Hearing #12 on Competition and Consumer Protection in the 21st Century

Federal Trade Commission

Constitution Center

April 9, 2019

Welcome

We Will Be Starting Shortly

Welcome and Introductory Remarks

Jim Trilling
Federal Trade Commission
Division of Privacy and
Identity Protection

Opening Remarks

Joseph J. Simons, Chairman
Federal Trade Commission

Goals of Privacy Protection

What Have We Heard So Far?

Session moderated by:

James C. Cooper

Federal Trade Commission

Bureau of Consumer Protection

What Should We Do About Privacy?

- What do consumers want?

- Is there some reason firms aren’t responding? (i.e., Is there a market failure?)

- Is there something government can do to improve things?

What Do Consumers Want?

- Survey evidence suggests privacy is very important

- Revealed preference suggests the opposite

- Privacy Paradox

Is There A Market Failure?

- Possible culprits:
  - Asymmetric information
  - Cognitive biases

- Is understanding endogenous?
  - Rational ignorance

What Should Government Do?

- Benefits
  - Provide privacy levels consumers want
- Costs
  - Retarding information flows can have negative impacts on market performance and innovation

What Have We Heard?

- Even with full information, some experiments show consumers choosing to reveal information for very little compensation

- Increasing trust can increase willingness to share data:
  - Allowing control over third-party sharing of genetic information increases genetic testing rates
  - Increases level of Health Information Exchange operation in states with consent requirements

What Have We Heard?

- Behavioral targeting tends to generate more revenue for content providers than contextual
  - But, need to be careful when interpreting results due to strong selection effects
  - Increased revenue to content providers from behavioral targeting tends to be larger than lift to advertiser
  - AI not very good at identifying consumer attributes

- Opt-in reduces the quality of matching and data collection

What Have We Heard?

- Negative impact on investment:
  - VC investment
  - HIT investment & health outcomes

- Theory suggests that privacy regulation can have negative impact on competition

Goals of Privacy Protection

Panel Discussion:

Neil Chilson, Alastair Mactaggart,
Paul Ohm

Moderator: James Cooper

Break

10:30-10:45 am

The Data Risk Spectrum: From De-Identified Data to Sensitive Individually Identifiable Data

Session moderated by:

Cora Han & Elisa Jillson

Federal Trade Commission

Division of Privacy and Identity Protection

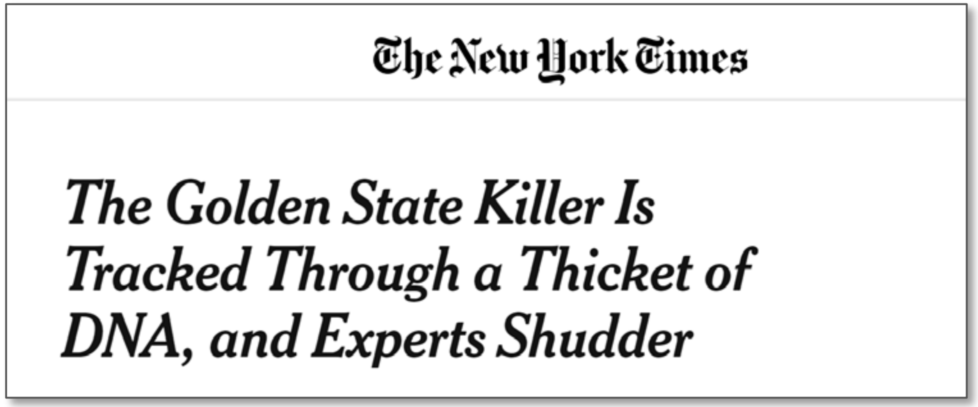
The Data Risk Spectrum: From De-Identified Data to Sensitive Individually Identifiable Data

Balancing Risk: De-Identification, Privacy and Precision

Jules Polonetsky

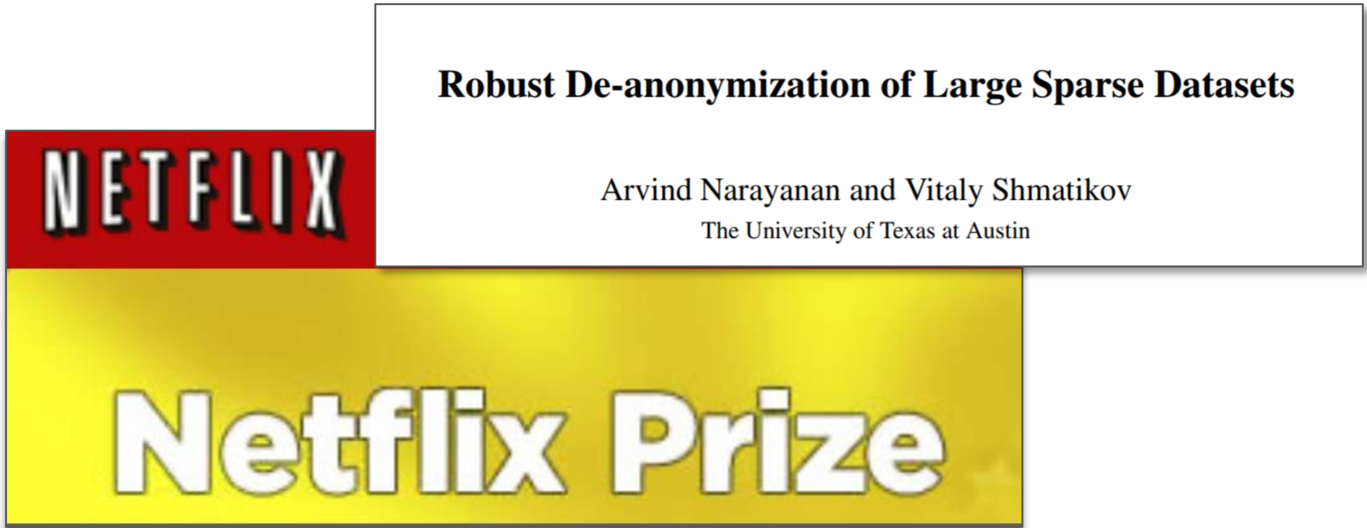
Future of Privacy Forum

Re-identification Attacks



The New York Times

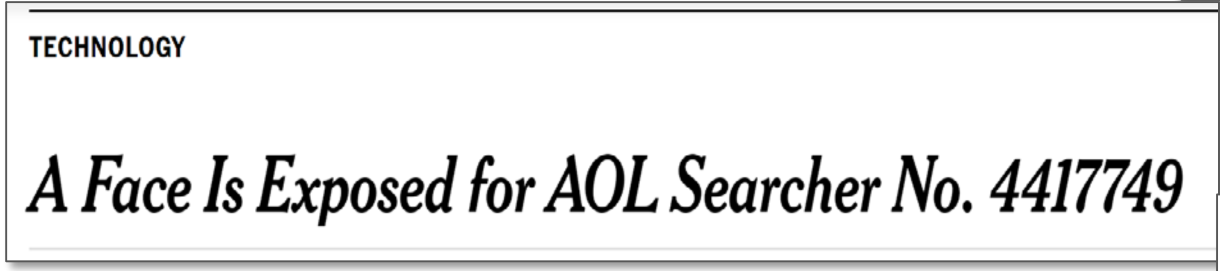
The Golden State Killer Is Tracked Through a Thicket of DNA, and Experts Shudder



Robust De-anonymization of Large Sparse Datasets

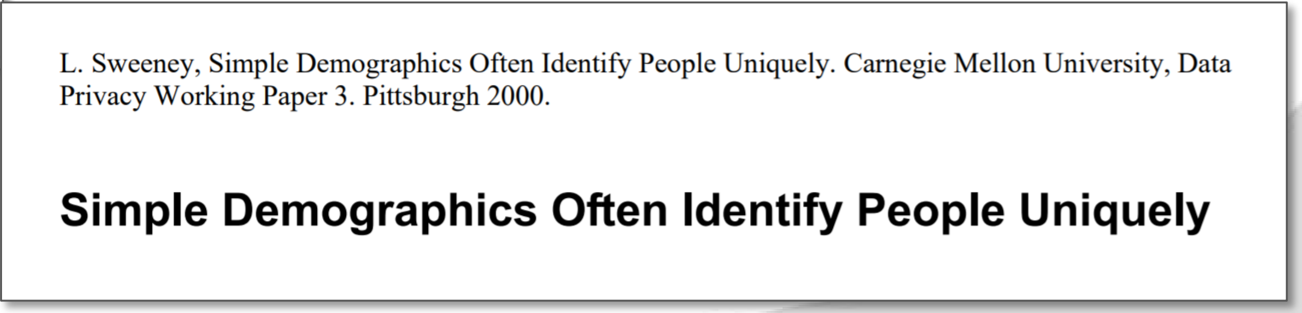
Arvind Narayanan and Vitaly Shmatikov
The University of Texas at Austin

NETFLIX
Netflix Prize



TECHNOLOGY

A Face Is Exposed for AOL Searcher No. 4417749



L. Sweeney, Simple Demographics Often Identify People Uniquely. Carnegie Mellon University, Data Privacy Working Paper 3. Pittsburgh 2000.

Simple Demographics Often Identify People Uniquely

Direct vs. Indirect Identifiers

- **Examples of direct identifiers:** Name, address, telephone number, fax number, health card number, health plan beneficiary number, license plate number, email address, photograph, biometrics, SSN.

- **Examples of indirect identifiers:** Sex, date of birth, age, geographic locations (postal codes, census geography, information about proximity to known or unique landmarks), language spoken at home, ethnic origin, total years of schooling, marital status, criminal history, total income, visible minority status, profession, event dates, number of children, high level diagnoses and procedures.

A VISUAL GUIDE TO PRACTICAL DATA DE-IDENTIFICATION

Produced by
FUTURE OF PRIVACY FORUM
FPF.ORG

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)

Privacy Shield

- **Key-coded Data.** Invariably, research data are uniquely key-coded at their origin by the principal investigator so as not to reveal the identity of individual data subjects. Pharmaceutical companies sponsoring such research do not receive the key. The unique key code is held only by the researcher, so that he or she can identify the research subject under special circumstances (e.g., if follow-up medical attention is required). **A transfer from the EU to the United States of data coded in this way would not constitute a transfer of personal data that would be subject to the Privacy Shield Principles.**

- Privacy Shield Supplemental Principles 14 (g),
https://www.privacyshield.gov/article?id=14-Pharmaceutical-and-Medical-Products

GDPR Pseudonymized Data

- Defined as Personal Information

- The GDPR requires controllers to collect data only for “specific, explicit and legitimate purposes.” Article 5 provides an exception to the purpose limitation principle, however, where data is further processed in a way that is “compatible” with the initial purposes for collection. Whether further processing is compatible depends on several factors outlined in Article 6(4), including the link between the processing activities, the context of the collection, the nature of the data, and the possible consequences for the data subject.

- **An additional factor to consider is “the existence of appropriate safeguards, which may include encryption or pseudonymization” (Article 6(4)(e)).**

- **GDPR allows controllers who pseudonymize personal data more leeway to process the data for a different purpose than the one for which they were collected.**

HIPAA Limited Data Set

- A LDS is **“protected health information”** that **excludes the following direct identifiers** of the individual or of relatives, employers, or household members of the individual:

- Names, Postal address information, other than town or city, State, and zip code, Telephone numbers, Fax numbers, Electronic mail addresses, Social Security numbers, Medical record numbers, Health-plan beneficiary numbers, Account numbers, Certificate and license numbers, Vehicle identifiers and serial numbers, including license plate numbers, Device identifiers and serial numbers, Web Universal Resource Locators (URLs), Internet Protocol (IP) address numbers, Biometric identifiers including fingerprints and voice prints, Full-face photographic images and any comparable image

- BUT a Limited Data Set MAY INCLUDE: Dates, city, town, state or full zip code (indirect identifiers)

- **May be shared but only for purposes of research, public health, or health care operations, if consistent with the original purpose for which it was disclosed**

Coded Data -Human Subjects Protection Under the Common Rule

- Office for Human Research Protections **does not consider research involving only coded private information or specimens to involve human subjects** if the following conditions are both met:

- (1) not collected specifically for the currently proposed research project through an interaction or intervention with living individuals; and
- (2) the investigator(s) cannot readily ascertain the identity of the individual(s) to whom the coded private information or specimens pertain because re-identification code is destroyed or held by an honest broker.

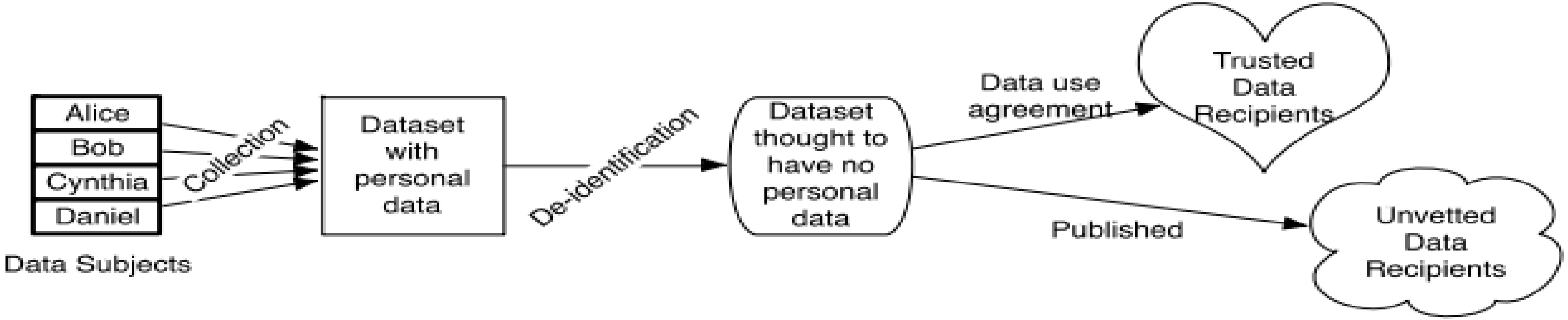
FERPA

- Studies Exception (see 20 U.S.C. §1232g(b)(1)(F) and §99.31(a)(6)): allows for the disclosure of PII from education records without consent to organizations conducting studies for, or on behalf of, schools, school districts, or postsecondary institutions. Studies can be for the purpose of developing, validating, or administering predictive tests; administering student aid programs; or improving instruction.

- Audit or Evaluation Exception (see 20 U.S.C. §1232g(b)(1)(C), (b)(3), and (b)(5) and §§99.31(a)(3) and 99.35): allows disclosure of PII from education records without consent to an organization for the purpose of conducting a study that compares program outcomes across school districts to further assess what programs provide the best instruction and then duplicate those results in other districts.

- **Requires the organization to conduct the study in a manner that does not permit the personal identification of parents and students by anyone other than representatives of the organization with legitimate interests. Contract must require the organization to conduct the study so as not to identify students or their parents.**

De-Identification Data Flow



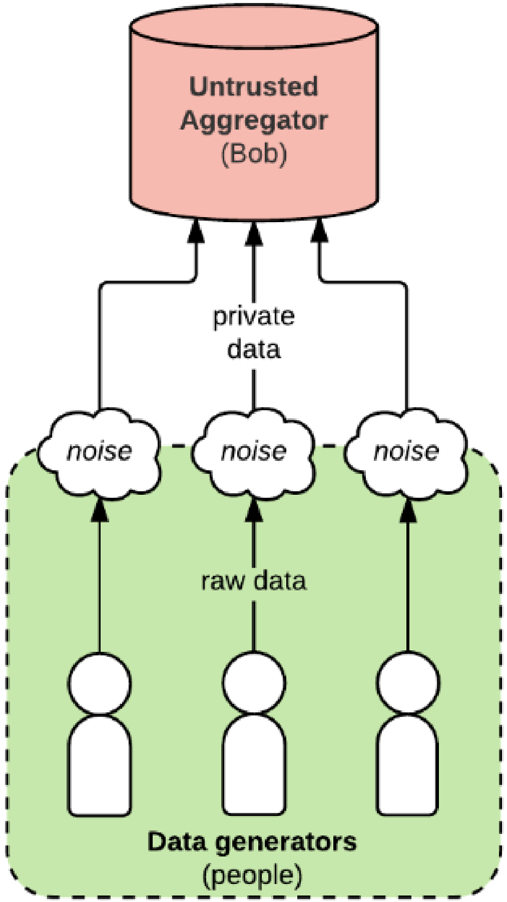
1. Remove direct identifiers;
2. Mask, transform, or de-identify indirect identifiers;
3. Test re-identification risk (e.g., motivated intruder test);
4. Make any further mitigations needed to adjust risk to an acceptable level;
5. Release with additional safeguards, as appropriate; and
6. Ongoing evaluation and testing for re-identification risk.

Assessing De-Identification Risks

- What risks are we concerned about? (Identity disclosure, attribute disclosure, inferences)
- Who are the attackers? (General public, expert attacker, company insider, information broker, neighbor)
- How much should we trust legal and administrative controls, such as contracts or commitments not to re-identify?
- How does sensitivity of the data heighten risk or affect our risks?
- Should pseudonymous data be treated more liberally? Should protected pseudonymous data, key coded pharma data be treated differently if strong controls exist?

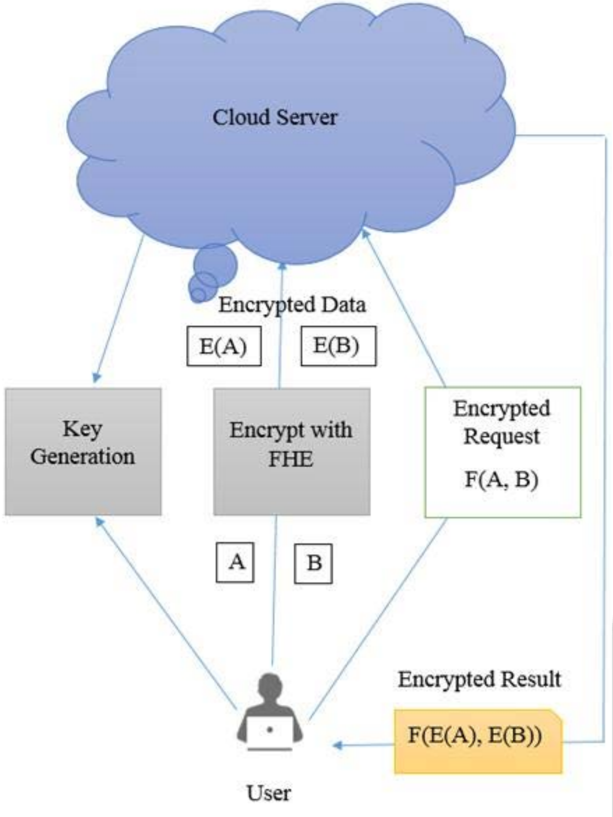
Differential Privacy

- **Definition:** A formal privacy concept that quantifies privacy loss. Data is considered differentially private when a data analyst always obtains the equivalent analytic result from a data set, *regardless of whether an individual's data is included in that data or whether it is excluded from it.*
- **Example:** Cheating on a test.
  1. We want to know how many users have ever cheated on a test.
  2. Ask users if they have ever cheated on a test by choosing either "yes" or "no."
  3. Add "noise" to the database containing user answers.
  4. If you know how the noise is distributed, you can generate a reasonably accurate count of how many people cheated on a test without knowing an individual's answer to the question.



Secure Multiparty Computation

- **Homomorphic encryption –** a method of encryption that computes values from encrypted data sources without revealing private individual's data.
- Datasets can be compared to analyze whether:
    - ➢ People provided homeless services end up in housing or holding jobs;
    - ➢ Student aid helps students succeed; or
    - ➢ Certain kinds of support can prevent people from being re-admitted to hospitals.



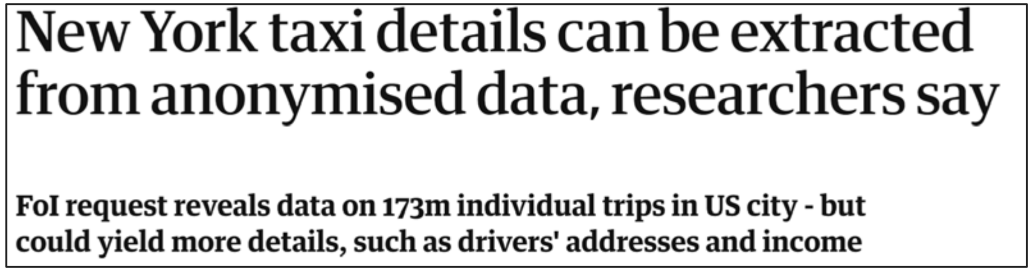
Two Risk Scenarios: De-Identification for Public and Non-Public Purposes

Public Release

Datasets where there are very serious efforts made to prevent re-identification, but where additional linking elements or additional data sets could lead to re-identification.

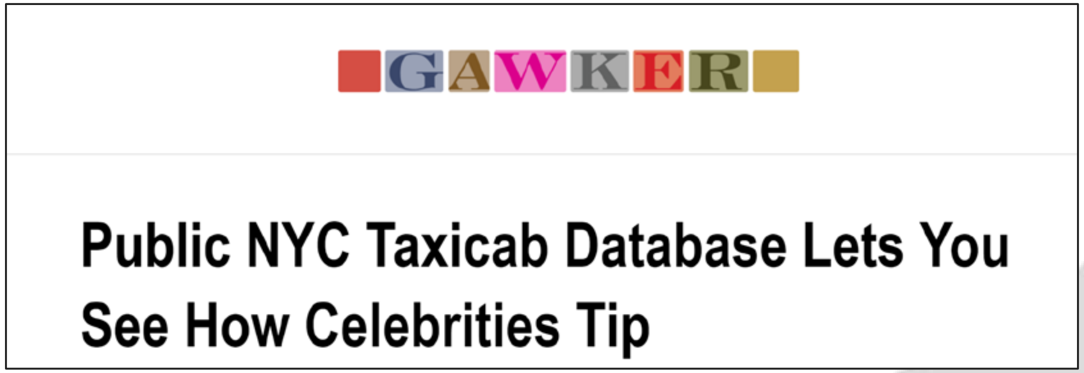
- Indirect identifiers are a significant concern.
- Legal and contractual controls are infeasible.
- Differential privacy is attractive in this scenario, but may be difficult to scale.

Uncontrolled Distribution or Weak Controls



New York taxi details can be extracted from anonymised data, researchers say

FoI request reveals data on 173m individual trips in US city - but could yield more details, such as drivers' addresses and income



GAWKER

Public NYC Taxicab Database Lets You See How Celebrities Tip

Controlled Release

Non-public

Datasets where de-identified or pseudonymous data are only shared with trusted parties, such as researchers, other organizational units or partners subject to controls.

- Indirect identifiers allow the performance of analytics, research, and fraud prevention.
- Legal and contractual controls may be relied on.

Controlled distribution or strong controls

Differential privacy is attractive for some use cases. Ex: Rappor, Apple (local differential privacy).

- If data set is small, adding noise may significantly impact accuracy.
- Not all calculations are supported.
- Company still holds original data set.
- Privacy budget limits number of queries.

Targeted Ads

Direct or Indirect Identifiers?

- IP Address
- Cookie
- Mobile Ad ID
- Location
- Lists of Web URLs or Apps used
- Demographics
- Customer IDs
- IDs linked to Direct Identifiers in the control of 3rd parties
- Profiles

Controls?

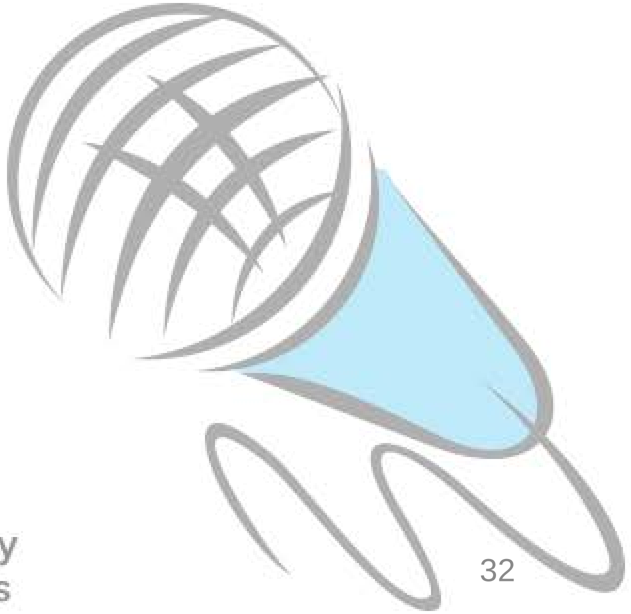
- Hashing
- Contracts
- Privacy Policies
- Self Regulatory Codes
- Cookie Expirations, Data Deletion
- Audits
- User Controls
- Header Bidding
- Exchanges
- Real Time Bidding

The Data Risk Spectrum: From De-Identified Data to Sensitive Individually Identifiable Data

Aoife Sexton

Trūata

TRŪATA.



Who We Are:

- We are a cloud-hosted, data anonymization, and analytics solution provider
- We are all about protecting privacy and powering results
- Our solution is ground-breaking and unique
- Our investors are Mastercard, IBM, and C3 IoT
- In production on IBM cloud with first customer

Implications of GDPR

Limits Data Collection and Retention

Analytics require historic data

GDPR principles of purpose limitation, data minimization and retention limit the use of data collected and how long that data may be retained

Tightens Criteria for Processing Personal Data

Stricter conditions for obtaining valid consent make finding a legal basis to process personal data more challenging

This leads to fragmented data sets, and an incomplete, and potentially biased, customer view

Raises Threshold for True Data Anonymization

Anonymization allows for further uses of data since data is no longer classified as personal data

Under GDPR, the bar has been raised on what will be classified as anonymized data

Data anonymized “in house” is likely to be classified as pseudonymized and not truly anonymized. It will still be considered to be personal data.

Businesses are looking for a way to protect their customers' privacy, maximize the value of their data assets, and minimize their risk of violating privacy regulations

Independent Anonymization

The Four Pillars of the Trūata Anonymization Solution

The strength and uniqueness of the Trūata Anonymization Solution lies in the combination of the four pillars:

Corporate Structural Safeguards

- A corporate Trust structure separates governance, assets and profit
- A Trust deed / constitution requires independence
- Changes to independence require notification to Supervisory Authority

Legal / Contractual Safeguards

- Formation documents prohibit use of data by Trūata other than for the customer
- Prohibits re-identification of data
- Customer Charter for ethical use of data
- Controller to Controller relationship

Organisational Controls

- Comprehensive privacy and information security compliance programs
- Data Protection Officer with staff of privacy experts
- Experienced data scientists and data analysts

Technology Platform

- Multi-step anonymization process
- Data sanitation and segregation
- Customised data modification techniques
- Outlier handling
- Differential privacy tests
- World-class analytics tools

The Data Journey

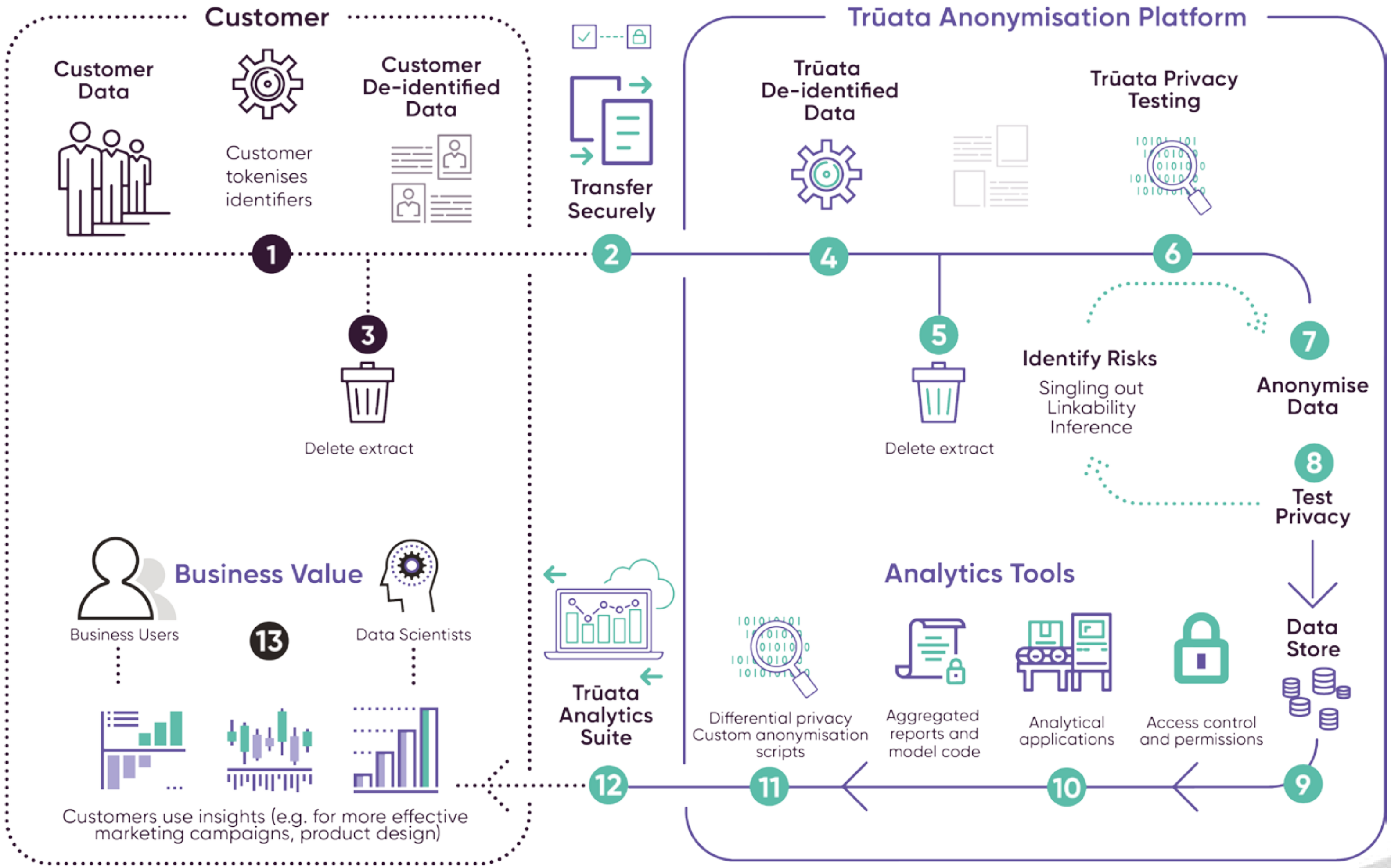
1-2 Customer removes personal data and tokenises direct and indirect identifiers then transfers to Trūata

4 Trūata tokenises direct and indirect identifiers

6-8 Trūata runs risk assessment routines to identify privacy risks and applies anonymization techniques to data

9-11 Trūata moves data to customer specific data storage and performs analytics to generate reports and model code

12-13 Aggregated reports and model code provided to customer. Customer can use those insights to improve their business processes and better serve their customers



Applicable for business critical modelling and analytics use cases across a wide range of industries

Marketing Campaign Insights

Customer Acquisition and Retention

Customer Segmentation and Modelling

Business Forecasting

Loyalty Program Earn and Burn

Business Experimentation

Performance Benchmarking

Customer Engagement

Retail

Media

Banking

Telcos

Hospitality

Automotive

Airlines

The Trūata Value

Protecting Privacy

Achieving true anonymization preserves privacy but is highly complex and difficult to achieve. It requires deep expertise in both data science and privacy.

Through the operation of the 4 pillars of the Trūata Anonymization Solution, Trūata can reduce the risk of re-identification to an insignificant level.

Anonymization can assist companies to act responsibly and ethically and to build trust with their end users.

Powering Results

Enabling access by companies of all sizes to the latest cloud based anonymization techniques and analytics technologies.

Enabling companies to analyse all of their customer data, not just subsets, resulting in less biased, more powerful analytics.

Enabling companies to maximise data utility and to take data driven decisions. This helps companies compete on a more level playing field.

The Data Risk Spectrum: From De-Identified Data to Sensitive Individually Identifiable Data

Panel Discussion:

Deven McGraw, Jules Polonetsky,
Michelle Richardson, Aoife Sexton,
Shane Wiley

Moderators: Cora Han & Elisa Jillson

FTC's 2012 Privacy Report

Data falls outside the scope of the framework (it is not reasonably linked to a specific consumer, computer, or other device) if:

(1) a given data set is not reasonably identifiable,

(2) the company publicly commits not to re-identify it,

(3) the company requires any downstream users of the data to keep it in de-identified form.

The Data Risk Spectrum: From De-Identified Data to Sensitive Individually Identifiable Data

Panel Discussion:

Deven McGraw, Jules Polonetsky,
Michelle Richardson, Aoife Sexton,
Shane Wiley

Moderators: Cora Han, Elisa Jillson

Break

12:00-1:00 pm

Remarks

Noah Joshua Phillips, Commissioner
Federal Trade Commission

Consumer Demand and Expectations for Privacy

Session moderated by:

Daniel Gilman
Federal Trade Commission
Office of Policy Planning

Laura Riposo VanDruff
Federal Trade Commission
Division of Privacy and Identity Protection

Consumer Demand and Expectations for Privacy

Panel Discussion:

Lorrie Faith Cranor, Avi Goldfarb, Ariel Fox Johnson,
Jason Kint, Laura Pirri,
Heather West

Moderators:
Daniel Gilman & Laura Riposo VanDruff

Current Approaches to Privacy, Part 1

Session moderated by:

Jared Ho & Laura Riposo VanDruff

Federal Trade Commission

Division of Privacy and Identity Protection

Current Approaches to Privacy, Part 1

Data Privacy Laws: Overview & Comparisons

Margot Kaminski

University of Colorado Law School

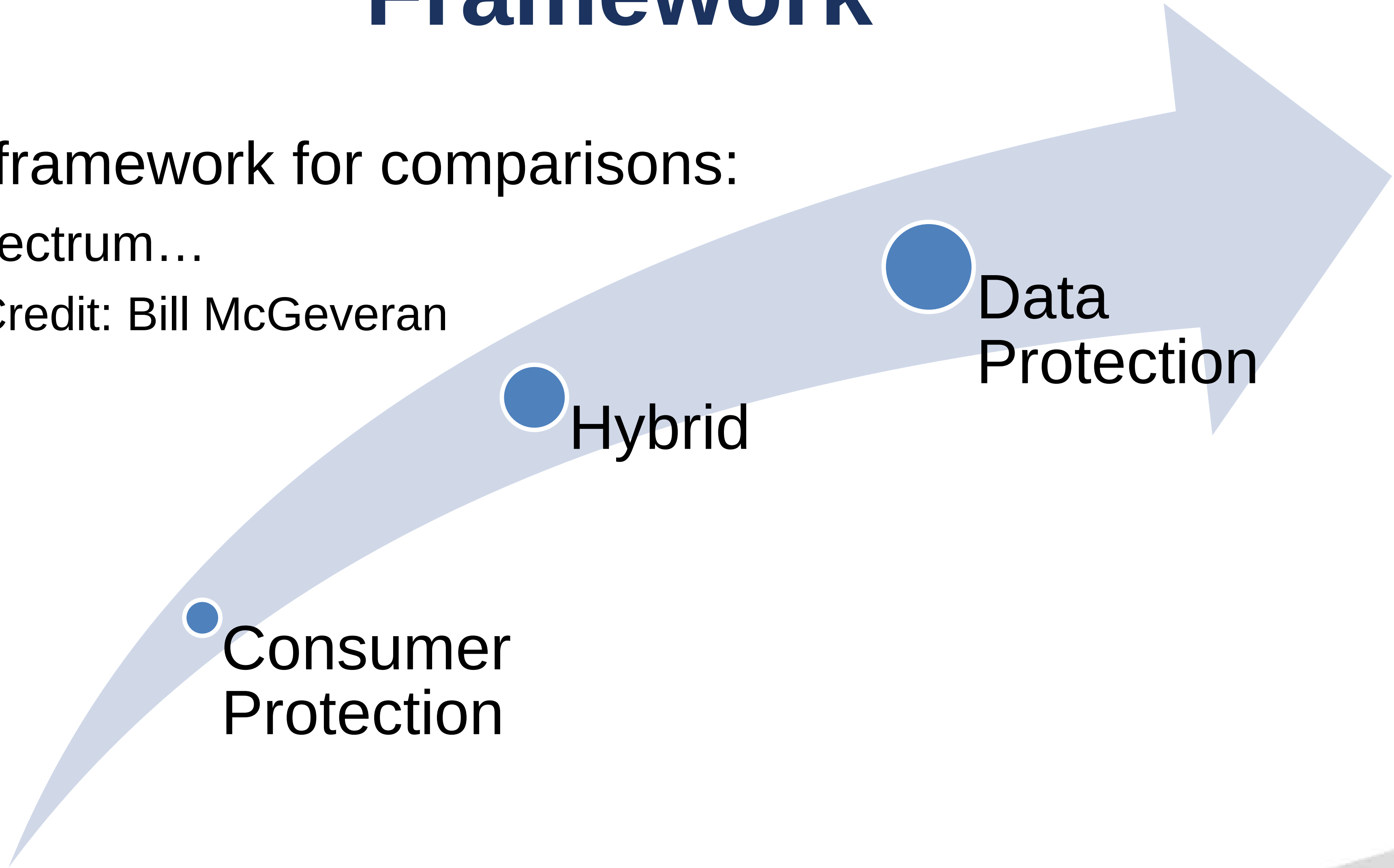
University of Colorado
Boulder

Comparing Data Privacy Laws

- Brief introduction, comparing approaches in:
  - U.S. Federal law(s)
  - The GDPR
  - Proposed and enacted state laws
- …all in 5-10 minutes!

Framework

- Basic framework for comparisons:
  - A spectrum…
    - Credit: Bill McGeveran



Consumer Protection

Hybrid

Data Protection

Additional Points of Comparison

- Omnibus vs. sectoral
- Notice & choice vs. …something else
- Individual rights vs. compliance
- Hard vs. soft law
  - enforcement vs. cooperation vs. self-regulation
  - rules vs. standards

Current federal law(s)

- FTC
  - Omnibus-ish
  - Largely consumer protection
    - Hard to reach third parties/data brokers
- Federal Sectoral Statutes
  - HIPAA, COPPA, GLBA
    - FIPPs-based, sometimes data-protection-like
      - notice & choice-centric

GDPR: High Level

- General Data Protection Regulation (GDPR)
  - Omnibus
    - Broad definition of personal data
  - Data protection par excellence
    - Follows the data
    - Including (especially) to govern third parties
  - Hard law
    - Significant fines
    - Both individual rights of enforcement & regulators & serious courts
  - …Combined with soft law/“collaborative governance”
    - Broad standards
    - private-public partnerships

What is in the GDPR?

- Core elements of the GDPR:
  - Individual Rights
  - Obligations for companies

What is in the GDPR?

- Individual Rights
  - FIPPs-based:
    - Notice, access, correction, erasure & others

What is in the GDPR?

- Obligations for companies:
  - Core principle of “accountability”
    - must be able to **demonstrate compliance with** GDPR.
  - Core principle of “lawfulness”
    - When a data controller processes personal data, there must be a **legitimate ground** for processing.
    - <u>Not just consent; in fact often **not** consent.</u>

What is in the GDPR?

- Obligations for companies, continued:
  - Transparency: affirmative notice
  - Documentation: keep records
  - Security obligations
  - Appoint a data protection officer (under certain circumstances)
  - Conduct impact assessments (under certain circumstances)
  - “Data protection by design and by default”

GDPR Wrap-up

- The GDPR is
  - A hard law data protection regime with soft law/collaborative features
  - That focuses on **both** individual rights and company compliance

California Consumer Privacy Act (CCPA)

- By comparison, the California Consumer Privacy Act (CCPA) is:
  - Somewhere between consumer protection & data protection
  - Omnibus-ish
    - Very broad definition of personal info
    - Limited to businesses (3-prong def)

What is in the CCPA?

- CCPA contains:
  - Notice & access rights
  - Limited deletion right
  - Limited opt-out right
- Enforceable by State AG
  - Who also promulgates rules

CCPA vs. GDPR

- Overlap on: transparency & individual control
- Diverge on: compliance/company obligations
- CCPA is missing many core elements of GDPR:
    - no “legal basis for processing”, purpose specification (maybe a little), use limitations, data minimization
- Vastly differing enforcement mechanisms
    - Private right of action
- Vastly differing court contexts
    - Different human rights backgrounds

Proposed state laws

- Proposed state laws largely mimic CCPA, **not** GDPR
    - But do evidence significant paradigm shift in U.S. data privacy laws:
        - Shift away from sectoral towards omnibus(-ish)
        - Shift towards data protection, from consumer protection model
- Variations:
    - Some add a private right of action
    - Some establish exploratory committees rather than law
    - Many focus on data security

Current Approaches to Privacy, Part 1

Panel Discussion:

Margot Kaminski, Fred Cate,
Markus Heyder, David LeDuc,
Laura Moy, Shaundra Watson

Moderators: Jared Ho & Laura Riposo VanDruff

Break

3:30-3:45 pm

Current Approaches to Privacy, Part 2

Session moderated by:

Andrea Arias & Elisa Jillson

Federal Trade Commission

Division of Privacy and Identity Protection

Current Approaches to Privacy, Part 2

Panel Discussion:

Lothar Determann, Jay Edelson,
Rebecca S. Engrav, Alan Raul,
Tracy Shapiro

Moderators: Andrea Arias & Elisa Jillson

Hypo #1

Company A, a U.S. start-up with a German subsidiary, offers a newsletter for cycling enthusiasts with information on safety, health and new cycling products, funded through ads. It is developing a new product that can sense dangers (e.g., weather changes, drunk drivers) and warn cyclists. Health insurance companies, auto makers, and city planners seek access to its data.

One day, an engineer inadvertently accesses a file containing name and health insurance provider for 200,000 employees and newsletter subscribers.

Hypo #2

Company B develops a free mobile app, with a location sharing opt-in, that offers shopping discounts based on location. City planners interested in making downtown shopping areas more “walkable” offer to pay for access to the app’s data.

Hypo #3

Company C sells fertility trackers, in which users can record the dates of sexual activity and diagnosis or treatment for a STD. Company C decides to provide access to de-identified data sets to pharmaceutical companies, public health advocates, and advertisers.

Carla Consumer doesn’t want her personal information to be sold. Frustrated that she can’t find a “Do Not Sell My Personal Information” link, she deletes the app. A year later, Carla asks Company C to delete all information about her.

Hypo #4

Company D sells smart coffee makers that can be connected to an alarm clock app. The company installs a microphone but does not disclose its presence. Three years later, Company D announces a software update that will activate the speaker so that it can respond to commands to make coffee. The company will also data-mine the voice recordings to improve the product.

Calvin Consumer is concerned that Company D may have recorded his conversations. He wants to access all data about him.

Hypo #5

Company E offers a free Internet browser to consumers. It mines browsing history and behavior to infer demographic information about consumers, which it sells to advertisers. It turns out that one popular data set is for females 10 to 12 years old.

Candace Consumer requests access to all data Company E stores about her so that she can correct any inaccurate data.

Closing Remarks

Jim Trilling

Federal Trade Commission

Division of Privacy and

Identity Protection

Thank You

Join Us Tomorrow at 9:00 am!