Today

- Privacy Issues

Homework

- Exercise Set 5 due
- Exercise Set 6 out
- Quiz next Wed
ACM/IEEE SE Ethics Clauses Related to Privacy ...

3.12 (Product): Work to develop software and related documents that respect the privacy of those who will be affected by that software.

6.07 (Profession): Be accurate in stating the characteristics of software on which they work, avoiding not only false claims but also claims that might reasonably be supposed to be speculative, vacuous, deceptive, misleading, or doubtful.

- ... think Terms of Service and Privacy Policies !!!

7.06 (Colleagues): Assist colleagues in being fully aware of current standard work practices including policies and procedures for protecting passwords, files, and other confidential information, and security measures in general.

8.05 (Self): Improve their knowledge of relevant standards and the law governing the software and related documents on which they work.
Basic Ideas in Privacy (Ch. 2)

“Personal Information”: Information related or traceable to an individual

- “Sensitive” information such as demographics, health, finance, politics, ...
- Personal identifying information such as phone, username, email, ...
- Non-text data such as images, videos, audio, ...

Exercise 1: What personal information are you (potentially) collecting in your senior design application?

Exercise 2: What personal information could you be collecting in your senior design app?

Inferring Personal Identity ...

- Even when direct PII is not stored (ids, names, etc.) ...
- It can sometimes be inferred (“traced” to an individual)
  - E.g., age + city + gender + ethnicity
  - Establishing “links” to other sources (like search engines)
- Similar issues with location tracking

⇒ What seems like innocuous data collection may not be!

Exercise 3: Revisit your answers to exercise 1 and 2 …
Who is interested in (our) personal information?

- Governments (foreign & domestic)
- Businesses/organizations (marketing, product development, pricing, ...)
- Companies that aggregate/combine/sell data ("legally") ... data brokers
- Thieves (collect and sell data “illegally”)

Exercise 4: How much do you value your data?
Your Data is Big Business!!! ... What it could actually be worth ...

It is unclear how much your data is really worth ... here is one take:\footnote{\url{medium.com/wibson/how-much-is-your-data-worth-at-least-240-per-year-likely-much-more-984e250c2ffa}}

- In 2017, US digital advertising revenues were around $83 Billion
  - Roughly 287 million US internet users
  - So, on average $289 was spent per US user (likely much higher!)
  - in 2017 Facebook made $39.9 Billion on ad revenue ($40 Billion total)
- In 2016, Equifax, Experian, and TransUnion made combined over $9 Billion
  - Roughly $10 per US adult
- “Stolen data” could be worth upwards of $300/yr (per person)
  - once source claims ...
    - a debit card, two credit cards, driver’s license, email and physical address worth roughly $100
    - passwords range from $3 (e.g., Netflix) to $1,000 (bank accounts)
    - complete medical records up to $1,000
- Insurance companies ...
  - e.g., purchase and track all types of health and non-health related data
  - a 2012 study found health care costs could be predicted on data like ethnicity, watching TV, and mail-order purchases\footnote{\url{www.propublica.org/article/health-insurers-are-vacuuming-up-details-about-you-and-it-could-raise-your-rates}}
- Data brokers ...
  - Multibillion dollar industry (e.g., AlianceData $7.5 billion in 2017 revenue for loyalty, marketing data)
  - Collect and infer (combine, connect, etc.) data for various companies
  - targeting data (new & existing customers), prediction analytics, id verification, fraud detection, people search,
Notes on the value of your data

• Some have argued (e.g., J. Lanier) that your data should not be free
  – E.g., would give you more legal rights and protections
  – And would potentially transform economies

• The real value of data is different than what is payed for it
  – e.g., if data was harder (more expensive) to get, would companies pay more?
  – the Target example does a nice job of capturing the real impact
  – ... which is that it can be used to manipulate people’s behaviors
  – e.g., Facebook, psychographics & microtargeting, the 2016 election
Three aspects of privacy

1. Freedom from intrusion (being left alone)
   - e.g., prying, trespassing, unauthorized access, search

2. Control of information about oneself

3. Freedom from surveillance
   - e.g., followed, tracked, watched, eavesdropped upon

Types of potential privacy threats we can face:

- Intentional, institutional uses of personal information
  - e.g., for law enforcement, taxation, marketing, etc.

- Unauthorized use or release by “insiders”

- Theft of information

- Inadvertent leakage through negligence or carelessness

- Our own actions (for gain and/or unaware of risks)
Q: What are current data privacy risks concerned with computing?

Personal data privacy risks (w.r.t. computing):

1. Everything done online can be recorded and linked to a device & location
2. Storage is cheap: companies, governments, ..., can store more & more data
3. Once stored, data seems to last “forever”
4. Most collection we’re unaware of (e.g., through mobile apps & websites)
5. Collection even that companies aren’t even aware of (e.g., libraries)
6. Keeping data secure can be hard
7. Even secured data can later be shared (policy changes, selling, gov. requests)
8. The “cloud”: dont control our own data, relying on others to for us

Other factors: Public vs Private Figures

- Privacy (generally) applies differently to “public figures”
  - e.g., public figures rely on publicity as a livelihood
- In general, actions in public aren’t private
  - e.g., going out in public implies permission to be photographed
  - freedom of speech, freedom of the press, ...

Exercise 5: What are “digital” examples?
Other factors: Informed Consent

- When company informs people about its data collection and use policies
- And a user can decide whether to use the product/service under the policies

Increasingly: opt-out or opt-in policies

- **Opt-out** requires user to explicitly not allow “secondary use” and collection of information (more common ... why?)

- **Opt-in** requires user to explicitly allow “secondary use” and collection of information

... see G.D.P.R.
Target Example from pre-2012

- Like most companies, uses “predictive analytics”
- I.e., to make predictions of customer behavior
- Used for direct marketing campaigns (e.g., send coupons)

Desire to influence (modify) customer shopping loyalty

- Many customers shopping “habits” are hard to change
- However, some big “life” events can change this
- One is during pregnancy
- Target wanted to know when customers were starting second trimester

After analyzing purchases of women who signed up for baby registries:

- Discovered increase in purchases of 25 products just prior to second trimester
- E.g., scent-free lotions and mineral supplements
- They wanted to identify and target ads (to modify their shopping habits)
- But without making it seem “creepy”

Exercise 6: Are there any privacy issues?