


# CS 4001: Computing, Society & Professionalism

Munmun De Choudhury | Assistant Professor | School of Interactive Computing

**Week 6: Information Privacy**  
**February 21, 2017**



With your permission, you give us more permission. If you give us information about who some of your friends are, we can probably use some of that information, again, with your permission, or improve the quality of our searches. We don't need you to type at all, because we know where you are, with your permission. We know where you have been, with your permission. We can more or less guess what you are thinking about. – *Eric Schmidt, Google CEO (The Atlantic)*

# Information Technology Erodes Privacy

- Information collection, exchange, combination, and distribution easier than ever means less privacy
- Scott McNealy (Sun Microsystems): “You have zero privacy anyway. Get over it.”
- This class: we will consider how we leave an “electronic trail” of information behind us and what others can do with this info





# Perspectives on Privacy

# Defining Privacy

- Privacy related to notion of access
  - Privacy is not “being alone”, but defining who has access to what
- Access
  - Physical proximity to a person
  - Knowledge about a person
- Privacy is a “zone of inaccessibility”
- Privacy violations are an affront to human dignity
  - You violate privacy when you treat a person as a means to an end.
  - Some things ought not be known – you look away when your friend is typing their password
- Too much individual privacy can harm society
- Where to draw the line?

# Harms of Privacy

- Cover for illegal or immoral activities
- Burden on the nuclear family
- Hidden dysfunctional families
  - Incidents of domestic violence
- Ignored people on society's fringes
  - People with disability e.g., with mental illness

# Benefits of Privacy

- Individual growth
  - Necessary to blossom into a unique individual
- Individual responsibility
- Freedom to be yourself
  - Nobody likes to be videotaped all the time
- Intellectual and spiritual growth
- Development of loving, trusting, caring, intimate relationships



# Is There a Natural Right to Privacy?

- Privacy rights stem from property rights: “a man’s home is his castle”
- Coercive Acts before American Revolution (that evolved from the Quartering Act) – allowed billeting of soldiers in private homes
- This led to 3<sup>rd</sup> Amendment to U.S. Constitution – principle of home as a sanctuary in the Bill of Rights:

**No Soldier shall, in time of peace be quartered in any house, without the consent of the Owner, nor in time of war, but in a manner to be prescribed by law.**

# Is There a Natural Right to Privacy?

- Samuel Warren (Harvard graduate businessman) and Louis Brandeis (Boston attorney; later Supreme Court justice): People have “the right to be let alone”
- Judith Jarvis Thomson: the above definition is problematic
  - Smith being monitored at his home with a video camera without his knowledge – he is left alone technically, but it is a privacy violation
- Judith Jarvis Thomson: “Privacy rights” overlap other rights
- *Conclusion: Privacy is not a natural right, but it is a prudential right*

# Privacy and Trust

- Perhaps modern life is actually more private than life centuries ago
  - Most people don't live with extended families
  - Automobile allows us to travel alone
  - Television v. public entertainment
- Challenge: we now live among strangers
- Remedy: establishing reputations
  - Ordeal, such as lie detector test or drug test
  - Credential, such as driver's license, key, ID card, college degree
- Establishing reputation is done at the cost of reducing privacy



# Class Activity 1: Secret Monitoring

# Rule Utilitarian Evaluation

- If everyone monitored nannies, it would not remain a secret for long
- Consequences
  - Nannies would be on best behavior in front of camera
  - Might reduce child abuse and parents' peace of mind
  - Would also increase stress and reduce job satisfaction of child care providers
  - Might result in higher turnover rate and less experienced pool of nannies, who would provide lower-quality care
- Harms appear greater than benefits, so we conclude action was wrong

# Social Contract Theory Evaluation

- It is reasonable for society to give people privacy in their own homes
- Nanny has a reasonable expectation that her interactions with baby inside home are private
- Sullivan's decision to secretly monitor the nanny is wrong because it violates her privacy

# Kantian Evaluation

- Imagine rule, “An employer may secretly monitor the work of an employee who works with vulnerable people”
- If universalized, there would be no expectation of privacy by employees, so secret monitoring would be impossible
- Proposed rule is self-defeating, so it is wrong for Sullivans to act according to the rule

# Summary

- Three analyses have concluded Sullivans were wrong to secretly monitor how well their nanny takes care of their baby
- Morally acceptable options
  - Conduct more comprehensive interview of nanny
  - More thoroughly check nanny's references
  - Spend a day or two at home observing nanny from a distance
  - Be up-front with nanny about desire to install and use surveillance software on laptop





# Information Disclosures

# Public Records

- Public record: information about an incident or action reported to a government agency for purpose of informing the public
- Examples: birth certificates, marriage licenses, motor vehicle records, criminal records, deeds to property
- Computerized databases and Internet have made public records much easier to access

# Records Held by Private Organizations

- Credit card purchases
- Purchases made with loyalty cards
- Voluntary disclosures
- Posts to social network sites

# Data Gathering and Privacy Implications

- Facebook tags
- Enhanced 911 services
- Rewards or loyalty programs
- Body scanners
- Implanted chips
- OnStar
- Automobile “black boxes”
- Medical records
- Digital video recorders
- Cookies and flash cookies

# Facebook Tags

- Facebook allows users to tag people who are on their list of friends
- New feature from couple of years ago – automatic tagging
- About 100 million tags added per day in Facebook
- Facebook uses facial recognition to suggest name of friend appearing in photo
- Does this feature increase risk of improper tagging?

# Enhanced 911 Services

- Cell phone providers in United States required to track locations of active cell phones to within 100 meters
- Allows emergency response teams to reach people in distress
- What if this information is sold or shared?

# Rewards or Loyalty Programs

- Shoppers who belong to store's rewards program can save money on many of their purchases
- Computers use information about buying habits to provide personalized service
  - ShopRite computerized shopping carts with pop-up ads
- Do card users pay less, or do non-users get overcharged?

# Body Scanners

- Some department stores have 3-D body scanners
- Computer can use this information to recommend clothes
- Scans can also be used to produce custom-made clothing
- Can body scanners be misused?



# Implanted Chips

- Taiwan: Every domesticated dog must have an implanted microchip
  - Size of a grain of rice; implanted into ear
  - Chip contains name, address of owner
  - Allows lost dogs to be returned to owners
- RFID tags approved for use in humans
  - Can be used to store medical information
  - Can be used as a “debit card”
- What kind of privacy violations are possible with implanted chips?

- OnStar manufactures communication system incorporated into rear-view mirror
- Emergency, security, navigation, and diagnostics services provided subscribers
- Two-way communication and GPS
- Automatic communication when airbags deploy
- Service center can even disable gas pedal

# Medical Records

- Advantages of changing from paper-based to electronic medical records
- Quicker and cheaper for information to be shared among caregivers
  - Lower medical costs
  - Improve quality of medical care
- Once information in a database, more difficult to control how it is disseminated
  - What are possible risks?

# Digital Video Recorders

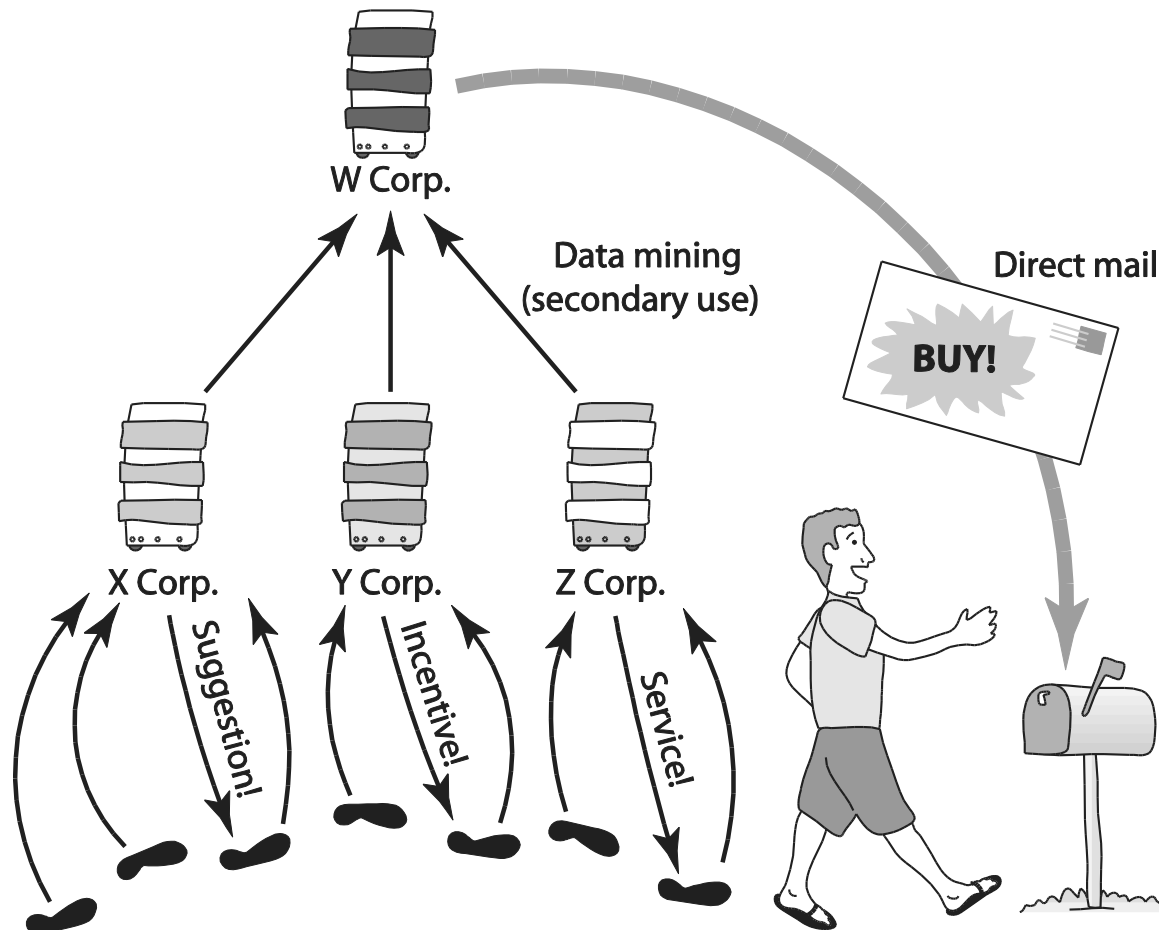
- TiVo service allows subscribers to record programs and watch them later
- TiVo collects detailed information about viewing habits of its subscribers
- Data collected second by second, making it valuable to advertisers and others interested in knowing viewing habits

# Data Mining

# Data Mining

- Searching records in one or more databases, looking for patterns or relationships
- Can be used to profiles of individuals
- Allows companies to build more personal relationships with customers

# Secondary Uses of Information




# How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did

- “[Pole] ran test after test, analyzing the data, and before long some useful patterns emerged. Lotions, for example. Lots of people buy lotion, but one of Pole’s colleagues noticed that women on the baby registry were buying larger quantities of unscented lotion around the beginning of their second trimester. Another analyst noted that sometime in the first 20 weeks, pregnant women loaded up on supplements like calcium, magnesium and zinc.”
- As Pole’s computers crawled through the data, he was able to identify about 25 products that, when analyzed together, allowed him to assign each shopper a “pregnancy prediction” score.
- More important, he could also estimate her due date to within a small window, so Target could send coupons timed to very specific stages of her pregnancy.





# Class Activity 2: Genetics and health insurance



How is secondary information  
used? Some examples...

# Google's Personalized Search

- Secondary use: Information collected for one purpose use for another purpose
- Google keeps track of your search queries and Web pages you have visited
  - It uses this information to infer your interests and determine which pages to return
  - Example: “bass” could refer to fishing or music
- Also used by retailers for direct marketing

# Collaborative Filtering

- Form of data mining
- Analyze information about preferences of large number of people to predict what one person may prefer
  - Explicit method: people rank preferences
  - Implicit method: keep track of purchases
- Used by online retailers and movie sites

# Microtargeting

- Started before 2004 US Presidential elections
- Political campaigns determine voters most likely to support particular candidates
  - Voter registration
  - Voting frequency
  - Consumer data
  - GIS data
- Target direct mailings, emails, text messages, home visits to most likely supporters

# Price Customization

- Data aggregation firms sell profiles of shopper to on-line retailers
- Retailers use this information to offer different prices to different people

# Credit Reports

- Example of how information about customers can itself become a commodity
- Credit bureaus
  - Keep track of an individual's assets, debts, and history of paying bills and repaying loans
  - Sell credit reports to banks, credit card companies, and other potential lenders
- System gives you more choices in where to borrow money
- Poor credit can hurt employment prospects

# Social Network Analysis

- Data mining now incorporating information collected from social networks
- Examples
  - Cell phone companies in India identify “influencers” – provide discounts
  - Police predict locations of big parties
  - Banks evaluate the riskiness of loans




# Ownership of Transaction Information

- Who controls transaction information?
  - Buyer?
  - Seller?
  - Both?
- Opt-in: Consumer must explicitly give permission before the organization can share info
- Opt-out: Organization can share info until consumer explicitly forbid it
- Opt-in is a barrier for new businesses, so direct marketing organizations prefer opt-out

# Class Discussion

- If you voluntarily have your body scanned at a departmental store, who should own that information: you or the store?
- Should the store have the right to sell your body measurements to other business? Explain your reasoning.



# Class Activity 3: Secondary information use



# Consumer Backlash

# Facebook Beacon

- Fandango, eBay, and 42 other online businesses paid Facebook to do “word of mouth” advertising
- Facebook users surprised to learn information about their purchases was shared with friends
- Beacon was based on an opt-out policy
- Beacon strongly criticized by various groups
- Facebook switched to an opt-in policy regarding Beacon

# Malls Track Shoppers' Cell Phones

- In 2011 two malls recorded movement of shopper by tracking locations of cell phones
  - How much time people spend in each store?
  - Do people who shop at X also shop at Y?
  - Are there unpopular areas of mall?
  - Small signs informed shoppers of study
- After protest, mall quickly halted study

# iPhone Apps Upload Address Books

- In 2012 a programmer discovered Path was uploading iPhone address books without permission
- Internet community pointed out this practice violated Apple's guidelines
- CEO of Path apologized; app rewritten
- Twitter, Foursquare, and Instagram also implicated for same practice

# Instagram's Proposed Change to Terms of Service

- Late 2012: Instagram announced changes
  - Privacy policy
  - Terms of service
- Legal experts: Instagram and Facebook would have right to use photos in ads without permission
- Instagram CEO: New policy misunderstood
- Changed advertising section of terms of service agreement back to original version





# Sharing of Anonymized Datasets

# Netflix Prize

- Netflix offered \$1 million prize to any group that could come up with a significantly better algorithm for predicting user ratings
- Released more than 100 million movie ratings from a half million customers
  - Stripped ratings of private information
- Researchers demonstrated that ratings not truly anonymous if a little more information from individuals was available
  - Movie ratings predicted political leanings and sexual orientation
- U.S. Federal Trade Commission complaint and lawsuit
- Netflix canceled sequel to Netflix Prize

# AOL Search Dataset

- In 2006, AOL research team released three months worth of search queries from 650K AOL users
  - Support university research
- Anonymization using a random integer identifier for each user
- But aggregation of queries by a single identifier revealed a lot about the person, even without PII
- Queries also contained personal info – address, SSN
- NYT identified several of the users
- Following public backlash, the dataset was taken down after 3 days
- Where did AOL go wrong?