TIMNIT GEBRU
Independent Community Rooted Research

2023 SymSys Distinguished Speaker Lecture
Stanford University
Wednesday, Feb. 15, 2023 | 5 – 6:30 PM
Building 420, Room 040
Today’s Agenda

1. A reminder of our predicament
2. The dominant approach to privacy in the U.S.: notice and consent
3. Can the market solve our privacy problem?
   • The role of preferences
   • The privacy paradox
   • Measuring the value of privacy
4. Alternatives to the market: government regulation
   • Preferences vs. politics
   • Europe vs. USA
   • What are the alternatives?
Perspectives on Data Privacy

- Data privacy often involves a balance of competing interests

- Making data available for meaningful analysis
  - For public goods
    - Auditing algorithmic decision-making for fairness
    - Medical research and health care improvement
    - Protecting national security and personal safety
  - For private goods
    - Personalized advertising

- Protecting individual privacy
  - Personal value of privacy and respect for individual
  - Freedom of speech and activity
  - Avoiding discrimination
  - Preventing access from “adversaries”
Going Too Far?

What we commonly accept right now:

• The ability of social networks to connect you with who they think you want to connect with
• The power of recommendation systems to know exactly what you want
• The ability of advertisers to precisely target you

These same technologies can be put to uses with less benign consequences:

• Targeting in political campaigns (by candidates, foreign actors)
• Government surveillance (e.g. terrorism, immigration, etc.)
Cambridge Analytica scandal breaks (March 2018)

Used a digital app “This is Your Digital Life” to mine information on people who consented to participate and all of the people in their network (87m people)

Facebook discovered this in 2015 and asked that the data be deleted

It was not
Facebook’s Mark Zuckerberg says sorry in full-page newspaper ads

By Sheena McKenzie, CNN
Updated 2:17 PM EDT, Sun March 25, 2018

“We have a responsibility to protect your information. If we can’t, we don’t deserve it.”
The Consequences: None!

Wall Street Has Moved On From Cambridge Analytica Scandal

Facebook’s share price from January 2 to May 10, 2018 (closing prices)

On March 17, The Guardian broke the story of Cambridge Analytica harvesting millions of Facebook profiles to manipulate people.
Not Just The Private Sector

Suicide hotline shares data with for-profit spinoff, raising ethical questions

The Crisis Text Line’s AI-driven chat service has gathered troves of data from its conversations with people suffering life’s toughest situations.
Europe’s Approach: GDPR

Strengthen the privacy rights of individuals

- **Valid Consent**: Stricter rules for obtaining consent as a legal basis for processing.
- **Transparency**: The right to clear information over what data is collected and how it is processed.
- **Correction**: The right to rectify inaccurate personal data.
- **Erasure**: The right in certain cases to have personal data erased.
- **Data Portability**: The right to move personal data from one service provider to another.
- **Automated Processing**: The right not to be subject to a decision based solely on automated processing.
Far-Reaching and With Teeth

- Complete overhaul of data protection framework
- Applies across all member states of EU
- Applies to all organizations processing the data of EU residents – wherever those organizations are based
- Administrative penalties for breach of up to 4% of global revenue
- Data subjects have a right to bring actions (in their home state) and to receive damages if their rights have been breached
Facebook and GDPR

• Perhaps it will create pressure for regulatory convergence?

Commenting on the GDPR:

“We’re still nailing down the details on this, but it should directionally be, in spirit, the whole thing.”

Mark Zuckerberg to Reuters (April 2018)

“We’re going to make all the same controls and settings available everywhere, not just in Europe.”

Mark Zuckerberg (in a call to clean up the prior comment)
## China’s New Privacy Law

<table>
<thead>
<tr>
<th>Effective Date</th>
<th>Consent</th>
<th>Data Subject Rights</th>
<th>Penalties</th>
</tr>
</thead>
<tbody>
<tr>
<td>November 1, 2021</td>
<td>Required for disclosure of personal information, data transfers, and processing of sensitive data</td>
<td>Right to access, correction, and deletion of personal information</td>
<td>Administrative fine of up to 50 million RMB (about $7.7 million) or 5% of the processor’s turnover in the last year</td>
</tr>
</tbody>
</table>
Today’s Agenda

1. A reminder of our predicament
2. The dominant approach to privacy in the U.S.: notice and consent
3. Can the market solve our privacy problem?
   • The role of preferences
   • The privacy paradox
   • Measuring the value of privacy
4. Alternatives to the market: government regulation
   • Preferences vs. politics
   • Europe vs. USA
   • What are the alternatives?
A (Brief) History of Privacy Legislation

Our existing legal framework developed as a response to specific concerns about the data in the hands of major institutions:

- Credit (FCRA)
- Health information (HIPAA)
- Government-held information (Privacy Act)
- Educational information (FERPA)
- Other legislation addresses: genetic information, information about children, financial information, and electronic communications

Enforcement happens through public and private litigation and FTC regulation.
Notice and Consent

More generally, the U.S. approach to data held by companies is based on a combination of transparency and choice.

U.S. entities inform individuals of their respective information-flow practices and provide people with a choice to consent or not.

This approach has a broad appeal given:

1. Our conception of privacy as the right to control information about oneself
2. Our commitment to notions of a competitive free market, in which people can go elsewhere if they don’t like the terms
But We Cannot Keep Up..

• Expansion of data universe
• Gaps in coverage of existing laws and policies
• Limits of a focus on personally identifiable information, given aggregation, linking, etc.
• Increasingly unrealistic to rely simply on those who hold the data
• And the list goes on...
Can Technology Save Us?

What did we learn from Mehran?

1. Anonymization is kind of dead end
2. Encryption has been pretty effective so far but
   • quantum computing may defeat the public key and
   • if we successfully encrypt, we lose other things we care about (e.g. the right of government agencies to access data to protect people’s safety and security)
3. Differential privacy is limited in its application (protecting privacy in the context of statistical analysis)
4. Plus it is getting worse every day: digital trails, data aggregation, facial recognition, etc.
So What Do We Do?

- No one can agree on a definition of privacy
- Technical solutions address only part of the challenge
- We are trying to balance a set of competing interests

Is this a problem for...
Today’s Agenda

1. A reminder of our predicament
2. The dominant approach to privacy in the U.S.: notice and consent
3. Can the market solve our privacy problem?
   - The role of preferences
   - The privacy paradox
   - Measuring the value of privacy
4. Alternatives to the market: government regulation
   - Preferences vs. politics
   - Europe vs. USA
   - What are the alternatives?
Notice and Consent: In Practice

For the current light-touch approach to privacy regulation to work, three things must be true:

1. Individuals must be able to make informed, rational choices about the costs and benefits of different privacy policies
2. The market must be able to deliver a diversity of products with different privacy settings
3. We must be able to achieve the societal balance that we want between privacy and other values via a set of decentralized decisions

Are these things true in practice?
Measuring Preferences

How can you measure preferences regarding privacy and information sharing?


Three main approaches in social science:

• Public opinion polls
• Privacy behaviors
• Willingness to pay for privacy
Early Studies

Harris-Equifax Consumer Privacy Survey (1990) asked:

• Are you very concerned about threats to your personal privacy today?
• Do you agree strongly that business organizations seek excessively personal information from consumers?
• Do you agree strongly that the government since Watergate is still invading the privacy of citizens?
• Do you agree that consumers have lost all control of the circulation of their information?

Index coded as high concern (3 or 4 yes), moderate concern (2), low concern (1)
Categorizing People

• **Fundamentalists** are generally distrustful of organizations that ask for their personal information, worried about the accuracy of computerized information and additional uses made of it, and are in favor of new laws and regulatory actions to spell out privacy rights and provide enforceable remedies...

• **Pragmatists** weigh the benefits to them of various consumer opportunities and services, protections of public safety or enforcement of personal morality against the degree of intrusiveness of personal information sought and the increase in government power involved....

• The **Unconcerned** are generally trustful of organizations collecting their information, comfortable with existing procedures and not in favor of new privacy laws or regulations.
The Privacy Index

- What score did you receive? What category are you in? Why?
Confidence in Institutions

Few express confidence that their records will remain private and secure

<table>
<thead>
<tr>
<th>Institution</th>
<th>Very confident</th>
<th>Somewhat confident</th>
<th>Not too confident</th>
<th>Not at all confident</th>
<th>Don't know</th>
</tr>
</thead>
<tbody>
<tr>
<td>Your credit card companies</td>
<td>9</td>
<td>29</td>
<td>21</td>
<td>25</td>
<td>12</td>
</tr>
<tr>
<td>Government agencies</td>
<td>6</td>
<td>25</td>
<td>23</td>
<td>31</td>
<td>11</td>
</tr>
<tr>
<td>Your landline telephone company</td>
<td>6</td>
<td>25</td>
<td>21</td>
<td>29</td>
<td>15</td>
</tr>
<tr>
<td>Your cellular telephone company</td>
<td>5</td>
<td>26</td>
<td>25</td>
<td>31</td>
<td>11</td>
</tr>
<tr>
<td>Your email provider(s)</td>
<td>3</td>
<td>26</td>
<td>26</td>
<td>30</td>
<td>11</td>
</tr>
<tr>
<td>Your cable TV company</td>
<td>5</td>
<td>23</td>
<td>24</td>
<td>29</td>
<td>16</td>
</tr>
<tr>
<td>Companies or retailers you do business with</td>
<td>4</td>
<td>22</td>
<td>28</td>
<td>33</td>
<td>10</td>
</tr>
<tr>
<td>Your search engine provider(s)</td>
<td>2</td>
<td>14</td>
<td>25</td>
<td>41</td>
<td>15</td>
</tr>
<tr>
<td>The online video sites you use</td>
<td>1</td>
<td>10</td>
<td>24</td>
<td>42</td>
<td>19</td>
</tr>
<tr>
<td>The social media sites you use</td>
<td>1</td>
<td>10</td>
<td>24</td>
<td>45</td>
<td>18</td>
</tr>
<tr>
<td>The online advertisers who place ads on websites you visit</td>
<td>1</td>
<td>6</td>
<td>23</td>
<td>53</td>
<td>13</td>
</tr>
</tbody>
</table>

Source: Survey conducted August 5, 2014-September 2, 2014. Refused responses are not shown.

PEW RESEARCH CENTER
Hypotheticals

% of adults who would find these different scenarios acceptable or not acceptable

Office surveillance cameras
Several co-workers of yours have recently had personal belongings stolen from your workplace, and the company is planning to install high-resolution security cameras that use facial recognition technology to help identify the thieves and make the workplace more secure. The footage would stay on file as long as the company wishes to retain it, and could be used to track various measures of employee attendance and performance.

Auto insurance
Your insurance company is offering a discount to you if you agree to place a device in your car that allows monitoring of your driving speed and location. After the company collects data about your driving habits, it may offer you further discounts to reward you for safe driving.

Smart thermostat
A new technology company has created an inexpensive thermostat sensor for your house that would learn about your temperature zone and movements around the house and potentially save you on your energy bill. It is programmable remotely in return for sharing data about some of the basic activities that take place in your house like when people are there and when they move from room to room.
Privacy and Information Sharing

- Which scenarios did you find acceptable? Not acceptable? Why?
Measuring Behavior

Privacy Paradox

Stated

Private Data

Actual

Incentive

Navigation

Encryption

Illusion

Private Data

HTTP://
Willingness to Pay

- Another strategy is to estimate how much people value privacy in monetary terms.
- One study estimated that students would be willing to pay between $30 and $44 to use a website that protected against improper access and secondary use of their data.
- But these estimates do not seem stable:
  - An experiment offered gift cards to people at a mall $10 for an anonymous gift card and $12 for one that was trackable.
  - They were all offered the opportunity to trade.
  - 50% of those who got the anonymous card kept it; only 9% of those who got the trackable card switched.
  - People value privacy more when they have it than when they do not.
Preferences Depend on Context

Public’s shifting concerns on security and civil liberties

Bigger concern about govt anti-terrorism policies? (%)

- Not gone far enough to protect country
- Gone too far restricting civil liberties

Post-Snowden, increased opposition to govt surveillance

The government’s collection of telephone and internet data as part of anti-terrorism efforts

Source: Survey conducted Dec. 8-13, 2015.
Don’t know responses not shown.

PEW RESEARCH CENTER
Preferences are Malleable

Default visibility settings in social media over time

- Visible (default setting)
- Not visible

2005:
- Entire Internet
- Facebook
- Friends
- User
- Contact information
- Networks
- Wall
- Basic profile data
- Birthday
- Names
- Photos

2014:
- Entire Internet
- Facebook
- Friends
- User
- Contact information
- Networks
- Wall
- Extended profile data
- Likes
- Names
- Photos
- Basic profile data
- Gender
- Picture
The Challenge with Preferences

- They are difficult to measure
- They do not correlate with behavior
- They are context dependent, malleable
- They are unstable

- This is a problem for models of consumer choice
  - A market model depends on individuals being able to choose the products that have the features they most want

- Note: this is also a problem for democratic politics
  - One job of politicians is to represent the preferences of the public
A Competitive Marketplace

“I’m not even going to try and compete on privacy...”

Image: Screenshot www.neeva.com; Wikimedia Commons CC BY-SA 4.0
A Competitive Marketplace?

Image: Flickr, CC BY 2.0
Value of Collective Deliberation

• Important social goals can only be achieved if we agree to sacrifice privacy at times
• Need to think about the costs and benefits of privacy “cumulatively and holistically”
• Obvious cases: national security, pandemic tracking/response, development of new disease treatments
• Calls for a regulatory approach
Today’s Agenda

1. A reminder of our predicament
2. The dominant approach to privacy in the U.S.: notice and consent
3. Can the market solve our privacy problem?
   • The role of preferences
   • The privacy paradox
   • Measuring the value of privacy
4. Alternatives to the market: government regulation
   • Preferences vs. politics
   • Europe vs. USA
   • What are the alternatives?
How Much Variation Is There?

- There is evidence of strong policy convergence around a set of core fair information principles (Bennett 1992)
  - Openness
  - Individual access and correction
  - Collection limitation
  - Use limitation
  - Disclosure limitation
  - Security

- 43 information privacy laws had been enacted by the early 2000s, mostly across Europe and other OECD countries

- By 2016, more than 120 countries had adopted similar laws
# What About Implementation?

<table>
<thead>
<tr>
<th>Self-Regulation</th>
<th>Privacy Self-Management</th>
<th>Regulation by Supervisory Bodies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information practices are implemented internally by organizations with limited external intervention</td>
<td>Control is ensured through the active participation and intervention of those whose data is held</td>
<td>Interaction between data subjects and controllers is mediated by a third party</td>
</tr>
</tbody>
</table>
Generating Hypotheses

Why does Europe adopt more a more assertive regulatory approach, while the United States relies on self-regulation and privacy self-management?

- Different values (as a function of experience, constitution, etc.)
- Pressure from organized interests, groups
- Partisan and electoral competition
- Bureaucratic interests
How are policies adopted in a democracy? It requires thinking about:

- Preferences
- Organized interests
- Aggregation
- Decision-making
- Implementation
The Role of Organized Interests

So if policy choices are unlikely to reflect some aggregate assessment of what the public wants, what do they reflect?

What kinds of interest groups organize to influence the policy process on data privacy? What do you think they want?
Corporations

INTEL'S APPROACH TO PRIVACY

Intel has a longstanding history of supporting privacy. Recent discussion about the need for a US federal law inspired us to draft a bill that will optimize for both innovation and protecting privacy.

We published a draft of that bill on this site in November of 2018, and invited some of the country’s top privacy experts to discuss the draft. We also wanted to hear from you.
Tech Companies

YOU’RE LOOKING AT ONE OF THE MOST POWERFUL MEN IN THE WORLD

(but this one won’t steal your data)

Meet Tim Cook, Apple CEO

Image: Apple advertisement
The Ad Tech Business Model

What's Ad Growth Going to Look Like in the Digital Arena?
Digital advertising revenue worldwide from 2015 to 2021, by format (in billion U.S. dollars)

- Banner Advertising
- Video Advertising
- Search Advertising
- Social Media Advertising
- Classifieds

* forecast

Source: Statista Digital Market Outlook
Associations

Images: U.S. Chamber of Commerce; AccessNow; Internet Association; EPIC.
Civil Society
# States and Municipalities

<table>
<thead>
<tr>
<th>CCPA</th>
<th>GDPR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Requires consent from individual</td>
<td>Has a legitimate interest component</td>
</tr>
<tr>
<td>Wide definition of personal information</td>
<td>Defines PII and sensitive information</td>
</tr>
<tr>
<td>Allows for opt-out</td>
<td>Defaults for opt-in</td>
</tr>
<tr>
<td>Fines up to millions of dollars</td>
<td>Fines up to millions of dollars</td>
</tr>
<tr>
<td>Private right to action, class suits</td>
<td>Public complaint to enforcement agency</td>
</tr>
</tbody>
</table>

Image: Flickr, CC BY 2.0
Breaking Governance Down

How are policies adopted in a democracy? It requires thinking about:

- Preferences
- Organized interests
- Aggregation
- Decision-making
- Implementation
Today’s Agenda

1. A reminder: our predicament
2. Contemporary policy debates around privacy
3. How do we govern privacy
   • The role of policy preferences
   • The privacy paradox
   • Measuring the value of privacy
   • Organized interests
4. How should we govern privacy
   • Issues with notice and consent
   • What are the alternatives?
Should the US Become More Like Europe?

• Supporters of notice and consent “see regulatory protection of privacy as interfering with the fundamentally benign trajectory of information technologies and the benefits they unlock” and argue that individuals should be able to make their own self-interested decisions about what to disclose.

• Others argue that “regulatory intervention may be needed to balance the interests of the subjects of data against the power of commercial entities and governments holding that data.”

• Where do you come down?
But GDPR is Far from Perfect

• The critics say that GDPR:
  • Has strengthened the largest companies, and penalized small and medium sized businesses
  • Is cost prohibitive to comply with
  • Threatens innovation and research
  • Has not generated greater trust online
  • Increases the power of government under the pretense of protecting the consumer
  • Doesn’t incentivize private sector innovation to protect privacy

Source: Testimony of Roslyn Layton, U.S. Senate Judiciary Committee, March 2019
How Far Do Pop-Ups Get Us?
Are individuals up to the challenge of navigating privacy in the information age?

Social scientists are skeptical (Acquisti et al 2015). Lawyers are concerned (Solove 2013). Information scientists doubt it (Nissenbaum 2011). Why?

1. People are uncertain about their preferences
2. Preferences are context dependent
3. Privacy preferences can be manipulated
4. Privacy self-management does not scale well
5. People cannot factor in aggregation
6. People cannot anticipate harm

An Honest Take
Alternatives?

1. Comprehensive privacy regulation, but the challenge is that this denies people the freedom to make choices (paternalism)
   - It is not always clear how to trade off privacy and data use
   - There are social benefits to data aggregation and benefits

2. Improving privacy self-management through:
   - Opt-in consent rather than opt-out consent
   - Managing privacy globally rather than locally
   - Shift to managing downstream use
   - Codify basic privacy norms about what is not acceptable
If bipartisan cooperation can be achieved on anything, comprehensive privacy legislation is the low-hanging fruit.

A super majority of registered voters want a bill in this Congress.

In the last Congress, there were bills on both sides of the aisle:

Consumer Online Privacy Rights Act (Cantwell)
SAFE Data Act (Wicker)

And House versions as well. Some version of each will be re-introduced.

Time to do your homework!