

6.805/6.806/STS.085, Ethics and Law on the Electronic Frontier
Lecture 9: Personal Information on the Web

Lecturer: Hal Abelson

Personal Information on the Web

1. Everyone can snoop
2. What they have on you
3. Having sensitive information about you
4. Identity theft

George Washington Camera exercise

We are sort of in the middle as it relates to privacy. It hasn't quite gelled yet, and you can put up cameras and put it up and run it through face recognition programs and put it on the Web.

- A. How many people used to think that it was creepy to go on the Internet and find the map of your house?
 1. This was about 10 years or so ago.
 2. Hal did this to someone with information found about her neighbors and honors on the Internet. She was totally freaked.
 3. As long as you stay away from medical records and TV video rental records. These are protected. This is very different from the policies in Europe.
- B. How many people believe MIT should put up web cams?
 1. Is there a benefit? Maybe. Maybe not. Just to show that it's cool, say for advertising that you should consider coming to a particular school.
 2. Some students don't see a problem with using just one web camera, but sees privacy issues when you have several because then you can track an individual's everyday movement.
 3. The problem comes in with asymmetric use. Say when you don't know that you are being watched. Transparency issues.
 4. Another problem is how is the video capture or web cam being used.
 5. Also the issue of integrity issue. For instance delaying the feed and manipulating the images.
- A. Desktop Spyware
 1. Desktop Snooper - www.snooperspyware.com
 2. Desktop spyware issues
 - i. hidden camera that looks at any computer you choose
 - ii. marketing says that this type of software is about protecting your children
 3. Latest news is SONY software that limits the number of copies and Sony didn't tell you that this is what was being done. The software is also very difficult to remove.
- C. Global Storage per person (graph)
 1. Latonya Sweeney, MIT grad, went on to document the amount of Storage needed to document a person's activities, etc.

2. In 2000, Basically enough storage around the world to store information every 3.5 minute about a person.
3. MA - Fields on an electronic birth certificate now (1999)
 - a. Also lists things like “number of birth terminations”
 - b. Up to 226 fields recorded for every birth in Massachusetts

D. O’Harra’s book

1. ChoicePoint
2. Axiom – does thing to help employers?

D. Harvard News

1. The way that Harvard bills for medicine is by your student ID. The company that used the student IDs also recorded the drugs they ordered from Harvard medical service.
2. Student organization had a survey which linked the names and IDs.
3. The company actually put up the ID information on a public server. Claimed they didn’t do anything wrong because they were de-identified.

E. De-Indentification

1. Confidential information – birthdate, sex, ethnicity
 - a. we don’t want this information to get out, so we remove the names
2. De-identified records with zip code information or overlapping fields actually narrow down the possibilities of the particular individual making it easier to identify.
3. Uniquely identifying individuals
 - a. Latonya Sweeney research
 - b. Date of birth, gender, and 5-digit zip code uniquely identifies almost 90% of the USA population
4. Arkansas Juvenile Offender Records
 - a. The encoding used can significantly narrow down your selection of possible individuals.

F. Identity Theft

1. Exercise – Applying for a SSN
2. Was created originally to use only for social security program (1935)
3. Later the SSN became a “unique” identifier
 - a. first 3 numbers tell the state you were born in
 - 000 will never start a valid SSN number
 - b. digits 4 and 5 are known as group codes – this tells you when the SSN was assigned
 - c. Name a problem with the SSN identifier
 - allows you to falsify one easily
 - you can buy a SSN (socialsecuritypeoplesearch.com). However you must have a valid reason for locating the SSN of the person.
 - Use Google. People post their own SSNs on the web.
 - Professors publish the last 6 digits of a student’s SSN with their respective grades. Bad idea.
 - EDGAR database
 - SEC database – Turned out to be very easy to find SSN of rich people including Bill Gates

4. Are birth records public?
 - a. On Vital Search in CA, you can look up birth records that include the mother's maiden name

G. Social Security Death Index

1. Turns out that dead people really don't have any rights

The above items are all on public records and can be used to examine identity theft.

- H. People don't get how their personal information spreads. There is a difference between the friendly, neighborhood pharmacist and the large databases.
- I. There is a transition that the government needs to make and consumers need to enforce in finding a more secure system for identifying individuals. The current system of SSNs and personal information is not secure enough.

For the rest of the class, we will have a Presidential Commission that will convene a summit to rectify the SSN problem.

Exercise: World Summit on Identity Theft

How should we fix the identity problem?

Panel 1: Consumers Union, ACLU, etc.

1. Very difficult to prove that you are not the alleged impersonator, and it is very difficult to correct false information in a database.
2. Problems stem from 70-year old system and reuse of SSNs.
3. Propose unique issue of SSN-like identifier (10 digits) that accompanies present SSN. Numbers will be randomly generated, linked to biometric data and not associated with personal info.
4. For information to be released, the individual must give consent. Any organization or company that doesn't comply, will be fined \$1000 for each form of personal information.

Panel 2: Massachusetts State Police, Federal Trade Commission, etc.

1. Problem that confounds law enforcement agencies and poses threat to interstate commerce
2. USB key that will be given with the license registration
3. This will be better for identifying individuals
4. If key is lost, there is a 24-hour hotline to call. This will be no more expensive than operating a credit card customer service.
5. To retrieve a new USB key, then you must go to DMV and present 3 forms of ID.

Panel 3: VISA, and other financial ad insurance industries

1. Integral part of nation's commerce
2. Need access to certain information to provide services accurately (credit report, birth date, etc.)
3. Propose centralized database with limited information, keeping competition open.

4. Goal of database is to use as a reference and correctly verify that a person is who they say they are.
5. Another party will maintain the database.

Panel 4: eBay, Amazon, WalMart and other private sectors

1. New unique identifier system needs to be cheap. We don't want to pay for them and the consumer shouldn't have to pay for them.
2. Those that issue the identifier should pay for it.
3. We don't want restrictive laws about how we can share data.
4. More or less, the system is fine as is. Any new solution should not hurt the consumer's purchasing power and our profits.

Panel 5: Equifax, Experian, Acxiom, and other background-check and employment records industry

1. Every individual has a cryptographic card that has a personal identifier where they are required to sign
2. Two layers of security. (1) No one knows the personal identifier in the card, and (2) if the card is stolen, then you have to answer a secret question
3. The government will issue the cards. Must have 3 separate forms of ID (similar to receiving driving license.)
4. Helps with accuracy and security of information.
5. Willing to lose some money in order to gain security.

Other concerns

1. Name changes but no identity theft
2. Changes in physical identity

Deliberations of committee occurs.