Claire Lobdell 10.30.2018

Ethics of Big Data talking points

Questions to ask when determining the ethics of data collection:

1. *How obvious is it that data is being collected and compiled*? Ex. Cross-site tracking with cookies, smart TVs that ping phones and other devices in the vicinity so that advertisers can link the data that's being generated across different accounts.

2. *How easy is it to opt out of that data collection?* Ex. Do you have to opt out or opt in? Can you find the privacy policy? Is there a straightforward way to submit a form or change your settings?

3. *Are the uses of the data collection made clear to the users*? Ex. Cambridge Analytica—Facebook users take a personality quiz, which is a smokescreen to give CA access to detailed user data that's otherwise hidden behind privacy settings—as many as 3K-5K "data points" per user. CA uses that to micro-target political ads for Brexit or to try to suppress African American votes, for example, sending ads to specific black voters about Hillary Clinton calling some underage criminals "super predators" while campaigning for Bill Clinton in 1996.

4. *How else could the data be used? And does the security of the system depend on the good intentions of the operator?*
   - Asking about citizenship on the US census
   - DACA database, Deferred Action for Childhood Arrivals, started in 2012 under Obama, asked young undocumented people to submit large amounts of personal data (immigration status, current and past addresses, biometric data)—feels a lot different in the hands of the current administration
   - Even people who are supposed to be really good at this stuff have trouble with it. E.g. Strava heat map. Australian college student Nathan Ruser dug into it and found heat maps around known US military bases in Iraq and Afghanistan—showing, for example, where soldiers go on runs; and also heat maps in the middle of deserts that are likely black sites.

5. *If you're already paying for a service, why should they be able to sell your data?*
   - Ex. As of March 2017, internet service providers are allowed to sell data about your internet use, including web browsing and app usage.
   - In 2017, iRobot, the company that sells Roomba, revealed that they store and plan to sell the house maps that Roombas create.
   - Summer 2018, GlaxoSmithKline, the pharmaceutical company, announced a $300 million dollar deal to access the genetic information from 23andMe customers

Other Talking Points

- Big tech companies = really ad companies. Google gets 84% of their revenue from ads, controls 40% of all US online advertising, Facebook controls another 20%
- Systems reflect the biases of their creators and biases of the ways data collected and which data is included. Ex. Child abuse prevention system in Allegheny County, PA (see *Automating Inequality* by Virginia Eubanks)—interactions with social safety net (food stamps, welfare

benefits, etc.) give subjects an automatically higher score, as do previous, unsubstantiated reports to CPS; facial recognition software that disproportionately misidentifies dark skinned people; LAPD’s data-driven policing to predict who will commit crimes: “the formula for determining whether someone's on the Chronic Offenders Bulletin is based partly on how often someone is interviewed by the police. But that's something that's simply more likely to happen in those places with heavier police presence.”

- Why do cookies and ad targeting matter? Dynamic/personalized pricing, “creepiness” factor; example of FB allowing ad targeting by race, hate group affiliation
- Data spill can’t be mopped up
- People in positions of authority (tech execs, legislators, etc.) sometimes say that because young people live so much of their lives online, they don’t care about privacy and sometimes use that to justify large-scale data harvesting and reuse. This feels like a convenient argument for big tech companies to make, belied by actual conversations with young people.

Harm reduction:

- Using search engine like DuckDuckGo that explicitly doesn’t track you
- Browser extensions like https everywhere and privacy badger (both from Electronic Frontier Foundation)
- Consider using a more secure browser, like Brave or Tor, or a VPN