Today’s Agenda

1. Technical considerations in data privacy
   • Data storage and encryption
   • Multi-party privacy
   • Information discovery
   • Issues in facial recognition
2. Regulatory considerations in data privacy
3. Methods for Resolving Value Trade-Offs:
   • Utilitarianism and/versus Rights
Today’s Agenda

1. Technical considerations in data privacy
   • Data storage and encryption
   • Multi-party privacy
   • Information discovery
   • Issues in facial recognition
2. Regulatory considerations in data privacy
3. Methods for Resolving Value Trade-Offs:
   • Utilitarianism and/versus Rights
• Recall that data privacy involves competing interests

• A conversation from a few years ago:
  • Grade data should not be stored on non-Stanford systems
  • E.g., Should not use Google spreadsheet to store students grades
  • Reason: security of the underlying data (FERPA)

• *Verifiably true story time!*
  • Where did my PhD thesis go?

• Are you concerned about your grades in this class being stored in a Google spreadsheet? If so, what might mitigate your concerns?
Encryption

• Perhaps we could encrypt all stored data
  • If there is a system breach, data is “safe”
  • Why might this not really be the case?
    • Data is often decrypted for analysis (and left that way)
    • Breach could lead to leak of decryption key

• Perhaps government should be able to decrypt data for security
  • Encryption systems could provide a backdoor
  • E.g., NSA can have access to “key escrow” to get decryption key

• Do you see any potential problems?
In 2021, Apple announced it would scan photos on iCloud for child sexual abuse material (CSAM)
- Apple pulls release of this feature after public outcry
- Objection: such scanning could be extended by different governments interesting in finding other content they find objectionable

Proposals also exist for in-app content moderation in end-to-end encrypted messaging systems
- Machine learning model in messaging app that detects harmful content
- E.g., Apple’s Communication safety in Messages feature

Messages analyzes image attachments and determines if a photo contains nudity, while maintaining the end-to-end encryption of the messages. The feature is designed so that no indication of the detection of nudity ever leaves the device. Apple does not get access to the messages, and no notifications are sent to the parent or anyone else.

Source: https://www.apple.com/child-safety/
Today’s Agenda

1. Technical considerations in data privacy
   • Data storage and encryption
   • **Multi-party privacy**
   • Information discovery
   • Issues in facial recognition
2. Regulatory considerations in data privacy
3. Methods for Resolving Value Trade-Offs:
   • Utilitarianism and/versus Rights
Multi-Party Privacy

- Traditional privacy set-up with a data subject and a data user
  - Platform wants to use your data for personalization and marketing
  - Medical researcher wants to use your data to improve healthcare
  - NSA needs to determine contents of phone for national security

- Increasingly live in multi-party privacy settings
  - Alice posts a picture (or a tweet) tagging Bob and Chris
  - Bob may untag himself (if he is notified about the tag)
    - Untagging may offend Alice, so Bob may not do it (modify behavior)
    - Bob’s friends may still see the post (if friends with Alice or Chris)
    - Untagging or even deleting the post may not matter, if post was copied
  - Bob can (and, research shows, will) change behavior with respect to Alice to avoid being in posts (i.e., unwanted privacy violations)

- Have you modified your behavior with others to prevent multi-party privacy breaches? Or worry about offending others?
Today’s Agenda

1. Technical considerations in data privacy
   • Data storage and encryption
   • Multi-party privacy
   • **Information discovery**
   • Issues in facial recognition

2. Regulatory considerations in data privacy

3. Methods for Resolving Value Trade-Offs:
   • Utilitarianism and/versus Rights
Information Discovery

• Data minimalism
  • Only collect the data needed at time of data gathering
  • View espoused by GDPR

• Data maximalism
  • Gather as much data as possible
  • Don’t know a priori what data may be useful later

• Is choice context dependent?
  • What would you want Google to do?
  • What would you want your medical practitioner to do?

• Bonus Story Time!
  • Unexpected automotive customers
  • Diapers and what?!
1. Technical considerations in data privacy
   • Data storage and encryption
   • Multi-party privacy
   • Information discovery
   • Issues in facial recognition
2. Regulatory considerations in data privacy
3. Methods for Resolving Value Trade-Offs:
   • Utilitarianism and/versus Rights
Facial recognition suffers from the same biases as other forms of algorithmic decision-making
  • Gender Shades project by Joy Buolamwini and Timnit Gebru
  • Video

Issue of classification accuracy
  • Less data for darker-skinned people, especially women

But, in criminal facial recognition, often have more data of darker-skinned people
Classification vs. Matching

- In classification task, we are learning parameters for model based on data
  - This is the case when we are determining if an image contains a face
  - More data often means better classification
  - Built in assumption that the testing (actual usage) data is same distribution as training data
  - So, if we don’t do a good job classifying groups with limited data in training set, we don’t expect to get penalized for this in the testing set (i.e., actual usage)

- Recall the classifier that is 99.5% correct if it predicts “negative” all the time and only 0.5% of population are positive for some condition
Classification vs. Matching

• In matching task, we are trying to find closest matching image in a dataset
  • This is the case when we have image of a face and are trying to identify whose face it is
  • More data often means higher likelihood of a match
  • Go ahead, upload more pictures of yourself to Facebook. I’ll wait.

• In criminal justice context, higher likelihood of match can mean higher chance of misidentification

• African-Americans often under-represented in datasets for classification of faces and can be over-represented in datasets for matching of faces
  • They are hurt in both the classification and matching contexts
“Nearly 40 percent of Rekognition’s false matches in our test were of people of color, even though they make up only 20 percent of Congress.”

“In a recent letter to Amazon CEO Jeff Bezos, the Congressional Black Caucus expressed concern about the ‘profound negative unintended consequences’ face surveillance could have for Black people, undocumented immigrants, and protesters.”
“London police trial of facial recognition technology generated 104 ‘alerts’, of which 102 were false.”

“Another trial by South Wales police returned 2,400 false positives from CCTV footage gathered at UEFA football matches and the like.”

“Legislation currently before the Australian Parliament would allow national security agencies to use driver's license photos and, potentially, social media images to match with CCTV footage.”

Okay, now go delete all those Facebook images. I’ll wait.
The Metropolitan Police, the U.K.’s biggest police department with jurisdiction over most of London, announced Friday it would begin rolling out new ‘live facial recognition’ cameras in London, making the capital one of the largest cities in the West to adopt the controversial technology.

Privacy activists immediately raised concerns, noting that independent reviews of trials of the technology showed a failure rate of 81%.
1. Technical considerations in data privacy
   - Data storage and encryption
   - Multi-party privacy
   - Information discovery
   - Issues in facial recognition

2. Regulatory considerations in data privacy

3. Methods for Resolving Value Trade-Offs:
   - Utilitarianism and/versus Rights
More Like Europe?
The EU General Data Protection Regulation (GDPR) is the most important change in data privacy regulation in 20 years.

The regulation will fundamentally reshape the way in which data is handled across every sector, from healthcare to banking and beyond.
<table>
<thead>
<tr>
<th>CLEAR LANGUAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TODAY</strong></td>
</tr>
<tr>
<td>Often businesses explain their privacy policies in lengthy and complicated terms</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CONSENT FROM USER</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TODAY</strong></td>
</tr>
<tr>
<td>Businesses sometimes assume that the user's silence means consent to data processing, or they hide a request for consent in long, legalistic, terms and conditions — that nobody reads</td>
</tr>
<tr>
<td>TODAY</td>
</tr>
<tr>
<td>-----------------------------------------------------------------------</td>
</tr>
<tr>
<td>The user might not be informed when his/her data is transferred outside the EU</td>
</tr>
<tr>
<td>Sometimes businesses collect and process personal data for different purposes than for the reason initially announced without informing the user about it</td>
</tr>
<tr>
<td>Businesses use algorithms to make decisions about the user based on his/her personal data (e.g. when applying for a loan); the user is often unaware about this</td>
</tr>
<tr>
<td>STRONGER RIGHTS</td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td><strong>TODAY</strong></td>
</tr>
<tr>
<td>Often businesses do not inform users when there is a data breach, for instance when the data is stolen</td>
</tr>
<tr>
<td>Often the user cannot take his/her data from a business and move it to another competing service</td>
</tr>
<tr>
<td>It can be difficult for the user to get a copy of the data businesses keep about him/her</td>
</tr>
<tr>
<td>It may be difficult for a user to have his/her data deleted</td>
</tr>
</tbody>
</table>
## Stronger Enforcement

<table>
<thead>
<tr>
<th>TODAY</th>
<th>TOMORROW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data protection authorities have limited means and powers to cooperate</td>
<td>The <strong>European Data Protection Board</strong> grouping all 28 data protection authorities, will have the powers to provide <strong>guidance</strong> and <strong>interpretation</strong> and adopt <strong>binding decisions</strong> in case several EU countries are concerned by the same case</td>
</tr>
<tr>
<td>Authorities have no or limited fines at their disposal in case a business violates the rules</td>
<td>The 28 data protection authorities will have harmonised powers and will be able to <strong>impose fines</strong> to businesses up to 20 million EUR or 4% of a company’s worldwide turnover</td>
</tr>
</tbody>
</table>
Developments in California

California Consumer Privacy Act comes into effect in January 2020. It is modeled on the GDPR.

California Privacy Rights Act passes (as proposition 24) to strengthen the CCPA in November 2020.

It establishes the California Privacy Protection Agency, the first dedicated privacy regulator in the United States.

23 states began considering their own privacy legislation in 2021. Colorado and Virginia each passed a bill.
Crafting a National Privacy Law

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

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

And House versions as well.

As these bills come up for the debate, the question is how far do you want to go.
Scenario for Discussion

• Imagine you are a Member of Congress and you need to take a position on the following four issues in the privacy legislative debate:

1. Whether to require that companies seek “meaningful, informed” consent from users around how their data will be used
2. Whether to mandate that consumers “opt-in” for the collection of personal data by private companies
3. Whether to obligate companies to ensure that consumers can take their data with them if they wish
4. Whether to empower a new federal agency with responsibility for protecting and enforcing privacy rights

• What position would you take on each of these issues? How would you decide? How might your position depend on the part of the country you represent?
• What might be most challenging in making these ideas operational?
Graham, Cotton, Blackburn Introduce Balanced Solution to Bolster National Security, End Use of Warrant-Proof Encryption that Shields Criminal Activity

WASHINGTON – Senate Judiciary Committee Chairman Lindsey Graham (R-South Carolina) and U.S. Senators Tom Cotton (R-Arkansas) and Marsha Blackburn (R-Tennessee) today introduced the Lawful Access to Encrypted Data Act, a bill to bolster national security interests and better protect communities across the country by ending the use of “warrant-proof” encrypted technology by terrorists and other bad actors to conceal illicit behavior.

“Terrorists and criminals routinely use technology, whether smartphones, apps, or other means, to coordinate and communicate their daily activities. In recent history, we have experienced numerous terrorism cases and serious criminal activity where vital information could not be accessed, even after a court order was issued. Unfortunately, tech companies have refused to honor these court orders and assist law enforcement in their investigations. My position is clear: After law enforcement obtains the necessary court authorizations, they should be able to retrieve information to assist in their investigations. Our legislation respects and protects the privacy rights of law-abiding Americans. It also puts the terrorists and criminals on notice that they will no longer be able to hide behind technology to cover their tracks,” said Graham.
Today’s Agenda

1. Technical considerations in data privacy
   • Data storage and encryption
   • Multi-party privacy
   • Information discovery
   • Issues in facial recognition
2. Regulatory considerations in data privacy
3. Methods for Resolving Value Trade-Offs:
   • Utilitarianism and/versus Rights
Utilitarianism

Sum up all the values of all the pleasures on the one side, and those of all the pains on the other. The balance, if it be on the side of pleasure, will give the good tendency of the act upon the whole, with respect to the interests of that individual person; if on the side of pain, the bad tendency of it upon the whole.

" " Jeremy Bentham
Utilitarianism

What is Utilitarianism? a **general theory of what is valuable** and of **right action**.

1. **A Theory of the Good?**

   A **Theory of what is valuable**: happiness (and only happiness) is intrinsically good.

   **Bentham**
   - **Hedonistic view.** Only pleasure or happiness is intrinsically good; only pain intrinsically bad. Happiness is constituted by **subjective** experiences.

   **J. S. Mill**
   - there are higher and lower pleasures. Happiness is constituted by enjoyment of higher pleasures. An **objective** account.
Utilitarianism

What is Utilitarianism? a general theory of what is valuable and of right action.

2. A Theory of the Right

The right action is that action, of all available alternatives, that produces the greatest net balance of happiness.
Utilitarianism

Utilitarianism has four core elements

**Consequentialism:** right and wrong to be determined by an action’s consequences, real world states of affair. Intentions unimportant.

**Hedonism:** what matters most fundamentally is pleasure and pain; or happiness.

**Impartiality:** each person counts in the utilitarian calculation, and each person counts equally.

**Aggregation:** objective function is to maximize is the total sum of happiness.
1. Do we sum happiness over all living people only, or over all people into the indefinite future? (what discount rate should we apply to the unborn/future generations?)
Core Challenges to Utilitarianism

2. Do we count the happiness or unhappiness of all sentient beings, or only humans?

(utilitarianism for animals?)
3. And of course:

How do we compare the happiness/unhappiness of one person to that of another?

The problem of interpersonal comparison of utility
4. What about the distribution of risk and harm?

Is it permissible to impose increased risks/harms on some persons in order to deliver large gains for others?
5. Remember John Rawls?

*A Theory of Justice* provides an alternative to utilitarianism, which “fails to take seriously the distinction between persons.”
Entertain the following hypothesis:

Machine Learning is underwritten by utilitarianism:

the right action is the one that maximizes or optimizes.

If true, then we need to evaluate utilitarianism as a moral view.
Privacy is a value that is necessary to protect our higher-order interest in autonomy and self-determination.

The claim of individuals to determine for themselves when, how, and to what extent information about them is shared with or communicated to others.

QUESTION: How should we institutionalize – or put into practice – this value? How should we balance it against other values?
### Back to Square One: What is Privacy?

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>An <strong>interest</strong>?</td>
<td>(one consideration among many; can be defeated by other values or rights?)</td>
</tr>
<tr>
<td>A <strong>right</strong>?</td>
<td>(strong presumption against infringement)</td>
</tr>
<tr>
<td>A <strong>constitutional right</strong>?</td>
<td>(ordinary lawmaking can’t remove; sits above majoritarian decisionmaking)</td>
</tr>
<tr>
<td>An <strong>inalienable constitutional right</strong>?</td>
<td>(ordinary lawmaking can’t remove and right holder can’t renounce – no notice and consent!)</td>
</tr>
<tr>
<td>A <strong>fundamental human right</strong>?</td>
<td>(no government can remove/abridge)</td>
</tr>
</tbody>
</table>
Rival Values

What other values (or interests, or rights) might be in tension with privacy (however we construe it)?

1. National Security (e.g., terrorism)
2. Public Safety (e.g., crime)
3. Innovation
4. Convenience
Opinion

You Are Now Remotely Controlled

Once capitalists control the science and the scientists, the secrets and the truth.

By Shoshana Zuboff
Ms. Zuboff is the author of “The Age of Surveillance Capitalism.”
We Need a Law to Save Us From Dystopia

It’s not too late. And it better be comprehensive.

By Charlie Warzel
Mr. Warzel is an Opinion writer at large.
Opinion | THE PRIVACY PROJECT

Be Paranoid About Privacy

We need to take back our privacy from tech companies — even if that means sacrificing convenience.

By Kara Swisher
Ms. Swisher covers technology and is a contributing opinion writer.

Dec. 24, 2019
EVERY MINUTE OF EVERY DAY, everywhere on the planet, dozens of companies — largely unregulated, little scrutinized — are logging the movements of tens of millions of people with mobile phones and storing the information in gigantic data files. The Times Privacy Project obtained one such file, by far the largest and most sensitive ever to be reviewed by journalists. It holds more than 50 billion location pings from the phones of more than 12 million Americans as they moved through several major cities, including Washington, New York, San Francisco and Los Angeles.

Each piece of information in this file represents the precise location of a single smartphone over a period of several months in 2016 and 2017. The data was provided to Times Opinion by sources who asked to remain anonymous because they were not authorized to discuss it and feared repercussions for doing so. The opinions expressed are those of the authors, not of The New York Times.
How Your Phone Betrays Democracy

By Charlie Warzel and Stuart A. Thompson

DEC. 21, 2019
Total Surveillance Is Not What America Signed Up For

By The Editorial Board

DEC. 21, 2019
It is a federal crime to open a piece of junk mail that’s addressed to someone else. Listening to someone else’s phone call without a court order can also be a federal crime.

The Supreme Court has ruled that the location data served up by mobile phones is also covered by constitutional protections. The government can’t request it without a warrant.

But the private sector doesn’t need a warrant to get hold of your data. There’s little to prevent companies from tracking the precise movements of hundreds of millions of Americans and selling copies of that dataset to anyone who can pay the price.
You Should Be Freaking Out About Privacy

Nothing to hide, nothing to fear? Think again.

Featuring Farhad Manjoo and Kara Swisher  Video by Adam Westbrook

Dec. 20, 2019
Rival Values

• How should we balance privacy against security/public safety/innovation/convenience?
• What weight, under what circumstances, should we attach to these interests?
• Who should decide the framework?
• Who should decide who decides?