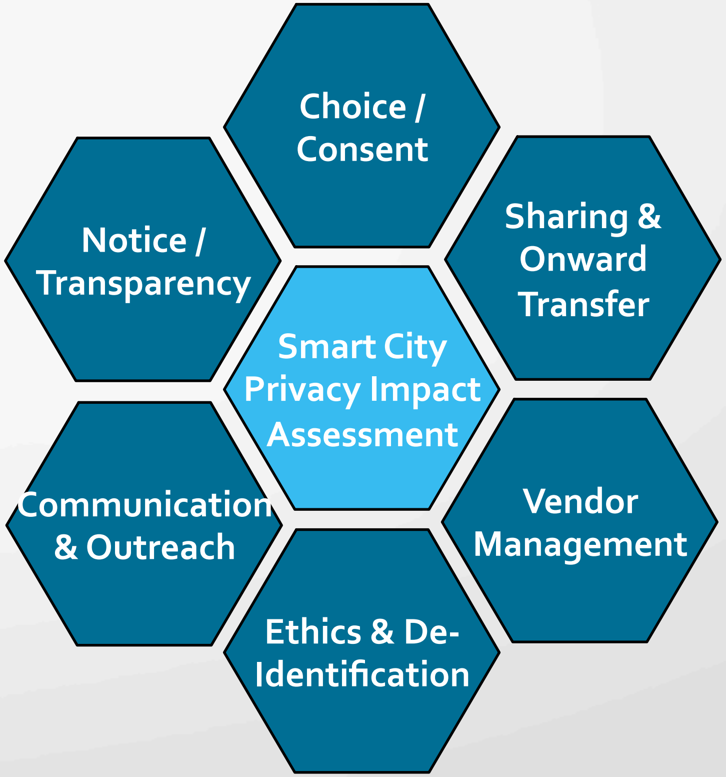
- Limited Choice
- Ethical & Societal Risks
- Open Data Requirements
- Vendor Management
- Public-Private Partnerships
- Equity

Solution

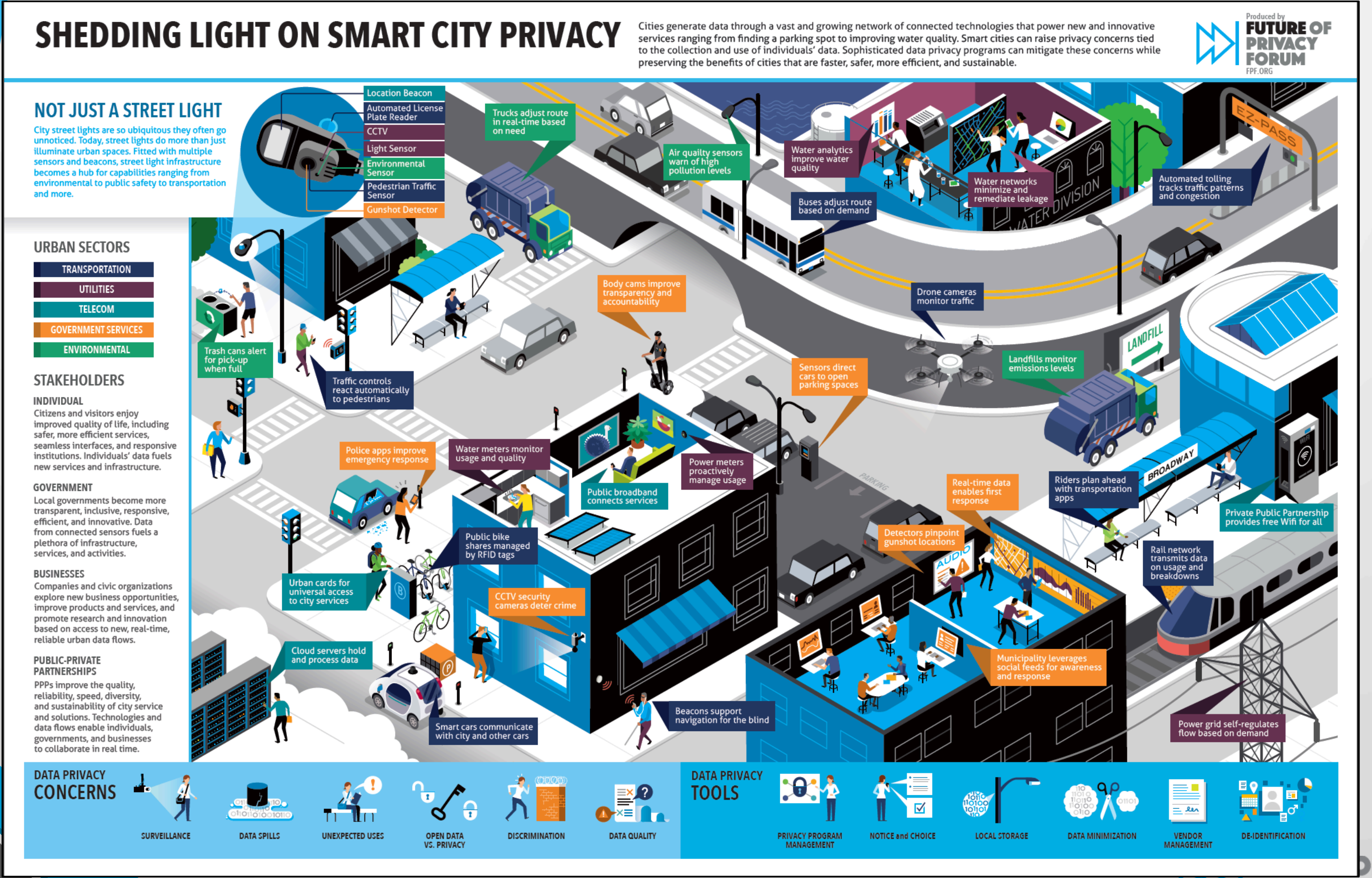
Develop a framework to help smart cities and technology partners identify privacy-related risks and proactively develop mitigation strategies

Scope

- City Connectivity
- Infrastructure Sensors
- Data Analytics
- Public Transportation
- Civic Identity Management
- And more…


Choice / Consent
Notice / Transparency
Sharing & Onward Transfer
Smart City Privacy Impact Assessment
Communication & Outreach
Vendor Management
Ethics & De-Identification

Smart Cities



SHEDDING LIGHT ON SMART CITY PRIVACY

Cities generate data through a vast and growing network of connected technologies that power new and innovative services ranging from finding a parking spot to improving water quality. Smart cities can raise privacy concerns tied to the collection and use of individuals' data. Sophisticated data privacy programs can mitigate these concerns while preserving the benefits of cities that are faster, safer, more efficient, and sustainable.

Produced by FUTURE OF PRIVACY FORUM
FPF.ORG

NOT JUST A STREET LIGHT

City street lights are so ubiquitous they often go unnoticed. Today, street lights do more than just illuminate urban spaces. Fitted with multiple sensors and beacons, street light infrastructure becomes a hub for capabilities ranging from environmental to public safety to transportation and more.

Location Beacon
Automated License Plate Reader
CCTV
Light Sensor
Environmental Sensor
Pedestrian Traffic Sensor
Gunshot Detector

URBAN SECTORS

- TRANSPORTATION
- UTILITIES
- TELECOM
- GOVERNMENT SERVICES
- ENVIRONMENTAL

STAKEHOLDERS

INDIVIDUAL
Citizens and visitors enjoy improved quality of life, including safer, more efficient services, seamless interfaces, and responsive institutions. Individuals' data fuels new services and infrastructure.

GOVERNMENT
Local governments become more transparent, inclusive, responsive, efficient, and innovative. Data from connected sensors fuels a plethora of infrastructure, services, and activities.

BUSINESSES
Companies and civic organizations explore new business opportunities, improve products and services, and promote research and innovation based on access to new, real-time, reliable urban data flows.

PUBLIC-PRIVATE PARTNERSHIPS
PPPs improve the quality, reliability, speed, diversity, and sustainability of city service and solutions. Technologies and data flows enable individuals, governments, and businesses to collaborate in real time.

Trucks adjust route in real-time based on need

Air quality sensors warn of high pollution levels

Water analytics improve water quality

Buses adjust route based on demand

Water networks minimize and remediate leakage

Automated tolling tracks traffic patterns and congestion

Trash cans alert for pick-up when full

Traffic controls react automatically to pedestrians

Body cams improve transparency and accountability

Drone cameras monitor traffic

Sensors direct cars to open parking spaces

Landfills monitor emissions levels

Police apps improve emergency response

Water meters monitor usage and quality

Power meters proactively manage usage

Public broadband connects services

Real-time data enables first response

Riders plan ahead with transportation apps

Private Public Partnership provides free Wifi for all

Public bike shares managed by RFID tags

Detectors pinpoint gunshot locations

Rail network transmits data on usage and breakdowns

Urban cards for universal access to city services

CCTV security cameras deter crime

Cloud servers hold and process data

Municipality leverages social feeds for awareness and response

Smart cars communicate with city and other cars

Beacons support navigation for the blind

Power grid self-regulates flow based on demand

DATA PRIVACY CONCERNS

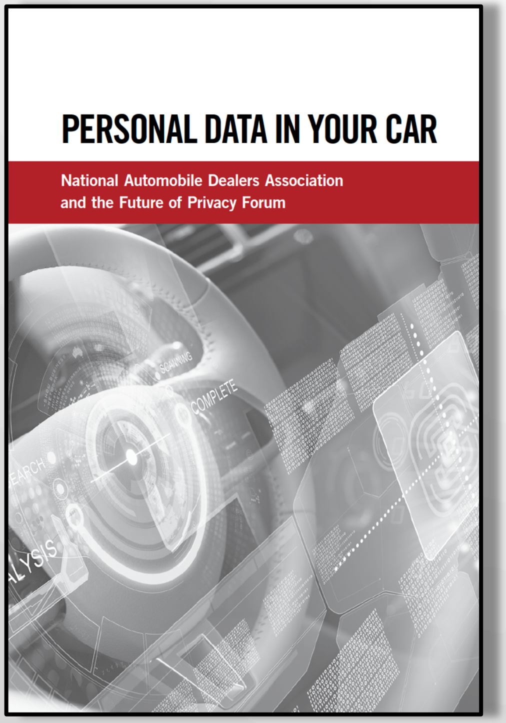
- SURVEILLANCE
- DATA SPILLS
- UNEXPECTED USES
- OPEN DATA VS. PRIVACY
- DISCRIMINATION
- DATA QUALITY

DATA PRIVACY TOOLS

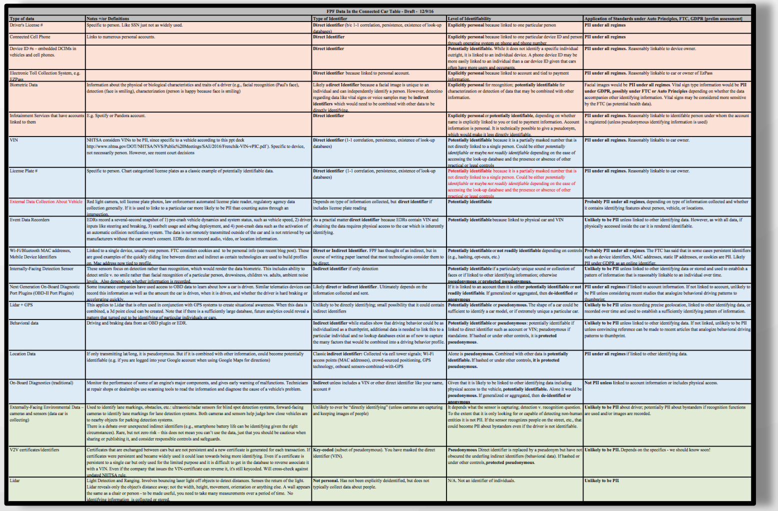
- PRIVACY PROGRAM MANAGEMENT
- NOTICE and CHOICE
- LOCAL STORAGE
- DATA MINIMIZATION
- VENDOR MANAGEMENT
- DE-IDENTIFICATION

Connected Cars

- Consumer Guide to the Connected Car
- Comments to NHTSA Automated Vehicle Guidance
- Analysis of PII in the car


PERSONAL DATA IN YOUR CAR

National Automobile Dealers Association and the Future of Privacy Forum


FPF Data In the Connected Car Table - Draft - 12/9/16

Type of data	Notes +/or Definitions	Type of Identifier	Level of Identifiability	Application of Standards under Auto Principles, FTC, GDPR [prelim assessment]
Driver's License #	Specific to person. Like SSN just not as widely used.	**Direct identifier** (b/c 1-1 correlation, persistence, existence of look-up databases)	**Explicitly personal** because linked to one particular person	**PII under all regimes**
Connected Cell Phone	Links to numerous personal accounts.	**Direct Identifier**	**Explicitly personal** because linked to one particular device ID and person through operating system on phone and phone number	**PII under all regimes**
Device ID #s – embedded DCIMs in vehicles and cell phones.		**Direct identifier**	**Potentially identifiable.** While it does not identify a specific individual outright, it is linked to an individual device. A phone device ID may be more easily linked to an individual than a car device ID given that cars often have more users and occupants.	**PII under all regimes.** Reasonably linkable to device owner.
Electronic Toll Collection System, e.g. EZPass		**Direct identifier** because linked to personal account.	**Explicitly personal** because linked to account and tied to payment information.	**PII under all regimes.** Reasonably linkable to car or owner of EzPass
Biometric Data	Information about the physical or biological characteristics and traits of a driver (e.g., facial recognition (Paul's face), detection (face is smiling), characterization (person is happy because face is smiling)	Likely a **direct Identifier** because a facial image is unique to an individual and can independently identify a person. However, detectino regarding data like vital signs or voice samples may be **indirect identifiers** which would need to be combined with other data to be directly identifying.	**Explicitly personal** for recognition; **potentially identifiable** for characterization or detection of data that may be combined with other information.	Facial images would be **PII under all regimes**. Vital sign type information would be **PII under GDPR, possibly under FTC or Auto Principles** depending on whether the data accompanies other identifying information. Vital signs may be considered more sensitive by the FTC (as potential health data).
Infotainment Services that have accounts linked to them	E.g. Spotify or Pandora account.	**Direct identifier**	**Explicitly personal** or **potentially identifiable**, depending on whether name is explicitly linked to you or tied to payment information. Account information is personal. It is technically possible to give a pseudonym, which would make it less directly identifiable.	**PII under all regimes.** Reasonably linkable to identifable person under whom the account is registered (unless pseudonymous identifying information is used)
VIN	NHTSA considers VINs to be PII, since specific to a vehicle according to this ppt deck http://www.nhtsa.gov/DOT/NHTSA/NVS/Public%20Meetings/SAE/2016/Frenchik-VIN-vPIC.pdf). Specific to device, not necessarily person. However, see recent court decisions	**Direct identifier** (1-1 correlation, persistence, existence of look-up databases)	**Potentially identifiable** because it is a partially masked number that is not directly linked to a single person. Could be either *potentially identifiable* or maybe *not readily identifiable* depending on the ease of accessing the look-up database and the presence or absence of other practical or legal controls	**PII under all regimes.** Reasonably linkable to car owner.
License Plate #	Specific to person. Chart categorized license plates as a classic example of potentially identifiable data.	**Direct identifier** (1-1 correlation, persistence, existence of look-up databases)	**Potentially identifiable** because it is a partially masked number that is not directly linked to a single person. Could be either *potentially identifiable* or maybe *not readily identifiable* depending on the ease of accessing the look-up database and the presence or absence of other practical or legal controls	**PII under all regimes.** Reasonably linkable to car owner.
External Data Collection About Vehicle	Red light camera, toll license plate photos, law enforcement automated license plate reader, regulatory agency data collection generally. If it is used to linke to a particular car more likely to be PII than counting autos through an intersection.	Depends on type of information collected, but **direct identifier** if includes license plate reading	**Potentially identifiable**	**Probably PII under all regimes**, depending on type of information collected and whether it contains identifying features about person, vehicle, or locations.
Event Data Recorders	EDRs record a several-second snapshot of 1) pre-crash vehicle dynamics and system status, such as vehicle speed, 2) driver inputs like steering and breaking, 3) seatbelt usage and airbag deployment, and 4) post-crash data such as the activation of an automatic collision notification system. The data is not remotely transmitted outside of the car and is not retrieved by car manufacturers without the car owner's consent. EDRs do not record audio, video, or location information.	As a practial matter **direct identifier** because EDRs contain VIN and obtaining the data requires physical access to the car which is inherently identifying.	**Potentially identifiable** because linked to physical car and VIN	**Unlikely to be PII** unless linked to other identifying data. However, as with all data, if physically accessed inside the car it is rendered identifiable.
Wi-Fi/Bluetooth MAC addresses, Mobile Device Identifiers	Linked to a single device, usually one person. FTC considers cookies and to be personal info (see recent blog post). These are good examples of the quickly eliding line between direct and indirect as certain technologies are used to build profiles on. Mac address now tied to profile.	**Direct or Indirect Identifier.** FPF has thought of as indirect, but in course of writing paper learned that most technologists consider them to be direct.	**Potentially identifiable** or **not readily identifiable** depending on controls (e.g., hashing, opt-outs, etc.)	**Probably PII under all regimes**. The FTC has said that in some cases persistent identifiers such as device identifiers, MAC addresses, static IP addresses, or cookies are PII. Likely PII under GDPR as an online identifier.
Internally-Facing Detection Sensor	These sensors focus on detection rather than recognition, which would render the data biometric. This includes ability to detect smile v. no smile rather than facial recognition of a particular person, drowsiness, children vs. adults, ambient noise levels. Also depends on whether information is recorded.	**Indirect identifier** if only detection	**Potentially identifiable** if a particularly unique sound or collection of faces or if linked to other identifying information; otherwise **pseudonymous** or **protected pseudonymous.**	**Unlikely to be PII** unless linked to other identifying data or stored and used to establish a pattern of information that is reasonably linkable to an individual over time.
Next Generation On-Board Diagnostic Port Plugins (OBD-II Port Plugins)	Some insurance companies have used access to OBD data to learn about how a car is driven. Similar telematics devices can record this information as well as the amount the car is driven, when it is driven, and whether the driver is hard braking or accelerating quickly.	Likely **direct** or **indirect identifier**. Ultimately depends on the information collected and sent.	If it is linked to an account then it is either **potentially identifiable** or **not readily identifiable.** If generalized or aggregated, then **de-identified** or **anonymous**	**PII under all regimes** if linked to account information. If not linked to account, unlikely to be PII unless considering recent studies that analogize behavioral driving patterns to thumbprint.
Lidar + GPS	This applies to Lidar that is often used in conjunction with GPS systems to create situational awareness. When this data is combined, a 3d point cloud can be created. Note that if there is a sufficiently large database, future analytics could reveal a pattern that turned out to be identifying of particular individuals or cars.	Unlikely to be directly identifying; small possibility that it could contain indirect identifiers	**Potentially identifiable** or **pseudonymous.** The shape of a car could be sufficient to identify a car model, or if extremely unique a particular car.	**Unlikely to be PII** unless recording precise geolocation, linked to other identifying data, or recorded over time and used to establish a sufficiently identifying pattern of information.
Behavioral data	Driving and braking data from an OBD plugin or EDR.	**Indirect identifier** while studies show that driving behavior could be as individualized as a thumbprint, additional data is needed to link this to a particular individual and no lookup databases exist as of now to capture the many factors that would be combined into a driving behavior profile.	**Potentially identifiable** or **pseudonymous**: potentially identifiable if linked to direct identifier such as account or VIN; pseudonymous if standalone. If hashed or under other controls, it is **protected pseudonymous**.	**Unlikely to be PII** unless linked to other identifying data. If not linked, unlikely to be PII unless convincing reference can be made to recent articles that analogize behavioral driving patterns to thumbprint.
Location Data	If only transmitting lat/long, it is pseudonymous. But if it is combined with other information, could become potentially identifiable (e.g. if you are logged into your Google account when using Google Maps for directions)	Classic **indirect identifier:** Collected via cell tower signals; Wi-Fi access points (MAC addresses), crowd-sourced positioning, GPS technology, onboard sensors-combined-with-GPS	Alone is **pseudonymous.** Combined with other data is **potentially identifiable.** If hashed or under other controls, **it is protected pseudonymous.**	**PII under all regimes** if linked to other identifying data.
On-Board Diagnostics (traditional)	Monitor the performance of some of an engine's major components, and gives early warning of malfunctions. Technicians at repair shops or dealerships use scanning tools to read the information and diagnose the cause of a vehicle's problem.	**Indirect** unless includes a VIN or other direct identifier like your name, account #	Given that it is likely to be linked to other identifying data including physical access to the vehicle, **potentially identifiable.** Alone it would be **pseudonymous.** If generalized or aggregated, then **de-identified or anonymous**	**Not PII unless** linked to account information or includes physical access.
Externally-Facing Environmental Data – cameras and sensors (data car is collecting)	Used to identify lane markings, obstacles, etc.: ultrasonic/radar sensors for blind spot detection systems, forward-facing cameras to identify lane markings for lane detection systems. Both cameras and sensors help judge how close vehicles are to nearby objects for parking detection systems. There is a debate over unexpected indirect identifiers (e.g., smartphone battery life can be identifying given the right circumstances). Rare, but not zero risk – this does not mean you can't use the data, just that you should be cautious when sharing or publishing it, and consider responsible controls and safeguards.	Unlikely to ever be "directly identifying" (unless cameras are capturing and keeping images of people)	It depends what the sensor is capturing; detection v. recognition question. To the extent that it is only looking for or capable of detecting non-human entities it is not PII. If the sensor recognizes people on the street, etc., that could become PII about bystanders even if the driver is not identifiable.	**Unlikely to be PII** about driver; potentially PII about bystanders if recognition functions are used and/or images are recorded.
V2V certificates/identifiers	Certificates that are exchanged between cars but are not persistent and a new certificate is generated for each transaction. If certificates were persistent and became widely used it could lean towards being more identifying. Even if a certificate is persistent to a single car but only used for the limited purpose and it is difficult to get in the database to reverse associate it with a VIN. Even if the company that issues the VIN-certificate can reverse it, it's still keycoded. Will cross-check against updated NHTSA rule.	**Key-coded** (subset of pseudonymous). You have masked the direct identifier (VIN).	**Pseudonymous** Direct identifier is replaced by a pseudonym but have not obscured the underling indirect identifiers (behavioral data). If hashed or under other controls, **protected pseudonymous**.	**Unlikely to be PII.** Depends on the specifics - we should know soon!
Lidar	Light Detection and Ranging. Involves bouncing laser light off objects to detect distances. Senses the return of the light. Lidar reveals only the object's distance away; not the width, height, movement, orientation or anything else. A wall appears the same as a chair or person - to be made useful, you need to take many measurements over a period of time. No identifying information is collected or stored.	**Not personal.** Has not been explicitly deidentified, but does not typically collect data about people.	N/A. Not an identifier of individuals.	**Unlikely to be PII**

FUTURE OF PRIVACY FORUM

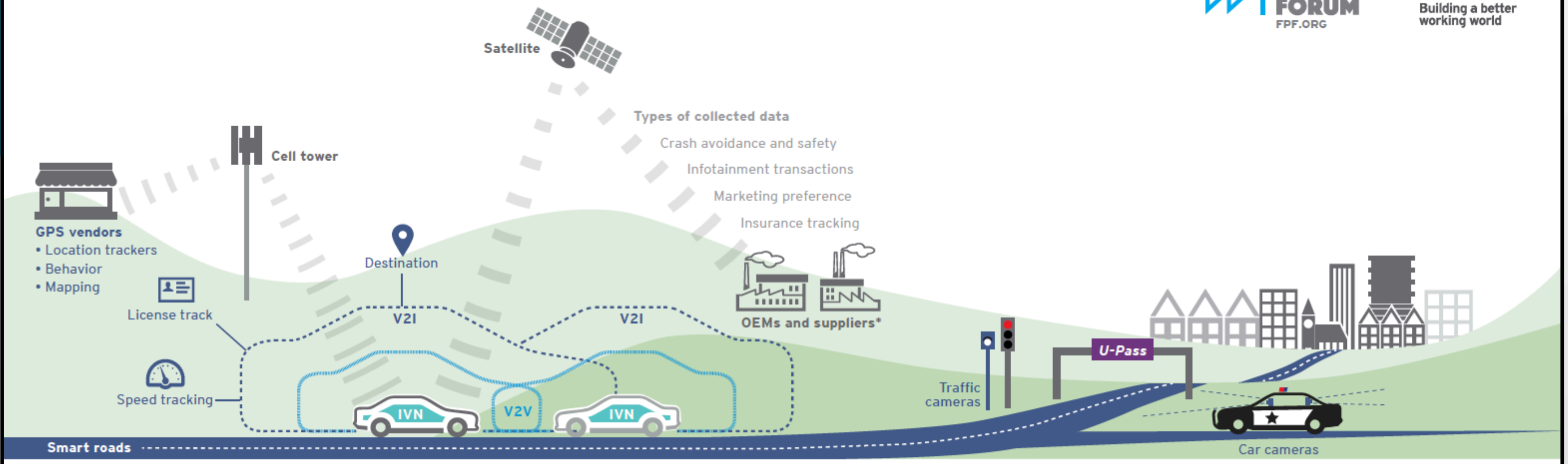
Connected Cars

A VISUAL GUIDE TO AUTOMOTIVE CONNECTIVITY

The connected vehicle ecosystem includes internal and external information processing systems that generate and process data related to vehicle and occupant use.

Produced by
FUTURE OF PRIVACY FORUM
FPF.ORG

In collaboration with
EY
Building a better working world



🟩 IVN = in-vehicle network ||| V2V = vehicle-to-vehicle network ▮▮ V2I = vehicle-to-infrastructure network

*Original equipment manufacturers and auto parts suppliers

Data sharing

Challenge – The connected vehicle ecosystem challenges traditional boundaries of data ownership. Vehicle systems generate a plethora of data, including location, speed, driving behavior and road conditions. This data is generated and transferred to various destinations, such as vendors, OEMs and suppliers, for numerous purposes, blurring established lines of choice and consent.

Data ownership – Establish an understanding of vendor-based processing and rights over data ownership that consider data origination, processing and sharing requirements. Guidance should consider data subject consent and vendor processing rights to establish ownership and avoid inappropriate processing or sharing.

Contract management – An inventory of vendors receiving identifiable information is maintained throughout the vendor life cycle. Contracts contain appropriate language to stipulate the vendors' responsibilities for protecting personal information during processing and expectations for actions during breaches.

Governance

Challenge – The diverse nature of the vehicle ecosystem creates challenges in obtaining alignment to a common data protection standard. Disparate technologies s,uch as the methods used to secure and transmit data in IVN, V2V and V2I, networks, and on to various vendors' networks, require the establishment of a common language of data protection principles and standards.

Unified approach – Develop a privacy framework and associated controls that consider all components of the vehicle ecosystem to validate compliance with regulatory and compliance data protection requirements.

Privacy Impact Assessments (PIAs) – By moving the assessment of the use of personal information to the beginning of the development life cycle, rework and inappropriate use of data can be avoided. PIAs can be used to assess privacy risk in the development of new technologies, third-party access, data retention and aggregation and/or anonymization of data to reduce risk.

Choice

Challenge – Data created in the vehicle ecosystem supports vehicle safety, personal preferences and behavioral use. Obtaining, monitoring and complying with user choice is complicated by the multitude of purposes for which data was collected, processed and disclosed to entities with which the data is shared.

In-vehicle notices – Create just-in-time notices that describe use and sharing of data and are served prior to processing. Just-in-time notices include a description of how the user's personal information will be used for each interaction involving personal information. The notice includes the data elements collected, for what purposes, who it is shared with and why, and how data is retained and includes the name of a privacy contact.

Data use option – A mechanism to enable the user to choose how data will be used when data is not required to fulfill a necessary vehicle function. The mechanism includes a contextual explanation of the purposes for collecting user data and enables users to provide informed consent.

Safeguards

Challenge – An effective implementation of a consistent control framework across the vehicle ecosystem that contains multi-vendor systems requires coordination of governance and technical standards, as well as controls.

Compliance Control Framework – Establish the control framework for remote, connected, local and authorized security and privacy for secure and authorized processing.

Govern safeguards by establishing accountability over operational support processes and resources by defining clear accountabilities.

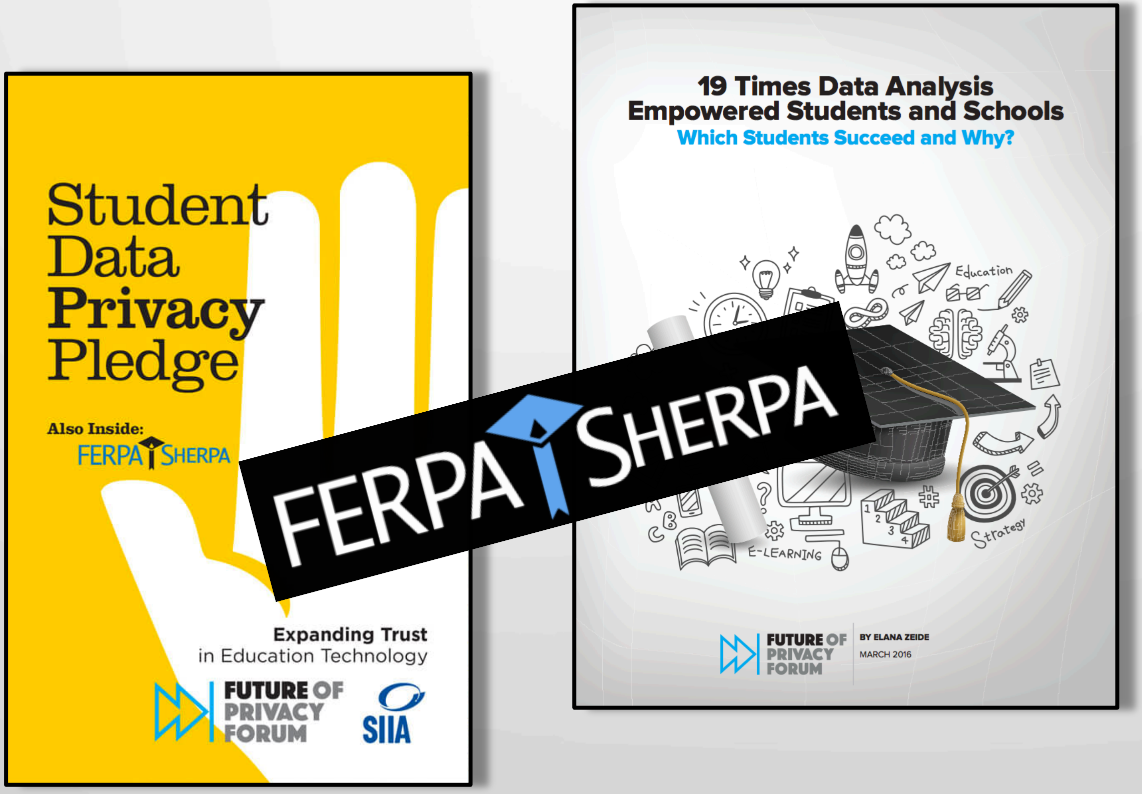
Lock down the essentials and manage the security of vehicle-to-vehicle communications.

Encrypted data transport – Personal information is stored and transmitted using industry-recognized encryption methods to prevent unauthorized access to personal information.

Logical access controls – Implementation of appropriate logical access controls for resources supporting back-end processing.

Student Data

- Student Data Privacy Pledge: 350+ companies commit to specific legal protection of student data
- FERPA Sherpa: Parents' and Educators' guides to student privacy
- K-12 and Higher Education Working Groups
- Ongoing studies regarding beneficial student data uses



Student Data Privacy Pledge

Also Inside: FERPA SHERPA

Expanding Trust
in Education Technology

FUTURE OF PRIVACY FORUM

SIIA

FERPA SHERPA

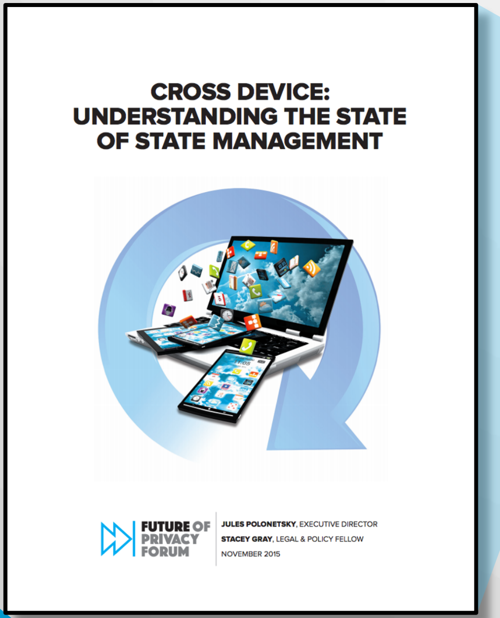
19 Times Data Analysis Empowered Students and Schools
Which Students Succeed and Why?

FUTURE OF PRIVACY FORUM

BY ELANA ZEIDE
MARCH 2016

Location & Ad Tech

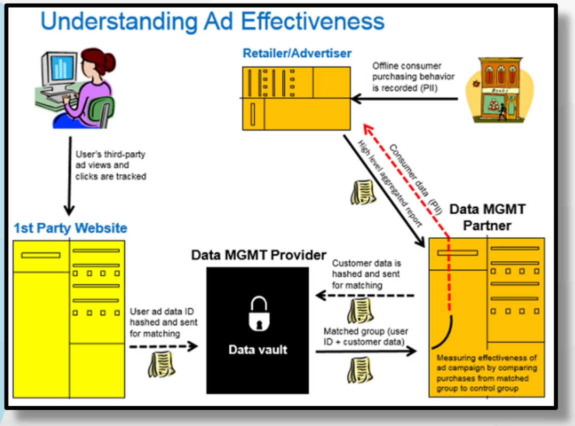
- Mobile Location Analytics Code of Conduct
- In progress: Ad Tech Due Diligence Guide
- Research on cross-device tracking, state management, and precise geolocation data collection



CROSS DEVICE: UNDERSTANDING THE STATE OF STATE MANAGEMENT

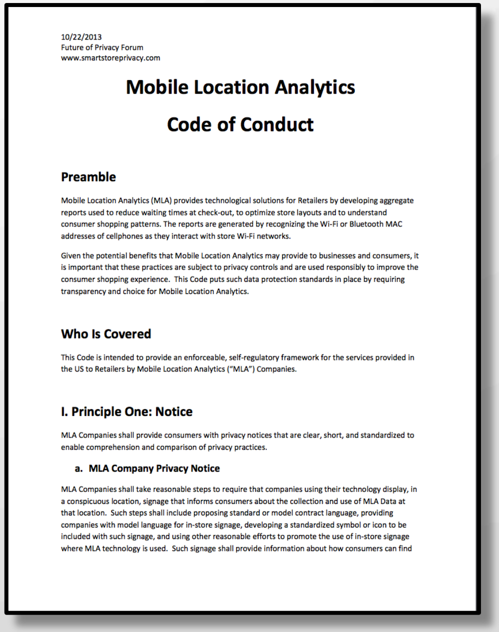
FUTURE OF PRIVACY FORUM

JULES POLONETSKY, EXECUTIVE DIRECTOR
STACEY GRAY, LEGAL & POLICY FELLOW
NOVEMBER 2015



Understanding Ad Effectiveness

Retailer/Advertiser
Offline consumer purchasing behavior is recorded (PII)
User's third-party ad views and clicks are tracked
Consumer data (PII)
High level aggregated report
Data MGMT Partner
1st Party Website
Data MGMT Provider
Customer data is hashed and sent for matching
User ad data ID hashed and sent for matching
Data vault
Matched group (user ID + customer data)
Measuring effectiveness of ad campaign by comparing purchases from matched group to control group



10/22/2013
Future of Privacy Forum
www.smartstoreprivacy.com

Mobile Location Analytics

Code of Conduct

Preamble

Mobile Location Analytics (MLA) provides technological solutions for Retailers by developing aggregate reports used to reduce waiting times at check-out, to optimize store layouts and to understand consumer shopping patterns. The reports are generated by recognizing the Wi-Fi or Bluetooth MAC addresses of cellphones as they interact with store Wi-Fi networks.

Given the potential benefits that Mobile Location Analytics may provide to businesses and consumers, it is important that these practices are subject to privacy controls and are used responsibly to improve the consumer shopping experience. This Code puts such data protection standards in place by requiring transparency and choice for Mobile Location Analytics.

Who Is Covered

This Code is intended to provide an enforceable, self-regulatory framework for the services provided in the US to Retailers by Mobile Location Analytics ("MLA") Companies.

I. Principle One: Notice

MLA Companies shall provide consumers with privacy notices that are clear, short, and standardized to enable comprehension and comparison of privacy practices.

a. MLA Company Privacy Notice

MLA Companies shall take reasonable steps to require that companies using their technology display, in a conspicuous location, signage that informs consumers about the collection and use of MLA Data at that location. Such steps shall include proposing standard or model contract language, providing companies with model language for in-store signage, developing a standardized symbol or icon to be included with such signage, and using other reasonable efforts to promote the use of in-store signage where MLA technology is used. Such signage shall provide information about how consumers can find

Internet of Things

Several reports and ongoing research projects:

- Best Practice Guide for Wearable Devices
- Privacy Implications of Microphones in the Home
- Drones & Privacy by Design
- Kids & the Connected Home

Best Practices for
Consumer Wearables &
Wellness Apps & Devices

August 17, 2016

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

Always On:
Privacy Implications of
Microphone-Enabled Devices

In collaboration with

BY STACEY GRAY
APRIL 2016

Drones and Privacy by Design:
Embedding Privacy Enhancing
Technology in Unmanned Aircraft

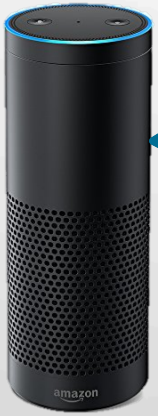
August 2, 2016

KIDS & THE CONNECTED HOME:
PRIVACY IN THE AGE OF CONNECTED
DOLLS, TALKING DINOSAURS,
AND BATTLING ROBOTS

DECEMBER 2016

Consumer-business dialogue leads to privacy tech innovations.

Alexa, are you spying on me?



I only send audio back to Amazon when I hear you say the wake word. For more information, and to view Amazon's privacy policy, visit the help section of your Alexa app.

Will Europe Set the Agenda?

- May 2018 Effective Date of the new General Data Protection Regulation
  - Major rights backed up by significant penalties
  - Extra-territorial
  - Platforms build services globally.

PLUS: May 2018 – ePrivacy regulation, still draft, regulating the Internet of Things, cookies, tracking devices, over the top services.

States and Cities Will Set The Agenda

Student Privacy Law

Facial Recognition and Biometrics

Location Track

Right to Be Forgotten

Attorneys General, Class Action Bar, and local commissions.

Thank You!

Jules Polonetsky
CEO, Future of Privacy Forum

Visit our site: **http://www.fpf.org**

- www.fpf.org
- @futureofprivacy
- @JulesPolonetsky

FUTURE OF PRIVACY FORUM